Original Research Article Molecular Biology

International Journal of Pharma and Bio Sciences

EVALUATION OF MITOCHONDRIAL D-LOOP AND 16SrRNA DIVERSITY AMONG CULTURED *PENAEUS MONODON* SAMPLES

DEBABRATA MONDAL[1] AND NRIPENDRANATH MANDAL[*2]

[1]Division of Molecular Medicine, Bose Institute, P-1/12 CIT Scheme VIIM, Kolkata – 700054, India.
[2] Professor, Division of Molecular Medicine, Bose Institute, P-1/12 CIT Scheme VIIM, Kolkata – 700054, India.*

ABSTRACT

The giant black tiger shrimp, *Penaeus monodon* (Fabricius, 1798) is one of the most widely captured and cultured marine shrimp species and has a huge global economic impact. The last three decades, shrimp aquaculture has been threatened by various major virus outbreaks, mainly white spot syndrome virus. Therefore, evaluation of genetic diversity, as well as population genetic differentiation is required to identify overexploited regions of high-quality pond-reared *P. monodon*. In this study, the population genetic diversity of *P. monodon* was determined by the molecular data of the mitochondrial control region (mt D-Loop) and the 16SrRNA gene. The mitochondrial DNA (mtDNA) regions were highly AT-rich and a very high level of genetic diversity was observed. The diversity of 16SrRNA haplotypes was (0.963 ± 0.023), and the diversity of D-Loop haplotypes was (0.972 ± 0.017). The mitochondrial control region (D-Loop) was highly polymorphic and should be the potential markers for intraspecific population diversity study of *P. monodon*. The 16SrRNA would be the best marker for phylogenetic studies in *P. monodon*. Linkage disequilibrium analysis revealed that the increment of rare mutations and haplotypes. This study has a great significance to maintain high-quality pond-reared stock diversity, and also in accomplishing selective breeding.

KEYWORDS: ***Cultured Penaeus monodon, haplotype, mitochondrial DNA, genetic diversity, linkage disequilibrium, phylogenetics***

NRIPENDRANATH MANDAL*

Professor, Division of Molecular Medicine, Bose Institute, P-1/12 CIT Scheme VIIM, Kolkata – 700054, India.

Corresponding Author

Received on: 23-07-2019
Revised and Accepted on: 03-09-2019
DOI: http://dx.doi.org/10.22376/ijpbs.2019.10.4.b51-68

INTRODUCTION

The giant black tiger shrimp, *Penaeus monodon* (Fabricius, 1798) is one of the most widely captured and cultured shrimp species throughout the world, including India. The last three decades, shrimp aquaculture has been threatened by various major virus outbreaks, mainly white spot syndrome virus.[11,12] Therefore, breeding in captivity is the major requirement to facilitate genetic upgrading for better growth as well as defense against various diseases. Identification of existing cultured stocks and evaluation of their growth performance under commercial conditions would be the most crucial step towards the selective breeding strategies of *P. monodon*.[1] Therefore, evaluation of genetic diversity, as well as population genetic differentiation is required to identify overexploited regions of high-quality pond-reared *P. monodon*.[2] In aquaculture species, mitochondrial DNA sequences are widely used to evaluate the genetic diversity, and these sequences are an extremely powerful tool to investigate the genetic variability and phylogenetic relationships among crustacean groups.[3-4] The mitochondrial genome of *P. monodon* contains thirteen protein-encoding genes, 22 tRNA genes, and two rRNA genes.[5] The mitochondrial DNA control region (mtDNA D-Loop) is a non-coding, hypervariable region that has been used for population genetics studies in crustaceans.[6-7] The size of the complete mitochondrial genome of *P. monodon* is about 16 kb[5], in it, a long stretch of AT-rich noncoding control region of 991 bp is known as the D-Loop.[9] D-Loop based genetic studies on penaeid shrimps have been well documented.[2,4,10] Similarly, the16SrRNA mitochondrial gene contains both fast- and slow-evolving regions, which can provide useful information about the family taxonomic level.[8] This study was aimed to evaluate the population genetic diversity and phylogeny of cultured *P. monodon*, and samples were collected from highly-virus infected ten different cultured ponds in West Bengal, India. To accomplish the above-mentioned aim, the partial sequence of the mtDNA control region (D-Loop) and the 16SrRNA gene were analyzed.

MATERIALS AND METHODS

Ethical Statement

P. monodon is not an endangered or protected species and did not belong to the IUCN red list. The area positions from where the samples were compiled are not protected and does not involve any specific licenses. All the work was done by using the approval from the Institutional Animal Ethics Committee, of the Bose Institute (Registration. No. 95/1999/CPCSEA).

Collection of samples

Total 60 mature (approximately 55 gm weight per sample) cultured *P. monodon* (Fabricius, 1798), samples were randomly collected from ten different cultured ponds of West Bengal, India which had been highly affected by White spot disease. Sample collection locations were North-24-Parganas, South-24-Parganas, and Purba Medinipur (Figure 1). Presence or absence of White spot syndrome virus was checked by following the same protocol as previously described by Chakrabarty *et al*.[11-12]

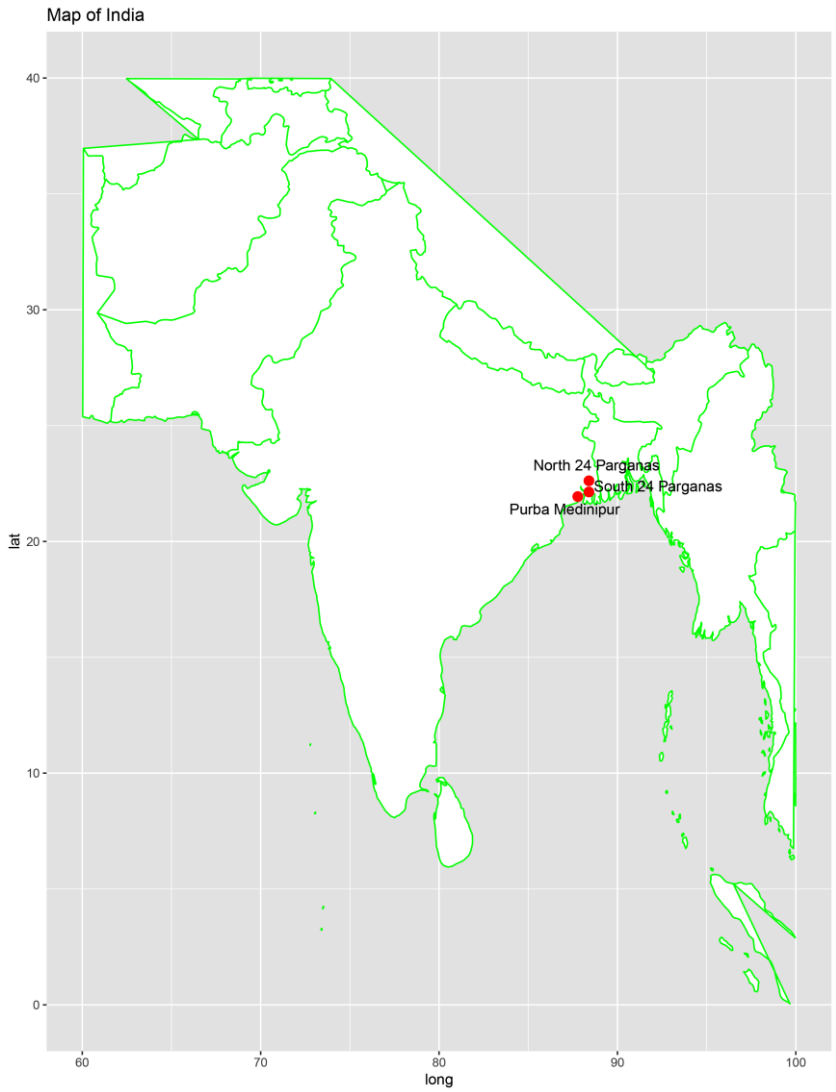

Figure 1
Map of India and sample collection locations were marked with red dots (North-24-Parganas, South-24-Parganas, and Purba Medinipur).

Preparation of Genomic DNA

Pleopod and tail tissues (average 100 mg) were dissected out of each shrimp sample, and subjected to genomic DNA isolation using phenol–chloroform method.[13] Each DNA sample was quantified by a Shimadzu UV160U spectrophotometer (Shimadzu Corp., Kyoto, Japan). Each of the isolated DNA samples were electrophoresed in a 0.8% agarose gel and subsequently, stained with ethidium bromide (10 mg mL^{-1}) to determine their quality before using in PCR amplification.

The PCR amplification of mitochondrial D-Loop and 16SrRNA fragment, and sequencing of PCR amplified fragments

The PCR amplification of two mtDNA segments was performed in a thermal cycler (Applied Biosystems, Veriti™ 96 Well Thermal Cycler # 9902, California, USA) with a reaction mixture of 25 µL containing 400 ng DNA, 200 µM deoxyribonucleotide triphosphates, 2.0 mM $MgCl_2$, 1X buffer, and 1.0 U *Taq* DNA polymerase (MP Biomedical, California, USA) and 30 pmol of each oligonucleotide primer [D-Loop fragment: 12S (F): 5' AAGAACCAGCTAGGATAAAACTTT 3' and 1R (R): 5' GATCAAAGAACATTCTTTAACTAC 3' [4]; 16SrRNA fragment: 16SAR (F): 5' CGCCTGTTTATCAAAAACAT 3' and 16SBR (R): 5' CCGGTTTGAACTCAGATCATG 3' [14]] from Integrated DNA Technologies (Integrated DNA Technologies, Inc., Iowa, USA). The thermal profile for PCR was as follows: 94˚C for 5 min, followed by 35 cycles of 94˚C for 1 min, experimental annealing temperature: 50-55˚C (D-Loop) and 60˚C (16SrRNA) for 1 min, an extension for 1 min at 72˚C and the final extension at 72˚C for 5 min. The amplified DNA was electrophoresed in a 2.0% agarose gel at 90 V for 2 h, and subsequently stained with ethidium bromide (10 mg mL^{-1}) and photographed in a Bio-Rad Molecular Imager® ChemiDoc ™ XRS+ Imaging System (Bio-Rad Laboratories, Inc., California, USA). The PCR-amplified single uniform band was then purified using the GeneJET Gel Extraction Kit (Thermo Scientific, Massachusetts, USA). The single fragments were bidirectionally sequenced using the ABI PRISM dye terminator ready reaction kit (Applied Biosystems, California, USA) followed by the manufacturer protocol in an automated DNA sequencing machine (Applied Biosystems Inc., California, USA). Each chromatogram was checked by Finch TV v. 1.4.0 (Geospiza, Inc., Washington, USA; http://www.geospiza.com) and BioEdit v. 7.2.5.[15] Only high-quality chromatograms

were considered for analysis. Sequences from both strands in each specimen were aligned with Codon Code Aligner v. 8.0.2 (Codon Code Corporation, Massachusetts, USA) and GeneDoc v. 2.7.[16] Individual consensus sequences were retrieved from both alignments and manually checked. The FASTA sequence format of chromatograms were subjected to the NCBI BLASTN program [17-18] to observe sequence similarity with previously submitted sequences.

