Journal of Entomology and Zoology Studies

Available online at www.entomoljournal.com

E-ISSN: 2320-7078
P-ISSN: 2349-6800
JEZS 2015; 3 (3): 98-110

Received: 17-04-2015
Accepted: 20-05-2015

R. Udayasuriyan
Department of Zoology
Bharathiar University
Coimbatore – 641046 Tamil Nadu, India.

P. Saravana Bhavan
Associate Professor Department of Zoology Bharathiar University Coimbatore – 641046 Tamil Nadu, India.

C. Vadivalagan
Department of Zoology
Bharathiar University
Coimbatore – 641046 Tamil Nadu, India.

G. Rajkumar
Department of Zoology
Bharathiar University
Coimbatore – 641046 Tamil Nadu, India.

Correspondence:
P. Saravana Bhavan
Associate Professor
Department of Zoology
Bharathiar University
Coimbatore – 641046 Tamil Nadu, India.

Efficiency of different COI markers in DNA barcoding of freshwater prawn species

R. Udayasuriyan, P. Saravana Bhavan, C. Vadivalagan and G. Rajkumar

Abstract
Five species of prawns (*Macrobrachium rosenbergii*, *Macrobrachium malcolmsonii*, *Macrobrachium lamarrei*, *Macrobrachium lamarrei lamarroids* and *Caridina gracilipes*) were identified morphologically, and molecularly discriminated by mt-COI gene with four different sets of universal primers (LCO1490 & HCO2198, COIa & COIf, CrustF1 & HCO2198 and CrustF2 & HCO2198) in order to find out the best suited primer set. The primer sets, LCO1490 & HCO2198 and COIa & COIf was worked well in all subjected species. In terms of number of sequences the former primer set was found to be better than that of the latter primer set. But in terms of quality of phylogenetic information (synonymous and non-synonymous substitutions, saturations, transitional and transvertional type substitutions, divergence rates and phylogenetic tree topology just the reverse was resulted. Therefore, the primer set COIa & COIf showed more significant results in non-synonymous substitution, saturations, transvertional type substitution and divergence rates when compared with LCO1490 & HCO2198.

Keywords: DNA barcoding, mt-COI gene, Freshwater prawns, Substitutions, Divergence, Phylogenetic tree

1. Introduction
Freshwater prawns occur in a vast range of habitats, from torrential mountain streams down to sluggish, oligohaline waters. Amongst the freshwater prawn families, the two most families [Atyidae and the Palaemonidae (Subfamily, Palaemoninae)] are the near exclusively freshwater, which also have brackish water and marine representatives [1]. A total of 655 freshwater prawn species (just over a quarter of all described Caridean, the infra order) are presently known **[2].** It is difficult to estimate the true species richness of freshwater shrimps, as every year new taxa continue to be described, mainly in the two most numerically dominant genera, *Caridina* and *Macrobrachium* [2]. *Macrobrachium* is characterized by the extreme enlargement of the second pair of pereiopods, at least in the male, which in many species may exceed the body length. Most members of the genus are easily recognizable by the well developed, often elongated second chelipeds.

1.2 Problem of species identification
Prawns, like most other crustaceans, are able to change color, depending upon growth, background coloration and time of day [3]. *Macrobrachium* species have a wide distribution, in heterogeneous or geographically isolated environments can have a phenotype variation, because they are prone to show plastic responses to different environmental influences [4]. On the other hand, morphological characters may often be undergoing convergent evolution as they are under similar selective pressure [5]. The freshwater shrimp genus *Caridina* is a challenge to taxonomists of all hues because of the confusion between intra- and interspecific variation of characters [6]. The phylogenetic relationships of the genus *Macrobrachium* remain comparatively poorly understood [7, 8]. Juvenile *Macrobrachium* can be notoriously difficult to identify accurately using traditional morphological methods [9] because many of the commonly used taxonomic characters are conserved between species in juveniles, and because most morphological identification requires adult males **[10].**

1.3 Importance of DNA barcoding
In the freshwater environment, a number of prawns has been described morphologically but it has lot of deceptive due to cryptic morphological characters such as dimorphism, larval and adult variation etc. DNA barcoding comes from the rapid and cost-efficient acquisition of molecular data, enabling large-scale species identification [11], whereas conventional taxonomy

is time consuming, and in some cases it is almost impossible to apply [12]. DNA-based species identification offers enormous potential benefits for the biological scientific community. It will help open the treasury of biological knowledge and increase community interest in conservation biology and understanding of evolution [14]. DNA barcoding becomes necessary when morphological traits do not adequately discriminate species [15] or if species have polymorphic life cycles and/or exhibit pronounced phenotypic plasticity [16]. As far as prawns are concerned, the DNA barcoding is a useful tool in species identification and plays an imperative role for assessing non-described and cryptic species [17, 18].

The cytochrome c-oxidase sub-unit I (COI) gene has several claims to be a suitable DNA barcode marker, including ease in amplification across a wide variety of organisms and provision of enough information to enable organisms to be identified to the species level. But it also has its drawbacks, including inherent risks due to presence of the maternal inheritance (noticeably failure in detecting hybridization), the presence of pseudogenes (NUMTS), and its inconsistent evolutionary rate among lineages [10]. These disadvantages continue to disappoint biologists hoping to rely on single gene as the sole marker for taxonomic identification [19]. However, in recent days many studies have been assessed with different COI markers, LCO140 & HCO2198, COIa & COIf, CrustF1 & HCO2198, and CrustF2 & HCO2198[20, 21], but still there is no specific primers for freshwater prawns. Therefore, this study was conducted to identify the most suited (the best possible) primer set for freshwater prawns among these primer sets. Actually, this study was accomplished with both classical and molecular taxonomy of few species of freshwater prawns, *Macrobrachium rosenbergii, M. malcolmsonii, M. lamarrei, M. lamarrei lamarroids* and *Caridina gracilipes* through mt-COI gene.

DNA sequences are characterized by summary statistics like length and base composition and when the nucleotide sequences are compared to one another prior to phylogenetic analysis, additional parameters like overall rate of nucleotide substitution, ratio of two specific instantaneous substitution rate at which transitions and transversions occurred and the rate of variation among sites play a significant role and are necessary for accurate reconstruction of phylogeny [22]. Therefore, in this study, the phylogenetic information, such as synonymous and non-synonymous substitutions, saturations, transitional and transvertional type substitutions, divergence rates and phylogenetic tree topology were all revealed.

2. Materials and Methods

2.1 Sample collection and Species Identification

Five different species of freshwater prawns was collected from five different places depending up on their availability during February to June 2014 (Plate 1). They are *M. rosenbergii* (culture pond, Marathakkara, Kerala, India), *M. malcolmsonii* (Cauvery river, Anaikarai, Kumbakonam, Tamil Nadu, India), *M. lamarrei* (Cauvery river, Manachanallur, Tiruchirappalli, Tamil Nadu, India), *M. lamarrei lamarroids* (lake, Sulur, Coimbatore, Tamil Nadu, India) and *C. gracilipes* (Aliyar river, Anaimalai, Pollachi, Tamil Nadu, India). The collected prawn species were identified based on morphological characters, included overlapping of the second segment over first and third segments, rostral structure, rostral teeth, periopods and telson by Mr. M. Kathirvel, Former Principal Scientist, Central Institute of Brackish water Aquaculture, ICAR, Chennai. These characters are presented in Table-1, Plate 1; Figs. 1-5b.

Table 1: Morphological characteristics of different freshwater prawn species

Species	Common name	Length (cm)	Rostral teeth	Periopods	Telson
Macrobrachium rosenbergii	Giant river prawn	7.5 cm	9/13	The carpus longer than the merus.	Telson extends up to the end of the uropods.
Macrobrachium malcolmsonii	Monsoon river prawn	5.8 cm	7-11/2-3	carpus slightly, but distinctly longer than merus.	Telson over reaching posterolateral spines.
Macrobrachium lamarrei	---	4.3 cm	7-9/5-8	Five pairs of periopods, or true legs present in cephalothorax region.	---
Macrobrachium lamarrei lamarroids	---	6.0 cm	4-6/2	Carpus two times longer than the chela.	Telson end with two spines
Caridina gracilipes	---	2.0 cm	24/11	Second chelipeds slender and extending up to 2nd segment of antennular peduncle.	Telson is posteriorly with a triangular median point

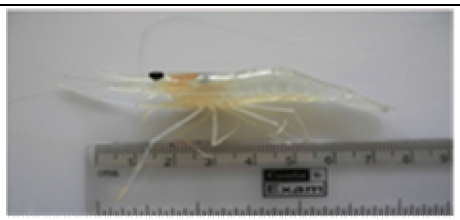


Fig 1:

Fig 1a:

Fig 1b:

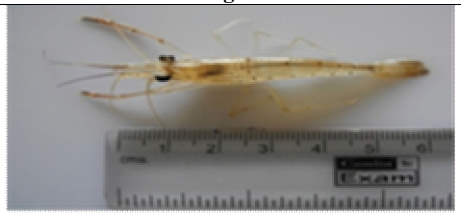


Fig 2:

Fig 2a:

Fig 2b:

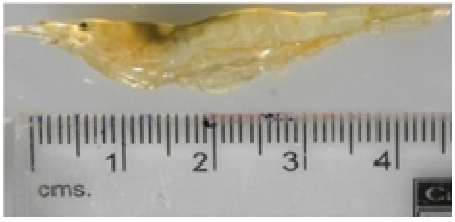


Fig 3:

