COMPENSATORY BASE CHANGES ILLUMINATE MORPHOLOGICALLY DIFFICULT TAXONOMY.

by

MICHAEL WILLIAM RUHL

(Under the Direction of Tracie M. Jenkins)

ABSTRACT

Compensatory base changes (CBCs) in the ribosomal RNA (rRNA) internal transcribed spacer 2 (ITS2) secondary structures have been used to successfully verify the taxonomy of closely related species. CBCs have never been used to distinguish morphologically indistinct species. Under the hypothesis that CBCs will differentiate species, novel software for CBC analysis was applied to morphologically indistinguishable insect species in the genus *Altica*. The analysis was species-specific for sympatric *Altica* beetles collected across four ecoregions and concordant with scanning electron microscopy data. This research shows that mining for CBCs in ITS2 rRNA secondary structures is an effective method for taxa analysis.

INDEX WORDS: ITS2; internal transcribed spacer 2; systematics; species delimitation; rDNA secondary transcript structure; speciation; taxonomy, *Altica litigata*, *Altica* sp.

COMPENSATORY BASE CHANGES ILLUMINATE MORPHOLOGICALLY DIFFICULT TAXONOMY.

by

MICHAEL WILLIAM RUHL

BS, University of Georgia, 2008

A Thesis Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

MASTER OF SCIENCE

ATHENS, GEORGIA

2009

COMPENSATORY BASE CHANGES ILLUMINATE MORPHOLOGICALLY DIFFICULT TAXONOMY.

by

MICHAEL WILLIAM RUHL

Major Professor: Tracie M. Jenkins

Committee: S. Kris Braman
Carol Robacker
Margie Paz

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
August 2009

DEDICATION

This thesis is dedicated to my wife Rebecca, daughter Kaitlyn, and son Trevor. Their personal sacrifice and support have allowed me to pursue and achieve my academic goals. Without their support I would never have reached this milestone in my education.

ACKNOWLEDGEMENTS

I first thank the members of my committee (Dr. Kris Braman, Dr. Carol Robacker, and Dr. Margie Paz). I can still remember the look on their faces when they first saw my proposed graduate plan. While I recognized the look of doubt, and words such as "ambitious" were tossed around the table, they were supportive. Instead of restraining me from pursuing my 'ambitious' goals, they always encouraged me to go after them. I thank them for their support and trust. You made me work even harder because I didn't want to let you down.

I acknowledge my major professor, Dr. Tracie M. Jenkins for encouraging me to expand my academic horizons, think outside the box, take risks, and to "go for it kiddo". Her tenacious nature, perfectionist attitude, and tireless work ethic were instrumental in my academic success. She set the standard high and taught me not to walk in the footsteps of giants, but to be the giant and carve out my own. Thank you for talking me into being your graduate student, taking me under your wing, and believing in me. You have been a positive influence in my life.

Dr. Matthias Wolf at the University of Würzburg, Germany helped me to understand the bioinformatic processes behind ITS2-based phylogenetic and CBC analyses. He has been an invaluable resource and a willing collaborator. I am not only grateful to him, but consider myself fortunate to have worked with him.

Dr. Laurent LeSage shared with me his immense knowledge of *Altica* biology, taxonomy, as well as gross and scanning electron (SEM) morphology. His generosity significantly contributed to the success of this work. I will be forever indebted to him.

I am thankful for the friendship and help from my fellow classmates, especially Glen Ramsey (the ultimate department resource), Su Yee Kim (my EndNote advisor and one of the nicest people I know), and Melanie Pawlish (moral support, especially on the highway). Also a big thanks goes to Whitney Boozer, Mark Galatowitsch, Evelyn Carr, Jose Adriano Giorgi, and Sonja Brannon who have all helped me at some point on this journey.

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

1.1. INTRODUCTION AND LITERATURE REVIEW

This purpose of this study was to evaluate the efficacy of species determination via compensatory base change (CBC) analysis. Compensatory base changes (CBCs) are mutations (Figure 1.1) that occur in both nucleotides of a paired structural position while retaining the paired nucleotide bond. A hemi-CBC (hCBC) is a mutation of a single nucleotide in a paired structural position while maintaining the nucleotide bond. Analysis of these CBCs in the ribosomal RNA (rRNA) internal transcribed spacer 2 (ITS2) secondary structure has been used to successfully verify the taxonomy of closely related species. These analyses have been performed primarily in plants and fungi (Muller et al., 2007) and have never been used to distinguish morphologically indistinct species in eukaryotes. This study, therefore, tests the hypothesis that a CBC analysis would differentiate species in a morphologically equivocal insect taxon, the beetle genus *Altica*.

The Evening Primrose Beetle (*Altica litigata)* (Coleoptera: Chrysomelidae) is an established landscape and nursery pest that has been reported to aggregately feed on Crape Myrtle (non-indigenous ornamental in the family Lythraceae) and Primrose (indigenous ornamental in the family Onagraceae) plant families (Pettis, 2007). Throughout the southeastern United States the entire life cycle of *A. litigata* has been described on Primrose but only adults have been described on Crape Myrtle (Cabrera et al., 2008; Pettis et al., 2004). Thus, *A. litigata* adults are assumed to migrate from

Primrose to Crape Myrtle (*Lagerstremia spp*.) in late spring/early summer to feed (Cabrera et al., 2008). However, recent mitochondrial and ITS array sequence data contradicted this assumption and revealed that flea beetles collected from Crape Myrtle and Primrose were not the same *Altica* species (Jenkins et al., 2009a).

1.2. Morphology

Correct insect taxonomy is crucial to understanding the interactions between the insect herbivore and its host plant. The flea beetle species in the genus *Altica* are morphologically so similar that accurate species identification has been challenging (Fernandez and Hilker, 2006; W. M, 1979) and, therefore, host-insect behavior difficult to study, particulary since many *Altica* species are polyphagous with multiple species feeding on the same plant at the same time (Jenkins et al., 2009a; Jenkins et al., 2009b; L. LeSage, Ontario Canada, personal communication).

The adult male aedaegus (intromittent organ) is commonly used to identify species within the genus *Altica* (Laroche et al., 1996; LeSage, 1995b, 2000, 2002). Electron scanning microscope (SEM) data from this study showed that the aedeagus of *A. litigata* and *Altica* sp. were morphologically distinct from each another (Figure 1.2). This character, however, has been reported to be unreliable due to intraspecific variation, parthenogenetic reproduction, and unrecognized sibling species (Laroche et al., 1996; W. M, 1979). A DNA molecular marker that is reliable, inexpensive, and provides relatively fast results would be useful for verification of species, especially in the genus *Altica*.

1.3. ITS2 structure

The ITS2 intergenic sequence is located between the 5.8S and 28S ribosomal genes. It can evolve 30x slower (Caccone et al., 2004) than mitochondrial DNA (mtDNA) and has been shown to be an important part of ribogenesis (Joseph *et al*., 1999). Despite variability in ITS2 sequence and length, its rRNA secondary transcript structure contains a highly conserved core (Coleman, 2003; Schultz et al., 2005). This highly conserved core appears to be common to all eukaryotes and should be useful for evaluating the relationships of closely related organisms (Coleman, 2009a; Schultz et al., 2006). Thus a CBC between two closely related species may infer that they can no longer successfully intercross (Muller et al., 2007).

1.4. Sequence homogenization

Jenkins et al. (2009a) showed that for species in three beetle genera (*Altica*, *Lysathia*, and *Phaedon*) the rDNA sequence array (18S partial, ITS1, 5.8S, ITS2, and 28S partial) was not only species specific but homogenized within species. Could this homogenization of the rDNA array translate into a CBC within one of the conserved helices of the ITS2 secondary transcript structure between *A. litigata* and *Altica* sp., as in Muller et al. (2007)? It is the overall purpose of this research to answer this question.

CHAPTER 2

COMPENSATORY BASE CHANGES ILLUMINATE MORPHOLOGICALLY DIFFICULT TAXONOMY.[1]

[1] Ruhl, M.W, Wolf, M., and Jenkins, T.M Revision submitted to *Molecular Phylogenetics & Evolution*, 07/02/2009.

2.1. Introduction

Compensatory base changes (CBCs) are mutations that occur in both nucleotides of a paired structural position which retain the paired nucleotide bond. A hemi-CBC (hCBC) is a mutation of a single nucleotide in a paired structural position which maintains the nucleotide bond. Analysis of these CBCs in the ribosomal RNA (rRNA) internal transcribed spacer 2 (ITS2) secondary structure has been used to successfully verify the taxonomy of closely related species. These analyses have been performed primarily in plants and fungi (Muller et al., 2007) and have never been used to distinguish morphologically indistinct species in eukaryotes. This study, therefore, tests the hypothesis that a CBC analysis would differentiate species in a morphologically equivocal insect taxon, the beetle genus *Altica*.

The Evening Primrose Beetle (*Altica litigata)* (Coleoptera: Chrysomelidae) is an established landscape and nursery pest that has been reported to aggregately feed on Crape Myrtle (non-indigenous North American ornamental in the family Lythraceae) and Primrose (indigenous North American ornamental in the family Onagraceae) plant families (Pettis et al., 2007). The entire life cycle of *A. litigata* has been described on Primrose (Pettis et al., 2004). Only adult beetles have been found on Crape Myrtle (*Lagerstroemia* sp.), which were assumed to migrate from Primrose to Crape Myrtle in late spring/early summer to feed (Pettis et al., 2004). Recent sequence data (Figure 2.1) appeared to contradict this assumption and revealed that flea beetles collected from Crape Myrtle and Primrose were not the same *Altica* species (Jenkins et al., 2009a; Jenkins et al., 2009b; L. LeSage, Ontario Canada, personal communication).

2.1.1. Morphology

Correct insect taxonomy is crucial to understanding the interactions between the insect herbivore and its host plant. Flea beetle species in the genus *Altica* are morphologically indistinct, making accurate species identification challenging (Fernandez and Hilker, 2006; Phillips, 1979). Thus, host-insect behavior has been difficult to study, particularly since many *Altica* species are polyphagous with multiple species feeding on the same plant at the same time (Jenkins et al., 2009a).

The adult male aedaegus (intromittent organ) is commonly used to identify species within the genus *Altica* (Laroche et al., 1996; LeSage, 1995). Scanning electron microscopy (SEM) data from this study showed that the aedeagus of *A. litigata* and *Altica* sp. were morphologically distinct from each other (Figure 1.2). This character, however, has been reported to be unreliable by some due to intraspecific variation, parthenogenetic reproduction, and unrecognized sibling species (Laroche et al., 1996; Phillips, 1979). A DNA molecular marker that is reliable, inexpensive, and provides relatively fast results would be useful for verification of species, especially in the genus *Altica*.

2.1.2. ITS2 structure

The ITS2 intergenic sequence is located between the 5.8S and 28S ribosomal genes. It can evolve 30x slower (Caccone et al., 2004) than mitochondrial DNA (mtDNA) and has been shown to be an important part of ribogenesis (Joseph et al., 1999). Despite variability in sequence and length, the rRNA secondary transcript structure of the ITS2 region contains a highly conserved core (Coleman, 2003; Schultz

et al., 2005). This highly conserved core is common to all eukaryotes and should be useful for evaluating the relationships of closely related organisms (Coleman, 2009; Schultz et al., 2006). Thus, a CBC between two closely related species may infer that they can no longer successfully intercross (Muller et al., 2007).

