The availability of mitochondrial DNA cytochrome oxidase I gene for the distinction of forensically important flies in China

CAI Ji-Feng 1 , LIU Min 1 , YING Bin-Wu 1 , DENG Reng-Li 2 , DONG Jian-Guo 3 , ZHANG Lin 1 , TAO Tao 1 , PAN Hong-Fu 1 , YAN Hong-Tao 1 , LIAO Zhi-Gang $^{1\,*}$

- (1. West China School of Preclinical and Forensic Medicine, Sichuan University, Chengdu 610041, China;
 - 2. West China School of Clinical Medicine, Sichuan University, Chengdu 610041, China;
 - 3. School of Preclinical Medicine, Shanghai Second Medical University, Shanghai 200025, China)

Abstract: Sarcosaphagous flies are frequently found on dead bodies shortly after death. Species within this family differ in their developmental times , so an accurate identification of every species is necessary for the correct estimation of the post mortem interval (PMI). Identification of specimens is traditionally performed using their morphological features. The morphological similarity of sarcosaphagous flies especially their eggs , larva and pupae , poses a challenge for forensic entomologists. Therefore a molecular method was established for species identification. In this study , A 278 base pair region of mitochondrial DNA (mtDNA) coding for cytochrome oxidase subunit one (COI) was investigated for identification of the following forensically important species of sarcosaphagous flies from western China , including *Ophyra capensis* (Wiedemann), *Chrysomya megacephala* (Fabricius), *Lucilia sericata* (Meigen), *Lucilia cuprina* (Wiedemann), and *Boettcherisca peregrina* (Robineau-Desvoidy). The results indicated that the COI region sequenced allowed identification of major species , providing separation of congeneric species with high support. In contrast , the data could not distinguish taxa from the same species group , *i.e.* the *L. sericata* and *L. cuprina* groups because of low sequence divergence.

Key words: Sarcosaphagous flies; species identification; mitochondrial DNA; cytochrome oxidase I

At the end of the 19th century, the French entomologist Megnin and Benecke used sarcosaphagous flies that had developed on dead horses lying in a field to gather forensic information (Amendt $et\ al\$., 1998). Today, forensic entomology, including the study of ecology and development of sarcosaphagous flies colonizing a corpse, is commonly used in investigations of death (Leclercq, 1978). In practice, the accuracy of this method is about 1 day for a 1-month-old corpse and 1 week for a 6-month-old corpse (Gaudry $et\ al\$., 2001).

The first step , however , is to identify accurately sarcosaphagous flies , which is conventionally done by using flies' morphological features. The morphological similarity poses a great challenge for forensic entomologists , especially in identifying larvae. The minuscule morphological differences between species of immature larvae are sometimes extremely hard to determine and requires an expert with specialized taxonomic knowledge in the field to assure correct identification (Catts and Goff , 1992). Although identification keys are available , only a few experts are able to identify the larvae of sarcosaphagous flies to species level. Up to now , all eggs and pupa and the

first to third instar larvae of sarcosaphagous flies could not be identified by morphologyical method (Wallman and Donnellan , 2001). Under these circumstances , species identification based on genetic examination is an option.

Such morphological challenges make a molecular approach to the identification of these species desirable. Recent molecular studies have demonstrated the applicability of mtDNA sequencing to the identification of important Old and New World species of carrionbreeding blowflies (Catts and Goff , 1992). When using mtDNA sequences for species identification, one must consider the possibility of a single sequence may not be representative of the entire species. However, previous studies of mtDNA restriction site variation within a wide variety of insect species indicate that the extent of variation within species is almost always small compared to differences between species (Sperling, 1993). For this study of divergence between species, people deliberately chose sequences that are known to show relatively low levels of variation within populations (Simon, 1991). The recent study indicated that analysis of the partial sequences focused on sections of the cytochrome oxidase b subunit one and two (COI

收稿日期 Received: 2004-07-21;接受日期 Accepted: 2004-12-21

^{*} 通讯作者 Author for correspondence , Fax: +86-28-85501544; E-mail: zhigangliao_1@hotmail.com

and CO $\rm II$) encoding region of mtDNA revealed abundant phylogenetically informative nucleotide substitutions that could identify sarcosaphagous flies at species to species group level. Our studies involved sequencing to examine the potential of DNA for use in the identification of blowfly in forensic investigation. The sequence variation within the CO $\rm II$ gene of the mitochondrial genome of several species was selected for examination. Analyses were based on a 278 base pair region of the gene coding for cytochrome oxidase subunit one ($\rm CO~II$).