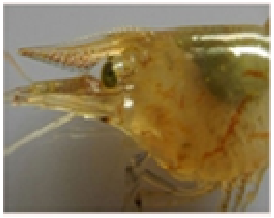


Fig 3a:

Fig 3b:

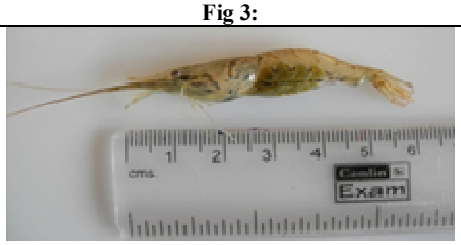


Fig 4:

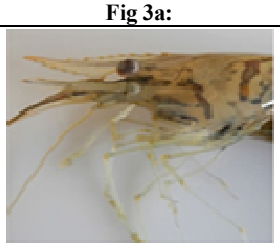


Fig 4a:

Fig 4b:

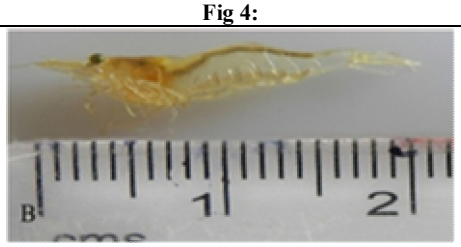


Fig 5:

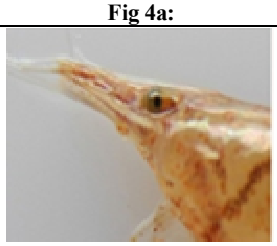


Fig 5a:

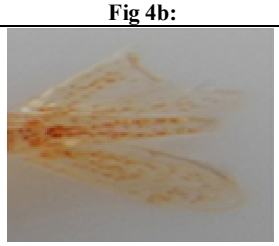


Fig 5b:

Plate 1: Morphological examination of freshwater prawn species: **Fig 1:** *Macrobrachium rosenbergii*, its cephalothorax (**Fig 1a**), and its uropod and telson (**Fig 1b**); **Fig 2:** *Macrobrachium malcolmsonii*, its cephalothorax (**Fig 1a**), and its uropod and telson (**Fig 2b**); **Fig 3:** *Macrobrachium lamarrei*, its cephalothorax **(Fig 3a)**, and its uropod and telson (**Fig 3b**); **Fig 4:** *Macrobrachium lamarrei lamarroids*, its cephalothorax (**Fig 4a**), and its uropod and telson (**Fig 4b**); **Fig 5:** *Caridina gracilipes*, its cephalothorax (**Fig 5a**), and its uropod and telson (**Fig 5b**).

2.2 Molecular identification

Genomic DNA was isolated from the abdominal muscle tissues by using Qiagen Dneasy Blood and Tissue Kit (Germany). 1% Agarose Gel Electrophoresis (GENEI, Bangalore, India) was performed and the genomic DNA was detected in a Gel documentation system (Medicare, India). DNA amplification of mt-COI gene was carried out in Eppendorf Thermo Cycler with four set of primers given in Table 2. Amplification was performed in a total volume of 50 µl containing 4 µl of DNA template, 20 p.mol of each primer, 400 µM of dNTP, and 0.4 µl of Taq DNA polymerase (Qiagen). The thermo cycler condition was as follows: 5 min at 95°C for pre-running, 35 cycles of 60 s each at 95°C for denaturation, 60 s at 49-52°C for annealing, and 90 s at 72°C for extension, followed by 5 min at 72°C for a final extension. The amplified product was resolved with 2% AGE (GENEI, Bangalore, India). Sequencing was done by using ABI 3500 XL Genetic Analyzer with manufacturer's protocol of Xcelris Genomics Ltd., Ahmadabad, India.

Table 2: Different primer sets for COI gene and their details

Primer Set (P)		Gene and primer name	Primer sequences	GC content (%)	Reference
P1	F	LCO1490	5'-GGTCAACAAATCATAAAGATATTG- 3'	29.2	Costa *et al*. (2007)
	R	HCO2198	5'TAAACTTCAGGGTGACCAAAAAATCA- 3'	34.6	
P2	F	COIa	5'-AGTATAAGCGTCTGGGTAGTC-3'	47.6	Rossi and Mantelatto, (2013)
	R	COIf	5'-CCT GCA GGAGGAGGA GAC CC-3'	70.0	
P3	F	CrustF1	5'-TTTTCTACAAATCATAAAGACATTGG-3'	26.9	Costa *et al*. (2007)
	R	HCO2198	5'TAAACTTCAGGGTGACCAAAAAATCA- 3'	34.6	
P4	F	CrustF2	5'GGTTCTTCTCCACCAACCACAARGAYATHGG-3'	49.5	
	R	HCO2198	5'TAAACTTCAGGGTGACCAAAAAATCA- 3'	34.6	

P1-P4, Primer sets; F, Forward; R, Reverse

2.3 Sequence statistical analysis

The sequences were aligned pair wise by using CAP3. The similarity between sequences was identified by BLAST. The internal stop codon was removed by using BLAST. The reading frame shift was deducted by ORF finder. The trimmed sequence was authenticated with GenBank. The multiple sequence alignment was done by using T-Coffee and the aligned sequence was highlighted (identical, similar and variable sites of amino acids) with multiple align show (MAS). Estimation of synonymous (Ks) and non-synonymous (Ka) substitutions was calculated by Li93 method of DAMBE for 3rd codon position. The maximum likelihood (ML) analysis for the synonymous and non synonymous substitutions was produced by joint reconstructions of ancestral states by Muse-Gaut model of codon substitution [23] and Felsenstein model of nucleotide substitution [24]. The probability of rejecting the null hypothesis of neutral evolution (P-value) was calculated [25]. Analysis of sequence saturation was done by using DAMBE

5.3.10[26] for calculating the transitional and transvertional substitutions against genetic distance (TN93). The substantial saturation of the sequence was checked by using the method of Xia *et al*., [27]; Xia & Lemey, [28] (DAMBE). Finally, the nucleotide divergence was calculated by Kimura-2-Parameter model (K2P) by using the software, MEGA v 6.01 and the phylogenetic relationship was analyzed [27].

The evolutionary relationship of *Macrobrachium* (four species studied in this study and ten retrieved species) and *Caridina* (one species studied in this study and four retrieved species), totally nineteen nucleotide sequences of freshwater prawns was inferred by the Maximum Likelihood (ML) method of Kimura 2-parameter model. The initial tree(s) for the heuristic search was obtained automatically by applying Neighbor-Joining (NJ) and BioNJ algorithms to a matrix of pair wise distance estimated using the Maximum Composite Likelihood (MCL) approach, the tree topology was selected with superior log likelihood value and the 3rd codon position was opted. Sequence position with more than 95% site coverage alone was allowed to read with 5% alignment gaps, missing data and ambiguous bases at any position. The likelihood-ratio test [29] and various model selection criteria such as AIC [30] and BIC [31] were applied to select appropriate substitution models.

3. Results and Discussion

3.1 Morphological characteristics

In this study five freshwater prawn species (of which four species, such as *M. rosenbergii*, *M. malcolmsonii*, *M. lamarrei* and *M. lamarrei lamarroids* belong to the family, Palaemonidae, and one species, *C. gracilipes* belongs to the family, Atyidae) were morphologically identified based on their rostral structure, rostral teeth, periopods and telson. *M. rosenbergii* (sub adult of about 7.5 cm length) was characterized by up-curved rostrum with 9 upper teeth and 13 lower teeth. The carpus of the periopods was longer than the merus, and the telson extends up to the end of the uropods. *M. malcolmsonii* (sub adult of about 5.7 cm length) rostrum was straight in comparison with *M. rosenbergii*, rostrum was with 7-11 upper teeth and 2-3 lower teeth. The carpus was slightly, but distinctly longer than merus in periopods, and the telson was over reaching the posterior lateral spines. *M. lamarrei* (adult of about 4.3 cm length) rostrum was straight with 7-9 upper teeth and 5-8 lower teeth. *M. lamarrei lamarrroids* (adult of about 6.0 cm length), rostrum was straight with 4-6 upper teeth and 2 lower teeth, the carpus was two times longer than the chela, and the telson end with two spines. *C. gracilipes* (adult of about 2.0cm length) rostrum was straight with 24 upper teeth and 11 lower teeth. It's 2nd chelipeds were slender and extending up to 2nd segment of antennular peduncle. Telson was posteriorly with a triangular median point (Table 1, Plate 1; Fig. 1 to Fig. 5b). Generally, *Macrobrachium* spp., were with smooth rounded dorsal body surface, while penaeids usually have a simple or complex ridge at the dorsal apex of the abdomen [32]. The rostral formula was varies from species to species and also in the same species during different life stages, like post larvae, juveniles and adults. Even though these species have exhibited some morphological differences, they were validated by molecular characteristics of mt-COI gene.

