Journal of Entomology and Zoology Studies

Available online at www.entomoljournal.com

E-ISSN: 2320-7078
P-ISSN: 2349-6800
JEZS 2016; 4(3): 27-31

Received: 04-03-2016
Accepted: 05-04-2016

Naina Khullar
Department of Zoology and Environmental Sciences, Punjabi University, Patiala-147002, India.

Devinder Singh
Department of Zoology and Environmental Sciences, Punjabi University, Patiala-147002, India.

Chandan Kumar Jha
Department of Human Genetics, Punjabi University, Patiala-147002, India.

Correspondence
Devinder Singh
Department of Zoology and Environmental Sciences, Punjabi University, Patiala-147002, India.

Short *COI* marker: A valuable tool for identification and phylogenetic analysis of 6 forensically important blow fly species from India

Naina Khullar, Devinder Singh, Chandan Kumar Jha

Abstract
There are complex morphological differences in the species of blow flies associated with the decomposing corpses in different habitats and environments. A thorough diligent examination of each specimen is required for identifying them, as species associated with an exposed corpse differ from those on the buried or burnt body or when the corpse is writhed after death. Molecular data serves as a lucrative tool in identifying insect specimens, suitably when the specimens are damaged and their morphological identification is arduous. Mitochondrial cytochrome oxidase subunit I (*COI*) DNA of 6 blow fly species were analysed to study its application value as the identification marker. The sequencing of a 350bp *COI* gene of these six species was done, and the data was eventually analyzed. Phylogenetic tree was constructed based on UPGMA method and which showed similitude in phylogram pattern. Also the sequences obtained for these species were compared with the sequences obtained for them in the other parts of the world for population studies.

Keywords: mtDNA; Calliphoridae; COI; phylogeny

1. Introduction
Molecular identification of species has now gained an awe inspiring status. DNA typing is an unbiased and a very reliable concept, as DNA is unique to every organism and they are the ultimate dowers of every phenotypic trait. Mitochondrial DNA is endowed with a transcendence to decipher phylogeny. Among all the genes known in mitochondrial DNA, the *COI* gene is considered to be at the acme. This gene has been very well studied at the biochemical level and its size and structure is conserved in all aerobic organisms [1]. Highly conserved and variable regions are closely associated in the *COI* gene, making this gene perfect for tracing phylogeny. In the present study, a short fragment of approximately 350bp of the *COI* gene has been chosen for the phylogenetic analysis. Forensic entomology is the science which deals with the application of the knowledge about insects in medico-legal investigations. A vast diversity of necrophagous insects and microbes as associated with the decomposing corpses, and a few serve as crucial circumstantial witness to the crime, thus by helping in estimating the post mortem interval (PMI). Among these insects, the blow flies (Diptera: Calliphoridae) are the most vital detectives, as they are the earliest and most abundant invaders on the corpse. The study of waves of succession of these insects and various aspects relating there developmental pattern have been well advocated [2-5]. The entomological evidences rely on the correct species identification which in turn is important for developmental studies and distribution. In this paper, the authors have attempted identification and phylogenetic analysis of 6 different species of blow flies following Singh and Sidhu, 2004 [6], namely Hemipyrellia ligurriens (Wiedemann, 1830), Hemipyrellia pulchra (Wiedemann, 1830), Lucilia ampullacea Villeneuve, 1922, Lucilia sericata (Meigen, 1826), Chrysomya nigripes Aubertin, 1932 and Chrysomya albiceps (Wiedemann, 1819). Forensic evidences are mostly fragmented and incomplete, thus there is a need for a short mtDNA marker which could resolve the identification problem accurately. All the species chosen are of forensic importance [7]. This study underpins the importance of such a short *COI* gene marker which is quite reliable for marking identity to the species. The Sarcophagid species, *Sarcophaga subvicina* was chosen as the out group.

Methodology

Collection of blow flies

The first step subsumed the collection of blow flies from different regions of North India (Himachal Pradesh, Punjab, Haryana and Uttarakhand), followed by sorting, pinning, labelling and identification of the collected specimens and finally preserving these identified samples in ethanol.

