1-1-2002

Phylogenetic Structure of Two Central Mexican Centruroides Species Complexes

William Ian Towler
ian@itowler.com

PHYLOGENETIC STRUCTURE OF TWO CENTRAL MEXICAN *CENTRUROIDES* SPECIES COMPLEXES

Thesis submitted to
The Graduate College of
Marshall University

In partial fulfillment of the
Requirements for the degree of
Master of Science
Biological Sciences

Submitted by

William Ian Towler

Marshall University

May 1,2002

This thesis was accepted on _____
Month Day Year

as meeting the research requirements for the master's degree.

Advisor_____

Department of _____

_____
Dean of the Graduate College

ABSTRACT

PHYLOGENETIC STRUCTURE OF TWO CENTRAL MEXICAN CENTRUROIDIES SPECIES COMPLEXES

By William Ian Towler

Central Mexico is home to numerous species of highly toxic *Centruroides* scorpions. Two species complexes *C. infamatus,* (C.L. Koch, 1844), and *C. limpidus* (Karsch, 1879) typify the complex relationships that exist between and within the complexes. Their existing taxonomic status is based on morphological features such as coloration and morphosculpture. A complete and modern study of these scorpions does not exist, and is needed. In an attempt to clarify the status and relationship between these complexes we initiated a molecular based approach applying mitochondrial gene markers (16S and CO1). This study confirms two divergent clades within *C. infamatus*; divergence rate estimates their common ancestor's age as 2-4 Ma for HKY+G+I divergence rate (11.7 ± 0.9 %) and 3-5 Ma for uncorrected *p* (7.2 ± 0.4 %). Further study is necessary with sampling all over the range of both taxa, to confirm existence of two independent study also suggests that more than one ancient monophyletic lineage (possibly, more than one species) exist within currently accepted *Centruroides limpidus limpidus.* The type locality of *Centruroides limpidus* is Puebla, which lies in the same geographic area as Guerrero. Thus, we might assume that the Querétaro/Guerrero lineage corresponds to 'true' *C. limpidus*, and that the Balsas Depression populations could belong to another, 'cryptic', or 'sibling' species. Further, detailed investigations should be done to test these preliminary conclusions: the need for many more populations from the entire range of *C. limpidus* is needed. Several data sets (mitochondrial and nuclear genes, allozymes, morphology, toxin structure/activity, etc.) could be analyzed to establish the true taxonomic and genetic structure of the populations and species of *Centruroides*.

DEDICATION

I would like to dedicate this to my parents, Rick and Karen Towler. My parents have been supportive of my academic pursuits from the beginning, and continue to do so. Without their love, understanding and support, this thesis would not be possible. Thanks mom and dad, this is dedicated to you.

ACKNOWLEDGMENTS

I would like to thank Dr. Evans and Dr. Little for their time and for the time spent serving on my committee. I would also like to acknowledge the NASA Space Grant Foundation for their generous support of the work in this thesis. I would also like to extend thanks to Dr. Joy for the use of his camera and microscope. Travis Baley and Harold Payton deserve thanks for the time spent in the lab. Many thanks extended to Mr. Javier Ponce Saavedra for the collection of the specimens used in this study. Sherrine Ibrahim, my future wife, deserves many thanks for her patience and commitment. I would also like to thank Dr. Sheldon Guttman for not giving up on me as an undergraduate student who was struggling.

Lastly, I would like to thank Dr. Victor Fet for his guidance and eagerness to teach and explain. Dr. Victor Fet served as an excellent source of inspiration and knowledge; he has served as a professor and a friend, and for that I thank him.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER I

Introduction

The Species Debate

What is a species? Few questions asked in biology today stir up more intense feelings than, the 'what is a species' question. Charles Darwin once stated "*No one definition has satisfied all naturalists; yet every naturalist knows vaguely what he means when he speaks of a species…In determining whether a form should be ranked as a species or as a variety, the opinion of naturalists having sound judgment and wide experience seems the only guide to follow"* (Darwin, 1859).

In 1758, Linnaeus defined the fundamentals of modern classification as we know it today. The system was based around the type concept of species, and stated the species was the lowest category of nomenclature, but did not explicitly define 'species' and what a 'species' entails. In this classification system, different species are grouped together to form genera. The genera then group together to make up families, families form orders, and so on. In 1940, Mayr modified the concept of the Linnaean species and introduced the idea of a biological species, after having found polytypic species, or species which are composed of several subspecies or geographical variants. The idea of a biological species stated that "a species are a group of actually or potentially interbreeding natural populations, which are reproductively isolated from other such groups" (Mayr, 1942). With introduction of the biological species concept came the question of subspecies. As defined by Mayr and Ashlock (1991), a subspecies is "an aggregate of local populations of a species inhabiting a geographic subdivision of the range of the species and differing taxonomically from other populations of the species." According to the International Code of Zoological Nomenclature, the subspecies is the lowest recognized category and is still in use (Jeffrey, 1973). Even though the subspecies stands as a valid taxonomic unit, it presents some problems in the usage of this unit. With the use of Mayr's

Biological species model as the definition of subspecies, how do you "draw the line" on what a subspecies is; how do you differentiate a zone of intergradation from a zone of hybridization? Because of this confusion many biologists and systematists would like to remove subspecies as a valid taxonomic category as it does not offer any utility in relation to questions of evolutionary processes and patterns. When talking about the systematics of organisms you inevitably end up talking about evolution as well. Evolution, no matter how you choose to look at the process, is just that, a process. The process of evolution is continuous and never ending. With that in mind we must remember that as we collect our data we are only getting a "snapshot in time" of this process of evolution. The species and subspecies debate continues, and recent advances in molecular biology have added to the debate. Since all living organisms contain DNA as their molecule that serves as a "blueprint for life", we can study the relationship of organisms by comparing their DNA using sequence analysis. This molecular approach has several advantages over the classical morphological approach. The first advantage is that DNA comparison, via sequence analysis, can be used to compare any organism that was living at one time. Secondly, since the evolutionary change of DNA follows a more or less regular pattern, we can use quantitative analysis of these changes, to compare the organisms.

Advances in DNA Technology

In the spring of 1983, Kary Mullis, an employee of Cetus Corporation, conceived the idea for the polymerase chain reaction or PCR, and changed the study of systematics forever (Mullis, 1990). The process of PCR allows users to target a specific section of DNA and make millions of copies in a matter of hours using an enzyme named DNA polymerase. Although there are several different DNA polymerases available for use, *Taq* polymerase is the most commonly used. *Thermus aquaticus* is a species of bacteria that lives in a hot spring, and therefore has a DNA polymerase that is able to withstand the high temperatures that are routinely reached in a PCR reaction. The PCR process employs the use of a machine called a thermocycler. The thermocycler is nothing more than a sophisticated heating block which changes the temperature of the block according to the

conditions preset by the experimenter via a computer. The standard PCR reaction consists of many cycles (anywhere from 15 – 40 cycles), each cycle having three different parts. The first part of a cycle is called denaturing. During denaturing, the double stranded DNA template (or target DNA) becomes single stranded due to the high temperature at this step. The second step is called annealing. The temperature is dropped and the primers (the oligonucleotides that target specific genes) find and stick to their corresponding sequences in the single stranded DNA template. In the last step of the process, the temperature is raised slightly to activate the enzyme *Taq* polymerase, which extends the sequence off of the primer in the 5' direction. This process is repeated over and over to produce the millions of copies. The process won Kary Mullis the Nobel Prize in 1993, and has changed biology and systematics forever. With PCR, systematists are able to target genes for amplification and sequence these genes directly and use these sequences as characters in the development of cladograms or phylogenetic trees. To obtain the sequence of a given target, the product of the PCR reaction is used in a process called the Sanger method of sequencing, which uses fluorescently labeled nucleotides. These labeled nucleotides are then read by a laser that is tunable to the specific wavelength of light emitted by these fluorescently labeled nucleotides (Hillis, 1996). This information is read after a polyacrylamide gel electrophoresis (PAGE) is performed. One reaction is loaded into one lane for a total of four lanes per sequence. In PAGE, the shortest fragments will migrate the farthest. Therefore, the bottom-most band indicates that its particular dideoxynucleotide was added first to the labeled primer.

DNA Systematics

During the early 1990's mitochondrial genes were the target of choice for phylogenetic studies. The success of the mitochondria genes is due in a large part because they have their own DNA separate from the nucleus, the ability to have multiple mitochondria per cell, and they are found in high abundances in muscle tissue. Multiple mitochondria per cell, makes extraction of the mitochondrial DNA (mtDNA) much easier than extracting nuclear DNA. The mitochondrial DNA found in animals, with one exception, is a duplex covalently closed circular molecule. This molecule in animals is completely

uninterrupted and the sequence contains no intronic regions (Moritz *et al.*, 1987). In almost all animal systems, the mtDNA is inherited buy the offspring via the female, and there is virtually no recombination. Due to the low amount of recombination, the mtDNA molecule contains highly conserved gene arrangement and product functionality across animals. The mitochondrial genome contains two ribosomal RNA genes, 22 transfer RNA genes, and 13 genes that are protein coding (Avise, 1994). The mitochondrial genes most often used in phylogenetic studies are 16S, 12S, ND4, COI and COII. The 16S ribosomal RNA gene has been used before in scorpions (Fet *et al.*, 1999), and is widely used in modern molecular evolutionary studies of various arthropods (Simon *et al.*,1994). The additional information that is provided by molecular data allows systematists, a new tool to look at organisms that have previously been hard to study due to fixed or conservative morphology. Molecular approaches to systematics have begun to play an important role in looking at the evolutionary history of scorpions.

Scorpions

Scorpions date back about 450 million years ago, where their appearance is first seen in the middle Silurian (Polis, 1990). Eurypterida, or the water scorpions, most probably served as the starting point for scorpion evolution (Savory, 1977). Marine and amphibious scorpions persisted well into the Carboniferous (250 – 300 million years ago). The first unmistakably terrestrial scorpion is *Palaeopisthacanthus*, which appeared in the Upper Carboniferous, and the stigmata are preserved (Rolfe, 1980). Today's scorpions are generally similar in appearance to the Paleozoic forms. With exception to the changes due to locomotion and breathing in a terrestrial environment, the basic body plan is externally similar to the scorpions that lived 425 million years ago (Polis, 1990). Today, scorpions can be found on all continents except Antarctica. Scorpions have come to occupy all nonboreal habitats, including deserts, temperate forests, rain forests, tropical forests, savanna, grasslands, the intertidal zone, and snow covered mountains with elevations over 5,500 meters. In some habitats, the densities of scorpions can be as high as 2-12 per square meter (Polis, 1990). In ideal habitats, as many as 13 species can be found sympatrically, and in most areas three to six species are common (Polis, 1990).

The success of these animals can be explained by their ability to adapt, and make use of habitats that are not commonly used by other organisms. Many scorpions spend 92-97 percent of their time inactive in burrows, allowing them to have very low metabolic rates, among the lowest arthropod rates ever recorded (Polis, 1990). While still primitive in morphology, scorpions do exhibit some advanced characteristics like their sensory organs. The scorpion eye is among the most sensitive of the arthropod eyes and may allow them to navigate and orient themselves using starlight (Fleissner, 1977). Scorpions have an extremely varied diet including insects and various other arachnids. As a very old group, scorpions have survived over 400 million years by incorporating a mix of both primitive and advanced characteristics. They have shown a great deal of plasticity in physiology, behavior and response to environmental stresses rather than plasticity in gross morphology, this is the most probable reason for their persistence as a taxon over time. Scorpions now number 1270 species, over 159 genera, in 18 different families (Fet *et al*., 2000).

CHAPTER II

Review of Literature

Centruroides

Centruroides Marx, 1890, with 41 species and 24 subspecies (Fet & Lowe, 2000) is one of the most diverse genera of Buthidae (second only to *Tityus* C. L. Koch with 46 species and 22 subspecies), and the only genus of this family found in North America (Fet *et al*., 2000). *Centruroides* are only found in the New world, with a center of radiation in Central America ((Fet *et al*., 2000). They are abundant in various natural habitats ranging from tropical forest to temperate deserts. *Centruroides* is especially diverse in Mexico (Lourenço & Sissom, 2000) and the Caribbean (Armas, 1988), and is also found in Central and South America (Sissom & Lourenço, 1987). A few species within the genus are toxic, and potentially lethal to humans. In Mexico alone, scorpions of this genus caused an average of 1696 deaths per year for the twelve years reported by Mazzotti and Bravo-Becherelle (1963). Between the years of 1890 and 1926, the Mexican city of Durango (pop. 40,000) had an average of 45 deaths per year (Baerg, 1961), and yet this rate only ranked 14th among all the Mexican states (Mazzotti and Bravo-Becherelle,1963). Because of their impact on human health scorpions of the genus *Centruroides* are a group of scorpions that receive a great deal of attention. Lourenço and Sissom (2000) reviewed scorpion diversity in Mexico, emphasizing the need for more investigation and novel approaches (including DNA techniques) to understand complex taxonomy, origin, and distribution of Mexican *Centruroides*. They say (Lourenço and Sissom, 2000, p. 117): "...it is likely that some of these species will prove to be synonyms, and equally likely that others will represent complexes of sibling species rather than single species".

The genus *Centruroides* was first described in 1890, by G. Marx. A complete and modern study of the genus does not exist, and is much needed. The taxonomy of many *Centruroides* species is confusing, and has traditionally been based mainly on morphological characters such as coloration and morphosculpture, with the only existing key (Stahnke & Calos, 1977) being outdated due to current research on species with in the genus.

Centruroides infamatus infamatus

The first species complex consists of *Centruroides infamatus infamatus* (Koch, 1844) and *Centruroides infamatus ornatus* (Pocock, 1902). These subspecies can be differentiated from each other by coloration in the field. Geography is another defining character, but does not serve as the best character because these two subspecies have overlapping ranges. In this study the samples that represent *Centruroides infamatus infamatus* (*Cii*) were taken in Zumpimito, which is in the center of the country at 1560 meters above sea level. The samples that represent *Centruroides infamatus ornatus* (*Cio*), were taken further North and West from Zumpimitio, at four different sites. The sites together have an average elevation of 1795 meters above sea level, with the highest sight at 2020 meters above sea level. Figure 2.1 below, shows a distribution map for both species (Beutelspacher, 2000). This species complex is a mountainous species that lives in a transitional area between mountains and warm land, with tropical deciduous forest, thorn-shrubs and xerophytic vegetation at the lower lands over the Balsas Basin. In 2000, these two subspecies were suggested to be only forms of the species, and that they should be collapsed (Beutelspacher, 2000).

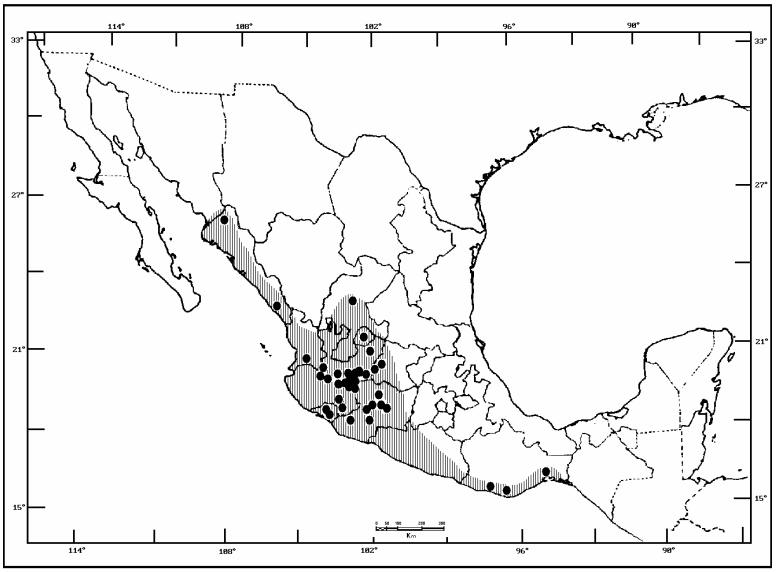


FIGURE 2.1 DISTRIBUTION MAP OF C. INFAMATUS INFAMATUS AND C. INFAMATUS ORNATUS, AS STATED BY BEUTELSPACHER IN, CATÁLOGO DE LOS ALACRANES DE MÉXICO, 2000.

Centruroides limpidus limpidus

The second species complex consists of *Centruroides limpidus limpidus* (*Cll*) (Karsch, 1879), and *Centruroides limpidus tecomanus* (*Clt)*(Hoffman, 1932) which are morphologically distant from *C. infamatus*. The data from existing litertature (Hoffman, 1932), these two subspecies can be differentiated by the presence or absence of a subacular tooth. This subacular tooth is just below the aculas on telson of the scorpion. Figure 2.2 below, shows a clear picture of this subacular tooth. The tooth can be found on *C. limpidus tecomanus*, but not as prevalent on *C. limpidus limpidus*. The use of the subacular tooth as a character to differentiate the two subspecies is not a good idea. In a later study, the subaculaer tooth was shown as a character that declines in size with an increase in age within the genus of *Centruroides* (Lourenco, 1982).

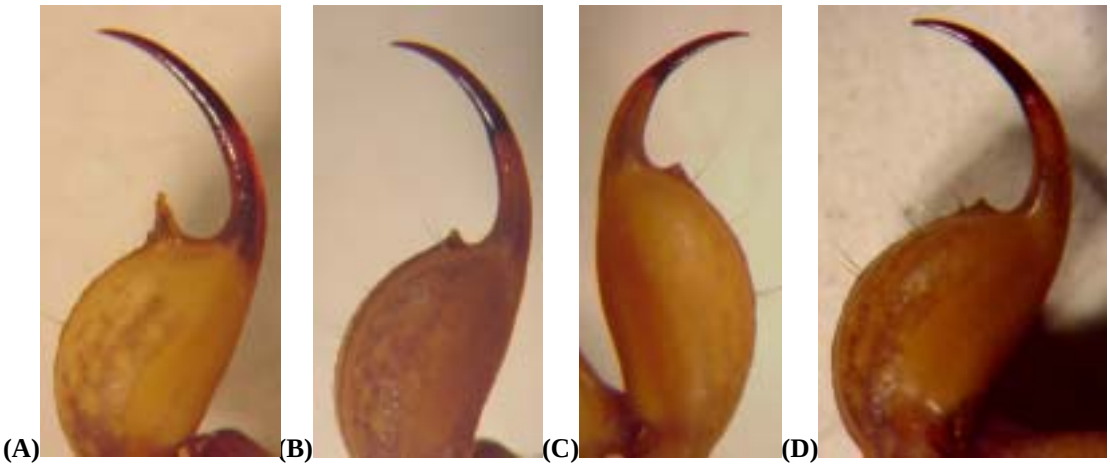


FIGURE 2.2 PICTURES OF THE SUBACULEAR TOOTH ON ALL FOUR SPECIES USED IN THIS STUDY. (A) C. LIMPIDUS TECOMANUS, (B) C. LIMPIDUS LIMPIDUS, (C) C. INFAMATUS INFAMATUS, AND (D) C. INFAMATUS ORNATUS. THE SUBACULEAR TOOTH IS MUCH MORE PROMINENT IN (A), THAN ANY OTHER SPECIES.

Another difference between the two subspecies is the geography. *C. limpidus limpidus* has a range that is more Southern and Western in comparison to that of *C. limpidus tecomanus.* The samples of *C. limpidus limpidus* represented in this study were all taken from the East side of the Sierra Madre del Sur Mountian range. The *C. limpidus tecomanus* samples used in this study were taken from the West side of the Mountian range, and were both near costal environments. Figure 2.3 below shows a distribution map of both *C. limpidus tecomanus* (open circles), and *C. limpidus limpidus* (filled circles) (Beutelspacher, 2000).

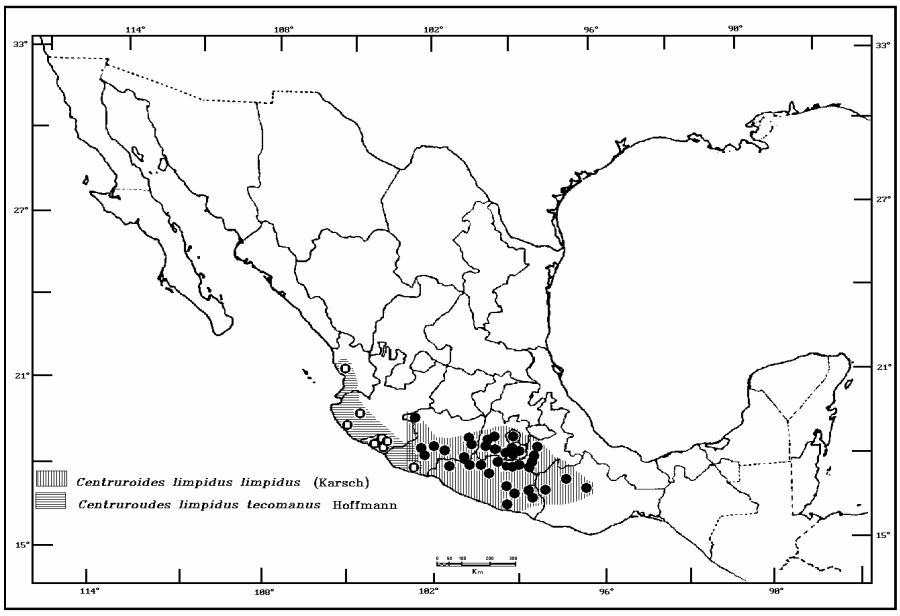


FIGURE 2.3 DISTRIBUTION MAP OF C. LIMPIDUS LIMPIDUS (FILLED CIRCLES) AND C. LIMPIDUS TECOMANUS (OPEN CIRCLES), AS STATED BY BEUTELSPACHER IN, CATÁLOGO DE LOS ALACRANES DE MÉXICO, 2000**.**

The Mexican deserts and mountains represent an ancient area of scorpion evolution. Although scorpions seem to preserve some of the most ancient arthropod features, they exhibit high genetic differentiation and active speciation (Fet *et al.*, 1999, 2001). *Centruroides* are one of a few extant scorpion genera for which Cenozoic fossils exist, from both the Miocene amber of the Dominican Republic (Schawaller, 1979) and the Miocene/Oligocene amber of Chiapas, Mexico (Santiago-Blay & Poinar, 1993). Thus, the age of the extant lineages can be very old, and their evolution could be considered against the geological events that occurred in the geographical area. The toxicological studies show that such speciation in a number of cases is paralleled by increased venom potency. Developing independent DNA-based phylogenies will allow us to investigate relatedness versus convergence of high toxicity within this diverse genus.

These characters are not conclusive for phylogenetic purposes. Using modern DNA-based methods of phylogenetic analysis (Maximum Parsimony, Maximum Likelihood, Neighbor-Joining algorithms; Gantenbein *et al.*, 1999; Fet *et al.*, 2001), we are able to get a better picture of what has happened in an organisms history.

Molecular Systematics

From the time of Charles Darwin, it has long been a dream for many biologists to reconstruct the evolutionary history for all organisms on the Earth, and express it in the form of a tree (Haeckel, 1866). The classical approach to this pursuit was to use fossils, but the fossil record presents some problems. The fossil record generally only preserves hard substances, and soft tissue is lost; rendering the fossil record incomplete and fragmented (Nei & Kumar, 2000). With the advances in molecular biology DNA has given us a new tool to construct these trees. The use of DNA based phylogenies requires the use of certain mathematical models or algorithms. In the past, the algorithms of choice in the past have been Neighbor Joining (NJ), Maximum Parsimony (MP), and Maximum Likelihood (ML) (Nei & Kumar, 2000). As the science of phylogenetics is advancing and becoming more refined, the tools are becoming more powerful and refined as well. This study will include a new arguably more powerful analysis tool called Bayesian Analysis. With the combination of these four methods, the data collected should reveal a good picture of the evolution of these scorpions.

Neighbor Joining

The neighbor joining (NJ) process is a distance matrix method used to compare the taxa. In a distance matrix method, evolutionary distances are computed for all pairs of taxa in the sample, and a phylogenetic tree is constructed from these distance values. The logic behind all distance matrix methods is that a pair of DNA sequences differing at 10% of their nucleotide sites are more closely related to each other than a pair differing 30%. The more time that has passed between two sequences diverged from a common ancestor, the more the sequences will differ. The NJ algorithm is very similar to another algorithm called the Minimum Evolution method (ME). Both the NJ and ME methods are distance

matrix methods, but the NJ method is used more frequently because it is more efficient, and produces the same tree topology the ME would. One of the important ideas with the NJ method is the idea of neighbors (Saitou & Nei, 1987). Neighbors are defined as two different taxa that are connected by a single node in an unrooted tree. It is important to note that the starting tree in the NJ method is unrooted and in a star formation where every taxa is connected to the center of the tree. From this star tree, the length of each branch on the tree (determined by the distance computed for all the pairs of taxa) is compared and neighbors are put together based on their branch lengths. The results of a NJ tree are often given a statistical value to show the confidence in the tree produced, this value is called the bootstrap value and is usually found the branch of the tree that the value corresponds with. The fact that you can put a statistical value to this tree makes it attractive and well accepted.

Maximum Parsimony

Originally developed for processing morphological characters, the Maximum Parsimony method has been in use since 1966 (Hennig, 1966). Since its introduction, the MP method for phylogenetic tree generation has been adapted to deal with molecular data (Fitch, 1971, Hartigan, 1973). A MP tree is based on the assumption that the tree that is most likely, is the tree that requires the fewest number of changes or steps, to explain the data in the sequences. When comparing the sequences, the MP method only uses variable sites for analysis. Sites in the sequence that do not vary between taxa are eliminated from the analysis. Sober states, “the less we need to know about the evolutionary process to make a phylogenetic inference, the more confidence we can have in our conclusions (Sober, 1988). MP methods have advantages over other methods because it is free from assumptions that are required in ML and NJ methods. Since every mathematical model used today is a crude approximation of reality, the model-free MP method may give more reliable trees than other methods (Miyamoto& Cracraft, 1991). Like the NJ method the MP method can be assigned bootstrap values to show the confidence in the tree produced.

Maximum Likelihood

The idea of using a Maximum Likelihood (ML) method for phylogenetic inference was first used in 1967 for gene frequency data (Cavalli-Sforza & Edwards, 1967). The ML method tries to infer an evolutionary tree by finding the tree that maximizes the probability of observing the data in that tree. The ML method is a well established statistical method (Nei & Kumar, 2000). When using any Maximum likelihood method, you have to have a model from which to work from. In phylogenetics, this model describes how the various rates and probabilities of a nucleotide substitution will occur. For example, the standard software used in for determining a model to use, tests 56 different models for the use in ML analysis. When the ML method is computing the probability of a given tree or topology, the accurateness of this computation is the based on how well the chosen model fits the data set. If a model is chosen that does not fit the data well, the tree will be flawed. Like the methods mentioned above, the ML method can be assigned bootstrap values to show the confidence in the tree produced.

Bayesian Analysis

The last method of analysis that will be used in this project is Bayesian Analysis. The use of Bayesian Analysis (BA) is somewhat new in phylogenetics, but its use offers some advantages. Unlike the ML method of analysis, which searches for the *single best* tree, BA searches for the best *set* of trees. The Bayesian analysis for this study will be carried out using the program called *Mr.Bayes* (Huelsenback & Ronquist, in press). BA differs from ML in that in BA, the program is looking at 4 different tree topologies and comparing them all at the same time. By doing this you increase the likelihood that you will actually get the most probable tree topology. The tree generated by the BA method is a strict consensus tree derived from each of the best trees picked per cycle. Unlike the previous methods of analysis, there is no bootstrapping on a BA tree. Bootstraping a BA tree is not necessary due to algorhythm that BA uses, not to mention that bootstrapping a BA tree would take roughly 30 days of solid computing time.

With four different methods, each generating their own tree, we are left to choose which tree to trust the most. If changes in the gene sequence we studied really represent the

history of evolution in our taxa, and the alignment of these sequences are robust, then, in theory, the more closely the trees produced by different methods should be in topology. The differences will be small and indicate a real uncertainty in the process of evolution (Hall, 2001).

Objective

Two species selected for this study are *Centruroides infamatus* (C. L. Koch, 1845) and *C. limpidus* (Karsch, 1879), which are widely distributed in Central Mexico. Each of those has two described subspecies but validity of these forms is not clear. We intend to study 11 populations of these two species complexes throughout Central Mexico, encompassing all ecological gradients and altitudinal levels. Figure 2.4 below shows a map of our sampling locations.

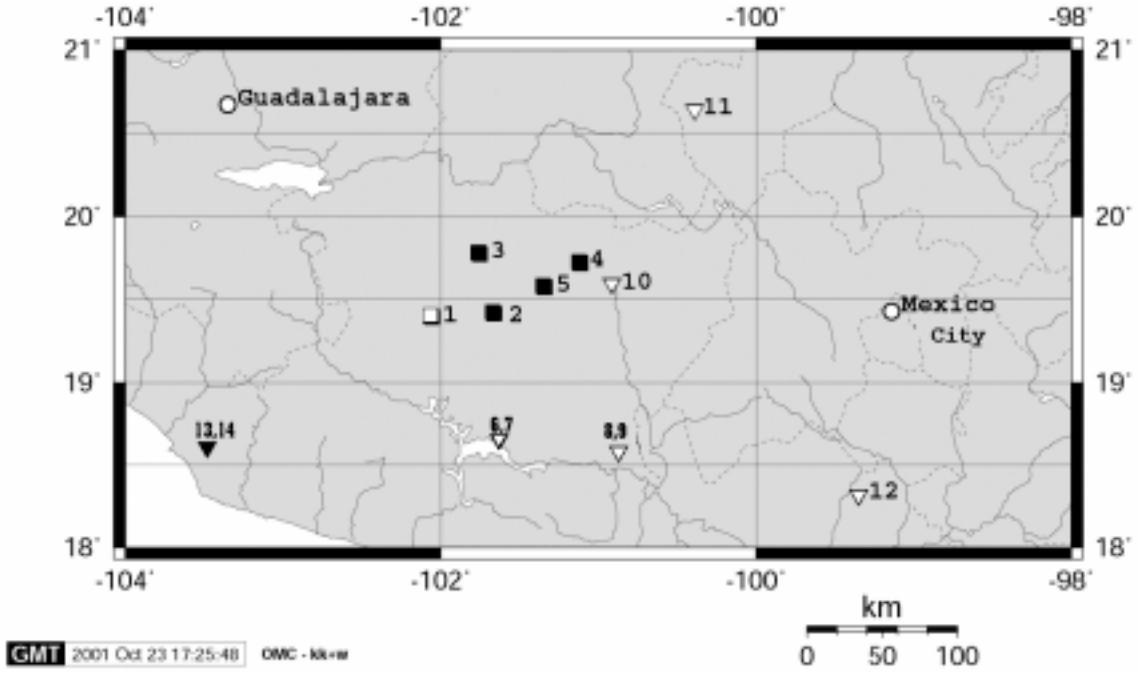


FIGURE 2.4 MAP OF THE SAMPLES OF CENTRUROIDES IN CENTRAL MEXICO (SEE DETAILED LOCALITIES DATA IN TEXT). C. INFAMATUS INFAMATUS: 1 – CIIZU1, ZUMPIMITO; C. INFAMATUS ORNATUS: 2 – CIOCO1, EL COBANO; 3 – CIOLC1, LA CARATACUA; 4 - CIIMO1, MORELIA; 5 – CIOTI1, TIRIPETIO; C. LIMPIDUS LIMPIDUS: 6 – CLLCH1 AND 7 – CLLCH2, CHURUMUCO; 8 – CLLTO1, TZIRANDARO; 9 – CLLAR1, ARUA; 10 - CLLTZ1, TZITZIO; 11 – CLLQU1, QUERÉTARO; 12 – CLLHU1, HUITZUCO; C. LIMPIDUS TECOMANUS: 13 – CLTFB1 AND 14 – CLTFB2, EL FARO DE BUCERIAS.

