Mitochondrial 16S rRNA Gene as a Tool for Scorpion Systematics

Thesis submitted to
The Graduate College
Marshall University

In partial fulfillment of the
Requirements for the Degree of
Masters of Science
Biological Science

by

Mark D. Barker

Marshall University
Huntington, WV
1 May, 2000

This thesis was accepted on May 2 2000
Month Day Year

as meeting the research requirements for the Master's Degree.

Advisor: Victor Fet

Committee: Michael Little

Committee: Dan K Evans

Department of Biological Sciences

Leonard J Deutsch

Dean of the Graduate College

Acknowledgements

I would like to first thank Dr. Evans and Dr. Little for being on my committee. I also would like to thank Dr. Little for the use of his lab and his computer on which much of my analysis was performed. I would like to thank B. Gantenbein who performed the analysis on the *Euscorpius* project, collected specimens, and obtained some of the sequences of the samples in the joint lab project.

I would like to thank the following people for without their collection of samples my study could have never been performed: M. Soleglad, S. Tallarvoric, R. Farley, J. Bigelow, G. Lowe, G. Polis, D. Huber, T. Yamashita, and B. Tomberlin.

Lastly I would like to thank Dr. Victor Fet for the use of his lab and his expertise over the last two years, and whose support made this study possible.

Table of Contents

Table of Figures

Table of Tables

Page

Abstract

The use of morphological analysis in the classification and phylogeny of scorpions often has led to confusion and controversy. In order to clarify these problems other character sets should be considered for analysis. One such tool is the use of molecular comparisons between organisms. This study used a section of the mitochondrial 16S ribosomal RNA gene. This technique was applied to investigate four scorpion genera belonging to three families: *Hadrurus*, *Anuroctonus* (Iuridae), *Euscorpius* (Euscorpiidae), and *Centruroides* (Buthidae). Total DNA was extracted with a standard phenol/chloroform method; a gene of interest was amplified using PCR; and resulting DNA sequence analysis was accomplished using phylogenetic software PAUP 4b. Analysis demonstrated a lack of divergence between the North American subspecies *Hadrurus arizonensis arizonensis* and *H.a.pallidus*, which were originally separated by coloration. Another North American genus, *Anuroctonus*, which possesses one described species, showed sufficient divergence to suggest multiple species. The genus *Euscorpius* demonstrated a species-level genetic diversity of the Alpine forms in *E. germanus* separated by the Adige River (Northern Italy) not

illustrated by morphological differences. Lastly, the current classification of the genus *Centruroides* (USA and Mexico) may not be sufficient considering the amount of genetic diversity present within the species *C. exilicauda*. The quality of information achieved using molecular techniques show it as a very useful tool in systematics.

1. Introduction

1.1 Taxonomy

Few subjects have been as debated as the definition for the terms, species and subspecies, and what these taxonomic divisions entail. The subspecies is the lowest category in the International Code of Zoological Nomenclature. The confusion over the term began with Linnaeus, who used the lowest category of nomenclature as species (Linnaeus 1758). Infraspecific forms were designated as varieties and these were described as two different entities: one as an individual within a population of wide variation and another as a population of polymorphic species. Attempts to clarify this point were then made in 1781 when subspecies were declared to be the geographically defined races and varieties were the individuals within a highly variable population (Esper 1781). Later the term variety was restricted to use in botany and once again the definition for this term was returned to the term "subspecies" giving it once again two entities. In 1942 Mayr modified the idea of the morphological species and introduced the concept of the biological species which stated that species are "groups of actually or potentially interbreeding natural populations, which are reproductively isolated from other such groups" (Mayr 1942). Modern definitions for species often consider these groups units of evolution with characteristics that are mutually exclusive to different species. Using this concept of species would lead to a subspecies as reflecting a change in frequencies of alleles and therefore, would be subject to rapid and minor changes (Gaines 1978; Crother, 1992). Also, this uncertainty, when dealing with the subspecies concept, can be abused for statements of biodiversity when dealing with conservation biology. Through manipulation of the terminology reports of biodiversity could be greatly over- or underestimated depending on the motivation of the sampler. With today's concerns over natural populations the ability to distinguish between so called subspecies of organisms must be settled on a genetic level to test the variability of a group of organisms. Molecular techniques offer a method to check the genetic variability within these groups which not even an undisputed definition exists. Scorpions represent a group in which the taxonomic debate continues due to conserved features and overlapping of identifiers serving as defining features.

It is known that scorpions date back 425 to 450 million years (Silurian) at which point they are believed to have diverged from the Eurypterida, or water scorpions. The earliest scorpions are thought to have been aquatic, and other than the change in form associated with adaptation to land existence the basic body plan has not changed much in the last 425 million years. Modern scorpions have wide geographic distribution including every continent of the world excluding Antarctica. Scorpions are currently divided into 18 families with 159 genera and 1270 species (Fet and Sissom 2000). Their diverse habitats include desert, sevannas, grasslands, temperate forests, tropical forests, rain forests, intertidal zones, caves, and mountains.

The success of scorpions can be credited to their adaptable ecology, behavior, physiology, and life history. Many are capable of supercooling and upon return to normal temperatures can resume normal activities within a few hours. Another adaptive character is their low metabolic rate allowing some to survive for over a year without food. These traits make for an organism that has survived on the earth for over 400 million years and explain the morphological conservation witnessed within this group. The controversy in classification of this ancient group of organisms can be seen clearly within the following genera: *Hadrurus* Thorell (Iuridae), *Anuroctonus* Thorell (Iuridae), *Euscorpius* Thorell (Euscorpiidae), and *Centruroides* Marx (Buthidae).

Species of the genus *Hadrurus* (Thorell 1876), the "giant hairy scorpion", are widespread in the American deserts. Body size ranging from 100 to 125 mm in length make it among the largest terrestrial invertebrates in North America. They are found in the states of Oregon, Idaho, California, Nevada, Arizona, Utah, and in Mexico (Fig. 1). The genus *Hadurus* currently contains eight species, three within the United States and five, within Mexico. These species have been intensively studied. Originally, the genus was included in the family Vaejovidae, subfamily Vejovinae, and was then placed in the family Iuridae (Francke and Soleglad 1980). In 1976, Soleglad divided eight species of *Hadrurus* into subgroups based on trichobothrial counts. *H. concolor* Stahnke, 1969, *H. hirsutus* (Wood, 1863), and *H. pinteri* Stahnke, 1969 were placed in the *hirsutus* group, while the U.S. species *H. arizonensis* Ewing, 1928, *H. spadix* Stahnke, 1940, and *H. obscurus* Williams, 1970 were placed in the arizonensis group. Also there exist within the species *H. arizonensis* two subspecies *H. a. arizonensis* and *H. a. pallidus* Williams, 1970 (Williams, 1970) whose differences in morphological classification are coloration of carapace, with that of *H. a. pallidus* being much lighter with reported intermediates in coloration (Willians 1970). Likewise based on coloration was the classification of *H. obscurus* by Williams (1970). Both of these groupings were found to have similar trichobothrial patterns within each pair (Soleglad 1976).

A second genus of Iuridae found in North America is *Anuroctonus*. This genus was also originally placed in the Vaejovidae family, subfamily Vaejovinae. Stahnke (1974) removed *Anuroctonus* from this subfamily and placed it in the subfamily Hadrurinae. In 1981 Francke and Soleglad rearranged the family dividing it into two subfamilies, Iurinae (containing *Iurus* and *Calchas*) and Caraboctoninae (containing tribe Hadrurini with *Hadrurus* and tribe Caraboctonini with *Hadruroides* and *Caraboctonus*). Stockwell (1992) suggested that genus *Anuroctonus* is the sister group of genus *Hadrurus*, and reinstated the subfamily Hadrurinae for these two genera (Fet and Sissom 2000). Currently there is only one species within the genus *Anuroctonus;* however, at least three taxa have been proposed to exist. As a part of this study, we tested the level of divergence for three proposed species: the previously established *A. phaiodactylus* and two new species currently under description by M. E. Soleglad (unpublished), *A. pococki* and *A. bajae* (Fig. 2).

Genus *Euscorpius* is found in the area surrounding the Mediterranean and southern Europe. The taxonomy of this genus is extremely complicated and not well resolved. Euscorpius currently contains six species with 42 subspecies (Fet and Sissom, 2000) which makes it the most polymorphic genus of scorpions at below-species level.

Even the number of species within *Euscorpius* is not clear. One of the *Euscorpius* lineages is referred to as the "*Euscorpius mingrelicus* (Kessler, 1874) complex", which was shown by Bonacina (1980) to be different from the taxa associated with *E. germanus* (Koch, 1837). Scherabon (1987) using morphological analysis described two forms of *E. germanus* from Austria, the typical (T-form) and the "Kaarawanken-Form" (K-form). Morphological comparison of the "K-Form" with *E. germanus gamma* (Di Caporiacco, 1950) and *E. germanus histrorum* (Di Caporiacco, 1950) demonstrated that these forms are identical. In 1980, Bonacina considered both of these forms to be subspecies of *E. mingrelicus* (Bonacina 1980).

E. germanus is found within mountainous areas from the Balkans, Austria, northern Italy and southern Switzerland (Capra 1939, Valle 1969, Di Caporiacco 1950, Bonacina 1980, Crucitti 1993, Fet and Braunwalder 1997, Gantebein et al., 1998). This group was originally classified using traditional features including trichobothrial counts and patterns (Birula 1900, Hadzi 1931, Vachon 1962, 1981) and number of teeth on pectinal organs (Capra 1939). Using these qualifications Capra (1939) divided *E. germanus* into four forms. These elements were: a "typical" form found in East Italy (Trentino, Alto-Adige and Cadore), the A-form, in the Lombardian Alps (North Italy), the B-form found in the Piemontesian Alps (Northwest Italy), and the C-form in the Selva-di-Tarnova-region (modern Slovenia) and the Karawanken Alps of Austria. Then in 1950 Di Caporiacco conformed these forms and gave them status of subspecies (*E. g. germanus*, *E. g. alpha, E. g. beta*, and *E. g. gamma*). Bonacina (1980) suggested, due to overlapping of trichobithria counts and pectinal teeth, hybrids of *E. g. alpha and E. g. germanus*. This has caused confusion over the status of these taxa.

The fourth genus included in the study was that of *Centruroides*. This genus belongs to the largest of the scorpion families, Buthidae. This group was recognized as a genus in 1890 (Marx 1890). *Centruroides* is a large genus with 41 species and 24 subspecies (including nine nominotypical forms)

(Fet & Lowe 2000). The genus is only found in the New World, centered in Central America, the Carribean and Mexico, and extending widely into northern South America and the southern United States (Fig. 5). Centruroides includes small to quite large species, among them some of the most toxic for humans (Mexico). The taxon inhabits diverse environments from hot arid deserts to tropical rainforests (Fet & Sissom 2000). This group like the others represented here required a comparative morphology-independent technique, which could allow a comprehensive study of its taxonomy and phylogeny.

1.2 Mitochondrial DNA Analysis

Since the late 1970's the use of mitochondrial DNA homology for evolutionary studies has provided a tool with many benefits. Using mitochondrial DNA allows monitoring of gene flow within populations, effects of hybridization, and phylogenetic relationships (Moritz 1987). The mitochondrial present in animals, with one exception, is a duplex covalently closed circular molecule. In animals there are no interrupted genes and it contains no introns (Moritz 1987). It possesses highly conserved gene arrangement and product functionality. Essentially, mitochondrial DNA is passed through the female to offspring and has as a result virtually no recombination. It contains two ribosomal RNA genes, 22 transfer RNA genes, and 13 protein coding genes (Avise 1994). Another key feature of this molecule is its high rate of mutation of about 5 to 10 times faster than that of single copy nuclear DNA. This rate is not in all regions of the molecule. The areas of high mutation rate known as nonconserved regions and regions with low mutation rate known as conserved regions (Avise 1994).

The technique of polymerase chain reaction (PCR) allows production of a million times the starting DNA sequence copies. This amplification of template sequence allows for the use of small quantities of DNA from samples that have degraded, dried, frozen, or preserved. This process was first developed for detection of human T-cells lymphoma leukemia viruses, hepatitis B virus, and bacterial pathogens in water samples (Lewin 1994). The process of PCR consists of three steps where the determining factors are time and temperature. The first step is denaturation in which the temperature is raised to a range of 92 and 96 C to allow for a dissociation of the two strands of DNA. The second step is known as the annealing step during which the single stranded DNA primers bind to the to the now denatured DNA template. This stage of the process requires a temperature of near 50 C. The third step of the process is the extension phase. In this phase the temperature is raised to 72 to 74 C and the *Taq* DNA polymerase adds 35 to 100 nucleotides per second to the template strands between the binding sites of the attached primers. This extension process is done in a five to three prime direction on the DNA strand (Hillis 1996). This process is then repeated for 25 to 30 cycles, giving high yield of the desired segment of DNA.

An automated process sequences the selected region of DNA. This is accomplished usually by using the Sanger method, using fluorescently labeled DNA segments. Another method incorporates the use of radioactively labeled DNA fragments and is know as the Maxam-Gilbert technique (Hillis 1996). These fragments are detected using a tunable laser, which records as a segment passes a fixed point. This information is recorded as a chromatograph, which can be visually read as a DNA sequence (Hillis 1996)

This study attempted to apply molecular biology techniques to the taxonomy of scorpions, and to offer a comparative technique, which can be applied to different species, genera, and other taxonomic levels. We demonstrate the actual number of changes in a selected DNA sequence of a certain gene, which was necessary in order to arrive at distinctly divergent organisms. These data allow the use of a method of cladistic analysis based on percentage differences (distances) as well as on phylogenetic trees.

2. Material and Methods

Extraction of DNA was performed by first removing a appendage from a specimen and allowing to air dry for two hours at room temperature to allow the evaporation of all ethanol (specimens were stored in 96 % ethanol). Sample was then placed in liquid nitrogen and ground with a ceramic mortar and pestle to achieve shearing of cell membranes. A phenol/chloroform extraction was then performed upon the tissue (Protocol 1).

Next, a one-microliter amount of the sample is placed within the PCR mixture (Protocol 2). This is run in the Perkin-Elmer 2400 thermocycler with settings of 40 seconds in the three steps of annealing, extension, and denaturation. The annealing temperature was 48 degrees Celsius. The primers used in this amplification have the sequences CGATTTGAACTCAGATCA (forward 18-mer) and GTGCAAAGGTAGCATAATCA (reverse 20-mer). The corresponding positions in the mitochondrial genome are 12,867-12,887 and 13,218-13,310 from the *Drosophila yakuba* genome (Clary and Wolstenholme, 1985). Forward primer is universal for arthropods; reverse primer was developed by V. Fet (Gantenbein et al. 1999). Cleaning of the PCR product is achieved using a Ultrafree MC 30,000 cellulose centrifuge filters (Millipore Inc.)(Protocol 3). The products were then sequenced by The University of Georgia's DNA Core Facility.

Sequences were then aligned by hand or by the use of ClustalX alignment software. The aligned sequences were then analyzed using PAUP 4.0 (Phylogenetic Analysis Using Parsimony) software to achieve relations among the sequences.

2.1 *Hadrurus*

Thirteen adult specimens were taken for this study from 5 species of the genus *Hadurus*. The specimen locations and numbers assigned to them are as follows:

1) *Hadurus pinteri* Stahnke, 1969 from Isla Danzante, Baja California Sur, Mexico (collected by Dr. Victor Fet) was abbreviated as Hpin.
2) *H. concolor* Stahnke, 1969 from Bahia de Los Angeles, Baja California Norte, Mexico (collected by Dr. Victor Fet) was abbreviated as Hcon.
3) *H. spadix* Stahnke, 1940 from the vicinity of Hawthorne, Mineral Co., Nevada, USA (collected by G. Lowe) was abbreviated Hspa

4) *H. obscuris* Williams, 1970 from Anza-Borrego Desert State Park, San Diego, California, USA (collected by M. Soleglad) two specimens abbreviated Hob1 and Hob2

5) *H. arizonensis* Ewing, 1928 consisted of 8 specimens

(5.1) near Death Valley National Monument, Inyo Co., California, USA (collected by G. Lowe) abbreviated Har1

(5.2) Yuma, Yuma Co., Arizona, USA (collected by J. Bigelow) abbreviated Har2

(5.3) Dateland, Yuma Co., Arizona, USA (collected by R. Farley) abbreviated Har3

(5.4), (5.5), and (5.6) Anza-Borrego Desert State Park, San Diego Co., California, USA (collected by M. Soleglad) abbreviated Har4, Har5, and Har6

(5.7) Indio, Riverside Co., California, USA (collected by S. Tallarovic) abbreviated Har7

(5.8) Soda Springs, near Baker, San bernandino Co., California, USA (collected by S. Tallarovic) abbreviated Har8

The outgroups for the analysis were *Euscorpius flavicaudis* an European species (Degeer, 1778) and

Centuroides exilicauda (Wood, 1863) a Mexican species of the family Buthidae (collected by Dr. Victor Fet). The choice of *C. exilicauda*, according to Sissom, represents the most remote extant scorpions to the family Iuridae (Sissom, 1990).

2.2 *Anuroctonus*

Representatives of the only recognized species of the genus, *A. phaidactylus*, the type species (AnUT), were tested against the proposed species of *A. bajae* and *A. pococki* mentioned above. Four specimens of the proposed species *A. bajae* were analyzed and were assigned names according to collection sites. The outgroup chosen for the comparison was *H. arizonensis*.

A. bajae: (suggested species)

AnJT	(Jasper Trail, San Diego Co.)
AnRV	(Anza, Riverside Co.)
AnCH	(Chihuahua Road, San Diego Co.)
AnPM	(Pinyon Mountain, San Diego Co.)

A. pococki (suggested species)

AnMO	(Monterey, Monterey Co., CA)

A.phaiodactylus (recognized species)

AnUT = (Garrison, Utah)

2.3 *Euscorpius*

The first of the two *Euscorpius* studies was performed with a total of 11 specimens, 8 of the suggested *E. germanus*, 2 of the proposed *E. gamma*, and one of the outgroup *E. flavicaudis*(Fig. 3). The *E. germanus* samples included specimens from Schluderns, from Brixen, from Bremer, from Vols, from Oberdrauberg, from Dobratsch, from Shytt, from Crnice. The E. gamma specimens included samples from Koschuta and from Trogerner-Klamm. The outgroup was formed of 11 specimens of E. falvicaudis from Lauris. The analysis of these sequences was performed by B. Gantenbein, and obtaining of the PCR products for sequencing was a combined effort between the labs of V. Fet and B. Gantenbein.

	E. germanus								*E. gamma*		*E. flavicaudis*
Population	Schluderns (4)	Brixen (3)	Bremer (1)	Vols (2)	Oberdrauburg (5)	Dobratsch (6)	Schnutt (7)	Crnice (8)	Koschuta (10)	Trogerner Klamm (9)	Lauris

In the second *Euscorpius* study, twenty-six specimens of *Euscrorpius germanus* were collected from various locations in Middle and Western Europe. The specimens were labeled according to the location at which they were collected and their corresponding numbers for figure 4 is as follows:

E. germanus

Western Clade

Abbreviation	Site	Map number (Fig.1)
EalGO	Gondo	(1)
EalFO	Fontainemore	(2)
EalRA	Rancate	(3)
EalSO	Sottoponte	(7)
EalSP	San Pellegrino	(12)
EalSG	San Giovan	(13)
EalOL	Olmo al Brembo	(15)

EalCA	Carona	(19)
EalTA	Tavernola	(23)
EalMO	Molina di Ledra	(26)
EalML	Marling	(27)

Eastern Clade

EgeSM	Sta Maria	(29)
EgeSH	Scluderns	(31)
EgeVO	Vols	(35)
EgeVE	Vetriolo	(37)
EgeME	Mezzocana	(41)
EgeBO	Borca	(42)
EgeST	Starkenbach	(44)
EgeTZ	Tarrenz	(45)
EgeKR	Kranzach	(46)
EgeDE	Dellach	(47)
EgeOB	Oberdrauburg	(48)
EgeCR	Crnice	(52)

E. gamma

EgaKO	Koschuta

E. flavicaudis

EflLA	Lauris

2.4 *Centrurcides*

The study of *Cetruroides* included 10 populations of *C. exilicauda* (Fig. 6) from Southwestern North America and Central America. The choices for outgroup comparison were *C. vittatus* from Arkansas and *C. bani* from the Dominican Republic. Each specimen is abbreviated according to its site of collection with *C. vittatus* and *C. bani* labeled as Cvit abd Cban.

