INNSPUB

International Journal of Biosciences | IJB |
ISSN: 2220-6655 (Print), 2222-5234 (Online)
http://www.innspub.net
Vol. 13, No. 4, p. 139-148, 2018

RESEARCH PAPER OPEN ACCESS

CO1 based DNA barcoding of some pentatomomorpha bugs (Hemiptera: Heteroptera) from Swat, Pakistan

Sana Ullah[1*], Habib Ahmad[2], Muhammad Ather Rafi[3]

[1]*Departement of Zoology, Hazara University Mansehra, Khyber Pakhtunkhwa, Pakistan*

[2]*Departement of Genetics, Hazara University Mansehra, Khyber Pakhtunkhwa, Pakistan*

[3]*Departement of Zoology, Women University Swabi, Khyber Pakhtunkhwa, Pakistan*

Key words: Heteroptera; DNA Barcoding; CO1 gene; Swat; K2P divergence.

http://dx.doi.org/10.12692/ijb/13.4.139-148 Article published on October 14, 2018

Abstract

DNA barcodes are useful for the identification of species. This study is an expansion of the database of barcodes by adding some cytochrome oxidase I (COI) gene sequences of Pentatomomorpha species of Pakistan. The COI gene sequences obtained in this study were analyzed for nucleotide composition and divergence at intraspecific, interspecific and intergeneric level. Phylogenetic trees were constructed by Neighbor Joining, Maximum Likelihood and Minimum Evolution methods using the MEGA 7 software. COI gene sequences of thirteen species were deposited to Barcode of Life Data System (BOLD). Out of the total sequences for eight species, namely, *Hotea curculionoides* Herrich-Schaeffer, *Adria parvula* Dallas, *Neohalys acuticornis* Ahmad and Perveen, *Scotinophara ochracea* Distant, *Lachnesthus singalensis* Dohrn, *Rhyparothesus dudgeoni* Distant, *Neophysopelta schlanbuschi* Ahmad & Abbas, *Homoeocerus sigilatus* Stal this is the first molecular study which has generated distinct barcodes for each species. Analysis of the COI gene sequences shows that barcode gap was distinct for between species, genera and at higher taxonomic levels. An increase in the mean K2P divergence across different taxonomic levels was observed. The mean intraspecific divergence was (0.7%). The mean interspecific divergence was (11.7%) and the mean divergence between genera was (21.1%). The mean A+T contents were (66.3%). From the Neighbor Joining, Maximum Likelihood and Minimum Evolution clustering analyses, we concluded that DNA Barcoding is quite helpful for identification of Pentatomomorpha species of the area. This study encourages using DNA barcodes as complement with a morphological taxonomy for easy identification of plant bug species.

* **Corresponding Author:** Sana Ullah ✉ sunkhan13@gmail.com

Introduction

The infraorder Pentatomomorpha (Plant Bugs) constitutes of phytophagus and predatory insects with more than 10,000 species worldwide (Schuh and Slater, 1995). Many classical taxonomic protocols rely on phenotypic characters and require a detailed inspection of the specimens. Identification of damaged specimens and the remains of an organism is difficult to identify on these methods. These facts are major impediments for the assessment, conservation and management of global diversity (Hanner *et al*., 2005; Costa and Carvalho, 2007). Hebert *et al*., (2003) introduced a new concept of DNA Barcoding based on a standard molecular approach for the identification of species to overcome the problems in conventional identification. The goal of DNA Barcoding is to develop a barcode library for all species on earth (Frezel and Leblois, 2008). Non-taxonomists can easily identify specimens with rapid, inexpensive methods and accessible Barcode of Life Data System (BOLD). This system also enables the acquisition, storage, analysis and publication of DNA barcode records (Ratnasingham and Hebert, 2007). COI gene has been accepted as a practical, standardized species level barcode for animal groups (Hogg and Hebert, 2004; Hebert *et al*., 2004a; Hebert *et al*., 2004b; Ward *et al*., 2005; Costa *et al*., 2007). DNA Barcoding is being successfully utilized for molecular identification of a vast variety of insect taxa, studies analyzing plant bugs are still rare (Michael *et al*., 2014). Many plant bugs are economically very important, but hardly any molecular data are available for these species from Pakistan. Therefore, this work was undertaken to check the utility of DNA Barcoding in identification of plant bugs using COI barcodes from Swat, Pakistan.

