DNA barcodes and insights into the phylogenetic relationships of butterflies of the genus *Eurema* (Pieridae) from Uttarakhand, India

Ankita Rajpoot1*; Ved Prakash Kumar2 and Archana Bahuguna1

Author (s) affiliations

1 Molecular Systematic laboratory**,** Zoological Survey of India, NRC, 218, Kaulagarh road, Dehradun 248195, Uttarakhand, India.

2 Wildlife Institute of India, Post Box # 18, Chandrabani, Dehradun 248001, Uttarakhand, India.

Corresponding author

Ankita Rajpoot
Zoological Survey of India, NRC, 218, Kaulagarh Road,
Dehradun 248195, Uttarakhand, India
Phone: +91-135- 0135-2758362 / 2754939
E.mail:anku050988@gmail.com

DNA barcodes and insights into the phylogenetic relationships of butterflies of the genus *Eurema* (Pieridae) from Uttarakhand, India

ABSTRACT

In DNA barcoding, mitochondrial gene cytochrome *c* oxidase I recommended as a tool for the rapid identification and discovery of species. Genus *Eurema*, Family Pieridae is a highly diverse and distributed along wide geographic ranges in the world as well as India including approximately 70 species throughout the world. The present study is preliminary approach, we included n=12 specimen (3 samples per species) of four different *Eurema* species, listed in IUCN as Least concern species, were collected from Uttarakhand (India), to give the DNA barcodes and examine patterns of gene evolution through molecular phylogenetics with publicly available sequences of other 17 *Eurema* species present in different countries.

The generated (n=12) COI sequences compared with the sequences of the conspecifics submitted from different geographic regions, all four species were correctly identified. The obtained COI barcodes clearly showed the intraspecific and interspecific distance among four *Eurema* species by using a K2P technique. In spite NJ, clustering analysis also successfully discriminated all four species.

In phylogenetic topology, included 21 *Eurema* species in the Bayesian and Maximum Likelihood analysis recovered all species in two monophyletic clades with strong support and resolved taxonomic position within species groups. Although higher-level relationships among other *Eurema* species groups require additional study.

Keywords: *Eurema;* Cytochromec oxidase (COI); DNA barcoding; Phylogeny; Uttarakhand

DNA barcoding (http://www.barcodinglife.org) has gained so much popularity during the last decade in documenting the global biodiversity (Hebert et al. 2003ab, Ebach and Holdrege 2005, Abdo and Golding 2007, Meusnier et al. 2008, Monaghan et al. 2009, Hajibabaei et al. 2011, Haye et al. 2012, Leray and Knowlton 2015). Cytochrome *c* oxidase subunit I (COI) is a widely used standard mitochondrial gene for DNA barcoding to facilitate identification of biological specimens (Hebert et al. 2003). DNA Barcoding is also using as a tool for discrimination between cryptic species and to discover new species based on nucleotide sequence divergence. Availability of large amount of sequences in public databases will indeed speed up species description and identification for taxonomist (Blaxter 2004, Ratnasingham and Hebert 2007). Overall the goal of DNA barcoding is explicitly to aid species identifica-

tion, it has frequently been used for phylogenetic inference at multiple taxonomic levels (Tautz et al. 2003, Savolainen et al. 2005) prompting many scientist to contemplate the phylogenetic value of DNA barcode datasets (Tautz et al. 2003, Savolainen et al. 2005, Hajibabaei et al. 2007, Wahlberg and Wheat 2008).

Uttarakhand lies in Central Himalaya at $77^0$45' East longitudes to 81^0 North longitudes, elevation ranges from 300 m to 7000 meters above sea level covering a geographical area of 53,485 sq km, has a great diversity of flora and fauna constitutes 65% of the total area of the state (Sundriyal and Sharma 2016, Uttarakhand Annual Plan 2011–12). The rich faunal and floral diversity of the state comprises 4907 faunal species (include 3948 invertebrate and 959 vertebrate species) and 5096 Floral species (included Angiosperms and Gymnosperms) (Arora and Kumar 1995, Tak and Sati 2010, Envis Newsletter 2013). High floral diversity indicates high butterfly diversity in state and butterflies is an important group of 'model' organisms used, for centuries, to investigate many areas of biological research, ecology, evolution, population genetics and developmental biology (Boggs et al. 2003).

Lepidoptera, as the second largest order of insects with more than 157,000 species, are of particular interest in systematic research. Among them, India is home to about 1800 species and subspecies of butterflies (Kunte et al. 2017) which is about 8.74% of total butterfly species in the world and constitute 65% of total Indian fauna. Uttarakhand harbours 407 species of butterflies (Singh and Sondhi 2016). Due to their beautiful colours and high diversity, butterflies are well studied through morphologically, systematically, and ecological perspective in India as well as Uttarakhand (Kunte 2000, Khan et al. 2004, Kumar 2008, Jain and Jain 2012, Bhardwaj et al. 2012, 2013, Kumaraswamy and Kunte 2013, Kunte et al. 2015). Family Pieridae is highly diverse and distributed along wide geographic ranges. Unfortunately, there are few molecular studies has been reported on butterflies from India (Ragupathy et al. 2009, Gaikwad et al. 2011).

Genus *Eurema (*Hübner 1819) is highly diverse genus of family Pieridae, distributed along wide geographic ranges in the world as well as India including approximately 70 species throughout the world (*Eurema*, page). *Eurema* butterflies are comparatively small butterflies with lemon yellow wings, bordered with black, especially on the upper side of both wings, commonly called grass yellow butterflies. The undersides of the wings, which are usually light in colour than the upper side, are marked with several black or faint brown markings (Osamu Yata 1991, Jeratthitikul et al. 2009). There are number of molecular studies available on different family of butterflies including DNA barcoding, molecular phylogeny and phylogeography aspect (Scott 1985, Brower and DeSalle 1998, Zimmermann et al. 2000,

Xue 2009, Xiangqun et al. 2015, Martin et al. 2017). Although few molecular studies available on family Pieridae (Braby et al. 2006, Solovyev et al. 2015).

In India there are eight species reported under the genus *Eurema* such as *Eurema brigitta* Cramer, 1780 (Small grass yellow); *E. laeta* Boisduval, 1836 (Spotless grass yellow); *E. lacteola* Distant,1886 (Scare grass yellow); *E. andersonii*, Moore (One-spot grass yellow); *E. hecabe* Linnaeus, 1758 (Common grass yellow); *E. blanda* Boisduval, 1836 (Three-spot grass yellow), *E. nilgiriensis* Menon, 1987 (Nilgiri grass yellow) and *E. simulatrix* Semper, 1891 (Scarce Changeable Grass Yellow) (Kunte 2017). Out of these, five species (*E. brigitta*, *E. laeta*, *E. andersonii*, *E. hecabe*, and *E. blanda)* are reported from Uttarakhand (Smetacek 2013, Singh and Sondhi 2016). Among, these two species *(E. brigitta and E. andersonii*) listed as least concern in IUCN (Larsen 2011, Muller and Tennent 2011). In present study, we included five *Eurema* species (4 from current study and 1 from Genbank) from Uttarakhand to develop an authentic genetics reference library based on DNA barcoding approach. Furthermore, we conducted phylogenetic analysis to clarify the relationships between the other species of genus *Eurema*, which is globally distributed through using the sequences available in GenBank.