Intraspecific nucleotide diversity

The D-Loop and 16SrRNA sequences were aligned with the Multiple Sequence Comparison by Log-Expectation (MUSCLE) algorithm.[19] The Maximum Composite Likelihood (MCL)[20] estimation of the pattern of nucleotide substitution, the Maximum Likelihood (ML) estimation of substitution matrix, and the ML estimate of transition/transversion bias (*R*) were analyzed in MEGA v. 7.0.26.[21] The number of haplotypes (h), the number of segregating sites (S), total number of mutations (Eta), haplotype diversity (Hd ± S.E.), and nucleotide diversity (Pi ± S.E.) were estimated by DnaSP v. 6.[22] The total number of transitions, total number of transversions and the total number of substitutions were calculated by Arlequin v. 3.5.[23]

Finding of unique haplotypes

Haplotypes distance matrix based on pairwise haplotypes sequence similarity was used for classical hierarchical cluster analysis in PAST v. 3.23.[24] In a classical hierarchical cluster analysis, the pairwise distance matrix was subjected to the UPGMA (unweighted pair group method with arithmetic mean) clustering algorithm and clustering was performed based on the Euclidian's correlation similarity index. Matrix plots of D-Loop and 16SrRNA haplotypes were generated to reveal the pairwise haplotype sequence similarity in PAST v. 3.23 [24]

Analysis of haplotype network and phylogenetic relationship

The Median-Joining (MJ) network based on D-Loop and 16SrRNA haplotype data were drawn using the program Network v. 5 (http://www.fuxus.engineering.com). Phylogenetic relationships among D-Loop and 16SrRNA haplotypes were revealed by PhyML v. 3.0.[25-26] and MAFFT v. 7 [27] program. PhyML v. 3.0 was run under the statistical test called Approximate Likelihood - Ratio Test (aLRT: SH-like) [28] for branch support applying the four substitution rate categories. The final PhyML tree was visualized in TreeDyn v. 198.3.[29] MAFFT v. 7 was run under the Neighbor-Joining method applying 1000 bootstrap resampling, and the tree was visualized in Phylo.io.[30]

STATISTICAL ANALYSIS

Linkage disequilibrium (LD) of the haplotypes was used to describe the degree to which recombination rearranges genetic diversity. [31] LD measured the association of alleles at different loci with particular traits or phenotypes to identify a single polymorphism within a gene. [32] In LD measurement, the squared correlation coefficient between polymorphic sites (r^2) and standardized disequilibrium coefficient (D') [33-34] was evaluated. The r^2 was covered by both recombination and mutation; whereas, D' was covered only by recombination. [35-36] The pairwise comparisons between polymorphic loci was determined by Fisher's exact test and Bonferroni Correction, and both measured the statistical significance ($P < 0.05$). The LD of D-Loop and 16SrRNA haplotypes was measured using DnaSP v. 6. [22] Indels were treated as single polymorphic sites, and decay was examined exploratorily by graphs of pairwise distances (bp) versus r^2.

RESULTS

The PCR amplification of mitochondrial DNA fragments of P. monodon

The PCR of the mitochondrial control region (D-Loop) and 16SrRNA were performed to determine the genetic diversity among cultured *P. monodon*. The 612 bp fragment of D-Loop (contained partial 12SrRNA sequence) was amplified by 12S (F) - 1R (R) primer set, and a 561 bp fragment of 16SrRNA gene was amplified by 16SAR (F) - 16SBR (R) primer set (Figure 2). The FASTA format of each sequence was subjected to the NCBI BLASTN program and showed the highest similarity with previously GenBank, NCBI submitted sequences of *P. monodon* (Figures 3 and 4).

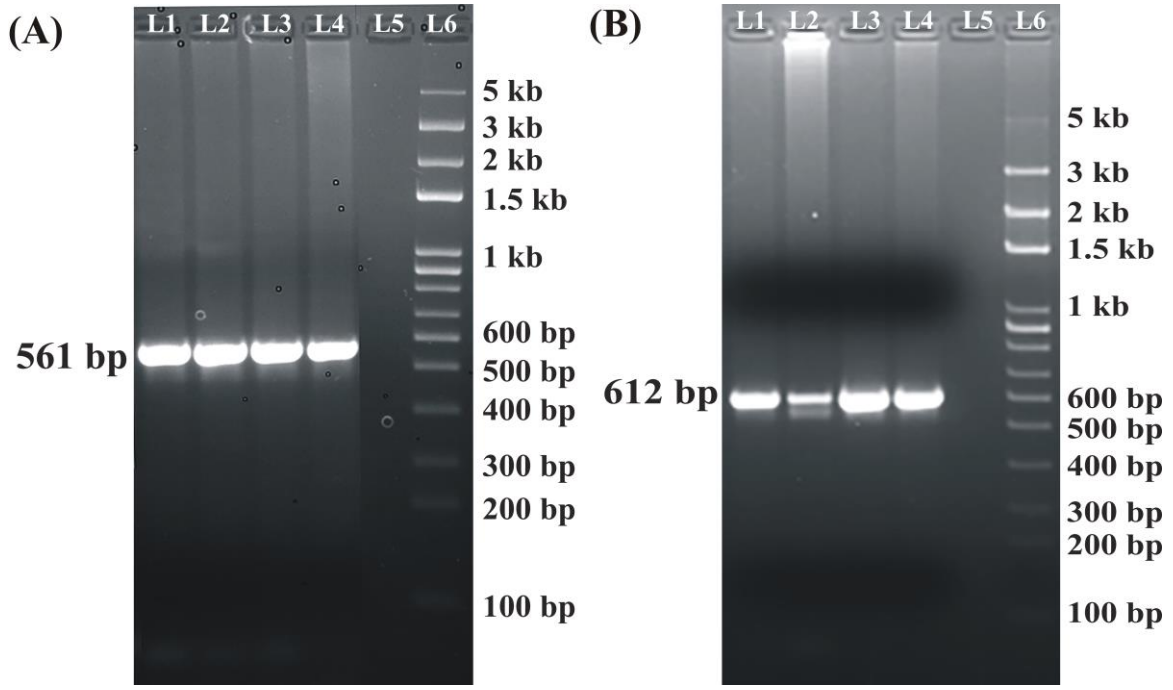

Figure 2
The PCR amplification of 16SrRNA and D-Loop mtDNA region.

(A) The PCR amplified 561 bp fragments of the partial 16SrRNA gene were shown in L1 – L4, whereas, L5 indicated a negative control and L6 indicated a molecular weight marker. (B) The PCR amplified 612 bp fragments of partial D-Loop region was shown in L1 – L4, whereas, L5 indicated a negative control and L6 indicated a molecular weight marker.

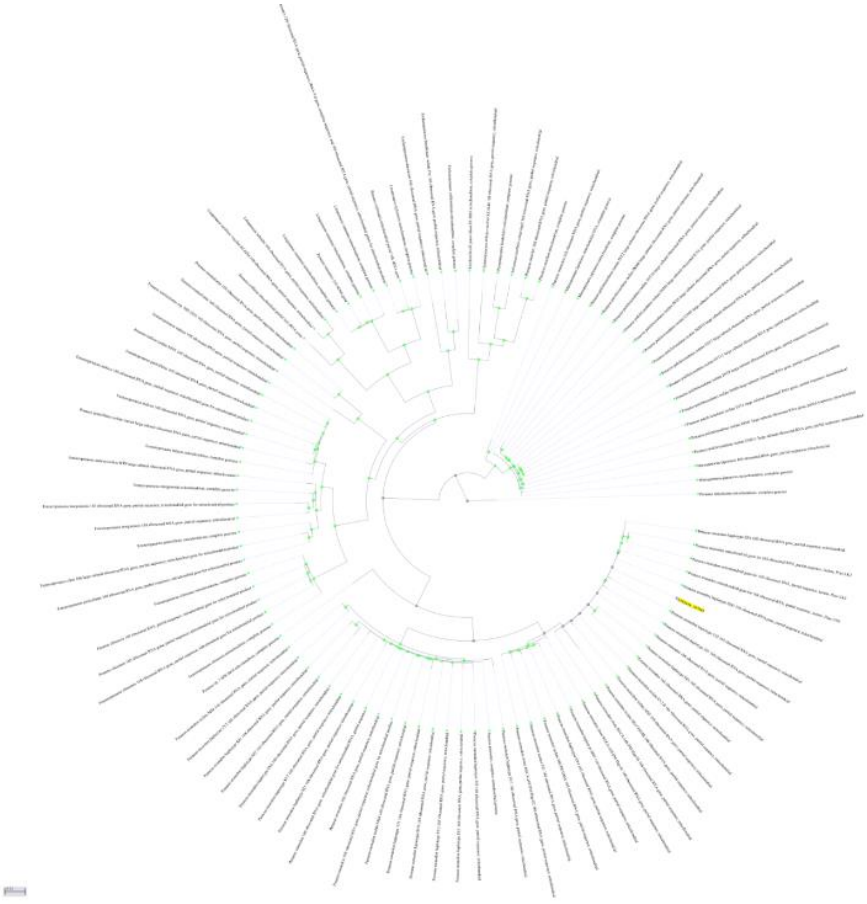

Figure 3
The NCBI BLASTN result of 16SrRNA gene (561 bp of Haplotype 1).

