**International Journal of Entomology Research**
**ISSN: 2455-4758; Impact Factor: RJIF 5.24**
**Received: 23-03-2020; Accepted: 12-04-2020; Published: 18-04-2020**
**www.entomologyjournals.com**
**Volume 5; Issue 2; 2020; Page No. 116-119**

# Analysis of the mitochondrial COI gene fragment and its informative potential for phylogenetic analysis in family pentatomidae (hemiptera: hetroptera)

**Ramneet Kaur[1*], Devinder Singh[2]**
[1, 2] Department of Zoology, Punjabi University, Patiala, Punjab, India

## Abstract

Pentatomidae is a widely diverse family represented by 4,722 species belonging to 896 genera. It is considered as one of the largest family within suborder Heteroptera. In the present study, partial mitochondrial COI gene fragment of approximately 600bp from seven species of family Pentatomidae collected from different localities of Northern India has been analysed. The data divulged an A+T content of 65.8% and an R value of 1.39. The COI sequences were added directly to Genbank NCBI. The database analysis shows mean K2P divergence of 0.7% at intraspecific level and 13.5% at interspecific level, indicating a hierarchal increase in K2P mean divergence across different taxonomic levels.

**Keywords:** pentatomidae, mitochondrial gene, COI

## Introduction

Family Pentatomidae, chosen for the present study, is one of the largest families within suborder Hetroptera (Rider 2006-2017) [8]. Most species in this family are economically important as agricultural pests, whereas some are used as biological control agents. Their high reproductive rates and an exceptional ability to transmit diseases make the hemipterans worst agricultural pests (Song and Liang, 2009) [9]. The 5' end of the COI gene has been proposed as a standardized DNA ''Barcode'' for the identification of species in the animal kingdom (Hajibabaei *et al.*, 2005; Floyd *et al*., 2009) [4, 2]. It has been used as an identification tool for the heteropteran insects by enabling the recognition of morphologically cryptic species and can be helpful in pest management (Park *et al.*, 2011) [6]. Moreover, COI has been used as a standardised tool for molecular taxonomy and identification (Ratnasingham and Herbert, 2007) [7]. COI is considered as the fastest evolving protein coding gene in the mitochondrial genome along with COII suited best for analysis of closely related species, subspecies and geographical populations (Zhao *et al.*, 2015) [12].

In the present study, we analysed a partial 600bp COI sequence of 7 species of family Pentatomidae collected from different localities of northern India. The aim of the present study was to draw the phylogenetic relationship of populations of family Pentatomidae collected from the Indian subcontinent and to study their relation with the out-group taxa based upon mitochondrial COI gene fragment. COI sequences of these species have been added to the existing database. The dataset has been analysed at different hierarchal level for base composition and sequence divergence analysis.

## Material and Methods

Adult specimens belonging to seven species of family Pentatomidae were collected from different regions of northern India and were preserved in ethanol. Specimens were identified using relevant literature and by comparing them with the identified specimens available in the Department of Zoology and Environmental Sciences, Punjabi University, Patiala. DNA was extracted from legs of the specimens following the method of Kambhampati and Rai (1991) [5] with minor modifications. A region of COI gene was amplified using primers LCO1490 and HCO2198 (Folmer *et al*., 1994) [3]. Amplification of the target DNA was done by incubating the samples at three steps (denaturation, annealing and extension) under the following conditions: 1 cycle, 95° C (5 min); 35 cycles, 95° C (1 min), 50° C (1 min) 72° C (90 s); 1cycle, 72° C (7 min). PCR products were visualized on 1% agarose gel with ethidium bromide staining under UV light. The amplified products were got sequenced from AgriGenome, Cochin (India). The sequence data of 510 bp was retrieved in the form of chromatograms. The sequences were checked manually for the exclusion of ambiguous nucleotides. Related sequences were retrieved from GenBank using Basic Local Alignment Search Tool (blast) algorithm (Altschul *et al*., 1997) [1]. Sequences were aligned using Clustal W and phylogenetic analysis was done in MEGA 6.06 software package (Tamura *et al*., 2013) [10]. Maximum Likelihood analysis was performed using Tamura-Nei parameter (TN93+G+I; BIC=5393.994; InL= -2505.427) selected as the best fit model for the analysis of the present data. 1000 pseudo replicates were generated to test the robustness of the tree nodes. Sequences deposited in GenBank by other workers of the congeneric specimens were taken for the alignment purpose. All the sequences were aligned in ClustalW and divergence at population and species levels was analysed by K2P model of base substitution. Phylogenetic analysis was carried out using, Maximum Likelihood (ML, Figure 1) Minimum Evolution (ME, Figure 2) and Neighbour-Joining (NJ, Figure 3) approaches in MEGA 6.0software (Tamura *et al*., 2013) [10].

