Felix A. H. Sperling,[1] *Ph.D.; Gail S. Anderson,*[2] *Ph.D.; and Donal A. Hickey,*[3] *Ph.D.*

A DNA-Based Approach to the Identification of Insect Species Used for Postmorten Interval Estimation

REFERENCE: Sperling, F. A. H., Anderson, G. S., and Hickey, D. A., **"A DNA-Based Approach to the Identification of Insect Species Used for Postmortem Interval Estimation,"** *Journal of Forensic Sciences*, JFSA, Vol. 39, No. 2, March 1994, pp. 418–427

ABSTRACT: Insect larvae found on a corpse can be used for estimating postmortem intervals. Here, we describe a molecular method for rapid identification of these insects. Specific insect DNA fragments were amplified using the polymerase chain reaction (PCR), followed by direct DNA sequencing of the amplification products. We sequenced 2300 base pairs of mitochondrial DNA from each of three blowfly species: *Phormia regina*, *Phaenicia sericata* and *Lucilia illustris*. All three species are important in forensic entomology. We found 118 nucleotide differences between the *L. illustris* and *P. sericata* sequences, 186 between *L. illustris* and *P. regina*, and 192 between *P. sericata* and *P. regina*. Based on these abundant DNA sequence differences, we can unambiguously identify the immature larval stages of these insects. These DNA sequence differences were also used to predict species-specific, diagnostic restriction sites in the amplified DNA, and these predictions were verified by digestion with nine restriction enzymes. The DNA sequences reported here encode the mitochondrial COI, COII and tRNA-leucine genes.

KEYWORDS: pathology and biology, postmortem interval, Calliphoridae, blowflies, mitochondrial DNA, cytochrome oxidase genes, polymerase chain reaction, restriction digests

Blowflies lay their eggs at predictable times in the decay cycle of a corpse [*1–3*]. The particular assemblage of insect species found on the remains, together with estimates of ambient temperature, can allow more precise estimates of time of death than are possible by any other means. One drawback in the use of insects as indicators, however, is that species differences are subtle in the immature stages [*4*] and therefore the larvae must often be reared to the adult stage before they can be reliably identified to species [*2,3*]. This means a delay of several days, or even weeks, before the time of death can be estimated. An even more serious problem is the fact that insect larvae may also die before identification is possible. Yet for criminal cases, early focusing of investigations can be crucial for their successful conclusion. Consequently, there is a need for a reliable means of identifying these insects immediately after collection at the scene of a crime.

Received for publication 6 May 1993; revised manuscript received 19 July 1993; accepted for publication 20 July 1993.

[1]Postdoctoral Fellow, Department of Biology, University of Ottawa, Ottawa, Ontario, Canada.
[2]Assistant Professor, Department of Biological Sciences, Simon Fraser University, Burnaby, British Columbia, Canada.
[3]Professor, Department of Biology, University of Ottawa, Ottawa, Ontario, Canada.

A variety of biochemical techniques have been used to identify insect species. Allozymes are widely used as diagnostic markers for species and races, but these markers are often expressed only in particular life stages and are easily degraded in material that is not live or deep-frozen. Immunological techniques such as ELISA have the disadvantage of being relatively insensitive to differences between closely related species. In contrast, DNA-based identification of organisms is possible using any life stage, and can greatly reduce the time necessary to obtain identifications. Mitochondrial DNA (mtDNA) is particularly well suited as a marker for such identifications, since it is a small molecule that is relatively resistant to degradation; hundreds of copies are generally present in each cell; and its mutation rate is high enough to provide numerous sequence differences between closely related species [*5*].

In this paper, we present DNA sequences for a 2.3 kilobase region of mtDNA in three blowfly species (Calliphoridae: *Phormia regina, Phaenicia sericata, Lucilia illustris*). These three species have different ecological associations and all three commonly serve as forensic indicators [*1*]. The molecular methods described here allow identification of immature larvae within one day of receipt of specimens. Moreover, identifications are possible even with dead insect material.

Materials and Methods

Samples

Live flies and larvae were all collected in the vicinity of Vancouver, British Columbia or taken from a lab colony at Simon Fraser University. The thorax of each fly was used for DNA extraction, and the abdomen, head, wings and legs were saved as vouchers. For large larvae, the middle third of the body was used for DNA extraction and the posterior and anterior ends were used as vouchers. Identifications of adult voucher material were provided by B. Cooper, Biosystematics Research Division, Agriculture Canada.

