12-1-2008

Conservation Genetics of the Comal Springs Riffle Beetle (Heterelmis comalensis) Populations in Central Texas, with Examination of Molecular and Morphological Variation in Heterelmis Sp. Throughout Texas

Tina Katherine Gonzales
Texas State University-San Marcos, Dept. of Biology, tixgirl3@hotmail.com

CONSERVATION GENETICS OF THE COMAL SPRINGS RIFFLE BEETLE (*HETERELMIS COMALENSIS)* POPULATIONS IN CENTRAL TEXAS, WITH EXAMINATION OF MOLECULAR AND MORPHOLOGICAL VARIATION IN *HETERELMIS* SP. THROUGHOUT TEXAS

Presented to the Graduate Council of
Texas State University-San Marcos
in Partial Fulfillment
of the Requirements

for the Degree

Master of SCIENCE

by

Tina K. Gonzales, B.S.

San Marcos, Texas
December 2008

CONSERVATION GENETICS OF THE COMAL SPRINGS RIFFLE BEETLE (*HETERELMIS COMALENSIS*) POPULATIONS IN CENTRAL TEXAS, WITH EXAMINATION OF MOLECULAR AND MORPHOLOGICAL VARIATION IN *HETERELMIS* SP. THROUGHOUT TEXAS

Committee Members Approved:

_____
Chris C. Nice, Chair

_____
James R. Ott

_____
Floyd W. Weckerly

_____
Yixin Zhang

Approved:

_____
J. Michael Willoughby
Dean of the Graduate College

ACKNOWLEDGEMENTS

I would first like to thank my graduate thesis advisor, Dr. Chris Nice, for his support throughout this project – my research and thesis would not have been possible without his expertise and guidance. Dr. Nice's enthusiasm for research and learning, and his commitment to his students inspired me every day. He is an incredible mentor and teacher, and I feel very fortunate to have been a part of his lab. I would also like to thank my other committee members, Dr. Jim Ott, Dr. Floyd Weckerly, and Dr. Yixin Zhang, not only for the time and effort that they committed to this project, but for helping me to become a better scientist. I want to say thank you to all of the professors I had at Texas State, each one of you has made a special impact on my life and I will always be grateful to you.

I would like to thank Randy Gibson (United States Fish and Wildlife Service), who was an integral part of my thesis research. It is hard to list all of the ways in which Randy contributed to this project, but I could not have done it without him. Not only did he help me to collect my endangered beetles, but he was also a constant source of knowledge about them. Randy's passion for aquatic insects made me passionate about them too, and thanks to Dr. Nice and Randy Gibson, I now have a deep appreciation for invertebrates. I also want to thank those who helped me collect beetles, including Joe Fries (U. S. Fish and Wildlife Service), Pete Diaz, Zach Gompert, Lauren Lucas, and countless others at the USFWS. Thank you also to Ethan Chappell for helping us to sample for beetles throughout Spring Lake, and for the use of equipment.

I also owe a debt of gratitude to my parents, Loraine Reed and Raul Gonzales, and my other family and friends for the encouragement and support they continuously provided, and to Jessica Hurley for always believing in me, supporting me, and listening to me talk about my thesis project for days at a time.

This study was funded by a Texas Parks and Wildlife Section 6 grant awarded to Dr. Chris C. Nice, and through scholarships received from Texas State University-San Marcos and the San Antonio Conservation Society. *H. comalensis* samples were collected

by USFWS personnel in compliance with a Federal Fish and Wildlife Permit (TE676811-1; Date of approval: 09-08-2005), and all other *Heterelmis* samples were collected in compliance with state and national permits (Texas Parks and Wildlife Permit 04-05, date of approval: 02-03-2005; National Park Service Permit GUMO-2006-SCI-0008, date of approval: 05-06-2006).

This manuscript was submitted on September 10, 2008

TABLE OF CONTENTS

Page

LIST OF TABLES

Table **Page**

LIST OF FIGURES

Figure **Page**

ABSTRACT

CONSERVATION GENETICS OF THE COMAL SPRINGS RIFFLE BEETLE (*HETERELMIS COMALENSIS*) POPULATIONS IN CENTRAL TEXAS, WITH EXAMINATION OF MOLECULAR AND MORPHOLOGICAL VARIATION IN *HETERELMIS* SP. THROUGHOUT TEXAS

by

Tina K. Gonzales, B.S.

Texas State University-San Marcos

December, 2008

SUPERVISING PROFESSOR: CHRIS C. NICE

The Comal Springs Riffle Beetle, *Heterelmis comalensis*, is an endangered endemic species, known to occur in only two spring complexes in the Texas Hill Country, Comal Springs and San Marcos Springs. We surveyed molecular genetic variation in *H. comalensis* and three congeners from Texas using mitochondrial and nuclear sequence markers and amplified fragment length polymorphisms (AFLPs), to determine whether *H. comalensis* is experiencing reduced genetic variability within and among populations in its limited geographic range, and to delineate evolutionarily significant units (ESU)s for the species. A morphological analysis of a critical character

used to distinguish *Heterelmis* species in the current taxonomic key was then conducted. The concordance between the taxonomic characters used to delineate *Heterelmis* species and the patterns of molecular genetic variation found within and among them was then assessed. Mitochondrial DNA (mtDNA) analyses revealed high levels of genetic variation within and differentiation among three of the seven *H. comalensis* localities sampled, particularly when compared to the common flight- and drift-dispersed species of *Heterelmis*, suggesting isolation among these populations, with little to no current gene flow. However, the other four *H. comalensis* localities sampled were genetically invariant according to mtDNA data. Partially supporting mtDNA results, AFLP analyses clustered the seven *H. comalensis* populations into two groups, those with high and those with low genetic diversity. Sequencing of the single-copy nuclear gene ITS revealed a single haplotype within *H. comalensis* and its most closely related congener, suggesting recent divergence of the two species. Finally, this study found that nominal species designations using the current taxonomic key are incongruent with morphological and molecular data, necessitating amendment to the key. Although genetic variation is exceptionally high in three of the *H. comalensis* populations, common summer droughts in Texas in combination with human withdrawal of water threaten the springs that *H. comalensis* inhabit, and may have caused severe population bottlenecks in the four invariant localities sampled.

CHAPTER I

INTRODUCTION

Statement of the Problem/Significance

One of the primary goals in conservation is maintaining genetic variation in natural populations. Genetic variation allows populations to adapt to changing environments and recover from disturbances (Lacy, 1997; Pearman, 2001). Small populations experiencing reductions in genetic variation are much less likely to be able to adapt to environmental changes than large populations, and therefore are more likely to go extinct (Bürger and Lynch, 1995). The process of genetic drift increases the risk of extinction, which further threatens small populations (Lacy, 1997; Frankham, 2005), and inbreeding depression may be exacerbated in populations when genetic variation is low (Lacy, 1997). Because they have limited ranges and relatively small populations compared to more common species, threatened or endangered species are more likely to experience population subdivision (Hedrick and Kalinowski, 2000) and subsequent reduced genetic variation. Therefore, fostering and maintaining variation in these populations is of critical concern (Lacy, 1997).

Although traditionally conservation units have been based on the biological species concept, recent attention has turned to maintaining evolutionarily significant units (ESU)s (Vogler and Desalle, 1994; Crandall et al., 2000). ESUs describe unique groups

of organisms that through either historical genetic isolation or independent adaptive responses to differing selection are on independent evolutionary trajectories. ESUs should be managed separately in order to conserve the variation that has developed over their separate evolutionary histories (Vogler and Desalle, 1994; Moritz, 1999, 2002; Crandall et al., 2000). Genetic variation is organized and maintained by ecological and evolutionary processes and management practices that consider these processes are an important objective of conservation efforts (Crandall et al., 2000; Moritz, 2002). If historical migration between populations has helped to maintain genetic variation through immigration of new alleles, then fostering this relationship should be a conservation goal (Crandall et al., 2000). However, if populations are not experiencing gene flow and genetic variation has developed through historical isolation and differentiation among populations, then maintaining isolation may be important to conserve populations with separate evolutionary histories. These differentiated populations are considered separate ESUs and may require separate management consideration (Haig, 1998).

Increasingly, population genetics is used to diagnose ESUs as genetic tools have become more reliable and available to researchers (Hedrick and Miller, 1992; Haig, 1998; Goldstein et al., 2000). In order to identify ESUs, genetic conservation efforts focus on 1) individual variation within populations, and 2) variation among populations (Rand, 1996). Genetic analysis has proved to be a reliable method for determining differentiation between groups and may be critical to the application of the ESU concept in conservation practice (Haig, 1998). For example, by taking a population genetics approach, molecular techniques can be used to identify unique populations that should be considered for

protection (Moritz, 1999; Crandall et al., 2000). Because funding for conservation is limited, genetically distinct populations may need to be identified and prioritized.

The U. S. Endangered Species Act seeks to identify and protect threatened and endangered species (USFWS, 1973), which includes maintaining diversity in those species based on the ESU concept (Haig, 1998). This study seeks to apply the ESU concept using a population genetics approach to the endangered Comal Springs Riffle Beetle, *Heterelmis comalensis*, which is known to occur in only two spring complexes in central Texas, Comal Springs and San Marcos Springs. Both the Comal and San Marcos Spring complexes are associated with the Edwards Aquifer of central Texas. The Edwards Aquifer is comprised mainly of limestone and dolomite, and is associated with the Balcones Fault zone. The aquifer is recharged through the Guadalupe River watershed (Figure 1) and water flows southwest to northeast through a network of openings in the rocks created by a series of steep-angled step faults (GBRA, 1988; LBG-Guyton Associates, 2004). Water from the aquifer spills naturally from artesian springs associated with the major faults. *Heterelmis comalensis* is reliant on these constant spring flows from the Edwards Aquifer for continued existence (Gibson et al., 2008).

Heterelmis comalensis, first described in 1988, and subsequently listed as a federally endangered species (USFWS, 1997), is endemic to central Texas. Although extensive sampling was conducted throughout the Comal Springs complex (Comal River and Landa Lake, New Braunfels, TX), *H. comalensis* was found to occur there in only six different locations. Springs at these separate locations are fed from different water sources (depths) from the Edwards Aquifer (LBG-Guyton Associates, 2004), therefore genetic differentiation among these riffle beetle “populations” could be linked to the

different water sources. In addition, in 2004, *H. comalensis* was confirmed in the San Marcos Springs complex of the San Marcos River, San Marcos, Texas, some 30 kilometers away. Although extensive sampling was also conducted throughout San Marcos Springs, to date *H. comalensis* has only been found in one location there (R. Gibson, pers. comm.). In order to define boundaries of ESUs for *H. comalensis*, it is essential to have an understanding of the genetic variation present within the species, as well as determine how that variation is partitioned. This includes determining the level of differentiation among the separate localities of *H. comalensis* within Comal Springs and between Comal and San Marcos Springs.

Determining genetic differentiation among the known *H. comalensis* localities can also be used to examine the geographic patterns of genetic variation within the species. If groups of localities within Comal Springs are experiencing current gene flow with little genetic differentiation, then they might be managed together. However, if individual localities are determined to be genetically isolated, then they may be considered separate conservation units and should be managed as such. Similarly, knowledge about genetic variation within the San Marcos Springs locality, as well as between the San Marcos Springs and Comal Springs localities, is necessary to evaluate possible ESUs within the species. Differentiation between the San Marcos and Comal Springs complexes will help to determine whether there is current gene flow between the two locations, and may indicate the dispersal capabilities of *H. comalensis*. If the locality at San Marcos Springs is not differentiated from the Comal Springs complex and there seems to be high levels of gene flow between the two complexes, then they may not represent separate ESUs, and therefore may be managed together. However, if the two spring complexes in which *H.*

comalensis are found are significantly differentiated from one another, or if the separate localities at San Marcos and Comal Springs are significantly differentiated, then they may require separate management consideration.

Heterelmis comalensis is the smallest of the *Heterelmis* species (n = 5) in the United States, possesses vestigial hind wings, and is flightless (Bosse et al., 1988). *Heterelmis comalensis* is the only species of *Heterelmis* with vestigial wings (Bosse et al., 1988), therefore dispersal capabilities may be limited in this as compared to the more widespread flight- and drift-dispersed species of *Heterelmis* that are found in the United States (Figure 2). Limited dispersal among the *H. comalensis* localities could result in reduced gene flow and reduced genetic variation in this species, which could be further exacerbated by genetic drift and inbreeding depression if population sizes are small (Lacy, 1997). Thus, comparing genetic variation in *H. comalensis* localities to that found in localities of the more common flight- and drift-dispersed species of *Heterelmis* sampled at similar geographic distances will help to determine whether *H. comalensis* is experiencing reduced genetic variation compared to the more common and more highly vagile species of *Heterelmis*. If species of *Heterelmis* show different patterns of dispersal and gene flow, knowledge about the phylogenetic relationships of *Heterelmis* species might help to determine whether gene flow patterns in *H. comalensis* are typical or restricted due to its limited flight capabilities. Phylogenetic data may also help to uncover the evolutionary history of the endangered *H. comalensis*. Finally, it is critical that molecular and morphological distinctions among species are reflected in nominal species designations so that informed conservation management decisions can be made concerning *H. comalensis*.

Thus, we used mitochondrial and nuclear DNA sequence data, amplified fragment length polymorphisms, and morphological data from *H. comalensis* and three other widespread *Heterelmis* species to answer the following seven questions: Are the six localities of *H. comalensis* within Comal Springs genetically differentiated from each other? Is genetic structure within Comal Springs' localities concordant with the different water sources feeding the springs? Do analyses indicate migration between Comal and San Marcos Springs and if so, what are the implications for the dispersal capabilities and migration routes of *H. comalensis* between habitats? How does genetic variation within and among the flightless insular localities of *H. comalensis* compare to the more common flight- and drift-dispersed species at similar geographic distances? What are the effective population sizes for each of the *Heterelmis* species? What is the phylogenetic relationship between *H. comalensis* and the more widespread *Heterelmis* species, and when did speciation occur? Is there concordance between nominal taxonomic distinctions, morphology, and molecular data? These data will allow management decisions to be made which reflect the natural genetic structure found among the *H. comalensis* populations, as well as preserve the genetic variation within and between populations.

Multiple genetic markers were examined in this study, including mtDNA and nuclear DNA sequence and AFLP markers, because the information gained from a single locus is limited based on the inherent qualities of each marker, therefore conservation genetics' results are most reliable when based on data from multiple loci (Forister et al., 2008). MtDNA data was examined because it is thought to have a faster mutation rate (Avise, 1994), making it sensitive enough for population-level studies. We also examined nuclear DNA sequence data because nuclear DNA is generally thought to have a slower

mutation rate (Avise, 1994); therefore, it reveals the more distant evolutionary history of species than mtDNA, and can either corroborate or refute mtDNA results. We also examined *H. comalensis* individuals using AFLPs, because they give a good picture of overall genomic similarity among individuals (Mueller and Wolfenbarger, 1999; Bensch and Åkesson, 2005).

CHAPTER II

MATERIALS AND METHODS

Beetle Biology

Aquatic beetles of the genus *Heterelmis* (Coleoptera: Elmidae), commonly called "riffle beetles," get their name from their propensity to be found living in the shallow riffles associated with fresh-water springs and seeps (Sanderson and Brown, 1956; Brown, 1987). *Heterelmis* species range in body size from a length of 1.7 mm to 2.6 mm, and a width of 0.8 mm to 1.20 mm (Brown, 1972). The riffle beetles occur in constant conditions of temperature, dissolved oxygen, and pH (Bosse, 1979). Firm substrates such as cobble, gravel, woody debris, and aquatic vegetation are essential components of the habitat of this genus (Bosse, 1979).