MATERIALS AND METHODS

Sampling for Genetic analysis:

N=12 samples (3 samples per species) from each four species of genus *Eurema*, were collected from various localities in the Mussoorie and Dehradun, Uttarakhand, India (Fig.1) Additionally, we also used another (n=78) sequences of 17 *Eurema* species including a Uttarakhand specie (*E. andersoni*) from NCBI Genbank for the phylogenetic analysis (Table 1). Butterflies were collected by using the sweeping net and dead or damaged animals then placed in a separate glassine envelope stored at -20°C until further use. One or more legs removed, kept in ethanol for DNA isolation. For initial identification, we evaluate the wing marking patterns of the all species found in Uttarakhand except *E. andersonii* (A rare species, Reported in 2003, 2004 and 2015 (Singh and Bhandari 2003, Singh and Sondhi 2016). *It* earlier reported by Mackinnon and de Nicéville (1899) from Mussoorie, Uttarakhand.

Laboratory procedure

The total genomic DNA was extracted from the leg samples (n=12) using Qiagen DNeasy Tissue Kit (QIAGEN, Germany) protocol; a partial fragment of COI (Folmer et al. 1994) was amplified in an Eppendorf Master Thermocycler. Polymerase chain reaction (PCR) master mix preparation and PCR cycling condition performed according to the previously published

article (Rajpoot et al. 2017). To check the contamination of the DNA, a negative control was set up with a PCR master mix. Sanger sequencing performed by a commercial using the same primers. All obtained sequences were good for both the reverse and forward primers. Chromatograms in both directions compared using Codon Code Aligner 3.9 (Codon Code Corporation) and automatic base calls checked along the sequence, both where the two sequences were in disagreement and elsewhere.

Data analysis

The sequences obtained were validated and cleaned using Cromas 2.6.4 (http://www.technelysium.com.au), multiple sequence alignment with high accuracy and high throughput sequences of equal length were generated after multiple sequence alignment (MSA) using MUSCLE (Edgar 2004), sequences characteristic were performed manually software MEGA 7 (Kumar et al. 2016). FASTA format of these species sequences conformed by BLAST search at NCBI. Moreover, recheck the BLAST result, neighbor-joining (NJ) (Saitou and Nei 1987) with bootstrap test (1000 replicates) (Felsenstein 1985) using to see the closet taxa. The Kimura 2-parameter (K2P) model (Kimura 1980) of base substitution was used to calculate pairwise genetic distance in MEGA 7 software (Kumar et al. 2016). Additionally, to check the performance of DNA barcoding, we downloaded sequences of same species from Genbank originated from different geographical areas (Table 1). The intra and interspecies nucleotide divergence and haplotypes variation calculated using MEGA 7 (Kumar et al. 2016).

Phylogenetic analysis was performed using MrBayes v3.1 (Ronquist et al. 2011) to establish relationships based on Bayesian Inferences (BI), while a Maximum Likelihood (ML) approach was implemented in PhyML v3.1 (Guindon et al. 2010). The Akaike Information Criterion (AIC) in jModelTest v2.1.3 (Darriba et al. 2012) employed to determine the best-fit model of sequence evolution. The same model selected for BI and ML analysis as it proved to be the best-fit model calculated for each analysis. Nodal support for the ML analyses carried out with 1000 bootstrap replications. The BI calculations were run over two million generations, after which 25% of the trees were discarded as burn-in. Tracer v1.5 (Drummond and Rambaut, 2007) software was used to assess trace files generated by MrBayes v3.1, in order to assess whether mixing was achieved and to choose a suitable percentage burn-in. Moreover to find out the more accurate tree topology *Pontia edusa (KY846094.1)* was used as an outgroup (Xiangqun et al. 2015).

RESULT AND DISCUSSION

Morphological identification

The collected 12 specimens of genus *Eurema* species first confirmed based on the morphological characteristics. Identification done by observing wing shape and colour pattern described in available keys/identification guides (Antram 1924, Peile 1937, Gunthilagaraj et al. 1998, Kunte 2000, Rangnekar 2007).

Analysis of molecular dataset

Isolation, PCR amplification and sequencing of DNA from leg of butterflies yielded good quality sequences. We generated DNA barcode sequences of 12 specimens representing four distinct, morphologically identifiable species, belonging to a single genus *Eurema*, the obtained length of these COI sequences were 565 bp. All sequences generated in this study submitted in GenBank (sequence submitted but waiting for Accession number) for future uses.

Most of our sequences 98%-100% matched with the COI sequences of respective species, which already deposited in NCBI GenBank databases (Table 2). The obtained sequences of 565 bp (COI gene) of 12 specimen contained 409 Conserved regions, 92 Variable sites, 75 Parsimony informative sites, 16 Singleton sites and 8 haplotypes (Table 2). The maximum likelihood estimate of transition /transversion bias (R) were 1.60 and nucleotide composition were A 38.6, T 30.6, C 13.6, G 17.1 (Fig.2). Furthermore, the overall haplotypes (hd) and nucleotide diversity (π) was 0.923 and 0.08807 respectively, while, the overall sequences divergence (k) among all sequences was 0.93% and overall nucleotide difference between four species was 43.765 (Table 2).

Determination of Intraspecific and Interspecific K2P distances

The obtained COI sequences clearly showed the intraspecific and interspecific distance among four *Eurema* species by using a K2P technique. The interspecific variation in *Eurema*, mean pairwise sequences divergences (d) of COI ranged from 0.061% (*E. blanda/ E. hecabe*) to 0.131% (*E. hecabe*/*E. leata*) with an average genetic distance of 0.84%. Although, *d* value 0.118% and 0.112% occur in *E. blanda*/*E. brigitta* and *E. blanda*/*E. leata* respectively, *d* value 0.114% and 0.119% occur in *E. hecabe*/*E. brigitta* and *E. brigitta*/*E. leata* correspondingly. In addition, observed overall interspecific mean diversity among four species was 0.06% (Table 3).

The intraspecific variation, *d* value ranged from 0.002% *(E. blanda*) to 0.051% (*E. hecabe*) with an average genetic distance 024%. However, d value 0.003% and 0.041% oc-

curred in *E. brigitta* and *E. leata* respectively. Moreover, observed overall intraspecific mean diversity was 0.09% (Table 3).

To reconfirmed, the interspecific and intraspecific evolutionary divergence within Uttarakhand *Eurema* species, K2P based NJ tree topology showed three well supported monophyletic clades, where *E. blanda* and *E. hecabe* clustered together in the first clade with 65% to 100% bootstrap value, *E. andersonii* present in second clade with 95% bootstrap value from clade first, while *E. brigitta* and *E. leata* clustered together in third clade with 64% to 100% bootstrap value. Overall, tree topology supported the interspecific evolutionary divergences result where all five species showed barcoding variation (Fig. 3). Moreover, variation of species-specific sites between five species given in Table 5.

Evolutionary divergence and phylogenetic relationship among twenty-one Eurema species

Calculation of mean pairwise sequences divergences (d) between different *Eurema* species showed that they ranged from 0.1% (*E. l .lacteola/E. hecabe/E. ada)* to 0.15% (*E. brigitta/E. lisa; E. nicippe/E. s. techmessa; E. s. techmessa/E. boisduvaliana*/ *E. lisa*) with an average genetic distance of 0.10% (Table 4).

The phylogenetic trees generated from COI sequences dataset by BI and ML approach tree were identical therefore; they merged into one tree depicting each tree individual support values. In tree topology, all twenty-one *Eurema* species recovered in two major clades (A and B) (Fig. 4). The clade A recovered as a monophyletic clade with strong bootstrap, although only seven of twenty-one species covered with two subclade (I and II) and was placed in the basal position as the sister to the rest of the clades of the *Eurema* species. In subclade I, only two species (*E. a. marginella* and *E. albula*) were present with strong bootstrap support (95%-100%), while subclade II covered five species (*E. nicippe, E. xantochlora, E. Mexicana, E. boisduvaliana* and *E. lisa*) with different sister clade although bootstrap support less (79%- 84%).