Adult flies were frozen at −70°C until used for DNA extraction. DNA was also extracted from three additional batches of specimens: 1) Third instar larvae that were placed live into 99% enthanol, kept at room temperature for less than a week, then kept at 4°C until used; 2) First and second instar larvae that were kept at room temperature in 75% ethanol for one month, then transferred to 99% ethanol and treated like the large larvae above; and 3) Adult flies that were dried for five days at room temperature.

DNA Extraction

Total genomic DNA was extracted from thoraces of individual adult flies by a modification of the technique of Harrison et al. [*6*]. Thoraces were ground into powder using disposable plastic pestles inside 1.5 mL microfuge tubes immersed in liquid nitrogen. Lifton buffer (800 μL of: 0.1 M Tris buffer, 0.2 M sucrose, 0.05 M ethylaminediaminetetraacetate (EDTA), 0.5% sodium dodecyl sulfate, pH 9.0) was added to the powder to lyse membranes and inhibit DNAses. The homogenate was lightly vortexed and incubated at room temperature for 15 min to 2 h. Then 120 μL of 8 M potassium acetate was added, the mixture was inverted to mix it, and the tubes were placed on ice for 15 min to 3 h. The resulting precipitate was spun down for 15 min in a microfuge, and the supernatant was decanted to a new tube. The supernatant was extracted with one volume of phenol, followed by extraction with one volume of chloroform/isoamyl alcohol (24:1). DNA was precipitated by addition of one volume of isopropanol, then cooled to −20°C and centrifuged for 25 min. The pellet was washed in 500 μL of 70% ethanol and resuspended in 200 μL of TE (10 mM TrisHCl, 1 mM EDTA, pH 8.0).

Ethanol-preserved and dried samples were processed as above, except: 1) Ethanol-preserved samples were dried in a vacuum to remove ethanol before grinding in liquid nitrogen; 2) Proteinase K (12 μg) was added to the homogenate immediately after addition of Lifton buffer, and the homogenate was incubated at 60°C for 2 h; 3) one additional phenol extraction was performed to ensure that the proteinase K was removed; and 4) the final DNA pellet was resuspended in 100 μL of TE buffer.

Amplification and Sequencing

DNA fragments were amplified using the polymerase chain reaction (PCR) as per the instructions contained in the Perkin Elmer Cetus GeneAmp© PCR Reagent Kit, using 1 μL of DNA in TE, in a total volume of 50 μL overlain by one drop of light mineral oil. Taq polymerase was added after an initial incubation at 95°C for 3 min in a Perkin Elmer Cetus DNA thermal cycler, during an annealing phase of 45°C for 2 min, and before an extension phase of 72°C for 1.5 min. This was followed by 30 to 35 cycles of 94°C for 1 min, 45°C for 1 min, and 72°C for 1.5 min and terminated with one cycle that was the same as the previous ones except for an extension of 5 min duration. An extension time of 2.5 min was used to amplify fragments of 2.4 kilobases length (2303–2306 base pairs plus primers).

Genomic DNA was subjected to PCR amplification [*7*] using general insect mtDNA primers developed by comparisons among sequences of *Drosophilia* [*8*], *Apis* [*9*] and *Anopheles* [*10*]. Other primers were designed directly from calliphorid fly sequences and positioned approximately every 300 base pairs. Double-stranded PCR product was cleaned with Centricon© 100 microconcentrators (Amicon, Inc.), and sequenced directly using the Taq DyeDeoxy™ Terminator Cycle Sequencing system (Applied Biosystems, Inc.). In all cases, sequence was confirmed from both sense and antisense strands.

Prediction of Restriction Sites

The fly mtDNA sequences were searched for matches to the recognition sites of 129 different restriction enzymes, using the MicroGenie™ sequence analysis program (Beckman Instruments, Inc.). Nine enzymes were chosen for further work, based on: 1) the presence of sites in most sequences, to give an internal control for enzyme effectiveness; 2) differences between species in the locations of restriction sites; 3) robustness of the enzyme under varying temperatures and buffer conditions; and 4) low or moderate cost of enzymes.

Digestions with Restriction Endonucleases

About 1 μg of the DNA fragment resulting from PCR amplification (usually 1 to 4 μL of the DNA suspension) was digested in a total volume of 15 μL according to the recommendations of suppliers of restriction endonucleases (generally New England Biolabs, Inc.). The DNA was electrophoresed in 1.5 to 2.0% agarose gels to separate fragments by size, then visualized with ethidium bromide [*11*].

