Biogeographica 77 (4) : 157-172 Paris, décembre 2001 ISSN 1165-6638

MITOCHONDRIAL DNA REVEALS A DIVERGENT PHYLOGENY IN TROPICAL *CENTRUROIDES* (SCORPIONES: BUTHIDAE) FROM MEXICO

W. Ian TOWLER[1], Javier PONCE SAAVEDRA[2], Benjamin GANTENBEIN[3] & Victor FET[1]

[1] *Department of Biological Sciences, Marshall University, Huntington, West Virginia 25755-2510,* USA; *email: fet@marshall.edu*

[2] *Laboratorio de Entomología, Facultad de Biología, Universidad Michoacana de San Nicolás de Hidalgo, Morelia, Michoacán, CP 58060,* MEXICO

[3] *Institute of Cell, Animal and Population Biology, Ashworth Laboratories, The University of Edinburgh, West Mains Road, Edinburgh EH9 3JT, Scotland,* UNITED KINGDOM

ABSTRACT. - Numerous species of highly toxic *Centruroides* scorpions inhabit the tropics of Mexico. Relationships within two common species *C. infamatus* (C. L. Koch, 1844) and *C. limpidus* (Karsch, 1879), and their subspecies, are not known; their existing taxonomy is based on a few morphological characters. In order to further clarify the status of these taxa, we initiated a molecular survey applying mitochondrial (16S rRNA mtDNA sequences) gene markers. Our pilot molecular data set recovers a clear divergent phylogeny among the populations of these species from Michoacán, Guerrero, and Querétaro. Four deep monophyletic clades (putative species) are revealed, with monophyly of *C. limpidus* is not confirmed; *C. limpidus tecomanus* Hoffmann, 1932 does not appear to belong to *C. limpidus*. Phylogenetic, taxonomic, and biogeographic implications are discussed.

KEY-WORDS. - Scorpion, Mexico, Michoacán, *Centruroides infamatus, C. limpidus*, mtDNA

RESUME.- Des nombreuses espèces très toxiques de scorpions du genre *Centruroides* habitent les régions tropicales du Mexique. Les relations entre deux espèces communes, *C. infamatus* (C. L. Koch, 1844) et *C. limpidus* (Karsch, 1879) et leurs sous-espèces ne sont pas connues ; la taxonomie existante est fondée sur quelques caractères morphologiques. Dans le but de clarifier le statut de ces taxa, nous avons débuté une analyse moléculaire avec utilisation de marqueurs géniques mitochondriaux (séquences 16S rARN mtADN). Nos données moléculaires pilotes démontrent une nette divergence phylogénétique entre les populations de ces espèces, originaires de Michoacán, Guerrero et Querétaro. Quatre clades monophylétiques (espèces putatives), sont révélées, avec la monophylie de *C. limpidus* non confirmée ; *C. limpidus tecomanus* Hoffmann, 1932 ne semble pas appartenir à *C. limpidus*. Des implications taxonomiques et biogéographiques sont discutées.

MOTS-CLES. - Scorpion, Mexico, Michoacán, *Centruroides infamatus, C. limpidus*, mtADN

INTRODUCTION

Centruroides Marx, 1890, with 41 species and 24 subspecies (FET & LOWE, 2000) is one of the most diverse genera of the scorpion family Buthidae (second only to *Tityus* C. L. Koch), and the only genus of this family found in North America. Scorpions of this genus (including many Central Mexican species) are among the most toxic in the world for humans and other mammals (MAZZOTTI & BRAVO-BECHERELLE, 1963; SIMARD &

WATT, 1990). They are abundant in various natural habitats ranging from tropical forest to temperate deserts. *Centruroides* is especially diverse in Mexico and the Caribbean, but is also found in Central and South America (SISSOM & LOURENÇO, 1987; FET & LOWE, 2000). Taxonomy of many *Centruroides* species is confusing, and was traditionally based mainly or exclusively on the features of morphosculpture and coloration characters. No modern taxonomic or phylogenetic analysis exists for this genus, and the only existing generic key (STAHNKE & CALOS, 1977) is outdated. The only faunal treatment of the Mexican fauna was that by HOFFMANN (1932). The keys and lists for 32 species and subspecies of *Centruroides* found in Mexico were produced by DÍAZ NÁJERA (1964, 1966, 1975), BEUTELSPACHER (2000); FET and LOWE (2000). Although a comparative study (DEHESA-DÁVILA *et al.*, 1996) on the primary structures of Na^+ channel-specific toxins in *C. infamatus infamatus* (C. L. Koch, 1844) and in *C. limpidus limpidus* (Karsch, 1879) has recently been done, the phylogenetic relationships among these species remain uncertain.

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 (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".

This paper is a part of an ongoing study of scorpion fauna, ecology, biology, and biogeography in the state of Michoacán (PONCE & BEUTELSPACHER, 2001). We report here the first pilot data based on the analysis of mitochondrial DNA (mtDNA) sequences of the 16S rRNA gene from two common species of *Centruroides* endemic to Central Mexico. 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, 1863) 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 was high enough for a good phylogenetic signal (HUELSENBECK & HILLIS, 1996).

MATERIAL AND METHODS

SPECIMENS ANALYSED

Total of 14 specimens of two species of *Centruroides* were analysed; each was represented by its two existing subspecies. These specimens were collected and identified by J.P.S. from the following 12 localities in Mexico (see Fig. 1; abbreviations as given below): *Centruroides infamatus infamatus* (C. L. Koch, 1844): Zumpimito, 1560 m a.s.l., Municipio de Uruapan, Michoacán, Mexico (*Cii*ZU1). – *C. infamatus ornatus* Pocock, 1902 (all Michoacán, Mexico): El Cobano, 1540 m, Municipio de Gabriel Zamora (*Cio*EC1); La Caratacua, 1680 m, Municipio de Coeneo (*Cio*LC1); near Morelia city, 1940 m, Municipio de Morelia (*Cio*MO1); Tiripetio, 2020 m, Municipio de Morelia (*Cio*TI1). – *C. limpidus limpidus* (Karsch, 1879): Churumuco, 300 m, Michoacán, Mexico (*Cll*CH1, *Cll*CH2); Arúa, 480 m, Municipio de Huetamo, Michoacán, Mexico (*Cll*AR1); Tzirandaro, 500 m, Guerrero, Mexico (*Cll*TO1); Tres Puentes, Municipio de

Tzitzio, Michoacan, Mexico, 1775 m (*Cll*TZ1); Huitzuco, 960 m, near Iguala, Guerrero, Mexico (*Cll*HU1); Cañon de los Cajones, 1820 m, Municipio de Querétaro, Querétaro, Mexico (*Cll*QU1). – *C. limpidus tecomanus* Hoffmann, 1932: El Faro de Bucerias, Municipio de Coahuayana, Michoacán, Mexico (*Clt*FB1, *Clt*FB2). These localities represented a broad geographic spectrum of habitats and climates characteristic for the study area, and varied in altitude from 300 to 1940 m a.s.l.