## Results and Discussion

COI sequences of about 600bp representing 200 amino acids were obtained from seven species of family Pentatomidae and all these sequences were submitted to GenBank database (Table 1).

**Table 1**: Details of species analysed in the present study

| S. No | Taxa | Number of sample | Collection Place | Collection Month/Year | Accession no. |
|---|---|---|---|---|---|
| 1. | *Palomena prasina* | 1 | Himachal Pradesh | June, 2014 | KY508431 |
| 2. | *Chlorochroa ligata* | 1 | Himachal Pradesh | October, 2015 | KY624495 |
| 3. | *Graphosoma lineatum* | 1 | Himachal Pradesh | March, 2013 | KY624496 |
| 4. | *Plautia viridicollis* | 1 | Himachal Pradesh | June, 2016 | KY624497 |
| 5. | *Halyomorpha picus* | 1<br>1 | Himachal Pradesh | May, 2014 | KY624498<br>KY508433 |
| 6. | *Dalpada neoclavata* | 1 | Himachal Pradesh | July, 2016 | KY508432 |
| 7. | *Pentatoma sp.* | 1 | Himachal Pradesh | September, 2015 | KY508434 |

No stop codons or frame shifts were detected, indicating that sequences were not pseudogenes (NUMTs). Thirteen sequences of thirteen different species submitted by other workers were procured directly from GenBank (Table 2). The final aligned data belonged to 21 COI sequences of 600 bp representing 13 species and 7 genera. Sequence of *Homaemus aeneifrons* belonging to family Scutelleridae was taken as out-group. The alignment showed 423 conserved sites, 263 variable sites and 200 parsimony informative sites. The average A+T content was 65.8% (Table 3).

**Table 2:** List of taxa whose sequences were downloaded from NCBI for alignment

| S.NO | Name | Accession number | country |
|---|---|---|---|
| 1. | *Palomena prasine* | KM021668.1 | Germany |
| 2. | *Chlorochroa ligata* | KR039777.1 | Canada |
| 3. | *Graphosoma lineatum* | KM021891.1 | Germany |
| 4. | *Plautia viridicollis* | KU163633.1 | India |
| 5. | *Halyomorpha picus* | KM226884.1 | India |
| 6. | *Dalpada neoclavata* | KT273901.1 | India |
| 7. | *Palomena angulosa* | GQ292252.1 | Korea |
| 8. | *Chlorochroa granulosa* | KR037841.1 | Canada |
| 9. | *Graphosoma italicum* | KX960060.1 | France |
| 10. | *Plautia crossota* | KU163631.1 | India |
| 11. | *Halyomorpha halys* | KU601521.1 | Japan |
| 12. | *Dalpada nigricollis* | KU377166.1 | India |
| 13. | *Pentatoma rufipes* | KM021597.1 | Germany |

This data was further analysed for sequence divergence at different taxonomic levels. Intraspecific divergence ranged from 0.0% to 1.5% and interspecific divergence ranged from 9.0% to 16.0%.

**Table 3**: Per cent base composition of the COI segment studied

| Base | A | T | G | C |
|---|---|---|---|---|
| Average percentage | 31.8 | 34.0 | 16.4 | 17.8 |

**Table 4:** Maximum Likelihood Pattern of Nucleotide Substitution

| | A | T | C | G |
|---|---|---|---|---|
| A | - | *7.86* | *4.24* | 4.87 |
| T | *7.26* | - | 13.78 | *3.69* |
| C | *7.26* | 25.64 | - | *3.69* |
| G | 9.57 | *7.88* | *4.24* | - |

The data also revealed higher number of transitional sites (53.86%) as compared to the total number of transversional sites (46.14%). The estimated Transition/Transversion ratio (*R*) for the dataset was found to be 1.39 (Table 4). Substitution pattern and rates were estimated under the Tamura-Nei (1993) [10] model. All the members showed distinct barcodes with no case of barcode sharing. COI marker proved to be useful in the diagnosis of the family Pentatomidae as during the BLAST search sequences match with sequences of congeneric species from GenBank. Nucleotide composition showed a bias towards A +T content (65.9%) as has also been recorded in other heteropterans by Zhang *et al*. (2013) [11].