Materials and methods

The lush green and historic Swat Valley lies between 35° 0′ 0" North latitude and 72° 30′ 0" East longitude and is part of the Khyber Pakhtunkhwa Province of Pakistan. Specimens of infraorder Pentatomomorpha were collected during June 2014 - October 2017. Extensive and intensive visits have been made for the collection of true bugs from their natural abodes by the straight line transect method. Specimens were identified using relevant literature. For the extraction of DNA Thermo scientific GeneJET Genomic DNA purification Kit (#K0721, Lot 00409902) was used. Legs were detached from ethanol-preserved specimens and grinded in liquid nitrogen using mortar and pestle. The grinded samples were collected into 1.5 ml Micro centrifuge tube and following the manufacturer's instruction the DNA was purified and stored at -20 ºC. For the amplification of COI gene, LCOI490 and HCO2198 primers were used (Folmer *et al*. 1994). PCR reaction was performed in a final volume of 20 µL containing 10 µL of Thermo Scientific Dream Taq Green PCR Master Mix (#K1081, Lot: 00355703), 1 µL forward primer, 1 µL reverse primer, 0.5 µL Taq polymerase, 1 µL template DNA and 6.5 µL nuclease free water. The amplification conditions were 1 cycle, 95 ºC for 5 minutes, 35 cycles, 94 ºC for 30 seconds, 48 ºC for 30 seconds, 72 ºC for 35 s, 1 cycles, 72 ºC for 10 minutes, hold at 4 ºC. PCR products were visualized on 1% agarose gel with EtBr staining under UV light. The confirmed PCR products were getting sequenced from BGI Bio Solutions (Hong Kong) Co., Limited. The sequences were used for the identification of species through NCBI, BLAST search and BOLD. New sequences were submitted to Barcode of Life Data System (BOLD) by utilizing the standard protocol for establishing a DNA barcode library. All the details are available on BOLD under project Heteroptera of Swat (Code: SDP602). Sequences of congeneric specimens deposited by other workers were taken from Gene Bank using the BLAST tool for comparisons and analysis. The sequences were aligned by a Clustal W option in the MEGA 7 software. All analyses were performed by using MEGA 7 software (Kumar *et al*., 2016).

Results

In this study COI sequences of thirteen species belonging to 12 genera representing five families of infraorder Pentatomomorpha were obtained. For eight species, namely, *Hotea curculionoides, Adria parvula, Neohalys acuticornis, Scotinophara ochracea, Lachnesthus singalensis Rhyparothesus*

dudgeoni, Neophysopelta schlanbuschi, Homoeocerus sigilatus this is the very first molecular study which has generated distinct barcodes for each species. CO1 gene sequences were submitted to the BOLD database (Table 1). No stop codon or frame shifts were detected, indicating that sequences were not pseudogenes (NUMTs). BLAST search was carried out which resulted in matching of our sequences with the congeneric species from GenBank with no instance of barcode sharing by members of different genera. CO1 gene of related species were obtained from Gene Bank Database for comparison. The final aligned data pertain to 38 COI sequences of 564 bp representing 32 species and 25 genera.

Sequence of *Rhynocoris marginatus* GQ229415 belonging to family Rediviidae, infraorder Cimicomorpha was taken as outgroup.

Table 1. Detail of species analyzed in the present study.

S. no.	Taxa	Collection site	Date of collection	Accession no.
1	*Hotea curculionoides*	Kuza Bandai	02 May 2016	MG298987
2	*Andrallus spinidens*	Kuza Bandai	08 August 2015	MG298982
3	*Eocanthecona furcellata*	Barikot	18 May 2014	MG298984
4	*Adria parvula*	Kuza Bandai	08 August 2015	MG298986
5	*Neohalys acuticornis*	Kuza Bandai	23 July 2017	MG298985
6	*Scotinophara ochracea*	Kuza Bandai	02 May 2016	MG298989
7	*Homoeocerus sigilatus*	Barikot	08 August 2015	MG298981
8	*Cletus bipunctatus*	Kuza Bandai	13 July 2016	MG299092
9	*Cletus punctiger*	Kuza Bandai	17 July 2016	MG299062
10	*Neophysopelta schlanbuschi*	Kuza Bandai	08 August 2015	MG298988
11	*Metochus uniguttatus*	Kuza Bandai	02 May 2016	MG298983
12	*Lachnesthus singalensis*	Kuza Bandai	12 June 2015	SDP602026-17
13	*Rhyparothesus dudgeoni*	Kuza Bandai	12 June 2016	MG299061

Table 2. Percentage of base composition in the CO1 gene sequences.