The clade B was a monophyletic clade and strongly supported bootstrap value, although fourteen *Eurema* species were also recovered into eight subclade (III-X) with supported bootstrap values. In subclade III, three of fourteen species (*E. hecabe, E. lecta* and *E.ada*) clustered with strong support value (95%-100%), although *E. lecta* and *E.ada* showed less support value (65%-79%), while in subclade V covered four species (*E. simulatrix, E. s. Sodalist, E. blanda* and *E. s. tecmessa*) with strong support value (95%-100%), though, two species (*E. simulatrix* and *E. blanda*) showed lesser support value (69%-75%). The subclade VIII and X consisted with four, two of each species i.e. (*E. puella* and *E. smilax*) and (*E.*

brigitta and *E. leta*) with strongly supported bootstrap value (80%-94% and 95%-100%) respectively, while subclade IV, VI, VII and IX consisted with each species with high support value (95%-100%) respectively (Fig.4).

Butterflies are affably the most attractive group of invertebrates and have been a source of inspiration for generations of natural historians and scientists. Subsequently, the classification of Lepidoptera (generic- and specific-level) is reasonably stable and the majority of taxa have named (Ackery et al. 1999). In the family Pieridae, *Eurema* with morphological synapomorphy is relatively unique and easy to be distinguished from the remaining genera. The five species (*E. hecabe*, *E. leata*, *E. brigitta*, *E. andersoni* and *E. blanda*) out of ca70 recognized species under the genus are distributed in Uttarakhand. In these DNA sequences of CO1 gene revealed that the obtained COI sequences are very helpful to discriminating the butterfly species of genus *Eurema*. The high interspecific and intraspecific variation between and within the Uttarakhand species support the unique DNA barcode of these species.

CONCLUSION

The DNA sequences of COI gene revealed that the obtained COI sequences are very helpful to discriminating the butterfly species of genus *Eurema*. The present study samples correctly fall with the respective *Eurema* species without any ambiguities. Moreover, NJ clustering analysis of five Uttarakhand species showed well supported monophyletic clade of the sequences belonging to the same species without any overlap, even though these sequences are from the specimens separated by a large geographic distances (Fig 2, 3).

Some previous studies support long distance dispersal of butterflies (Craft et al. 2010, Gaikwad et al. 2011). Low sequence divergence and the sequences from the conspecifics butterflies of different biogeographically originated samples lies in monophyletic clades suggest long distance dispersal in Lepidoptera as well as in subjected genus in our study. This explains feasibility of DNA barcoding for species identification even when the species are separate by large geographic areas. Phylogenetic study of this genus, the two major clade (A and B) were confirmed, where the clade A represent the species of the Palearctic eco zone and the clade B represent species of the Indio-Malayan eco zone. Genus *Eurema* have large number of species (~70 species) and we included 21 species due to the lack of sequences, so there is need for further study of this genus to estimate the evolutionary pattern and deep phylogenetic relationship.

COMPETING INTEREST

The authors declare that they have no competing interest.

ACKNOWLEDGMENT

The authors are grateful to Dr. K. Chandra, Director, Zoological Survey of India, Kolkata for their support and encouragement. We are very grateful to Mr. Pramod Kumar, Senior Zoologist, NRC, ZSI for helping in the initial identification of species. Authors also acknowledge the support provided by Officers and staff of NRC, ZSI, Dehradun. The authors also acknowledge the Department of Science and Technology (DST), Government of India, New Delhi for financial support under the Women Scientist Scheme (WOS-B) (DST/Disha/SoRF/022/2013).

LITERATURE CITED

Abdo, Z., and G.B.Golding. 2007. A step toward barcoding life: a model-based, decision-theoretic method to assign genes to preexisting species groups. Syst Biol. 56: 44–56.

Ackery, P.R., R.de Jong, and R.I. Vane-Wright. 1999. The butterflies: Hedyloidea, Hesperoidea and Papilionoidea. In Lepidoptera, moths and butterflies.1. Evolution, systematics and biogeography. Handbook of zoology 4 (35), Lepidoptera (ed. N. P. Kristensen), 263–300. Berlin: de Gruyter.

Arora, G.S., and Kumar A. 1995. Fauna of western Himalaya, U.P. Pt. 1. Zoological Survey of India, Himalayan ecosystem series.

Antram, C.B. 1924. Butterflies of India. Mittal Publication, New Delhi.

Bhardwaj, M., and V.P. Uniyal. 2013. High-altitude butterfly fauna of Gangotri National Park, Uttarakhand: Patterns in species abundance, composition and similarity. ENVIS Bulletin on Wildlife and Protected Areas-Arthropods and their Conservation in India (Insects & Spiders), 38–48.

Bhardwaj, M., V.P., Uniyal, A.K., Sanyal, and A.P. Singh. 2012. Butterfly communities along an elevation gradient in the Tons valley, Western Himalayas: Implications of rapid assessment of insect conservation. J. Asia-Pacific Entoml. 15(1): 207–217.

Braby, M.F., R. Vila, and N.E. Pierce. 2006. Molecular phylogeny and systematics of the pieridae (lepidoptera: papilionoidea): higher classification and biogeography. *Zool. J Linnean Society* , 2006, 147 , 239–275.

Blaxter, M.L. 2004. The promise of DNA taxonomy. Philos Trans R. Soc. Lond. Biol. Sci. 359:669–679.

Boggs, C.L., W.B.Watt, and P. R. Ehrlich. 2003. Butterflies: evolution and ecology taking flight. Rocky mountain biological lab symposium series. University of Chicago Press.

Brower, A.V.Z. and R. DeSalle. 1998. Patterns of mitochondrial versus nuclear DNA sequence divergence among nymphalid butterflies: the utility of wingless as a source of characters for phylogenetic inference. *Insect Mol. Biol* **7:** 73-82.

Bagolini, B., R.L. Dal., A. Lippert, and H. Nothdurfter. 1995. The Man in the Ice: Funding 1992 at the Tisejoch, Gem. Schnals, South Tyrol. In: Spindler K, Rastbichler-Zissernig

E, Wilfing H, for the Nedden D, Nothdurfter H, editors. The man in the ice. New findings and results.Wien: Springer-Verlag. 3-22.

Chevenet, F., C. Brun., A.L. Banuls, B. Jacq, and R. Chisten. 2006. TreeDyn: towards dynamic graphics and annotations for analyses of trees. BMC Bioinformatics.7:439. (PubMed)

Castresana, J. 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol. Biol. Evol. (4):540-52. (PubMed)

Envis Newsletter. 2013. Diversity of "Flora and Fauna of Uttarakhand" Vol.-9 Issue-2 Oct-Dec., 2013.

Ebach, M.C., and C. Holdrege. 2005. DNA barcoding is no substitute for taxonomy. Nature 434: 697.

Edgar, R.C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32(5):1792-7. (PubMed)

Eurema,page on Markku Savela's butterfly website http://www.nic.funet.fi/pub/sci/bio/life/insecta/lepidoptera/ditrysia/papilionoidea/pieridae/coliadinae/eurema/index.html

Folmer, O., M. Black., W. Hoeh., R. Lutz, and R. Vrijenhoek. 1994. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Mol. Mar. Biol. Biotechnol, 3:294-299.

Felsenstein, J. 1985. Confidence limits on phylogenies: An approach using the bootstrap. Evolution 39:783-791.

Gaikwad, S.S., H.V. Ghaskadbi., M.S. Patole, and Y.S. Shouche. 2011. DNA barcoding of nymphalid butterflies (Nymphalidae: Lepidoptera) from Western Ghats of India. Mol. Biol. Rep. 39:2375-2383. doi: 10.1007/s11033-011-0988-7.