3.2 Molecular characteristics

The isolated genomic DNA showed greater than 10 kb nucleotides (Plate 2; Fig. 1). The primer sets, P1 (LCO1490 & HCO2198) and P2 (COIa and COIf) worked well in all species, and the remaining primer sets, P3 and P4 (Crust F1 & HCO2198 and Crust F1 & HCO2198 respectively) were not successfully amplified in all freshwater prawn species (Plate 2; Fig. 2 to Fig. 6). Regarding the primer sets, all four combinations were worked well with *M. rosenbergii, M. malcolmsonii* and *C. gracilipes*. However, the sequence length of *M. malcolmsonii* (P3 & P4) was low. So these sequences were not included in this study. In the cases of *M. lamarrei* and *M. lamarrei lamarroids* only two sets of primers (P1 & P2) were worked well. The sequences generated in this study were submitted to BLAST and the results were shown in Table-3. The sequences generated in this study were authenticated by GenBank and the accession numbers are given in Table-4.

Table 3: BLAST identification of COI gene sequences of freshwater prawns generated with P1 and P2 for similarity with different *Macrobrachium* and *Caridina* species available in the NCBI database

Primer set	Species	Bits (score)	Expected	Identity (%)	Gap (%)	Strand	Matched Species	Matched Country
P1	*Macrobrachium rosenbergii*	531 (287)	5e-147	83	14/633 (2%)	Plus/ Plus	*Macrobrachium rosenbergii*	China
P2	*Macrobrachium rosenbergii*	1107 (599)	0.0	99	0/605 (0%)	Plus/ Plus	*Macrobrachium rosenbergii*	Taiwan
P1	*Macrobrachium malcolmsonii*	182 (98)	6e-42	87	0/164 (0%)	Plus/ Plus	*Macrobrachium villosimanus*	India
P2	*Macrobrachium malcolmsonii*	704 (381)	0.0	86	3/644 (0%)	Plus/ Minus	*Macrobrachium rosenbergii*	Australia
P1	*Macrobrachium lamarrei*	436 (236)	9e-119	89	421/509 (83%)	Plus/ Plus	*Caridina gracilipes*	India
P2	*Macrobrachium lamarrei*	604 (327)	2e-169	94	5/408 (1%)	Plus/ Minus	*Caridina gracilirostris*	Australia
P1	*Macrobrachium lamarrei lamarroids*	521 (282)	3e-144	90	4/596 (0%)	Plus/ Plus	*Macrobrachium asperulum*	Taiwan
P2	*Macrobrachium lamarrei lamarroids*	741 (401)	0.0	90	2/582 (0%)	Plus/ Plus	*Macrobrachium canarae*	India
P1	*Caridina gracilipes*	279 (151)	2e-71	75	27/651 (4%)	Plus/ Plus	*Macrobrachium lanchesteri*	Thailand
P2	*Caridina gracilipes*	662 (358)	0.0	89	4/540 (0%)	Plus/ Minus	*Caridina peninsularis*	Australia

P1, LCO1490 & HCO2198; P2, COIa & COIf.

Table 4: Details of COI gene sequence accession numbers with GenBank for the sequences generated with P1, P2, P3 and P4 in different freshwater prawns and retrieved species from the NCBI database

Species Name	Country	Author's name	Year	Accession numbers
Macrobrachium rosenbergii	India	Paper's Authors	2014	KJ652338
Macrobrachium rosenbergii	India	Paper's Authors	2014	KJ708773
Macrobrachium rosenbergii	India	Paper's Authors	2014	KJ708774
Macrobrachium rosenbergii	India	Paper's Authors	2014	KJ708775
Macrobrachium malcolmsonii	India	Paper's Authors	2014	KJ708779
Macrobrachium malcolmsonii	India	Paper's Authors	2014	KJ708780
Macrobrachium lamarrei	India	Paper's Authors	2014	KJ708781
Macrobrachium lamarrei	India	Paper's Authors	2014	KJ708782
Macrobrachium lamarrei lamarroids	India	Paper's Authors	2014	KJ708783
Macrobrachium lamarrei lamarroids	India	Paper's Authors	2014	KJ708784
Caridina gracilipes	India	Paper's Authors	2014	KJ652339
Caridina gracilipes	India	Paper's Authors	2014	KJ708776
Caridina gracilipes	India	Paper's Authors	2014	KJ708777
Caridina gracilipes	India	Paper's Authors	2014	KJ708778
Macrobrachium australiense	Australia	Sharma and Hughes	2008	EU787489
Macrobrachium asperulum	China	Liu and Tzeng	2006	AB250551
Macrobrachium idella	India	Jayaraj *et al.,*	2011	JF774069
Macrobrachium idea	France	Zimmermann *et al.,*	2009	GU205063
Macrobrachium equidens	India	Saravana bhavan *et al.,*	2013	KF882423
Macrobrachium olfersii	Panama	Rossi and Mantelatto	2012	JQ805960
Macrobrachium crenalutum	Costa rica	Rossi and Mantelatto	2012	JQ805902
Macrobrachium tolmerum	Australia	Page *et al.,*	2011	JF487978
Macrobrachium nipponense	China	Fang *et al.,*	2009	JN874545
Macrobrachium lar	Taiwan	Liu *et al.,*	2005	AB235270
Caridina nilotica	Viet Nam	Page *et al.*,	2006	DQ478473
Caridina sumatrensis	Indonesia	Page *et al.*,	2006	DQ478472
Caridina multidentata	Taiwan	Page *et al.*,	2006	DQ478457
Caridina indistinct	Australia	Page *et al.*,	2011	JF487977
Caridina rubella	Japan	Weese *et al.*,	2011	JF926149
Caridina lanceolata	Indonesia	Roy *et al.*,	2005	DQ155590
Penaeus monodon	Thailand	Khamnamtong *et al.*,	2007	EF646261

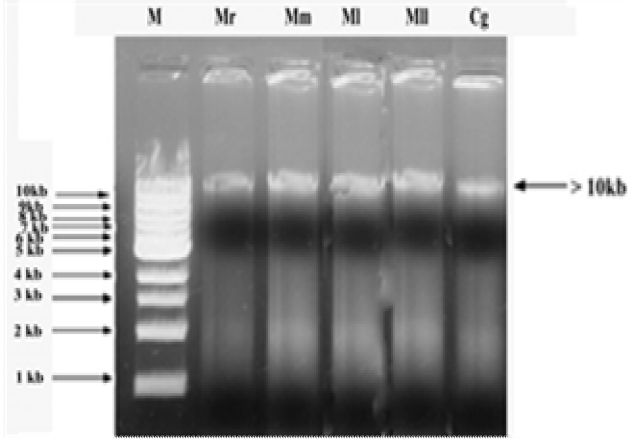


Fig 1: 1% AGE shows greater than 10kb of genomic DNA from freshwater prawns.
M, Marker (1kb); *Mr, M. rosenbergii*, Mm, *M. malcolmsonii*; Ml, *M. lamarrei*; Mll, *M. lamarrei lamerroids;* Cg, *C. gracilipes*

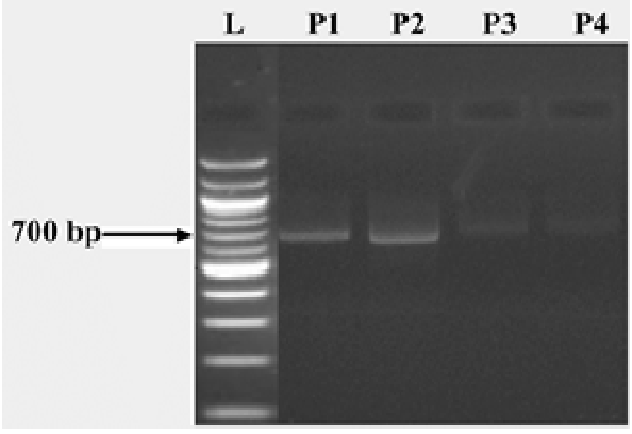


Fig 2: 2% AGE shows PCR amplified product COI gene from *Macrobrachium rosenbergii.*
L, Ladder (100bp); P1, LCO1490 & HCO2198; P2, COIa & COIf; P3, CrustF1 & HCO2198; P4, CrustF2 & HCO2198.

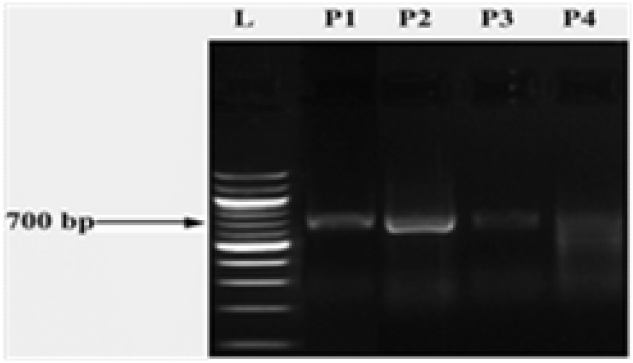


Fig 3: 2% AGE shows PCR amplified product COI gene from *Macrobrachium malcolmsonii.*
L, Ladder (100bp); P1, LCO1490 & HCO2198; P2, COIa & COIf; P3, CrustF1 & HCO2198; P4, CrustF2 & HCO2198.