1 MATERIALS AND METHODS

1.1 Insects

Four adults of Chrysomya megacephala (Fabricius), 2 adults of *Lucilia sericata* (Meigen) and their eggs, 2 adults of Boettcherisca peregrina (Robineau Desvoidy) and its 1 larva and 1 pupa, 4 adults of Lucilia cuprina (Wiedemann), 2 adults of Ophyra capensis (Widemann) were obtained from Huhhot district , SA ($39^{\circ}35' \, \mathrm{S}$, $110^{\circ}46' \, \mathrm{E}$) , western China. In addition, specimens from Chengdu and Dunhuang district of western China were included. These were Chrysomya megacephala (4 adults), Lucilia sericata (2 adults), Boettcherisca peregrina (2 adults), Musca domestica vicina (2 adults) were obtained from Chengdu district, SA (30°05'S, 102°54' E) and Boettcherisca peregrina (1 adult) from Dunhuang district, SA (37°56′S,92°07′E).

1.2 DNA extraction

DNA was extracted using Chelex method (Walsh and Petzger, 1991). DNA was quantified through Primate-specific alpha-satellite probe assay (Waye *et al.*, 1989).

1.3 PCR primers

A portion of the CO I gene was amplified and sequenced using primers C1-J-2495 (5'-CAG CTA CTT TAT GAG CTT TAG G-3') (sense) and C1-N-2800 (5'-CAT TTC AAG CTG TGT AAG CAT C-3') (antisense) (Sperling *et al.*, 1994; Wells and Sperling, 1999).

1.4 PCR conditions

The PCR reaction volume was 37.5 μ L, containing 20 – 40 ng DNA, 6 μ L dNTP (1 mmol/mL),1.5 U Taq polymerase, 3.75 μ L 10 × buffer (Mg²⁺ 1.5 mmol/L),0.3 μ L primers (50 nmol/mL).

PCR amplifications were performed in a thermocycler (Perkin-Elmer 9600), with initiative denaturing for 1 min at $94\,^\circ\!\!\mathrm{C}$, followed by 38 cycles of $94\,^\circ\!\!\mathrm{C}$ for 1 min , $48\,^\circ\!\!\mathrm{C}$ for 1 min and $72\,^\circ\!\!\mathrm{C}$ for 2 min .

1.5 Sequncing

After purification of the PCR products with QIAquick columns cycle sequencing was performed on both forward and reverse strands using ABI PRISM Big Dye Terminator Cycle Sequencing Ready Reaction Kit (Applied Biosystems). Removal of excess dyedeoxyterminators primers and buffer was accomplished by DYE-EX spincolumns (Qiagen).

1.6 Phylogenetic analysis

Phylogenetic analysis is the study of the relationships between different groups of organisms, and is thus based on similarity and difference in characters chosen (Hillis and Moritz, 1990). As the sequences were protein coding and did not contain any insertions or deletions, alignments were carried out using the multiple-alignment program DNAMAN, version 4.0 analyses (Lynnon BioSoft, Copyright 1994-98). We tested whether the sequences were of mitochondrial origin or represented paralogous sequences resident in the nucleus (Hillis et al., 1996) in two ways using MEGA (Kumar et al., 1993). Sequences were translated with mitochondrial genetic code and , since nuclear paralogues can lose their coding function, the predicted amino acid sequences were inspected for inappropriate stop codons. We also compared the base composition of the individual sequences since nuclear paralogues can have divergent base compositions relative to mitochondrial genes. A neighbour-joining tree using the Tamura and Nei (Tamura and Mei, 1993) model of nucleotide substitution was constructed using the MEGA package (Kumar et al., 1993). Maximum parsimony analysis was also conducted to calculate the tree based on the least number of evolutionary steps. Five hundred bootstrap replications were performed. Bootstrap values are calculated by repeated random sampling of the data to provide an indication of the confidence limited for a particular group.