Population Sampling

A total of 152 *H. comalensis* individuals were collected from the six sites at Comal Springs (Figure 3; Table 1), and the population at San Marcos Springs (Figure 4; Table 1). In addition, at least twenty specimens each were collected from three of the more common *Heterelmis* species, *H. glabra* (n = 77), *H. vulnerata* (n = 70), and *H. obesa* (n = 21), from 14 different spring and riverine systems across central and western Texas (Figure 5; Table 1). It should be noted that the Fessenden Springs population

(Figure 2) was initially identified as *H. vulnerata* using the current taxonomic key, but because it fell within *H. glabra* according to mtDNA and nuclear DNA sequence data, it was included as part of *H. glabra* for all statistical analyses. Upon capture, specimens were placed in 95% ethanol. The two front tarsi were removed from all specimens from each locality, and genitalia were removed from male specimens, and all were mounted on microscope slides and photographed at 20x, 40x, and/or 100x magnification using an Olympus BX45 camera; these slides and photo records were kept as vouchers. After photo records were taken, DNA extraction was performed with the Purgene DNA Isolation Kit (Gentra Systems, Minneapolis, MN, USA). All other tissue was used for DNA extraction because of the beetles' small size. In all, DNA was extracted from 320 specimens from 21 localities (Table 1).

Molecular Techniques

Mitochondrial DNA

We sequenced a region of two mtDNA genes, cytochrome oxidase subunits I and II (COI and COII) for 320 specimens from the 21 *Heterelmis* localities (Table 1). The COI and COII regions of mtDNA have been used extensively in prior sequencing studies with invertebrates (Juan et al., 1998; Caterino and Sperling, 1999; Austin et al., 2002; Szalanski and Owens, 2003). This yielded a 457 base pair fragment for COI and 557 base pair fragment for COII. Polymerase chain reaction (PCR) for the COI fragment was performed using the two primers BONNIE (5'-ACA AAC CAC AAG GAC ATC GG-3') and CLYDE (5'-GGG AGA GAC AGG AGA AGG AG-3'). These primers were developed by cloning PCR fragments from several *Heterelmis* individuals. PCR for the

COII fragment was performed using the two primers PIERRE (5'-AGA GCC TCT CCT TTA ATA GAA CA-3') and EVA (5'-GAG ACC ATT ACT TGC TTT CAG TCA-3') (Caterino and Sperling, 1999). General thermocycler parameters for PCR were: 94˚ C for 2 min., 35 cycles of 94˚ C for 1 min., 60˚ C for 1 min., and 72˚ C for 1 min., followed by 72˚ C for 10 min., and a 4˚ C hold, with optimization (annealing temp. of 48˚- 60˚ C) for each locality.

Fluorescently labeled dideoxy terminators were used for single stranded sequencing reactions according to Beckman Coulter, Inc. specifications. Labeled amplicons were separated and visualized using a capillary DNA sequencer (CEQ model 8800). Strands were replicated in both directions with the original primers, with partial to complete overlap of the region for most individuals. However, to sequence the COII region in some individuals, internal primers IFAH (5'-GTT AGA TGA AAT TAA TAA CCC-3') and IRBH (5'-TCG GTT ATC TAC ATC CAA AAG-3') were also used. Sequences were aligned using the program Sequencher 4.8 (Gene Codes Corp., Ann Arbor, MI). A partition homogeneity test between the COI and COII sequence data sets, conducted using PAUP* 4.0b10 (Swofford, 2002), detected no conflicting signal between the two mtDNA regions (P = 0.4700), thus the two sequence data sets were combined (= 1014 bp) for all statistical analyses. All unique mtDNA sequences produced in this study were deposited in GenBank under accession numbers EU827315 - EU827361.

Data Analysis and Statistical Methods using mtDNA Data

Haplotype frequencies and the average number of nucleotide differences among haplotypes (π) were estimated using Arlequin 3.11 (Excoffier et al., 2005). To compare

genetic differentiation among *H. comalensis* populations, pairwise Φ-statistics (Φ_{ST}; Excoffier et al., 1992) were calculated for all seven *H. comalensis* localities using Arlequin 3.11. Secondly, the Raymond and Rousset (1995) exact test of population differentiation was performed for all pairwise combinations of *H. comalensis* localities using Arlequin 3.11. To determine if the species-wide genetic variation found in *H. comalensis* is explained by differences between the two spring complexes in which it is found, an analysis of molecular variance (AMOVA; Excoffier et al., 1992) was conducted using Arlequin 3.11. The AMOVA partitioned the total molecular variance into a nested hierarchy of between spring complexes (Comal and San Marcos Springs), among populations within spring complexes, and within populations (localities).

To determine whether haplotype variation is reduced in *H. comalensis* compared to the more widespread flight- and drift-dispersed *Heterelmis* species, haplotype diversity (h) (Nei, 1987) was calculated for all localities using Arlequin 3.11 (Excoffier et al., 2005). In addition, a maximum parsimony haplotype network was created with the mtDNA dataset using TCS 1.2.1 (Clement et al., 2000), which employs the statistical parsimony algorithms of Templeton et al. (1992), including the *H. comalensis* and *H. glabra* haplotypes. Only these two species were included in this analysis because the networks of the other two species were too distantly related and beyond the limits of statistical parsimony.

To compare population differentiation among the *Heterelmis* species, pairwise Φ-statistics were calculated for all localities within each species, again using Arlequin 3.11 (Excoffier et al., 2005), and then averaged. In addition, isolation by distance (IBD) was tested for among the localities within the three *Heterelmis* species, *H. comalensis*, *H.*

glabra, and *H. vulnerata*, using a Mantel test (1000 replications), with the program IBDWS 3.15 (Jensen et al., 2005). The Mantel test determined the significance of the correlation (r) between log-transformed genetic distance (Slatkins Similarity Index M = ((1/PhiST) - 1)/4)) and log-transformed geographic distance. The measures of IBD were calculated to detect whether observed patterns of genetic variation resulted from geographic distance between localities, and to help compare dispersal patterns among the different *Heterelmis* species.

To reveal how distantly related haplotypes are within and among species, and indicate the amount of genetic variation present within and among the species, sequence divergences within and among the four *Heterelmis* species were determined using uncorrected p-distances, or the uncorrected proportional differences among haplotypes, calculated using DNAsp 4.0 (Rozas et al., 2003). Sequence divergence within *H. comalensis* was then compared to that found in the more widespread flight- and drift-dispersed species to determine whether *H. comalensis* is experiencing reduced variation due to its limited range and potentially limited dispersal capabilities.

To determine phylogenetic relationships among the four *Heterelmis* species, Neighbor-Joining and Bayesian Maximum Likelihood (BML) methods were employed. In the first analysis a tree was computed with estimates based on Neighbor-Joining (1000 bootstraps) using PAUP* 4.0b10 (Swofford, 2002). *H. obesa*, the most distantly related species to *H. comalensis* sampled in this study based on morphological data, was used as the outgroup. For the BML analysis, we used a partitioned model based on codon position. The mtDNA data was partitioned using McClade 3.0 (Maddison and Maddison, 1992). To combine the COI (457 bp) and COII (557 bp) mtDNA fragments for the BML

analysis, the partitioned data was edited so that only whole codons were included in the combined sequences. This reduced the entire fragment length to 1008bp but allowed the data to be aligned in the first position. However, this did eliminate one variable site between *H. obesa* and the other three species (2^{nd} nucleotide). We evaluated 56 potential models of DNA evolution with MODELTEST 3.7 (Posada and Crandall, 1998) and the best-fit model for each codon position according to AIC estimates was determined. The models chosen for the 1^{st}, 2^{nd}, and 3^{rd} positions were GTR + I (N_{ST} = 6; rates = equal), F81 + I (N_{ST} = 1; rates = equal), and TrN + G (N_{ST} = 6; rates = gamma). These models were then used in the Bayesian Maximum Likelihood analysis in MrBayes (Hall, 2001) to determine the most likely tree.

MtDNA data was also used to estimate female effective population sizes for all populations with at least one polymorphic site using FLUCTUATE 1.5 (Kuhner et al., 2005). This program uses a Metropolis-Hastings Markov Chain Monte Carlo technique to analyze the molecular sequences from individuals in a population and makes maximum likelihood estimates of the parameter Θ (= $2N_e\mu$). Θ was calculated assuming constant population size and using an initial estimate of Θ based on Watterson's (1975) method. The search strategy used 10 short chains with 200 steps and a sampling increment of 20, followed by 2 long chains with 20,000 steps and a sampling increment of 20. Female effective population size estimates were then calculated using Watterson's (1975) equation $2N_e\mu = \Theta$, and assuming a constant mutation rate for arthropods of 1.1% per million years per lineage (Brower, 1994).

Nuclear DNA

We also sequenced a single copy nuclear gene, the ribosomal DNA internal transcribed spacer region (ITS) for 30 *Heterelmis* individuals across 15 localities, including *H. comalensis* (n = 15), *H. glabra* (n = 9), and *H. vulnerata* (n = 6) (Table 1). The ITS has been used extensively for sequencing of invertebrates (Powers et al., 1997; Adams et al.,1998; Szalanski et al., 2000). PCR primers were ITSF2 (5'-TTG AAC ATC GAC ATT TCG AAC GCA CAT-3') and ITSR2 (5'-TTC TTT TCC TCC SCT TAY TRA TAT GCT TAA-3') (Z. Gompert; pers. comm.). General thermocycler parameters for PCR were: 94˚ C for 4 min., 35 cycles of 94˚ C for 50 sec., 50˚ C for 50 sec., and 72˚ C for 1 min., followed by 72˚ C for 10 min., and a 4˚ C hold. PCR yielded a ~500 base pair fragment for all 30 individuals. Single stranded sequencing reactions, and separation and visualization of labeled amplicons were conducted using the previously outlined methods. Sequences were aligned using the program Sequencher 4.8, resulting in a 502 bp fragment for all 30 individuals (Table 1).

Amplified Fragment Length Polymorphisms (AFLPs)

AFLP marker profiles were produced for 5-14 individuals from each of the seven populations of *H. comalensis* sampled. AFLP analysis was performed as described by Vos et al. (1995), with some modifications made by Gompert et al. (2006a, b) genomic DNA (1.5 µl) was digested with the restriction enzymes *EcoRI* and *MseI*; adapter oligonucleotides were ligated onto the ends of the digested fragments. This reaction was carried out in a solution containing 1.1 µl of 10X T4 DNA Ligase buffer, 1.1 µl of NaCl (0.5 M), 0.55 µl of BSA (1mg/ml), 0.1 µl *MseI*, 0.42 µl *EcoRI*, 0.33 µl T4 DNA Ligase,

1 µl of the *MseI* adapters (~0.62 g/l), and 1 µl of the *EcoRI* adapters (~0.064 g/l). This was diluted to a total volume of 11 µl. DNA was incubated in the above solution for 2 hours at 37˚ C, and then diluted with 60 µl of water. The dilute digested DNA with ligated adapters (~30 µl) was then purified using a Qiaquick PCR Purification Kit (Qiagen) to remove small DNA fragments that might interfere with the amplification of larger, scoreable fragments in later amplifications. Purification was performed according to the manufacturer's protocol, except to elute with 30 µl of elution buffer. The restriction/ligation product was then diluted further with 60 µl of water.

The purified digested DNA was then subjected to pre-selective PCR amplification, performed in a total volume of 20 µl (pH 9.0) containing 4 µl of clean restriction/ligation product, 0.5 µl of each of the pre-selective primers (20 µM), 0.5 µl of dNTPs, and 0.5 µl of *Taq* polymerase. Thermocycler parameters for pre-selective amplification were as follows: 72˚ C for 2 min., 25 cycles of 94˚ C for 20 sec., 56˚ C for 30 sec., and 72˚ C for 2 min., followed by 60˚ C for 30 min., and a 4˚ C hold. The pre-selective PCR product was then diluted with 40 µl of water, and used as the template for selective PCR amplification. Selective PCR was performed to amplify a subset of the fragments amplified in pre-selective PCR, in a total volume of 20 µl (pH 9.0) containing 3 µl of pre-selective PCR product (1:10 dilution), 1 µl of a *MseI*-CTTG selective primer (5 µM), 1 µl of and *EcoRI*-ACA selective primer (1 µM; fluorescently labeled with D3 or D4), 4 µl of buffer E, 0.5 µl of dNTPs, and 0.5 µl of *Taq* polymerase. Thermocycler parameters for selective amplification were as follows: 94˚ C for 20 sec., 50 cycles at 94˚ C for 20 sec., 66-56˚ C for 30 sec., decreasing by 1˚ C each cycle until 56˚ C was

reached. The remaining cycles were performed at 56˚ C, then 72˚ C for 2 min., ending with 60˚ C for 30 min., and a 4˚ C hold.

Amplified fragments were separated and visualized (Beckman Coulter CEQ 8800 automated DNA sequencer) using a 6% denaturing polyacrylamide gel, with 0.3 µl of DNA size standard-600 (Beckman Coulter), 40 µl of deionized formamide (per sample), and 1 µl of selective PCR product added to each lane. Beckman Coulter fragment analysis software was used to visualize AFLP bands, and band selection and quality control were performed following the previously described methods of Gompert et al. (2006a, b). Only fragments with a peak height of 150 fluorescent units or greater were scored. All fragments less than 70 base pairs in length were excluded from analysis. Of those fragments greater than 70 base pairs in length, fragments appearing in less than 10 individuals were excluded to reduce the probability of analyzing homoplastic and aberrant fragments, which might dilute the signal present in the data (Gompert et al. 2006a, b).

Data Analysis and Statistical Methods using AFLP Data

Program STRUCTURE v2.2 (Pritchard et al., 2000; Falush et al., 2007) was used to identify (cluster) localities of genomically similar individuals based on their AFLP banding profiles. STRUCTURE employs a model-based Bayesian clustering algorithm to assign individuals probabilistically to clusters to minimize deviations from linkage equilibrium. Analyses were conducted using a model allowing for recessive alleles, appropriate for dominant molecular markers such as AFLPs (Falush et al., 2007). The admixture model was used, which allows for gene exchange among populations, and runs

were conducted using a Markov chain Monte Carlo (MCMC) with 500,000 generations with an initial burn-in of 50,000 generations. Prior information regarding the locality from which an individual was sampled was ignored. The number of clusters (*k*) was evaluated from one to seven, and 10 independent MCMC runs were conducted for each *k* to provide an assessment of the variance in likelihood estimates for each value of *k*. We then plotted *k* versus the mean log likelihood for each *k* to aid in selecting the number(s) of clusters that best explained the AFLP data, which is assumed to be the value of *k* at which the log likelihood reaches an asymptote (Pritchard et al., 2000).

Morphological Comparison among Species

The taxonomic key used to identify species of *Heterelmis* in North and Central America was first introduced by Brown (1972) and was later amended when *H. comalensis* was discovered (7-7A'; Bosse et al., 1988). The key first distinguishes *H. tarsalis* from other *Heterelmis* species based on a "strongly, longitudinally concave" metasternal disk, clavate tibiae on front legs, and "conspicuous fringe of erect hairs" on the "basal four segments of the front and middle tarsi" (1; Brown, 1972; Bosse et al., 1988). All other species have a "metasternal disk [that is] only feebly concave posteriorly," the "tibiae of front legs is not clavate" in other species, and the "basal four segments of front and middle tarsi [are] without a conspicuous fringe of erect hairs" (1'; Brown, 1972; Bosse et al., 1988). Next, the key delimits species based on the presence of "two close, stout, short spines" on the "inner apex" of the "basal segment of each tarsus" (2; Brown, 1972; Bosse et al., 1988). If the two spines are present, the length of the prothorax and the slope of the median lobe of the male genitalia are examined. If the

length of the prothorax is “more than 0.925 mm” and the “median lobe of [the] male genitalia with the basal median projection seen from the side [is] gradually sloping behind” then it is identified as *H. longula,* from “high mountain streams of central Mexico” (3; Brown, 1972; Bosse et al., 1988). However, if the length of the prothorax is “less than 0.925 mm” and the “median lobe of male genitalia when seen from the side [is] strongly declivous behind,” then the disk of the pronotum must be examined (3’; Brown, 1972; Bosse et al., 1988). If the “disk of the pronotum [has] impressions” then it is identified as *H. obesa,* from Arizona, New Mexico, Mexico, and Guatemala (4; Brown, 1972; Bosse et al., 1988). However, if the “disk of [the] pronotum [is] evenly convex” then the individual is identified as *H. obesa plana* from central Mexico (4’; Brown, 1972; Bosse et al., 1988).