2.1.3. Sequence homogenization

Jenkins et al. (2009a) showed that for species in three beetle genera (*Altica*, *Lysathia*, and *Phaedon*) the rDNA sequence array (18S partial, ITS1, 5.8S, ITS2, and 28S partial) was not only species specific but homogenized within a species. Could this homogenization of the rDNA array translate into a CBC within one of the conserved helices of the ITS2 secondary transcript structure between *A. litigata* and *Altica* sp., as in Muller et al. (2007)? The purpose of our research is to answer this question.

2.2. Materials and Methods

2.2.1. GenBank sequence query

The National Center for Biotechnology Information (NCBI) database (Benson et al., 2008) was queried for sequences containing the 5.8S (partial or complete genes), Internal Transcribed Spacer 2 (ITS2 complete), and 28S (partial or complete genes). All sequences from the Family Chrysomelidae meeting this criterion (Table 1.1) were analyzed.

To identify Coleopteran species having helix IV in the ITS2, GenBank was queried with the search string "Coleoptera Internal Transcribed Spacer 2 Complete" on June 22, 2009. The query produced 709 accessions. Each accession was individually

analyzed for a complete ITS2 fragment utilizing the "Annotate" feature (default settings) on the ITS2 website (http://its2.bioapps.biozentrum.uni-wuerzburg.de/cgi-bin/index.pl?annotator). If neither the 5.8S rRNA nor 28S rRNA ends were identified, the Annotate process was repeated with the "Analyze reverse complement (exclusively)" option enabled. Each complete ITS2 fragment was then folded using Mfold (http://mobyle.pasteur.fr/cgi-bin/portal.py?form=mfold) with default settings and analyzed for the presence of helix IV.

2.2.2. DNA extraction, polymerase chain reaction (PCR), and sequencing

A total of 129 individual samples were analyzed for ITS2 sequence. Specimens from previous studies totaled 119 (Jenkins et al., 2009a; Jenkins et al., 2009b). New DNA samples were extracted from an additional 10 individuals according to Jenkins et al. (2007). These include three *Altica litigata* collected in Tifton, GA by Will Hudson, four *Altica corni* collected in Quebec, Canada by Laurent LeSage, two *Altica chalybea* museum specimens collected in Quebec, Canada and mounted by Laurent LeSage, and one *A. chalybea* collected in Griffin, GA by Tyler D. Eaton.

ITS2 fragments were amplified according to Jenkins et al. (2007) using novel primers anchored in the 5.8S (FB5.8SFWD: CTGGACCACTCCTGGCT) and 28S rDNA region (FB28SREV: GGTAGTCTCACCTGCTCTG).

Amplified ITS2 DNA fragments were sequenced in both directions by the Sequencing and Synthesis Facility (SSF) at MWG (High Point, NC). Sequences were verified for accuracy using Sequencher v.3.1.1 (Gene Codes Corp., Ann Arbor, MI) software. Contigs representing consensus sequences were merged into single

consensus sequences.

2.2.3. 5.8S-28S rRNA interaction and HMM-based ITS2 annotation

All ITS2 sequences from Jenkins et al. (2009a, 2009b), unique to this study, or accessed from GenBank (Table 1.1) were annotated by identifying a 25 nucleotide interaction of the 5' 5.8S rDNA subunit end with 25 nucleotides of the 28S rDNA subunit 3' end (Keller et al., 2009). Sequences without this interaction were discarded as they are either incomplete, contain errors, or lacked the required 25 nucleotides on each end for the interaction to occur.

2.2.4. Homology modeling

An ITS2 template structure was made from *A. litigata* (GenBank No. EU682395) using RNA structure (Mathews et al., 2004) (default settings at 37^{o} C) and visually verified to be the most correct structure having the lowest energy (-63.4 dH, third lowest energy form). *A. litigata* was chosen for the homology model because it had the closest genetic relationship with our unknown beetle and had a shorter sequence than other beetles in the study. Modeling longer sequences after shorter sequences will fail to produce high quality models (set to "Identity" matrix with a threshold value >50%, gap costs: gap open 15, gap extension 2).

Homology modeling (Coleman, 2007; Schultz et al., 2005; Schultz et al., 2006; Selig et al., 2008; Wolf et al., 2005a) using the visually verified *A. litigata* template structure was performed on all remaining ITS2 sequences. Sequences with structure information were saved in FASTA format. Mean helix variance among *Altica* sequences

was calculated (standard deviation (SD) calculated using the formula: $SD=\sqrt{[\sum(X-M)^2/(n-1)}$, standard error of the mean (SEM) calculated using: $SEM=SD/(\sqrt{N})$. Confidence interval (CI), standard error of mean (SEM), degrees of freedom (df), and P values calculated with InStat software (GraphPad Software, 2236 Avenida de la Playa, La Jolla, CA 92037 USA).

2.2.5. 4SALE Alignment and CBC Analysis

The ITS2 sequences with homology structures were synchronously aligned using 4SALE (Seibel et al., 2006; Seibel et al., 2008). CBCAnalyzer (Wolf et al., 2005b) was used to confirm initial results produced by the CBC Matrix feature in 4SALE. An analysis of hemi-CBCs was done using CBCAnalyzer version 1.1. The alignment with structural information (Figure 2.2) was exported. The resultant filename was changed to a filename with the extension ".xfasta" (Schultz and Wolf, 2009).

2.2.6. Phylogenetic analysis using ProfDistS

The alignment output file, with '.xfasta' extension, from 4SALE was imported into ProfDistS (Friedrich et al., 2005; Wolf et al., 2008). "RNA/DNA structure Profile Neighbor Joining" was selected from the "Run" menu (Bootstrap= 1000, Distance Correction Model= General Time Reversible, Ratematrix Q= Q_ITS2.txt (an ITS2 specific model included as a supplemental file in ProfDistS)). The resultant tree file with node strengths was viewed in ProfDistS and then visualized and reproduced in Dendroscope (Huson et al., 2007) (annotated in Microsoft PowerPoint).

2.2.7. Morphological characterization

Aedeagus from each species was photographed by scanning electron microscopy (SEM) using the Zeiss 1450EP variable pressure SEM (Carl Zeiss MicroImaging, Inc. One Zeiss Drive, Thornwood, NY 10594, and Oxford Instruments X-Ray Technology, Inc., 275 Technology Circle, Scotts Valley, CA 95066). Specimens were first photographed in variable pressure mode and then coated in a 2-micron layer of gold using the SPI Module Sputter Coater (Structure Probe, Inc., 569 East Gay Street, West Chester, PA 19380). Coated specimens were photographed in high vacuum (2600 Pa) mode.

The length and width of each aedeagi was digitally measured from seven *A. litigata* and six *Altica* species using the measurement tool within the Zeiss 1450EP SEM operating program. The mean length and width was calculated for each species and standard deviations were calculated per section 2.4.

2.3 Results

2.3.1. GenBank sequence query

There were 709 GenBank accessions, all of which were analyzed for Helix IV. Seven of these accessions contained helix IV and four accessions represented Chrysomelids used in this study (Table 1.1). The other three non-Chrysomelid species, one Hymenopteran (a parasitoid wasp on *Altica*), and two Curculionid beetles, were not relevant to this study and, therefore, not included.

2.3.2. DNA extraction, polymerase chain reaction (PCR), and sequencing

The primer pair FB5.8SFWD and FB28SREV effectively amplified the complete requisite ITS2 sequence with a minimum of 25 nucleotides from the 5' 5.8S rDNA and 25 nucleotides from the 3' 28S rDNA regions.

2.3.3. 5.8S-28S rRNA interaction and HMM-based ITS2 annotation

The presence of a 5.8S-28S rDNA gene interaction (Figure 2.3) is evidenced by the ability to produce an HMM-based ITS2 annotation (not shown). Since ITS2 sequence was homogenized across all individuals within a species, annotated ITS2 sequences were deposited in GenBank under the accession numbers *Altica corni* (GQ325432), *Altica chalybea* (GQ325433), *Altica litigata,* haplotype ITSB (EU682395), *Altica* sp., haplotype ITSC (EU682396), and *Lysathia ludoviciana,* haplotype ITSA (EU682397).

2.3.4. Homology modeling

Homology modeling *Altica* sp. after *A. litigata* (Figure 2.4) produced a high mean structural identity of 98.41667% (SD=2.90637%, SEM=1.6779935%, 95% CI= 91.1968499 to 105.6364901, df=2, two-tailed P value equals 0.4450 (statistically insignificant)) considering all helices of all *Altica* samples. In addition, when considering the homology modeling of all Chrysomelid species in this study (Table 1.1), there is a high mean structural identity of 84.39516% (SD=14.4587%, SEM=2.5968592%, 95% CI=79.0916650 to 89.6986550, df=30, two-tailed P value is less than 0.0001) observed.

2.3.5. 4SALE alignment and CBC analysis

Aligned sequences (Figure 2.2) produced a phylogenetic tree (Figure 2.5) that separated the *Altica* beetle species in concordance to the phylogenetic tree shown in Jenkins et al. (2009a). It also separated all Chrysomelid species used in this study according to genera. As hypothesized, CBC analysis (Figure 2.6) using novel software (Keller et al., 2009; Seibel et al., 2008) revealed the presence of CBCs between *A. litigata* and *Altica* sp. that distinguishes *Altica* species.

2.3.6. Phylogenetic analysis using ProfDistS

Tree topology from the phylogenetic analysis using synchronous ITS2 secondary structure and sequence separated *A. litigata* and *Altica* sp. into well-supported subclades (80% bootstrap strength) (Figure 2.5).

CBC Analysis located one CBC in the third helix and three additional hemi-CBCs, also in the third helix, between *A. litigata* and *Altica* sp. (Figure 2.6). CBCs were also present between all four *Altica* species represented in the study (Table 1.1). Only *Nisotra* sp. of the Chrysomelid beetles in tribe Alticini, which included the four *Altica* species in this study, had helix IV (Table 1.1).

2.3.7. Morphological characterization

The aedeagus of *A. litigata* was distinctly different in shape, length, and width in comparison to *Altica* sp. (Figure 1.2). Mean (M) length of *A. litigata* was M=1.81329 mm (SD=0.07695 mm, SEM=0.02908437 mm, 95% CI=1.74212312 to 1.88445688 mm, df=6, two-tailed P value is less than 0.0001) and width M=375.61429 µm (SD=15.50607 µm, SEM=5.86074358 µm, 95% CI=361.27356720 to 389.95501280 µm, df=6, two-

tailed P value is less than 0.0001). Mean length of *Altica* sp. M=1.26833 mm (SD=0.03549 mm, SEM=0.01448873 mm, 95% CI=1.23108552 to 1.30557448 mm, df=5, two-tailed P value is less than 0.0001) and width M=305.51667 µm (SD=10.30602 µm, SEM=4.2074150 µm, 95% CI=294.7011638 to 316.3321762 µm, df=5, two-tailed P value is less than 0.0001).