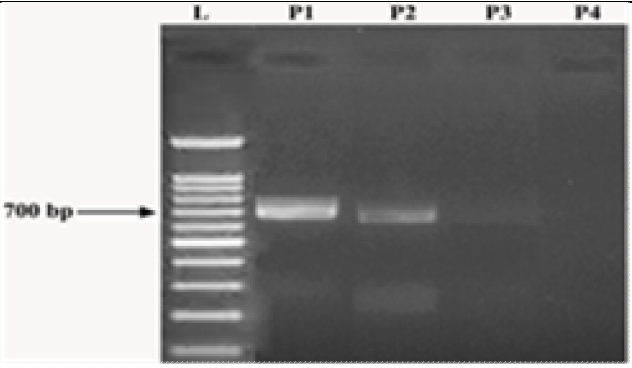


Fig 4: 2% AGE shows PCR amplified product COI gene from *Macrobrachium lamarrei.*
L, Ladder (100bp); P1, LCO1490 & HCO2198; P2, COIa & COIf; P3, CrustF1 & HCO2198; P4, CrustF2 & HCO2198.

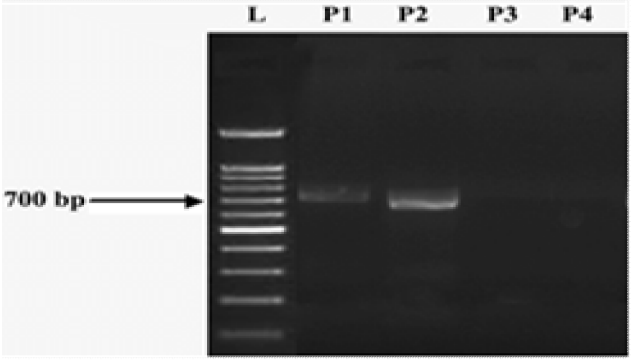


Fig 5: 2% AGE shows PCR amplified product COI gene from *Macrobrachium lamarrei lamarroids.*
L, Ladder (100bp); P1, LCO1490 & HCO2198; P2, COIa & COIf; P3, CrustF1 & HCO2198; P4, CrustF2 & HCO2198.

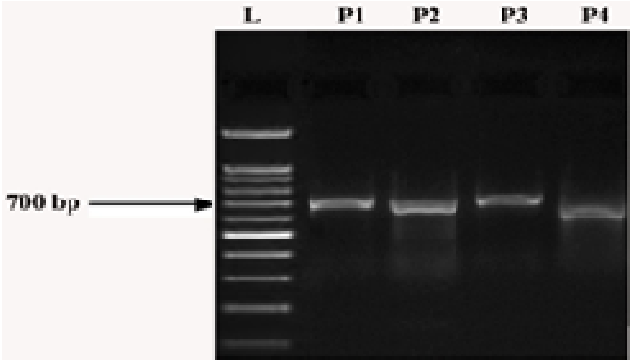


Fig 6: 2% AGE shows PCR amplified product COI gene from *Caridina gracilipes.*
L, Ladder (100bp); P1, LCO1490 & HCO2198; P2, COIa & COIf; P3, CrustF1 & HCO2198; P4, CrustF2 & HCO2198.

Plate 2: Qualitative analysis of genomic DNA and PCR amplified products of COI gene of different species of freshwater prawns.

3.3 General properties of sequences

The analyzed set of sequences with P1 showed 148 identical amino acid residues, 60 similar amino acid residues and 495 variable sites, and P2 showed 251 identical amino acid residues, 70 similar amino acid residues and 339 variable sites (Table 5, Plate 3; Fig. 1 and 2).

Table 5: Number of identical, similar and variable sites of amino acid residues in the COI gene sequences generated with P1 and P2 in different freshwater prawn species

Primers set	Primer Name	Number identical amino acid residues	Number of similar amino acid residues	Variable sites
P1	LCO1490 & HCO2198	148	60	495
P2	COIa & COIf	251	70	339

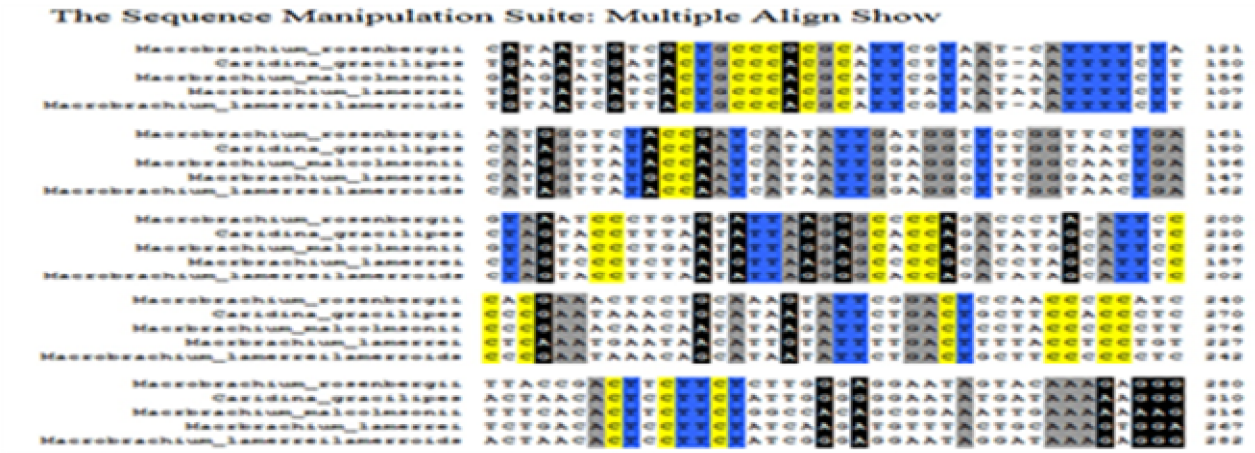


Fig 1: Multiple sequence alignment of COI gene sequences generated in different freshwater prawn species with P1 (LCO1490 & HCO2198). An alignment is formatted by using multiple align show (MAS) with coloured background and a consensus setting of 100%. Identical residues are indicated by amino acid colour and similar residues are black in colour. Gaps and other residues are given in white background.

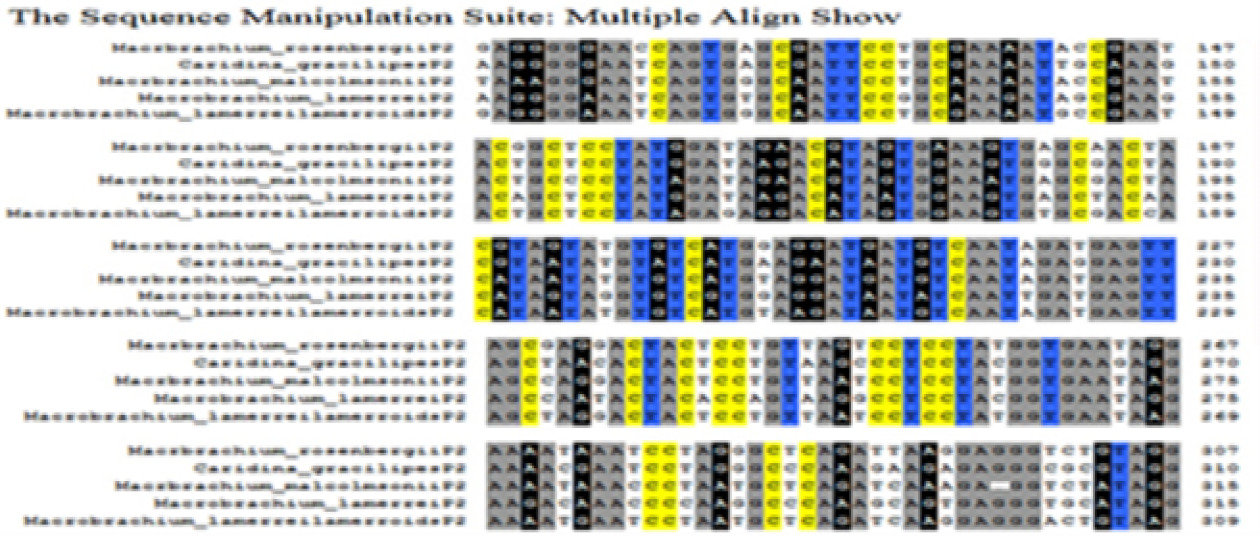


Fig 2: Multiple sequence alignment of COI gene sequences generated in different freshwater prawn species with P2 (COIa & COIf). An alignment is formatted by using multiple align show (MAS) with coloured background and a consensus setting of 100%. Identical residues are indicated by amino acid colour and similar residues are black in colour. Gaps and other residues are given in white background.

Plate 3: Multiple sequence alignment.

Generally, like other protein coding genes in the COI gene also, majority of the substitutions occurred in the 3rd codon position of the sequences generated. In the sequence generated with P2, the transitional and transvertional type substitutions was higher in A and T nucleotides in 3rd codon position followed by 2nd and 1st codon positions (62.09%, 59.77% and 57.38% respectively). In the case of P1, these substitutional rates were 58.3%, 58.0% and 57.0% for 3rd, 2nd and 1st codon position respectively. In this study, the base composition of the COI gene fragment varied among the species, but it was commonly demonstrated with an overall AT bias of 57.72% and GC of 41.63% for P1, and 60.20% and 39.78% for AT and GC biases respectively for P2 (Table 6).

