Journal of Advanced Zoology

ISSN: 0253-7214

Volume ***45 Special*** *Issue* ***02 Year 2023*** *Page* ***1222:1239***

DNA barcoding and Phylogenetic analysis for bio-surveillance of insect fauna and their identification in Seri ecosystem

K.Haripriya[1], S.Kalpana[2], D.M.Mamatha [3]

Department of Biosciences & Sericulture
Sri Padmavati Mahila Visvavidyalayam (Women's University), Tirupati, Andhra Pradesh-517502

(Corresponding Author Email Id: drjharipriya7@gmail.com)

Article History
Received: 11 February 2023
Revised: 21 August 2023
Accepted: 05 October 2023

Abstract

Molecular Studies on insect fauna of Seri ecosystem is vital to develop new strategies for pest control in sericulture. The growth of Seri-industry and the quality of silk production are directly interlinked with the pests' control as it damages both Mulberry and Silkworm crops. In the current study, 36 different Mulberry pests of orders Lepidoptera, Hemiptera, Coleoptera, Orthoptera, Diptera, Hymenoptera, Araneae, and Mantodea were collected and reared. The total genomic DNA was isolated from individual insect species. Partial cytochrome oxidase I gene was amplified and molecular signatures were created in the form of DNA barcodes for correct identification and authentication of insect fauna related to the Seri ecosystem. This study is focused on Insect Biodiversity conservation to evolve sustainable measures. Biodiversity conservation studies are essential in Agriculture and its allied areas for biological control of pests. This multi-functional research study contributes to the scientific understanding of molecular systematics with flawless identification of insect fauna and a novel way of pest control and crop protection strategy.

Keywords: Mulberry, Silkworm, DNA barcoding, Insect fauna, Seri-ecosystem, Cytochrome oxidase I gene

1. Introduction

Recent research studies are focused on the modernization of taxonomy due to its slow pace to identify and describe the vast existence of rich biodiversity (Wilson, 2003; Nicole, 2010). Advancement of Insect taxonomic studies implies the characterization of

new species particularly those that damage different crops (Hoagland, 1996). The existence of Insect fauna is poorly documented in Seri-ecosystem. Identification of species is a necessary process in pest control management. Recognition of species using external morphological features is a long and time-consuming process (Sipek & Ahrens, 2011; Hebert et al., 2003) and needs more taxonomical knowledge. DNA barcoding technique attracts the attention of researchers and emerged as a standard tool for recognition and identification of insect species (Raupach et al., 2010; Butcher et al., 2012). DNA barcoding has acquired many applications in the conservation of biodiversity like the identification of new species and helps to understand their description, to identify immature life stages, cryptic species the and explanation of sexual dimorphism (Riedel et al., 2013).

From the past 20 years, creation of molecular signatures to the group of organisms are being used for their identification and classification (Baker & Palumbi, 1994; Sperling et al., 1994; Degalle & Birstein 1996). Identification and classification of unknown species by combining genetic information with taxonomical features has become an advanced research area to be focused on to study vast diversity of insect populations in the animal kingdom. Selection of standard DNA marker is an essential task for the creation of useful DNA barcode (Kress & Erickson 2007). In this regard, the core barcode COI gene, for insects belonging to different orders have shown the best performance in PCR amplification and sequencing. The complete sequencing of mitochondrial DNA has been done in 19 species including two coleopterans and thousands of partial COI gene sequences of insects were found in NCBI database (Bae et al., 2001). However, the samples specimen collected for this study is unique and not covered earlier.

To create a DNA barcode for an unknown species, it is necessary to convert DNA sequences into clean DNA barcodes by editing and aligning of obtained sequences. Standard tools like BioEdit sequence alignment editor (Hall, 1999), MULTALIN (Combot et al., 2000), NCBI-BLAST (Altschul et al., 1997) and MEGA software are used to align and analyse the sequences. The aligned sequences are submitted to BOLD for generation of DNA Barcodes for each species. Therefore, the DNA barcoding data of a particular species are deposited as a project in BOLD database for public access which has been used for Molecular systematics study since many years.

Mitochondrial cytochrome oxidase I (COI) partial gene sequence act as identification fragment in DNA barcoding (Hebert et al., 2003). COI region has the ability to separate closely related species (Harvey et al., 2003), allocation of unknown species to known species and accelerate identification of species having complex external morphology (Frezal & Leblois, 2008). Generally, DNA barcodes contain DNA sequence having 400 to 650 base pairs which can be amplified by polymerase chain reaction. DNA barcoding is an advanced taxonomic approach used to identify insect species by using a unique fragment of a DNA sequence across species. DNA barcoding is a potential taxonomy tool used to recognize known and unknown insect specimens based on the pattern of DNA sequence fragment of the partial mitochondrial COI gene.

Sericulture industry is the most vulnerable sector susceptible to insect pests. Sustainability of mulberry cultivation and silkworm rearing requires observation, recognition, and distinguishing of different species. Due to its perennial nature, mulberry engages many insects' fauna as pests and predators throughout the year based on disparity in seasons (Rangaswamy et al., 1976). Usage of chemical pesticides for managing complicated pest problems in sericulture has shown noticeable results in disease control but left toxicity residues to environment and humankind. Spraying of chemical insecticide on mulberry crops is harmful to silkworms as it causes various diseases. For the proper management of those insect fauna, diversity studies are needed. Conservation of insect biodiversity plays a significant role in the maintenance of the ecosystem. Biodiversity assessment is used to study the species richness in particular area. Accurate identification and authentication of insects is the primary step in biodiversity studies (Narendran, 2001).

Henceforward, collection of various insects is carried out in Seri-ecosystem and their molecular diversity is studied by generating DNA barcodes using partial mitochondrial cytochrome oxidase I gene which further pave ways to generate alternate control methods on various sericulture insect pests. The diversity of insect's fauna in Seri-ecosystem which refers to the variability in insect fauna and their ecologically important host plants, has not yet been recognized completely. The current study focused on molecular conservation of Sericulture related insect species is first of its kind which advances the Sericulture Industry in terms of efficient Mulberry cultivation and advanced Silkworm rearing techniques to produce quality silk.

2. Materials and Methods

2.1. Collection and preservation of Insect fauna

In South India, thirty-six live insect fauna belonging to the orders Araneae, Coleoptera, Diptera, Hemiptera, Hymenoptera, Lepidoptera, Mantodea, Orthoptera were collected from various Sericulture research institutes and farms like Central Sericultural Research Training Institute-Mysore, Regional Sericulture research station-Salem, and from mulberry gardens of Department of Sericulture, Sri Padmavati Women's university, Tirupati. Insects were collected by hand picking, net sweeping and visual searching methods. The insect species considered for this study are pests and predators on mulberry and silkworm crops. Photographs of all insect specimens were taken. Insect species are morphologically identified with the help of available keys and catalogues (Chanthy et al., 2010; Gupta & Singh 2013; Murthy et al., 2015). Collected samples were washed with alcohol and placed in 95% alcohol which were labeled legibly with details of collection place, date and collector names. All samples were preserved at -20^{O}C for further molecular processes. The insects thus collected are morphologically recognized and used for DNA barcoding studies.