Base	A	T	C	G
Mean percentage	31.5	34.8	18.0	15.7

The final alignment showed 312 conserved sites, 252 variable sites and 233 parsimony informative sites without out group. The alignment showed 298 conserved sites, 266 variable sites and 236 parsimony informative sites including out group. The mean A +T content was 66.3% (Table 2). The data were analyzed for sequence divergence at different taxonomic levels. A hierarchical increase in K2P mean divergence among specimens at various taxonomic levels was observed Table 5 summarize divergence data (K2P distance).

Table 3. Pairwise K2P intraspecific divergence percentage.

S. no.	Taxa (Accession no.)	Taxa (Accession no.)	Divergence %
1	*Andrallus spinidens* KX351377	*Andrallus spinidens* MG298982	1.1
2	*Eocanthecona furcellata* KJ459922	*Eocanthecona furcellata* MG298984	0.7
3	*Cletus bipunctatus* KP753491	*Cletus bipunctatus* MG299092	1.1
4	*Cletus punctiger* KP753849	*Cletus punctiger* MG299062	1.1
5	*Metochus uniguttatus* GU247506	*Metochus uniguttatus* MG298983	0.0
6	*Physopelta schlanbuschi KU242578*	*Neophysopelta schlanbuschi MG298988*	0.2

Intraspecific divergence ranged from 0.0% to 1.1% with a mean of 0.7% (S.D. 0.49), while interspecific divergence ranged from 4.1% to 17.1% with a mean of 11.7% (S.D. 4.59) and intergeneric divergence ranged from 13.0% to 27.8% with a mean of 21.1% (S.D. 4.5). To resolve further the barcode gap phylogenetic trees were constructed based on three different methods. In neighbor joining (NJ), maximum likelihood (ML)

and minimum evolution (ME) trees congeneric species cluster together with closely related genera cohesively clustered together. Taxa belonging to a particular family and subfamily normally formed a coherent cluster. These analyses indicated that CO1 gene sequences are useful in identification of species.

Table 4. Pairwise K2P interspecific divergence percentage.

S. no.	Taxa (Accession no.)	Taxa (Accession no.)	Divergence %
1	*Eurygaster integriceps* KU760764	*Eurygaster maura* KM022776	4.1
2	*Eurydema dominulus* KP190206	*Eurydema ventralis* KJ541625	8.4
3	*Nezara antennata* KC135971	*Nezara viridula* KX467340	5.3
4	*Plautia stali* JN087440	*Plautia viridicollis* KU163633	17.1
5	*Homoeocerus dilatatus* GQ292215	*Homoeocerus sigilatus* MG298981	9.6
6	*Cletus bipunctatus* MG299092	*Cletus punctiger* KP753849	15.9
7	*Cletus bipunctatus* MG299092	*Cletus punctiger* MG299062	15.2
8	*Cletus bipunctatus* KP753491	*Cletus punctiger* KP753849	16.3
9	*Cletus bipunctatus* KP753491	*Cletus punctiger* MG299062	15.6
10	*Physopelta schlanbuschi* KU242578	*Physopelta gutta* KP233791	10.9
11	*Neophysopelta schlanbuschi MG298988*	*Physopelta gutta* KP233791	10.7

Discussion

DNA Barcoding is useful for quick identification and discrimination of morphologically similar species belonging to different groups. In this study, eight new barcodes were developed on the BOLD database. Habeeb and Sanjayan (2011) barcoded *Oxycarenus laetus* (Lygaeidae) from India on the bases of COI gene sequence for easy identification of the species. Kaur and Sharma (2017) barcoded 14 species of the family Pentatomidae of these seven species were barcoded for the first time. New barcodes generated in this study will be used for easier identification and authentication of these species in the future.

Table 5. Comparative data of divergence percentage.

K2P Divergence	Range %	Mean ± SD %
Intraspecific	0.0-1.1	0.7±0.49
Interspecific	4.1-17.1	11.7±4.59
Intergeneric	13.0-27.8	21.1±4.5

All the taxa of the present study analyzed showed distinct barcode gaps. For identification at species level intraspecific divergence must be less than interspecific divergence. In this study mean K2P intraspecific divergence was found lower than mean interspecific divergence. The mean intraspecific divergence (0.7%) of the present study is in agreement with the intraspecific divergence reported in other animal taxa, such as 0.27% is Birds (Hebert *et al.*, 2004b), 0.39% in marine fish (Ward *et al.*, 2005), 0.11% in mayflies (Ball *et al.*, 2005), and 0.06% in bats (Clare *et al.*, 2007). Kaur and Sharma (2017) reported 2.5 % mean intraspecific divergence with a minimum value of 0.0% for species of *Halyomorpha picus*. Jung *et al.*, (2011) analyzed COI sequences of 139 true bugs and recorded mean intraspecific divergence of 0.4 %. Tembe *et al.*, (2014) obtained 80 COI gene sequences belonging to 43 species and reported 0.4 % intraspecific divergence. The mean intrespecific divergence was (11.7%). Hebert *et al.*, (2003) suggested minimum interspecific divergence of 3% for the diagnosis of insect species. Tembe *et al.*, (2014) reported the mean interspecific divergence of 11.7% in has study on

