MITOCHONDRIAL DNA PHYLOGENETIC ANALYSIS OF MYRMECOLACIDAE (INSECTA: STREPSIPTERA)

by

Natalie Rose Halbert

Submitted to the Office of Honors Programs and Academic Scholarships
Texas A&M University
in partial fulfillment of the requirements for

1998-99 UNIVERSITY UNDERGRADUATE RESEARCH FELLOWS PROGRAM

April 15, 1999

Approved as to style and content by:

James B. Woolley (Faculty Advisor)
Department of Entomology

Susanna Finnell, Executive Director
Honors Programs and Academic Scholarships

Fellows Group: Molecular Biology

ABSTRACT

Mitochondrial DNA Phylogenetic Analysis of Myrmecolacidae (Insecta: Strepsiptera)

Natalie Rose Halbert (Dr. James B. Woolley)

University Undergraduate Fellow 1998-1999

Texas A&M University, Department of Entomology

The relationship between *Caenocholax fenyesi* Pierce (Myrmecolacidae) males (host: *Solenopsis invicta*) and females (host: *Hapithus agitator*) is examined. Further, the relationships between several unidentified myrmecolacid females from Papua New Guinea are determined.

Methods were developed for extraction, amplification, and sequencing of mitochondrial DNA (mtDNA) from Strepsiptera. High-resolution mtDNA sequences obtained by these methods were used in the phylogenetic analysis of Myrmecolacidae. Phylogenetic Analysis Using Parsimony (PAUP) was employed in the analysis of myrmecolacid family structure using segments of the mitochondrial genes Cytochrome Oxidase Subunit I (COI, 284 bases) and 12S rRNA (12S, 272 bases).

Analysis of the COI region showed 30% variation between the *C. fenyesi* male and putative female, 0.7% variation between two unknown myrmecolacid from West New Britain (WNB-A and WNB-B), and 20% variation between the WNB unknown species and an unknown myrmecolacid species from Popondetta (POP). *Elenchus japonicus* (Strepsiptera: Elenchidae) was used as an outgroup to generate a phylogenetic tree of the COI region.

Analysis of the 12S region showed 37% variation between the *C. fenyesi* male and putative female, 0.3% variation between WNB-A and WNB-B, and 19% variation between a WNB-A/B and a

third unknown species from West New Britain (WNB-C). *Stichotrema dallatoraneum* (Strepsiptera: Myrmecolacidae) was also included in the 12S analysis for confirmation of a monophyletic myrmecolacid family.

The high degree of variation between the *C. fenyesi* male and putative female sequences indicate these individuals are not the same species. Additionally, this data provides evidence of at least two unidentified species from Papua New Guinea (WNB-A/B, WNB-C, and POP).

ACKNOWLEDGEMENTS

I must thank and acknowledge many people for their help with this project, as I did not even know what Strepsiptera and PCR were when I started. I thank Dr. Spencer Johnston first and foremost for the last two years of ideas, suggestions, and motivation. Without his unending faith in me, I might never have finished this project. Additionally, I thank Dr. Jeyaraney Kathirithamby for generously donating samples, for her hospitality, and for teaching us all about the wonderfully exotic Strepsiptera. Prior to this project, Rebecca Saff and I performed preliminary PCR and sequencing of some strepsipteran species at Oxford University with Dr. Kathirithamby. Though the results of our trial-and-error work were not used in the final data analysis, the knowledge and experiences gained at Oxford are irreplaceable. Therefore, I thank Rebecca for her friendship, honesty, and help getting this project off the ground. I am grateful to Dr. Jim Derr and his laboratory for teaching me a thing or two about DNA extractions and PCR. I must also mention Andrea Jensen and Larry Ross, who have truly kept me sane this year and helped me to overcome a multitude of research problems. Further, I owe thanks to Dr. Jerry Cook for unselfishly donating samples. Last, but certainly not least, I thank Dr. James Woolley for serving as my advisor, for unending advice on phylogenetic analysis, and for his dedication to entomology.

DEDICATION

This thesis is dedicated to my parents, who have given me 22 years of unending, unselfish, and sometimes unreciprocated love and support; and to Judy Crews and Andrew Dunn, who have instilled in me a passion to pursue the unknown.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Strepsiptera Background

Strepsiptera (Insecta), often referred to as "stylops," are curious entomophagous parasitoids of cosmopolitan distribution. They are curious not only for their biology but also for their host-parasitoid relationships. Kathirithamby (1989) reports that Strepsiptera are exclusively parasitic in members of seven insect orders such as Orthoptera, Hemiptera, Hymenoptera, and Diptera. Strepsiptera exhibit extreme sexual dimorphism and unusual life cycles. The order Strepsiptera is divided into two suborders: Mengenillidia and Stylopidia (Kinzelbach 1978).

Strepsiptera have only two free-living stages: the small, active adult winged male and the first instar larva. The adult males have prominent compound eyes, elegantly branched antennae, "haltere-like" fore-wings, and fan-shaped hind wings. They are short-lived (some live for only two hours), they do not feed, and a breeding flight is their sole mission. With the exception of one family of Mengenillidia, the bizarre adult females are neotenic. Mengenillidia females pupate outside their host, while Stylopidia females remain permanently endoparasitic. Therefore, one must look for the host to find Stylopidia females. Strepsipteran females produce live young called first instar larvae (or triungulins), who emerge from the female and her host through a brood canal. Strepsipteran hosts carry the triungulins to their nest of eggs or larvae. Once in the nest, the first instar Strepsiptera can enter the eggs or larvae of their host (Kathirithamby and Johnston 1992). The endoparasitc Strepsiptera and the strepsipteran host mature simultaneously; Strepsiptera generally emerge from adult hosts. Triungulins are microscopic, ranging in size from 0.08-0.30 mm. The size of adult male Strepsiptera is 1.6-6mm, while adult female Strepsiptera are 2mm-2.6cm (Kathirithamby 1989).

Myrmecolacidae Background

Within the suborder Stylopidia, the family Myrmecolacidae is unique. Myremecolacid males and females not only have contrasting morphology and development, but they also parasitize different hosts. Myrmecolacid males parasitize Formicidae (ants) while females infect Orthoptera (grasshoppers, crickets). As a result of this extreme sexual dimorphism, it is difficult to correctly match males with females of each species (Kathirithamby 1989, Kathirithamby & Hamilton 1992). Of the 98 species of Myrmecolacidae described, only five have known females. Of these females, only 2 have been positively identified and matched to males. Free-living males caught in light traps have been used to identify 87 myrmecolacid species. The remaining 6 species were identified from males found within their hosts (Kathirithamby 1998). One approach used to pair males and females has been to compare the respective first instar larvae found in the hosts. This evaluation gives tenuous and uncertain relationships since there is little taxonomic data on first instars. Information is additionally limited because the first instar larvae are rarely found in the host. On host entry, first instars moult to apodous second instar larvae (Kathirithamby et. al. 1984). Therefore, comparisons of this nature are based on the shed cuticles of the first instar rather than the entire organism.