Table 6: DNA sequence of COI gene generated in different species of freshwater prawns with P1 and P2

1	Information	Mr	Mm	Ml	Mll	Cg
P1	Length	655	598	492	594	649
	Weight (single-stranded)	201.291	185.867	151.538	182.774	199.704
	Weight (double-stranded)	404.71	369.505	304.001	367.012	400.991
	Adenine (A)	176	167	120	172	188
	Cytosine (C)	164	106	110	139	152
	Guanine (G)	116	161	97	104	114
	Thymine (T)	199	164	165	179	195
	C + G	280	267	207	243	266
	A + T	375	331	285	351	383
P2	Length	608	262	408	626	540
	Weight (single-stranded)	189.214	80.707	126.773	194.188	167.181
	Weight (double-stranded)	375.678	161.894	252.096	386.764	333.641
	Adenine (A)	177	81	143	208	189
	Cytosine (C)	100	56	68	102	102
	Guanine (G)	166	44	95	136	109
	Thymine (T)	165	81	102	180	140
	C + G	266	100	163	238	211
	A + T	342	162	245	388	329

P1, LCO1490 & HCO2198; P2, COIa & COIf.
Mr, Macrobrachium rosenbergii; Mm, *Macrobrachium malcolmsonii*; Ml, *Macrobrachium lamarrei*; Mll, *Macrobrachium lamarrei lamarroids;* Cg, *Caridina gracilipes*

3.4 Synonymous and non-synonymous substitutions

Assessment of the synonymous (Ks) and non-synonymous substitutions (Ka) is an essential component for understanding whether evolutionary changes occur in a particular DNA sequence. In this study, the overall Ks rate was higher than Ka for the sequence generated with P1 (0.957 and 0.460 respectively). Whereas, in the sequences generated with P2, the Ks rate was lower than the Ka (0.470 and 0.488 respectively). Moreover, the overall Ka rate was higher in the sequence generated with P2 than that of P1. It indicates the fact that the sequence generated with P2 has more evolutionary significance than P1 (Table 7).

Table 7: Synonymous (Ks) and Non-synonymous (Ka) substitutions, and divergence value of the COI gene sequences between different species of freshwater prawns

Between species	Synonymous (Ks)		Non-Synonymous (Ka)		Divergence (%)	
	P1	P2	P1	P2	P1	P2
Mr-Mm	1.622	1.276	0.889	0.805	1.353	NA
Mr-Ml	0.613	0.110	0.397	0.328	1.080	0.598
Mr-Mll	0.690	0	0.321	0.226	0.500	11.340
Mr-Cg	0.690	0.042	0.336	0.369	0.500*	NA*
Mm-Ml	1.071	0.276	0.784	0.805	1.015	0.894
Mm-Mll	2.207	1.245	0.762	0.711	1.917	2.328
Mm-Cg	2.207	0.908	0.762	0.733	1.917*	2.328*
Ml-Mll	0.464	0.115	0.172	0.275	2.039	1.115
Ml-Cg	0.005	0.690	0.172	0.336	2.039*	3.556*
Mll-Cg	0	0.035	0	0.298	2.697*	3.748*
Over all	0.957	0.470	0.460	0.488	1.505	2.878
	--	--	--	--	1.788*	3.210*

P1, LCO1490 & HCO2198; P2, COIa & COIf.
Mr, Macrobrachium rosenbergii; Mm, *Macrobrachium malcolmsonii*; Ml, *Macrobrachium lamarrei*; Mll, *Macrobrachium lamarrei lamerroids;* Cg, *Caridina gracilipes*
* Divergence rate of *Macrobrachium* species against *Caridina gracilipes*

The synonymous and non-synonymous substitutions present for each nucleotide at specific site, that is inferred synonymous (dS) and inferred non-synonymous (dN) substitutions along the length of nucleotide was calculated. The number of appearance of positive value for inferred non-synonymous substitution was found to be significantly higher ($P<0.05$). Therefore, the sequence generated with P2 explains more biological changes (natural selection) than that of the sequence generated with P1 (Plate 4; Fig. 1 and 2).

DNA substitutions of protein-coding sequences are categorized into synonymous and non-synonymous codon substitutions. Because non-synonymous substitutions induce amino acid replacement, they may have been subjected to Darwinian selection during the evolutionary process. The relative occurrence of non-synonymous substitutions with respect to synonymous substitutions is used to represent the strength of selective pressure [21, 33].

3.5 Saturation

It is generally stated that if sequences exhibit substitutional saturation, they possess less phylogenetic information and therefore could not be able to reveal deeper insights for phylogenetic signal. In this study, the test of substantial saturation revealed that the value of Iss (Index of substantial saturation) is lower than the Iss.c (Critical value of Index of substantial saturation) (Iss: P1 = 0.543 and P2 = 0.580; Iss.c: P1 = 0.767 and P2 = 0.832. Hence it is strongly evident that the sequences generated with P1 and P2 did not undergo substantial saturation because Iss.c value is greater than Iss value. In a study of phylogenetic analysis of *Macrobrachium* species, the reported Iss value was 0.228, and the Iss.c value was 0.714, and thus, there was little substitutional saturation [34].

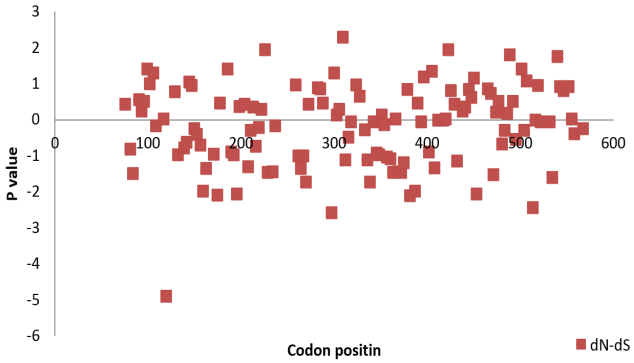


Fig 1: Number of inferred synonymous (dS) and non-synonymous (dN) substitutions of nucleotides at 3rd codon position for COI gene sequences generated in different species of freshwater prawns with P1(LCO1490 & HCO2198).

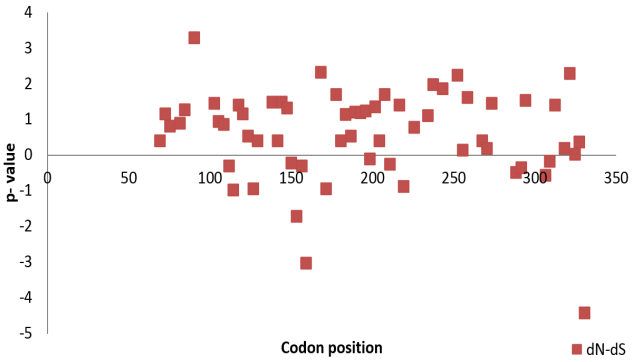


Fig 2: Number of inferred synonymous (dS) and non-synonymous (dN) substitutions of nucleotides at 3rd codon position for COI gene sequences generated in different species of freshwater prawns with P2 (COIa & COIf).

Plate 4: Inferred synonymous (dS) and non-synonymous (dN) substitutions of nucleotides.

3.6 Transitional and transvertional type substitutions

The transitional (Ts) and transvertional (Tv) type substitutions occurred at 3rd codon positions of the COI gene sequences generated with P1 and P2 were plotted in the scattergram and the results are presented in Plate 5; Figs. 1 & 2.

In this study, the Tv was higher than the Ts in sequences generated with both P1 and P2 (Ts: P1 = 0.30 and P2 = 0.30; Tv: P1 = 0.39 and P2 = 0.42). This indicates the fact that Tv contain stronger phylogenetic signal than Ts. These substitutions would perhaps lead to less saturation of the sequences.

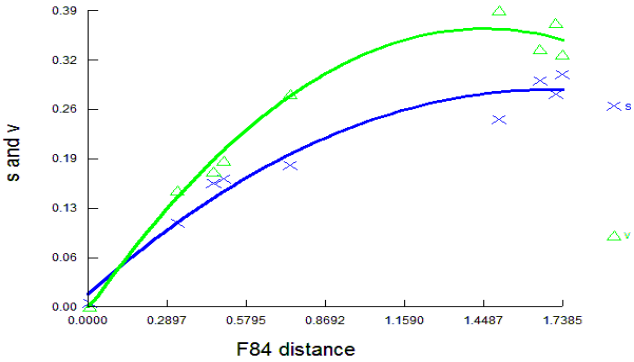


Fig 1: Scattergram shows transitional and tranversional type substitutions occurred in COI gene sequences generated in different species of freshwater prawns with P1 (LCO1490 & HCO2198).

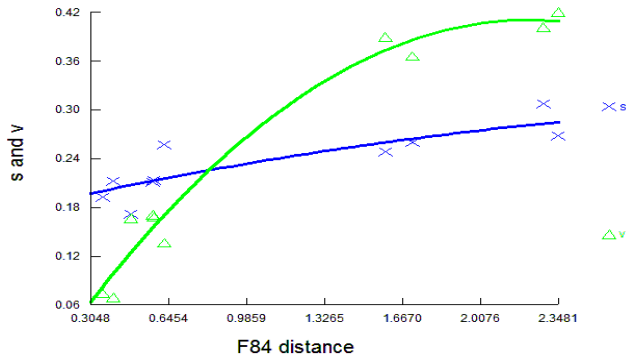


Fig 2: Scattergram shows transitional and tranversional substitutions occurred in COI gene sequences generated in different species of freshwater prawns with P2 (COIa & COIf).

Plate 5: Transitional and tranvertional type substitutions of nucleotides.