The figure showed the highest sequence similarity with previously submitted sequences. Query sequence (Haplotype 1 from this investigation) was highlighted in yellow color.

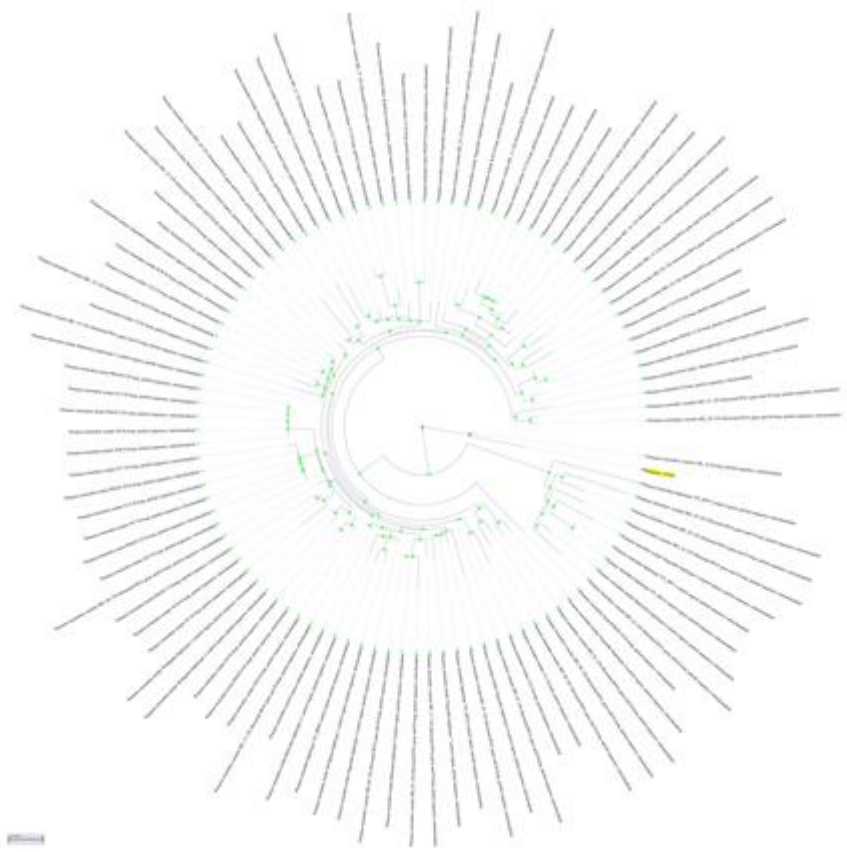

Figure 4
The NCBI BLASTN result of D-Loop region (612 bp of Haplotype 1)

The figure showed the highest sequence similarity with previously submitted sequences. Query sequence (Haplotype 1 from this investigation) was highlighted in yellow color.

Intraspecific nucleotide diversity

The 561 bp 16SrRNA gene fragments were AT-rich and equal amount of A (33.8%) and T (33.6%) was observed (Figure 5 and 7; Table 1). The 612 bp D-Loop fragments were AT-rich and frequency of A (39.9%) and T (40.5%) was almost equal (Figure 6 and 7; Table 1). For 16SrRNA gene, the number of haplotypes (h) was 18, and haplotype diversity (Hd ± S.E.) and nucleotide diversity (Pi ± S.E.) was (0.963 ± 0.023) and (0.01093 ± 0.00128), respectively (Table 2 and 3). For D-Loop sequences, the number of haplotypes (h) was 22, and haplotype diversity (Hd ± S.E.) and nucleotide diversity (Pi ± S.E.) was (0.972 ± 0.017) and (0.01103 ± 0.00170), respectively (Table 4 and 5). The total number of mutations (Eta) or variable/polymorphic sites (S) of D-Loop and 16SrRNA haplotypes were 60 (56) and, 25 respectively. The conservation (C) and homozygosity among D-Loop and 16SrRNA haplotype sequences were shown in Table 3 and 5. The conservation (C) and homozygosity among 16SrRNA haplotype sequences were higher than D-Loop haplotypes. The number of transitions, number of transversions, and the number of substitutions of D-Loop and 16SrRNA haplotypes were shown in Table 2 and 4. The ML estimate of transition/transversion bias (*R*) of 16SrRNA and D-Loop sequences was 3.91 and 3.41, respectively. The transition/transversion bias (*R*) was highest among 16SrRNA sequences. The number of invariable [+*I*] sites of 16SrRNA and D-Loop sequences were 48% and 7%, respectively. The number of invariable [+*I*] sites was also highest in 16SrRNA sequences. The ML estimate of substitution rates depicted that four types of transversional substitutions and two types of transitional substitutions have occurred within all sequences. The base substitution rate for four types of transversional substitutions and two types of transitional substitutions was the same. The transitional substitutions rate was highest in 16SrRNA sequences (19.91), whereas, the transversional substitutions rate was highest in D-Loop sequences (2.83) (Table 6). The Maximum Composite Likelihood estimate of the pattern and rate of nucleotide substitution depicted that four types of transversional substitutions and four types of transitional substitutions have occurred within all sequences. For D-Loop sequences, the overall transition/transversion bias (*R*) was 2.023. For 16SrRNA sequences, the overall transition/transversion bias (*R*) was 3.317. Overall, the rate of transitions was higher than the rate transversions [T to C transition was highest for 16SrRNA sequences (45.6) as well as D-Loop sequences (39.67)] (Table 7).

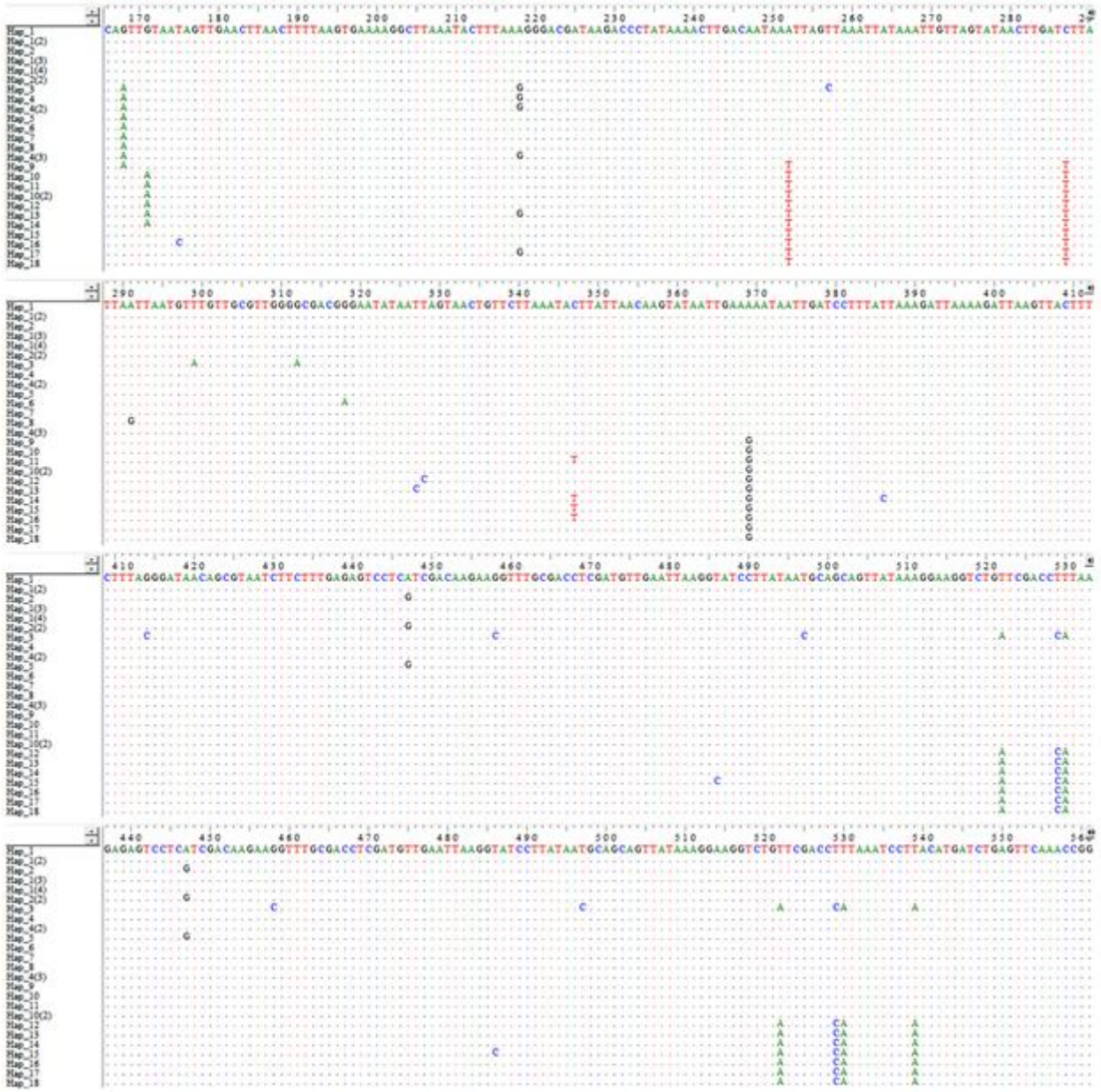

Figure 5
The alignment result of 16SrRNA haplotypes by the MUSCLE algorithm.

The identical nucleotides were represented by color dots; whereas, single polymorphic sites of each haplotype were shown by specific nucleotides along with their position. Single nucleotide polymorphic sites were valuated against Haplotype 1 of 16SrRNA gene.