The fitness of any stylopized strepsipteran host is drastically affected as the male pupates and the female nurtures her young. Examples of decreased fitness of the host include reduced wing length, egg number, gonadal weight, ovarian development, and feeding activity (Solulu et al. 1998). Unlike other entomophagous parasitoids, host death occurs only after the free-living male emerges from the puparium, or after all the first instar larvae have emerged from the neotenic female.

Caenocholax fenyesi (Myrmecolacidae)

Pierce (1909) originally described *Caenocholax fenyesi* from four males collected in Mexico by Dr. A. Fenyes. Light trap collections of *C. fenyesi* in the United States show this species is found in 7 of

the southern states. *C. fenyesi* were first recorded stylopizing *Solenopsis invicta* Buren (red imported fire ant) sexuals and workers by Kathirithamby and Johnston (1992) in College Station, Texas.

Imported fire ants were introduced into the United States in Mobile, Alabama in 1918 (Creighton 1930). Buren (1972) reported that *S. richteri* Forel was most likely introduced before *S. invicta*. Though *S. richteri* is now only found in a small region of Misissippi and Alabama, *S. invicta* can be found in most of the Southern United States. *S. invicta* has a devastatingly large economic impact on agriculture every year, despite numerous chemical control methods. Stylopization of *S. invicta* by pupal stages of *C. fenyesi* causes abnormal social behaviors such as isolation, unusual posture, and decreased aggressiveness (Cook 1996). The findings of Kathirithamby and Johnston (1992) and Cook (1996) suggest potential for biological control of *S. invicta* by *C. fenyesi*.

Five strepsipteran females were found within two *Hapithus agitator* (Orthoptera, collected by Dr. R. Wharton) hosts in College Station, Texas. The females were described as members of the family Myrmecolacidae based on external morphological characters (Cook 1996). Cook (1996, et al. 1998) further identified the females as *Caenocolax fenyesi*. This identification was based on comparisons of male *C. fenyesi* first instars and first instars found in the putative females.

Stichotrema dallatorreanum (Myrmecolacidae)

Hofeneder (1910) was first to describe the female and first instar larvae of *Stichotrema dallatorreanum*. Though the male member of this species has yet to be identified, female *S. dallatorreanum* have been found in numerous leaf-eating tettigoniid hosts in Papua New Guinea (PNG), as shown in Table 1. These hosts are economic pests of the oil palm (*Elaeis guineensis*), a significant cash crop of PNG (Solulu et al. 1998). *S. dallatorreanum* is primarily, if not exclusively, parthenogenetic. If males do exist, they play a major role in genetic variation, since the species must reproduce both parthenogenetically and sexually.

Table 1. Known *S. dallatorreanum* Tettigoniid Hosts (Orthoptera) in Papua New Guinea

(As shown in Solulu 1997)

Host Species	**Distribution in PNG**
Sexava nubilis	Pak Island, Manus Province (Admeralty Island)
Sexava sp.	Maprik, Shouten Island, East Sapik Province
Segestes decoratus	Bubia, Morobe Province
Segestidea novaeguinea	Popondetta, Oro Province
Phyllophora lanceolata	Levage, West New Britain
Lesser katydid sp.	Dami, West New Britain

An unknown myrmecolacid female whose orthopteran host is still undetermined has been collected from Popondetta (POP). A similar orthopteran host stylopized with a myrmecolacid female was found in West New Britain (WNB-C). Kathirithamby (unpublished) hypothesizes that these two females are the same species of Myrmecolacidae within the same species of orthopteran host. Another undetermined orthopteran from West New Britain was found on two occasions to be stylopized with unknown myrmecolacid females (WNB-A and WNB-B). These two individuals are believed to be the same species (but different from POP and WNB-C) with a different host than POP/WNB-C (Kathirithamby unpublished). The unknown species and their hosts are summarized in Table 2.

Table 2. Summary of Species of Strepsiptera Analyzed

Strepsipteran Species (family, sex)	**Host Species (order)**
Stichotrema dallatorreanum (Myrmecolacidae, female)	*Segestidea novaeguineae* (Orthoptera)
Elenchus japonicus[1] (Elenchidae, male and female)	*Sogatella furcifera* (Delphacidae)
Caenocolax fenyesi (Myrmecolacidae, proposed female)	*Hapithus aggitatus* (Orthoptera)
Caenocolax fenyesi (Myrmecolacidae, male)	*Solenopsis invicta* (Hymenoptera)
Species from Popondetta (Myrmecolacidae, female)	Known but undetermined, same as WNB-C (Orthoptera)
Species from West New Britain-A&B (Myrmecolacidae, female)	Known but undetermined (Orthoptera)
Species from West New Britain-C (Myrmecolacidae, female)	Known but undetermined, same as Popendetta (Orthoptera)

[1] *E. japonicus* used as outgroup for analysis of the COI region (see Introduction, Phylogenetic Analysis)

Mitochondrial DNA in Population Studies

The preferential use of mitochondrial genes as opposed to nuclear genes in population and phylogenetic studies is the result of many factors. The mitochondrial genome evolves 1-10 times faster than the nuclear genome (Vawter and Brown 1986) and does not exhibit recombination. Therefore, genetic variation is frequently detectable between closely related species and even individuals of the same species (Simon et al. 1994). Additional advantageous characteristics of the mitochondrial genome include maternal inheritance, small size, and highly conserved gene order and content. Further, mtDNA tends to vary more between populations and less within populations than nuclear genes (Moritz et al. 1987). Figure 1 details the mitochondrial genome.

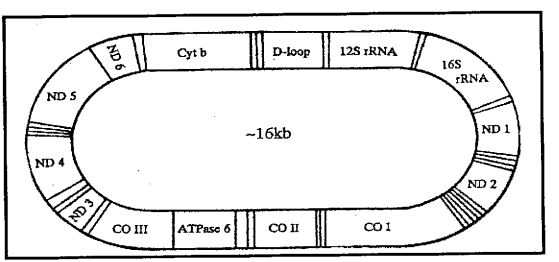

Figure 1. Structure of the Animal Mitochondrial Genome

The mitochondria encodes 13 proteins, 22 transfer RNAs, and 2 ribosomal RNAs.

