International Journal of Zoology Studies
www.zoologyjournals.com
ISSN: 2455-7269
Received: 19-04-2026, Accepted: 20-05-2026, Published: 04-06-2026
Volume 11, Issue 2, 2026, Page No. 115-118

Molecular phylogenetic analysis of eight *Nymphalidae* butterfly species based on mitochondrial COI gene

Monika Jagarwal, Priyanka Verma, Ghanshyam Kachhawa*
Department of Zoology, University of Rajasthan, Jaipur, Rajasthan, India
Corresponding Author: Ghanshyam Kachhawa
DOI: https://doi.org/10.66856/ijzs.2026.11.2.11052

Abstract
The mitochondrial cytochrome c oxidase subunit I (COI) gene is widely acknowledged as a reliable molecular marker for phylogenetic research and species identification. This study investigated the molecular phylogenetics of eight butterfly species belonging to the family *Nymphalidae*: *Danaus chrysippus, Danaus genutia, Hypolimnas bolina, Hypolimnas misippus, Junonia lemonias, Junonia orithya, Melanitis leda,* and *Phalanta phalantha.* The nucleotide composition showed strong AT bias (70.32%) compatible with mitochondrial DNA characteristics of *Lepidoptera*. Pairwise genetic distance analysis showed lowest divergence between *Junonia lemonias* and *Junonia orithya* (0.0074) and the highest between *Melanitis leda* and *Phalanta phalantha* (0.1903). The mean intra-generic distance (0.050) was significantly lower than the mean inter-generic distance. The transition/transversion bias (R = 0.82) indicated conserved mitochondrial evolution in insects. The Neighbour Joining phylogenetic tree showed clear genus-level clustering with strong bootstrap support. These results confirmed the effectiveness of the COI gene as a reliable DNA barcode. It is useful for species identification and phylogenetic analysis in *Nymphalidae* butterflies.

Keywords: *Nymphalidae*, COI gene, DNA barcoding, phylogenetic analysis, genetic distance, neighbour joining, MEGA12, *Lepidoptera*

Introduction
Butterflies of the family *Nymphalidae* represent one of the largest and most diverse groups within the order *Lepidoptera*. It comprises over 6,000 species distributed across all major biogeographical regions (Wahlberg *et al*., 2009) [16]. They serve as ecological significant organisms. They act as pollinators and bioindicators. The accurate species identification within this family is fundamental for biodiversity assessments, conservation planning and evolutionary studies. Traditional taxonomic approaches based on morphological characters are important. However, these can be limited by phenotypic plasticity, sexual dimorphism and convergent evolution (Hebert *et al*., 2003a) [4]. DNA barcoding is a molecular technique that uses mitochondrial cytochrome c oxidase subunit I (COI) gene. This technique has emerged as a powerful tool for species identification and phylogenetic inference (Hebert *et al*., 2003a, 2003b) [4, 5]. The standard 648 bp barcode region of the COI gene offers several advantages. It evolves rapidly enough to distinguish closely related species while also retaining conserved regions for universal primer design (Folmer *et al*., 1994) [3]. The COI gene has been widely used in different invertebrate groups including *Lepidoptera* for identifying species boundaries and studying evolutionary relationships (Ward *et al*., 2005; Meier *et al*., 2006) [10, 15]. The concept of a barcode gap refers to the difference between intraspecific and interspecific genetic variation. It forms the theoretical basis of DNA barcoding for species identification (Hebert *et al*., 2003b; Ratnasingham & Hebert, 2007) [5, 11]. When interspecific distances are greater than intraspecific distances the COI gene acts as a reliable diagnostic marker. Previous studies on *Lepidoptera* have shown that COI based barcoding provides high accuracy in species identification (Zhou *et al*., 2009) [17]. The present study aimed to assess the molecular phylogenetics of eight *Nymphalidae* butterfly species using COI gene sequences retrieved from NCBI. The objectives of the study were to analyse nucleotide composition and substitution patterns to evaluate intra and inter generic genetic distances to test the barcode gap and to reconstruct a phylogenetic tree using the Neighbour Joining method. The study provides insights into the usefulness of COI based barcoding for species identification in Indian *Nymphalidae* and adds to the growing global barcode database.