Three congeneric species were used for comparison, *C. exilicauda* (Wood, 1863) from Tempe, Arizona, USA (coll. J. Bigelow) (*Cex*TE), *C. vittatus* (Say, 1821) from Arkansas, USA (coll. T. Yamashita) (*Cvi*AR), and *C. bani* Armas & Marcano Fondeur, 1987 from Dominican Republic (coll. D. Huber) (*Cba*DO). Animals were shipped on dry ice to Marshall University (West Virginia, USA) and preserved in 96 % ethanol for DNA analysis. Voucher specimens are deposited in V.F.'s collection at Marshall University.

DNA EXTRACTION AND SEQUENCING TECHNIQUES

Comparative analysis of the mitochondrial (mt) 16S ribosomal RNA is widely used in modern molecular evolutionary studies of various arthropods (SIMON *et al.*, 1994). For detailed techniques and protocols see GANTENBEIN *et al.* (1999, 2000). Total DNA was extracted from preserved (in 95 % ethanol) muscle tissue using a standard extraction method (Qiagen™). An approximately 400 bp fragment of the 16S rRNA gene was amplified by the polymerase chain reaction (PCR) using the universal forward primer LR-J-12887 (SIMON *et al.*, 1994) and a scorpion-specific reverse primer (GANTENBEIN *et al.*, 1999).

PHYLOGENETIC ANALYSES

17 mtDNA sequences representing different haplotypes (including ambiguities) were aligned using Clustal X (HIGGINS *et al.*, 1991) and by eye. All ambiguities and gaps were omitted for further phylogenetic analyses as described in SWOFFORD *et al.* (1996), resulting in 366 characters, and 15 unique haplotypes, which were deposited in the GenBank database. For the reconstruction of phylogenetic relationships we applied Maximum Likelihood (FELSENSTEIN, 1981) analysis and Maximum Parsimony to our DNA data. In order to select the most appropriate DNA model of nucleotide substitution we calculated hierarchic likelihood ratio test statistics using the program Modeltest 3.06 Posada which is implemented in PAUP* 4.0b8 (SWOFFORD, 1998) and tests 56 different substitution models based on a NJ tree using Jukes-Cantor (JUKES & CANTOR, 1969) distances. Details about likelihood ratio tests are given in HUELSENBECK and RANNALA (1997) and in HUELSENBECK and CRANDALL (1997). The HKY85 + Γ model (HASEGAWA *et al.*, 1985) was selected. The rate heterogeneity among sites was assumed to follow a gamma distribution (shape parameter α was ML-estimated) with four categories, each represented by its mean (YANG, 1996). In a further step, the molecular clock hypothesis (i.e., equal rates across all sequences) was tested using the χ^2 approximated likelihood ratio test statistics with OTUs-2 degrees of freedom (*df* = 17 minus 2 = 15) which was not rejected with a *P*-value of 0.09. Therefore, we explored the tree space by 100 heuristic tree searches and by randomising the order of the sequence input using the clock enforcement option in PAUP*. Transitions (ti) were down-weighted relative to transversions (tv) according to the ML-estimated ti/tv ratio, which was about 7:1 in favour of ti (see Results). Tree stability of best trees was evaluated by calculating the consistency index excluding uninformative sites (CIu), and the retention index (RI) (KITCHING *et al.*, 1998).We also applied the genetic distance approach, i.e. the comparison of pairwise genetic distances between haplotypes. We used both the

uncorrected sequence divergence and the corrected divergence using the HKY+Γ model with the parameters estimated from ML.

In all phylogenetic trees *Centruroides vittatus* (Say, 1821) and *C. bani* Armas & Marcano Fondeur, 1987 were preset as outgroups. *C. exilicauda* (Wood, 1863) was tested as an outgroup to *C. limpidus* and *C. infamatus*, and thereafter treated as an ingroup. All tree searches were performed by the branch-and-bound algorithm. Reliability of inferred topologies of trees were assessed using the bootstrap procedure (FELSENSTEIN, 1985), and resampled 1,000 pseudo-replicates except for ML where we resampled 100 pseudo-replicates because of computational time.

DNA SEQUENCE AVAILABILITY

All 14 new sequences were deposited to the GenBank Nucleotide Sequence Database with the following accession numbers: C. infamatus infamatus: CiiZU1 = AF439753; C. infamatus ornatus: CioEC1 = AF439754; CioLC1 = AF439755; CioMO1 = AF439756; CiiTI1 = AF439757; C. limpidus limpidus: CllCH1 = AF439758; CllCH2= AF439759; CllAR1 = AF439760; CllTO1 = AF439761; CllTZ1 = AF439762; CllHU1= AF439763; CllQU1= AF439764;; C. limpidus tecomanus: CltFB1 = AF439765; CltFB2 = AF439766. Three outgroup sequences were published previously (GANTENBEIN et al., 2001): C. exilicauda: CexTE = AJ288639; C. vittatus: CviAR = AJ288643; C. bani: CbaDO = AJ288644

RESULTS

The calculated pairwise genetic distances between the 17 sequences are presented in Table I.

The two tree topologies inferred from Maximum Likelihood and Maximum Parsimony (Fig. 2) are almost congruent except for the phylogenetic postion of the *Centruroides exilicauda* sequence. The analyses revealed four four well supported main clusters, with the following results:

(a) Monophyly of *C. infamatus* was confirmed, with a sole *C. infamatus infamatus* specimen from Zumpimito (*Cii*ZU1) forming a sister clade to all four sampled populations of *C. infamatus ornatus* (*Cio*CO1, *Cio*LC1, *Cio*TZ1, *Cio*TI), which all grouped closely together.

