Molecular Phylogeny and Sequence Divergence in Stink Bugs of family Pentatomidae (Hemiptera: Heteroptera) inferred from analysis of Mitochondrial Cyt B Gene in India

Ramneet Kaur* 1; Devinder Singh2
1Department of Zoology, Punjabi University, Patiala, India
2Department of Zoology, GGDSD College, Chandigarh, India

Article History: Received 02nd July, 2020; Accepted 16th July, 2020; Published 23rd July, 2020

ABSTRACT

The present work was aimed to investigate the phylogenetic analysis of different species of Indian stink bugs belonging to family Pentatomidae based on mitochondrial Cyt b DNA sequences. We sequenced and analysed a 440 bp fragment of the mitochondrial Cyt b gene to assess the level of sequence divergence in different species of Pentatomid bugs collected from distant localities of Northern India. The data divulged an R value of 1.19 suggesting transitions to be more than transversions. 18 sequences representing 10 species and 5 genera were finally analysed using partial Cyt b gene region. The intraspecific divergence ranged from 0.2 to 7.2% with an average of 3.75 ± 3.0% whilst interspecific divergence ranged from 3.8 to 10.3% with an average of 7.49 ± 1.7% and intergeneric distance ranged from 12.1 to 20.2% with an average of 16.37 ± 2.2%. The percentage divergence values revealed a clear pattern of increased nucleotide diversity from conspecific to congeneric to intra subfamily level using Cyt b gene fragment.

Keywords: Pentatomidae, Stink Bugs, Mitochondrial Cyt b, Phylogenetic analysis

INTRODUCTION

Hemiptera is a very large and diverse order of hemimetabolous insects which is further divided into five suborders: Heteroptera, Cicadomorpha, Fulgoromorpha, Sternorrhyncha and Coelorrhyncha (Schuh and Slater, 1995). Out of these five suborders, Heteroptera is a highly diverse taxon with more than 42,000 species in over 5800 genera and 140 families described worldwide (Henry, 2009; Park et al., 2011). Infaordo Pentatomomorpha, which consists of 40 families representing more than 14,000 species, is the second largest among the seven infraorders of Heteroptera and constitutes the most recognizable group of true bugs. Pentatomomorpha is further divided into six superfamilies viz., Pentatomoidea, Coreoidea, Pyrrhocoroidea, Idiostoloidea, Aradoide and Lygaeoidea (Dolling, 1981; Schaefer, 1993; Schuh and Slater, 1995; Weirauch and Schuh, 2011; Tian et al., 2011). Superfamily Pentatomoidea consists of about 7,000 known species in 15 families, of which family Pentatomidae is a large family containing 4,722 species in about 900 genera (Rider 2019). Pentatomids are generally known as “Stink Bugs”, because of the emission of disagreeable pungent odour from a pair of scent glands that open in the region of the metapleura. They produce large quantities of strong smelling and irritating defensive chemicals, which are released when they are disturbed or molested. The high reproductive rates and an exceptional ability to transmit diseases make them the worst agricultural pests (Song and Liang, 2009).

