International Journal of Ecology and Environmental Sciences
www.ecologyjournal.in
Online ISSN: 2664-7133, Print ISSN: 2664-7125
Received: 17-02-2021, Accepted: 18-03-2021, Published: 21-04-2021
Volume 3, Issue 2, 2021, Page No. 69-72

Phylogeny and molecular studies of twelve Butterflies species genus using RAPD-PCR technique from the Tiruchirappalli district, Tamil Nadu region in India

Nesarajan J
PhD, Assistant Professor, Department of Zoology, Bishop Heber College, Tiruchirappalli, Tamil Nadu, India

Abstract
The present study reveals that RAPD-PCR is extremely useful for rapid identification of genetic polymorphisms in lepidopteron because of reproducibility of the result for each of the species. Mitochondrial genome is one of the most frequently used loci in phylogenetic and phylogeographic analysis, and it is becoming increasingly possible to sequence and analyse this genome in its entirety from diverse taxa. The mitochondrial gene coding for Cytochrome oxidase subunit I (COI) was extensively used for studying the molecular diversity of eukaryotic organisms, and hence the same has been adopted in this study as well. From the 12 selected butterfly's species, the genomic DNA was isolated by CTAB-NaCl method, and their quality and quantity were tested through agarose gel electrophoresis. The COI gene was amplified using the custom-designed degenerate primer set which yielded around ~650 bp amplicons. The genus and species names of the closest match from the database were identified based on the BLAST parameters like identity, query coverage and E value. The identity of all the sequences were close to 100 % (99-98 %) and it is obvious as the COI gene is highly conserved in nature for the same purpose it has been used as a marker to study evolutionary relationships. In this study, a phylogenetic tree was constructed for the 12 COI gene sequences of different butterflies by Neighbor-Joining method employing Bootstrap type. Dendogram constructed using the UPGMA of NTSYS spc2.2 software divided the butterflies species into two clades. There is a difference in the branching pattern between the morphological and molecular data, which signifies the need for using molecular tools for taxonomic classification as well as in understanding the evolutionary relationship.

Keywords: butterfly, phylogeny, blast, RFLP; RAPD-PCR

Introduction

DNA taxonomy is now used in harmony but in addition to other classical morphological data to delimit species (Tautz *et al.,* 2002) [1]. Although it has been well accepted that DNA taxonomy can solve many taxonomic problems, it has still not got a central role in it. Presently, scientists are working on phylogeny and phylogeography of different species using the DNA as the central theme of their analysis. Although the morphological attributes are going to play the major role in the taxonomic description, DNA can be given a better position than what it has today. We believe the best way to give DNA its fair chance in taxonomy will be to implement DNA barcoding as an international unit for identification of species. Paul Hebert from University of Guelph in Canada developed the use of part of the mitochondrial gene as a universal identification marker for living organisms (Savolainen *et al.,* 2005) [2]. The whole notion of product barcode is based on the arrangement of 10 alternate numbers in 11 positions to create 100 billion unique numbers, which can then be used as an individual product identity (Hebert *et al.,* 2003) [3]. Genomic DNA can be used in the same manner, but the problem faced is that only 4 bases are there to work with. This problem is solved when we look at the enormous size of the available DNA in the animal cell. It has been calculated by (Hebert *et al.,* 2003) [4] just taking 15 sites of nucleotide positions can create a possibility of 15 codes, which is huge compared to the artificial barcode system. Some of the 4 sequences are very highly conserved, and other regions provide diversity to be checked for at least intra specific levels.

Mitochondrial genome is one of the most frequently used loci in phylogenetic and phylogeographic analysis, and it is becoming increasingly possible to sequence and analyse this genome in its entirety from diverse taxa (Muller, 2006) [5]. As the COI gene chosen for the present study is part of mitochondrial DNA, The complete mitochondrial genome 2rRNAs, 22tRNAs and 13 protein-coding genes with two control regions. Interestingly, the fishes' mitochondrial genome has two control regions (Lee *et al.,* 2001) [6]. In the mitochondrial genome, COI gene is an approximate 656 bp region. This gene encodes part of the terminal enzyme of the respiratory chain of mitochondria. Within species variation for this gene is low compared with between species variation (Ward, 2009) [7]; with the use of this~600 base pair of sequence from the 5'terminus of the gene (Hebert *et al.,* 2003) [4]. Even though shorter fragments of COI have also been found to be effective for the identification of specimens with degraded DNA, however, where a 650 base sequences is not easily obtainable (Hajibabaei *et al.,* 2007) [8]. The present study reveals that RAPD-PCR is extremely useful for rapid identification of genetic polymorphisms in lepidopteron because of reproducibility of the result for each of the species. Mitochondrial genome is one of the most frequently used loci in phylogenetic and phylogeographic analysis, and it is becoming

increasingly possible to sequence and analyse this genome in its entirety from diverse taxa