Materials and Methods
Sequence Retrieval and Alignment
This study assembled DNA barcode sequences for eight butterfly species of family *Nymphalidae*. The mitochondrial DNA barcode sequences were retrieved from the National Center for Biotechnology Information (NCBI) database. BLAST was used to identify and download authenticated cytochrome c oxidase subunit I (COI) gene sequences for each species. The selected species were *Danaus chrysippus, Danaus genutia, Hypolimnas bolina, Hypolimnas misippus, Junonia lemonias, Junonia orithya, Melanitis leda* and *Phalanta phalantha.* The sequences were selected within the 650–750 base pair range. It is done to ensure consistency with the standard DNA barcode region for *Lepidoptera*. The sequences were then used for subsequent alignment and molecular analyses. All retrieved COI sequences were aligned using the CLUSTALW algorithm implemented in MEGA12 (Kumar *et al*., 2024) [7].

Genetic and Phylogenetic Analysis

The nucleotide composition was calculated using MEGA12 for each species. Pairwise genetic distances were estimated using the Maximum Composite Likelihood (MCL) model (Tamura *et al*., 2004) [14] with pairwise deletion applied for handling missing data. The transition/transversion bias (R) was estimated using the Kimura 2-parameter (K2P) model (Kimura, 1980) [6]. The phylogenetic relationships among eight *Nymphalidae* species were reconstructed using Neighbour Joining (NJ) method (Saitou & Nei, 1987) [12] in MEGA12. The reliability of internal branches was assessed using bootstrap analysis with 1,000 replicates. The final dataset comprised 724 nucleotide positions after pairwise deletion of ambiguous sites.

Results

Nucleotide Composition of COI Gene Sequences

The nucleotide composition of the COI gene was analysed in eight *Nymphalidae* butterfly species. The average nucleotide percentages were 39.13% for Thymine (T) 31.19% for Adenine (A) 15.36% for Cytosine (C) and 14.32% for Guanine (G) (Table 1). The total AT content was 70.32% while the GC content was 29.68%. At the species level Thymine content ranged from 37.07% (*Hypolimnas bolina*) to 40.18% (*Melanitis leda*). Adenine from 29.11% (*Phalanta phalantha*) to 32.90% (*Hypolimnas misippus*). Cytosine ranged from 14.08% (*Hypolimnas misippus*) to 16.59% (*Phalanta phalantha*) and Guanine ranged from 13.53% (*Hypolimnas misippus*) to 15.12% (*Junonia orithya*).

Table 1: Percentage nucleotide composition of COI gene sequences in eight *Nymphalidae* butterfly species

Species name	T(U)	C	A	G
Danaus chrysippus	39.12	15.22	31.51	14.16
Danaus genutia	39.77	16.08	29.68	14.47
Hypolimnas bolina	37.07	16.09	33.05	14.20
Hypolimnas misippus	39.49	14.08	32.90	13.53
Junonia lemonias	38.34	15.81	30.83	15.02
Junonia orithya	39.07	14.37	31.44	15.12
Melanitis leda	40.18	16.29	29.38	14.16
Phalanta phalantha	39.44	16.59	29.11	14.87
Average	39.13	15.36	31.19	14.32

Pairwise Genetic Distance Analysis

The standard error estimates are shown above the diagonal while the genetic distance values are shown below the diagonal (Table 2). The lowest genetic distance was observed between *Junonia lemonias* and *Junonia orithya* with a value of 0.0074. The genetic distance between *Hypolimnas bolina* and *Hypolimnas misippus* was 0.0641 while the distance between *Danaus chrysippus* and *Danaus genutia* was 0.0775. The highest genetic distance was recorded between *Melanitis leda* and *Phalanta phalantha* with a value of 0.1903. The inter - generic genetic distances ranged from 0.0788 to 0.1903 while the intra - generic genetic distances ranged from 0.0074 to 0.0775.