**Table 5:** Pairwise K2P intraspecific divergence

| S.NO | Species (accession no.) | Species (accession no.) | Divergence (%) |
|---|---|---|---|
| 1. | *Palomena prasina* ( KY508431) | *Palomena prasina* ( (KM021668.1)* | 0.9 |
| 2. | *Chlorochroa ligata(* KY624495*)* | *Chlorochroa ligata (*KR039777.1)* | 1.5 |
| 3. | *Graphosoma lineatum* ( KY624496) | *Graphosoma lineatum* (KM021891.1)* | 1.5 |
| 4. | *Plautia viridicollis* ( KY624497) | *Plautia viridicollis* (KU163633.1)* | 0.0 |
| 5. | *Halyomorpha picus* ( KY624498) | *Halyomorpha picus* (KM226884.1)* | 0.0 |
| 6. | *Halyomorpha picus* ( KY508433) | *Halyomorpha picus* (KM226884.1)* | 0.0 |
| 7. | *Dalpada neoclavata* ( KY508432) | *Dalpada neoclavata* (KT273901.1)* | 0.6 |
| 8. | *Pentatoma sp.* ( KY508434) | *Pentatoma rufipes* (KM021597.1)* | 1.2 |

Average K2P intraspecific divergence was found to be 0.7% with minimum of 0.0% in *Plautia viridicollis* to a maximum of 1.5% in *Graphosoma lineatum* (Table 4). The K2P value was 0.6% for *Dalpada neoclavata*, 0.9% for *Palomena prasine*, 1.2% for *Pentatoma rufipes* and 1.5% for *Chlorochroa ligata* (Table 5).

**Table 6:** Pairwise K2P interspecific divergence

| S. No. | Taxa (accession no.) | Taxa (accession no.) | Divergence (%) |
|---|---|---|---|
| 1. | *Palomena prasina* (KY508431) | *Palomena angulosa* (GQ292252.1)* | 11.1 |
| 2. | *Chlorochroa ligata* (KY624495) | *Chlorochroa granulosa* (KR037841.1)* | 15.0 |
| 3. | *Graphosoma lineatum* (KY624496) | *Graphosoma italicum* (KX960060.11)* | 16.0 |
| 4. | *Plautia viridicollis* (KY624497) | *Plautia crossota* (KU163631.1)* | 14.0 |
| 5. | *Halyomorpha picus* ( (KY624498) | *Halyomorpha halys* (KU601521.1)* | 15.6 |

| | | | |
|---|---|---|---|
| 6. | *Halyomorpha picus* (KY508433) | *Halyomorpha halys* (KU601521.1)* | 15.6 |
| 7. | *Dalpada neoclavata* (KY508432) | *Dalpada nigricollis* (KU377166.1)* | 9.0 |

Average interspecific divergence was found to be 13.5% with minimum of 9.0% for *Dalpada neoclavata* and *Dalpada nigricollis* pair and maximum of 16.0% for *Graphosoma lineatum* and *Graphosoma italicum* pair (Table 6).

We endeavour to generate phylogenetic trees for the present work. In all the trees, namely Maximum likelihood, Minimum evolution and Neighbour joining, different genera clustered separately and congeners clustered together. In maximum likelihood tree, *Homaemus aeneifrons* belonging to Family Scutelleridae taken as outgroup showed a close relation to *Chlorochora ligata* and *Chlorochora granulosa*..

