FINAL PERFORMANCE REPORT

As Required by

THE ENDANGERED SPECIES PROGRAM

TEXAS

Grant No. TX E-161-R

(F13AP00690)

Endangered and Threatened Species Conservation

Distinguishing the Neches River Rose Mallow (*Hibiscus dasycalyx*) from its congeners using genetic and niche modeling methods

Prepared by:

Dr. Josh Banta

Carter Smith
Executive Director

Clayton Wolf
Director, Wildlife

8 September 2017

INTERIM REPORT

STATE: Texas GRANT NUMBER: TX E-161-R-1

GRANT TITLE: Distinguishing the Neches River Rose Mallow (*Hibiscus dasycalyx*) from its congeners using genetic and niche modeling methods.

REPORTING PERIOD: 1 September 2013 to 31 Auguts 2017

OBJECTIVE(S). To resolve the taxonomic relationships among *Hibiscus dasycalyx* and its congeners (*H. laevis* and *H. moscheutos*), quantify the hybridization threat posed by *H. laevis* and *H. moscheutos* to *H. dasycalyx*, and create ground-truthed, geo-referenced maps of East Texas, showing the areas of suitable habitat for *H. dasycalyx* versus its congeners.

Segment Objectives:

Task #1. August 2013 – October 2013: Intensive (non-destructive) leaf sampling of H. dasycalyx and its congeners in the field.

Task #2. October 2013 – August 2015: Phylogenetic and population genetic analysis of H. dasycalyx and its congeners using modern molecular methods.

Task #3. October 2014 – July 2015: Creation of ecological niche models.

Task #4. July – August 2015: Refinement of the ecological niche models and analysis of niche separation among species.

Significant Deviations:

None.

Summary Of Progress:

Please see Attachment A.

Location: Angelina, Trinity, and Neches river watersheds in Cherokee, Harrison, Houston, Trinity, Angelina, Anderson, and Neches counties, Texas.

Cost: Costs were not available at time of this report, they will be available upon completion of the Final Report and conclusion of the project.

Prepared by: Craig Farquhar Date: 8 September 2017

Approved by:  Date: 8 September 2017

C. Craig Farquhar

ATTACHMENT A

Distinguishing the Neches River Rose Mallow (*Hibiscus dasycalyx*) from its congeners using genetic and niche modeling methods

Principal Investigator

Dr. Josh Banta, Assistant Professor

Department of Biology

University of Texas at Tyler

3900 University Blvd.

Tyler, TX 75799

(903) 565-5655

jbanta@uttyler.edu

Co-PIs

John Placyk (Assistant Professor)[1], Marsha Williams (Research Associate)[1], Lance Williams (Associate Professor)[1], and Randall Small (Associate Professor)[2]

[1]Department of Biology, University of Texas at Tyler

[2]Department of Ecology and Evolutionary Biology, University of Tennessee – Knoxville