DNA extraction

The alcohol preserved samples were washed in distilled water to remove any foreign DNA or microbes. These samples were then dissected using sterilized surgical blade. Only legs were dissected out and used for DNA isolation. The tissue was crushed using liquid nitrogen. Then cell lysing (0.01M EDTA, 0.03 M Tris- HCl, 0.01 M sucrose) solution and proteinase K was added and incubated overnight at 55 °C in 1.5µL Eppendorf tubes [8]. These were then subjected to centrifugation at 13,000 rpm for 7 minutes. After this, the pellet containing the debris was discarded and the supernatant was carefully procured and used for further extraction. Equal proportions of isopropanol and chilled ethanol were added approximately double in volume to the supernatant. To this supernatant 1µl of glycogen solution (Qiagen) was added which is a vital ingredient for the precipitation of traces of DNA. These were then left in the deep freeze (-20 °C) overnight for DNA precipitation. Then these were centrifuged at 13,000 rpm for 10 minutes. The supernatant was discarded and pellet was washed with ethanol and air dried. Phenol: chloro form: iso-amyl alcohol (PCI) treatment was done for further purification. This purified pellet was then dissolved in tris-EDTA (TE) and stored at -20 °C.

Amplification

A specific region of *COI* gene of mitochondrial DNA was chosen as the target for amplification. Amplification was the most challenging part as optimum annealing temperature was required to be set for every sample. All polymerase chain reactions (PCR) were performed using Bio-rad T100 TM thermal cycler. The thermal cycler conditions were the following: initial denaturation at 98 °C for 2 minutes followed by 40 cycles at 98 °C for 30 seconds, annealing at 47.3 °C for 30 seconds, elongation at 75 °C for 30 seconds and final elongation at 75 °C for 10 minutes. 50µL PCR cocktail constituted of Phusion DNA polymerase enzyme 1U/50µL reaction, 5X Buffer 10µL, 10pm dNTP, 50Mm $MgCl_2$ 1µL, 10pm primers 1µL each and MQ water (Thermo Fisher Scientific, India). This fragment of *CO1* gene was amplified using primers C1-J-2495 (5′ CAGCTACTTTAT GAGCTTTAGG) [9]. And C1-N-2800 (5′ CATTTCAA GYTGTGTAAGCRTC) [10].

Electrophoresis

After the PCR was completed, the products were checked on 2% Agarose by Gel Electrophoresis and amplicon size was compared using reference gene ruler 100bp Ladder. 2% agarose gel spiked with Ethidium bromide at a final concentration of 0.5µg/mL was prepared. 5.0 µL of PCR product was mixed with 1 µL of Gel tracking dye. The PCR product of size 350bp was generated through this reaction.

Sequencing

Using the gene specific sequencing primers C1-J-2495(5′ CAGCTACTTTAT GAGCTTTAGG); C1-N-2800 (5′ CATTTCAA GYTGTGTAAGCRTC) and ABI BigDye® Terminator v3.1 Cycle Sequencing reaction kit (Applied Biosystems, USA), the purified PCR amplicons were sequenced. Sequencing was performed on 3130 Genetic analyser Automated DNA sequencing machine. The softwares used for sequence analysis were Sequencing Analysis 5.1; Chromas Pro v3.1. Sequences were submitted to Genbank (accession numbers in table1). Additional *CO1*sequenes for relevant Calliphorids were obtained from the genbank for population studies at www.ncbi.nlm.nih.gov (Table 1).

Phylogenetic analysis

The genotyping and phylogenetic methods used in this paper are widely and most commonly practiced. The evolutionary history was inferred using the UPGMA method [11]. The bootstrap consensus tree was inferred from 1000 replicates (Figure1.) The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches [12]. The evolutionary distances were computed using the number of differences method [13]. And are in the units of the number of base differences per sequence. The analysis involved 23 nucleotide sequences. There were a total of 15943 positions in the final dataset.

Table 1: Species marked* are the sequences submitted to the Genbank by the authors while those unmarked are sequences picked from Genbank.

