Hexapoda 30&31: 2(2023-2024): 1–12
DOI: 10.55446/hexa.2023-2024.05

OPEN ACCESS

GENETIC VARIABILITY AMONG THE GEOGRAPHIC POPULATION OF *LEPTOCORISA ORATORIA* (HEMIPTERA: ALYDIDAE) IN INDIA AS INFERRED FROM MITOCHONDRIAL COI GENE SEQUENCE

S. Lahari, C. M. Kalleshwaraswamy*, B. C. Hanumanthaswamy and N. L. Naveena

Insect Systematics Laboratory, Department of Entomology, College of Agriculture,
Keladi Shivappa Nayaka University of Agricultural and Horticultural Sciences,
Shivamogga – 577204, Karnataka, India
**E-mail*: kalleshwaraswamycm@uahs.edu.in (Corresponding author)

ABSTRACT

The paddy gundhi bug, *Leptocorisa oratoria* Fab. is a common pest species found throughout India. The genetic diversity studies of these bugs from India were conducted during 2022-2023. Morphologically identified representative specimens were selected for DNA isolation. The *L. oratoria* from 14 locations in India were analysed for variations in the *mtCOI* gene to identify the genetic diversity. The nucleotide variations showed maximum variations at 21 sites in the population collected from Odisha. Variations in the amino acids revealed the maximum composition of Leucine in all the 14 sequences, with higher A+T bias. Haplotype analysis yielded 12 haplotypes from 14 sequences of *L. oratoria.* The analysis of haplotype divergence, assessed through Tajima's D statistics indicated the presence of low-frequency polymorphisms. The genetic distance and nucleotide diversity were low probably due to the absence of geographical barrier, gene flow and migration. This is the first attempt to analyse molecular data for this well-known and serious pest of paddy from India.

KEYWORDS: Genetic Diversity Analysis, Haplotype Analysis, *Leptocorisa oratoria*, Mitochondrial Cytochrome C Oxidase I, Molecular Characterization

(Received: 01 October 2024; Accepted: 20 October 2024; Published: 01 November, 2024)

INTRODUCTION

Rice bugs are one of the few pests that feed directly on developing rice spikelets in the field causing a considerable economic loss. In addition, they release a pungent odour that is imparted to the rice (Dutta and Roy, 2016). Rice bugs usually concentrate on small-scale upland rice fields and are also common in rainfed lowland rice environments. Highly variable populations and damage of rice bugs are seen as the infestation is only at a particular period of crop growth. Among various rice bug species *Leptocorisa* spp., known as paddy gundhi bugs, are commonly found in the Indian subcontinent. *Leptocorisa* (Latreille) belongs to the superfamily Coreoidea, family Alydidae and tribe Leptocorisini. Commonly referred to as slender rice bugs, these occur in almost all rice-growing regions of the world. Adults and nymphs suck the plant sap and attack the rice grains, starting from the flowering stage and rendering the grains sterile. The milky stage damage results in empty or half-filled grains, hampering the quality. Grains thus infested become emptied or pecked (Patel et al., 2006).

Fourteen species of *Leptocorisa* genus are reported worldwide of which nine species are considered to cause severe injury in rice plants (Hasegawa, 1971, Litsinger et al., 2015). Two species, *L.* Cobblah *oratoria* (Fabricius) and *L. acuta* (Thunberg) are commonly found to cause considerable damage to rice (Cobblah and Hollander, 1992; Hosamani et al., 2009). Lahari et al. (2024) sampled and studied the species diversity across India and provided the illustrated identification key for three species comprising *L. oratoria* Fabricius, *Leptocorisa acuta* Thunberg and *Leptocorisa lepida* Breddin. *L. oratoria* is predominantly found feeding on rice (Kumar and Goswami, 2020; Lahari et al., 2024). However, the geographic variability of this important species is not studied in India. There is much scope to use DNA sequences for understanding ecology and thereby using them in pest management (Shashank et al., 2022). Understanding the genetic diversity of this species helps to understand the reasons for variation in terms of geographic location, environmental factors and population genetic structure.

Citation: Lahari S, Kalleshwaraswamy C M, Hanumanthaswamy B C and Naveena N L. 2024. Genetic Variability among the geographic population of *Leptocorisa oratoria* (Hemiptera: Alydidae) in India as inferred from mitochondrial COI gene sequence. Hexapoda: 30&31: 2(2023-2024) : 1-12; DOI: 10.55446/hexa.2023-2024.05

MATERIALS AND METHODS

Insect collection

The specimen collection was done by using sweep nets in the paddy fields and also by hand collection. Attempts were also made to obtain preserved specimens through requests from different entomological organizations across India and specimens were procured. A representative specimen from a selected locality was selected for the present study comprising of 14 sites.

DNA isolation

Specimens after being morphologically identified as *L. oratoria* were molecularly analyzed by following the modified CTAB method (Hunt, 1997). The specimens were washed in distilled water and crushed using the extraction buffer [Five per cent SDS, 50mM Tris base and 10mM EDTA] in the autoclaved mortar and incubated at 65ºC for 10 min in the hot water bath. The heated suspension was treated with 10 per cent CTAB and 5M NaCl to be again incubated at 65ºC for 10 min. The samples after incubation were cooled and centrifuged at 11,000 rpm, 4 ºC for 10 min with the addition of an equal volume of chloroform: isoamyl alcohol (24:1). The supernatant was carefully pipetted out and transferred to fresh microcentrifuge tube, to which pre-chilled isopropanol and 3M Sodium acetate was added and stored at -20 °C overnight. On the following day, samples were centrifuged at 11,000 rpm for 10 min followed by 70 per cent ethyl alcohol wash. Later, the air-dried pellet was dissolved in Tris EDTA (Tris-HCl 10Mm and EDTA-Na_2 1Mm) buffer and stored at -20°C for future use.