PCR and DNA Sequencing

Polymerase Chain Reaction (PCR) is an *in vitro* enzymatic method of DNA amplification that takes advantage of the temperature dependent properties of DNA replication. When DNA polymerase is added to a buffered solution of synthetic oligonucleotide primers, the four deoxyribonucleotide triphosphates (dNTPs: dATP, dTTP, dGTP, and dCTP), and template DNA, exponential synthesis of the target fragment is possible. Modern PCR methods employ a thermostable DNA polymerase, such as *Taq* polymerase. *Taq* polymerase is isolated from the thermophilic hot springs bacterium *Thermus aquaticus* (Hillis et al. 1996). The thermostable properties of *Taq* polymerase enable it to withstand temperatures of 96C necessary to denature the template DNA.

PCR occurs in three stages: denaturation, annealing, and extension. During the first stage, DNA is denatured at high temperatures (~94C). Denaturation exposes the priming site for each strand. Next, the temperature is lowered (typically 45-60C) to allow annealing of the primers to their complementary site on the templates. The annealing temperature is variable and depends on the length and base composition of the primers. An increase in temperature to 72C usually follows annealing, which allows for DNA-primer hybrid stabilization and primer extension. During the extension phase, *Taq* polymerase incorporates dNTPs into an extension of the primers, thus creating a nascent strand complementary to the template. Each cycle theoretically doubles the quantity of DNA in the PCR solution. A typical PCR includes 25 to 40 cycles. In theory, PCR allows an immense quantity of a specific region to be produced from a small amount of template DNA (Hillis et al. 1996).

Chain termination sequencing of the amplified target fragment by standard dideoxy chain termination reactions involves incorporating dNTPs and labelled dideoxynucleotides (ddNTPs) in primer-directed DNA extensions. Dideoxynucleotides are identical to dNTPs except they lack a 3' hydroxyl group. The polymerase terminates chain extension when a ddNTP is incorporated into the nascent chain. In this study, each ddNTP was labeled with a different dRhodamine dye so that individual nucleotides could be detected by an argon laser.

Phylogenetic Analysis

Molecular data allows direct comparisons of relative levels of genetic differentiation among groups of organisms. Phylogenetic analysis of such molecular data involves complex theories and statistics (Hillis et al. 1996), but the underlying logic is quite simple. Parsimony analysis, the estimation of a parameter based on the principle of minimizing the number of events necessary to explain the data, is generally employed as part of phylogenetic studies. Further, methods such as bootstrapping are used to determine the strength of particular phylogenetic trees. Bootstrapping measures the "robustness" of individual regions of phylogenetic trees by resampling the data.

Outgroups, or those taxa assumed to be phylogenetically outside of an ingroup, are commonly employed in phylogenetic analysis. Use of an outgroup gives directionality to character state changes. Outgroup comparison is only possible, however, if the resulting ingroup is monophyletic. The species examined in this study all belong to the family Myrmecolacidae. To make the myrmecolacid sequences a monophyletic ingroup, *Elenchus japonicus* was chosen as the outgroup for COI sequence analysis (see Table 2). *E. japonicus* serves exceptionally well as an outgroup because this species is outside, but closely related to, the family Myrmecolacidae. *E. japonicus* parasitizes *Sogatella furcifera* (Delphacidae) and is a member of the family Elenchidae. The classification of Elenchidae as a sister group to Myrmecolacidae is based on morphological characters such as antennal joints, female cephalothorax, and brood canal openings (Kathirithamby 1989). Since a 12S sequence for *E. japonicus* was unavailable, the solutions for 12S are unrooted.

OBJECTIVES

I. To adapt existing molecular genetic techniques for use in extraction, amplification, and sequencing of mitochondrial DNA from Strepsiptera;

II. To generate mtDNA sequences from

- *C. fenyesi* male and putative female
- *S. dallatorreanum*
- Unknown species from Popendetta and West New Britain
- *E. japonicus*;

III. To determine if the putative female and known male *C. fenyesi* are indeed the same species;

IV. To examine phylogenetic relationships among unknown species from Popendetta and West New Britain.

MATERIALS AND METHODS

Species Collection and Preservation

Fire ant (*S. invicta*) colonies from Bee Creek Park in College Station, Texas in June and July 1998 were analyzed for the presence of male *C. fenyesi*. A minimum of 100 fire ants were dissected from each colony. Male *C. fenyesi* were found in four of the six sampled colonies in larval (late fourth instar), early pupal (with headcap formation), or late pupal (with full puparium) stages (Johnston unpublished). Males and each host were preserved separately in 95% ethanol.

Dr. Jerry Cook of Texas A&M University generously donated a putative female *C. fenyesi* and unstylopized *H. aggitatus* (host of putative female) for our studies. The putative female *C. fenyesi* was captured in a malaise trap in Lick Creek Park, College Station, Texas. Samples of *Elenchus japonicus*, *Sogatella furcifera* (host of *E. japonicus*), *Stichotrema dallatoraneum*, *Segestidea novaeguineae* (host of *S. dallatoraneum*), the species from Popondetta, and the species from West New Britain were all generously donated by Dr. Jeyaraney Kathirithamby of Oxford University. All samples were preserved in 95% ethanol.

DNA Extraction

Optimal sample sizes for DNA extraction of each specimen are shown in Table 3. DNA was easily extracted from female Strepsiptera, but DNA extractions from males proved more difficult. These differences are likely due to the dimorphism between strepsipteran females and males. The limited size and proteinaceous exoskeleton of the males explains the difficulties in DNA extraction. Females are larger and composed of soft tissue, which more readily permits protein degradation and nucleic acid extraction.

TABLE 3. Optimal Sample Sizes for DNA Extraction

Sample	Size[1]
Male *C. fenyesi*	All
Proposed Female *C. fenyesi*	25-35 mg
Female *E. japonicus*	25-35 mg
Male *E. japonicus*	All
Female *S. dallatoraneum*	25-35 mg
Female Popondetta species	25-35 mg
Female West New Britain species (A, B, &C)	25-35 mg
Strepsipteran Hosts	30-40mg

[1] Sample size of "All" reflects entire organism; defined sizes are approximate

DNA extractions were performed using a QIAamp® Tissue Kit (QIAGEN® #29306). The tissue protocol was used with the following exceptions and clarifications:

1) Samples were rinsed to eliminate excess ethanol and air dried in 1.5 mL microfuge tubes.
2) Samples were frozen in liquid nitrogen and pulverized (maximum 5 times) with pellet pestle.
3) Proteinase K digestion was performed overnight (minimum 12 hours).
4) A 5 minute, 70C incubation of the QIAamp® spin column loaded with 100 µL deionized H_2O preheated to 70C preceded elution.
5) DNA was eluted twice with 100 µL deionized H_2O preheated to 70C (total elution volume of 200 µL).
6) Extracted DNA was stored at -20C.