3.7 Inter and intra species divergence

The divergence rate of freshwater prawn species for the sequences generated with P1 was ranged from 0.500-2.697 with the highest divergence of 2.697 between *M. lamarrei lamarroids* and *C. gracilipes*, and the lowest divergence of 0.500 between *M. rosenbergii* and *M. lamarrei lamarroids*. The average divergence of *Macrobrachium* species against *C. gracilipes* was 1.788 with the highest divergence of 2.697 between *M. lamarrei lamarroids* and *C. gracilipes*, the lowest divergence of 0.500 between *M. rosenbergii* and *C. gracilipes*, *M. lamarrei* formed divergence value of 2.039 with *C. gracilipes* and the divergence of 1.917 was calculated between *M. malcolmsonii* and *C. gracilipes* (Table 7).

In the cases with P2, the divergence rate of freshwater prawn species was ranged between 0.598-11.344 with the highest divergence value of 11.344 between *M. rosenbergii* and *M. lamarrei lamerroids,* and the lowest divergence value of 0.598 between *M. rosenbergii* and *M. lamarrei*. The average divergence of *Macrobrachium* species against *C. gracilipes* was 3.210 with the highest divergence of 3.748 between *M. lamarrei lamarroids* and *C. gracilipes*, the lowest divergence of 2.328 between *M. malcolmsonii* and *C. gracilipes*, *M. lamarrei* formed divergence value of 3.556 with *C. gracilipes* (Table 7).

The divergence rate of sequences generated with P1 for freshwater prawn species were calculated against one out group, *Penaeus monodon* retrieved from NCBI was ranged between 0.666-3.062 (minimum with *M. lamarrei,* and maximum with *M. rosenbergii*). In the cases with P2 the divergence value was ranged between 0.646-24.731(minimum with *M. malcolmsonii* and maximum with *M. lamarrei lamarroids*) (Table 8).

The overall divergence value between the species of freshwater prawns, with *Caridina*, and with the marine prawn, *P. monodon* was higher in the sequences generated with P2 (2.878, 3.210 and 10.696) when compared with the sequences generated with P1 (1.505, 1.788 and 1.868) (Tables 7 and 8). Therefore, sequences generated with P2 more delineates the *Macrobrachium* species, *Macrobrachium* species from *C. gracilipes,* and *Macrobrachium* species from *P. monodon* than that of the sequence generated with P1. Therefore it is suggested that the P2 may be used to discriminate the cryptic species as well.

3.7.1 Intra species divergence

There is no intra-species divergence data available with NCBI database for freshwater prawn species except *M. rosenbergii*. Therefore, the sequence generated for *M. rosenbergii* in this study was compared with available information in database for the same species and the divergence rate was calculated. The divergence rate of sequence generated with P1 was ranged between 1.387-1.611 (minimum with *M. rosenbergii* (Brazil) and maximum with *M. rosenbergii* (India). The divergence rate of sequence generated with P2 was ranged between 0.777-1.342 (minimum with *M. rosenbergii* (Brazil) and maximum with *M. rosenbergii* (India) (Table 8).

It has been reported that the inter-specific divergences of different *Macrobrachium* spp., (*M. acanthurus, M. amazonicum, M. brasiliense, M. carcinus, M. iheringi, M. jelskii, M. olfersi, M. potiuna, and M. grandimanus*) was ranged from 5.5 to 17.5% for 16S r-RNA and 15.1 to 25.5% for COI gene; the intra-specific divergence of these *Macrobrachium* spp., was ranged from 0 to 3.2% for 16S r-RNA and 0 to 12.6% (different population of *M. grandimanus*) for COI gene[8, 35] .

Table 8: Inter species divergence of different freshwater prawns with the marine prawn, *Penaeus monodon* against the sequences generated withP1 and P2, and intra species divergence of *Macrobrachium rosenbergii* available in the NCBI data base against the sequences generated in the present study

Species		Divergence (%)	
		P1	**P2**
Inter species	*Macrobrachium rosenbergii*	3.062	NA
	Macrobrachium malcolmsonii	1.081	0.646
	Macrobrachium lamarrei	0.666	10.244
	Macrobrachium lamarrei lamarroids	2.266	24.731
	Caridina gracilipes	2.266	17.86
	Over all	**1.8682**	**10.6962**
Intra species	*Macrobrachium rosenbergii* (Australia)	1.479	0.900
	Macrobrachium rosenbergii (China)	1.479	0.777
	Macrobrachium rosenbergii (India)	1.611	1.342
	Macrobrachium rosenbergii (Brazil)	1.387	0.777
	Over all	**1.489**	**0.949**

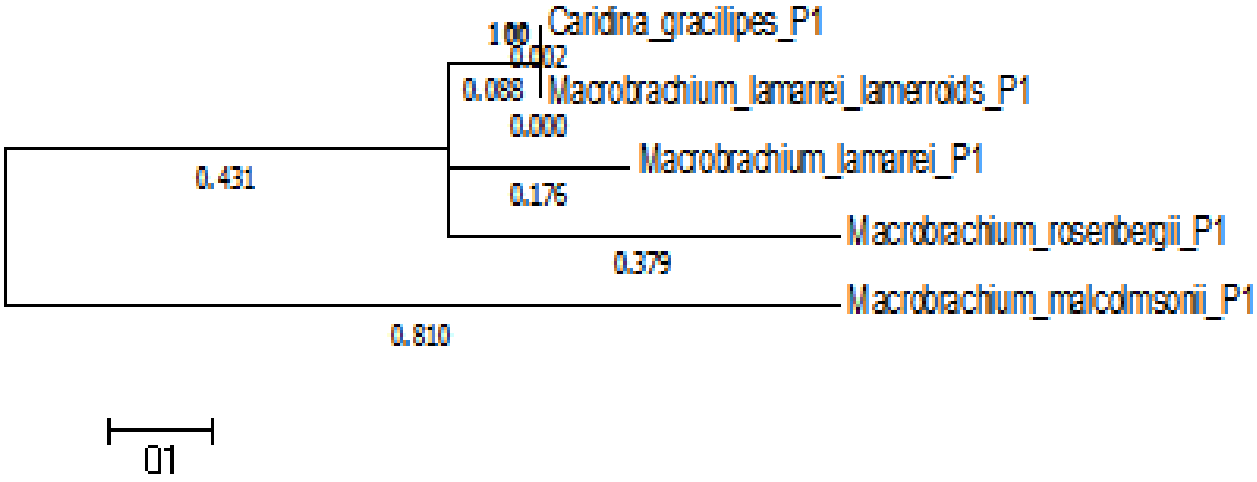
P1, LCO1490 & HCO2198; P2, COIa & COIf.

Note: No data available for other freshwater prawn species in the NCBI data base other then *Macrobrachium rosenbergii*

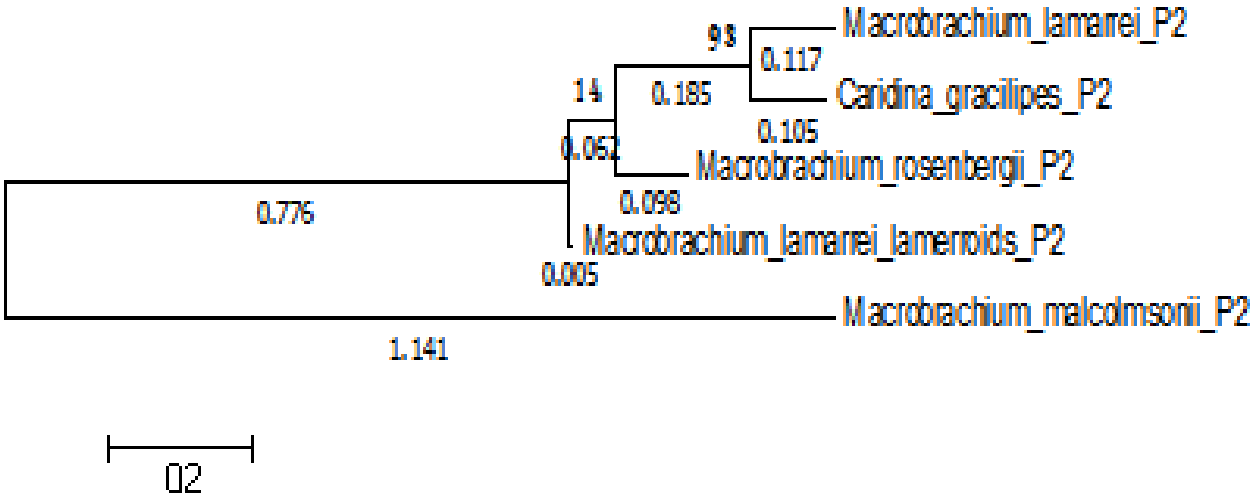
3.8 Evolutionary relationship of freshwater prawn

Based on the lowest AICc value (Akaike Information Criterion, corrected), 1411.435 and the lowest BIC score (Bayesian Information Criterion), 1448.470 for sequences generated with P1, and AICc value, 588.137and BIC score, 624.283 for sequences generated with P2 in partial COI gene the best fit model (K2P) of ML was calculated. The bootstrap length was observed to be >19 with the branch length of >0.000.