Quantification of DNA

The concentration of the DNA was estimated based on the intensity of ethidium bromide fluorescence against DNA concentration standards. Agarose gel of 0.8 per cent was used to confirm the integrity of the DNA. The absorbance readings in the spectrophotometer (Nanodrop, Thermo Scientific) were used to assess the quality and quantity of the extracted genomic DNA.

PCR amplification

A fragment of ~658 base pairs of the *mtCOI* gene fragment was amplified using the universal primers LCO1490 and HCO2198 (Folmer et al., 1994). The PCR (polymerase chain reaction) amplification was carried out in a Master Cycler (Applied Biosystem Proflex PCR system) with a reaction volume of 30 µl comprising 6 µl of template DNA/genomic DNA (25-50 ng), 3 µl of Taq polymerase buffer with $MgCl_2$ (10X), 2.4 µl of dNTPs (2 mM), 0.9 µl of forward primer (10 pmol), 0.9 µl of reverse primer (10 pmol), 0.9 µl Taq polymerase enzyme (1U/µl) and 15.9 µl of nuclease-free water was used for amplification of the specific gene in the thermal cycler. The PCR conditions for the amplification of the specific region of the *mtCOI* followed a single cycle of initial denaturation at 94°C for 5 min, 35 cycles of denaturation at 94°C for 40 s, annealing at 51°C for 45 s, initial extension at 72°C for 1 minute, a single cycle of final extension at 72°C for 10 min and a final hold at 4°C. PCR included both negative control and positive control. PCR amplicons after gel electrophoresis on a 1.5 per cent agarose gel stained with ethidium bromide were visualised in a gel documentation system (Major Science: SmartView Pro 2100 Imager System, CCD UVCI-2100).

Sequencing of the mtCOI gene and phylogenetic analysis

The amplified PCR product of the *mtCOI* gene was sent for sequencing at Barcode Biosciences, Bengaluru. The chromatograms of the gene sequences obtained were edited in Bioedit version 7.2.5 to remove ambiguous bases. The edited sequences were checked for homology by blasting in NCBI-BLAST and were submitted to NCBI's GenBank to obtain the accession number of gundhi bug species for different locations.

Haplotype analysis and genetic variations

Possible haplotypes in the study were analysed using the algorithms from DnaSP ver. 5.0 software. Statistics like the number of variable sites, number of haplotypes, haplotype diversity and nucleotide diversity were calculated and computed in the DnaSP software. The haplotype network tree was computed using Popart v1.7 (Population Analysis with Reticulate Trees) software (Leigh and Bryant, 2015) by dividing the 14 sequences in the study into two groups, North and South, based on their geographical similarities. Median-joining network was created to understand the intraspecific phylogeny within these 14 sequences (Bandelt et al*.,* 1999).

Analysis of Molecular variance (AMOVA) as given by Excoffier et al*.,* 1992 was conducted using the DnaSP ver. 5.0 software (Librado and Rozas, 2009) which also computes the possible gene flow and genetic distance between the groups and haplotypes based on the *mtCOI* sequence data.

RESULTS AND DISCUSSION

DNA barcoding

Morphological identification can be confirmed using molecular analysis. In the current study *mitochondrial*

cytochrome oxidase I (*mtCOI)* gene was used for the identification and interpretation of the evolutionary relationship of the *L. oratoria* across India. Modified CTAB method to isolate the genomic DNA from *L. oratoria* from fourteen different locations from the states of Karnataka, Kerala, Tamil Nadu, Andhra Pradesh, West Bengal and Gujarat and accession numbers were obtained. These accession numbers are OR295916, OR297688, OR296321, OR296707, OR333540, OR294965, OR319147, OR298283, OR322045, OR342097, OR342099, OR342102 and OR342108. Using the universal primers LCO1490 and HCO2140 approximately ~658 base pair regions of the *mtCOI* gene were amplified and sequenced using Sanger dideoxy sequencing method. Ambiguous bases, insertions and deletions were removed using Bioedit version 7.2.5. Accession numbers were obtained by submitting these sequences to NCBI's GenBank.

Nucleotide polymorphism

The sequences acquired ranged in length from around 580 to 640 base pairs. Nucleotide polymorphism, variations at different sites within the fourteen sequences were examined (Figure 1). The analysis revealed that 520 out of 557 sites were conserved, while 37 nucleotide sites displayed variability, leading to a total of 38 mutation sites.

At position 144 within the sequence, a consistent substitution of cytosine with thymine was observed across all populations, with the exception of UAHS, Bhyranahalli, and Mandya. Similarly, at site 546, all sequences except UAHS, TNLO, Pattambi, Bantwal, Vellayani, Rayalcheruvu, and Majilabail exhibited a substitution of cytosine in place of thymine. The maximum nucleotide differences were detected in the Cuttack population from Odisha, with variations present at 21 sites. These sites were situated at positions 10, 35, 41, 80, 84, 87, 90, 93, 144, 202, 213, 216, 218, 223, 238, 253, 254, 255, 256, 345, and 546.

Within the 557 bps nucleotide sequences, specific sites were identified where parsimony informative variations occurred, specifically at positions 144, 345, 498, and 546. These regions signify potential areas of distinction among populations within the context of a conserved gene. Additionally, a total of 10 conserved sequence sets were recognized across the 557 base pairs, which were consistently present in all 14 populations. These conserved sequence sets exhibited varying lengths. The minimum segment length observed was 15 base pairs, spanning positions 65 to 79, while another segment spanned 76 base pairs, encompassing positions 259 to 335 (Table 1).

