TEXAS TECH UNIVERSITY
Natural Science Research Laboratory

Occasional Papers

Museum of Texas Tech University
Number 381 24 March 2022




Mitochondrial Perspective on Species Identification and Delimitation for Troglobitic *Cicurina* (Arachnida: Araneae: Hahniidae) from Central Texas

Front cover: Photographs of meshweaver spiders and of biologists exploring the spiders' habitat in the karst features of central Texas. Top left) Adult female *Cicurina varians* amid her web; top right) Jean Krejca of Zara Environmental LLC rappelling into a pit in northern Bexar County in search of endangered karst invertebrates; bottom left) Stirling Robertson of TxDOT negotiating a karst feature in Bexar County; and bottom right) an adult female *C. baronia* for which typical troglobitic characteristics of little pigmentation and eyelessness are evident. Photographs courtesy of Jean Krejca.

Mitochondrial Perspective on Species Identification and Delimitation for Troglobitic *Cicurina* (Arachnida: Araneae: Hahniidae) from Central Texas

Preston J. McDonald, Julie A. Parlos, James C. Cokendolpher, Stirling J. Robertson, Jean K. Krejca, Jennifer C. Girón, Robert J. Baker, and Caleb D. Phillips

Abstract

Central Texas is home to a diverse fauna of endemic species found in the karst areas along the Balcones Fault Line, the Edwards Aquifer region, and associated springs. The fauna occurring in Bexar County experience especially high anthropogenic pressure due to urban sprawl and suburban development in and around San Antonio, one of the largest cities in the United States. Among local fauna are numerous troglobitic spider species of the genus *Cicurina* Menge (subgenus *Cicurella* Chamberlin and Ivie, 1940). Many species of this genus are thought to have small distributions and are often represented in museums and datasets by very few specimens. Species taxonomy for this group has been defined primarily by differences in the reproductive anatomy of adult females, which are rare in comparison to the number of immature individuals found in the wild. Prior studies have shown that non-morphologically identifiable immature specimens, in conjunction with adult morphology, aid in illuminating species distributions through incorporation of genetic data. The phylogenetic assessment of the area's diverse species of *Cicurina*, which currently includes three federally listed species (*C. madla* Gertsch, 1992, *C. vespera* Gertsch, 1992, and *C. baronia* Gertsch, 1992), can benefit from a framework upon which to test species boundaries and identify priority areas for further investigations. The species delimitation analyses reported herein provide an updated and expanded understanding of currently recognized species relationships and distributions. Statistical support was obtained for many recognized species, but hypotheses invalidating some species are also proposed. In addition, detections of potentially undescribed species only known from genetics of immature specimens are presented. Finally, significant divergences within federally endangered species were also identified, and priorities for future research are suggested.

Key words: *Cicurina*, endangered species, species delimitation

Supplementary material related to this manuscript is available online at https://doi.org/10.5061/dryad.3bk3j9kmg

Introduction

Cicurina Menge, 1871 (Araneae: Hahniidae; Wheeler et al. 2017) is a large genus of spiders that includes 140 species with a Holarctic distribution (World Spider Catalog 2021). Within *Cicurina*, the subgenus *Cicurella* Chamberlin and Ivie, 1940 includes approximately 80 species, many of which are eyeless and show some degree of troglomorphy (Hedin 2015). Many of these species are endemic to the karst faunal regions (KFRs; geographically and biologically defined areas of subterranean distinction) of the southern US, and *Cicurina* is especially species-rich in central Texas (Gertsch 1992; Paquin and Dupérré 2009). White et al. (2009) hypothesized how the diversification of this group is possibly tied to the temporal progression of the karst geology of the region. Many species face substantial encroachment from human populations

(USFWS 1993). For example, federally listed endangered species of *Cicurina* are found both inside and outside the city limits of San Antonio, Bexar County, Texas (USFWS 2011), which is now the seventh largest city in the United States. Currently recognized endangered species of *Cicurina* found in Bexar County include *C. baronia*, *C. vespera*, and *C. madla* Chamberlin and Ivie, 1940.

In addition to difficulties associated with cave access, which makes studying this group challenging, troglobitic *Cicurina* typify a common invertebrate scenario in which adult specimens are rarely observed. Paquin and Dupérré (2009) reported the ratio of encounter rates with immature, adult female, and adult male specimens was around 100:10:1. Yet, many species are described from the reproductive anatomy of only one or a few female specimens (Cokendolpher 2004). The general absence of museum vouchers for a series of adults for a given species, as well as commonly lacking adult representatives from both sexes, precludes an understanding of intra- and interspecific variability for the morphological characters upon which species designations are based.

Integrating genetic techniques to expand knowledge of *Cicurina* species and their distributions is appealing for multiple reasons. For example, whereas morphological identifications are extremely valuable and have been the basis of species naming, distinguishing species based on morphology is performed by a few skilled experts, and levels of intra- versus interspecific morphological variability are uncertain (Paquin and Dupérré 2009). The commonly encountered immature specimens are of less use for morphological examination, but the same comparative DNA sequence information can be obtained from adults and immatures alike. In general, the perspective that genetic data provides through phylogenetic estimation and sequence divergence is a powerful tool for species identification (Baker and Bradley 2006).

Multiple studies have investigated central Texas *Cicurina* from a genetic perspective. For example, Paquin and Hedin (2004) used mitochondrial COI sequences and a phylogenetic approach to show that examined species of *Cicurina* almost always comprise easily identifiable and discrete clades; for the subset of species examined by Paquin and Hedin (2004), the only instances of genetically indistinguishable conspecifics were for the non-listed species pairs *C. caliga* Cokendolpher and Reddell, 2001, with *C. hoodensis* Cokendolpher and Reddell, 2001, and *C. puentecilla* Gertsch, 1992, with *C. platypus* Cokendolpher, 2004. As noted in Paquin and Hedin (2004), disagreement between morphology and single-locus genetics could reflect some combination of incorrect morphological taxonomy, ancestral reticulation, incomplete lineage sorting, and ongoing hybridization.

Species designations are central to ecological, evolutionary and conservation interpretation (Reid and Carstens 2012). Species are typically recognized based on the combined interpretation of different data types involving thresholds that are often researcher and taxon specific. However, evolutionary model-based methods that provide probabilities about species boundaries have gained favor and are now routinely employed due to their advantages (Pons et al. 2006; Luo et al. 2018). Methods of species delimitation based on single-locus multispecies coalescent models have been noted to over-split species, especially in taxa with significant population structure such as troglobitic spiders (see Hedin 2015 for a review of this problem as applied to cave spiders); however, these methods provide a reasonable framework for developing hypotheses when working with mitochondrial barcode data. The distributions provided by species delimitation statistical approaches describe the odds that divergences among a set of lineages being considered arose via an intraspecific coalescent or an interspecific speciation process, and as such provide a framework for hypotheses about species to be followed up with by more extensive data and investigation (see Hedin 2015 for a delimitation analysis focusing on a specific group of *Cicurina*). For example, model-based species hypotheses in *Cicurina* would statistically justify follow up efforts similar to those presented in Hedin et al. (2018). That study used the joint perspective of a multi-marker approach involving morphology and nuclear and mitochondrial loci to propose the synonymy of federally listed *C. madla* and *C. venii* Gertsch, 1992, into *C. madla*, as well as subsuming non-listed *C. loftini* Cokendolpher, 2004, into listed *C. vespera*.

Previous research focusing on central Texas *Cicurina* has resulted in many unique DNA sequences associated with immature specimens, as well as mor-

phologically defined adult specimens, distributed across the numerous karst features throughout central Texas. The goal of this study was to gain new insight into the diversity of *Cicurina* and to statistically define areas of taxonomic uncertainty. This goal was accomplished by leveraging previously reported and newly generated COI gene sequences into a phylogenetic analysis incorporating species delimitation analyses and clade comparisons.

Materials and Methods

A total of 170 *Cicurina* specimens collected from across central Texas were included in this study (Supplementary Table 1). Of these, 57 were originally reported in Paquin and Hedin (2004), 71 were reported in Hedin et al. (2018), and 43 are being reported for the first time here. Specimens in this study included 72 adults with morphological identifications either from a previous publication (Paquin and Hedin 2004; Hedin et al. 2018) or were identified by one of the authors (JCC) following literature guidelines (Cokendolpher 2004; Paquin and Dupérré 2009). In addition to eyeless *Cicurina* specimens, specimens of *C. pampa* Chamberlin and Ivie, 1940 (six-eyed) and *C. varians* Chamberlin and Ivie, 1940 (eight-eyed) were included. Collection of new specimens for this study was performed by Zara Environmental, LLC and Pape-Dawson Engineers, Inc. following US Fish and Wildlife Service protocols and permits. Specimens were deposited at the Natural Science Research Laboratory of the Museum of Texas Tech University (NSRL) either frozen or in 80% ethanol at room temperature. These specimens were later received as loans from the NSRL for inclusion in this study. One *Dictyna* sp. collected in Lubbock, Texas, also was included as an outgroup.