Selection of Gene Regions

Preliminary attempts at obtaining sequences by PCR of a region of the mtDNA 16S rRNA gene produced unusable sequences, possibly due to nonspecific primers. Regions of the mtDNA 12S rRNA

and Cytochrome Oxidase Subunit I (COI) genes were chosen on the basis of their effectiveness in preliminary studies. The 12S rRNA primers were designed by Dr. Scott Davis of Texas A&M University for *Polistes belligreans* (yellow jacket) and cover a region of 400 bp.

12S rRNA forward	5'-TAC TAT GTT ACG ACT TAT-3'
12S rRNA reverse	5'-AAA CTA GGA TTA GAT ACC C-3'

The COI primer sequences are universal and have been shown effective for lepidopterans, dipterans, coleopterans, thysanopterans, hemipterans, and homopterans (Simon et al. 1994). These primers span a 440 bp region of the COI gene. The standardized primer names given by Simon are shown in parentheses.

COI forward (C1-J-1751)	5'-GGA TCA CCT GAT ATA GCA TTC CC-3'
COI reverse (C1-N-2191)	5'-CCC GGT AAA ATT AAA ATA TAA ACT TC-3'

Polymerase Chain Reaction

The PCR protocols for 12S and COI primers are shown in Table 4. All reactions had a total volume of 50 µL. Table 5 illustrates the PCR thermal cycling protocols for 12S and COI primers performed on a GeneAmp® PCR System 9700 by PE Applied Biosystems. The 12S annealing temperature is considerably lower than the COI annealing temperature because the 12S primers have a lower GC content and are shorter than the COI primers.

Table 4. PCR Protocols for 12S and COI Gene Regions

Reagent	12S Volume (μL)[1]	COI Volume (μL)[1]
Qiagen® Taq PCR Master Mix[2]	25	25
Template DNA[3]	5	5
Forward Primer, 5 pmol/μL	5	2.5
Reverse Primer, 5 pmol/μL	5	2.5
$MgCl_2$, 25mM	3	3
Water, distilled	7	12

[1] Total reaction volume is 50 μL.
[2] Contains: 1.25 U *Taq* polymerase, 0.5XPCR Buffer (1.5 mM $MgCl_2$), 200 μM each dNTP
[3] Final amount of DNA/reaction: 0.5-1μg

Table 5. PCR Thermal Cycling Protocols for 12S and COI Gene Regions

Cycles	12S	COI
1	94C, 3 min	94C, 3 min
40	94C, 1 min 46C, 1 min 72C, 1 min	94C, 1 min 52C, 1 min 72C, 1 min
1	72C, 10 min	72C, 10 min

DNA Sequencing

ABI Prism® BigDye™ Terminatior Cycle Sequencing Ready Reaction Mix was used in the sequencing reaction. PCR products were ethanol precipitated and resuspended in distilled water to a final concentration of 20 ng/μL. The sequencing reaction was performed using a GeneAmp® PCR System 9700 with the protocol given in Table 6. The total volume of each reaction was 6 μL. Products of the sequencing reaction were ethanol precipitated and sequenced according to standard instructions for the ABI Prism® 277 Sequencer.

Table 6. Sequencing Reaction Protocol

Reagent	Volume (μL)
Big Dye®	2
Template DNA, 20 ng/μL[1]	2
Primer, 5 pmol/μL	2

[1] Required template: 10 ng/100 bp DNA

Data Analysis

Forward and reverse sequences were aligned and corrected using Sequencher 3.0 (Gene Codes Corp. 1995). This generated a consensus sequence for each sample. These sequences were aligned with each other using CLUSTAL X (improved version of CLUSTAL V, Higgins et al. 1992) using a gap opening of 10, a gap extension of 0.5, and a transition weight of 0.5. *E. japonicus* sequences were included in the COI study as an outgroup in phylogenetic analysis. Phylogenetic Analysis Using Parsimony (PAUP* 4.0b2, Swofford 1999) was employed for parsimony analysis of the sequences. PAUP identifies informative positions (i.e. base substitutions) and calculates hypothetical relationships between taxa. A maximum parsimony method was used, which minimized the number of nucleotide

substitutions necessary to explain the data set. Maximum parsimony solutions were found by branch-and-bound searches followed by bootstrapping to test the strength of relationships between taxa.

RESULTS

The abbreviations used for this study are given in Table 7. Fragment length microvariation was observed for both gene regions. Out of 440 bp for the COI gene region, 284 bases were sequenced and analyzed. Of these, the only sequence insertion/deletion sites were for bases 126-132 (see Table 8) between *E. japonicus* (family Elenchidae) and all other sequences (family Myrmecolacidae). Table 8 shows CLUSTAL sequence alignments for the COI region.

Table 7. Legend for Sequence Analysis

Abbreviation	Species (sex)
WNB-A	Species from West New Britain, undetermined (female)
WNB-B	Species from West New Britain, undetermined (female)
WNB-C	Species from West New Britain, undetermined (female)
POP	Species from Popondetta, undetermined (female)
Sd	*Stichotrema dallatoraneum* (female)
Female Cf	*Caenocolax fenyesi* (putative female)
Male Cf	*Caenocolax fenyesi* (male)
Female Ej	*Elenchus japonicus* (female)
Male Ej	*Elenchus japonicus* (male)