2 RESULTS

2.1 Morphological identification of specimens

Eggs obtained from abdomen of adults of *Lucilia sericata* were definitely known for their species identity. But one pupa and larva (first instar) from a rabbit body (in the experiment) couldn't be identified by morphological method (Hu and Min, 2000). All adults of specimens were identified easily by morphological method except the two most closely related species, *i*. e. *Lucilia sericata* and *Lucilia cuprina*.

2.2 CO I sequence divergence

Six species were sequenced. The sequences have been deposited in GenBank with accession number AY818091 - 818125.

2.3 CO I sequence analysis

The neighbor-joining (NJ) tree based on the CO I sequence data using Kimura's two-parameter model was constructed (Fig. 1). Invariant positions in the sequence were removed and the remaining variant positions numbered relative to the original sequence to

indicate the number of substitutions supporting the divergence of the taxa. The neighbour-joining tree was identical to the tree obtained using maximum parsimony method. Three distinct congeneric clusters were formed based on the sequence data. High bootstrap values support the three nodes. Bootstrap values provide an

indication of the percentage support for a grouping by randomly resampling the data.

2.4 Phylogenetic analysis

A total of 278 aligned sites for the 29 CO I sequences were included in the analyses (Fig. 1).

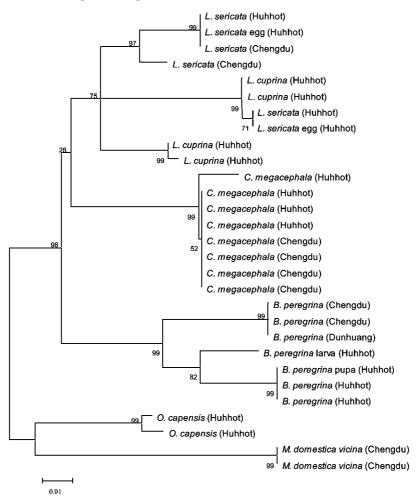


Fig. 1 The neighbor-joining (NJ) tree based on the $CO\ I$ sequence data using Kimura's two-parameter model. The bar indicates 0.01 substitutions per site.

All specimens were grouped into 4 clusters with high bootstrap support. The first and second clusters were formed by 2 groups respectively, and the third and fourth by 1 group respectively. At the clusters level, there was 98% support between the first and the second clusters. At the group level within the first cluster, there was 28% support between the first and second groups. At the species level, the first group of the first cluster was formed by 3 subgroups of specimens of Lucilia sericata and Lucilia cuprina. Within the first subgroup, the first 3 specimens of Lucilia sericata shared 99% support, whereas the fourth specimen deviated a little as showed as 97%. Within the second subgroup containing 2 species, i.e. Lucilia sericata and Lucilia cuprina, the support between species is

surprisingly greater than the support within species as shown as 71% vs 99%. Within the third subgroup there was 99% bootstrap support between the two specimens of *Lucilia cuprina* from Huhhot. The bootstrap support among all three sub-groups was 75%. In the second group of the first cluster , there were 8 specimens of *Chrysomya megacephala* with approximately 52% support among them.

In the second cluster , there were 2 groups , with 99% support between them. The support within the first group was as high as 99% whereas the support within the second group of adult specimens was 99% , and the support between adult specimens as a whole and that of larva was 82% .

In the third and fourth clusters, the support was

99% respectively at the species level.

Table 1 shows that the level of CO I gene nucleotide divergence between species groups within sarcosaphagous flies taxa. A total of 278 aligned sites for the 29 CO I sequences were included in the phylogenetic analyses by Kimura 2-parameter distance method. Uncorrected percentage sequence divergence among all taxa ranged from 0 to 17.8%. The overall average was 9.9%, and the average level of divergence within the *Chrysomya megacephala*, *Lucilia sericata*, *Lucilia cuprina*, *Boettcherisca peregrina*, *Ophyra capensis*, *Musca domestica vicina* group was 0.29%,

2.8%, 2.86%, 3.59%, 0.37%, and 0 respectively. While the maximum and minimum level of divergence between the *Chrysomya megacephala* and *Lucilia sericata* was 12.6% and 6.4% respectively (average 8.9%); that of *Lucilia sericata* and *Lucilia cuprina* was 8.5% and 0.4% respectively (average 5.1%); that of *Chrysomya megacephala* and *Boettcherisca peregrina* was 13.4% and 11.3% respectively (average 12.7%); that of *Chrysomya megacephala* and *Musca domestica vicina* was 17.0% and 15.2% respectively (average 15.4%).