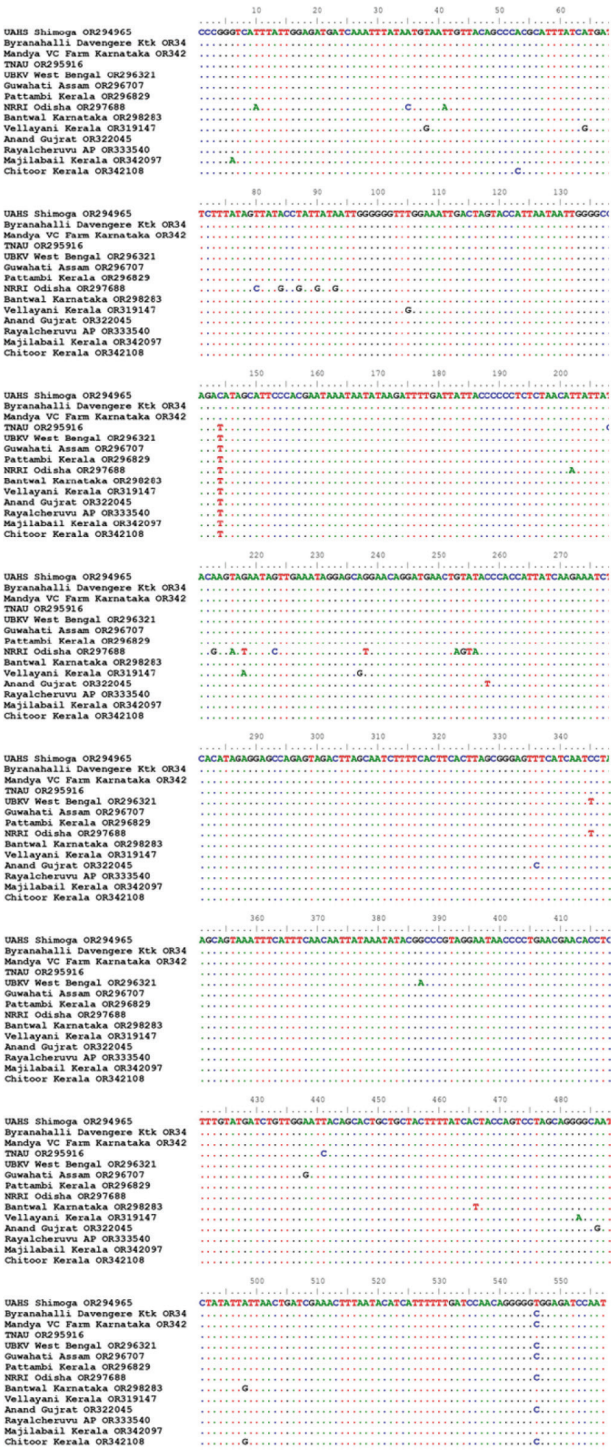

Fig. 1. Nucleotide dissimilarity in the mtCOI gene in the different populations of *L. oratoria.*

Amino acid differences

Amino acid compositions derived from the sequences in the study (Table 2) showed amino acid cysteine to be absent across all sequences. Lysine was present solely in the Vellayani sequence, accounting

Table 1. Conserved nucleotide sequence segments in 14 populations of *L. oratoria*

Regions	Position	Consensus sequences
1	11-34	TTATTGGAGATGATCAAATTTATA
2	65 -79	TGATTTTCTTTATAG
3	106- 143	GGAAATTGACTAGTACCATTAATAATTGGGGCCCCAGA
4	145-201	ATAGCATTCCCACGAATAAATAATATAAGATTTTGATTATTACCCCCCTCTCTAACA
5	259-335	CCACCATTATCAAGAAATCTATCACATAGAGGAGCCAGAGTAGACTTA GCAATCTTTTCACTTCACTTAGCGGGAGT
6	346-386	CTAGGAGCAGTAAATTTCATTTCAACAATTATAAATATACG
7	388-437	CCCGTAGGAATAACCCCTGAACGAACACCTCTATTTGTATGATCTGTTGG
8	442-465	ACAGCACTGCTGCTACTTTTATCA
9	467-482	TACCAGTCCTAGCAGG
10	499-545	TTAACTGATCGAAACTTTAATACATCATTTTTTGATCCAACAGGGGG

Note: A-Adenine, C-Cytosine, G-Guanine, T-Thymine.

for approximately 0.541 per cent, while it was absent in the remaining sequences. Leucine displayed the highest percentage range, spanning from 13.514 per cent to 12.973 per cent, across the sequences. Conversely, glutamine exhibited the lowest percentage at 0.541 per cent. Upon calculation, all 14 sequence mean amino acid compositions converged around 185 units. This analysis offered insights into the variation in amino acid compositions within the *L. oratoria* sequences, underscoring the significance of specific amino acids in the genetic makeup of these bugs.

Molecular phylogeny of *L. oratoria*

The analysis focused on the *mtCOI* gene sequence, comprising around 557 base pairs, to assess genetic divergence. A set of 14 *L. oratoria* specimens from different geographic locations in India were examined to understand their interrelationships. The nucleotide frequencies within the sequence were computed following sequence alignment using CLUSTAL W in Bioedit version 7.2.5. The composition of nucleotides within the sequence was as follows: Adenine (A) accounted for 32.47 per cent, Thymine (T) for 31.02 per cent, Cytosine (C) for 19.38 per cent, and Guanine (G) for 17.13 per cent. Notably, the average A+T content was higher at 63.521 per cent compared to G+C content at 36.541 per cent, (Table 3). This observation indicated a prevailing bias towards A and T nucleotides in the *mtCOI* gene sequence.

This study also examined the probabilities of transition and transversion substitutions using maximum likelihood estimates based on a substitution model. The analysis revealed that the overall likelihood of transition substitutions was greater than that of transversion substitutions. The overall transition-to-transversion bias was quantified as 4.885 (Table 4). Further, transition and transversion substitutions in purines were more pronounced at 21.153 compared to pyrimidines at 0.738. These analyses provide insights into the patterns of nucleotide substitutions, shedding light on the genetic dynamics within the *L. oratoria* population.