Traditionally, identification has been done on the basis of morphological diagnoses provided by taxonomical studies. But only experts such as taxonomists and trained technicians can identify taxa accurately, as it requires special skills acquired through extensive experience (Jinbo et al., 2011). Moreover, small size of some insects, availability of damaged specimens or partial fragments of body parts or immature stages also act as a barrier in the identification process. To overcome these limitations molecular techniques are commonly used nowadays in which molecular data have proven to be very effective for understanding phylogenetic relationships (Blair and Hedges, 2005; Regier et al., 2005; Roe and Sperling, 2007) and diagnosing and delimiting cryptic species (Sperling and Hickey, 1994; Goetze, 2003; Blair et al., 2005; Hendrixson and Bond, 2005; Roe and Sperling, 2007). The degree of genetic relatedness between species reveals their relationships with one another. The more the genetic relatedness, the more will be the similarity in DNA fragments in any two individuals. The techniques on the basis of which we can perform the DNA typing include PCR, AFLP, RAPD, RFLP, RT-PCR and microsatellite technology (Mullis and Falooana, 1987; Arawal and Shrivastava, 2014). Mitochondrial DNA is widely used as a valuable molecular marker for phylogenetic studies in animals, because of its simple genomic structure (Avise, 1994). It is small, circular DNA with length ranging from 14,000 to 17,000 bp that comes predominantly from maternal sources (Lessinger et al., 2000; Boone, 1999; Cameron et al., 2007). It has relatively fast mutation rate, lacks introns and shows no recombination (Brown et al., 1979; Hebert et al., 2003). Thus, mtDNA suits as the best device to trace phyloegeny (Simon et al., 1994; Rubinoff and Holland, 2005). The genomic DNA is first subjected to PCR amplification and then analysed by sequencing. By comparing the DNA sequences, it is possible to create phylogenetic trees which show the degree of relatedness between the species. This technique was developed by Sanger et al. (1977) and is
considered as the best way to directly estimate genetic variations of specific genes among taxa. The datasets can be aligned and analysed with different computer-based algorithms and methods of phylogenetic analyses. The most commonly used methods are multiple alignment of nucleotide or protein sequences and their analyses by maximum parsimony, maximum likelihood, Bayesian inference and distance algorithms, implemented in programs such as Clustal W, MUSCLE, MEGA 6.0 (Thompson et al., 1994; Huelsenbeck et al., 2001; Swofford, 2002; Tamura et al., 2013).

**MATERIALS AND METHODS**

**Taxon sampling**

The first step subsumed the collection of stink bugs from different regions of North India (Himachal Pradesh, Punjab and Uttarakhand), followed by sorting, pinning, labelling and identification of the collected specimens. The collection of stink bugs involved both day and night collection. After collection few specimens were pinned while rest of them were preserved in absolute alcohol and stored at -20°C for molecular studies. For a molecular study, DNA isolation is the first and crucial step. DNA isolation is a multistep procedure which involves cell lysis by treatment with lytic enzymes and/or detergents, extraction with organic solvents and then the recovery of DNA by alcohol precipitations. Here, liquid nitrogen was used which allowed access to DNA by breaking cell walls while other cellular enzymes and chemicals remained inactivated, thus giving the higher yield of DNA. DNA was extracted from thorax region using modified phenol: chloroform method (Kambhamptati and Rai, 1991). Genomic DNA was eluted in 100 µl of TE buffer and stored at -20°C. Integrity of extracted DNA was checked on 0.8% agarose gel by horizontal gel electrophoresis. Quantification of DNA was done by nanodrop-spectrophotometer. The Cyt b gene was amplified using forward primer 5'- TAGGATATGTTTTACCTTGAGGACA-3' and reverse primer 5'- TCCTCCTAATTATTAGGAATTG-3' (Muraji et al., 2000). The protocol followed for amplification of DNA was shown in Table 1. PCR was carried out at an annealing temperture of 55°C for 40 sec and an extension at 72°C for 40 sec.

The most commonly used method of sequencing DNA, the dideoxy method or chain termination method, was developed by Sanger et al. (1977). The amplified products of Cyt b gene fragments were got sequenced directly from Agrigenome Pvt. Ltd, Cochin (India). The edited sequences were compared with related sequences from the nucleotide database of the National Centre for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov), using Basic Local Alignment Search Tool (BLAST) algorithm (Altschul et al., 1997). Cyt b nucleotide sequences generated from this study were converted into amino acid sequences using MEGA software package (Kumar et al., 1994; 2001) and checked for the occurrence of stop codons. Nucleotide sequences were further aligned in MUSCLE tool (Edgar, 2004) of Codon Code Aligner v. 3.7.1. Nucleotide composition and substitution statistics were analysed in MEGA version 6.06 (Tamura et al., 2013) for inferring phylogenetic trees and for estimating divergence times. Percentage divergence matrix was drawn using Kimura-2-parameter model (Kimura, 1980) in MEGA 6.06 (Tamura et al., 2013). Codon positions included were 1st+2nd+3rd Noncoding.

