Genomic Assessment of Taxonomic Status of Central Texas *Eurycea* Salamanders

Principal Investigator: David M. Hillis
Department of Integrative Biology
University of Texas
Austin, TX 78712
Phone: 512-471-5792
E-mail: dhillis@austin.utexas.edu

Co-Principal Investigator: David C. Cannatella
Department of Integrative Biology
University of Texas
Austin, TX 78712
Phone: 512-232-4862
E-mail: catfish@utexas.edu

Additional Personnel: Thomas J. Devitt
Department of Integrative Biology
University of Texas
Austin, TX 78712

current address:
Environmental Scientist
Watershed Protection Department
City of Austin
Austin, Texas 78704

April M. Wright
Department of Integrative Biology
University of Texas
Austin, TX 78712

Abstract

We conducted phylogenetic analyses of all known described and undescribed species of central Texas neotenic *Eurycea* (based on analyses of 34,518 loci), and population STRUCTURE analyses of all known populations of *E. tonkawae*, *E. naufragia*, and *E. chisholmensis* (based on analyses of 75,296 loci). Phylogenetic analysis supported the monophyly and genetic distinctiveness of all described species in the group; in particular, *E. tonkawae*, *E. naufragia*, and *E. chisholmensis* from north of the Colorado River were each reciprocally monophyletic, significantly supported, and genetically distinct. Population STRUCTURE analyses of the species north of the Colorado River also supported the species boundaries between *E. tonkawae*, *E. naufragia*, and *E. chisholmensis*. A few populations that have been provisionally assigned to *E. naufragia* were determined here to belong to *E. tonkawae* (Georgetown Springs) or *E. chisholmensis* (all known populations north of Lake Georgetown). The analyses show strong support for three additional undescribed species from south of the Colorado River. One of these has been reported previously as an undescribed species (the Pedernales Salamander); we extend the range of this undescribed species to include a cave on the west side of the Pedernales River. The two remaining undescribed species occur in west-central and western portions of the Edwards Plateau. We present georeferenced locality data and distribution maps for all the species of Texas neotenic *Eurycea*.

Introduction

The spring and cave salamanders of central Texas (genus *Eurycea*) represent a major radiation of distinct and endemic Texas amphibians. The species in this group have highly restricted habitat requirements, and each described species is restricted to specific aquifers and river systems. Many of these aquifers are under pressure from development, which can result in decreased water quality and quantity in surface flows. As a result of these pressures, several of the endemic Texas *Eurycea* are listed as Federal Threatened or Endangered species by the U.S. Fish and Wildlife Service.

Although the taxonomic status of many Texas *Eurycea* is well established (Chippindale et al., 2000; Hillis et al., 2001), the species boundaries among several of the protected species have recently been questioned (Forstner, 2012). The systematic revisions of this group have been based on a combination of morphological, allozyme, and DNA sequencing studies. However, these salamanders are subject to rapid morphological change in response to environmental factors, especially related to the differences between surface and cave habitats (Wiens et al., 2003). For example, the enormous morphological differences between the subclade *Typhlomolge* (Blind Salamanders) and epigean species (those commonly found at surface springs) led Wake (1966) to suggest that these two groups were not closely related, but rather the result of two independent dispersal events into Texas. However, molecular studies have made it clear that *Typhlomolge* is actually much

2

more closely related to epigean species south of the Colorado River than it is to the epigean species north of the river (Hillis et al., 2001).

The three described species north of the Colorado River are highly distinct from the remaining species in studies of allozymes and mitochondrial DNA, but a few populations have been assigned to described species only provisionally (Chippindale et al., 2000). In addition, mitochondrial DNA sequences represent only a single genetic locus, and capture of mitochondrial DNA between species can make use of this marker for species assignment problematic at the individual and population level (Avise, 2004). This is especially true in groups in which hybridization is known to occur, as it is in Texas *Eurycea* (Sweet, 1984). Although the great genetic divergence among *Eurycea* species north of the Colorado River (in Travis, Williamson, and Bell counties) clearly shows that at least three species inhabit this area, additional data are needed to clarify the status of some of the populations. This is especially true for numerous cave populations in this area, which are often morphologically divergent. Whether this morphological diverence represents local adaptation or additional speciation is currently unresolved.

There have been rapid advances in genomic sequencing technologies since the taxonomic revision of this group by Chippindale et al. (2000). In addition, many new populations (and perhaps new species) of Texas *Eurycea* have been collected in the past decade. Here we use these new technologies to examine thousands of nuclear DNA loci across the range of Texas *Eurycea*, and also conduct detailed population analyses of the federally threatened species of *Eurycea* north of the Colorado River (*E. tonkawae, E. naufragia*, and *E. chisholmensis*). This analysis tests the current taxonomy of the group, and examines the status and species boundaries of the three described species that occur north of the Colorado River. We also use these analyses to examine the status of several cave populations that have been provisionally assigned to these species.

Objective

To examine and test the taxonomic status the described species of Texas *Eurycea*, with particular emphasis on populations of the proposed endangered [or threatened] species, *E. tonkawae, E. naufragia*, and *E. chisholmensis*, using a broad survey of nuclear DNA markers from hundreds of independent loci.

[Note: the Objective statement above is taken directly form the original grant proposal, as requested in the Report Guidelines. We proposed to examine at least "hundreds" of loci. In our final analyses, we were able to examine 34,518 loci for the phylogenetic analyses, and 75,296 loci for the STRUCTURE analyses.]

Geographic Location

See Appendix Table 1 for a georeferenced list of all specimens examined. These

included samples from all known species, three undescribed species (here called *Eurycea* sp. 1, sp. 2, and sp. 3), and samples from all available localities of *E. tonkawae*, *E. naufragia*, and *E. chisholmensis* (the three species that occur north of the Colorado River, which were the focus of the study). We examined samples from throughout the range of the Texas neotenic *Eurycea* clade, from Bell County in the northeast to Val Verde County in the southwest. We excluded two individuals from the genomic analyses due to low data yield (Pecan Springs and Kickapoo Springs) and one because of problematic labeling of the sample (Panther Canyon).

Methods

Our focus was on the species north of the Colorado River: *E. chisholmensis*, *E. naufragia*, and *E. tonkawae*, although we compared all known species in the Texas neotenic *Eurycea* clade to compare in a phylogenetic analysis. All localities have been georeferenced (Appendix Table 1), and all associated data, specimens, and remaining tissue samples are deposited in the Genetic Diversity Collection of the Biodiversity Collections, Department of Integrative Biology, University of Texas.

Molecular Methods

We used double-digest restriction-site associated DNA sequencing (ddRADseq) (Peterson et al., 2012). This approach employed two restriction enzymes to cleave DNA from the entire genome of one individual into fragments. These are then separated by size into a library that is a reduced representation of the genome containing fragments of a specified length (in this case, 300–400 nucleotides). These DNA fragments were then sequenced (Illumina platform) in millions of parallel reactions, producing many copies of each fragment that were then aligned and analyzed using a software pipeline specifically designed to handle RAD data (Catchen et al., 2013). The data examined were single nucleotide polymorphisms (SNPs) sampled from throughout the genome of respective species.

Bioinformatics Pipeline

We examined a range of values for five key parameters in the STACKS 1.30 pipeline. These parameters are presented in Appendix Table 2. For each parameter, we tested at least three values, including the default value for values that have recommended default. We selected parameter values for use in the final pipeline if that value both minimized the total missing data and the number of taxon pairs for which there is no data in common. The final step in the STACKS pipeline involved exporting the matrix of retained SNPs. At this step, it was possible to exclude incomplete SNPs (SNPs for which one or more individuals do not have data) from the final phylogenetic data matrix. Our final data matrices included 34,518 loci for the phylogenetic analyses (all populations), and 75,296 loci for the population Structure analyses (populations north of the Colorado River).