2.4. Discussion

It is important to note that the CBC criterion usually works in only one direction. For example, if there is a CBC then there are two species; if there is no CBC there still could be two species (M. Wolf, University of Würzburg, personal communication). Muller et al. (2007) noted intraspecific CBCs might occur in 5.25% of samples when too few individuals from a diverse population were studied. In this study 129 individuals were examined and no intraspecific variation was seen. All specimens from a single species were observed to have no sequence variation regardless of geographic origin. The homogenization of the ITS array in *Altica* translated into one complete CBC as well as three hCBCs within the conserved helix III (Figure 2.6) of the ITS2 secondary transcript structure. This result was concordant with the ITS2 sequence/structure phylogeny (Figure 2.5) and the SEM morphological data sets (Tables 2.1-2.2), which attest to the strength of the CBC analysis.

The ITS2 sequences with their homology-based structure must be synchronously aligned if ITS2 CBC analysis is to be effectively applied. When ITS2 sequence alone is used, as in DNA barcoding of ITS2 sequences, nucleotide changes in conserved pairing positions that comprise a CBC are unaccounted for. This is why sequence alone may

indicate separate species but cannot rule out the possibility of sibling species or intraspecific variation within a single species when nested within the same phylogenetic clade.

The ITS2 helix IV was not observed in most of the beetles in this study (Table 1.1). This does not appear to be an Order level characteristic as the helix IV was observed in a few species within Coleoptera (Table 1.1) including one from the tribe Alticini, but not genus *Altica*, one from the tribe Lemini and two from the tribe Bruchini (Table 1.1). The etiology of the fourth helix reduction is not clear, nor is the role, if any, of helix IV in biogenesis, protein translation and adaptive feeding behavior. What is clear is that biogenesis still occurs. Interestingly, most beetles without a helix IV in the rDNA secondary transcript structure were observed to have additional, less defined helices located between helices III and I. Four helices were therefore spatially maintained within the structure. This tendency toward four helices may have implications for the binding of proteins integral to ribosomal biogenesis. The region of helix IV could be a target site for natural selection. As the adaptive needs of a beetle change, natural selection may act within the area of helix IV to facilitate a change in the direction of a more defined helix. This change may serve to more efficiently recruit non-ribosomal protein factors. Previous work suggests a framework for studying ribosomal biogenesis in yeast (Côté et al., 2002), and may offer a framework from which to study ribosomal biogenesis in insects.

The significance of this study to species determination is threefold. First, CBC analysis is capable of distinguishing morphologically indistinguishable eukaryotes. The CBC process is relatively quick (with known ITS2 sequences it takes less than two

hours), inexpensive and the tools are available online for free. Morphological analysis of the aedeagus using SEM for 13 specimens (two species) used approximately four technician hours for aedeagus SEM preparation, approximately three-and-a-half hours of SEM time (including 30 minutes for sputter coating) for a total cost[2] of approximately $480 USD. The same project at UGA commercial rates would have cost approximately $1550 USD.

In contrast, the total cost for a CBC analysis of the same 13 specimens (including DNA extraction, PCR, sequencing, annotating the ITS2, and folding the ITS2 structure, homology modeling, synchronous sequence/structure alignment, and tree building) was approximately $315[3].

Secondly, this study showed the value of CBC analyses for insects: within species, between species, and between genera. CBC analysis was species specific. It differentiated two morphologically indistinguishable species sympatric across four ecoregions[4]. This research further demonstrates the possible universality of CBC analysis in eukaryotic organisms.

Lastly, based on this study and sequence data from Jenkins et al. (2009a, 2009b), it is clear (although the gross morphology is indistinguishable) that *A. litigata*

---

[2] Incurred cost estimate based on non-UGA rate for academic purposes. Fees are accurate as of 6/29/2009 as noted by the University of Georgia Center for Advanced Ultrastructural Research lab (http://www.uga.edu/caur/equipment/fees.htm).

[3] Technician time at $20/hr. Four technician hours to process DNA, $15/specimen for DNA extraction, PCR (including primers), and sequencing, and two technician hours to perform computerized CBC analysis.

[4] Collections from four ecoregions (Southeastern Plains, Southern Coastal Plains, Interior Plateau, and Piedmont) across four US southern states (Louisiana, Alabama, Mississippi, and Georgia)(Jenkins et al (2007, 2009a, 2009b)). Ecoregion map is available from ftp://ftp.epa.gov/wed/ecoregions/us/useco.pdf. We also added 4 *Altica litigata* specimens from Oenothera bordering a Pecan orchard in Tifton, GA (a Southeastern Plains ecoregion).

and *Altica* sp. are separate species that were easily differentiated by CBC analysis.

2.5. Acknowledgements.

We are grateful to Tyler D. Eaton, UGA-Griffin Entomology, for invaluable work in DNA extraction, sequencing, and for his review. We also thank Noelle Barkley, USDA for reviewing and providing helpful comments. We are tremendously grateful for the two anonymous reviewers that helped make this manuscript significantly better. We acknowledge Dr. Laurent LeSage for generously sharing his time and immense knowledge of *Altica* biology. Funding for this project was provided by Dr. Tracie M. Jenkins's lab at the University of Georgia Department of Entomology.

CHAPTER 3

3.1. DISCUSSION

It is important to note that the CBC criterion usually works in only one direction. for example, if there is a CBC then there are two species (M. Wolf, University of Würzburg, person communication); if there is no CBC there still could be two species. Muller et al. (2007) noted intraspecific CBCs might occur in 5.25% of samples when too few individuals from a diverse population were studied. In this study, 129 individuals were examined and no intraspecific variation was seen. All specimens from a single species were observed to have no sequence variation regardless of geographic origin. The homogenization of the ITS array (Figure 2.1) in *Altica* translated into one complete CBC as well as three hCBCs within the conserved helix III (Figure 2.6) of the ITS2 secondary transcript structure. This result was concordant with the ITS2 sequence/structure phylogeny (Figure 2.5) and the SEM morphological data sets (Tables 2.1-2.2), which attest to the strength of the CBC analysis.

The ITS2 sequences with their homology-based structure must be synchronously aligned if ITS2 CBC analysis is to be effectively applied. When ITS2 sequence alone is used, as in DNA barcoding of ITS2 sequences, nucleotide changes in conserved pairing positions that comprise a CBC are unaccounted for. Therefore, no assumption can be made of the biological relationship between species analyzed with DNA barcoding. However, a CBC analysis between species may illuminate the biological relationship. It is also important to understand that the ITS2 sequence contains more variability than its

associated secondary transcript structure. Therefore, a CBC analysis within the ITS2 structure should contain less intraspecific variation than DNA barcoding while providing increased confidence in delimiting biological species.

The ITS2 helix IV was not observed in most of the beetles in this study (Table 1.1). This does not appear to be an Order level characteristic as the helix IV was observed in a few species within Coleoptera including one from the tribe Alticini, but not genus *Altica*, one from the tribe Lemini and two from the tribe Bruchini (Table 1.1). The etiology of the fourth helix reduction is not clear, nor is the role, if any, of helix IV in biogenesis, protein translation and adaptive feeding behavior. What is clear is that biogenesis still occurs. Interestingly most beetles without a helix IV in the rDNA secondary transcript structure were observed to have additional helices located between helices III and I. Four helices were therefore spatially maintained within the structure. This tendency toward four helices may have implications for the binding of proteins integral to ribosome biogenesis. The region of helix IV could be a target site for natural selection. As the adaptive needs of a beetle change natural selection may act within the area of helix IV to facilitate a change in the direction of a more defined helix. This change may serve to more efficiently recruit non-ribosomal protein factors. Like previous work which suggest a framework for studying ribosomal biogenesis in yeast (Côté et al., 2002), this research may offer a framework from which to study ribosomal biogenesis in insects.

An interesting observation was found when the CBC tree (Figure 3.1) was compared to the PNJ tree (Figure 2.5) in that the *Altica* clades were concordant in both. The CBC tree was provided in the figures for discussion only. It does, however,

demonstrate the ability to use CBCs as character states for a phylogeny. In this study, the CBC and PNJ tree topologies were concordant to the Genus level.

Morphological data (Figures 3.2 to 3.6) were provided for comparative purposes. This data was important for the following reasons. It established that the aedeagus of *Altica* sp. were distinct from *Altica litigata. Altica* sp. was also distinct from other *Altica* species defined by the aedeagus in this study, and in other studies (LeSage, L., 1995, 2000, 2002; L. LeSage, Ontario Canada, personal communication). The validity of the *Altica litigata* species designation, however, has been called into question (L. LeSage, Ontario Canada, personal communication). It appears as if the aedeagus of *A. litigata* (Fall, 1910) and *Altica foliacea* (LeConte, 1858) are the same. It will be up to future DNA analyses to determine the correct species designation. This study demonstrates the importance of having a one-to-one correlation between morphology and DNA. It also established aedeagi morphometric data on *Altica* sp. for future research.

The significance of this study to species determination is threefold. First, CBC analysis is capable of distinguishing morphologically indistinguishable eukaryotes. The CBC process is relatively quick (with known ITS2 sequences it takes less than two hours), inexpensive and the tools are available online for free. Morphological analysis of the aedeagus using SEM for 13 specimens (two species) used approximately four technician hours for aedeagus SEM preparation, approximately three-and-a-half hours of SEM time (including 30 minutes for sputter coating) for a total cost[5] of approximately $480 USD. The same project at UGA commercial rates would have cost approximately

[5] Incurred cost estimate based on non-UGA rate for academic purposes. Fees are accurate as of 6/29/2009 as noted by the University of Georgia Center for Advanced Ultrastructural Research lab (http://www.uga.edu/caur/equipment/fees.htm).

$1550 USD.

In contrast, the total cost for a CBC analysis of the same 13 specimens was approximately $315[6]. Also, since processing DNA for DNA barcoding is essentially the same (DNA Barcoding, 2006), the cost difference between DNA barcoding and CBC analysis is marginal. However, DNA barcoding is not as reliable when using a single genetic locus (Spooner, 2009). Analyzing multiple loci for DNA barcoding is therefore more costly then CBC analysis. Therefore, the ability to use a single locus to identify species, as shown in this study, could represent significant cost savings.

Secondly, this study showed the value of CBC analyses for insects: within species, between species, and between genera. CBC analysis was species specific. It differentiated two morphologically indistinguishable species sympatric across four ecoregions[7]. This research further demonstrates the possible universality of CBC analysis in eukaryotic organisms.

Lastly, based on this study and sequence data from Jenkins et al. (2009a, 2009b), it is clear that, although the gross morphology is indistinguishable, *A. litigata* and *Altica* sp. are separate species that were easily differentiated by CBC analysis.

---

[6] Technician time at $20/hr. Four technician hours to process DNA, $15/specimen for DNA extraction, PCR (including primers), and sequencing, and two technician hours to perform computerized CBC analysis.

[7] Collections from four ecoregions (Southeastern Plains, Southern Coastal Plains, Interior Plateau, and Piedmont) across four US southern states (Louisiana, Alabama, Mississippi, and Georgia)(Jenkins et al, (2007, 2009a, 2009b)). Ecoregion map is available from ftp://ftp.epa.gov/wed/ecoregions/us/useco.pdf. We also added 4 *Altica litigata* specimens from *Oenothera* bordering a Pecan orchard in Tifton, GA (a Southeastern Plains ecoregion).

Table 1.1. ITS2 species data (Coleoptera: Chrysomelidae: Galerucinae) including GenBank No., species, tribe and presence or absence of Helix IV. GQ325432, GQ325433, EU682395, EU682396, EU682397 sequence each represent a consensus haplotype since the ITS2 is homogenized within species (refer to text).