(b) The subspecies *C. limpidus tecomanus* was represented in our samples by two specimens from El Faro de Bucerias (*Clt*FB1, *Clt*FB2) at the coastal zone of southwestern Michoacán. Two close DNA haplotypes from this population did not form a monophyletic group with specimens of *C. limpidus limpidus*, but instead formed an unexpected but highly supported sister clade to *C. infamatus*. Therefore, monophyly of *C. limpidus* was not confirmed.

(c) Within the seven studied specimens of *C. limpidus limpidus*, two deeply divergent, well-supported clades were observed. One clade was represented by four populations at the Balsas Depression area in Michoacán (*Cll*CH1, *Cll*CH2, *Cll*AR1, *Cll*TZ1, *Cll*MO1) which grouped closely together. The second clade recovered within *C. limpidus limpidus* included two specimens from two remote mountainous areas further northwest (Querétaro, *Cll*QU1) and southwest (Guerrero, *Cll*HU1).

DISCUSSION

Centruroides is one of a few extant scorpion genera for which Cenozoic fossils exist, including the Miocene/Oligocene amber of Chiapas, Mexico, with minimum age 22.5 - 26 million years (Ma) (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. GANTENBEIN *et al.* (2001) demonstrated recently for *Centruroides* from Sonora and Baja California, that speciation events in the deserts of northwestern Mexico could be of at least the Miocene age (5-3 Ma). For 16S rRNA gene in *Centruroides*, the evolutionary rate (sequence divergence per time unit; LI, 1997) was estimated to vary from 1.4-2.3% uncorrected sequence divergence (*p*) and from 3-5% HKY + Γ distance per Ma (GANTENBEIN *et al.*, 2001). These estimates of molecular evolutionary rates in mitochondria lie in the range of the clock calibration in primates by BROWN *et al.* (1979) based on mtDNA RFLP data. This study predicts an evolutionary rate of mitochondrial DNA of about 2% uncorrected divergence per Ma and was cited many times for organisms where no fossil data or dated vicariant events are available. More important for invertebrates, however, is the estimated divergence rate for the mitochondrial 16S gene in Hawaiian *Drosophila* species that was again estimated to about 2% uncorrected divergence per Ma (DESALLE *et al.*,1987). Another clock calibration used pairwise genetic divergences of sister taxa of snapping shrimps of the genus *Alpheus* across the Isthmus of Panama and revealed a rate of 1.4% corrected sequence divergence for the mitochondrial cytochrome oxidase I gene (KNOWLTON & WEIGHT, 1998). These rates imply that the divergence rate in the scorpion mitochondrial 16S rRNA gene is comparable to other rates estimated in other invertebrates and even in vertebrates although scorpions have a much longer (mammalian-like) generation time (~2 yrs, for buthids; POLIS and SISSOM, 1990) compared to *Drosophila* sp. (~14 days), and have one of the lowest known metabolism rates in the animal kingdom (HADLEY, 1990). Consequently, we expect a lower rate in scorpion genes than in genes of other animal groups assuming that variation in evolutionary rates across organism groups is mainly due to differences in generation time or metabolism rate (MARTIN & PALUMBI, 1993).

Our pilot molecular data set analysis, 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; new preliminary data support this conclusion (Table II). Moreover, 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). Our pilot DNA data confirm 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. In our opinion 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*, 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. 2). 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 %).

We thus suspect 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 & CALOS, 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 *et al.* (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 (tree not shown) 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 & BEUTELSPACHER, 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). We sampled its populations from 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+Γ divergence (23.7 ± 3.5 %) and uncorrected *p* (10.9 ± 1.0 %) (Fig. 2). 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+Γ rate (11 %) and 3-5 Ma for uncorrected *p* (7 %).

Our results, therefore, suggest that more than one ancient monophyletic lineage (possibly, more than one species) exist within currently accepted *Centruroides limpidus limpidus*. Moreover, our preliminary morphological data indicate that there is some difference between two clades identified by this DNA analysis. The number of the setae on the anterior margin of the carapace differs significantly between the Balsas Depression populations (8-11, usually 8 with 4 on each side, n=29) from that in Querétaro and Huitzuco, Guerrero (6-7 setae, usually 6 with 3 on each side, n=13). 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. Our DNA analyses revealed similar situations in *C. exilicauda* (GANTENBEIN *et al.*, 2001) and in the Alpine scorpion species *Euscorpius germanus* (C.L. Koch) (Euscorpiidae) (GANTENBEIN *et al.*, 2000). Local speciation and endemism are common in the studied area, which forms a biogeographic boundary for many faunal groups between Nearctic and Neotropics (FA & MORALES, 1993). Further, detailed investigations should be done to test our preliminary conclusions: one has to study many more populations from the entire range of *C. limpidus*. 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*.

ACKNOWLEDGEMENTS

We are grateful to Carlos R. Beutelspacher Baigts for his advise on the taxonomic work. We are thankful to the office for Coordination of the Scientific Research, Universidad Michoacana de San Nicolás de Hidalgo, for their support of this work in Mexico; to Erwin P. Miranda L., Rocío J. Moreno B., Marco A. Villaseñor R., Ramon Cancino M., A. Leticia Escalante J., Erasto Hernández C., and Elvia Lemus O. for their support in the field work; and to Mr. Rubén Sosa in Churumuco and Mr. Gelasio Hernandez in Arúa, Huetamo Michoacán, for their hospitality. We thank Dietmar Huber, Joe Bigelow, and Tsunemi Yamashita who donated the specimens used in this study as outgroups; Travis Bailey and Harold Payton for their help in DNA analysis; and Wilson R. Lourenço for his assistance in reviewing and publishing this paper. We gladly acknowledge "kk+w - digital cartography" (Kiel, Germany) whose Generic Mapping Tools (GMT) OMC software (http://www.aquarius.geomar.de/omc/) was used to create Fig.1 map.

As LOURENÇO & SISSOM (2000, P. 124) stated, Mexico has a significant tradition in the study of scorpions, related to the important public health problems due to scorpion stings, and "training of Mexican specialists and/or collaborative research between scientists in Mexico and those of other countries would be highly desirable". We hope that our international collaboration will be one of the steps toward this goal.

REFERENCES

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

BEUTELSPACHER, C.R., 2000. Catálogo de los alacranes de México. Universidad Michoacana de San Nicholás de Hidalgo, Morelia, Michoacán, 175p.

BROWN, W. M., M. GEORGE, JR. & A.C. WILSON, 1979. Rapid evolution of animal mitochondrial DNA. Proc. Nat. Acad. Sc. USA, 44: 1967-1971.

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.