Phylogenetic Analysis

We fit a model of DNA sequence evolution to each of the data matrices using jModelTest (Posada and Buckley, 2004). The best-fit model was the General Time-Reversible (GTR) model with Gamma-distributed rate variation. We performed phylogenetic estimation in four different software packages on each of the five final SNP matrices. The first two analyses were maximum likelihood methods. We used Garli (Zwickl, 2006) to obtain a point estimate of topology for each data matrix. Because SNP datasets inherently exclude invariant sites, we also used RAxML (Stamatakis, 2006) to estimate a tree, as RAxML has recently implemented corrections for this type of ascertainment bias (Leache et al., 2015). We estimated a sample of trees in MrBayes (Ronquist et al., 2012), and summarized this sample with a majority-rule consensus tree to produce posterior probability support vales for each branch. The MCMC chain was run for 10 million generations, and checked for convergence using Tracer (Rambaut et al., 2014). Finally, we used PAUP* version 4.0a146 (Swofford, 2003; updated 2015) to conduct the SVD Quartets method of Chifman and Kubatko (2014) for constructing species trees from gene trees that evolve under the coalescent process.

Population STRUCTURE Analysis

We used Bayesian clustering models implemented in STRUCTURE 2.3.4 (Pritchard et al., 2000; Falush, 2003; Falush et al., 2007) to infer the number of populations (K) and estimate individual population assignments. For populations north of the river, we explored values of K from 1–5. We assumed correlated allele frequencies and admixture among populations using the default settings. Three replicate runs were conducted for each value of K in STRUCTURE, with each run consisting of 50,000 sweeps after a burn-in period of 10,000 sweeps. We assessed MCMC convergence by comparing K across replicate runs. We determined the best K by examining both the log probability of the data (log Pr($X|K$)) and the ΔK statistic following Evanno et al. (2005). Results were summarized and visualized using the web program CLUMPAK (Kopelman et al., 2015).

Results

Data Matrix Assembly

Our phylogenetic matrix contained 34,518 SNPs. When assembling a phylogenetic matrix, STACKS distills the dataset down to a single representative from each population. Therefore, some sites were not utilized in the phylogenetic analysis if they were polymorphic within populations. Output for STRUCTURE did not include this filtering, so each individual in the data matrix was equivalent to one individual organism. Therefore, the STRUCTURE matrices are much larger (75,296 loci) than

the phylogenetic matrices. All data matrices will be released in public data repositories upon final publication of this study.

Phylogenetic Analyses

All of our phylogenetic analyses produced very similar results, and differed only in the relationships of populations within species. Here we show two representative trees: The tree from the SVD Quartet analysis to show of summary species tree (Figure 1), and a tree from the Bayesian analysis to show the posterior probability support values for each branch (Figure 2). All phylogenetic analyses provided strong support for the recognized species of *Eurycea*; in particular, each of the three described species from north of the Colorado River was significantly supported (Figure 2). As in previous studies (e.g., Chippindale et al., 2000; Hillis et al., 2001), our phylogenetic trees indicated the deepest deep split in the Texas neotenic *Eurycea* from north and south of the Colorado River. North of the Colorado River, where our sampling was the most thorough, all analyses showed strong and significant support for groups referable to *E. tonkawae*, *E. naufragia*, and *E. chisholmensis* (Figures 1 and 2).
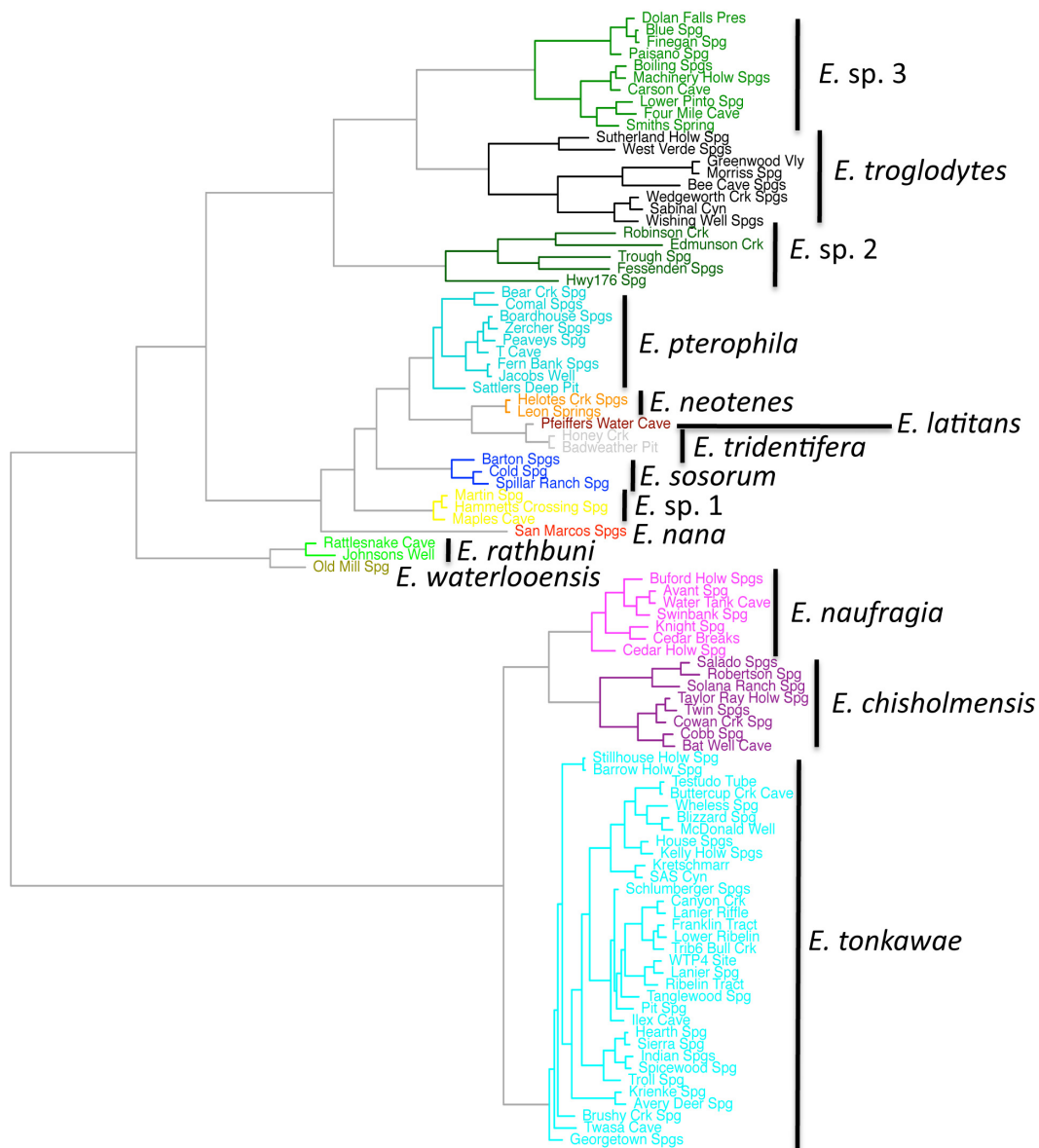
Figure 1. Species tree from the SVD Quartets analysis. Species are indicated by colors. The scale bar indicates 100 nucleotide changes.
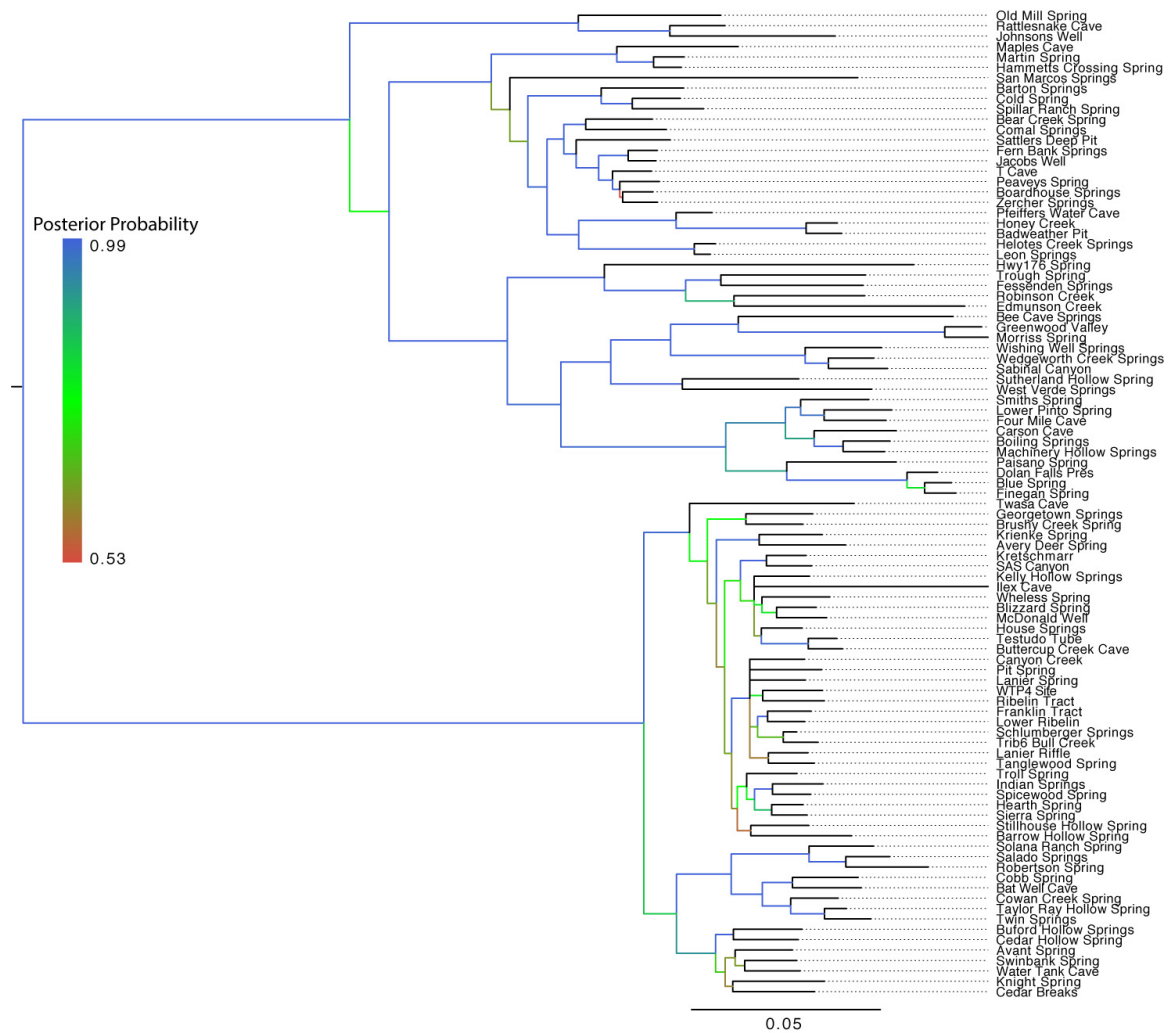
Figure 2. Phylogenetic tree from the Bayesian analysis. Posterior probability support values for each branch are indicated by the color scale; blue branches show statistically significant support (PP > 95%).