The pairwise genetic differences were calculated for the 14 sequences (Table 5). The study revealed an average mean genetic distance of approximately 1.145 per cent with a standard error of 0.00198. Notably, the smallest genetic distance was observed between the Mandya and Bhyranahalli populations, indicating a pronounced similarity between these two species. In contrast, the most significant genetic distance of 4.629 per cent was identified between the Cuttack and Vellayani populations, indicating substantial nucleotide sequence variability between the two. This disparity highlights the genetic diversity within these *Leptocorisa* species, reflecting their varying evolutionary trajectories and genetic backgrounds.

Haplotype studies and genetic variations

The presence of haplotypes was assessed across the 14 sequences, utilising DnaSP Ver. 6.12.03. The analysis identified a total of 12 distinct haplotypes (Table 6). Among these, Hap_7, comprising populations from Cuttack, Odisha, exhibited the highest degree of segregation. Following this, Hap_9 from Vellayani, Kerala, and Hap_10 from Anand, Gujarat displayed significant segregation (Figure 2).

Despite the grouping of sequences based on their geographical collection locations into South and North India, the analysis uncovered minor nucleotide

Table 2. Percent composition of amino acids in the 14 sequences of *L. oratoria*

Sequences vs amino acids	Ala	Cys	Asp	Glu	Phe	Gly	His	Ile	Lys	Leu	Met	Asn	Pro	Gln	Arg	Ser	Thr	Val	Trp	Tyr	Total
UAHS Shimoga OR294965	6.486	0.000	3.784	1.081	7.027	9.730	1.622	7.568	0.000	12.973	7.568	4.865	8.108	0.541	2.162	9.189	7.027	7.027	2.162	1.081	185
Byranahalli OR342102	6.486	0.000	3.784	1.081	7.027	9.730	1.622	7.568	0.000	12.973	7.568	4.865	8.108	0.541	2.162	9.189	7.027	7.027	2.162	1.081	185
Mandya VC Farm Karnataka OR342099	6.486	0.000	3.784	1.081	7.027	9.730	1.622	7.568	0.000	12.973	7.568	4.865	8.108	0.541	2.162	9.189	7.027	7.027	2.162	1.081	185
TNLO OR295916	6.486	0.000	3.784	1.081	7.027	9.730	1.622	7.568	0.000	12.973	7.568	4.865	8.108	0.541	2.162	9.189	7.027	7.027	2.162	1.081	185
UBKV West Bengal OR296321	6.486	0.000	3.784	1.081	7.027	9.730	1.622	7.568	0.000	12.973	7.568	4.865	8.108	0.541	2.162	9.189	7.027	7.027	2.162	1.081	185
Guwahati Assam OR296707	6.486	0.000	3.784	1.081	7.027	9.730	1.622	7.568	0.000	12.973	7.568	4.865	8.108	0.541	2.162	9.189	7.027	7.027	2.162	1.081	185
Pattambi Kerala OR296829	6.486	0.000	3.784	1.081	7.027	9.730	1.622	7.568	0.000	12.973	7.568	4.865	8.108	0.541	2.162	9.189	7.027	7.027	2.162	1.081	185
ODILO OR297688	7.027	0.000	3.784	1.081	6.486	9.189	1.622	7.027	0.000	12.973	9.189	5.405	8.108	0.541	2.162	9.189	7.568	5.405	2.703	0.541	185
Bantwal Karnataka OR298283	6.486	0.000	3.784	1.081	7.027	9.730	1.622	7.568	0.000	12.973	7.568	4.865	8.108	0.541	2.162	9.189	7.027	7.027	2.162	1.081	185
Vellayani Kerala OR319147	6.486	0.000	3.784	1.081	6.486	10.270	1.622	7.568	0.541	13.514	7.027	4.865	8.108	0.541	2.162	8.649	7.027	7.027	2.162	1.081	185
Anand Gujrat OR322045	6.486	0.000	3.784	1.081	7.027	9.730	1.622	7.568	0.000	12.973	7.568	4.865	8.108	0.541	2.162	9.189	7.027	7.027	2.162	1.081	185

Table 2. Continued...

Sequences vs amino acids	Ala	Cys	Asp	Glu	Phe	Gly	His	Ile	Lys	Leu	Met	Asn	Pro	Gln	Arg	Ser	Thr	Val	Trp	Tyr	To tal
Rayal cheruvu AP OR333540	6.486	0.000	3.784	1.081	7.027	9.730	1.622	7.568	0.000	12.973	7.568	4.865	8.108	0.541	2.162	9.189	7.027	7.027	2.162	1.081	185
Majilabail Kerala OR342097	6.486	0.000	3.784	1.081	7.027	9.730	1.622	7.568	0.000	12.973	7.568	4.865	8.108	0.541	2.162	9.189	7.027	7.027	2.162	1.081	185
Chitoor Kerala OR342108	6.486	0.000	3.784	1.081	7.027	9.730	1.081	7.568	0.000	12.973	7.568	4.865	8.649	0.541	2.162	9.189	7.027	7.027	2.162	1.081	185
Avg.	6.525	0.000	3.784	1.081	6.950	9.730	1.583	7.529	0.039	13.012	7.645	4.903	8.147	0.541	2.162	9.151	7.066	6.911	2.201	1.042	185

Note: Ala-Alanine, Cys-Cysteine, Asp-Aspartic acid, Glu- Glutamic acid, Phe- Phenylalanine, Gly- Glycine, His- Histidine, Ile- Isoleucine, Lys-Lysine, Leu-Leucine, Met-Methionine, Asn- Asparagine, Pro- Proline, Gln- Glutamine, Arg- Arginine, Ser-Serine, Thr- Threonine, Val- Valine, Trp- Tryptophan, Tyr- Tyrosine