Materials and Methods
Collection of butterfly species
The present study twelve different species of butterflies representing five different families were collected by the Bishop Heber College campus, Trichy-17, Tamil Nadu. INDIA. Species identification was made using various field guides and other available literature.

Molecular Methods
DNA isolation and sequencing
From the 12 selected butterfly's species, the genomic DNA was isolated by CTAB-NaCl method, and their quality and quantity were tested through agarose gel electrophoresis. The gel image (Figure 1) shows the DNA are of significant quality and could be used for any molecular techniques. The approximate concentration of the DNA was calculated to be around 100 mg/ml. All 12 DNA samples were aseptically sealed and shipped to Macrogen Inc. (Seoul, South Korea), where the amplification and sequencing were done. The COI gene was amplified using the custom-designed degenerate primer set which yielded around ~650 bp amplicons. The PCR amplicons were purified and then sequenced by the automated DNA sequence employing dye-termination method.

DNA sequence analysis
The sequences obtained out of sequencing runs were aligned using the CAP3 program. CAP3 program aligns 2 different reads into a single contig based on the overlapping sequences present in both the reads. The obtained contigs were then searched against the non-redundant public database, Nucleotide database of NCBI using the BLAST program. Based on the search results, the end sequences were trimmed to get high quality contigs, which were analysed further.

BLAST analysis
BLAST (Basic local alignment search tool) is a Bioinformatic tool for local alignment which employs different algorithms for identification of sequences similar to that of the query sequences in the database. The sequences given as an input are called query sequences, and the sequences matched from the database are called subject sequences. The genus and species names of the closest match from the database were identified based on the BLAST parameters like identity, query coverage and E value. The identity of all the sequences were close to 100 % (99-98 %) and it is obvious as the COI gene is highly conserved in nature for the same purpose it has been used as a marker to study evolutionary relationships. The names of the database hits along with their maximum identity match were given results section (Table 2).

Submission of DNA sequences in GenBank
All 12 sequences were submitted in GenBank for public access using the tool Bank It. The release date was chosen as immediate. The accession numbers provided by GenBank for the sequences are tabulated in results section (Table 3).

RFLP analysis
RFLP is a modified method in restriction analysis which employs one or two restriction enzymes to differentiate the given DNA fragments.

Multiple sequence alignment
Multiple sequence alignment (MSA) is a global alignment program where each one of the given sequences is aligned from start to end with all the other sequences. ClustalW algorithm was used for this multiple sequence alignment, which comes as a part of MEGA6 software package.

Phylogenetic analysis
In this study, a phylogenetic tree was constructed for the 12 COI gene sequences of different butterflies by Neighbor-Joining method employing Bootstrap type.

Results

Table 1: The selected Butterfly species

Sl. No.	Common Name
Family: Pieridae	
1	Common Grass yellow
2	3 Spot grass yellow
3	Mottled emigrant
Family: Nymphalidae	
4	Lemon pansy
5	Tawny Coster
Family: Papilionidae	
6	Common rose
7	Common Mormon
8	Crimson rose
Family: Lycaenidae	
9	Common Cerulean
10	Forget me not
11	Common Jezebel
Family: Hesperiidae	
12	Rice swift