Population Structure Analysis

The optimal value of *K* (number of genetically distinct groups) for the populations of *Eurycea* north of the Colorado River was 3, which corresponds to the three described species *E. tonkawae*, *E. naufragia*, and *E. chisholmensis* (Figure 3). A few populations showed the presence of genes that were otherwise found primarily in another species; these genes are likely ancestral retained polymorphisms (polymorphisms retained from a common ancestor of the two respective species). However, these genes could be also be indicative of low levels of gene flow between some adjacent populations. To distinguish between these two possibilities, it would be necessary to see if the "misplaced" genes in a population form linked blocks in the genome. We did not examine linkage among any of the loci, so our data are insufficient to distinguish these two possibilities.

8

Figure 3. Structure analysis (*K* = 3) of *Eurycea* from north of the Colorado River, based on 75,296 loci. Vertical bars represent individuals; black lines separate populations. The three colors represent the three major genetic divisions, and correspond to populations of the three described species. Gray = *E. tonkawae*; Red = *E. naufragia*; Blue = *E. chisholmensis*.

Species Status and Geographic Distribution

Figure 4 presents the distribution of all populations of central Texas neotenic *Eurycea* examined in this study. Figure 5 presents the details of the distribution of *E. tonkawae*, *E. naufragia*, and *E. chisholmensis* from north of the Colorado River.
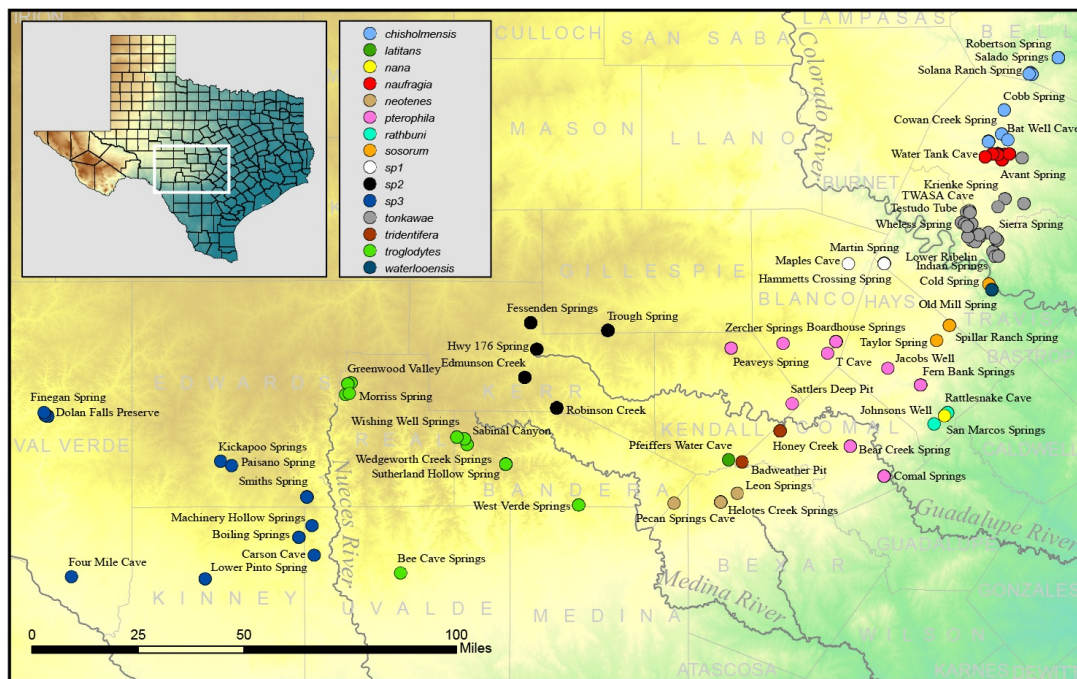

Figure 4. Distribution of all described and known but undescribed species of Texas neotenic *Eurycea*. Georeferenced data for all sample localities are presented in Appendix Table 1.
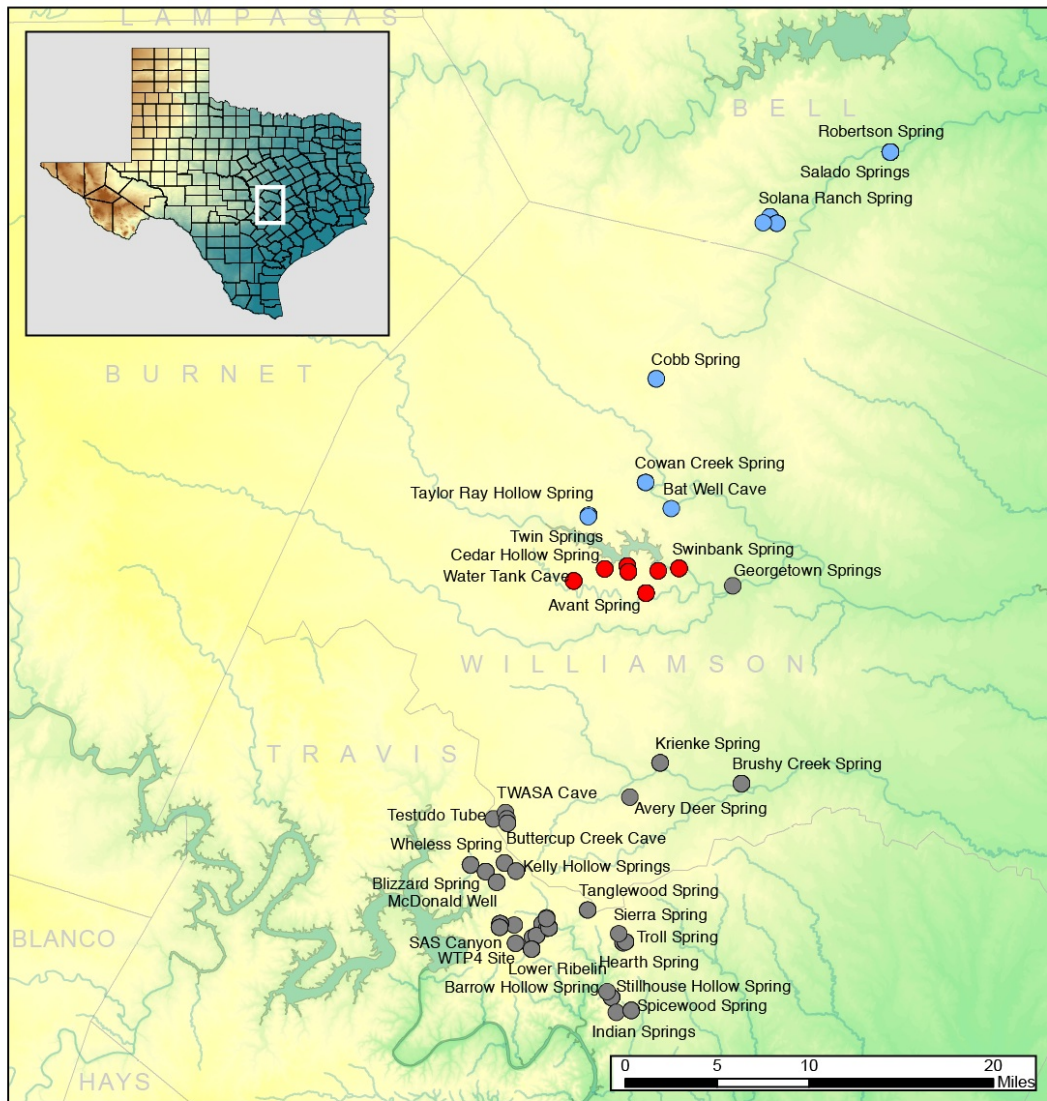
Figure 5. Distribution of *E. tonkawae*, *E. naufragia*, and *E. chisholmensis* localities (the species that occur north of the Colorado River). Georeferenced data for all sample localities are presented in Appendix Table 1.