In the phylogenetic tree topology of the sequences generated with P1 for the freshwater prawn, *M. malcolmsonii* formed a single distinct clade and the remaining species (*M. rosenbergii*, *M. lamarrei*, *M. lamarrei lamarroids* and *C. gracilipes*) aligned in a separate cluster, in which *M. lamarrei lamarroids* and *C. gracilipes* formed as sister taxon with bootstrap value of 100 (Plate 6; Tree. 1). In the cases with P2 also *M. malcolmsonii* formed a single distinct clade and the remaining species grouped in a single clusters, in which *C. gracilipes* and *M. lamarrei* formed as sister taxon with bootstrap value of 93 (Plate 6; Tree. 2). However, differences were seen in the branch formation with P1, and P2.



Tree 1: Phylogenetic significance of COI gene sequences generated in different species of freshwater prawns with P1 (LCO1490 & HCO2198).



Tree 2: Phylogenetic significance of COI gene sequences generated in different species of freshwater prawns with P2 (COIa & COIf).

Plate 6: Tree topology.

In the phylogenetic tree topology (PTT) for sequences generated with P1 along with retrieved species formed two major clusters with branch length of 0.356 and 0.447. The first cluster consisted of all retrieved species *Caridina* (*Caridina multidentata*, *Caridina sumatrensis* and *Caridina indistincta*) and *Macrobrachium* (*M. olfersii, M. idella* and *M. idae*), and in which *C. multidentata* and *C. sumatrensis* formed as a sister taxon with bootstrap value of 59. In the second cluster, *M. malcolmsonii* (studied in this study) appeared as a single distinct clade (bootstrap value of 100), and the remaining four species (three *Macrobrachium* spp., *M. lamarrei, M. rosenbergii* and *M. lamarrei lamarroids,* and a *Caridina* sp., *C. gracilipes*) studied in this study along with eight retrieved species (two *Caridina* spp., *Caridina rubella* and *Caridina lanceolata*) and six *Macrobrachium* spp., (*Macrobrachium equidens*, *Macrobrachium tolmerum*, *Macrobrachium asperulum*, *Macrobrachium crenalatum, Macrobrachium nipponense* and *Macrobrachium austreliense*) were formed as a large cluster with bootstrap value of 85 (Plate 7).

In the PTT for sequences generated with P2 along with retrieved species formed two major clusters with branch length of 0.207 and 0.269, in the first cluster, *C. gracilipes* formed as an out group of *Macrobrachium* species studied in this study (*M. lamarrei, M. rosenbergii* and *M. lamarrei lamarroids*), in which *M. rosenbergii* and *M. lamarrei lamarroids* formed as a sister taxon with the bootstrap value of 50. The second cluster was divided into two sub-clusters with branch length of 0.570 and 0.348. In the first sub-cluster (all formed with retrieved species), *M. olfersii* formed as an out group of *C. multidentata*, *C. indistincta*, *C. sumatrensis*, *M. idella* and *M. idae*, in which *M. idella* and *M. idae* formed as a sister taxon with the bootstrap value of 83. Thus, *Caridina* formed as a separate clade. The second sub-cluster (all formed with retrieved species except one species, i.e., *M. malcolmsonii* studied in the present study) consisted of *M. tolmerum, C. lanceolata*, *M. asperulum*, *M. austreliense*, *M. malcolmsonii, M. equidens, M. crenalatum, M. nipponense* and *C. rubella* was formed without branch length (Plate 8).

In this study, the sequences generated with both P1 and P2 formed polyphyletic (non-monophyletic) tree topology of freshwater prawns. However, in the tree constructed for P2, the studied species of the present study mostly appeared in a single cluster and *Caridina* appeared as an out group of *Macrobrachium*, whereas, in the case of P1 the studied species are scatterly appeared all along the tree. This indicates the fact that the tree is well resolved with P2 than that of P1. Murphy & Austin [8] also suggested that many of the Palaemonid genera are non-monophyletic.

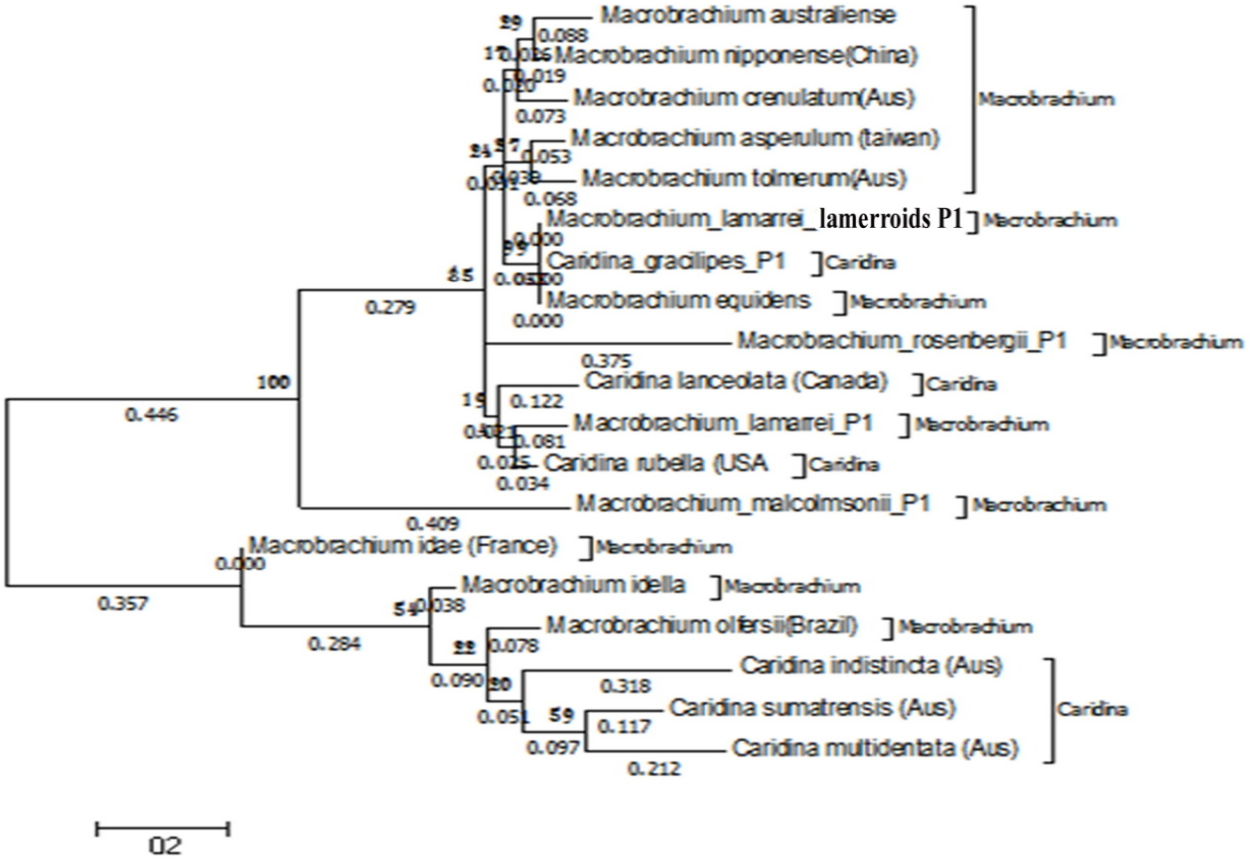


Plate 7: Phylogenetic significance of COI gene sequences generated in different species of freshwater prawns with P1 (LCO1490 & HCO2198), and retrieved species.