2.2. Genomic DNA isolation

Total Genomic DNA samples were extracted from frozen insect specimens stored in 100% ethyl alcohol. The total DNA was isolated from each insect sample using DNeasy kit (Qiagen), following the manufacturer's instructions. A total of thirty-six insect species belonging to eight orders were applied for total genomic DNA preparations. The integrity of DNA was checked on 1% agarose gel by gel electrophoresis and visualized in gel doc system. The quantification of extracted DNA was measured using nano-spectrophotometer values at 260/280 nm absorbance.

2.3. Amplification of partial mitochondrial COI gene

Partial mitochondrial cytochrome oxidase I gene was amplified from the extracted DNA of each insect sample through polymerase chain reaction (PCR) carried out in Gradient Master cycler Nexus (Applied Biosystems). Universal standard primer pairs viz., LepF, LepR, & LCO, HCO (Table 1) were used to amplify the known region of cytochrome oxidase I gene. The sequence information of primers was retrieved from Barcode of Life data system database and the synthesized oligos were purchased (Eurofins Scientifics). The volume of PCR reaction was made up to 50µl reaction containing 5µl (300µg/µl) of genomic DNA, 25µl of EmeraldAmp MAX PCR Master mix (Clontech), 2.5µl (10pm/µl) forward primer, 2.5µl (10pm/µl) reverse primer, 15µl of PCR grade water. The PCR reaction conditions were standardized and used. The range of annealing temperatures for amplification was from 45^0C -55^0C for all collected insect samples. 5 µl of each PCR sample was analyzed on 1.5% agarose gel in Tris-Acetic acid-EDTA buffer stained with ethidium bromide (0.5µg/ml) to check the successful amplification. 200bps DNA ladder (TAKARA) was used as DNA standard. The gel was visualized and gel pictures were taken using Gel documentation system. The PCR-generated COI amplicon of ~650 bps purified and eluted using NucleoSpin Gel and PCR Cleanup kit. Most clear and high integrity PCR products were selected for sequencing with the Big Dye Terminator cycle sequencing reaction kit and carried out in ABI DNA sequencer. Amplified products were sequenced in both directions with the same primers used for PCR to get accurate COI gene sequence information.

Table 1: Primers used for COI gene amplification

S. No.	Primer name	Primer sequence	Direction
1	Lep_F	5' ATTCAACCAATCATAAAGATATTGG 3'	Forward
2	Lep_R	5'TAAACTTCTGGATGTCCAAAAAATCA3'	Reverse
3	LCO1480	5'GGTCAACAAATCATAAAGATATTGG 3'	Forward
4	HCO2198	5'TAAACTTCAGGGTGACCAAAAAATCA3'	Reverse

2.4. Data Interpretation

The trace files of COI gene sequences were visualized and edited using 'Codoncode aligner' software (www.codoncode.com). The primer sequences,

ambiguous bases, noisy peaks were removed from each raw sequence and combined both forward and reverse sequences for the consensus sequence formation. All consensus sequences were verified for presence of any gaps and stop codons by translating them as protein sequences and the sequences were exported in FASTA format. The conservity among the COI gene sequences of insect species was studied using ClustalW (Kumar et al., 2018). CD-search tool of NCBI (National Centre for Biotechnology Information) was used to identify the domain regions in COI gene sequences of insect species. Further, sequence analysis was done in MEGA server. The complete aligned sequences were submitted to BLASTN (Basic Local Alignment Search Tool-Nucleotide) (Altschul et al., 1990) search to identify the closely related species and to assign them to the known species. The files to generate DNA barcodes for the edited COI gene sequences namely specimen data file, taxonomic data file, collection data files were prepared and submitted to the BOLD database (Barcode of life data system) (Ratnasingham & P.D.N. Hebert 2002). Basic nucleotide sequence statistics like nucleotide composition, frequencies of A, T, G, C base pairs, AT pairs, GC pairs overall transition/transversion ratio (ts/tv), pairwise nucleotide distances using kimura-2 parameter (K2P) (Kimura, 1980) were calculated using MEGAX (Kumar et al., 2018). Nucleotide polymorphism features such as conserved sites, variable sites, parsimony informative sites, singleton sites, insertion-deletion polymorphism (Indels), haplotype diversity (Hd), nucleotide diversity (Pi) were determined to evaluate genetic variability among the insect species using DnaSP (Librado et al., 2009) software to study the genetic distance across the insect species.

2.5. Phylogeny studies

To demonstrate the genetic relation across the collected insect species three types of phylogenetic trees were constructed using NJ (Neighbour Joining) method, Maximum likelihood (ML) and Bayesian Inference (BI) methods respectively. One thousand boot strap replications and P-distance model parameters were used to create NJ (Neighbour Joining) (Saitou & Nei, 1987) tree. One thousand boot strap replications and Tamura 3-parameter method was used to form the Maximum likelihood (ML) phylogenetic tree (Tamura et al., 2011). The branch swap filter values were set as weak (40-60%), moderate (61-75%), good (76-88%) and strong support value >89%. Both trees were generated in MEGA software. Bayesian inference analysis of partial COI genes was performed in MrBayes ver. 3.2.7 Server (Ronquist et al., 2003). General time reversible (GTR) þG method was set as the best fitting model. Markov Chain Monte Carlo (MCMC) algorithm was repeated twice upto 2 million generations and sampling of trees was done every hundredth generations with four independent chains running concurrently and first 5,000 trees were abolished as burn-in. The posterior side of Bayesian tree branching pattern probabilities were calculated by using remaining trees with the help of 50% majority-rule consensus tree. FigTree (http://tree.bio.ed.ac.uk/ software/figtree/) software package was used to check and visualize the Bayesian tree.

3. Results

3.1 Sample Collection and COI gene Amplification

The thirty-six insect species collected from various sericulture research centres and mulberry gardens were processed for molecular studies. Among the insect fauna, 10 Lepidoptera, 10 Hemiptera, 2 Diptera, 8 Coleoptera, 3 Orthoptera, 1 Hymenoptera, 1 Araneae and 1 Mantodea species (Table 2) were identified based on their morphometric features. 1% agarose gel analysis showed high quality genomic DNA bands of all insect samples. Nano-spec readings of genomic DNA at 260/280 absorbance were found in the range of 1.6 to 2.0.