DESALLE, R., T. FREEDMAN, E. M. PRAGER & A. C.WILSON, 1987. Tempo and mode of sequence evolution in mitochondrial DNA of Hawaiian *Drosophila*. J. Mol. Evol., 26: 157-164.

DÍAZ NÁJERA, A., 1964. Alacranes de la República Mexicana. Identificación de ejemplares capturados en 235 localidades. Rev. Inst. Salubr. Enferm. Trop., México, 24: 15-30.

DIAZ NAJERA, A., 1966. Alacranes de la República Mexicana. Clave para identificar especies de *Centrurus*. Rev. Invest. Salud Publ., 26: 109-123.

DIAZ NAJERA, A., 1975. Listas y datos de distribución geográfica de los alacranes de México (Scorpionida). Rev. Inst. Salubr. Enferm. Trop., 35: 1-36.

FA, J. E. & L. M. MORALES, 1993. Patterns of mammalian diversity in Mexico. Pp.319-361, *In*: Ramamoorthy, T. P., R. Bye, A. Lot & J. Fa (Eds.), Biological Diversity of Mexico: Origins and Distribution.. Oxford University Press, New York-Oxford.

FELSENSTEIN, J., 1981. Evolutionary trees from DNA sequences: a maximum likelihood approach. J. Mol. Evol., 17: 368-376.

FELSENSTEIN, J., 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution, 39 (4): 783-791.

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. & G. LOWE, 2000. Family Buthidae. Pp. 54-286, *In*: Fet, V., W. D. Sissom, G. Lowe & M. E. Braunwalder. Catalog of the Scorpions of the World (1758-1998). New York Entomol. Soc., New York.

FET, V. & B. J. PONDEXTER, 1992. Genetic variation of mitochondrial DNA in *Centruroides exilicauda* (Scorpiones: Buthidae) as revealed by the polymerase chain reaction. American Zool., 32: 136A.

FET, V., M. E. SOLEGLAD & M. D. BARKER, 2001. Phylogeny of the "hirsutus" group of the genus *Hadrurus* Thorell, 1876 based on morphology and mitochondrial DNA (Scorpiones: Iuridae). Pp. 139-160, *In* Fet, V. & P. A. Selden (Eds.), Scorpions 2001. In Memoriam Gary A. Polis. British Arachnological Society, Burnham Beeches, Bucks.

GANTENBEIN, B., V. FET, C. R. LARGIADÈR & A. SCHOLL, 1999. First DNA phylogeny of *Euscorpius* Thorell, 1876 (Scorpiones: Euscorpiidae) and its bearing on taxonomy and biogeography of this genus. Biogeographica, 75: 49-65.

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.

HADLEY, N. F., 1990. Environmental physiology. Pp. 321-340, *In*: Polis, G. A. (Ed.), The Biology of scorpions.. Stanford University Press, Stanford, CA.

HASEGAWA, M., K. KISHINO & T. YANO, 1985. Dating the human-ape splitting by a molecular clock of mitochondrial DNA. J. Mol. Evol., 22: 160-174.

HIGGINS, D. G., A. J. BLEASBY & R. FUCHS, 1991. CLUSTAL V: improved software for multiple sequence alignment. Cabios, 8: 189-191.

HOFFMANN, C. C., 1932. Monografías para la entomología médica de México. Monografía Num. 2, Los escorpiones de México. Segunda parte: Buthidae. Anal. Inst. Biol. Univ. Nac. Autón. México, 3 (3): 243-282; (4): 283-361.

HUELSENBECK, J. P. & K. A. CRANDALL, 1997. Phylogeny estimation and hypothesis testing using maximum likelihood. Ann. Rev. Ecol. Syst., 28: 437-466.

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. & B. RANNALA, 1997. Phylogenetic methods come of age: testing hypothesis in an evolutionary context. Science, 276: 227-232.

JUKES, T. H. & C. R. CANTOR, 1969. Evolution of protein molecules. Pp.21-132, *In*: Munro, H. N. (Ed.), Mammalian Protein Metabolism. Academic Press, New York.

KITCHING, I. J., P. L. FOREY, C. J. HUMPHRIES & D. M. WILLIAMS, 1998. Cladistics. 2nd edition. The Systematics Association, Oxford University Press, New York.

KNOWLTON, N. & L. A. WEIGHT, 1998. New dates and new rates for divergence across the Isthmus of Panama. Proc. Roy. Soc., London, B., 265: 2257-2263.

LI, W.-H., 1997. Molecular evolution. Sinauer Associates, Sunderland, MA.

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.

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.

MARTIN, A. P., & S. R. PALUMBI, 1993. Body size, metabolic rate, generation time, and the molecular clock. Proc. Natl. Acad. Sci. USA, 90: 4087-4091.

MAZZOTTI, L. & M. A. BRAVO-BECHERELLE, 1963. Scorpionism in the Mexican Republic. Pp. 119-131, *In*: Keegan, H. L. & W. V. MacFarlane (Eds.), Venomous and poisonous animals and noxious plants of the Pacific region. New York: Pergamon Press.

POLIS, G. A. & W. D. SISSOM, 1990. Life history. Pp. 161-223, *In*: Polis, G. A. (Ed.), The Biology of scorpions. Stanford University Press, Stanford, CA.

PONCE, J. & C. R. BEUTELSPACHER, 2001. Alacranes de Michoacán. Universidad Michoacana de San Nicholás de Hidalgo, Morelia, Michoacán. 103p.

PONCE, J., R. J. MORENO & C. R. BEUTELSPACHER, 1999. Características morfológicas de *Centruroides infamatus* Koch y *C. limpidus* Karsch (Scorpiones: Buthidae), especies comunes en el Estado de Michoacán, México. Memorias del XXXIV Congreso Nacional de Entomología. Soc. Mex. de Entomol: 9-14.

POSADA, D. & K. A. CRANDALL, 1998. Modeltest: testing the model of DNA substitution. Bioinformatics, 14: 817-818.

SANTIAGO-BLAY, J. A. & G. O. POINAR JR., 1993. First scorpion (Buthidae: *Centruroides*) from Mexican amber (Lower Miocene to Upper Oligocene). J. Arachnol., 21: 147-151.

SIMARD, M. & D. D. WATT, 1990. Venoms and toxins. Pp. 414-444, *In*: Polis, G. A. (Ed.), The Biology of scorpions. Stanford University Press, Stanford, CA.