Species	Country/Locality	Accession Number	Group
**Hemipyrellia ligurriens*	India: Kutani, Bhimtal, Uttarakhand	KP325711	4
**Hemipyrellia pulchra*	India: Dehradun, Uttarakhand	KP325709	5
**Lucilia ampullacea*	India: Afzalgarh, Uttarakhand	KP325712	1
**Lucilia sericata*	India: Nalagarh, Himachal Pradesh	KP325713	3
**Chrysomya nigripes*	India:	KP644240	2
**Chrysomya albiceps*	India: Bilaspur, Himachal Pradesh	KP325710	6
Chrysomya albiceps	Belgium	KF919011	6
Chrysomya albiceps	Australia	AB112839	6
Chrysomya nigripes	Malaysia	JN229028	2
Chrysomya nigripes	Malaysia	GU174026	2
Chrysomya nigripes	Malaysia	JX187367	2
Lucilia sericata	Australia	AB112843	3
Lucilia sericata	Denmark	EF531193	3
Lucilia sericata	China	DQ328671	3
Lucilia sericata	China	DQ328674	3
Hemipyrellia ligurriens	Australia	JX913759	4
Hemipyrellia ligurriens	The people`s Republic of China	DQ345092	4
Hemipyrellia pulchra	The people`s Republic of China	DQ345091	5
Lucilia ampullacea	South Korea	EU925394	
Lucilia ampullacea	USA	DQ453487	1
Lucilia ampullacea	Spain	KF225230	1
Lucilia ampullacea	Spain	KF225234	1
Sarcophaga subvicina	Belgium	JQ413461	Out group

This table enlists the individuals chosen for the analysis along with their accession number and country (Locality in case of Indian specimens).

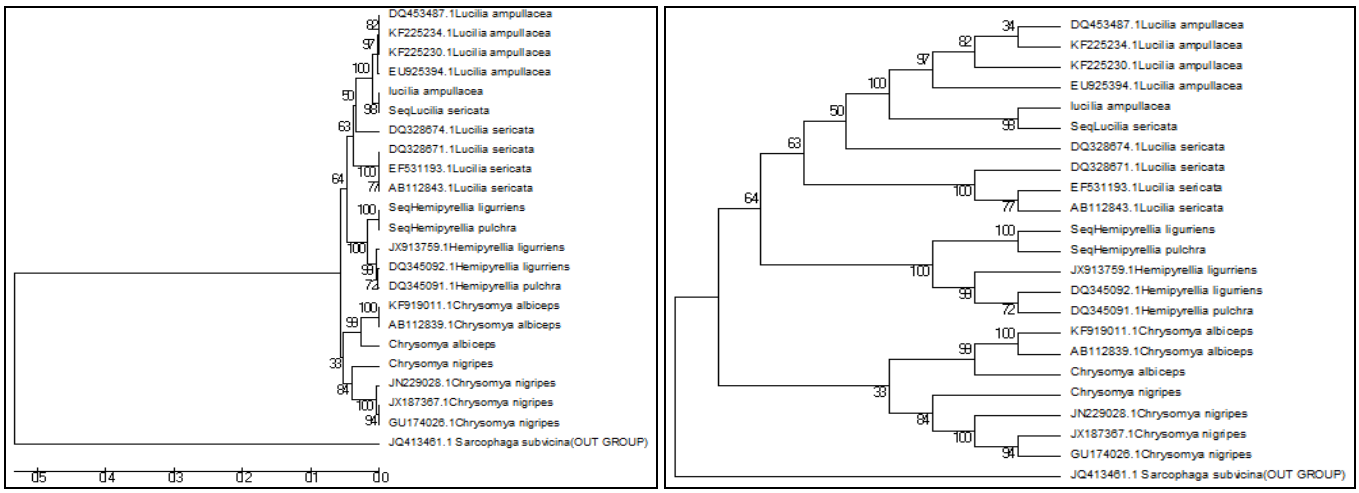

Fig 1: Evolutionary relationships among the taxa inferred using UPGMA method. Posterior probabilities are depicted on nodes.

Results and Discussion

Sequencing and alignment of a total of 6 species was carried over 350bp region and additional 16 sequences obtained from Genbank were compared. The sequences correspond to position 2482 to 2810 of *Drosophila yakuba* mitochondrial genome (accession number KF824901). Out of 214 variable locations 74 sites were considered parsimoniously informative.