Table 2: Taxonomic details of 36 Insect fauna of Seri-ecosystem

S.NO	Sample ID	Order	Family	Genus	Species
1	DP	Lepidoptera	Crambidae	*Diaphania*	*pulverulentalis*
2	SM	Lepidoptera	Pieridae	*Delias*	*eucharis*
3	SUK2	Lepidoptera	Nymphalidae	*Acraea*	*violae*
4	HP2	Lepidoptera	Pyralidae	*Etiella*	*behrii*
5	WF	Lepidoptera	Erebidae	*Eressa*	*strepsimeris*
6	SMU4	Lepidoptera	Geometridae	*Cleora*	*repulsaria*
7	SMU10	Lepidoptera	Erebidae	*Olene*	*mendosa*
8	PAB3	Lepidoptera	Pyralidae	*Pyla*	*impostor*
9	SUK3	Lepidoptera	Xyloryctidae	*Tymbophora*	*peltastis*
10	SB	Lepidoptera	Noctuidae	*Spodoptera*	*litura*
11	PI	Hemiptera	Pseudococcidae	*Paracoccus*	*marginatua*
12	HP3	Hemiptera	Aphididae	*Toxoptera*	*odinae*
13	SUK6	Hemiptera	Pentatomidae	*Acrosternum*	*gramineum*
14	SUK1	Hemiptera	Pentatomidae	*Eocanthecona*	*furcellata*
15	MD	Hemiptera	Pentatomidae	*Halyomorpha*	*picus*
16	RC	Hemiptera	Rhyparochromidae	*Dieuches*	*schmitzi*
17	SMU8	Hemiptera	Pentatomidae	*Nezara*	*viridula*
18	PAB8	Hemiptera	Aphididae	*Aphis*	*glycines*
19	PA	Hemiptera	Pseudococcidae	*Maconellicoccus*	*hirsutus*
20	SMU12	Hemiptera	Aleyrodidae	*Aleurodicus*	*dispersus*
21	PAB5	Diptera	Sphaeroceridae	*Spelobia*	*bifrons*
22	UF	Diptera	Tachinidae	*Exorista*	*sorbillans*
23	PAB7	Coleoptera	Coccinellidae	*Calvia*	*quatuordecimguttata*
24	WG	Coleoptera	Coccinellidae	*Cryptolaemus*	*montrouzieri*
25	S2MB	Coleoptera	Coccinellidae	*Calvia*	*championorum*
26	2MB	Coleoptera	Coccinellidae	*Calvia*	*punctata*

27	SMU7	Coleoptera	Curculionidae	*Myllocerus*	*viridanus*
28	SMU1	Coleoptera	Coccinellidae	*Coccinellidae*	-
29	CO	Coleoptera	Dermestidae	*Attagenus*	*fasciatus*
30	SMU5	Coleoptera	Coccinellidae	*Illeis*	*cincta*
31	PAB1	Orthoptera	Tettigoniidae	*Hexacentrus*	*japonicus*
32	PAB6	Orthoptera	Pyrgomorphidae	*Neorthacris*	*acuticeps*
33	SMU6	Orthoptera	Pyrgomorphidae	*Sphenacris*	*crassicornis*
34	SMU9	Hymenoptera	Formicidae	*Meranoplus*	*magrettii*
35	SMU2	Araneae	Salticidae	*Telamonia*	*dimidiata*
36	MC	Mantodea	Hymenopodidae	*Odontomantis*	*pulchra*

Partial COI gene amplification in Insect DNA samples was confirmed based on agarose gel image analysis. In 1.5% agarose gel, high integrity amplified bands were detected. In each sample, the amplification band was found nearly at 650 bps. Further, PCR products were extracted from agarose gel and purified using Gel extraction kit (TAKARA). Gel clean-up step was done to remove excess primers, unincorporated base pairs and to avoid errors during sequencing process. Later, the clean-up samples were submitted to Sanger's Dideoxy sequencing to obtain partial COI gene sequence information.

Seventy-two trace files were obtained from sequencing. Each sample was sequenced in both the ends i.e., forward and reverse. The trace files of COI gene sequences were edited and combined to form consensus sequences. The exported FASTA format of COI gene sequences of all insects along with their details were uploaded to Barcode of life data systems (BOLD) database. After processing, finally, COI gene sequence of each insect sample got BIN number and DNA barcode. COI gene sequences were translated and conserved regions were identified using Clustal-W in MEGA software.

Comparative studies concluded that among the lepidopterans, *Diaphania pulverulentalis*, *Olene mendosa*, *Cleora repulsaria*, *Etiella behrii*, were close to each other as their alignment score was above 70% among them. *Acrea violae*, *Pyla impostor*, *Delias eucharis* has shown above 50% conserved regions between them. Between *Spodoptera litura* and *Tymbophora peltastis*, the alignment was found to be greater than 50%.

Among the Hemipterans, more conserved regions (above >50) were found between *Aleurodicus dispersus*, *Toxoptera odinae*, *Maconellicoccus hirsutus*, *Eocanthecona furcellata*. *Nezara viridula* was more related to *Dieuches schmitzi* (77.7%), *Halyomorpha picus* (72.8%). *Aphis glycines* and *Toxoptera odinae* were conserved with 65% alignment score. *Paracoccus marginatus*, *Acrosternum gramineum*, *Eocanthecona furcellata* has shown less similarity with remaining hemipterans.

Among the Coleopterans, *Cryptolaemus montrouzieri* has shown more similarity with *Coccinellidae species*. *Clavia punctata* has shown more similar regions with *Myllocerus viridanus* (67.8%) and *Calvia championorum* (54.6%). High conserved regions (>60%) were observed between Coccinellidae species, *Calvia quaturodecimguttata, Illeis cincta*. *Attagenus fasciatus* has shown less similar regions with remaining Coleopteran species.

Meranoplus magrettii of Hymenoptera has shown 50% conserved regions with *Nezara viridula* of Hemiptera. *Sphenacris crassicornis* and *Neorthecris acuticeps* of Orthopterans were close to each other with 82.7% similarity. *Hexacentrus japonicus* of Orthoptera has shown 70% similar regions with *Macrobrachium rosenbergii* of Crustacean which was taken as reference. *Odontomantis pulchra* of Mantodea species has shown 73% conserved regions with *Diaphania pulverulentalis* of lepidopteran species.

Exorista sorbillans of Diptera has 65% alignment with *Marcobrachium rosenbergii* of Crustacea. *Spelobia bifrons* of Diptera and *Telamonia dimidiata* of Araneae were shown more than 70% similar regions with *Diaphanaia pulverulentalis* of lepidopteran species. *Macrobrachium rosenbergii* of Crustacea, an out-group member was showing 72% similar regions with *Hexacentrus japonicus* of Orthoptera. The conserved regions of COI gene sequences of all insect species were studied.