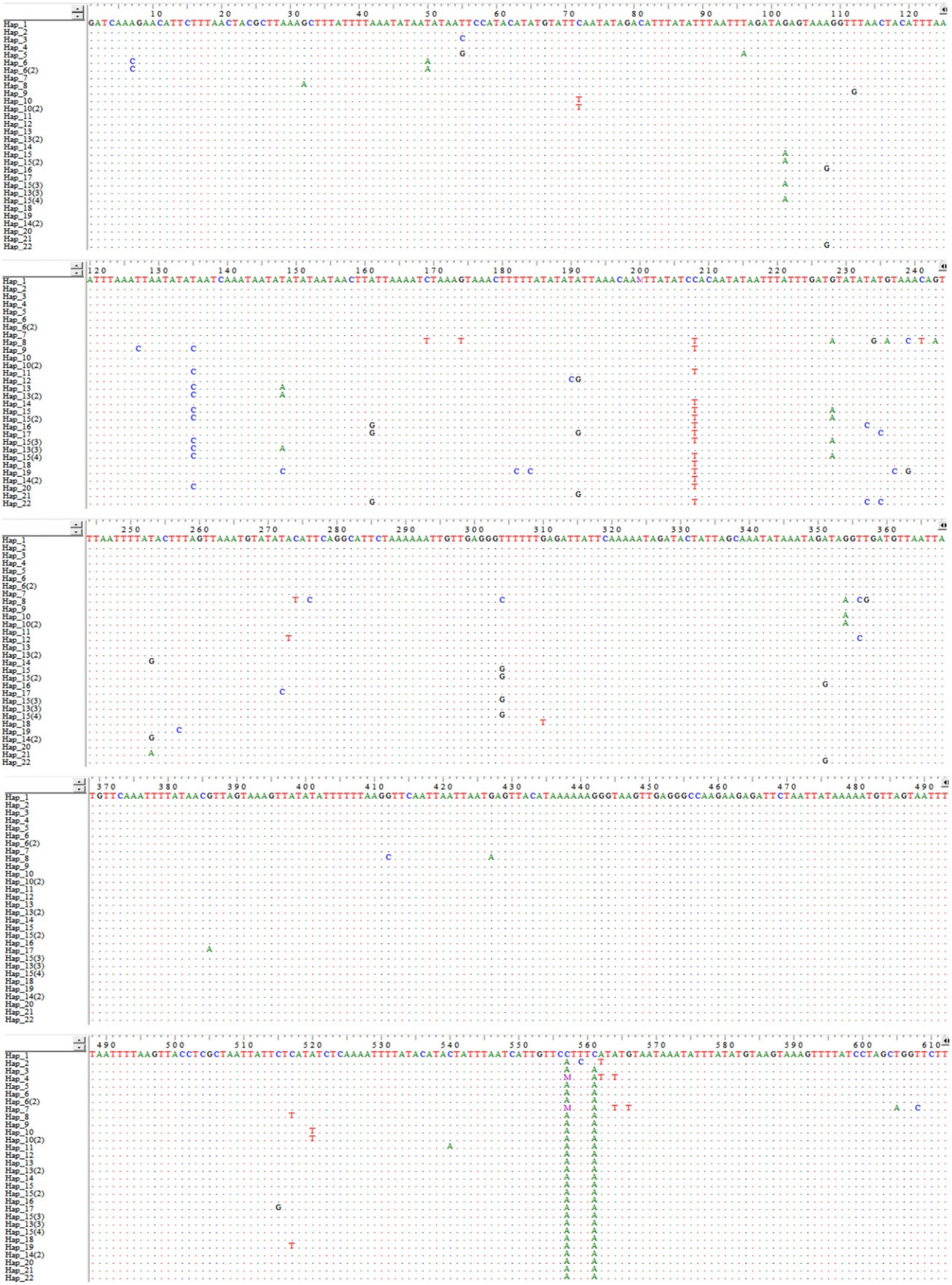

Figure 6
The alignment result of D-Loop haplotypes by the MUSCLE algorithm.

The identical nucleotides among haplotypes were represented by color dots; whereas, single polymorphic sites were shown by specific nucleotides along with their position. Single nucleotide polymorphic sites were valuated against Haplotype 1 of D-Loop region.

>Haplotype_1_16SrRNA_561 bp

CGCCTGTTTATCAAAAACATGTCTATATGATTGTTATATAAAGTCTAGCCTGCCCACTGAATTATTTTTAAAGGGCCGCGGTATACTG
ACCGTGCGAAGGTAGCATAATCATTAGTCTTTTAATTGAAGGCTTGTATGAATGGTTGGACAAAAAGTAATCTGTCTCAGTTGTAAT
AGTTGAACTTAACTTTTAAGTGAAAAGGCTTAAATACTTTAAAGGGACGATAAGACCCTATAAAACTTGACAATAAATTAGTTAAAT
TATAAATTGTTAGTATAACTTGATCTTAATTAATGTTTGTTGCGTTGGGGCGACGGGAATATAATTAGTAACTGTTCTTAAATACTTA
TTAACAAGTATAATTGAAAAATAATTGATCCTTTATTAAAGATTAAAAGATTAAGTTACTTTAGGGATAACAGCGTAATCTTCTTTGA
GAGTCCTCATCGACAAGAAGGTTTGCGACCTCGATGTTGAATTAAGGTATCCTTATAATGCAGCAGTTATAAAGGAAGGTCTGTTC
GACCTTTAAATCCTTACATGATCTGAGTTCAAACCGG

>Haplotype_1_D Loop_612 bp

GATCAAAGAACATTCTTTAACTACGCTTAAAGCTTTATTTTAAATATAATATAATTCCATACATATGTATTCAATATAGACATTTATAT
TTAATTTAGATAGAGTAAAGGTTTAACTACATTTAAATTAATATATAATCAAATAATATATATAATAACTTATTAAAATCTAAAGTAA
ACTTTTTATATATATTAAACAAMTTATATCCACAATATAATTTATTTGATGTATATATGTAAACAGTTAATTTTATACTTTAGTTAAATG
TATATACATTCAGGCATTCTAAAAAATTGTTGAGGGTTTTTTGAGATTATTCAAAAATAGATACTATTAGCAAATATAAATAGATAG
GTTGATGTTAATTATGTTCAAATTTTATAACGTTAGTAAAGTTATATATTTTTTAAGGTTCAATTAATTAATGAGTTACATAAAAAAG
GGTAAGTTGAGGGCCAAGAAGAGATTCTAATTATAAAAATGTTAGTAATTTTAAGTTACCTCGCTAATTATTCTCATATCTCAAAAT
TTTATACATACTATTTAATCATTGTTCCTTTCATATGTAATAAATATTTATATGTAAGTAAAGTTTTATCCTAGCTGGTTCTT

Figure 7
The FASTA sequence format of partial fragment of Haplotype 1 of the 16SrRNA gene as well as D-Loop region.

These two sequences were taken as a standard for SNP evaluation and other analysis.

Table 1
Location and frequency of nucleotide polymorphism within unique haplotypes.

No. of Haplotypes	Nucleotide Position	Nucleotide Shift (nucleotide change respect to haplotype 1)	Frequency (no of times presents within total no of haplotypes)
16SrRNA (h = 18)	168, 171, 312, 318	G to A	14
	175, 257, 327, 328, 386, 486, 497, 529	T to C	15
	218, 291, 369, 447	A to G	19
	287, 347	C to T	14
	252	A to T	10
	299, 522, 530, 539	**T to A**	**25**
	414, 458	G to C	2
D-Loop (h = 22)	7, 239	A to C	2
	32, 102, 228, 236, 243, 354, 386, 427	G to A	10
	50, 96, 148, 253, 605	T to A	5
	273, 520, 562, 564	A to T	6
	55, 127, 135, 148, 182, 184, 190, 233, 237, 257, 272, 276, 304, 356, 559, 608	T to C	22
	55, 112, 253, 304, 357	T to G	5
	72, 169, 208, 241, 274, 517	C to T	17
	108, 161, 191, 234, 239, 351	A to G	12
	174, 310, 566	G to T	3
	412	G to C	1
	515	C to G	1
	540, 557, 561	**C to A**	**42**

The most frequent polymorphic sites were shown in bold.

Table 2
Molecular diversity indices of 16SrRNA haplotypes.

No. of Sequence	No. of Haplotypes	No. of polymorphic sites	No. of transitions	No. of transversions	No. of substitutions	Gene diversity	Nucleotide composition
25	18	25	18	7	25	0.9633 ± 0.0235	C: 13.92% T: 33.59% A: 33.76% G: 18.73% Total:100.00%

Table 3
Haplotype diversity and LD analysis of 16SrRNA haplotypes.

Number of Haplotypes (h)	Haplotype (gene) diversity (Hd ± S.E.)	Nucleotide diversity (Pi ± S.E.)	Total number of mutations (Eta) or Number of variable/polymorphic sites (S)	Sequence conservation (C)	Homozygosity	Linkage Disequilibrium						
						Number of pairwise comparisons	Number of significant pairwise comparisons by Fisher's exact test	Number of significant comparisons using the Bonferroni procedure	Number of significant pairwise comparisons by chi-square test	Number of significant comparisons using the Bonferroni procedure	|D'| values	R2 values
18	0.963 ± 0.023	0.01093 ± 0.00128	25	0.955	0.989	300	45	9	53	24	Y = 1.0287 - 0.6203X	Y = 0.1771 - 0.2535X

Table 4
Molecular diversity indices of D-Loop haplotypes.

No. of Sequence	No. of Haplotypes	No. of polymorphic sites	No. of transitions	No. of transversions	No. of substitutions	Gene diversity	Nucleotide composition
31	22	56	38	23	61	0.9724 ± 0.0172	C: 9.04% T: 40.46% A: 39.87% G: 10.63% Total:100.00%

Table 5
Haplotype diversity and LD analysis of D-Loop haplotypes

Number of Haplotypes (h)	Haplotype (gene) diversity (Hd ± S.E.)	Nucleotide diversity (Pi ± S.E.)	Total number of mutations (Eta) or Number of variable/polymorphic sites (S)	Sequence conservation (C)	Homozygosity	Linkage Disequilibrium						
						Number of pairwise comparisons	Number of significant pairwise comparisons by Fisher's exact test	Number of significant comparisons using the Bonferroni procedure	Number of significant pairwise comparisons by chi-square test	Number of significant comparisons using the Bonferroni procedure	|D'| values	R^2 values
22	0.972 ± 0.017	0.01103 ± 0.00170	60 (Eta) or 56 (S)	0.908	0.985	1225	95	0	172	92	Y = 0.9824 + 0.0431X	Y = 0.1634 - 0.3166X

Table 6
Maximum Likelihood Estimate of Substitution Matrix of 16SrRNA and D-Loop sequences.