(Ce = *Centruroides exilicauda*)

CeBA = Bahia de Los Angeles, Baja California, Mexico, coll. V. Fet
CeSI = San Ignacio, Baja California, Mexico, coll. V. Fet
CeJU = Juncalito, Baja California, Mexico, coll. V. Fet
CeCA = Cabo San Lucas, Baja California, Mexico (Cape area, coll. M. E. Soleglad)
CeBC = Black Canyon City, Arizona, USA, coll. B. Tomberlin
CeTU = Tucson, Arizona, USA, , coll. B. Tomberlin
CeTE = Tempe, Arizona, USA, coll. J. Bigelow
CeYU = Yuma, Arizona, USA, coll. J. Bigelow
CePP = Puerto Penasco, Sonora, Mexico,), coll. G. A. Polis
CeGU = Guaymas Islands, Sonora, Mexico, coll. G. A. Polis
CvAR = *C. vittatus*, Russelville, Arkansas, coll. T. Yamashita
Cban = *C. bani*, Dominican Republic, coll. D. Huber

Protocol 1 (DNA Extraction)

- ground tissue is placed in 900 ul of Lifton buffer at a pH of 9 and 40ul of proteinase K.
- then incubated in a water bath overnight at 47 to 51 degrees Celsius.
- place in ice bath for 15 minutes until sediment forms and centrifuge for 15 minutes at 14,000 rpm.
- the top liquid layer is then removed to another tube and the sediment discarded.
- 2 ul of RNAse is then added and incubated at room temperature for 20 minutes (this step is optional).
- 150ul of equilibrated phenol is added

- 150ul of CIA (24 parts chloroform and 1 part isoamyl alcohol)
- mixture is then vortex mixed and spun at 14000 rpm for 5 minutes
- top aqueous layer is removed and transferred to another tube, while the lower layer is discarded
- repeat previous 3 steps twice more
- add 150 ul 3M sodium acetate
- add 300 ul 100% isopropanol
- invert several times and place in freezer for at least two hours
- spin at 14000 rpm for 15 minutes
- remove all liquid leaving pellet
- wash with 150 ul of 70% ethanol
- spin for 5 minutes at 14000 rpm and remove liquid
- allow to air dry
- suspend DNA in 150 ul of ultrapure water

Protocol 2 (PCR reaction)

Master Mix of 12 reactions each 50 ul

-276 ul Ultrapure water

-72ul Perkin Elmer MgCl2

-60 ul Perkin Elmer PCR buffer

-60 ul Perkin Elmer Nucleotide Mix

-3-4 ul Perkin Elmer Amplitaq

Protocol 3 (PCR product cleaning)

-remaining product after electrophoretic verification of DNA amplification

-an additional 350 ul ultrapure water is added to filter tube

-spin at 7000 rpm for 5 minutes

-filtered watered is discarded

-add 250 ul ultrapure water to filter tube

-spin at 7000 rpm for 5 minutes

-filtrate is then suspended in 30 ul of ultrapure water and stored for sequencing

Figure 1

Ranges of the species of the genus *Hadrurus* (Williams 1980)

(*H.a. pallidus*= closed circle)

(*H. concolorous*= closed squares)

(*H. hirsutus*= open triangles)

(*H. obscurus*= open circles)

(*H. pinteri*= open squares)

Figure 1

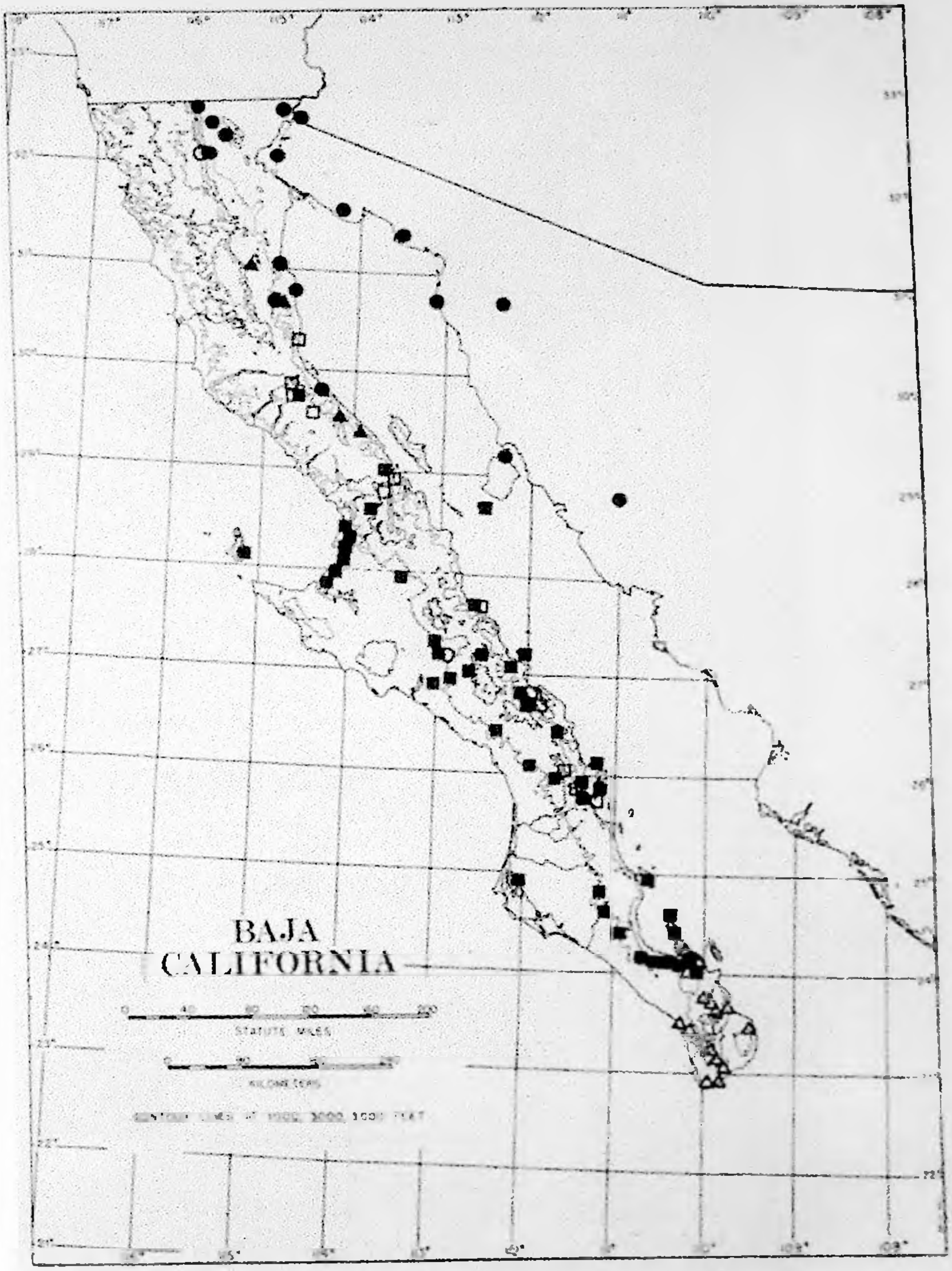

BAJA
CALIFORNIA
0 40 80 120 160 200
STATUTE MILES
KILOMETERS
CONTOUR LINES AT 1000, 3000, 5000 FEET

Figure 2

Anuroctonus ranges

(suggested species by M. Solglad)

Figure 2

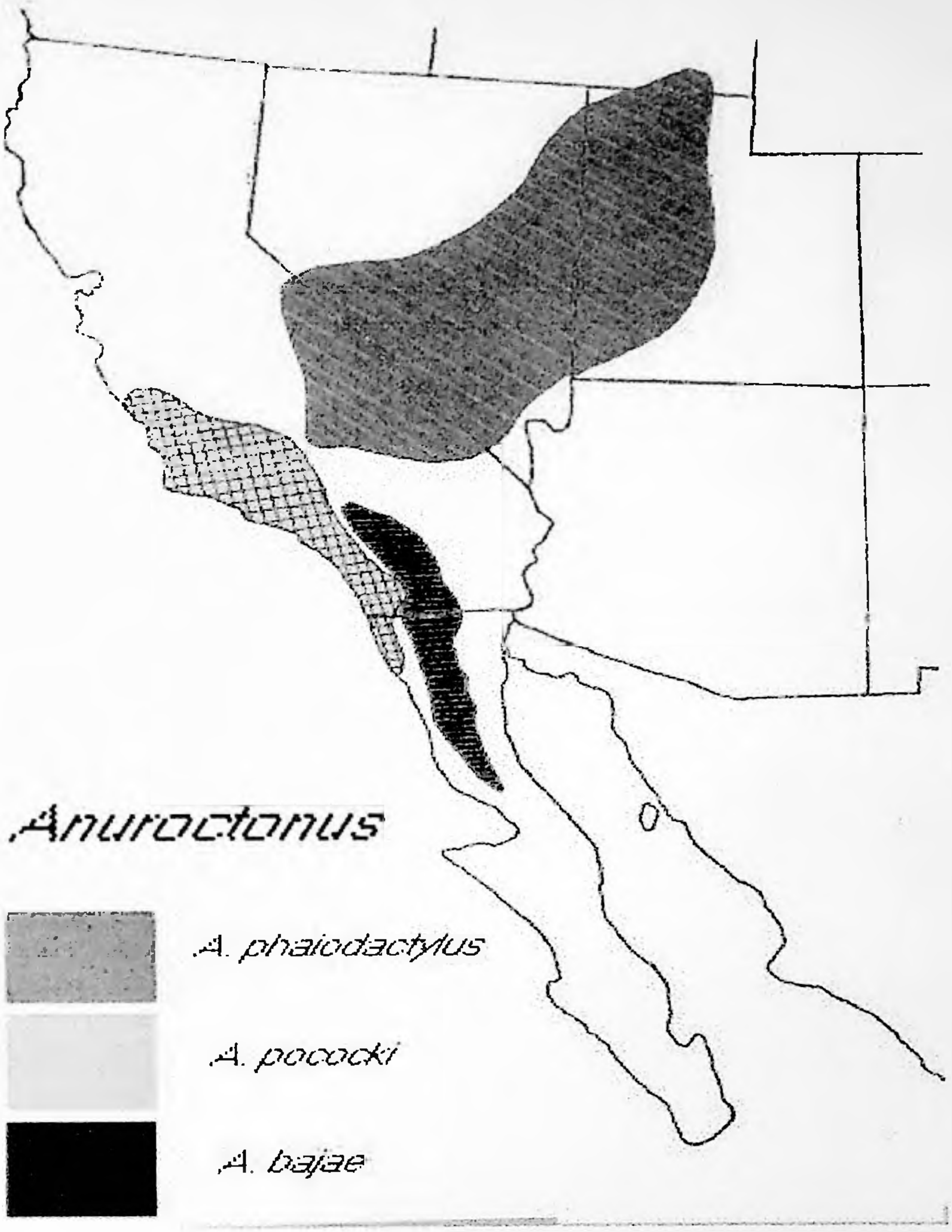


Figure 3

Euscorpius gamma study sites

1. Bremer
2. Vols
3. Brixen
4. Schulderns
5. Oberdrauberg
6. Dobratsch
7. Schutt
8. Crnice
9. Trogerner-Klamm
10. Koschuta

Figure 3

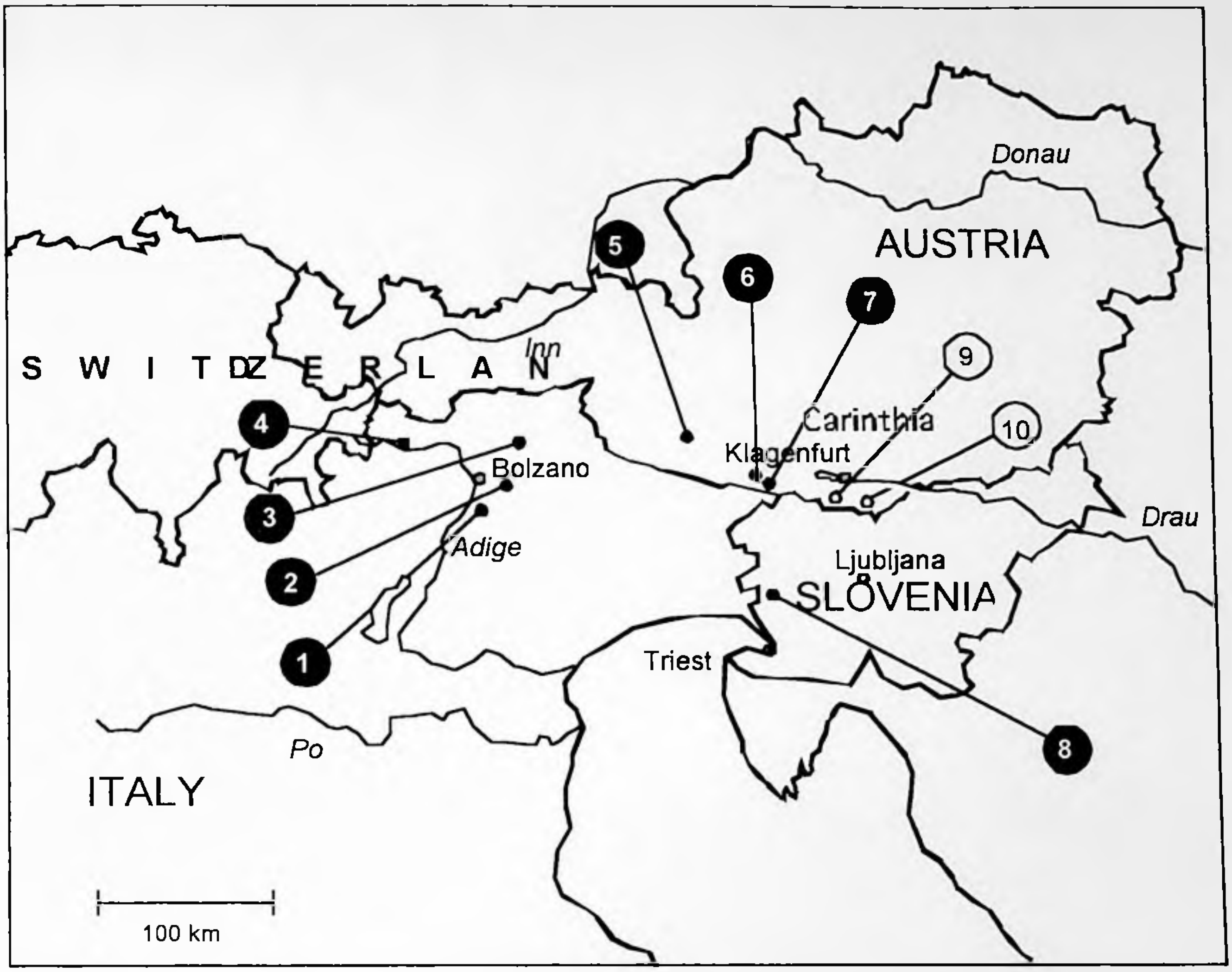

Donau
AUSTRIA
5
6
7
9
10
Inn
S W I T Z E R L A N D
4
Carinthia
Klagenfurt
Bolzano
3
Drau
Adige
2
Ljubljana
SLOVENIA
1
Triest
8
Po
ITALY
100 km

Figure 4

Euscorpius alpha study sites

Figure 4

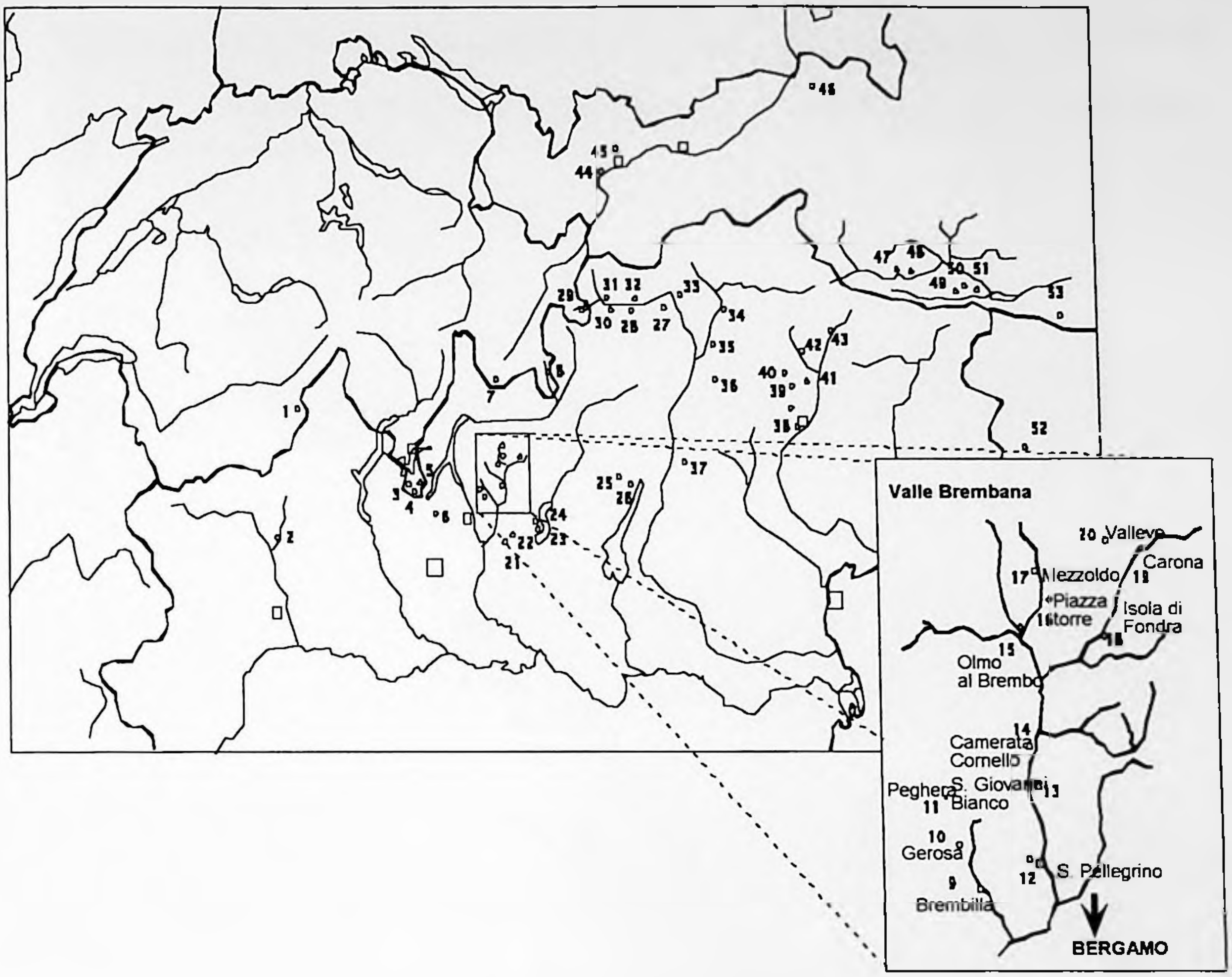

Valle Brembana
20 Valleve
Carona
19
17 Mezzoldo
Piazza
16 torre
Isola di
Fondra
18
15
Olmo
al Brembo
14
Camerata
Cornello
S. Giovanni
Bianco
13
Peghera
11
10
Gerosa
9
Brembilla
12
S. Pellegrino
BERGAMO
1
2
3
4
5
6
7
8
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53

Figure 5

Range of *Centruroides exilicauda* (Williams 1980)