Guindon, S., J. F. Dufayard., V. Lefort., M. Anisimova., W. Hordijk, and O. Gascuel. 2010. New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. Syst. Biol. 59(3):307-21. (PubMed).

Gunthilagaraj, K., T.N.A. Perumal., K. Jayaram., and M. Kumar. 1998. Some South Indian Butterflies Field Guide Published under project lifescape. Indian Acad. Scie, Bangalore.

Haye, P.A., N.I. Segovia., R. Vera., I.A. MdGallardo, and C. Gallardo-Escárate. 2012. Authentication of commercialized crab-meat in Chile using DNA barcoding. Food Control, 25: 239-244. doi: 10.1016/j.foodcont. 2011.10.034

Hajibabaei, M., S. Shokralla., X. Zhou., G.A.C.Singer, and D.J. Baird. 2011. Environmental barcoding: a next-generation sequencing approach for biomonitoring applications using river benthos. PLoS One, 6: e17497. doi:10.1371/journal.pone.0017497 (PMID: 21533287).

Hajibabaei, M., G.A.C. Singer., P.D.N., Hebert, and D.A. Hickey. 2007. DNA barcoding: how it complements taxonomy, molecular phylogenetics and population genetics. Trends in Genetics 23: 167-172.

Hebert, P.D.N., A. Cywinska., S.L. Ball, and J.R. DeWaard. 2003a. Biological identifications through DNA barcodes. Proc. R. Soc. Lond. B. 270: 313-321.

Hebert, P.D.N., S. Ratnasingham and J.R. deWaard. 2003b. Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. Proc. R. Soc. B. 270 (Suppl): 96-99.

Huelsenbeck, J.P. and F. Ronquist. 2001. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinfor*. **17,** 754–755.

Jain, N. and A. Jain. 2012. Butterfly diversity of Hadoti Region, Rajasthan, India. Flora Fauna. 18(2): 274-276.

Jeratthitikul, E., A. Lewvanich., B.A. Butcher and C. Lekprayoon . 2009. A Taxonomic Study of the Genus *Eurema* Hübner, [1819] (Lepidoptera: Pieridae) in Thailand. The Natural History Journal of Chulalongkorn University 9(1): 1-20.

Kunte, K., S. Sondhi, and P. Roy. (eds.) 2017. Butterflies of India, *v. 2.30.* Indian Foundation for Butterflies.

Kunte, K., P. Roy., S. Kalesh and U. Kodandaramaiah. 2015. Butterflies of India, v. 2.10. Indian Foundation for Butterflies.

Kunte, K. 2000. Butterflies of peninsular India. Universities Press (India) Limited, Hyderabad.

Kumar. S., G. Stecher and K. Tamura. 2016. MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol. Biol. and Evol. 33:1870-1874.

Kumar, P. 2008. Handbook on Common Butterflies of Uttarakhand . 1-136. (Published by the Director, Zool. Surv. India, Kolkata), ISBN 97-81-8171-198-4.

Kumaraswamy, S and K, Kunte. 2013. Intergrating biodiversity and conservation with modern agricultural landscapes. Biodivers Conserv. 22: 2735-2750.

Khan, M.R., A. Khurshid., B. Ikram, A.I. Malik, and A. Mir. 2004. Biodiversity of butterflies from district Pooch and Sudhnoti, Azad Kashmir. Asian J. Plant Sci. 3(5): 556-560.

Kimura, M. 1980. A simple method to estimating evolutionary rate of base substitution through comparative study of nucleotide substitution. Mol. Biol. Evol. 16, 1111-1120.

List of butterflies of India (Pieridae).2017. https://en.wikipedia.org/wiki

Leray, M. and N. Knowlton. 2015. DNA barcoding and metabarcoding of standardized samples reveal patterns of marine benthic diversity. Proc. Nat. Acad. Sci. USA, 112: 2076-2081. doi: 10.1073/pnas. 1424997112 (PMID: 25646458).

Larsen, T. 2011. *Eurema brigitta*.The IUCN Red List of Threatened Species 2011:m e.T159946A5344963. Reference1.docx. Downloaded on 12 September 2017.

Martin, S.H., M. Möst, W.J. Palmer, C. Salazar" W.O. McMillan., F.M. Jiggins, and C.D. Jiggins. 2017**.** Natural Selection and Genetic Diversity in the Butterfly *Heliconius melpomene***.** Genetics, 203(1): 525-541. doi: 10.1534/genetics.115.183285

Muller, C.J. and W.J. Tennent. 2011. *Eurema andersoni*. The IUCN Red List of Threatened Species 2011: e.T159801A5338505. http://dx.doi.org/10.2305/IUCN.UK.2011-2.RLTS.T159801A5338505.en. Downloaded on 12 September 2017.

Monaghan, M.T., R. Wild., M. Elliot., T. Fujisawa., M. Balke et al. 2009. Accelerated species inventory on Madagascar using coalescent-based models of species delineation. Syst. Biol. 58(3): 298-311.

Meusnier, I., G.A. Singer., J.F. Landry., D,A. Hickey.,, P.D.N. Hebert, et al. 2008. A universal DNA mini-barcode for biodiversity analysis. BMC Genomics. 9:214.

Mackinnon, P.W. and L, De Niceville. 1899. List of butterflies of Mussoorie in the Western Himalayas and neighbouring regions. J. Bombay Natu. Hist. Soci. 11: 205-221, 368-389, 585-605.

Peile, H.D. 1937. A guide to collecting butterflies of India. John Bale Sons & Danielsson Ltd, London.

Rajpoot, A., V.P. Kumar., A. Bahuguna., P. Maity, and D. Kumar. 2017. Araniella cucurbitina: the first molecular evidence of a Palearctic species of genus Araniella inhabiting India, Mitochondrial DNA Part A. 28: 1-9. doi:10.1080/24701394.2017.1373105.

Ragupathy S., S.G. Newmaster, M. Murugesan, and V. Balasubramaniam. 2009. DNA barcoding discriminates a new cryptic grass species revealed in an ethanobotany study by the hill tribes of the Western Ghats in Southern India. Mol. Ecol. Res. 9:164–171.

Rangnekar, P. 2007. The Butterflies of Goa. Mineral Foundation of Goa, Goa.

Ratnasingham, S. and P.D.N. Hebert. 2007. BOLD: the barcode of life data system (wwwbarcodinglifeorg). Mol. Ecol. Notes. 7:355-364.

Singh, A.P. and S. Sondhi. 2016. Butterflies of Garhwal, Uttarakhand, western Himalaya, India. J. Threat. Taxa, Volume: 8(4): 8666-8697.

Sundriyal, M. and B. Sharma. 2016. Status of biodiversity in central Himalaya. Appl. Ecol. and Envier. Sci. Vol. 4, No. 2, pp 37-43. Reference1.docx

Solovyev, V.I., Y. Ilinsky and O.E. Kosterin. 2015. Genetic integrity of four species of *Leptidea* (Pieridae, Lepidoptera) as sampled in sympatry in West Siberia. Comp. Cytogen, 9 (3), 299-324.http://doi.org/10.3897/CompCytogen.v9i3.4636

Smetacek ,P. 2013. Butterfly diversity of Uttarakhand: Checklist, Recent Work and future goals. http//www.flutters.org/home/docs/list_of_butterflies_of_uttarakhand_peter_smetacek.pdf.

Singh, A.P. and R.S. Bhandari. 2003. Butterfly diversity in tropical moist deciduous sal (Shorea robusta) forests of Dehradun Valley: the lower western Himalayas. Indian Fore. 129(10): 1257-1269.