A set of molecular (mtDNA) sequences will be produced through total DNA extraction and PCR amplification. These molecular markers have recently become a powerful tool for evaluating the taxonomic status of animal populations, subspecies, and species. The first information on applicability of mtDNA analysis to the species-level taxonomy of *Centruroides* was reported by Fet and Poindexter (1992). Recently, comparisons of 16S rRNA mtDNA sequences allowed to clarify phylogeny at the species level among the populations of *Centruroides exilicauda* (Wood) from Baja California and Sonora (Mexico) and Arizona (USA) (Fet *et al.*, 1999; Gantenbein *et al.*, 2001) as well as other scorpion genera and families (Fet *et al.*, 2001; Gantenbein *et al.*, 1999,2000). The number of base pairs that are sequenced from this gene fragment in this study are high enough for a good phylogenetic signal (Huelsenbeck & Hillis, 1996). Analysis of DNA sequences from mitochondrial genes, which are routinely used for phylogenetic studies at the species level (Fet *et al.*, 1999) will be done using standard alignment and phylogenetic software. We will create and analyze phylogenies of species, subspecies, and populations. As outgroups, *C. exilicauda* (Wood) (Fet & Poindexter, 1992) from Baja California, *C. vittatus* (Say) from Arkansas, and *C. bani* (Armas & Fondeur) from the Dominican Republic, will be used.

The goal is to analyze historical origin of these species, to test various biogeographic hypotheses of the origin (otherwise untestable), and to resolve confusing taxonomy that exists in the scorpion genus *Centruroides* of Central Mexico. This study will help to understand unique features of Central Mexican biodiversity and the ways of its historical formation. Since species of *Centruroides* are among the deadliest scorpions in the world, understanding of their taxonomy is also necessary for species delineation in practical studies such as toxicology, and public health.

CHAPTER III

Methods

Sample Collection

Two different species complexes, the *Centruroides infamatus* (C. L. Koch, 1844) complex that includes *Centruroides infamatus ornatus* and *Centruroides infamatus infamatus* and the *Centruroides limpidus* (Karsch, 1879) complex that includes *Centruroides limpidus limpidus,* and *Centruroides limpidus tecomanus* were chosen and collected. Our collaborator Mr. Javier Ponce Saavedra, in central Mexico, collected all samples. Mr. Ponce Saavedra is a Ph. D student at the Universidad Michoacana de San Nicolás de Hidalgo in Michoacan, Mexico, and has all the appropriate permits required by the Mexican government to collect these samples. The samples were collected by Mr. Saavedra using standard black light techniques as described in previous literature (Stahnke, 1972). Once collected the samples were placed on dry ice and shipped to Marshall University. Upon arrival they were removed from the dry ice and placed in individual labeled vials filled with 97% ethanol. The ethanol serves as a preservative for the DNA. Below is a list of the samples by label, with a description of the habitat and locality they were collected from. Table 3.1 below, presents a summary of this information.

Centruroides infamatus infamatus

CiiZU1 This specimen was found in Zumpimito, Municipal of Uruapan, Michoacán. The area is located in the transition between the temperate zone and the warmer areas. This locality is just west, but near the center of the state. The elevation is 5118 feet above sea level. Female

Centruroides infamatus ornatus

CioLC1 From “La Caratacua”, Municipal of Coeneo, Michoacán. This area is located in the temperate zone, near the mountains of the Neovolcanic belt. This area is in the Northwest portion of the state, and has an elevation of 5511 feet above sea level. Male.

CioTI1 This specimen was found in Tiripetío, Municipal of Morelia, Michoacán. Found in the temperate zone of Morelia valley, this area lies in the north central part of the state. The elevation is 6627 feet above sea level. Female

CioMO1 From within Morelia City, Municipal of Morelia, Michoacán. This area is in the northern-central portion of the state with an elevation of 6364 feet above sea level. Female

CioEC1 Found at “El Cóbano”, Municipal of Gabriel Zamora, Michoacán. This area has warmer weather near the Neovolcanic belt with an elevation of 5052 feet above sea level. Male

Centruroides limpidus limpidus

CllTZ1 From “Tres Puentes” Tzitzio, Municipal of Tzitzio, Michoacán. This region lies in the north central part of the state with warm weather and an altitude of 5823 feet above sea level. This is just on the edge of the depression of the Balsas. Female

CllQU1 This specimen was found in Querétaro, Querétaro. This population was found within the city limits at an elevation of 5971 feet above sea level. Female

CllTO1 From Tzirándaro, Guerrero, which is located in the middle of the balsas depression, near the balsas river. The weather is very hot with an altitude of only 1607 feet above sea level. Female

CllAR1 Found at Arúa, Municipal of Huetamo, Michoacán. This locality is situated on the boarder of the Michoacán and Guerrero States. With an altitude of 1607 feet above sea level, it is very hot, but has a very high relative humidity. The conditions here are similar to Tzirándaro, except it is more humid. Male

CllHU1 From Huitzuco, Guerrero, near Iguala. This is the type locality for this species. It has an elevation of 3149 feet above sea level. Female

CllCH1 This is an example from Churumuco, Michoacán. This locality is located in the Balsas depression, and has the lowest elevation at 984 feet above sea level. Very hot and dry. Female

CllCH2 Same as above. Female

Centruroides limpidus tecomanus

CltFB1 This specimen was found at "El Faro de Bucerías", Municipal of Aquila, Michoacán. This locality is within the zone Hoffman reports the type examples are found. Male

CltFB2 Same as above. Female

Table 3.1 Species labels with Gender, species name and GPS cordinates. GPS cordinates are given in UTM.

Label	***GPS***	***Species***	***Sex***
CioTI1	-101.3333,19.5500	*C. infamatus ornatus*	Female
CiiZU1	-102.0500,19.3667	*C. infamatus infamatus*	Female
CioEC1	-102.0333,19.1500	*C. infamatus ornatus*	Male
CioLC1	-101.5833,19.8167	*C. infamatus ornatus*	Male
CioMo1	-101.1167,19.7000	*C. infamatus ornatus*	Female
CllAR1	-100.8833,18.5833	*C. limpidus limpidus*	Male
CllCH1	-101.6333,18.6167	*C. limpidus limpidus*	Female
CllCH2	-101.6333,18.6167	*C. limpidus limpidus*	Female
CllHU1	-99.3500,18.3000	*C. limpidus limpidus*	Female
CllQU1	-100.3833, 20.6000	*C. limpidus limpidus*	Female
CllTO1	-100.8633,18.5633	*C. limpidus limpidus*	Female
CllTZ1	-100.9167,19.5667	*C. limpidus limpidus*	Female
CltFB1	-103.5000,18.6000	*C. limpidus tecomanus*	Female
CltFB2	-103.5000,18.6010	*C. limpidus tecomanus*	Male

As outgroups for phylogenetic analysis, *C. exilicauda* (Wood), from Baja California, *C. vittatus* (Say) from Arkansas, and *C. bani* (Armas & Fondeur) from the Dominican Republic, will be used. These are all samples that have been used in previous studies in this lab.

DNA Extraction, Amplification and Sequencing.

Total DNA was extracted from the legs of the samples, one at a time using a commercially available DNA extraction kit from Quiagen Inc. The extraction was carried out using the Qiagen Dneasy tissue kit protocol with slight modifications. For further information see appendix B.

This study amplified two different mtDNA genes, 16S and COI. All amplification reactions were carried out on a Perkin Elmer Geneamp ® PCR 2400, version 2.10 thermocycler. The protocol used in the 16S and COI amplification can be found in table 3.2 below.

Table 3.2 Temperature and time settings for each step of the PCR reaction used in the amplification of all target sequences. Pre-cycling conditions, number of cycles, and post-cycling conditions given below.

16S & CO1	Time	Temperature
Denaturing	0:45	94
Annealing	0:45	48
Extension	0:45	72
Pre-cycling	5:00	94
Cycles	30	
Post-cycling	10:00	72

The primers used in amplification of the 16S gene fragment have the following sequences, CGA TTT GAA CTC AGA TCA (forward 18-mer, also known as 38) and GTG CAA AGG TAG CAT AAT CA (reverse 20-mer, also known as 40). The primers used in the CO1 amplification have the sequence of CCC GGT AAA ATT AAA ATA TAA ACT TC (forward 25-mer, also known as Nancy) and GGT CAA CAA ATC ATA AAG ATA TTG (reverse 24-mer, also known as LCO). For more information on the primers please see appendix G. The PCR product is then checked using gel electrophoresis. If the product is good, then it is cleaned using an Ultrafree MC 30,000 cellulose centrifuge filter from Millipore Inc. The cleaned and filtered product is then sent to the Molecular Genetics Instrumentation Facility, Department of Genetics, University of Georgia, for sequencing. All sequences have been submitted to GENBANK, further information can be found in Appendix F.

Data Analysis

All data analysis was performed on an IBM compatiable PC, with an AMD K6-450 processor, 128 megs of RAM, running Windows® 98SE operating system.

After the sequences were obtained for all samples, they were aligned using the default settings in CLUSTALX version 1.81 for windows (Thompson *et al*., 1997). The alignment of the sequences was done by software to help maximize the repeatability of the study, as the computer program uses algorithms and will align the same sequences in the same manner every time where as a human may not. For further information about the CLUSTALX program and its use, please see Appendix E.

The testing of different evolutionary models for the ML method was done using a Software package MODELTEST 3.06 (Posada & Crandall, 1998). The program tested 56 different models of nucleotide substitution. The appropriate model chosen by MODELTEST was then implemented in the ML analysis.

All estimates and trees of ML, NJ and MP methods, were computed using the PAUP* 4.10b software (Swofford, 1998). The parameters for ML were estimated to π_A = 0.34, π_C = 0.15, π_G = 0.12, and π_T = 0.37, α = 0.15, transition (ti)/transversion (tv) ratio = 7.45 (κ = 18.41), -ln L = 1411.76. For MP, ti's were down-weighted relative to tv's by factor seven. The tree topology of the MP tree is a strict concensus of 12 trees, with 209 steps. NJ analysis was done with default settings.

Bayesian Analysis was done using the software package entitled Mr.Bayes (Huelsenbeck & Ronquist, In Press). The tree generated is a strict concensus tree of 10,000 trees generated after running 1,000,000 repititons.

All trees generated by both PAUP* and MrBayes, were imported into TREEVIEW (Page, 1996) for processing and enhancement. The trees were then saved and ready for use in this article.

CHAPTER IV

Results

Table 4.1 shows the log likelihood scores from the MODELTEST program for 56 different evolutionary models. The model that was chosen was HKY + I + G with a score of 1,936.20.

Table 4.1 Log of the likelihood scores computed by MODELTEST for 56 evolutionary models for use in analysis.

		+I	*+G*	*+I+G*
JC	2,087.26	2,087.26	2,080.96	2,077.52
F81	2,056.11	2,056.11	2,050.27	2,047.10
K80	1,996.69	1,996.69	1,988.46	1,984.11
HKY	1,947.49	1,947.49	1,939.77	1,936.20
TrNef	1,995.59	1,995.59	1,987.05	1,982.60
TrN	1,947.33	1,947.33	1,939.58	1,936.14
K81	1,996.17	1,996.17	1,988.02	1,983.67
K81uf	1,947.14	1,947.14	1,939.15	1,935.59
TIMef	1,995.08	1,995.08	1,986.63	1,982.18
TIM	1,946.98	1,946.98	1,938.92	1,935.52
TVMef	1,994.90	1,994.90	1,986.71	1,982.13
TVM	1,943.51	1,943.51	1,934.88	1,931.74
SYM	1,993.83	1,993.83	1,985.53	1,980.89
GTR	1,943.36	1,943.36	1,934.67	1,931.69

Table 4.2 shows the Maximum Likelihood distance matrix (upper right) and uncorrected *p* distance matrix (lower left) of *Centruroides* 16S mtDNA sequences. Parameters for ML-distances were set to the HKY85 +I +G model; transition/transversion ratio = 7.44 (kappa = 18.41); assumed nucleotide frequencies (set by user): A=0.35, C=0.15, G=0.12, T=0.38; shape parameter α = 0.15. 58 characters are excluded, 366 characters remaining.

Table 4.2 Maximum Likelihood distance matrix (upper right) and uncorrected *p* distance matrix (lower left) of *Centruroides* 16S mtDNA sequences.

		1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*
1	*CioLC1*	-	0.00	0.00	0.01	0.11	0.14	0.16	0.22	0.23	0.21	0.22	0.22	0.50	0.41	0.20	0.28	0.37
2	*CioEC1*	0.00	-	0.00	0.01	0.11	0.14	0.16	0.22	0.23	0.21	0.22	0.22	0.50	0.41	0.20	0.28	0.37
3	*CioMO1*	0.00	0.00	-	0.01	0.12	0.15	0.16	0.22	0.22	0.22	0.23	0.23	0.51	0.42	0.21	0.29	0.36
4	*CioTI1*	0.01	0.01	0.01	-	0.13	0.16	0.19	0.25	0.26	0.24	0.25	0.25	0.54	0.42	0.22	0.31	0.38
5	*CiiZU1*	0.07	0.07	0.07	0.08	-	0.12	0.14	0.21	0.22	0.20	0.24	0.24	0.53	0.46	0.23	0.29	0.40
6	*CltFB1*	0.08	0.08	0.08	0.09	0.07	-	0.01	0.22	0.23	0.21	0.20	0.20	0.41	0.40	0.20	0.28	0.36
7	*CltFB2*	0.09	0.09	0.09	0.10	0.08	0.01	-	0.22	0.23	0.24	0.23	0.23	0.44	0.44	0.23	0.31	0.35
8	*CllCH1*	0.10	0.10	0.10	0.11	0.10	0.10	0.09	-	0.00	0.03	0.08	0.08	0.54	0.39	0.19	0.20	0.23
9	*CllCH2*	0.11	0.11	0.10	0.11	0.10	0.10	0.10	0.00	-	0.03	0.09	0.09	0.56	0.40	0.18	0.21	0.22
10	*CllTO1*	0.10	0.10	0.10	0.11	0.09	0.09	0.10	0.02	0.03	-	0.06	0.06	0.46	0.40	0.20	0.19	0.26
11	*CllTZ1*	0.10	0.10	0.10	0.11	0.10	0.09	0.10	0.05	0.06	0.04	-	0.00	0.44	0.40	0.23	0.24	0.29
12	*CllAR1*	0.10	0.10	0.10	0.11	0.10	0.09	0.10	0.05	0.06	0.04	0.00	-	0.44	0.40	0.23	0.24	0.29
13	*Cbani127*	0.14	0.14	0.14	0.15	0.14	0.13	0.13	0.14	0.14	0.12	0.12	0.12	-	0.46	0.40	0.50	0.60
14	*Cvit126*	0.15	0.15	0.15	0.15	0.15	0.14	0.15	0.14	0.14	0.14	0.14	0.14	0.14	-	0.36	0.49	0.38
15	*Cexil83*	0.10	0.10	0.11	0.11	0.11	0.10	0.11	0.09	0.09	0.10	0.11	0.11	0.13	0.14	-	0.29	0.27
16	*CllQU1*	0.12	0.12	0.13	0.13	0.12	0.12	0.13	0.10	0.10	0.09	0.11	0.11	0.15	0.16	0.12	-	0.11
17	*CllHU1*	0.15	0.15	0.15	0.16	0.15	0.14	0.14	0.11	0.11	0.12	0.12	0.12	0.18	0.14	0.12	0.07	-

Four different trees from the different analysis's were constructed. The ML and NJ trees were combined, with bootstraps for each method at the nodes of each branch. Figure 4.1 shows 3 of the four trees constructed. The ML/NJ tree were placed opposite the MP tree for the ease of comparison.

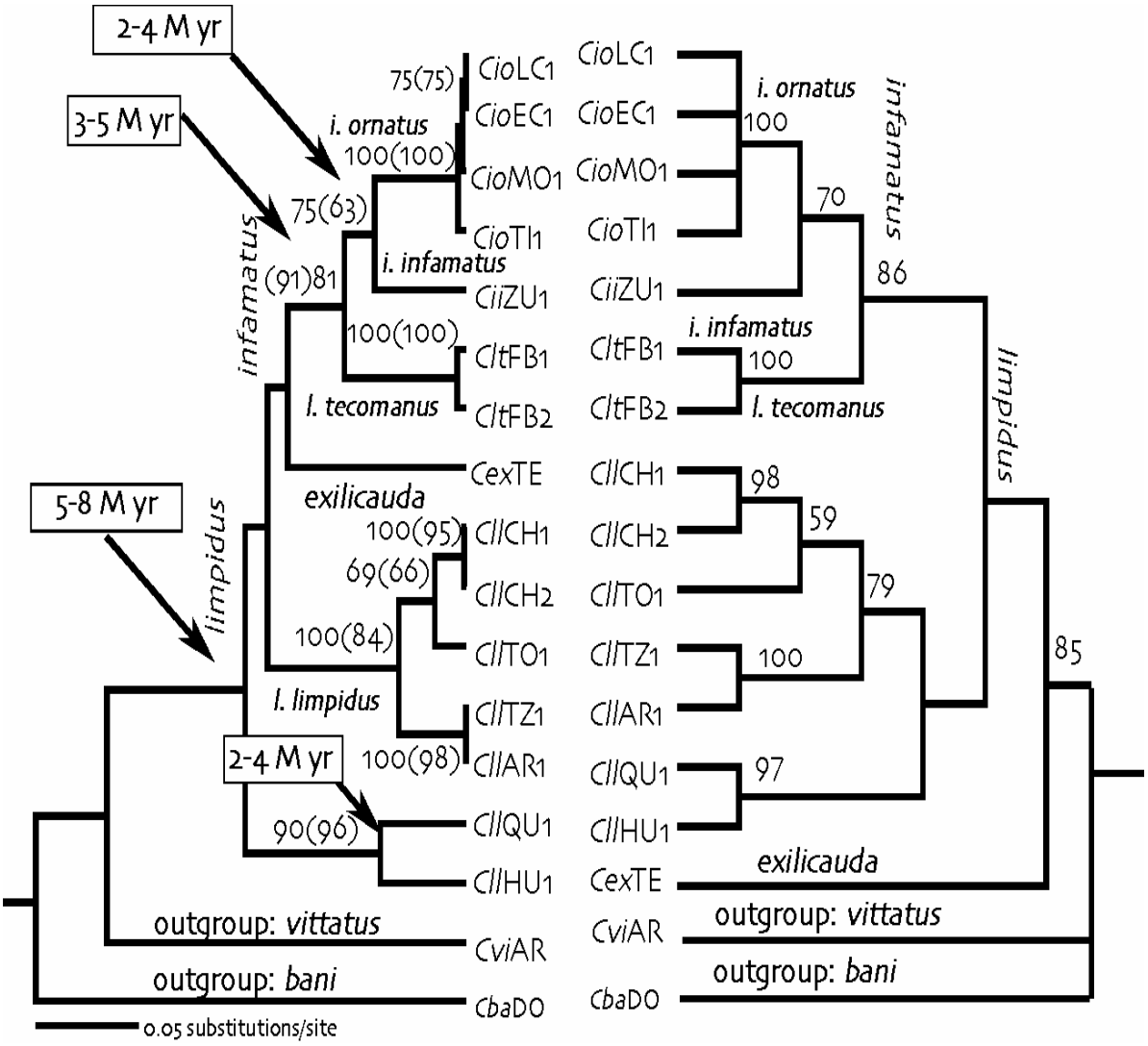


FIGURE 4.1 16S MTDNA GENE GENEALOGY (366 BP) OF CENTRUROIDES FROM CENTRAL MEXICO. TREE TOPOLOGY WAS REVEALED BY MAXIMUM LIKELIHOOD (ML) ANALYSIS (LEFT) USING THE HKY85 + I +G MODEL AND BY WEIGHTED MAXIMUM PARSIMONY (MP) ANALYSIS (RIGHT). THE PARAMETERS FOR ML ARE LISTED IN THE METHODS SECTION. THE SINGLE MPTREE NEEDED 209 STEPS, AND THE CONSISTENCY INDEX EXCLUDING UNINFORMATIVE CHARACTERS (CIU) AND THE RETENTION INDEX (RI) WERE 0.56 AND 0.70, RESPECTIVELY. THE TREE TOPOLOGY OF THE CLADOGRAM ESTIMATED BY ML DIFFERS FROM THE TOPOLOGY REVEALED BY MP ONLY IN THE PHYLOGENETIC POSITION OF THE SEQUENCE OF C. EXILICAUDA. BOOTSTRAP VALUES CORRESPOND TO BOOTSTRAPPING (IN PARENTHESES FROM NEIGHBOUR-JOINING ANALYSIS).

The fourth tree generated was the Bayesian analysis tree shown in figure 4.2, which is a concensus tree generated from 10,000 sampled trees.

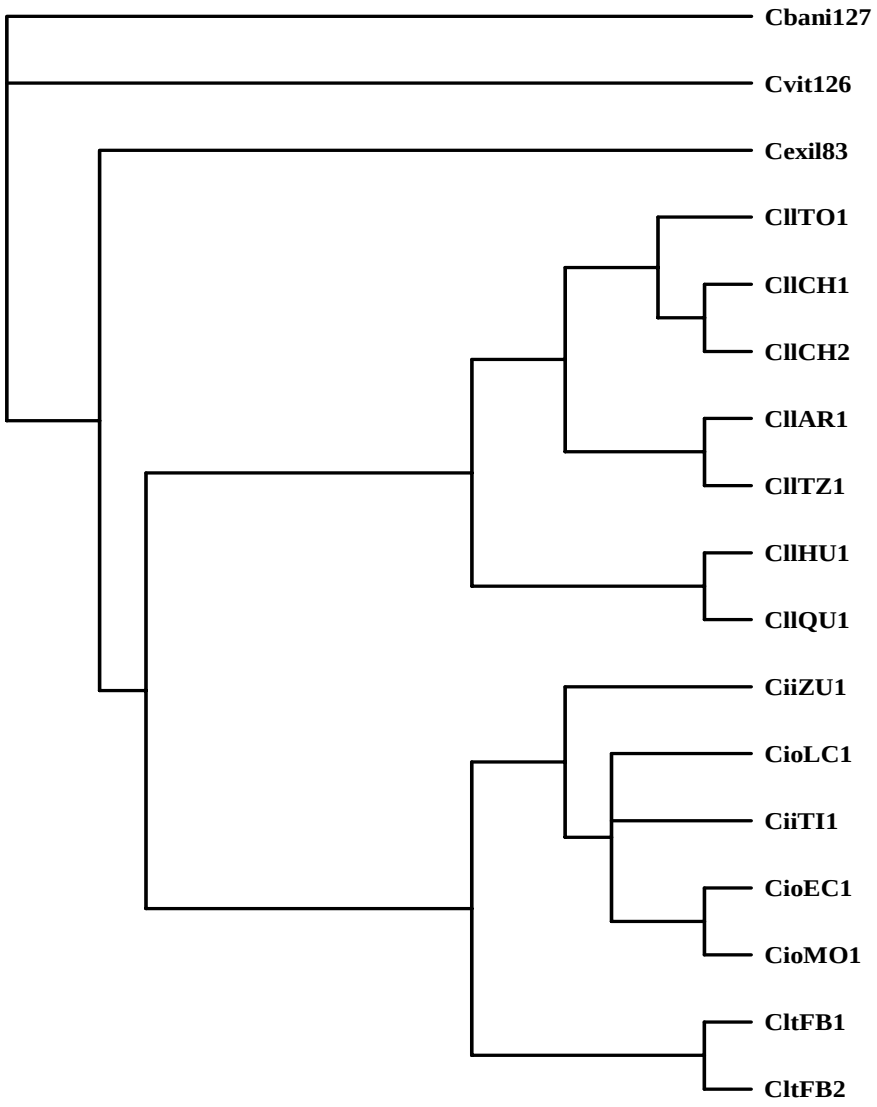


FIGURE 4.2 BAYESIAN ANALYSIS CONCENSUS TREE OF 10,000 TREES SAMPLED FROM 1,000,000 REPETITIONS.

The results from the CPU time needed in the analysis are found in table 4.3. The table shows the amount of time need for each of the data analysis methods used in hours: minuets: seconds format.

Table 4.3 CPU time needed in the analysis of 17 DNA sequences in Central Mexican *Centruroides* species. Data shown in hours: minuets: seconds format.

	Neighbor Joining	Maximum Parsimony	Maximum Likelihood	Bayesian
CPU time used	0:00:48	0:02:35	0:18:51	18:36:24

CHAPTER V

Summary and Conclusion

This study, first of all, confirmed monophyly of the studied populations of *Centruroides infamatus*. There were clearly observed two clade within this species. One was represented by the sole *C. infamatus infamatus* specimen taken from a humid pine-oak forest of Zumpimito, on the southern slope of the Trans-Mexican Volcanic Belt mountain range (these mountains are also known as "Corredor Tarasco"). Another clade included four specimens identifed as *C. infamatus ornatus* collected from the same landscape as Zumpimito (El Cobano, 1540 m) and three colder and drier highland localities (La Caratacua, 1680 m; Morelia city, 1940 m; and Tiripetio, 2020 m) of the northern slope. The taxonomic status of these two subspecies is unresolved. They were originally described as separate species, and *C. ornatus* (type locality Jalisco, Mexico) was later downgraded to a subspecies of *C. infamatus* by Hoffmann (1932). Most recently, Beutelspacher (2000) synonymized this taxon with *C. infamatus* (treating it as a “forma *ornatus*” is not in accordance with the International Code of Zoological Nomenclature, which does not allow categories below subspecies rank). However, Ponce and Beutelspacher (2001) suggested that these two subspecies do in fact have certain morphometric differences. Further support is offered by the fact that, at least in Michoacán these two forms were never found sympatrically (observations of J.P.S.; the data from literature showing sympatry of *C. i. infamatus* and *C. i. ornatus* are based on misidentifications). This study confirms two divergent clades; divergence rate estimates their common ancestor’s age as 2-4 Ma for HKY+Γ divergence rate (11.7 ± 0.9 %) and 3-5 Ma for uncorrected *p* (7.2 ± 0.4 %). Further study is necessary with sampling all over the range of both taxa, to confirm existence of two independent lineages. If such are confirmed, *C. i. ornatus* could be elevated to a species rank. It appears that the

nomenclatural category of subspecies is not consistent with the viewpoint of the phylogenetic species concept (PSC), which defines species based on monophyly.

For *Centruroides limpidus*, the situation was more complicated, since monophyly of this species was not confirmed by the DNA data. The coastal population from El Faro de Bucerias, identified as the subspecies *C. l. tecomanus*, in all analyses, formed an unexpected but highly supported sister clade to *C. infamatus* rather than to *C. l. limpidus*! Moreover, at least in one of our analyses *C. exilicauda* from Arizona formed an ingroup, cutting between *C. l. limpidus* clade and *C. limpidus tecomanus* (Fig. 4). The evolutionary rate estimate of the age of common ancestor of *C. infamatus* and "*C. limpidus tecomanus*" is ca. 3-5 Ma for HKY+Γ divergence rate (15.3 ± 1.9 %), and 4-7 Ma for uncorrected *p* (10.1 ± 0.8 %).

It is suspected that "*C. l. tecomanus*" is, in fact, a separate species, which may not be closely related to *C. limpidus* at all. This enigmatic taxon was not studied in detail since it its description by Hoffmann (1932) from several localities in the Mexican state of Colima (Colima, Tecoman, and Manzanillo). It is differentiated from the nominotypic subspecies essentially only by presence of a developed subaculear tooth (Stahnke, 1977) – a character which is known to vary even within developmental stages in *Centruroides* (Lourenço, 1982); some juvenile specimens identified formerly as *C. l. tecomanus* appear in fact to be juveniles of *C. l. limpidus* (Ponce & Beutelspacher, 2001). Additional morphological data (Ponce, in preparation) show that this form also is distinguished from *C. limpidus limpidus* by the shape of female pectinal teeth, the color pattern on the carapace, and the carination of mesosomal tergites. The recent comparative study by Dehesa-Dávila, (1996) on the primary amino acid structure of the Na^+-channel-specific β-toxins supports our view that *C. l. tecomanus* is highly divergent from *C. limpidus*. Comparing the amino acid sequences, in their study, *i.e.* the *Cii*-1 toxin (= *C. i. infamanus*) and two *Cll* toxins (= *C. l. limpidus*) with the *Clt*-1 toxin (= *C. l. tecomanus*) revealed that 3-4 % amino acid replacements were found between the *Cii* and the *Cll*

sequences, whereas 4-7 % substitutions were found between *Clt*-1 and all other toxins. Moreover, a phylogenetic tree reconstructed from their alignment by Maximum Parsimony revealed that *Clt*-1 clearly was found separated from the two *Cll*1-2 and *Cii*-1 sequences. Although these sequences might represent actually paralogous genes of a closely related toxin gene family, our phylogenetic approach indicates that toxin genes harbor considerably phylogenetic information despite the high selective constraint. In this context, especially synonymous mutations (expected to be neutral) at the DNA level would be more informative since amino acid replacements might be of adaptive nature. The geographic range of *C. l. tecomanus* is confined to a rather narrow ecological area along the Pacific coast (Región de la Costa) within the states of Nayarit, Colima, Michoacán, and Jalisco. *C. l. tecomanus* does not cross the mountains of Sierra Madre del Sur into the Región de Sierras del Centro in Michoacán (Ponce, 2001) and thus is practically allopatric with *C. l. limpidus*.

The nominotypic subspecies *C. limpidus limpidus* has a rather wide range in the south-central Mexico (central Guerrero, Morelos, Michoacán, Distrito Federal, Mexico, Querétaro, and parts of Oaxaca and Puebla). Samples were taken from its populations at Michoacán, Guerrero, and Querétaro. Within this material, the DNA phylogeny reveals two clear lineages of deep divergence. One lineage (well supported at 79 to 84 % bootstrap) includes specimens from Churumuco, Arua, Tzirandaro, and Tzitzio, all found within the boundaries of the Balsas Depression in Michoacán, along the entire altitudinal profile (300 to 1775 m a. s. l.). Another well-supported lineage (96 to 97 % bootstrap) includes two specimens sampled from the mountains further to the west (two distant localities in Querétaro and Guerrero). Time of divergence between two lineages can be estimated at ca. 5-8 Ma BP for both HKY +I+G divergence (23.7 ± 3.5 %) and uncorrected *p* (10.9 ± 1.0 %) (Table 3). This estimate is consistent with the geological age of Balsas Depression, which was formed from 4 to 6 Ma BP. The mountains of Sierra Madre del Sur in the south, and the Mexican Transverse Volcanic Belt in the north and west have been rising since the Eocene-Oligocene time. The tectonic events could have isolated the populations in the area of modern Balsas Depression, which possibly could speciate *in situ* and then disperse toward the adjacent mountain slopes of the

transitional zone (e.g. Tzitzio, 1775 m). The divergence time between Querétaro and Guerrero samples was estimated as 2-4 Ma BP for HKY +I +G rate (11 %) and 3-5 Ma for uncorrected *p* (7 %).