Genomic DNA was extracted from an entire immature specimen or three to four legs of an adult specimen following the Qiagen DNeasy Blood and Tissue Kit (Qiagen Inc., California) protocol. To increase DNA yield, 40 µL 1M Dithiothreitol and 40 µL Proteinase K were added during the DNA extraction following recommendations by Campos and Gilbert (2012). An additional 40 µL of Proteinase K was added three hours after starting the incubation of the DNA extraction. DNA extractions were incubated at 50°C overnight with agitation.

To provide an improved resource for COI primer assay development, DNA extracted from a single specimen (TK188501, *C. madla*) was used for shotgun sequencing for mitochondrial genome assembly. Approximately 100 ng of DNA was prepared for sequencing using Illumina TruSeq LT Sample Prep kit and protocol (Illumina Inc., San Diego, California) and sequenced on an Illumina MiSeq at RTLGenomics (Lubbock, Texas). Sequencing resulted in 85.9 million 150 bp paired-end reads. Reads were de novo assembled using Velvet (Zerbino 2008). Through comparison of coverages across a range of k-mer sizes, a k-mer size of 131 was used for assembly. The resulting contigs were sorted by length, through which the largest contig was found to be 14,684 bp. This contig was then compared to the NCBI nucleotide collection using BLAST, through which the best match was found to be *Araneae* sp. MT-2014 isolate CL113 mitochondrion, partial genome (accession number KM244672.1, E-value ≈ 0). The mitochondrial genome contig for TK158801 was then annotated using the MITOS server (Bernt et al. 2013). Read depth of coverage and mapping quality scores were recovered using Bowtie2-2.3.1 (Langmead and Salzberg 2012) by mapping raw reads back to the original contig under default settings.

Primers used in this study to amplify COI were either drawn from studies of other spider families and tested for amplification in *Cicurina* or were developed in this study using previously available *Cicurina* COI sequences and the *C. madla* mitochondrial genome assembled herein (Table 1). All loci were amplified and sequenced following thermal profiles with an initial denaturation of 94°C, 120 s; 35 cycles with a denaturation of 94°C for 45 s, annealing at 50°C for 60 s, an extension of 72°C for 75 s; and a final extension of 72°C for 10 min. Fragments were amplified in 25 µL reactions with 10 ng DNA, 0.30 µM of selected primers, 1.25 mM $MgCl_2$, 0.2 mM deoxynucleoside triphosphates, 1X reaction buffer, and 1.25 U Taq DNA polymerase (Promega Corporation, Wisconsin). Amplified fragments were purified using ExoSAP-IT (USB Corporations, Ohio) with a thermoprofile following

Table 1. PCR and sequencing primers. Abbreviations: Amp = used in PCR amplification, Seq = sequencing primer.

Primer	Pair	(F/R) Sequence	Citation	Use
COI1490	A	CWA CAA AYC ATA RRG ATA TTG G	Ledford et al. (2011)	Amp & Seq
C1-N-2776	A	GGA TAA TCA GAA TAN CGN CGA GG	Hedin and Maddison (2001)	Amp & Seq
COI_F1	B	ACT AGT TTA CTG CGA TGG TT	This study	Amp & Seq
COI_R1	B	TCC GGA TAA TCT GAA TAC CG	This study	Amp & Seq
COI_F2	C	ACT TTC AAG GTT AAG AGT GGT	This study	Amp & Seq
COI_R2	C	GGA TAA TCT GAA TAC CGA CGA	This study	Amp & Seq
COI_F3	D	TCA AGG TTA AGA GTG GTA CT	This study	Amp & Seq
COI_R3	D	AAA CAT CCG GAT AAT CTG AA	This study	Amp & Seq
C1-J-2309	NA	TTT ATG CTA TAG TTG GAA TTG G	Hedin and Maddison (2001)	Seq
C1-J-1751	NA	GAG CTC CTG ATA TAG CTT TTC C	Hedin and Maddison (2001)	Seq
C1-J-1718	NA	GGA GGA TTT GGA AAT TGA TTA GTT CC	Chang et al. (2007)	Seq
COI400F	NA	TTT TAC CTG GAT TTG GGA TTG T	This study	Seq
COI750R	NA	ACA AAN CCA ATA CAY CAC ATY	This study	Seq

manufacturer's specifications. If multiple fragments were amplified, reactions were increased to 50 µL and appropriately sized fragments were extracted from 0.8% agarose gels using the Qiagen Gel Extraction Kit (Qiagen Inc., California). Sequences were generated with an ABI Prism 3730 DNA Analyzer with Big Dye version 3.1 Cycle Sequencing chemistry and protocol (Applied Biosystems™, New York). Sequences were aligned using Sequencher 4.9 (Gene Codes Corporation, Michigan), chromatograms for each base call were manually inspected, and absence of premature stop codons was verified.

Bayesian phylogenetic trees were constructed using BEAST 2.6.3 (Bouckaert et al. 2019). Two parallel runs of ten million generations each were conducted. After discarding a 10% burn-in on each run, a total of 18 million states were combined and assessed in Tracer 1.7.1 (Rambaut et al. 2018). The effective sample size for all parameters was greater than 200. Stationarity and convergence were assessed by visual examination of parameter traces. Model averaging was conducted automatically using bModelTest as implemented in BEAST (Bouckaert and Drummond 2017). The analyses were partitioned on codon position, with linked clock and tree models, and unlinked site models. A strict clock rate was estimated with a log-normal prior LogN (1,1.25). A calibrated Yule model was used. The ingroup taxa were constrained to be monophyletic, with the most recent common ancestor node date prior set to LogN (10,0.3) to reflect the assumption that included *Cicurina* diversified no earlier than the formation of the Balcones Fault Zone ca. 15 million years ago (Woodruff and Abbott 1979). Bayesian species delimitation under a generalized mixed Yule coalescent model was conducted using bGMYC in R 3.5.3 (Reid and Carstens 2012). The trees produced by BEAST were annotated and combined using TreeAnnotator 2.6.3 and subsampled at an interval of 100 thousand to produce 181 trees. These trees were input to bGMYC, which was run for 2.2×10^6 generations, with a burn-in of 2×10^5 and a thinning interval of 2×10^4. The threshold parameters were set to t1 = 1 and t2 = 50. A point estimate of the species delimitation was produced using bgmyc.point with conspecificity probability threshold > 0.5 to represent the posterior mean of the species delimitation distribution.

A maximum likelihood phylogenetic tree was constructed using RaxML-NG 1.0.0 (Kozlov et al. 2019). The partitions were set as in the Bayesian tree. Automated model selection was conducted using ModelTest-NG 0.1.6 and the Akaike Information Criterion (Darriba et al. 2020). Bootstrap replicates were computed with automated bootstrapping, which converged after 750 replicates. Maximum likelihood species delimitation under a multi-rate Poisson tree process using mPTP 0.2.4 (Kapli et al. 2019) was also conducted which used the maximum likelihood tree as input. A total of four parallel runs of 6.25×10^6 generations each were run with a minimum branch length of 1.754×10^{-3}. Convergence was assessed by examination of traces of log-likelihood scores. The best-scoring model likelihood was approximately 567, while the null model representing a single species spanning all tips scored approximately 411.

All specimen-level pairwise comparisons of number of raw base differences per site (p-distance) were calculated in MEGA (Kumar et al. 2016). Intraclade divergence values were calculated and reported as the mean of p-distances of relevant specimens. Clade assignments for divergence calculations were based on the results of the species delimitation analyses described above. Tree and species delimitation visualizations were produced in R using ggtree 2.1.2 (Yu et al. 2017). Geographic maps were produced using QGIS 3.4.13, with additional karst zone and roadway data (Veni 2002). Sequence data generated as part of this study are available through Genbank accession numbers KU552265–KU552332.

Results

De novo assembly of shotgun sequence data for specimen TK188501 recovered a 14,648 bp mitochondrial genome assembly. The mean coverage depth across this genome was 588 (1st quartile = 511, 3rd quartile = 686). The consensus assembly and annotation for this genome is available on Dryad at https://doi.org/10.5061/dryad.3bk3j9kmg. All primers used in this study, including new COI primers based on the mitochondrial genome reference, are available in Table 1.