Saitou, N. and M. Nei. 1987. The neighbor-joining method: A new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4:406-425.

Scott, J.A. 1985. The phylogeny of butterflies (Papilionoidea and Hesperioidea). J. Reser. on Lepidoptera 23: 241–281.

Tamura, K., G. Stecher., D. Peterson., A. Filipski, and S. Kumar. 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0 Mol. Biol. Evol. 30, 2725-2729.

Tak, P.C. and J.P. Sati. 2010. Aves. In: Fauna of Uttarakhand, State Fauna Series, (ZSI, Kolkata), 18(1), 505-532.

Tautz, D., P. Arctander., A. Minelli., R.H. Thomas and A.P. Vogler (2003) A plea for DNA taxonomy. Trend in Eco. and Evol. 18: 70-74.

Uttarakhand Annual Plan. 2011–12. Finalisation Meeting Between Hon'ble Dputy Chairman, Planning Commission & Hon'ble Chief Minister, Uttarakhand". http//:www.planningcommission.nic.in Retrieved 3 July 2017.

Ward, R.D., R. Hanner., and P.D.N Hebert. 2009. The campaign to DNA barcode all fishes, FISH-BOL J. Fish Biol. 74: 329–356.

Wahlberg, N. and C.W. Wheat. 2008. Genomic outposts serve the phylogenomic pioneers: Designing novel nuclear markers for genomic DNA extractions of Lepidoptera. Syste. Biol. 57: 231–242.

Xiangqun ,Y., K. Gao, F. Yuan., P.Wang, and Y. Zhang. 2015. Phylogenetic relationships of subfamilies in the family nHesperiidae (Lepidoptera: Hesperioidea) from China. Sci. Rep. 5, 11140; doi: 10.1038/srep11140.

Xue, G. X. 2009. Taxonomy and Phylogeny of Hesperiidae (Lepidoptera: Hesperioidea) from China. (Dissertation for Doctoral Degree.

Yata, O. 1991. A Revision of the Old World Species of the Genus *Eurema* Hubner (Lepidoptera, Pieridae). Part II. Description of the *smilax*, the hapale, the ada and the *sari* (part) groups. Bull. Kitakyushu Mas. Nat. Hisl. 10: 1-51.

Zimmermann, M., N. Wahlberg and H. Descimon. 2000. Phylogeny of *Euphydryas* checkerspot butterflies (Lepidoptera: Nymphalidae) based on mitochondrial DNA sequences data. Ann. Entomol. Soc. Amer. **93,** 347–355.

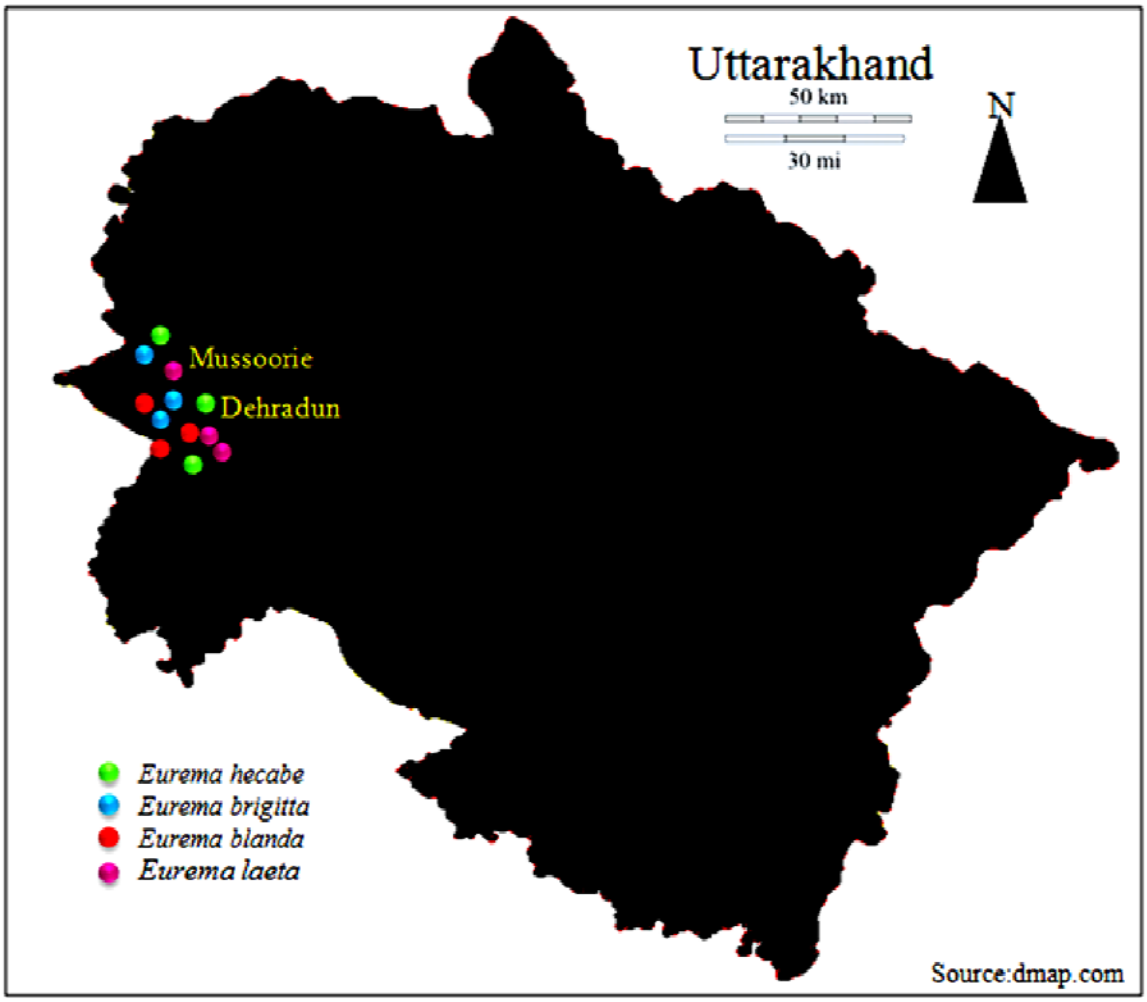

Figure 1. Sampling location of four *Eurema* species included in present study from Uttarakhand.