Based on the conserved regions, Domain analysis was done. Both Heme_Cu_Oxidase_I_super family and D-pathway domains were observed in COI translated sequences of *Delias eucharis, Acraea violae, Etiella behrii, Pyla impostor, Tymbophora peltastis, Spodoptera litura, Paracoccus marginatus, Toxoptera odinae, Eocanthecona furcellata, Aphis glycines, Maconellicoccus hirsutus, Aleurodicus dispersus, Exorista sorbillans, Calvia quaturodecimguttata, Cryptolaemus montrouzieri, Coccinellidae sps., Illeis cincta, Hexacentrus japonicus.*

Only Heme_Cu_Oxidase_I_super family domain was identified in COI protein sequences of *Acrosternum gramineum, Attagenus fasciatus*. In remaining species none of the domains were identified. BLAST analysis was done to judge the partial COI gene sequences and to assign them to the known species in the reference databases. The sequence comparison studies revealed that the partial COI gene sequences are in close resemblance (>95%) to the concerned mentioned species. Genetic variability studies were carried out with all these respective partial gene sequences.

The cumulative nucleotide frequencies of these insect fauna were 34.5% (A), 33.7% (T), 15.8%(G), 16.0%(C). AT bases are rich when compared to GC in all COI gene sequences as standard as insect mitochondrial DNA. The translational substitutions i.e., rate of inter change of purines or pyramidines are dominant when compared to the transversion substitutions i.e., rate of interchange between purines and pyramidines in COI gene sequences. The overall rate of transition / transversion ratio R=0.80%. No Indels were found in COI gene sequences. A total of 187 polymorphic sites were recognized. 7 variable sites, 14 singleton variable sites and 173 parsimony informative

sites were identified in partial COI gene sequences. The total haplotype diversity (Hd) and nucleotide diversity (Pi) were calculated as 0.997 and 0.44366 respectively.

The higher values obtained in the pairwise distances as shown in fig.3 revealed the distance between two sequences and low values concluded the close relation between two sequences indicating the common ancestry. As per the P-distance calculations among the COI genes, it is reported that *Delias eucharis*, *Pyla impostor*, *Acraea violae* were divergent with other lepidopterans as those species contains high P-distance values. Among Hemipterans *Paracoccus marginatus*, *Nezara viridula*, *Dieuches schmitzi*, *Eocanthecona furcellata*, *Acrosternum gramineum*, *Halyomorpha picus* got high P-distance values. *Aphis glycines* was closely related to *Toxoptera odinae* as the P-distance value was very less among those two species (0.000).

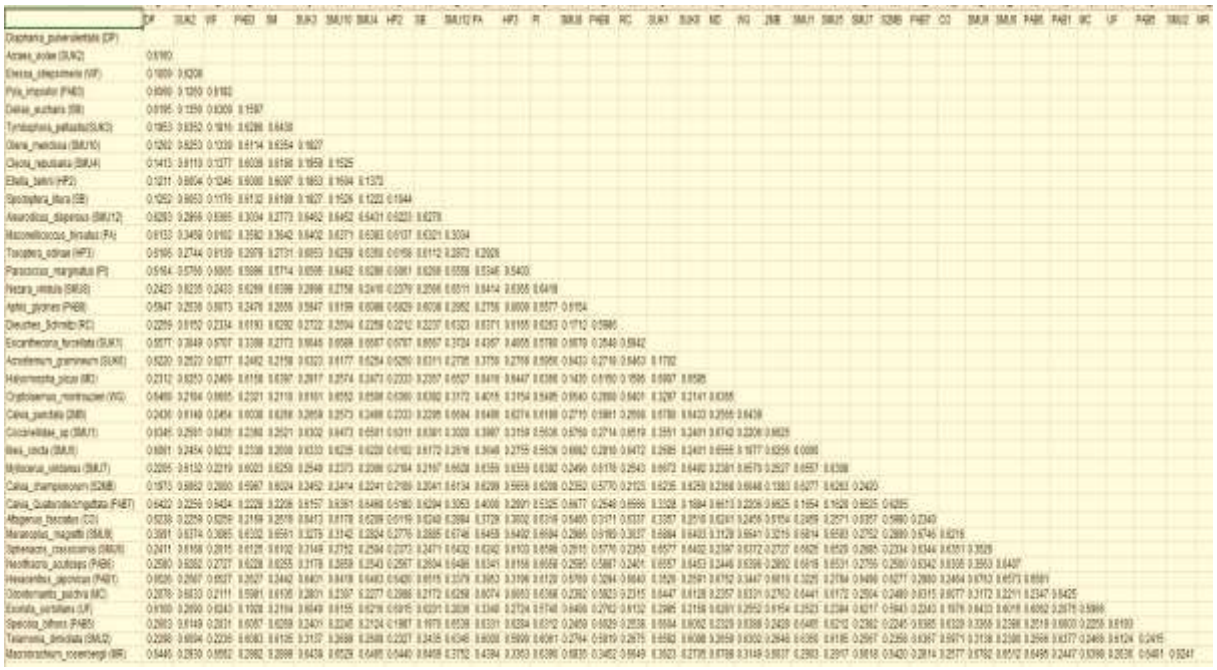

Table Pairwise genetic distances among COI sequences of Insect fauna

In Coleopterans, *Illeis cincta* and Coccinellidae species are more closely related species as their P-distance was 0.000. The two dipteran species having high P-distance value and concluded their divergence. In three orthopteran species *Sphenacris crassicornis* has shown more similarity with *Neorthacris acuticeps* as it shows low p-distance. *Neorthacris acuticeps* and *Hexacentrus japonicus* are divergent to each other as per their pairwise distance values. *Odontomantis pulchra* has shown high divergent with *Eocanthecona furcellata* of Hemipteran species and less divergent with *Diaphania pulverulentalis* of Lepidoptera. *Telamonia dimidiata* of Araneae was more divergent with *Eocanthecona furcellata* of Hemiptera and it is closely related to *Eressa strepsimeris* of Lepidoptera.

3.2 Phylogenetic analysis

Phylogenetic studies were carried out using partial mitochondrial COI gene sequence of Insect species by NJ (Neighbour Joining) method, Maximum likelihood

(ML) and Bayesian Inference (BI) methods. The generated phylogenetic trees were shown almost same results. In NJ tree, most of the clades are having above 75% branch support which concluded that those species are strongest clades in the entire tree. The consistency index and length of the consensus NJ (Fig.1.) tree was 0.050 and 100. All insect species of the current study are identical or closely related and emerged as monophyletic groups.