This study suggests that more than one ancient monophyletic lineage (possibly, more than one species) exist within currently accepted *Centruroides limpidus limpidus.* The type locality of *Centruroides limpidus* is Puebla, which lies in the same geographic area as Guerrero. Thus, we might assume that the Querétaro/Guerrero lineage corresponds to 'true' *C. limpidus*, and that the Balsas Depression populations could belong to another, 'cryptic', or 'sibling' species. Further, detailed investigations should be done to test these preliminary conclusions: the need for many more populations from the entire range of *C. limpidus* is needed. Several data sets (mitochondrial and nuclear genes, allozymes, morphology, toxin structure/activity, etc.) could be analysed to establish the true taxonomic and genetic structure of the populations and species of *Centruroides*.

This study used four different data analysis methods, including three established methods, neighbor joining, maximum parsimony, and maximum likelihood. The fourth method, bayesian analysis was used in addition to the previous methods mentioned. When looking at the resulting trees from each of the four methods they are all very similar, so similar that 3 of them NJ, MP, and ML can be shown together with out any problem. The BA tree is again similar in overall tree topology, and for the most part gives the same results. This says a few things, first, the gene we have chosen for analysis is a good gene to use within these animals. Second, it says that all methods of data analysis are valid forms, and should be considered for use in future studies. With that being said, because the trees generated by these four different analysis methods are so similar to each other, BA may not need to be done. When looking at table 4.3, you can see that the BA program, at 18 plus hours, takes a significantly longer time to generate the 'final tree' that we are looking for, and since the four trees do not differ in topology drastically there is really no need to render the computer useless for 18 plus hours.

Future studies should use NJ, ML, and MP in analysis until the BA becomes more efficient in regards to time and resource uses.

Works Cited

Avise, J.C. 1994. *Molecular Markers, Natural History and Evolution*. Chapman and Hall. London.

Armas, L. F., 1988. *Sinopsis de los escorpiones antillanos*. Editorial Científico-Técnica, La Habana, 102p.

Baerg, W.J. 1961. *Scorpions: Biology and effects of their venom.* University of Arkansas Agriculture Experimental Station, Bulletin 649. 34 pp.

Beutelspacher, C.R., 2000. *Catálogo de los alacranes de México.* Universidad de San Nicholás de Hidalgo, Morelia, Michoacán, 175pp.

Cavalli-Sforza, L.L. & A.W.F. Edwards. 1964. Phylogenetic analysis: Models and Estimation procedures. *Am. J. Hum. Genet.* 19: 233-257.

Darwin, C. R. 1859. *On the Origin of Species by means of natural selection*. J. Murray, London.

Dehesa-Dávila, M., A. N. Ramírez, F. Z. Zamudio, G. Gurrola-Briones, A. Liévano, A. Darzon & L. D. Possani, 1996. Structural and functional comparison of toxins from the venom of the scorpions *Centruroides infamatus infamatus, Centruroides limpidus limpidus* and *Centruroides noxius*. *Comp. Biochem. Physiol*., 113B 2: 331-339.

Fet, V., M. Barker, & B. Gantenbein. 1999. Species-level variation of the mitochondrial 16S rRNA gene sequence: use in molecular systematics and biogeography. Proc. *West Virginia Acad. Sci.*, 71: 15.

Fet, V. & B. J. Poindexter. 1992. Genetic variation of mitochondrial DNA in *Centruroides exilicauda* (Scorpiones : Buthidae) as reveled by the polymerase chain reaction. *Amer. Zool.* 32 5: 136A.

Fet, V. & Lowe, G. 2000. Family Buthidae C. L. Koch, 1837. *In* V. Fet, W. D. Sissom, G. Lowe & M. E. Braunwalder. *Catalog of the scorpions of the world (1758–1998).* New York, NY: The New York Entomological Society: 54–286.

Fet, V., W. D. Sissom, G. Lowe & M. E. Braunwalder. 2000. *Catalog of the scorpions of the world (1758–1998).* New York, NY: The New York Entomological Society

Fet, V., M. E. Soleglad & M. Barker. 2001. Phylogeny of the 'hirsutus' group of the genus *Hadrurus* Thorell, 1876 (Scorpiones: Iuridae) based on morphology and mitochondrial DNA. *In* V. Fet & P. A. Selden (eds.). *Scorpions 2001: In Memoriam Gary A. Polis*. Burnham Beeches, UK: The British Arachnological Society (in press).

Fitch, W.M. 1971. Toward defining the course of evolution: Minimum change for a specific tree topology. *Syst. Zool.* 20: 406-416.

Fleissner, G. 1977. The absolute sensitivity of the median and lateral eyes of the scorpion, *Androctonus australis* L. (Buthidae, Scorpiones). *Journal of Comparative Physiology.* 118A: 109-120.

Gantenbein, B., V. Fet, C. R. Largiader & A. Scholl. 1999. First DNA phylogeny of the *Euscorpius* Thorell 1876 (Scorpiones: Euscorpiidae) and its bearing on the taxonomic and biogeography of this genus. *Biogeographica (Paris),* **75**: 59–72.

Gantenbein, B., V. Fet, M. Barker & A. Scholl, 2000. Nuclear and mitochondrial markers reveal the existence of two parapatric scorpion species in the Alps: *Euscorpius germanus* (C. L. Koch, 1837) and *E. alpha* Caporiacco, 1950, stat. nov. (Scorpiones, Euscorpiidae). Rev. Suisse Zool., 107 (4): 843-869.

Gantenbein, B., V. Fet & M. D. Barker, 2001. Mitochondrial DNA markers reveal a deep, divergent phylogeny in *Centruroides exilicauda* (Wood) (Scorpiones: Buthidae). Pp. 235-244, *In*: Fet, V. & P. A. Selden (Eds.), Scorpions 2001. In Memoriam Gary A. Polis. British Arachnological Society, Burnham Beeches, Bucks.

Haeckel, E. 1866. *Generelle Morphologie der Organismen.* Georg Riemer, Berlin.

Hall, B. G., 2001. *Phylogenetic Trees Made Easy*. Sinauer Assoc. Inc. Publishers

Hartigan, J.A. 1973. Minimum evolution fits to a given tree. *Biometrics* 29: 53-65.

Hennig, W. 1966. *Phylogenetic systematics.* University of Illinois Press, Urbana.

Hillis, D.M., C. Moritz, and B.K. Mable. 1996. *Molecular Systematics*. Sinauer Associates, Inc., Mass., US.

Hoffmann, C.C. 1932. Monografias para la entomologiá médica de México. Monografia Num. 2, Los escorpiones de México. Sequnda parte: Buthidae. *Anales del Instituto de Biologia Unversidad Nacional Autónoma de México,* 3 (3): 243-282; (4): 283-361.

Huelsenbeck, J. P. & D. M. Hillis, 1996. Parametric bootstrapping in molecular phylogenetics: Applications and performance. Pp.19-45, *In*: Ferraris, J. D. & S. R. Palumbi (Eds.), *Molecular Zoology: Advances, Strategies, and Protocols*. Wiley-Liss, Inc., New York.

Huelsenbeck, J.P, & F.R.Ronquist. InPress. MRBAYES:Bayesian inference of phylogeny. *Biometrics*.

Jeanmougin,F., Thompson,J.D., Gouy,M., Higgins,D.G. and Gibson,T.J. 1998 Multiple sequence alignment with Clustal X. *Trends Biochem Sci*, 23, 403-5.

Jeffrey, C. 1973. *Biological Nomenclature*. Edward Arnold Limited.

Karsch, F. 1879. Scorpionologische Beiträge. Part I. *Mitteilungen des Münchener Entomologischen Vereins,* 3: 6-22.

Koch, C.L. 1844. *Die Arachniden*. NürnBerg: C.H. Zeh'sche Buchhandlung, 11:1-174.

Linnaeus, C. 1758. *Systema Naturae*. Tenth Edition

Lourenço, W. R. & W. D. Sissom, 2000. Scorpiones. Pp.115-135, *In*: Llorente Bousquets, J., E. González Soriano & N. Papavero (Eds.), *Biodiversidad, taxonomía y biogeografía de artropodos de México,* vol. 2. Universidad Nacional Autónoma de México, México.

Lourenço, W. R., 1982. Utilisation de l'épine sous-aiguillonnaire dans la taxonomie des Scorpions de la famille des Buthidae (région néotropicale). *Boll. Mus. Zool. Univ. Torino,* 5: 73-78.

Marx, G. 1890. Arachnida. *In* L.O. Howard (ed.), Scientific results of the explorations by the U.S. Fish Comission Steamer Albatross. No. V. Annoted Catalogue of the Insects collected in 1887-88'. *Proceedings of the United States National Museum*, 12(1) 201-211.

Mayr, E. 1942. *Systematics and the Orgin of Species.* Columbia University Press. Cambridge.

Mayr, E. & Peter D. Ashlock. 1991. *Principles of Systematic Zoology* (2nd edition). McGraw Hill Press, New York.

Mazzotti, L. & M.A. Bravo-Becherelle. 1963. Scorpionism in he Mexican Republic. In H.L. Keegan and W.W. McFarland, eds., *Venomous and poisonous animals and noxious plants of the pacific area.,* pp 119-131. New York: Pergamon.

Miyamoto, M. M. and J. Cracraft. 1991. Phylogenetic inferance, DNA sequence analysis, and the future of molecular systematics. In *Phylogenetic analysis of DNA sequences.* (M.M. Miyamoto and J. Cracraft, eds.), pp. 3-17. Oxford University Press, New York.

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

Mullis, K. 1990. The unusual origin of the polymerase chain reaction. *Scientific American* April 56-65

Nei, M. & Kumar, S. 2000. *Molecular Evolution and Phylogenetics*. Oxford University Press.

Page, R. D. M. 1996. TREEVIEW: An application to display phylogenetic trees on personal computers. *Computer Applications in the Biosciences* 12: 357-358.

Pocock, R.I. 1902. A contribution to the systematics of scorpions. I. Some corrections in nomenclature. II. Notes on some species of *Parabuthus* contained in the British Museum. III. Descriptions of some new and old species. *Annals and Magazine of Natural History,* (7), 10:364-380.

Polis, G.A. *The Biology of Scorpions*. 1990. Stanford University Press. Stanford, California.

Ponce, J. & C. R. Beutelspacher, 2001. *Alacranes de Michoacán*. Universidad Michoacana de San Nicholás de Hidalgo, Morelia, Michoacán. 103p

Posada, D. & K. A. Crandall. 1998. MODELTEST: testing the model of DNA substitution. *Bioinformatics* 14 (9): 817-818

Rolfe, W.D. 1980. Early invertebrate terrestrial faunas. In A. Panchen, Ed, *The terrestrial environments and the origin of land vertebrates.*, pp. 117-157. Systematics Association Special Volume No. 15 London : Academic.

Saitou, N. & M. Nei. 1987. The neighbor joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol Evol.* 24: 189-204.

Santiago-Blay, J. A. & Poinar, G. O., Jr. 1993. First scorpion (Buthidae: *Centruroides*) from Mexican amber (Lower Miocene to Upper Oligocene). *J. Arachnol.* **21**: 147–151.

Savory, T.H. 1977. *The Arachnida. 2nd Edition.* London : Academic. 340 pp.

Schawaller, W. 1979. Erst Nachweis eines Skorpions im dominikanischen Bernstein. *Stuttgarter Beitr. Naturk.*, ser. B, **45:** 1–15.

Simard, M. & Watt, D. D. 1990. Venoms and toxins. *In* G. A. Polis (ed.). *The biology of scorpions.* Stanford, CA: Stanford University Press: 414–444.

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

Sissom, W.D. & W.R. Lourenco. 1987 The Genus *Centruroides* in South America (Scorpiones, Buthidae). *Journal of Arachnology*, 15: 11-28.

Swofford, D. L., 1998. PAUP*. Phylogenetic analysis using parsimony (*and other methods). Version 4. Sunderland: Sinauer Associates.

Stahnke, H.L. 1972. UV light, a useful field tool. *Bioscience* 22: 604-607.

Stahnke, H.L & M. Calos. 1977. A key to the species of the genus *Centruroides* Marx (Scorpionida Buthidae). *Entomological News* 88: 5 & 6 : 111-120, May & June

Sober, E. 1988. *Reconstructing the past: Parsimony, evolution, and inferance.* MIT Press, Cambridge, M.A.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F. and Higgins, D.G. (1997) The *ClustalX* windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research*, 24:4876-4882.

Appendix

Appendix A

Recipes and Formulas

50x TAE (Tris, Acetic Acid and EDTA) Buffer
For 1 Liter

40 mM Tris Base	193.8g
40 mM Acetic Acid	45.7g
1 mM EDTA	14.8g

PH=8.0
Stored at room temperature

Ethidium Bromide (EtBr)

10mg/ml concentration

H_2O	10mL
Ethidium bromide	0.1 gram

Stored at room temperature

Primers

100μM Stock
10μM use
1/9 One part primer/ Nine part H_2O

Store in freezer

Appendix B

Extracting DNA from Animal Tissue

1. Remove specimen from ethanol. Using forceps (and gloved hands), place the forceps where the legs join the main body, and pull two legs off of the scorpion.
2. Pour some liquid nitrogen into a mortar, and add the legs. Using the pestle slowly rock / grind the legs into small pieces.
3. The rest of the protocol is taken from the Qiagen Dneasy tissue kit and is used as a modified and listed below. Using the now ground tissue, place in a 1.5-ml micro-centrifuge tube, and add 180 µl Buffer ATL. Ensure the correct amount of starting material is used (see page 8). For tissues such as spleen with a very high number of cells for a given mass of tissue, no more than 10 mg starting material should be used. It is advisable to cut the tissue into small pieces to enable more efficient lysis.
4. Add 20 µl Proteinase K, mix by vortexing, and incubate at 55°C until the tissue is completely lysed. Vortex occasionally during incubation to disperse the sample, or place in a shaking water bath or on a rocking platform. Lysis time varies depending on the type of tissue processed. Lysis is usually complete in 1–3 h. If it is more convenient, samples can be lysed overnight; this will not affect them adversely. After incubation, the lysate may appear viscous but should not be gelatinous as it may clog the DNeasy mini column.
5. Add 4 µl of RNase A (100 mg/ml), mix by vortexing, and incubate for 2 min at room temperature. Transcriptionally active tissues such as liver and kidney contain high levels of RNA, which will copurify with genomic DNA. If RNA-free genomic DNA is required, carry out this optional step. If residual RNA is not a concern, omit this step.
6. Vortex for 15 sec. Add 200 µl Buffer AL to the sample, mix thoroughly by vortexing, and incubate at 70°C for 10 min. It is essential that the sample and Buffer AL are mixed immediately and thoroughly by vortexing or pipetting to yield a homogeneous solution. A white precipitate may form on addition of Buffer AL, which in most cases will dissolve during the incubation at 70°C. The precipitate does not interfere with the DNeasy procedure. Some tissue types (e.g., spleen, lung) may form a gelatinous lysate after

addition of Buffer AL. In this case, vigorously shaking or vortexing the preparation before addition of ethanol in step 7 is recommended.

7. Add 200 µl ethanol (96–100%) to the sample, and mix thoroughly by vortexing. It is important that the sample, Buffer AL, and the ethanol are mixed thoroughly to yield a homogeneous solution. A white precipitate may form on addition of ethanol. It is essential to apply all of the precipitate to the DNeasy spin column.
8. Pipette the mixture from step 4 into the DNeasy mini column sitting in a 2-ml collection tube. Centrifuge at ³6000 x g(8000 rpm) for 1 min. Discard flow-through and collection tube.
9. Place the DNeasy mini column in a new 2-ml collection tube, add 500 µl Buffer AW1, and centrifuge for 1 min at ³6000 x g (8000 rpm). Discard flow-through and collection tube.
10. Place the DNeasy mini column in a 2-ml collection tube, add 500 µl Buffer AW2, and centrifuge for 3 min at full speed to dry the DNeasy membrane. Discard flow-through and collection tube. This centrifugation step ensures that no residual ethanol is carried over during the following elution. Discard flow-through and collection tube. Following the centrifugation step, remove the DNeasy spin column carefully so that the column does not contact the flow-through, since this will result in carryover of ethanol. (If carryover of ethanol occurs, empty the collection tube and reuse it in another centrifugation step for 1 min at full speed.)
11. Place the DNeasy mini column in a clean 1.5-ml or 2-ml microcentrifuge, and pipette 200 µl Buffer AE directly onto the DNeasy membrane. Incubate at room temperature for 1 min, and then centrifuge for 1 min at 6000 G (8000 rpm) to elute. Elution with 100 µl (instead of 200 µl) increases the final DNA concentration in the elute, but also decreases the overall DNA yield (see Figure 3 on page 12).
12. Repeat elution once as described in step 11.

Appendix C

Setting up and running a Thermocycler Reaction
Wear Gloves at all times

1. Pull needed ingredients out of the freezer and start them thawing.
2. Prepare an ice bucket to keep the master mix in while mixing.
3. Measure out the following ingredients in the respective proportions. All ingredients are listed in µl. (For Eight different samples)

H_2O	300
$MgCl_2$	48
10x buffer	41
Primer 1	8
Primer 2	8
DNTP	8
Taq polymerase	2.0

4. Label a 1.5 ml microcentrafugue tube “master mix” in which all of the above ingredients will be placed.
5. Label all thin wall tubes that will be placed in the thermocycler.
6. Starting with the largest volume, add all ingredients except for the *Taq* polymerase, which will be added at the very end. Be sure to change pipette tips between each ingredient.
7. Add taq polymerase, be sure to pipette up and down to make sure that you have all of the *Taq* out of the pipette tip and into the master mix.
8. Vortex master mix for 5 seconds.
9. Pipette 46.5 µl out of the master mix tube into each of the waiting thin-walled tubes that are going into the thermocycler.
10. Add 3.5 µl of template DNA to the each of the corresponding thin-walled tubes, and close the lid on the tube. Tap the side gently to knock down anything that may be on the sidewall. When adding the template DNA be sure that you get the tip of the pipette all the way down to the bottom of the thin-walled tube.
11. Place the thin-walled tubes into the thermocycler and press start.
12. When the thermocycler is finished remove the thin-walled tubes from the machine and secure them in a rack and place the rack in a refrigerator.

Appendix D

How to prepare for and run a gel electrophoresis

Pouring the Gel

1. Measure out 70 ml of TAE buffer
2. Weigh out .70 grams of Agarose Low EED (Fisher Biotech)
3. Add both TAE buffer and .70 grams of Agarose to a small Earlameirer flask.
4. Microwave flask for 45 seconds until agarose is in solution. This usually happened when the contents of the flask have come to a boil.
5. Remove from microwave, and wait until the flask is cool enough to touch with the bare hand; at this point add 2.2 ul of Ethidium Bromide Solution to the mixture.
6. Stir
7. Prepare casting tray by putting stoppers on the each end, and pour gel.
8. After having poured the gel, insert comb that is most appropriate (8, 12 or two combs), and wait for the gel to solidify. This can be determined by the opaqueness of the gel. When the gel is warm and still liquid it is clear, and solid it is opaque.

Running the Gel

1. Carefully remove the stoppers at the ends of the gel tray. Place the tray with the gel on it, into the electrophoresis box. Be sure that the well side of the gel is to the right, or the side with the negative electrode.
2. Fill the electrophoresis box with buffer until the gel is completely submersed. There should be a thin layer over the gel.
3. Remove the comb in the gel by gently moving the comb side to side, rocking it out of the wells. This prevents air bubbles in the wells. In the event that an air bubble occurs, use an empty pipette tip to dislodge the bubble.
4. On a piece parafilm, place N dots of 3µl of loading dye, (N is the number of PCR samples that are going to be examined.
5. With a new tip each time, add 4µl of your PCR product to one of the dots of loading dye, and mix by pipetting up and down. The solution should turn blue in color.

6. With the mixture in the pipette tip, carefully position the tip over the well. Release the mixture into the well.
7. Change tips and repeat. Once all samples are loaded, place electrodes / lid on electrophoresis box, and connect the box to the power supply.
8. Turn power supply on, and set to about 90 Milliamps. Let the gel run for 25 min, or 2/3 of the length of the gel.
9. Turn the power off and unplug leads from power supply. Remove lid and take gel out of the buffer.
10. Place gel on the transilluminator box. Put on eye protection, and turn on the UV light tray.
11. Look at your results, and take a picture. To take a picture, place camera over the gel and hold trigger for 2 seconds.
12. Pull the tab on the film (near the top of the camera) and wait for 30 seconds. Remove picture from the chemical sheet.

Appendix E

Using ClustalX for sequence alignment 1.81

1. Open notepad. You cannot use Microsoft Word for the preparation of the input file. The file needs to be saved as a *.txt file.
2. Each sequence needs to entered as follows:

Start:	Each sequence starts with the ">" symbol.
Label:	The next ten character spaces directly after the start symbol are assigned to the sequence as the name or label. I suggest that you do not use anything longer than 8 characters.
Space:	This space separates the data label from any other information that you may want to include.
Extra data:	It is important to note that this information will not be shown anywhere other than the input file. PAUP, and CLUSTALX will not use this data. This is the chance to enter any other data that you would like to see.
Hard Return:	Enter a hard return. This separates the label information from the actual sequence itself.
Sequence data:	Here is the sequence data
Hard return:	Hard return and start again for the next sequence.

There is no symbol or character needed to indicate the end of a file. Please see the example input file below.

>123asdf Location data
GTCGTCGTCGTCGTCGTCGTCGTCGTCGT……..etc.

>234asdf location data
CTGCTGCTGCTGCTGCTGCTGCTGCTGCT……..etc.

3. Once the input file has been created, save the file as *.txt.
4. Open CLUSTALX 1.81. Open the "File" menu, and select "load sequences". This will bring up a browsing window, find your file and click open.
5. Open the "Alignment" menu and select "Output format options". Select the nexus format, and click close

Open the "Alignment" menu and select "Do complete alignment" At this point CLUSTALX will generate the alignment and save it to the disk. The file generated can be used right away by PAUP.

Appendix F

DNA Sequences

All sequences used or generated for this thesis are listed in the following appendix. The sequences have also been submitted to GenBank, a public database for sequences, and can be found at the following address :

http://www.ncbi.nlm.nih.gov/

16S Sequence accension numbers :

CiiZU1	AF439753	CllAR1	AF439760
CioEC1	AF439754	CllTO1	AF439761
CioLC1	AF439755	CllTZ1	AF439762
CioMo1	AF439756	CllHU1	AF439763
CiiTI1	AF439757	CllQU1	AF439764
CllCH1	AF439758	CltFB1	AF439765
CllCH2	AF439759	CltFB2	AF439766

Cytochrome Oxidase 1 accension Numbers:

CiiZU1	CllAR1
CioEC1	CllTO1
CioLC1	CllTZ1
CioMo1	CllHU1
CiiTI1	CllQU1
CllCH1	CltFB1
CllCH2	CltFB2

Paper label	our label	ponce	Paper label	our label	ponce
CiiZU1	(171)	[A]	CllAR1	(181)	[I]
CioEC1	(176)	[E]	CllTO1	(177)	[H]
CioLC1	(172)	[B]	CllTZ1	(180)	[F]
CioMo1	(175)	[D]	CllHU1	(182)	[K]
CiiTI1	(179)	[C]	CllQU1	(173)	[G]
CllCH1	(178)	[J]	CltFB1	(174)	[M]
CllCH2	(184)	[N]	CltFB2	(183)	[L]

Appendix G

Primer Information

CO I primers
Ron

Length : 23
Sequence: GGA TCA CCT GAT ATA GCA TTC CC
Td : 66.0
Tm: 71.8 (%GC method)
Tm: 68.0 (2 * (A + T) + 4*(GC))

A + T = 12 (52.2%)
G + C = 11 (47.8%)

Nancy

Length: 26
Sequence: CCC GGT AAA ATT AAA ATA TAA ACT TC
Td: 64.1 (nearest neighbor method)
Tm: 66.6 (%GC method)
Tm: 66.0 (2 * (A + T) + 4*(GC))

A+T = 19 (73.1 %)
G+C = 7 (26.0 %)

LCO I

Length: 24
Sequence: GGT CAA CAA ATC ATA AAG ATA TTG
Td: 60.2 (nearest neighbor method)
Tm: 65.3 (%GC Method)
Tm: 62.0 (2* (A + T) + 4*(GC))

A + T = 17 (70.8 %)
G + C = 7 (29.2 %)

16S Primers

38

Length:	18
Sequence:	CGA TTT GAA CTC AGA TCA
Td:	52.3 (nearest neighbor method)
Tm:	59.9 (%GC method)
Tm:	50.0 (2* (A + T) + 4*(GC))

A + T = 11 (61.1 %)
G + C = 7 (38.9 %)

40

Length:	20
Sequence:	GTG CAA AGG TAG CAT AAT CA
Td:	58.0 (nearest neighbor method)
Tm:	64.2 (%GC method)
Tm:	56.0 (2* (A + T) + 4*(GC))

A + T = 12 (60.0 %)
G + C = 8 (40.0 %)

Curriculum Vitae

William Ian Towler
1226 7th Street
Huntington, Wv, 25701
(304)-697-5471
ian.towler@marshall.edu

Education

1996 - 2000	B.S., Zoology, Miami University, Oxford, Ohio.
2000 - 2002 (anticipated)	M.S., Biology, Marshall University, Huntington, West Virginia.

Publications & Presentations

Fet, V., M. E. Soleglad, B. Gantenbein, W. Ian Towler, & Elizabeth V. Fet. DNA analysis indicates a new phylogeny of *Euscorpius* (Scorpiones: Euscorpiidae) from Greece. *Association of Southeastern Biologists 63rd Annual Meeting, April 2002.* Applichian State University.

Towler, W. I., Ponce Saavedra, J., B. Gantenbein, B. & V. Fet. 2001. Mitochondrial DNA reveals a divergent phylogeny in Tropical *Centruroides* (Scorpiones: Buthidae) from Mexico. Biogeographica, 77(4): 157-172.

Towler, W. I., V. Fet, Ponce Saavedra, J., & B. Gantenbein, B. 2001. Genetic Divergence in Central Mexican *Centruroides limpidus* and *Centruroides infamatus* (Scorpiones: Buthidae) as revealed by 16S mitochondrial DNA. *American Arachnological Society 25th Annual Meeting, July 2001.* Keene State College, Keene, New Hampshire.

Towler, W. I., Ponce Saavedra, J., B. Gantenbein, B. & V. Fet. 2001. Mitochondrial DNA systematics of Central Mexican *Centruroides* (Scorpiones: Buthidae). *Association of Southeastern Biologists 62nd Annual Meeting, April 2001.* Tulane / Loyola University, New Orleans, Louisiana. Published in Southeastern Biology Vol 48, Number 2, 4/2001, Pg.100.

Towler, W. I., Lushnikova, T., Bailey, T., Payton, H. & V. Fet. 2001. Phylogeny of the Eurasian species of *Mustela* (Carnivora: Mustelidae) derived from 16S rRNA gene of mitochondrial DNA. *21st Annual Midwest Ecology & Evolution Conference, March 2001.* Miami University, Oxford, Ohio.

In preparation:
Gantenbein, B., Fet, V., Towler, W. I. & Gromov, A. V. Mitochondrial DNA indicates high divergence in *Mesobuthus* (Scorpiones: Buthidae) from Western and Central Asia.

Peles, J. D., Towler, W. I., and S. I. Guttman population Genetic Structure of Earthworms (Lumbricus rubellus) in Soils Contaminated by Heavy Metals from Sewage Sludge.

Grants & Awards

2001 **Marshall University NASA Space Grant**. DNA evolution of Mexican scorpions. Marshall University, $1,000.00

1999 **Undergraduate Research Award.** A genetic key for species identification with the mayfly genus *Stenonema*: mitochondrial DNA Restriction Fragment Length Polymorphism Analysis. Miami University; $2,100.

1998-99 **Undergraduate Summer Scholar.** A genetic key for species identification with the mayfly genus *Stenonema*: mitochondrial DNA Restriction Fragment Length Polymorphism Analysis. Miami University; $400.00.

Relevant Experience

2000-Present **Research Assistant,** Department of Biology, Marshall University.
(Dr Victor Fet)

- Conducted Phylogenetic research on scorpions using mitochondrial DNA sequence analysis
- Conducted phylogenetic research on mammals in the mink family (Mustelidae), using mitochondrial and nuclear DNA sequence analysis
- Conducted phylogenetic research on spiders using mitochondrial DNA sequence analysis

Summer 2000 **Laboratory Manager**, Department of Zoology, Miami University.
(Dr. Sheldon Guttman)

- Organized researchers and research equipment, workers.
- Maintained inventory of stock chemicals and materials.

1997-2000 **Field Assistant**, Department of Zoology, Miami University
(Dr. Sheldon Guttman)

- Conducted field research to measure loss of genetic variability in fish species in Clermount County watersheds.

1997-2000 **Laboratory Assistant,** Miami University
(Dr. Sheldon Guttman)

- Participated in research measuring genetic diversity in Ohio fishes, insects, and freshwater mussels.

Courses Taught

Lab Instructor for :

Spring 2002 Principles of Biology, BSC 121, Marshall University
Fall 2001 Principles of Genetics, BSC 324, Marshall University
Spring 2001 Principles of Genetics, BSC 324, Marshall University
Fall 2000 Introduction to Biology, BSC 104, Marshall University

Scientific Skills

- Starch Gel Electrophoresis & Staining
- Cellulose Acetate Gel Electrophoresis & Staining
- DNA Extraction & PCR Amplification
- DNA Restriction Fragment Length Polymorphism Analysis
- Radiation worker certification, NRC
- Primer Designing
- DNA Sequencing
- Photodocumentation
- Electrofishing

Professional Societies

American Arachnological Society
Association of Southern Biologists

Other

Boy Scouts of America

- 1992 Eagle Scout
- 1995 Vigil Honor, Order of the Arrow

Alternative Spring Break

- 1999 Miami University. Participated in a week of service in Seattle, Washington through America-Corps City Year.
- 2000 Miami University. Planned and organized a week of service in Atlanta, Georgia through Hands on Atlanta.