If in step number two there are not “two stout spines on inner apex” of the “basal segment of tarsi” then the disk of the pronotum is examined (Brown, 1972; Bosse et al., 1988). If the “disk of [the] pronotum” has a “distinct transverse impression [and] usually with distinct oblique impressions as well,” then the parameres of the male genitalia have to be examined (5; Brown, 1972; Bosse et al., 1988). If the “parameres of male genitalia [have] an inner apical fringe of hairs” then the width of the “median lobe of male genitalia is examined (6; Brown, 1972; Bosse et al., 1988). On the other hand, if the parameres of the “male genitalia [are] without an inner apical fringe of hairs” and are “1.9-2.4 mm long [and] 0.9-1.1 mm wide,” then they are identified as *H. vulnerata,* known from Texas and Oklahoma (6’; Brown, 1972; Bosse et al., 1988). If in step 5 the “disk of [the] pronotum [is] without impressions or with very feeble ones,” then the “basal segment of each tarsus” and body size are examined (5’; Brown, 1972; Bosse et

al., 1988). If the “basal segment of each tarsus [has] 3 short, stout spines,” the body is “elongate, sides subparallel, [the] elytra [is] barely wider than [the] thorax,” and the body is “2.3-2.6 mm long, 1.05-1.2 mm wide” then the individual is identified as *H. stephani* from the Santa Rita Mountains of Arizona (8; Brown, 1972; Bosse et al., 1988). If however, the “basal segment of [the] tarsus is without such spines,” the body is “plump, sides arcuate,” the “elytra [is] noticeably wider than [the] thorax” and the body is “1.7-2.1 mm long, 0.8-1.0 mm wide” then the individual is indentified as *H. simplex* from Guatemala, Trinidad, and Tobago (8’; Brown, 1972; Bosse et al., 1988).

Finally, if in number 7 the “median lobe of [the] male genitalia [is] broad (0.05 mm wide between apices of parameres)” and the body is “1.85-2.2 mm long, 0.9-1.1 mm wide,” then the individual is identified as *H. obscura* from Mexico and Central America (7; Brown, 1972; Bosse et al., 1988). If on the other hand the “median lobe (penis) of male genitalia [is] narrow (about 0.017 mm) between apices of parameres,” then one has to look at size and color to make a final distinction between species (7’; Bosse et al., 1988). If the individual is “small and relatively slender (1.7 to 2.1 mm long, 0.8 to 0.91 mm wide),” and is “pale or light in color,” then it is identified as *H. comalensis*, known only in Comal Springs according to the current taxonomic key (7A; Bosse et al., 1988). However, if the individual is “larger and more robust (1.9 to 2.35 mm long, 1.0 to 1.17 mm wide)” and darker in color then it is identified as *H. glabra*, according to the current key found in Arizona, Mexico, Central America, and the Big Bend region of Texas (7A’; Bosse et al., 1988). Thus, the key character for distinguishing *H. vulnerata* from *H. comalensis* and *H. glabra* is a fringe of hairs on the parameres of the male genitalia (6-6’; Bosse et al., 1988).

There is individual variation in number of hairs found on the parameres of male genitalia across species found within Texas (personal observation). Because this character is critical for distinguishing *H. comalensis* individuals, we counted the number of hairs found on male genitalia for 160 specimens, from 21 populations including all four *Heterelmis* species collected for this study. Individuals from the Fessenden Spring population were included with *H. glabra* based on mtDNA and nuclear DNA data. To test for significant differences, the number of hairs was compared across species using a nonparametric Kruskal-Wallis test, following the method outlined in Conover (1999). In addition, a multiple comparison procedure was performed to specifically determine significant differences between pairs of species (Conover, 1999). The number of hairs counted was then compared to the current taxonomic key and the mtDNA and nuclear DNA data produced in this study to determine whether current nominal taxonomic distinctions are concordant with morphological and molecular data.

CHAPTER III

RESULTS

Mitochondrial DNA Data

DNA sequencing of the COI and COII region of mtDNA produced 37 unique haplotypes (Table 1). A total of 211 variable sites (20.8%) were found among the 37 haplotypes. Among the 152 *H. comalensis* individuals, 8 haplotypes were found in the seven localities, with four private alleles at West Shoreline, one private allele at Spring Island, and one at San Marcos. The most common haplotype occurred in all seven localities (Table 1). The average number of nucleotide differences (π) among the *H. comalensis* haplotypes was 1.857. In 77 *H. glabra* individuals, 20 haplotypes were found in the five localities sampled, with no shared haplotypes among any locality (Table 1). Among *H. glabra* haplotypes, π was 5.0. In 70 *H. vulnerata* individuals, four haplotypes were found across the seven localities. Two haplotypes were common with either or both occurring in all localities, and two haplotypes were private (Table 1). Among *H. vulnerata* haplotypes, π was 1.286. Finally, in 21 *H. obesa* individuals, five haplotypes were present in the two sampled localities. One haplotype was shared and four were private (Table 1). Among *H. ob*esa haplotypes, π was 2.5.

Differentiation among *H. comalensis* Populations

Inspection of pairwise Φ-statistics showed that Spring Runs 1, 2, and 3, and the Backwater Spring localities at Comal Springs are all identical with no variation within or between them, as they are all fixed for the common haplotype (Table 2). However, the Spring Island and West Shoreline populations at Comal Springs, and the San Marcos Springs population all exhibit statistically significant genetic differentiation with respect to all other localities (Table 2). In addition, the populations Spring Island, West Shoreline, and San Marcos Springs all have unique haplotypes, suggesting that they are divergent from one another. The Raymond and Rousset (1995) exact test of population differentiation confirmed the above results (Table 2). The AMOVA partitioned the total molecular genetic variance into a hierarchy of between spring complexes (Comal Springs vs. San Marcos Springs; -2.72%), among populations within spring complexes (35.5%), and within populations (67.2%) (Table 3). Thus, although localities within Comal Springs and San Marcos Springs are differentiated, the combined populations are not differentiated as separate spring complexes.

Genetic Differentiation among *Heterelmis* Species

The genetic variation (indexed by haplotype diversity h) in *H. comalensis* was first compared to genetic variation among the more common species for each population (Table 1). No haplotype diversity for *H. comalensis* was present at Spring Runs 1, 2, and 3, and Backwater Spring in the Comal Springs complex, however at West Shoreline, Spring Island, and San Marcos Springs, h was comparable to that found in the sampled localities of *H. glabra*, *H. obesa*, and H. *vulnerata* (Table 1). In *H. comalensis*, h was

highest at West Shoreline (h = 0.7463), and only three among all other localities had similar or higher values (2 *H. glabra* localities and 1 *H. obesa* locality; Table 1). Spring Island at the Comal Springs complex and San Marcos Springs had relatively high haplotype diversity as well (h = 0.5286 and h = 0.5000, respectively) compared to *H. glabra* (mean h = 0.59238), *H. vulnerata* (mean h = 0.40545) and *H. obesa* (mean h = 0.6101) (Table 1).

A haplotype network restricted to *H. comalensis* and *H. glabra* was created to look specifically at the number of unique haplotypes within and among these closely related species (Figure 6). The most common haplotype in *H. comalensis*, haplotype 1, is found in all seven localities (Table 1), and haplotype 3 is common between Comal and San Marcos Springs (Figure 6). However, both Comal and San Marcos Springs also have unique haplotypes, suggesting that the two localities are differentiated and have not experienced recent or substantial gene flow; demonstrated as well by the pairwise comparisons of localities shown above. *Heterelmis glabra* shows the most sequence variation, and that variation is tied to the geographic variation among the five sampled populations of *H. glabra.* There were no shared haplotypes among any of the five populations of *H. glabra* and the distances (number of nucleotide differences) among *H. glabra* haplotypes were as large as the distance between *H. glabra* and *H. comalensis* haplotypes, suggesting that there is little to no gene flow among the separate *H. glabra* populations (Figure 6). In fact, the geographically isolated Fern Bank population, while falling within the nominal species *H. glabra*, has haplotypes that are 10 nucleotide differences to the other *H. glabra* haplotypes. In comparison, *H. comalensis* and *H. glabra* are distinguished by four nucleotide differences (Figure 6).

Inspection of Φ-statistics calculated among all localities within a species and then averaged, demonstrated population differentiation within and among the different *Heterelmis* species (Table 4). All five *H. glabra* populations were significantly differentiated from one another, with an average Φ-statistic of 0.811 (s.d. = 0.107) across localities, while *H. comalensis* localities had an average Φ-statistic of 0.209 (s.d. = 0.156) across populations (Table 4). In contrast, *H. vulnerata* localities had an average Φ-statistic of 0.09 (including negative values; s.d. = 0.335). The *H. obesa* populations were found to be significantly differentiated with a Φ-statistic of 0.39 (Table 4), but since only two *H. obesa* populations were sampled, the results for this species should be considered cautiously. Mantel tests showed no significant isolation by distance (IBD) among localities of *H. comalensis* ($r = 0.4491$, $p = 0.9327$) or *H. glabra* ($r = 0.4667$, $p = 0.9164$). However, marginally significant IBD was found among the *H. vulnerata* localities ($r = -0.4676$, $p = 0.0521$).

Sequence Divergences among *Heterelmis* Species

Uncorrected sequence divergences showed that *H. comalensis* has lower sequence divergence within it than *H. glabra*, but more than the common *H. vulnerata* and *H. obesa* (Table 6). Without the Fern Bank locality included, *H. glabra* populations still show the highest sequence divergence (0.66%). Comparison of among-species divergence showed that *H. comalensis* and *H. glabra* are 1.1% divergent on average, and both species are 12% divergent from *H. vulnerata*, and 13% divergent from *H. obesa*. *H. vulnerata* is also 13% divergent from *H. obesa* (Table 6). Comparison of sequence divergence (uncorrected p-distance) between the Fern Bank and the other *H. glabra*

haplotypes showed that Fern Bank is about 1.5% divergent from the rest of *H. glabra*. This is higher than the average sequence divergence between *H. comalensis* and *H. glabra*, not including Fern Bank (1.0%), again suggesting that Fern Bank is a cryptic species within *H. glabra*.

Phylogenetic Analysis of *Heterelmis* Species

Phylogenetic analysis showed *H. comalensis* to be most closely related to *H. glabra*, but to be monophyletic with moderate bootstrap support (75) and high Bayesian maximum likelihood support (1.00), suggesting that it is significantly divergent from *H. glabra* (Figure 7). However, *H. glabra* is paraphyletic with respect to *H. comalensis*. *Heterelmis glabra* as taxonomically defined by morphology was also shown to have a high degree of geographic and genetic structure within it, with each locality defined by unique haplotypes. High divergence was found among each of the *H. glabra* populations with moderate bootstrap support and high Bayesian maximum likelihood support (Figure 7). Although the Fern Bank population falls within nominal taxon *H. glabra*, it shows significant phylogenetic distance from all other *H. glabra* populations (Figure 7). Finally, the phylogenetic analysis showed that *H. vulnerata* and *H. obesa* are very distantly related to the other two species. There was much less geographic structure found in *H. vulnerata* and *H. obesa* than in *H. glabra*, but the structure found in both was similar to that found in *H. comalensis* (Figure 7).

Rough estimates of divergence times based on Brower's (1994) estimate of the mitochondrial mutation rate for arthropods showed that *H. comalensis* and *H. glabra* split roughly 0.50 million years ago. Fern Bank may have split from the rest of *H. glabra* 0.68

million years ago, while the ancestor of *H. comalensis* and *H. glabra* diverged from *H. vulnerata* roughly 5.5 million years ago and from *H. obesa* about 6.1 million years ago. *Heterelmis vulnerata* also split from *H. obesa* about 6 million years ago based on sequence divergence estimates (Table 6).

Population Sizes and Growth Rates

Estimates of female effective population sizes (N_e) were obtained for all localities with at least one polymorphism using the mtDNA data. For *H. comalensis, H. glabra,* and *H. obesa* estimated N_e were on average larger than the common *H. vulnerata*, although N_e varied dramatically among localities (Table 5). Within *H. comalensis,* West Shoreline and Spring Island had the largest estimated N_e, while N_e for San Marcos was less than half that size (Table 5). For *H. comalensis* and *H. glabra* N_e was similar, although the Caroline Springs estimated N_e was significantly larger than any other *H. glabra* N_e sampled (Table 5). Although one *H. vulnerata* locality showed a higher estimated N_e, all of the other *H. vulnerata* localities had relatively small N_e (Table 5). The two *H. obesa* populations had large differences in their estimated N_e (Table 5). Because of their dependency on fresh-flowing water, it seems probable that N_e within all of the *Heterelmis* species found in Texas may vary over time due to periodic droughts and consequent bottleneck events. All invariant localities had to be left out of this analysis, therefore we cannot estimate their N_e.

Nuclear DNA Data

Surprisingly, there was no sequence variation in the 502 bp ITS region within or across any of the *H. comalensis* localities (Table 1). There was also no sequence variation in this gene between *H. comalensis* and *H. glabra* individuals (Table 1). This analysis included the Fessenden Springs locality, providing further support for the mtDNA designation of Fessenden Springs' individuals as *H. glabra* and not *H. vulnerata*. Additionally, no sequence variation was found in the ITS gene within or among any of the *H. vulnerata* localities (Table 1). However, the *H. vulnerata* sequence was highly variable compared to that found in *H. comalensis* and *H. glabra*. In fact, there was so much insertion/deletion variation between the two haplotypes that they could not be properly aligned. Because of the genetic invariance revealed in the ITS region within and between *Heterelmis* species, no statistical analyses were conducted using these data. The unexpected genetic invariance found in ITS within and between *H. comalensis* and *H. glabra* prompted a search to find a single-copy nuclear gene that is more genetically informative. Several different primer pairs for multiple genes, including *wingless*, *elongation factor 1-alpha*, *tektin*, *tubulin*, and *white*, were used in PCR reactions, but these either did not produce a PCR product or did not produce a readable sequencing product.

Amplified Fragment Length Polymorphism Data

Data Analysis and Statistical Methods using AFLP data

A total of 159 polymorphic AFLP markers were generated. From 10 independent MCMC runs each, the number of clusters (*k*) 1-3 had no variance in their likelihood

estimation, while all other *k* did show substantial variance (Figure 8). When plotted, $k = 3$ is where the log likelihood reached asymptote, thus the number of clusters *(k*) that best explained the data for *H. comalensis* was three clusters (Figure 8). However, this Bayesian clustering analysis at $k = 3$ revealed two main groups with individuals scattered across several localities being assigned to a third group with low probability. The first major group corresponded to the four localities at the Comal Springs complex with low mtDNA variation (Spring Runs 1-3 and Backwater Spring), while the other group consisted of the three high diversity populations (West Shoreline, Spring Island, and San Marcos Springs) (Figure 8). These data further support results from the mtDNA analysis, and reveal more recent divergences among *H. comalensis* localities than the nuclear DNA does.