For a total of 171 individuals the consensus COI alignment was 1,119 bp. The average number of bp sequenced per individual was 993 bp with a standard deviation of 85 bp. Thirty-six percent of the alignment (401 bp) was found to be variable, with 87% (350 bp) of variable positions found to be parsimony informative.

Both Bayesian and Maximum Likelihood phylogenetic reconstructions tended to recover clades exclusive to previously identified species (Fig. 1, Fig. 2, Supplementary Fig. 1, Supplementary Fig. 2). This included recovery of clades for each of *C. madla*, *C. vibora* Chamberlin and Ivie, 1940, *C. travisae* Chamberlin and Ivie, 1940, *C. selecta* Chamberlin and Ivie, 1940, *C. vespera*, *C. baronia*, *C. mirifica* Chamberlin and Ivie, 1940, and *C. varians* Chamberlin and Ivie, 1940. Bayesian posterior probabilities and Maximum Likelihood bootstrap support values were typically high (i.e., greater that 95%) for most of these clades. Instances in which morphologically based species identifications and mitochondrial genetic relationships disagreed included the polyphyletic relationships involving *C. caliga* within *C. hoodensis* Cokendolpher and Reddell, 2001, *C. platypus* Chamberlin and Ivie, 1940, within *C. puentecilla* Gertsch, 1992, as well as *C. neovespera* Cokendolpher, 2004, within *C. bullis* and *C. bullis* Cokendolpher, 2004, within *C. brunsi* Cokendolpher, 2004.

Support values for nodes uniting aforementioned species level clades commonly had low support values; whereas one or the other phylogenetic estimation approaches would at times report high support, the other tended not to. Also, the relationship of eyeless *Cicurina* to the eyed-species (*C. pampa* and *C. varians*) was poorly resolved. In addition to relationships for clades involving morphologically identified specimens, other well supported but unnamed clades were recovered. This included 1) a clade sister to *C. madla* formed by TK190947 (Fern Cave, Medina County) and TK190936 (King Toad Cave, Bexar County) referred to as *C. cf. madla* in Hedin et al. (2018), 2) a clade sister to *C.*

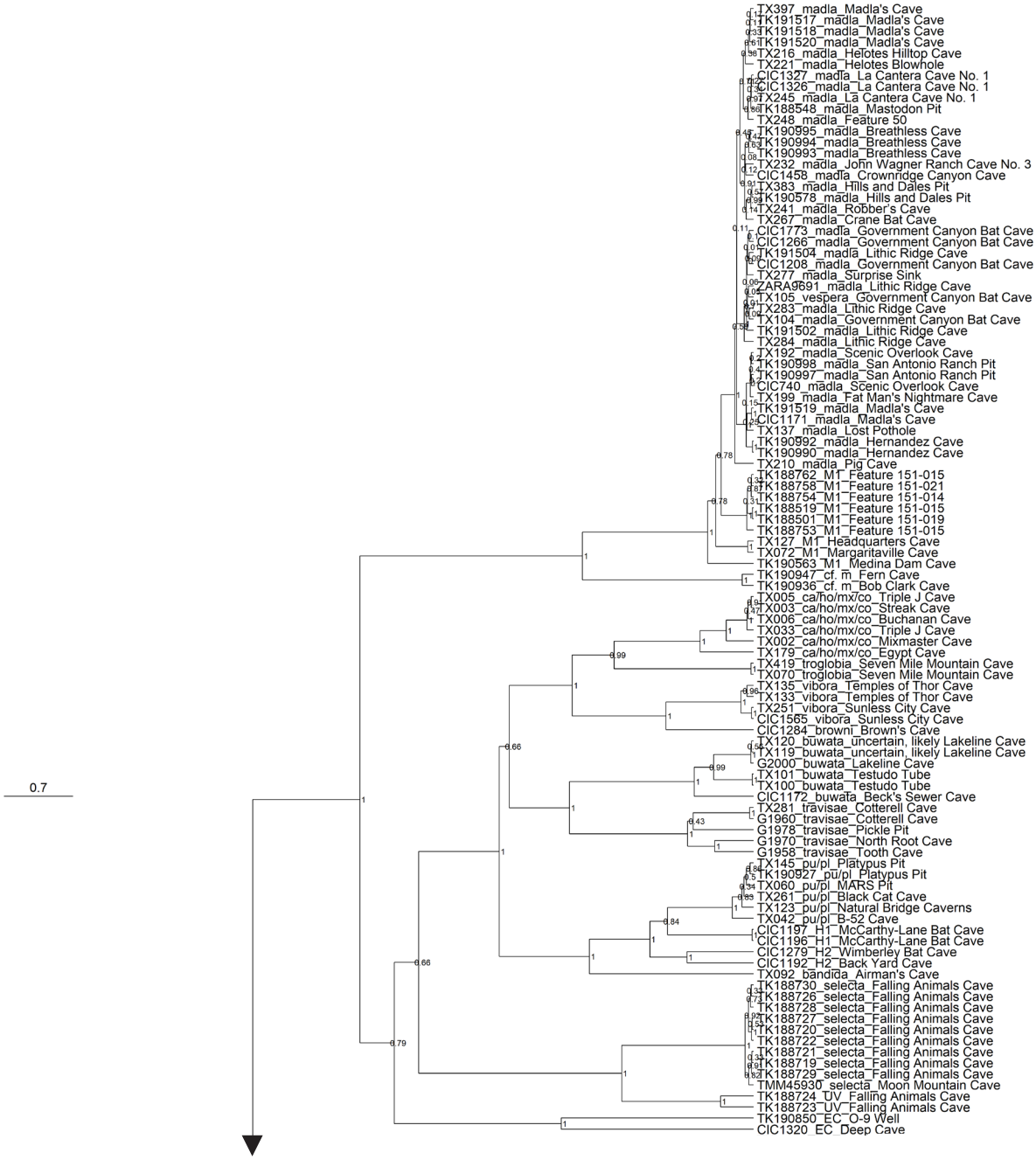

Figure 1. Bayesian phylogenetic reconstruction displayed as an ultrametric tree of eyeless *Cicurina* from central Texas based on the mitochondrial COI locus. Tip names include specimen accession numbers, posterior species/clade designations, and locality names. Node values are posterior support values. The scale bar is in units of substitutions/site.