Computer Skills

- Proficiency with Microsoft Office 2000 professional (Word, Excel, PowerPoint, Access)
- Problem Solving / Trouble shooting skills with Windows 95, 98, and ME.
- Basic system skills in with Linux (general trouble shooting)
- C++ Programming skills.
- PAUP 4.08B
- ClustalX 1.81

Appendix H

MEXICO COI raw data

CiiZU1 (171) [A] {835}

AGNGGTTCGCCCCATAGAATAACNTATAGCCGNATGGCTGANNATNNAATAGGTATGCCNCT
CCTCCAGTCTAGGGATCAAAAAATAGTAGGTATGTAAAATTCCGATCAGTCAACANCATNGT
AATNGCCCCAGCCAACACAGGAAGAGACAACAAAAGCAAAACAGCAGTAACCATTACAGATC
AAACAAACAATGGAAGACGATCCAACCCCATACCACTTCTCCGCATATTAATAATAGTGGTA
ATAAAATTAATAGCCCCCAAAATAGAAGAAACACCAGCTAAATGCAAAGAAAAAATAGTAAG
GTCAACCGAACCCCCAACATGAGCCAAAGAAGAAGAAAGAGGGGGGTAGACTGTTCAGCCAG
TCCCCGCCCCTCTTTCTAATGCTGCAGAAGAGAGCAAAAGAAAAAAAGCAGGAGGAAGTAGT
CAAAAACTCATATTATTTATCCGAGGGAAAGCTATATCTGGAGCACCTAACATCAAAGGAAC
CAATCAATTACCAAATCCACCAATTATAATAGGTATAACCATAAAAAAAATCATAACAAAAG
CATGAGCCGTCACAACTACATTATAAACTTGGTCATCCCCAATCAAAGATCCCGGCATTCCA
ATCTCCCCACGAACAATCAAACTCAAAGCAGTCCCAACCATAGAAGCTCAAATCCCTAACAC
CAAATACATAGTACCAATATCTTTATGTTGGNNNNGNCCNANNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN

CioLC1 (172) [B] {836}

TAGTGGTNNGCTCCATAGAATACCNNTGTNAATGCTGANTTNNNATAGTGGTNCCCCCCCAC
CAGTCAGGGATCAAAANAAGAAGTATNAAAGATTCCGGTCCGTCATACAACATAGGAATNGC
TCCAGCTCAACACAGGAAGAGTCNACAAANGTTANACNGCNGTTTCNATCACNGATNAAACA
AACAAAGGGNGTCGATCCAACCCCATACCTACTTCTACGTATATTAATAATAGTATGTAATA
AAATTAATATGCCCCCAAAATAGTATGAAACACCAGCTAAATGTAAAGAAAAAATAGTAAGA
TCAACCGAACCCCCCCACATGAGCCAAAGAAGAAGAAAGAGGGGGGTAGACTGTTCAGCCAGT
CCCCGCCCCTCTTTCTAGTGCCGCAGAAGAAAGCAAAAGGAAAAAAGCAGGAGGAAGAAGCC
NGAAACTCATATTATTCATCCGAGGAAAAGCCATGTCCGGAGCACCTAACATCAAAGGAACC
AACCAATTACCAAATCCCCCAATTATAACAGGCATAACCNTAAAAAAAATCATAACAAAAGC
ATGAGCCGTNACAACTACATTATAAACCTGATCATCCCCAATCAAAGATCCCGGCATTCCAA
TCTCTCCACGAACAATTAATCTTAAAGCAGNTCCAACCATAGAAGCCCAAATTCCCAACACC
AAATACATAGTACCAATATCTTTATGATTNGNTNGNCCANNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN

CllQU1 (173) [G] {837}

AGATGNTGNTCTCCATAGTATNNGTATAGCTCNCATGGCTGGCNCNTCACGAAGAGTGTNGC
TGCCCCACACAGACGCTGTGTATAGAATTAAGGNANGTCTTAAAGTTCCGTATGCTGTCATA
CATCATAGTANTCGCCCCNGCCAACTACAGGCAAGGGATANTCNAAAGCTAAACAGCAGTAA
CCATCACAGACCAAGCAAACAAAGGATGACGATCCAAACCTATCCCACTTCTCCGCATATTA
ATAATAGTAGNCAATAAAATTAATAGCCCCTAAAATAGAAGAAACACCAGCCAAATGCAAAG
AAAAAATAGTAAGATCAACCGACCCCCCAACATGAGCCAAAGAAGAAGAAAGAGGGGGGTAG
ACTGTTCAGCCAGTCCCCGCCCCTCTTTCCAATGCAGCAGAGGAAAGCAAAAGAAAAAAAGC
CGGCGGAAGCAACCAAAAACTCATATTATTCATACGAGGAAAAGCCATGTCCGGAGCACCCA
ACATCAAGGGCACCAACCAATTCCCAAACCCACCAATTATAATGGGCATAACCATAAAAAAA

ATTATCACAAAAGCATGAGCTGTCACAACTACATTATAAACCTGATCATCCCCAATCAAAGA
GCCAGGCATACCAATCTCACCACGAACAATCAAGCTCAAAGCAGTCCCAACCATAGAAGCCC
AAATCCCTAATACTAAATACATAGTACCAATATCTTTATATNGGGNGGGAACCAANNNNNNN

CltFB1 (174) [M] {838}

GTANTGCTCATGATANCNANATAGNTNCATGTTGATATAAAATAGGATCCCCCCCACCAGCA
GGATCAAAAAAAGAAGTATTAAAATTCCGATCTGTTAACAATATGGTAATAGCCCCAGCCAA
CACAGGAAGAGACAATAAAAGCAAAACAGCAGTAACCATCACAGACCATACAAATAAAGGAA
GACGATCCAACCCCATTCCACTTCTCCGTATATTAATAATAGTAGTGATAAAATTAATAGCC
CCTAAAATAGAAGAAACACCAGCTAAATGCAAAGAAAAAATAGTAAGATCAACCGAACCCCC
AACATGAGCTAAAGAAGAAGAAAGAGGGGGGTAAACTGTTCAACCAGTCCCCGCCCCTCTTT
CTAATGCCGCAGAAGAAAGCAAAAGAAAAAAAGCAGGGGGAAGCAGCCAAAAACTCATATTA
TTCATTCGAGGAAAAGCCATATCAGGAGCACCTAACATTAAAGGAACTAATCAATTACCAAA
CCCCCCAATTATAATAGGCATAACCATAAAAAAAATTATAACAAAAGCATGAGCCGTCACAA
CCACGTTATAAACTTGATCATCTCCAATCAAAGACCCCGGCATCCCAATCTCCCCACGAACA
ATCAATCTCAAAGCAGTTCCAACCATAGAAGCTCAAATCCCTAAAACCAAATATATAGTACC
AATATCTTTATATTGGGNNNGNNCCANNN

CioMO1 (175) [D] {839}

TGATGATAGTCTCNATAGATNTGTTGTNCATGCTGAGTTAANATAGGGTGCCCCCCCACCAG
GCAGGNATCAAAAAAGGANGTATTAAAGATTCCGGTCCGTCGTACAACATAGTAATNGCTCC
AGCCAACACAGGAAGAGACAACAAAAGTATAGACTGCAGTAACCATCACAGATCAGACAAAC
AAAGGANGACGATCCAACTCGCATACCACTTCTACGTNTATTAATAATAGTAGTANATAAAA
TTAATAGCCCCCAAAATAGAAGAANCACCNGCTAAATGTAAAGAAAAAATAGTAAGATCAAC
CGAACCCCCCACATGAGCCAAAGAAGAAGAAAGAGGGGGGTAGACTGTTCAGCCAGTCCCCG
CCCCTCTTTCTAGTGCCGCAGAAGAAAGCAAAAGGAAAAAAGCAGGAGGAAGAAGCCAGAAA
CTCATATTATTCATCCGAGGAAAAGCCATATCCGGAGCACCTAACATCAAAGGAACCAACCA
ATTACCAAATCCCCCAATTATAATAGGCATAACCATAAAAAAAATCATAACAAAAGCATGAG
CCGTAACAACTACATTATAAACCTGATCATCCCCAATCAAAGATNCCGGCATTCCAATCTCT
CCACGAACAATTAATCTTAAAGCAGTCCAAACCATAGAAGCCCAAATTCCCAACACCAAATA
CATAGTACCAATATCTTTATGATTNGGNNNGANCNANNNNN

CioEC1 (176) [E] {840}

TGTNTNGTTNATGATTTGTTGTNCATGCTGAGTTAGANATAGGGTGCCCCCCCACCAGCAGG
ATCAAAAAAAGANGTATTAGAAATTCCGGTCCGTCAACAACATAGTAATAGCTCCAGCCAAC
ACAGGAAGAGACAACAAAAGTATAGACAGCAGTAACCATCACAGATCAAACAATCAAAGGAA
GACGATCCAACCCCATACCACTTCTACGTATATTAATAATAGTAGTAATAAAATTAATAGCC
CCCAAAATAGAAGAAACACCAGCTAAATGTAAAGAAAAAATAGTAAGATCAACCGAACCTCC
CACATGAGCCAAAGAAGAAGAAAGAGGGGGGTAGACTGTTCAGCCAGTCCCCGCCCCTCTTT
CTAGTGCCGCAGAAGAAAGCAACAGGAAAAAAGCAGGAGGAAGAAGCCNGAAACTCATATTA
TTCATCCGAGGAAAAGCCATGTGCCGGAGCACCTAACATCAAAGGAACCAACCAATTACCAA
ATCCCCCAATTATANTAGGCATANCCNTAAAAAAAATCATNACAGNAGCATGAGCCGTAACA
ACTACATTATAAACCTGATCATCCCCAATCAAAGATNCCGGCATTCCAATCTCTCCACGAAC
AATTAATCTTAAAGCAGTTCAANCCATAGAAGCCCAAATTCCCAACACCAAATACATAGTAC
CAATATCTTTATGATTGGGGTGGACCNANNNNN

ClITO1 (177) [H] {841}

TTACANGGTGCGTCCCCTAGGATCATTGTTAGTTTCATAGCNTTGTAGTATCGNNCACAAGG
TATNCCNTCNACNACGCAGACTGNGTGCTNNANATAGAGNAGGNTCTTAAAATTCCGATTCC
GTTTNTAACATAGTAATAGCTCCAGCACAGATACTGGCAAGAAGGACAACAAAAGCTNNACA
GCNGTNACCATTACAGATCAAACATTCAAAGGNTGACGATCCAACCCTATNCCACTACTTCG
CGTATTAATAATAGTAGTAATAAAATTAATAGCCCCCAAAATAGAAGAAACACCAGCCAAAT
GTAAAGAAAAAATAGTAAGATCAACCGACCCTCCTACATGAGCCAAAGATGAAGAAAGAGGG
GGGTATACTGTTCAACCAGTCCCCGCCCCTCTTTCCAATGCCGCAGAAGAAAGTAAAAGAAA
AAAAGCAGGAGGAAGCAACCAAAAACTCATATTATTCATACGAGGAAAAGCCATATCCGGAG
CCCCCAACATCAAAGGAACTAACCAATTCCCAAACCCTCCAATCATAATCGGCATAACTATA
AAAAAAATTATAACAAAAGCATGAGCCGTAACAACTACATTATAAACCTGATCATCCCCAAT
CAAAGACCCCGGCATCCCAATCTCACCTCGAACAATTAAACTCAAAGCAGTTCCAACTATAG
AAGCCCAAATTCCTAACACCAAATATATAGTCCCAATATCTTTATGATTGGGNTNGACCNAN
NNN

ClICH1 (178) [J] {842}

AGNNNNNTTTTACNGGTGGCTCCCNCNAGAANATATTGTATNGTNACACATCGCCTTGCACT
ATCTANTAANTANGTANNCCGNACACNNCACAGANTGTGTGATATNANAACAGAGNAGGCCT
TAAAATTCCGATCCGTCTATAACATAGNAATAGCTCCAGCCAATACTGGAAAACACAACAAA
AGCAAAACAGCAGTAACCATTACAGATCAAACAAACAAAGGAAGACGATCCAACCCTATNCC
ACTACTTCGCNTATTAATAATAGTAGTAATAAAATTAATAGCTCCTAAAATAGAAGAAACAC
CAGCCAAATGTAAAGAAAAAATAGTAAGATCAACCGACCCTCCTACATGAGCCAAGGATGAA
GAAAGAGGGGGGTATACTGTTCAACCAGTCCCCGCCCCTNTTTNCAATGCCGNANAANAAAN
CNAAANAAANNNNCNNNNNGNANNNNCCNAANNCTNATNTTTNNATNCNANNNNNNNNNNT
NNNNNNNNNNNNNNNNNNNNNNNTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNTNN
NNN

CiiTI1 (179) [C] {843}

GTGTNGCCCAAAAAATCAAAATNAATGCTGATATAAAATAGGGTCCCCCCCACCAGCAGGAT
CAAAAAAAGAAGTATTAAAATTCCGGTCCGTCAACAACATAGTAATAGCTCCAGCCAACACT
AGGAAGAGACAACAAAAGTAAGACAGCAGTAACTCATCACAGATCAAACAAACAAAGGAAGA
CGATCCAACCCCATACCACTTCTACGTATATTAATAATAGTAGTAATAAAATTAATAGCCCC
CAAAATAGAAGAAACACCAGCTAAATGTAAAGAAAAAATAGTAAGATCAACCGAACCCCCCA
CATGAGCCAAAGAAGAAGAAAGAGGGGGGTAGACTGTTCAGCCAGTCCCCGCCCCTCTTTCT
AGTGCCGTAGAAGAAAGCAAAAGGAAAAAAGCAGGAGGAAGAAGCCAGAAACTCATATTATT
CATCCGAGGAAAAGCCATGTCCGGAGCACCTAACATCAAAGGAACCAACCAATTACCAAATC
CCCCAATTATAATAGGCATAACCATAAAAAAAATCATAACAAAAGCATGAGCCGTAACAACT
ACATTATAAACCTGATCATCCCCAATCAAAGATCCCGGCATTCCAATCTCTCCACGAACAAT
TAATCTTAAAGCAGTTCCAACCATAGAAGCCCAAATTCCCAACACCAAATACATAGTACCAA
TATCTTTTTNGGNGGGGNCNCNANNN

ClITZ1 (180) [F] {844}

GGTTGCCCAAAAAATCAAAACAAATGCTGATACAAAATAGGATCCCCCCCACCAGCAGGATC
AAAAAAAGAAGTATTAAAATTCCGATCCGTCAACAACATAGTAATAGGCCCCAGCCAATACA
GGAAGAGACAACAAAAGCAAAACAGCAGTAACCATTACAGATCAAACAAACAAAGGAAGACG
ATCTAACCCTATTCCTCTACTTCGCATATTAATAATAGTAGTAATAAAGTTAATAGCCCCTA
AAATAGAAGAGACACCAGCCAAATGTAAAGAAAAAATAGTAAGATCAACCGACCCCCCTACA

TGAGCTAAAGAAGAAGAAAGAGGGGGGTACACTGTTCAACCAGTCCCCGCCCCTCTTTCCAA
TGCTGCAGAAGAAAGTAAAAGAAAAAAAGCAGGAGGAAGCAACCAAAAACTCATATTATTCA
TACGAGGAAAAGCCATATCTGGAGCCCCCAACATCAAAGGAACTAACCAATTTCCAAACCCC
CCAATCATAATTGGTATAACCATAAAAAAAATTATAACAAAAGCATGAGCCGTAACAACTAC
ATTATAAACCTGATCATCCCCAATTAAAGACCCCGGCATCCCAATCTCACCCCGAACAATTA
AACTCAAAGCAGTCCCAACTATAGAAGCCCAAATTCCCAACACCAAATATATAGTACCAATA
TCTTTTTNNNNNGGGNNCCCN

ClIAR1 (181) [I] {845}

TGGTNGCCCAAAAAATCAAAACNAATGCTGATACAAAATAGGATCCCCCCCACCAGCAGGAT
CAAAAAAAGAAGTATTAAAATTCCGATCCGTCAACAACATAGTAATGGCCCCAGCCAATACA
GGAAGAGACAACAAAAGCAAAACAGCAGTAACCATTACAGATCAAACAAACAAAGGAAGACG
ATCTAACCCTATTCCTCTACTTCGCATATTAATAATAGTAGTAATAAAGTTAATAGCCCCTA
AAATAGAAGAGACACCAGCCAAATGTAAAGAAAAAATAGTAAGATCAACCGACCCCCCTACA
TGAGCTAAAGAAGAAGAAAGAGGGGGGTACACTGTTCAACCAGTCCCCGCCCCTCTTTCCAA
TGCTGCAGAAGAAAGTAAAAGAAAAAAAGCAGGAGGAAGCAACCAAAAACTCATATTATTCA
TACGAGGAAAAGCCATATCTGGAGCCCCCAACATCAAAGGAACTAACCAATTTCCAAACCCC
CCAATCATAATTGGTATAACCATAAAAAAAATTATAACAAAAGCATGAGCCGTAACAACTAC
ATTATAAACCTGATCATCCCCAATTAAAGACCCCGGCATCCCAATCTCACCCCGAACAATTA
AACTCAAAGCAGTCCCAACTATAGAAGCCCAAATTCCCAACACCAATATATAGTACCAATAT
CTTATATTGGGGGGGNACCCAANNN

CltFB2 (183) [L] {846}

GGGNNAGTTCTACTNTTATCGCATTNCGNANGGCCAGATATAAAATAGGATCTCCCCCACCA
GCAGGATCAAAAAAAGAAGTATTAAAATTTCGATCTGTTAACAATATGGTAATAGCCCCAGC
CAACACAGGAAGAGACAATAAAAGCAAAACAGCAGTAACCATCACAGACCATACAAATAAAG
GAAGACGATCCAACCCCATTCCACTTCTTCGTATGTTAATAATAGTAGTGATAAAATTAATA
GCCCCTAAAATAGAAGAAACACCAGCTAAATGCAAAGAAAAAATAGTAAGATCAACCGAACC
CCCAACATGAGCTAAAGAAGAAGAAAGAGGGGGGTAAACTGTTCAACCAGTCCCCGCCCCTC
TTTCTAATGCCGCAGAAGAAAGCAAAAGAAAAAAAGCAGGGGGAAGCAACCAAAAACTCATA
TTATTCATTCGAGGAAAAGCCATATCAGGAGCACCTAACATTAAAGGAACTAATCAATTACC
AAACCCCCCAATTATAATAGGCATAACCATAAAAAAAATTATAACAAAAGCATGAGCCGTCA
CAACCACGTTATAAACTTGATCATCTCCAATCAAAGACCCCGGCATCCCAATCTCCCCACGA
ACAATCAATCTCAAAGCAGTTCCAACCATAGAAGCTCAAATTCCTAAAACCAAATATATAGT
ACCAATATCTTTTTNGNGGGGGNCCCCNNNNN

ClICH2 (184) [N] {847}

GGGTGCCCANAAATCAAAGTCAATGCTGATATAAAATAGGATCCCCCCCACCAGCAGGATCA
AAAAAAGAAGTATTAAAATTCCGATCCGTCAATAACATAGTAATAGCTCCAGCCAATACTGG
AAGAGACAACAAAAGCAAAACAGCAGTAACCATTACAGATCAAACAAACAAAGGAAGACGAT
CCAACCCTATCCCACTACTTCGCATATTAATAATAGTAGTAATAAAATTAATAGCTCCTAAA
ATAGAAGAAACACCAGCCAAATGTAAAGAAAAAATAGTAAGATCAACCGACCCTCCTACATG
AGCCAAAGATGAAGAAAGAGGGGGGTATACTGTTCAACCAGTCCCCGCCCCTCTTTCCAATG
CCGCAGAAGAAAGTAAAAGAAAAAAAGCAGGAGGAAGCAACCAAAAACTCATATTATTCATA
CGAGGAAAAGCCATATCCGGAGCCCCCAACATCAAAGGAACTAACCAATTCCCAAACCCTCC
AATCATAATCGGCATAACCATAAAAAAAATTATAACAAAAGCATGAGCCGTAACAACTACAT
TATAAACCTGATCATCCCCAATCAAAGACCCCGGCATTCCAATCTCACCTCGAACAATTAAA

CTCAAAGCAGTTCCAACTATAGAAGCCCAAATTCCTAATACCAAATATATAGTCCNATATCT
TNTTNGGGGGGGNNCCCCNNNNN

(83) {848}

GGGTGTCCAAAAAACCAAAACAAATGCTGATATAAAATAGGATCTCCGCCCCCANCAGGGTC
AAAAAAAGANGTATTAAAATTCCGATCCGTCAACAACATAGTAATAGCCCCTGCTAACACAG
GAAGAGACAATAACAACAAAACAGCAGTAACTATCACAGATCAAACAAACAAAGGAAGACGA
TCTAATCCCATACCACTTCTCCGTATATTAATAATAGTAGTAATAAAATTAATAGCCCCCAA
AATAGAAGAAACACCAGCCAAATGAAGAGAAAAAATAGTAAGATCAACTGAACCCCCAACAT
GAGCCAAAGAGGAAGAAAGAGGAGGATAGACCGTTCACCCCGTTCCTGCCCCTCTCTCCAGT
GCAGCAGAAGAAAGTAAGAGAAAAAAAGCAGGAGGAAGTAACCAAAAACTCATATTATTCAT
CCGAGGAAAGGCCATATCAGGAGCACCTAACATCAAGGGGACTAACCAGTTCCCAAAACCCC
CAATCATAATAGGCATGACTATAAAAAAAATTATAACAAAAGCATGGGCCGTCACAACAACA
TTATAAACTTGATCATCCCCAATCAAAGAACCTGGTATCCCAATCTCACCCCGAACAATTAA
ACTCAAAGCAATCCCAACCATAGAAGCCCAAATACCCAACACCAAATACATAGTACCAATAT
CTTTTTTNGGNGGGNNCCNCN

16 S data raw

***Cii*ZU1 (171) [A] {685}**

TNTTNNCGCCTNTTTTGGGTCCGAACAGACCCCCTTTTACTCCTCTTGCGGAATAGAGGAAATCTAAT
CCAACATCGAGGTCGCAAACATATTCGTCAATTTGAGCTTTAAGAATATATTACGCTGTTATCCCTAA
AGTAACTTGTTTAAGCTTCAAAAGTTTTGGGTATCAAAATAATGTTATCTTAATATTATGAAAGTGTT
TATCTTTCCACCGCCCCAGTAAAACATACTTTTAATTTATTAAATTATTATATGTAAAGCTTTATAGG
GTCTTCTTGTCTAAAAGACATATTTTAGCCTTTTTACTAAAAAGTAAATTTTAAAGAAAAAAGTTAAT
AAAGAAACTTTCTAGTTTATCCTTTCATTCCAGTCTTAAATTACAAGACTAATGATTATGCTACCTTT
GCACA
Bases :413

***Cio*EC1 (176) [E] {691}**

NGCCTTTTTACAGGTCGAACAGACCTCCTTTTACTCCTCTTGCGGAGTGAAGGAAATCTAATCCAACA
TCGAGGTCGCAAACGTATTTGTCAATTTGAACTTTAAAAATACATTACGCTGTTATCCCTAAAGTAAC
TTATTTAAGCTTCAAAAATTTTGGGTATCAAAATAATACTATTTTAATACTTTGAAAGTGTTTATCTT
TCCACCGCCCCAGTAAAACACACTTTTAATTTACTAAATTATTATATGTAAAGCTTTATAGGGTCTTC
TTGTCTAAAAGAGGTATTTTAGCCTTTTTACTAAAAGGTAAATTTTGAAGAAAAAAGCTAAGAAAGAA
ACTCTCTAGTTTATCCTTTCATTCCAGTCTTAAATTATAAGACTAATGATTATGCTACCTTTGCACAC
Bases :408

***Cio*LC1 (172) [B] {686}**

AGNCCACTTTACAGGTCCGAACAGACCTCCTTTTACTCCTCTTGCGGAGTGAAGGAAATCTAATCCAA
CATCGAGGTCGCAAACGTATTTGTCAATTTGAACTTTAAAAATACATTACGCTGTTATCCCTAAAGTA
ACTTATTTAAGCTTCAAAAATTTTGGGTATCAAAATAATACTATTTTAATACTTTGAAAGTGTTTATC
TTTCCACCGCCCCAGTAAAACACACTTTTAATTTACTAAATTATTATATGTAAAGCTTTATAGGGTCT
TCTTGTCTAAAAGAGGTATTTTAGCCTTTTTACTAAAAGGTAAATTTTGAAGAAAAAAGCTAAGAAAG
AAACTCTCTAGTTTATCCTTTCATTCCAGTCTTAAATTATAAGACTAATGATTATGCTACCTTTGCAC
A
Bases :409

***Cio*MO1 (175) [D] {809}**

GGTTTTTTTACAGGTCGAACAGACCTCCTTTTACTCCTCTTGCGGAGTGAAGGAAATCTAATCCAACA
TCGAGGTCGCAAACGTATTTGTCAATTTGAACTTTAAAAATACATTACGCTGTTATCCCTAAAGTAAC
TTATTTAAGCTTCAAAAATTTTGGGTATCAAAATAATACTATTTTAATACTTTGAAAGTGTTTATCTT
TCCACCGCCCCAGTAAAACACACTTTTAATTTACTAAATTATTATATGTAAAGCTTTATAGGGTCTTC
TTGTCTAAAAGAGGTATTTTAGCCTTTTTACTAAAAGGTAAATTTTGAAGAAAAAAGCTAAGAAAGAA
ACTCTCTAGTTTATCCTTTCATTCCAGTCTTAAATTATAAGACTAATGATTATGCTCTCTTTTNGCCA
CA
Bases :410

***Cii*TI1 (179) [C] {747}**

TNTTGGANACCNNTTNAGGACGCCAGACCTCCTTTTACTCCTCTTGCGGAGTGAAGGAAATCTAATCC
AACATCGAGGTCACAAACGTATTTGTCAATTTGAACTTTAAAAATACATTACGCTGTTATCCCTAAAG
TAACTTATTTAAGCTTCAAAAATTTTGGGTATCAAAATAGTACTATTTTGATACTTTGAAAGTGTTTA
TCTTTCCACCGCCCCAGTAAAACACACTTTTAATTTACTAAATTATTATATGTAAAGCTTTATAGGGT
CTTCTTGTCTAAAAGAGGTATTTTAGCCTTTTTACTAAAAGGTAAATTTTGAAGAAAAAAGCTAAGAA
AGAAACTCTCTAGTTTATCCTTTCATTCCAGTCTTAAATTATAAGACTAATGATTATGCTACCTTTGC
ACA

Bases :411

***Cll*CH1 (178) [J] {810}**

TGGTTTTATTTTATGNGTCGAACAGNACCCCCTTTTATTCCTCTTGCGGAATAGTGGAAATCTAATCC
AACATCGAGGTCGCAAACATATTTGTCAATTTGAACTTTCAAAATATATTACGCTGTTATCCCTAAAG
TAACTTATTTAAGCTTCAAAATTTTTGGGTATTAAAATGATGATATTTTAATATTGTAAAGGTGTTTT
ATCCTTCCACCGCCCCAGTGAAACATATTTTTAATCTATTAAATTATTTTATGTAAAGCTTTATAGGG
TCTTCTTGTCTAAAAGAAACATTTTAGCCTTTTTACTAAAAAGTAAAATTTAAAAGAAAAAGCTAAGA
AACAAGCTCTCTAGTTTATCCTTTCATTCCAGTCTTAAATTATAAGACTAATGATTATGCTCTTTTTT
GNCACA
Bases :414

***Cll*CH2 (184) [N] {813}**

GGTTTATTTTATGNGTCGAACAGNCCCCCTTTTATTCCTCTTGCGGAATAGTGGAAATCTAATCCAAC
ATCGAGGTCGCAAACATATTTGTCAATTTGGACTTTCAAAATATATTACGCTGTTATCCCTAAAGTAA
CTTATTTAAGCTTCAAAATTTTTGGGTATTAAAATGATGATATTTTAATATTGTAAAGGTGTTTTATC
CTTCCACCGCCCCAGTGAAACATATTTTTAATCTATTAAATTATTTTATGTAAAGCTTTATAGGGTCT
TCTTGTCTAAAAGAAACATTTTAGCCTTTTTACTAAAAAGTAAAATTTAAAAGAAAAAGCTAAGAAAC
AAGCTCTCTAGTTTATCCTTTCATTCCAGTCTTAAATTATAAGACTAATGATTATGCTCTTTTTTGGC
ACA
Bases :411

***Cll*AR1 (181) [I] {751,748}, [sequence is identical to 748]**

TTTTTNGAAGCCTTTATGGGTCGCCAGACCCCCTTTTATTCCTCTTGCGGAATAATGGAAATCTAATC
CAACATCGAGGTCGCAAACATATTCGTCAATTTGAACTTTCAAAATACATTACGCTGTTATCCCTAAA
GTAACTTATTTAATCTTCAAAATTTTTGGGTATTAAAATAATGATATTTTAATATTATAGAGGTGTTT
AATCCTCTCACCGCCCCAGTGAAACATAATTTTAATTTATTAAATTATTTTATGTAAAGCTTTATAGG
GTCTTCTTGTCTAAAAGAAGCATTTTAGCCTTTTTACTAAAAGGTAAAATTTAAAAGAAAAAATTAAG
AAAGAAGCTCTCTAGTTTACCCTTTCATTCCAGTCTTAAATTATAAGACTAATGATTATGCTACCTTT
GCACANCTC
Bases :417

***Cll*TO1 (177) [H] {692}**

CCANTCATGGGTCCGAACAGACCCCCTTTTATTCCTCTTGCGGAATAGTGGAAATCTAATCCAACA
TCGAGGTCGCAAACATATTTGTCAATTTGAACTTTCAAAATATATTACGCTGTTATCCCTAAAGTA
ACTTATTTAAGCTTCAAAATTTTTGGGTATTAAAATGATGGTATCTTAATATTGTAAAGGTGTTTC
ATCCTCCCACCGCCCCAGTGAAACATATTTTTAATTTATTAAATTATTTTATGTAAAGCTTTATAG
GGTCTTCTTGTCTAAAAGAAGCATTTTAGCCTTTTTACTAAAAAGTAAAATTTAAAAGAAAAAACT
AAGAAACAAGCTCTCTAGTTTATCCTTTCATTCCAGTCTTAAATTATAAGACTAATGATTATGCTN
CCTTTGC
Bases :403

***Cll*TZ1 (180) [F] {748}**

GAGCCCTTTGGGTCGNCAGACCCCCTTTTATTCCTCTTGCGGAATAATGGAAATCTAATCCAACATCG
AGGTCGCAAACATATTCGTCAATTTGAACTTTCAAAATACATTACGCTGTTATCCCTAAAGTAACTTA
TTTAATCTTCAAAATTTTTGGGTATTAAAATAATGATATTTTAATATTATAGAGGTGTTTAATCCTCT
CACCGCCCCAGTGAAACATAATTTTAATTTATTAAATTATTTTATGTAAAGCTTTATAGGGTCTTCTT
GTCTAAAAGAAGCATTTTAGCCTTTTTACTAAAAGGTAAAATTTAAAAGAAAAAATTAAGAAAGAAGC
TCTCTAGTTTACCCTTTCATTCCAGTCTTAAATTATAAGACTAATGATTATGCTACCTTTGCACA
Bases :405

CllHU1 (182) [K] {811}

GGTCTTTTTATCGNCNTCGACAGNACCTTCTTTTATCCCTCTTGCGGGATAGAGGAAGCTTAATCCAA
CATCGAGGTCGCAAACATGTTTGTCAATTTGGACTTTCGAAACACATTACGCTGTTATCCCTAAAGTA
ACTTATTTAAATTTCAAAATTTTTGGGTATTAAAATAGTTTTATTTTAGCCCTCTAAAAGTGTTTTAT
CTTTCCGCCGCCCCAGTGAAACATGTTTTTAATTTTTTAAATTATTTTATGTAAAGCTTTATAGGGTC
TTCTTGTCTAAAAGAGACATTTTAGCCTTTTTACTAAAAAGTAAAATTCAAAAGAAAAAACTAAGAAA
GAAACTTCCTGGTTTACCCCTTCATCCCAGTCTTAAATTATAAGACTAATGATTATGCTCTTTTTGNC
ACA
Bases :411