Heteroptera from India. In the study of Kaur and Sharma (2017) the mean interspecific divergence was 11.9%. The interspecific divergence was in accordance with that reported in other studies. Hebert *et al.*, (2004b) reported 7.93% interspecific divergence in birds. Similarly, 9.93% divergence was reported in fish (Ward *et al.*, 2005), and 7.8% in bats (Clare *et al.*, 2007).

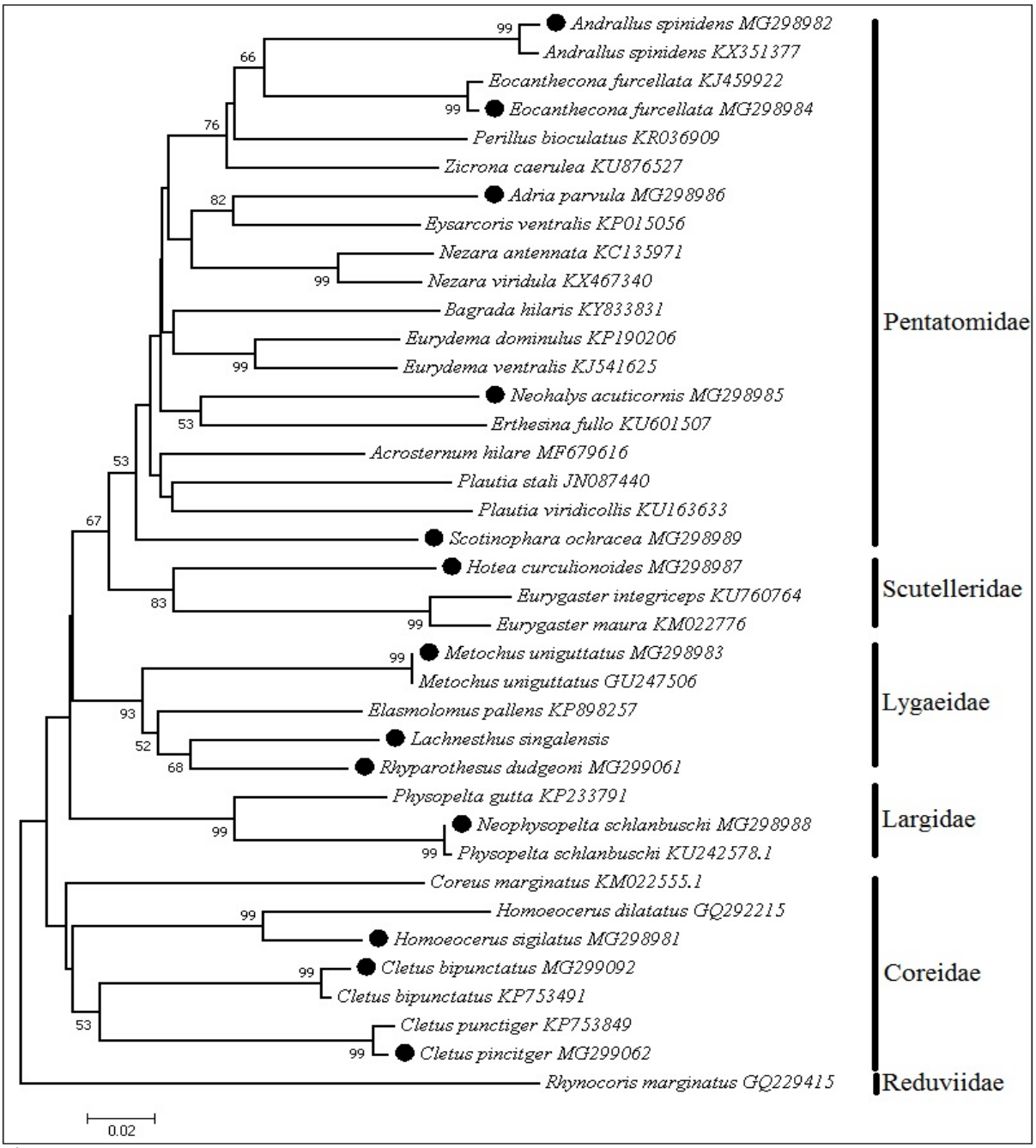

Fig. 1. Neighbor joining tree based on COI (K2P model). Numbers indicate the percentage of 1000 bootstrap replicates greater than 50. ● indicates sequences of the present study.