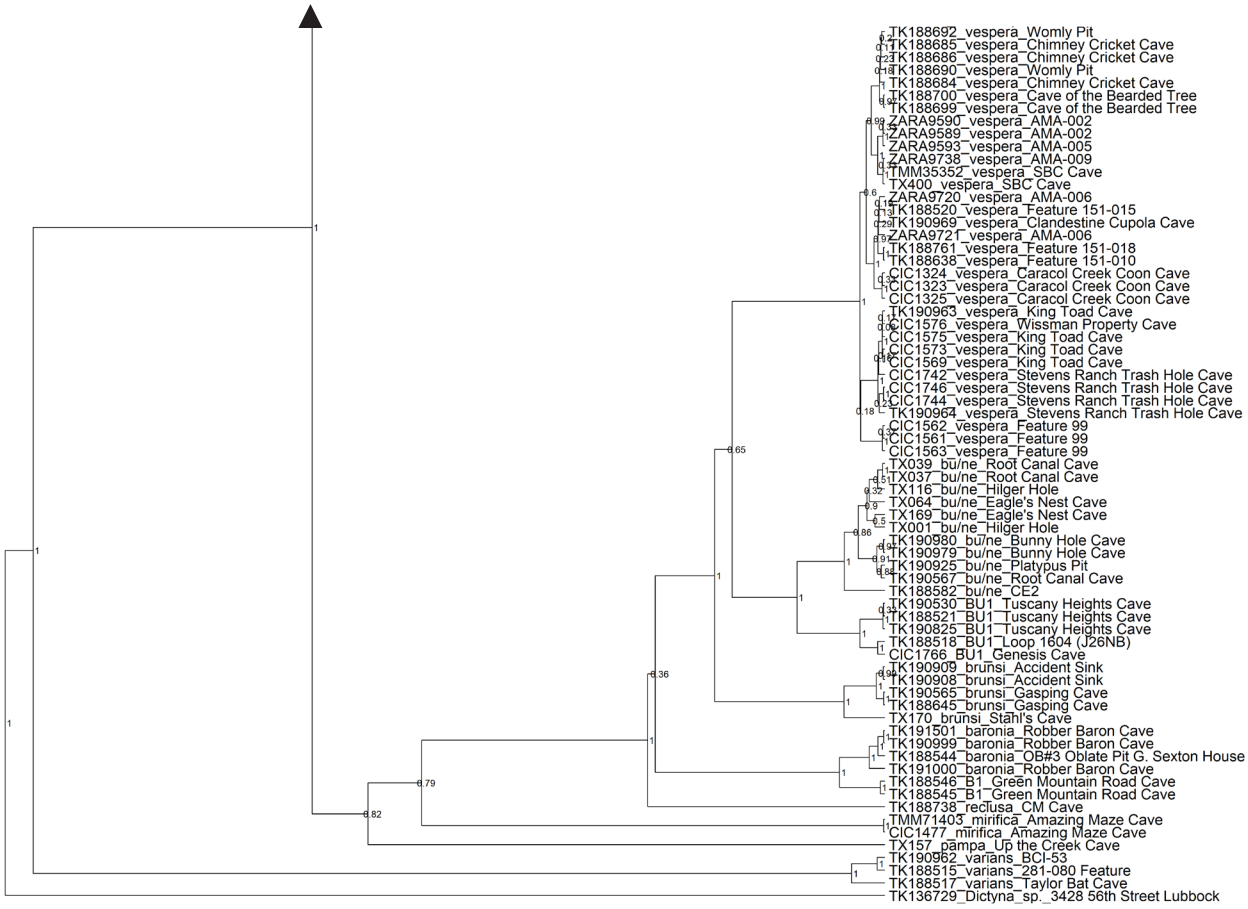

Figure 1. (cont.)

Figure 2. Maximum likelihood phylogenetic reconstruction of eyeless *Cicurina* from central Texas based on the mitochondrial COI locus. Tip names include specimen accession numbers, posterior species/clade designations, and locality names. Node values are bootstrap support values. The scale bar is in units of substitutions/site.

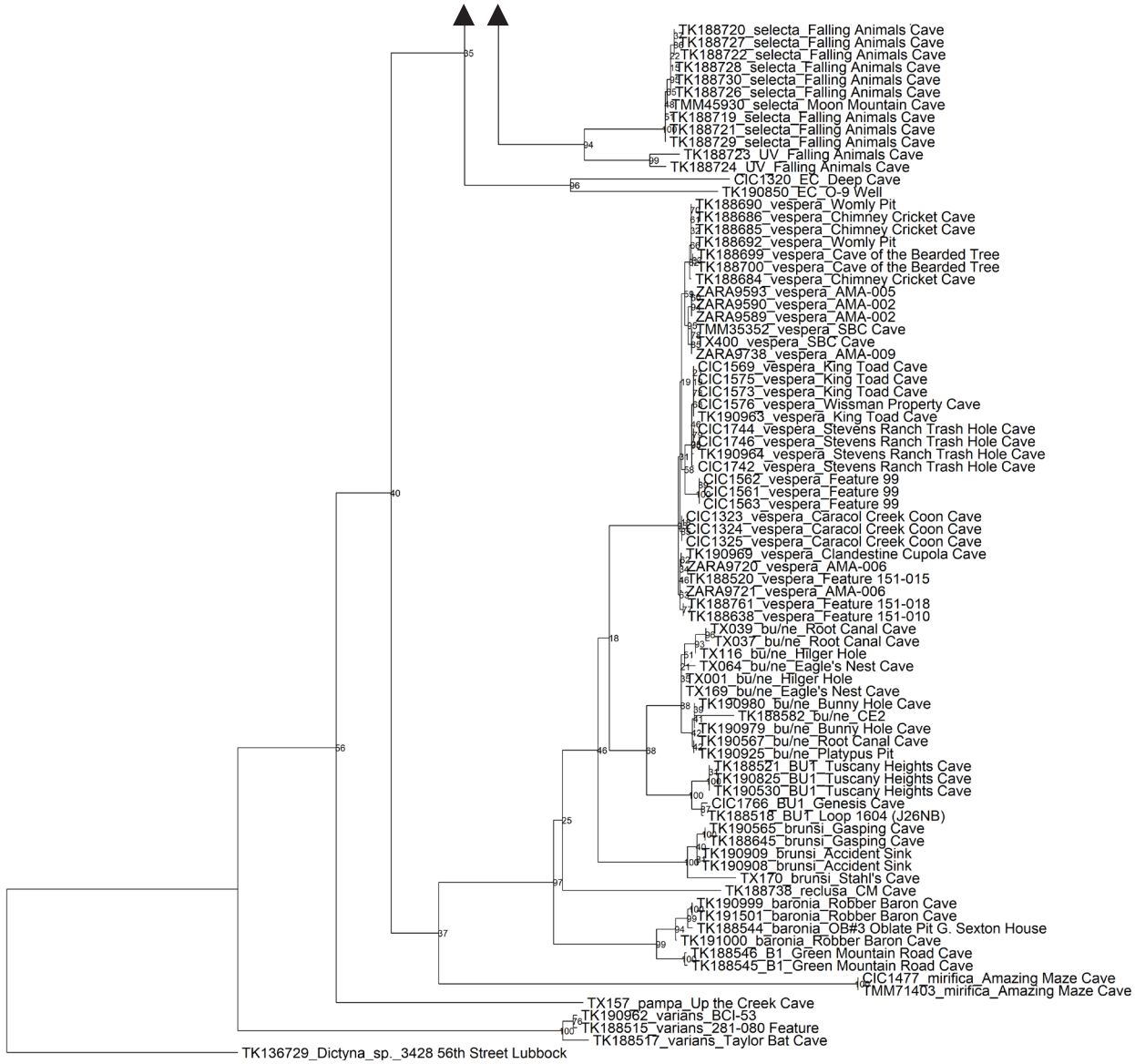

Figure 2. (cont.)

selecta formed by TK188723 and TK188724 (both from Falling Animals Cave, Uvalde County) referred herein as UV clade, (3) a clade formed by TK190850 (0-9 Well, Crocket County) and CIC1320 (Deep Cave, Edwards County) referred herein as EC Clade, (4) a clade formed by CIC1196 and CIC1197 (both McCarthy-Lane Bat Cave, Hays County) referred herein as H1 clade, and (5) a clade formed by CIC1192 (Backyard Cave, Hays County) and CIC1279 (Wimberley Cave, Hays County) referred herein as H2 clade (Fig. 3, Fig. 4). In each instance, the best estimated node uniting the clade to the rest of the phylogeny was generally at least as deep as those for clades of currently recognized species.

Species delimitation analyses using Bayesian and Maximum Likelihood approaches were largely in agreement (Fig. 3). Both methods support *C. cf. madla*, *C. troglobia* Cokendolpher, 2004, *C. vibora*, *C. travisae*, *C. selecta*, *C. vespera*, *C. mirifica*, and *C. varians*. Both delimitation methods also suggest synonymy of some currently recognized species. This included support for synonymy of *C. caliga*, *C. hoodensis*, *C. mixmaster* Cokendolpher and Reddell, 2001, and I Gertsch, 1992. Synonymy of *C. neovespera* and *C. bullis* was also supported, but morphologically identified *C. bullis* specimens also were placed within a clade containing *C. brunsi*. One or both methods also reported species level divergences for all unnamed clades described above.

Notable areas of disagreement between delimitation methods concern federally listed species. The first occurs in *C. madla* in which the Bayesian method suggests a clade inclusive of several specimens morphologically identified as *C. madla* and referred to herein as M1 clade. Although the Maximum Likelihood method did not recover this clade, this potential clade includes specimens from new localities in the west (Medina Dam Cave, Medina County and Margaritaville Cave, Bandera County) and to the northeast of the main *C. madla* distribution in Bexar County. For *C. baronia*, Maximum Likelihood delimitation also suggested a species level divergence at the deepest node for this group. In this instance, the divergence separates two immature specimens (TK188545 and TK188546; labeled as B1 clade) from Green Mountain Road Cave from the other four *C. baronia* specimens (TK188544 from Oblate Pit (G. Sexton's house) and TK190999, TK191000 and TK191501 from Robber Baron Cave).

The eight-eyed (*C. varians*) and six-eyed (*C. pampa*) species were on average 14.9% and 13.1% divergent from all eyeless *Cicurina*, respectively. An average clade divergence of 11.8% was found among clades supported by both species delimitation boundaries reported above. The smallest divergence of 5% was observed between the clade containing both *C. neovespera* and *C. bullis* (bu/ne clade) and a second *C. bullis* clade (BU1 clade). The largest divergence was observed between *C. madla* and *C. mirifica* (Supplementary Table 2). The intra-clade divergence values ranged from zero (*C. mirifica* and H1 clades) to 9.5% (EC clade; Table 2). Both *C. madla* and *C. baronia* were found to have bimodal pairwise genetic divergence distributions (Fig. 5). For both species, the larger divergence values were from comparisons between individuals distributed across the deepest node in each species' clade.