Table 3. Nucleotide composition in the populations of *L. oratoria*

Sequences vs nucleotide	A	C	G	T	A+T	G+C
UAHS OR294965	32.50	19.39	17.06	31.06	63.55	36.45
Byranahalli OR342102	32.50	19.57	17.06	30.88	63.38	36.62
Mandya OR342099	32.50	19.57	17.06	30.88	63.38	36.62
TNAU OR295916	32.50	19.57	17.06	30.88	63.38	36.62
UBKV OR296321	32.68	19.21	16.88	31.24	63.91	36.09
Guwahati OR296707	32.32	19.39	17.24	31.06	63.38	36.62
Pattambi OR296829	32.50	19.21	17.06	31.24	63.73	36.73
NRRI OR297688	32.68	19.75	17.41	30.16	62.84	37.16
Bantwal OR298283	32.32	19.03	17.24	31.42	63.73	36.27
Vellayani OR319147	32.50	19.21	17.41	30.88	63.38	36.62
Anand OR322045	32.32	19.39	17.24	31.06	63.38	36.62
Rayalcheruvu OR333540	32.50	19.21	17.06	31.24	63.73	36.27
Majilabail OR342097	32.68	19.21	16.88	31.24	63.91	36.09
Chitoor OR342108	32.14	19.57	17.24	31.06	63.20	36.8
Average	32.402	19.277	17.207	31.124	63.521	36.541

Note: A-Adenine, C-Cytosine, G-Guanine, T-Thymine.

Table 4. Maximum composite likelihood estimate of the pattern of nucleotide substitution for 14 sequences of *L. oratoria*

Nucleotides	A	T	C	G
A	-	2.41	1.51	28.17
T	2.52	-	1.11	1.33
C	2.52	1.78	-	1.33
G	53.39	2.41	1.51	-

Note:-- Each entry shows the probability of substitution (r) from one base (row) to another base (column)[1]. For simplicity, the sum of r values is made equal to 100. Rates of different transitional substitutions are shown in bold and those of transversional substitutions are shown in italics. The nucleotide frequencies are 32.47% (A), 31.02% (T/U), 19.38% (C), and 17.13% (G). The transition/transversion rate ratios are k1 = 21.153 (purines) and k2 = 0.738 (pyrimidines). The overall transition/transversion bias is R = **4.885**, where R = [A*G*k1 + T*C*k2]/[(A+G)*(T+C)]. This analysis involved 14 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All ambiguous positions were removed for each sequence pair (pairwise deletion option). There were a total of 557 positions in the final dataset. Evolutionary analyses were conducted in MEGA11.

differences across all sequences. Notably, a common haplotype, Hap_2, was shared by the collections from Mandya and Bhyranahalli (Karnataka). Similarly, the collections from Pattambi (Kerala) and Rayalcheruvu (Andhra Pradesh) formed Hap_6, as depicted by Popart v1.7.

The analysis of haplotype divergence, assessed through Tajima's D statistics, yielded a value of -1.99704. This value indicates the presence of low-frequency polymorphisms. The nucleotide diversity (pi) was calculated as 0.0112652. In addition, the Analysis of Molecular Variance (AMOVA) was conducted to assess genetic variation among the haplotypes. The correlation of random genetic accessions within groups and in the population (ϕST) was computed as 0.15009 (Table 7). A lower value of ϕST suggested that there was limited partitioning of the populations into subgroups.

Furthermore, gene flow and genetic differences were analysed for two groups: the North and South Indian populations consisting of 12 haplotypes (Table 8). Despite the presence of only four haplotypes, the Northern group exhibited a higher number of segregating sites, indicating more significant nucleotide polymorphism and nucleotide diversity. The notably high haplotype diversity within the Northern group signifies substantial genetic differences among populations, implying a unique genetic makeup for each population. The gene flow (Nm) was computed to be approximately 5.61, indicating a lack of significant genetic drift. Nei's genetic distance (Da) and the Fixation index (F_{ST}) were calculated as 0.00125 and 0.08179, respectively, suggesting the similarity in genetics between the two groups.

The utilization of DNA barcodes expedites the taxonomic identification process and serves as an alternative approach to evaluating biodiversity. DNA barcoding emerges as a practical method for distinguishing

Table 5. Pairwise distance between the sequences

Sequences	UAHS	Byra nahalli	Mandya	TNAU	UBKV	Guwahati	Pattambi	NRRI	Bantwal	Vella yani	Anand	Rayal cheruvu	Majilabail	Chit toor
UAHS OR294965														
Byranahalli OR342102	0.00180													
Mandya OR342099	0.00180	0.00000												
TNAU OR295916	0.00542	0.00723	0.00723											
UBKV OR296321	0.00723	0.00542	0.00542	0.00906										
Guwahati OR296707	0.00542	0.00360	0.00360	0.00723	0.00542									
Pattambi OR296829	0.00180	0.00360	0.00360	0.00360	0.00542	0.00360								
NRRI OR297688	0.03869	0.03679	0.03679	0.04058	0.03491	0.03679	0.03679							
Bantwal OR298283	0.00542	0.00723	0.00723	0.00723	0.00906	0.00723	0.00360	0.04058						
Vellayani OR319147	0.01269	0.01453	0.01453	0.01453	0.01637	0.01453	0.01086	0.04629	0.01453					
Anand OR322045	0.00906	0.00723	0.00723	0.01089	0.00906	0.00723	0.00723	0.04058	0.01089	0.01823				
Rayalcheruvu OR333540	0.00180	0.00360	0.00360	0.00360	0.00542	0.00360	0.00000	0.03679	0.00360	0.01086	0.00723			
Majilabail OR342097	0.00360	0.00542	0.00542	0.00542	0.00723	0.00542	0.00180	0.03869	0.00542	0.01269	0.00906	0.00180		
Chitoor OR342108	0.00722	0.00541	0.00541	0.00904	0.00722	0.00541	0.00541	0.03868	0.00541	0.01636	0.00904	0.00541	0.00722	

Note: The number of base substitutions per site from between sequences is shown. Analyses were conducted using the Kimura 2-parameter model. This analysis involved 14 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All ambiguous positions were removed for each sequence pair (pairwise deletion option). There were a total of 711 positions in the final dataset. Evolutionary analyses were conducted in MEGA11.