Results

The mtDNA region sequenced in this study includes the cytochrome oxidase b subunit I and II genes (COI and COII) and the tRNA leucine (UUR) gene. One individual per species was sequenced over this region. The sequences extend from the tRNA tyrosine gene to the tRNA lysine gene, and correspond to bases 1468 to 3771 in the published *Drosophlia yakuba* mitochrondrial sequence [*8*]. A comparison of these DNA sequences

for the three blowfly species is shown in Appendix 1. The sequences have been deposited in GenBank under accession numbers L14945-7 [*12*].

We found that all three DNA sequences were distinctly different from one another (Table 1 and Appendix 1). The two most closely related species, *Lucilia illustris* and *Phaenicia sericata*, showed a sequence divergence of 5.12% from each other; this represents more than one hundred nucleotide substitutions between the two sequences. Both of these species showed sequence divergences of greater than 8% from *Phormia regina*. All three blowfly sequences showed divergences of greater than 12% from the *Drosophila yakuba* sequence [*8*]. These levels of DNA sequence divergence are consistent with the accepted phylogenetic relationships between these species [*13*]. Further comparisons of mtDNA evolution in these flies will be presented later as part of a larger survey.

As expected, nucleotide substitutions among the blowflies were not uniformly distributed throughout the length of the DNA, that is, some regions showed a higher degree of sequence divergence than others. For instance, there were many substitutions scattered throughout most of the COI and COII genes, but the tRNA leucine gene (1567–1632 in Appendix 1) and the beginning of the COII gene (1641–1708 in Appendix 1) showed no differences among the blowfly species. In contrast to this, the short intervening sequence between the tRNA leucine and COII genes was highly variable, showing three nucleotide insertions or deletions.

Using these sequences, we predicted the location of short diagnostic DNA sites that are recognized and cleaved by restriction endonucleases. Computer-generated restriction maps were produced from each of the three DNA sequences. Predicted restriction sites for nine restriction endonucleases are mapped in Fig. 1. Subsequently, the presence of each of these sites was successfully tested by enzyme digestion of PCR-amplified fragments. Differences among the species were verified by: 1) amplification of mtDNA using PCR; 2) digestion of PCR-generated fragments with restriction endonucleases that cleave the DNA at different locations in different species; and 3) separation of DNA fragments by size in agarose gels to allow visualization of diagnostic patterns of stained DNA fragments. The results are shown in Fig. 2 (and see the following).

In addition to verifying the predicted restriction patterns, we also tested the facility with which DNA could be amplified from specimens with partially degraded DNA. This is important in evaluating the potential of the technique for application in non-ideal, real situations. We found that amplification of the entire 2.4 kilobase fragment was routinely possible using undegraded template DNA obtained from adult flies frozen at −70°C. Slightly degraded samples, such as large larvae preserved in 99% ethanol for six months, could generally be used for the amplification of fragments up to 1400 base pairs length; they were not suitable, however, for the amplification of longer fragments. We also tested DNA that was substantially more degraded. This was extracted from two dried flies and two small larvae that were initially preserved in 75% ethanol. Even in these cases, DNA fragments up to 350 bp in length could be amplified reliably. This means that molecular data can be collected from insect evidence of very variable quality.

TABLE 1—*Percent DNA sequence divergence (above diagonal) and number of nucleotide substitutions (below diagonal) across a 2.3 kilobase region of mtDNA.* Drosophila yakuba *is included for outgroup comparison.*

Fly Species	*P. sericata*	*L. illustris*	*P. regina*	*D. yakuba*
P. sericata	...	5.12	8.34	12.32
L. illustris	118	...	8.07	12.74
P. regina	192	186	...	12.97
D. yakuba	284	294	299	...