Morphological Comparison among Species and Locations

All *H. vulnerata* (n = 29) individuals lacked hairs on their parameres (Table 7). Most *H. obesa* individuals (21/28) also had no hairs, but this trait expression ranged to eight hairs, which constitutes a deviation from the current key (Brown, 1972; Bosse et al., 1988). *H. glabra* (n = 46) exhibited a wide range of values overlapping those of the three other species, with number of hairs ranging from zero to 32, which again deviates from the current key. Finally, for *H. comalensis* (n = 57) the number of hairs ranged from 16 to 35, overlapping only *H. glabra* (Table 7). Significant differences (Kruskal-Wallis: $T = 110.86$, d.f. = 3, $p = <0.0001$) were found among *Heterelmis* species in number of hairs on the male genitalia. The multiple comparison procedure supported a significant difference in number of hairs among all species-pairs except *H. obesa* and *H. vulnerata*.

However, because the distributions of number of hairs among the four *Heterelmis* species overlap considerably, this is not a character that can be used for species identification within *Heterelmis*. The morphological data were also compared to the current taxonomic key (Bosse et al., 1988), and both the mtDNA and nuclear DNA data produced in this study to determine whether these data are congruent.

The first problem encountered was that the species distinctions in the taxonomic key are based on slight morphological differences that are difficult to identify without direct comparison, including size differences among species (personal observation). This is especially true with respect to distinguishing *H. comalensis* and *H. glabra* individuals, which are distinguished from one another based solely on slight size and color differences in the key (7A-7A'; Bosse et al., 1988). Moreover, these closely related species overlap in size, both in length and in width (Bosse et al., 1988), with a great deal of individual morphological variation within species (personal observation). Furthermore, the key describes *H. comalensis* as only occurring in Comal Springs, and *H. glabra* as only occurring in the Big Bend region of Texas, so it would seem easy to distinguish between them, but this study shows that *H. glabra* is in fact found across west and central Texas. Additionally, nominal species designations and morphological data are incongruent with the molecular data presented in this study, suggesting that species may be misidentified based on the current key. Specifically, in step number 6' of the taxonomic key, if the "parameres of [the] male genitalia [are] without an inner apical fringe of hairs" within the body size range presented, the individual is identified as *H. vulnerata* (Brown, 1972; Bosse et al., 1988). However, mtDNA and nuclear DNA data showed that the Fessenden Springs population, which has individuals without an inner

apical fringe of hair (9 out of 23), are in fact *H. glabra*. This is currently the only character that distinguishes *H. vulnerata* from *H. glabra* and *H. comalensis*. In this way, the Fessenden Springs population was initially identified as *H. vulnerata* using the current taxonomic key, but surprisingly mtDNA and nuclear DNA data showed it to be clearly part of *H. glabra*.

CHAPTER IV

DISCUSSION

Genetic Differentiation within *H. comalensis*

This study examined genetic variation within and among localities of *H. comalensis* in order to define the boundaries of ESUs for this endangered, endemic species. Analysis of mtDNA data revealed surprisingly large amounts of genetic variation in three of the seven *H. comalensis* localities sampled: West Shoreline and Spring Island populations at Comal Springs and San Marcos Springs. The localities at Comal Springs, Spring Runs 1, 2, 3 and Backwater Spring, were all genetically invariant – fixed for haplotype 1. West Shoreline, Spring Island, and San Marcos Springs are all significantly differentiated from all other localities based on pairwise Φ-statistics, with pairwise values greater than those typically found between populations of other invertebrate species, even those separated by greater geographic distances (e.g. Nice et al., 2005). The differentiation among these localities was also supported by the Raymond and Rousset (1995) exact test, which revealed significant differentiation between all population pairs including West Shoreline, Spring Island or San Marcos Springs, despite the invariance found in the other four localities. In addition, the AFLP data set identified two general groups of populations – those with high diversity (San Marcos Springs, West Shoreline,

and Spring Island) and those with low diversity (Springs Runs 1-3 and Backwater Spring), partially corroborating results from the mtDNA analyses.

A history of genetic isolation among the West Shoreline, Spring Island, and San Marcos Springs populations, with little to no recent gene flow is suggested. Although these three genetically diverse *H. comalensis* populations contain unique haplotypes as shown in Figure 6, the most common haplotype found in all seven *H. comalensis* localities suggests that they shared gene flow at some point in their distant history, or that haplotype 1 was present in the common ancestral population of all *H. comalensis*. In addition, San Marcos Springs and the Spring Island populations share two haplotypes, which is more than is shared between any two localities within Comal Springs. Nevertheless, the high levels of divergence among the three diverse populations and all other localities suggest that recent gene flow has not occurred among any of the *H. comalensis* localities. Because West Shoreline, Spring Island, and San Marcos Springs have all experienced historical genetic isolation with little to no recent gene flow, they may each represent a separate ESU within *H. comalensis*, and therefore may require separate consideration when making conservation management decisions for this endangered species (Vogler and Desalle, 1994; Moritz, 1999, 2002; Crandall et al., 2000). Conserving genetic variation in these three potential ESUs is critical considering that the other four *H. comalensis* localities, Springs Runs 1, 2, and 3 and Backwater Spring, are genetically invariant, at least at the mitochondrial level.

Why are four out of the six localities in the Comal Springs complex genetically invariant when the other two show such high levels of genetic differentiation? The answer may be linked to the hydrology at this location and the water sources that feed the

separate springs. Springs within the Comal Springs complex are supplied from different water sources (depths) from the western part of the Edwards Aquifer (LBG-Guyton Associates, 2004). Spring Runs 1, 2, and 3 are fed from a single water source from the up-thrown fault block of the Balcones Fault Zone, while West Shoreline (Spring #7; LBG-Guyton Associates, 2004) and Spring Island are fed from a deeper water source from the down-thrown fault block (LBG-Guyton Associates, 2004). Lower water levels in the up-thrown block means that reduced flow due to drought and over-pumping from the Edwards Aquifer affects flow to the springs supplied from the up-thrown fault block at Comal Springs before it affects flows on the down-thrown block. This, in addition to higher elevations of the spring orifices at Spring Runs 1, 2, and 3 has resulted in these springs going dry more often and for longer periods than the lower elevation springs at West Shoreline and Spring Island that are supplied from the down-thrown fault block (LBG-Guyton Associates, 2004). This was observed when a drought in central Texas in the 1950s, in addition to pumping of wells from the Edwards Aquifer, caused all six of the springs at Comal Springs to dry up (GBRA, 1988; USFWS, 1997).

Localities of *H. comalensis* apparently survived the drought when Comal Springs ceased flowing in 1956 by moving down into the substrate beneath the receding water level (USFWS, 1997). The Spring Runs were dry for considerably longer than the springs in and around Landa Lake (including West Shoreline and Spring Island), and pools of water remained in Landa Lake until the springs began to flow again in September of 1956 (LBG-Guyton Associates, 2004). Although Backwater Spring was not included in the 2004 hydrologic survey by LBG-Guyton Associates, it can be assumed that this locality shares a water source with West Shoreline and Spring Island, as do the other springs in

the northern part of Landa Lake. However, Backwater Spring is found in a shoreline seep just north of the Spring Island area, and at that elevation was almost certainly dry during the drought in the 1950s. In addition, the only haplotype found at Backwater Spring is the same common haplotype that defines populations at Spring Runs 1, 2, and 3 (found in all populations of *H. comalensis*). These observations in conjunction with molecular genetics data suggest that the invariance found in the Spring Run and Backwater Spring localities is due to either bottleneck and/or founder effects. In addition, clustering of AFLP marker data indicates that the Spring Run and Backwater Spring localities cluster as one group, and because AFLPs are better indicators of total variation throughout the genome (Mueller and Wolfenbarger, 1999; Bensch and Åkesson, 2005), this clustering lends support to the hypothesis of recent and severe bottleneck events in the Spring Run and Backwater Spring localities.

It is possible that the Spring Run localities and Backwater Spring regularly go extinct when the springs dry up and are then repopulated with migrants from West Shoreline and Spring Island, but both are not probable because the level of differentiation between West Shoreline and Spring Island suggests that there is little migration between springs within the Comal Springs complex. Additionally, all invariant localities share the same haplotype, which is not the most common haplotype found in West Shoreline (closest geographically to Spring Runs 1, 2, and 3). Alternatively, individuals of one haplotype may have founded all of the populations, and the West Shoreline and Spring Island populations have been able to differentiate over time. On the other hand, the Spring Run springs and Backwater Spring may have reduced variation due to genetic drift through periodic or repeated population reductions. Although variation remains in

West Shoreline and Spring Island, these two populations were certainly negatively impacted when Comal Springs dried up in the 1950s, and could again be severely affected if the springs went dry for longer periods of time.

In addition, LBG-Guyton Associates (2004) found that although some of the groundwater moving from west to east in the aquifer discharges at Comal Springs, some groundwater moves through the aquifer beneath Comal Springs and discharges at San Marcos Springs (LBG-Guyton Associates, 2004; Johnson and Schindel, 2008). Additional groundwater to the north flows south and discharges at San Marcos Springs, but in times of low flow, more than ninety percent of the discharge at San Marcos Springs is water from the western part of the aquifer that has bypassed Comal Springs (Johnson and Schindel, 2008). In fact, this is why flow at San Marcos Springs did not cease in the 1950s, although Comal Springs did – water from the up-thrown fault block that would normally be expelled through the Spring Runs instead moved under Comal Springs and discharged at San Marcos Springs (LBG-Guyton Associates, 2004). If *H. comalensis* does indeed move down into the substrate as the water level recedes, it seems probable riffle beetles have moved (migrated) with the water from Comal to San Marcos Springs during times of very low or no flow at Comal Springs, at some time in the distant past. Testing of *H. comalensis* confirmed that they do indeed move downward through the substrate to remain in moving water, and they consistently move toward the current in experiments (Edwards Aquifer Authority, 2007). This may explain the shared haplotypes between Comal and San Marcos Springs, and may even explain how the San Marcos Springs population was initially founded. Because San Marcos Springs has not gone dry in recorded history (LBG-Guyton Associates, 2004) this population is less likely to be

affected by drought and pumping of wells from the aquifer. Nevertheless, because it represents a separate evolutionarily significant unit, variation at this locality should be maintained through conservation efforts.

Finally, the analysis of molecular variance (AMOVA) showed that while the majority of the genetic variance is found within populations (localities), a substantial amount is found among populations within spring complexes. However, essentially none of the variation was explained by separating localities by spring complexes. This means that although San Marcos Springs is significantly differentiated from localities at Comal Springs, the variation is found within and among populations, and not between spring complexes. Therefore, conservation management decisions should address population-level conservation efforts, with focus on the three evolutionarily significant units at Comal and San Marcos Springs.

Genetic Differentiation among *Heterelmis* Species

Population genetics theory predicts that common species with continuous ranges should have more genetic variation within, and less variation among populations than rare species (Premoli, 1997; Rymer and Ayre, 2006). This is because populations of widespread species are generally larger and more interconnected, and can thereby maintain more genetic variation (Frankham, 1996) than the smaller, isolated populations of rare species (Premoli, 1997). This is especially true for populations of endangered species, such as *H. comalensis*, for whom only a few localities exist, with little or no gene flow between them. In addition, although the common species of *Heterelmis* are thought to use flight and drift dispersal, *H. comalensis* has vestigial wings (Bosse et al., 1988) and

consequently cannot fly to disperse, which makes its restricted range even more vital. The best way to determine whether rare species are experiencing reduced genetic variation due to small population size or restricted range is to compare the genetic variation found in the rare species to that found in a closely related, widespread congener (Rymer and Ayre, 2006). So, we compared the genetic variation in *H. comalensis* to that found in the more widespread flight- and drift-dispersed species of *Heterelmis* at similar geographic distances.

Surprisingly, mtDNA results revealed large amounts of genetic variation in the three variable *H. comalensis* populations when compared to the widespread congeners, specifically when compared to the most common *Heterelmis* species, *H. vulnerata*. For example, West Shoreline, Spring Island, and San Marcos Springs each have haplotype diversity (h) values that equal or exceed those of *H. glabra, H. obesa,* and *H. vulnerata,* and in fact only three other localities sampled had a similar or higher haplotype diversity than the West Shoreline locality. Sequence divergences within the endangered *H. comalensis* were greater than those found in the common *H. vulnerata* and *H. obesa*. Moreover, the within-species sequence divergence for *H. comalensis* is comparable to that found in other common taxa with larger geographic/sampling ranges, looking at the same regions of mtDNA (Austin et al., 2002; Brower and Jeansonne, 2004; Monaghan et al., 2005). Therefore, with respect to closely related congeners and with respect to other invertebrate species, *H. comalensis* is not experiencing reduced genetic variation due to its limited range and limited mobility. These data point to potentially large population sizes in the three genetically variable *H. comalensis* populations.

Calculating female effective population sizes (N_e) as shown in Table 5 revealed that within the three variable *H. comalensis* populations, N_e are as high or higher than the N_e found in *H. glabra*, *H. vulnerata*, and *H. obesa*. In fact, the two variable populations at Comal Springs, West Shoreline and Spring Island, have two of the highest N_e estimated. In addition, although the San Marcos Springs population has a smaller N_e than the two estimated at Comal Springs, it is still larger than most of the *H. vulnerata* N_e, and several of the *H. glabra* and *H. obesa* N_e. However, because recent and drastic reductions in population size may not yet be detectable using mtDNA (Lavery et al., 1996), these results should be accepted cautiously until nuclear sequence markers have corroborated the findings.

The unexpected genetic variation within, and differentiation among, the *H. comalensis* localities is similar to that found in its closely related congener *H. glabra*. Limited phenotypic (Bosse et al., 1988) and among-species sequence divergences between *H. comalensis* and *H. glabra* are explained by recent speciation of the two groups (Hebert et al., 2004). The phylogenetic analysis conducted herein confirmed that *H. glabra* is the nearest relative of *H. comalensis*; however, *H. comalensis* was shown to be monophyletic with moderate bootstrap and high Bayesian support, suggesting that it is significantly divergent from *H. glabra* and does in fact represent a separate species. In addition, the phylogenetic analysis revealed that *H. glabra* is paraphyletic with respect to *H. comalensis*, and the two probably represent sister species, although more sampling within *H. glabra* is needed to confirm this. Estimated divergence times indicate that *H. comalensis* and *H. glabra* diverged about one-half million years ago; however, within a short time each one of the *H. glabra* localities must have become isolated.

Like *H. comalensis*, the variation within, and differentiation among, *H. glabra* populations suggests a history of genetic isolation, with little or no recent gene flow. Some of the highest haplotype diversity values were observed in the *H. glabra* populations. The haplotype network revealed that variation in *H. glabra* is highly geographically structured, with unique haplotypes defining each population. The distances (number of nucleotide differences) among the *H. glabra* haplotypes were as large as, or larger than, the distance between the *H. glabra* and *H. comalensis* haplotypes. All of the *H. glabra* populations are significantly differentiated from one another with moderate bootstrap and high Bayesian support in the phylogenetic analysis conducted herein, and with an average pairwise Φ-statistic of 0.81 across populations. In addition, within-species sequence divergences (p-distances) were highest in *H. glabra*, and because each locality is defined by unique haplotypes this again points to the historical isolation of these populations. These data suggest that each of the *H. glabra* populations has evolved as a separate group.

Multiple lines of evidence suggest that the Fern Bank population is so divergent from the rest of *H. glabra* that while falling into nominal species *H. glabra* according to morphology, it may represent a cryptic species (Pfenninger and Schwenk, 2007). One, Fern Bank is geographically isolated. Two, Fern Bank haplotypes are more than twice as divergent from the other *H. glabra* haplotypes as *H. comalensis* haplotypes are to *H. glabra* haplotypes. Three, sequence divergence between Fern Bank and the other *H. glabra* haplotypes is higher than the average divergence between *H. comalensis* and *H. glabra*, specifically when the same distance is calculated without including Fern Bank. Four, the phylogenetic distance between Fern Bank and the rest of *H. glabra* is greater

than the distance between *H. glabra* and *H. comalensis*, suggesting that it diverged from *H. glabra* before *H. comalensis*. Further sampling from throughout its range is required to elucidate possible cryptic species variation within *H. glabra*.