JP: (903) 566-7147; jplacyk@uttyler.edu

MW: (903) 565-5883; mwilliams@uttyler.edu

LW: (903) 565-5878; lwilliams@uttyler.edu

RS: (865) 974-2371; rsmall@utk.edu

Texas Parks and Wildlife Consultant

Jason Singhurst, Botanist, Nongame and Rare Species Program

Final Report

August 31, 2017

Abstract

The Neches River Rose Mallow (*Hibiscus dasycalyx*) is an endemic Texas wildflower that has been listed as threatened under the Endangered Species Act. This study focused on the taxonomy and distribution of *H. dasycalyx* using integrative methods that combined genetics, genomics, population genetics, and ecology. The goal was to resolve the taxonomic status of *H. dasycalyx* relative to its co-occurring and widespread congeners, *H. laevis* and *H. moscheutos*, and then combine this information with ecological information to model their distributions on the landscape. We used two different DNA sequencing approaches: Sanger-based sequencing of the nuclear gene *GRANULE-BOUND STARCH SYNTHASE I* (*GBSSI*) together with genome-wide restriction site-associated DNA sequencing (RAD-seq) to provide an overwhelming number (thousands) of informative loci. Our findings suggest that *H. dasycalyx* is a separate taxon from *H. laevis* and *H. moscheutos*, and that hybridization with *H. laevis* is occurring. New populations of *H. dasycalyx* were documented in the Neches River floodplain, on the border of Trinity and Angelina Counties (in an area known as Boggy Slough). We found that *H. dasycalyx* is predicted to generally be closer to the banks of waterways than its congeners, and relegated to very flat, broad, frequently-flooded areas with highly erodible alluvial deposits. The ecological niche models developed by this project are available in the form of georeferenced raster maps. They can be used to choose the best sites for *H. dasycalyx* reintroduction and/or habitat restoration projects. Overall, our study supports *H. dasycalyx* as a genetically and ecologically distinct species, and suggests that hybridization with *H. laevis* is a pervasive threat. We recommend mitigating this threat by removing *H. laevis* from sites where *H. dasycalyx* occurs.

Introduction

When studying rare species, there are two main scientific concerns that must be addressed: (1) determining if the taxonomic standing of the target species is upheld (i.e., if it is a real, rare species) and (2) deciding where the target species occurs or is likely to occur. A robust genetic understanding of the target species and its congeners is required to accomplish this task. The other concern is deciding where the target species occurs or is likely to occur, which confirms the rarity of the species, identifies potentially undiscovered populations, and highlights promising locations for reintroductions. Ecological information can also help with species delimitation: despite subtle genetic and morphological differentiation between incipient species, ecological differentiation may be more pronounced (due to character displacement; Coyne and Orr 2004). Therefore, ecological differentiation may be an important leading indicator of when to treat two entities as separate taxa, before genetic approaches can pick up on these differences to the same extent (Raxworthy et al. 2007). Ecologically-based species delimitation is an especially important ingredient in taxonomy from a conservation perspective, because a conservation plan that ignores ecological differences among two distinct groups will wash out the important habitat differences among them, and therefore and do a poorer job of conserving either of the two groups than a conservation plan that treats those entities as separate and caters to their individual ecological needs

The Neches River Rose Mallow (*Hibiscus dasycalyx*) has recently been listed as a threatened species under the Endangered Species Act by the US Fish Wildlife Service (USFWS 2013). *H. dasycalyx* is a shrubby perennial marsh plant that is endemic to East Texas (Klips, 1995; Mendoza, 2004). *H. dasycalyx* is located in very few counties in East Texas (Cherokee, Houston, and Trinity), occurring within three watersheds (Angelina, Neches, and Trinity). It grows in seasonally wet alluvial soils that are flooded in late winter and early spring but that dry out in summer (TPWD 2011). Several of the historical habitat locations of the species have declining or extinct populations (Klips 1995; Mendoza 2004; TPWD 2011). Some documented threats to the species survival include: mowing/grazing, herbicide usage,

collections for horticulture, alterations of hydrology, and habitat encroachment by exotic and native species (Mendoza 2004; TPWD 2011).

Hybridization with co-occurring species is another potential threat to *H. dasycalyx*. Several studies have found either morphological or molecular evidence of hybridization between *H. dasycalyx* and its congeners, *H. moscheutos* and *H. laevis* (Blanchard 1976; Klips 1995; Mendoza 2004; Small 2004). All three species have the same ploidy (diploid) and are cross-fertile in the laboratory (Klips 1995). In addition, *H. dasycalyx* often co-occurs with *H. laevis* and *H. moscheutos* without any obvious barriers to interspecific reproduction (Correll and Correll 1972; Blanchard 1976; Klips 1995; TPWD 2011). Despite its potential importance as a conservation issue, the magnitude of the hybridization threat to *H. dasycalyx* is not understood. Molecular genetic data are necessary to address this issue, because morphological delineations of species and hybrids can be inconclusive and even misleading (Scotland et al. 2003; Hörandl and Emadzade 2012; Thompson et al. 2012).