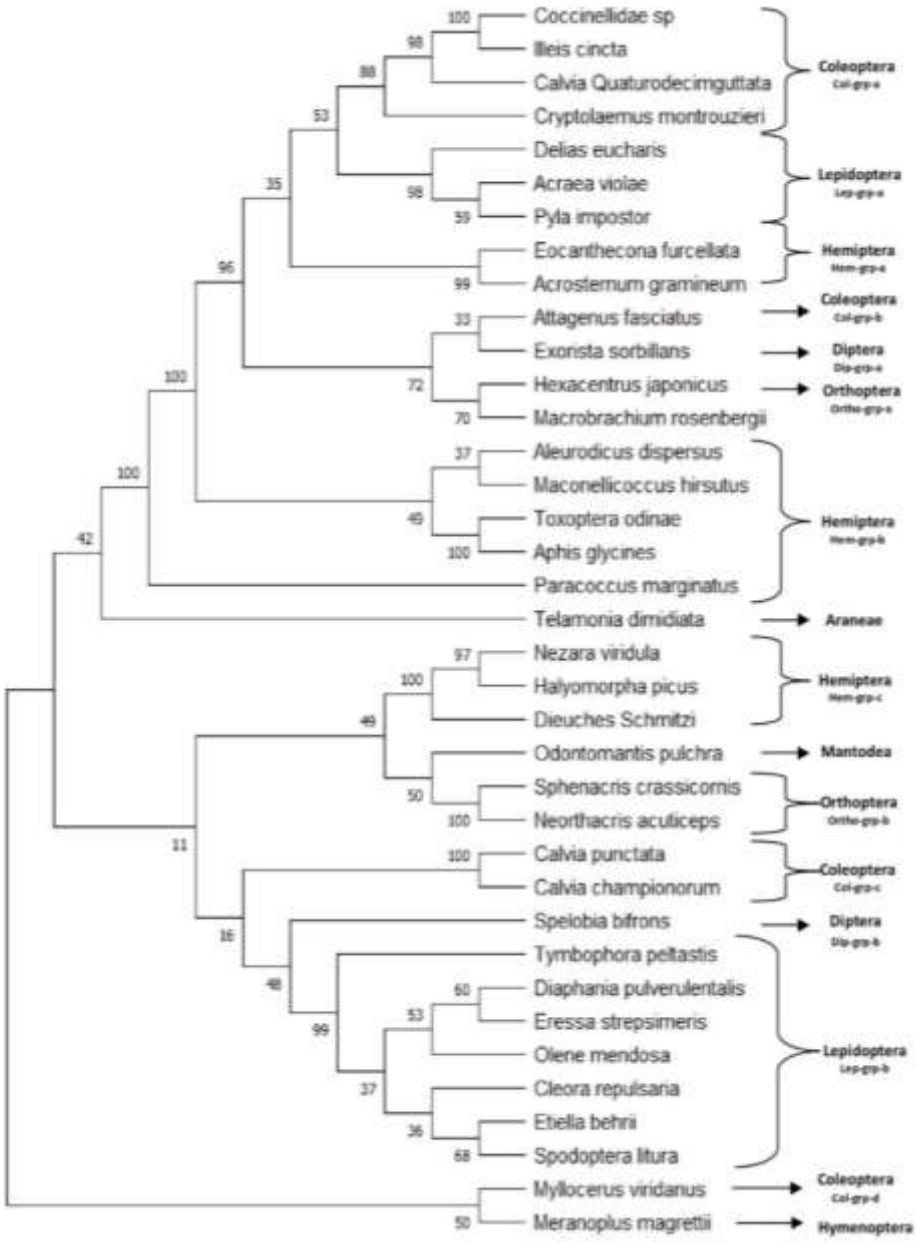

Fig:1. Phylogenetic tree explaining distances between insect fauna based on COI gene sequences by Neighbour Joining method. The bootstrap values and branch lengths were shown.

Among lepidopterans *Diaphania pulverulentalis*, *Eressa strepsimeris*, *Olene mendosa*, *Cleora repulsaria*, *Etiella behrii*, *Spodoptera litura* were formed as one cluster. But the branch length was below 75 and came under below moderate support tree. *Delias*

eucharis, *Acraea violae* and *Pyla impostor* were formed as one cluster with strong branch support value above 95 and *Tymbophora peltastis* has emerged separately and exhibited more divergence among the lepidopterans. *Paracoccus marginatus* of Hemiptera, *Telamonia dimidiata* of Araneae and *Spelobia bifrons* of Diptera species were emerged as separate branches but not as monophyletic groups and revealed that these species are more divergent among thirty-six insect species selected for the study.

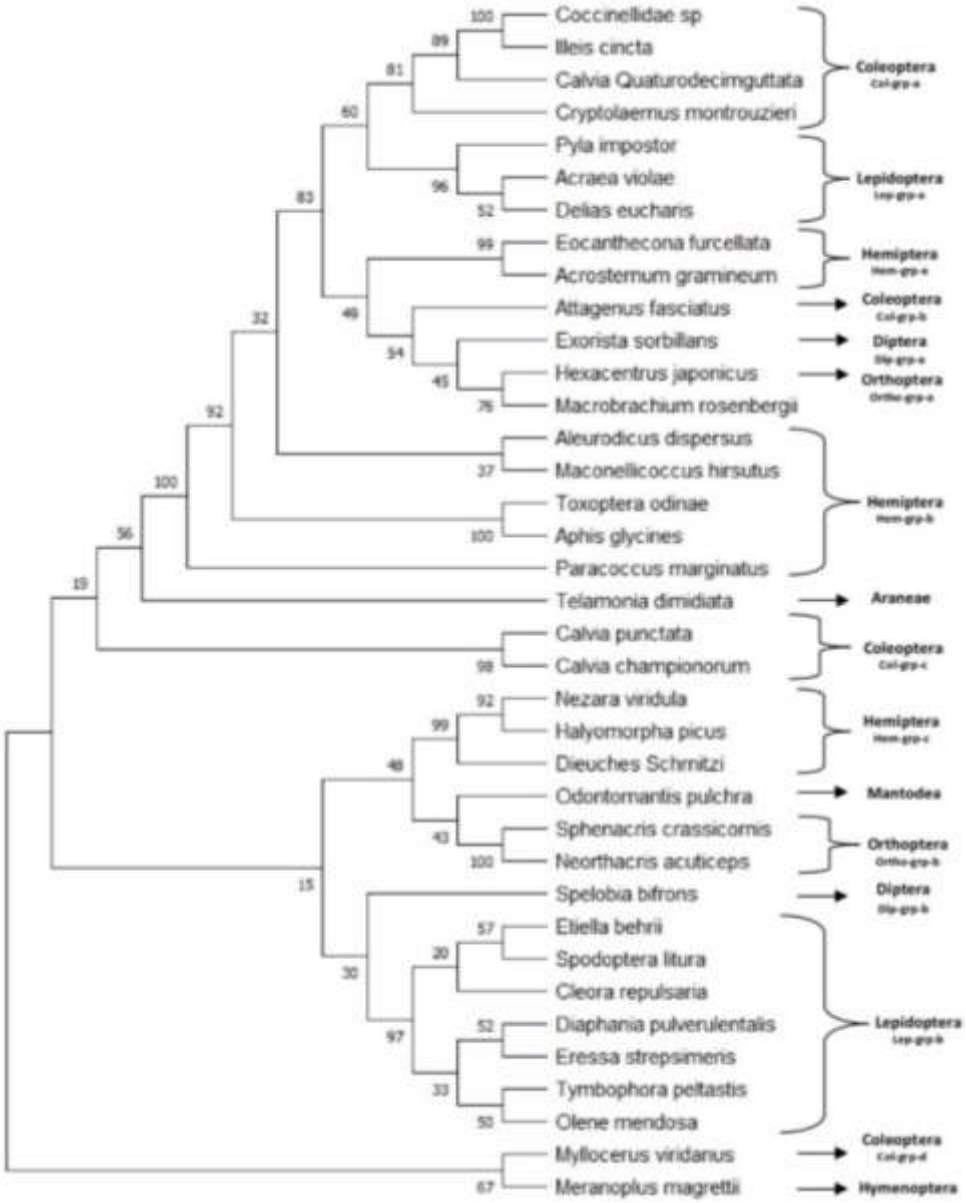

Fig:2. Phylogenetic tree explaining distances between insect fauna based on COI gene sequences by Maximum likelihood criterion. The bootstrap values and branch lengths were shown.