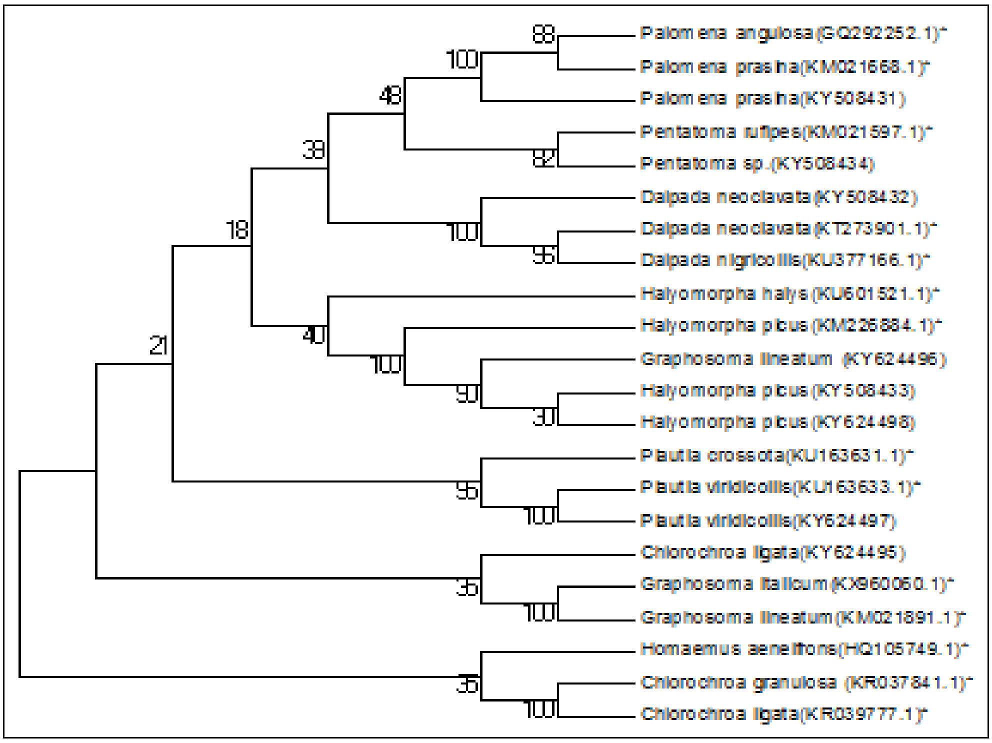

**Fig 1:** Maximum likelihood tree based on (K2P). Numbers indicate the percentage of 1000 bootstrap replicate.

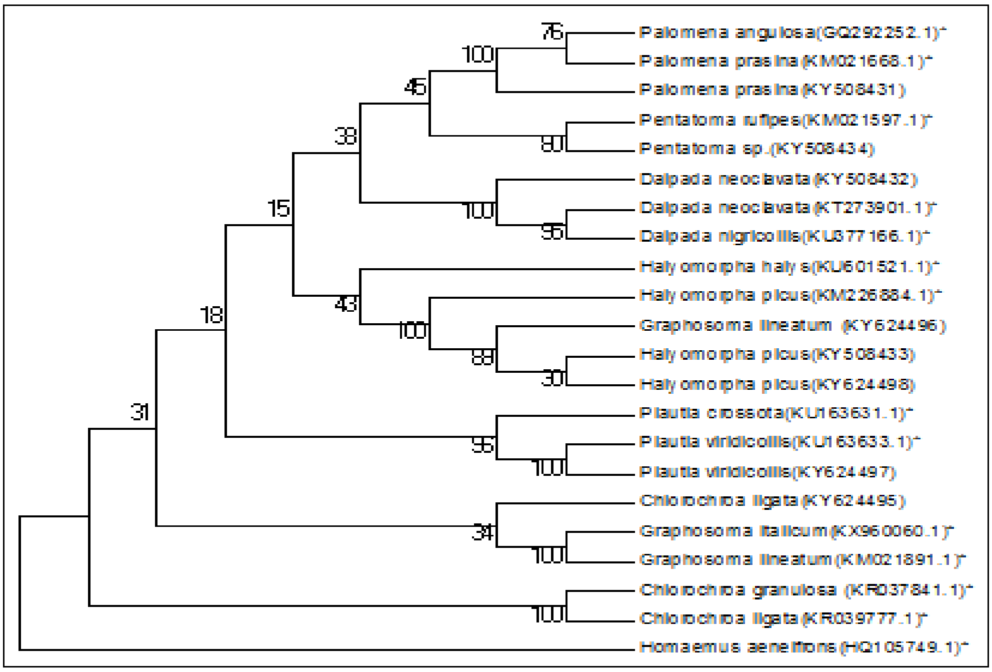

**Fig 2:** Minimum Evolution tree based on (K2P). Numbers indicate the percentage of 1000 bootstrapreplicate

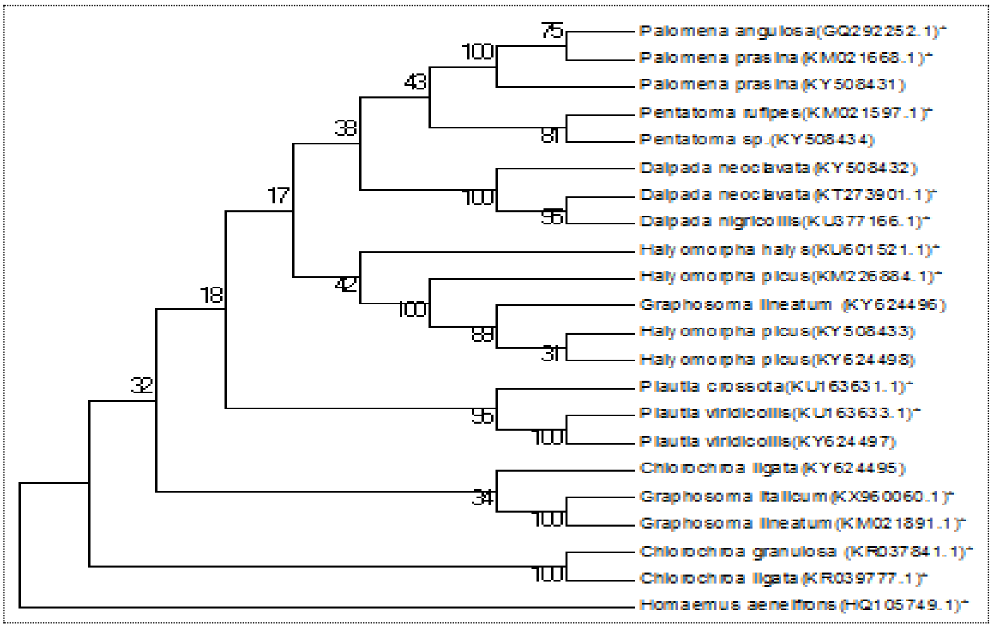

**Fig 3:** Neighbour-Joining tree based on (K2P). Numbers indicate the percentage of 1000 bootstrap replicate.

## Conclusion

In the present study, COI sequences of seven different species of family Pentatomidae was aligned with the reference sequences downloaded from NCBI for phylogenetic analysis. The database analysis shows hierarchical increase in K2P mean divergence both at intraspecific and interspecific levels.

## Acknowledgement

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

## References

1. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Research. 1997; 25(17):3389-3402.
2. Floyd RM, Wilson JJ, Hebert PDN. In *DNA Barcodes and Insect Biodiversity*, R.G. Foottit and P.H. Adler (eds.). Wiley-Blackwell, Oxford, 2009, 417-432p.
3. Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Molecular Biology and Biotechnology. 1994; 3:294-299.
4. Hajibabaei M, deWaard JR, Ivanova NV, Ranasingham S, Dooh RT. Critical factors for assembling a high volume of DNA barcodes. Philosophical Transactions of the Royal Society of London. 2005; 360:1959-1967.