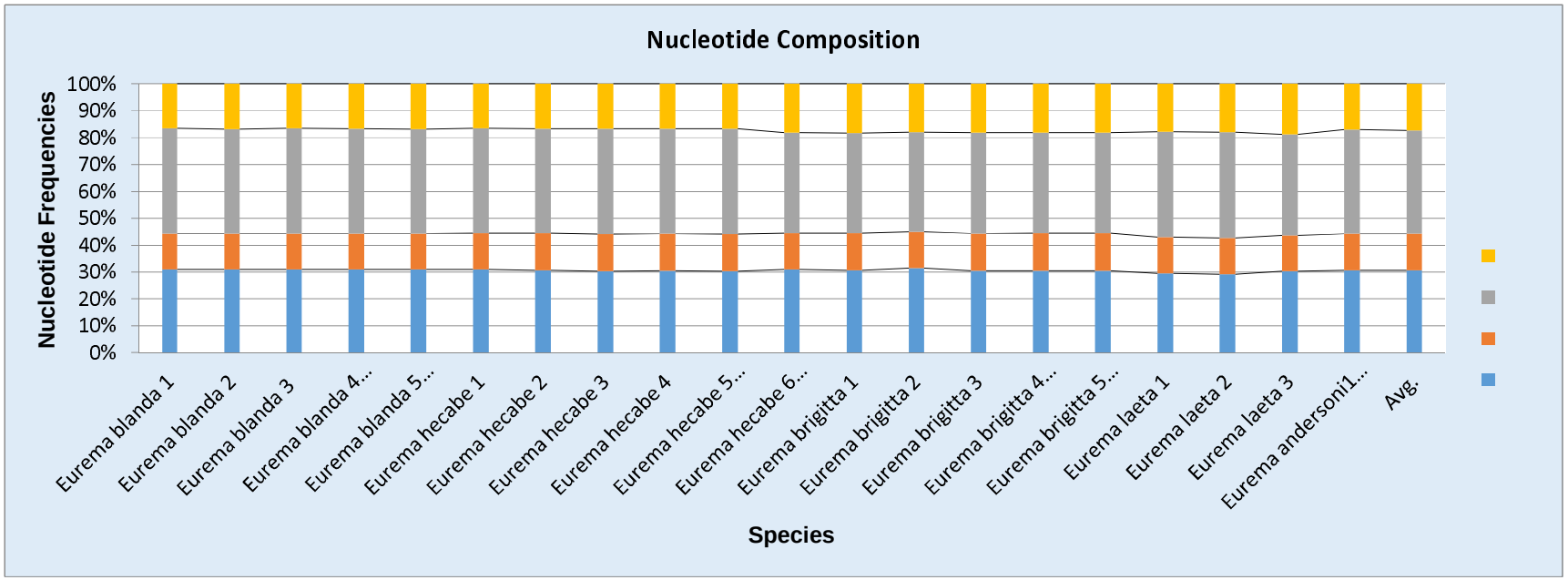

Figure 2: Observed nucleotide frequencies among 5 *Eurema* species (from Uttarakhand) in mitochondrial COI gene

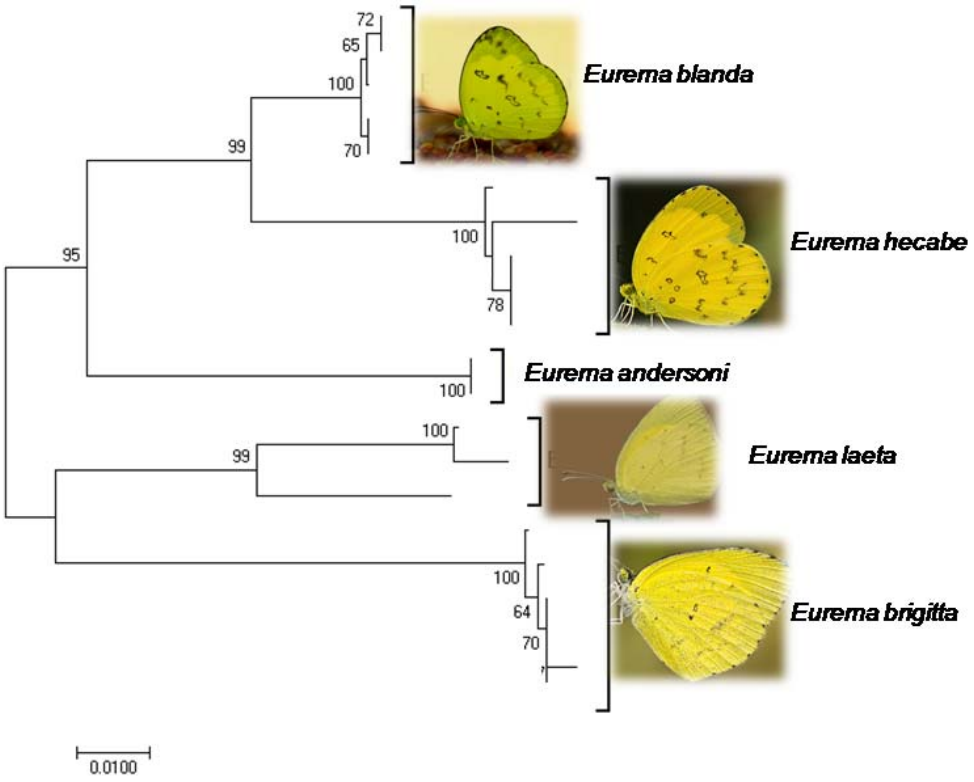

Figure 3: The COI sequences based Neighbor-Joining (NJ) tree among 5 *Eurema* species from Uttarakhand. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) shown next to the branches. The evolutionary distances were computed using the Kimura 2-parameter method and are in the units of the number of base substitutions per site.

.

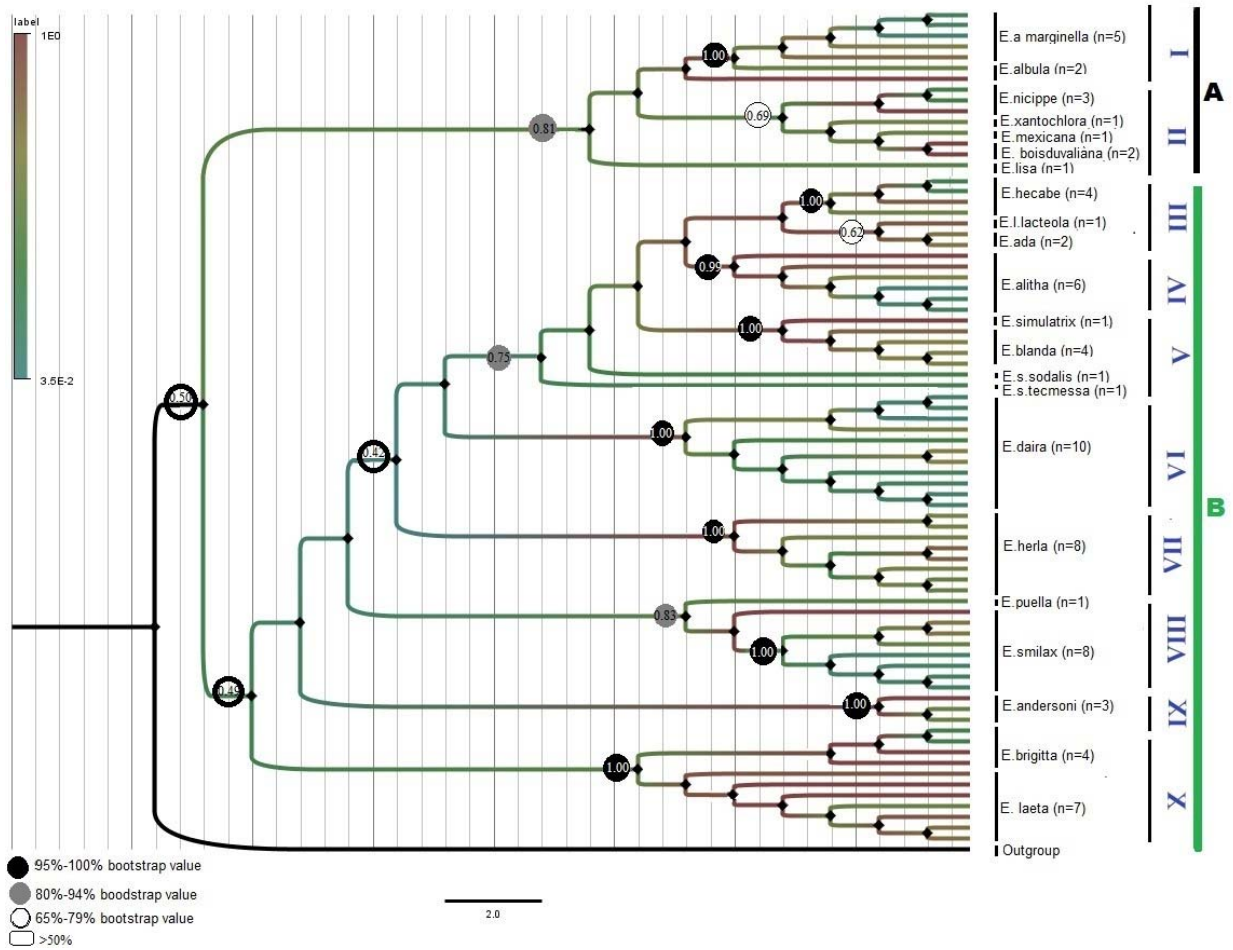

Figure 4: The ML and BI based phylogenetic tree showing the relationship between 21 species present under the genus *Eurema*. The ML tree reconstructed using the maximum likelihood (ML) method implemented in the PhyML program (v3.0) (Guindon, et al.,2010) and BI tree reconstruct using the MrBayes 3.1.231(Huelsenbeck JP & Ronquist F,2001).The circle with different colour showing the bootstrap value and number under circles represent BI clade posterior probability.

