Mitochondrial DNA-Based Identification of Developmental Stages and Empty Puparia of Forensically Important Flies (Diptera) in Egypt.
Submit Your Paper Today!

Sign up for e-Table of Contents
Please register here to receive an 'e-Table of Contents' directly to your inbox for every issue of Journal of Forensic Science and Medicine as soon as they are published: www.jfsmonline.com/login.asp

Editor Contact Details
Email: editor@jfsmonline.com; lingl@ocmemd.org

For further details please visit: www.jfsmonline.com

Sponsored by: The "111 Plan" of China – Evidence Science Innovation and Talent Base

The JFSM iOS App for iPad is here!
Mitochondrial DNA-Based Identification of Developmental Stages and Empty Puparia of Forensically Important Flies (Diptera) in Egypt

Hedayat A. Abdel Ghaffar1, Marie Z. Moftah2, Alexandre Favereaux2,3, Mohamed Swidan1, Omar Shalaby1, Sara El Ramah1,2, Reda Gamar4

1Department of Zoology, Faculty of Science, Alexandria University, Alexandria, Egypt, 2Department of Biology and Geology, Faculty of Education, Alexandria University, Alexandria, Egypt, 3Interdisciplinary Institute of Neuroscience, UMR 5297, CNRS, Bordeaux, France, 4Department of Zoology, Faculty of Science, Alexandria University, Alexandria, Egypt

Abstract

Relying on morphology in the estimation of postmortem intervals is not always accurate for all life phases of flies, especially for immature stages. The present study demonstrates the application of cytochrome oxidase I sequences, a partial mitochondrial (mt) gene region, to differentiate forensically important flies in Alexandria, Egypt. Thirty-three adult flies, larvae, and pupae were collected from rabbit carcasses. Nineteen were used for genotyping. Sequence analysis revealed no significant intraspecific divergence in Diptera species. Accordingly, a neighbor-joining tree using the Kimura 2-parameter model illustrated reciprocal morphology between species. Specimens represented five species, four genera, four subfamilies, two families, and one order. We herein identify five different Diptera species, Chrysomya albiceps, Chrysomya megacephala, Calliphora vicina, Lucilia sericata, and Ophyra capensis, using mt DNA as a species-specific marker for identification in a local database set-up.

Keywords: Cytochrome oxidase, Diptera, mitochondrial DNA, phylogenetic analysis

Introduction

Forensic entomology is a powerful tool that can aid in estimating the minimum postmortem interval (PMI) of human remains. However, homicide victims may not be discovered for days, weeks, or months.1-4 The challenging step for forensic entomologists is larval species identification because many morphological similarities among closely related species make it difficult to distinguish them with precision.5 Published studies of forensically relevant flies mainly focused on Calliphoridae (Wells and Sperling, 2001).6-9 as they are generally the first arthropods to locate and oviposit onto a corpse. Their common remnants in crime investigations are pupae and puparial cases.7 These can be useful for PMI estimation.6 Thus, crime investigations can benefit from molecular genotyping methods. DNA analysis can provide independent morphology verification and may even discover the presence of cryptic species when this would not otherwise be possible.9

Mitochondrial DNA (mt DNA) has a higher mutation rate than nuclear DNA, which increases the chance of generating species-specific markers.3 It also possesses both conserved and variable segments. In addition, mt DNA may be easier to isolate than nuclear DNA.10 A barcoding approach based on mt DNA may be useful for the identification of taxa, for which the use of morphology or the association of different life stages is problematic.11 Therefore, the present study involved mt DNA analysis of different developmental stages of forensically important flies, including pupae and empty puparia left behind after adult emergence. Among all the genes known in mt DNA, the cytochrome oxidase I (COI) gene is considered the best to use. This gene is very well studied, and its size and structure are conserved in all aerobic organisms.12 It is often used in species identification.13 Highly conserved and variable regions are closely associated in the COI gene, making it perfect for phylogeny tracing.14

Address for correspondence: Prof. Marie Z. Moftah, Department of Zoology, Faculty of Science, Alexandria University, Alexandria, Egypt. E-mail: marie.moftah@alexu.edu.eg

The goal of this study was to demonstrate that genetic identification of flies can be reliable, even from fragments or incomplete empty puparia, which is not achievable with morphological examination.