Table 8. CLUSTAL Sequence Alignment for COI

Sequences for a 284 bp COI region are shown for 7 samples. Dashes represent site losses or gains and are automatically inserted by CLUSTAL to maintain sequence alignment.

```
                                 1111111111222222222233333333344444444455555555556666666666777777
Taxon/Node       12345678901234567890123456789012345678901234567890123456789012345678901234567890123456789012345678901234
-------------------------------------------------------------------------------------------------------------------------
WNB-B            GTATTTAAGTTTCGGTCTGTTAAAAGCATAGTAATTGCTCCAGCTAGTACTGGTAATGATAATAAAAGGAGAAG
WNB-A            GTATTTAAGTTTCGGTCTGTTAAAAGTATAGTAATTGCTCCAGCTAGTACTGGTAATGATAATAAAAGGAGAAG
POP              GTATTTAAATTACGGTCTGTTAAGAGTATTGTAATTGCTCCTGCTAGTACAGGAAGTGATAATAATAAAAGGAAG
Female C.f.      GTATTTAAGTTTCGATCAGTTAATAATATTGTAATTGCTCCTGCTAGTACAGGAAGTGATAATAATAAAAGGAAG
Male C.f.        GTATTTAAATTACGGTCAGTTAATAATATTGTAATTGCTCCTGCAAGAACTGGGAGTGAAAGTAAAAGAAGAAT
Male E.j.        GAATTAAAATTTCGATCAAATAATAGTATAGTAATTGCCCCAGCTAATACTGGTAAAGATAAAATTAATAAGAT
Female E.j.      GAATTAAAATTTCGATCAAATAATAGTATAGTAATTGCCCCAGCTAATACTGGTAAAGATAAAATTAATAAGAT

                                                                                                 1111111111111111111111111111111111111111111111111111
                 77777888888888899999999900000000011111111122222222223333333334444444444
Taxon/Node       56789012345678901234567890123456789012345678901234567890123456789012345678
-------------------------------------------------------------------------------------------
WNB-B            TGCTGTAATTGCTACAGATCAAACAAATAAAGGGGTTTGATCTAGTGATATTTGTGGAGCTCGTATATTGATAG
WNB-A            TGCTGTAATTGCTACAGATCAAACAAATAAAGGGGTTTGATCTAGTGATATTTGTGGAGCTCGTATATTGATAG
POP              TGCAGTGATTGCGACAGATCAAACGAATAAAGGTGTTTGATCTAGTGATATCCCGGGGGATCGTATATTAATTG
Female C.f.      TGCTGTAATAGCTACTGATCAAACAAATAATGGTGTTTGGTCTAATGTTATTCTTTCTGATCGTATATTAATTG
Male C.f.        AGCTGTAATAAGAATAGATCAAACTAGGAGGGGAATTTTATCTATAGTAAAATTTTTATGGTGTATGTTAATGA
Male E.j.        TGAAGTTAATAAAATAGATCAAATAAATAAAGATAATTTTTTAATTGGGAT----T--GTTCTCATATTGTAAA
Female E.j.      TGAAGTTAATAAAATAGATCAAATAAATAAAGATAATTTTTTAATTGGGAT----T--GTTCTCATATTGTAAA

                 11111111111111111111111111111111111111111111111111222222222222222222222
                 45555555556666666666777777777788888888899999999900000000011111111112222
Taxon/Node       90123456789012345678901234567890123456789012345678901234567890123456789012
-------------------------------------------------------------------------------------------
WNB-B            TTGTAGTAATGAAGTTTACGGCACCTAGAATAGAGGAAATTCCGGCTAAATGAAGGGAGAAAATTGCTAAGTCT
WNB-A            TTGTAGTAATGAAGTTTACGGCACCTAGAATAGAGGAAATTCCGGCTAAATGAAGGGAGAAAATTGCTAAGTCT
POP              TAGTTGTAATAAAATTTACTGCCCCTAAGATAGAGGAAATTCCCGCTAGATCAAAGGAAAAAATTGCTAAGTCT
Female C.f.      CTGTTGTAATAAAATTAACTGCTCCTAGAATTGATGAACACCAGCTAAGTGAAGAGAAAAATAGCTAAATCA
Male C.f.        TTGTAGAAATAAAATTAATAGCCCCTAAAATTGATGATATTCCAGCGATATGTAAGGAAAAAATAGATAGATCG
Male E.j.        TAGTAGTGATAAAATTAATAGATCTTAAAATTGAAGAAATACCAGCTAGGTGTAATCTAAAGATTATTAATTCT
Female E.j.      TAGTAGTGATAAAATTAATAGATCTTAAAATTGAAGAAATACCAGCTAGGTGTAATCTAAAGATTATTAATTCT

                 22222222222222222222222222222222222222222222222222222222222222222222222
                 22222223333333333444444444455555555566666666677777777788888
Taxon/Node       34567890123456789012345678901234567890123456789012345678901234
-------------------------------------------------------------------------------
WNB-B            ACAGATGCCCCTCCATGGGCAATACCTGATGATAGTGGAGGATACACAGTTCATCCTGTTCC
WNB-A            ACAGATGCTCCTCCATGGGCAATACCTGATGATAGTGGAGGATACACAGTTCATCCTGTTCC
POP              ACTGAAGCCCCTCCATGGGCAATTCCTGCTGAGAGGGGAGGGTAGACTGTTCATCCAGTACC
Female C.f.      ACTGAGGCTCCTCCATGTGCAATTGCTCCAGCGAGCGGAGGGTAAACTGTTCATCCTGGACC
Male C.f.        ATAGAGGCTCCTCTGTGAAAAATGTTGGAGGCTAAAGGTGGATAAATAGTTCATCCTGTTCC
Male E.j.        ATTGAATAGTCTGAGTGTCTTTGAGTTAAAGTAAATGGAGGATATAATGTTCACCCTGACCT
Female E.j.      ATTGAATAGTCTGAGTGTCTTTGAGTTAAAGTAAATGGAGGATATAATGTTCACCCTGACCT
```

There were 86 total base differences between the male and putative female *C. fenyesi*, representing a total variation of 30% (adjusted for missing data). In comparison, there were 0 base differences between the male and female *E. japonicus* and 2 base differences between the WNB-A and WNB-B samples. Base differences and total variation between samples for the COI region are shown in Table 9.

Table 9. Base Differences and Total Variation Between Taxa for a 284 bp COI Region

Base differences are shown below the diagonal, while total variation, or mean character differences, are shown above the diagonal (adjusted for missing data).

		1	2	3	4	5	6	7
1	**WNB-B**	-	0.00704	0.20423	0.22183	0.31690	0.37410	0.37410
2	**WNB-A**	2	-	0.20423	0.21479	0.30986	0.37050	0.37050
3	**POP**	58	58	-	0.21479	0.33099	0.37410	0.37410
4	**Female Cf**	63	61	61	-	0.30282	0.37410	0.37410
5	**Male Cf**	90	88	94	86	-	0.39928	0.39928
6	**Male Ej**	104	103	104	104	111	-	0.00000
7	**Female Ej**	104	103	104	104	111	0	-

Branch-and-bound search by PAUP resulted in the tree shown in Figure 2. *E. japonicus* male and female sequences were defined as a monophyletic sister-group to the ingroup. Bootstrapping support values were generated for 1000 replicates using the 50% majority-rule and are also shown in Figure 2. The consistency index for this tree is 0.8313.