The high genetic variation and structuring among the *H. glabra* and *H. comalensis* populations suggests that they have all been genetically isolated for an extended time, despite *H. glabra*'s potential ability to fly for dispersal. In general, the *H. comalensis* and *H. glabra* N_e estimates were similar and relatively large, but variable. Finally, there was no evidence of isolation by distance (IBD) among localities within *H. comalensis* and *H. glabra*. AFLP data also showed no indication of isolation by distance among *H. comalensis* populations, because the localities that were clustered together into groups were not the ones that are closest geographically. These data suggest that the *H. comalensis* populations are likely not exchanging migrant individuals, which supports results suggesting limited gene flow among populations (Jensen et al., 2005). In both species, random drift and mutation acting independently in each locality over a long period of time would allow the populations to diverge from one another, resulting in the high levels of divergence among populations that were revealed in this study.

One probable explanation for the differences in genetic divergence between the nuclear ITS and the mitochondrial COI and COII genes is simply differences in evolutionary rates between the two types of DNA. Because mtDNA evolves at a faster rate than nuclear DNA (Avise, 1994), it reflects the more recent evolutionary changes between *H. glabra* and *H. comalensis*, specifically differentiation that has accumulated since they diverged from one another, about 500,000 years ago. However, the slower evolving nuclear ITS region is more reflective of the distant past and genetic differences

within and between *H. glabra* and *H. comalensis* simply have not had enough time to accumulate in this gene.

The other two common *Heterelmis* congeners sampled in this study, *H. obesa* and *H. vulnerata*, are considered to be much more distantly related to *H. comalensis* and *H. glabra* based on current taxonomy (Brown, 1972; Bosse et al., 1988). The phylogenetic analysis confirmed that *H. vulnerata* and *H. obesa* are distantly related to the above species, and to each other. Higher among-species sequence divergences suggest that these two groups have not shared gene flow in a very long time, supporting results from the phylogenetic analysis. In addition, estimated divergence times suggest that *H. comalensis* and *H. glabra* both split from *H. vulnerata* about 5.5 million years ago and from *H. obesa* about 6.1 million years ago, and that *H. vulnerata* also split from *H. obesa* about six million years ago. However, the patterns of genetic variation within *H. obesa* are similar to those found in *H. comalensis* and *H. glabra*. For example, haplotype diversity values in both of the *H. obesa* populations were comparable to those found in *H. comalensis* and *H. glabra*, and the two *H. obesa* populations were significantly differentiated according to pairwise Φ-statistics, suggesting little recent gene flow between them. In addition, estimates of N_e for *H. obesa* were varied, but comparable to estimated N_e for *H. comalensis* and *H. glabra*. However, the lower within-species sequence divergence in *H. obesa* suggests that isolation among populations of this species is a more recent occurrence than in *H. comalensis* and *H. glabra*. Because only two populations of *H. obesa* were sampled, these results must remain tentative until more sampling is done.

Conversely, for localities of the most common species *H. vulnerata*, there was less genetic variation within and differentiation among, than in and among localities of

the other three species, including the endangered *H. comalensis*. Haplotype diversity (h) values for the *H. vulnerata* localities were on average the lowest among the four species, excluding the invariant *H. comalensis* localities. In fact, *H. vulnerata* is the only other species in which some localities lacked mtDNA variation. This is probably due to low population sizes, as estimates revealed that *H. vulnerata* has the smallest N_e on average. The average differentiation among localities of *H. vulnerata* determined by pairwise Φ-statistics was less than half that found among the *H. comalensis* populations, including the four invariant localities. Additionally, within-species sequence divergence was lower in *H. vulnerata* than in *H. comalensis*. These data suggest significantly high levels of gene flow among the widespread *H. vulnerata* localities (Rymer and Ayre, 2006). Unlike *H. comalensis*, *H. vulnerata* is known to use flight dispersal (R. Gibson, pers. comm.), therefore high levels of gene flow among distant localities seems probable. In addition, because *H. vulnerata* localities share high levels of gene flow across a broad geographic range, random drift among these localities would not be as strong as that found in isolated populations and could explain the lower levels of population differentiation found in *H. vulnerata* (Rymer and Ayre, 2006). It should be noted that all samples of *H. vulnerata* collected in this study came from within the Guadalupe River drainage basin, Texas, although the species' range spans parts of Texas, Oklahoma, and possibly Missouri (Brown, 1972; Bosse, 1979; Bosse et al., 1988; Missouri Dept. of Conservation, 2005). Because of the high levels of gene flow among the *H. vulnerata* localities, the Guadalupe River basin could be acting like one large, moderately subdivided population – evident by the low levels of variation within it (Zickovich and Bohonak, 2007). However, moderate

IBD among the *H. vulnerata* localities suggests that despite high gene flow, there are constraints on migration within this species.

Although population genetics theory predicts that widespread species will have more genetic variation within, and less variation among populations than rare species (Frankham, 1996), only half of that prediction was met in this study. There was indeed less variation found among localities of the widespread *H. vulnerata* than among populations of the endangered *H. comalensis*. However, more variation was found within the isolated populations of *H. comalensis* than in the widespread, interconnected localities of *H. vulnerata*. A similar pattern among rare and widespread species has been uncovered in plants (Rymer and Ayre, 2006) and passerine birds (Hansson and Richardson, 2005). The differences in the genetic variation found within and among localities of *Heterelmis* point to strong differences in gene flow patterns among the different species. It seems that *H. comalensis* and *H. glabra* populations have been historically isolated from one another, with little to no recent gene flow. In addition, these species exhibit high levels of genetic variation within populations (excluding the invariant *H. comalensis* localities), which points to large N_e for both *H. comalensis* and *H. glabra*. On the other hand, the common *H. vulnerata* exhibited very low population differentiation, with less genetic variation within localities, suggesting high levels of gene flow among relatively smaller N_e of *H. vulnerata*.

One possible explanation for the differences in gene-flow patterns among the *Heterelmis* species could be linked to differences in niche specialization. Specifically, *H. vulnerata* seem to be found in riverine habitats, which may make drift migration easier and more effective for them since they can locate potential habitat downstream.

Conversely, *H. comalensis* and *H. glabra* seem to be restricted to headwater springs, and might find it difficult to locate suitable springs or seeps in which to inhabit via downstream dispersal, which might limit the amount of gene flow exhibited among populations of these potentially spring-dependent species. The perplexing part of this observation is *H. obesa*, which seems to be found in riverine habitats similar to *H. vulnerata*, but whose patterns of genetic variation within and among populations are similar to *H. comalensis* and *H. glabra*. Although *H. obesa* can probably use drift for dispersal, it is possible that other barriers to gene flow exist for this widespread species.

Morphological Comparison among *Heterelmis* Species

A comparison of the number of hairs found on the inside of the parameres of male genitalia showed that morphological and molecular data are incongruent with the taxonomic key used for species identification. Specifically, the character that distinguishes *H. vulnerata* from *H. glabra* and *H. comalensis* is the lack of an "inner apical fringe of hairs" on the parameres of the male genitalia (6-6'; Bosse et al., 1988). Yet in this study, we found one population of *H. glabra* (according to mtDNA and nuclear data), Fessenden Springs, that also lacked hairs on the male genitalia. In fact, the Fessenden springs population was initially misidentified as *H. vulnerata*. Because the size ranges of these three species overlap, this is the only defining character that distinguishes *H. vulnerata* from *H. glabra*. Therefore, any other *H. glabra* individuals that lack hairs on the inside parameres of the male genitalia would probably be misidentified as well. In addition, although *H. obesa* are easily identified since they are the only species in their size category found in the U.S., the taxonomic key identifies

them as having no hairs on the inside parameres of the male genitalia (Brown, 1972), but this study found that some individuals do indeed possess such hairs.

Potentially even more critical is that the size and color differences between *H. comalensis* and *H. glabra*, used to distinguish the two species (Bosse et al., 1988), are almost impossible to evaluate without using direct comparisons of many samples (personal observation). The only additional difference between the two species in the taxonomic key is that *H. comalensis* is listed as only occurring in Comal Springs, Texas. If *H. comalensis* actually does occur outside of its known range, it may be misidentified as *H. glabra* simply because it could not be distinguished. Data collected in this study demonstrate that the current taxonomic key used for identification of *Heterelmis* species in the U.S. (Brown, 1972; Bosse et al., 1988) may lead to misidentification, especially of *H. glabra* localities. At this time, the only method by which to reliably identify or discriminate between some species is to use molecular analyses. Additional discrete quantitative characters should be identified, potentially in coordination with molecular analyses, to create a reliable taxonomic key for species identification within the *Heterelmis* genus.

One additional note should be made about flight among the *Heterelmis* species. It was noted in this study that although *H. vulnerata*, *H. obesa*, and *H. glabra* are all thought to have normal wings, their wings degrade very quickly after eclosion (personal observation). Only one (*H. obesa*) individual was noted as having fully intact normal wings, and only about 15 out of 152 *H. comalensis* individuals were noted as having intact vestigial wings. Unless a large sample is examined, the chances of recovering individuals with intact wings is low, therefore this character is not generally reliable for

differentiating between *H. glabra* and *H. comalensis*. In addition, although flight may be an important mode of dispersal for the common *H. vulnerata,* the low levels of gene flow (migration) evident among even geographically closer *H. glabra* populations suggests that flight may not be an efficient mode of dispersal for this species, despite its potential ability to fly.

Conclusions

This study showed that despite its limited range and mobility, *H. comalensis* is not genetically depauperate as population genetics theory might predict. In fact, mtDNA variation is exceptionally high in the West Shoreline, Spring Island, and San Marcos Springs populations, specifically when compared to more widespread congeners, suggesting little gene flow among these populations. For this reason, and because of the high levels of differentiation among these three populations, each one should be considered a separate ESU within *H. comalensis*. This means that each one of these populations may need separate management consideration, which may include maintaining genetic isolation among them. However, evidence points to recent and severe bottlenecks in Spring Runs 1 - 3 and Backwater Spring, potentially due to the loss of flow at these localities when water levels are reduced in the Edwards Aquifer. If the springs were to go dry at Comal or San Marcos Springs for extended periods of time, genetic variation in the remaining variable populations could be lost.

Table 1. Population data. Species name, location collected, and sample size (N; for mtDNA samples). MtDNA haplotypes are designated by number, with number of individuals possessing haplotype in parentheses. Haplotype diversity (h; Nei, 1987) is provided for mtDNA data. Haplotypes for the single-copy nuclear gene ITS are designated by letter, with number of individuals possessing haplotype in parentheses. Gene diversity is not reported for ITS because there was no variation detected within populations.

Nominal taxonomic designation	**Population/locality**	**N**	**mtDNA haplotypes (number of individuals)**	**h**	**nuclear ITS haplotypes (number of individuals)**
H. comalensis	Run 1, Comal Springs, Comal Co.	20	1(20)	0.0000	A(2)
	Run 2, Comal Springs, Comal Co.	21	1(21)	0.0000	A(2)
	Run 3, Comal Springs, Comal Co.	21	1(21)	0.0000	A(5)
	Backwater Spring, Comal Springs, Comal Co.	12	1(12)	0.0000	
	Spring Island, Comal Springs, Comal Co.	21	1(13), 2(7), 3(1)	0.5286	A(2)
	West Shoreline, Landa Lake, Comal Co.	29	1(9), 4(7), 5(10), 6(2), 7(1)	0.7463	A(2)
	San Marcos Springs, Hays Co.	28	1(18), 3(1), 8(9)	0.5000	A(2)
H. glabra	Fern Bank, Little Arkansas Spring, Hays Co.	25	9(9), 10(1), 11(14), 12(1)	0.5767	A(2)
	Fessenden (Stockman's) Spring, Kerr Co.	11	13(5), 14(3), 15(2), 16(1)	0.7455	A(2)
	Caroline Spring, Independence Creek, Terrell Co.	10	17(1), 18(3), 19(1), 20(2), 21(1), 22(1), 23(1)	0.9111	A(2)
	Finegan Spring, Devils R., Val Verde Co.	10	24(9), 25(1)	0.2000	A(1)
	Dolan Springs, Val Verde Co.	21	26(7), 27(13), 28(1)	0.5286	A(2)
H. vulnerata	Old Channel, Comal River, Comal Co.	10	29(6), 30(4)	0.5333	B(1)
	San Marcos River, WWTP, Hays Co.	10	29(5), 30(4), 31(1)	0.6444	B(1)

(Table 1 – Continued)

Nominal taxonomic designation	**Population/locality**	**N**	**mtDNA haplotypes (number of individuals)**	**h**	**nuclear ITS haplotypes (number of individuals)**
H. vulnerata	Plum Creek at Hwy 183, Caldwell Co.	19	29(4), 30(14), 32(1)	0.4327	B(2)
	Guadalupe R. off Hwy 183, Gonzales Co.	10	29(3), 30(7)	0.4667	B(2)
	Guadalupe R. off Hwy 474, Kendall Co.	10	29(10)	0.0000	
	Guadalupe R. EAST, off Hwy 474, Kendall Co.	10	29(8), 30(2)	0.3556	
	East Dam, Spring Lake, Hays Co.	1	29(1)	0.0000	
H. obesa	McKittrick Creek Groto, GMNP, Culberson Co.	10	33(5), 34(1), 35(3), 36(1)	0.7111	
	Smith Spring, GMNP, Culberson Co.	11	33(4), 37(7)	0.5091	

Table 2. Pairwise measures of population differentiation within *H. comalensis*. Pairwise Φ – statistics are below the diagonal, and Raymond and Rousset (1995) exact test probabilities (p-values) are above the diagonal, for all *H. comalensis* populations.

	Spring Run 1	Spring Run 2	Spring Run 3	Backwater Spring	Spring Island	West Shoreline	San Marcos
Spring Run 1	0.000	N/A	N/A	N/A	**0.00395**	**<0.0001**	**0.00455**
Spring Run 2	0.000	0.000	N/A	N/A	**0.00350**	**<0.0001**	**0.00225**
Spring Run 3	0.000	0.000	0.000	N/A	**0.00610**	**<0.0001**	**0.00270**
Backwater Spring	0.000	0.000	0.000	0.000	**0.03865**	**0.00115**	**0.04845**
Spring Island	**0.280**	**0.286**	**0.286**	**0.220**	0.000	**<0.0001**	**<0.0001**
West Shoreline	**0.409**	**0.415**	**0.415**	**0.355**	**0.129**	0.000	**<0.0001**
San Marcos	**0.229**	**0.234**	**0.234**	**0.181**	**0.298**	**0.419**	0.000

Bold indicates p ≤ 0.05

Table 3. AMOVA analysis of the seven *H. comalensis* populations sampled within Comal Springs and San Marcos Springs, identifying that genetic variation is partitioned within and among populations, rather than between the two spring complexes. This suggests that conservation efforts may have to focus on population level concerns, rather than focusing on differences between the two locations.

Source of Variation	Df	Sum of Squares	Variance Components	Percentage of Variation
Between Spring Complexes	1	4.245	-0.01290	0.0% (n.s.)
Among Populations within Spring Complexes	5	18.765	0.16809	34.50% *
Within Populations	145	46.162	0.31836	65.5% *
Total	151	69.171	0.47355	

* = $p < 0.05$

Table 4. Average pairwise Φ-statistic and ± s.d. among populations within each of the four *Heterelmis* species.

Species	**Avg. pairwise Φ-statistic**	**Standard Deviation**
H. comalensis	0.20909	± 0.15588
H. glabra	0.81110	± 0.10701
H. vulnerata	0.09330	± 0.33519
*H. obesa**	0.39166	N/A

* only 1 comparison

Table 5. Demographic parameter estimates. Θ (= $2N_e\mu$) with 95% CI, female effective population sizes with 95% CI (N_e, calculated from Θ assuming μ = 1.1% per million years per lineage; Brower, 1994) for all populations with at least one polymorphism using mtDNA data (COI and COII), calculated with FLUCTUATE 1.5 (Kuhner et al., 1995, 1998).