Coccinellidae species, *Illeis cincta*, *Calvia quaturodecimguttata*, *Cryptolaemus montrouzieri* of Coleoptera were formed as a monophyletic group. Hemipteran species such as *Eocanthecona furcellata* and *Acrosternum gramineum* are formed as a clade with

strong branch support value. *Attagenus fasciatus* of Coleoptera along with *Exorista sorbillans* of Diptera formed as a same clade with poor branch support.

Hexacentrus japonicus of Orthoptera formed the clade with *Macrobrachium rosenbergii* which was selected as outgroup. *Aleurodicus dispersus*, *Maconellicoccus hirsutus* formed the sister clade with *Toxoptera odinae* and *Aphis glycines* among Hemipterans.

Nezera viridula, *Halyomorpha picus*, *Dieuches schmitzi* of Coleoptera, *Odontomantis pulchra* of Mantodea, *Sphenacris crassicornis* and *Neorthacris acuticeps* of Orthoptera formed as a single cluster. *Calvia punctata*, *Calvia championorum* of Coleoptera are formed as a single clade. *Myllocerus viridanus* of Coleoptera and *Meranoplus magrettii* of Hymenoptera emerged as separate monophyletic group directly from ancestral node.

The phylogenetic tree formed by ML (Fig.2.) tree was as same as NJ tree. Alike in NJ tree, the species *Tymbophora peltastis* forms a clade with remaining lepidopteran species in ML tree. Branch lengths, cluster formations, divergence between species and monophyletic groups are likewise both in NJ and ML trees.

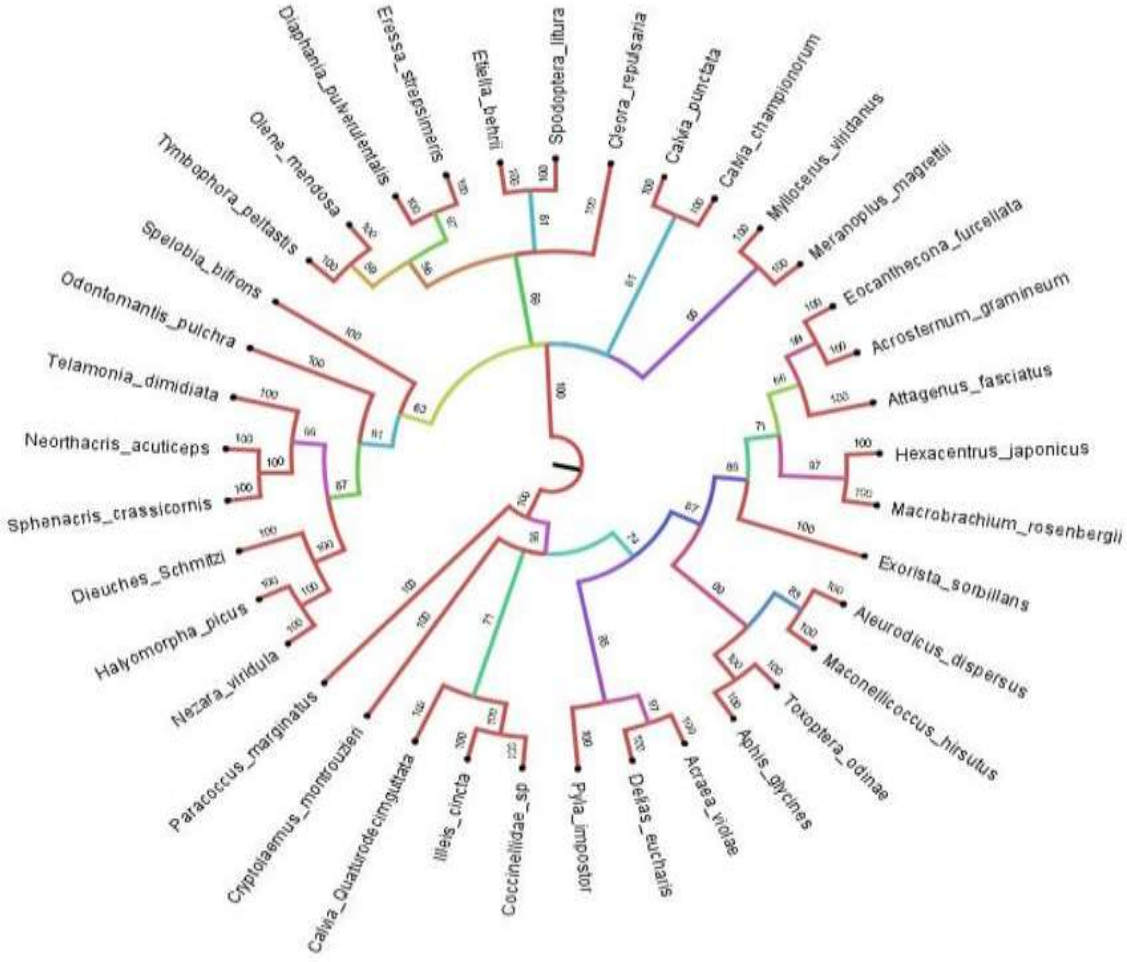

Fig:3. Phylogenetic tree - COI gene sequence by Bayesian inference method

In BI tree analysis (Fig.3.), eight monophyletic groups are identified. *Paracoccus marginatus* of Hemiptera, *Cryptolaemus montrouzieri* of Coleoptera species formed as individual clades and concluded that those two species are more divergent to other insect fauna. Remaining insect species were formed as clades as similar as in NJ and ML trees. But the probability percent of most of the branches have very strong support as the values are above 90 and 100% percent.

Number of analysed indel sites were also calculated based on the clusters formed in phylogenetic trees. The pairwise distances and indel sites strongly supports the phylogenetic studies of COI gene sequences of insect species.

4. Discussion

DNA barcoding is a molecular tool essentially helpful for rapid assessment of biodiversity, molecular phylogeny and evolution (Pei Nancei et al., 2017). The methodology involves sequencing of a short region of mitochondrial cytochrome c oxidase I (COI) gene, which is termed as a most conserved region. Due to larger taxonomic content of DNA and rapid sequencing technologies, there is a need to construct the tree at even larger scales (Hibbett et al., 2005; Michelle et al., 2006). However, traditional phylogenetic methods may face difficulties to maintain larger data sets. Hence, newly developed phylogenetic methods are more efficient to solve the above problem and it became a way for detailed studies on the pattern of evolution of mitochondrial and nuclear genes (Simon et al., 1994) and regenerate the phylogenetic tree based on DNA sequences to determine their evolutionary relationships at various taxonomic levels (Swafford 1996).