Table 1 Percent uncorrected sequence divergence between sarcosaphagous flies taxa for the CO I gene^a

```
28
                                                          29
[1]
[2] 1.50
[3] 1.50 0.00
[4] 1.50 0.00 0.00
[5] 1.50 0.00 0.00 0.00
[6] 1.50 0.00 0.00 0.00 0.00
[7] 1.50 0.00 0.00 0.00 0.00 0.00
[8] 1.50 0.00 0.00 0.00 0.00 0.00 0.00
[9] 9.20 8.00 8.00 8.00 8.00 8.00 8.00 8.00
[10]7.60 6.40 6.40 6.40 6.40 6.40 6.40 6.40 3.00
[11] 12.6 11.3 11.3 11.3 11.3 11.3 11.3 11.3 8.90 6.10
[12] 9.20 8.00 8.00 8.00 8.00 8.00 8.00 8.00 0.00 3.00 8.90
[ 13 ] 12.6 11.3 11.3 11.3 11.3 11.3 11.3 11.3 8.90 6.10 0.00 8.90
[14] 9.20 8.00 8.00 8.00 8.00 8.00 8.00 8.00 0.00 3.00 8.90 0.00 8.90
[\ 15\ ]\ 8.00\ 6.80\ 6.80\ 6.80\ 6.80\ 6.80\ 6.80\ 5.30\ 3.70\ 7.30\ 5.30\ 7.30\ 5.30
[ 16 ] 12.2 10.9 10.9 10.9 10.9 10.9 10.9 10.9 8.50 5.70 0.40 8.50 0.40 8.50 6.80
[\ 17\ ]\ 7.60\ 7.02\ 7.02\ 7.20\ 7.20\ 7.20\ 7.20\ 5.70\ 4.10\ 7.70\ 5.70\ 7.70\ 5.70\ 0.40\ 7.30
[18] 12.2 10.9 10.9 10.9 10.9 10.9 10.9 10.9 8.50 5.70 0.40 8.50 0.40 8.50 6.80 0.00 7.30
```

^aFor species that diverged from all others by a relatively large percentage and where within-species variation was small, the minimum sequence divergence percentages are given that distinguish any individual within that species from any other species. Species/specimen codes: 1 - 4: Chrysomya megacephala (Huhhot); 5 - 8: Chrysomya megacephala (Chengdu); 9,10: Lucilia sericata (Chengdu); 11,12: Lucilia sericata (Huhhot); 13,14: Lucilia sericata egg(Huhhot); 15 - 18: Lucilia cuprina (Huhhot); 19: Boettcherisca peregrina larva (Huhhot); 20: Boettcherisca peregrina pupa (Huhhot); 21: Boettcherisca peregrina (Dunhuang); 22,23: Boettcherisca peregrina (Huhhot); 24,25: Boettcherisca peregrina larva (Chengdu); 26,27: Ophyra capensis (Huhhot); 28,29: Musca domestica vicina (Chengdu).

3 DISCUSSION

In this study, the DNA extraction and sequencing of all specimens of six species of sarcosaphagous flies were finished. Both specimens from Huhhot that were stored in 75% alcohol for 3 – 6 months and from Chengdu that were stored by air-dried condition for over 2 years can be extracted and sequenced.

The results indicate that the morphological method is as effective as mtDNA sequence method in species identification of adult sarcosaphagous flies and can solve

the difficulty in differentiating the closely related species shows *Lucilia sericata* and *Lucilia cuprina*. However, the morphological method requires expertise in specialized taxonomy to assure correct identification (Leclercq, 1978), whereas the technology of mtDNA sequence method turns to be mature and the concerned facilities are easily accessible even in ordinary laboratories.