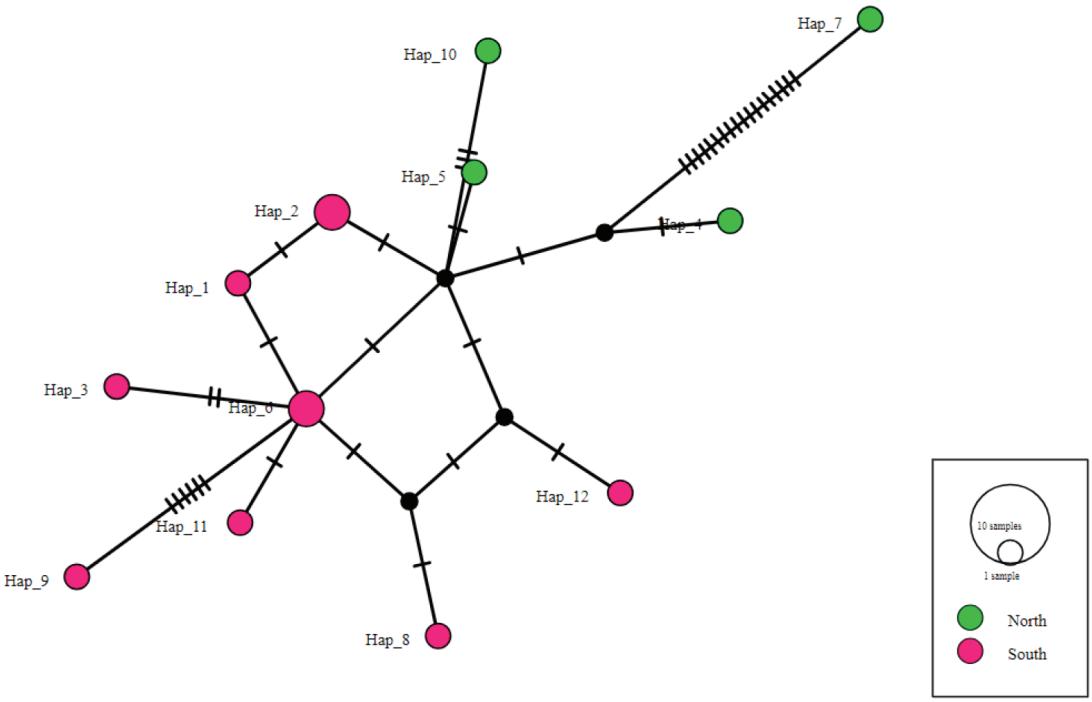

Figure 2. Median Joining networks of 12 haplotypes.

Table 6. Haplotypes from 14 populations of *L. oratoria*

Groups	Sl. No.	Haplotype	Populations
South India	1	Hap_1	UAHS
	2	Hap_2	Byranahalli
	3	Hap_2	Mandya
	4	Hap_3	TNAU
	5	Hap_6	Pattambi
	6	Hap_6	Rayalcheruvu
	7	Hap_8	Bantwal
	8	Hap_9	Vellayani
	9	Hap_11	Majilabail
	10	Hap_12	Chitoor
North India	11	Hap_4	UBKV
	12	Hap_5	Guwahati
	13	Hap_7	NRRI
	14	Hap_10	Anand

species such as *Leptocorisa acuta*, *L. oratoria*, and *L. chinensis*, which exhibit striking morphological similarities that can result in misidentification (Dong et al., 2021; Lahari et al., 2024). Mitochondrial genes have shown remarkable conservation and maternal inheritance,

Table 7. AMOVA analysis of genetic variation in *L. oratoria* haplotype populations (at Mean FST = 0.15009)

Variation	df	Sum of squares	Variance	Variation (%)
Among populations	1	116.39	5.92	15.01
Within populations	12	402.25	33.52	84.99
TOTAL	13	518.64	39.44	100

Note: df- Degrees of freedom

Table 8. Analysis of gene flow in the groups of haplotypes

Populations	South	North	Total
Number of sequences	10	4	14
Number of segregating sites, S	14	24	37
Number of Haplotypes, h	8	4	12
Haplotype diversity, Hd	0.95556	1	0.97802
The average number of differences, K	3.48889	12 .16667	6.27473
Nucleotide diversity, Pi	0.00626	0 .02184	0.01127

establishing them as a standard for insect identification. Within these mitochondrial genes, this study focused on a fragment of the *mtCOI* gene, which has been established as a valuable tool for barcoding insects. The *mtCOI* gene proves to be effective in identifying a wide array of metazoans (Folmer et al., 1994). In the present investigation, the *mtCOI* DNA barcodes of 14 *L. oratoria* populations were obtained using universal primers. Comparable barcodes were obtained from hemipterans of Tamil Nadu (Sangeetha and Alugachamy, 2019). The *COI* region typically spans around 500 to 658 base pairs (Jalali et al., 2015), while this study employed approximately 557 non-ambiguous base pairs in the sequencing process.

Variations at the nucleotide level were detected among the 14 sequences of *L. oratoria* spanning a length of 557 base pairs. These variations were identified at 37 positions with ten consistent sequence segments. Studies on the genetic diversity of *Riptortus pedestris* in Korea showed similar findings (Park et al.,2018). In that study, a segment of the *mtCOI* gene, 402 base pairs in length, exhibited approximately 36 variable sites. Similarly, the nucleotide composition of the *mtCOI* gene displayed a significant bias of 63.52% towards adenine (A) and thymine (T) content commonly seen in heteropterans (Kaur and Sharma, 2017).