In the current study, DNA barcodes were generated by amplifying the standard COI gene marker that can be given its role in taxonomy, as molecular sequencing of DNA as a universal unit for recognition, which is exact and fast reliable method for identification of organisms at species level. Domain analysis was performed using CD-search tool (NCBI) based on the conserved regions based on multiple sequence Alignment. Domain analysis revealed that, among lepidopterans, *Delias eucharis, Acraea violae, Etiella behrii, Pyla impostor, Tymbophora peltastis* show one domain region belonging to Heme-C4-Oxidase-1, super family (cl00275) with two proton channels. In *Delias eucharis* and *Acraea violae*, second domain region D- pathway proton channel was identified between 20-40 amino acids. In *Etiella behrii* and *Spodoptera litura*, D-pathway channel was identified at first amino acid of the protein sequence. In *Pyla impostor* and *Tymbophora peltastis* D-pathway channels were found in between 30-40 amino acids. In *Diaphania pulverulentalis, Eressa strepsimeris, Cleora repulsaria, Olene mendosa* no conserved regions were observed.

Among Hemipterans, *Paracoccus marginatus, Toxoptera odinae, Acrosternum gramineum, Eocanthecona furcellata, Aphis glycines, Maconellicoccus hirsutus, Aleurodicus dispersus* showed conserved domain region belonging to Heme-C4-Oxidase-super family. D-pathway proton channel was identified in *Paracoccus marginatus,*

Toxoptera odinae, Eocanthecona furcellata, Aphis glycines, Maconellicoccus hirstus and *Aleurodicus dispersus*. No conserved regions were found in *Halyomorpha picus, Dieuches schmitzi, Nezara viridula*. Between two Dipteran species, *Exorista sorbillans* has shown the domain region of Heme-C4-Oxidase-1 Super family and D-Pathway, but no region was found in *Spelobia bifrons.* Among Coleopterans, *Calvia quaturodecimguttata, Cryptolaemus monotrouzieri, Coccinellidae sps, Attagenus fasciatus, Illeis cincta* contains a domain region of Heme-C4-1 oxidase-I super family. D-Pathway proton channel was also present in above all species except in *Attagenus fasciatus*.

In orthopterans, *Hexacentrus japonicus* has shown the conserved domain regions of Heme-C4-oxidase-I super family and D-pathway proton channel. No domain region was reported in *Neorthacris acuticeps*, *Sphenacris crassicornis. Telamonia dimidiata* of Araneae does not display any conserved domain region in its partial COI gene sequence. The partial COI gene sequence of *Meranoplus magrettii* (Hymenoptera) and *Odontomantis pulchra* (Mantodeae) did not show any domain. In the above all species, COI gene sequences are compared with partial COI gene sequence of *Macrobrachium rosenbergii* and complete mitochondrial genome of *Aedes aegypti* (Diptera), *Spodoptera litura* (Lepidoptera) as reference. The regions which were conserved in above species are transmembrane complexes which participate in the respiration process of prokaryotes and mitochondria catalysing the reduction of oxygen and distributing protons across the membrane. Two proton channels D-pathway and K-pathway have been found in subunit 1 of Cytochrome c Oxidase (CcO) and ubiquinol oxidase, plays an important role as well defined pathway to transfer the pumped protons.

DNA-based identification of insect fauna should be a part of procedure of integrated pest management in providing information regarding pest biology, ecology and behaviour to study the ecological balance of particular habitat. Furthermore, identification of species answers the queries related to species differentiation based on the mitochondrial COI gene.

The created phylogenetic trees based on partial COI gene sequences in the present study have explained the cluster formation, closely related and separated the divergent species among the Insect fauna. The phylogenetic analysis of Insect fauna collected from Seri-ecosystem was resoluted by the NJ, ML and BI methods. This study signifies a wide-ranging phylogenetic analysis and evolutionary relationships of Insect fauna of Seri-ecosystem belonging to different orders of phylum Arthropoda based on molecular data. Phylogenetic studies results showed that the species belonging to Lepidoptera, Hemiptera, Coleoptera, Hymenoptera, Orthoptera, Diptera, Araneae and Mantodea were formed as monophyletic groups with strong branch support values. The more divergent species among 36 Insect fauna are mentioned in results. The comparative assessment of phylogenetic studies of insect fauna of Seri-ecosystem is first of its kind. Molecular systematic studies of rare species in Seri-ecosystem were also discussed.

Henceforward, this study is a major area to be focused on Insect Biodiversity conservation to evolve sustainable measures. Biodiversity conservation studies are

essential in Agriculture and its allied areas for biological control of pests. This multi-functional research study contributes to the body of knowledge of scientific understanding on the molecular systematics flawless identification of insect fauna and a novel way of pest control and crop protection strategy, which can contribute towards increased productivity through crop management for the farmers in averting and mitigating the crop losses due to insect pest attack.

List of Abbreviations

COI gene - Mitochondrial Cytochrome C Oxidase I gene
PCR - Polymerase Chain Reaction
NJ - Neighbour Joining method
ML - Maximum Likelihood
BI - Bayesian Inference
GTR - General Time Reversible
MCMC - Markov Chain Monte Carlo
BOLD - Barcode Of Life Data systems

Acknowledgments

The study was carried out in the DNA Barcoding Laboratory of Dept. of Biosciences and Sericulture, Sri Padmavati Mahila Visvavidyalayam, Andhra Pradesh, India. The authors are thankful to the University for providing the necessary laboratory facilities.

References

A. Riedel, K. Sagata, Y. R. Suhardjono, R. Tänzler, M. Blake, Integrative taxonomy on the fast track—towards more sustainability in biodiversity research, Front. Zool, 1186(2013)10-15.

B. A. Butcher, M. A. Smith, M. J. Sharkey, D. L. J. Quicke, A turbo-taxonomic study of Thai Aleiodes (Aleiodes) and Aleiodes (Arcaleiodes) (Hymenoptera: Braconidae: Rogadinae) based largely on COI barcoded specimens, with rapid descriptions of 179 new species, Zootaxa 3457(2012)1–232.

C. Combet, C. Blanchet, C. Geourjon, G. Deléage, NPS@: network protein sequence analysis, Trends Biochem. Sci, 25(2000)147–150.

C. S. Baker, S. R. Palumbi, Which whales are hunted – a molecular genetic approach to monitoring whaling, Science 265(1994)1538–1539.

C. Simon, F. Frati, A. Beckenbach, B. Crespi, H. Liu, P. Flook, Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers, Ann. Entomol. Soc. Am, 87(1994)651–701.