Discussion

Our results support the taxonomic status of the described species, and of the known but undescribed species of Texas neotenic *Eurycea* (see Appendix Table 1). Each of the currently recognized species forms a monophyletic group in the analyses shown in Figures 1 and 2, and the levels of genetic divergence are consistent with the recognition of the current breaks among species. In particular, we found strong support from both the phylogenetic analyses (Figures 1 and 2) and the population STRUCTURE analysis (Figure 3) for the three described species from north of the Colorado River: *E. tonkawae*, *E. naufragia*, and *E. chisholmensis* (see distributions in

10

Figure 5).

The boundaries among *E. tonkawae*, *E. naufragia*, and *E. chisholmensis* have been clarified by our analyses. Chippindale et al. (2000) provisionally assigned the population of *Eurycea* at Georgetown Springs to *E. naufragia*, but they only had access to a single juvenile specimen and did not have enough tissue to include this population in their allozyme analyses. Our analyses (Figures 1-3) indicate that this population is actually *E. tonkawae*. Although Georgetown Springs in geographically proximate to the range of *E. naufragia*, the Balcones Fault extends northward and separates the populations of *E. naufragia* to the west from *E. tonkawae* populations to the east of the fault (and Interstate 35) at Brushy Creek Spring and Georgetown Springs. Thus, the range of E*. tonkawae* extends along and east of the Balcones Fault north to Georgetown. Georgetown Springs have been severely impacted by water intake for the City of Georgetown, and we have been unable to locate any additional specimens of *Eurycea* at this locality since collection of the single juvenile in 1991. The single specimen from Georgetown Springs does show some indication of shared alleles with *E. naufragia* populations (see Figure 3), but not more than several other populations of *E. tonkawae* that are geographically well separated from the range of *E. naufragia* (including the type locality of *E. tonkawae* at Stillhouse Hollow, and several cave populations in the Buttercup Creek region). Given these patterns, these shared alleles likely represent retained ancestral polymorphisms rather than evidence of ongoing gene flow. The Georgetown Springs population clusters strongly and significantly with the other *E. tonkawae* populations in the phylogenetic analyses (Figure 1-2).

Chippindale et al. (2000) also provisionally assigned a cave population of *Eurycea* from Bat Well Cave in Williamson County to *E. naufragia*; this population was located between other known populations of *E. naufragia* and known populations of *E. chisholmensis*. Since then, several spring populations have been discovered in this general area, and they usually have been treated as populations of *E. naufragia* as well. Our analyses indicate that these populations form the sister-group to known populations of *E. chisholmensis* (from the springs along Salado Creek), and should be referred to that species. Although these populations are genetically divergent and distinct from the populations of *E. chisholmensis* in the Salado Creek drainage, we recommend that they be included in *E. chisholmensis*, rather than described as a distinct species. This restricts the range of *E. naufragia* to springs south and east of Lake Georgetown in the San Gabriel Creek watershed; all known populations of *Eurycea* north of Lake Georgetown are here included within *E. chisholmensis* (Figures 1-5).

Other cave populations from north of the Colorado River are closely related to adjacent surface populations. Although some individuals from these cave populations appear to be morphologically distinct from surface-dwelling salamanders, all cave populations appear to be assignable to one of the described species. Most of the cave populations (i.e., all of the populations in the Buttercup Creek area) fall within known genetic variation of *E. tonkawae*. As noted above, the

Bat Well Cave population is assignable to *E. chisholmensis* (as here delimited), and the population from Water Tank Cave is assignable to *E. naufragia* (see Figure 5 and Appendix Table 1).

Acknowledgments

Literature Cited

Avise, J. C. 2004. Molecular markers, natural history, and evolution. Sinauer Associates, Sunderland, MA.

Catchen, J., P. A. Hohenlohe, S. Bassham, A. Amores, and W. A. Cresko. 2013. Stacks: an analysis tool set for population genomics. *Molecular Ecology* 22:3124–3140.

Chifman, J. and L. Kubatko. 2014. Quartet inferences from SNP data under the coalescent model. *Bioinformatics* 30:3317–3324.

Chippindale, P. T., A. H. Price, J. J. Wiens, and D. M. Hillis. 2000. Phylogenetic relationships of central Texas hemidactyliine plethodontid salamanders, genus *Eurycea*, and a taxonomic revision of the group. *Herpetological Monographs* 14:1-80.

Evanno, G., S. Regnaut, and J. Goudet. 2005. Detecting the number of clusters of individuals using the software structure: a simulation study. *Molecular Ecology* 8:2611-2620.

Falush D. 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* 164:1567–1587.

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.

Forstner, M. 2012. An evaluation of the existing scientific evidence for the currently proposed hyperdiversity of salamanders (*Eurycea sp.*) in central Texas. Unpublished Report, sent to U.S. Fish and Wildlife Service.

Hillis, D. M., D. A. Chamberlain, T. P. Wilcox, and P. T. Chippindale. 2001. A new species of subterranean blind salamander (Plethodontidae: Hemidactyliini: *Eurycea: Typhlomolge*) from Austin, Texas, and a systematic revision of central Texas paedomorphic salamanders. *Herpetologica* 57:266-280.

Kopelman, N. M., J. Mayzel, N. A. Rosenberg, and I. Mayrose. 2015. CLUMPAK: a program for identifying clustering modes and packaging population structure inferences across K. *Molecular Ecology* 15:1179-1191.

Leaché, A. D., Banbury, B. L., Felsenstein, J., Nieto-Montes de Oca, A., and A. Stamatakis. 2015. Short tree, long tree, right tree, wrong tree: new acquisition bias corrections for inferring SNP phylogenies. *Systematic Biology* DOI: 10.1093/sysbio/syv053.

Peterson B. K., J. N. Weber, E. H. Kay, H. S. Fisher, and H. E. Hoekstra. 2012. Double digest RADseq: An inexpensive method for *de novo* SNP discovery and genotyping in model and non-Model species. *PLoS ONE* 7: e37135.

Posada, D., and T. Buckley. 2004. Model selection and model averaging in phylogenetics: advantages of Akaike information criterion and Bayesian approaches over likelihood ratio tests. *Systematic Biology* 53:793-808.

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

Rambaut, A., M. A. Suchard, D. Xie, and A. J. Drummond. 2014. Tracer v1.6, available from http://beast.bio.ed.ac.uk/Tracer.