Out of 400 bp for the 12S gene region, 272 bases were sequenced and analyzed. Table 10 shows CLUSTAL sequence alignments for the 12S region.

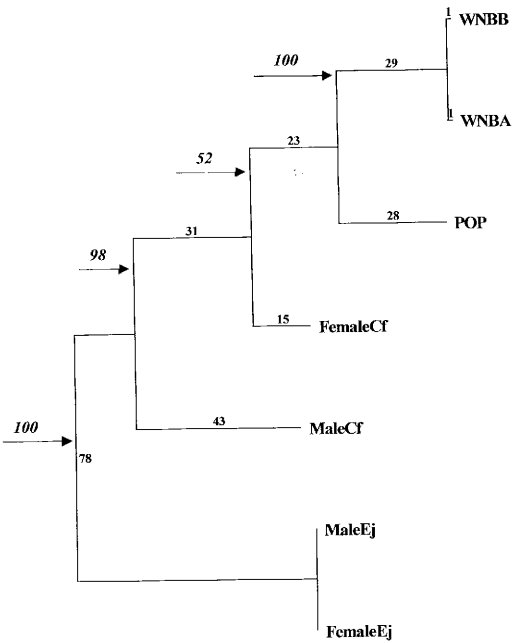

Figure 2. COI Phylogenetic Tree Generated by PAUP

Branch lengths are shown as numbers on branches. Arrows represent bootstrapping support values.

Table 10. CLUSTAL Sequence Alignment for 12S

Sequences for a 272 bp 12S region are shown for 6 samples. Dashes represent site losses or gains and are automatically inserted by CLUSTAL to maintain sequence alignment.

```
                    1111111111222222222233333333344444444455555555566666666677777777
Taxon/Node          12345678901234567890123456789012345678901234567890123456789012345678901234567
-------------------------------------------------------------------------------------------------
WNB-A               GTTTTAGTCTATTCAGAGGAATCTGTCCTATAATTGATAATACACGTTATATTATPACTTGATTTTGTAAATTCAGTT
WNB-B               GTTTTAGTCTATTCAGAGGAATCTGTCCTATAATTGATAATACACGTTATATTATACTTGATTTTGTAAATTCAGTT
WNB-C               GTCTTAATCTTTTCAGAGGAATCTGCCCCGTAATCGATAGTACACGTTAAATGCTACATAGTTTTGTAAAT-CAGCT
Female Cf           ATCTCATTCCATTCAGAGGAACCTACCCCGTGATTGATAGTACACGATTAGCTTTACTTAATTTATTT--G----TT
Sd                  ATTTTAATCAGTTTAGAGGATTCTGTTATTTA---GATAATTCACGTTATAAATTAATTTAATTTAATAA------TT
Male Cf             GTTATAATCAATTTAGAGGATTCTGTTATTTA---GAAAATACACGATTAATTTAATTTAATTTTATTA------TT

                                    1111111111111111111111111111111111111111111111111111111
                    7788888888889999999990000000001111111112222222222333333333444444444455555
Taxon/Node          8901234567890123456789012345678901234567890123456789012345678901234567890234
-------------------------------------------------------------------------------------------------
WNB-A               TGTATACCTCCGTCATAAGATTATCCTTGGAGGGTAATAATTTTCTTAATTTTTGATAGAAGAATATGTCAGGTCAA
WNB-B               TGTATACCTCCGTCATAAGATTATCCTTGGAGGGTAATAATTTTCTTAATTTTTGATAGAAGAATATGTCAGGTCAA
WNB-C               TGTATACCGCCGTCATAAGATTATCTTTTTAGAGAAGGAATTTCTTAATATTT--TACGATAGTATGTCAGGTCAA
Female Cf           TGTATATCTCCGTTATCAGAAAATCTTTTTAGAGTTATAATTTTCTTGATTTATAAATTATAAAGTATTTCAGGTCAA
Sd                  TTTATATCGTTGTTATAAAAAAATATTTTAATATTTAAAAATTTTTAATAAATT--TAATTTAATATTTCAAATCAA
Male Cf             TTTATATCGTTGTTATAAGAAAATATTTTAATATTTTAAATTTCTAA-AAATT--TTTATTAATATTTCAAATCAA

                    1111111111111111111111111111111111111111111222222222222222222222222222222
                    5555566666666667777777777888888888899999999900000000011111111112222222233
Taxon/Node          5678901234567890123456789012345678901234567890123456789012345678901234567890
-------------------------------------------------------------------------------------------------
WNB-A               GGTGCAGCTAATAATTAAGTTTGAGATGGATTACAATATAATTATATA---CATGGAAGTTATACTGAAAAT-TGTA
WNB-B               GGTGCAGCTAATAGTTAAGTTTGAGATGGATTACAATATAATTATATATG-CATGGAAGTTATACTGAAAAT-TGTA
WNB-C               GGTGCAGTTATAATTATGAG-GAGGTGGATTACAATAAAATTATTTA---TATGGATTNTATGTTGTAATA-TATA
Female Cf           GGTGCAGCTTATAATTAAGAG-GAGGTGGGTTCCAATAGGTTTTTGTTTACAACGGATTTGGTAGTGAAATACTATT
Sd                  AATAAAATTTATAATTAAAAATTA-ATAGATTATATTA--CTTAAATT----ATTTA-ATACTAATGAATTTAAATT
Male Cf             GATTTAATTTATAATTAAAGATTA-ATGGATTACAATA--AATATAAA----ATTGGCATTAAAATTTAATATTTTA

                    2222222222222222222222222222222222222222
                    3333333344444444455555555566666666667777
Taxon/Node          2345678901234567890123456789012345678901
-------------------------------------------
WNB-A               ACTGAAGGTGGATTTGAAAGTAAAATTGTTATATTATGATA
WNB-B               ACTGAAGGTGGATTTGAAAGTAAAATTGTTATATTATGATA
WNB-C               AATGAAGGTGGATTTGAAAGAAAACCTGGTAAATTATGATA
Female Cf           AATGAAGGTGGATTTGATAGTAAT-TTAACATATTTAGTTT
Sd                  AATAAAAAAAAATTTAAAAGTAATAATAAATTATATTTTAT
Male Cf             AATAAAATTGGATTTAATAGTAATTATATTAAATATTTTAT
```

The region from bases 203-206 was omitted from PAUP analysis due to alignment ambiguity. Elimination of this region, however, did not change the tree generated and had a minimal impact on the analysis of the 12S region. Of the remaining 268 bases, there were 93 total base differences between the male and putative female *C. fenyesi*, representing a total variation of 37% (adjusted for missing data). Base differences and total variation between samples for the 12S region are shown in Table 11.