Identification

The Sarcophagid out group, *Sarcophaga subvicina* got clearly separated in the phylogenetic tree. Calliphorid species were rightly allotted to the subfamilies Chrysomyinae and Luciliinae with each of the subfamilies monophyletic. The aim was to choose conspecific specimens from different geographical localities to evaluate the extent of variations in various species and to evaluate the conservation of *COI* gene over geographical ranges. Genus level arrangement precisely revealed the accession of the species with posterior probability values greater than 98% except *Chrysomya nigripes* and *Chrysomya albiceps*. Posterior probability (PP) reflects for a sister group relationship between the two genera *Lucilia* and *Hemipyrellia* under the subfamily Luciliinae seem perfect with PP value 64%. At species level, the overall PP support is above 92%, proving excellent (*Lucilia ampullacea* and *Lucilia sericata* 99%; *Hemipyrellia pulchra* and *Hemipyrellia ligurriens* 100%). Majority of the specimens were correctly allotted their respective species. The exceptions were *Hemipyrellia ligurriens* and *Hemipyrellia pulchra*, which were intermingled. *Chrysomya nigripes* and *Chrysomya albiceps* come up as clearly two different species belonging to subfamily Chrysomyinae. One clade of *Lucilia ampullacea* includes specimens from USA and Spain and the other clade consists of specimen from Korea. While Indian *L. ampullacea* specimen is represented as a separate clade showing closer similarity with Korean specimen than with those of USA and Spain which throws some light on the geographical distribution.

Intra specific variation

Average evolutionary divergence over sequence pairs within groups was estimated. The numbers of base substitutions per site from averaging over all sequence pairs within each group is shown in Table 2 with maximum value of 0.06 and were obtained by a bootstrap procedure (1000 replicates). Analyses were conducted using the Tamura-Nei model [14].

Table 2: The table represents maximum intraspecific variation.

Species	Average evolutionary divergence
Lucilia ampullaceal	0.01
Chrysomya nigripes	0.04
Lucilia sericata	0.06
Hemipyrellia ligurriens	0.03
Hemipyrellia pulchra	0.03
Chrysomya albiceps	0.04

Interspecific variation

Evolutionary divergence over sequence pairs between groups was estimated. The number of base substitutions per site so calculated between groups is shown in the Table 3 with maximum and minimum values lying between of 0.02 to 0.12. Analyses were conducted using the Tamura-Nei model. The two species *Lucilia ampullacea* and *Lucilia sericata* seem to be quite closely related showing only 0.06 divergence value among different clades; *Hemipyrellia ligurriens* and *Hemipyrellia pulchra* show negligible divergence of 0.02 among different clades; while *Chrysomya nigripes and Chrysomya albiceps* have been found to be relatively quiet diverged with divergence value of 0.12 among different clades. Inter specific variation among different species of Calliphoridae ranges from 0.1% to 13.5% (Figure 2).

Cryptic species and geographical variation

The different clades, representing world population of *Lucilia sericata* and *Lucilia ampullacea* are placed separately from those in India, which throws some light on the geographical patterns of distribution of populations as these also, differ phenotypically. Intra specific variation is observed to be quite low (highest being observed in *Lucilia sericata*) which eliminates possibility of any cryptic species among these studied species.

Table 3

	Lucilia ampullacea **Group 1**	***Chrysomya nigripes*** **Group 2**	***Lucilia sericata*** **Group 3**	***Hemipyrellia ligurriens*** **Group 4**	***Hemipyrellia pulchra*** **Group 5**	***Chrysomya albiceps*** **Group 6**
Lucilia ampullacea Group 1						
Chrysomya nigripes Group 2	0.12					
Lucilia sericata Group 3	0.06	0.11				
Hemipyrellia ligurriens Group 4	0.10	0.11	0.09			
Hemipyrellia pulchra Group 5	0.10	0.11	0.09	0.02		
Chrysomya albiceps Group 6	0.12	0.11	0.11	0.11	0.11	

Maximum likelihood estimate of gamma parameter for site rates

The estimated value of the shape parameter for the discrete Gamma Distribution was 1.3114. Substitution pattern and rates were estimated under the model (+G). A discrete Gamma distribution was used to model evolutionary rate differences among sites (5 categories, [+G]). Mean evolutionary rates in these categories were 0.16, 0.44, 0.77, 1.23, 2.40 substitutions per site. The nucleotide frequencies are A = 28.42%, T/U = 41.12%, C = 15.33%, and G = 15.13%. For estimating ML values, a tree topology was automatically computed. The maximum Log likelihood for this computation was -1324.142. The analysis involved 23 nucleotide sequences. In the final test there were 289 positions.