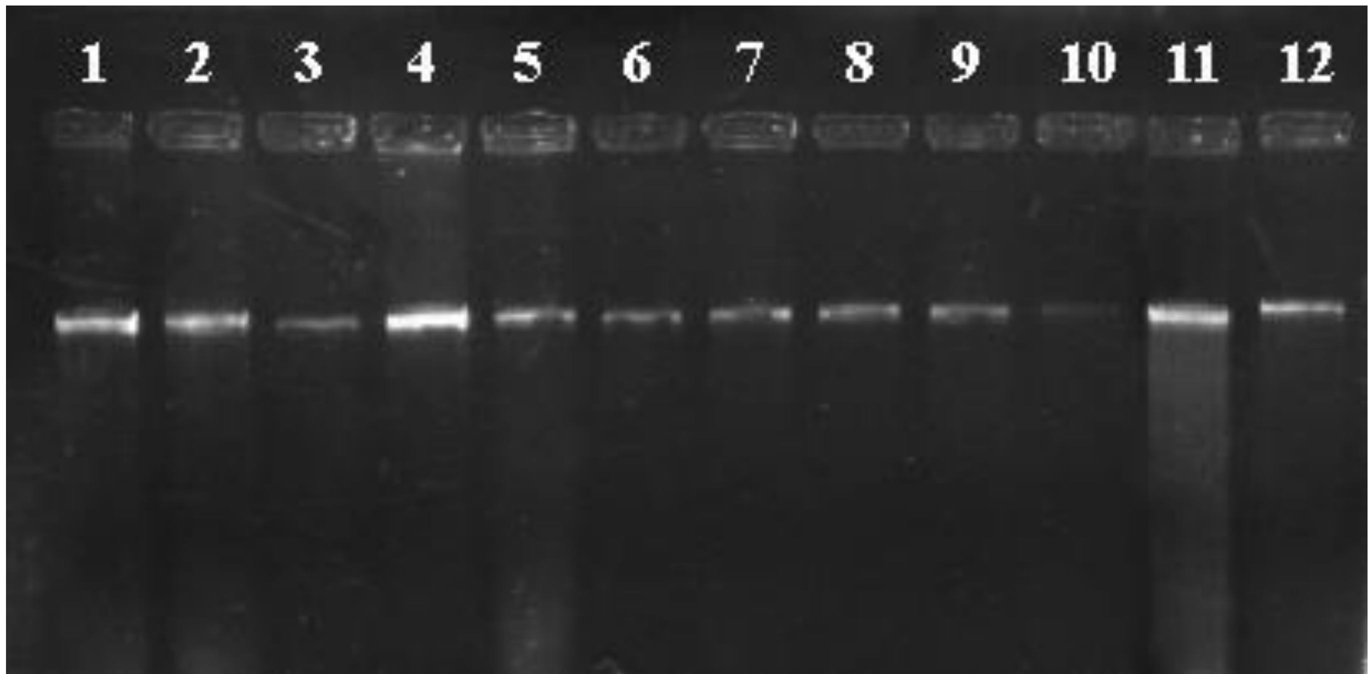

Fig 1: Agarose gel (0.8 %) with genomic DNAs isolated from butterflies

Table 2: BLAST analysis

Sl. No.	Common Name	Database Hit	Identity
1	Common Grass yellow	*Eurema hecabe*	99 %
2	3 Spot grass yellow	*Eurema blanda*	99 %
3	Mottled emigrant	*Catopsilia pyranthe*	99 %
4	Lemon pansy	*Junonia lemonias*	99 %
5	Tawny Coster	*Acraea sp.*	99 %
6	Common rose	*Pachliopta aristolochiae*	99 %
7	Common Mormon	*Papilio polytes*	99 %
8	Crimson rose	*Pachliopta hector*	99 %
9	Common Cerulean	*Jamides celeno*	98 %
10	Forget me not	*Catochrysops Strabo*	98 - 99 %
11	Common Jezebel	*Delias eucharis*	99 %
12	Rice swift	*Borbo cinnara*	99 %

Table 3: GenBank accession numbers for the submitted sequences

Sl. No.	Common Name	Best match	GenBank Accession Number
1	Common Grass yellow	*Eurema hecabe*	KT123299
2	3 Spot grass yellow	*Eurema blanda*	KT123300
3	Mottled emigrant	*Catopsilia pyranthe*	KT123301
4	Lemon pansy	*Junonia lemonias*	KT123302
5	Tawny Coster	*Acraea sp.*	KT123303
6	Common rose	*Pachliopta aristolochiae*	KT123304
7	Common Mormon	*Papilio polytes*	KT123305
8	Crimson rose	*Pachliopta hector*	KT123306
9	Common Cerulean	*Jamides celeno*	KT123307
10	Forget me not	*Catochrysops Strabo*	KT123308
11	Common Jezebel	*Delias eucharis*	KT123309
12	Rice swift	*Borbo cinnara*	KT123310