CllQU1 (173) [G] {687}
GCCCCTTTTATGGGTCGAACAGACCCCCATTTATTCCTCTTGCGGAATAGTGGAAGCTTAATCCAACA
TCGAGGTCGCAAACATATTTGTCAATTTGAGCTTTCAAAATACATTACGCTGTTATCCCTAAAGTAAC
TTATTTAAGCTTCAAAATTTTTGGGTATTAAAATAATTCTATTTTAGCCCTCTAAAAGTGTTTATCTT
TCCACCGCCCCAGTGAAACATATTCCTAATTTTTTAGATTATTATATGTAAAGCTTTATAGGGTCTTC
TTGTCTGAAAGAAATATTTTAGCCTTTTTACTAAAAAGTAAAATTCAAAAGGAAAAACTAAGAAAGAA
ACTTCCTGGTTTATCCCTTCATCCCAGTCTTAAATTATAAGACTAATGATTATGCTACCTTTGCACA
Bases :407

CltFB1 (174) [M] {688}
CTTNGCCACTTTATGGGTCGAACAGACCCCCTTTTTACTCCTCTTGCGGAATAAAGGAAATCTAAT
CCAACATCGAGGTCGCAAACATATTTGTCAATTTGAGCTTTAAAAATATATTACGCTGTTATCCCT
AAAGTAACTTATTTAAAGTTCAAAAATTTTGGGTATTAAAATAATATTATTTTAATGTTTTAGAAG
TGTTAGTCTTCTCACCGCCCCAGTGAAACATATTTTTAATTTATTAAATTATTATATGTAAAGCTT
TATAGGGTCTTCTTGTCTAAAAGACATATTTTAGCCTTTTTACTAAAAAGTAAAATTTAAAGAAAA
AAGTCNAGAAAGAAACTCTCCAGTTTATCCTTTCATTCCAGTCTTAAATTACAAGACTAATGATTA
TGCTACCTTTGC
Bases :408

CltFB2 (183) [L] {812}
GGTTTCTTTTATGNGTCGAACAGNCCNCCTTTTTACTCCTCTTGCGGAATAAAGGAAATCTAATCCAA
CATCGAGGTCGCAAACATATTTGTCAATTTGAGCTTTAAAAATATATTACGCTGTTATCCCTAAAGTA
ACTTATTTAAAGTTCAAAAATTTTGGGTATTAAAATAATATTATTTTAATGTTTTAGAAGTGTTAGTC
TTCTCACCGCCCCAGTGAAACATATTTTTAATTTATTAAATTATTATATGTAAAGCTTTATAGGGTCT
TCTTGTCTAAAAGACATATTTTAGCCTTTTTACTAAAAAGTAAAATTTAAAGAAAAAAGTCAAGAAAG
AAACTCTCCAGTTTATCCTTTCATTCCAGTCTTGAATTACAAGACTAATGATTATGCTCTTTTTNGNC
AC
Bases :410

16S Aligned

```
CllCH1      ---TGGTTTTATTTTATGNGTC-GAACAGNACCCCCTTTTATTCCTCTTG
CllCH2      -----GGTTTATTTTATGNGTC-GAACAGN-CCCCCTTTTATTCCTCTTG
```

```
CllTO1      ---------CCANTCATGGGTCCGAACAGA-CCCCCTTTTATTCCTCTTG
CllAR1      TTTTTNGAAGCCTTTATGGGTC--GCCAGA-CCCCCTTTTATTCCTCTTG
CllTZ1      -------GAGCCCTT-TGGGTC--GNCAGA-CCCCCTTTTATTCCTCTTG
CioEC1      ------NGCCTTTTTACAGGTC-GAACAGACCTCCTTTT-ACTCCTCTTG
CioMO1      ------GGTTTTTTTACAGGTC-GAACAGACCTCCTTTT-ACTCCTCTTG
CioLC1      -----AGNCCACTTTACAGGTCCGAACAGACCTCCTTTT-ACTCCTCTTG
CiiTI1      -TNTTGGANACCNNTTNAGGAC--GCCAGACCTCCTTTT-ACTCCTCTTG
CiiZU1      -TNTTNNCGCCTNTTTTGGGTCCGAACAGACCCCCTTTT-ACTCCTCTTG
CltFB1      ---CTTNGCCACTTTATGGGTC-GAACAGACCCCCTTTTTACTCCTCTTG
CltFB2      -----GGTTTCTTTTATGNGTC-GAACAGNCCNCCTTTTTACTCCTCTTG
CllHU1      ------GGTCTTTTTATCGNCNTCGACAGNACCTTCTTTTATCCCTCTTG
CllQU1      ------GCCCCTTTTATGG--GTCGAACAGACCCCCATTTATTCCTCTTG

CllCH1      CGGAATAGTGGAAATCTAATCCAACATCGAGGTCGCAAACATATTTGTCA
CllCH2      CGGAATAGTGGAAATCTAATCCAACATCGAGGTCGCAAACATATTTGTCA
CllTO1      CGGAATAGTGGAAATCTAATCCAACATCGAGGTCGCAAACATATTTGTCA
CllAR1      CGGAATAATGGAAATCTAATCCAACATCGAGGTCGCAAACATATTCGTCA
CllTZ1      CGGAATAATGGAAATCTAATCCAACATCGAGGTCGCAAACATATTCGTCA
CioEC1      CGGAGTGAAGGAAATCTAATCCAACATCGAGGTCGCAAACGTATTTGTCA
CioMO1      CGGAGTGAAGGAAATCTAATCCAACATCGAGGTCGCAAACGTATTTGTCA
CioLC1      CGGAGTGAAGGAAATCTAATCCAACATCGAGGTCGCAAACGTATTTGTCA
CiiTI1      CGGAGTGAAGGAAATCTAATCCAACATCGAGGTCACAAACGTATTTGTCA
CiiZU1      CGGAATAGAGGAAATCTAATCCAACATCGAGGTCGCAAACATATTCGTCA
CltFB1      CGGAATAAAGGAAATCTAATCCAACATCGAGGTCGCAAACATATTTGTCA
CltFB2      CGGAATAAAGGAAATCTAATCCAACATCGAGGTCGCAAACATATTTGTCA
CllHU1      CGGGATAGAGGAAGCTTAATCCAACATCGAGGTCGCAAACATGTTTGTCA
CllQU1      CGGAATAGTGGAAGCTTAATCCAACATCGAGGTCGCAAACATATTTGTCA

CllCH1      ATTTGAACTTTCAAAATATATTACGCTGTTATCCCTAAAGTAACTTATTT
CllCH2      ATTTGGACTTTCAAAATATATTACGCTGTTATCCCTAAAGTAACTTATTT
CllTO1      ATTTGAACTTTCAAAATATATTACGCTGTTATCCCTAAAGTAACTTATTT
CllAR1      ATTTGAACTTTCAAAATACATTACGCTGTTATCCCTAAAGTAACTTATTT
CllTZ1      ATTTGAACTTTCAAAATACATTACGCTGTTATCCCTAAAGTAACTTATTT
CioEC1      ATTTGAACTTTAAAAATACATTACGCTGTTATCCCTAAAGTAACTTATTT
CioMO1      ATTTGAACTTTAAAAATACATTACGCTGTTATCCCTAAAGTAACTTATTT
CioLC1      ATTTGAACTTTAAAAATACATTACGCTGTTATCCCTAAAGTAACTTATTT
CiiTI1      ATTTGAACTTTAAAAATACATTACGCTGTTATCCCTAAAGTAACTTATTT
CiiZU1      ATTTGAGCTTTAAGAATATATTACGCTGTTATCCCTAAAGTAACTTGTTT
CltFB1      ATTTGAGCTTTAAAAATATATTACGCTGTTATCCCTAAAGTAACTTATTT
CltFB2      ATTTGAGCTTTAAAAATATATTACGCTGTTATCCCTAAAGTAACTTATTT
CllHU1      ATTTGGACTTTCGAAACACATTACGCTGTTATCCCTAAAGTAACTTATTT
CllQU1      ATTTGAGCTTTCAAAATACATTACGCTGTTATCCCTAAAGTAACTTATTT

CllCH1      AAGCTTCAAAATTTTTGGGTATTAAAATGATGATATTTTAATATTGTAAA
CllCH2      AAGCTTCAAAATTTTTGGGTATTAAAATGATGATATTTTAATATTGTAAA
CllTO1      AAGCTTCAAAATTTTTGGGTATTAAAATGATGGTATCTTAATATTGTAAA
CllAR1      AATCTTCAAAATTTTTGGGTATTAAAATAATGATATTTTAATATTATAGA
CllTZ1      AATCTTCAAAATTTTTGGGTATTAAAATAATGATATTTTAATATTATAGA
CioEC1      AAGCTTCAAAAATTTTGGGTATCAAAATAATACTATTTTAATACTTTGAA
CioMO1      AAGCTTCAAAAATTTTGGGTATCAAAATAATACTATTTTAATACTTTGAA
CioLC1      AAGCTTCAAAAATTTTGGGTATCAAAATAATACTATTTTAATACTTTGAA
CiiTI1      AAGCTTCAAAAATTTTGGGTATCAAAATAGTACTATTTTGATACTTTGAA
CiiZU1      AAGCTTCAAAAGTTTTGGGTATCAAAATAATGTTATCTTAATATTATGAA
CltFB1      AAAGTTCAAAAATTTTGGGTATTAAAATAATATTATTTTAATGTTTTAGA
CltFB2      AAAGTTCAAAAATTTTGGGTATTAAAATAATATTATTTTAATGTTTTAGA
CllHU1      AAATTTCAAAATTTTTGGGTATTAAAATAGTTTTATTTTAGCCCTCTAAA
CllQU1      AAGCTTCAAAATTTTTGGGTATTAAAATAATTCTATTTTAGCCCTCTAAA
```

```
CllCH1      GGTGTTTTATCCTTCCACCGCCCCAGTGAAACATATTTTTAATCTATTAA
CllCH2      GGTGTTTTATCCTTCCACCGCCCCAGTGAAACATATTTTTAATCTATTAA
CllTO1      GGTGTTTCATCCTCCCACCGCCCCAGTGAAACATATTTTTAATTTATTAA
CllAR1      GGTGTTTAATCCTCTCACCGCCCCAGTGAAACATAATTTTAATTTATTAA
CllTZ1      GGTGTTTAATCCTCTCACCGCCCCAGTGAAACATAATTTTAATTTATTAA
CioEC1      AGTGTTT-ATCTTTCCACCGCCCCAGTAAAACACACTTTTAATTTACTAA
CioMO1      AGTGTTT-ATCTTTCCACCGCCCCAGTAAAACACACTTTTAATTTACTAA
CioLC1      AGTGTTT-ATCTTTCCACCGCCCCAGTAAAACACACTTTTAATTTACTAA
CiiTI1      AGTGTTT-ATCTTTCCACCGCCCCAGTAAAACACACTTTTAATTTACTAA
CiiZU1      AGTGTTT-ATCTTTCCACCGCCCCAGTAAAACATACTTTTAATTTATTAA
CltFB1      AGTGTTA-GTCTTCTCACCGCCCCAGTGAAACATATTTTTAATTTATTAA
CltFB2      AGTGTTA-GTCTTCTCACCGCCCCAGTGAAACATATTTTTAATTTATTAA
CllHU1      AGTGTTTTATCTTTCCGCCGCCCCAGTGAAACATGTTTTTAATTTTTTAA
CllQU1      AGTGTTT-ATCTTTCCACCGCCCCAGTGAAACATATTCCTAATTTTTTAG

CllCH1      ATTATTTTATGTAAAGCTTTATAGGGTCTTCTTGTCTAAAAGAAACATTT
CllCH2      ATTATTTTATGTAAAGCTTTATAGGGTCTTCTTGTCTAAAAGAAACATTT
CllTO1      ATTATTTTATGTAAAGCTTTATAGGGTCTTCTTGTCTAAAAGAAGCATTT
CllAR1      ATTATTTTATGTAAAGCTTTATAGGGTCTTCTTGTCTAAAAGAAGCATTT
CllTZ1      ATTATTTTATGTAAAGCTTTATAGGGTCTTCTTGTCTAAAAGAAGCATTT
CioEC1      ATTATTATATGTAAAGCTTTATAGGGTCTTCTTGTCTAAAAGAGGTATTT
CioMO1      ATTATTATATGTAAAGCTTTATAGGGTCTTCTTGTCTAAAAGAGGTATTT
CioLC1      ATTATTATATGTAAAGCTTTATAGGGTCTTCTTGTCTAAAAGAGGTATTT
CiiTI1      ATTATTATATGTAAAGCTTTATAGGGTCTTCTTGTCTAAAAGAGGTATTT
CiiZU1      ATTATTATATGTAAAGCTTTATAGGGTCTTCTTGTCTAAAAGACATATTT
CltFB1      ATTATTATATGTAAAGCTTTATAGGGTCTTCTTGTCTAAAAGACATATTT
CltFB2      ATTATTATATGTAAAGCTTTATAGGGTCTTCTTGTCTAAAAGACATATTT
CllHU1      ATTATTTTATGTAAAGCTTTATAGGGTCTTCTTGTCTAAAAGAGACATTT
CllQU1      ATTATTATATGTAAAGCTTTATAGGGTCTTCTTGTCTGAAAGAAATATTT

CllCH1      TAGCCTTTTTACTAAAAAGTAAAATTTAAAAGAAAAAGCTAAGAAACAAG
CllCH2      TAGCCTTTTTACTAAAAAGTAAAATTTAAAAGAAAAAGCTAAGAAACAAG
CllTO1      TAGCCTTTTTACTAAAAAGTAAAATTTAAAAGAAAAAACTAAGAAACAAG
CllAR1      TAGCCTTTTTACTAAAAGGTAAAATTTAAAAGAAAAAATTAAGAAAGAAG
CllTZ1      TAGCCTTTTTACTAAAAGGTAAAATTTAAAAGAAAAAATTAAGAAAGAAG
CioEC1      TAGCCTTTTTACTAAAAGGTAAATTTTGAAGAAAAAAGCTAAGAAAGAAA
CioMO1      TAGCCTTTTTACTAAAAGGTAAATTTTGAAGAAAAAAGCTAAGAAAGAAA
CioLC1      TAGCCTTTTTACTAAAAGGTAAATTTTGAAGAAAAAAGCTAAGAAAGAAA
CiiTI1      TAGCCTTTTTACTAAAAGGTAAATTTTGAAGAAAAAAGCTAAGAAAGAAA
CiiZU1      TAGCCTTTTTACTAAAAAGTAAATTTTAAAGAAAAAAGTTAATAAAGAAA
CltFB1      TAGCCTTTTTACTAAAAAGTAAAATTTAAAGAAAAAAGTCNAGAAAGAAA
CltFB2      TAGCCTTTTTACTAAAAAGTAAAATTTAAAGAAAAAAGTCAAGAAAGAAA
CllHU1      TAGCCTTTTTACTAAAAAGTAAAATTCAAAAGAAAAAACTAAGAAAGAAA
CllQU1      TAGCCTTTTTACTAAAAAGTAAAATTCAAAAGGAAAAACTAAGAAAGAAA

CllCH1      CTCTCTAGTTTATCCTTTCATTCCAGTCTTAAATTATAAGACTAATGATT
CllCH2      CTCTCTAGTTTATCCTTTCATTCCAGTCTTAAATTATAAGACTAATGATT
CllTO1      CTCTCTAGTTTATCCTTTCATTCCAGTCTTAAATTATAAGACTAATGATT
CllAR1      CTCTCTAGTTTACCCTTTCATTCCAGTCTTAAATTATAAGACTAATGATT
CllTZ1      CTCTCTAGTTTACCCTTTCATTCCAGTCTTAAATTATAAGACTAATGATT
CioEC1      CTCTCTAGTTTATCCTTTCATTCCAGTCTTAAATTATAAGACTAATGATT
CioMO1      CTCTCTAGTTTATCCTTTCATTCCAGTCTTAAATTATAAGACTAATGATT
CioLC1      CTCTCTAGTTTATCCTTTCATTCCAGTCTTAAATTATAAGACTAATGATT
CiiTI1      CTCTCTAGTTTATCCTTTCATTCCAGTCTTAAATTATAAGACTAATGATT
CiiZU1      CTTTCTAGTTTATCCTTTCATTCCAGTCTTAAATTACAAGACTAATGATT
CltFB1      CTCTCCAGTTTATCCTTTCATTCCAGTCTTAAATTACAAGACTAATGATT
```

```
CltFB2      CTCTCCAGTTTATCCTTTCATTCCAGTCTTGAATTACAAGACTAATGATT
CllHU1      CTTCCTGGTTTACCCCTTCATCCCAGTCTTAAATTATAAGACTAATGATT
CllQU1      CTTCCTGGTTTATCCCTTCATCCCAGTCTTAAATTATAAGACTAATGATT

CllCH1      ATGCTCTTTTTTG-NCACA--
CllCH2      ATGCTCTTTTTTG-GCACA--
CllTO1      ATGCTNCCTTTGC--------
CllAR1      ATGCTACCTTTGC-ACANCTC
CllTZ1      ATGCTACCTTTGC-ACA----
CioEC1      ATGCTACCTTTGC-AC-AC--
CioMO1      ATGCTCTCTTTTN-GCCACA-
CioLC1      ATGCTACCTTTGC-ACA----
CiiTI1      ATGCTACCTTTGC-ACA----
CiiZU1      ATGCTACCTTTGC-ACA----
CltFB1      ATGCTACCTTTGC--------
CltFB2      ATGCTCTTTTTNG-NCAC---
CllHU1      ATGCTCTTTTTGNCACA----
CllQU1      ATGCTACCTTTGC-ACA----
```

Label	*Sex*	*GPS*	*Sequence 16S*	*16S* accession *Number*	*Sequence CO1*
CiiTI1	Female	-101.3333,19.5500	747	AF439757	
CiiZU1	Female	-102.0500,19.3667	685	AF439753	
CioEC1	Male	-102.0333,19.1500	691	AF439754	
CioLC1	Male	-101.5833,19.8167	686	AF439755	
CioMo1	Female	-101.1167,19.7000	809	AF439756	
CllAR1	Male	-100.8833,18.5833	751	AF439760	
CllCH1	Female	-101.6333,18.6167	810	AF439758	
CllCH2	Female	-101.6333,18.6167	814	AF439759	
CllHU1	Female	-99.3500,18.3000	811	AF439763	
CllQU1	Female	-100.3833, 20.6000	687	AF439764	
CllTO1	Female	-100.8633,18.5633	692	AF439761	
CllTZ1	Female	-100.9167,19.5667	748	AF439762	
CltFB1	Female	-103.5000,18.6000	688	AF439765	
CltFB2	Male	-103.5000,18.6010	812	AF439766	

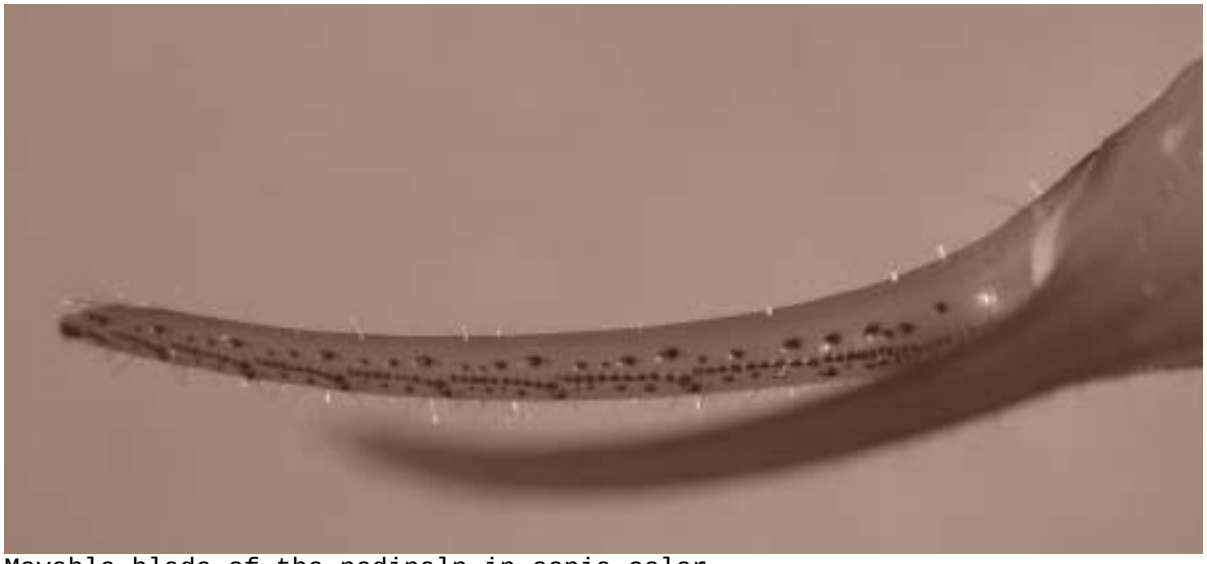
Appendix I

Pictures

Centruroides exilicauda

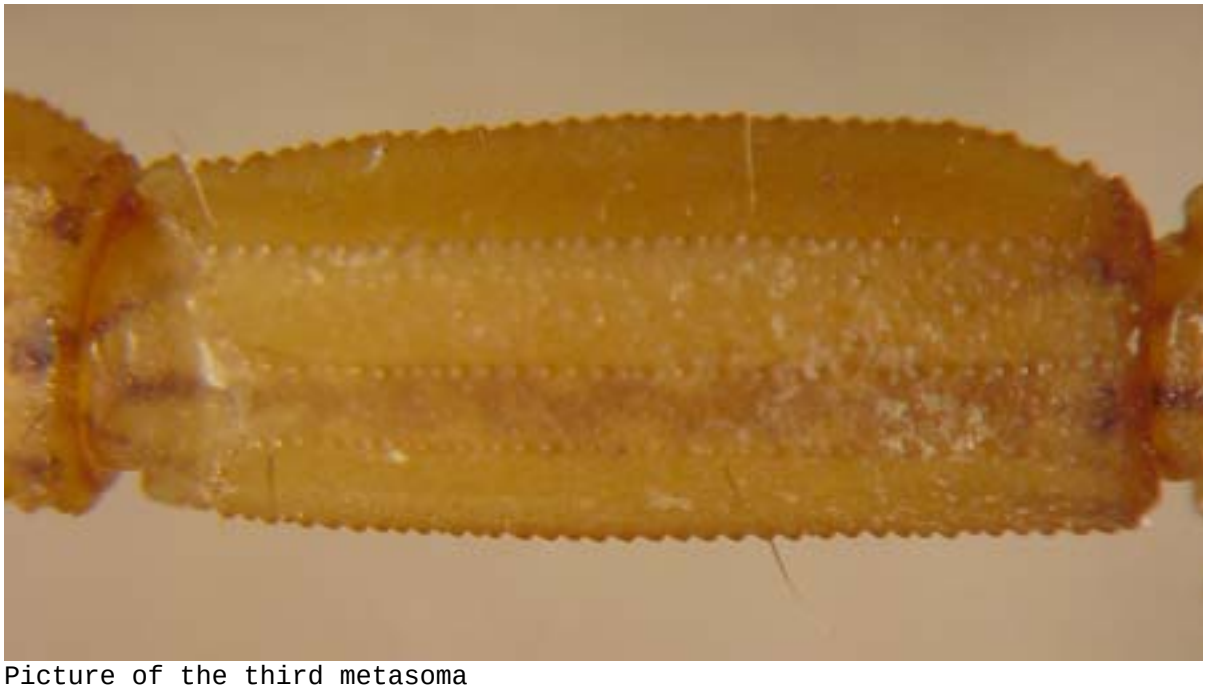
Collected Mexico, Baja del Sur.

Movable blade of the pedipalp in full color



Movable blade of the pedipalp in sepia color

Picture of the face showing carination and the eyes



Picture of the third metasoma

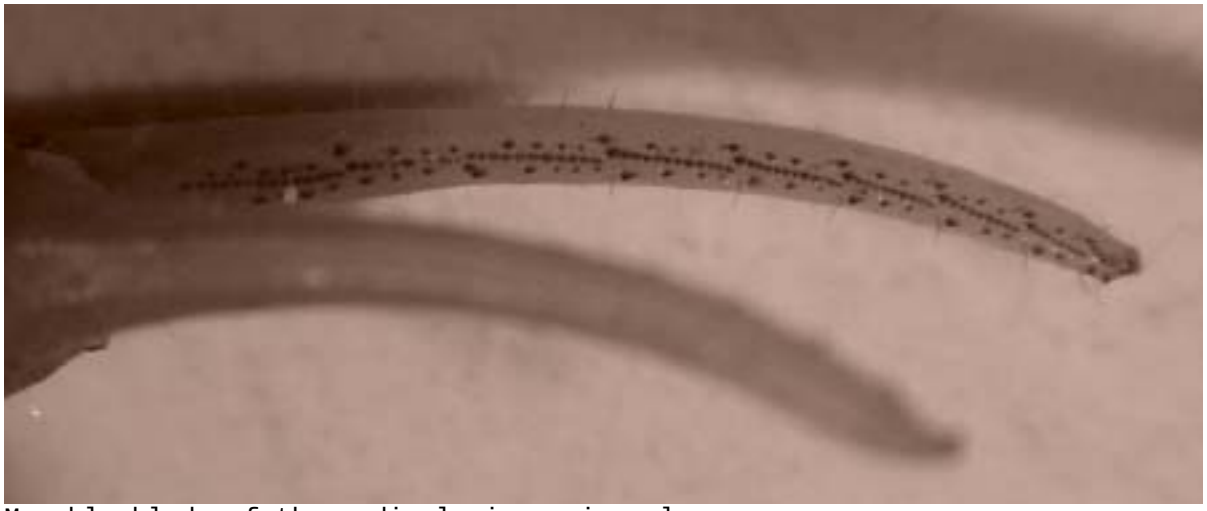
Picture of the telson including vesicle and aculeus

Another view of the face

Centruroides vittatus

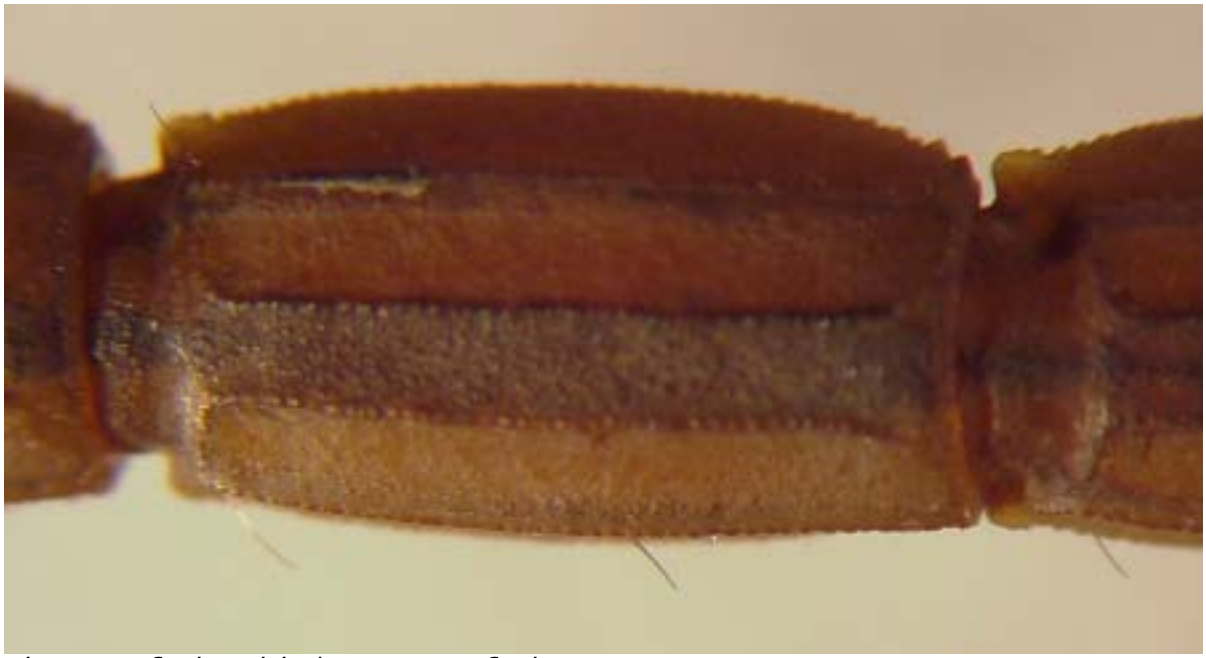
Collected Arkansas, Russelville

Movable blade of the pedipalp in full color



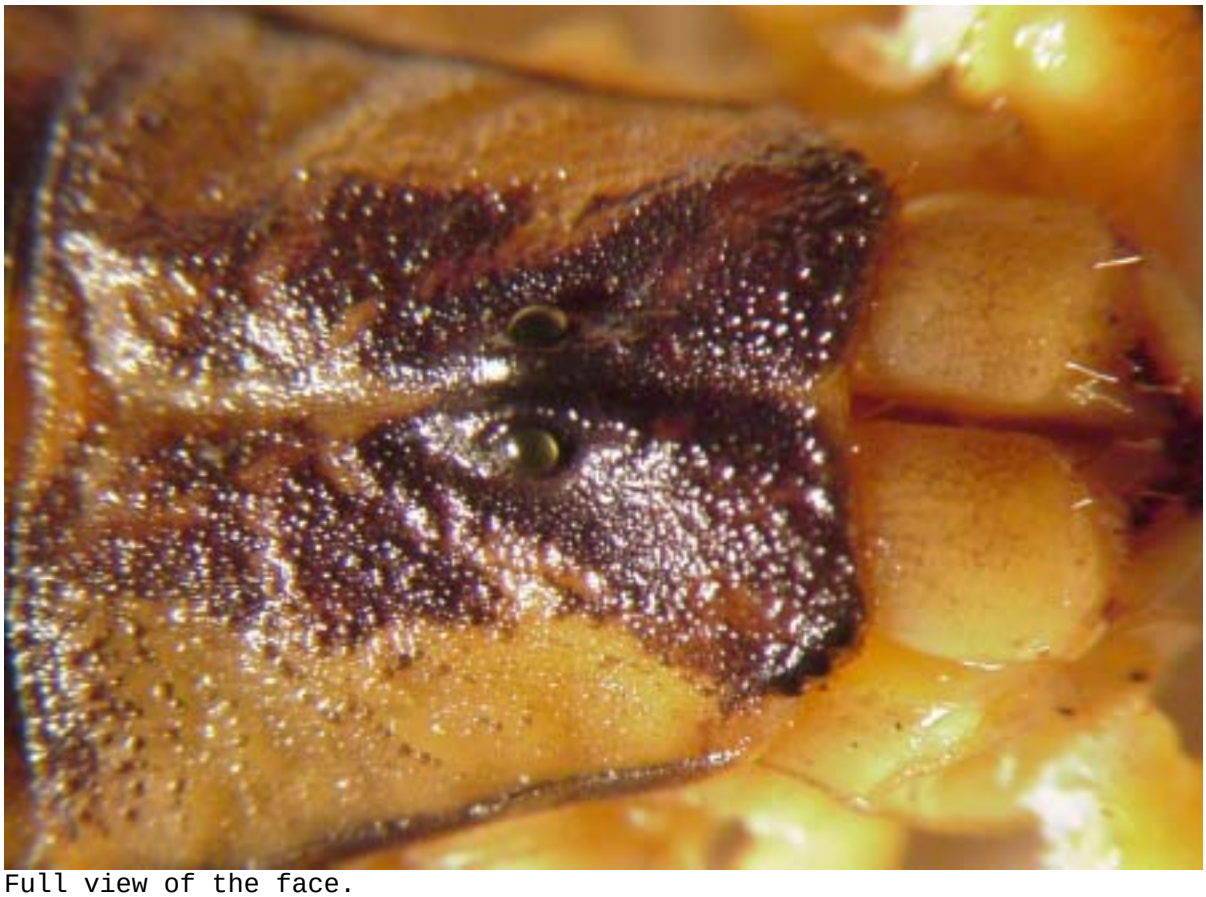
Movable blade of the pedipalp in sepia color

Picture of the face showing carination and eyes.