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. LOURENÇO, 1987. The genus *Centruroides* in South America (Scorpiones, Buthidae). J. Arachnol., 15: 11-28.

STAHNKE, H. & M. CALOS, 1977. A key to the species of the genus *Centruroides* Marx (Buthidae, Scorpionida). Entomol. News, 88: 111-120.

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

SWOFFORD, D. L., G. J. OLSEN, P. J. WADDELL & D. M. HILLIS, 1996. Phylogenetic inference. Chapter 4. Pp. 407-425, *In*: Hillis, D. M., C. Moritz & B. K. Mable (Eds.), Molecular Systematics, 2nd Ed. Sinauer Associates, Sunderland, MA.

YANG, Z., 1996. Among-site rate variation and its impact on phylogenetic analyses. Tren. Ecol. Evol., 11: 367-372.

Reçu en octobre 2001
Accepté en novembre 2001

Table I. 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 + Γ 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.

		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	*CitZU1*	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	*CbaDO*	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	*CviAR*	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	*CeXTE*	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	-

Table II. Morphometric comparison of adult *C. infamatus infamatus* (21 females, 8 males) and *C. i. ornatus* (total 10 females, 10 males).

Character	*C. i. infamatus*	*C. i. ornatus*
Pectinal tooth number, female	21– 25 (ave. 22.0)	17 – 22 (ave. 18.5)
Pectinal tooth number, male	22 – 26 (ave. 24.0)	20 – 22 (ave. 21.0)
Movable finger length, female (mm)	5.60 – 6.75 (ave. 6.09)	4.70 – 5.80 (ave. 5.13)
Movable finger length, male (mm)	5.15 – 6.35 (ave. 5.71)	4.25 – 5.70 (ave. 5.22)
Chela width, female (mm)	1.90 – 2.50 (ave. 2.13)	1.55 – 1.95 (ave. 1.72)
Chela width, male (mm)	1.85 – 2.30 (ave. 2.01)	1.80 – 1.90 (ave. 1.83)

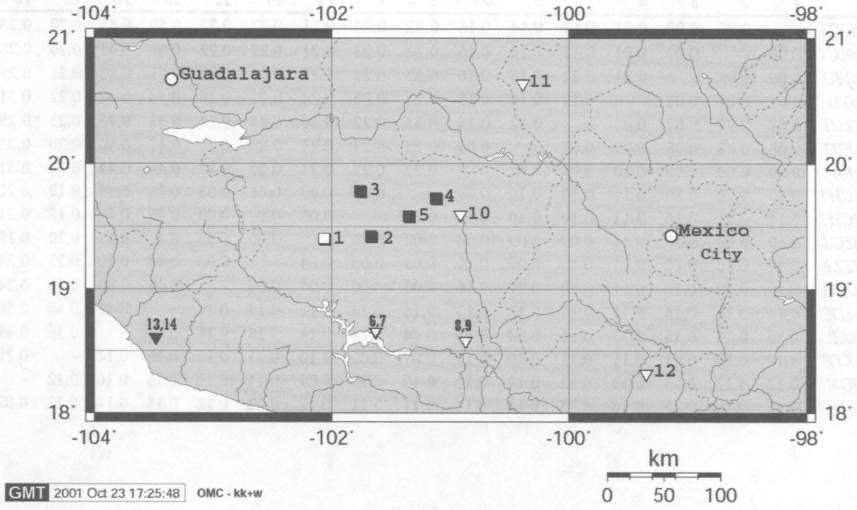

Fig. 1. Map of the samples of *Centruroides* in Central Mexico (see detailed localities data in text). *C. infamatus infamatus*: 1 – *Cii*ZU1, Zumpimito; *C. infamatus ornatus*: 2 – *Cio*CO1, El Cobano; 3 – *Cio*LC1, La Caratacua; 4 - *Cii*MO1, Morelia; 5 – *Cio*TI1, Tiripetio; *C. limpidus limpidus*: 6 – *Cll*CH1 and 7 – *Cll*CH2, Churumuco; 8 – *Cll*TO1, Tzirandaro; 9 – *Cll*AR1, Arua; 10 - *Cll*TZ1, Tzitzio; 11 – *Cll*QU1, Querétaro; 12 – *Cll*HU1, Huitzuco; *C. limpidus tecomanus*: 13 – *Clt*FB1 and 14 – *Clt*FB2, El Faro de Bucerias.

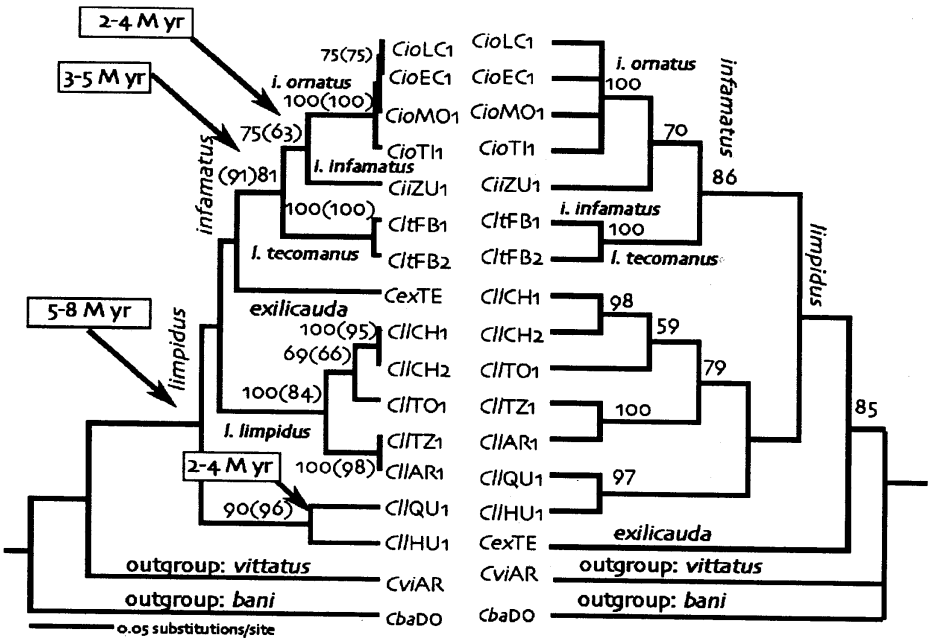

Fig. 2. 16S mtDNA gene genealogy (366 bp) of highly toxic *Centruroides* from central Mexico. Included are the subspecies *C. infamatus infamatus, C. i. ornatus, C. limpidus limpidus,* and *C. l. tecomanus*. Tree topology was revealed by Maximum Likelihood (ML) analysis (left) using the HKY85 + Γ model and by weighted Maximum Parsimony (MP) analysis (right). The parameters for ML were estimated to $\pi_A = 0.34$, $\pi_C = 0.15$, $\pi_G = 0.12$, and $\pi_T = 0.37$, $\alpha = 0.15$, transition (ti)/transversion (tv) ratio = 7.45 ($\kappa = 18.41$), -ln L = 1411.76. For MP, ti's were down-weighted relative to tv's by factor seven. 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). Divergence time estimates are given for ML tree under HKY85 + Γ model (see 'Discussion' for details). Abbreviations for haplotypes are given in Materials & Methods.