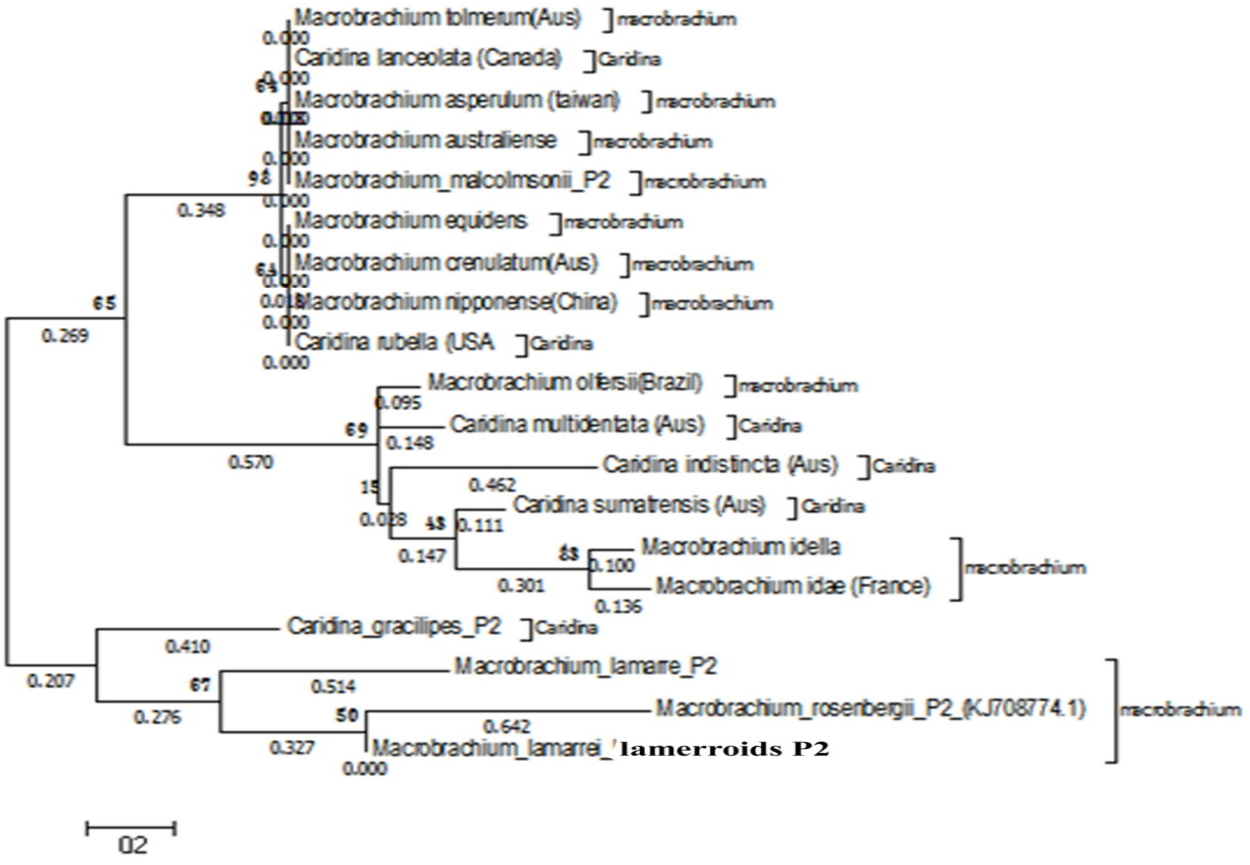


Plate 8: Phylogenetic significance of COI gene sequences generated in different species of freshwater prawns with P2 (COIa & COIf), and retrieved species.

4. Conclusion

Among the four primer sets tested in this study, two primer sets (P1, LCO1490 & HCO 2198 and P2, COIa & COIf) were worked well in all aspects. Among these two, P2 was better than P1 with more AT bias, more non-synonymous, more critical substantial saturation value, more transvertional substitution, more divergent, and well resolved polyphyletic tree. Therefore, the sequences generated with P2 were clearly delineated the studied freshwater prawns than that of P1. Thus, the primer set with COIa & COIf is more reliable than the primer set with LCO1490 & HCO 2198 as far as freshwater prawns are concerned.

5. Acknowledgements

The Science and Engineering Research Board, Department of Science and Technology, Government of India, New Delhi is gratefully acknowledged for the financial support provided in the form of research project (SB/EMEQ-291/2013, dt. 01.08.2013 of SERB, New Delhi). The authors are sincerely thanking Mr. M. Kathirvel, Former Principal Scientist, Central Institute of Brackishwater Aquaculture (ICAR), Chennai - 600028, Tamil Nadu, India, for species identification.

6. References

1. Jayachandran KV, Indira B. Sustainable exploitation of freshwater prawn diversity of India for food and livelihood security with emphasis on planning. Indian J Sci Res 2010; 1(2):127-132.
2. De Grave S, Cai Y, Anker A. Global diversity of shrimps (Crustacea: Decapoda: Caridea) in freshwater. Hydrobiologia 2008; 595:287-293.
3. Montgomery S. Biology and life cycles of prawns, Cronulla, Fisheries Research Centre. Primefact No. 268, 2010, 1-8.
4. Schwander T, Leimar O. Genes as leaders and followers in evolution. Trends Ecol Evol 2011; 26(3):143-151.
5. Yang Z, Rannala B. Bayesian species delimitation using multi locus sequence data. Proc Natl Acad Sci USA, 2010; 107(20):9264-9269.
6. Smith MJ, Williams WD. Intraspecific variation within the Atyidae: A study of morphological variation within a population of *Paratya australiensis* (Crustacea: Decapoda), Aust J Mar Freshwat Res 1980; 31(3):397-407.
7. Liu MY, Cai YX, Tzeng CS. Molecular systematic of the freshwater prawn genus *Macrobrachium* Bate, 1868 (Crustacea: Decapoda: Palaemondiae) inferred from mtDNA sequences, with emphasis on East Asian species. Zool Stud 2007; 46(3):272-289.
8. Murphy NP, Austin CM. Phylogenetic relationships of the globally distributed freshwater prawn genus *Macrobrachium* (Crustacea: Decapoda: Palaemonidae): biogeography, taxonomy and the convergent evolution of abbreviated larval development. Zool Scr 2005; 34(2):187-197.
9. Holthuis LB. The Paleamonidae collected by the siboga and snellius expeditions, with remarks on other species. The Decapoda of the siboga expedition. Part X. The Palaemonidae. I. subfamily Palaemoninae. Siboga-Expeditie Leiden 1950; 39:1-268.
10. Short J. Systematics and biogeography of Australian *Macrobrachium* (Crustacea: Decapoda: Palaemonidae) with descriptions of other new freshwater Decapoda. Ph.D. Thesis, The University of Queensland, 2000.
11. Frezal L, Leblois R. Four years of DNA barcoding: Current advances and prospects. Infect Genet Evol 2008; 8(5):27-36.
12. Rusch DB, Halpern AL, Sutton G, Heidelberg KB, Williamson S, Yooseph S *et al.* The sorcerer II global ocean sampling expedition: Northwest Atlantic through Eastern Tropical Pacific. PLoS Biol 2007; 5(3):e77.
13. Stoeckle M. Taxonomy, DNA, and the Barcode of Life. BioScience 2003; 53:796-797.
14. Guo HY, Wang WW, Yang N, Guo BL, Zhang S, Yang RJ *et al.* DNA barcoding provides distinction between *Radix astragali* and its adulterants. Sci China-Life Sci 2010; 53(8):992-999.
15. Pegg GG, Sinclair B, Briskey L, Aspden WJ. MtDNA barcode identification of fish larvae in the southern Great Barrier Reef, Australia. Sci Mar 2006; 70(2):7-12.
16. Randrianiaina RD, Glaw F, Thomas M, Glos J, Raminosoa N, Vences M. Descriptions of the tadpoles of two species of *Gephyromantis*, with a discussion of the phylogenetic origin of direct development in mantellid frogs. Zootaxa 2007; 1401:53-61.
17. Bucklin A, Wiebe PH, Smolenack SB. DNA barcodes for species identification of euphausiids (Euphausiacea, Crustacea). J Plankton Res 2007; 29(6):483-493.
18. Costa FO, deWaard JR, Boutillier J, Ratnasingham S, Dooh RT, Hajibabaei M, Hebert PDN. Biological identifications through DNA barcodes: The case of the Crustacea. Canadian J Fish Aquat Sci 2007; 64:272-295.
19. Kress WJ, Wurdack KJ, Zimmer EA, Weigt LA, Janzen DH. Use of DNA barcodes to identify flowering plants. Proc Natl Acad Sci USA 2005; 102*(*23):8369-8374.
20. Rossi N, Mantelatto FL. Molecular analysis of the freshwater prawn *Macrobrachium olfersii* (Decapoda, Palaemonidae) supports the existence of a single species throughout its distribution. PLoS ONE 2013; 8(1):e54698.
21. Dwivedi B, Gadagkar SR. The impact of sequence parameter values on phylogenetic accuracy. Biol Med 2009; 1(3):50-62.
22. Muse SV, Gaut BS. A likelihood approach for comparing synonymous and non-synonymous nucleotide substitution rates, with application to the chloroplast genome. Mol Biol Evol 1994; 11(5):715-724.
23. Felsenstein J. Evolutionary trees from DNA sequences: a maximum likelihood approach. J Mol Evol 1981; 17(6):368-376.
24. Kosakovsky Pond SL, Frost SDW, Muse SV. HyPhy: Hypothesis testing using phylogenies. Bioinformatics 2005; 21(5):676-679.
25. Xia X. DAMBE5: A comprehensive software package for data analysis in molecular biology and evolution. Mol Biol Evol 2013; 30(7):1720-1728.
26. Xia X, Xie Z, Salemi M, Chen L, Wang Y. An index of substitution saturation and its application. Mol Phylogenet Evol 2003; 26(1):1-7.
27. Xia X, Lemey P. Assessing substitution saturation with DAMBE. The phylogenetic handbook: A practical approach to DNA and protein phylogeny. eds. by Lemey, P., Salemi, M. & Vandamme, A., 2nd edition Cambridge University Press, 2009, 615-630.
28. Kimura M. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. J Mol Evol 1980; 16(2):111-120.
29. Stuart A, Ord K. Kendall's advanced theory of statistics, Volume 2, 5th edition, Edward Arnold, London. (KA2), 1991.
30. Akaike H. A new look at the statistical model identification. IEEE Trans. Autom Contr 1974; 19:716-723.

31. Schwarz G. Estimating the dimension of a model. Ann Stat 1978; 6(2):461-464.
32. Fincham AA, Wickins JF. Identification of commercial prawns and shrimps. Brit Mus Publ 1976; 779:7.
33. Whelan S, Goldman N. A general empirical model of protein evolution derived from multiple protein families using a maximum likelihood approach. Mol Biol Evol 2001; 18(5):691-699.
34. Zhang Q, Cheng Q, Guan W. Mitochondrial COI gene sequence variation and taxonomic status of three *Macrobrachium* species. Zool Res 2009; 30(6):613-619.
35. Pileggi LG, Mantelatto FL. Molecular phylogeny of the freshwater prawn genus *Macrobrachium* (Decapoda, Palaemonidae), with emphasis on the relationships among selected American species. Invertebr Syst 2010; 24(1):194-208.