	16SrRNA Sequences				D-Loop Sequences			
	A	**T/U**	**C**	**G**	**A**	**T/U**	**C**	**G**
A	--	*2.55*	*2.55*	**19.91**	--	*2.83*	*2.83*	**19.34**
T/U	*2.55*	--	**19.91**	*2.55*	*2.83*	--	**19.34**	*2.83*
C	*2.55*	**19.91**	--	*2.55*	*2.83*	**19.34**	--	*2.83*
G	**19.91**	*2.55*	*2.55*	--	**19.34**	*2.83*	*2.83*	--

Each entry was the probability of substitution (r) from one base (row) to another base (column). Substitution patterns and rates were estimated under the Kimura (1980) 2-parameter model (+*G*+*I*). [37] A discrete Gamma distribution was used to model evolutionary rate differences among sites (4 categories, [+*G*], parameter = 0.0500). The rate variation model allowed for some sites to be evolutionarily invariable [+*I*] sites (48% sites for 16SrRNA and 7% sites for D-Loop). Rates of different transitional substitutions were shown in **bold** and those of transversional substitutions were shown in *italics*. All positions containing gaps and missing data were eliminated. There was a total of 561 positions (for 16SrRNA) and 610 positions (for D-Loop) in the final dataset.

Table 7
Maximum Composite Likelihood Estimate of the Pattern of Nucleotide substitution of 16SrRNA and D-Loop sequences.

	16SrRNA Sequences				D-Loop Sequences			
	A	**T**	**C**	**G**	**A**	**T**	**C**	**G**
A	--	*3.34*	*1.38*	**5.57**	--	*4.74*	*1.04*	**5.96**
T	*3.36*	--	**18.89**	*1.86*	*4.66*	--	**8.71**	*1.25*
C	*3.36*	**45.6**	--	*1.86*	*4.66*	**39.67**	--	*1.25*
G	**10.04**	*3.34*	*1.38*	--	**22.28**	*4.74*	*1.04*	--

Each entry showed the probability of substitution (r) from one base (row) to another base (column). [20] For simplicity, the sum of r values was made equal to 100. The rates of different transitional substitutions were shown in **bold** and those of transversional substitutions were shown in *italics*. The nucleotide frequencies of 16SrRNA sequences were 33.76% (A), 33.59% (T/U), 13.92% (C), and 18.73% (G). The nucleotide frequencies of D-Loop sequences were 39.85% (A), 40.58% (T/U), 8.91% (C), and 10.66% (G). All positions containing gaps and missing data were eliminated. The transition/transversion rate ratios of 16SrRNA sequences were *k1* = 2.991 (purines) and *k2* = 13.65 (pyrimidines). The overall transition/ transversion bias was *R* = 3.317, where *R* = [A*G**k1* + T*C**k2*]/[(A+G) *(T+C)]. Similarly, the transition/transversion rate ratios of D-Loop sequences were *k1* = 4.784 (purines) and *k2* = 8.364 (pyrimidines). The overall transition/transversion bias was *R* = 2.023, where *R* = [A*G**k1* + T*C**k2*]/[(A+G) *(T+C)].

Finding of unique haplotypes

In a sequence pairwise similarity matrix, the Cophenetic correlation value of D-Loop and 16SrRNA haplotypes were 0.9663 and 0.6032, respectively; and were found to be suitable for hierarchical clustering. Among 22 haplotypes of D-Loop, haplotype 8 showed the lowest similarity with another paired haplotype and clustered in a separate branch of the dendrogram. The scale bar indicates the evolutionary distance, and the unit of scale expressed the homology as well as divergence. Among 18 haplotypes of 16SrRNA, haplotype 3 showed a lower similarity (Figure 8). In the matrix plot of D-Loop haplotypes, haplotype 8 showed the lowest similarity with another paired haplotypes (dissimilarity range: 0.0273 – 0.041); similarly, among 16SrRNA haplotypes, haplotype 3 showed the lower similarity (dissimilarity range: 0.0207 – 0.031) (Figure 8).

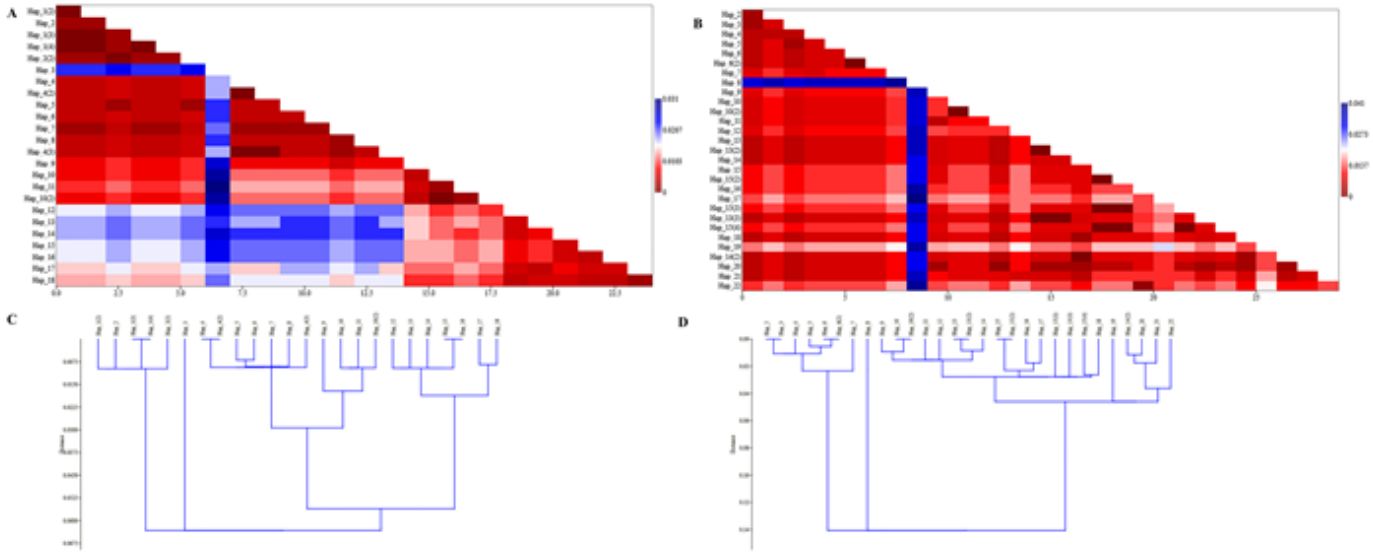

Figure 8
The matrix plot and classical hierarchical cluster analysis of 16SrRNA and D-Loop haplotypes.

(A) The sequence pairwise similarity matrix of 16SrRNA haplotypes. The colors indicated the similarity index. Dark red color denoted the highest sequence similarity; whereas, dark blue denoted the lowest sequence similarity. (B) The sequence pairwise similarity matrix of D-Loop haplotypes. Dissimilarity range was represented by similarity index. Dark red color denoted the highest sequence similarity; whereas, dark blue denoted the lowest sequence similarity. (C) The classical hierarchical cluster analysis of 16SrRNA haplotypes using the UPGMA method. (D) The classical hierarchical cluster analysis of D-Loop haplotypes using the UPGMA method.

Analysis of haplotype network and phylogenetic relationship

Haplotype networks reveal the intraspecific genealogical relationship and can be used to infer the biogeography and history of populations. The D-Loop haplotype network depicted that haplotype 15 was the most frequently observed haplotype; which was followed by haplotype 13. These two haplotypes were connected with other haplotypes through a very low number of nucleotide change or substitution; whereas, haplotype 8 showed the highest number of nucleotide substitutions (Figure 9). In 16SrRNA haplotype network, haplotype 1 and haplotype 4 were the most frequently observed haplotype and connected with other haplotypes through a very low number of nucleotide substitutions; whereas, haplotype 3 showed the highest number of nucleotide substitutions (Figure 9). In PhyML and MAFFT maximum-likelihood phylogenetic tree of D-Loop and 16SrRNA haplotypes, branch length represented the substitution or residue per site, and which is proportional to the evolutionary distances between haplotypes. In MAFFT maximum-likelihood phylogenetic tree, the bootstrap value of each node was shown and which denoted the confidence of topology of each node of the tree. Higher bootstrap values denoted a higher level of statistical support of node. In PhyML and MAFFT maximum-likelihood phylogenetic tree, haplotype 8 (D-Loop) and haplotype 3 (16SrRNA) were less closely related to other haplotypes. These two haplotypes represented as an outgroup in the phylogenetic analysis (Figure 10).

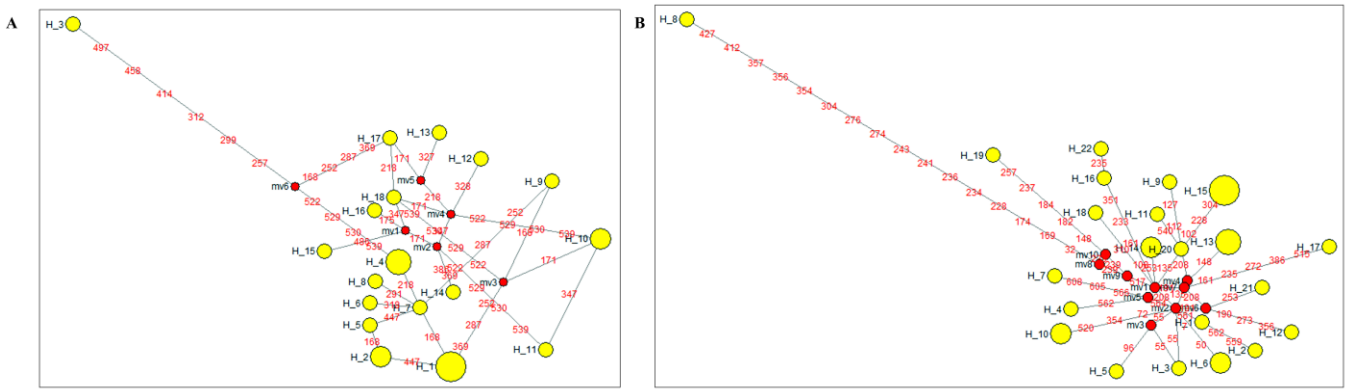

Figure 9

The Median-Joining network of 16SrRNA haplotypes (A) and D-Loop haplotypes (B).