Figure 5

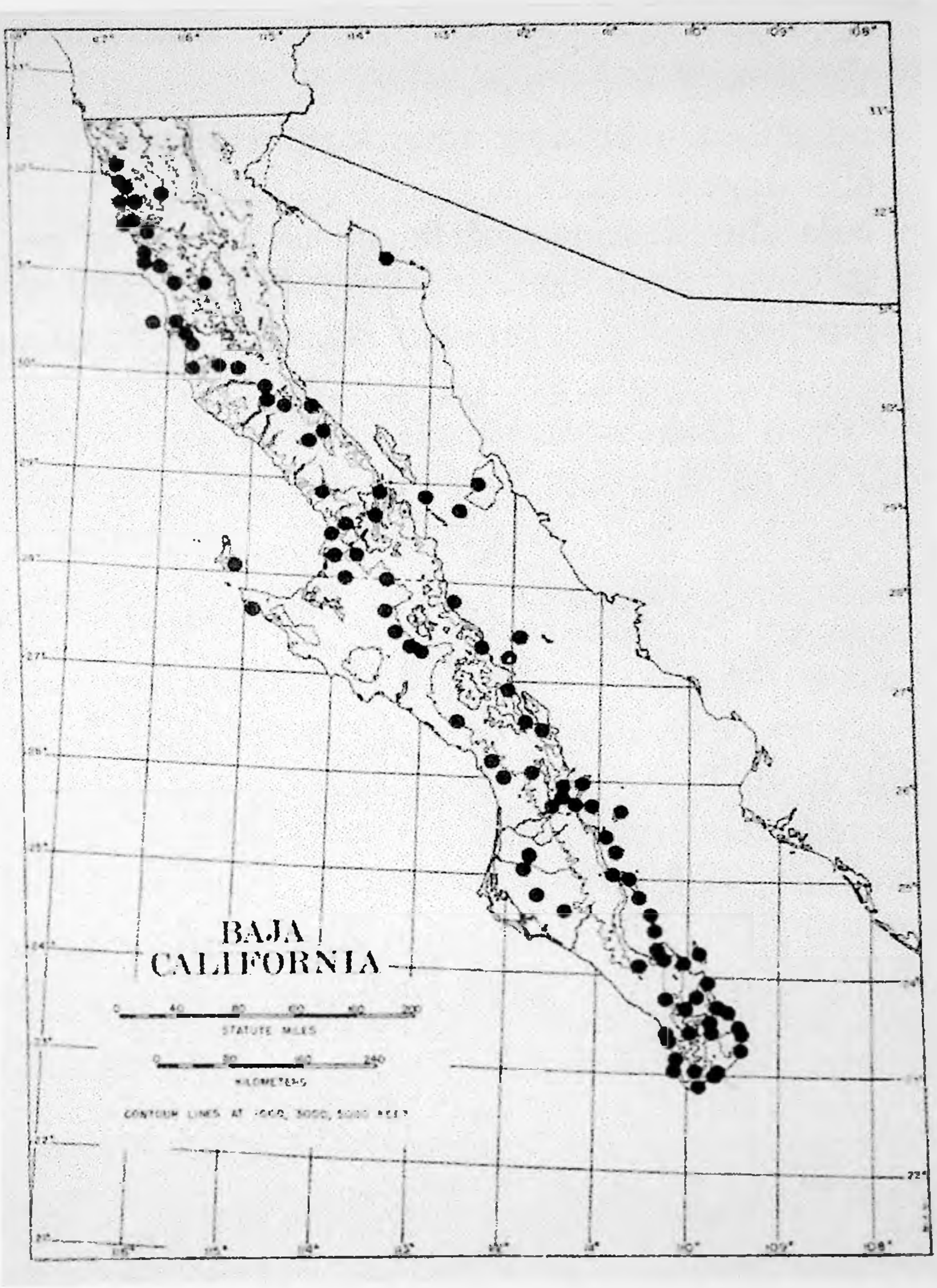
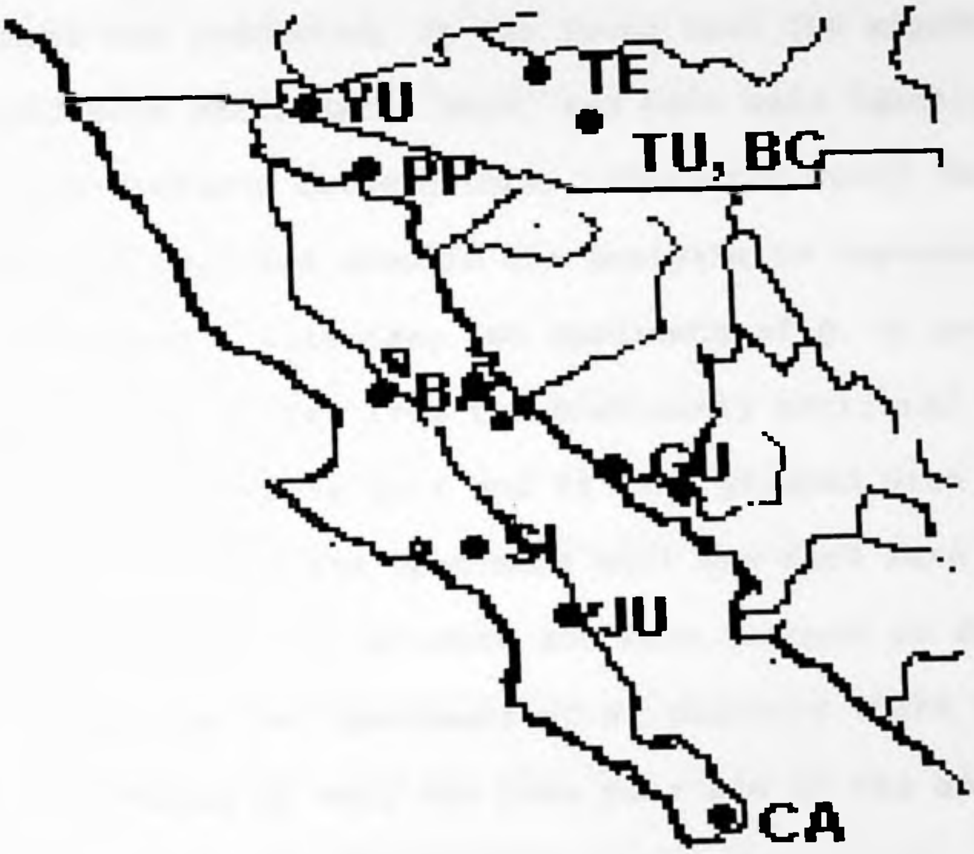

BAJA CALIFORNIA
0 40 80 120 160 200
STATUTE MILES
0 80 160 240
KILOMETERS
CONTOUR LINES AT 1000, 3000, 5000 FEET

Figure 6

Anuroctonus map of collection sites

CeBA = Bahia de Los Angeles, Baja California, Mexico, coll. V. Fet
CeSI = San Ignacio, Baja California, Mexico, coll. V. Fet
CeJU = Juncalito, Baja California, Mexico, coll. V. Fet
CeCA = Cabo San Lucas, Baja California, Mexico (Cape area, coll. M. E. Soleglad)
CeBC = Black Canyon City, Arizona, USA, coll. B. Tomberlin
CeTU = Tucson, Arizona, USA, , coll. B. Tomberlin
CeTE = Tempe, Arizona, USA, coll. J. Bigelow
CeYU = Yuma, Arizona, USA, coll. J. Bigelow
CePP = Puerto Penasco, Sonora, Mexico, coll. G. A. Polis
CeGU = Guaymas Islands, Sonora, Mexico, coll. G. A. Polis

Figure 6



3. Results

3.1 *Hadrurus*

When the analysis of the sequences obtained from the specimens was conducted, it was found that the sequences of the specimens Har2, Har3, Har4, and Har4 were identical, with zero distance between them. Therefore, only the sequence of Har2 was used in the analysis to represent all four specimens. Likewise, two specimens of *H. a. pallidus*, Har6 and Har7, differ from the previously mentioned sequence by only base pair and is thus grouped with the others into Har2. The specimens Har1 and Har8 were found to have an identical sequence and were grouped as Har1. And finally the two specimens of *H. obscuris* (both from San Diego) differed in only one base pair and in the analysis were grouped into group Hob1.

The distance matrix (Table 1) consists of six analyzed groups and two outgroups. The genetic distance between *Hadrurus* specimens ranges from 1.1% to 12.5%. When compared to outgroups genetic distance was 21% to 35% confirming all *Hadrurus* sequences are in fact *Hadrurus*.

The phylogenetic tree formed internal clades of *H. concolor* with *H. arizonensis* (Har1 and Har2) and a second clade of Hob1 (*H. obscurus*) and Hspa (*H. spadix*)(Fig. 9).

Also another internal clade is formed by the two haplotypes of *H. arizonensis*. *H. pinteri* is supported by bootstrap resampling as a sister group with the other clades (100%) and the clade of Har1 and Har2 as two haplotypes is also supported 100% (Fig. 10). The clade of *H. spadix* and *H. obscurus* are supported 71% by bootstrap (Fig.10).

3.2 *Anuroctonus*

Analysis of the sequences from the *Anuroctonus* samples resulted in the formation of distinct clades (Fig. 11). Two major clades were formed, for which bootstrap analysis yielded a 99% and 93% values (Fig.12). One clade is formed by the samples from Monterey Co., California (AnMo) and Garrison, Utah (AnUt), the type locality (93%), while the other is formed by those from southern California (99%) (proposed *A. bajae* species). Also clearly divergent are the AnCH and the clade containing AnJT and AnPM, with an 80% bootstrap value. The distance matrix gives 4% to 6% difference with organisms found within a single terminal clade and 7% to 9% with those outside this clade (Table 2).

3.3 *Euscorpius*

Analyses of the DNA sequences of the Euscorpius study were done by Benjamin Gantenbein, as this study was a joint project between the two laboratories. There were two parts in this project: (a) determining status of *E. gamma* (Austria and Slovenia); and (b) investigating divergence within *E. germanus* (Italy, Switzerland, Austria, and Slovenia).

The first of the two *Euscorpius* studies resulted in a single phylogenetic tree. The five in-group *Euscorpius* specimens were split clearly into two clades (Fig. 13). These clades correspond to the "T-form" and the "K-form" of Scherabon (1987). Within the "T-Form clade" the distance matrix showed a genetic distance of 0.8% to 3.5%. Within the "K-Form" there is a genetic distance of 1.6%. The distance between the two clades range from 5.6% to 7.5% suggesting different morphospecies within the same genus. The distance between the two clades and the outgroup, *E. flavicaudis*, was from 9.8% to 11.5%.

In the second study, a phylogeny within *E. germanus* was revealed: the entire set of samples fell into two clear deep clades. When analyzed, the DNA variation in *E. alpha* ("western clade") showed 45 out of 353 bases were polymorphic and within the *E. germanus* ("eastern clade") sequence 7 out of 369 were found to be polymorphic. The phylogenetic analysis of these sequences resulted in one tree indicating two clades (Fig.14) within *E. germanus* having bootstrap values of 96% and 90%. The Adige River in North Italy divides these population groups, and the genetic distance between these clades is comparable to that between *E. g. germanus* and *E. gamma*. As a result of this *Euscorpius* study, *E. gamma* and *E. alpha* were raised to the species rank.

3.4 *Centuroides*

Centruroides analyis gave three possible trees (Figs. 15-17). All three show two distinct clades with minor rearrangement of the component specimens. None of these rearrangements created a new supported clade different from those on any tree. The separation of the two clades has a bootstrap value of 90% (Fig 19). The separation of the clade containing CeBA, CeSI, CeJU has a bootstrap value of 98% while the other larger clade shows a bootstrap value of 73% with its separation (Fig. 19). Also CeSI anjd CeJU are separated with a value of 84 and CePP and CeCA are equally different form the clades of CeBC and CeTU and the clade

CeTE and CeGU as from each other with a bootstrap value of 78%. The distance matrix shows distance within the main clades is less than 6% while that between members of different clades is greater than 7%.

Figure 7

Hadrurus phylogentic tree 1

Figure 7

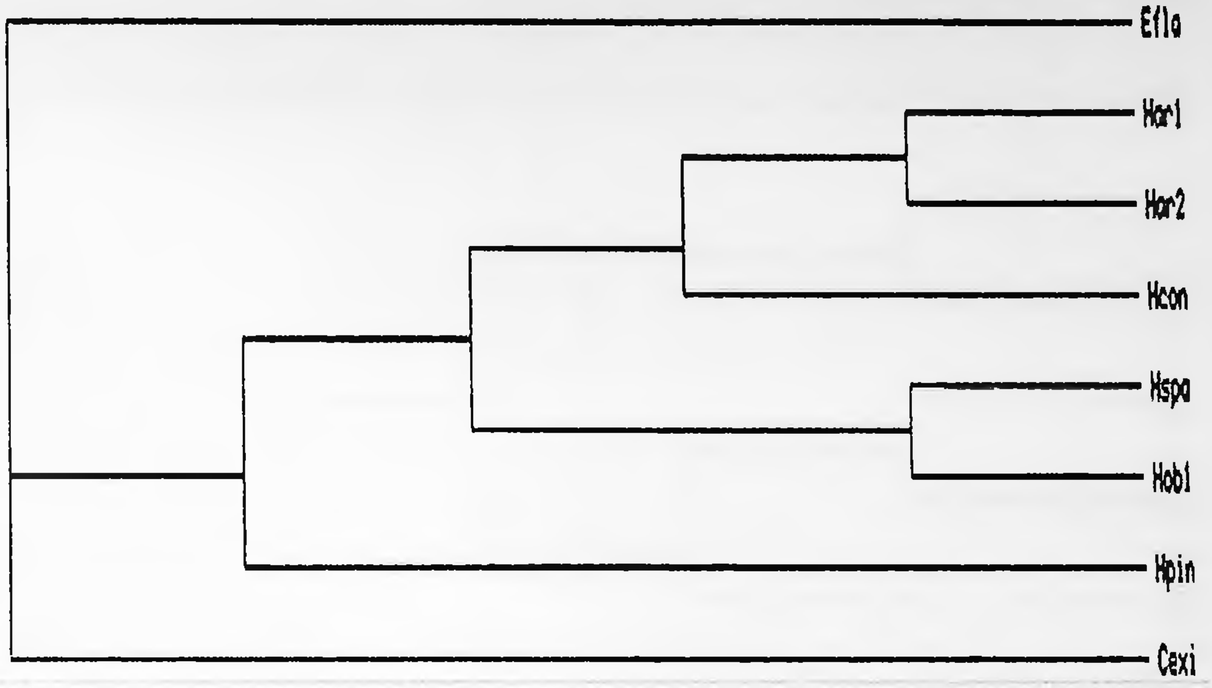


Figure 8

Hadrurus phylogenetic tree 2

Figure 8

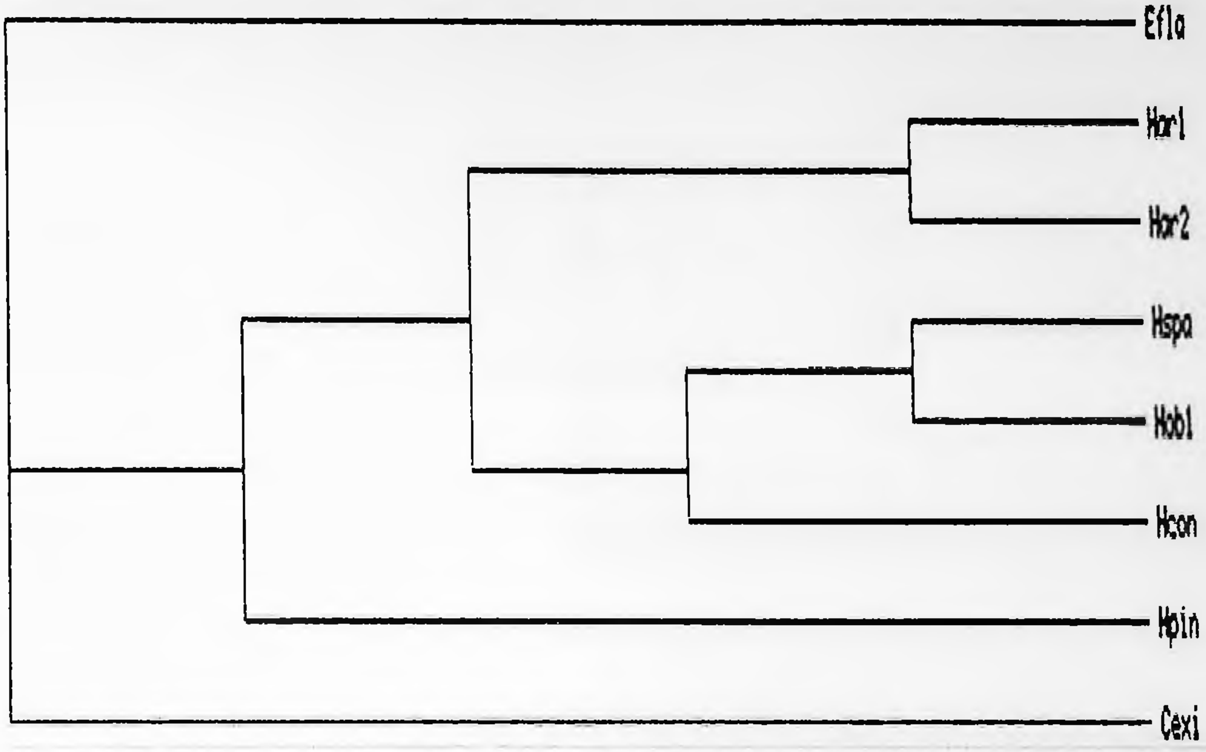


Figure 9

Hadrurus phylogenetic consensus tree

Figure 9

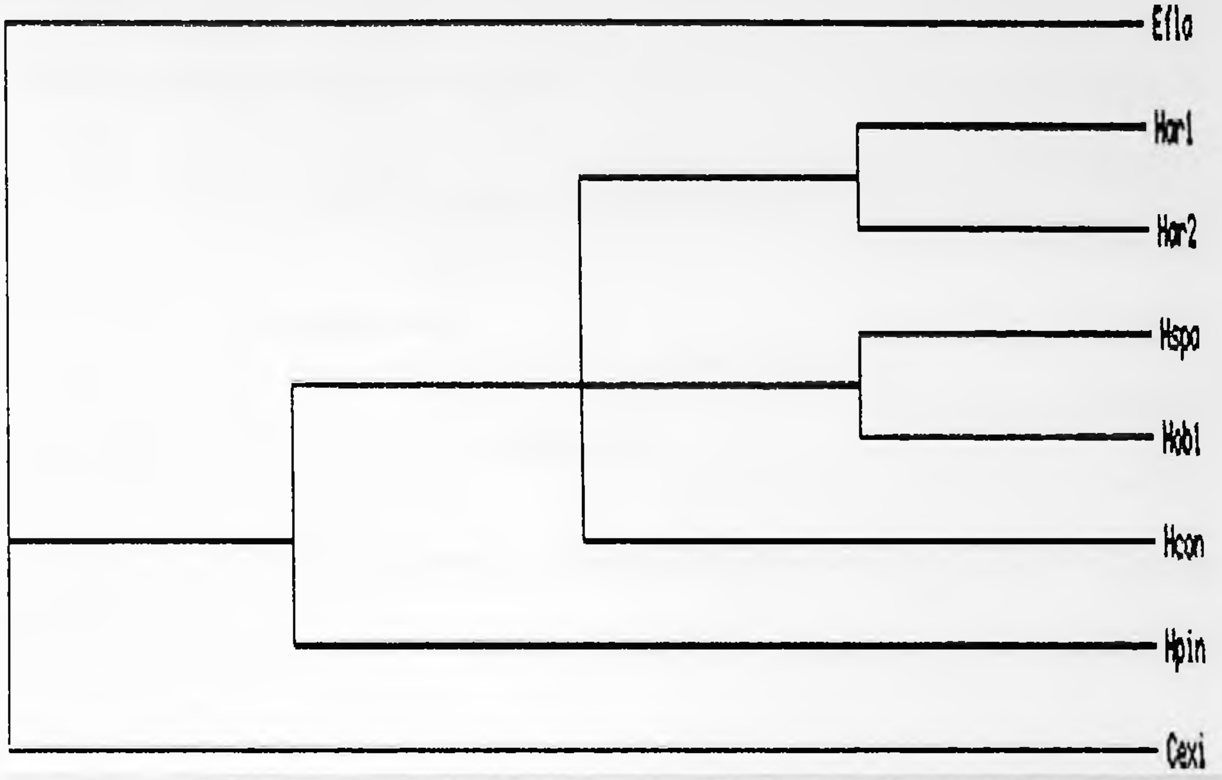


Figure 10

Hadrurus phylogenetic bootstrap analysis

Figure 10

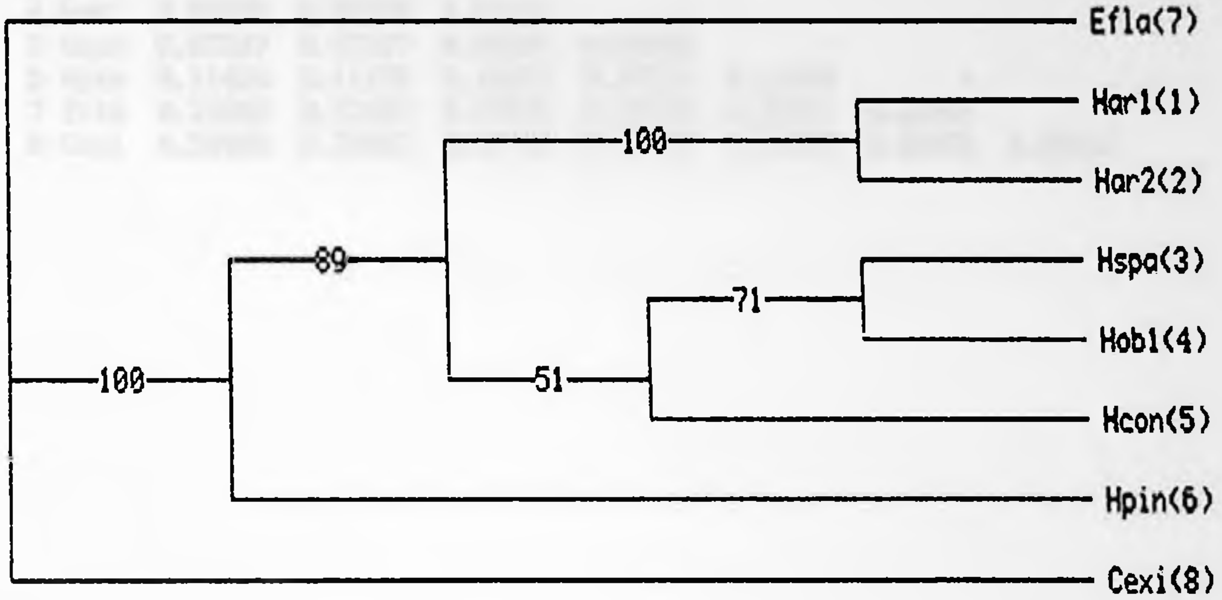

Efla(7)
Har1(1)
Har2(2)
Hspa(3)
Hob1(4)
Hcon(5)
Hpin(6)
Cexi(8)
100
89
100
71
51

Table 1

Distance Matrix *Hadrurus*

	1	2	3	4	5	6	7
1 Har1	-						
2 Har2	0.01087	-					
3 Hspa	0.05707	0.06250	-				
4 Hobl	0.06250	0.06793	0.02446	-			
5 Hcon	0.07337	0.07337	0.06250	0.06793	-		
6 Hpin	0.11429	0.11156	0.10067	0.09516	0.12264	-	
7 Efla	0.22408	0.22408	0.21574	0.21565	0.21311	0.22397	-
8 Cexi	0.28996	0.29283	0.30708	0.30715	0.29857	0.31063	0.28112