Picture of the third segment of the metasoma

Picture of the telson including vesicle and aculeus



Full view of the face.

Centruroides bani
Collected Hispaniola


Movable blade of the pedipalp in full color


Movable blade of the pedipalp in sepia color

Picture of the face showing carination and eyes

Picture showing the tarsus & ungues from the side

Picture showing the tarsus & ungues from above

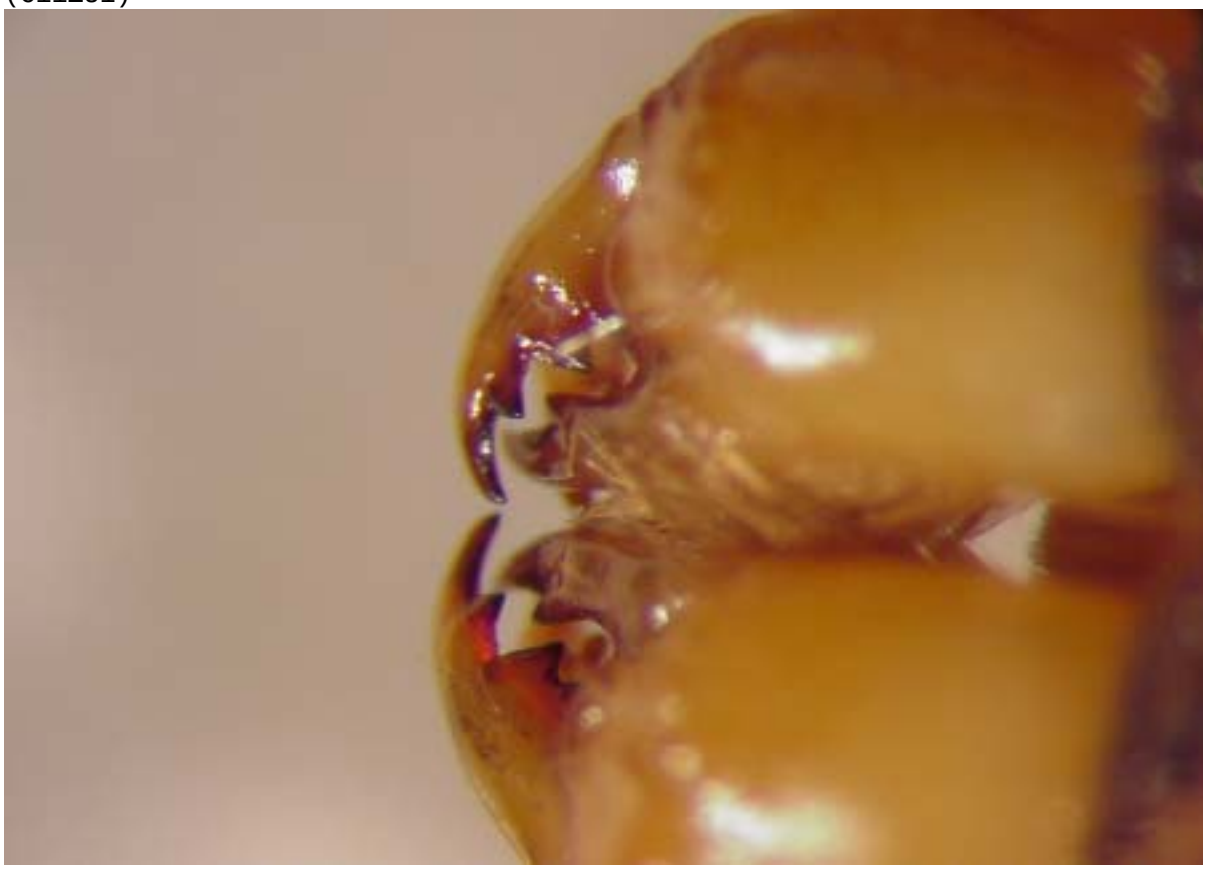
Picture of the third segment of the metasoma

Picture of the telson including vesicle and aculeus

Full view of the face.

Centruroides infamatus infamatus

(CiiZU1)



Picture of the chelicerae

Movable blade of the pedipalp in full color

Movable blade of the pedipalp in sepia color.

Picture if the spiracles and pectines

Picture of the telson including vesicle and aculeus

Full view of the carapace

Centruroides infamatus ornatus

(CioLC1)

Picture of the chelicerae

Movable blade of the pedipalp in full color

Movable blade of the pedipalp in sepia color

Picture of the spiracles and pectines

Picture of the third & fourth segment of the metasoma

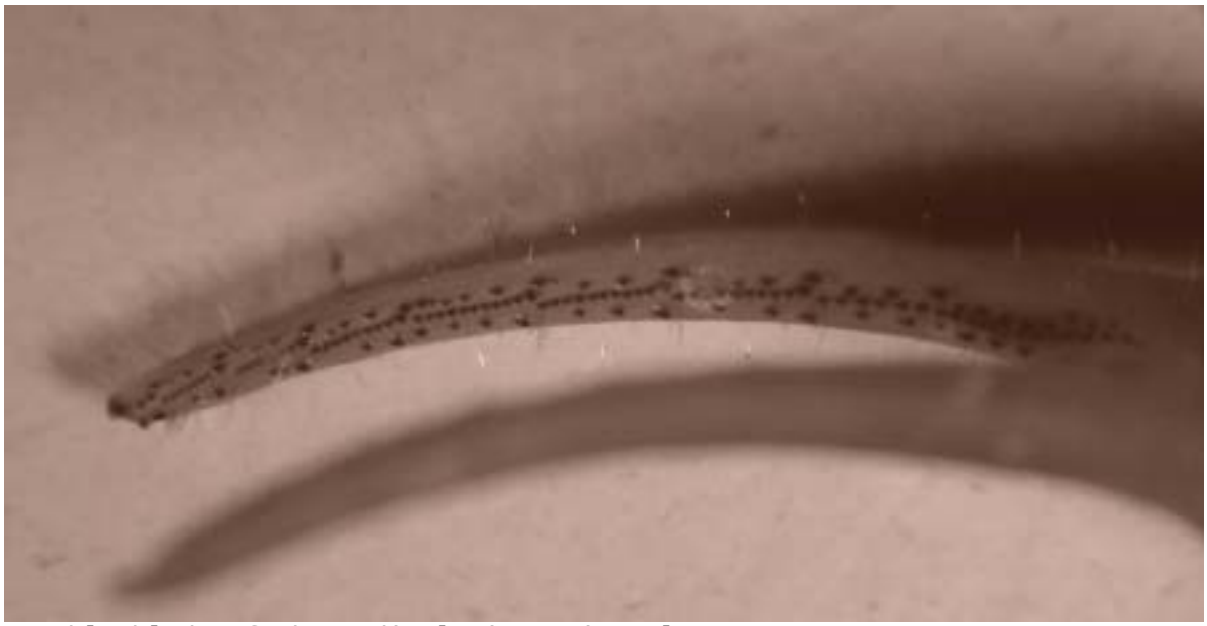
Picture of the telson including vesicle and aculeus

Full view of the face

Centruroides limpidus limpidus

(CllQU1)

Movable blade of the pedipalp in full color



Movable blade of the pedipalp in sepia color

Picture of the chelicerae

Picture of the third segment of the metasoma

Picture of the telson including vesicle and aculeus

Full view of the face

Centruroides limpidus tecomanus

(CltFB1)

Movable blade of the pedipalp in full color

Movable blade of the pedipalp in sepia color

Picture of the spiracles and pectines

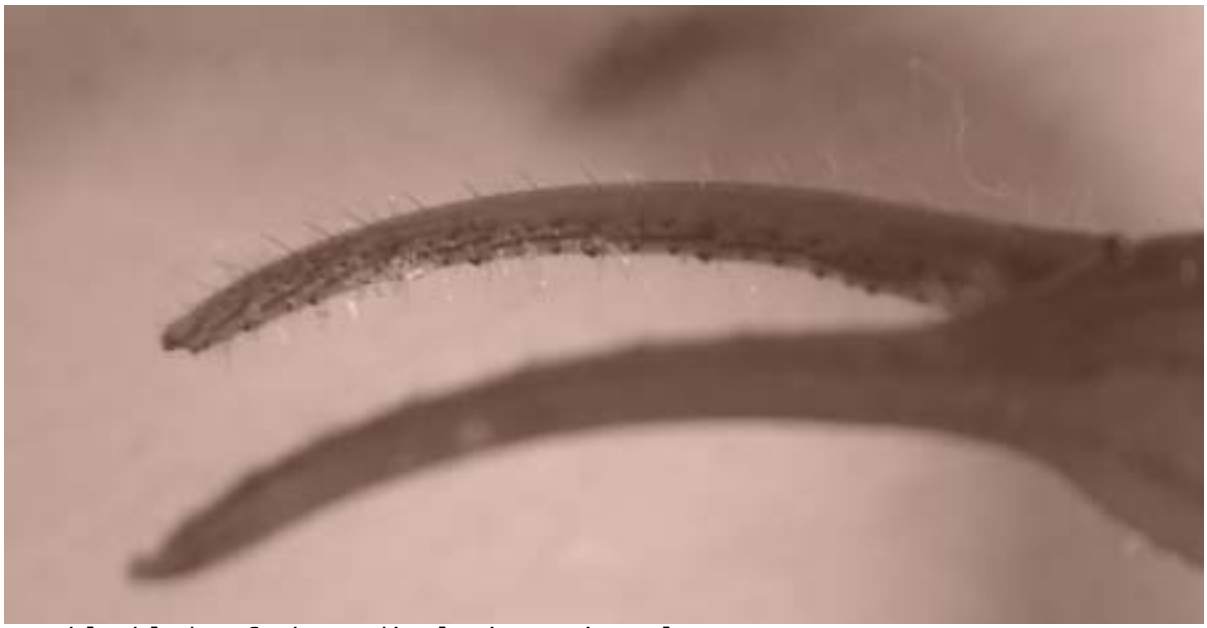
Picture of the telson including vesicle and aculeus

Full view of the face

Centruroides infamatus ornatus

(CioMo1)

Movable blade of the pedipalp in full color



Movable blade of the pedipalp in sepia color

Picture of the chelicerae and eyes

Picture of the third segment of the metasoma

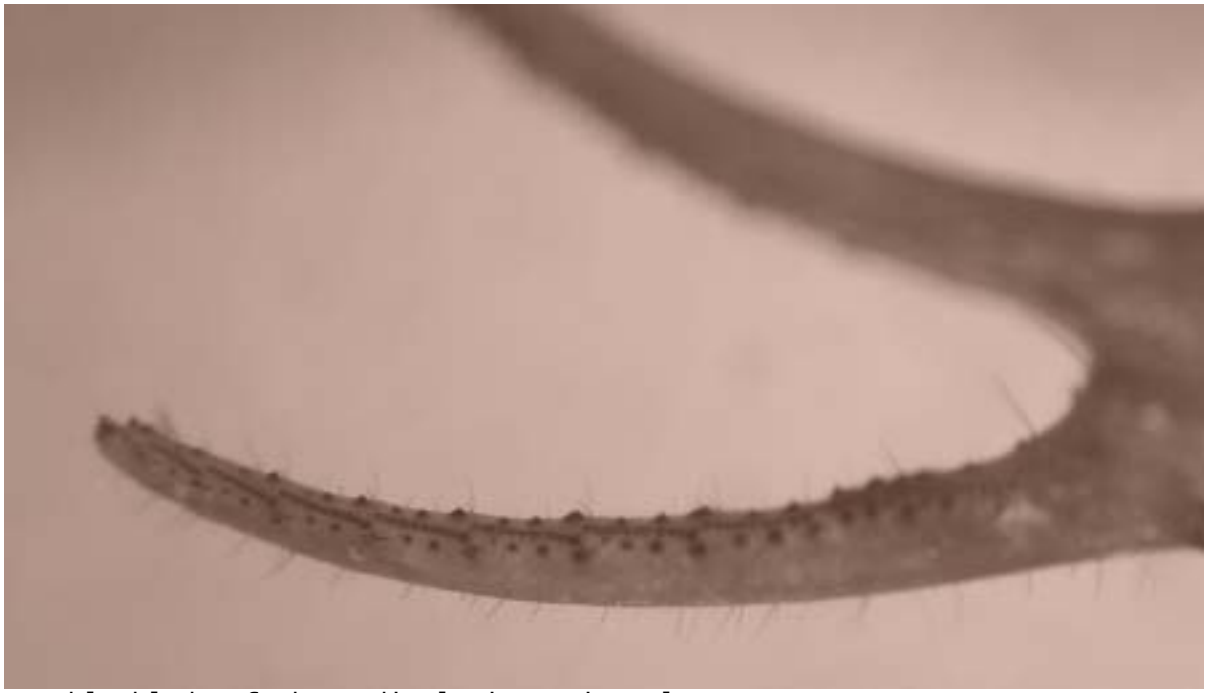
Picture of the telson including vesicle and aculeus

Full view of the face

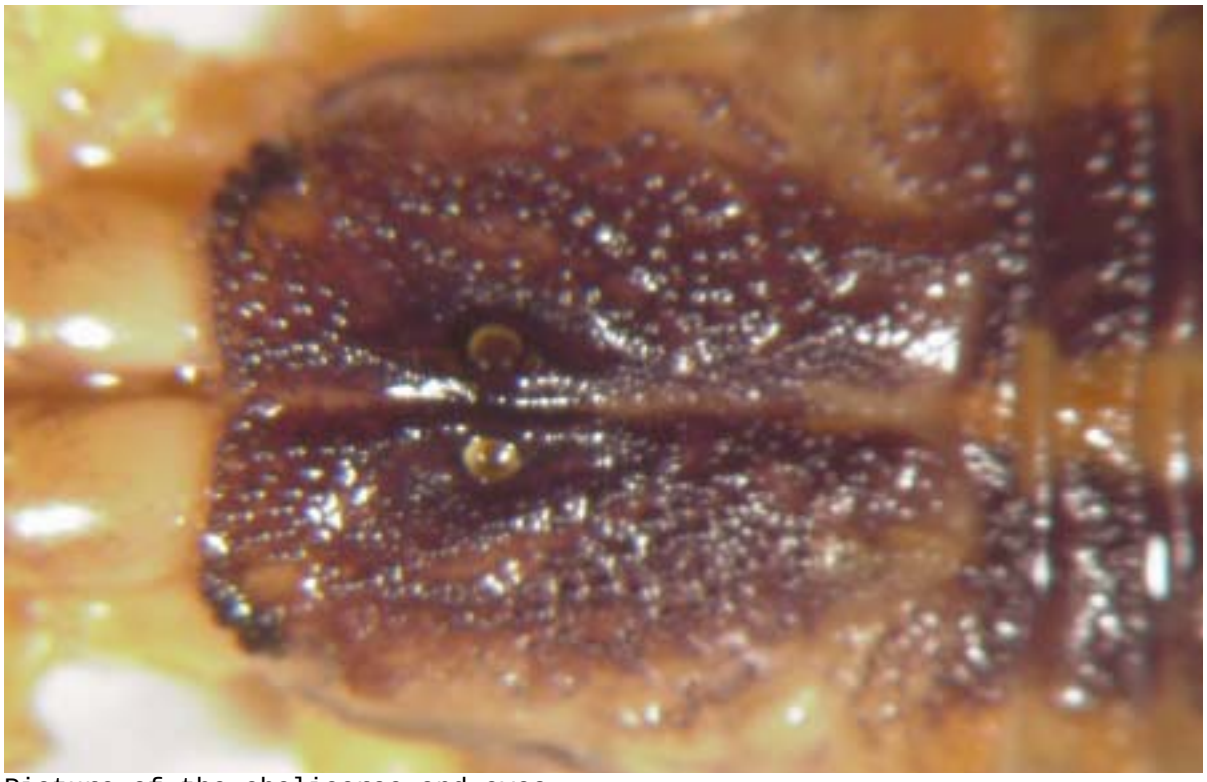
Centruroides infamatus ornatus

(CioEC1)

Movable blade of the pedipalp in full color



Movable blade of the pedipalp in sepia color



Picture of the chelicerae and eyes

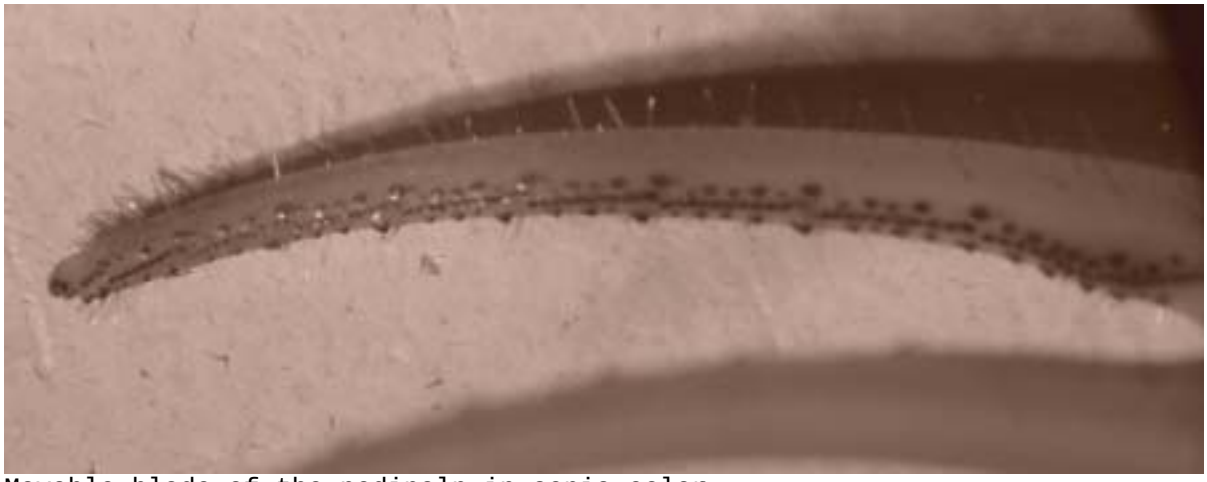
Picture of the third segment of the metasoma

Picture of the telson including vesicle and aculeus

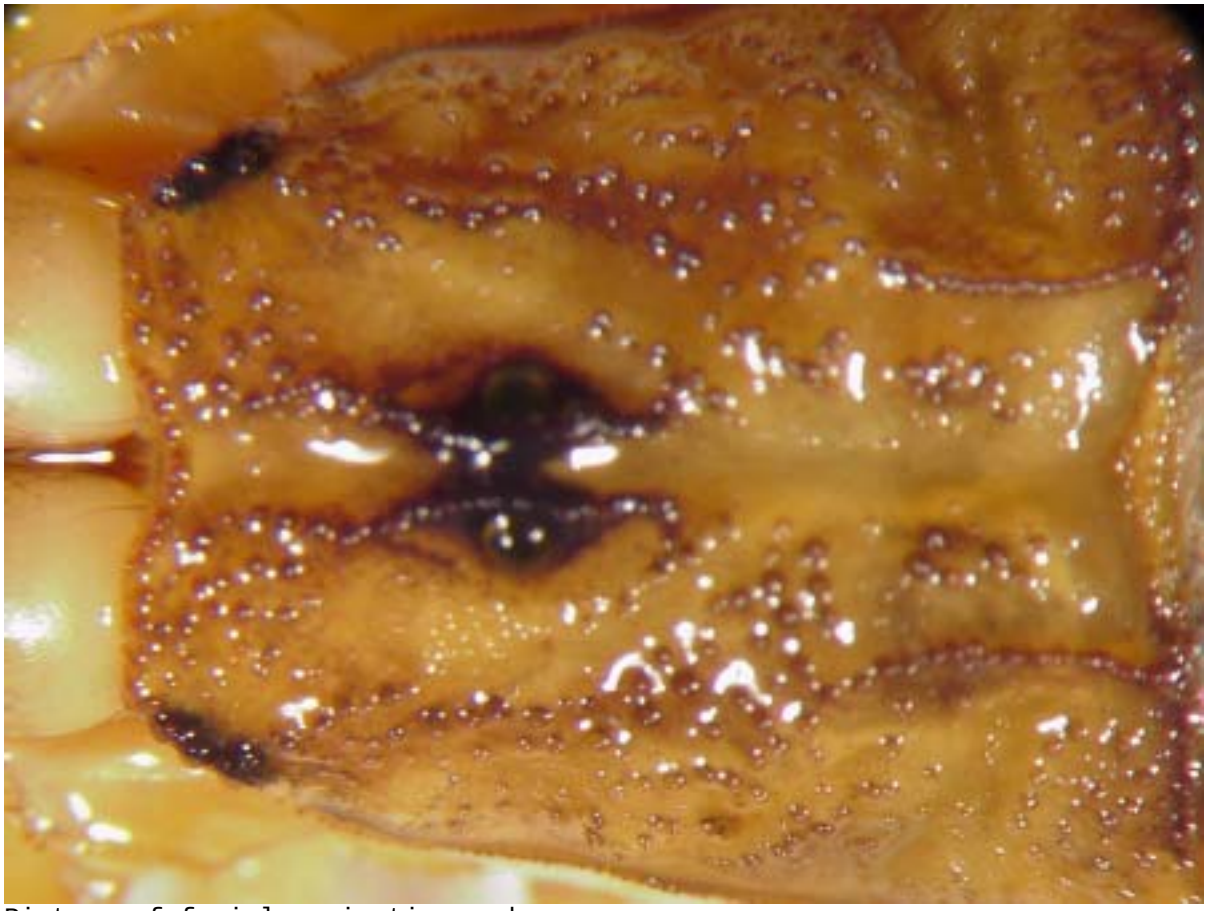
Full view of the face

Centruroides limpidus limpidus
(CllTO1)

Movable blade of the pedipalp in full color



Movable blade of the pedipalp in sepia color



Picture of facial carination and eyes

Picture showing the tarsus ungues from the bottom

Picture of the third segment of the metasoma

Picture of the telson including vesicle and aculeus

Full view of the face

Centruroides limpidus limpidus
(CllCH1)

Movable blade of the pedipalp in full color

Movable blade of the pedipalp in sepia color

Picture of facial carination and eyes

Picture showing the tarsus ungues from the bottom

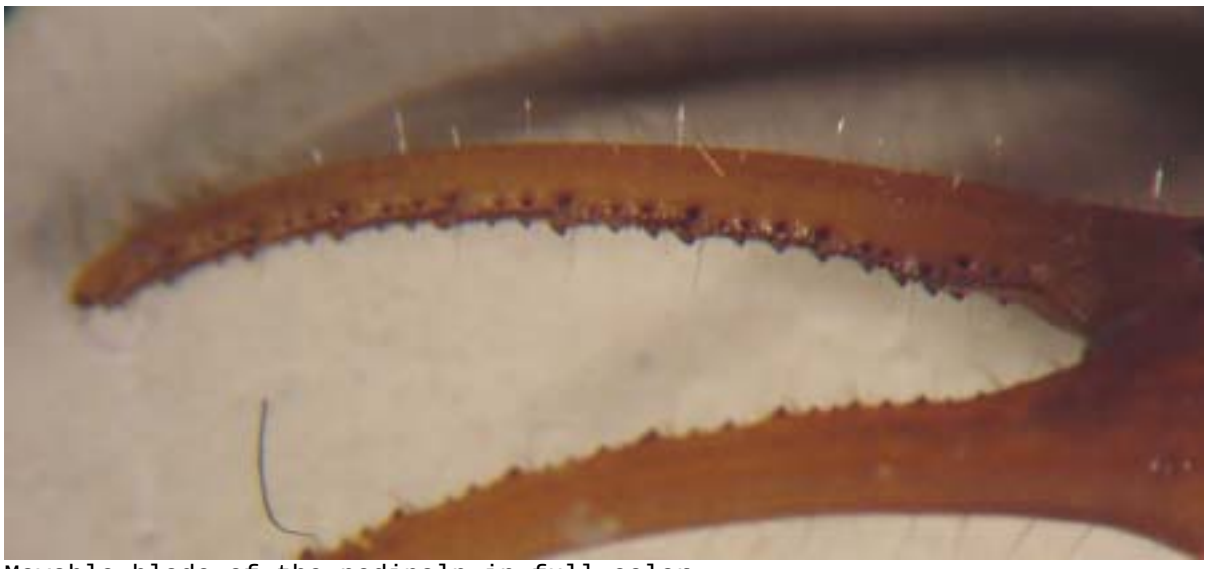
Picture of the third segment of the metasoma

Picture of the telson including vesicle and aculeus

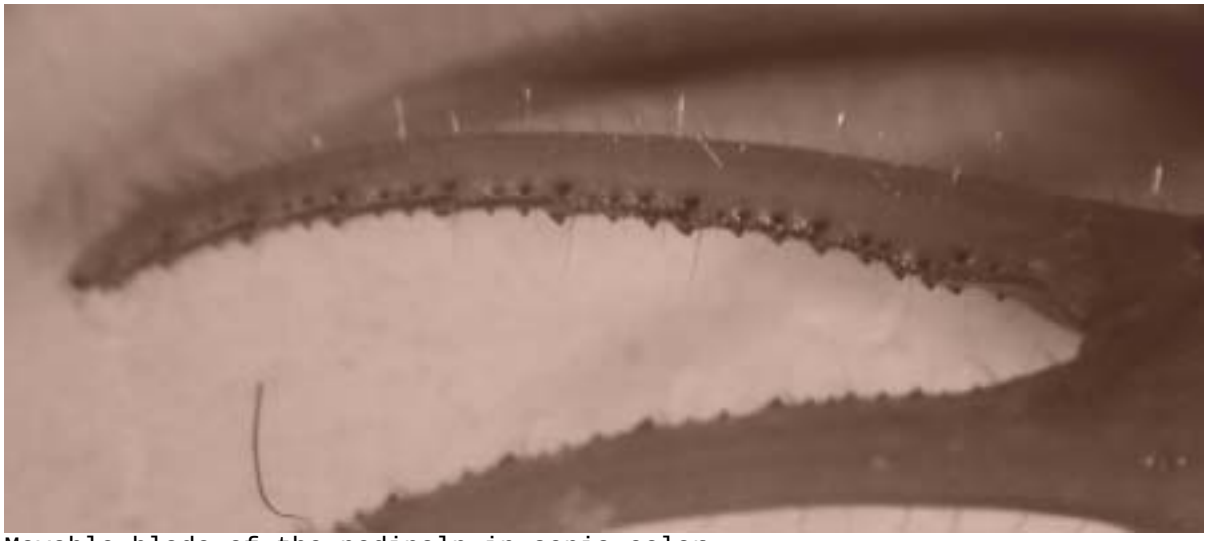
Full view of the face

Centruroides limpidus limpidus

(CllCH2)



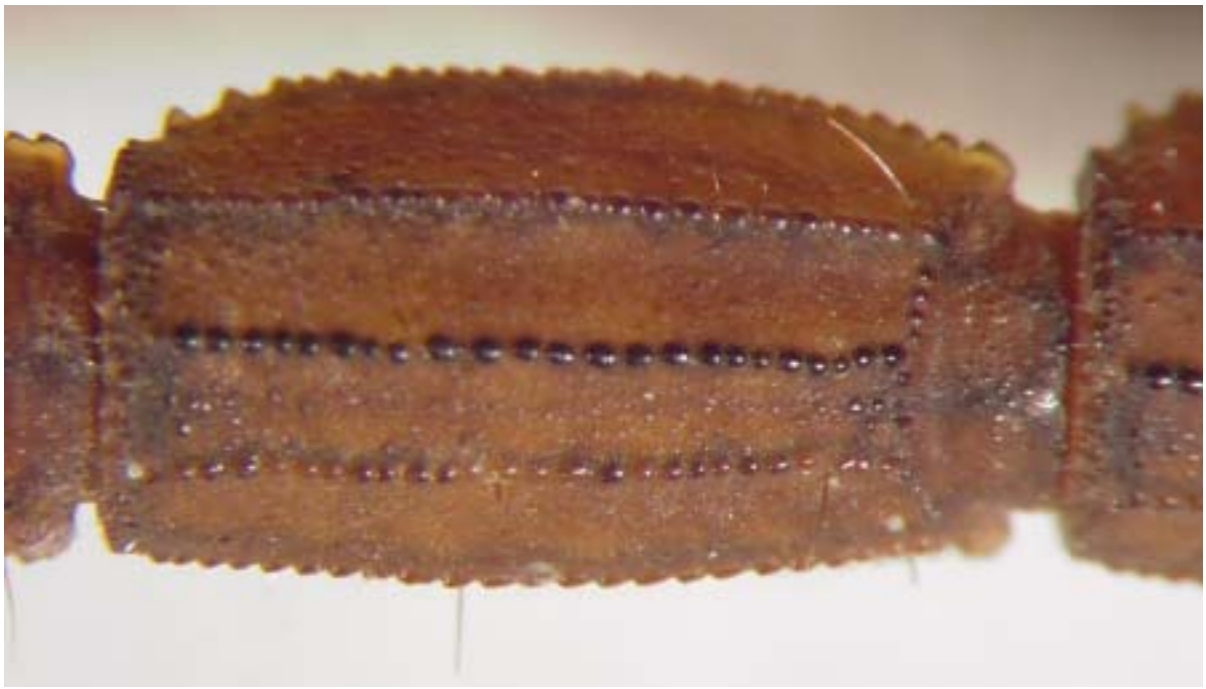
Movable blade of the pedipalp in full color



Movable blade of the pedipalp in sepia color

Picture of facial carination and the eyes.