Ronquist F., M. Teslenko, P. van der Mark, D. L. Ayres, A. Darling, S. Hohna, B. Larget, L. Liu, M. A. Suchard, J. P. Huelsenbeck. 2012. MrBayes 3.2: Efficient bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology* 61:539–542.

Stamatakis, A. 2006. RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22:2688–2690.

Sweet, S. S. 1984. Secondary contact and hybridization in the Texas cave salamanders *Eurycea neotenes* and *E. tridentifera*. *Copeia* 1984:428-441.

Swofford, D. L. 2003. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sinauer Associates, Sunderland, Massachusetts.

USFWS, 2012. Endangered and threatened wildlife and plants; endangered status for four central Texas salamanders and designation of critical habitat. Federal Register, August 22, 2012; http://federalregister.gov/a/2012-19659

Wake, D. B. 1966. Comparative osteology and evolution of the lungless salamanders, family Plethodontidae. *Memoirs of the Southern California Academy of Sciences* 4:1-111.

Wiens, J. J., P. T. Chippindale, and D. M. Hillis. 2003. When are phylogenetic analyses misled by convergence? A case study in Texas cave salamanders. *Systematic Biology* 52:501-514.

Zwickl, D. J. 2006. Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion. Ph.D. dissertation, The University of Texas at Austin.

Appendix Table 1. Samples examined: Field numbers, localities, latitude and longitude, and identification. Field and specimen number abbreviations: Nathan F. Bendik (NFB), Paul Chippindale (PC), Thomas J. Devitt (TJD), Andy G. Guelsenkamp (AGG), David M. Hillis (DMH), John K. Karges (JKK), Lisa O'Donnell (LOD), Colin E. Peden (CEP), Andy Price (AHP), and Texas Natural History Collection (TNHC).

SampleID	Latitude	Longitude	Locality name	Taxon
DMH_91.91	30.660839	-97.726674	Buford Hollow Springs	*E. naufragia*
DMH_91.89	30.660833	-97.726667	Buford Hollow Springs	*E. naufragia*
DMH_88.151	30.371796	-97.763522	Stillhouse Hollow Spring	*E. tonkawae*
DMH_88.131	30.371796	-97.763522	Stillhouse Hollow Spring	*E. tonkawae*
DMH_90.10	30.371796	-97.763522	Stillhouse Hollow Spring	*E. tonkawae*
DMH_90.8	30.371796	-97.763522	Stillhouse Hollow Spring	*E. tonkawae*
DMH_88.130	30.371796	-97.763522	Stillhouse Hollow Spring	*E. tonkawae*
DMH_88.149	30.371796	-97.763522	Stillhouse Hollow Spring	*E. tonkawae*
DMH_88.148	30.371796	-97.763522	Stillhouse Hollow Spring	*E. tonkawae*
DMH_91.96	30.650494	-97.668045	Georgetown Springs	*E. tonkawae*
DMH_91.4	30.663894	-97.751119	Knight Spring	*E. naufragia*
DMH_91.6	30.663894	-97.751119	Knight Spring	*E. naufragia*
DMH_91.8	30.663894	-97.751119	Knight Spring	*E. naufragia*
DMH_91.3	30.663894	-97.751119	Knight Spring	*E. naufragia*

DMH_91.7	30.663889	-97.751111	Knight Spring	*E. naufragia*
DMH_90.24	30.516757	-97.661295	Brushy Creek Spring	*E. tonkawae*
DMH_90.13	30.516757	-97.661295	Brushy Creek Spring	*E. tonkawae*
DMH_90.20	30.516757	-97.661295	Brushy Creek Spring	*E. tonkawae*
DMH_90.17	30.516757	-97.661295	Brushy Creek Spring	*E. tonkawae*
DMH_91.74	30.492730	-97.856637	Testudo Tube	*E. tonkawae*
AHP_3331	30.645561	-97.736397	Avant Spring	*E. naufragia*
AHP_3325	30.645561	-97.736397	Avant Spring	*E. naufragia*
AHP_3334	30.645561	-97.736397	Avant Spring	*E. naufragia*
AHP_3322	30.645561	-97.736397	Avant Spring	*E. naufragia*
AHP_3324	30.645561	-97.736397	Avant Spring	*E. naufragia*
AHP_3255	30.660005	-97.750563	Cedar Breaks	*E. naufragia*
AHP_3254	30.660005	-97.750563	Cedar Breaks	*E. naufragia*
AHP_3246	30.660000	-97.750556	Cedar Breaks	*E. naufragia*
AHP_3252	30.660000	-97.750556	Cedar Breaks	*E. naufragia*
AHP_3253	30.660000	-97.750556	Cedar Breaks	*E. naufragia*
AHP_3255	30.660000	-97.750556	Cedar Breaks	*E. naufragia*
DMH_90.312	30.420847	-97.840243	Schlumberger Springs	*E. tonkawae*
AHP_3288	30.375833	-97.767222	Barrow Hollow Spring	*E. tonkawae*
AHP_3438	30.425985	-97.814481	Canyon Creek	*E. tonkawae*
AGG_1718	30.421454	-97.818068	Canyon Creek	*E. tonkawae*
AGG_1719	30.421454	-97.818068	Canyon Creek	*E. tonkawae*

AGG_1720	30.421454	-97.818068	Canyon Creek	*E. tonkawae*
DMH_90.417	30.943616	-97.544174	Salado Springs	*E. chisholmensis*
AGG_1919	30.943611	-97.544167	Salado Springs	*E. chisholmensis*
DMH_90.1	30.943611	-97.544167	Salado Springs	*E. chisholmensis*
DMH_90.2	30.943611	-97.544167	Salado Springs	*E. chisholmensis*
DMH_90.23	30.943611	-97.544167	Salado Springs	*E. chisholmensis*
DMH_90.3	30.943611	-97.544167	Salado Springs	*E. chisholmensis*
DMH_90.322	30.943611	-97.544167	Salado Springs	*E. chisholmensis*
PC_031	30.943611	-97.544167	Salado Springs	*E. chisholmensis*
PC_032	30.943611	-97.544167	Salado Springs	*E. chisholmensis*
AGG_1134	30.418630	-97.814135	Pit Spring	*E. tonkawae*
AGG_1135	30.418630	-97.814135	Pit Spring	*E. tonkawae*
AGG_1175	30.698386	-97.781728	House Springs	*E. tonkawae*
AGG_1176	30.698386	-97.781728	House Springs	*E. tonkawae*
AGG_1741	30.463037	-97.847873	House Springs	*E. tonkawae*
AGG_1742	30.463037	-97.847873	House Springs	*E. tonkawae*
AGG_1444	30.408555	-97.838956	WTP4 Site	*E. tonkawae*
AGG_1446	30.408555	-97.838956	WTP4 Site	*E. tonkawae*
AGG_1447	30.530746	-97.725400	Krienke Spring	*E. tonkawae*
AGG_1448	30.530746	-97.725400	Krienke Spring	*E. tonkawae*

AGG_1449	30.530746	-97.725400	Krienke Spring	*E. tonkawae*
AGG_1450	30.530746	-97.725400	Krienke Spring	*E. tonkawae*
AGG_1454	30.412557	-97.825715	Lanier Riffle	*E. tonkawae*
AGG_1477	30.412557	-97.825715	Lanier Riffle	*E. tonkawae*
AGG_1482	30.461708	-97.874325	Wheless Spring	*E. tonkawae*
AGG_1483	30.461708	-97.874325	Wheless Spring	*E. tonkawae*
AGG_1492	30.461708	-97.874325	Wheless Spring	*E. tonkawae*
AGG_1511	30.461708	-97.874325	Wheless Spring	*E. tonkawae*
AGG_1534	30.413647	-97.822471	Lanier Spring	*E. tonkawae*
AGG_1549	30.413647	-97.822471	Lanier Spring	*E. tonkawae*
AGG_1621	30.404649	-97.826447	Ribelin Tract	*E. tonkawae*
AGG_1630	30.404649	-97.826447	Ribelin Tract	*E. tonkawae*
AGG_1734	30.408727	-97.754535	Hearth Spring	*E. tonkawae*
AGG_1735	30.408727	-97.754535	Hearth Spring	*E. tonkawae*
AGG_1737	30.409597	-97.752569	Troll Spring	*E. tonkawae*
AGG_1738	30.409597	-97.752569	Troll Spring	*E. tonkawae*
AGG_1739	30.409597	-97.752569	Troll Spring	*E. tonkawae*
AGG_1754	30.430984	-97.782260	Tanglewood Spring	*E. tonkawae*
AGG_1756	30.430984	-97.782260	Tanglewood Spring	*E. tonkawae*
AGG_1766	30.457373	-97.838669	Kelly Hollow Springs	*E. tonkawae*
AGG_1767	30.457373	-97.838669	Kelly Hollow	*E. tonkawae*