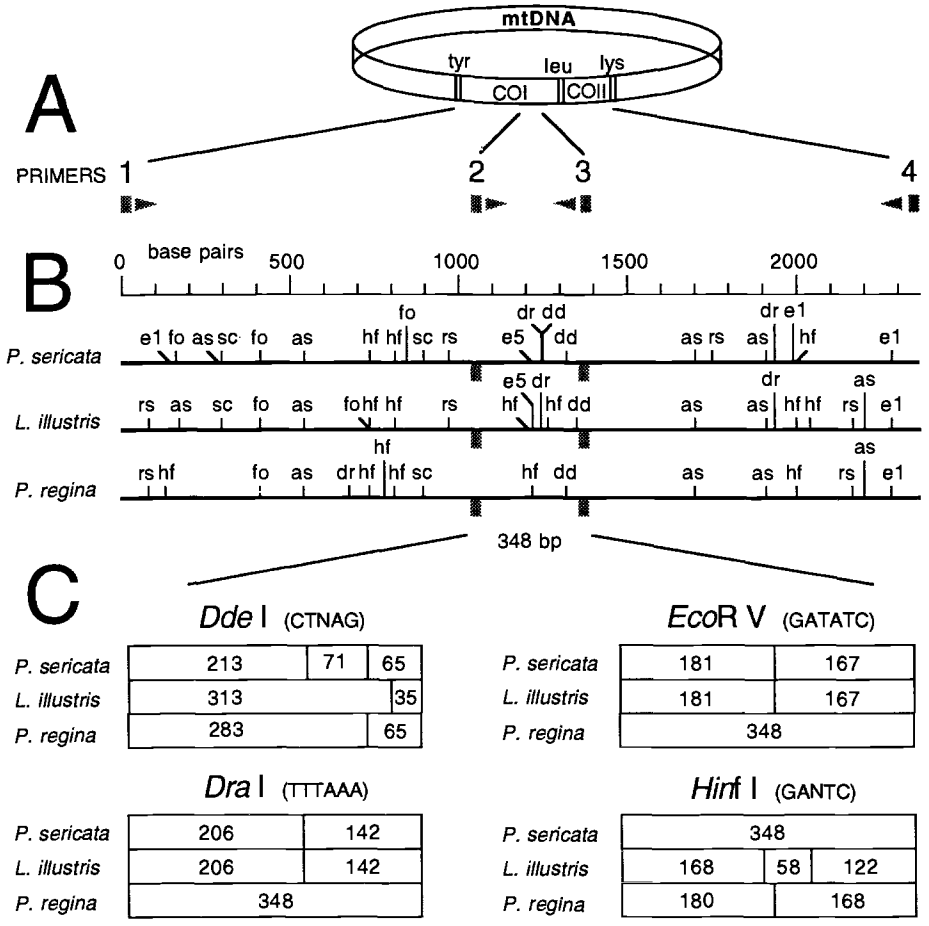

FIG. 1A—*Schematic showing region of mtDNA amplified and location of primers. Primer 1 = 5' TACAATTTATCGCCTAAACTTCAGCC 3' (bp 1-26 in Appendix 1); 2 = 5' CAGCTACTTTATGAGCTTTAGG 3' (bp 1033 - 1054); 3 = 5' CATTTCAAGC/TTGTGTAAGCATC 3' (bp 1359 - 1480); 4 = 5' GAGACCATTACTTGCTTTCAGTCATCT 3' (bp 2333 - 2359).* B. *Locations of restriction sites for nine enzymes across entire sequences: as = Ase I; dd = Dde I; dr = Dra I; e1 = EcoR I; e5 = EcoR V; fo = Fok I; hf = Hinf I; rs = Rsa I; sc = Sac I.* C. *Fragments produced by digestion of a 348 base pair region of COI amplified using primers 2 and 3.*

We reasoned that, for the application of this technique in real forensic situations, it would be most efficient to amplify a relatively small DNA fragment from a particularly variable segment of the sequenced region. Therefore, based on the sequences shown in Appendix 1, we designed oligonuleotide primers 2 and 3 to amplify across a small but informative region of 348 bp, which provides at least 5 restriction site differences between each pair of species (Figs. 1 and 2). All of our samples were readily amplified across this region. Four diagnostic enzymes (Dde I, Dra I, EcoR V and Hinf I) were useful for distinguishing *P. regina* from the two other blowflies, whereas Dde I and Hinf I distinguished *P. sericata* from *L. illustris*. In fact, Hinf I provided readily-distinguishable restriction fragment patterns for each species, and was also the least expensive of the four enzymes. The ability to infer the diagnostic sequence differences by restriction enzyme digestion therefore means that it is not necessary to repeat the complete DNA sequence for each specimen.

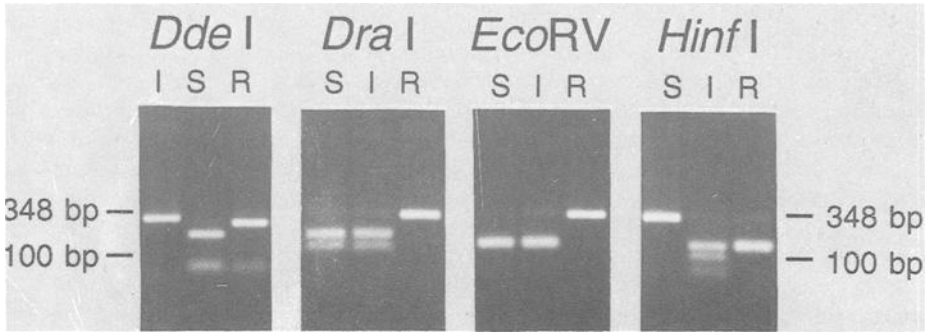

FIG. 2—*Agarose gels (2%) showing diagnostic fragment patterns produced by digestion of 348 base pair PCR product: i* = Lucillia illustris; *s* = Phaenicia sericata; *r* = Phormia regina.