Figure 11

Anuroctonus phylogenetic tree

Figure 11

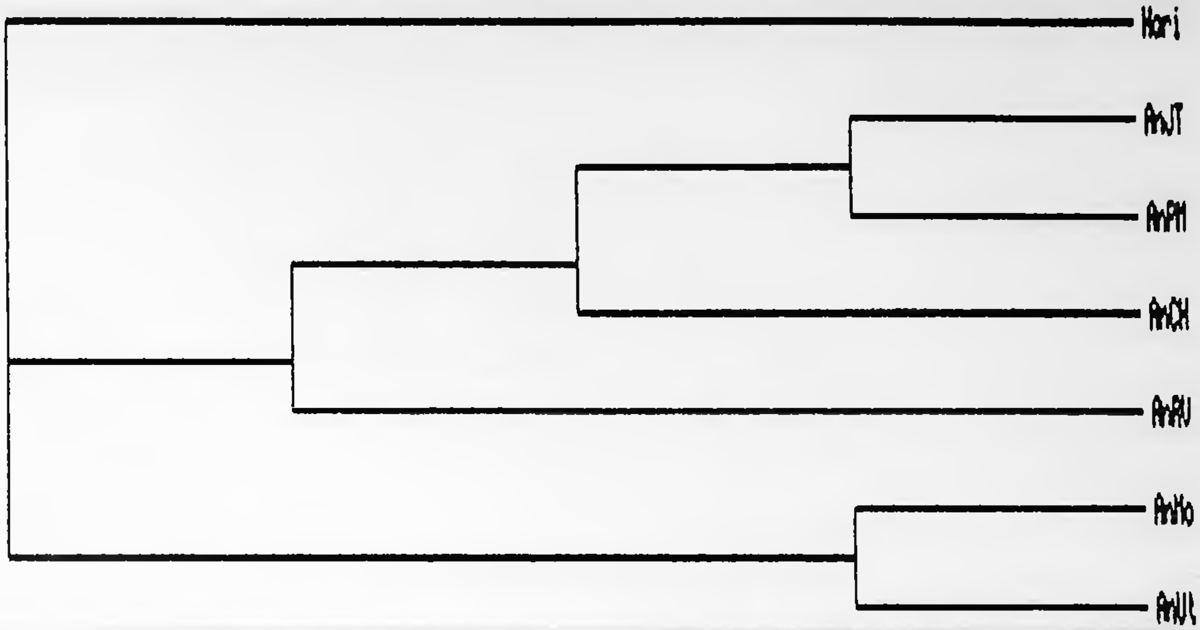


Figure 12

Anuroctonus phylogenetic bootstrap analysis

Figure 12

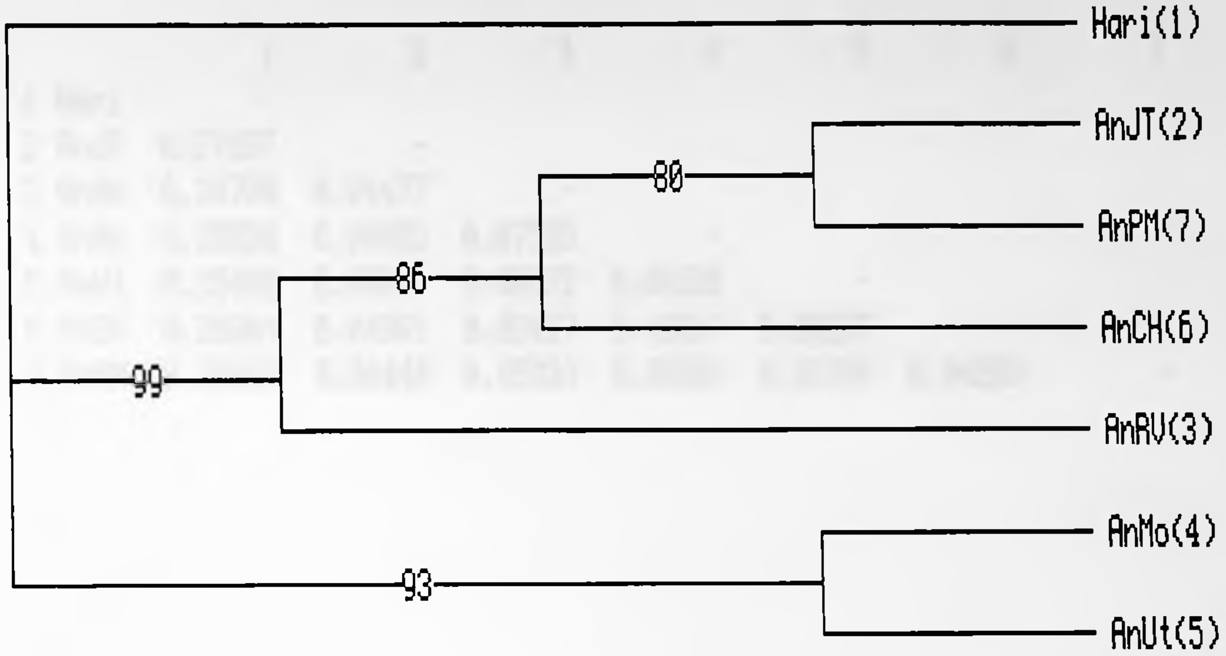


Table 2
Distance Matrix *Anuroctonus*

	1	2	3	4	5	6	7
1 Hari	-						
2 AnJT	0.27637	-					
3 AnRV	0.24780	0.04477	-				
4 AnMo	0.25020	0.09095	0.07505	-			
5 AnUt	0.25486	0.08871	0.08072	0.04536	-		
6 AnCH	0.26961	0.04391	0.02417	0.10317	0.08292	-	
7 AnPM	0.26891	0.04448	0.05923	0.09893	0.08306	0.04293	-

Figure 13

Euscorpius phylogenetic tree showing only the T-Form and K-Form

Figure 13

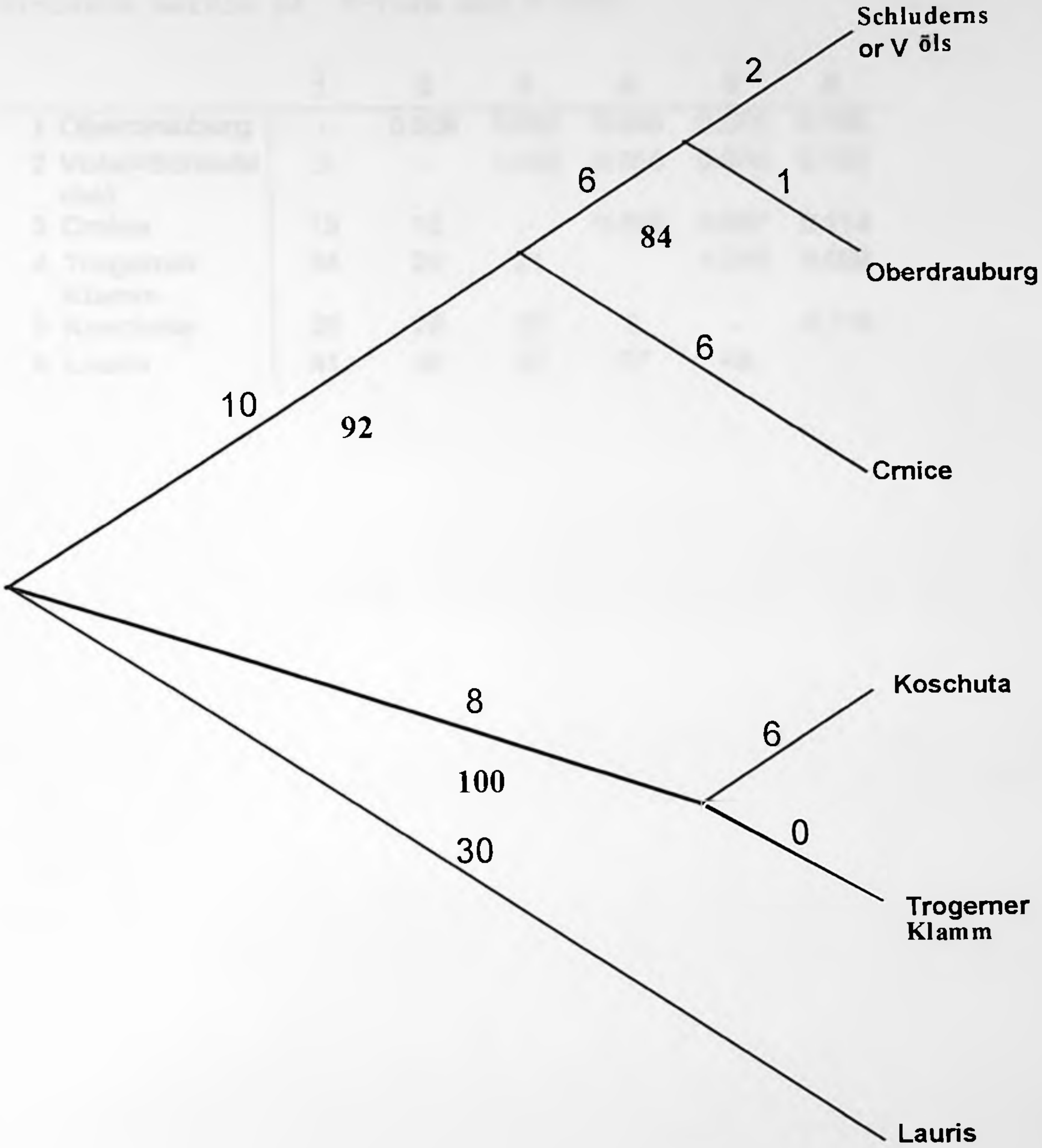

Schluderns or Völs
2
1
Oberdrauburg
6
84
6
Crnice
10
92
Koschuta
8
6
100
0
30
Trogerner Klamm
Lauris

Table 3
Distance matrix of T-Form and K-Form

	1	2	3	4	5	6
1 Oberdrauburg	-	0.008	0.035	0.064	0.075	0.109
2 Vols(=Schluderns)	3	-	0.035	0.058	0.075	0.103
3 Crnice	13	13	-	0.056	0.067	0.114
4 Trogerner Klamm	24	22	21	-	0.016	0.098
5 Koschuta	28	28	25	6	-	0.115
6 Lauris	41	39	43	37	43	-

Figure 14

Euscorpius phylogentic tree for *E. alpha* study

Figure 14

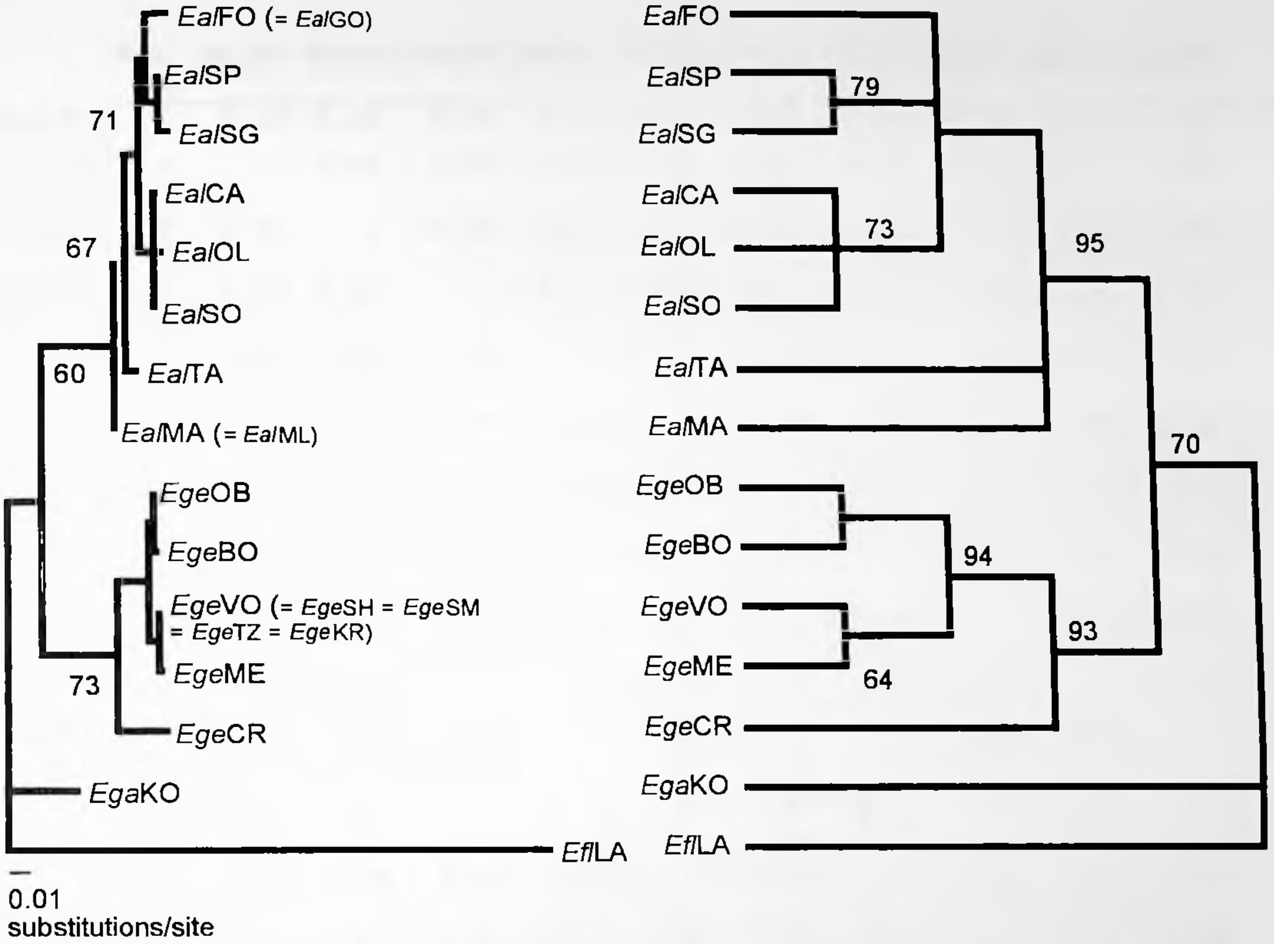

*Eal*FO (= *Eal*GO)
*Eal*SP
*Eal*SG
71
*Eal*CA
67
*Eal*OL
*Eal*SO
60
*Eal*TA
*Eal*MA (= *Eal*ML)
*Ege*OB
*Ege*BO
*Ege*VO (= *Ege*SH = *Ege*SM = *Ege*TZ = *Ege*KR)
*Ege*ME
73
*Ege*CR
*Ega*KO
*Efl*LA
0.01 substitutions/site
*Eal*FO
*Eal*SP
79
*Eal*SG
*Eal*CA
73
*Eal*OL
*Eal*SO
95
*Eal*TA
*Eal*MA
70
*Ege*OB
*Ege*BO
94
*Ege*VO
93
*Ege*ME
64
*Ege*CR
*Ega*KO
*Efl*LA

Table 4
Euscorpius alpha study
Distance matrix

	EalFO	EalSP	EalSC	EalCA	EalOL	EalSO	EalTA	EalMA	EgeOB	EgeBO	EgeVO
EalFO	–	0.02	0.02	0.02	0.02	0.02	0.03	0.02	0.06	0.06	0.06
EalSP	0.02	–	0.01	0.02	0.02	0.01	0.02	0.02	0.06	0.06	0.06
EalSC	0.03	0.01	–	0.02	0.02	0.02	0.02	0.02	0.07	0.06	0.07
EalCA	0.02	0.03	0.04	–	0.01	0.00	0.02	0.02	0.07	0.06	0.07
EalOL	0.03	0.02	0.03	0.01	–	0.01	0.02	0.02	0.07	0.07	0.07
EalSO	0.03	0.02	0.03	0.00	0.01	–	0.02	0.02	0.06	0.06	0.06
EalTA	0.05	0.05	0.05	0.05	0.03	0.03	–	0.01	0.06	0.06	0.06
EalMA	0.04	0.04	0.04	0.04	0.03	0.03	0.01	–	0.06	0.05	0.06
EgeOB	0.41	0.49	0.54	0.64	0.56	0.49	0.31	0.31	–	0.00	0.01
EgeBO	0.39	0.47	0.52	0.62	0.54	0.47	0.30	0.30	0.00	–	0.01
EgeVO	0.41	0.49	0.54	0.64	0.56	0.49	0.31	0.31	0.01	0.01	–
EgeME	0.51	0.62	0.69	0.81	0.70	0.61	0.39	0.39	0.01	0.01	0.00
EgeCR	0.38	0.46	0.50	0.36	0.40	0.45	0.19	0.18	0.06	0.07	0.07
EgaKO	0.25	0.20	0.22	0.25	0.19	0.20	0.10	0.09	0.43	0.36	0.43
EflLA	0.45	0.79	0.86	0.65	0.62	0.65	0.48	0.45	0.65	0.59	0.62