Picture showing the tarsus ungues from the bottom



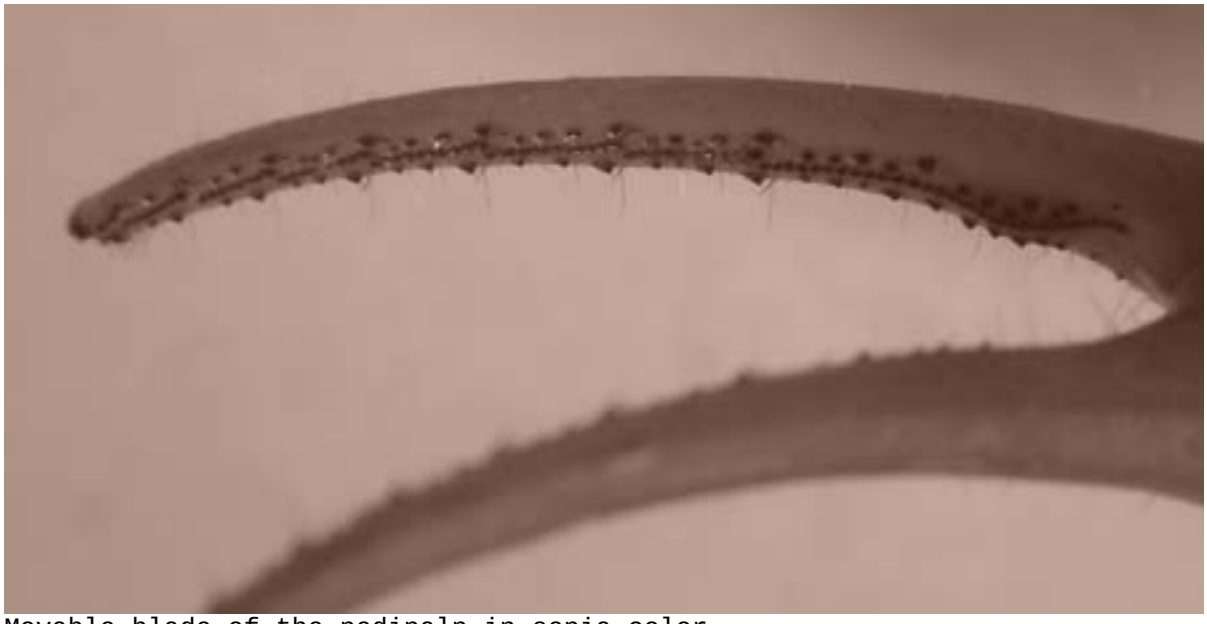
Picture of the third segment of the metasoma

Picture of the telson including vesicle and aculeus

Full view of the face

Centruroides limpidus limpidus
(CllTZ1)

Movable blade of the pedipalp in full color



Movable blade of the pedipalp in sepia color

Picture of facial carination and the eyes.

Picture showing the tarsus ungues from below

Picture of the third segment of the metasoma

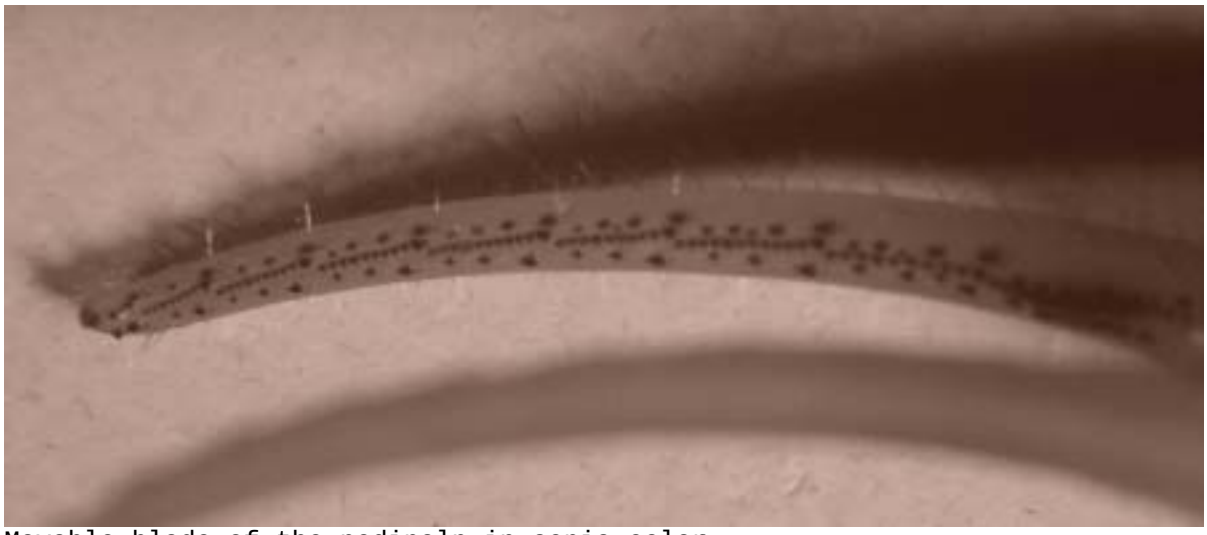
Picture of the telson including vesicle and aculeus

Full view of the face

Centruroides limpidus limpidus

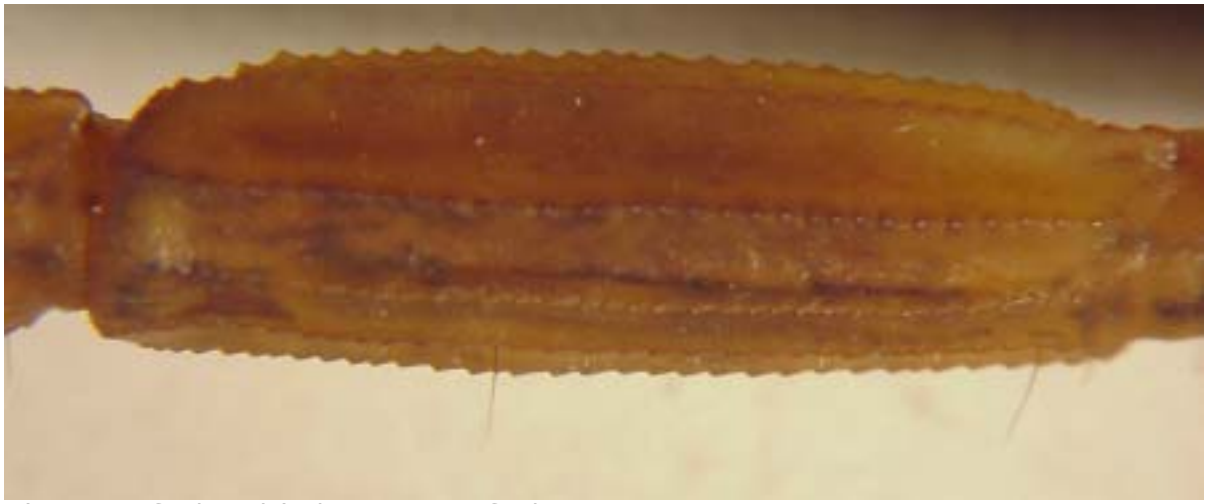
(CllAR1)

Movable blade of the pedipalp in full color



Movable blade of the pedipalp in sepia color

Picture of facial carination



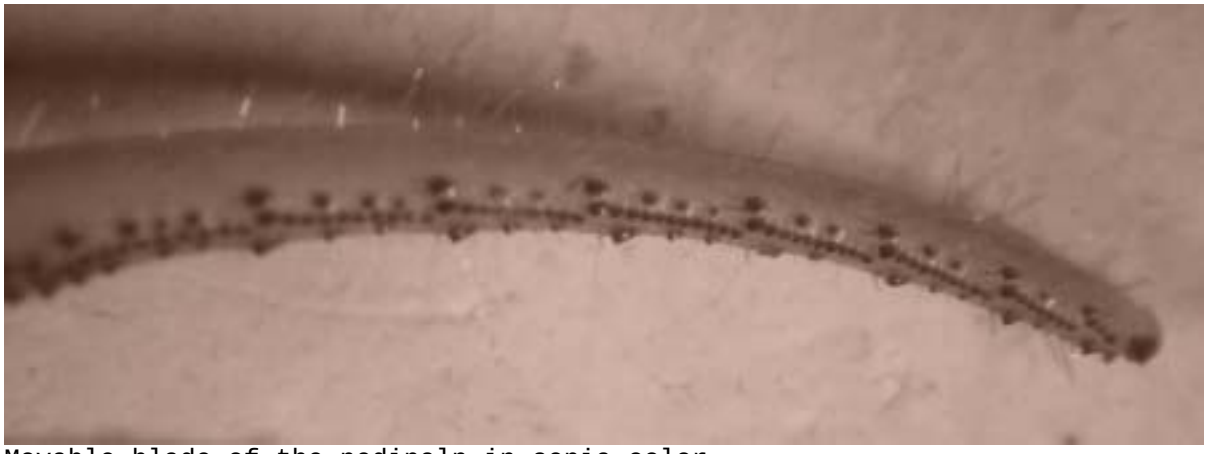
Picture of the third segment of the metasoma

Picture of the telson including vesicle and aculeus

Full view of the face

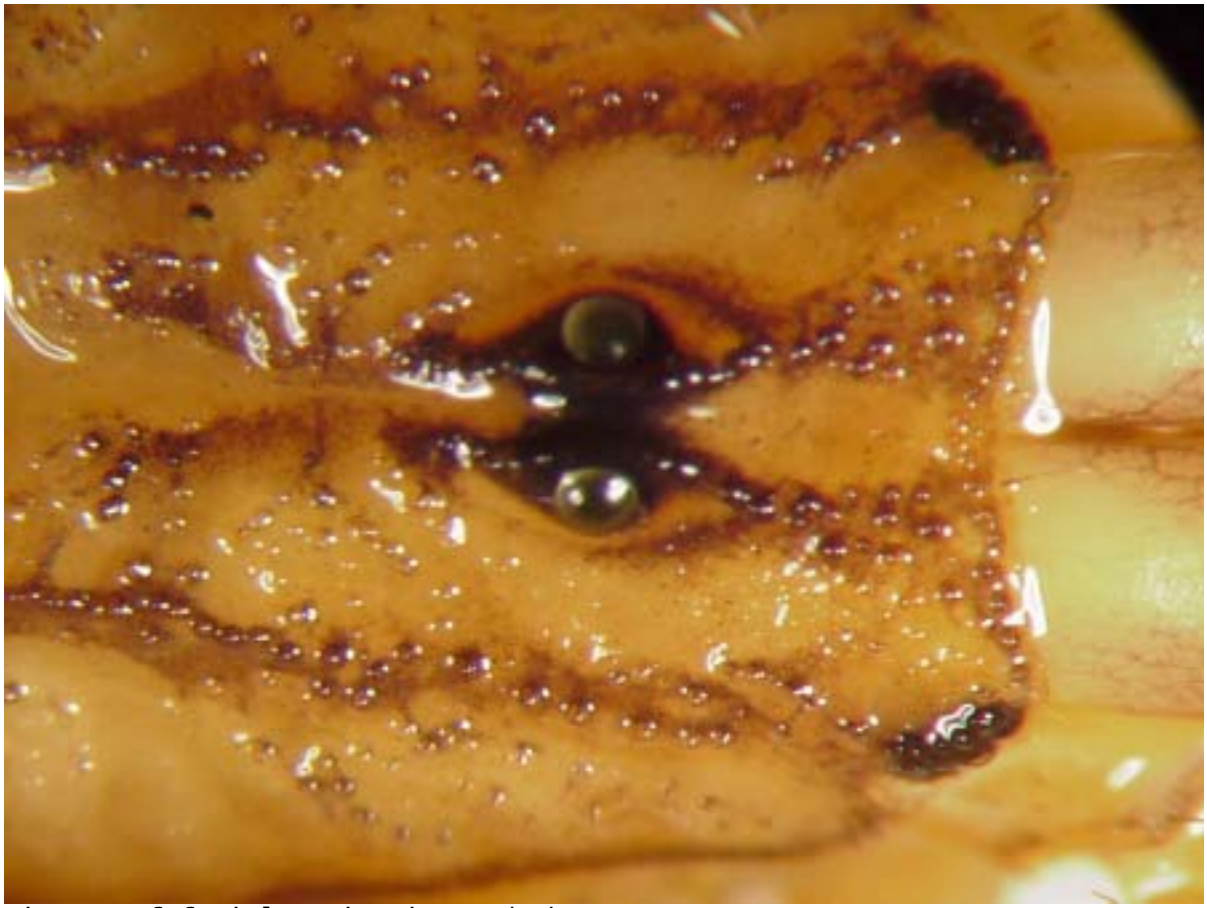
Centruroides limpidus limpidus
(CllHU1)

Movable blade of the pedipalp in full color



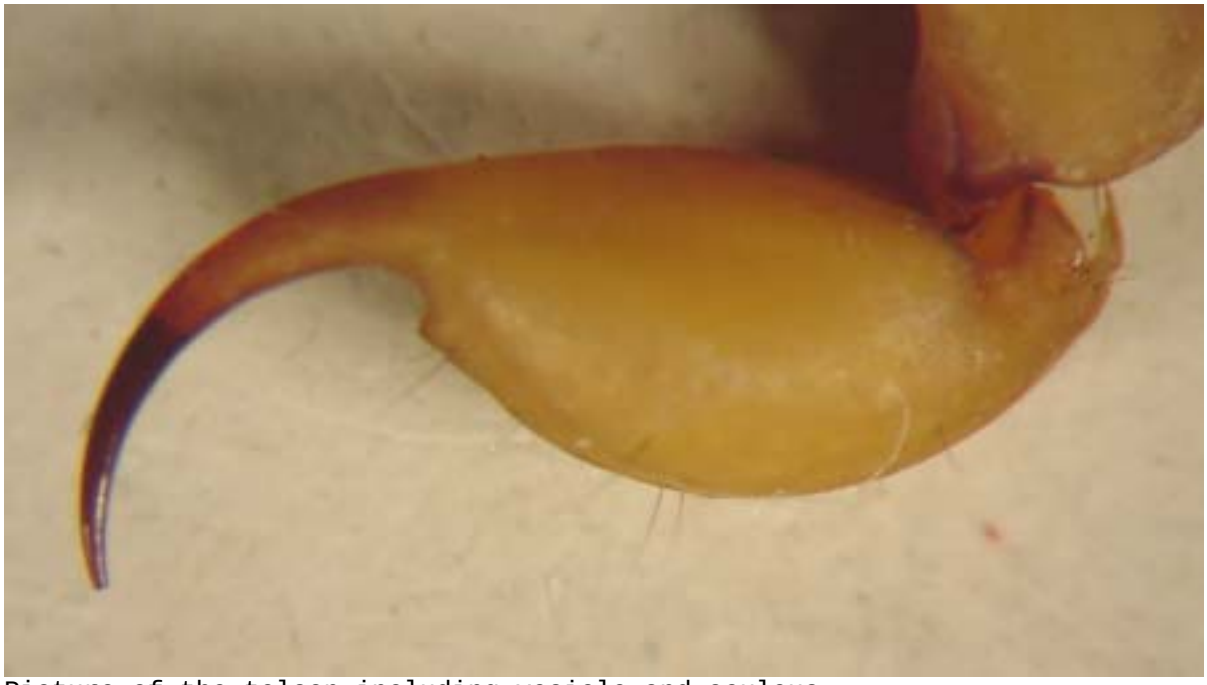
Movable blade of the pedipalp in sepia color

Picture of the chelicerae



Picture of facial carination and the eyes.

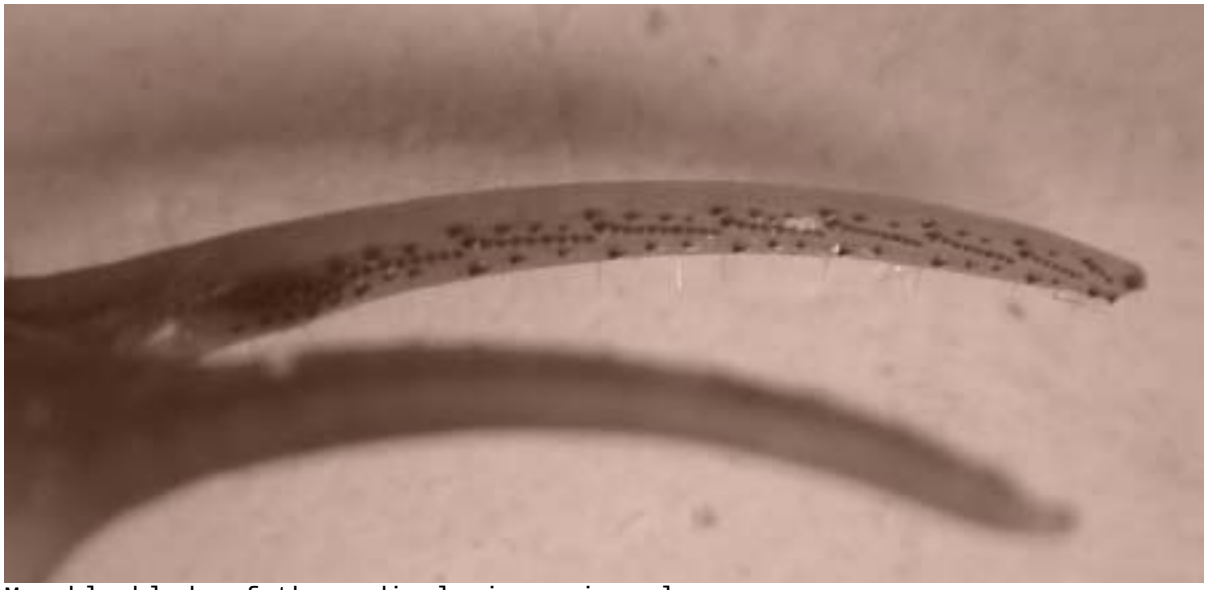
Picture of the third segment of the metasoma



Picture of the telson including vesicle and aculeus

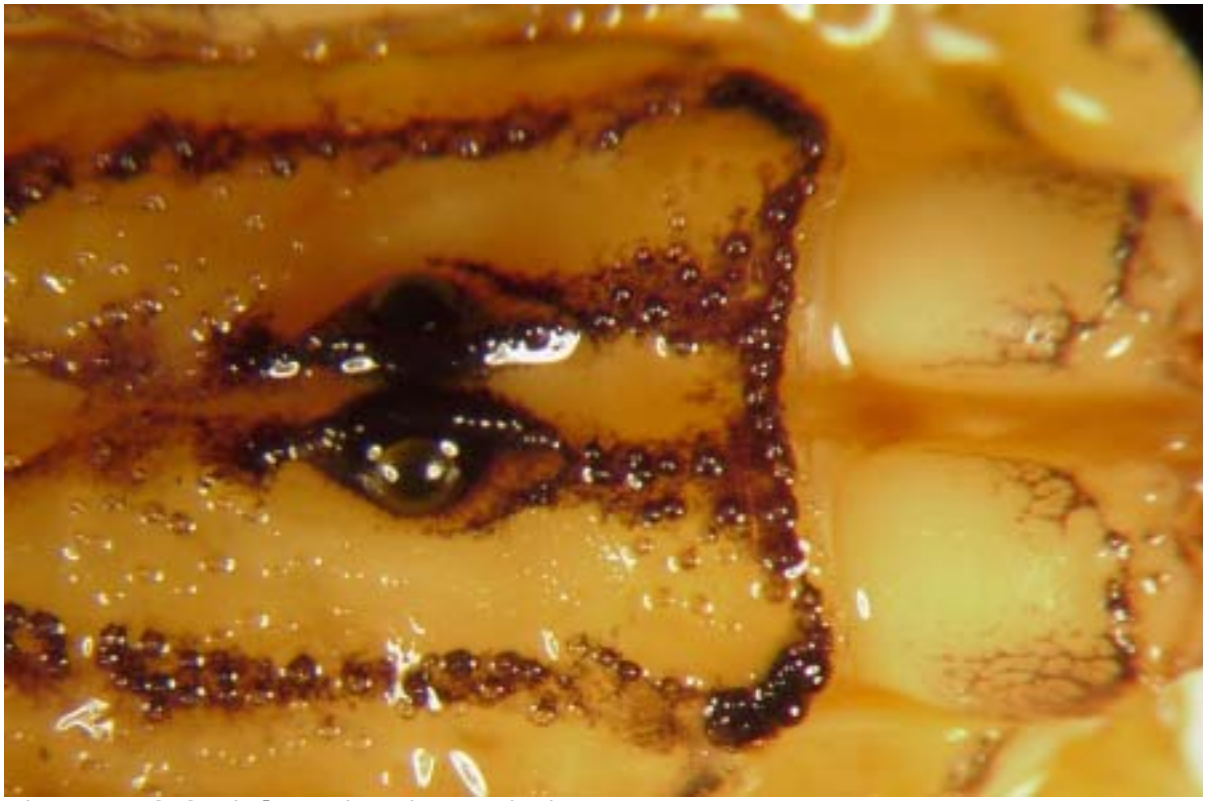
Centruroides limpidus tecomanus
(CltFB2)

Movable blade of the pedipalp in full color



Movable blade of the pedipalp in sepia color

Picture of the chelicerae



Picture of facial carination and the eyes

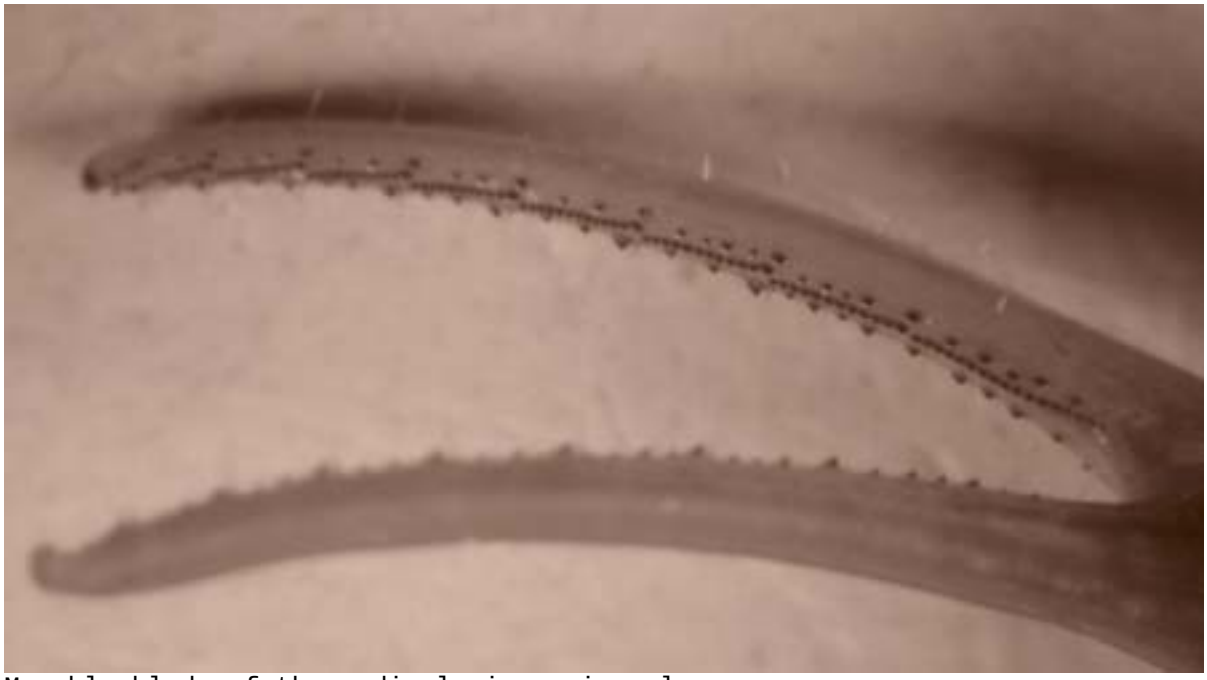
Picture of the third segment of the metasoma

Picture of the telson including vesicle and aculeus

Full view of the animal

Centruroides limpidus limpidus
(CllCH2)

Movable blade of the pedipalp in full color



Movable blade of the pedipalp in sepia color

Picture of the chelicerae

Picture of the chelicerae and facial carination

Picture of the third segment of the metasoma

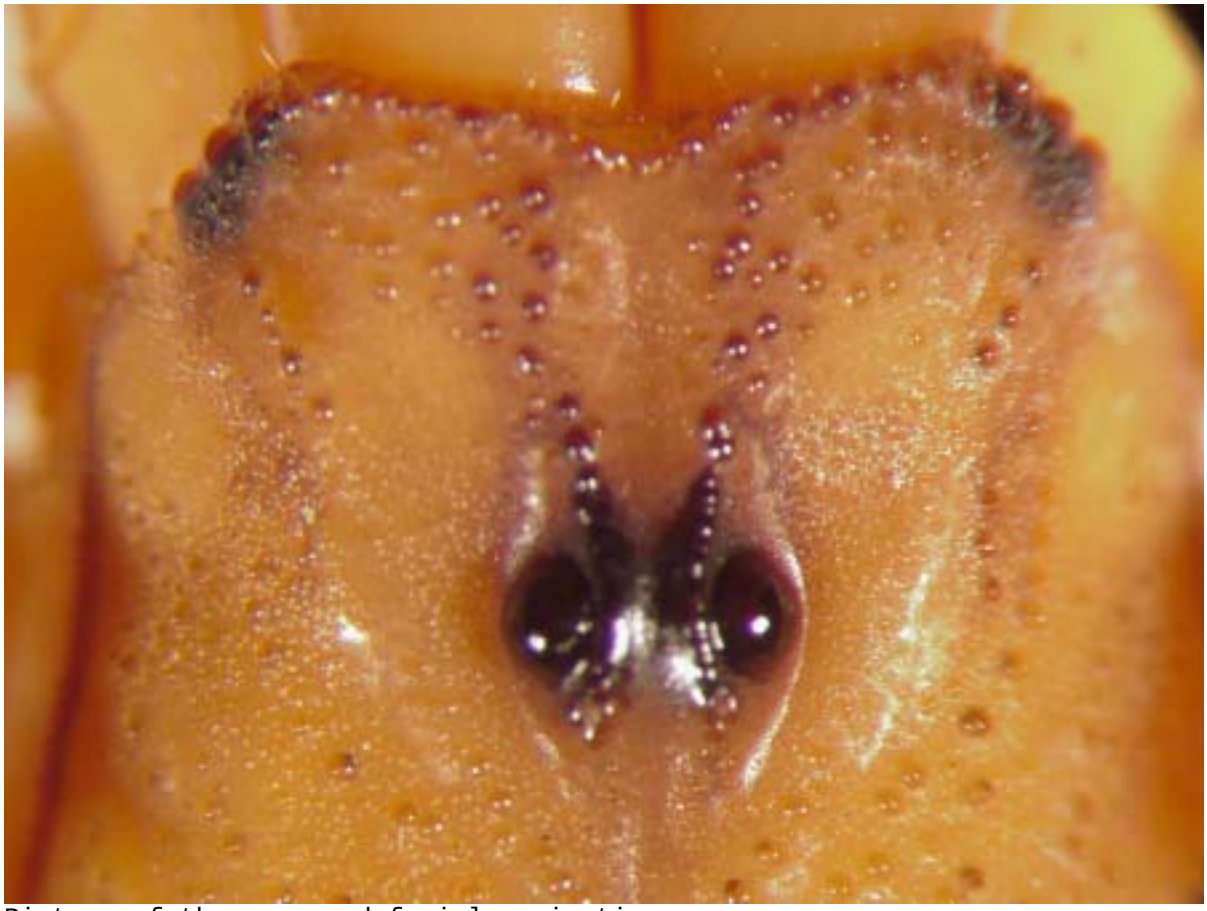
Picture of the telson including vesicle and aculeus

Full view of the animal

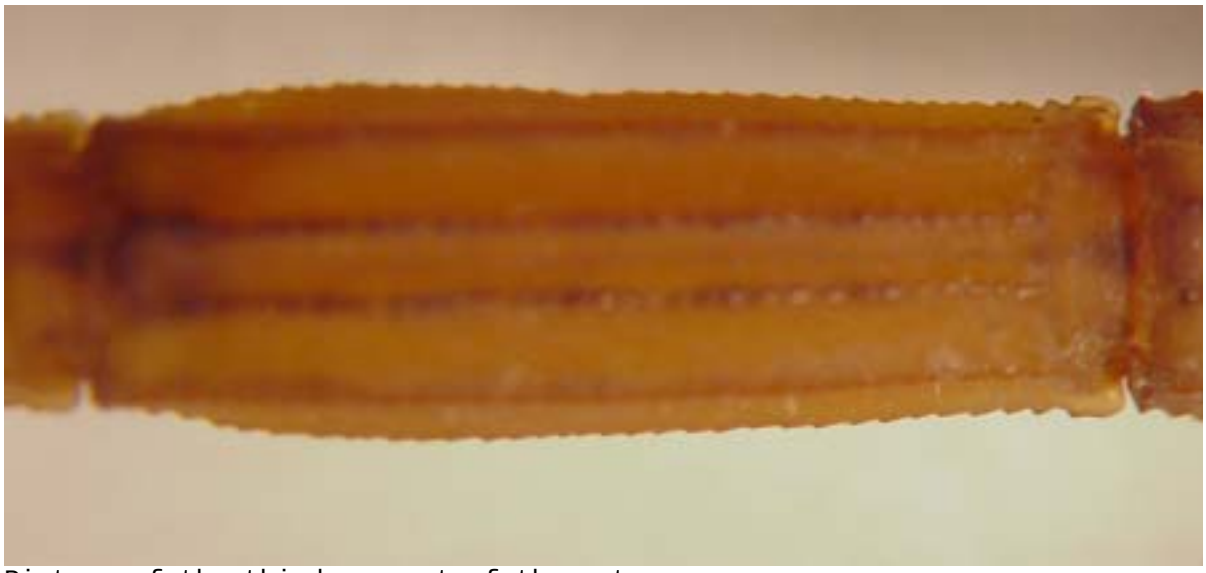
Centruroides elegans

Movable blade of the pedipalp in full color

Movable blade of the pedipalp in sepia color



Picture of the eyes and facial carination



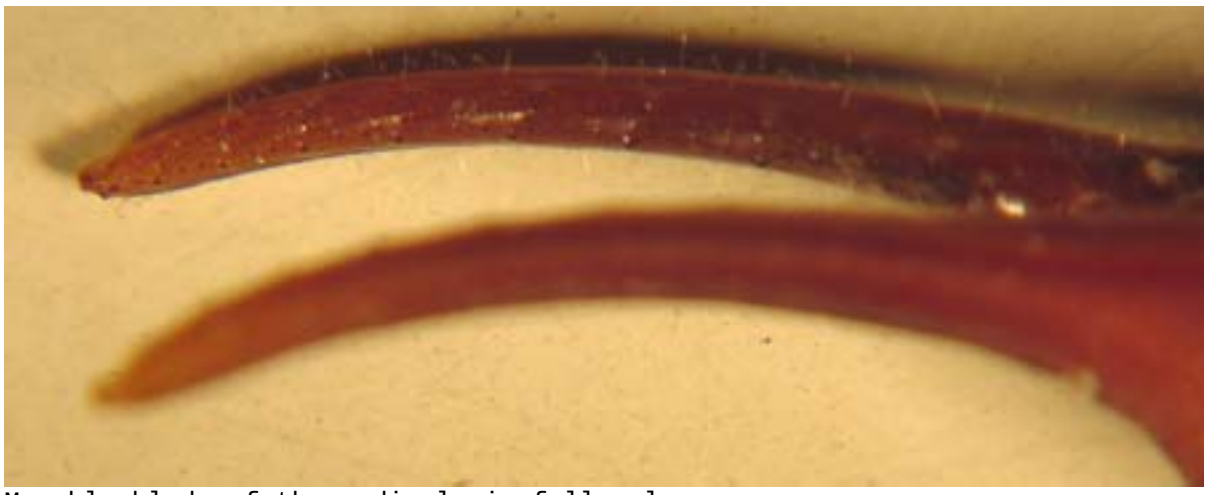
Picture of the third segment of the metasoma