GenBank No.	Species[a]	Tribe	Helix IV
GQ325432	Altica corni	Alticini	No
GQ325433	Altica chalybea	Alticini	No
EU682395	Altica litigata	Alticini	No
EU682396	Altica species	Alticini	No
EU682397	Lysathia ludoviciana	Alticini	No
EU110864	Apteropeda orbiculata	Alticini	No
EU110865	Longitarsus membranaceus	Alticini	No
AY116111	Nisotra sp.	Alticini	Yes
EU110866	Psylliodes affinis	Alticini	No
EU110868	Psylliodes crambicola	Alticini	No
EU110869	Psylliodes chalcomerus	Alticini	No
EU110870	Psylliodes chrysocephalus	Alticini	No
EU110873	Psylliodes cupreus	Alticini	No
EU110875	Psylliodes fusiformis	Alticini	No
EU110877	Psylliodes gibbosus	Alticini	No
EU110881	Psylliodes hospes	Alticini	No
EU110880	Psylliodes hispanus	Alticini	No
EU110882	Psylliodes luteolus	Alticini	No
EU110885	Psylliodes luridipennis	Alticini	No
EU110889	Psylliodes marcidus	Alticini	No
EU110891	Psylliodes napi	Alticini	No
EU110894	Psylliodes laticollis	Alticini	No
EU110893	Psylliodes sophiae	Alticini	No
AY116107	Lema sp.	Lemini	Yes
AF278563	Diabrotica porracea	Luperini	No
AF278565	Diabrotica virgifera zeae	Luperini	No
AF278564	Diabrotica virgifera virgifera	Luperini	No
AJ622023	Timarcha olivieri parnassia	Luperini	No
AJ622022	Timarcha erosa vermiculata	Timarchini	No
AJ622025	Timarcha erosa vermiculata	Timarchini	No
EF484408	Callosobruchus chinensis	Bruchini	Yes
EF484418	Callosobruchus maculatus	Bruchini	Yes

[a]Species used in this study accessed from GenBank between 8/2008 and 6/2009.

Table 2.1. Aedeagus[a] length (mm) and width (μm) morphometrics for *Altica litigata* and *Altica* species (Refer to Fig.1.2)

Altica litigata		***Altica* species**	
Length (mm)	Width (μm)	Length (mm)	Width (μm)
1.787	367.5	1.319	306.6
1.949	373.7	1.228	320.7
1.872	349.3	1.269	306.5
1.831	383.7	1.253	306.5
1.744	397.9	1.240	288.3
1.732	385.7	1.301	304.5
1.778	371.5		

[a]The aedeagus from *Altica litigata* and *Altica* species were digitally measured from seven *A. litigata* and six *Altica* sp. specimens using the measurement tool within the Zeiss 1450EP SEM operating program (Refer to text).

Table 2.2. Aedeagus morphometrics. The aedeagus from *Altica corni, Altica sublicata* and *Lysathia ludoviciana* species were digitally measured using the measurement tool within the Zeiss 1450EP SEM operating program.

Altica corni		***Altica sublicata***		***Lysathia ludoviciana***	
Length	Width	Length	Width	Length	Width
1.407	361.4	1.636	335.0	1.758	282.4
1.346	324.8	1.833	379.9	1.632	270.2
1.382	351.8	1.829	349.3	1.713	270.2
				1.642	290.5
				1.561	266.1
				1.669	259.9

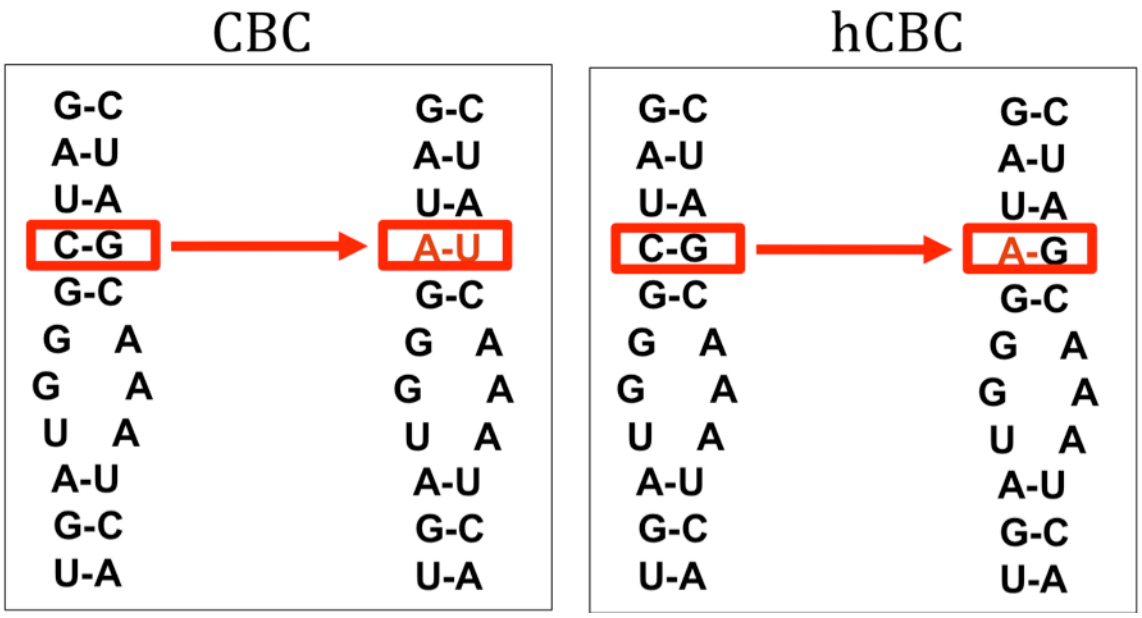
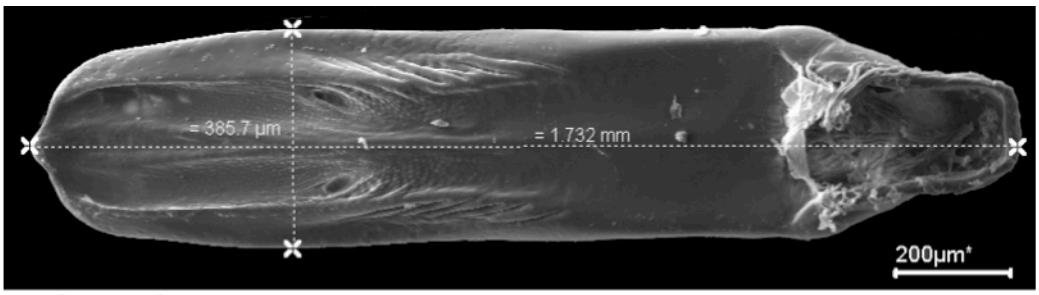
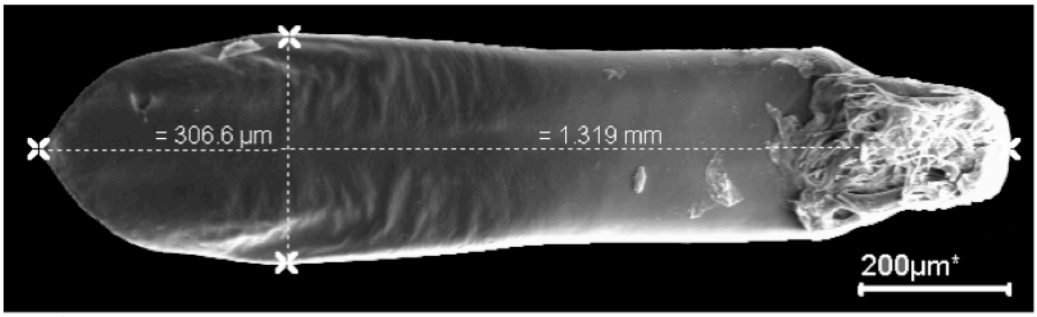


Figure 1.1. Visual depiction of compensatory base changes (CBCs). Cartoon depicts a conserved helix segment from the internal transcribed spacer 2 in which a CBC and a hemi-CBC are demonstrated; left and right respectively.



Altica litigata



Altica species

Figure 1.2. Male intromittent organs (aedeagus). Scanning electron microscopy (SEM) examination revealed morphological differences (shape, mean length, and mean width) between *Altica litigata* and *Altica* species. Both aedeagus in this figure, and those listed in Tables 2.1-2.2, were taken from specimens identified as *A. litigata* by external morphology.

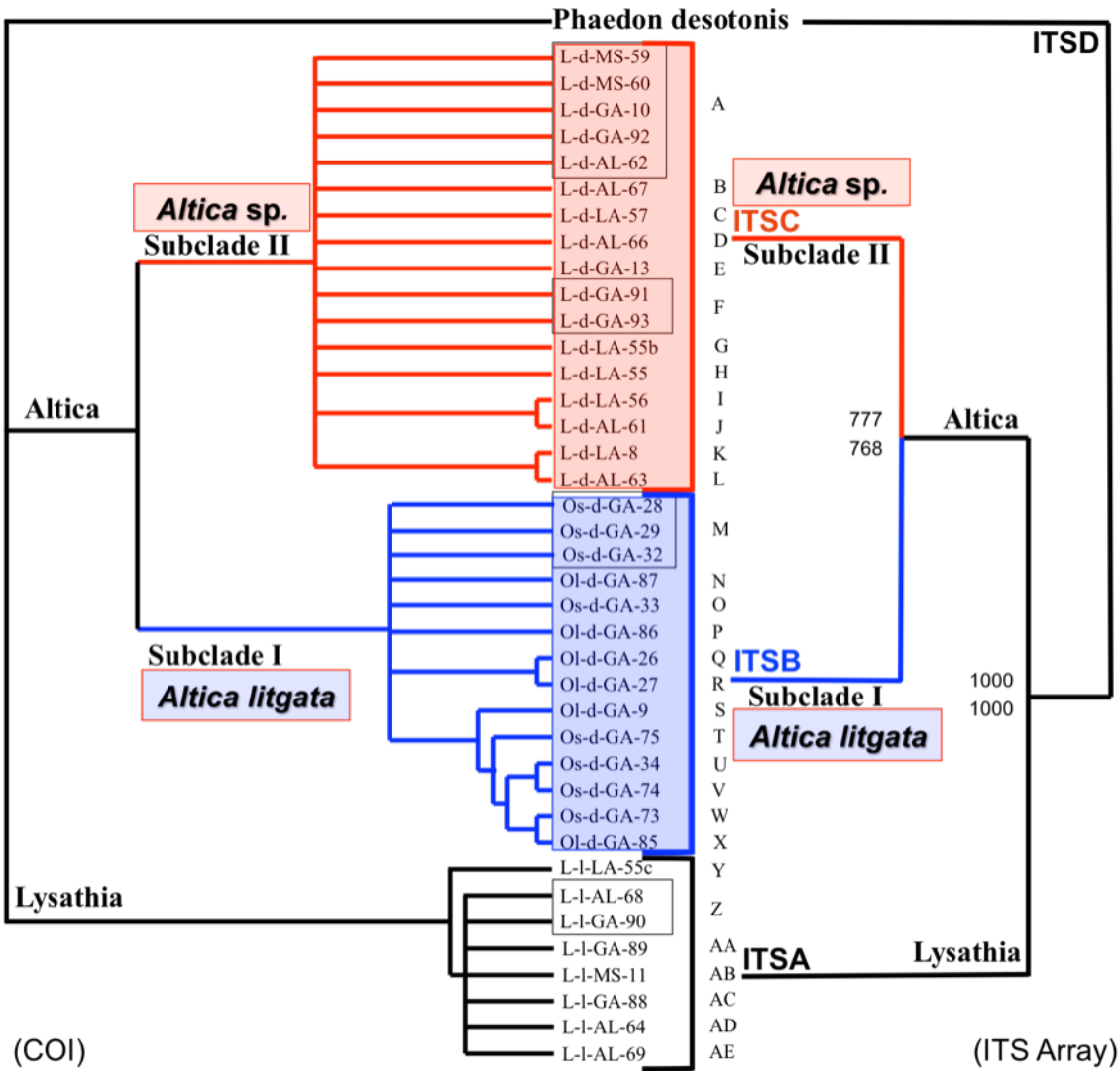


Figure 2.1. COI, COII, and ITS Array tree (adapted with permission from Jenkins et al., 2009a). Tree shows *Altica litigata* and *Altica* sp. as two distinct subclades.