Figure 15

Centruroides phylogenetic tree 1

Figure 15

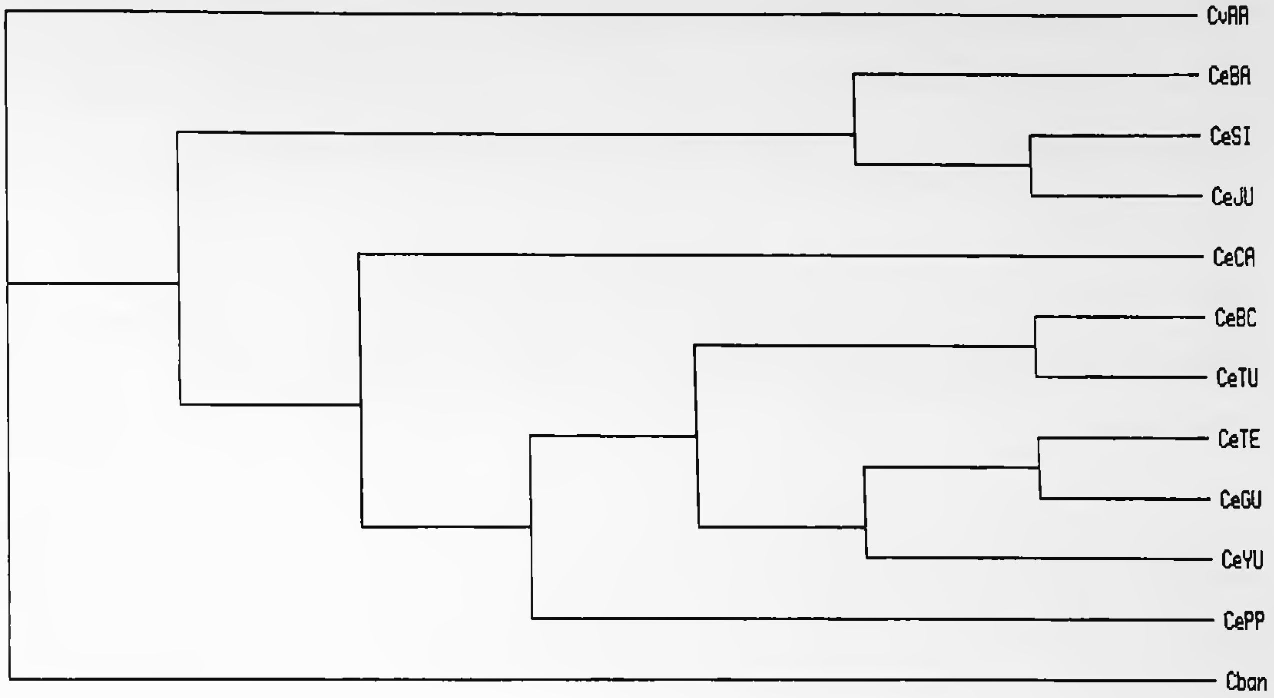


Figure 16

Centruroides phylogenetic tree 2

Figure 16

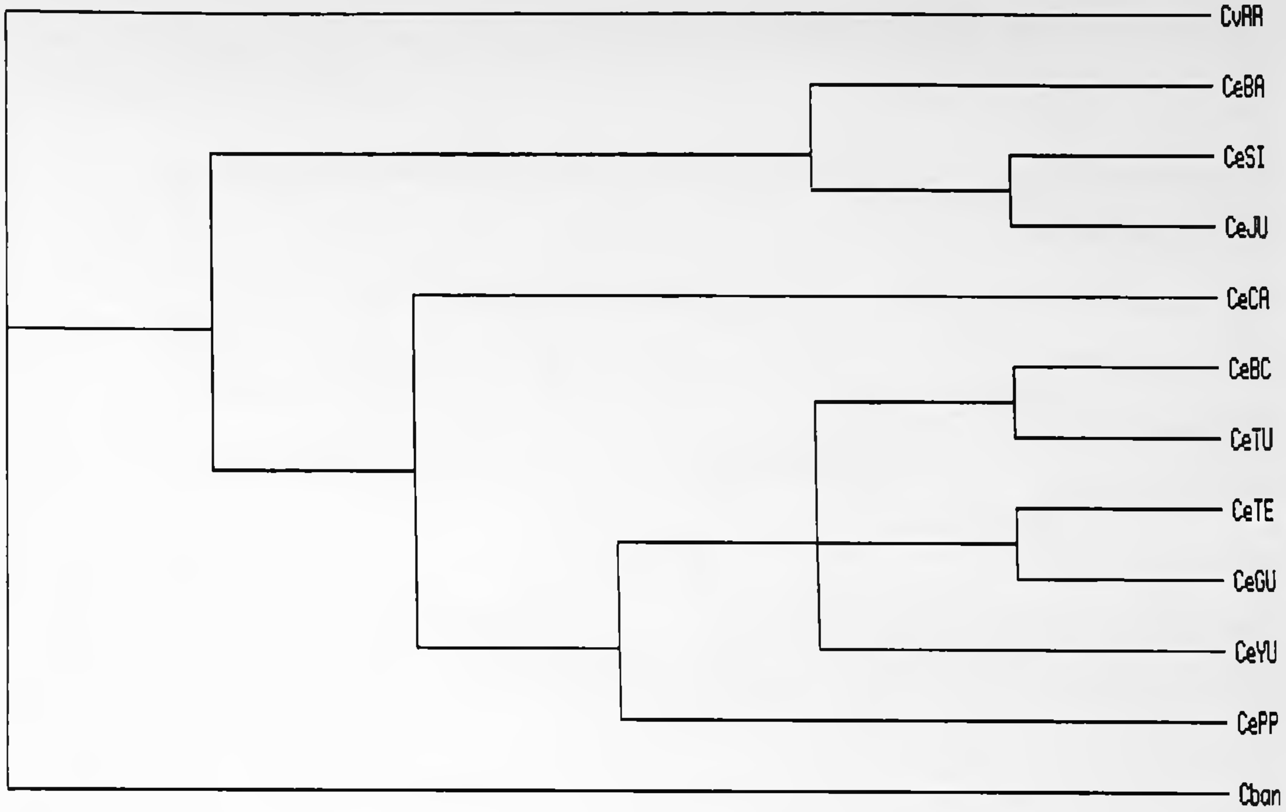


Figure 17

Centruroides phylogenetic tree 3

Figure 17

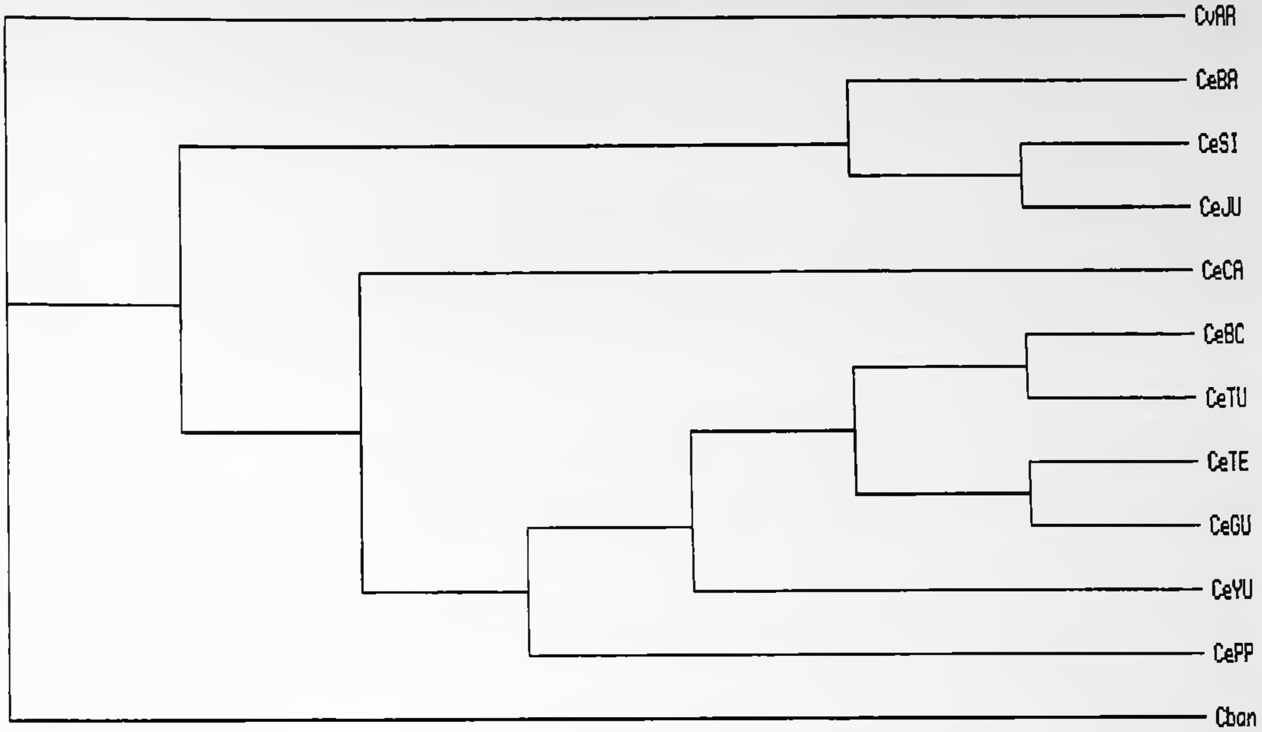


Figure 18

Centruroides phylogenetic consensus tree

Figure 18

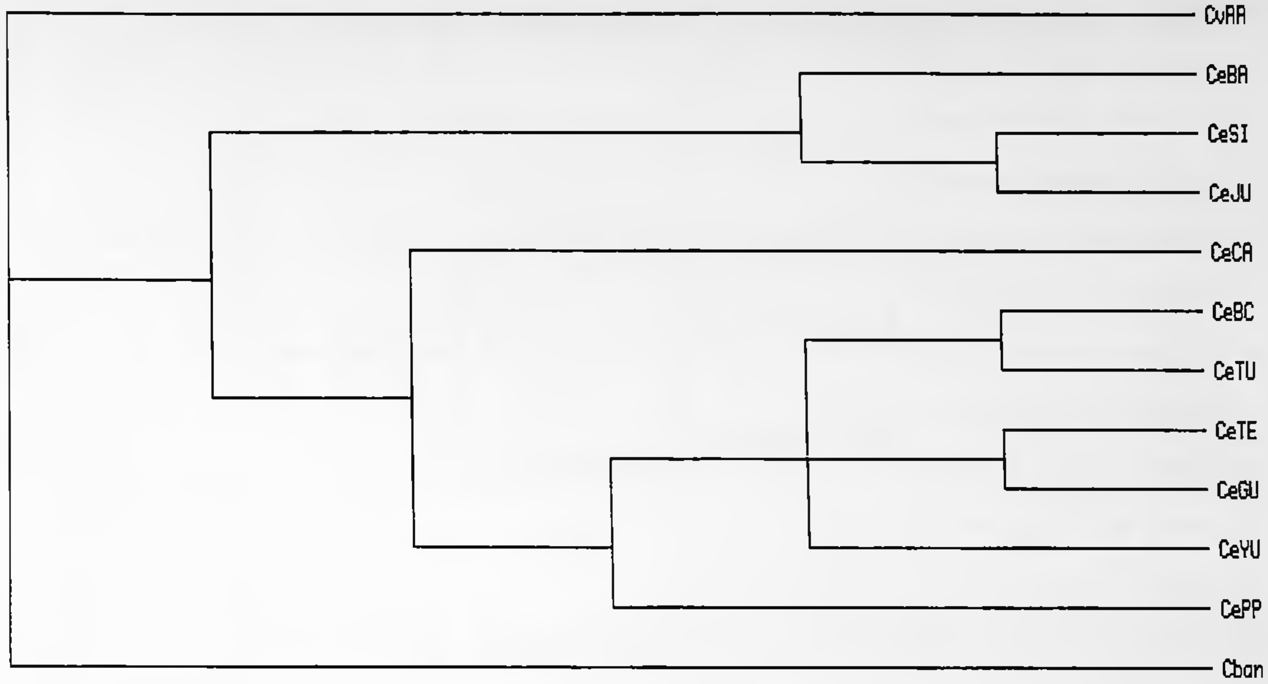


Figure 19

Centruroides phylogenetic bootstrap analysis

Figure 19

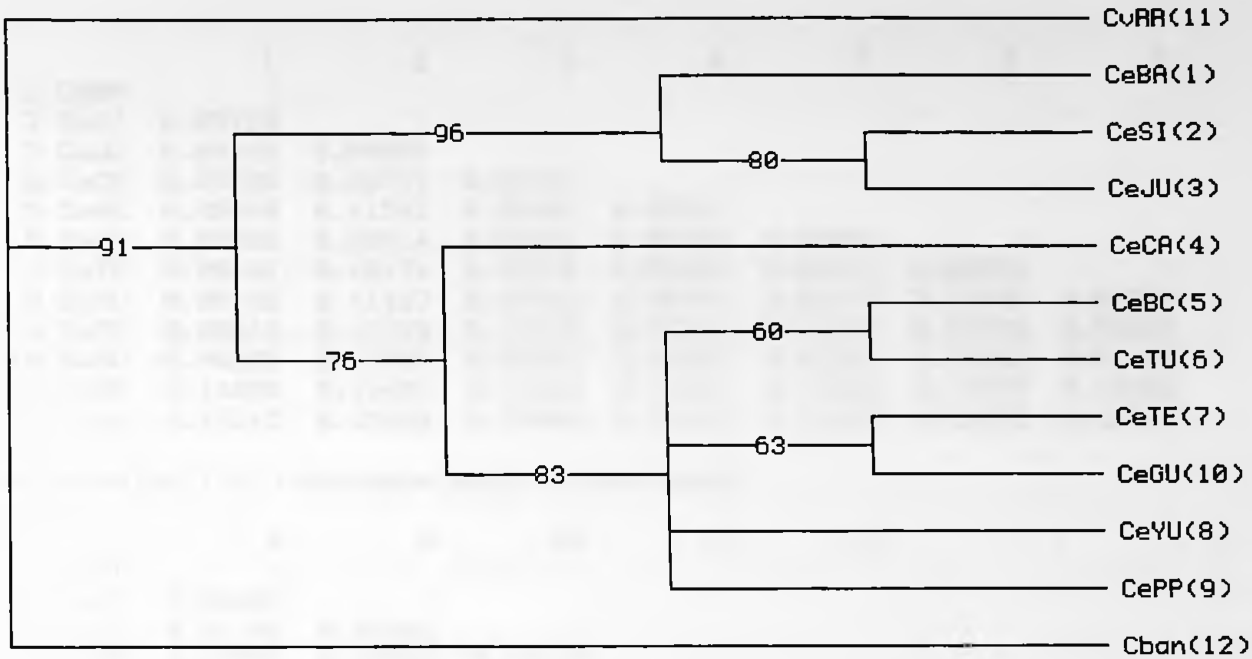

CvAA(11)
CeBA(1)
96
CeSI(2)
80
CeJU(3)
91
CeCA(4)
CeBC(5)
60
76
CeTU(6)
CeTE(7)
63
83
CeGU(10)
CeYU(8)
CePP(9)
Cban(12)

Table 5
Distance Matrix *Centruroides*

	1	2	3	4	5	6	7
1 CeBA	-						
2 CeSI	0.05719	-					
3 CeJU	0.05416	0.04066	-				
4 CeCA	0.09590	0.09747	0.07327	-			
5 CeBC	0.08968	0.11562	0.09496	0.05919	-		
6 CeTU	0.08860	0.09014	0.08021	0.06220	0.02486	-	
7 CeTE	0.09101	0.10174	0.07678	0.05048	0.02013	0.02323	-
8 CeYU	0.08146	0.11127	0.09181	0.06196	0.02179	0.03040	0.01997
9 CePP	0.09513	0.12219	0.10270	0.05349	0.02453	0.03320	0.02016
10 CeGU	0.09226	0.11936	0.09992	0.06466	0.02990	0.03862	0.01138
11 CvAA	0.14486	0.15462	0.14433	0.14290	0.14483	0.13957	0.13266
12 Cban	0.15315	0.15834	0.15004	0.14032	0.14190	0.14788	0.13842

Uncorrected ("p") distance matrix (continued)

	8	9	10	11	12
8 CeYU	-				
9 CePP	0.02432	-			
10 CeGU	0.02156	0.03241	-		
11 CvAR	0.13557	0.13842	0.14110	-	
12 Cban	0.13884	0.13386	0.14178	0.15478	-

4. Discussion

4.1 *Hadrurus*

The dark colored *H. pinteri* (Hpin), from volcanic areas of the Baja California peninsula, varies greatly from the other *Hadrurus* with a genetic distance 9.5% to 12.3%. *H. concolor*, another Baja peninsula group, likewise is distant from all other *Hadrurus* species with distances ranging from 6.2% to 7.3% (Table 4). These two species show distance values that have been observed in other "good" morphological species. *H. obscurus* (Hob1) and *H. spadix* (Hspa) form a well supported terminal clade but only have a genetic distance of 2.4%(matrix). The original distinction between the species was based on coloration (Williams 1970) and trichobothrial patterns between the two species were shown to be similar (Soleglad 1976). Fieldwork by Dr. Victor Fet and associates have suggested intermediates *between H. obscurus* and *H. spadix*. Continued research will be necessary to determnine if these are indeed two different species.

Some of the DNA sequences *from H. arozinensis arizonensis* (Har 2 and Har3), from Yuma and Dateland,Arizona, were identical to those *of H. a. pallidus* (Har4 and Har5), from Anza Borrego Desert State Park, and both were very close to the sequence for Har 7 a dark nominotypic population from Indio. *While H. a. arizonensis* sequence from Death Valley (Har1) was slightly distant at a genetic distance of 1.1% and was identical to a specimen of *H. a. arizonensis* from Soda Springs, California (Har8). The differentiation as stated earlier was vague between these subspecies and was based mainly on coloration further studies will be necessary to tell if these are merely color variants or if the two subspecies are valid.

4.2 *Anuroctonus*

The genus *Anuroctonus* currently contains only one accepted species (however, at least two others have been suggested). The current existing species AnUT from Garrison, Utah and AnMO from Monterey County, CA form a divergence with a bootstrap value of 93% indicating a well supported terminal clade but shows a genetic distance of 4.5% which is not a very high level of divergence for morphospecies. These two specimens vary from 8.2% to 9.8% in genetic distance from the other specimens of *Anuroctonus*. However, the second clade of the *Anuroctonus* is highly supported with a bootstrap value of 99%. This and the genetic distances indicate possible different species at least represented by these two major clades.

All specimens within the second clade likewise fall within a 5 % genetic difference from one another. The suggestion of AnMO (suggested *A. pococki*) as a different species may not be found on the grounds it groups with the recognized species *A. phaiodactylus* and has a genetic distance of about 4.5%. However the grouping and genetic distance of the proposed *A. bajae* seems to be well supported. The high divergence, quite unexpected from morphology (M. Soleglad, pers. comm.) could be due to sessile way of life of these animals, highly bound to substrate in which they burrow (Williams 1980), unlike three other genera in our study. Further analysis will be necessary to determine the exact number and boundaries of species, and proper classification of this genus. However, already this study indicates at least two species within the genus *Anuroctonus* will be necessary, and the old opinion on the monotypy of this genus is wrong.

4.3 *Euscorpius*

Molecular analysis reveals a deep divergence between the Karawanken Form (K-Form) and the Typical Form (T-Form). The K-Form was found to be morphologically identical to *E.*

mingrelicus gamma and *E. m. histrorum*, but is not identical to the nominotypic *E. mingrelicus* (Kessler 1874) from Caucasus. Therefore, it is suggested that this represents a different species, separate from both *E. germanus* and *E. mingrelicus*, based on morphological as well as molecular data.

DNA data does not agree with hybrid suggestions made by Bonacina (1980), or Kinzelbach (1975). Analysis showed two deep well-supported clades within the traditional species *E. germanus.* These clades correspond to the two described subspecies *E. g. germanus* and *E. g. alpha*, though the geographical ranges of these clades does not conform to that reported by morphological analysis of subspecies (Bonacina 1980, Caporiacco 1950).

4.4 *Centruroides*

C. vittatus (CvAR) from Russelville, Arkansas and *C. bani* (Cban) from the Dominican Republic show a large genetic distance from other specimens (populations of of *C. exilicauda*) with a genetic distance of 13.2% to 15.8. This confirms that morphospecies *C. vittatus* and *C. bani* are well-supported species. The divergence of the Baja California clade of *C. exilicauda* (CeBA, CeSI, and CeJU), with genetic distances of 4.0% to 5.4%, from the other clade containing specimens from Sonora, Arizona, and one population from Cabo San Lucas, Baja California (CeCA, CeBC, CeTu, CeTE, CeGU, CeYU, and CePP) is well supported with a bootstrap value of 90%. This suggests more than one species within *C. exilicauda* and the genetic distances of 7.3% to 15.4% seem to further support this idea. The type locality of this species is Cabo San Lucas. The Cabo San Lucas clade is divergent with a bootstrap value of 73% showing a moderately supported divergence; *C. exilicauda* specimen CeCA shows an intermediate distance between the other Baja specimens and the Arizona specimens. Further studies are necessary to exactly classify the discovered two clades within *C. exilicauda*.

5. Conclusion

Using mtDNA techniques has shown a possible method of clarifying problems, which arise from traditional morphological analysis. This two-year study demonstrated that coloration differences as seen in *Hadrurus* are not necessarily indicative of different subspecies and probably represent phenotypic characteristics within the same species. Through the use of molecular techniques the classification of the genus *Anuroctonus* may be promoted as it is lacking sufficient prior taxonomic division. Molecular analysis has demonstrated differences among the genus *Euscorpius* which were not apparent through morphological analysis, and showing the genetic diversity of the forms of Euscorpius separated by the Adige River in the Alps as well as the genetic diversity found within the species *E. germanus*. Lastly, the current classification of the genus *Centruroides* may not be sufficient considering the amount of genetic diversity present within the species *C. exilicauda*.

6. Literature Cited

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

Birula, A. A. (Byalynitsky-Birula, A.A.) 1917. Arachnoidea Arthrogastra Caucasica. Pars I. Scorpiones. Mémoires du Musée du Caucase, Tiflis: Imprimerie de la Chancellerie du Comité pour la Transcaucasie, A(5), 253 pp. (in Russian). English translation: 1964. Arthrogastric Arachnids of Caucasia. 1. Scorpions. Jerusalem: Israel Program for Scientific Translations, 170 pp.

Bonacina, A. 1980. Sistematica specifica e sottospecifica del complesso "*Euscorpius germanus*" (Scorpiones, Chactidae). Riv. Mus. civ. Sci. nat. „E. Caffi", 2: 47- 100.

Caporiacco, L. di. 1950. Le specie e sottospecie del genre „*Euscorpius*" viventi in Italia ed in alcune zone confinanti. Atti Accad. naz. Lincei Memorie, (ser. 8) 2: 159-230.

Capra, F. 1939. L'Euscorpius germanus(C.L. Koch) in Italia (Arachnida, Scorpiones) Memorie della Societa Entomologica Italiana. 18(2): 199-213.

Clary, D.O. and D.r. Wolstenholme. 1985. The mitochondrial DNA molecule of *Drosophila yakuba*: Nucleotide sequence, gene organization and genetic code. Journal of Molecular Evolution 22: 252-271.

Crother, B.I. 1992. Genetic characters, species concepts, and conservation biology. *Conservation Biology* 6: 314.