Table 2: Pairwise genetic distances (below diagonal) and standard errors (above diagonal) based on COI gene sequences

	D. chrysippus	***D. genutia***	***H. bolina***	***H. misippus***	***J. lemonias***	***J. orithya***	***M. leda***	***P. phalantha***
D. chrysippus	-	0.0116	0.0176	0.0163	0.0170	0.0159	0.0168	0.0180
D. genutia	0.0775	-	0.0179	0.0166	0.0166	0.0155	0.0170	0.0186
H. bolina	0.1528	0.1645	-	0.0107	0.0146	0.0140	0.0195	0.0197
H. misippus	0.1370	0.1494	0.0641	-	0.0139	0.0119	0.0172	0.0184
J. lemonias	0.1359	0.1352	0.1036	0.0955	-	0.0074	0.0196	0.0183
J. orithya	0.1312	0.1372	0.0991	0.0788	0.0074	-	0.0176	0.0177
M. leda	0.1535	0.1559	0.1794	0.1551	0.1689	0.1503	-	0.0196
P. phalantha	0.1579	0.1731	0.1801	0.1600	0.1593	0.1572	0.1903	-

Genetic Divergence Within and Between Genera

The mean intra-generic genetic distance across the three genera represented in this study was 0.050 whereas the mean inter-generic distance was 0.151 (Table 2). The higher inter-generic divergence relative to intra-generic divergence indicates a distinct barcode gap among the studied taxa.

Substitution Pattern Analysis

Analysis of the substitution pattern showed transitions (purine–purine or pyrimidine–pyrimidine substitutions) were more frequent than transversions in the COI gene sequences. The transition/transversion bias value of R = 0.82. The maximum log likelihood estimated for the maximum likelihood (ML) computation was –3,596.948, based on 724 positions in the final aligned dataset.

Phylogenetic Analysis Using the Neighbour Joining Method

The Neighbour Joining phylogenetic tree base on COI gene sequences showed clear genus level clustering among the eight *Nymphalidae* species (Fig. 1). Three major clades were identified and each corresponded to a recognised genus.

1. *Danaus chrysippus* and *Danaus genutia* formed a well-supported clade, reflecting their shared genus-level classification and moderate genetic distance (0.0775).
2. *Hypolimnas bolina* and *Hypolimnas misippus* clustered together with strong bootstrap support, consistent with their relatively low intra-generic distance (0.0641).
3. *Junonia lemonias* and *Junonia orithya* formed the most closely related pair (distance = 0.0074), appearing as sister taxa within a distinct clade.

Melanitis leda and *Phalanta phalantha* appeared as separate, more divergent lineages, consistent with their high inter-generic distances. The topology of the tree corroborated the expected subfamily-level relationships within *Nymphalidae*, with Danainae (*Danaus* spp.), Nymphalinae (*Hypolimnas, Junonia*), Satyrinae (*Melanitis*) and Heliconiinae (*Phalanta*) grouping accordingly.

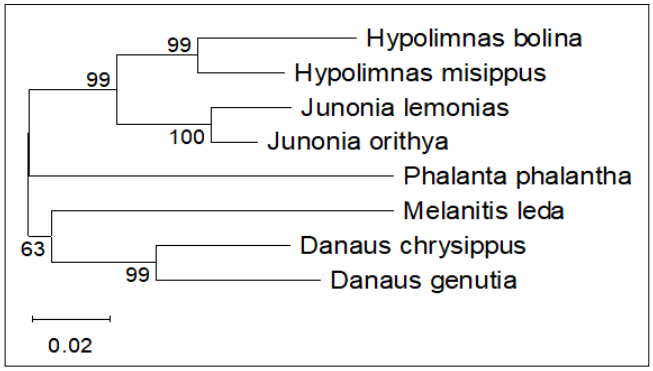

Fig 1: Neighbour Joining phylogenetic tree of eight *Nymphalidae* species based on COI gene sequences

Discussion

The present study used the mitochondrial COI gene as a molecular marker to evaluate phylogenetic relationships and genetic diversity among eight *Nymphalidae* butterfly species. The results obtained were generally consistent with previous studies on DNA barcoding and molecular phylogenetics in *Lepidoptera*.