Picture of the telson including vesicle and aculeus

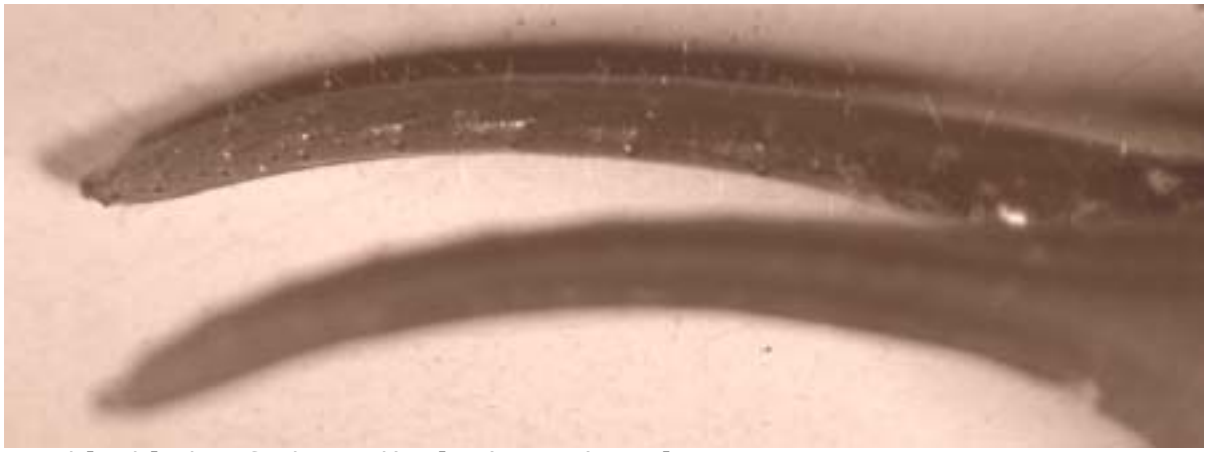
Full view of the face

Centruroides gracilis

Collected Monroe County, Florida, Dry Tortugas Islands, Garden Key

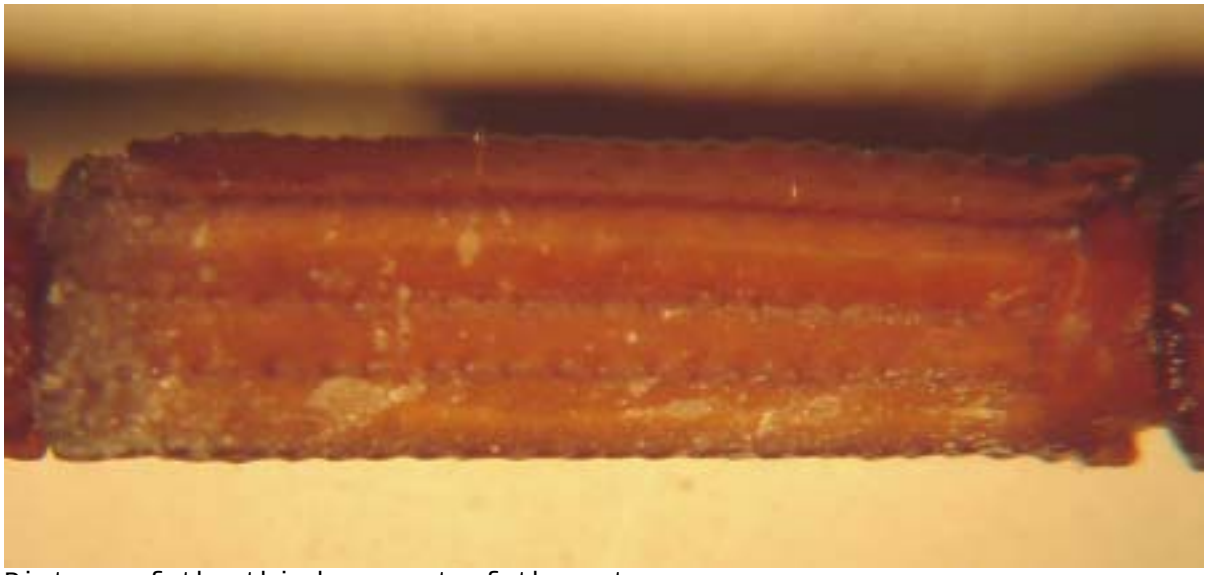


Movable blade of the pedipalp in full color



Movable blade of the pedipalp in sepia color

Picture of the chelicerae



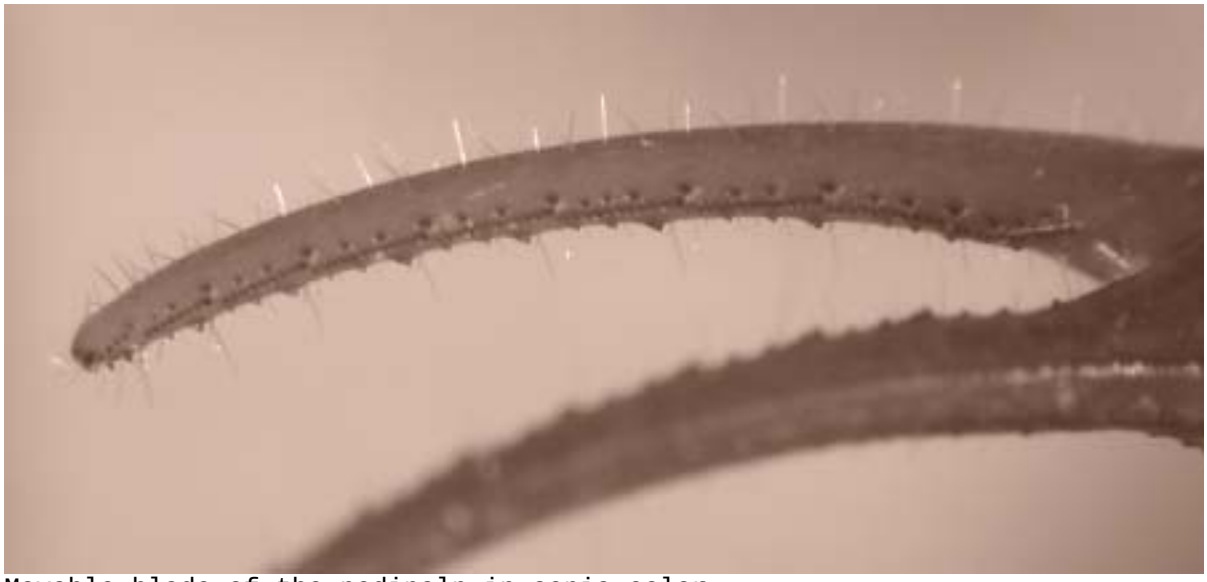
Picture of the third segment of the metasoma

Picture of the telson including vesicle and aculeus

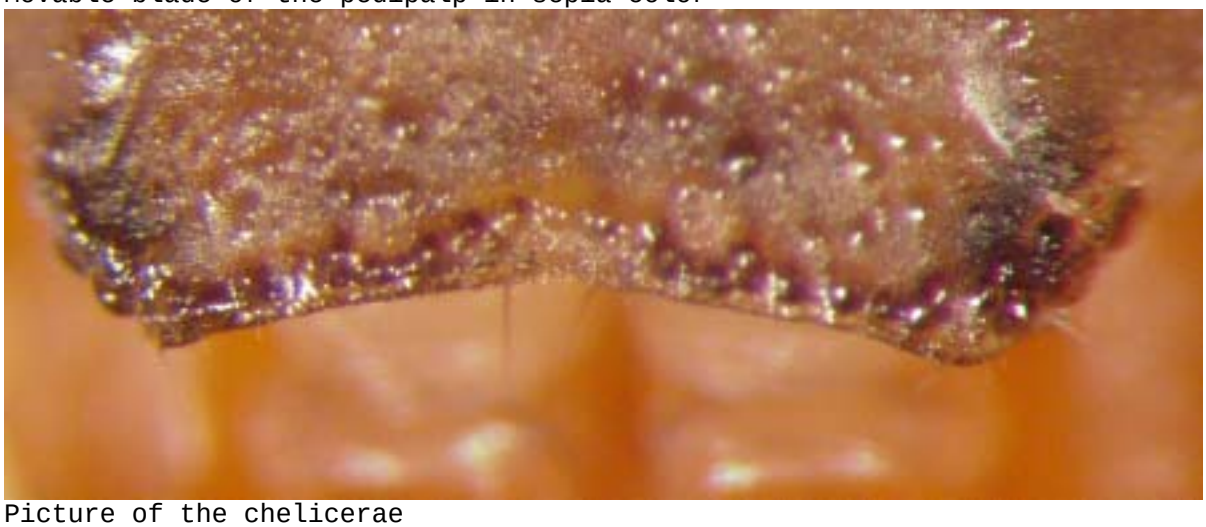
Full view of the face

Centruroides hentzi
Collected Alachua County Florida, Gainsville

Movable blade of the pedipalp in full color



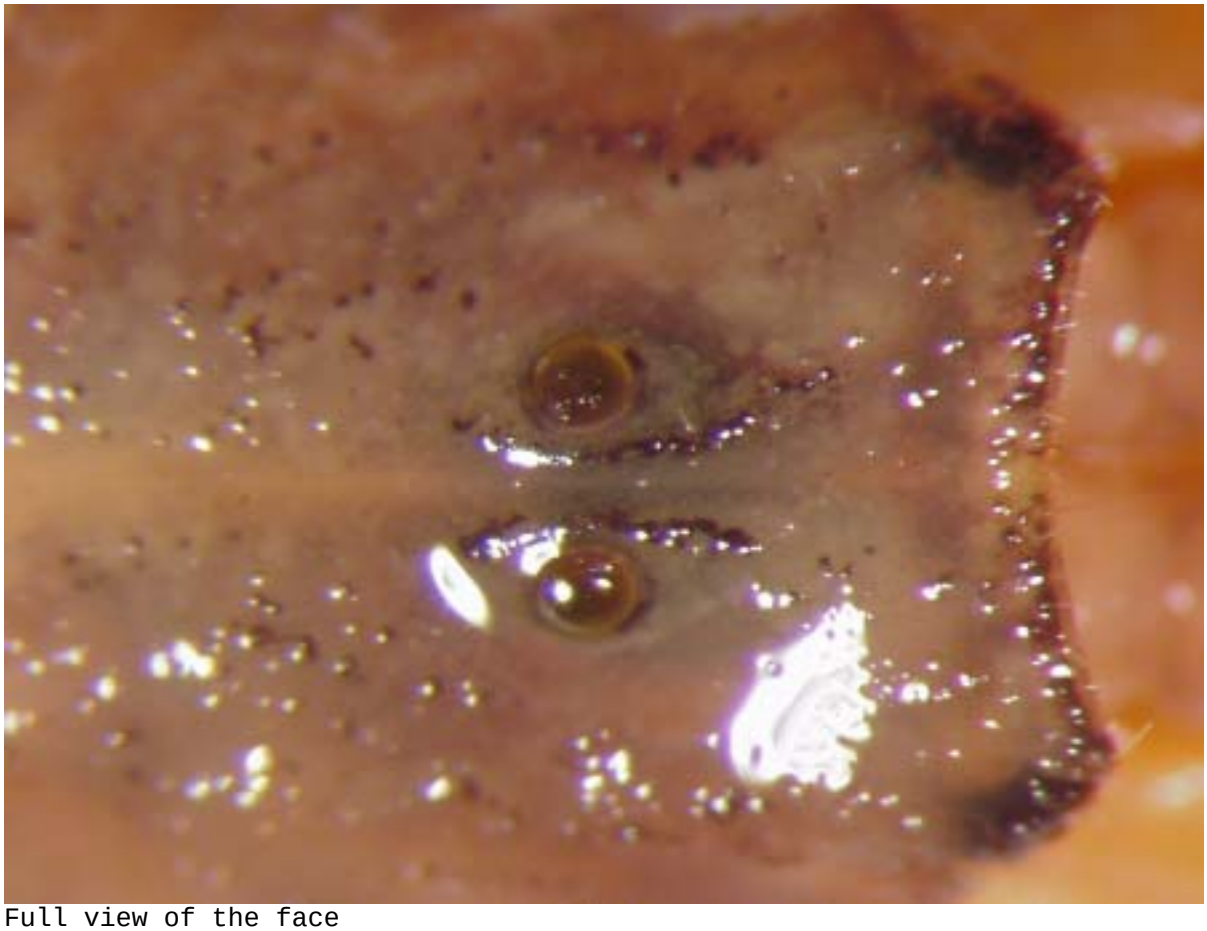
Movable blade of the pedipalp in sepia color



Picture of the chelicerae

Picture of the third segment of the metasoma

Picture of the telson including vesicle and aculeus



Full view of the face

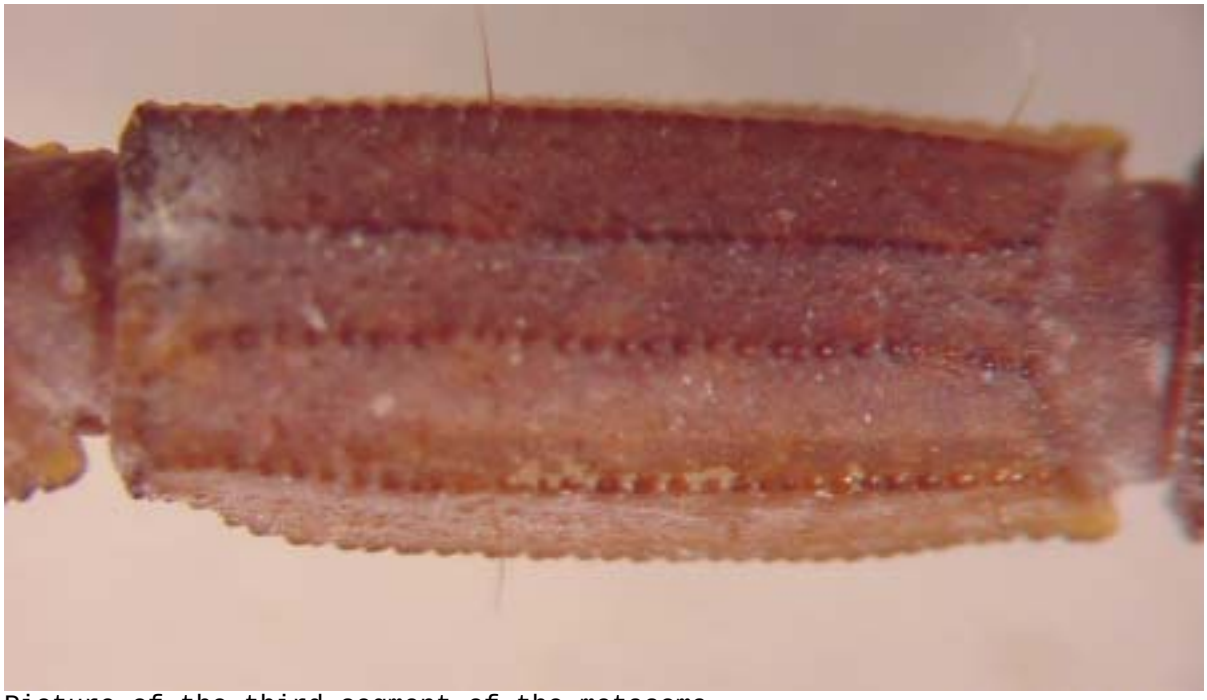
Centruroides keysi

Collected in Monroe County Florida, Key largo

Movable blade of the pedipalp in full color

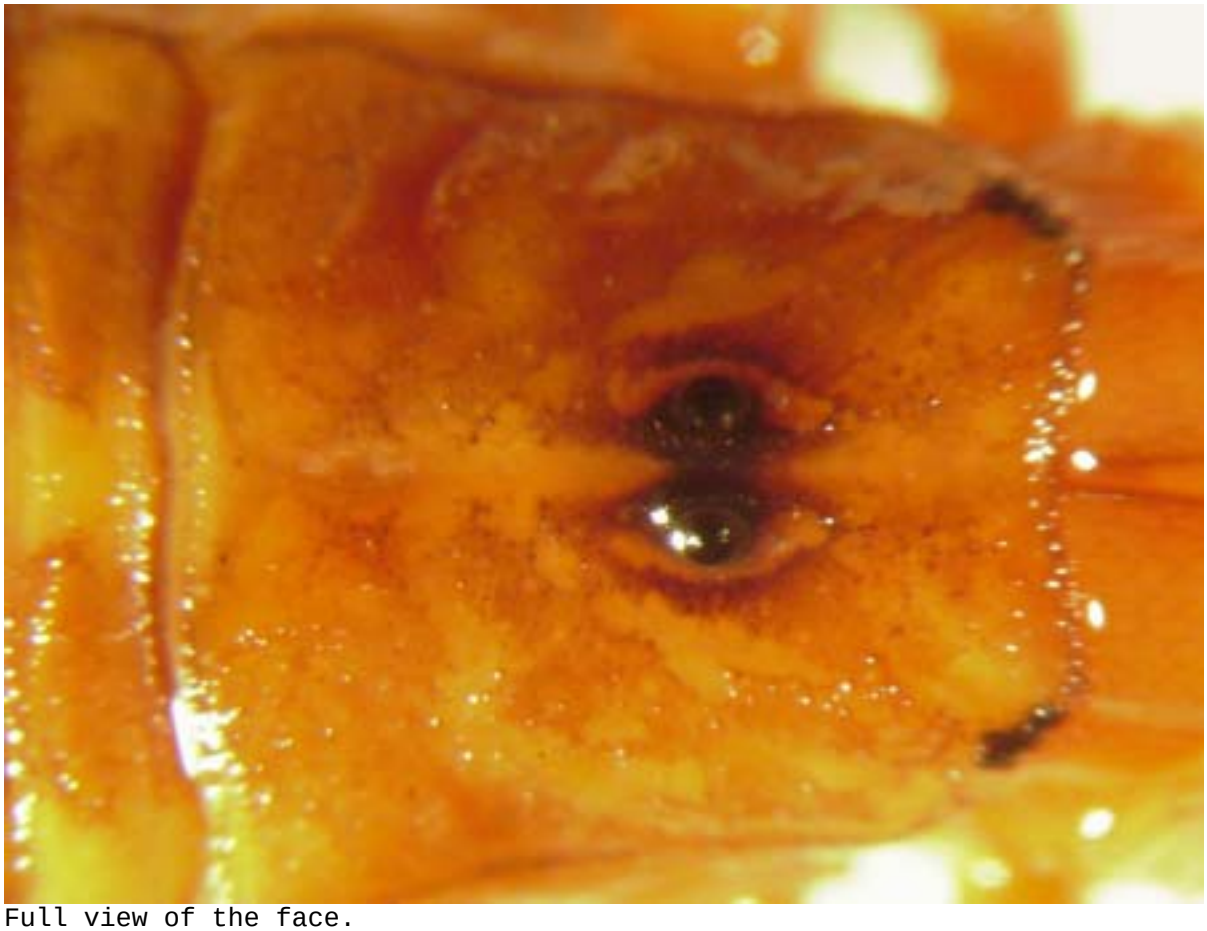
Movable blade of the pedipalp in sepia color

Picture of the chelicerae



Picture of the third segment of the metasoma

Picture of the telson including vesicle and aculeus



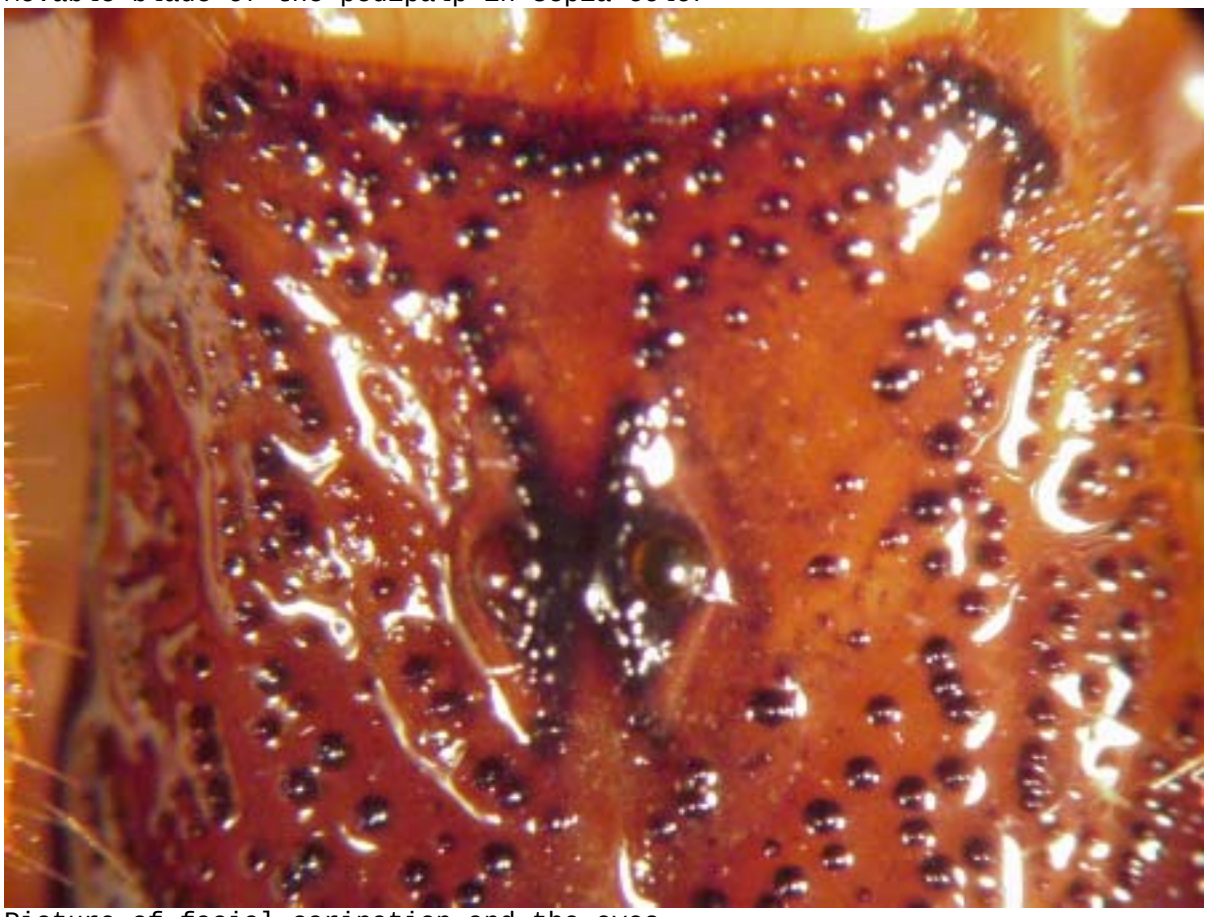
Full view of the face.

Centruroides margintesis

Collected

Movable blade of the pedipalp in full color

Movable blade of the pedipalp in sepia color



Picture of facial carination and the eyes

Picture of the third segment of the metasoma

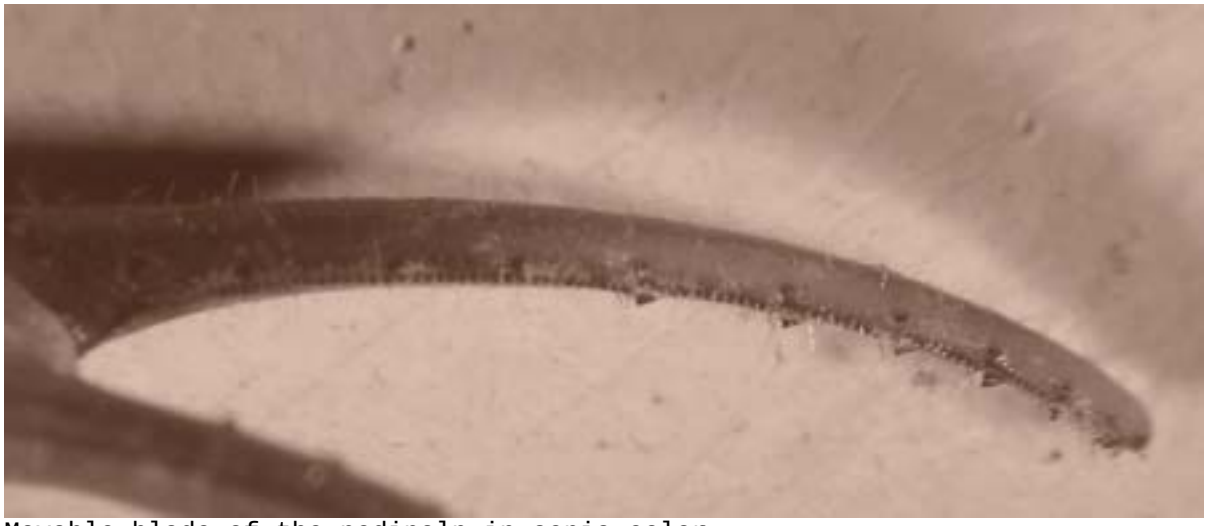
Picture of the telson including vesicle and aculeus

Full view of the face

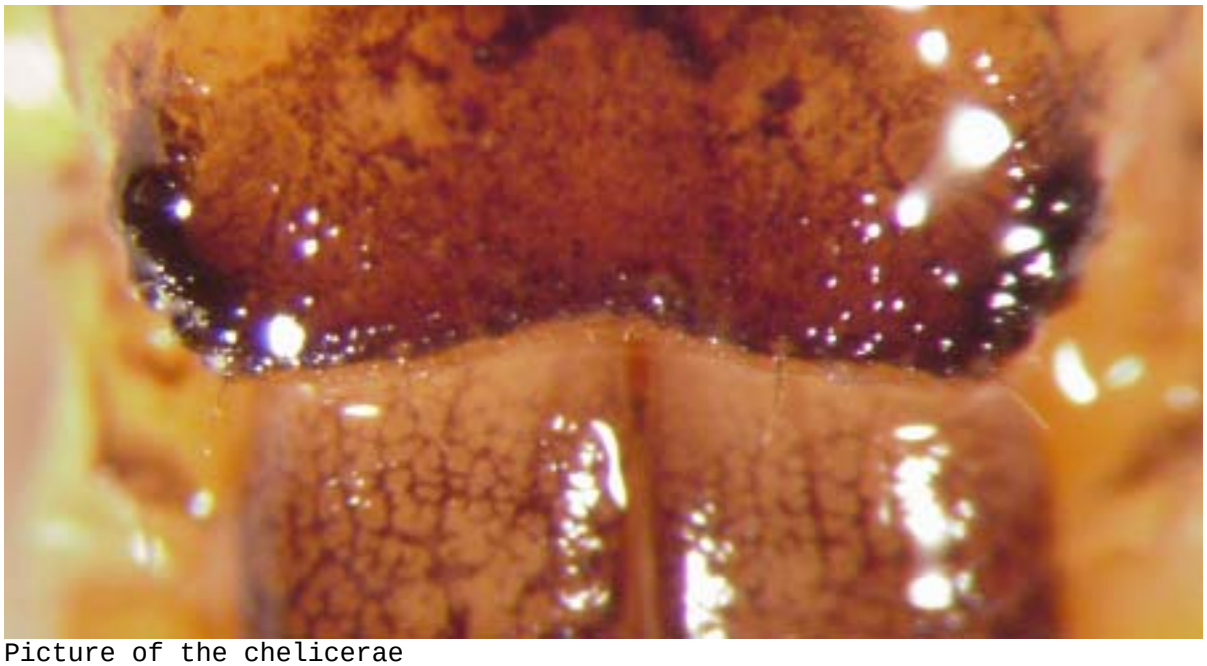
Lychas Mucronatus

Collected Maumere, Flores Island, Indonesia

Movable blade of the pedipalp in full color



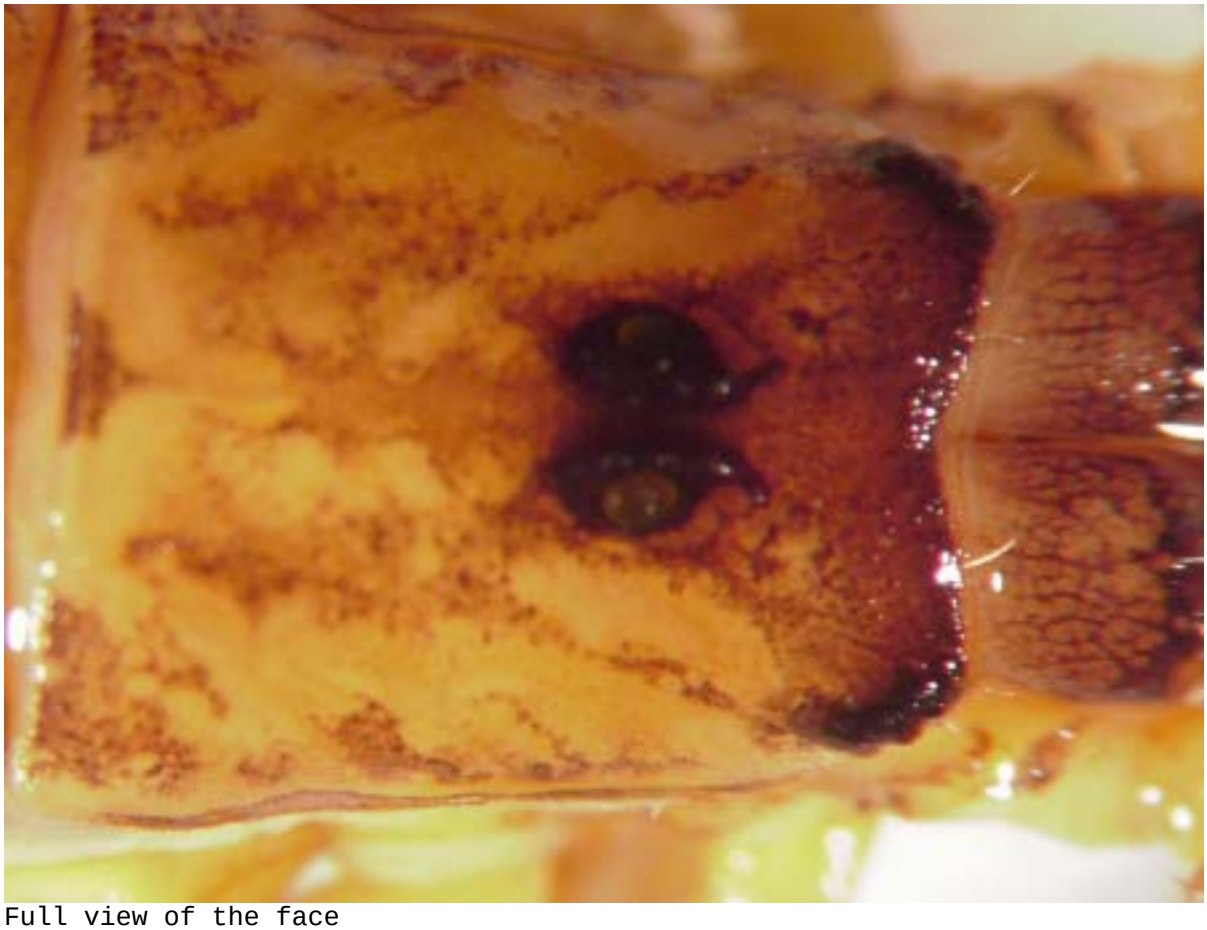
Movable blade of the pedipalp in sepia color



Picture of the chelicerae

Picture of the third segment of the metasoma

Picture of the telson including vesicle and aculeus



Full view of the face