Our research resulted in leucine (12.73%), phenylalanine (7.03%), isoleucine (7.56%), serine (9.19%), methionine (7.57 %), threonine (7.027%) and valine (7.027%) along with glysine (9.73%) and proline (8.1%) to be abundant in the *mtCOI* gene of *L. oratoria* (Pandi et al., 2023). The lysine (Lys) was exclusively present in the Vellayani populations and absent in other populations. This discrepancy could potentially be attributed to nucleotide loss during the elimination of ambiguous base pairs, resulting in a modification of nucleotide positioning within the third codon region.

The nucleotide composition observed in the acquired sequences of the *mtCOI* gene resembled the heteropterans in general (Li et al., 2012). However, slight variations were noted, with slightly elevated cytosine values and correspondingly decreased thymine values. Additionally, this study found a higher prevalence of nucleotide transitions in comparison to transversions. The overall transition-to-transversion bias, denoted as R and quantified at 4.885 indicated conservation of mutational changes that tend to occur more frequently due to transitions and transversions. This conservation was potentially attributed to the reduced occurrence of transversions in the 3rd codon region, which implies possible alterations in the amino acid sequence (Ojha et al., 2016).

The analysis of pairwise distances among the 14 populations of *L. oratoria* revealed a range of distances from 0 to 3.8 per cent. A clear trend emerged where the distances between populations within the same state were generally lower, spanning from 0.0018 to 0.0072. In contrast, the greatest distances were observed between populations originating from the Northern and Southern regions owing to their geographical distances (Yaakop et al., 2022). Their study examined the molecular aspects of *mtCOI* in five *L. oratoria* specimens and found pairwise distances ranging from 0 to 0.0457 per cent. This suggests a similar trend of genetic diversity and interconnectedness within these populations.

Haplotypes refer to variations in DNA along a single chromosome that are commonly inherited in conjunction. In this study of 14 *L. oratoria* sequences, 12 unique haplotypes emerged. These were categorized into two clusters based on their geographic origin. The evaluation of genetic diversity and potential population segregation was conducted by utilizing Tajima's D statistics, revealing a score of -1.99704, which suggests the presence of infrequent polymorphisms. Nucleotide diversity (pi) was calculated at 0.01126, similar to other coreids (Joyce et al., 2017). The exploration of Molecular Variance (AMOVA) resulted in a significant portion of variations existing within the populations, accounting for 84.9914 per cent as observed in other heteropterans like aphids (Nam et al., 2019).

The average coefficient of gene differentiation (Gst) was determined to be 0.0997, indicating that 9.97 per cent of the genetic variation was present between populations. This outcome suggests a noteworthy level of gene flow alongside moderate genetic differentiation among certain populations. The Fst (Fixation index) indicated a low to moderate level of genetic differentiation. A similar pattern emerged from this study of two groups (North and South) of 12 haplotypes, with a gene flow (Nm) of 5.61. This value suggests a reduced impact of genetic drift due to the population's genetic diversity. This phenomenon might be attributed to the absence of geographical barriers, fostering either low or moderate genetic differentiation and unrestricted gene flow among populations within the same species (Guo et al., 2023). This genetic divergence among populations could also stem from migration, contributing to an increase in gene flow hence their stability in the environment, owing to their pest status.

CONCLUSION

Molecular identification using the *mtCOI gene* was performed and phylogenetic analysis for *L. oratoria*

from different locations was assessed. The genetic variability of *L. oratoria* populations was determined and possible nucleotide differences were ascertained along with the population variations which described the presence of infrequent polymorphism and reduced genetic drift, but a good amount of gene flow between the populations. This is the first attempt to analyze mtCOI for this widely distributed pest of paddy from India. Further studies on the geographic variability of this species across the world need to be targeted.

ACKNOWLEDGMENTS

The authors thank Dr. C.A. Viraktamath, Emeritus Professor, Department of Entomology, UAS, GKVK, Bengaluru for guidance and suggestions. Authors acknowledge Dr PR Shashank, Senior Scientist, IARI, Pusa Campus, New Delhi and Dr M Raghuraman Professor, Banaras Hindu University, Varanasi and a few other collaborators across India for sharing gundhi bug specimens.

FINANCIAL SUPPORT

There is no financial support for the study.

AUTHOR CONTRIBUTION STATEMENT

LS conducted experiments. CMK conceived and designed research. LS and CMK wrote the manuscript. LS, CMK, NNL and BCH collected samples. All authors read and approved the manuscript.

CONFLICTS OF INTEREST/COMPETING INTERESTS

The authors declare no conflict of interest.

REFERENCES

Bandelt H, Forster P, **Röhl A**. 1999. Median-joining networks for inferring intraspecific phylogenies. Molecular Biology and Evolution 16(1): 37-48. https://doi.org/10.1093/oxfordjournals.molbev.a026036

Cobblah M A, Hollander D J. 1992. Specific differences in immature stages, oviposition sites and hatching patterns in two rice pests, *Leptocorisa oratorius* (Fabricius) and *L. acuta* (Thunberg) (Heteroptera: Alydidae). International Journal of Tropical Insect Science 13(1): 1-6. https://doi.org/10.1017/S1742758400013825

Dong X, Yi W, Zheng C, Zhu X, Wang S, Xue H, Ye Z, Bu W. 2021. Species delimitation of rice seed bugs complex: Insights from mitochondrial genomes and ddRAD seq data. Zoologica Scripta 51(2): 185-198. https://doi.org/10.1111/zsc.12523

Dutta S, Roy N. 2016. Life table and population dynamics of a major pest, *Leptocorisa acuta* (Thunb.)(Hemiptera: Alydidae), on rice and non-rice system. International Journal of Pure and Applied Bioscience 4(1): 199-207. https://doi.org/10.18782/2320-7051.2202

Excoffier L, Smouse PE, Quattro JM. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics Jun; 131(2): 479-91. doi: 10.1093/genetics/131.2.479. PMID: 1644282; PMCID: PMC1205020.

Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek A R. 1994. DNA primers for amplification of mitochondrial cytochrome C oxidase subunit I from diverse metazoan invertebrates. Molecular marine Biology and Biotechnology 3(5): 294-299.

Pandi G G P, Mahendiran A, Jaipal C, Gowda G B, Totan A, Naiyar N, …, Prakash, R. 2023. Genome organization and comparative evolutionary mitochondriomics of rice earhead bug *Leptocorisa oratoria* (Fabricius). Current Science 125: 407-415.

Guo J, Li J, Massart S, He K, Francis F, Wang Z. 2023. Analysis of the genetic diversity of Two *Rhopalosiphum* species from China and Europe based on nuclear and mitochondrial genes. Insects 14(1): 57. https://doi.org/10.3390/insects14010057

Hasegawa H. 1971. Distribution and taxonomy of rice bugs in Southeast Asia. Proceedings of the Tropical Agriculture Research Series: Proceedings of the Symposium on Tropical Agriculture Researches, Tokyo, Japan.

Hosamani V, Pradeep S, Sridhara S, Kalleshwaraswamy C M. 2009. Biological studies on paddy ear head bug, *Leptocorisa oratorius* Fabricius (Hemiptera: Alydidae). Academic Journal of Entomology 2(2): 52-55.

Hunt G J. 1997. Insect DNA extraction protocol. Fingerprinting methods based on arbitrarily primed PCR. Springer Lab Manuals (pp. 21-24), Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-60441-6_3

Jalali S K, Ojha R, Venkatesan T. 2015. DNA barcoding for identification of agriculturally important insects. New horizons in insect science: Towards sustainable pest management (pp. 13-23). Springer Private Ltd, India. https://doi.org/10.1007/978-81-322-2089-3_2

Joyce A L, Higbee B S, Haviland D R, Brailovsky H. 2017. Genetic variability of two leaf-footed bugs, *Leptoglossus clypealis* and *Leptoglossus zonatus* (Hemiptera: Coreidae) in the Central Valley of California. Journal of Economic Entomology 110(6): 2576-2589. https://doi.org/10.1093/jee/tox222

Kaur H, Sharma K. 2017. *COI*-based DNA barcoding of some species of Pentatomidae from North India (Hemiptera: Heteroptera). Mitochondrial DNA Part A 28(5): 756-76. https://doi.org/10.1080/24701394.2016.1180513

Kumar V, Goswami T N. 2020. Report on diversity of *Leptocorisa* spp. and their egg parasitoids in rice at Sabour, Bihar. International Journal of Environmental Science and Technology 9(5): 741-751.

Lahari S, Kalleshwaraswamy C M, Hanumanthaswamy B C, Naveena N L. 2024. Taxonomy and genetic diversity of rice earhead bugs, *Leptocorisa* spp. (Hemiptera: Alydidae) from India. Phytoparasitica 52(3): 1-14. https://doi.org/10.1007/s12600-024-01161-7

Leigh J W, Bryant D. 2015. PopART: Full-feature software for haplotype network construction. Methods in Ecology and Evolution 6(9):1110–1116. https://doi.org/10.1111/2041-210X.12410

Li M, Tian Y, Zhao Y, Bu W. 2012. Higher level phylogeny and the first divergence time estimation of Heteroptera (Insecta: Hemiptera) based on multiple genes. PLoS One **7**(2): 32152. https://doi.org/10.1371/journal.pone.0032152

Librado P, Rozas J. 2009. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25(11): 1451-1452. https://doi.org/10.1093/bioinformatics/btp187

Litsinger J A, Barrion A T, Canapi B L, Libetario E M, Pantua P C, Cruz C G, … Macatula R. F, 2015, *Leptocorisa* rice seed bugs (Hemiptera: Alydidae) in Asia: A review. Philippine Entomologist 29(1): 1-103.

Nam H Y, Park Y, Lee J H. 2019. Population genetic structure of *Aphis gossypii* glover (Hemiptera: Aphididae) in Korea. Insects 10(10): 319. https://doi.org/10.3390/insects10100319

Ojha R, Jalali S K, Shivalingaswamy T M, Verghese A. 2016. Mitochondrial COX1 gene-based delineation of putative insect species from Salt Lake of Great Rann of Kutch, India. Journal of Entomology and Zoology Studies 4(5): 166-170.

Park C G, Lee S K, Park M, Lee W. 2018. Genetic diversity of *Riptortus pedestris* (Hemiptera: Alydidae) populations in the Republic of Korea based on COI sequence. Entomological Research 48(2): 122-127. https://doi.org/10.1111/1748-5967.12247

Patel D T, Stout M J, Fuxa J R. 2006. Effects of rice panicle age on quantitative and qualitative injury by the rice stink bug (Hemiptera: Pentatomidae). Florida Entomologist 89(3): 321-327. https://doi.org/10.1653/0015-4040(2006)89[321:EORPAO]2.0.CO;2

Sangeetha M, Alaguchamy N. 2019. Morphological and molecular characterization of selected Hemipteran insect Bugs from Sivaganga, South India. Research Dimensions 7(1): 56-62.

Shashank P R, Naveena N L, Rajgopal N N, Elliott T A, Sreedevi K, Sunil S, Meshram N M. 2022. DNA barcoding of insects from India: Current status and future perspectives. Molecular Biology Reports 49(11), 10617-10626. https://doi.org/10.1007/s11033-022-07628-2

Yaakop S, Sabri S, Kamaruddin N A K, Norizam N N, Mohammed M A. 2022. Species composition and DNA barcoding of Hemipteran assemblages throughout paddy growing seasons. Pertanika Journal of Tropical Agricultural Science 45(3): 631- 648. https://doi.org/10.47836/pjtas.45.3.06