**MATERIALS AND METHODS**

**Samples**

Fly samples were collected from rabbit carcasses. The rabbits had been killed by cervical dislocation to avoid pain, in accordance with the Alexandria University Ethics Committee and NIH guidelines. Samples were divided into three groups based on morphology. Experiments were performed in the Faculty of Science, Alexandria University, Egypt. Group “A” contained adult flies, larvae, and pupae that we placed in killing jars. After immobilization, insects were dissected to isolate flight muscles that were transferred into Hood’s solution (95 ml of 80% ethanol and 5 ml of glycerol). Some nondissected insects were placed into dry vials for direct pinning and further morphological identification. Adult flies, larvae, pupae, and empty puparia of group “B” were preserved immediately at −80°C for subsequent molecular analysis. Immature flies belonging to group “C” were kept alive and reared to the adult stage for species confirmation.

**Morphological identification**

We examined adult flies using a model U-TV1X-2 stereomicroscope (Olympus, Japan). Accurate adult identification was based on the work of Hinton,[15] Smith,[16] Mound,[17] Greenberg and Kunich,[18] and Castner.[19]

**DNA extraction**

DNA was extracted from insects of different developmental stages using two extraction protocols. First, the conventional phenol/chloroform/isoamyl alcohol method was partially modified by adding 20 mg/ml of proteinase K and increasing the digestion time to 48 h to improve the yield. Second, the E. Z. N. A.™ Insect DNA Kit (Omega Bio-Tech, Norcross, GA, USA) was used according to the manufacturer’s instructions. A total of 5–15 μl DNA with an absorbance ratio (A260/A280) of 1.7–1.9 was obtained. The extracted DNA was stored for 2 weeks at −20°C then re-suspended in 20 μl of RNase-free water. The DNA extracted fraction was spectrophotometrically quantified and diluted to a concentration of 100 ng/μl before polymerase chain reaction (PCR) amplification. The diluted stock was stored at 4°C to avoid DNA damage.

**Polymerase chain reaction amplification of the COI gene**

Two attempts to amplify the mt gene were performed. First, we tried to amplify the whole region of COI[20] using primers designed using Perprimer free software (Perprimer v1.1.17, Owen and Marshall copyright © 2003–2008).

Second, we chose a region of COI according to previously published primer sequences on Calliphoridae[21,22] (Simon et al., 1994)[23] [Table 1]. PCR was done in a final volume of 25 μl containing 1× PCR buffer, 2.5 mM MgCl₂, 0.1 μM of each primer, 0.2 mM dNTPs, 1.25 units of Taq polymerase, and 1 μg DNA template. Amplification was performed using an Eppendorf Mastercycler (AG 22331; Eppendorf, Hamburg, Germany). The thermal sequence consisted of an initial step of 2 min at 95°C followed by 35 cycles of 30 s at 94°C, 30 s at 54°C, and 1 min at 72°C. This was followed by 10 min at 72°C and the final product was stored at 4°C. PCR products were visualized on a 1.2% agarose gel (1× TBE with ethidium bromide using a GeneRuler™ 100 bp DNA ladder). The most intensely stained products were selected for sequencing.

**Sequencing**

Products were labeled using Big Dye® Terminator v.1.1 Sequencing Kit (Thermo Fisher Scientific, Waltham, MA, USA), and sequenced using the ABI 3130x1 automated sequencer (Thermo Fisher Scientific).

**Data and phylogenetic analysis**

Nucleotide sequences were prepared and analyzed using Finch TV software (Version 1.4.0, copyright© 2004–2006, Geospiza Inc. Seattle, Washington, USA), BioEdite Sequence Alignment Editor (copyright© 1997–2007, Tom Hall, 1999),[24] and Molecular Evolutionary Genetics Analysis (MEGA, version 4.0).[25] DNA sequence reads were edited manually using BioEdite and were aligned using Clustal W in MEGA software. The phylogenetic tree was generated by the neighbor-joining (NJ) technique using the Kimura 2-parameter (K2P) model of substitution.[26] Ophyra capensis was the out-group in all analyses. Maximum likelihood analysis was performed using the K2P model to compare the genetic distance between species using MEGA.

**Results**

COI data were obtained for almost all species, resulting in a nucleotide matrix that consisted of 1500 bp (http://www.ncbi.nlm.nih.gov/) and 300 bp.[3] The sequence analysis was based on 300 bp. Only the smaller fragments (SFs) were amplified and sequenced, using published primers.[3] When the amplified DNA was present, some amplifications of 1500-bp fragments were repeated. It was still difficult to obtain a sequence indicating that the region of interest had been amplified. Failures were associated with either a low amount of PCR products (indicated by a faint band in the ethidium bromide gel) or interference from nonspecific amplified products, resulting in a high background. Moreover, the nucleotide peaks were overlapping and were not clearly recognizable to be considered.