According to the neighbor-joining (NJ) tree based on the CO I sequence data by using Kimura's two-parameter model (Fig. 1), Calliphoridae, Sarcophagidae and Muscidae of Diptera clustered in

divergent branches. The high support for the congeneric grouping of species illustrates the potential of the CO I in interspecific distinction. The ability to clearly distinguish between these forensically prominent genera based on such a small region provides a strong indication of economy in choosing the region of the CO I. The difficulty we encountered in identifying the specimens of Lucilia sericata and Lucilia cuprina by mtDNA sequence method is corresponding to that reported by Wallman and Donnellan (Wallman et al., 2001), who encountered the difficulty in distinction between C. augur and C. dubia based on CO Isequence data in their study, and also the report of Stevens and Wall (Stevens and Wall, 1996), who examined the mitochondrial encoded 12S rRNA, CO I and CO I sequences of the two blowflies Lucilia cuprina and L. sericata originated from Hawaii. The two endemic Chinese species in our study were closely related to each other and further sequencing is desired to separate them. The intraspecific variation within Lucilia sericata and Lucilia cuprina should deserve further consideration.

The results of our study about CO I phylogenetic systematic data are in agreement with Wells and Sperling's (Wells and Sperling , 1999; Wells et al., 2001), who demonstrated, by examining CO I sequences, that the blowflies Chrysomya rufifacies and Chrysomya albiceps exhibited less than 1% intraspecific difference and about 3% interspecific one. However, the results of our study are in disagreement with their results concerning Lucilia sericata, Lucilia cuprina, Boettcherisca peregrina groups. The difference may be due to geographical separation of populations, which needs further study.

Though some studies reported the differentiation of samples from different fields, our study showed no differentiation between specimens from Dunhuang and Huhhot and lent little support to the differentiation between specimens from Chengdu and Huhhot in western China. So in the future, we should sample more specimens from more species and from different geographical regions throughout the world.

This report shows that , with future development , mitochondrial nucleotide sequences can provide a means of identifying forensically important flies. Sequencing this region of mtDNA has provided data that appear to indicate the potential of mtDNA in development of an identification protocol for use in forensic entomological studies. Future works include the examination of more variable mitochondrial genes , if necessary , and the analysis of additional species of flies from other parts of China .

Biotechnology Co. Ltd for mtDNA primers. We also thank Prof. DENG An-Xiao Deng (Parasite Department, School of Preclinical and Forensic Medicine, Sichuan University) for identifications of flies and Prof. LI Ying-Bi for valuable assistance with technical aspects of the study. We are also grateful for technical assistance and facilities provided by Forensic Physical Evidence Department of our school.