Addendum: 16S mtDNA sequence alignment for subspecies *C. infamatus infamatus*, *C. i. ornatus*, *C. limpidus limpidus*, and *C. l. tecomanus*.

```
CioLC1      NNNNNAGNCCACTTTACAGGTCCGAACAGACCTCCTTTTNACTCCTCTTG
CioEC1      NNNNNNNGCCTTTTTACAGGTCNGAACAGACCTCCTTTTNACTCCTCTTG
CioMO1      NNNNNNGGTTTTTTTACAGGTCNGAACAGACCTCCTTTTNACTCCTCTTG
CioTI1      NTNTTGGANACCNNTTNAGGACNNGCCAGACCTCCTTTTNACTCCTCTTG
CiiZU1      NTNTTNNCGCCTNTTTTGGGTCCGAACAGACCCCCTTTTNACTCCTCTTG
CltFB1      NNNCTTNGCCACTTTATGGGTCNGAACAGACCCCCTTTTTACTCCTCTTG
CltFB2      NNNNNGGTTTCTTTTATGNGTCNGAACAGNCCNCCTTTTTACTCCTCTTG
CllCH1      NNNTGGTTTTATTTTATGNGTCNGAACAGNACCCCCTTTTATTCCTCTTG
CllCH2      NNNNNGGTTTATTTTATGNGTCNGAACAGNNCCCCCTTTTATTCCTCTTG
CllTO1      NNNNNNGNCCANTCATGGGTCCNGAACAGANCCCCCTTTTATTCCTCTTG
CllTZ1      NNNNNNNGAGCCCTTNTGGGTCNGNNCAGANCCCCCTTTTATTCCTCTTG
CllAR1      TTTTTNGAAGCCTTTATGGGTCNGCNCAGANCCCCCTTTTATTCCTCTTG
CllQU1      NNNNNNNGCCCCTTTTATGGNGTCGAACAGACCCCCATTTATTCCTCTTG
CllHU1      NNNNNNGGTCTTTTTATCGNCNTCGACAGNACCTTCTTTTATCCCTCTTG
CbaDO       NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNCCCCTATTAATACTCTTG
CviAR       NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNCTTCCTATTATTACTCTTG
CexTE       NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNTCCCCTCCTCTTG

CioLC1      CGGAGTGAAGGAAATCTAATCCAACATCGAGGTCGCAAACGTATTTGTCA
CioEC1      CGGAGTGAAGGAAATCTAATCCAACATCGAGGTCGCAAACGTATTTGTCA
CioMO1      CGGAGTGAAGGAAATCTAATCCAACATCGAGGTCGCAAACGTATTTGTCA
CioTI1      CGGAGTGAAGGAAATCTAATCCAACATCGAGGTCACAAACGTATTTGTCA
CiiZU1      CGGAATAGAGGAAATCTAATCCAACATCGAGGTCGCAAACATATTCGTCA
CltFB1      CGGAATAAAGGAAATCTAATCCAACATCGAGGTCGCAAACATATTTGTCA
CltFB2      CGGAATAAAGGAAATCTAATCCAACATCGAGGTCGCAAACATATTTGTCA
CllCH1      CGGAATAGTGGAAATCTAATCCAACATCGAGGTCGCAAACATATTTGTCA
CllCH2      CGGAATAGTGGAAATCTAATCCAACATCGAGGTCGCAAACATATTTGTCA
CllTO1      CGGAATAGTGGAAATCTAATCCAACATCGAGGTCGCAAACATATTTGTCA
CllTZ1      CGGAATAATGGAAATCTAATCCAACATCGAGGTCGCAAACATATTCGTCA
CllAR1      CGGAATAATGGAAATCTAATCCAACATCGAGGTCGCAAACATATTCGTCA
CllQU1      CGGAATAGTGGAAGCTTAATCCAACATCGAGGTCGCAAACATATTTGTCA
CllHU1      CGGGATAGAGGAAGCTTAATCCAACATCGAGGTCGCAAACATGTTTGTCA
CbaDO       CATAATAAAGGAAAACTAATCCAACATCGAGGTCGCAAACACATTTGTCA
CviAR       CGTAATACAGGGAATTTAATCCAACATCGAGGTCACAAACACGTTTATCA
CexTE       CGGAGTGGAGGAAACTTAATCCAACATCGAGGTCGCAAACATATTTGTCA

CioLC1      ATTTGAACTTTAAAAATACATTACGCTGTTATCCCTAAAGTAACTTATTT
CioEC1      ATTTGAACTTTAAAAATACATTACGCTGTTATCCCTAAAGTAACTTATTT
CioMO1      ATTTGAACTTTAAAAATACATTACGCTGTTATCCCTAAAGTAACTTATTT
CioTI1      ATTTGAACTTTAAAAATACATTACGCTGTTATCCCTAAAGTAACTTATTT
CiiZU1      ATTTGAGCTTTAAGAATATATTACGCTGTTATCCCTAAAGTAACTTGTTT
CltFB1      ATTTGAGCTTTAAAAATATATTACGCTGTTATCCCTAAAGTAACTTATTT
CltFB2      ATTTGAGCTTTAAAAATATATTACGCTGTTATCCCTAAAGTAACTTATTT
CllCH1      ATTTGAACTTTCAAAATATATTACGCTGTTATCCCTAAAGTAACTTATTT
CllCH2      ATTTGGACTTTCAAAATATATTACGCTGTTATCCCTAAAGTAACTTATTT
CllTO1      ATTTGAACTTTCAAAATATATTACGCTGTTATCCCTAAAGTAACTTATTT
CllTZ1      ATTTGAACTTTCAAAATACATTACGCTGTTATCCCTAAAGTAACTTATTT
CllAR1      ATTTGAACTTTCAAAATACATTACGCTGTTATCCCTAAAGTAACTTATTT