Population	**Θ**	**Lower 95% CI of Θ**	**Upper 95% CI of Θ**	**N_e**	**Lower 95% CI of N_e**	**Upper 95% CI of N_e**
West Shoreline	0.00141758	0.00074451	0.00439605	64435	33841	199820
Spring Island	0.00104512	0.00022869	0.00223712	52256	10395	101687
San Marcos	0.00062014	0.00016513	0.00110141	28188	7506	50064
Fern Bank	0.00088869	0.00049733	0.00290762	40395	22606	132165
Fessenden	0.00151353	0.00046150	0.00585137	68797	20977	265971
Caroline	0.00522482	0.00209226	0.01350684	237492	95103	613947
Finegan	0.00045002	0.00004955	0.00134565	20455	2252	61166
Dolan	0.00136255	0.00054188	0.00256241	61934	24631	116473
Old Channel	0.00026590	0.00004761	0.00069567	12086	2164	31621
WWTP	0.00134667	0.00037266	0.00394039	61212	16939	179109
G.R. Gonzales	0.00034425	0.00001790	0.00130527	15648	814	59330
G.R. East	0.00032675	0.00007242	0.00108202	14852	3292	49183
Plum Creek	0.00063177	0.00013926	0.00131523	28717	6330	59783
McKittrick	0.00162001	0.00049911	0.00492659	73637	22687	223936
Smith	0.00035567	0.00004908	0.00075510	16167	2231	34323

Table 6. Sequence divergence within and among species, using uncorrected p-distances.

Species	***H. comalensis***	***H. glabra***	***H. vulnerata***	***H. obesa***
H. comalensis	0.00303			
H. glabra	0.01092	0.00928		
H. vulnerata	0.12019	0.12194	0.00263	
H. obesa	0.13417	0.13609	0.12998	0.00217

Values on diagonal are within-species sequence divergence.

TABLE 7—Descriptive statistics for the four *Heterelmis* samples (species) on number of hairs found on the inside of the parameres of the male genitalia. Significant differences were found among the four *Heterelmis* species (Kruskal-Wallis: T = 110.86, d.f. = 3, p = <0.0001), and using a multiple-comparison procedure, all species-pairs were shown to be significantly different except *H. vulnerata* and *H. obesa*.

Statistic	***H. vulnerata***	***H. obesa***	***H. glabra***	***H. comalensis***
N	29	28	46	57
Mean	0.0	1.0	13	25
Standard Deviation	0.0	2.0	11	5.0
Maximum	0.0	8.0	32.0	35.0
Minimum	0.0	0.0	0.0	16.0

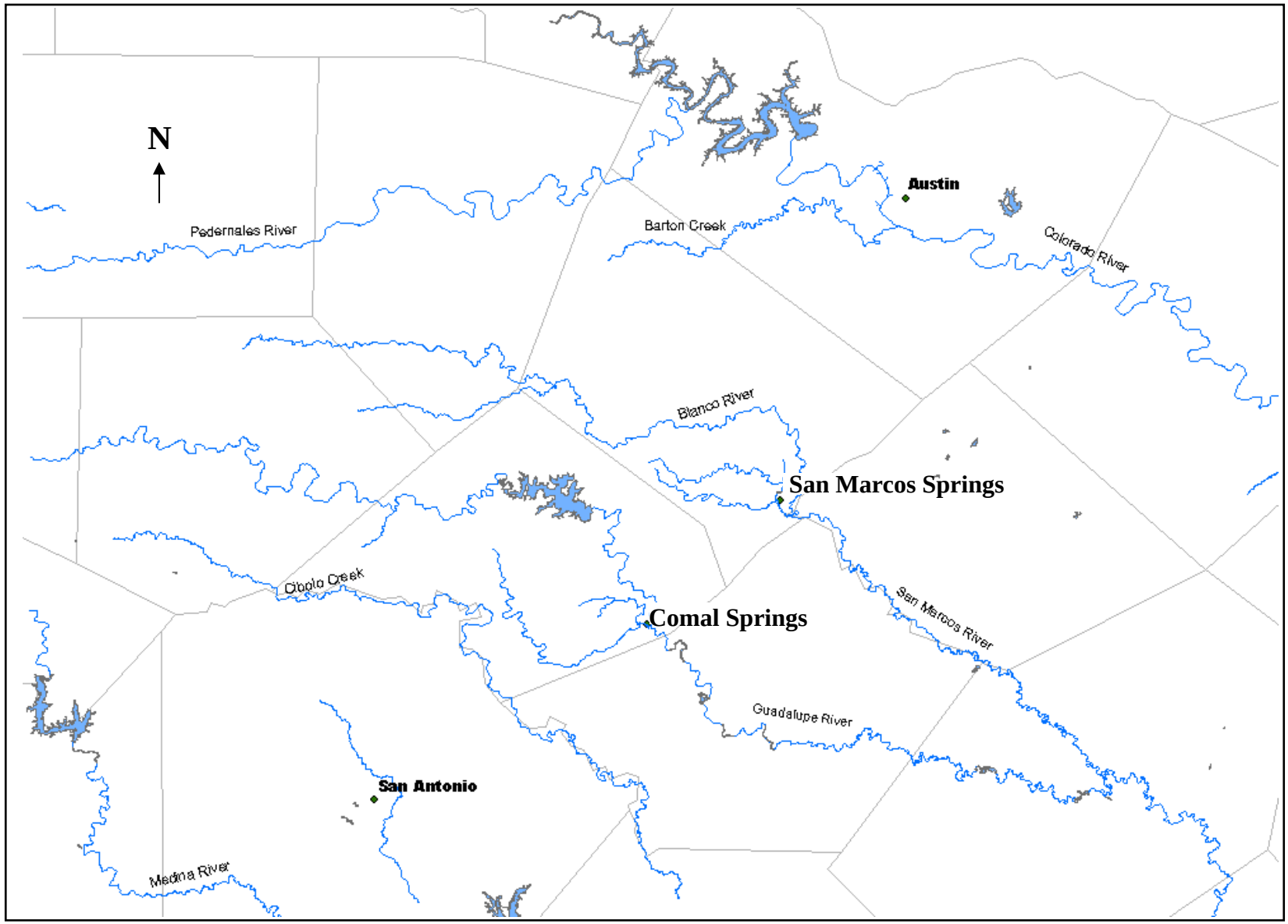

Figure 1. Guadalupe River drainage basin for the Edwards Aquifer, showing Comal Springs and San Marcos Springs locations.

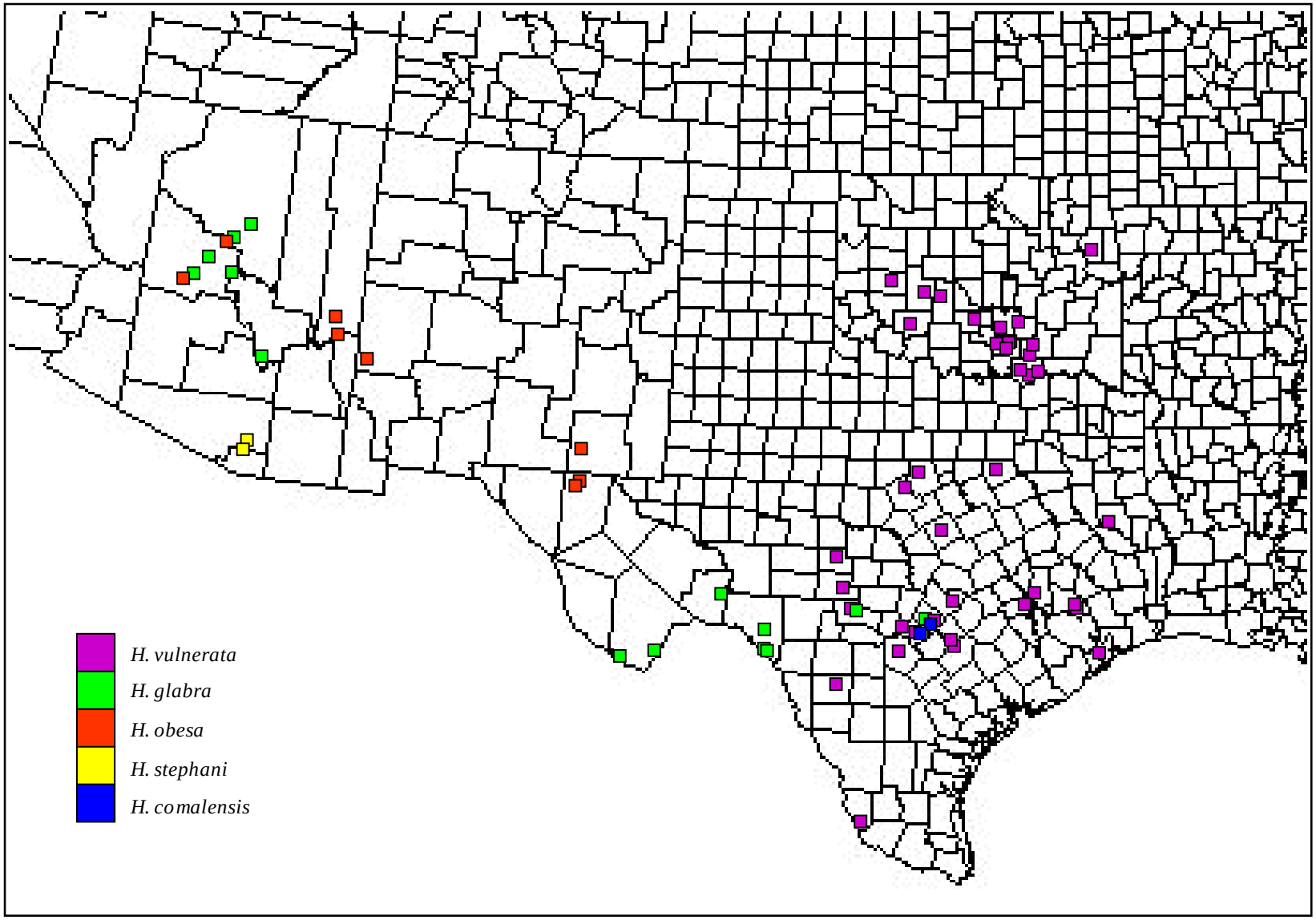

Figure 2. Known habitat locations of *Heterelmis* species within the United States. The most abundant species, *H. vulnerata*, occurs throughout Texas and Oklahoma.

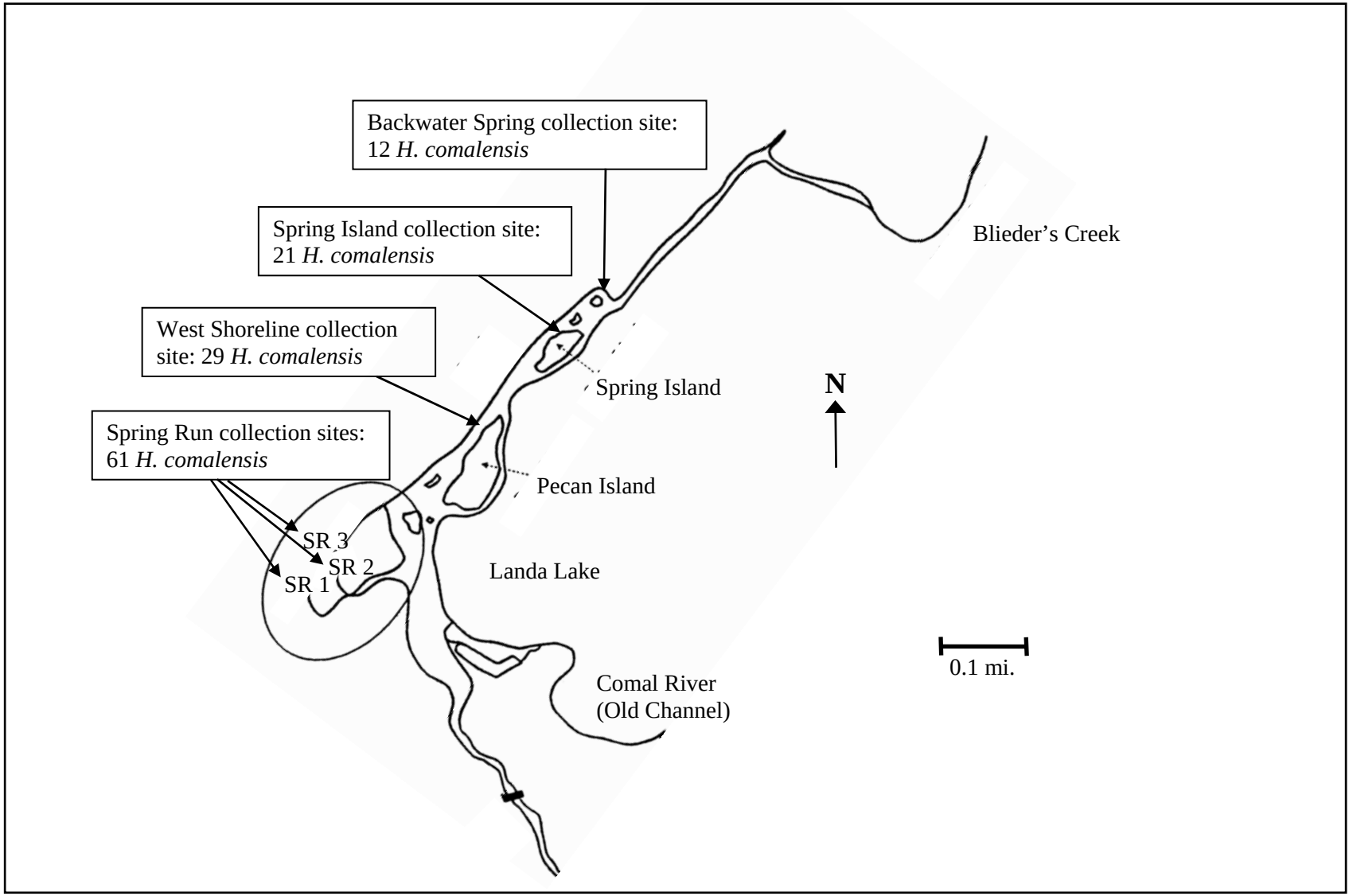

Figure 3. Collection locations of *H. comalensis* and sample sizes at Comal Springs, New Braunfels, Comal County, Texas. Although extensive sampling was conducted throughout Comal Springs and Landa Lake, *H. comalensis* was only found in these six locations.