Table 11. Base Differences and Total Variation Between Taxa for a 268 bp 12S Region

Base differences are shown below the diagonal, while total variation, or mean character differences, are shown above the diagonal (adjusted for missing data).

	1	2	3	4	5	6
1 **WNBA**	-	0.00375	0.19011	0.28571	0.38889	0.34127
2 **WNBB**	1	-	0.19392	0.28958	0.39286	0.34524
3 **WNBC**	50	51	-	0.26848	0.41434	0.35458
4 **FemaleCf**	74	75	69	-	0.39200	0.37200
5 **Sd**	98	99	104	98	-	0.21825
6 **MaleCf**	86	87	89	93	55	-

Branch-and-bound search by PAUP resulted in the unrooted tree shown in Figure 3. Bootstrapping support values were generated for 1000 replicates using the 50% majority-rule and are also shown in Figure 3. The consistency index for this tree is 0.8815.

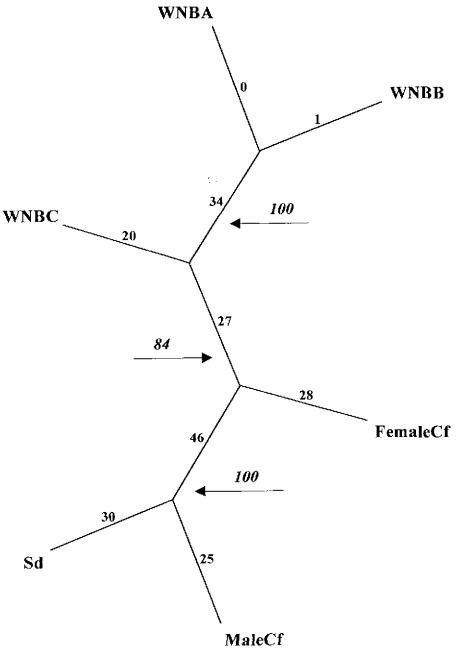

Figure 3. 12S Phylogenetic Tree Generated by PAUP

Branch lengths are shown as numbers on branches. Arrows represent bootstrapping support values.

CONCLUSIONS

General Considerations

Phylogenetic analysis showed the sequence data to be clean and the trees stable. The consistency index (CI) is a measure of the signal/noise ratio in the data (sum of the minimum number of steps required divided by the length of the tree). Therefore, a high CI indicates a tree based on low homoplasy, or very consistent data. The consistency index for COI was 0.8313, while that of 12S was 0.8815. These are high CI values, showing that the data is relatively free of conflicting phylogenetic signals.

Bootstrapping measures the robustness of phylogenetic trees, or how strong the evidence is for two branches to lay in a given orientation. High bootstrapping values indicate a robust tree. For instance, the boostrapping support value for COI between the male *C. fenyesi* and the upper branch is 98. Therefore, 98% of the random samplings gave a tree with the male *C. fenyesi* separate from the upper-portion of the tree (containing the putative female *C.fenyesi*). The bootstrapping values were high for every pair of branches except that for COI between the putative female *C. fenyesi* and the upper branch (bootstrapping support value of 52). Therefore, only 52% of the time did random sampling give the orientation of the two involved branches (lengths 23 and 15) as shown.

As in all sequence analysis, the validity of this analysis is compromised by the possibility of contamination. Though sterile techniques were used when handling these samples, there is always a risk of contaminating samples with human, viral, or other DNA. In the case of parasitoid organisms, the most likely cause of contamination is the host. For this reason, DNA from hosts of *E. japonicus*, *C. fenyesi* male, *C. fenyesi* putative female, and *S. dallatorreanum* was amplified and sequenced. The sequencing results were not clean enough at the time of

publication to include in this study. However, comparisons of host and respective parasite partial sequences proved that the samples were not contaminated by host DNA.

If the putative female *C. fenyesi* (or any other sample) were contaminated, we would expect the branch to be highly diverged from the phylogenetic tree (long branch). We do not see such a branch, however. The fact that the putative female *C. fenyesi* branch lies in the middle, not the outside, of both trees adds proof of true, uncontaminated sequences.

The 12S alignment was not as clean as that of the COI region. The alignment might be a real artifact of the high variability of the 12S region. However, to avoid the possibility of ambiguous phylogenetic branch positions and lengths, a four base pair region was omitted from the PAUP analysis (bases 203-206, see Table 10). The branch positions were unaffected by this deletion, and the change in branch lengths was minimal.

Caenocolax fenyesi

Comparisons of total nucleotide variation reveal that the male *C. fenyesi* and the putative female are probably not the same species. The male C. fenyesi and putative female sequences vary at 86 bases in the COI region (30% of the total) and at 93 bases in the 12S region (37% of the total), as shown on Tables 9 and 11. In addition, the male *C. fenyesi* and putative female do not appear as a monophyletic group in either the COI or 12S phylogenies (Figures 2 and 3); thus, there is no evidence that they share nucleotide substitutions that the other species do not. The discrepancy in total percent variation between the two mitochondrial genes results from multiple factors. Although the 12S gene evolves at about the same rate as the average for the entire mitochondrial genome (Simon et al. 1990), the 12S primers used in this study were specifically designed to cover a region of high variability. Further, the cytochrome oxidase I gene is known to be highly conserved among mtDNA regions (Hillis et al. 1996). These factors are consistent with the higher overall variation observed in the 12S region for the male *C. fenyesi* and putative female.

The high amount of variation between the male *C. fenyesi* and the putative female is consistent with variation between other species in this study. For instance, the variation between *S. dallatorreanum* and WNB-A is 39% for the 12S region, which is of the same magnitude as the 37% variation observed between the male *C. fenyesi* and the putative female. This suggests that the putative female is not of the species *C. fenyesi*.

The phylogenetic analysis does, however, support the hypothesis that the female in question is of the family Myrmecolacidae. This is especially evident in the COI tree (Figure 2), since the sequence of the female (putative) falls between the male *C. fenyesi* and the POP/WNB branch. Since we know that both *C. fenyesi* males and POP/WNB are myrmecolacid, we deduce that any species which falls within this part of the tree will be a myrmecolacid. The 12S tree (Figure 3) further demonstrates that the female (putative) falls in the middle of the known myrmecolacid species.