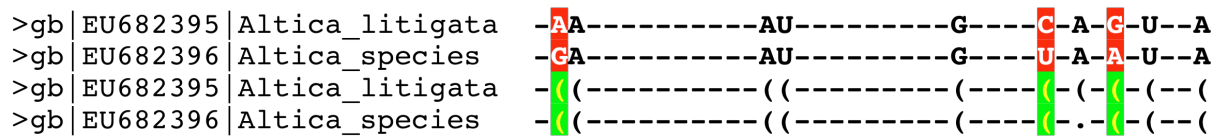


Figure 2.2. ITS2 synchronous sequence and structural alignment. Shaded red areas represent sequence variation while areas shaded in green represent the associated structural variation. Structural alignment is annotated in bracket-dot-bracket notation. The figure graphically depicts that the rDNA secondary structure is more conserved than its associated sequence.

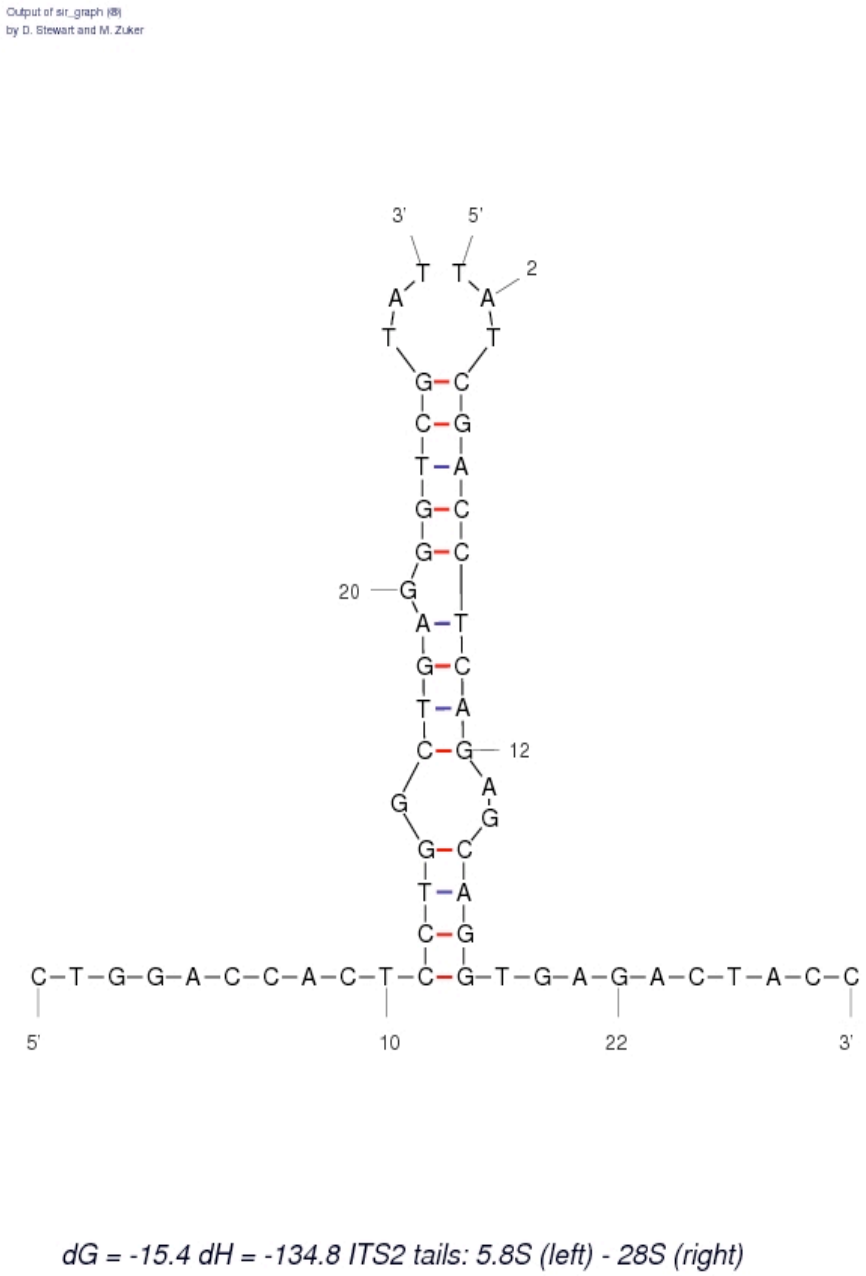


Figure 2.3. 5.8S & 28S rRNA interaction and HMM-based ITS2 annotation of *Altica litigata* (gb|EU682395). Annotation shows interaction of 25 nucleotides from the 5' end of 5.8S rRNA sequence with 25 nucleotides from the 3' end of 28S rRNA sequence.

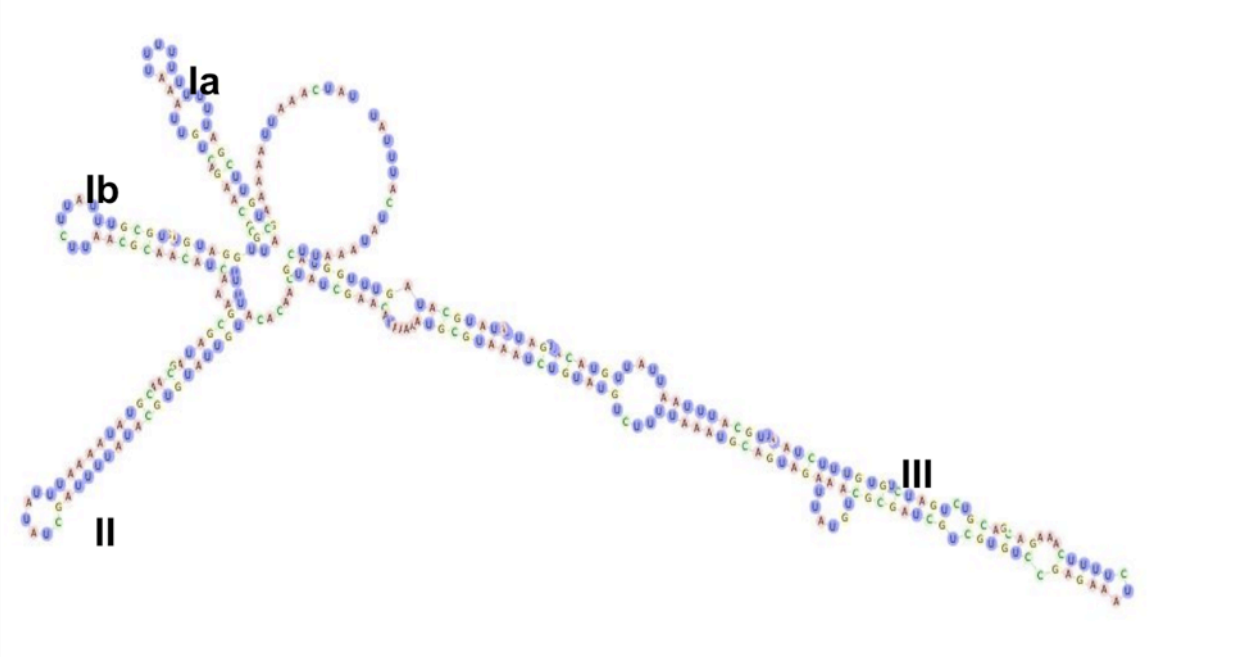


Figure 2.4. Internal transcribed spacer 2 (ITS2) rDNA secondary structure of *Altica litigata* (gb|EU682395). Roman numerals represent the helix number on the ITS2 secondary rDNA structure. *Altica litigata*, like many species, <1% (only 7 of 709 accessions in GenBank as of June 22, 2009) in the order Coleoptera, does not have a helix IV. As shown, however, the ITS2 rDNA secondary structure has an additional helix between II and I (labeled Ib), and therefore maintains a structure with four helices.

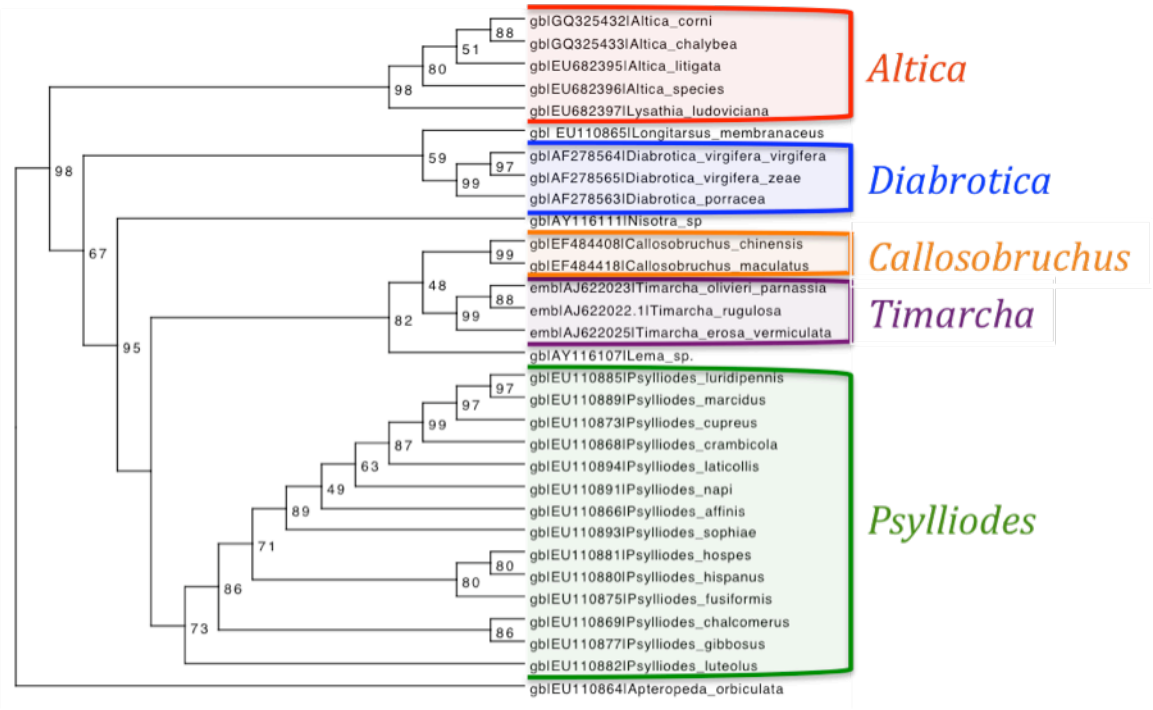


Figure 2.5. Profile Neighbor-Joining (PNJ) tree obtained by ProfDistS, synchronously calculated on ITS2 sequence and secondary structure information, using an ITS2 specific general time reversible (GTR) substitution model. (Refer to section 2.3.6). Bootstrap support values from 1000 pseudo-replicates.