Table 1 List of *Eurema* species, number of individuals, accession numbers and population trend of each used in the present study

S. no	Species name	Genus	Family	No. of individuals in present study/Genbank	Accession no	Population trend
1.	*Eurema hecabe*	*Eurema*	Pieridae	3 /5	KF881047,KC257480, KT273566,KT 880649, KJ423054	Asia or Africa, Australia Throughout India
2.	*Eurema brigitta*	*Eurema*	Pieridae	3/ 4	KJ422991, KJ422993, EF584862, KT273575	Asia, Australia and Africa.
3.	*Eurema laeta*	*Eurema*	Pieridae	3/4	HM175717,KT273577, KT600170, GU372560	India, Sri Lanka, China, Indochina, Japan and Australia
4.	*Eurema blanda*	*Eurema*	Pieridae	3/5	HM17571, EF58486, KJ423049, KJ422891, AB969804	India and Southeast Asia.
5.	*Eurema andersonii*	*Eurema*	Pieridae	-/3	HM175721 HM395582, KJ422890	India and other part of Asia
6.	*Eurema smilax*	*Eurema*	Pieridae	-/8	KT273561.1, KT273560.1, KT273559.1, KT273558.1, KT273556.1, KT273557.1, KT273555.1, KT273554.1	Australia
7.	*Eurema herla*	*Eurema*	Pieridae	-/8	KT273587.1, KT273586.1, KT273585.1, KT273584.1, KT273583.1, KT273582.1, KT273581.1, KT273579.1	North and Eastern Australia
8.	*Eurema ada*	*Eurema*	Pieridae	-/2	KF226437.1, KF226438.1	Indo-China to Sunda land (Malay Peninsula, Borneo, Sumatra and Java)
9.	*Eurema alitha*	*Eurema*	Pieridae	-/6	KT273568.1,KF400206.1,KF396139.1, KKF404627.1, T273569.1, KT273567.1	South east Asia
10.	*Eurema daira*	*Eurema*	Pieridae	-/12	KP848907.1, KP848906.1, KF491736.1,JQ526833.1, JQ526832.1, JQ526831.1, JQ526830.1, JQ526829.1, JQ526822.1, JQ526823.1, JQ526824.1, JQ526826.1,	Argentina , North to the Southern America
11.	*Eurema mexicana*	*Eurema*	Pieridae	-/1	AY954568.1	Mexico, Central and South western America and rarely in Canada.
12.	*Eurema nicippe*	*Eurema*	Pieridae	-/3	GU089559.1, GU089560.1, KF491738.1	West Indies, Costa Rica and Belize
13.	*Eurema albula*	*Eurema*	Pieridae	-/2	JQ578170.1, JQ578168.1,	Central American countries
14.	*Eurema sari sodalis*	*Eurema*	Pieridae	-/1	KF226448.1	Southeast Asia
15.	*Eurema marginella*	*Eurema*	Pieridae	-/5	HM416519.1, HM416471.1, HM416456.1KP848904.1, KP848905.1	Texas, Venezuela, Brazil
16.	*Eurema simulatrix*	*Eurema*	Pieridae	-/2	HQ962276.1, HQ962261.1,	Burma to Sunda land and the Philippines
17.	*E. s.tecmessa (simultarix subspecies)*	*Eurema*	Pieridae	-/1	KF226451.1	Burma to Sunda land and the Philippines
18	*Eurema xantochlora*	*Eurema*	Pieridae	-/1	GU164425.1	Mexico to Colombia & Venezuela
19	*Eurema puella*	*Eurema*	Pieridae	-/1	HM906322.1	Australia
20	*Eurema boisduvaliana*	*Eurema*	Pieridae	-/2	GU164407.1, GU164405.1	North America
21	*Eurema lisa*	*Eurema*	Pieridae	-/1	KF491737.1	Central America and the southern part of North America.

Table 2. Species cross verification and diversity estimates of four *Eurema* species of Uttarakhand .Variable sites (V); Parsimony Informative sites (PI); Haplotypes (Hap); Nucleotide diversity (π); Haplotype diversity (hd); overall sequences divergence (k)

	BLAST Result			**Diversity**					
Specimen	**Species with highest similarity (%)**	**Query coverage (%)**	**Similarity (100%)**	**V**	**PI**	**Hap**	**π**	**hd**	**k**
E. blanda	*E. blanda*	100	100	8	6	3	0.00241	1.000	1.33
E. hecabe	*E. hecabe*	100	100	17	15	1	0.06420	0.500	35.50
E. brigitta	*E. brigitta*	100	99	28	23	1	-	-	-
E. laeta	*E. laeta*	100	98	39	31	3	0.03922	1.000	21.333
Overall	-	-	-	92	75	8	0.08807	0.923	43.765

Table 3. Interspecific evolutionary divergence between four *Eurema* species of Uttarakhand. The number of base substitutions per site from averaging over all sequence pairs between groups are shown. Standard error estimate(s) are shown above the diagonal and were obtained by a bootstrap procedure (1000 replicates). Analysis were conducted using the Kimura 2-parameter model**.**

Species	**Interspecific**				**Intraspecific**	
	1	2	3	4	D	SE
Eurema blanda		0.008	0.014	0.010	0.002	0.001
Eurema hecabe	0.061		0.013	0.012	0.051	0.005
Eurema brigitta	0.118	0.114		0.014	0.003	0.001
Eurema leata	0.112	0.131	0.119		0.041	0.007

Table 4. Estimates of evolutionary divergence over sequence pairs between 21 *Eurema* species. Maximum and minimum divergences highlighted in table 1