Papua New Guinea Myrmecolacid Species

WNB-A and WNB-B are the same species of Myrmecolacidae, as seen in the 12S and COI variance tables and trees. The COI region shows a 0.704% variance, while the 12S region shows a 0.375% variance between WNB-A and WNB-B. Kathirithamby's hypothesis (unpublished) that these two organisms are the same species is strongly supported by this data.

The 12S unrooted phylogenetic tree illustrates strong divergence of *S. dallatorreanum*, WNB-C, and WNB-A/B. It appears that these three groups are separate species (as thought by Kathirithamby) and that WNB-C is more closely related to WNB-A/B than *S. dallatorreanum*. Further, we note that *S. dallatorreanum* from PNG is more closely related to *C. fenyesi* of North America than to other Myrmecolacidae from PNG.

As predicted by Kathirithamby (unpublished), POP and WNB-A/B are different species. The COI tree illustrates this point. The total variance between POP and WNB-A is 20.4%, as is

the variance between POP and WNB-B. At this point, there is not sufficient data to support the relationship of POP and WNB-C.

Future Research

To learn more about possible biological control of *S. invicta* by *C. fenyesi*, we must identify the *C. fenyesi* female. Since *C. fenyesi* is a myrmecolacid species, we expect to find the female within an orthopteran host. We also need to identify the male of our putative myrmecolacid female (originally presumed *C. fenyesi* female). The later may be a more feasible endeavor, since free-living males can be caught in light traps.

The use of both *C. fenyesi* and *S. dallatorreanum* as biological control agents is an undeniable possibility. The negative effects of these parasites on their host (decrease fitness) provide a natural means of pest control. However, proposing to use these parasites as biological control agents is presumptuous without knowledge of the mate, mating process, and sexual characteristics of each.

Currently, there are at least two undescribed species of myrmecolacid from Papua New Guinea (WNB-A/B, POP, WNB-C). Further mtDNA phylogenetic analysis will unravel the relationship between POP and WNB-C. These new species need further investigation and taxonomic identification.

This endeavor began as a pilot project to determine if DNA could be effectively extracted, amplified, and quantitatively compared in Strepsiptera. The results satisfactorily prove that mtDNA phylogenetic analysis is a suitable means of determining relationships between and among species of Strepsiptera. Likewise, the sole use of morphological characters to identify members of this peculiar order of insects is not necessarily an accurate means of classification.

REFERENCES

Buren, W.F. 1972. Revisionary studies on the taxonomy of the imported fire ants. J. Geo. Entomol. Soc. 7(1): 1-26.

Cook, J.L. 1996. A study of the relationship between *Caenocholax fenyesi* Pierce (Strepsiptera, Myrmecolacidae) and the red imported fire ant, *Solenopsis invicta* Buren (Hymenoptera, Formicidae). Ph.D. dissertation, Texas A&M University, College Station.

Cook, J.L., S.B. Vinson, and R. Gold. 1998. Developmental Stages of *Caenocholax fenyesi* Pierce (Strepsiptera: Myrmecolacidae): Descriptions and significance to the higher taxonomy of Strepsiptera. Int. J. Insect Morophol. & Embryol. 27: 21-26.

Creighton, W.S. 1930. The new world species of the genus *Solenopsis* (Hymenoptera: Formicidae). Proc. Am. Acad. Arts Sci. 66: 39-151.

Higgins, D.G., A.J. Bleasby, and R. Fuchs. 1992. CLUSTAL V: Improved software for multiple sequence alignment. Comput. Appl. Biosci. 8: 189-191.

Hillis, D.M., C. Moritz, and B.K. Mable [eds.]. 1996. Molecular Systematics, 2nd ed. Sinauer, Sunderland, Massachusetts.

Hofeneder, K. 1910. *Stichotrema* n.g. *dallatorreanum* n.sp. eine in einer Orthoptere libende Strepsiptere. Zool. Anz. 36: 47-49.

Kathirithamby, J. 1989. Review of the Order of Strepsiptera. Syst. Entomol. 14: 41-92.

Kathirithamby, J. 1998. Host-parasite associations: anatomical and developmental consequences. Int. J. Insect Morphol. Embryol. 27(1): 39-51.

Kathirithamby, J., and W.D. Hamilton. 1992. More covert sex: the elusive females of Myrmecolacidae. Trends Ecol. Evol. 7(10): 349-351.

Kathirithamby, J., and J.S. Johnston. 1992. Stylopization of *Solenopsis invicta* (Hymenoptera: Formicidae) by *Caenocholax fenyesi* (Strepsiptera: Myrmecolacidae) in Texas. Ann. Entomol. Soc. Am. 85(3): 293-297.

Kathirithamby, J., D.S. Smith, M.B. Lomas, and B.M. Luke. 1984. Apolysis without ecdysis in larval development of a strepsipteran, *Elenchus tenuicornis* (Kirby). Zool. J. Linnean Soc. 82: 335-343.

Kinzelbach, R.K. 1978. Strepsiptera. Die Tierwelt Deutchlands. 65: 166 pp.

Moritz, C., T.E. Dowling, and W.M. Brown. 1987. Evolution of animal mitochondrial DNA: Relevance for population biology and systematics. Ann. Rev. Ecol. Sys. 18: 269-292.

Pierce, W.D. 1909. A monographic revision of the twisted winged insects comprising the order Strepsiptera Kirby. U.S. Nat. Mus., Bull. 66: 1-232.

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

Simon, C., S. Paabo, T.D. Kocher, and A.C. Wilson. 1990. Evolution of mitochondrial ribosomal RNA in insects as shown by the polymerase chain reaction, pp. 235-244. *In* M. Clegg and S. O'Brien [eds.], Molecular Evolution. UCLA Symposium on Molecular and Cellular Biology, New Series. Vol. 122. Wiley-Liss, New York.

Solulu, T.M. 1997. Influence of *Stichotrema dallatorreanum* Hofender (Strepsiptera: Myrmecolacidae) on the performance of *Segestidea novaeguinea* (Bransisk) (Orthoptera: Tettigoniidae) in Papua New Guinea. M.S. thesis. Oxford University, Oxford.

Solulu, T.M., S.J. Simpson, and J. Kathirithamby. 1998. The effect of strepsipteran parasitism on a tettigoniid pest of oil palm in Papua New Guinea. Phys. Ento. 23: 388-398.

Swofford, D.L. 1999. PAUP*. Phylogenetic Analysis Using Parsimony (*and other methods). Version 4. Sinauer Associates, Sunderland, Massachusetts.

Vawter, L., and W.M. Brown. 1986. Nuclear and mitochondrial DNA comparisons reveal extreme rate variation in the molecular clock. Science 234:194-196.