References

- Amendt J , Krettek R , Niess C , Zehner R , Bratzke H , 1998. Forensic entomology from southern Italy. *J . Forensic Sci .* , 43(1):210 214. Catts EP , Goff ML , 1992. Forensic entomology in criminal investigations.
- Annu . Rev . Entomol . , 37 : 253 272 .
- Gaudry E, Myskowiak JB, Chauver B, Pasquerault T, Lefebvre F, Malgom Y, 2001. Activity of the forensic entomology department of the French Gendarmerie. Forensic Sci. Int., 120(1-2):68-71.
- Hillis DM, Moritz C, 1990. Molecular Systematics. USA: Sinauer Associates.
- Hu C, Min JX, 2000. Forensic Entomology. Chongqing: Chongqing Publishing House. 66 – 160.
- Kumar S , Tamura K , Nei M , 1993. MEGA: Molecular Evolutionary Genetics Analysis. Version 1.0. The Pennsylvania State University , University Park , Pennsylvania.
- Landry BS, Dextraze L, Boivin G, 1993. Random amplified polymorphic DNA fingerprinting and genetic variability assessment of minute parasitic wasp species (Hymenoptera: Mymaridae and Trichogrammatidae) used in biological control program of phytophagous insects. Genome, 36:580 – 587.
- Leclercq M , 1978. Entomologie et Médecine Légale. Masson , Paris. 100.
 Simon C , 1991. Molecular systematics at the species boundary: exploiting conserved and variable regions of the mitochondrial genome of animals via direct sequencing from amplified DNA. In: Hewitt GM , Johnston AWB , Young JPW eds. Molecular Techniques in Taxonomy. Berlin: Springer Verlag. 33 72.
- Sperling FAH, 1993. Mitochondrial DNA variation and Haldane's rule in the *Papilio glaucus* and *P. troilus* species groups. *Heredity*, 70:227 233.
- Sperling FAH , Anderson GS , Hickey DA , 1994. A DNA-based approach to the identification of insect species used for postmortem interval estimation. *J. Forensic Sci.* , 39(2):418-427.
- Stevens JR, Wall R, 1996. Species, sub-species and hybrid populations of the blowflies *Lucilia cuprina* and *Lucilia sericata* (Diptera: Calliphoridae). *Proc. R. Soc. Lond B*, 263:1335-1341.
- Tamura K , Nei M ,1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol* . *Biol* . *Evolut* . ,10:512-526.
- Wallman JF, 2001. Third-instar larvae of common carrion-breeding blowflies of the genus *Calliphora* in South Australia. *Invertebr*. *Taxon*, 15:37 51.
- Wallman JF , Donnellan SC , 2001. The utility of mitochondrial DNA sequences for the identification of forensically important blowflies (Diptera: Calliphoridae) in southeastern Australia. Forensic Sci. Int. , 120:60-67.
- Walsh BS, Petzger DA, 1991. R. Chelex-100 as medium for simple extraction of DNA for PCR-based typing from forensic material. Biotechniques, 10(4):506-510.
- Waye JS , Presley LA , Budowle B , Shutler GG , Fourney RM , 1989. A simple and sensitive method for quantifying human genomic DNA in forensic specimen extracts. *Biotechniques* , 7:852 855.
- Wells JD , Pape T , Sperling FAH , 2001. DNA-based identification and iolecular sydtematics of forensiclly important Sarcophagidae (Diptera).
 J. Forensic Sci., 46:1098-1102.
- Wells JD, Sperling FAH, 1999. Molecular phylogeny of Chrysomya albiceps and C. rufifacies. J. Med. Entomol., 36 (3):222 – 226.
- Zhang DX , Hewitt GM , 1996. Nuclear integrations: challenges for mitochondrial DNA markers. $Tr.\ Ecol.\ Evo.$, 11:247-251.

(责任编辑:袁德成)

mtDNA 中 CO I 分子标记在常见食尸性蝇类 鉴定中的应用

蔡继峰¹,刘 敏¹,应斌武¹,邓仁丽²,董建国³,张 林¹, 陶 涛¹,潘洪富¹,闫红涛¹,廖志钢¹

(1. 四川大学华西基础与法医学院 成都 610041; 2. 四川大学华西临床医学院 成都 610041; 3. 上海第二医科大学基础医学院 上海 200025)

摘要:死后不同时间,在尸体上出现不同种类食尸性蝇类的演替规律,可用于准确推断死亡时间。传统上仅依据蝇类形态学特征来判断种属,但由于蝇类的形态结构复杂和种间形态差异微小等特点,对蝇类尤其是对蝇类幼虫的种属鉴别很难。因此应用分子生物学方法对食尸性蝇类及其幼虫进行种属鉴定非常重要。本研究主要是利用此方法对我国西部部分地区常见双翅目食尸性蝇类包括,开普黑蝇、大头金蝇、丝光绿蝇及部分卵,铜绿蝇、棕尾别麻蝇及部分幼虫和蛹的线粒体 DNA(mtDNA)上细胞色素氧化酶辅酶 I(COI)中278 bp的基因序列进行鉴别。除个别蝇类如丝光绿蝇与铜绿蝇外,该方法均能有效地将上述食尸性蝇类鉴定到种属水平。在我国,它将成为法医鉴别食尸性蝇类种属的可靠依据。

关键词:食尸性蝇类;物种鉴定;线粒体 DNA;细胞色素氧化酶辅酶 I

中图分类号: Q969 文献标识码: A 文章编号: 0454-629((2005)03-0380-06