L_a																						
Ch_n	13.1																					
JX187367.1Ch_n	13.5	8.04																				
GU174026.1Ch_n	13.5	8.04	0.1																			
JN229028.1Ch_n	12.1	7.76	0.9	0.92																		
DQ328671.1L_s	7.24	11.1	10	9.97	9.9																	
EU925394.1L_a	2.74	12.8	13	13	13	7.6																
DQ453487.1L_a	2.74	12.4	13	12.7	12	7.6	0.6															
KF225234.1L_a	1.32	12.4	12	11.6	12	7.2	0.7	0														
KF225230.1L_a	1.32	12.4	12	11.6	12	7.2	0.7	0	0													
SeqL_s	0	13.1	13	13.5	12	7.2	2.7	2.7	1.3	1.3												
DQ328674.1L_s	6.38	13.1	12	11.7	12	7.8	7.4	7.4	7	7	6.38											
EF531193.1L_s	8.76	11.4	10	10.4	9.9	0.3	7.6	7.6	7.2	7.2	8.76	8.07										
AB112843.1L_s	8.76	11.4	10	10.4	9.9	0.3	7.6	7.6	7.2	7.2	8.76	8.07	0									
SeqH_l	10.2	11.1	11	10.8	10	9.6	9.4	9.4	8.7	8.7	10.2	8.56	11	11								
SeqH_p	9.54	11.7	10	10.5	11	10	9.9	9.9	8.9	8.9	9.54	9.27	10	10	0							
JX913759.1H_l	12	11.4	9.8	9.95	12	7.7	11	11	9.9	9.9	12	9.65	7.6	8	4	3.13						
DQ345092.1H_l	11.3	10.6	11	10.9	11	7.7	9.7	9.7	9.1	9.1	11.3	9.65	8	8.1	4	3.13	0.848					
DQ345091.1H_p	11.3	11.4	11	10.9	11	7	9.7	9.7	9.1	9.1	11.3	8.89	7.7	7.8	4	3.13	0.77	0.5				
Ch_a	11.6	10.2	12	12.2	12	9.5	12	12	11	11	11.6	12.6	9.8	9.8	10	10.7	10.43	10	10			
KF919011.1Ch_a	13.5	12.7	7.5	7.54	13	9.7	12	11	11	11	13.5	12.8	11	11	12	11.5	10.97	11	11	5		
AB112839.1Ch_a	13.5	12.7	8.3	8.49	13	9.7	12	11	11	11	13.5	12.8	12	11	12	11.5	12.18	12	12	5	0	
JQ413461.1_S_s	97.6	101	124	124	100	98	95	94	97	97	97.6	100	117	122	97	98.4	119.9	121	122	90	118	118

Fig 2

The figure represents interspecific distances using neighbour joining approach (Tamura Nei Model of substitution). Here L-_a is *Lucilia ampullacea*, L_s is *Lucilia sericata*, H_l is *Hemipyrellia ligurriens*, H_p is *Hemipyrellia pulchra*, Ch_a is *Chrysomya albiceps*, Ch_n is *Chrysomya nigripes* and S_s is *Sarcophaga subvicina.*

Conclusion

Our results indicate that identification of these Calliphorids using this 350bp small *COI* gene fragment is in good agreement with the phylogenetic trees generated by traditional approaches. The study shows that *COI* successfully distinguished 6 different species of Calliphoridae along with 16 conspecific specimens from different parts of the world. Also this short region has been proved to be useful for other Calliphorid species [15]. Sequence as short as 278bp [16]. And 304bp [17]. Have also proved fruitful. Species and genera were clearly separated and the subfamilial arrangement on the tree matched up with the one based on morphological characters. Also subfamily Chrysomyinae showed an ancestral position in the phylogenetic tree, being closely placed to the out group *Sarcophaga,* while the subfamily Luciliinae clearly showed an advanced position in the phylogenetic tree reflecting its recent origin. Also some geographical pattern is also observed in the distribution of populations, showing genetic dissimilarity in different clades. One interesting finding of this research was that DNA sequence as short as 350bp is sufficient enough for identification and tracking correct phylogeny, which is of great value in forensic entomology where most of the times only short fragmented remains of the insects are available. This paper also confirms the accuracy of using MEGA 5.2.2 in phylogenetic analysis. Also it is important to extensively survey various sequences (from a wide range of localities) and increasing the sampling size is must for the validation of this