Discussion

Species Identifications

We have shown that the divergence in DNA sequence between these calliphorid insect species is sufficient for clear species identification of their larvae. The results described here for three species can easily be extended to include other species that are likely to be found as evidence at any particular location. Fortunately, only a few dozen species of higher flies are likely to be relevant to forensic investigations in a single region, and there is considerable overlap in the faunal assemblages that occur across North America. We suggest that one or a few DNA sequences can serve as a reference for each species, and surveys within species can be performed using a battery of restriction sites that can be mapped on the reference sequence.

In addition to their utility in species identification, these sequences also provide basic information on the biological relationships between these species. *Phaenicia* and *Lucilia* are considered closely related genera by North American taxonomists [*13*], and both are generally subsumed under the name *Lucilia* by European researchers [*14*]. In contrast, *Phormia* is considered to belong to a separate tribe. The levels of DNA sequence divergence between the blowfly species used in this study clearly reflect these taxonomic groupings (Table 1).

Identification of Poorly Preserved Material

At present, a major inconvenience in forensic entomology is the fact that insect larvae have to be collected and maintained in a living state for identification. This requires some knowledge of insect biology, and this may not always be available when the evidence is being collected. As stated, a big advantage of the molecular approach to identification is that it is relatively insensitive to the state of preservation of the sample. Certainly, it is not necessary to maintain the larvae in a living state. A situation where this may be especially valuable is where the insect evidence was collected in the past as part of a general sample of the evidence. These old samples can now be identified based on their amplified DNA sequences. The method can also be further improved by optimizing DNA preservation and PCR amplification protocols (for example, Ref *15* for blackfly DNA preservation).

Variation Within Populations

When using DNA sequences (or any other character) for species identification, one must consider the possibility of variation within the species. In other words, 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 (reviewed in Ref *17*, see also Ref *18*) indicate that the extent of variation within species is almost always small compared to differences between species (for example, Ref *18*, Sperling and Hickey unpublished). For this study of divergence between species, we deliberately chose sequences that are known to show relatively low levels of variation within populations [*19*]. For a study of variation within the species, we would choose a region of the mitochondrial DNA that is known to be more polymorphic, such as the region containing the origin of replication. In blowflies, this region is approximately 2.5 kb in length [*20*].

There have been no direct surveys of the extent of DNA sequence variation within *P. sericata*, *L. illustris* and *P. regina*. In fact, this paper constitutes the first report of DNA sequences from these species. Another calliphorid, the New World Screwworm (*Cochliomya hominovorax*) shows some mtDNA variation across its range, and this variation has been used to determine that a new infestation in Libya probably originated from South America or the Caribbean islands [*16*]. We plan to look at population variation within forensically-important species in a future study. In a pilot study at our Vancouver locality, we found no restriction site differences among individuals within *P. sericata* (n = 5), *P. regina* (n = 5) or *L. illustris* (n = 2). This fits with the observation that there is generally much less DNA sequence divergence within species than between even very closely related sibling species. Nonetheless, until population data is established for DNA-based identification, we would advise forensic entomologists to use this method as a complement to, rather than a substitute for, standard taxonomic methods to identify species.

Advantages of the Molecular Approach

Like DNA fingerprint data for humans, the use of sequence and restriction site data to characterize insects allows convenient storage and manipulation of computerized information, as well as quantitative operations such as statistical estimates of the probability of correct identification.

In addition, it will be informative to study DNA sequences in order to understand mitochondrial gene evolution, since mitochondrial genes are crucial to oxidative respiration. The metabolic rates and development times of maggots are one of the most important aspects of their usefulness as forensic indicators, and it is conceivable that information on genetic relatedness of mtDNA lineages will prove useful in extrapolating development times for fly populations that have not been studied in the lab.

DNA sequencing and PCR technology are becoming widely used in medicine, and this technology may be easier to learn and apply by forensic investigators than is insect identification using more traditional morphological characters. Although a substantial investment is required to obtain initial sequences for a species, identification of species is relatively simple once the sequence is known. By converting taxonomic knowledge into a DNA-based format, we can make it easier for nonentomologists to use insects in forensic investigations.

Acknowledgments

We are grateful to V. Walker for suggestions leading to this study, and the lab of R. Harrison for mtDNA primer 4. We also thank B. Cooper (Biosystematics Research Di-

vision, Agriculture Canada) for identifications of flies and voucher material. Jo Leibovitz provided excellent technical assistance. Research funding was provided by an NSERC Canada Operating Grant to D.A.H. and by a grant from the NCE Program to Insect Biotech Canada.