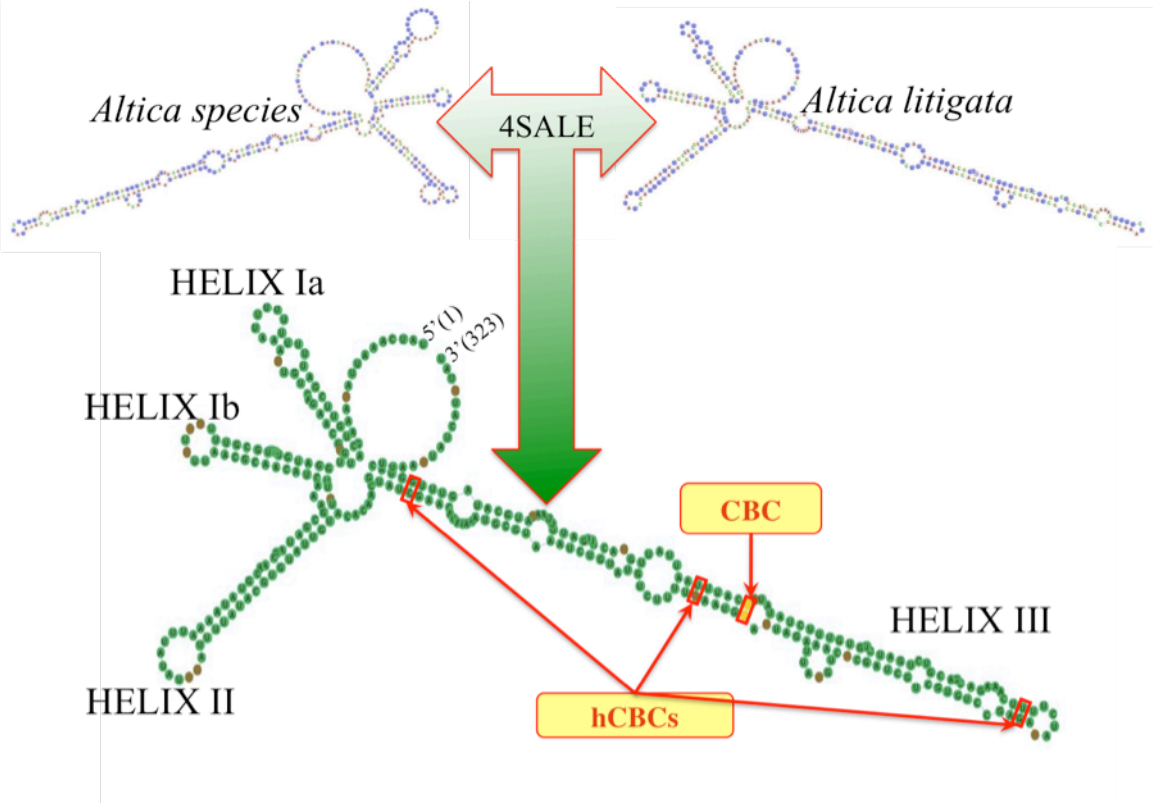


Figure 2.6. Analysis showing presence of compensatory base changes (CBCs) in conserved helix III of the Internal transcribed spacer 2 (ITS2) rDNA secondary transcript structure. Annotated structure is the ITS2 consensus of *Altica litigata* and *Altica* species. Green areas represent highly conserved base pairs. One significant CBC was found in the conserved helix III. Three additional hemi-CBCs were found within the conserved helix III, thus, providing further evidence for distinguishing species.

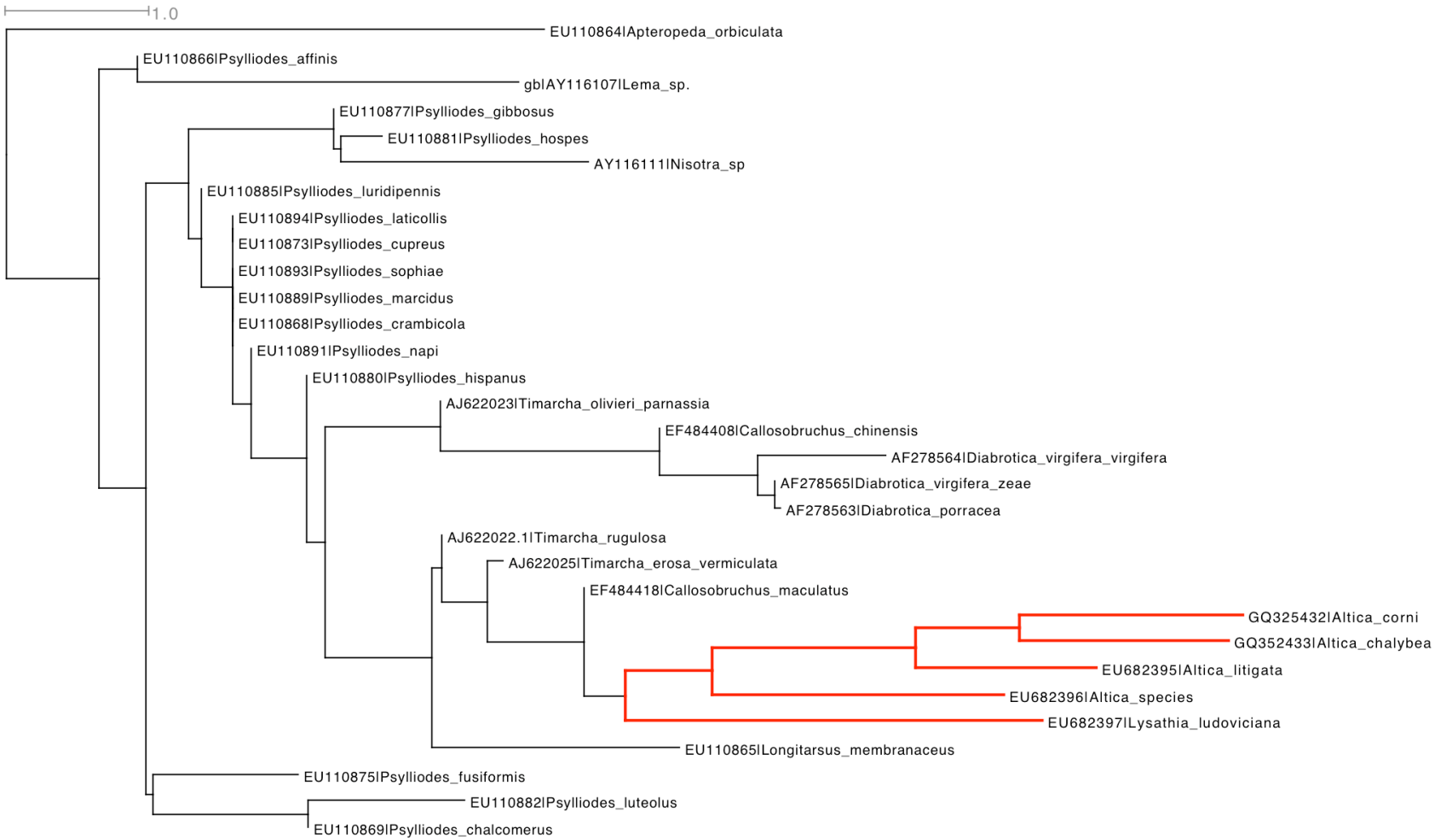


Figure 3.1. Phenogram based on compensatory base changes (CBCs) (Refer to section 3.1). While the *Altica* clade is concordant with other trees in this study, the basis in which the tree is built is vastly different. It uses CBCs as character state changes. Branch lengths are representative of the number of CBCs between each species.

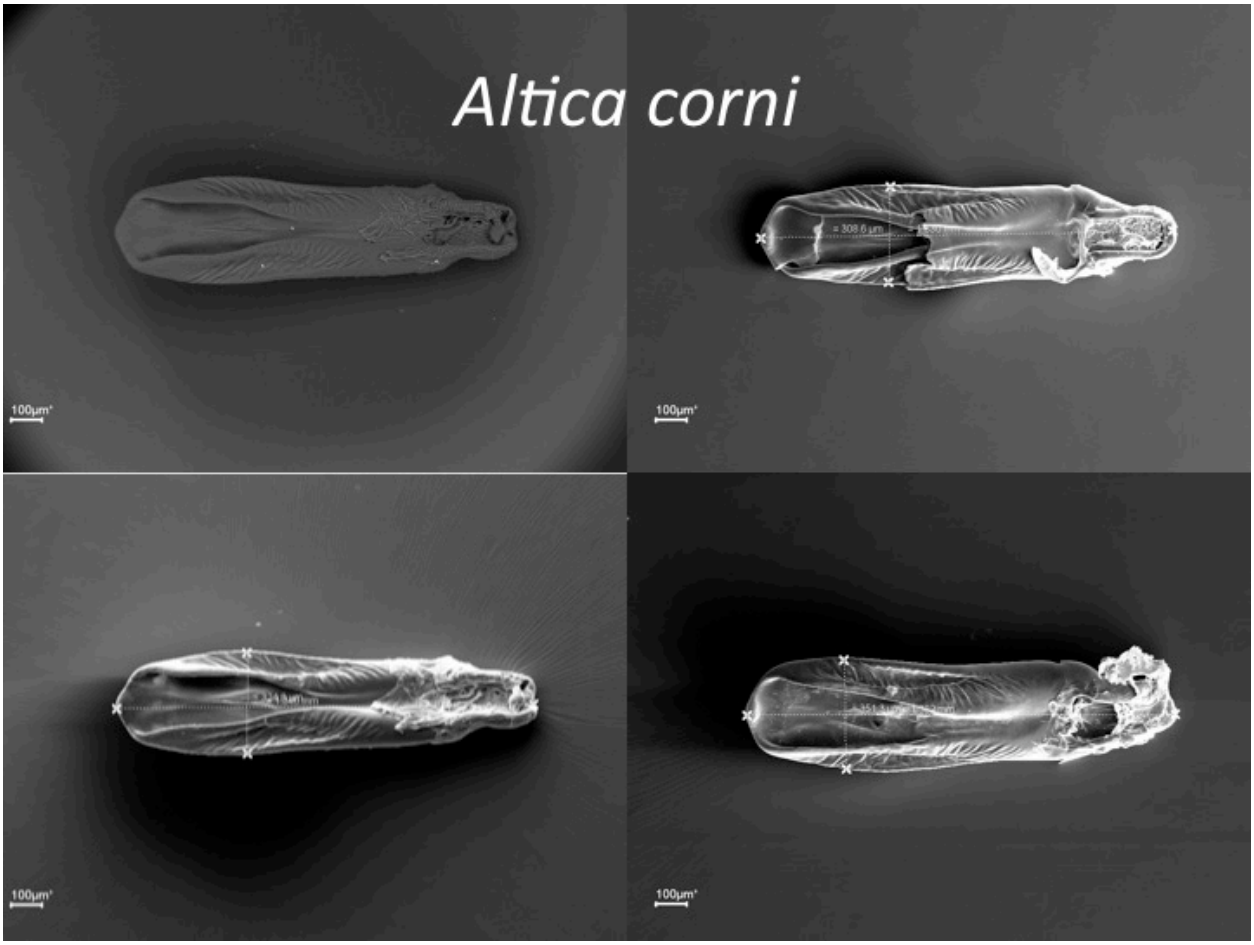


Figure 3.2. Male intromittent organs (aedeagus) of *Altica corni* viewed by scanning electron microscopy (SEM).