Each haplotype was represented by a yellow circle and the area of each circle was proportional to the relative frequency of haplotypes; length of connecting lines between haplotypes depicted the number of nucleotide differences between them and numbers cross the connecting lines represent the sites of nucleotide substitutions. In median joining networks, median vectors were shown in red circles.

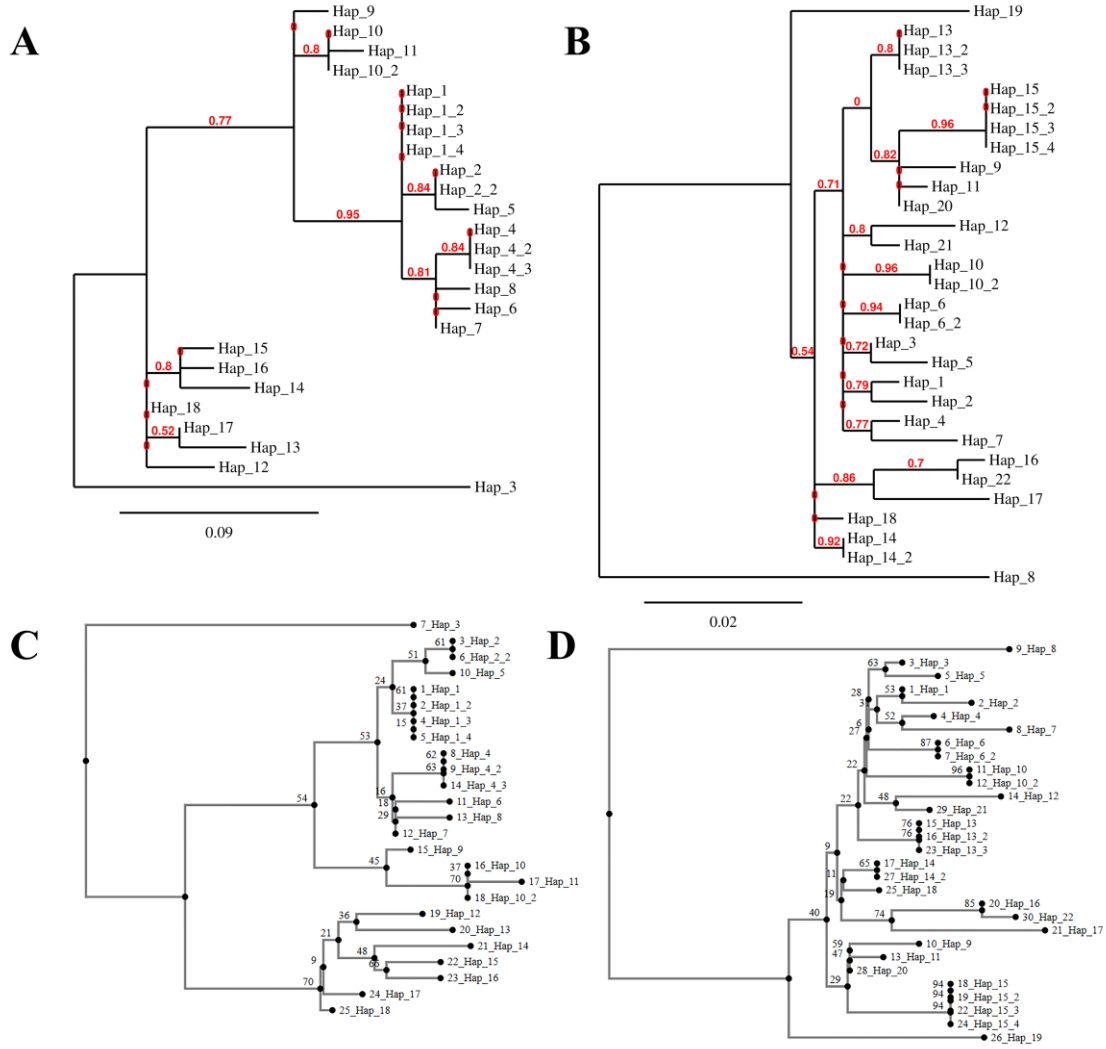

Figure 10

The Maximum-Likelihood (ML) phylogenetic tree of D-Loop and 16SrRNA haplotypes.

(A) The ML phylogenetic tree of 16SrRNA haplotypes generated by PhyML program, where scale 0.09 denoted the unit of nucleotide substitution per site. The reliability of internal branches was measured by nonparametric bootstrap method and the bootstrap values were displayed on the maximum likelihood phylogeny estimated from the haplotype dataset. (B) The ML phylogenetic tree of D-Loop haplotypes

generated by PhyML program, where scale 0.02 denoted the unit of nucleotide substitution per site. The nonparametric bootstrap method was applied to measure the reliability of internal branches as well as the whole phylogenetic tree. The bootstrap values were shown before the starting of each internal branch. (C) The ML-Neighbor-Joining phylogenetic tree of 16SrRNA haplotypes generated by MAFFT program, where bootstrap value of each node was shown before the node and sequence serial number (according to the MUSCLE algorithm) was shown before the name of haplotype or terminal node or leaf. Higher bootstrap values indicated high-level of reliability of the phylogenetic tree. (D) The ML-Neighbor-Joining phylogenetic tree of D-Loop haplotypes generated by MAFFT program, where bootstrap value of each node was shown before the node, indicated the reliability of internal branches; and sequence serial number (according to the MUSCLE algorithm) was shown before the name of haplotype or terminal node or leaf.

Linkage disequilibrium (LD) analysis

In LD measurement, the r^2 and D' was calculated based on single nucleotide substitutions identified for each locus. In case of 16SrRNA and D-Loop haplotypes, a total of 300 and 1225 pairwise comparison have been evaluated and plotted. The number of significant pairwise comparisons using the Bonferroni procedure of 16SrRNA and D-Loop haplotypes were shown in Table 8. A perfect linkage disequilibrium (total disequilibrium or complete LD; D' = 1.000 and r^2 = 1.000) was observed during pairwise SNP comparison and which are not separated by recombination among 16SrRNA and D-Loop haplotypes (Table 8) except some pairwise SNP comparison. Some pairwise SNPs of 16SrRNA showed fairly high LD (D' < 1) and the SNPs can't substitute each other were [252 - 522; 252 - 529; 252 - 530; 252 - 539; 287 - 522; 287 - 529; 287 - 530; 287 - 539; 369 - 522; 369 - 529; 369 - 530; 369 - 539; (D' = 0.777 and r^2 = 0.36)] and [347 - 522; 347 - 529; 347 - 530; 347 – 539 (D' = 0.632 and r^2 = 0.16)]. Some pairwise SNPs of 16SrRNA haplotypes showed negative D' values (168 – 171, 168 - 252; 168 - 287; 168 – 369), and these might be the indicator of inbreeding. [38] Some pairwise SNPs of D-Loop haplotypes showed moderate LD were [108 - 235; 233 - 235; 235 - 351; 356 - 517; 561 - 562; 562 – 564 (D' = 0.464 and r^2 = 0.21)]. D-Loop SNPs with fairly high LD (D' < 1) were [135 – 228 (D' = 0.700 and r^2 = 0.20)].

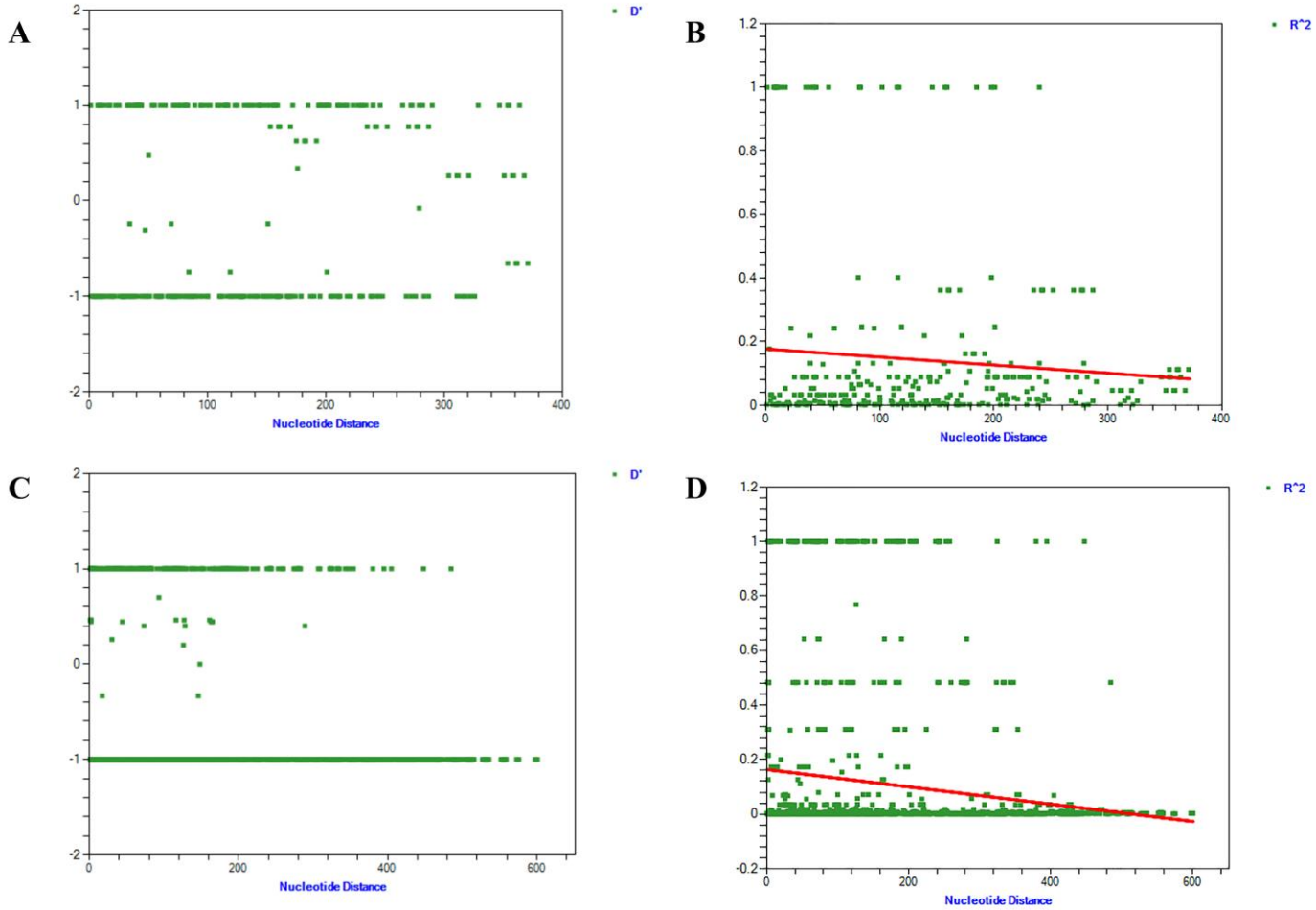

Figure 11
Linkage disequilibrium (LD) scatter plot for 16SrRNA and D-Loop haplotypes.