The high AT bias (70.32%) observed in the COI gene is a common feature of insect mitochondrial DNA (Simon *et al*., 1994) [13]. Such AT richness has been associated with mutational pressure, strand asymmetry during replication and codon usage bias (Avise, 2000) [1]. The nucleotide composition pattern of T > A > C > G reported is also similar to patterns reported from other *Nymphalidae* studies. This reflected and indicated the conserved nature of the COI gene in *Lepidoptera* (Brower, 1994) [2].

The pairwise genetic distances showed clear differences between intra-generic and inter-generic comparisons. The lowest genetic distance observed between *Junonia lemonias* and *Junonia orithya* (0.0074). This low divergence may indicate recent evolutionary separation between the two species. The genetic distance between *Danaus chrysippus* and *Danaus genutia* (0.0775) and between *Hypolimnas bolina* and *Hypolimnas misippus* (0.0641) are moderate and consistent with congeneric species that have undergone sufficient divergence for distinct morphological recognition but retain molecular evidence of shared ancestry. The highest divergence was observed between *Melanitis leda* and *Phalanta phalantha* (0.1903) reflecting substantial evolutionary differentiation between these genera.

The distinct barcode gap observed in the present study supports the effectiveness of the mitochondrial COI gene as a reliable marker for species-level identification within *Nymphalidae*. The mean inter-generic distance (0.151) was approximately 3.0 times higher than the mean intra-generic distance (0.050). It showed substantial genetic differentiation among genera. It is a clear barcode gap which is considered essential for accurate molecular identification and taxonomic discrimination (Hebert *et al*., 2003b; Ratnasingham & Hebert, 2007) [5, 11]. The divergence pattern observed is consistent with previous studies on *Lepidoptera* where intraspecific divergence generally remains below 2% whereas interspecific divergence exceeds 3% (Hebert *et al*., 2003b) [5]. Similar barcode gaps reported in other *Lepidopteran* studies further confirm the reliability and broad applicability of COI-based DNA barcoding for insect identification (Zhou *et al*., 2009; Meier *et al*., 2006) [10, 17].

The transition/transversion bias (R = 0.82) observed in the present study is consistent with the typical evolutionary pattern of mitochondrial genes in insects. Transitions are generally more common because they are biochemically more probable and less likely to produce drastic structural changes in DNA sequences. Such substitutions are often favoured during the early stages of evolutionary divergence whereas transversions tend to accumulate over longer evolutionary periods (Martin & Wang, 2011) [9]. The moderate R value obtained in this study indicated different levels of evolutionary divergence among the analysed taxa. Closely related species were observed within *Junonia* to more distantly related genera such as *Melanitis* and *Phalanta*. The observed substitution pattern also suggested that mutational saturation had minimal effect on the dataset. This supports the reliability of COI sequences for phylogenetic analysis and species identification.

The Neighbour Joining phylogenetic tree clearly resolved genus-level groupings. High bootstrap values supported these clusters. The results showed that COI sequences contain sufficient phylogenetic signal. They successfully recovered the expected taxonomic relationships. The clustering of *Danaus* species within Danainae, *Hypolimnas* and *Junonia* within Nymphalinae, *Melanitis* within Satyrinae, and *Phalanta* within Heliconiinae is broadly consistent with established subfamily classifications (Lamas, 2004; Wahlberg *et al*., 2009) [8, 16]. The positions of *Melanitis leda* and *Phalanta phalantha* as more divergent lineages in phylogenetic tree. This is expected given the high inter-generic distances observed between these taxa and other species in the dataset.

Overall, this study confirms that the mitochondrial COI gene is a reliable and effective marker for molecular phylogenetic analysis of Indian *Nymphalidae* butterflies. The findings contribute to the baseline genetic data for eight ecologically important species and support the ongoing development of a comprehensive DNA barcode library for South Asian *Lepidoptera*.