Genotyping sequences were successful in only 19 of 33 specimens when SF-COI was analyzed, likely due to poor

| Table 1: Primers of polymerase chain reaction and sequencing of cytochrome oxidase I |
|-----------------|-----------------|-------|
| Primer          | Primer sequences | Length |
| Forward         | C1-J-2495 (5'-CAG CTA CTT TAG G-3') | 22 |
| Reverse         | C1-N-2800 (5'-CAT TTC AAG CTG TGT AAG CAT C-3') | 22 |


DNA quality. Well-defined nucleotide peaks were obtained, and electrophoresis of PCR products was performed, yielding an average length of 300 bp. The nucleotide composition showed notably higher frequencies of thymine (T) and adenine (A) (41% and 32%, respectively) when compared to cytosine (C) and guanine (G) (13% and 14%, respectively). The selected sequence of each sample, based on bands’ sharpness and clarity, was chosen to undergo further investigations.

Morphological identification of fly species matched the results based on genotyping. Accordingly, the specimens represented five species, four genera, four subfamilies, two families, and one order [Table 2].

**GenBank matching**

For all tested samples of flies \( n = 19 \), exact matching was aligned with the GenBank database (http://www.ncbi.nlm.nih.gov/BLAST/) with an average of 97.68% maximum identity. Newly determined nucleotide sequences, based on the sequence analysis and BLAST search of partial CO1 sequences, led to classify the flies into five species: *Chrysomya albiceps* \( (n = 7) \), *Chrysomya megacephala* \( (n = 2) \), *Lucilia sericata* \( (n = 5) \), *Calliphora vicina* \( (n = 1) \), and *O. capensis* \( (n = 4) \) [Table 2].

**Divergence distance**

Pair-wise sequence differences observed for individual loci of CO1 are shown in Table 3. Within-group variation for all genera ranged from 0.0% to 0.6% with a mean of 0.24% for combined datasets [Table 3]. *Calliphoridae* between-group variation ranged from 8.6% to 16.4% with a mean of 11.4%. Variation between the *Muscidae* outgroup and the *Calliphoridae* genera ranged from 11.6% to 13.4% with a mean of 12.82%. The rate of increase declined in the higher taxonomic categories.

We then analyzed the genetic distance observed as measured by the percentage divergence between sequences. The most notable finding was the very small divergence (8.2%) separating *Ch. megacephala* and *L. sericata*. This was lesser than the amount of interspecific divergence (9%) separating *Ch. megacephala* and *Ch. albiceps*, which is considered the lowest interspecific variation.

Levels of interspecific variation between *Calliphoridae* species varied from 8.2% to 16.8%. Species pairs such as *Ch. albiceps*/ *Ch. megacephala* were separated by 9%. At a higher level, the subfamilies Chrysomyinae and Calliphoridae displayed 13%–16.8% variation while Chrysomyinae and Lucilinae displayed 8.2%–10.3% variation. Subfamilies *Calliphoridae* and *Lucilinae* displayed a difference of only 10.9% [Table 4].

**Bootstrapping and neighbor-joining tree**

The tree produced using the distance analysis method yielded highly similar topologies. Values were provided on the same K2P-NJ tree [Figure 1]. The subfamily status was well supported, with *Lucilina*, Chrysomyinae, and *Muscidae* displaying 100% support in all analyses. Chrysomyinae species was grouped together with 58% support and *C. vicina* was represented by 50% bootstrap. In almost all analyses, individuals of each species received 100% bootstrap support, indicating a strong basis for species distinction. For instance, in the phylogenetic tree, Chrysomyinae species were determined as the monophyletic group having strong bootstrap support. A major discrepancy between results was observed in the grouping of *O. capensis* individuals having strong bootstrap support. A major discrepancy between results was observed in the grouping of *O. capensis* individuals (different developmental stages), within which the subbranching between adults, larvae, and pupae was supported by a 52% bootstrapping value.

Intraspecific values were relatively low. Maximum values observed between individuals of each species are displayed in Table 4. The highest level was observed in *Ch. megacephala* and *O. capensis* at 0.6% in individuals collected at different developmental stages. Very little intraspecific variations were observed in each of the *Ch. albiceps* (at 0.3%) and *L. sericata* (at 0.0%) individuals collected at different developmental stages.

**Discussion**

The morphological similarity, especially between immature stages, poses a great challenge for forensic entomologists. The identification of immature Diptera by their DNA sequence has the advantage of speed and simplicity. Furthermore, species can be identified from partial remains lacking morphological character. Contrary to the long DNA fragments that may minimize stochastic variation across taxa and are more likely to reflect broader patterns of nucleotide divergence, shorter fragments have many advantages, such as being rapid, easy, and economical to analyze.
To the best of our knowledge, no large-scale analyses of the COI sequences of Diptera in Egypt have been done. In the present study, the use of molecular techniques provided a DNA extraction method capable of extracting amplifiable mt DNA from immature stages, adult thoracic muscles, and empty puparia.