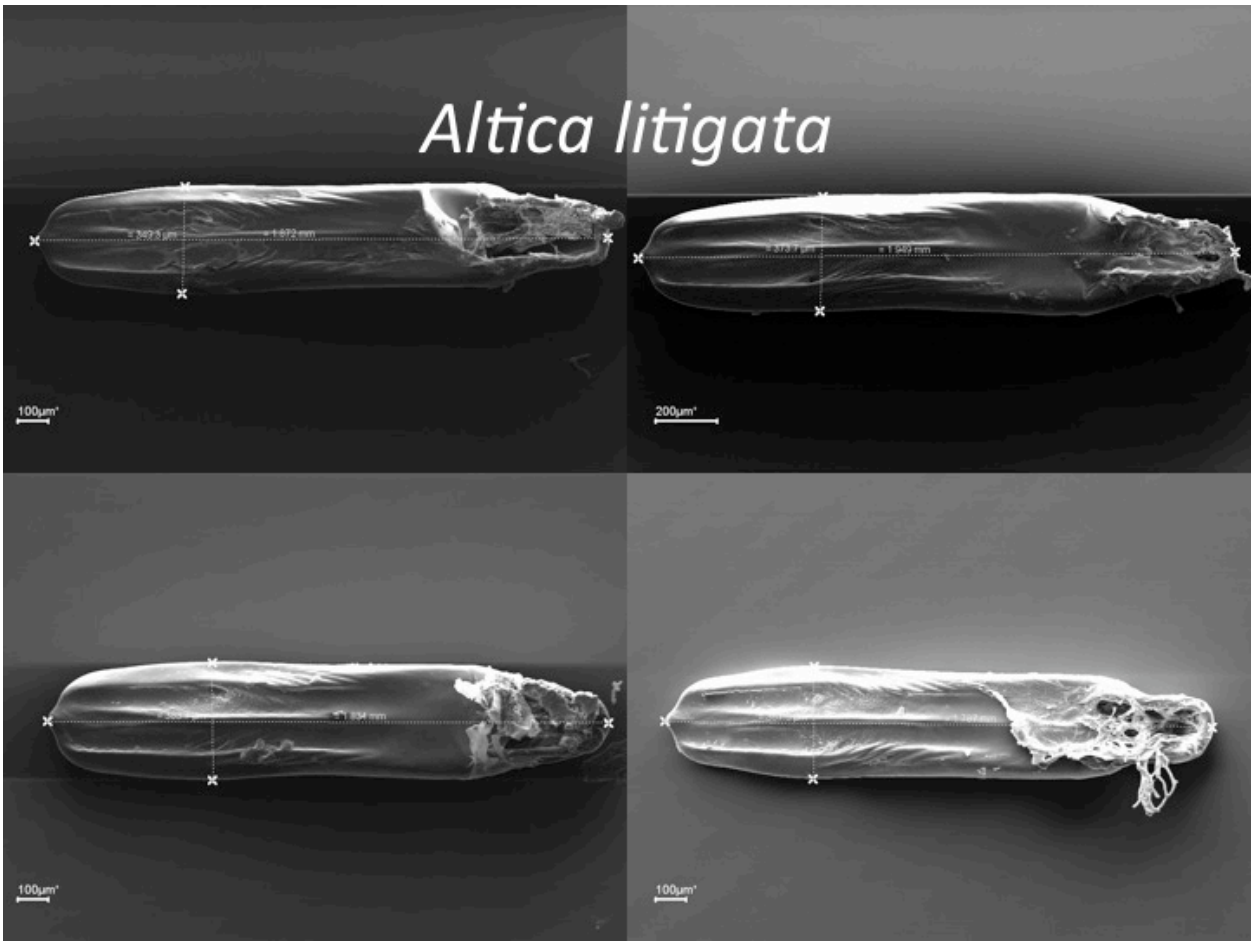


Figure 3.3. Male intromittent organs (aedeagus) of *Altica litigata* viewed by scanning electron microscopy (SEM).

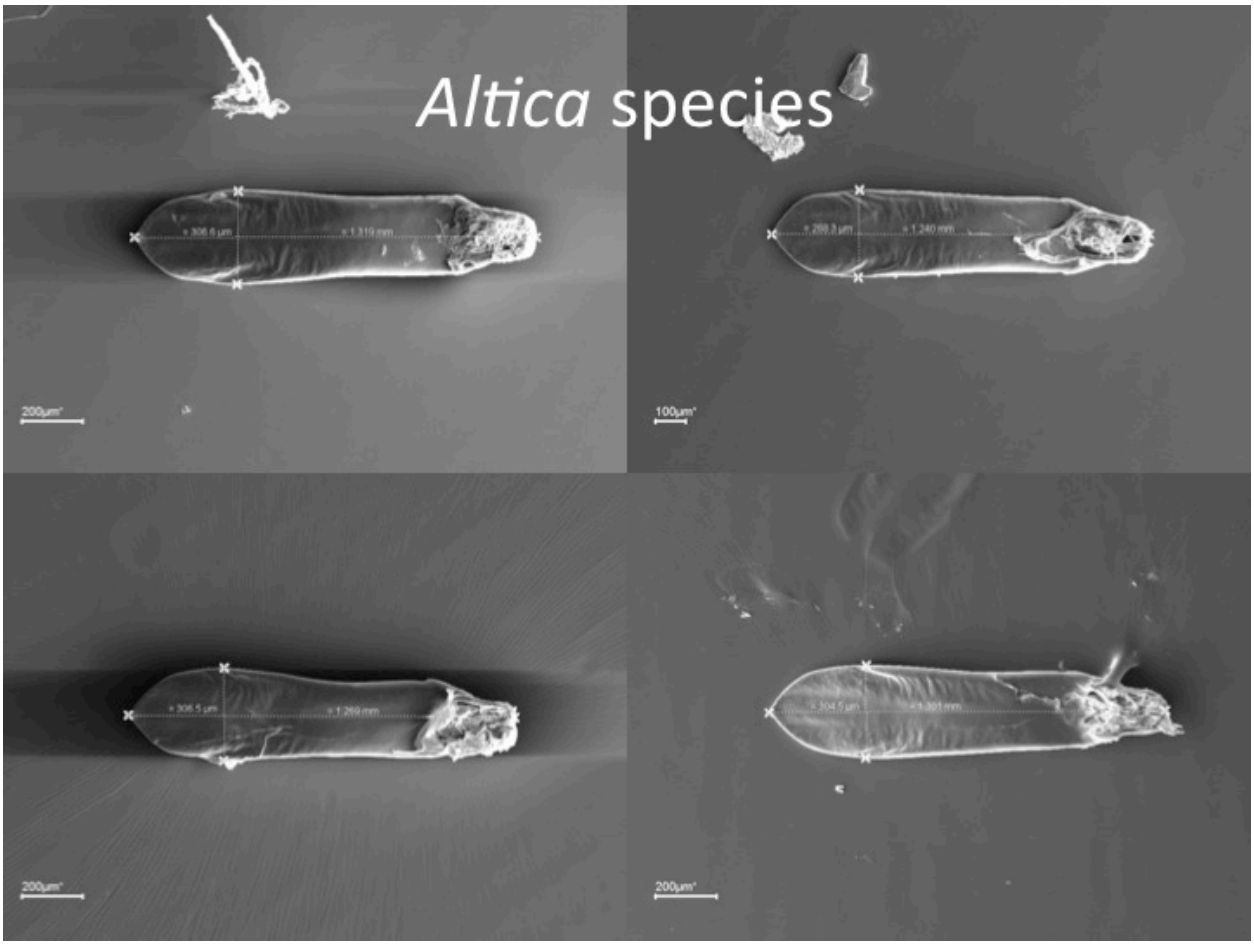


Figure 3.4. Male intromittent organs (aedeagus) of *Altica* species viewed by scanning electron microscopy (SEM).

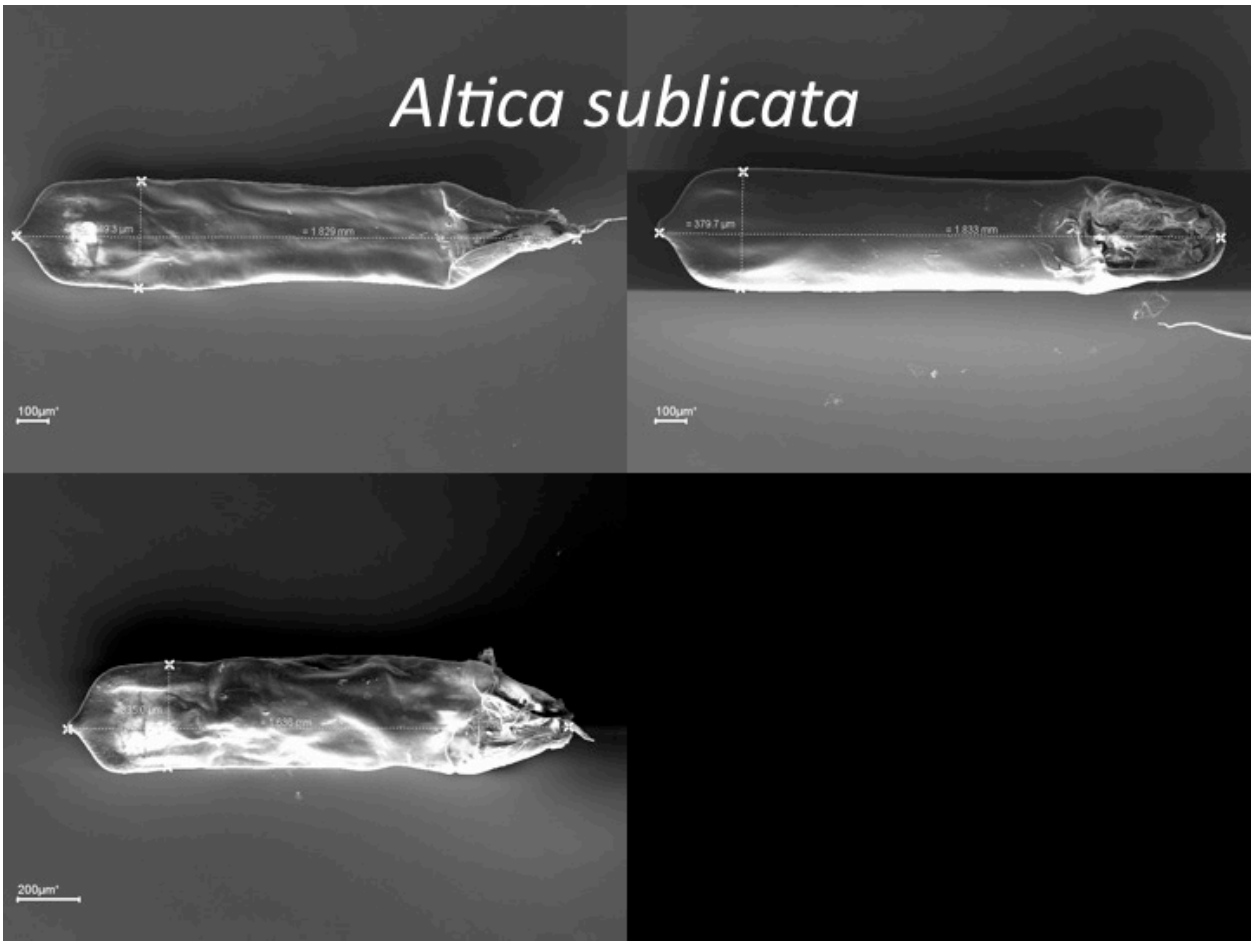


Figure 3.5. Male intromittent organs (aedeagus) of *Altica sublicata* viewed by scanning electron microscopy (SEM).