			Springs	
AGG_1776	30.421819	-97.850954	Kretschmarr	*E. tonkawae*
AGG_1777	30.421819	-97.850954	Kretschmarr	*E. tonkawae*
AGG_1778	30.421819	-97.850954	Kretschmarr	*E. tonkawae*
AGG_1797	30.414890	-97.757822	Sierra Spring	*E. tonkawae*
AGG_1817	30.361560	-97.759886	Indian Springs	*E. tonkawae*
AGG_1823	30.698386	-97.781728	Taylor Ray Hollow Spring	*E. chisholmensis*
AGG_1824	30.698386	-97.781728	Taylor Ray Hollow Spring	*E. chisholmensis*
AGG_1825	30.698386	-97.781728	Taylor Ray Hollow Spring	*E. chisholmensis*
AGG_1826	30.698386	-97.781728	Taylor Ray Hollow Spring	*E. chisholmensis*
AGG_1828	30.662087	-97.768887	Cedar Hollow Spring	*E. naufragia*
AGG_1829	30.662087	-97.768887	Cedar Hollow Spring	*E. naufragia*
AGG_1830	30.662087	-97.768887	Cedar Hollow Spring	*E. naufragia*
AGG_1837	30.790375	-97.728266	Cobb Spring	*E. chisholmensis*
AGG_1838	30.790375	-97.728266	Cobb Spring	*E. chisholmensis*
AGG_1839	30.720469	-97.736801	Cowan Creek Spring	*E. chisholmensis*
AGG_1840	30.720469	-97.736801	Cowan Creek Spring	*E. chisholmensis*
AGG_1842	30.720469	-97.736801	Cowan Creek Spring	*E. chisholmensis*
AGG_1850	30.362957	-97.748072	Spicewood	*E. tonkawae*

			Spring	
AGG_1851	30.362957	-97.748072	Spicewood Spring	*E. tonkawae*
AGG_1852	30.362957	-97.748072	Spicewood Spring	*E. tonkawae*
AGG_1885	30.899673	-97.639036	Solana Ranch Spring	*E. chisholmensis*
AGG_1886	30.899673	-97.639036	Solana Ranch Spring	*E. chisholmensis*
AGG_1887	30.899673	-97.639036	Solana Ranch Spring	*E. chisholmensis*
AGG_1895	30.895577	-97.633658	Solana Ranch Spring	*E. chisholmensis*
AGG_1896	30.895577	-97.633658	Solana Ranch Spring	*E. chisholmensis*
AGG_1897	30.895577	-97.633658	Solana Ranch Spring	*E. chisholmensis*
AGG_1900	30.895873	-97.644216	Solana Ranch Spring	*E. chisholmensis*
AGG_1893	30.943616	-97.544174	Robertson Spring	*E. chisholmensis*
AGG_1920	30.943616	-97.544174	Robertson Spring	*E. chisholmensis*
AGG_1914	30.662200	-97.710500	Swinbank Spring	*E. naufragia*
AGG_1915	30.662200	-97.710500	Swinbank Spring	*E. naufragia*
AGG_1916	30.662200	-97.710500	Swinbank Spring	*E. naufragia*
AGG_1917	30.662200	-97.710500	Swinbank Spring	*E. naufragia*
AGG_1930	30.697268	-97.781929	Twin Springs	*E. chisholmensis*

AGG_1937	30.457106	-97.862539	Blizzard Spring	*E. tonkawae*
AGG_1938	30.457106	-97.862539	Blizzard Spring	*E. tonkawae*
AGG_1939	30.457106	-97.862539	Blizzard Spring	*E. tonkawae*
AGG_2020	30.653951	-97.793219	Water Tank Cave	*E. naufragia*
AGG_2022	30.653951	-97.793219	Water Tank Cave	*E. naufragia*
AGG_2024	30.653951	-97.793219	Water Tank Cave	*E. naufragia*
DMH_89.3	30.497086	-97.846977	Twasa Cave	*E. tonkawae*
DMH_89.4	30.492652	-97.845872	Ilex Cave	*E. tonkawae*
DMH_91.69	30.489836	-97.845272	Buttercup Creek Cave	*E. tonkawae*
DMH_91.71	30.489836	-97.845272	Buttercup Creek Cave	*E. tonkawae*
NFB_4.28.10	30.507504	-97.749315	Avery Deer Spring	*E. tonkawae*
NFBLOD1	30.419009	-97.812697	Franklin Tract	*E. tonkawae*
NFBLOD2	30.419009	-97.812697	Franklin Tract	*E. tonkawae*
NFBLOD3	30.419009	-97.812697	Franklin Tract	*E. tonkawae*
NFBLOD4	30.419009	-97.812697	Franklin Tract	*E. tonkawae*
NFBLOD8	30.404644	-97.826440	Lower Ribelin	*E. tonkawae*
NFBLOD19	30.404644	-97.826440	Lower Ribelin	*E. tonkawae*
NFBLOD18	30.425222	-97.814706	Trib6 Bull Creek	*E. tonkawae*
NFBLOD21	30.419450	-97.851433	SAS Canyon	*E. tonkawae*
NFBLOD22	30.419450	-97.851433	SAS Canyon	*E. tonkawae*

NFBLOD23	30.419450	-97.851433	SAS Canyon	*E. tonkawae*
NFBLOD31	30.449846	-97.853812	McDonald Well	*E. tonkawae*
NFBLOD32	30.449846	-97.853812	McDonald Well	*E. tonkawae*
NFBLOD33	30.449846	-97.853812	McDonald Well	*E. tonkawae*
PCDMH_92.64	30.702778	-97.716389	Bat Well Cave	*E. chisholmensis*
PCDMH_94.12	29.890350	-100.986140	Dolan Falls Pres	*E.* sp. 3
PCDMH_94.10	29.890350	-100.986140	Dolan Falls Pres	*E.* sp. 3
Dolan_Falls_JKK	29.890350	-100.986140	Dolan Falls Pres	*E.* sp. 3
TJD_988	29.890350	-100.986140	Dolan Falls Pres	*E.* sp. 3
TJD_989	29.890350	-100.986140	Dolan Falls Pres	*E.* sp. 3
DMH_90.124	29.802005	-98.252848	Bear Creek Spring	*E. pterophila*
DMH_90.119	29.802005	-98.252848	Bear Creek Spring	*E. pterophila*
DMH_90.118	29.802005	-98.252848	Bear Creek Spring	*E. pterophila*
DMH_90.120	29.802005	-98.252848	Bear Creek Spring	*E. pterophila*
DMH_90.210	30.111116	-98.301952	Boardhouse Springs	*E. pterophila*
DMH_90.211	30.111116	-98.301952	Boardhouse Springs	*E. pterophila*
DMH_90.205	30.111116	-98.301952	Boardhouse Springs	*E. pterophila*

DMH_90.209	30.111116	-98.301952	Boardhouse Springs	*E. pterophila*
DMH_88.23	30.111116	-98.301952	Boardhouse Springs	*E. pterophila*
DMH_88.28	30.111116	-98.301952	Boardhouse Springs	*E. pterophila*
DMH_90.206	30.111116	-98.301952	Boardhouse Springs	*E. pterophila*
DMH_88.29	30.111116	-98.301952	Boardhouse Springs	*E. pterophila*
DMH_90.208	30.111116	-98.301952	Boardhouse Springs	*E. pterophila*
DMH_88.26	30.111116	-98.301952	Boardhouse Springs	*E. pterophila*
PC_016	29.983338	-98.013619	Fern Bank Springs	*E. pterophila*
PC_015	29.983338	-98.013619	Fern Bank Springs	*E. pterophila*
PC_013	29.983338	-98.013619	Fern Bank Springs	*E. pterophila*
PC_012	29.983338	-98.013619	Fern Bank Springs	*E. pterophila*
PC_010	29.983338	-98.013619	Fern Bank Springs	*E. pterophila*
DMH_90.75	29.749450	-99.426675	Sutherland Hollow Spring	*E. troglodytes*
DMH_90.80	29.749450	-99.426675	Sutherland Hollow Spring	*E. troglodytes*
DMH_90.74	29.749450	-99.426675	Sutherland Hollow Spring	*E. troglodytes*
DMH_90.77	29.749450	-99.426675	Sutherland Hollow Spring	*E. troglodytes*
DMH_90.76	29.749450	-99.426675	Sutherland	*E. troglodytes*