5. Kambhampati S, Rai KS. Mitochondrial DNA variation within and among populations of the mosquito *Aedes albopictus*. Genome. 1991; 34(2):288-292.
6. Park DS, Foottit R, Maw E, Hebert PD. Barcoding Bugs: DNA based identification of the true bugs (Insecta: Hemiptera: Heteroptera). PLoS One. 2011; 6(4):1-9.
7. Ratnasingham S, Hebert PDN. BOLD: The Barcode of Life Data System (www.barcodinglife.org). Molecular Ecology Notes. 2007; 7:355-364.
8. Rider DA, 2006-2017. www.ndsu.nodak.edu/ndsu/rider/pentatomoidea/
9. Song N, Liang A. The complete mitochondrial genome sequence of *Geisha distinctissima* (Hemiptera: Flatidae) and comparison with other hemipteran insects. Acta Biochimica et Biophysica Sinica. 2009; 41(3):206-216.
10. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Molecular Biology and Evolution. 2013; 30(12):2725-2729.
11. Zhang QL, Yuan ML, Shen YY. The complete mitochondrial genome of *Dolycoris baccarum* (Insecta: Hemiptera: Pentatomidae). Mitochondrial DNA. 2013; 24(5):469-471.
12. Zhao G, Li H, Zhao P, Cai W. Comparative mitogenomics of the assassin bug genus *Peirates* (Hemiptera: Reduviidae: Peiratinae) reveal conserved mitochondrial genome organization of *P. atromaculatus*, *P. fulvescens* and *P. turpis*. PloS One. 2015; 10(2):e0117862.