DISCUSSION

Paquin and Dupérré (2009) describe *Cicurina* of central Texas as a taxonomically diverse but poorly understood and understudied group. Much of the current knowledge gap and lack of specimen representation for this group reflects difficulties in physically accessing the karst habitats in central Texas, and the widespread and fragmented nature of the karst system is most certainly fundamental to the diversification of this group (White et al. 2009). Somewhat paradoxically, the species that are most abundantly represented by voucher specimens are also federally listed as Endangered (*C. madla*, *C. vespera*), situations arising probably more from conservation-directed sampling efforts (*C. madla*) and recent taxonomic revisions (*C. vespera*) rather than true species abundances, which are still unknown for this group. Amid the large number of standing uncertainties for this clade the motivation for current work was to identify taxonomic and phylogenetic subsets

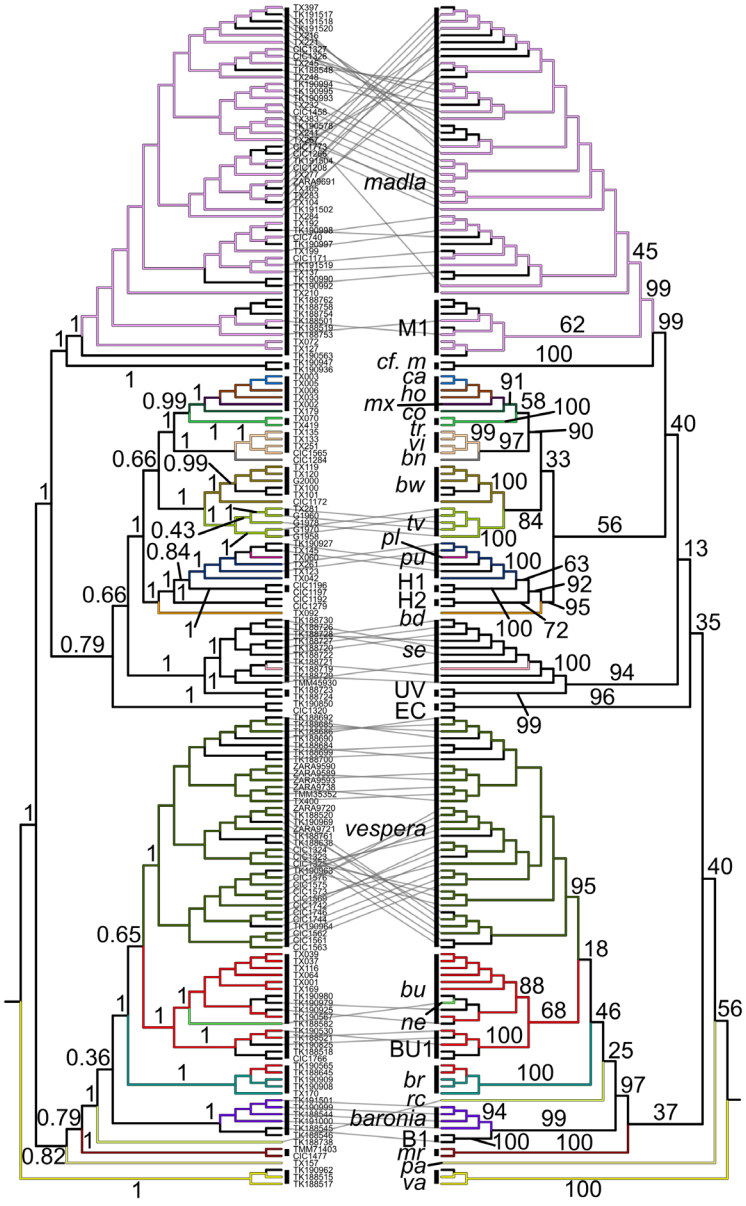

Figure 3. Species delimitations based on mitochondrial COI. On the left is a Bayesian cladogram (same topology as in Figure 1) with posterior probability node support values. On the right is a Maximum Likelihood cladogram (same topology as in Figure 2) with bootstrap node support values. Vertical black bars along tree tips indicate proposed species delimitations from Bayesian (left) and Maximum Likelihood (right) approaches. Branch colors indicate prior species identifications (Paquin and Hedin 2004 or Hedin et al. 2018); black branches lack prior identifications. Colors and abbreviation as in Figure 4. Notation: *C. browni* = bn, *C. brunsi* = br, *C. bullis* = bu, *C. buwata* = bw, *C. caliga* = ca, *C. coryelli* = co, *C. ezelli* = ez, *C. hoodensis* = ho, *C. mirifica* = mr, *C. travisae* = tv, *C. mixmaster* = mx, *C. neovespera* = ne, *C. platypus* = pl, *C. puentecilla* = pu, *C. reclusa* = rc, *C. selecta* = se, *C. troglobia* = tr, *C. varians* = va, *C. vibora* = vi, *C. cf. madla* = cf.m, *C. bandida* = bd, Hays County, McCarthy-Lane Bat Cave locality = H1, Hays County, Wimberley Cave and Back Yard Cave localities = H2, Uvalde County = UV, Edwards and Crocket Counties = EC, M1 = referenced *C. madla* clade, BU1 = referenced clade containing morphologically identified *C. bullis*.