			Hollow Spring	
DMH_90.79	29.749450	-99.426675	Sutherland Hollow Spring	*E. troglodytes*
DMH_90.109	29.749450	-99.426675	Sutherland Hollow Spring	*E. troglodytes*
DMH_90.110	29.749450	-99.426675	Sutherland Hollow Spring	*E. troglodytes*
DMH_90.111	29.749450	-99.426675	Sutherland Hollow Spring	*E. troglodytes*
DMH_90.105	29.749450	-99.426675	Sutherland Hollow Spring	*E. troglodytes*
DMH_90.96	29.749450	-99.426675	Sutherland Hollow Spring	*E. troglodytes*
DMH_90.139	29.637505	-98.694452	Helotes Creek Springs	*E. neotenes*
DMH_90.138	29.637505	-98.694452	Helotes Creek Springs	*E. neotenes*
DMH_90.147	29.637505	-98.694452	Helotes Creek Springs	*E. neotenes*
DMH_90.149	29.637505	-98.694452	Helotes Creek Springs	*E. neotenes*
DMH_90.150	29.637505	-98.694452	Helotes Creek Springs	*E. neotenes*
DMH_159.162	29.637505	-98.694452	Helotes Creek Springs	*E. neotenes*
DMH_90.144	29.637505	-98.694452	Helotes Creek Springs	*E. neotenes*
DMH_90.137	29.637505	-98.694452	Helotes Creek Springs	*E. neotenes*
DMH_90.146	29.637505	-98.694452	Helotes Creek Springs	*E. neotenes*
PCDMH_94.4	29.928205	-98.451048	Sattlers Deep Pit	*E. pterophila*

AHP_3467	29.989444	-99.953611	Greenwood Vly	*E. troglodytes*
AHP_3441	29.955556	-99.971389	Greenwood Vly	*E. troglodytes*
AHP_3465	29.989444	-99.953611	Greenwood Vly	*E. troglodytes*
AHP_3449	29.986389	-99.964167	Greenwood Vly	*E. troglodytes*
AHP_3453	29.986389	-99.964167	Greenwood Vly	*E. troglodytes*
TNHC_62772	29.807778	-99.558000	Wedgeworth Creek Springs	*E. troglodytes*
TNHC_62773	29.808000	-99.558000	Wedgeworth Creek Springs	*E. troglodytes*
DMH_88.192	29.893061	-97.930563	San Marcos Springs	*E. nana*
DMH_90.179	29.893061	-97.930563	San Marcos Springs	*E. nana*
DMH_88.200	29.893061	-97.930563	San Marcos Springs	*E. nana*
DMH_88.196	29.893061	-97.930563	San Marcos Springs	*E. nana*
DMH_88.203	29.893061	-97.930563	San Marcos Springs	*E. nana*
DMH_88.199	29.893061	-97.930563	San Marcos Springs	*E. nana*
DMH_90.183	29.893061	-97.930563	San Marcos Springs	*E. nana*
AHP_3077	29.893061	-97.930563	San Marcos Springs	*E. nana*
AHP_3076	29.893061	-97.930563	San Marcos Springs	*E. nana*
AHP_3074	29.893061	-97.930563	San Marcos	*E. nana*

			Springs	
AHP_3075	29.893061	-97.930563	San Marcos Springs	*E. nana*
DMH_88.15	30.106180	-98.481959	Zercher Springs	*E. pterophila*
DMH_93.4	29.634170	-98.852702	Pecan Springs Cave	*E. neotenes*
TNHC_62774	29.629000	-99.178000	West Verde Springs	*E. troglodytes*
TNHC_62775	29.628694	-99.177972	West Verde Springs	*E. troglodytes*
DMH_90.176	29.713616	-98.136952	Comal Springs	*E. pterophila*
DMH_90.174	29.713616	-98.136952	Comal Springs	*E. pterophila*
AHP_3072	29.713616	-98.136952	Comal Springs	*E. pterophila*
AHP_3071	29.713616	-98.136952	Comal Springs	*E. pterophila*
TNHC_62780	29.533000	-100.131000	Boiling Springs	*E.* sp. 3
DMH_89.70	29.823894	-99.566953	Sabinal Canyon	*E. troglodytes*
DMH_89.69	29.823894	-99.566953	Sabinal Canyon	*E. troglodytes*
PCDMH_92.101	29.762222	-98.666389	Pfeiffers Water Cave	*E. latitans*
DMH_91.41	30.341506	-98.137589	Martin Spring	*E.* sp. 1
DMH_91.43	30.341506	-98.137589	Martin Spring	*E.* sp. 1
DMH_91.52	30.341506	-98.137589	Martin Spring	*E.* sp. 1
DMH_91.39	30.341506	-98.137589	Martin Spring	*E.* sp. 1
DMH_91.51	30.341506	-98.137589	Martin Spring	*E.* sp. 1
DMH_91.42	30.341506	-98.137589	Martin Spring	*E.* sp. 1
DMH_90.56	29.480561	-100.078897	Carson Cave	*E.* sp. 3

DMH_91.55	30.339727	-98.137508	Hammetts Crossing Spring	*E. sp. 1*
DMH_91.54	30.339727	-98.137508	Hammetts Crossing Spring	*E. sp. 1*
DMH_91.53	30.339727	-98.137508	Hammetts Crossing Spring	*E. sp. 1*
DMH_90.6	30.339727	-98.137508	Hammetts Crossing Spring	*E. sp. 1*
DMH_91.12	29.901950	-97.921397	Rattlesnake Cave	*E. rathbuni*
DMH_91.13	29.901944	-97.921389	Rattlesnake Cave	*E. rathbuni*
TNHC_62781	29.758000	-100.397000	Kickapoo Springs	*E. sp. 3*
DMH_92.5	29.662783	-98.637230	Leon Springs	*E. neotenes*
CEP_030	29.410000	-100.450000	Lower Pinto Spring	*E. sp. 3*
CEP_039	29.410000	-100.450000	Lower Pinto Spring	*E. sp. 3*
AHP_3317	29.915283	-99.252230	Robinson Creek	*E. sp. 2*
AHP_3013	29.915283	-99.252230	Robinson Creek	*E. sp. 2*
AHP_3319	29.915283	-99.252230	Robinson Creek	*E. sp. 2*
AHP_3320	29.915283	-99.252230	Robinson Creek	*E. sp. 2*
AHP_3016	29.915283	-99.252230	Robinson Creek	*E. sp. 2*
AHP_3022	29.915283	-99.252230	Robinson	*E. sp. 2*

			Creek	
AHP_3020	29.915283	-99.252230	Robinson Creek	*E.* sp. 2
AHP_3023	29.915283	-99.252230	Robinson Creek	*E.* sp. 2
AHP_3318	29.915283	-99.252230	Robinson Creek	*E.* sp. 2
AHP_3017	29.915283	-99.252230	Robinson Creek	*E.* sp. 2
AHP_3019	29.915283	-99.252230	Robinson Creek	*E.* sp. 2
DMH_90.31	29.652505	-100.103342	Smith's Spring	*E.* sp. 3
DMH_90.33	29.652505	-100.103342	Smith's Spring	*E.* sp. 3
DMH_90.29	29.652505	-100.103342	Smith's Spring	*E.* sp. 3
DMH_90.32	29.652505	-100.103342	Smith's Spring	*E.* sp. 3
DMH_90.52	29.652783	-100.103342	Smith's Spring	*E.* sp. 3
DMH_90.30	29.652505	-100.103342	Smith's Spring	*E.* sp. 3
AHP_2989	30.143338	-99.077786	Trough Spring	*E.* sp. 2
AHP_2985	30.143338	-99.077786	Trough Spring	*E.* sp. 2
AHP_2982	30.143338	-99.077786	Trough Spring	*E.* sp. 2
AHP_2986	30.143338	-99.077786	Trough Spring	*E.* sp. 2
AHP_2979	30.143338	-99.077786	Trough Spring	*E.* sp. 2
AHP_3006	30.143338	-99.077786	Trough Spring	*E.* sp. 2
AHP_2988	30.143338	-99.077786	Trough Spring	*E.* sp. 2
AHP_2990	30.143338	-99.077786	Trough Spring	*E.* sp. 2
AHP_2991	30.143338	-99.077786	Trough Spring	*E.* sp. 2
AHP_2837	30.166672	-99.342230	Fessenden Springs	*E.* sp. 2