The present intraspecific divergence and interspecific divergence, support the results of previous studies and indicates that DNA Barcoding is useful for species identification of true bugs. The nucleotide composition of the sequences showed higher contents of A+T (66.3%). The higher contents of A+T of this study are in agreement with the reports of other authors. Habeeb and Sanjayan (2011) reported A+T contents of 67.2 % in COI gene sequence of *Oxycarenus laetus* (Hemiptera: Lygaeidae). Kaur and

Sharma (2017) analyzed COI gene sequences of 14 species belonging to family Pentatomidae and reported 65 % A+T contents. The higher contents of A+T have also been reported in other Heteroptera by Hebert *et al.*, (2003), Zhang *et al.*, (2013) and Raupach *et al.*, (2014). Baskar *et al.*, (2014) analyzed COI gene sequences of *Rhynocoris* species (Family Reduviidae) and reported higher contents of A+T (70%).

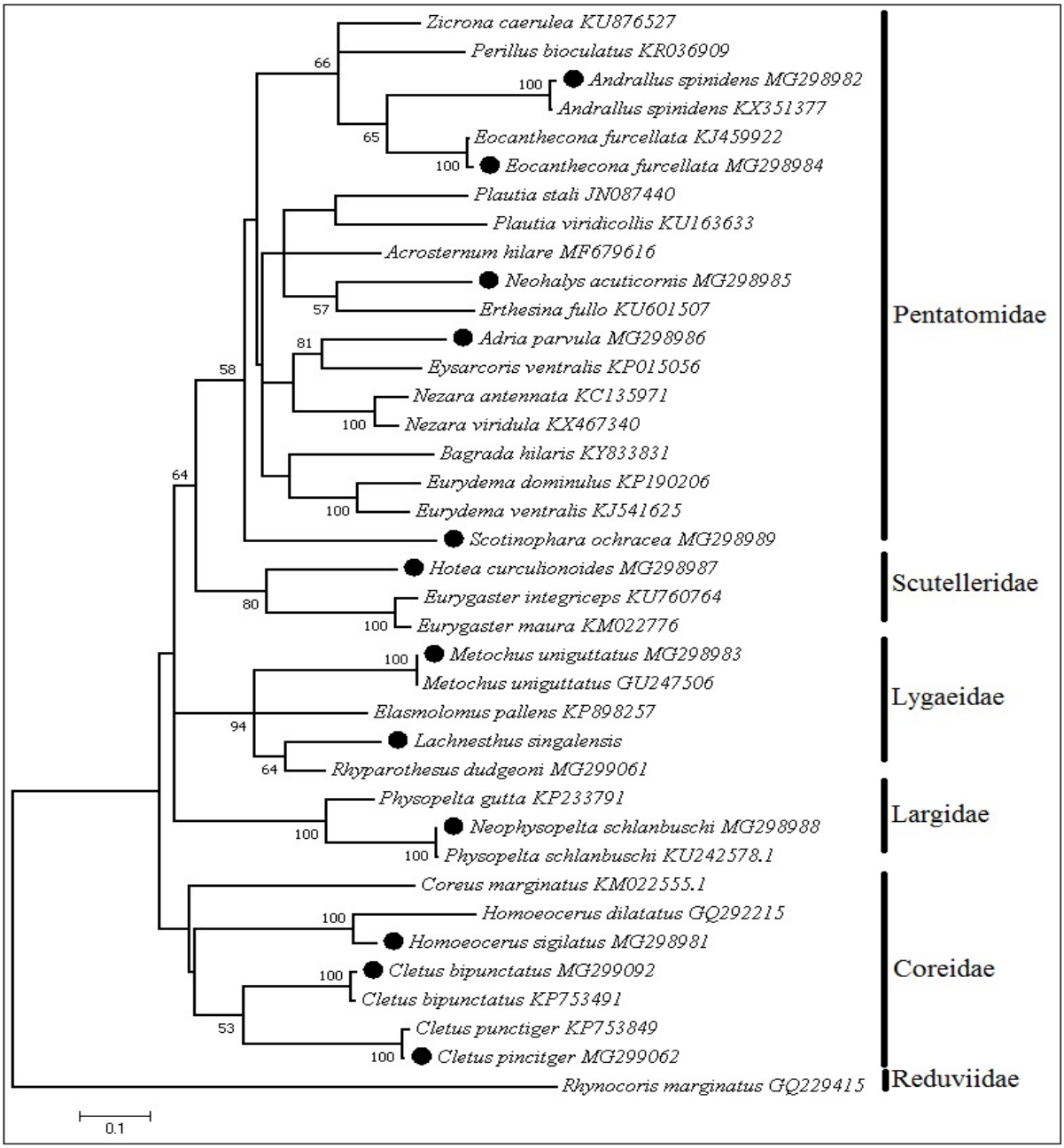

Fig. 2. Maximum Likelihood tree based on COI gene sequences. Numbers indicate the percentage of 1000 bootstrap replicates greater than 50. ● indicates sequences of the present study.