Conclusion

This study demonstrated the effectiveness of the mitochondrial COI gene for molecular phylogenetic analysis of eight *Nymphalidae* butterfly species. The pronounced AT bias (70.32%) and transition/transversion bias (R = 0.82) observed were consistent with established patterns of mitochondrial evolution in *Lepidoptera*. A clear barcode gap was identified with mean intra-generic distance (0.050) substantially lower than mean inter-generic distance (0.151), validating COI as a robust molecular marker for species discrimination in *Nymphalidae*. The Neighbour Joining phylogenetic tree successfully resolved genus-level clustering with all congeneric species pairs forming distinct, well-supported clades. The highest inter-generic divergence between *Melanitis leda* and *Phalanta phalantha* (0.1903) and the lowest intra-generic divergence between *Junonia lemonias* and *Junonia orithya* (0.0074) highlight the varying degrees of evolutionary differentiation within this family. These results provide valuable baseline data for the molecular taxonomy of Indian *Nymphalidae*. The study also demonstrates the usefulness of COI-based DNA barcoding as a complementary tool for butterfly species identification and biodiversity monitoring in South Asia.

Acknowledgement
The authors are grateful to the National Center for Biotechnology Information (NCBI) for providing access to publicly available COI gene sequences used in this study.

References

1. Avise JC. *Phylogeography:* The History and Formation of Species. Harvard University Press, Cambridge, Massachusetts, 2000.
2. Brower AVZ. Rapid morphological radiation and convergence among races of the butterfly *Heliconius erato* inferred from patterns of mitochondrial DNA evolution. Proceedings of the National Academy of Sciences,1994:91(14):6491–6495.
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 Marine Biology and Biotechnology,1994:3(5):294–299.
4. Hebert PDN, Cywinska A, Ball SL, deWaard JR. Biological identifications through DNA barcodes. Proceedings of the Royal Society B: Biological Sciences,2003:270(1512):313–321.
5. Hebert PDN, Ratnasingham S, deWaard JR. Barcoding animal life: Cytochrome c oxidase subunit 1 divergences among closely related species. Proceedings of the Royal Society B: Biological Sciences,2003:270(1):96–99.
6. Kimura M. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. Journal of Molecular Evolution,1980:16(2):111–120.
7. Kumar S, Stecher G, Suleski M, Sanderford M, Sharma S, Tamura K *et al*. Molecular Evolutionary Genetics Analysis version 12 for adaptive and green computing. Molecular Biology and Evolution,2024:41(6):015.
8. Lamas G. Checklist: Part 4A. *Hesperioidea–Papilionoidea*. In: Heppner JB, editor. Atlas of Neotropical *Lepidoptera*. Association for Tropical *Lepidoptera*, Gainesville, Florida, 2004, 5.
9. Martin JA, Wang Z. Next-generation transcriptome assembly. Nature Reviews Genetics,2011:12(10):671–682.
10. Meier R, Shiyang K, Vaidya G, Ng PKL. DNA barcoding and taxonomy in *Diptera*: A tale of high intraspecific variability and low identification success. Systematic Biology,2006:55(5):715–728.
11. Ratnasingham S, Hebert PDN. BOLD: The Barcode of Life Data System. Molecular Ecology Notes,2007:7(3):355–364.
12. Saitou N, Nei M. The neighbour-joining method: A new method for reconstructing phylogenetic trees. Molecular Biology and Evolution,1987:4(4):406–425.
13. Simon C, Frati F, Beckenbach A, Crespi B, Liu H, Flook P *et al*. Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. Annals of the Entomological Society of America,1994:87(6):651–701.
14. Tamura K, Nei M, Kumar S. Prospects for inferring very large phylogenies by using the neighbour-joining method. Proceedings of the National Academy of Sciences,2004:101(30):11030–11035.
15. Ward RD, Zemlak TS, Innes BH, Last PR, Hebert PDN. DNA barcoding Australia's fish species. Philosophical Transactions of the Royal Society B: Biological Sciences,2005:360(1462):1847–1857.
16. Wahlberg N, Leneveu J, Kodandaramaiah U, Peña C, Nylin S, Freitas AVL *et al*. Nymphalid butterflies diversify following near demise at the Cretaceous/Tertiary boundary. Proceedings of the Royal Society B: Biological Sciences,2009:276(1677):4295–4302.
17. Zhou X, Adamowicz SJ, Jacobus LM, DeWalt RE, Hebert PDN. Towards a comprehensive barcode library for arctic life — *Ephemeroptera*, *Plecoptera*, and *Trichoptera* of Churchill, Manitoba, Canada. Frontiers in Zoology,2009:6:30.