CllQU1      ATTTGAGCTTTCAAAATACATTACGCTGTTATCCCTAAAGTAACTTATTT
CllHU1      ATTTGGACTTTCGAAACACATTACGCTGTTATCCCTAAAGTAACTTATTT
CbaDO       ATTTGAACTTTCAAAATATATTACGCTGTTATCCCTAAAGTAACTTATTT
CviAR       ATTTGAACTTTCAAAACATATCACGCTGTTATCCCTAAAGTAACTTATTT
CexTE       ATTTGGACTTTCAAAATATATTACGCTGTTATCCCTAAAGTAACTTATTT

CioLC1      AAGCTTCAAAAATTTTGGGTATCAAAATAATACTATTT-TAATACTTTGA
CioEC1      AAGCTTCAAAAATTTTGGGTATCAAAATAATACTATTT-TAATACTTTGA
CioMO1      AAGCTTCAAAAATTTTGGGTATCAAAATAATACTATTT-TAATACTTTGA
CioTI1      AAGCTTCAAAAATTTTGGGTATCAAAATAGTACTATTT-TGATACTTTGA
```

```
CiiZU1    AAGCTTCAAAAGTTTTGGGTATCAAAATAATGTTATCT-TAATATTATGA
CltFB1    AAAGTTCAAAAATTTTGGGTATTAAAATAATATTATTT-TAATGTTTTAG
CltFB2    AAAGTTCAAAAATTTTGGGTATTAAAATAATATTATTT-TAATGTTTTAG
CllCH1    AAGCTTCAAAATTTTTGGGTATTAAAATGATGATATTT-TAATATTGTAA
CllCH2    AAGCTTCAAAATTTTTGGGTATTAAAATGATGATATTT-TAATATTGTAA
CllTO1    AAGCTTCAAAATTTTTGGGTATTAAAATGATGGTATCT-TAATATTGTAA
CllTZ1    AATCTTCAAAATTTTTGGGTATTAAAATAATGATATTT-TAATATTATAG
CllAR1    AATCTTCAAAATTTTTGGGTATTAAAATAATGATATTT-TAATATTATAG
CllQU1    AAGCTTCAAAATTTTTGGGTATTAAAATAATTCTATTT-TAGCCCTCTAA
CllHU1    AAATTTCAAAATTTTTGGGTATTAAAATAGTTTTATTT-TAGCCCTCTAA
CbaDO     AATATTCAAAAGTTTTGGGTATTAAAACTATCCTATTTATAATAATTTAA
CviAR     AAATTTCAAAATTTTTGGGTATTAAAATAATATTATTTTTAATA-TTTAA
CexTE     AATTTTCAAAAATTTTGGGTATTAAAATAATGTCATTTTTA-TATTTTGA

CioLC1    AAGTGTTT-ATCTTTCCACCGCCCCAGTAAAACACACTTTTAATTTACTA
CioEC1    AAGTGTTT-ATCTTTCCACCGCCCCAGTAAAACACACTTTTAATTTACTA
CioMO1    AAGTGTTT-ATCTTTCCACCGCCCCAGTAAAACACACTTTTAATTTACTA
CioTI1    AAGTGTTT-ATCTTTCCACCGCCCCAGTAAAACACACTTTTAATTTACTA
CiiZU1    AAGTGTTT-ATCTTTCCACCGCCCCAGTAAAACATACTTTTAATTTATTA
CltFB1    AAGTGTTA-GTCTTCTCACCGCCCCAGTGAAACATATTTTTAATTTATTA
CltFB2    AAGTGTTA-GTCTTCTCACCGCCCCAGTGAAACATATTTTTAATTTATTA
CllCH1    AGGTGTTTTATCCTTCCACCGCCCCAGTGAAACATATTTTTAATCTATTA
CllCH2    AGGTGTTTTATCCTTCCACCGCCCCAGTGAAACATATTTTTAATCTATTA
CllTO1    AGGTGTTTCATCCTCCCACCGCCCCAGTGAAACATATTTTTAATTTATTA
CllTZ1    AGGTGTTTAATCCTCTCACCGCCCCAGTGAAACATAATTTTAATTTATTA
CllAR1    AGGTGTTTAATCCTCTCACCGCCCCAGTGAAACATAATTTTAATTTATTA
CllQU1    AAGTGTTT-ATCTTTCCACCGCCCCAGTGAAACATATTCCTAATTTTTTA
CllHU1    AAGTGTTTTATCTTTCCGCCGCCCCAGTGAAACATGTTTTTAATTTTTTA
CbaDO     AAGTGTTTT-TCTTCTTACCGCCCCAGTAAAACATATTCTCAATTTATTA
CviAR     AAGTGTTTT-TCTTTCTGCCGCCCCAGCAAAACATATTCTCAATTTATTA
CexTE     AAGTGTTTTATCTTCCTACCGCCCCAGTAAAACATATTTTCAATCTGTTA

CioLC1    AATTATTATATGTAAAGCTTTATAGGGTCTTCTTGTCTAAAAGAGGTATT
CioEC1    AATTATTATATGTAAAGCTTTATAGGGTCTTCTTGTCTAAAAGAGGTATT
CioMO1    AATTATTATATGTAAAGCTTTATAGGGTCTTCTTGTCTAAAAGAGGTATT
CioTI1    AATTATTATATGTAAAGCTTTATAGGGTCTTCTTGTCTAAAAGAGGTATT
CiiZU1    AATTATTATATGTAAAGCTTTATAGGGTCTTCTTGTCTAAAAGACATATT
CltFB1    AATTATTATATGTAAAGCTTTATAGGGTCTTCTTGTCTAAAAGACATATT
CltFB2    AATTATTATATGTAAAGCTTTATAGGGTCTTCTTGTCTAAAAGACATATT
CllCH1    AATTATTTTATGTAAAGCTTTATAGGGTCTTCTTGTCTAAAAGAAACATT
CllCH2    AATTATTTTATGTAAAGCTTTATAGGGTCTTCTTGTCTAAAAGAAACATT
CllTO1    AATTATTTTATGTAAAGCTTTATAGGGTCTTCTTGTCTAAAAGAAGCATT
CllTZ1    AATTATTTTATGTAAAGCTTTATAGGGTCTTCTTGTCTAAAAGAAGCATT