RFLP analysis

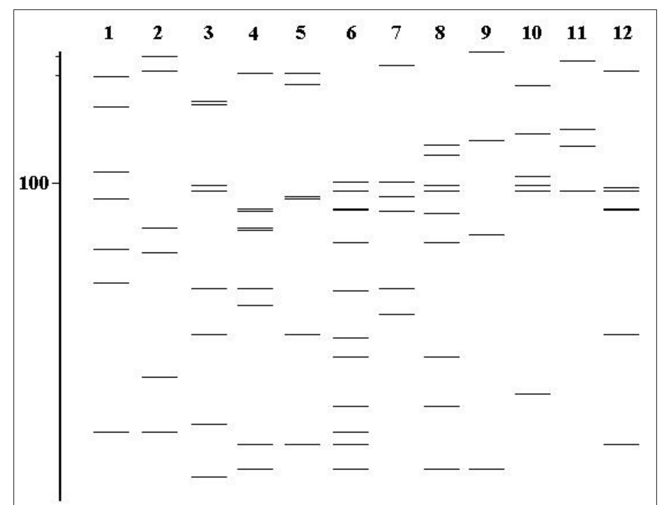

Fig 2: RFLP analysis showing the band patterns of different species

Legends: The image represents a typical 2 % agarose gel and the scale bar at the left end denotes the ladder with respective size of the bands. The lanes 1-12 represent individual species in the order as follows: 1.*Eurema hecabe;* 2.*Eurema blanda;* 3.*Catopsilia pyranthe;* 4.*Junonia lemonias;* 5.*Acraea* species; 6.*Pachliopta aristolochiae;* 7.*Papilio polytes;* 8.*Pachliopta hector;* 9.*Jamides celeno;* 10.*Catochrysops Strabo;* 11.*Delias eucharis;* 12.*Borbo cinnara.*

Multiple sequence alignment

Fig 3: Multiple sequence alignment of COI gene sequences from 12 Different butterflies.

Phylogenetic analysis

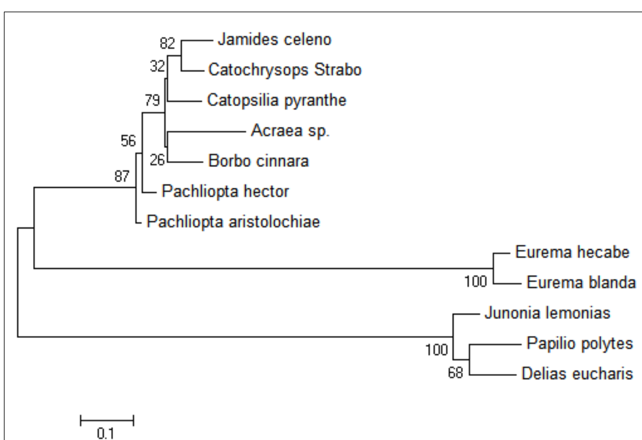

Fig 4: Phylogenetic tree of the COI gene sequences of 12 different butterflies

Discussion

Phylogenetics deals with the evolutionary relationships or differences between organisms based on their genetic makeup. The difference or relationships are usually represented by a tree-like structure called Phylogenetic tree whose branches represent points of relatedness/divergence. The numbers given at the branch points are called as the bootstrap values and the higher value represents more relatedness. For example, both the species of the genus *Eurema* (*Eurema hecabe* and *Eurema blanda*) falls under a same branch and has the bootstrap value of 100. On the other hand, species like *Junonia lemonias* and *Acraea sp* though they belong to the same family Nymphalidae, they are placed under different branches, which explains the differences in their genetic makeup. Similarly, the species *Borbo cinnara* commonly known as Rice swift, which belongs to the family Hesperiidae, has been seen to have close relations with that of *Acraea sp.* of the family Nymphalidae.

In this study, as a representative of each family, 1 – 3 butterflies were selected for the molecular phylogenetic studies. DNA was isolated from the butterfly thorax by CTAB – NaCl method with slight modifications. The quality and relative quantity of the DNA were assessed by agarose gel electrophoresis. After destaining, the gel was observed under UV using a gel documentation system, and the image was captured using the attached CCD camera.