Crucitti, P. 1993. Distribution and diversity of Italian scorpions. Redia, 76(2): 281-300.

Esper, E. J. C. 1781. *De varietatibus specierum in naturae productis*. Erlangen.

Fet, V. & M. E. Braunwalder. 1997. On the true authorship and taxonomic history of *Euscorpius germanus* (C. L. Koch, 1837) (nec C. L. Koch, 1836; nec Schaeffer, 1766)

(Scorpiones: Chactidae). Bull. Br. Arachnol. Soc., 10(8): 308-310.

Fet, V. & G. Lowe. 2000. Family Buthidae. Pp. 54-286 *In*: Fet, V. et al. Catalog of the Scorpions of the World (1758-1998). *New York Entomological Society, New York*, 690 pp.

Fet, V., W.D. Sissom, G. Lowe & M.E. Braunwalder. 2000. *Catalog of the Scorpions of the World (1758-1998)*. New York Entomological Society, New York, 690 pp.

Fet, V. and W.D. Sissom. 2000. Family Euscorpiidae. In Fet, V. et al. Catalog of Scorpions of the World (1758-1998). New York Entomological Society, New York.

Francke, O. F. and M. E. Soleglad. 1981. The family Iuridae Thorell (Arachnida, Scorpiones). Journal of Arachnology 9: 233-258.

Gaines, M.S., L.S. Mcleneghan, Jr. and R.K. Rose. 1978. Temporal patterns of allozymic variation in fluctuating populations of *Microtus ochrogaster*. Evolution 32: 723-39.

Gantenbein, B., L. Büchi, M. E. Braunwalder & A. Scholl 1998. The genetic population structure of *Euscorpius germanus* (C. L. Koch) (Scorpiones: Chactidae) in Switzerland. Pp. 33-40. In: Selden, P. A. (ed.), Proceedings of the 17th European Colloquium of Arachnology, Edinburgh 1997, 14-18 July, 1997.

Gantenbein, B., V. Fet, C. Largiader & A. Scholl. 1999a. First DNA phylogeny of the genus *Euscorpius* Thorell 1876 (Scorpiones, Euscorpiidae) and its bearing on the taxonomy and biogeography of this genus. *Biogeographica* (Paris), 75(3): 59-72.

Hadzi, J. 1929. Skorpije Schmidtove zbirke. *Euscorpius italicus polytrichus* n. ssp. i ostale nove rase (Die Skorpione der Schmidt'schen Sammlung: *Euscorpius italicus polytrichus* n. ssp und andere neue Rassen). Glasnik.

Hillis, D.M., C. Moritz, and B.K. Mable. 1996. Molecualr Systematics. Sinauer Associates, Inc., Mass., U.S.

Kinzelbach, R. 1975. Die Skorpione der Ägäis. Beiträge zur Systematik, Phylogenie und Biogeographie. *Zoologische Jahrbücher, Abtheilung für Systematik* 102(1): 12-50.

Linnaeus, C. 1758.Systema Naturae. Tenth Edition

Lewin, B. 1994. Genes V. Oxford University Press. New York

Marx, G. 1890 (Additional remarks regarding the types of the Scorpionidae desribed by Wood). Notes presented at the Meeting. Proceeedings of the Entomological Society of Washington, 1(1884-1887): 90-91.

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

Moritz, C., T.E. Dowling, and W.M. Brown. 1987. Evolution of Animal mitochondrial DNA: relevance for population biology and systematics. Ann. Rev. Ecol. Syst.: 18. 269-92.

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

Scherabon, B. 1987. Die Skorpione Österreichs in vergleichender Sicht unter besonderer Berücksichtigung Kärntens. *Carinthia II. / Naturwiss. Beiträge zur Heimatkunde Kärntens / Mitteilungen des Naturwissenschaftlichen Vereins für Kärnten. Klagenfurt, Verlag des Naturwissenschaftlichen Vereins für Kärnten* 45: 78-158.

Sissom, W. D. 1990. Systematics, biogeography and paleontology. *In* G. A. Polis (ed.), Biology of scorpions, pp. 64-160. Stanford, California: Stanford University Press.

Sissom, W. D. & V. Fet. 2000. Family Iuridae. Pp. 409-420 *In*: Fet, V. et al. Catalog of the Scorpions of the World (1758-1998). *New York Entomological Society, New York*, 690 pp.

Soleglad, M. E. 1976. The taxonomy of the genus Hadrurus based on chela trichobothria (Scorpionida: Vejovidae). Journal of Arachnology 3: 113-134.

Stahnke, H. L. 1969. A review of Hadrurus scorpions (Vejovidae). Entomological News 80: 57-65.

Stahnke, H. L. 1972. A key to the genera of Buthidae (Scorpionida). Entomological News, 83(5): 121-133.

Stahnke, H. L. 1974. Revision and keys to the higher categories of Vejovidae. Journal of Arachnology, 1(2). 107-141.

Stockwell, S. A. 1992. Systematic observations on North American Scorpionida with a key and checklist of the families and genera. Journal of Medical Entomology, 29(3): 407-422.

Vachon, M. 1962. Remarques sur l'utilisation en systématique, des soies sensorielles (Trichobothries) chez les scorpions du genre Euscorpius Thorell (Chactidae). Bull. Mus. Hist. nat., Paris, série 2, 34 (5): 345-354.

Vachon, M. 1981. Remarques sur la classification sous-spécifique des espèces appartenant au genre Euscorpius Thorell, 1876 (Scorpionida, Chactidae). Atti Soc. Tosc. Sci. Nat., Mem., ser. B, 88, suppl. (C.R. Vième Coll. Arachnol. express. franç., Modena-Pisa). 193-203.

Valle, A., M. T. Berizzi, M. Bonino, R. Gorio, E. Gimmilaro-Negri & A. Percassi. 1971. Le popolazioni italiane di *Euscorpius germanus* (C. L. Koch) (Scorpiones, Chactidae). Atti dell'VIII Congresso Nazionale Italiano di Entomologia (Firenze, 4-7 Settembre 1969): 93-96.

Williams, S. C. 1970. A systematic revision of the giant hairy scorpion genus *Hadrurus.* Occasional Papers of the California Academy of Sciences. 87: 1-62.

Williams, S. C. 1980. Scorpions of Baja California, Mexico, and Adjacent islands. Occasional Papers of the California Academy of Sciences. 135: 1-127.

Appendix 1 (Sequences)

Hadrurus

```
Hpin             TTCCTTGGTAAACTATTGCGCCAACCAGGACTTTTAATCCAACAT-CGAG
Hcon             TTCCTTAGTAAGCTATTGCGCCAACTAGGACTTTTAATCCAACAT-CGAG
Hhir             NNNNNNNGTAAGCTATTGNGCCANCTAGGACTTTTAATCCAACATGGGAG
Hspa             TTCCTTAGTAAGCTATTGCGCCAACTAGGACTTTTAATCCAACAT-CGAG
Hob1             TTCCTTAGTAAGCTATTGCGCCAACTAGGACTTTTAATCCAACAT-CGTG
Hob2             TTCCTTAGTAAGCTATTGCGCCAACTAGGACTTTTAATCCAACAT-CNAG
Hob3             TTCCTTAGTAAGCTATTGCGCCAACTAGGACTTTTAATCCAACAT-CGAG
Hob4             TTTCTTAGTAAGCTACTGCGCCAACTAGGACTTTTAATCCAACAT-CGAG
Har1             TTCCTTAGTAAGCTATTGCGCCAACTAGGACTTTTAATCCAACAT-CGAG
Har2             TTCCTTAGTAAGCTATTGCGCCAACTAGGACTTTTAATCCAACAT-CGAG
Har3             TTCCTTAGTAAGCTATTGCGCCAACTAGGACTTTTAATCCAACAT-CGNG
Har4             TTCCTTAGTAAGCTATTGCGCCAACTAGGACTTTTAATCCAACAT-CGNG
Har5             TTCCTTAGTAAGCTATTGCGCCAACTAGGACTTTTAATCCAACAT-CNAG
Har6             TTCCTTAGTAAGCTATTGCGCCAACTAGGACTTTTAATCCAACAT-CNAG
Har7             TTCCTTAGTAAGCTATTGCGCCAACTAGGACTTTTAATCCAACAT-CNAG
Har8             TTCCTTAGTAAGCTATTGCGCCAACTAGGACTTTTAATCCAACAT-NNAG
Har9             TTCCTTAGTAAGCTATTGCGCCAACTAGGACTTTTAATCCAACAT-CGAG
Ha10             TTCCTTAGTAAGCTATTGCGCCAACTAGGACTTTTAATCCAACAT-CGAG
Ha11             TTCCTTAGTAAGCTATTGCGCCAACTAGGACTTTTAATCCAACAT-CGAG

Hpin   GTCACAAAATTTTTTAATGATAAGAACTCTTAAAAAAAATTATGCTGTTATCCCTACAGT
Hcon   GTCACAAAAATTTTTGATGATAAGAACTCTTAAAAAAAATTATGCTGTTATCCCTACAGT
Hhir   GTCACAAAATTTGTTGATGANNAGAACTCTTAAAAAAAATNATGCTGTTATNCCTACAGC
Hspa   GTCACAAAATTTTTTGATGATAAGAACTCTTAAAAAAAATTATGCTGTTATCCCTACAGT
Hob1   GTCACAAAATTTTTTGATGATAAGAACTCTTAAAAAAAATTATGCTGTTATCCCTACAGT
Hob2   GTCACAAAATTTTTTGATGATAAGAACTCTTAAAAAAAATTATGCTGTTATCCCTACAGT
Hob3   GTCACAAAATTTTTTGATGATAAGAACTCTTAAAAAAAATTATGCTGTTATCCCTACAGT
Hob4   GTCACAAAATTTTTTGATGATAAGAACTCTTAAAAAAAATTATGCTGTTATCCCTACAGT
Har1   GTCACAAAATTCTTTGATGATAAGAACTCTTAAAAGAAATTATGCTGTTATCCCTACAGT
Har2   GTCACAAAATTCTTTGATGATAAGAACTCTTAAAAGAAATTATGCTGTTATCCCTACAGT
Har3   GTCACAAAATTCTTTGATGATAAGAACTCTTAAAAGAAATTATGCTGTTATCCCTACAGT
Har4   GTCACAAAATTCTTTGATGATAAGAACTCTTAAAAGAAATTATGCTGTTATCCCTACAGT
Har5   GTCACAAAATTCTTTGATGATAAGAACTCTTAAAAGAAATTATGCTGTTATCCCTACAGT
Har6   GTCACAAAATTCTTTGATGATAAGAACTCTTAAAAGAAATTATGCTGTTATCCCTACAGT
Har7   GTCACAAAATTCTTTGATGATAAGAACTCTTAAAAGAAATTATGCTGTTATCCCTACAGT
Har8   GTCACAAAATTCTTTGATGATAAGAACTCTTAAAAGAAATTATGCTGTTATCCCTACAGT
Har9   GTCACAAAATTCTTTGATGATAAGAACTCTTAAAAGAAATTATGCTGTTATCCCTACAGT
Ha10   GTCACAAAATTCTTTGATGATAAGAACTCTTAAAAGAAATTATGCTGTTATCCCTACAGT
Ha11   GTCACAAAATTCTTTGATGATAAGAACTCTTAAAAGAAATTATGCTGTTATCCCTACAGT

Hpin   AACTTATTCCTCCTTAAAAATTTCTCGATTCTACAAGACTGGTAAGCCACAATATCAAAG
Hcon   AACTTATTCCTTATTAAAAACTCCTCGATTTTACAAGATTAG-AAATCACAATATAAAAA
Hhir   AACTTATTACTCTTTTAAAATTNNTCGATTTTACAAGNCTAG-AAATCACAATATATANA
Hspa   AACTTATTCCTCCTTAAAAACTTTTCGATTTTACAAGATAAG-AACTCACAATACAAAAA
Hob1   AACTTATTCCTCTTTAAAAACTTTTCGATTTTACAAGATGAG-AACTCACAATACAAAAA
Hob2   AACTTATTCCTCTTTAAAAACTTTTCGATTTTACAAGATGAG-AACTCACAATACAAAAA
Hob3   AACTTATTCCTCTTTAAAAACTTTTCGATTTTACAAGATGAG-AACTCACAATACAAAAA
Hob4   AACTTATTCCTCTTTAAAAACTTTTCGATTTTACAAGATGAG-AACTCACAATACAAAAA
Har1   AACTTATTCCTCATTTAAAATTTTTCGATTTTACAAGACTAG-AACTCACAATATAAAAA
```

```
Har2  AACTTATTCCTCATTTAAAATTTTTCGATTTTACAAGACTAG-AACTCACAATATAAAAA
Har3  AACTTATTCCTCATTTAAAATTTTTCGATTTTACAAGACTAG-AACTCACAATATAAAAA
Har4  AACTTATTCCTCATTTAAAATTTTTCGATTTTACAAGACTAG-AACTCACAATATAAAAA
Har5  AACTTATTCCTCATTTAAAATTTTTCGATTTTACAAGACTAG-AACTCACAATATAAAAA
Har6  AACTTATTCCTCATTTAAAATTTTTCGATCTTACAAGACTAG-AACTCACAATATAAAAA
Har7  AACTTATTCCTCATTTAAAATTTTTCGATTTTACAAGATTAG-AACTCACAATATAAAAA
Har8  AACTTATTCCTCATTTAAAATTTTTCGATTTTACAAGACTAG-AACTCACAATATAAAAA
Har9  AACTTATTCCTCATTTAAAATTTTTCGATTTTACAAGACTAG-AACTCACAATATAAAAA
Ha10  AACTTATTCCTTATTTAAAATTTTTCGATTTTACAAGACTAG-AACTCACAATATAAAAA
Ha11  AACTTATTCCTGATTTAAAATTTTTCGNTTTTACAAGACTAG-AACTCACAATATAAAAA

Hpin  AAAAATTAATTTCCCAACTGCCCCAGTTAAACAATTTAGTTTCTGATTTGAAGACAGGA-
Hcon  TGATATTAACTCATTAACTGCCCCAGTTAAACAATTTCACCCTTGATTTGAAAACAAAAG
Hhir  GGATNTTAATTCNTCANCNGCCCCAGATAAACAATTTNATTTNTGACTTGAAAACAANAA
Hspa  TGATATTAACTCATTAACTGCCCCAGTTAAACAATTTAATTTCTGATTTGAAAACAAAAA
Hob1  TGATATTAATTCACTAACTGCCCCAGTTAAACAATTTAATTTCTGATTTGAAAACAAAAA
Hob2  TGATATTAATTCACTAACTGCCCCAGTTAAACAATTTAATTTCTGATTTGAAAACAAAAA
Hob3  TGATATTAATTCACTAACTGCCCCAGTTAAACAATTTAATTTCTGATTTGAAAACAAAAA
Hob4  TGATATTAATTCACTAACTGCCCCAGTTAAACAATTTAATTTCTGATTTGAAAACAAAAA
Har1  TGATATTAACTCATCAACTGCCCCAGTTAAACAATCTTATTCTTGATTTGAAAACAAAAA
Har2  TGACATTAACTCATCAACTGCCCCAGTTAAACAATCTTATTCTTGATTTGAAAACAAAAA
Har3  TGACATTAACTCATCAACTGCCCCAGTTAAACAATCTTATTCTTGATTTGAAAACAAAAA
Har4  TGACATTAACTCATCAACTGCCCCAGTTAAACAATCTTATTCTTGATTTGAAAACAAAAA
Har5  TGACATTAACTCATCAACTGCCCCAGTTAAACAATCTTATTCTTGATTTGAAAACAAAAA
Har6  TGACATTAACTCATCAACTGCCCCAGTTAAACAATCTTATTCTTGATTTGAAAACAAAAA
Har7  TGACATTAACTCATCAACTGCCCCAGTTAAACAATCTTATTCTTGATTTGAAAACAAAAA
Har8  TGATATTAACTCATCAACTGCCCCAGTTAAACAATCTTATTCTTGATTTGAAAACAAAAA
Har9  TGACATTAACTCATCAACTGCCCCAGTTAAACAATCTTATTCTTGATNTGAAAACAAAAA
Ha10  TGACATTAACTCATCAACTGCCCCAGTTAAACAATCTTATTCTTGATTTGAAAACAAAAA
Ha11  TGACATTAACTCATNAACTGGCCCAGNTAAACAATCTTATTCTTGATTTGGAAACAAAAA

Hpin  ACTAATTGTAAAGCTTGATAGGGTCTTCTCGTCCCCAAAAAAAATTTAGCTTTTTTACT
Hcon  ATTAATTGTAAAGCTTGATAGGGTCTTCTCGTCCCTAAACAAAATTTAGCTTTTTTACT
Hhir  ATNAATTGTAANGCTTGATAGGGTNTTCTNGNCCCCAAACAAAATTTAGCTTTTTTACT
Hspa  ATCAATTGTGAAGCTTGATAGGGTCTTCTCGTCCCCAAATAAAATTTAGCTTTTTTACT
Hob1  ATTAATTGTTAAGCTTGATAGGGTCTTCTCGTCCCCAAACAAAATTTAGCTTTTTTACT
Hob2  ATTAATTGTTAAGCTTGATAGGGTCTTCTCGTCCCCAAACAAAATTTAGCTTTTTTACT
Hob3  ATTAATTGTTAAGCTTGATAGGGTCTTCTCGTCCCCAAACAAAATTTAGCTTTTTTACT
Hob4  ATTAATTGTTAAGCTTGATAGGGTCTTCTCGTCCCCAAACAAAATTTAGCTTTTTTACT
Har1  ATCAATTGTAAAGCTTGATAGGGTCTTCTCGTCCCCAAACAAAATCTTAGCCTTTTTACT
Har2  ATCAATTGTAAAGCTTGATAGGGTCTTCTCGTCCCCAAACAAAATTTAGCCTTTTTACT
Har3  ATCAATTGTAAAGCTTGATAGGGTCTTCTCGTCCCCAAACAAAATTTAGCCTTTTTACT
Har4  ATCAATTGTAAAGCTTGATAGGGTCTTCTCGTCCCCAAACAAAATTTAGCCTTTTTACT
Har5  ATCAATTGTAAAGCTTGATAGGGTCTTCTCGTCCCCAAACAAAATTTAGCCTTTTTACT
Har6  ATCAATTGTAAAGCTTGATAGGGTCTTCTCGTCCCCAAACAAAATTTAGCCTTTTTACT
Har7  ATCAATTGTAAAGCTTGATAGGGTCTTCTCGTCCCCAAACAAAATTTAGCCTTTTTACT
Har8  ATCAATTGTAAAGCTTGATAGGGTCTTCTCGTCCCCAAACAAAATCTTAGCCTTTTTACT
Har9  ATCAATTGTAAAGCTTGATAGGGTCTTCTCGTCCCCAAACAAAATTTAGCCTTTTTACT
Ha10  ATCAAATGTAAAGCTTGATAGGGTCTTCTCGTCCCCAAACAAAATTTAGCCTTTTTACT
Ha11  ATCAATTGTAAAGCTTGATAGGGTCTTCTCGGCCCCAAACNANATTTAGCCTTTTTACT

Hpin  AAATTATAAAATTCAAAATATTTAAATAAGACAAAAACATTTAGTGAGACCCTTCATTCC
Hcon  AAATGATAAAATTCAAAATATTTAAATAAGACAGAAATATTCAGTAAAACCCTTCATTCC
Hhir  AAATGATAAAATTCAAAATATTTAAATAAGACAGAAATATTCAGTGAAACCCTTCATTCC
Hspa  AAATGATAAAATTCAAAATATTTAAGTAAGACAAAAATATTTAGTGAAACCCTTCATTCC
Hob1  AAATGATAAAATTCAAAATATTTAAGTAAGACAAAAATATTTAGTGAAACCCTTCATTCC
Hob2  AAATGATAAAATTCAAAATATTTAAGTAAGACAAAAATATTTAGTGAAACCCTTCATTCC
```

```
Hob3  AAATGATAAAATTCAAAATATTTAAGTAAGACAAAAATATTTAGTGAAACCCTTCATTCC
Hob4  AAATGATAAAATTCAAAATATTTAAGTAAGACAAAAATATTTAGTGAAACCCTTCATTCC
Har1  AAACAATAAAATTCAAAATATTTAAGTAAGACAGAAATATTTAGTGAAACCCTTCATTCC
Har2  AAACAATAAAATTCAAAATGTTTAAATAAGACAGAAATATTTAGTGAAACCCTTCATTCC
Har3  AAACAATAAAATTCAAAATGTTTAAATAAGACAGAAATATTTAGTGAAACCCTTCATTCC
Har4  AAACAATAAAATTCAAAATGTTTAAATAAGACAGAAATATTTAGTGAAACCCTTCATTCC
Har5  AAACAATAAAATTCAAAATGTTTAAATAAGACAGAAATATTTAGTGAAACCCTTCATTCC
Har6  AAACAATAAAATTCAAAATGTTTAAATAAGACAGAAATATTTAGTGAAACCCTTCATTCC
Har7  AAACAATAAAATTCAAAATGTTTAAATAAGACAGAAATATTTAGTGAAACCCTTCATTCC
Har8  AAACAATAAAATTCAAAATATTTAAGTAAGACAGAAATATTTAGTGAAACCCTTCATTCC
Har9  AAACAATAAAATTCAAAATGTTTAAATAAGACAGAAATATTTAGTGAAACCCTTCATTCC
Ha10  AAACAATAAAATTCAAAATGTTTAAATAAGACAGAAATATTTAGTGAAACCCTTCATTCC
Ha11  NAACAATNAAATTCANAATGTTTAAATAAGACAGANATATTTAGTGAAACCCTTCATTCC
```

```
Hpin            AGCCCTAAATTAAAAGACTA
Hcon            AGTCCCAAATTAAGAGACTA
Hhir            AGTCCCAAATT---------
Hspa            AGTCCCAAATTAAAAGACTA
Hob1            AGTCCTAAATTAAAAGACTA
Hob2            AGTCCTAAATTAAAAGACTA
Hob3            AGTCCTAAATTAAAAGACTA
Hob4            AGTCCTAAATTAAAAGACTA
Har1            AGTCCTAAATTAAAAGACTA
Har2            AGTCCTAAATTAAAAGACTA
Har3            AGTCCTAAATTAAAAGACTA
Har4            AGTCCTAAATTAAAAGACTA
Har5            AGTCCTAAATTAAAAGACTA
Har6            AGTCCTAAATTAAAAGACTA
Har7            AGTCCTAAATTAAAAGACTA
Har8            AGTCCTAAATTAAAAGACTA
Har9            AGTCCTAAATTAAAAGACTA
Ha10            AGTCCTAAATTAAAAGACTA
Ha11            AGNCCTAAATTAANAGACTA
```