**RESULTS AND DISCUSSION**

The first step was extraction of DNA from both dry old pinned and alcohol preserved fresh samples. It was observed that the quality as well as the quantity of DNA was far better in case of alcohol preserved fresh samples in comparison to old samples. Only insect legs and thorax were used for extraction. The integrity of the extracted DNA was checked qualitatively by horizontal agarose gel electrophoresis. PCR amplification was performed for 7 species representing 4 genera belonging to two subfamilies (Pentatominae and Podopinae) of family Pentatomidae. The PCR products were run on 2% agarose gel along with a 100 and 500 bp DNA ladder to determine the size of the amplified products. The bands obtained were clean and intense. Specific primers were used in the present study and measures were taken to avoid any chances of amplifying numts (nuclear mitochondrial DNA). The samples containing the required amplicons (Figures 1-3) were sequenced by both forward and reverse sequencing by chain termination method. Once sequences were obtained, these were analysed, edited and trimmed accordingly by Codon Code Aligner.

Lane M: 100 bp DNA ladder 

Lanes: Different species of family Pentatomidae taken for the present study

### Multiple sequence alignments

11 sequences of 440 bp representing seven species of four genera belonging to two subfamilies (Pentatominae and Podopinae) were finally taken for further analysis (Table 2). 7 corresponding sequences of five species deposited by other workers were procured directly from Genbank (Table 3). The final aligned data pertained to 18 Cyt b sequences of 440 bp representing 10 species and 5 genera. Sequence of Aeschyntelusnotatus, belonging to family Rhopalidae, was included in the analysis as the outgroup taxon.

### Nucleotide substitution

Nucleotide substitution patterns may play an important role in obtaining information about phylogenetic relationships and the

<table>
<thead>
<tr>
<th>S. No.</th>
<th>Reagents</th>
<th>Volumes (µl/10µl rxn)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>PCR Buffer (5X)</td>
<td>1.20 ml</td>
</tr>
<tr>
<td>2.</td>
<td>Magnesium Chloride (50mM)</td>
<td>0.20 ml</td>
</tr>
<tr>
<td>3.</td>
<td>dNTPs Mix(2.5mM)</td>
<td>1.20 ml</td>
</tr>
<tr>
<td>4.</td>
<td>Primer-Forward (10pM)</td>
<td>0.25 ml</td>
</tr>
<tr>
<td>5.</td>
<td>Primer-Reverse (10pM)</td>
<td>0.25 ml</td>
</tr>
<tr>
<td>6.</td>
<td>Template DNA (ng/µl)</td>
<td>0.50 ml</td>
</tr>
<tr>
<td>7.</td>
<td>Phusion DNA polymerase</td>
<td>0.10 ml</td>
</tr>
<tr>
<td>8.</td>
<td>Double Distilled water</td>
<td>6.30 ml</td>
</tr>
<tr>
<td></td>
<td>Total volume</td>
<td>10 ml</td>
</tr>
</tbody>
</table>
Table 1: Various components of the reaction mixture with the volume used.

Table 2: Details of species analysed for Cyt b in the present study.
population structure of different organisms and such variations may show increased diversification among independent lineages and increasingly divergent taxa (Galtier et al., 2005; Souza et al., 2016). Transitional substitutions are known to occur more readily than transversional substitutions (Brown et al., 1979; Muraji et al., 2000). The pattern of nucleotide substitutions was estimated using Maximum Composite Likelihood method. The estimated transition/transversion bias was \( R = 1.19 \) Cyt b sequences respectively, suggesting transitions to be more than transversions. Maximum number of transitions is the result of A+T rich regions in the insect genome.

In several insect species, it has been reported that the frequency of base changes was different among different types of nucleotide substitution (Dowton et al., 1998; Muraji et al., 2000). In Cyt b gene region, A-T transversions showed a lower frequency than T-C transition. In particular, in the second positions of codons, T-C transitions showed the highest frequency which is 3.5 times higher than the number of A-T transversions (Table 4). Muraji et al. (2000) also found similar results showing the high rates of T-C transitions than A-T transversions in partial Cyt b gene fragment. According to their study, nucleotide substitution may occur more frequently between pyrimidine bases than between pyrimidine and purine bases respectively.