The mitochondrial gene coding for Cytochrome oxidase subunit I (COI) was extensively used for studying the molecular diversity of eukaryotic organisms, and hence the same has been adopted in this study as well. Since, the conventional primers failed to amplify the COI gene in butterfly samples, a set of degenerate primers were designed and used for amplification and sequencing. The gene was amplified (~650 bp) and was sequenced by automatic DNA sequencing. Processed sequences were assembled using the Contig Assembly Program version 3 (CAP3). The contigs obtained were compared to the sequences in NCBI database using Basic Local Alignment Search Tool (BLAST).

The processed sequences were then submitted to the GenBank using the submission portal, BankIt for getting accession numbers for each sequence. Phylogenetic analysis was performed using the MEGA6 package. Phylogenetic tree from the multiple aligned sequences were constructed by Neighbor-Joining (NJ) statistical method. Bootstrap type was followed with the number of Bootstrap replications set at 500. Maximum composite-likelihood method was chosen as a default substitution model for the construction of phylogenetic tree. All 12 sequences were aligned primarily with pair-wise alignment followed by multiple alignments. For the alignment of sequences, MEGA6 software package, which employs ClustalW algorithm, was used.

Twelve different species of butterflies representing 5 different families were selected for the molecular diversity analysis. From the 12 selected species, the genomic DNA was isolated by CTAB-NaCl method, and their quality and quantity were tested through agarose gel electrophoresis. The COI gene was amplified using the custom-designed degenerate primer set, which yielded around ~650 bp amplicons. The PCR amplicons were purified and then sequenced by the automated DNA sequence employing dye-termination method. The sequences obtained out of sequencing runs were aligned using the CAP3 program. Nucleotide database of NCBI using the BLAST program. The genus and species names of the closest match from the database were identified based on the BLAST parameters like identity, query coverage and E value. All 12 sequences were submitted in GenBank for public access using the tool BankIt. The release date was chosen as immediate. The accession numbers were provided by GenBank for the sequences.

Conclusions

The present study reveals that RAPD-PCR is extremely useful for rapid identification of genetic polymorphisms in lepidopteron because of reproducibility of the result for each of the species. Since no DNA sequence information is required, RAPD-PCR can be widely used in identification and differentiation of closely related insect species, although large number of random primers is often required.

Thus the results besides showing the relationships and difference between the selected species, also signifies the merits of molecular phylogenetic studies based on evolutionarily conserved genes such as Cytochrome oxidase which clearly predicts the phylogeny in contrast to the conventional methods which are based on visible traits like morphology, wing pattern, colour, etc.

Acknowledgements

This work was financially supported by a grant from the University Grants Commission (UGC-SERO) Hyderabad, India is gratefully acknowledged.

References

1. Tautz D, Arctander P, Minelli A, Thomas RH, Vogler AP. DNA points the way ahead in taxonomy. Nature. 418(6897): 479, 2002.
2. Savolainen V, Cowan R, Vogler A, Roderick G, Lane R. Towards writing the encyclopaedia of life: an introduction to DNA barcoding. Philosophical Transactions of the Royal Society B 360:, 2005, 1805-1811.
3. Hebert PDN, Cywinska A, Ball SL, de Waard JR. Biological identifications through DNA barcodes. Proceedings of the Royal Society of London. Series B, 2003a:270:313-321.
4. Hebert PDN, Ratna singham S, Waard JR. Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. Proceedings of the Royal Society B: Biological Sciences,2003b:270:96-99.
5. Muller RL. Evolutionary Rates, Divergence Dates, and the Performance of Mitochondrial Genes in Bayesian Phylogenetic Analysis. Systematic Biology,2006:55(2):289-300.
6. Lee J, Miya M, Lee Y, Kim C, Park E, Aoki Y *et al*. The complete DNA sequence of the mitochondrial genome of the self-fertilizing fish Rivulus marmoratus (Cyprinodontiformes, Rivulidae) and the first description of duplication of a control region in fish. Gene,2001:280:1-7.
7. Ward RD, Hanner R, Hebert PDN. The campaign to DNA barcode all fishes, FISH-BOL. Journal of Fish Biologym,2009:74:329-356.
8. Hajibabaei M, Singer GAC, Hebert PDN, Hickey DA. DNA barcoding: how it complements taxonomy, molecular phylogenetics and population genetics. Trends In Genetics, 2007:23(4):167-172.