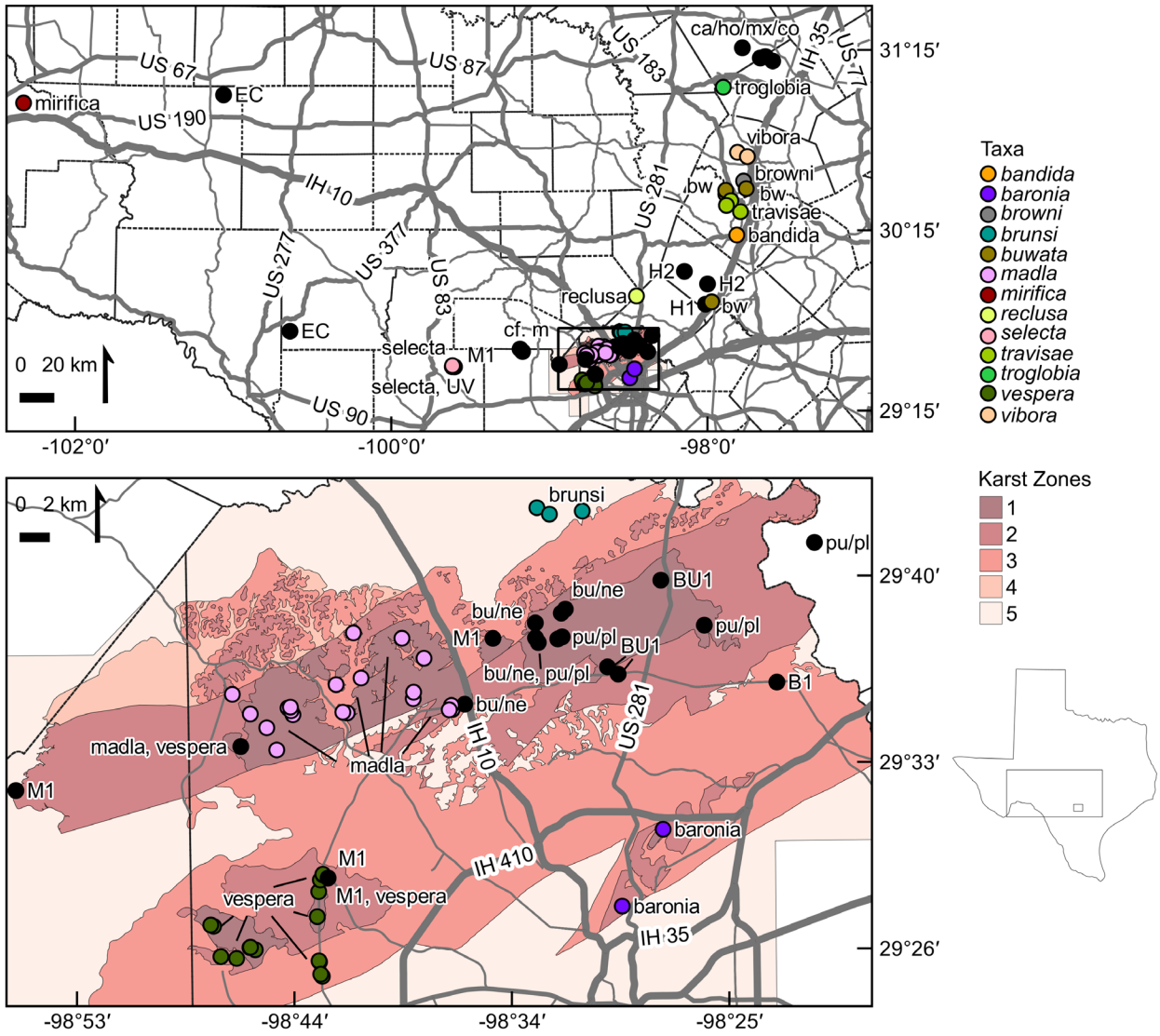

Figure 4. Detailed map of study area with the lower right image of Texas detailing the locations of the two inset maps. The larger region depicted in the top figure includes the total distribution of specimens included in this study. The lower panel focuses on the area of Bexar County where many species included in this study, including federally listed *C. madla*, *C. vespera*, and *C. baronia* are known to occur. Each dot represents a collection locality. Colored (not black) dots denote the single species that occurs at a given location. Black dots are used for instances of sympatry and clade names separated by commas lists sympatric lineages. Instances in which the appropriate species name is in question are denoted by species abbreviations separated by a forward slash. Figure legend provides color key to species/clades and matches branch coloring included in Figure 3. Karst Zone distributions are also shaded. Abbreviations follow those used in Figure 3.

Table 2. Intra-clade average percent divergence for the COI alignment. Species with no value reported were represented by single individuals. Notations: *cf. madla* = clade sister to *C. madla*, naming following Hedin et al. (2018); ccmh = clade containing specimens morphologically identified as either *C. coryelli*, *C. caliga*, *C. mixmaster*, or *C. hoodensis*; pu/pl = clade containing specimens morphologically identified as either *C. platypus* or *C. puentecilla*; bu/ne clade containing specimens morphologically identified as either *C. bullis* or *C. neovespera*; H1 = Hays County clade containing McCarthy-Lane Bat Cave locality; H2 = Hays County clade containing Wimberley Cave and Back Yard Cave localities; UV = Uvalde County Clade; EC = clade containing Edwards and Crocket County specimens.

Clade	Mean Divergence
EC	9.5
H2	4.7
travisae	3.5
buwata	2.5
UV	2.4
baronia	1.9
ca/ha/mx/co	1.9
brunsi	1.4
varians	1.4
madla	1.3
bu/ne	1.2
BU1	1.1
cf. madla	1.0
vespera	0.9
pu/pl	0.8
vibora	0.6
selecta	0.3
troglobia	0.1
H1	0.0
mirifica	0.0
browni	--
reclusa	--
bandida	--
pampa	--

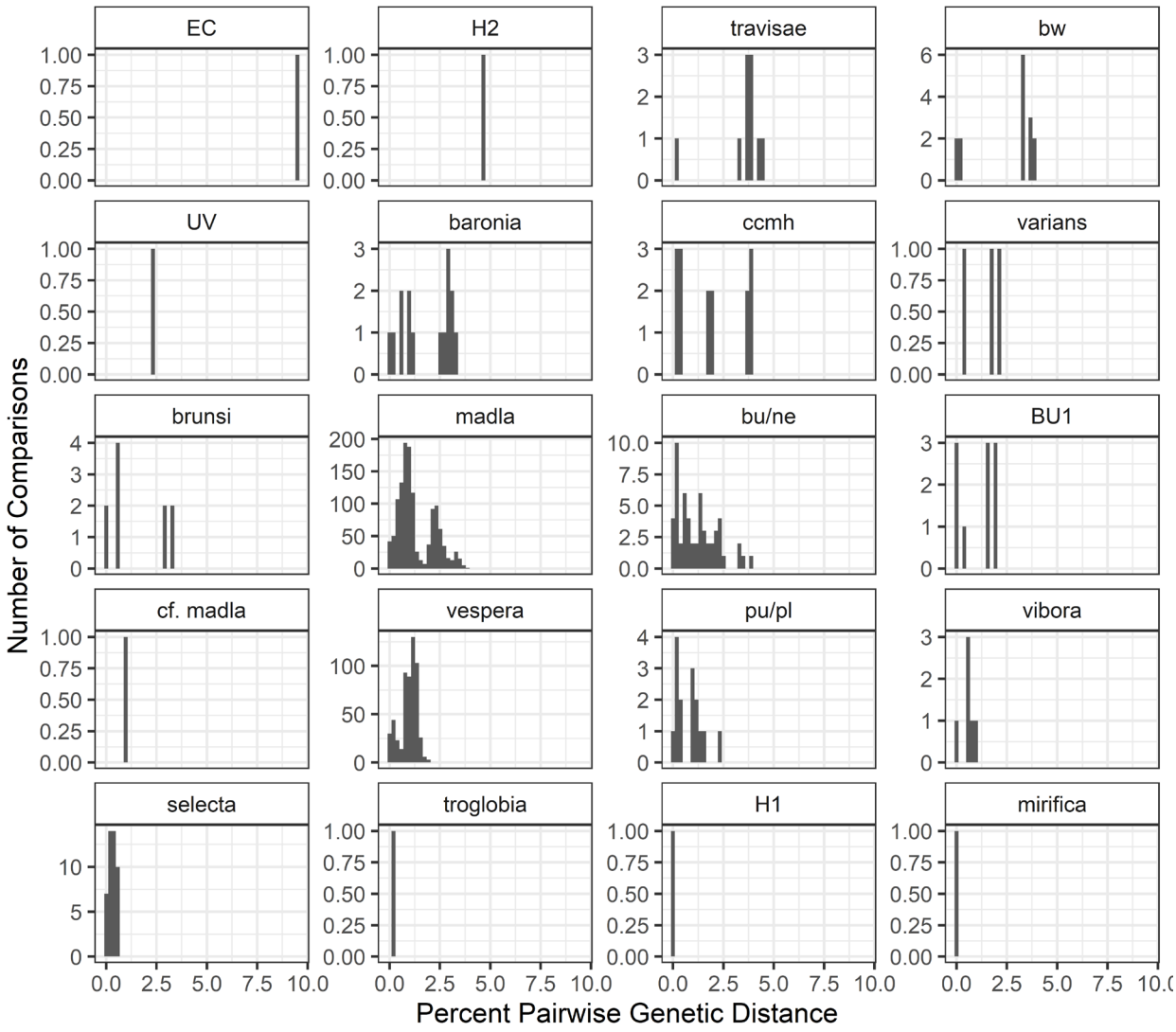

Figure 5. Specimen pairwise raw percent sequence divergence distributions at COI organized to show intra-clade divergences identified through species delimitation analyses. Plots are organized from top-left to bottom-right by decreasing mean intra-clade percent divergence. Abbreviations follow those used in Table 2.

that, based on available information, may be given priority in future efforts to better understand and plan the conservation management of this group.

With 51 samples assigned to its clade *C. madla* is the best represented species in this study. Previously, interpretation of locality information led to the suggestion that *C. madla* and *C. vespera* were highly similar (Paquin and Hedin 2004; Hedin 2015). However, additional follow up work comparing genitalic characteristics indicate this was an error and very likely not the case (Hedin et al. 2018). Also, Hedin et al. (2018) synonymized federally listed *C. venii* into *C. madla*. Although the current study cannot provide additional genetic insight into this relationship, this revision is supported by best available data outlined in Hedin et al. (2018). Other important observations regarding *C. madla* are the identities of specimens forming a poorly supported relationship to the rest of *C. madla*. This includes TK190563, an immature specimen from Medina County; perhaps because of long branch attraction this specimen's relationship to the rest of *C. madla* is not congruent between phylogenies presented here. If TK190563 is to be considered *C. madla*, as only one species delimitation analysis supports, then this specimen represents a new locality (Medina Dam Cave) for *C. madla*. Another instance of data deficiency requiring future work involves the status of the clade formed by TK190947 and TK190936 (referred to herein as *C. cf. madla* following Hedin et al. 2018); both delimitation methods support this clade as distinct from *C. madla*. Moreover, the average genetic divergence of this clade to *C. madla* clade was found to be 9%, which is greater than that observed between other species of *Cicurina* included here. As discussed in Hedin et al. (2018), the current lack of morphological data for this clade results in its specific status being ambiguous.