Although there are many studies in which phylogenetic relationships between families have been analyzed. Studies analyzing phylogenetic relationships at lower taxonomic units are very rare. In this study Neighbor Joining, Maximum Likelihood and Minimum Evolution analyses of plant bugs on the bases of COI gene sequences were presented to elaborate interspecific phylogenetic relationships. The NJ, ML and ME analyses showed that clustering of congeneric species were significant on COI bases.

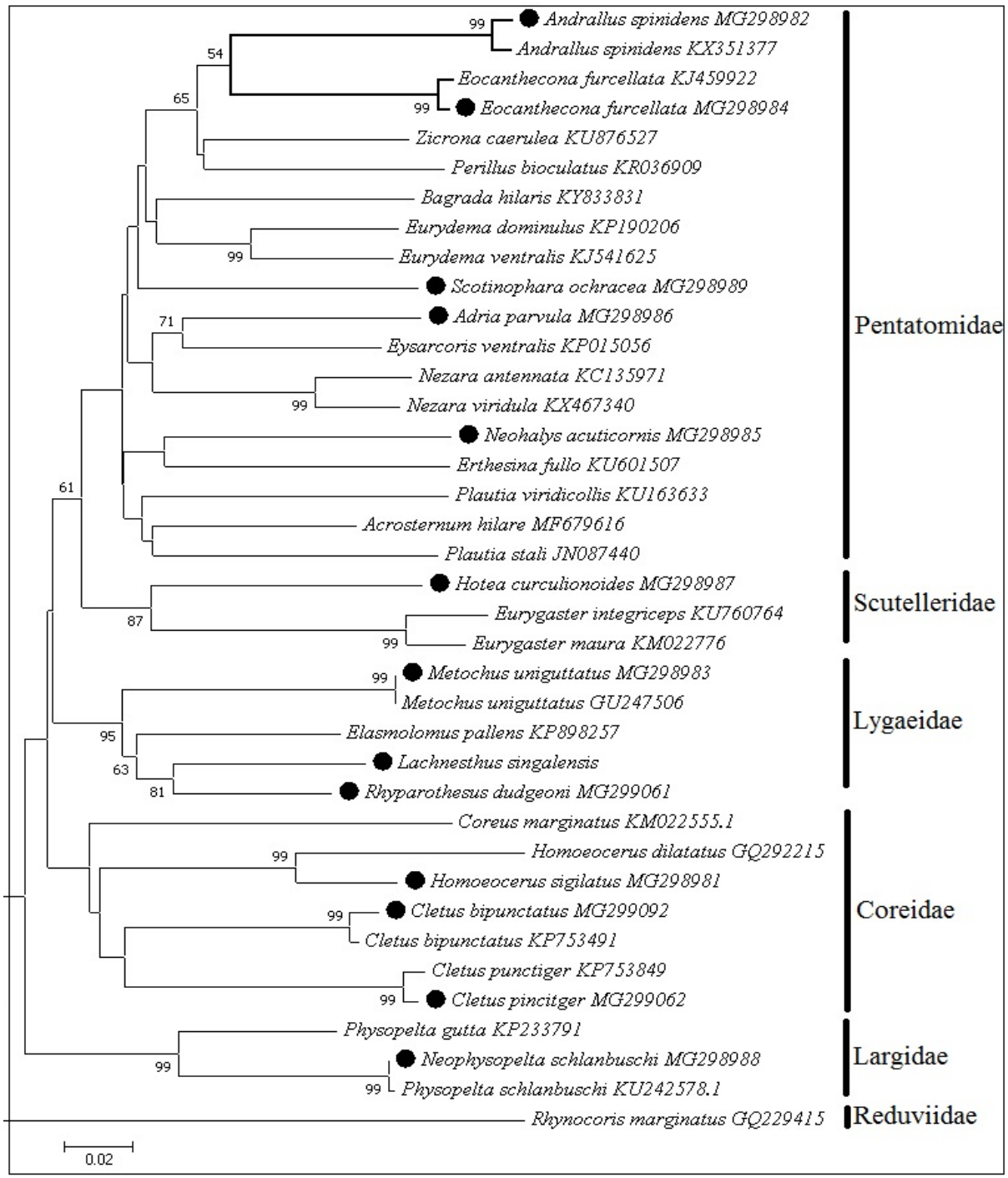

Fig. 3. Minimum Evolution tree based on COI gene sequences. Numbers indicate the percentage of 1000 bootstrap replicates greater than 50. ● indicates sequences of the present study.