CllAR1    AATTATTTTATGTAAAGCTTTATAGGGTCTTCTTGTCTAAAAGAAGCATT
CllQU1    GATTATTATATGTAAAGCTTTATAGGGTCTTCTTGTCTGAAAGAAATATT
CllHU1    AATTATTTTATGTAAAGCTTTATAGGGTCTTCTTGTCTAAAAGAGACATT
CbaDO     GATTATTTTATGTAAAGCTTTATAGGGTCTTCTTGTCTAAAAGATGTATT
CviAR     AATTTTCTTATGTAAAGTTTTATAGGGTCTTCTTGTCTAAAAGAGACATT
CexTE     AATTATTTTATGTAAAGCTTTATAGGGTCTTCTTGTCTAAAAGAAACATT

CioLC1    TTAGCCTTTTTACTAAAAGGTAAATTTTGAAGAAAAAAGCTAAGAAAGAA
CioEC1    TTAGCCTTTTTACTAAAAGGTAAATTTTGAAGAAAAAAGCTAAGAAAGAA
CioMO1    TTAGCCTTTTTACTAAAAGGTAAATTTTGAAGAAAAAAGCTAAGAAAGAA
CioTI1    TTAGCCTTTTTACTAAAAGGTAAATTTTGAAGAAAAAAGCTAAGAAAGAA
CiiZU1    TTAGCCTTTTTACTAAAAAGTAAATTTTAAAGAAAAAAGTTAATAAAGAA
CltFB1    TTAGCCTTTTTACTAAAAAGTAAAATTTAAAGAAAAAAGTCNAGAAAGAA
CltFB2    TTAGCCTTTTTACTAAAAAGTAAAATTTAAAGAAAAAAGTCAAGAAAGAA
CllCH1    TTAGCCTTTTTACTAAAAAGTAAAATTTAAAAGAAAAAAGCTAAGAAACAA
CllCH2    TTAGCCTTTTTACTAAAAAGTAAAATTTAAAAGAAAAAGCTAAGAAACAA
CllTO1    TTAGCCTTTTTACTAAAAAGTAAAATTTAAAAGAAAAAACTAAGAAACAA
CllTZ1    TTAGCCTTTTTACTAAAAGGTAAAATTTAAAAGAAAAAATTAAGAAAGAA
CllAR1    TTAGCCTTTTTACTAAAAGGTAAAATTTAAAAGAAAAAATTAAGAAAGAA
```

```
CllQU1      TTAGCCTTTTTACTAAAAAGTAAAATTCAAAAGGAAAAACTAAGAAAGAA
CllHU1      TTAGCCTTTTTACTAAAAAGTAAAATTCAAAAGAAAAAACTAAGAAAGAA
CbaDO       TTAGCCTTTTTACTAAAAAGTAAAGTTAAAAAGTAA-AATTAAAAAAGAG
CviAR       TTAGCCTTCTTACTAAAAGGTGAAATTTGAAATTAATAGTTAAGAAAGAA
CexTE       TTAGCCTTTTTACTAAAAAGTAAAGTTCAAAAGAAAGAGCCAAGAAAGAA

CioLC1      ACTCTCTAGTTTATCCTTTCATTCCAGTCTTAAATTATAAGACTAATGAT
CioEC1      ACTCTCTAGTTTATCCTTTCATTCCAGTCTTAAATTATAAGACTAATGAT
CioMO1      ACTCTCTAGTTTATCCTTTCATTCCAGTCTTAAATTATAAGACTAATGAT
CioTI1      ACTCTCTAGTTTATCCTTTCATTCCAGTCTTAAATTATAAGACTAATGAT
CiiZU1      ACTTTCTAGTTTATCCTTTCATTCCAGTCTTAAATTACAAGACTAATGAT
CltFB1      ACTCTCCAGTTTATCCTTTCATTCCAGTCTTAAATTACAAGACTAATGAT
CltFB2      ACTCTCCAGTTTATCCTTTCATTCCAGTCTTGAATTACAAGACTAATGAT
CllCH1      GCTCTCTAGTTTATCCTTTCATTCCAGTCTTAAATTATAAGACTAATGAT
CllCH2      GCTCTCTAGTTTATCCTTTCATTCCAGTCTTAAATTATAAGACTAATGAT
CllTO1      GCTCTCTAGTTTATCCTTTCATTCCAGTCTTAAATTATAAGACTAATGAT
CllTZ1      GCTCTCTAGTTTACCCTTTCATTCCAGTCTTAAATTATAAGACTAATGAT
CllAR1      GCTCTCTAGTTTACCCTTTCATTCCAGTCTTAAATTATAAGACTAATGAT
CllQU1      ACTTCCTGGTTTATCCCTTCATCCCAGTCTTAAATTATAAGACTAATGAT
CllHU1      ACTTCCTGGTTTACCCCTTCATCCCAGTCTTAAATTATAAGACTAATGAT
CbaDO       CTTTTCCAGTCTATCCTTTCATTCCAGTCTTCAATTATAAGACTAATGAT
CviAR       ATTTTCTAGTTTACCCTTTCATTCCAGTCTTAAATTACAAGACTAATGAT
CexTE       ACTCTCTAGTTTATCCTTTCATTCCAGTCTTAAATTACAAGACTAATGAT

CioLC1      TATGCTACCTTTGCNACANNNNNN
CioEC1      TATGCTACCTTTGCNACACNNNNN
CioMO1      TATGCTCTCTTTTNNGCCACANNN
CioTI1      TATGCTACCTTTGCNACANNNNNN
CiiZU1      TATGCTACCTTTGCNACANNNNNN
CltFB1      TATGCTACCTTTGCNNNNNNNNNN
CltFB2      TATGCTCTTTTTNGNNCACANNNN
CllCH1      TATGCTCTTTTTTGNNCACANNNN
CllCH2      TATGCTCTTTTTTGNGCACANNNN
CllTO1      TATGCTNCCTTTGCNNNNNNNNNN
CllTZ1      TATGCTACCTTTGCNACANNNNNN
CllAR1      TATGCTACCTTTGCNACANCTCNN
CllQU1      TATGCTACCTTTGCNACANNNNNN
CllHU1      TATGCTCTTTTTGNCACANNNNNN
CbaDO       TATGCTACCTTTGCNACACNNCCN
CviAR       TATGCTACCTTTGCNACANNNCTN
CexTE       TATGCTACCTTTGCNACANNNNNN
```