(A) The D' plot for 16SrRNA haplotypes. (B) The LD of 16SrRNA haplotypes was measured as a function of the distance between polymorphic sites based upon the r^2 statistical measurement (Y-axis). Significant comparisons were determined by Fisher's exact test and Bonferroni correction. X-axis measurements indicated the nucleotide distance across the 16SrRNA gene. The logarithmic trend line, indicating the magnitude and

extent of LD, was fitted to data by the DnaSP program. (C) The D' plot for D-Loop haplotypes. (D) The LD of D-Loop haplotypes was measured as a function of the distance between polymorphic sites based upon the r^2 statistical measurement (Y-axis). Significant comparisons were determined by Fisher's exact test and Bonferroni correction. X-axis measurements indicated the nucleotide distance across the D-Loop sequences. The logarithmic trend line, indicating the magnitude and extent of LD, was fitted to data by the DnaSP program.

Table 8
The LD analysis results of 16SrRNA and D-Loop haplotypes.

Site1 - Site2	r^2 value	Fisher's Exact Test	Chi-square Test
16SrRNA Haplotypes			
168 – 171	0.18	0.057	4.441*
168 - 252; 168 - 287; 168 – 369	0.25	0.033*	6.173*
171 - 252; 171 - 287; 171 - 369	0.40	0.003**	10.048**
175 - 347; 347 - 386; 347 - 486	0.22	0.160	5.469*
252 - 287; 252 - 369; 287 - 369; 522 - 529; 522 - 530; 522 - 539; 529 - 530; 529 - 539; 530 – 539	1.000	0.000***B	25.000***B
252 - 347	0.24	0.026*	6.061*
252 - 522; 252 - 529; 252 - 530; 252 - 539; 287 - 522; 287 - 529; 287 - 530; 287 - 539; 369 - 522; 369 - 529; 369 - 530; 369 - 539	0.36	0.007**	9.035**
257 - 299; 257 - 312; 257 - 414; 257 - 458; 257 - 497; 299 - 312; 299 - 414; 299 - 458; 299 - 497; 312 - 414; 312 - 458; 312 - 497; 414 - 458; 414 - 497; 458 - 497	1.000	0.040*	25.000***B
347 - 369	0.24	0.026*	6.061*
347 - 522; 347 - 529; 347 - 530; 347 – 539	0.16	0.081	4.046*
D-Loop Haplotypes			
7 - 50; 72 - 520; 108 - 233; 108 - 351; 233 - 351	1.000	0.002**	30.000***B
32 - 169; 32 - 174; 32 - 234; 32 - 236; 32 - 241; 32 - 243; 32 - 274; 32 - 276; 32 - 357; 32 - 412; 32 - 427; 112 - 127; 169 - 174; 169 - 234; 169 - 236; 169 - 241; 169 - 243; 169 - 274; 169 - 276; 169 - 357; 169 - 412; 169 - 427; 174 - 234; 174 - 236; 174 - 241; 174 - 243; 174 - 274; 174 - 276; 174 - 357; 174 - 412; 174 - 427; 182 - 184; 182 - 237; 182 - 257; 184 - 237; 184 - 257; 190 - 273; 234 - 236; 234 - 241; 234 - 243; 234 - 274; 234 - 276; 234 - 357; 234 - 412; 234 - 427; 236 - 241; 236 - 243; 236 - 274; 236 - 276; 236 - 357; 236 - 412; 236 - 427; 237 - 257; 241 - 243; 241 - 274; 241 - 276; 241 - 375; 241 - 412; 241 - 427; 243 - 274; 243 - 276; 243 - 357; 243 - 412; 243 - 427; 272 - 386; 272 - 515; 274 - 276; 274 - 357; 274 - 412; 274 - 427; 276 - 357; 276 - 412; 276 - 427; 357 - 412; 357 - 427; 386 - 515; 412 - 427; 566 - 605; 566 - 608; 605 - 608	1.000	0.033*	30.000***B
32 - 228; 169 - 228; 174 - 228; 228 - 234; 228 - 236; 228 - 241; 228 - 243; 228 - 274; 228 - 276; 228 - 357; 228 - 412; 228 - 427	0.17	0.167	5.172*
32 - 354; 161 - 272; 161 - 386; 161 - 515; 169 - 354; 174 - 354; 190 - 191; 191 - 272; 191 - 273; 191 - 386; 191 - 515; 234 - 354; 236 - 354; 241 - 354; 243 - 354; 274 - 354; 276 - 354; 354 - 357; 354 - 412; 354 - 427	0.31	0.100	9.310**
32 - 356; 32 - 517; 169 - 356; 169 - 517; 174 - 356; 174 - 517; 182 - 517; 184 - 517; 190 - 356; 234 - 356; 234 - 517; 235 - 272; 235 - 386; 235 - 515; 236 - 356; 236 - 517; 237 - 517; 241 - 356; 241 - 517; 243 - 356; 243 - 517; 257 - 517; 273 - 356; 274 - 356; 274 - 517; 276 - 356; 276 - 517; 356 - 357; 356 - 412; 356 - 427; 357 - 517; 412 - 517; 427 - 517; 559 - 561; 559 - 562; 564 - 566; 564 - 605; 564 - 608	0.48	0.067	14.483***
72 - 354; 108 - 161; 161 - 233; 161 - 235; 161 - 351; 354 - 520	0.64	0.007**	19.286***B
102 - 135	0.31	0.008**	9.231**
102 - 208	0.15	0.100	4.615*
102 - 228	0.77	0.000***	23.077***B
108 - 235; 233 - 235; 235 - 351; 356 - 517; 561 - 562; 562 – 564	0.21	0.131	6.467*
135 – 228	0.20	0.031*	5.880*
208 - 228	0.20	0.042*	6.000*

**** 0.01 <P < 0.05; ** 0.001 < P< 0.01; *** P < 0.001; B, significant by the Bonferroni procedure.***

DISCUSSION

Mitochondrial DNA polymorphism data provides vital information about the intraspecific population genetic diversity, and also reveals the molecular phylogeographic relationships and intraspecific genealogical relationship. In this study, population genetic diversity among *P. monodon* samples using D-Loop and 16SrRNA haplotypes were measured. Among D-Loop and 16SrRNA sequences, the average frequency of A and T nucleotide was higher than G and C, as per the common features for Arthropods and

Crustaceans mitochondrial genome.[5,39] The ML estimate of transition/transversion bias suggested that the rate of transitions was higher than the rate of transversions, and a very high level of transitions were observed in 16SrRNA gene segments with a higher percentage of invariable sites. The MCL estimate of overall transition/transversion bias of 16SrRNA genes was highest. The highest haplotype diversity was observed in the mitochondrial control region or D-Loop sequences. Previously, Benzie [40] had reported the maximum haplotype diversity of *P. monodon* and that was 0.682 ± 0.002, and Klinbunga *et al* [41] had reported the maximum nucleotide diversity of *P. monodon* and that was 0.00334 ± 0.00003. This study evaluated a recent haplotype diversity of *P. monodon* among cultured pond samples which were higher than the previously reported diversity. The higher level of genetic diversity has been observed among cultured populations that might be the result of highly infectious virus (especially White spot syndrome virus) outbreaks each year and horizontal gene transfer by pathogens. Many decapods crustaceans are characterized by a high level of genetic diversity due to high mutation rate in the mitochondrial control region.[7] Time-dependent genetic monitoring and evaluation of black tiger shrimp are necessary to identify any negative forces on genetic diversity due to aquaculture.[42-43] The S and Eta were higher in D-Loop haplotypes as it was the most hypervariable region among other mitochondrial DNA segment as well as due to very high rate of mutations. Moreover, the number of transitions, number of transversions, and the number of substitutions of D-Loop haplotypes were higher than 16SrRNA haplotypes. The results of this study suggested that D-Loop mtDNA segments were highly polymorphic in nature. The Median-Joining network of 16SrRNA and D-Loop haplotypes showed the star-like intraspecific genealogical relationship. The number of substitutions was represented as a unique number for each site, which crossed the linking lines. The most frequently observed and common D-Loop haplotype was haplotype 15. The most frequently observed 16SrRNA haplotypes were haplotype 1 and haplotype 4. From the result section and based on the cluster, matrix, network, and phylogenetic analysis; it was very clear that most unique haplotypes were haplotype 8 (D-Loop) and haplotype 3 (16SrRNA). These two haplotypes are rare haplotypes and contain rare mutations and polymorphisms. The LD analysis indicated the occurrence and integration of rare mutations, and which is responsible for the generation of rare haplotypes.

CONCLUSION

The mitochondrial control region (D-Loop) was found to be highly polymorphic and could prove to be a potential marker for the study of intraspecific population diversity study of *P. monodon*. The 16SrRNA would prove to be an excellent marker for phylogenetic studies in *P. monodon*. This study has a great significance to maintain high-quality pond-reared stock diversity, and also in accomplishing selective breeding. A high level of genetic diversity of *P. monodon* would be helpful in preventing deadly diseases as a high level of mutations which acts as a buffer against detrimental diseases.

FUNDING ACKNOWLEDGMENTS

The authors acknowledge the Ministry of Earth Sciences, Government of India for funding support for this research work. Mr. Debabrata Mondal is grateful to DST INSPIRE, Government of India for his fellowship. The authors would like to thank Mr. Ranjit Kumar Das for technical assistance in sample preparation and handling of lab wares.