Sequencing of approximately 300 bp of the mt COI region appeared to show five species-specific differences. Sequences did not vary in the same species between immature stages and adults. Hence, they were useful in the identification of all life stages.

The data obtained would allow an entomologist to distinguish between specimens of Ch. albiceps, Ch. megacephala, L. sericata, C. vicina, and O. capensis in cases where morphological differentiation is difficult. This agrees with the study of Sperling et al.,[21] who demonstrated that mt DNA COI sequences (along with COII and tRNA leu) could be used to identify the immature larvae of other forensically important Diptera (such as Lucilia illustris, Phormia regina, and Phaenicia sericata).

The present study showed interspecific variation of 9% between Ch. megacephala and C. albiceps, which is higher than the 3% variation reported by Wells and Sperling (2001).[4] The maximum intraspecific variability was 0.6%, showing a distinct separation between individuals of different populations of a species, and a separate species altogether. This supports the finding of Wells and Sperling (2001) of an interspecific divergence of ≤3% and intraspecific divergence of ≤1% in a number of forensically significant Calliphoridae species. Hence, calculations of pairwise differences between individuals should provide a useful indication of the validity of a grouping in an analysis.

The maximum level of intraspecific variability was observed for Ch. megacephala (0.6%), which is not overly elevated compared to the variability of the worldwide Ch. megacephala...
specimens reported in previous studies.\textsuperscript{[6,29]} Moreover, a higher variability value is observed for a congeneric species with forensic importance. Wells and Sperling reported <1\% intraspecific level of variation within the COI and COII sequence for Chryssoninae flies. Likewise, Ottanto et al.\textsuperscript{[30]} used COI to distinguish between Hypoderma larvae and reported a rate of intraspecific variation ranging from 1.14\% to 1.59\%.

The phylogenetic analyses herein support the subfamily separation at the species level. Each species formed distinct conspecific and monophyletic clusters. In the phylogenetic tree, the monophyletic separation of C. albiceps, C. megacephala, L. sericata, and O. capensis supported by strong bootstrapping values confirmed the sufficient resolution of the genetic marker. This indicated the robust nature of the region that can be used in distinguishing between species. It also reflected apomorphic character states for each group and the level of nucleotide divergence between groups (overall differences).

This is consistent with the findings of Yadong et al. (2010) and Liu et al.\textsuperscript{[31]}

The forensic utility of any method also depends on a number of factors. The most important are likely accuracy, speed, and cost. The technical accuracy of DNA sequence analysis is known. The appropriate choice of specific primers can readily circumvent common criticisms concerning problems of bacterial and other forms of DNA contamination. DNA extraction, PCR amplification of target fragment(s), and sequencing are routinely completed in <48 h, particularly when short and definitive fragments can be identified and sequenced in only one run.\textsuperscript{[12]} Similarly, the cost for double-strand sequences, such as a short fragment, is low. Hence, this study suggests the short well-conserved mt DNA markers may be reliable, rapidly analyzed, and inexpensive molecular markers used to identify forensically important flies.

From the Egyptian perspective, the present study provides a strong foundation for molecular-based phylogenetic analysis of forensically important Diptera species. This is important especially in the case of insect fragments and poorly preserved DNA samples.

Egypt’s generic taxonomy clearly requires further investigations of intraspecific variations using different markers combinations to consider other forensically important species, such as Sarcophagidae. Some entomologists recommend the combined analysis of different mt DNA fragments for Diptera species identification.\textsuperscript{[13]} The combined analysis of COI and COII fragments is a more accurate approach for Diptera species identification than single mt DNA fragment analysis.\textsuperscript{[14,32]}

Finally, technical replicates (parallel PCR and sequencing from the same sample) in barcoding experimental designs need to be performed. Data reproducibility should be determined empirically, as it will depend on the sequencing depth, type of samples, sequence analysis pipeline, and number of replicates.\textsuperscript{[36]}

The present study is a step toward the development of molecular tools used in forensic investigations in Egypt. COI sequence data appear to provide a strong basis for species identification and a valuable investigative tool in forensic entomology.

**Financial support and sponsorship**

STDF project number 513 and FP7 project number 245807 “Neuromed.”

**Conflicts of interest**

There are no conflicts of interest.

**References**

Abdel Ghaffar, et al.: Mt DNA-based identification of Egyptian Diptera