In the phylogenetic trees taxa belonging to the same family clustered together with 53-97% bootstrap support. Thus intraspecific and interspecific phylogenetic relationships were clearly established. The phylogenetic analyses showed that identification of species on morphological characters and molecular bases are extremely consistent. Monophyly of family Pentatomidae was supported in all the phylogenetic trees. Xia and Zheng (2004) analyzed 432 bp sequences of Cyt b gene of eight species of the family Pentatomidae for phylogenetic relationships. They used Neighbor joining, Minimum Evolution, Maximum Like hood, Maximum parsimony models for phylogenetic relationships. They reported that all the trees showed similar topology. Barman *et al.,* (2017) analyzed genetic variation in four species of

Pentatomidae on the basis of COI gene sequences. In this study the phylogenetic analysis of the family Pentatomidae on the basis of COI gene sequences are in agreement with morphology based identification of the species. Li (2008) obtained COI gene sequences of Coreidae to elaborate phylogenetic relationships. Sequences of Reduviidae and Pyrrhocoroidea were taken as out group. He reported close relationships between genus *Coreus* and Cletus. This study contributes to the assembly of a DNA barcode library for the plant bugs.

The results of these analyses are congruent with previous studies and morphology based identification. These analyses confirm the utility of COI gene sequence for identification of true bugs and to elaborate their phylogenetic relationships. Further studies are needed to add more data to the BOLD database for authentic identification of plant bugs species.

Conclusions

The patterns of divergence and clustering of species in the phylogenetic trees confirms the utility of Barcoding in species level identification of Heteroptera species. DNA barcodes may be used for quick identification of unexplored species from Pakistan. Additional efforts are needed to make a more comprehensive reference library. This study encourages using DNA barcodes as complement with a morphological taxonomy for authentic identification of Heteroptera species.

Acknowledgement

We are thankful to Prof. Dr. Iqbal Munir, Director, Institute of Biotechnology and Genetic Engineering, The University of Agriculture Peshawar-Pakistan for providing facilities for molecular work.

References

Ball SL, Hebert PDN. 2005. Biological identifications of mayflies (Ephemeroptera) using DNA barcodes. Journal of the North American Benthological Society **24,** 508-524. http://dx.doi/abs/10.1899/04-142.1

Barman AK, Joyce AL, Torres R, Higbee BS. 2017. Assessing Genetic Diversity in Four Stink Bug Species, Chinavia hilaris, Chlorochroa uhleri, Chlorochroa sayi, and Thyanta pallidovirens (Hemiptera: Pentatomidae), Using DNA Barcodes. Journal of Economic Entomology **110,** 2590–2598. http://dx.doi.org/10.1093/jee/tox227

Baskar A, Dhanasekaran S, Tirumurugaan KG, Vignesh AR, Prakash B. 2014. Molecular Genetic Variation and Phylogeny of Genus Rhynocoris Based on mitochondrial Cytochrome c Oxidase Subunit I Gene (Heteroptera: Reduviidae). Indian Journal of Natural Science **4,** 1474-1484.

Clare EL, Lim BK, Engstrom MD, Eger JL, Hebert PN. 2007. DNA barcoding of Neotropical bats: species identification and discovery within Guyana. Molecular Ecology Notes **7,** 184-190. http://dx.doi.org/10.1111/j.1471-8286.2006.01657.x

Costa FO, Carvalho GR. 2007. The Barcode of Life Initiative: synopsis and prospective societal impacts of DNA barcoding of Fish. Genomics, Society and Policy **3,** 52-56. http://dx.doi.org/10.1186/1746-5354-3-2-29

Felsenstein J. 1985. Confidence limits on phylogenies: An approach using the bootstrap. Evolution **39,** 783-791. http://dx.doi.org/10.1111/j.15585646.1985.tb00420.x

Frezal L, Leblois R. 2008. Four years of DNA barcoding: current advances and prospects. Infection, Genetics and Evolution **8,** 727-736. http://dx.doi.org/10.1016/j.meegid.2008.05.005.

Habeeb SKM, Sanjayan KP. 2011. Sequencing and phylogenetic analysis of the mitochondrial cytochrome c oxidase subunit i of Oxycarenus laetus (Hemiptera: Lygaeidae). International Journal of Plant, Animal and Environmental Science **1,** 85-92.