D. L. Swofford, G. J. Olsen, P. J. Waddell, D. M. Hillis, Phylogenetic Inference, In: D. M. Hillis, C. Moritz, B. K. Mable, Eds, Molecular Systematics, 2nd Edition, Sinauer Associates, Sunderland (MA), (1996)407-514.

D. S. Hibbett, R. H. Nilsson, M. Snyder, M. Fonseca, J. Costanzo, M. Shonfeld, Automated phylogenetic taxonomy: an example in the homobasidiomycetes (mushroom-forming fungi), Syst Biol,54(4) :(2005)660-8.

E. O. Wilson, The encyclopaedia of life, Trends in Ecology and Evolution, 18(2003) 77–80.

F. A. H. Sperling, G. S. Anderson, D. A. Hickey, A DNA-based approach to the identification of insect species used for postmortem interval estimation, J Forensic Sci 39(1994)418–427.

F. Ronquist, J. P. Huelsenbeck, MrBayes 3: Bayesian phylogenetic inference under mixed models, Bioinformatics, 19 (2003) 1572–1574.

G. Rangaswami, M. N. Narsimhanna, K. Kashiviswanathan, C. R. Sastry, M. S. Jolly, Mulberry cultivation, Sericulture manual 1, FAO, Rome (1976).

I. Kim, J. S. Bae, H. D. Sohn, P. D. Kang, K. S. Ryu, B. H. Sohn, W. B. Jeong, B. R. Jin, Genetic homogeneity in the domestic silkworm, Bombyx mori, and phylogenetic relationship between B. mori and the wild silkworm moth, B. mandarina, using mitochondrial COI gene sequences. International Journal of Industrial Entomology, 1(2000)9–17.

J. S. Bae, I. Kim, S. R. Kim, B. R. Jin, H. D. Sohn, Mitochondrial DNA sequence variation of the mushroom pest flies, *Lycoriella mali* (Diptera: Sciaridae) and *Coboldia fuscipes* (Diptera: Scatopsidae), in Korea, Applied Entomological Zoology, 36(2001) 451–457.

K.E. Hoagland, The taxonomic impediment and the convention on biodiversity, Assoc. Syst. Coll. News, 24(1996), 62-67.

Koichiro Tamura, Daniel Peterson, Nicholas Peterson, Glen Stecher, Masatoshi Nei, Sudhir Kumar, MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods, Molecular Biology and Evolution, 28(10) :(2011)2731–2739.

L. Frézal, R. Leblois, Four years of DNA barcoding: current advances and prospects, Infect Genet Evol, 8(2008)727–736.

M. A. Kimura, Simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences, J Mol Evol,16(1980)111–120.

M. J. Raupach, J.J. Astrin, K. Hannig, M. K. Peters, M. Y. Stoeckle, J. W. Wägele, Molecular species identification of central European ground beetles (Coleoptera: Carabidae) using nuclear rDNA expansion segments and DNA barcodes, Frontiers in Zoology, 7(1) :(2010) 26.

M. L. Harvey, I. R. Dadour, G. Silvana, Mitochondrial DNA cytochrome oxidase I gene: potential for distinction between immature stages of some forensically important fly species (Diptera) in Western Australia, Forensic Sci Int, 131(2003)134–139.

M. S. Murthy, S. K. Nagaraj, A. Prabhuraj, C. M. Kalleswaraswamy, Rice leaf folder *Cnaphalocrocis medinalis* (Lepidoptera: Crambidae) on wheat (*Triticum aestivum*; Poales: Poaceae) in India, Florida Entomol, 98(4) :(2015)1269–1270.

Michelle. M. McMahon, Michael. J. Sanderson, Phylogenetic Supermatrix Analysis of GenBank Sequences from 2228 Papilionoid Legumes, Systematic Biology,55(5) :(2006)818–836.

N. Saitou, M. Nei, The neighbour-joining method: a new method for reconstructing phylogenetic trees, Mol. Biol. Evol. 4 (1987) 406-425.

O. Folmer, M. Black, W. Hoeh, R. Lutz, R. Vrijenhoek, DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates, Mol Mar Bio Biotech, 3(1994)294–299.

P. Chanthy, S.Belfield, R.Martin, Insects of upland crops in Cambodia, Australian Centre Int Agric Res,132(2010)ISBN: 978-1-921615-88-7.

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

P. D. N. Hebert, A. Cywinska, S.L. Ball, J.R. deWaard, Biological identifications through DNA barcodes, Proc. Biol. Sci. 270 (2003) 313-321.

P. Librado, J. Rozas, DnaSP v5: a software for comprehensive analysis of DNA polymorphism data, Bioinformatics, 25 (2009) 1451-1452.

P. Sipek, D. Ahrens, Inferring larval taxonomy and morphology in Maladera species (Coleoptera: Scarabaeidae: Sericini) using DNA taxonomy tools, Systematic Entomology, 36(2011) 628-643.

Pei. Nancai, Chen. Bufeng, Kress. W, Advances of Community-Level Plant DNA Barcoding in China. Frontiers in Plant Science, 8(2017).

R. DeSalle, V. J. Birstein, PCR identification of black caviar, Nature 381(1996)197–198.

R. Gupta, D. Singh, Taxonomic notes on five species of the genus *Cletus Stal* (Heteroptera: Coreidae) from northern India with particular reference to their female genitalia, J Entomol Zoo Stud,1(2013)44–51.

S. F. Altschul, T. L. Madden, A. A. Schäffer, J. Zhang, Z. Zhang, Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, Nucleic Acids Res,25(1997)3389–3402.

S. F. Altschul, W. Gish, W. Miller, E. W. Myers, D. J. Lipman, Basic local alignment search tool, J Mol Biol,215(3) :(1990)403–410.

S. Kumar, G. Stecher, M. Li, C. Knyaz, K. Tamura, MEGA X: molecular evolutionary genetics analysis across computing platforms, Mol Biol Evol, 35(6) :(2018)1547–1549.

S. Nicole, Biodiversity analysis through DNA barcoding; application in agri-food and seafoods products [dissertation], Padova: Universita' Degli Studi Di Padova (2010).

S. Ratnasingham, P. D. N. Hebert, Bold: The Barcode of Life Data System (http:// www.barcodinglife.org), Molecular Ecology Notes 7 (3) :(2007)355-364.

T. A. Hall, BioEdit: A User-Friendly Biological Sequence Alignment Editor and Analysis Program for Windows 95/98/NT. Nucleic Acids Symposium Series, 41(1999), 95-98.

T. C. Narendran, Taxonomic entomology: Research and education· in India, Current Science, 81 (5) :(2001)445-447.

W. J. Kress, D. L. Erickson, A two-locus global DNA barcode for land plants: the coding rbcL gene complements the non-coding trnH-psbA spacer region. PloS one, 2(6) :(2007)508.

URLs

www.codoncode.com

http:// www.boldsystems.org

http://tree.bio.ed.ac.uk/software/figtree/