References

[1] Greenberg, B., "Flies As Forensic Indicators," *Annals of the Entomological Society for America*, Vol. 28, 1991, pp. 565–577.

[2] Turner, B. D., "Forensic Entomology," *Forensic Science Progress*, Vol. 5, 1991, pp. 129–151.

[3] Catts, E. P. and Goff, M. L., "Forensic Entomology in Criminal Investigations," *Annual Review of Entomology*, Vol. 37, 1992, pp. 253–272.

[4] Liu, D. and Greenberg, B., "Immature Stages of Some Flies of Forensic Importance," *Annals of the Entomological Society of America*, Vol. 82, 1989, pp. 80–93.

[5] Avise, J. C., "Ten Unorthodox Perspectives on Evolution Prompted by Comparative Population Genetic Findings on Mitochondrial DNA," *Annual Review of Genetics*, Vol. 25, 1991, pp. 45–69.

[6] Harrison, R. G., Rand, D. M., and Wheeler, W. C., "Mitochondrial DNA Variation in Field Crickets Across a Narrow Hybrid Zone," *Molecular Biology and Evolution*, Vol. 4, 1987, pp. 144–158.

[7] Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B., and Ehrlich, H. A., "Primer-Directed Enzymatic Amplification of DNA with a Thermostable DNA Polymerase," *Science*, Vol. 239, 1988, pp. 487–491.

[8] Clary, D. O. and Wolstenholme, D. R., "The Mitochondrial DNA Molecule of *Drosophila yakuba*: Nucleotide Sequence, Gene Organization, and Genetic Code," *Journal of Molecular Evolution*, Vol. 22, 1985, pp. 252–271.

[9] Crozier, R. H. and Crozier, Y. C., "The Mitochondrial Genome of the Honeybee *Apis mellifera*: Complete Sequence and Genome Organization," *Genetics*, Vol. 133, 1993, pp. 97–117.

[10] Mitchell, S. E., Cockburn, A. F., and Seawright, J. A., "Complete Sequence of the Mitochondrial Genome of *Anopheles quadrimaculatus*," GenBank, Los Alamos National Laboratory, Group T-10, Mail Stop K710, Los Alamos, NM 87545.

[11] Sambrook, J., Fritsch, E. F., and Maniatis, T., "*Molecular Cloning, Second Edition*," Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1989.

[12] *GenBank*, Los Alamos National Laboratory, Group T-10, Mail Stop K710, Los Alamos, NM 87545.

[13] Shewell, G. E., "Calliphoridae," *Manual of Nearctic Diptera*, Vol. 2, J. F. McAlpine, Ed., Agriculture Canada Monograph No. 28, 1987, pp. 1133–1145.

[14] Rognes, K., "Blowflies (Diptera, Calliphoridae) of Fennoscandia and Denmark," *Fauna Entomologica Scandinavica*, Vol. 24, 1991, pp. 1–272.

[15] Post, R. J., Flook, P. K., and Millest, A. L., "Methods for the Preservation of Insects for DNA Studies," *Biochemical Systematics and Ecology*, Vol. 21, 1993, pp. 85–92.

[16] Taylor, D. B., Hammack, L., and Roehrdanz, R. L., Reproductive Compatibility and Mitochondrial DNA Restriction Site Analysis of New World Screwworm, *Cochliomya hominovorax*, from North Africa and Central America," *Medical and Veterinary Entomology*, Vol. 5, 1991, pp. 145–151.

[17] Vogler, A. P., DeSalle, R., Assman, T., Knisley, G. B., and Schultz, T. D., "Molecular Population Genetics of the Endangered Tiger Beetle, *Cicindela dorsalis* (Coleoptera: Cicindelidae)," *Annals of the Entomological Society of America*, Vol. 86, 1993, pp. 134–141.

[18] Sperling, F. A. H., "Mitochondrial DNA Variation and Haldane's Rule in the *Papilio glaucus* and *P. troilus* Species Groups," *Heredity*, Vol. 70, 1993, pp. 227–233.

[19] Simon, C., "Molecular Systematics at the Species Boundary: Exploiting Conserved and Variable Regions of the Mitochondrial Genome of Animals via Direct Sequencing from Amplified DNA," *Molecular Techniques in Taxonomy*, G. M. Hewitt, A. W. B. Johnston and J. P. W. Young, Eds., Springer Verlag, Berlin, 1991, pp. 33–72.