Hanner R, Schindel D, Ward B, Hebert P. 2005. Fish Barcode of Life (FISH-BOL). Workshop

report University of Guelph, Ontario Canada, 5-8 June 2005. p 21.

Hebert PDN, Cywinska A, Ball SL, Waard JD. 2003. Biological identifications through DNA barcodes. Proceedings of Royal Society, London, B. **270,** 313–321. http://dx.doi.org/10.1098/rspb.2002.2218

Hebert PDN, Penton EH, Burns JM, Janzen DH, Hallwachs W. 2004a. Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly Astraptes fulgerator. Proceedings of the National Academy of Sciences USA **101,** 14812-14817. http://dx.doi.org/10.1073/pnas.0406166101

Hebert PDN, Stoeckle MY, Zemlak TS, Francis CM. 2004b. Identification of birds through DNA barcodes. PLoS Biology **2,** 1657-1663. http://dx.doi.org/10.1371/journal.pbio.0020312

Hogg ID, Hebert PDN. 2004. Biological identification of springtails (Collembola: Hexapoda) from the Canadian Arctic, using mitochondrial DNA barcodes. Canadian Journal of Zoology **82,** 749-754. http://dx.doi.org/10.1139/z04-041

Jung S, Duwal RK, Lee S. 2011. COI barcoding of true bugs (Insecta, Heteroptera). Molecular Ecology Resources **11,** 266–270. http://dx.doi.org/10.1111/j.1755-0998.2010.02945.x.

Kaur H, Sharma K. 2017. COI-based DNA barcoding of some species of Pentatomidae from North India (Hemiptera: Heteroptera). Mitochondrial DNA A. **28(5),** 756-761. http://doi.org/10.1080/24701394.2016.1180513

Kumar S, Stecher G Tamura K. 2016. MEGA 7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. Molecular Biology Evolution **33,** 1870-1874. http://dx.doi.org/10.1093/molbev/msw054.

Li P. 2008. Molecular Phylogenetic Research of Partial Species of Coreidae (Hemiptera: Heteroptera). Shaanxi Normal University.

Michael JR, Lars H, Stefan MK, Fabian D, Jerome M, Martin MG. 2014. Building-Up of a DNA Barcode Library for True Bugs (Insecta: Hemiptera: Heteroptera) of Germany Reveals Taxonomic Uncertainties and Surprises. Plos One **9,** 1-13. http://doi.org/10.1371/journal.pone.0106940.

Ratnasingahm S, Hebert PDN. 2007. BOLD: The barcode of life data system www.barcodelife.org). Molecular Ecology **7,** 355–364. http://dx.doi.org/10.1111/j.1471-8286.2007.01678.x

Raupach MJ, Hendrich L, Kuchler SM, Deister F, Moriniere J, Gossner MM. 2014. Building-up of a DNA barcode library for true bugs (Insecta: Hemiptera: Heteroptera) of Germany reveals taxonomic uncertainties and surprises. PLoS One. **9,** e106940. http://dx.doi.org/10.1371/journal.pone.0106940.

Saitou N, Nei M. 1987. The neighbor-joining method. A new method for reconstructing phylogenetic trees. Molecular Biology and Evolution. **4(4),** 406-425. http://dx.doi.org/10.1093/oxfordjournals.molbev.a040454

Schuh RT, Slater JA. 1995. True Bugs of the World (Hemiptera: Heteroptera). Classification and Natural History. Ithaca, Cornell University Press, 1-336.

Tembe S, Shouche Y, Ghate HV. 2014. DNA barcoding of Pentatomomorpha bugs (Hemiptera: Heteroptera) from Western Ghats of India. Meta Gene **2,** 737–745. http://dx.doi.org/10.1016/j.mgene.2014.09.006.

Ward RD, Zemlak TS, Innes BH, Last PR, Hebert PDN. 2005. DNA barcoding Australia's fish

species. Philosophical Transactions of the Royal Society. B **360,** 1847-1857.
http://dx.doi.org/10.1098/rstb.2005.1716

Xia DJ, Zheng ZM. 2004. Discussion on Phylogenetic relationships of partial Pentatomidae insects based on sequences of cytochrome b gene. Zoological Research **25,** 397-402.

Zhang Q, Yuan M, Shen Y. 2013. The complete mitochondrial genome of Dolycoris baccarum (Insecta: Hemiptera: idae). Mitochondrial DNA. **24,** 469–471.
http://dx.doi.org/10.3109/19401736.2013.766182.