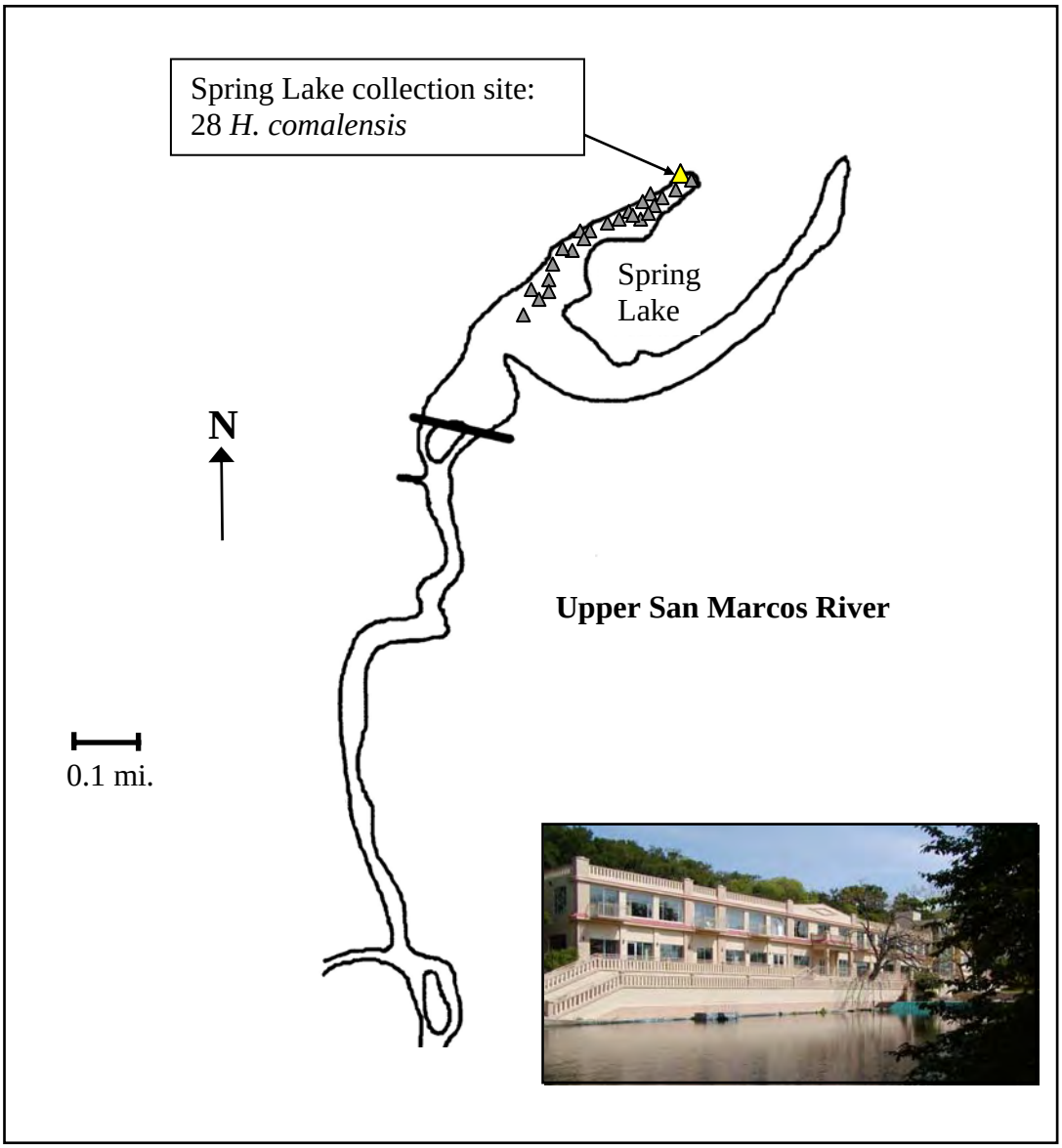

Figure 4. Collection location of *H. comalensis* at San Marcos Springs, and sample size from San Marcos, Hays County, Texas. Although sampling was conducted throughout Spring Lake (gray triangles), *H. comalensis* was only found in one location (yellow triangle). Inset: picture of collection site in front of the San Marcos Hotel and Conference center.

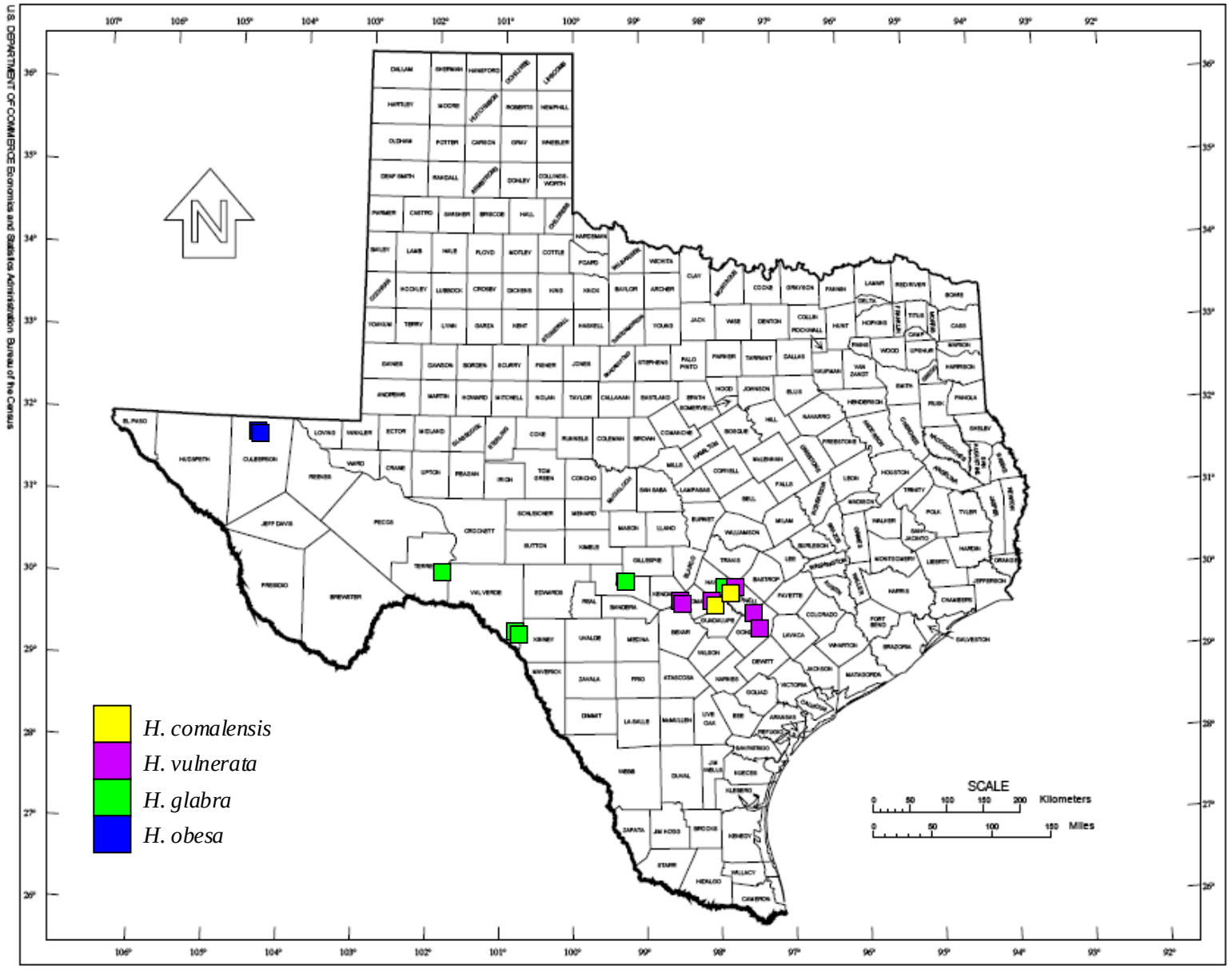

Figure 5. Map of collection sites for *Heterelmis* species across central and western Texas.

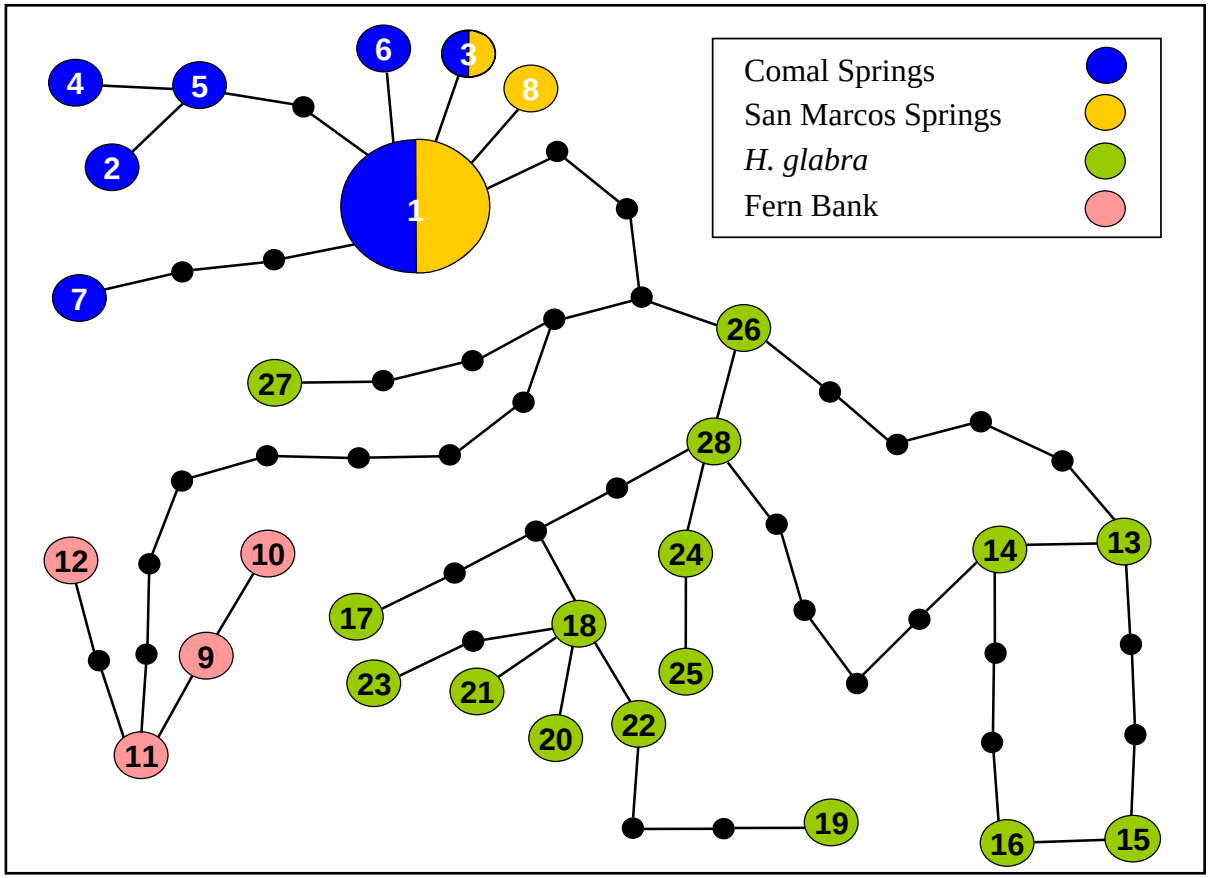

Figure 6. Haplotype network for *H. comalensis* and *H. glabra* populations. In a haplotype network each unique haplotype is represented by a numbered circle (size somewhat proportional to frequency). Haplotypes are connected by lines of equal length. The lines each represent one nucleotide difference among haplotypes, and the small black circles represent a missing or unsampled haplotype. The circles representing haplotypes are colored differently for Comal and San Marcos Springs to show which haplotypes are differentiated among locations (unique versus shared).The Fern Bank haplotypes have also been colored differently from the rest of *H. glabra* to show the distance (number of nucleotide changes) between that population and the other *H. glabra* populations.

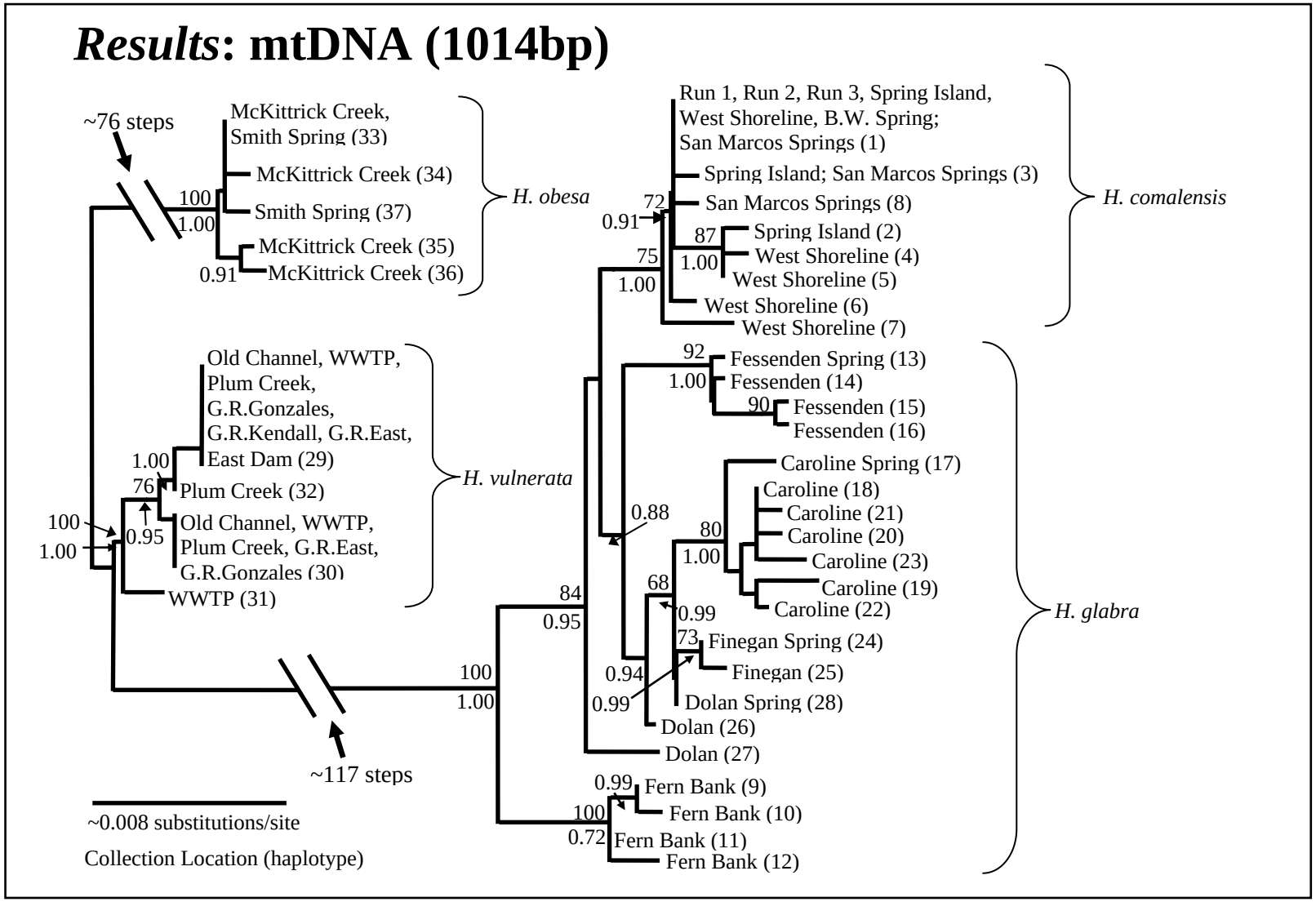

Figure 7. Phylogenetic hypothesis for *Heterelmis.* Numbers in parentheses indicate haplotype number. Above branches are the neighbor-joining bootstrap support values, below branches are the Bayesian Maximum likelihood posterior probabilities.