AUTHORS CONTRIBUTION STATEMENT

D. M. designed the work, performed experimental works, analyzed the data, acquisition of the data, interpreted the results, wrote the manuscript, and collected samples. N. M. designed the work, analyzed the data, wrote the manuscript, and provided funds for this study. Both authors have given final approval for the publication of the article.

CONFLICT OF INTEREST

Conflict of interest declared none.

REFERENCES

1. Benzie JAH. Genetic and reproduction research on giant tiger prawns (*Penaeus monodon*): pond reared spawners achieved in Australia. Aust Biotech. 1994;4:222–4.
2. Kumar N, Lakra WS, Majumdar KC, Goswami M, Ravinder K. Genetic diversity in the Indian population of *Penaeus monodon* (Fabricius, 1798) as revealed by mtDNA sequence analysis. Aquac Res. 2007;38(8):862–9. DOI: 10.1111/j.1365-2109.2007.01740.x
3. Cunningham CW, Blackstone NW, Buss LW. Evolution of king crabs from hermit crab ancestors. Nature. 1992;355(6360):539–42. DOI: 10.1038/355539a0
4. Chu KH, Li CP, Tam YK, Lavery S. Application of mitochondrial control region in population genetic studies of the shrimp Penaeus. Mol Ecol Notes. 2003;3(1):120–2. DOI: 10.1046/j.1471-8286.2003.00376.x
5. Wilson K, Cahill V, Ballment E, Benzie J. The Complete Sequence of the Mitochondrial Genome of the Crustacean *Penaeus monodon*: Are Malacostracan Crustaceans More Closely Related to Insects than to Branchiopods? Mol Biol Evol. 2000;17(6):863–74. DOI: 10.1093/oxfordjournals.molbev.a026366
6. Mcmillen-Jackson AL, Bert TM. Disparate patterns of population genetic structure and population history in two sympatric penaeid shrimp species (*Farfantepenaeus aztecus* and *Litopenaeus setiferus*) in the eastern United States. Mol Ecol. 2003;12(11):2895–905. DOI: 10.1046/j.1365-294x.2003.01955.x
7. McMillen-Jackson AL, Bert TM. Genetic Diversity in the mtDNA Control Region and Population Structure in the Pink Shrimp *Farfantepenaeus Duorarum*. J Crustac Biol. 2004;24(1):101–9.

DOI: 10.1651/c-2372

8. Murphy NP, Austin CM. Molecular Taxonomy and Phylogenetics of Some Species of Australian Palaemonid Shrimps. J Crustac Biol. 2003;23(1):169–77. DOI: 10.1163/20021975-99990324
9. Khedkar GD, Reddy AC, Ron TB, Haymer D. High levels of genetic diversity in *Penaeus monodon* populations from the east coast of India. Springerplus. 2013;2(1). DOI: 10.1186/2193-1801-2-671
10. Tzeng T-D, Yeh S-Y, Hui C-F. Population genetic structure of the kuruma prawn (*Penaeus japonicus*) in East Asia inferred from mitochondrial DNA sequences. ICES J Mar Sci. 2004;61(6):913–20. DOI: 10.1016/j.icesjms.2004.06.015
11. Chakrabarty U, Dutta S, Mallik A, Mandal N. White spot syndrome virus (WSSV) and prevalence of disease resistance in a commercially cultured population of *Penaeus monodon* Fabricius, 1798 (Decapoda, Dendrobranchiata). Crustaceana. 2014;87(14):1593–605. DOI: 10.1163/15685403-00003382
12. Chakrabarty U, Dutta S, Mallik A, Mondal D, Mandal N. Identification and characterisation of microsatellite DNA markers in order to recognise the WSSV susceptible populations of marine giant black tiger shrimp, *Penaeus monodon*. Vet Res. 2015;46(1). DOI: 10.1186/s13567-015-0248-2
13. Sambrook J, Russel DW. Molecular Cloning – A laboratory Manual. 3rd ed. In: Argentine J, editor. New York: Cold Spring Harbor Laboratory Press; 2001. 69–611 p.
14. Palumbi SR Benzie JAH. Large mitochondrial DNA differences between morphologically similar penaeid shrimp. Mol Mar Biol Biotechnol. 1991;1(1):27–34.
15. Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. InNucleic acids symposium series; 1999. 95–8 p. Available from: http://jwbrown.mbio.ncsu.edu/JWB/papers/1999 Hall1.pdf
16. Nicholas KB, Nicholas HBJ. GeneDoc: a tool for editing and annotating multiple sequence alignments. 1997. Available from: https://www.scienceopen.com/document?vid=c8a87cd1-255f-4129-802b-d2382bb0fb37
17. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990;215(3):403–10. DOI: 10.1016/s0022-2836(05)80360-2
18. Altschul S, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 1997;25(17):3389–402. DOI: 10.1093/nar/25.17.3389
19. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 2004;32(5):1792–7. DOI: 10.1093/nar/gkh340
20. Tamura K, Nei M, Kumar S. Prospects for inferring very large phylogenies by using the neighbor-joining method. Proc Natl Acad Sci. 2004;101(30):11030–5. DOI: 10.1073/pnas.0404206101
21. Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol Biol Evol. 2016;33(7):1870–4. DOI: 10.1093/molbev/msw054
22. Rozas J, Ferrer-Mata A, Sánchez-DelBarrio JC, Guirao-Rico S, Librado P, Ramos-Onsins SE, et al. DnaSP 6: DNA Sequence Polymorphism Analysis of Large Data Sets. Mol Biol Evol. 2017;34(12):3299–302. DOI: 10.1093/molbev/msx248
23. Excoffier L, Lischer HEL. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Resour. 2010;10(3):564–7. DOI: 10.1111/j.1755-0998.2010.02847.x
24. Hammer Ø, Harper DAT, Ryan PD. PAST: paleontological statistics software package for education and data analysis. Palaeontol Electron. 2001;4(1):9. Available from: https://paleo.carleton.ca/2001_1/past/past.pdf
25. Dereeper A, Guignon V, Blanc G, Audic S, Buffet S, Chevenet F, et al. Phylogeny.fr: robust phylogenetic analysis for the non-specialist. Nucleic Acids Res. 2008;36(Web Server):W465–9. DOI: 10.1093/nar/gkn180
26. Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O. New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. Syst Biol. 2010;59(3):307–21. DOI: 10.1093/sysbio/syq010
27. Katoh K, Misawa K, Kuma K, Miyata T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. 2002;30(14):3059–66. DOI: 10.1093/nar/gkf436
28. Anisimova M, Gascuel O. Approximate Likelihood-Ratio Test for Branches: A Fast, Accurate, and Powerful Alternative. Syst Biol. 2006;55(4):539–52. DOI: 10.1080/10635150600755453
29. Chevenet F, Brun C, Bañuls A-L, Jacq B, Christen R. TreeDyn: towards dynamic graphics and annotations for analyses of trees. BMC Bioinformatics. 2006;7(1). DOI: 10.1186/1471-2105-7-439
30. Robinson O, Dylus D, Dessimoz C. Phylo.io: Interactive Viewing and Comparison of Large Phylogenetic Trees on the Web. Mol Biol Evol. 2016;33(8):2163–6. DOI: 10.1093/molbev/msw080
31. Smith SM, Yuan Y, Doust AN, Bennetzen JL. Haplotype Analysis and Linkage Disequilibrium at Five Loci in *Eragrostis tef*. G3: Genes|Genomes|Genetics. 2012;2(3):407–19. DOI: 10.1534/g3.111.001511
32. Flint-Garcia SA, Thornsberry JM, Buckler ES. Structure of Linkage Disequilibrium in Plants. Annu Rev Plant Biol. 2003;54(1):357–74. DOI: 10.1146/annurev.arplant.54.031902.134907

33. Hedrick PW. Gametic disequilibrium measures: proceed with caution. Genetics. 1987;117(2):331–41. Available from: https://www.genetics.org/content/117/2/331.short
34. Wier BS. Genetic Data Analysis II. 2 sub edit. Sinauer Associates: Sunderland; 1996.
35. Weiss KM, Clark AG. Linkage disequilibrium and the mapping of complex human traits. Trends Genet. 2002;18(1):19–24. DOI: 10.1016/s0168-9525(01)02550-1
36. Carlson CS, Eberle MA, Kruglyak L, Nickerson DA. Mapping complex disease loci in whole-genome association studies. Nature. 2004;429(6990):446–52. DOI: 10.1038/nature02623
37. Kimura M. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J Mol Evol. 1980;16(2):111–20. DOI: 10.1007/bf01731581
38. Alvarenga AB, Rovadoscki GA, Petrini J, Coutinho LL, Morota G, Spangler ML, et al. Linkage disequilibrium in Brazilian Santa Inês breed, *Ovis aries*. Sci Rep. 2018;8(1). DOI: 10.1038/s41598-018-27259-7
39. Cook CE. The complete mitochondrial genome of the stomatopod crustacean *Squilla mantis*. BMC Genomics. 2005;6(1):105. Available from: https://bmcgenomics.biomedcentral.com/articles/10.1186/1471-2164-6-105
40. Benzie JAH. Population genetic structure in penaeid prawns. Aquac Res. 2000;31(1):95–119. DOI: 10.1046/j.1365-2109.2000.00412.x
41. Klinbunga S, Penman DJ, McAndrew BJ, Tassanakajon A, Jarayabhand P. Genetic variation, population differentiation, and gene flow of the giant tiger shrimp (*P. monodon*) inferred from mtDNA RFLP data. In: Flegel TW editor. Advances in Shrimp Biotechnology, National Center for Genetic Engineering and Biotechnology, Bangkok, Thailand; 1998. p.51-9.
42. Naylor RL, Goldburg RJ, Primavera JH, Kautsky N, Beveridge MCM, Clay J, et al. Effect of aquaculture on world fish supplies. Nature. 2000;405(6790):1017–24. DOI: 10.1038/35016500
43. Benzie JAH. Integration of quantitative and molecular genetics in shrimp breeding. Asian Fish Sci. 2010;23(4):497–523. Available from: https://www.cabdirect.org/cabdirect/abstract/20113203970