[20] Goldenthal, M. J., McKenna, K. A., and Joslyn, D. J., "Mitochondrial DNA of the Blowfly *Phormia regina*: Restriction Analysis and Gene Location," *Biochemical Genetics*, Vol. 29, 1991, pp. 1–11.

Address requests for reprints or additional information to
Felix Sperling, Ph.D.
Dept. of Biology
University of Ottawa
Ottawa, Ontario K1N 6N5
Canada

APPENDIX I

DNA sequences for *Phaenicia sericata* (Ps), *Lucillia illustris* (Li), and *Phormia regina* (Pr), across mitochondrial COI, COII and tRNA leucine genes. End primers (lower case) are included to make fragment sizes consistent with Fig. 1. Dots indicate identity to sequence of *P. sericata*.

```
Ps tacaatttatcgcctaaacttcagccATTTAATCGCAACAATGGTTATTTTCAACTAATCATAAAGATATTGGAACTTTA   80
Li                           .........................................................T......
Pr                           ..........G......A........T....................T......

Ps TATTTTATTTTTGGAGCTTGATCCGGAATAATTGGAACTTCTTTAAGAATTCTAATTCGAGCTGAATTAGGACATCCTGG  160
Li ..C........................T....C........A.........T......................C.....
Pr .....C.....C..................G.A.................C...........C........G..C.....

Ps AGCTTTAATTGGAGATGATCAAATTTATAATGTAATTGTTACAGCTCATGCTTTTTATTATAATTTTTTTATAGTAATGC  240
Li T..A..............C...............................................C.............
Pr ...AC.............C....................A..G.....C.................C...........A.

Ps CAATTATAATTGGAGGATTTGGAAATTGATTAGTTCCATTAATACTAGGAGCTCCAGATATAGCATTCCCTCGAATAAAT  320
Li .....................................T......T...................................
Pr .G...........................C.......CCT...T....G.....T........................

Ps AATATAAGTTTTTGACTTTTACCTCCTGCATTAACTTTATTATTAGTTAGTAGTATAGTAGAAAACGGAGCTGGAACAGG  400
Li ........A..............................C.......A................................
Pr ...........C...................................A.................T..G...........

Ps ATGAACAGTTTACCCTCCTCTATCTTCTAATATTGCTCATGGAGGAGCTTCTGTTGATTTAGCTATTTTCTCTCTTCATT  480
Li ...................T....A........C...................................T..........
Pr ......T........A...T...................C........A.............................C.

Ps TAGCAGGAATTTCTTCAATTTTAGGAGCTGTAAATTTTATTACTACAGTTATTAATATACGATCAACAGGAATTACTTTT  560
Li .............A........................................................G.........
Pr .G..C................................C.....A..T..A..............T...........A...

Ps GATCGAATACCTTTATTTGTTTGATCAGTAGTAATTACAGCTTTATTACTTTTATTATCATTACCAGTATTAGCAGGAGC  640
Li ........G........C........................................................T.....
Pr ..........................T.....T.....T...C.T..............T.................T..

Ps TATTACAATACTTTTAACAGACCGAAATCTTAATACATCATTCTTTGACCCTGCAGGAGGAGGTGATCCAATTTTATACC  720
Li ......T...........T..T..............T..............A...........A..C.............
Pr ......T...T.A.....T.........T.A.....T.....T.....T..A..T........A..C..T..........

Ps AACATTTATTTTGATTCTTTGGACACCCTGAAGTTTATATTTTAATTTTACCTGGATTTGGAATAATTTCTCATATTATT  800
Li ......................T..T......................................................
Pr ...................C..T...........A.......................C.....................

Ps AGTCAAGAATCAGGTAAAAAGGAAACATTCGGTTCATTAGGGATGATTTATGCCATATTAGCTATTGGATTATTAGGATT  880
Li ..............A..G..............A...C....A..A........T..........................
Pr ..............A..G........T.....A..T.....A..A.....C..T..........................

Ps TATTGTTTGAGCTCATCATATATTTACAGTAGGAATAGACGTTGATACACGAGCTTACTTTACTTCAGCTACTATAATTA  960
Li ......A.....A.....C........T.....G..............T.......................A.......
Pr ........................C..C....................T....................C..A.......

Ps TTGCTGTACCAACTGGAATTAAGATTTTTAGTTGATTAGCAACTCTTTATGGAACTCAATTAAACTATTCCCCTGCTACT 1040
Li ......................A.........................................T.....T.........
Pr ......T...............A............C............................T..C.....A......

Ps TTATGAGCTTTAGGATTTGTATTTTTATTCACTGTAGGAGGTTTAACTGGAGTTGTTTTAGCTAACTCTTCAGTTGATAT 1120