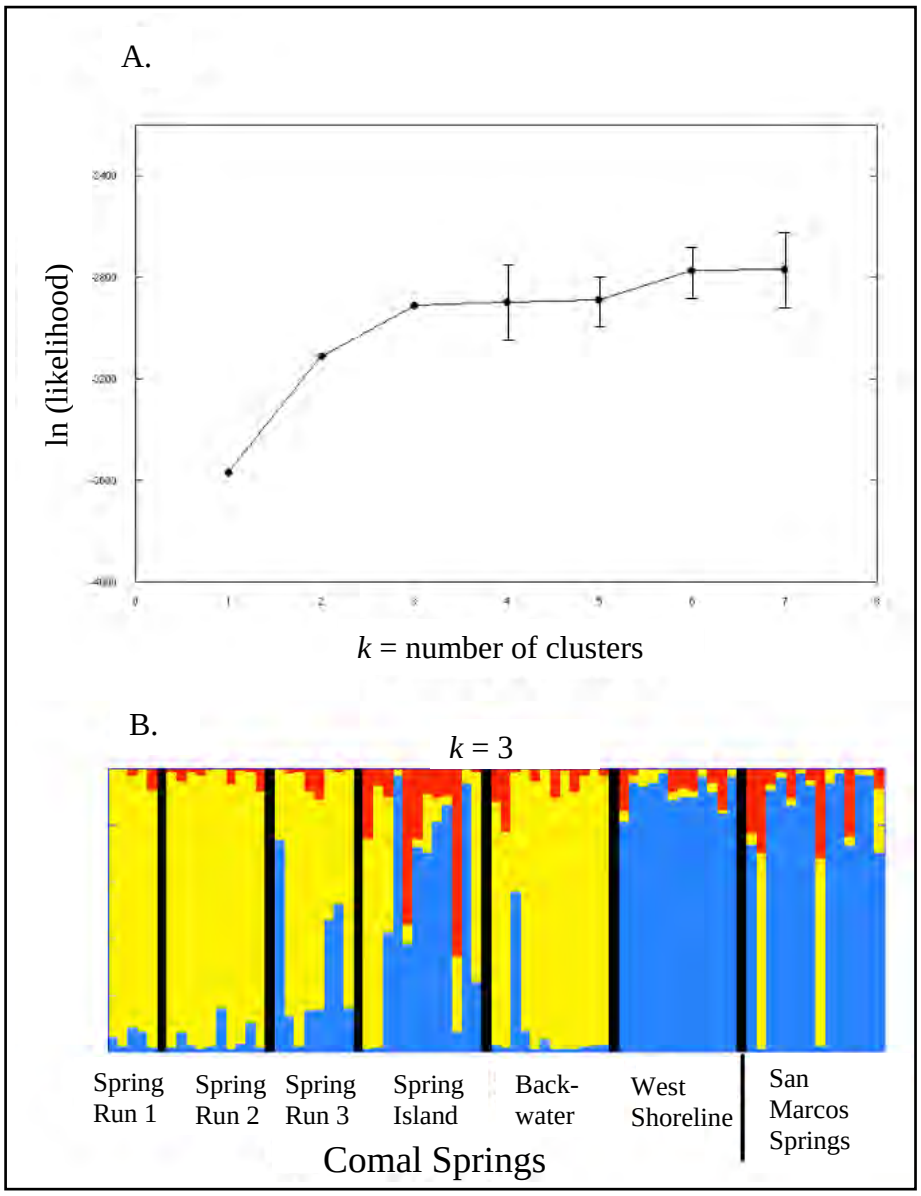

Figure 8: AFLP data. A) Plot of mean log likelihood from 10 MCMC runs vs. *k* = the number of inferred clusters. Lines represent standard deviations. B) Barplot for three clusters. Each vertical bar represents a single individual and is colored in accordance with the Bayesian estimate of the proportion of that individuals genome that originated in a given cluster or population based on STRUCTURE v2.2 under the admixture model.

LITERATURE CITED

Adams, B. J., A. M. Burnell, and T. O. Powers. 1998. A phylogenetic analysis of the genus Heterorhabditis (Nemata: Rhabditidae) based on internal transcribed spacer 1 DNA sequence data. Journal of Nematology 30:22-39.

Austin, J. W., A. L. Szalanski, P. Uva, A. Bagnères, and A. Kence. 2002. A comparative genetic analysis of the subterranean termite genus *Reticulitermes* (Isoptera: Rhinotermitidae). Entomological Society of America 95:753-760.

Avise, J. C. 1994. Molecular markers, natural history and evolution. Chapman & Hall, New York.

Bensch, S., and M. Åkesson. 2005. Ten years of AFLP in ecology and evolution: why so few animals? Molecular Ecology 14:2899-2914.

Bosse, L. S. 1979. A survey of the adult Dryopoids (Coleoptera) in the San Marcos and Comal Rivers in central Texas—Thesis. Southwest Texas State University, San Marcos, Texas.

Bosse, L. S., D. W. Tuff, and H. P. Brown. 1988. A new species of *Heterelmis* from Texas (Coleoptera: Elmidae). Southwestern Naturalist 33:199-203.

Brower, A. V. Z. 1994. Rapid morphological radiation and convergence among races of the butterfly *Heliconius erato* inferred from patterns of mitochondrial DNA evolution. Proceedings of the National Academy of Sciences of the United States of America 91:6491-6495.

Brower, A. V. Z., and M. M. Jeansonne. 2004. Geographical populations and "subspecies" of New World Monarch butterflies (Nymphalidae) share a recent origin and are not phylogenetically distinct. Annals of the Entomological Society of America 97:519-523.

Brown, H. P. 1972. Synopsis of the genus *Heterelmis* Sharp in the United States, with description of a new species from Arizona (Coleoptera, Dryopoidea, Elmidae). Entomological News 83:229-238.

Brown, H. P. 1987. Biology of riffle beetles. Annual Review of Entomology 32:253-273.

Bürger, R., and M. Lynch. 1995. Evolution and extinction in a changing environment: a quantitative-genetic analysis. Evolution 49:151-163.

Caterino, M. S., and F. A. H. Sperling. 1999. *Papilio* phylogeny based on mitochondrial cytochrome oxidase I and II genes. Molecular Phylogenetics and Evolution 11:122-137.

Clement, M., D. Posada, and K. A. Crandall. 2000. TCS: a computer program to estimate gene genealogies. Molecular Ecology 9:1657-1659.

Conover, W. J. 1999. Practical nonparametric statistics, 3rd ed. Wiley, New York.

Crandall, K. A., O. R. P. Bininda-Emonds, G. M. Mace, and R. K. Wayne. 2000. Considering evolutionary processes in conservation biology. Trends in Ecology and Evolution 15:290-295.

Edwards Aquifer Authority. 2007. Variable flow study seven years of monitoring and applied research. August report.

Excoffier, L., P. E. Smouse, and J. M. Quattro. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131:479-491.

Excoffier, L., G. Laval, and S. Schneider. 2005. Arlequin ver. 3.0: An integrated software package for population genetics data analysis. Evolutionary Bioinformatics Online 1:47-50.

Falush, D., M. Stephens, and J. K. Pritchard. 2007. Inference of population structure using multilocus genotype data: dominant markers and null alleles. Molecular Ecology Notes 7:574-578.

Forister, M. L., C. C. Nice, J. A. Fordyce, Z. Gompert, and A. M. Shapiro. 2008. Considering evolutionary processes in the use of single-locus genetic data for conservation, with examples from the Lepidoptera. Journal of Insect Conservation 12:37-51.

Frankham, R. 1996. Relationship of genetic variation to population size in wildlife. Conservation Biology 10:1500-1508.

Frankham, R. 2005. Genetics and extinction. Biological Conservation 126:131-140.

Gibson, J. R., S. J. Harden, and J. N. Fries. 2008. Survey and distribution of invertebrates from selected springs of the Edwards Aquifer in Comal and Hays Counties, Texas. Southwestern Naturalist 53:74-84.

Goldstein, P. Z., R. DeSalle, G. Amato, and A. P. Vogler. 2000. Conservation genetics at the species boundary. Conservation Biology 14:120-131.

Gompert, Z., C. C. Nice, J. A. Fordyce, M. L. Forister, and A. M. Shapiro. 2006a. Identifying units for conservation using molecular systematics: the cautionary tale of the Karner blue butterfly. Molecular Ecology 15:1759-1768.

Gompert, Z., J. A. Fordyce, M. L. Forister, A. M. Shapiro, C. C. Nice. 2006b. Homoploid hybrid speciation in an extreme habitat. Science 314:1923-1925.

Haig, S. M. 1998. Molecular contributions to conservation. Ecology 79:413-425.

Hall, B. G. 2001. Phylogenetic trees made easy. A how-to manual for molecular biologists. Sinauer Associates, Sunderland, Massachusetts.

Hansson, B., and D. S. Richardson. 2005. Genetic variation in two endangered *Acrocephalus* species compared to a widespread congener: estimates based on functional and random loci. Animal Conservation 8:83-90.

Hebert, P. D. N., E. H. Penton, J. M. Burns, D. H. Janzen, and W. Hallwachs. 2004. Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. Proceedings of the National Academy of Sciences 101:14812-14817.

Hedrick, P. W., and P. S. Miller. 1992. Conservation genetics: techniques and fundamentals. Ecological Applications 2:30-46.

Hedrick, P. W., and S. T. Kalinowski. 2000. Inbreeding depression in conservation biology. Annual Review of Ecology and Systematics 31:139-162.

Jensen, J. L., A. J. Bohonak, and S. T. Kelley. 2005. Isolation by distance, web service. BMC Genetics 6:13. v3.15 http://ibdws.sdsu.edu/

Johnson, S. B., and G. M. Schindel. 2008. Evaluation of the option to designate a separate San Marcos pool for critical period management. Edwards Aquifer Authority report number 08-01.

Juan, C., K. M. Ibrahim, P. Oromi, and G. M. Hewitt. 1998. They phylogeography of the darkling beetle, *Hegeter politus*, in the eastern Canary Islands. Proceedings of the Royal Society of London B 265:135-140.

Kuhner, M. K., J. Yamato, and J. Felsenstein. 1995. Estimating effective population size and mutation rate for sequence data using Metropolis-Hastings sampling. Genetics 140:1421-1430.

Kuhner, M. K., J. Yamato, and J. Felsenstein. 1998. Maximum likelihood estimation of population growth rates based on coalescent. Genetics 149:429-434.

Kuhner, M. K., J. Yamato, and P. Beerli. 2005. Fluctuate documentation. Accessed at http://evolution.genetics.washington.edu/lamarc/fluctuate.html

Lacy, R. C. 1997. Importance of genetic variation to the viability of mammalian populations. Journal of Mammalogy 78:320-335.

Lavery, S., C. Moritz, and D. R. Fielder. 1996. Indo-Pacific population structure and evolutionary history of the coconut crab *Birgus latro*. Molecular Ecology 5:557-570.

LBG-Guyton Associates. 2004. Evaluation of augmentation methodologies in support of in-situ refugia at Comal and San Marcos Springs, Texas. Edwards Aquifer Authority report, accessed April 29, 2008, at http://www.edwardsaquifer.org/pdfs/Reports/AS%20Reports/Recharge%20Enhancement%20Studies/Spring%20Aug%20Final%20Report/Spring%20Aug%20Final%20Report.pdf

Maddison, W. P., and D. R. Maddison. 1992. MacClade, version 3.0. Sinauer, Sunderland, Mass.

Missouri Department of Conservation. 2005. Fisheries research benthic collection, Table 30. Benthic macroinvertebrate collections for the Gasconade River from 1962-92. Accessed April 15, 2006, from http://mdc.mo.gov/fish/watershed/gascon/biotic/130bct30.htm

Monaghan, M. T., M. Balke, T. R. Gregory, and A. P. Vogler. 2005. DNA-based species delineation in tropical beetles using mitochondrial and nuclear markers. Philosophical Transactions of the Royal Society of London B 360:1925-1933.

Moritz, C. 1999. Conservation units and translocations: strategies for conserving evolutionary processes. Hereditas 130:217-228.

Moritz, C. 2002. Strategies to protect biological diversity and the evolutionary processes that sustain it. Systematic Biology 51:238-254.

Mueller, U. G., and L. R. Wolfenbarger. 1999. AFLP genotyping and fingerprinting. Trends in Ecology and Evolution 14:389-394.

Nei, M. 1987. *Molecular evolutionary genetics*. Columbia University Press, New York.

Nice, C. C., N. Anthony, G. Gelembiuk, D. Raterman, and R. ffrench-Constant. 2005. The history and geography of diversification within the butterfly genus *Lycaeides* in North America. Molecular Ecology 14:1741-1754.

Pearman, P. B. 2001. Conservation value of independently evolving units: sacred cow or testable hypothesis? Conservation Biology 15:780-783.

Pfenninger, M., and K. Schwenk. 2007. Cryptic animal species are homogeneously distributed among taxa and biogeographical regions. BMC Evolutionary Biology 7:121-126.

Posada, D., and K. A. Crandall. 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics 14:817-818.

Powers, T. O., T. C. Todd, A. M. Burnell, P. C. B. Murray, C. C. Fleming, A. L. Szalanski, B. A. Adams, and T. S. Harris. 1997. The internal transcribed spacer region as a taxonomic marker for nematodes. Journal of Nematology 29:441-450.

Premoli, A. C. 1997. Genetic variation in a geographically restricted and two widespread species of South American *Nothofagus*. Journal of Biogeography 24:883-892.

Pritchard, J. K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. Genetics 155:945-959.

Rand, D. M. 1996. Neutrality tests of molecular markers and the connection between DNA, polymorphism, demography, and conservation biology. Conservation Biology 10:665-671.

Raymond, M., and F. Rousset. 1995. An exact test for population differentiation. Evolution 49:1280-1283.

Rozas, J., J. C. Sánchez-DelBarrio, X. Messeguer, and R. Rozas. 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics 19:2496 2497.

Rymer, P. D., and D. J. Ayre. 2006. Does genetic variation and gene flow vary with rarity in obligate seeding *Persoonia* species (Proteaceae). Conservation Genetics 7:919-930.

Sanderson, M. W., and H. P. Brown. 1956. Records of riffle beetles and a description of the larva of *Hexacylloepus* (Coleoptera; Dryopoidea). Proceedings of the Oklahoma Academy of Sciences 37:67-76.

Swofford, D. L. 2002. PAUP* Phylogenetic analysis using parsimony (*and other methods), Version 4.0b10. Sinauer Associates, Sunderland, Massachusetts.

Szalanski, A. L., D. S. Sikes, R. Bischof, and M. Fritz. 2000. Population genetics and phylogenetics of the endangered American Burying Beetle, *Nicrophorus americanus* (Coleoptera: Silphidae). Annals of the Entomological Society of America 93:589594.

Szalanski, A. L., and C. B. Owens. 2003. Genetic variation of the southern corn rootworm, (Coleoptera: Chrysomelidae). Florida Entomologist 86:329-333.

Templeton, A. R., K. A. Crandall, and C. F. Sing. 1992. A cladistics analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. Genetics 132:619-633.

United States Fish and Wildlife Service (USFWS). 1973. The endangered species act of 1973. Accessed on October 19, 2004, at http://endangered.fws.gov/esa.html

United States Fish and Wildlife Service (USFWS). 1997. Endangered and threatened wildlife and plants; final rule to list three aquatic invertebrates in Comal and Hays Counties, TX, as endangered. Accessed April 15, 2006, at http://www.fws.gov/endangered/r/fr97629.html

Vogler, A. P., and R. DeSalle. 1994. Diagnosing units of conservation management. Conservation Biology 8:354-363.

Vos, P., R. Hogers, M. Bleeker, M. Reijans, T. van de Lee, M. Hornes, A. Friters, J. Pot, J. Palemena, M. Kuiper, and M. Zabeau. 1995. AFLP: a new technique for DNA fingerprinting. Nucleic Acids Research 23:4407-4414.

Watterson, G. A. 1975. On the number of segregating sites in genetical models without recombination. Theoretical Population Biology 7:256-276.

Zickovich, J. M., and A. J. Bohonak. 2007. Dispersal ability and genetic structure in aquatic invertebrates: a comparative study in southern California streams and reservoirs. Freshwater Biology 52:1982-1996.

VITA

Tina Katherine Gonzales was born in Beloit, Wisconsin on December 25, 1975, the daughter of Raul Gonzales and Loraine Reed. She graduated from Beloit Memorial High School in 1994. She received the degree of Bachelor of Science at Texas State University-San Marcos in May 2005. In January 2006, she entered the Graduate College of Texas State.

Permanent Address: 1213 Elm St.

Beloit, WI 53511

This thesis was typed by Tina K. Gonzales.