Anuroctonus

```
Hari T?GAACAGA-TTCCTTAGT-AAGCTATTGCGC
AnJT ?????????????CT?ATT-?TACTTTTGCAT
AnRV T?TAACA?ACTTCCTTATT-TTACTTTTGCAT
AnMo TCGAACAGACTTCCTTGTT-TTACTTTTACAC
AnUt NNNNNNNNNNNNNNNNNNN-NNNNCTNNNGCC
AnCH AGTCCCCCGACTTCTTATT-TTACTTTTGCAT
AnPM AAACCCCNGACTTTTTATT-TTACTTNN-CAT
```

```
Hari CAACTAGGACTTTTAATCCAACATC-GAGGTCACAAAATTCTTTGATGATA
AnJT AAATTAGGGATCTTAATCCAACATC-GAGGTCACAAACTTTTTTGATGATA
AnRV AAATTAGGAATCTTAATCCAACATC-GAGGTCACAAACTTTTTTGATGATA
AnMo AAGTTAGGTTTCTTAATCCAACATC-GAGGTCACAAACTTTTTTGATGATA
AnUt ANACTAGGNTTCTTAATCCAACATC-GAGGTCACAAACTTTTTTGATGATA
AnCH AAATTAGGAATCTTAATCCAACATC-GAGGTCACAAACTTTTTTGATGATA
AnPM AAATTAGGGATCTTAATCCAACATC-GAGGTCACAAACTTTTTTGATGATA
```

```
Hari AGAACTCTTAAAAGAAATTATGCTGTTATCCCTACAGTAACTTATTCCTC
AnJT AGGACTCTTTAAAAAAATTATGCTGTTATCCCTACAGTAACTTGTTCTTT
AnRV AGAACTCTTTAAAAAAATTATGCTGTTATCCCTACAGTAACTTGTT-TTT
AnMo AGAACTCTTAAAAAAAAATTATGCTGTTATCCCTACAGTAACTTGTTCTTT
AnUt AGGACTCTTAAAAAAAAATTATGCTGTTATCCCTACAGTAACTTGTTCTTT
AnCH AGAACTCTTTAAAAAAATTATGCTGTTATCCCTACAGTAACTTGTTTTTT
AnPM AGGACTCTTTAAAAAAATTATGCTGTTATCCCTACAGTAACTTGTTATTT


Hari ATTTAAAATTTTTCGATTTTACAAGACTAGAAC-TCACAATATAAAAAT
AnJT TTTAGGAATTTCTTGGTTGCTCTAGAACGATAATTCATAATATTAAAAC
AnRV TTTAGGAATTTCTTGATTTTTCTAGAATGATAATTCACGATATTAAAAC
AnMo TTTAGGAATTTTTTGATTTTTCTAGAAGAATGATTCACAATGTAAAAAC
AnUt TTTAGGAATTTTTTGATTTTTCTAGAAGTATAGTTCACAATGTAAAAAC
AnCH TTTAGGAATTTCTTGATTTTTCTAGAATGATAATTCACGATATTAAAAC
AnPM TTTAAGAATTTCTTGATTTCTCTAGAATGATAATTCACAATATTAAAAC


Hari GA--TATTAACTCATCAACTGCCCCAGTTAAACAATCTTATTCTTGATTTGA
AnJT AA--AAACTTTTTGTTTGCTGCCCCAGCAGAATAATTTTTTTATTAGGTTG-
AnRV AA--AAAATTTTTGTTTGCTGCCCCAGCA?AATAACTTTTTTATTAGGTTG-
AnMo AA--AAAGTTTTTGTTTGCTGCCCCAGCAAAATAATTTTTTTATTAGATTG-
AnUt AA--AAACTTTTTGTCTGCTGCCCCAGCAAAATAATTTTTTTACTAGGTTG-
AnCH AA--AAAATTTTTGTTTGCTGCCCCAGCAGAATAATTTTTTTATTAGGTTG-
AnPM AA--AAAGTTTTTGTTTGCTGCCCCAGCAAAATAATTTTTTTATTAGATTG-



Hari AAACAAAAAATCAA--TTGTAAAGCTTGATAGGGTCTTCTCGTCCCCAAACA
AnJT TAGTTAATTTAAAA--TTATGAAGCTTGATAGGGTCTTCTTGTCTTTAAAGT
AnRV TAGTTAATTAAAAA--TTATGAAGCTTGATAGGGTCTTCTTGTCTTTAAAGT
AnMo TAGTTAATTTAAAA--TTATAAAGCTTGATAGGGTCTTCTTGTCTTTAAAGA
AnUt TAGTTAGTTTAAAA--TTATAAAGCTTGATAGGGTCTTCTTGTCTTTAAAGT
AnCH TAGTTAATTAAAAA--TTATGAAGCTTGATAGGGTCTTCTTGTCTTTAAAGT
AnPM TAGTTAATTAGAAA--TTATAAAGCTTGATAGGGTCTTCTTGTCTTTAAAGT


Hari -AAATCTTAGCCTTTTTACTAAACAATAAAATTCAAAATATTTAAGTAAGA
AnJT -GAATTTTAGCTTTTTTACTAAAATATAAATTTCAAAATATTATTATGAGA
AnRV -GAATTTTAGCTTTTTTACTAAAATATAAATTTCAAAATATTACTATGAGA
AnMo -GTATTTTAGATTTTTTACTAAAATATAAATTTCAAAATATTATCATGAGA
AnUt -GTATTTTAGCTTTTTTACTAAAATATAAATTTCAAAATATTATCATGAGA
AnCH -GAATTTTAGCTTTTTTACTAAAATATAAATTTCAAAATATTACTATGAGA
AnPM -GAATTTTAGCTTTTTTACTAAAATATAAATTTCAAAATATTATTATGAGA

Hari CA-GAAATATTTAGTGAAACCCTTCATTCCAGTCCTAAATTAAAAGACTA
AnJT CA-TTTTTAATAAGTCAGACCTTTCATTCCAGTCCTAAATTAAAGGGCTA
AnRV CA-TTTTTAATAAGTTAAACCTTTCATTCCAGTCCTAAATTAAAAGACTA
AnMo CA-TTTTTAGTAAGTTAAACCTTTCATTCCAGTCTCAAATTAAAAGACTA
AnUt CA-TTTTTAGTAAGTTAAACCTTTCATTCCAGTCTCAAATTAAAAGACTA
AnCH CA-TTTTTAATAAGTTAAACCTTTCATTCCAGTCCTAAATTAAAAGACTA
AnPM CA-TTTTTAATAAGTCAAACCTTTCATTCCAGTCCTAAATTAAAAGACTA
```

Euscorpius

```
EalFO  CTTCCTTATTTTATTATTGCATGAAATAGGTATCTTAATCCAACATCGAGGTCACAAACT
EalGO  CTTCCTTATTTTATTATTGCATGAAATAGGTATCTTAATCCAACATCGAGGTCACAAACT
EalSP  CTTCCTTATTTTATTATTGCATGAAATAGGTATCTTAATCCAACATCGAGGTCACAAACT
EalSC  CTTCCTTATTTTATTATTGCATGAAATAGGTATCTTAATCCAACATCGAGGTCACAAACT
EalRA  CTTCCTTATTTTATTATTGCATGAAATAGGTA-CTTAATCCAACATCGAGGTCACAAACT
EalCA  CTTCCTTATTTTATTATTGCATGAAATAGGTATCTTAATCCAACATCGAGGTCACAAACT
EalOL  CTTCCTTATTTTATTATTGCATGAAATAGGTATCTTAATCCAACATCGAGGTCACAAACT
EalSO  CTTCCTTATTTTATTATTGCATGAAATAGGTATCTTAATCCAACATCGNGGTCACAAACT
EalTA  CTTCCTTATTTTATTATTGCATGAAATAGGTATCTTAATCCAACATCGAGGTCACAAACT
EalML  CTTCCTTATTTTATTATTGCATGAAATAGGTATCTTAATCCAACATCGAGGTCACAAACT
EalMA  CTTCCTTATTTTATTATTGCATGAAATAGGTATCTTAATCCAACATCGAGGTCACAAACT
EgeOB  CTTCCTTATTTTATTATTGCATGAAATAGGTATCTTAATCCAACATCGAGGTCACAAACT
EgeDE  CTTCCTTATTTTATTATTGCATGAAATAGGTATCTTAATCCAACATCGAGGTCACAAACT
EgeST  CTTCCTTATTTTATTATTGCATGAAATAGGTATCTTAATCCAACATCGAGGTCACAAACT
EgeBO  CTTCCTTATTTTATTATTGCATGAAATAGGTATCTTAATCCAACATCGAGGTCACAAACT
EgeSM  CTTCCTTATTTTATTATTGCATGAAATAGGTATCTTAATCCAACATCGAGGTCACAAACT
EgeVE  CTTCCTTATTTTATTATTGCATGAAATAGGTATCTTAATCCAACATCGAGGTCACAAACT
EgeSH  CTTCCTTATTTTATTATTGCATGAAATAGGTATCTTAATCCAACATCGAGGTCACAAACT
EgeVO  CTTCCTTATTTTATTATTGCATGAAATAGGTATCTTAATCCAACATCGAGGTCACAAACT
EgeME  CTTCCTTATTTTATTATTGCATGAAATAGGTATCTTAATCCAACATCGAGGTCACAAACT
EgeKR  CTTCCTTATTTTATTATTGCATGAAATAGGTATCTTAATCCAACATCGAGGTCACAAACT
EgeTZ  CTTCCTTATTTTATTATTGCATGAAATAGGTATCTTAATCCAACATCGAGGTCACAAACT
EgeCR  CTTCCTTATTTTATTATTgCATgAAATAGGTATCTTAATCCAACATCGAGGTCACAAACT
EgeBE  CTTCCTTGTTTTATTATTGCATGAAATAGGTATCTTAATCCAACATCGAGGTCACAAACT
EgaKO  CTTCCTTATTTTATTATTGCATGAAATAGGTATCTTAATCCAACATCGAGGTCACAAACT
EflLA  CTCCCTTATTTTATTATTGCATGAAATAGGGCTCTTAATCCAACATCGAGGTCACAAACT

EalFO TTCTTGATGATAAGAACTCTTTAAGAAAATTATGCTGTTATCCCTATAGTAACTTGTTCT
EalGO TTCTTGATGATAAGAACTCTTTAAGAAAATTATGCTGTTATCCCTATAGTAACTTGTTCT
EalSP TTCTTGATGATAAGAACTCTTTAAGAAAATTATGCTGTTATCCCTATAGTAACTTGTTCT
EalSC TTCTTGATGATAAGAACTCTTTAAGAAAATTATGCTGTTATCCCTATAGTAACTTGTTCT
EalRA TTCTTGATGATAAGAACTCTTTAAGAAAATTATGCTGTTATCCCTATAGTAACTTGTTCT
EalCA TTCTTGATGATAAGAACTCTTTAAGAAAATTATGCTGTTATCCCTATAGTAACTTGTTCT
EalOL TTCTTGATGATAAGAACTCTTTAAGAAAATTATGCTGTTATCCCTATAGTAACTTGTTCT
EalSO TTCTTGATGATAAGAACTCTTTAAGAAAATTATGCTGTTATCCCTATAGTAACTTGTTCT
EalTA TTCTTGATGATAAGAACTCTTTAAGAAAATTATGCTGTTATCCCTATAGTAACTTATTCT
EalML TTCTTGATGATAAGAACTCTTTAAGAAAATTATGCTGTTATCCCTATAGTAACTTATTCT
EalMA TTCTTGATGATAAGAACTCTTTAAGAAAATTATGCTGTTATCCCTATAGTAACTTATTCT
EgeOB TTCTTGATGATAAGAACTCTTGAAGAAAATTATGCTGTTATCCCTATAGTAACTTATTCT
EgeDE TTCTTGATGATAAGAACTCTTGAAGAAAATTATGCTGTTATCCCTATAGTAACTTATTCT
EgeST TTCTTGATGATAAGAACTCTTGAAGAAAATTATGCTGTTATCCCTATAGTAACTTATTCT
EgeBO TTCTTGATGATAAGAACTCTTTAAGAAAATTATGCTGTTATCCCTATAGTAACTTATTCT
EgeSM TTCTTGATGATAAGAACTCTTGAAGAAAATTATGCTGTTATCCCTATAGTAACTTATTCT
EgeVE TTCTTGATGATAAGAACTCTTGAAGAAAATTATGCTGTTATCCCTATAGTAACTTATTCT
EgeSH TTCTTGATGATAAGAACTCTTGAAGAAAATTATGCTGTTATTCTTGATGATAAGAACTCT
EgeVO TTCTTGATGATAAGAACTCTTGAAGAAAATTATGCTGTTATCCCTATAGTAACTTATTCT
EgeME TTCTTGATGATAAGAACTCTTGAAGAAAATTATGCTGTTATCCCTATAGTAACTTATTCT
EgeKR TTCTTGATGATAAGAACTCTTGAAGAAAATTATGCTGTTATCCCTATAGTAACTTATTCT
EgeTZ TTCTTGATGATAAGAACTCTTGAAGAAAATTATGCTGTTATCCCTATAGTAACTTATTCT
EgeCR TTCTTGATGATAAGGACTCTTGAAGAAAATTATGCTGTTATCCCTATAGTAACTTATTCT
EgeBE TTCTTGATGATAAGAACTCTTAAAGAAAATTATGCTGTTATCCCTATAGTAACTTATTCT
EgaKO TTCTTGATGATAAGAACTCTTAAAGAAAATTATGCTGTTATCCCTATAGTAACTTATTCT
EflLA TTCTTGATGATAAGGACTCTTAAAGAAAATTATGCTGTTATCCCTACAGTAACTTGTTCC

EalFO TTTTTAAAAACTTTTTGGTTTTTCAAGAGTTATCCTCACATTATTAAAATAAAATTTTA
EalGO TTTTTAAAAACTTTTTGGTTTTTCAAGAGTTATCCTCACATTATTAAAATAAAATTTTA
```

```
EalSP TTTTTAAAAACTTTTTGGTTTTTCAAGAGTTATTCTCATATTATTAAAATAAAATTTTA
EalSC TTTTTAAAAACTTTTTGGTTTTTCAAGAGTTATTCTCATATTATTAAAATAAAATTTTA
EalRA TTTTTAAAAACTTTTTGGTTTTTCAAGAGTTATTCTCATATTATTAAAATAAAATTTTA
EalCA TTTTTAAAAACTTTTTGGTTTTTCAAGAGTTATTCTCATATTATTAAAATAAAATTTTA
EalOL TTTTTAAAAACTTTTTGGTTTTTCAAGAGTTATTCTCATATTATTAAAATAAAATTTTA
EalSO TTTTTAAAAACTTTTTGGTTTTTCAAGAGTTATTCTCATATTATTAAAATAAAATTTTA
EalTA TTTTTAAAAACTTTTTGGTTTTTCAAGAGTTATTCTCATATTATTAAAATAAAATTTTA
EalML TTTTTAAAAACTTTTTGGTTTTTCAAGAGTTATTCTCACATTATTAAAATAAAATTTTA
EalMA TTTTTAAAAACTTTTTGGTTTTTCAAGAGTTATTCTCACATTATTAAAATAAAATTTTA
EgeOB TTTTTAAAAACTTTTTGATTGTTCAAGAGTTACTCCCACGTTATTGAAATAAAATTTTA
EgeDE TTTTTAAAAACTTTTTGATTGTTCAAGAGTTACTCCCACGTTATTGAAATAAAATTTTA
EgeST TTTTTAAAAACTTTTTGATTGTTCAAGAGTTACTCCCACGTTATTGAAATAAAATTTTA
EgeBO TTTTTAAAAACTTTTTGATTGTTCAAGAGTTACTCCCACGTTATTGAAATAAAATTTTA
EgeSM TTTTTAAAAATTTTTTGATTGTTCAAGAGTTATTCCCACGTTATTGAAATAAAATTTTA
EgeVE TTTTTAAAAATTTTTTGATTGTTCAAGAGTTATTCCCACGTTATTGAAATAAAATTTTA
EgeSH TTTTTAAAAATTTTTTGATTGTTCAAGAGTTATTCCCACGTTATTGAAATAAAATTTTA
EgeVO TTTTTAAAAATTTTTTGATTGTTCAAGAGTTATTCCCACGTTATTGAAATAAAATTTTA
EgeME TTTTTAAAAATTTTTTGATTGTTCAAGAGTTATTCCCACGTTATTGAAATAAAATTTTA
EgeKR TTTTTAAAAATTTTTTGATTGTTCAAGAGTTATTCCCACGTTATTGAAATAAAATTTTA
EgeTZ TTTTTAAAAATTTTTTGATTGTTCAAGAGTTATTCCCACGTTATTGAAATAAAATTTTA
EgeCR TTTTTAAAAACTTTTTGATTGGTCAAGAGTTAATCCCACGTTATTAAAATAAAATTTTA
EgeBE TTTTTAAAAACTTTTTGATTATTCAAGAGTTATTCCCACGTTATTGAAATAAAATTTTA
EgaKO TGTTTAAGAATTTTTTGGTTTTTCAAGAATTAAATTCACATTATAAAAATAAAATTTTA
EflLA TTCTTAAGAAATTTTTGATTTTTCAAGAGTTGTACTCACATTATATAAATAAAATTTTA
```

```
EalFO TTTATTTACTGCCCCAGTAAA
EalGO TTTATTTACTGCCCCAGTAAA
EalSP TTTATTTACTGCCCCAGTAAA
EalSC TTTATTTACTGCCCCAGTAAA
EalRA TTTATTTACTGCCCCAGTAGA
EalCA TTTATTTACTGCCCCAGTAGA
EalOL TTTATTTACTGCCCCAGTATA
EalSO TTTATTTACTGCCCCAGTAGA
EalTA TTTATTTACTGCCCCAGTAAA
EalML TTTATTTACTGCCCCAGTAAA
EalMA TTTATTTACTGCCCCAGTAAA
EgeOB TTTATTTACTGCCCCAGTAAA
EgeDE TTTATTTACTGCCCCAGTAGA
EgeST TTTATTTACTGCCCCAGTAGA
EgeBO TTTATTTACTGCCCCAGTAGA
EgeSM TTTATTTACTGCCCCAGTAGA
EgeVE TTTATTTACTGCCCCAGTAGA
EgeSH TTTATTTACTGCCCCAGTAGA
EgeVO TTTATTTACTGCCCCAGTAGA
EgeME TTTATTTACTGCCCCAGTAGA
EgeKR TTTATTTACTGCCCCAGTAGA
EegTZ TTTATTTACTGCCCCAGTAGA
EgeCR TTTATCTACTGCCCCAGTAAA
EgeBE TTTATTTACTGCCCCAGTAGA
EgaKO TTTATCTACTGCCCCAGTAAA
EflLA TTTATTTACTGCCCCAGTAAA
```

```
EalFO  ATAATTTGTGGTTTATATTGA-TGTTAAATTAGAATTATAAAGCTTAATAGGGTCTTCTT
EalGO  ATAATTTGTGGTTTATATTGA-TGTTAAATTAGAATTATAAAGCTTAATAGGGTCTTCTT
EalSP  ATAATTTGTGGTTTATATTGA-TGTTAAATTGGAATTATAAAGCTTAATAGGGTCTTCTT
```

```
EalSC   ATAATTTGTGGTTTATATTGA-TGTTAAATTGGAATTATAAAGCTTAATAGGGTCTTCTT
EalRA   ATAATTTGTGATTTATATTGA-TGTTAAATTGGAATTATAAAGCTTAATAGGGTCTTCTT
EalCA   ATAATTTGTGATTTATATTGA-TGTTAAATTAGAATTATAAAGCTTAATAGGGTCTTCTT
EalOL   ATAATTTGTGATTTATATTGA-TGTTAAATTGGAATTATAAAGCTTAATAGGGTCTTCTT
EalSO   ATAATTTGTGATTTATATTGA-TGTTAAATTGGAATTATAAAGCTTAATAGGGTCTTCTT
EalTA   ATAATTTGTAATTTATATTGA-TGTTAAGTTGGAATTATAAAGCTTAATAGGGTCTTCTT
EalML   ATAATTTATGATTTATATTGA-TGTTAAATTGGAATTATAAAGCTTAATAGGGTCTTCTT
EalMA   ATAATTTATGATTTATATTGA-TGTTAAATTGGAATTATAAAGCTTAATAGGGTCTTCTT
EgeOB   ATAATTTTTAGTTTATGTTGA-TATTAAATTGAGATTATAAAGCTTAATAGGGTCTTCTT
EgeDE   ATAATTTTTAGTTTATGTTGA-TATTAAATTGAGATTATAAAGCTTAATAGGGTCTTCTT
EgeST   ATAATTTTTAGTTTATGTTGA-TATTAAATTGAGATTATAAAGCTTAATAGGGTCTTCTT
EgeBO   ATAATTTTTAGTTTATGTTGA-TATTAAATTGAGATTATAAAGCTTAATAGGGTCTTCTT
EgeSM   ATAATTTTTAGTTTATGTTGA-TATTAAATTGAGATTATAAAGCTTAATAGGGTCTTCTT
EgeVE   ATAATTTTTAGTTTATGTTGA-TATTAAATTGAGATTATAAAGCTTAATAGGGTCTTCTT
EgeSH   ATAATTTTTAGTTTATGTTGA-TATTAAATTGAGATTATAAAGCTTAATAGGGTCTTCTT
EgeVO   ATAATTTTTAGTTTATGTTGA-TATTAAATTGAGATTATAAAGCTTAATAGGGTCTTCTT
EgeME   ATAATTTTTAGTTTATGTTGA-TATTAAATTGAGATTATAAAGCTTGATAGGGTCTTCTT
EgeKR   ATAATTTTTAGTTTATGTTGA-TATTAAATTGAGATTATAAAGCTTAATAGGGTCTTCTT
EgeTZ   ATAATTTTTAGTTTATGTTGA-TATTAAATTGAGATTATAAAGCTTAATAGGGTCTTCTT
EgeCR   ATAATTTTTAATTTATATTGA-TATTAAACTAAGATTATAAAGCTTAATAGGGTCTTCTT
EgeBE   ATAATTTTTGATTTATGTTGA-TATTAAATTGAGATTATAAAGCTTAATAGGGTCTTCTT
EgaKO   ATAATTTTTAATTTATATTGA-TATTAAATTGGAATTATAAAGCTTAATAGGGTCTTCTT
EflLA   ATAATTTTTAACTATGTTGATGTTAGATTTTAATTATAAAGCTTGATAGGGTCTTCTT
```

```
EalFO GTCTTTAAAATAAATTTTAGCCTTTTTACTAAAATGTAAAATTTGAAG
EalGO GTCTTTAAAATAAATTTTAGCCTTTTTACTAAAATGTAAAATTTGAAG
EalSP GTCTTTAAAGTAAATTTTAGCCTTTTTACTAAAATGTAAAATTTGAAG
EalSC GTCTTTAAAGTAAATTTTAGCCTTTTTACTAAAATGTAAAATTTGAAG
EalRA GTCTTTAAAGTAAATTTTAGCCTTTTTACTAGAATGTAAAATTTGAAG
EalCA GTCTTTAAAGTAAATTTTAGCCTTTTTACTAGAATGTAAAATTTGAAG
EalOL GTCTTTAAAGTAAATTTTAGCCTTTTTACTAGAATGTAAAATTTGAAG
EalSO GTCTTTAAAGTAAATTTTAGCCTTTTTACTAGAATGTAAAATTTGAAG
EalTA GTCTTTAAAGTGAATTTTAGCCTTTTTACTAAAATGTAAAATTTGAAG
EalML GTCTTTAAAGTGAATTTTAGCCTTTTTACTAAAATGTAAAATTTGAAG
EalMA GTCTTTAAAGTGAATTTTAGCCTTTTTACTAAAATGTAAAATTTGAAG
EgeOB GTCTTTAAAATGGATTTTAGCTTTTTTACTAAAATATAAAATTTGAAG
EgeDE GTCTTTAAAATGGATTTTAGCTTTTTTACTAAAATATAAAATTTGAAG
EgeST GTCTTTAAAATGGATTTTAGCTTTTTTACTAAAATATAAAATTTGAAG
EgeBO GTCTTTAAAATGGATTTTAGCTTTTTTACTAAAATATAAAATTTGAAG
EgeSM GTCTTTAAAATGGATTTTAGCTTTTTTACTAAAATATAAAATTTGAAG
EgeVE GTCTTTAAAATGGATTTTAGCTTTTTTACTAAAATATAAAATTTGAAG
EgeSH GTCTTTAAAATGGATTTTAGCTTTTTTACTAAAATATAAAATTTGAAG
EgeVO GTCTTTAAAATGGATTTTAGCTTTTTTACTAAAATATAAAATTTGAAG
EgeME GTCTTTAAAATGGATTTTAGCTTTTTTACTAAAATATAAAATTTGAAG
EgeKR -TCTTTAAAATGGATTTTAGCTTTTTTACTAAAATATAAAATTTGAAG
EegTZ GTCTTTAAAATGGATTTTAGCTTTTTTACTAAAATATAAAATTTGAAG
EgeCR GTCTTTAAAGTGGATTTTAGCTTTTTTACTAAAATATAAAATTTGAAG
EgeBE GTCTTTAAAGTGGATTTTAGCTTTTTTACTAAAATATAAAGTTTGAAG
EgaKO GTCTTTAAAGTGGATTTTAGCTTTTTTACTAAAATATAAAATTTGAAG
EflLA GTCTTTAATAAGAATTTTAGCTTTTTTACTAAAATATAAAATTTGAAG
```

```
EalFO   TATAATAATAAGACAGGATTATTTAGTTAAACCATTCATTCCAGTCCTAAATTACAAG
EalGO   TATAATAATAAGACAGGATTATTTAGTTAAACCATTCATTCCAGTCCTAAATTACAAG
```

```
EalSP  TATAATAATAAGACAAGATTATTTAGTTAAACCATTCATTCCNGNCCTAAGTTNCNNN
EalSC  TATAATAATAAGACAAGATTATTTAGTTAAACCATTCATTCCAANCCTAATTTNCANG
EalRA  TATAATAATAAGACAGGATTATTTAGTTAAACCATTCATTCCAGTCCTAAATTACAAG
EalCA  TATAATAATAAGACAGGATTATTTAGTTAAACCATTCATTCCAGTCCTAAATTACAAG
EalOL  TATAATAATAAGACAGGATTATTTAGTTAAACCATTCATTCAAGNCCTAAATTNCANG
EalSO  TATAATAATAAGACAGGATTATTTAGTTAAACCATTCATTCCAGTCCTAAATTACAAG
EalTA  TATAATAATAAGACAGGATTATTTAGTTAAACCATTCATTCCAGTCCTAAATTACAAG
EalML  TATAATAATAAGACAGGATTATTTAGTTAAACCATTCATTCCAGTCCTAAATTACAAN
EalMA  TATAATAATAAGACAGGATTATTTAGTTAAACCATTCATTCCAGTCCTAAATTACAAG
EgeOB  TATAATAATAAGACATGATTATTTAGTTAAACCATTCATTCCAGTCCTAAATTACAAG
EgeDE  TATAATAATAAGACATGATTATTTAGTTAAACCATTCATTCCAGTCCTAAATTACAAG
EgeST  TATAATAATAAGACATGATTATTTAGTTAAACCATTCATTCCAGTCCTAAATTACAAG
EgeBO  TATAATAATAAGACATGATTATTTAGTTAAACCATTCATTCCAGTCCTAAATTACAAG
EgeSM  TATAATAATAAGACATGATTATTTAGTTAAACCATTCATTCCAGTCCTAAATTACAAG
EgeVE  TATAATAATAAGACATGATTATTTAGTTAAACCATTCATTCCAGTCCTAAATTACAAG
EgeSH  TATAATAATAAGACATGATTATTTAGTTAAACCATTCATTCCAGTCCTAAATTACAAG
EgeVO  TATAATAATAAGACATGATTATTTAGTTAAACCATTCATTCCAGTCCTAAATTACAAG
EgeME  TATAATAATAAGACATGATTATTTAGTTAAACCATTCATTCCAGTCCTAAATTACAAG
EgeKR  TATAATAATAAGACATGATTATTTAGTTAAACCATTCATTCCAGTCCTAAATTACAAG
EgeTZ  TATAATAATAAGACATGATTATTTAGTTAAACCATTCATTCCAGTCCTAAATTACAAG
EgeCR  TATAATAATAAGACATGATTATTTAGTTAAACCATTCATTCCAGTCCTAAATTACAAG
EgeBE  TATAATAATAAGACATAATTATTTAGTTAAACCATTCATTCCAGTCCTAAATTACAAG
EgaKO  TATAATAATAAGACATGATTATTTAGTTAAACCATTCATTCCAGTCCTAAATTACAAG
EflLA  TATAATAATAAGACAGGAGTACTCAGTTAAACCATTCATTCCAGTCCTAAATTACAAG
```