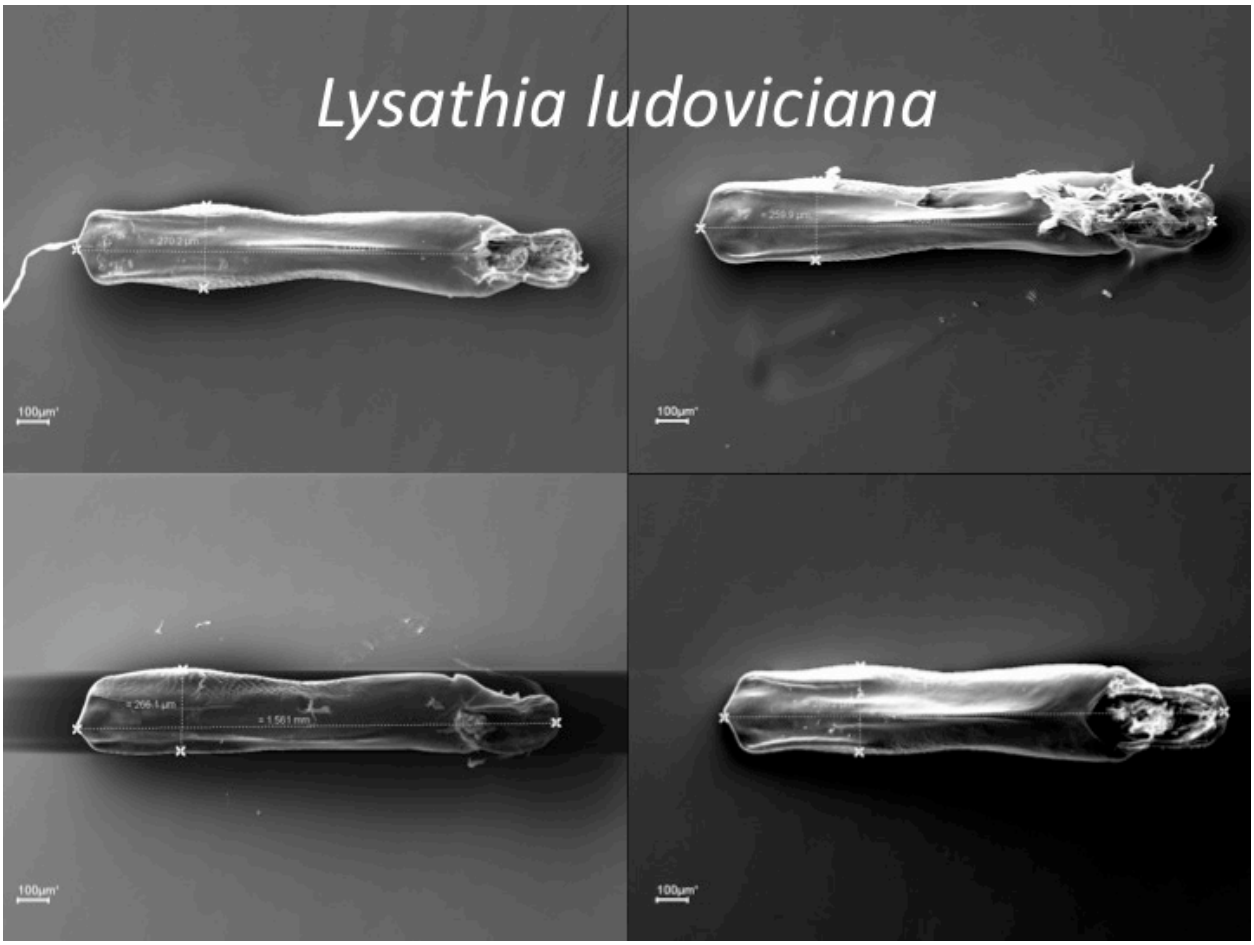


Figure 3.6. Male intromittent organs (aedeagus) of *Lysathia ludoviciana* viewed by scanning electron microscopy (SEM).

REFERENCES

Benson, D.A., et al., 2008. GenBank. Nucleic Acids Res 36, D25-30.

Cabrera, R.I., et al., 2008. Differential Resistance Among Crape Myrtle (Lagerstroemia) Species, Hybrids, and Cultivars to Foliar Feeding by Adult Flea Beetles (Altica litigata). HortScience 43, 403-407.

Caccone, A., et al., 2004. Extreme difference in rate of mitochondrial and nuclear DNA evolution in a large ectotherm, Gal·pagos tortoises. Mol. Phylogenet. Evol. 31, 794-798.

Coleman, A.W., 2003. ITS2 is a double-edged tool for eukaryote evolutionary comparisons. Genet. 19, 370-375.

Coleman, A.W., 2007. Pan-eukaryote ITS2 homologies revealed by RNA secondary structure. Nucl. Acids Res. 35, 3322-3329.

Coleman, A.W., 2009. Is there a molecular key to the level of “biological species” in eukaryotes? A DNA guide. Mol. Phylogenet. Evol. 50, 197-203.

Côté, C.A., et al., 2002. Dynamic conformational model for the role of ITS2 in pre-rRNA processing in yeast. RNA 8, 786-797.

Fernandez, P., Hilker, M., 2006. Host plant location by Chrysomelidae Basic and Appl. Ecol. 8, 97-116.

Friedrich, J., et al., 2005. ProfDist: a tool for the construction of large phylogenetic trees based on profile distances. Bioinformatics 21, 2108-2109.

Huson, D., et al., 2007. Dendroscope: An interactive viewer for large phylogenetic trees. BMC Bioinformatics 8, 460.

Jenkins, T., et al., 2007. Phylogeography used to illuminate maternal origins of exotic invasions of Coptotermes gestroi (Isoptera: *Rhinotermitidae*). Mol. Phylogenet. Evol. 42, 612-621.

Jenkins, T.M., et al., 2009a. Insights into Flea Beetle (Coleoptera: Chrysomelidae: Galerucinae) Host Specificity from Concordant Mitochondrial and Nuclear DNA Phylogenies. Ann. Entomol. Soc. Am. In press.

Jenkins, T.M., et al., 2009b. *Altica litigata* (Coleoptera: Chyrsomelidae: Galerucinae): A DNA Approach to Species Verification. Res. on Chrysomelidae Vol. II. In Press.

Joseph, N., et al., 1999. Ribosomal internal transcribed spacer 2 (ITS2) exhibits a common core of secondary structure in vertebrates and yeast. Nucleic Acids Res. 27, 4533-4540.

Keller, A., et al., 2009. 5.8S-28S rRNA interaction and HMM-based ITS2 annotation. Gene 430, 50-57.

Laroche, A., et al., 1996. Are Altica carduorum and Altica cirsicola (Coleoptera: Chrysomelidae) Different Species? Implications for the Release of A. cirsicola for the Biocontrol of Canada Thistle in Canada. Biol. Cont. 6, 306-314.

LeSage, L., 1995. Revision of the costate Species of *Altica Muller* of North-America north of Mexico (Coleoptera, Chrysomelidae). Can. Entomol. 127, 295-411.

LeSage, L., 2000. On the type series of *Altica chalybaea* (Coleoptera : Chrysomelidae). Entomol. News 111, 233-237.

LeSage, L., 2002. Flea Beetles of the Genus *Altica* Found on Grape in Northeastern North America (Coleoptera: Chrysomelidae). J. Entomol. Soci. Ont. 133, 3-46.

Mathews, D.H., et al., 2004. Incorporating chemical modification constraints into a dynamic programming algorithm for prediction of RNA secondary structure. Proc. Nat. Acad. Sci. of the U.S.A. 101, 7287-7292.

Muller, T., et al., 2007. Distinguishing species. RNA 13, 1469-1472.

Pettis, G.V., et al., 2004. Potential Resistance of Crape Myrtle Cultivars to Flea Beetle (Coleoptera: Chrysomelidae) and Japanese Beetle (Coleoptera: Scarabaeidae) Damage. J. Econ. Entomol. 97, 981-992.

Pettis, G.V., Braman, S. K., 2007. Effect of temperature and host plant on survival and development of Altica litigata fall. J. Entomol. Sci. 42, 66-73.

Phillips, W.M., 1979. A contribution to the study of species relations within the chrysomelid genus *Altica müller* in Britain. Zool. J. Linn. Soc. 66, 289-308.

Schultz, J., et al., 2005. A common core of secondary structure of the internal transcribed spacer 2 (ITS2) throughout the Eukaryota. RNA 11, 361-364.

Schultz, J., et al., 2006. The internal transcribed spacer 2 database-a web server for (not only) low level phylogenetic analyses. Nucleic Acids Res 34, W704-W707.

Schultz, J., Wolf, M., 2009. ITS2 Sequence-Structure Analysis in Phylogenetics: A How-to Manual for Molecular Systematics. Mol. Phylogenet. Evol. In Press, Corrected Proof.

Seibel, P.N., et al., 2006. 4SALE- a tool for synchronous RNA sequence and secondary structure alignment and editing. BMC Bioinformatics 7, 11.

Seibel, P.N., et al., 2008. Synchronous visual analysis and editing of RNA sequence and secondary structure alignments using 4SALE. BMC Res Notes 1, 91.

Selig, C., et al., 2008. The ITS2 Database II: homology modelling RNA structure for molecular systematics. Nucleic Acids Res 36, 377.

Wolf, M., et al., 2005a. Homology modeling revealed more than 20,000 rRNA internal transcribed spacer 2 (ITS2) secondary structures. RNA 11, 1616-1623.

Wolf, M., et al., 2005b. CBCAnalyzer: inferring phylogenies based on compensatory base changes in RNA secondary structures. In Silico Biol 5, 291-294.

Wolf, M., et al., 2008. ProfDistS: profile-distance based phylogeny on sequence-structure alignments. Bioinformatics 24, 2401-2402.