Li ..............G.................A........A.......................T..............
Pr .............................T...........A...............C.............TA.......
```

```
Ps TATTTTACATGATACATACTATGTAGTAGCTCACTTCCATTATGTTTTATCAATGGGAGCTGTATTTGCTATTATAGCAG 1200
Li ..................T..............T..T..C.....A........A.........................
Pr .........C........T..............T.....C.....A........A........C..............T.

Ps GATTTGTTCACTGATATCCTTTATTTACAGGATTAACTTTAAATACTAAGATATTAAAAAGTCAATTTGCTATTATATTT 1280
Li ....C.....T.........C...........C...........G.A.....G.....G.........A...........
Pr ..........T....TC...........T...C....A.......AC...T.............................

Ps ATTGGGGTAAATTTAACATTCTTCCCTCAACATTTCTTAGGATTAGCAGGAATACCACGACGATATTCAGACTACCCAGA 1360
Li .....A...........T.................T......C.............G........C.....T........
Pr .....T...........C.............................T........T........C..T.....T.....

Ps TGCTTACACAACTTGAAATGTAATTTCTACAATTGGGTCAACAATTTCTTTATTAGGAATTTTATTCTTCTTCTTTATTA 1440
Li ................................................CC................T.....T..C....
Pr ..........G...................T.....T...........A..C.............T.....T..C....

Ps TTTGAGAAAGTCTTGTATCTCAACGTCAAGTTTTATTCCCTGTTCAATTAAATTCATCAATTGAATGATTACAAAATACT 1520
Li .................A..................................................C...........
Pr ...........T.A.....A.......................C....................................

Ps CCACCAGCTGAACATAGTTATTCTGAATTGCCTTTATTAACTAATTTCTAATATGGCAGATTAGTGCAATGGATTTAAGC 1600
Li .............................A..................................................
Pr ..............C..C...AG......A..............C...................................

Ps TCCATATATAAAGTATTTTACTTTTATTAGAAT--AATAAATGTCAACATGAGCAAATTTAGGTTTACAAGATAGTTCTT 1680
Li .................................TA.............................................
Pr ................................---.............................................

Ps CTCCTTTAATAGAACAATTAATCTTTTTCCATGATCACGCACTTTTAATTTTAGTAATAATTACTGTACTTGTAGGATAC 1760
Li ...............................C...................................T...........T
Pr ............................T.....C..T.............................T..A.....T..T

Ps TTAATGTTTATATTATTTTTTAACAAATATGTAAATCGATATTTATTACACGGACAAACTATTGAAATTATTTGAACAAT 1840
Li .....A............................................T.............................
Pr .....A......C.......C....................C...C.T..T.............................

Ps TTTACCTGCAATTATTTTATTATTTATTGCTTTCCCTTCTCTTCGACTTTTATATTTACTTGATGAAATTAATGAACCTT 1920
Li ......A..........................T............T.A.....C.........................
Pr .................................T........A...........C...T.A...................

Ps CAATTACTTTAAAGGCAATTGGTCATCAATGATATTGAAGTTATGAATATTCAGATTTTGCAAATATTGAATTCGATTCA 2000
Li ......................A...........C....................C...A.............T......
Pr .......C..............A..C........C......................................T.....T

Ps TATATAATTCCTACTAACGAATTATCAATTGATAGCTTTCGTTTATTAGATGTAGATAATCGAGTAGTTTTACCAATAAA 2080
Li ..C...........A..T.................A..C...........C..T.....C....................
Pr ..............A..T.................T..C...........C..T..........................

Ps TTCTCAAGTTCGAATTTTAGTAACTGCTGCTGATGTAATTCATTCTTGAACTATCCCAGCACTAGGAGTAAAGGTAGATG 2160
Li .......A...................C..A..............................T..................
Pr ..A....A................A..A.................A..............TT.......T..........

Ps GAACTCCTGGTCGACTAAATCAAACAAATTTTTTAATTAACCGACCAGGTTTATTTTACGGACAATGTTCAGAAATTTGT 2240
Li .T.................C....................T.....T.................................
Pr .T..A..C..A...T....C........C...........T.................T.....................

Ps GGAGCTAATCATAGTTTTATACCAATTGTAATTGAAAGAATTCCAGTAAATTACTTTATTAAGTGAATTTCTAATAATAT 2320
Li ................................................................................
Pr ....................G......................................C..A...............G.

Ps AAATTCTTCATTagatgactgaaagcaagtaatggtctc 2359
Li ............
Pr ...C........
```