Among current federally listed species of *Cicurina*, *C. baronia* is known from the fewest localities and has the smallest specimen representation. As for *Cicurina* in general, it is not clear if this reflects concentration of past field effort or true species abundance. In addition to the type locality of *C. baronia* (Robber Barron Cave, Bexar County), this species has also recently been found to occur at Oblate Pit (G. Sexton's house), Bexar County (Hedin et al. 2018). That study also identified an immature specimen (TK188545) from Green Mountain Road Cave, Bexar County, to be mitochondrially similar to other specimens previously identified as *C. baronia*. Here, a second immature specimen (TK188546) of this mitochondrial lineage from Green Mountain Road Cave is reported. Notably, the clade formed by these specimens was proposed through the Maximum Likelihood delimitation analysis to represent a separate species, and these specimens differ from other *C. baronia* by an average of 3%. This divergence is most accurately described as subspecific divergence within *C. baronia*. However somewhat more complicating, nuclear data presented in Hedin et al. (2018) for the one specimen known at the time (TK188545) reported a sister relationship to TK190530, which is located in the BU1 mitochondrial clade and was morphologically identified as *C. bullis*. Although the deeper relationships among species in the COI phylogeny are at locations poorly supported, both the *C. baronia* and BU1 clades are well supported and unlikely to form a sister relationship. Joint consideration of available information suggests that phylogenetic estimation error, ancestral reticulation, incomplete lineage sorting, or ongoing hybridization are the basis for the observed phylogenetic patterns. Even so, the close mitochondrial relationship of TK188545 and TK188546 to specimens morphologically identified as *C. baronia* support they are currently best described as *C. baronia* and Green Mountain Road Cave as a new occurrence locality for *C. baronia*. Clearly, *C. baronia* needs to be better understood through additional field work and comparative research.

The polyphyletic relationship involving *C. caliga* and *C. hoodensis* recovered here is consistent with the relationship reported in Paquin and Hedin (2004). Cokendolpher and Reddell (2001) reported *C. caliga* to be found in sympatry with *C. hoodensis* in Buchanan Cave, the type locality of *C. hoodensis*. As discussed in other works (Paquin and Hedin 2004; Paquin and Dupérré 2009), observed sympatry and close genetic relationships in this group suggest synonymy may be warranted. However, in the current work both species delimitation analyses provide an expanded hypothesis about species synonymy involving *C. caliga*, *C. hoodensis*, *C. mixmaster*, and *C. coryelli*. The largest divergences found within this group are about the same as the largest values observed within currently recognized *C. madla* and *C. bullis*, two clades with reasonable sample sizes. The three discrete peaks in the pairwise genetic divergence distribution for this

clade is partially congruent with morphology-based species identifications; the largest pairwise divergences for this clade correspond to comparisons to the *C. coryelli* specimen, the second largest divergences to the *C. mixmaster* specimen, and the smallest divergences occurring between *C. caliga* and *C. hoodensis* specimens. Unfortunately, no nuclear data for any of these specimens is currently available but would be useful for making a more informed decision regarding whether observed variability is best described as intra- or interspecific.

The need to consider synonymizing *C. puentecilla* and *C. platypus* has been previously discussed (Paquin and Hedin 2004; Paquin and Dupérré 2009). Here, the *C. platypus* specimen (TX060) is a morphologically identified female from Mars Pit originally reported in Paquin and Hedin (2004). TX060 is sister to two immature specimens (TX145 and TK190927) presumed as *C. platypus* based on their genetic relationship and on their collection from the *C. platypus* type locality, Platypus Pit. However, the multiple instances of sympatry accumulating for *Cicurina* indicate locality data alone is insufficient for identification. The genitalic morphology of *C. platypus* and *C. puentecilla* have been described as very similar (Cokendolpher 2004; Paquin and Dupérré 2009). All available data suggest synonymy is warranted, however a nuclear genetic perspective about these species is currently absent.

Mitochondrial phylogenetic relationships and morphological identifications are in general disagreement among *C. bullis*, *C. neovespera*, and *C. brunsi*. Both delimitation analyses support conspecificity of *C. neovespera* (represented by single specimen, TK188582, identified in Hedin et al. 2018) and several morphologically identified *C. bullis*. However, two additional specimens are identified as *C. bullis* for the first time here (TK190565 and TK188645) but are within a clade otherwise consisting of *C. brunsi*. Given this confusion, as well as the nuclear similarity of one *C. baronia* (TK188545) to one *C. bullis* specimen (TK190530) reported in Hedin et al. (2018), more work on this group is also needed.

Several other clades were identified through delimitation analyses that represent either previously undescribed species or are unidentified immature specimens belonging to previously named species. One such clade, referred to herein as H1, is represented by two specimens originally reported in Hedin et al. (2018) (CIC1196 and CIC1197) and were collected from McCarthy-Lane Bat Cave, Hays County, a locality not associated with any named specimen examined in Paquin and Dupérré (2009). Also, in Hays County the situation is similar for clade H2 that is formed by CIC1192 and CIC1279, which are reported as collected at Backyard Cave and Wimberley Cave, respectively. The divergence between these specimens is close to 5%, and neither locality is referenced in Paquin and Dupérré (2009) as a known locality for a named species. A third clade was formed by TK188723 and TK188724, which were collected from Falling Animals Cave in Uvalde County. These specimens are more than 7% divergent from their sister clade (*C. selecta*) and this locality is also not referenced in Paquin and Dupérré (2009). Finally, the EC clade formed by two specimens (TK190850 from O-9 Well, Crocket County and CIC1320 from Deep Cave, Edwards County) are at least 11% divergent from all other *Cicurina* species and these localities also do not occur in Paquin and Dupérré (2009). Moreover, these two EC clade specimens are 9.5% divergent from each other so may each be distinct species as well. Several thousand potentially habitable caves and a network of pores and fissures exist in central Texas. The frequency at which unnamed lineages are identified despite relatively modest sample coverage, in combination with a number of potential synonymies, suggests the true species diversity of central Texas *Cicurina* may be different than currently recognized.

Findings discussed above suggest that many areas are in need of additional collection and laboratory study. Given its federal listing as endangered and small number of known occurrences, one immediate conservation priority should focus on the distribution and diversity of *C. baronia*. Also, the sequence divergence among *C. madla* and sister lineages needs further study to determine the best taxonomic treatment and distribution record for *C. madla*. In general, the current and any future federal listings of *Cicurina* would be much better informed by genetically based estimates of effective population sizes for each species in conjunction with more knowledge about species distributions.

Acknowledgments

Preston J. McDonald and Julie A. Parlos contributed equally to first authorship of this work. This work was funded by Texas Department of Transportation under Contract Number 57-3XXIA003. Thanks to Cindy Watson and Jenny Wilson of the US Fish and Wildlife Service for issuing permits so that known endangered species habitats could be recollected. We thank Clover Clamons (now with TxDOT Environmental Affairs Division) and Wendy Dickerson with Zara Environmental, LLC, for their contributions to data curation. We thank Zara Environmental and Pape-Dawson Engineers for specimen collection. The Museum of Texas Tech University Invertebrate Zoology Collection provided loans and use of voucher specimens and tissues in this project. Sequences were generated at the Center for Biotechnology & Genomics, Texas Tech University. Special thanks to Joel Ledford and Marshal Hedin for careful and constructive reviews of this manuscript.

Literature Cited

Baker, R. J., and R. D. Bradley. 2006. Speciation in mammals and the Genetic Species Concept. Journal of Mammalogy 87:643–662.

Bernt, M., et al. 2013. MITOS: Improved de novo Metazoan Mitochondrial Genome Annotation. Molecular Phylogenetics and Evolution 69:313–319.

Bouckaert, R., et al. 2019. BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. PLOS Computational Biology 15(4):e1006650.

Bouckaert, R. R., and A. J. Drummond. 2017. bModelTest: Bayesian phylogenetic site model averaging and model comparison. BMC Evolutionary Biology 17, article 42. doi:10.1186/s12862-017-0890-6.

Campos, P. F., and T. M. P. Gilbert. 2012. DNA extraction from keratin and chitin. Methods in Molecular Biology 1963:57–63. doi:10.1007/978-1-4939-9176-1_7.

Chang, J., D. Song, and K. Zhou. 2007. Incongruous nuclear and mitochondrial phylogeographic patterns in two sympatric lineages of the wolf spider Par*dosa astrigera* (Araneae: Lycosidae) from China. Molecular Phylogenetics and Evolution 42:104–121.

Cokendolpher, J. C. 2004. *Cicurina* spiders from caves in Bexar County, Texas (Araneae: Dictynidae). Texas Memorial Museum Speleological Monographs 6:13–58.