**Sequence divergence**

The genetic divergence among the selected species of family Pentatomidae was estimated using partial Cyt b gene sequences. All the species exhibited discriminative values of intra and interspecific and intergeneric divergence using Cyt b gene fragment. The percent divergence values revealed a clear pattern of increased nucleotide diversity from conspecific to congeneric to intra subfamily level. The intraspecific divergence ranged from 0.2 to 7.2% with an average of 3.75 ± 3.0% whilst interspecific divergence ranged from 3.8 to 10.3% with an average of 7.49 ± 1.7% and intergeneric distance ranged from 12.1 to 20.2% with an average of 16.37 ± 2.2%. Intergeneric divergence was quite high in the present study for all the genera and no ambiguity was observed as it was above 10% for all. For Halyomorphaicus interspecific divergence was quite higher than intraspecific divergence indicating the species to be well separated.

**Phylogenetic tree**

18 sequences of 440 bp representing 9 species of four genera belonging to two subfamilies (Pentatominae and Podopinae) were finally taken along with 1 species of family Rhopalidae (taken as an out group) using partial Cyt b gene fragment for phylogenetic analysis in family Pentatomidae with a bootstrap value of 1000 replications. The NJ and ML trees for the 440 bp region of Cyt b gene of mtDNA isolated from the nine species of Pentatomidae and their world sequences are shown in Figures 5 and 6. It is clear from the data that all the samples of the nine species are well separated from the other species and grouped together with their world sequences with high bootstrap values. In all trees, Graphosomalineatum and Graphosomarubrilineatum belonging to subfamily Podopinae acted as a sister taxon of genus Nezara and Plautia of subfamily Pentatominae with 81% BT in NJ, 82% in ME tree. Aeschyntelusnotatus (out group) with Gen Bank accession number EU427333.1 was most basal followed by Halyomorphapicus belonging to subfamily Pentatominae in all the trees. Figures 4 and 5 clearly showed that all the nine species are well separated with high bootstrap support.

Species *Plautiastali*, *P. crossoita* and *P. viridicollis* have no Gen Bank sequence for comparison and they are new for the Gen Bank database from India. All the present sequences belonging to *Nezara viridula* were clustered together showing 99% bootstrap value in NJ tree. The three species of genus Plautia whose world sequences were not available grouped together within their own samples and separated from other species with 82% bootstrap value in NJ trees. However, in ML tree species of subfamily Pentatominae appeared as a sister clade to genus Graphosoma.
of subfamily Podopinae with 78% bootstrap value (Figure 5). The analysis of phylogenetic trees clearly showed that all the pentatomid species considered for the present research based on Cyt b got rightly allotted to the three subfamilies namely, Pentatominae, Podopinae and Asopinae.

**CONCLUSION**

During phylogenetic analysis it was observed that subfamily Pentatominae might not be monophyletic or subfamily Asopinae and Podopinae showed very close relationship with Pentatominae. The three subfamilies showed a close relationship...
with each other in each tree. The species taken as an outgroup remained separated from the studied family in all the trees. All the species studied revealed a similar pattern with high bootstrap value forming well separated clades. Different genera clustered separately and congeners clustered together in all the trees. On the basis of above discussion it can be concluded that Cyt b gene fragment of mitochondrial DNA are useful in the identification and phylogenetic analysis of Indian Pentatomid bugs. Cyt b acts as wonderful markers for both phylogenetic analysis describing monophyly and polyphyly of the groups and at the same time it also throws some light on the apparent geographical population distribution pattern with the genetic variation in these gene fragments

**ACKNOWLEDGEMENT**

We acknowledges Rajiv Gandhi National Fellowship for financial support and DBT-IPLS Project for required laboratory facilities at Sophisticated Instrumentation Centre and Zoology Department, Punjabi University, Patiala.

**REFERENCES**