Species	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**
E.blanda																					
E.hecabe	0.05																				
E.brigitta	0.10	0.10																			
E.laeta	0.10	0.12	0.11																		
E.andersoni	0.08	0.09	0.11	0.12																	
E.smilax	0.08	0.08	0.12	0.11	0.10																
E.herla	0.08	0.08	0.11	0.09	0.10	0.08															
E.ada	0.05	0.01	0.10	0.12	0.09	0.08	0.08														
E.alitha	0.05	0.04	0.11	0.12	0.09	0.09	0.09	0.04													
E.daira	0.09	0.10	0.12	0.13	0.11	0.10	0.10	0.10	0.10												
E.mexicana	0.12	0.13	0.12	0.14	0.12	0.12	0.11	0.13	0.13	0.13											
E.nicippe	0.12	0.13	0.13	0.13	0.11	0.12	0.12	0.13	0.13	0.11	0.09										
E.albula	0.12	0.12	0.13	0.12	0.13	0.12	0.10	0.13	0.13	0.12	0.09	0.10									
E.a. marginella	0.12	0.12	0.13	0.12	0.13	0.11	0.10	0.12	0.13	0.11	0.09	0.10	0.01								
E.s.sodalis	0.06	0.06	0.10	0.12	0.07	0.09	0.09	0.05	0.06	0.09	0.12	0.11	0.13	0.12							
E.simulatrix	0.05	0.07	0.11	0.11	0.05	0.09	0.09	0.07	0.07	0.10	0.12	0.11	0.12	0.12	0.07						
E.s.tecmessa	0.09	0.10	0.12	0.13	0.11	0.12	0.12	0.10	0.10	0.12	0.13	0.15	0.14	0.14	0.09	0.10					
E.boisduvaliana	0.11	0.12	0.13	0.14	0.12	0.14	0.11	0.12	0.12	0.12	0.08	0.09	0.12	0.11	0.12	0.12	0.15				
E.l.lacteola	0.05	0.01	0.10	0.12	0.09	0.08	0.08	0.01	0.05	0.10	0.13	0.12	0.12	0.12	0.05	0.07	0.11	0.12			
E.lisa	0.12	0.13	0.15	0.13	0.12	0.12	0.11	0.13	0.14	0.14	0.13	0.13	0.11	0.11	0.11	0.12	0.15	0.13	0.13		
E.puella	0.07	0.08	0.11	0.12	0.09	0.07	0.09	0.08	0.08	0.10	0.11	0.12	0.13	0.13	0.08	0.08	0.12	0.12	0.08	0.13	
E. xantochlora	0.11	0.12	0.13	0.13	0.12	0.11	0.10	0.12	0.12	0.12	0.08	0.10	0.09	0.09	0.11	0.11	0.13	0.07	0.12	0.13	0.11

Table 5. Observed COI sequnce based variable sites among four *Eurema* species present in Uttarakhand

	Variable position																														
Nt Position ➤	26	29	32	35	38	47	50	52	53	56	59	61	62	64	65	67	68	70	71	74	83	86	89	92	94	98	100	101	110	115
Specimen																														
*E. blanda*1	A	A	T	T	T	A	T	A	T	A	A	G	T	A	T	A	T	A	G	A	T	A	T	A	C	T	A	A	A	A
*E. blanda*2		·										G							G											
*E. blanda*3		·										G							G									G		
*E. hecabe*1	G	·		A								G							A	T										
*E. hecabe*2	G	·		A								G							A	T										
*E. hecabe*3	G	·		A								G							A	T										
*E. brigitta*1	T	G	A				A	G	A		T	A	A	G	A	G			A	T							G	G	G	
*E. brigitta*2	T	G	A				A	G	A		T	A	A	G	A	G			A	T							G	G	G	
*E. brigitta*3	T	G	A				A	G	A		T	A	A	G	A	G			A	T							G	G	G	
*E. laeta*1			A	A			A	G	A	G	T	A	A	G		G	A	G	A	T	C			T		A	G			
*E. laeta*2			A	A			A	G	A	G	T	A	A	G		G	A	G	A	T	C		C	T		A	G			
*E. laeta*3			A		C		A	G	A	G	T	A		G		G	A	G	A	T	T	G	T		T					G
	11	11	11	11	11	11	11	11	11	11	12	12	12	12	12	12	12	12	12	12	12	12	12	12	13	13	13	13	13	13
Nt Position ➤	16	31	49	52	70	79	85	88	97	99	03	09	14	15	21	24	30	45	51	60	63	81	87	96	13	14	16	17	20	22
Specimen																														
*E. blanda*1	A	T	A	A	T	T	T	A	C	A	A	T	A	A	A	A	T	A	A	G	T	A	A	G	A	A	G	A	T	A
*E. blanda*2																														
*E. blanda*3																														
*E. hecabe*1					C				T					G		T		T				G						G		
*E. hecabe*2					C				T					G		T		T				G						G		
*E. hecabe*3					C				T					G		T		T				G						G		
*E. brigitta*1	T			G		A	A	T	A	G	G	A	G		T		G		C		C	T	G		G	T	A			
*E. brigitta*2	T			G		A	A	T	A	G	G	A	G		T		G		C		C	T	G		G	T	A			
*E. brigitta*3	T			G		A	A	T	A	G	G	A	G		T		G		C		C	T	G		G	T	A			
*E. laeta*1	T	G	T				A	A	A	G									G	A	G				G				A	G
*E. laeta*2	T	G	T				A	A	A	G									G	A	G				G				A	G
*E. laeta*3	T	A	T				A	G	G	G											A								A	G

Continued

	Variable position																														
	3	*3*	*3*	*3*	*3*	*3*	*3*	*3*	*3*	*3*	*3*	*3*	*3*	*3*	*3*	*3*	*3*	*3*	*3*	*3*	*3*	*3*	*3*	*3*	*4*	*4*	*4*	*4*	*4*	*4*	*4*
Nt . Position ➔	*2*	*2*	*2*	*3*	*3*	*3*	*3*	*4*	*4*	*5*	*5*	*6*	*6*	*6*	*7*	*8*	*8*	*8*	*8*	*9*	*9*	*9*	*9*	*1*	*1*	*3*	*3*	*4*	*5*	*6*
Specimen	*3*	*6*	*9*	*2*	*4*	*5*	*7*	*1*	*4*	*3*	*9*	*2*	*5*	*8*	*4*	*0*	*3*	*8*	*9*	*2*	*4*	*5*	*8*	*0*	*9*	*1*	*7*	*3*	*5*	*1*
*E. blanda*1	**T**	**A**	**G**	**A**	**G**	**A**	**G**	**G**	**A**	**G**	**T**	**G**	**G**	**T**	**A**	**G**	**T**	**A**	**T**	**T**	**A**	**A**	**A**	**T**	**A**	**A**	**T**	**G**	**A**	**A**
*E. blanda*2																									G					
*E. blanda*3																									G					
*E. hecabe*1		G						A		A			A				C		C					A	C	G		A	G	
*E. hecabe*2		G						A		A			A				C		C					A	C	G		A	G	
*E. hecabe*3		G						A		A			A				C		C					A	C	G		A	G	
*E. brigitta*1		G		T	A			A	T	A	G	A		G	G	A								A	T	T	A	A		T
*E. brigitta*2		G		T	A			A	T	A	G	A		G	G	A								A	T	T	A	A		T
*E. brigitta*3		G		T	A			A	T	A	G	A		G	G	A								A	T	T	A	A		T
*E. laeta*1	A		A	T	A	T	A	G		A										A	G		G		T	T	A	A		G
*E. laeta*2	A		A	T	A	T	A	G		A								G		A	G	G	G		T	T	A	A		G
*E. laeta*3	A			T	A	T	A	A		A										A	G		G		T	T		A		G

	4	*4*	*4*	*4*	*4*	*4*	*5*	*5*	*5*	*5*	*5*	*5*	*5*
Nt. Position ➔	*6*	*7*	*7*	*8*	*8*	*9*	*0*	*0*	*0*	*1*	*1*	*2*	*3*
Specimen	*4*	*3*	*6*	*2*	*5*	*7*	*0*	*6*	*8*	*5*	*8*	*1*	*0*
*E. blanda*1	**A**	**T**	**A**	**A**	**A**	**A**	**T**	**T**	**A**	**T**	**A**	**T**	**A**
*E. blanda*2													
*E. blanda*3													
*E. hecabe*1	G				G								
*E. hecabe*2	G				G								
*E. hecabe*3	G				G								
*E. brigitta*1		A				G	C				G	A	T
*E. brigitta*2		A				G	C				G	A	T
*E. brigitta*3		A				G	C				G	A	T
*E. laeta*1		A								G	G	A	
*E. laeta*2		A								A	G	A	
*E. laeta*3	G	G	G	G				G	G	A	G	A	