Cokendolpher, J. C., and J. R. Reddell. 2001. Cave spiders (Araneae) of Fort Hood, Texas, with descriptions of new species of *Cicurina* (Dictynidae) and *Neoleptoneta* (Leptonetidae). Texas Memorial Museum Speleological Monographs 5:35–55.

Darriba, D., D. Posada, A. M. Kozlov, A. Stamatakis, B. Morel, and T. Flouri. 2020. ModelTest-NG: A new and scalable tool for the selection of DNA and protein evolutionary models. Molecular Biology and Evolution 37:291–294.

Gertsch, W. J. 1992. Distribution patterns and speciation in North American cave spiders with a list of the troglobites and revision of the cicurinas of the subgenus *Cicurella*. Texas Memorial Museum Speleological Monographs 3:75–122.

Hedin, M. 2015. High-stakes species delimitation in eyeless cave spiders (*Cicurina*, Dictynidae, Araneae) from central Texas. Molecular Ecology 24:346–361.

Hedin, M., S. Derkarabetian, J. Blair, and P. Paquin. 2018. Sequence capture phylogenomics of eyeless *Cicurina* spiders from Texas caves, with emphasis on US federally-endangered species from Bexar County (Araneae, Hahniidae). Zookeys 769:49–76.

Hedin, M., and W. P. Maddison. 2001. A combined molecular approach to phylogeny of the jumping spider subfamily Dendryphantinae (Araneae: Salticidae). Molecular Phylogenetics and Evolution 18:386–403.

Kapli, P., et al. 2019. Multi-rate Poisson tree processes for single-locus species delimitation under maximum likelihood and Markov Chain Monte Carlo. Bioinformatics 33:1630–1638.

Kozlov, A. M., D. Darriba, T. Flouri, B. Morel, and A. Stamatakis. 2019. RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. Bioinformatics 35:4453–4455.

Kumar, S., G. Stecher, and K. Tamura. 2016. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. Molecular Biology and Evolution 33:1870–1874.

Langmead, B., and S. L. Salzberg. 2012. Fast gapped-read alignment with Bowtie 2. Nature Methods 9:357–359.

Ledford. J., P. Paquin, J. Cokendolpher, J. Campbell, and C. Griswold. 2011. Systematics of the spider genus *Neoleptoneta* Brignoli, 1972 (Araneae: Leptonetidae) with discussion of the morphology and relationships for the North American Leptonetidae. Invertebrate Systematics 25:334–388.

Luo, A., C. Ling, S. Y. W. Ho, and C. D. Zhu. 2018. Comparison of methods for molecular species delimitation across a range of speciation scenarios. Systematic Biology 67:830–846.

Paquin, P., and N. Dupérré. 2009. A first step towards the revision of *Cicurina*: redescription of type specimens of 60 troglobitic species of the subgenus *Cicurella* (Araneae: Dictynidae), and a first visual assessment of their distribution. Magnolia Press, Auckland, New Zealand.

Paquin, P., and M. Hedin. 2004. The power and perils of 'molecular taxonomy': a case study of eyeless and endangered *Cicurina* (Araneae: Dictynidae) from Texas caves. Molecular Ecology 13:3239–3255.

Pons, J., et al. 2006. Sequence-based species delimitation for the DNA taxonomy of undescribed insects. Systematic Biology 55:595–609.

Rambaut, A., A. J. Drummond, D. Xie, G. Baele, and M. A. Suchard. 2018. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. Systematic Biology 67:901–904.

Reid, N. M., and B. C. Carstens. 2012. Phylogenetic estimation error can decrease the accuracy of species delimitation: a Bayesian implementation of the general mixed Yule-coalescent model. BMC Evolutionary Biology 12:196.

USFWS (United States Fish and Wildlife Service). 1993. Endangered and threatened wildlife and plants; 90-day finding on a petition to list nine Bexar County, TX, invertebrates. Federal Register 58(229):63328–63329.

USFWS (United States Fish and Wildlife Service). 2011. Bexar County karst invertebrate distribution. Report dated 28 July 2011. United States Fish and Wildlife Service, Austin Ecological Services Field Office, Austin, Texas.

Veni, G. 2002. Delineation of hydrogeologic areas and zones for the management and recovery of endangered karst invertebrate species in Bexar County, Texas. A report prepared by George Veni & Associates for the US Fish and Wildlife Service, Austin, Texas.

Wheeler, W. C., et al. 2017. The spider tree of life: phylogeny of Araneae based on target-gene analyses from an extensive taxon sampling. Cladistics 33:574–616.

White, K., G. R. Davidson, and P. Paquin. 2009. Hydrologic evolution of the Edwards Aquifer recharge zone (Balcones fault zone) as recorded in the DNA of eyeless *Cicurina* cave spiders, south-central Texas. Geology 37:339–342.

Woodruff, C. M., and P. L. Abbott. 1979. Drainage-basin evolution and aquifer development in a karstic limestone terrain South-Central Texas, U.S.A. Earth Surface Processes and Landforms 4:319–334.

World Spider Catalog. 2021. World Spider Catalog. Version 22.5. Natural History Museum Bern, online at http://wsc.nmbe.ch. doi:10.24436/24432.

Yu, G., D. K. Smith, H. Zhu, Y. Guan, and T. Tsan-Yuk Lam. 2017. ggtree: an r package for visualization and annotation of phylogenetic trees with their covariates and other associated data. Methods in Ecology and Evolution 8:28–36. doi:10.1111/2041-1210x.12628.

Zerbino, D. R. 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Research 18:821–829.

Addresses of authors:

PRESTON J. MCDONALD

Department of Biological Sciences
Texas Tech University
Lubbock, TX 79409-3131 USA
mcdonaldp@usf.edu

JULIE A. PARLOS

Department of Biological Sciences
Texas Tech University at Waco
Waco, TX 76708 USA
julie.parlos@ttu.edu

James C. Cokendolpher

Natural Science Research Laboratory
Museum of Texas Tech University
Lubbock, TX 79409-3191 USA
cokendolpher@aol.com

Stirling J. Robertson

Texas Department of Transportation
Natural Resources Management Section
Austin, TX 78701 USA
and
Natural Science Research Laboratory
Museum of Texas Tech University
Lubbock ,TX 79409-3191 USA
stirling.robertson@txdot.gov

Jean K. Krejca

Zara Environmental, LLC
Manchaca, Texas 78652 USA
jean@zaraenvironmental.com

Jennifer C. Girón

Department of Entomology
Purdue University
West Lafayette, IN 47907 USA
and
Natural Science Research Laboratory
Museum of Texas Tech University
Lubbock, TX 79409-3191 USA
jennifer.giron@ttu.edu

Robert J. Baker [Deceased]

Department of Biological Sciences and
Natural Science Research Laboratory
Museum of Texas Tech University
Lubbock, TX 79409-3131 USA

Caleb D. Phillips

Department of Biological Sciences and
Natural Science Research Laboratory
Museum of Texas Tech University
Lubbock, TX 79409-3131 USA
caleb.phillips@ttu.edu

Publications of the Museum of Texas Tech University

Institutional subscriptions are available through the Museum of Texas Tech University, attn. NSRL Publications Secretary, Box 43191, Lubbock, TX 79409-3191. Individuals may also purchase separate numbers of the Occasional Papers directly from the Museum of Texas Tech University.

The Museum of Texas Tech University has a catalog of Occasional Papers which may be viewed online at www.depts.ttu.edu/nsrl. To do so, you must have Adobe Acrobat installed on your computer. If you have difficulty downloading Occasional Papers, please contact the Webmaster. If there is continued difficulty, contact the Webmaster and a single hard copy can be provided to you via mail at no charge.

Layout and Design: Lisa Bradley
Cover Design: Lisa Bradley. Photographs by Jean Krejca.
Production Editor: Lisa Bradley

This book was set in Times New Roman and printed on acid-free paper that meets the guidelines for permanence and durability of the Committee on Production Guidelines for Book Longevity of the Council on Library Resources.

Printed: 24 March 2022

Library of Congress Cataloging-in-Publication Data

Occasional Papers of the Museum of Texas Tech University, Number 381
Series Editor: Robert D. Bradley

Mitochondrial Perspective on Species Identification and Delimitation for Troglobitic *Cicurina* (Arachnida: Araneae: Hahniidae) from Central Texas

Preston J. McDonald, Julie A. Parlos, James C. Cokendolpher, Stirling J. Robertson, Jean K. Krejca, Jennifer C. Girón, Robert J. Baker, and Caleb D. Phillips

ISSN 0149-175X

Museum of Texas Tech University
Lubbock, TX 79409-3191 USA
(806)742-2442