AHP_3050	30.166672	-99.342230	Fessenden Springs	*E.* sp. 2
AHP_2978	30.166672	-99.342230	Fessenden Springs	*E.* sp. 2
AHP_3052	30.166672	-99.342230	Fessenden Springs	*E.* sp. 2
AHP_3055	30.166672	-99.342230	Fessenden Springs	*E.* sp. 2
AHP_3065	30.166672	-99.342230	Fessenden Springs	*E.* sp. 2
AHP_3053	30.166672	-99.342230	Fessenden Springs	*E.* sp. 2
AHP_3054	30.166672	-99.342230	Fessenden Springs	*E.* sp. 2
AHP_2839	30.166672	-99.342230	Fessenden Springs	*E.* sp. 2
PCDMH_92.121	30.263688	-97.770827	Barton Springs	*E. sosorum*
DMH_88.18	30.263688	-97.770827	Barton Springs	*E. sosorum*
DMH_88.137	30.263688	-97.770827	Barton Springs	*E. sosorum*
DMH_88.135	30.263688	-97.770827	Barton Springs	*E. sosorum*
DMH_88.8	30.263688	-97.770827	Barton Springs	*E. sosorum*
DMH_88.138	30.263688	-97.770827	Barton Springs	*E. sosorum*
DMH_88.140	30.263688	-97.770827	Barton Springs	*E. sosorum*
DMH_88.142	30.263688	-97.770827	Barton Springs	*E. sosorum*
JKK_2004.1	29.416802	-100.905223	Four Mile	*E.* sp. 3

			Cave	
AHP_3043	30.005838	-99.361952	Edmunson Creek	*E*. sp. 2
AHP_3042	30.005838	-99.361952	Edmunson Creek	*E*. sp. 2
AHP_3039	30.005838	-99.361952	Edmunson Creek	*E*. sp. 2
AHP_3040	30.005838	-99.361952	Edmunson Creek	*E*. sp. 2
DMH_90.265	30.076672	-98.329452	T Cave	*E. pterophila*
TNHC_62776	29.427500	-99.784611	Bee Cave Springs	*E. troglodytes*
DMH_90.270	30.091672	-98.658341	Peaveys Spring	*E. pterophila*
DMH_90.272	30.091672	-98.658341	Peaveys Spring	*E. pterophila*
TNHC_62777	29.829722	-99.592806	Wishing Well Springs	*E. troglodytes*
TNHC_62778	29.829722	-99.592806	Wishing Well Springs	*E. troglodytes*
JKK_2003.1	30.032789	-98.125214	Jacobs Well	*E. pterophila*
DMH_91.57	29.847227	-98.491674	Honey Creek	*E. tridentifera*
AGG_2062	29.847227	-98.491674	Honey Creek	*E. tridentifera*
CEP_034	29.745000	-100.360000	Paisano Spring	*E*. sp. 3
AHP_3009	30.088338	-99.320564	Hwy176 Spring	*E*. sp. 2
AHP_3007	30.088338	-99.320564	Hwy176 Spring	*E*. sp. 2
AHP_3010	30.088338	-99.320564	Hwy176 Spring	*E*. sp. 2
CEP_038	29.958636	-99.958675	Morriss Spring	*E. troglodytes*

TNHC_62779	29.568500	-100.086306	Machinery Hollow Springs	*E.* sp. 3
DMH_90.115	29.755833	-98.620278	Badweather Pit	*E. tridentifera*
AGG_1818	29.893833	-100.994528	Blue Spring	*E.* sp. 3
AGG_1924	30.279593	-97.780434	Cold Spring	*E. sosorum*
Tailclip_76	30.279593	-97.780434	Cold Spring	*E. sosorum*
AGG_1981	29.868294	-97.967597	Johnsons Well	*E. rathbuni*
AGG_1989	29.901088	-100.998815	Finegan Spring	*E.* sp. 3
AGG_1993	30.263683	-97.770820	Old Mill Spring	*E. waterlooensis*
AGG_1994	30.263683	-97.770820	Old Mill Spring	*E. waterlooensis*
AGG_2061	30.340256	-98.258956	Maples Cave	*E.* sp. 1

Appendix Table 2. Choice of Analytical Parameters for STACKS 1.3 Pipeline. In the 'Values Tested' column, the default value is indicated in bold lettering.

Parameter	Pipeline Step	Parameter Function	Values Tested	Value Selected
-M	ustacks	Minimum stack depth (in number of reads) to record a locus	**3**, 5, 7	3
-m	ustacks	Maximum mismatches allowed between reads in a stack	**2,** 3, 4	2
-n	cstacks	Maximum mismatches allowed between a proposed locus and a locus in the catalog	**0**, 1, 2, 3	2,3
-m	populations	Minimum stack depth required at a locus for individual to be added to population	3, 6, 9	3
-r	populations	Percent of individuals in a population that must possess a locus to include locus	20, 40, 60, 80, 90	

Notes on Appendix Table 2:

1. We examined three values for the minimum stack depth parameter (-M): M=3, M=5 and M=7. A stack depth of M=3 produced a consistent amount of missing data across most individuals in the dataset. M=5 and M=7 produced missing data that

were biased; populations from the second sequencing run had about twice as much missing data, in proportion to the total number of sites, as those from the first. Therefore, we selected M=3 as the optimal value.

2. We examined three values for the maximum within-individual mismatch parameter (-m; ustacks): m=2, m=3 and m=4. m=2 is the default and maximizes the number of loci in the data matrix. m=2 and m=3 result in the same average amount of missing data per taxon (35%), whereas m=4 produces an average missing data of 47%. In the distance matrix, m=2 minimizes the number of taxon pairs for which there is no data for a comparison (n=177).

3. We examined at four values of the maximum catalog mismatch parameter (-n): n=0, n=1, n=2, and n=3. The default value of this parameter is n=0. The default value for our data severely limited the stacks depth values for the **-m** and **-r** parameters of the populations step: We did not retrieve any loci that had a stack depth of more than 2, or were present in more than 30% of populations. This suggests a value of n=0 is overly-conservative for our data. There was no value that always minimized missing data across samples. Both n=1 and n=3 maximized and minimized missing data for different individuals. n=2 tended to perform intermediately, never producing either the most or least missing data. However, n=3 produced the fewest taxon pairs for which there is no missing data. Therefore, both n=2 and n=3 were used in the fitting of the final two parameters. When fitting -m, using n=3 retained more loci, but the distribution of missing data and number of taxon pairs for which there were no data in common was the same for n=2 and n=3. When the final parameter, r, was included, n=2 consistently produced more taxon pairs for which no data were available. We therefore selected n=3 to produce our final data matrix.

4. We analyzed the data using three values for the (-m, populations) parameter: m=3, m=6 and m=9. Much like the -n parameter, this parameter made little difference to the overall amount of missing data. However, increasing this value increases the amount of taxon pairs between which no data are shared (n=3 gives 174 taxon pairs for which there is no basis to compare, n=9 gives 354). Increasing this parameter also halves the number of loci retained in the phylogenetic matrix. We, therefore, chose to maximize data examined by using a value of m=3.

5. The r parameter governs the percent of individuals that must be represented at a locus to use that locus in the final phylogenetic matrix. We examined five values of this parameter: r=20, r=40, r=60, r=80, r=90. As the value of r increases, the number of taxon pairs for which there are no shared data points increases. However, for taxa where data is present, the completeness of the data increases with the value of r. We

analyzed all five matrices, and found that phylogenetic results remained consistent across all runs, with greater resolution among populations from the largest matrix. Therefore, we present results from this largest matrix (34,518 loci for the phylogenetic analyses of all populations; 75,296 loci for the Structure analysis of individuals north of the Colorado River).