Centruroides

```
CeBA CGAACAGA-CTCCTTTCATTCCTCTTGCGGAATGAAGGAAATTTAATCCA
CeSI GCGCACCCCCNCCTTTCATTCCTCTTGCGGAATGGAGGAAATTTAATCCA
CeJU TCGAACAGACTCCTTCCATTCCTCTTGCGGAATGGAGGAAATTTAATCCA
CeLP NNNNNNNNNNNNNNTCCCCTCCTCTTGCGGAGTGGAGGAAACTTAATCCA
CeBC CGA-CAGA-CCCCCTCCCTTCCTCTTGCGGAATGGCGGAAATCTAATCCA
CeTU NNNNNNCCNNCCCCTCCCTTCCTCTTGCGGAATGGAGGAAATCTAATCCA
CeTE NNNNNNNNNNNNNNNCCCCTCCTCTTGCGGAGTGGAGGAAATCTAATCCA
CeYU CGAACAGACCCCCTCCCATTCCTCTTGCGGAATGGAGGAAATCTAATCCA
CePP CGAACAGACCCCCTTCCCTTCCTCTTGCGGAATGGAGGAAACCTAATCCA
CeGU CGAACAGACCCCCTACCACTCCTCTTGCGGGGTGGAGGAAATCTAATCCA
CvAR NNNNNNNNCTTCCTATTATTACTCTTGCGTAATACAGGGAATTTAATCCA
Cban NNNNNNNNNCCCCTATTAATACTCTTGCATAATAAAGGAAAACTAATCCA

CeBA ACATCGAGGTCGCAAACATACTTGTCGATTTGAGCTTTCGAAGTATATTA
CeSI ACATCGAGGTCGCAAACATACTTGTCGATTCGAGCTTTCAAAGTATATTA
CeJU ACATCGAGGTCGCAAACATACTTGTCGATTTGAGCTTTCAAAGTATATTA
CeLP ACATCGAGGTCGCAAACATATTTGTCAATTTGGACTTTCAAAATATATTA
CeBC ACATCGAGGTCGCAAACATGTTTGTCAATTTGAGCTTTCAAAACATATTA
CeTU ACATCGAGGTCGCAAACGTGTTTGTCAATTTGAGCTTTCAAAACACATTA
CeTE ACATCGAGGTCGCAAACATGTTTGTCAATTTGAGCTTTCAAAACACATTA
CeYU ACATCGAGGTCGCAAACATATTTGTCAATTTGAGCTTTCAAAACACATTA
CePP ACATCGAGGTCGCAAACATGTTTGTCAATTTGAGCTTTCAAAACACATTA
CeGU ACATCGAGGTCGCAAACATGTTTGTCAATTTGAGCTTTCAAAACACATTA
CvAR ACATCGAGGTCACAAACACGTTTATCAATTTGAACTTTCAAAACATATCA
Cban ACATCGAGGTCGCAAACACATTTGTCAATTTGAACTTTCAAAATATATTA
```

```
CeBA CGCTGTTATCCCTAAAGTAACTTATTTAAACTTCAAAAATTTTGGGTATT
CeSI CGCTGTTATCCCTAAAGTAACTTATTTAAGCTTCAAAAATTTTGGGTATT
CeJU CGCTGTTATCCCTAAAGTAACTTATTTAAGTTTCAAAAATTTTGGGTATT
CeCA CGCTGTTATCCCTAAAGTAACTTATTTAATTTTCAAAAATTTTGGGTATT
CeBC CGCTGTTATCCCTAAAGTAACTTATTTAACTTTCAAAAATTTTGGGTATT
CeTU CGCTGTTATCCCTAAAGTAACTTATTTAACCTTCAAAAATTTTGGGTATT
CeTE CGCTGTTATCCCTAAAGTAACTTATTTAACTTTCAAAAATTTTGGGTATT
CeYU CGCTGTTATCCCTAAAGTAACTTATTTAACTTTCAAAAATTTTGGGTATT
CePP CGCTGTTATCCCTAAAGTAACTTATTTAATTTTCAAAAATTTTGGGTATT
CeGU CGCTGTTATCCCTAAAGTAACTTATTTAACTTTCAAAAATTTTGGGTATT
CvAR CGCTGTTATCCCTAAAGTAACTTATTTAAATTTCAAAATTTTTGGGTATT
Cban  CGCTGTTATCCCTAAAGTAACTTATTTAATATTCAAAAGTTTTGGGTATT

CeBA AAGA-TAATGTTATCC-TCATGCTCTAAAAGTGTTTT-TCTTTCTACCGC
CeSI AAGA-TGATGTTATCC-TTATACTCTAAAAGTGTTTT-TCTTTCTACCGC
CeJU AAAA-TAATGTTATCT-TTATACTCTAAAAGTGTTTT-TCTTTCTACCGC
CeCA AAAAATAATGTCATTT-TTATATTTGAAAGTGTTTTATCTTCCTACCGC
CeBC AAAAATAATGTCATTT-ATATATTTTAAAAGTGTTTT--CTTCTTACCGC
CeTU AAAA-TAATGTCATTT-ATATATTTTAAAAGTGTTTT-TCTTCCTACCGC
CeTE AAAA-TAATGTCATTT-ATATATTTTAAAAGTGTTTT-TCTTCTTACCGC
CeYU AAAAATAATGTCATTT-ATATTTTTTAAAAGTGTTTT-TCTTCTTACCGC
CePP AAAA-TAATGTCATTT-ATATATTTTAAAAGTGTTTT-TCTTCTTACCGC
CeGU AAAAATAATGTCATTT-ATATATTTTAAAAGTGTTTT-TCTTCTTACCGC
CvAR AAAAATAATATTATTT-TTAATATTTAAAAGTGTTTT-TCTTTCTGCCGC
Cban AAAAACTATCCTATTTATAATAATTTAAAAGTGTTTT-TCTTCTTACCGC

CeBA CCCAGTAAAACATATTTTTAATTTATTAAATTATTTTATGTAAAGCTTTA
CeSI CCCAGTAAAACATATTTCCAATTTATTAAATTATTTTATGTAAAGCTTTA
CeJU CCCAGTAAAACATATTTCTAATTTGTTAAATTATTTTATGTAAAGCTTTA
CeCA CCCAGTAAAACATATTTTCAATCTGTTAAATTATTTTATGTAAAGCTTTA
CeBC CCCAGTAAAACATATTTTTAATTTATTAAATTATTTTATGTAAAGCTTTA
CeTU CCCAGTAAAACATATTTTTAATTTATTAAATTATTTTATGTAAAGCTTTA
CeTE CCCAGTAAAACATATTTTTAATTTATTAAATTATTTTATGTAAAGCTTTA
CeYU CCCAGTAAAACATATTTTTAATTTATTAAATTATTTTATGTAAAGCTTTA
CePP CCCAGTAAAACATATTTTTAATTTATTAAATTATTTTATGTAAAGCTTTA
CeGU CCCAGTAAAACATATTTTTAATTTATTAAATTATTTTATGTAAAGCTTTA
CvAR CCCAGCAAAACATATTCTCAATTTATTAAATTTCTTATGTAAAGTTTTA
Cban CCCAGTAAAACATATTTTCAATTTTCTAGATTATTTTATGTAAAGCTTTA

CeBA TAGGGTCTTCTTGTCTAAAAGAAACATTTTAGCCTTTTTACTAAAAAGTA
CeSI TAGGGTCTTCTTGTCTAAAAGGAACATTTTAGCCTTTTCACTAAAAGGTA
CeJU TAGGGTCTTCTTGTCTAAAAGAAACATTTTAGCCTTTTTACTAAAAGGTA
CeCA TAGGGTCTTCTTGTCTAAAAGAAACATTTTAGCCTTTTTACTAAAAAGTA
CeBC TAGGGTCTTCTTGTCTAAAAGAAACATTTTAGCCTTTTTACTAAAAAGTA
CeTU TAGGGTCTTCTTGTCTAAAAGAAACATTTTAGCCTTTTTACTAAAAAGTA
CeTE TAGGGTCTTCTTGTCTAAAAGAAACATTTTAGCCTTTTTACTAAAAAGTA
CeYU TAGGGTCTTCTTGTCTAAAAGAAACATTTTAGCCTTTTTACTAAAAAGTA
CePP TAGGGTCTTCTTGTCTAAAAGAGACATTTTAGCCTTTTTACTAAAAAGTA
CeGU TAGGGTCTTCTTGTCTAAAAGAAACATTTTAGCCTTTTTACTAAAAAGTA
CvAR TAGGGTCTTCTTGTCTAAAAGAGACATTTTAGCCTTCTTACTAAAAGGTG
Cban TAGGGTCTTCTTGTCTAAAAGATGTATTTTAGCCTTTTTACTAAAAAGTA

CeBA AAATTCAAAAGAAAAAGTTAAGAAAGAAACTCTTCAGTTTATCCCTTCAT
CeSI AAGTTCAAAAGAAAAAGTTAAGAAAGAAACTCTTCAGTTTATCCCTTCAT
CeJU AAGTTCAAAAGAAAAAGTTAAGAAAGAAACTCTTCAGTTTATCCCTTCAT
CeCA AAGTTCAAAAGAAAGAGCCAAGAAAGAAACTCTCTAGTTTATCCTTTCAT
```

```
CeBC AATTTCAAAAGAAAAAGCCAAGAAAGAAACTTTCTAGTTTATCCCTTCAT
CeTU AATTTCAAAAGAAAAAGCTAAGAAAGAAACTTTCTAGTTTATCCCTTCAT
CeTE AATTTCAAAAGAAAAAGTCAAGAAAGAAACTTTCTAGTTTATCCTTTCAT
CeYU AATTTCAAAAGAAAAAGCCAAGAAAGAAACTTTCTAGTTTATCCCTTCAT
CePP AATTTCAAAAGAAAAAGCCAAGAAAGAAACTTTCTAGTCTATCCTTTCAT
CeGU AATTTCAAAAGAAAAAGTCAAGAAAGAAACTTTCTAGTTTATTCCTTCAT
CvAR AAATTTGAAATTAATAGTTAAGAAAGAAATTTTCTAGTTTACCCTTTCAT
Cban AAGTTAAAAAAGTAAAATTAAAAAAGAGCTTTTCCAGTCTATCCTTTCAT

CeBA TCCAGTCTTAAATTACAAGACTA
CeSI TCCAGTCTTAAATTACAAGACTA
CeJU TCCAGTCTTAAATTACAAGACTA
CeCA TCCAGTCTTAAATTACAAGACTA
CeBC TCCAGTCTTAAATTATAAGACTA
CeTU TCCAGTCTTAAATTACAAGACTA
CeTE TCCAGTCTTAAATTACAAGACTA
CeYU TCCAGTCTTAAATTACAAGACTA
CePP TCCAGTCTTAAATTACAAGACTA
CeGU TCCAGTCTTAAATTACAAGACTA
CvAR TCCAGTCTTAAATTACAAGACTA
Cban TCCAGTCTTCAATTATAAGACTA
```