research in the field of forensic entomology as most carrion flies are synanthropic. Though, this gene has appropriately identified all the species correctly and has rightly defined the paraphyletic groups (may be due to some species process) but it may be also because of very short sequence chosen for study. This ambiguity, however, can be resolved by further choice of this small gene fragment on large samples in future.

Acknowledgements
This study was funded by Innovation in Science Pursuit for Inspired Research (INSPIRE), an innovative programme sponsored and managed by the Department of Science & Technology, New Delhi.

Disclosure
Devinder Singh and Naina Khullar are not conversant of any memberships, financial holdings, affiliations that could raise a conflict of interest.

References
1. Gupta MP, Sharma K, Khan BU, Kapoor BC, Bajpai PK, Kumar A *et al*. Tissue specific esterase isozyme variation in *Clarias batrachus* and *C. gariepinus*. Global Journal of Pharmacology. 2009; 3(1):1-5.
2. Saraste M. Structural features of cytochrome oxidase. Quarterly Reviews of Biophysics 1990*;* 23:331-366.
3. Singh D, Greenberg B. Survival after the submergence in the pupae of five species of blow flies (Diptera: Calliphoridae). Journal of Medical Entomology. 1994; 31:757-759.
4. Singh D, Bharti M. Further observations on the nocturnal oviposition behaviour of blow flies (Diptera: Calliphoridae). Forensic Science International 2001; 120:124-126.
5. Bharti M, Singh D. Insect faunal successionon decaying rabbit carcasses in Punjab, India. Journal of Forensic Science. 2003; 48:1133-1143.
6. Singh D, Bala M. The Effect of Starvation on the larval behavior of two forensically important species of blow flies (Diptera: Calliphoridae). Forensic Science International 2009; 193:118-121.
7. Singh D, Sidhu IS. A check list of blow flies (Diptera: Calliphoridae) from North-west of India. Uttar Pradesh Journal of Zoology. 2004; 24(1):63-71.
8. Sukontason KL, Sribanditmongkol P, Chaiwong T, Vogtsberger RC, Piangjai S, Sukontason K. Morphology of immature stages of *Hemipyrellia ligurriens* (Wiedemann) (Diptera : Calliphoridae) for use in forensic entomology applications. Parasitology Research 2008*;* 103:877-887.
9. Kambhampati S, Rai KS. Mitrochondrial DNA variation within and among populations of the mosquito *Aedes albo-pictus*. Genome 1991; 34:288-292.
10. Sperling FAH, Anderson GS, Hickey DA. A DNA-based approach to the identification of insect species used for postmortem interval estimation. Journal of Forensic Science. 1994; 39:418-427.
11. Wells JD, Sperling FAH. DNA-based identification of forensically important Chrysomyinae (Diptera: Calliphoridae). Forensic Science International 2001; 120:109-114.
12. Sneath PHA and Sokal RR. *Numerical Taxonomy*. Freeman, San Francisco, 1973.
13. Felsenstein J. Confidence limits on phylogenies: An approach using the bootstrap. Evolution 1985; 39:783-791.
14. Nei M, Kumar S. Molecular Evolution and Phylogenetics. Oxford University Press, New York, 2000.
15. Tamura K, Nei M. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Molecular Biology and Evolution 1993; 10:512-526.
16. Wells JD, Sperling FAH. DNA-based identification of forensically important Chrysomyinae (Diptera: Calliphoridae). Forensic Science International 2001; 120:109-114.
17. Liu QL, Cai JF, Chang YF, Gu Y, Guo YD, Wang XH *et al*. Identification of forensically important blow fly species (Diptera: Calliphoridae) in China by mitochondrial cytochrome oxidase I gene differentiation. Insect Science 2011; 18:554-564.
18. Desmyter S, Gosselin M. COI sequence variability between Chrysomyinae of forensic interest. Forensic Science International: Genetics 2008; 3:89-95.