The Rose Mallows, *Hibiscus* L. sect. *Muenchhusia* (Heister ex Fabricium) O. Blanchard (Malvaceae), are a North American taxon that consist of five closely related species (Blanchard 1976). The five species included in this taxon are *Hibiscus coccineus* Walter, *Hibiscus dasycalyx* Blake & Shiller, *Hibiscus grandiflorus* Michaux, *Hibiscus laevis* Allioni, and *Hibiscus moscheutos* L. There are two subspecies within *H. moscheutos* that Blanchard (1976) recognized as *H. moscheutos* subsp. *moscheutos* (synonymous with *H. moscheutos* subsp. *palustris* L.) of the northeastern United States and *H. moscheutos* subsp. *Lasiocarpos* (Cavanilles) O. J. Blanchard (synonymous with *H. moscheutos* subsp. *Lasiocarpos*) of the southeastern coastal plains (Blanchard 1976). *Hibiscus* sect. *Muenchhusia* was separated from the large *Hibiscus* sect. *Trionum* by Blanchard (Fryxell 1988). The separation of sect. *Muenchhusia* as a monophyletic group was proposed from the taxon's overall shared chromosome number (n = 19; Wise and Menzel 1971), ecological similarities of being primarily wetland species, similar morphological characteristics of individuals, a shared growth habit, and common geographic distribution of eastern and central North America for natural growing populations (Blanchard 1976).

The focus of this study is on three of these mallows: *H. dasycalyx*, the threatened endemic Texas species, and the two sympatric, congeneric species *H. laevis* and *H. moscheutos* (Correll and Correll 1972; Blanchard 1976). There are indicative morphological characteristics among the three taxa as described by Blanchard's (1976) classification:

Hibiscus laevis possess vegetative parts that are completely glabrous, and leaves that are triangular-hastately three-lobed. The middle leaf lobe is two to six times as long as the width of the leaf and long acuminate. Calyces and capsules are also glabrous or nearly glabrous, and petals moderately spread beyond the calyx tube and are of pink or white color with a red base. Seeds tend to have a reddish-pubescent appearance.

Hibiscus moscheutos subsp. *lasiocarpos* possess vegetative parts that are pubescent, and leaves that are unlobed, lanceolate or elliptic-lanceolate to broadly triangular-ovate with a lower surface that is densely stellate-pubescent (with occasional simple hairs). Calyces are stellate-tomentose and capsules are pubescent, occurring in variations of simple, stellate, or glandular hairs, with variations appearing singly or in combination. Petals moderately spread beyond the calyx tube and are of white (most commonly) or pink color with a red base. Seeds are glabrous in appearance.

Hibiscus dasycalyx possess vegetative parts that are glabrous, and leaves that are deeply and narrowly three-lobed. Calyces and capsules are densely hirsute, and petals moderately spread beyond the calyx tube and are of white color with a red base. Seeds tend to have a reddish-pubescent appearance. Overall, *Hibiscus dasycalyx* is very similar to *H. laevis*, except for its highly pubescent calyx and fruit and extremely more narrowly and deeply lobed leaves.

Although Blanchard's (1976) descriptive taxonomic work set forth a foundation on furthering systematic work among the *Hibiscus* sect. *Muenchhusia* by segregating it from sect. *Trionum*, it did not provide evidence of the phylogenetic relationship of the species in sect. *Muenchhusia*. Blanchard (1976)

was only able to note that *H. dasycalyx* had strong similarities with *H. laevis* and that the plants that were observed at the type location (Apple Springs in Trinity county; Blake 1958) resembled the type specimen, and that wild-type specimen seeds were grown and consistent with the description of *H. dasycalyx* and produced viable seed. Wise and Menzel (1971) also added that within sect. *Muenchhusia* there were two distinct groups that consisted of Group I, *H. grandifloras* and *H. moscheutos*, and Group II, *H. coccineus* and *H. laevis*, and that crosses within groups produced fertile hybrids whereas between group crosses produced hybrids that were in general unable to produce fruiting bodies. This data did not include an *H. dasycalyx* specimen, but it did give support of two distinctive and naturally occurring groups yet did not provide much evidence of evolutionary trajectory (Wise and Menzel, 1971).

Further evolutionary work on the rare endemic Texas rose mallow and its two co-occurring species was carried out by Klips (1995) and Small (2004). Klips (1995) sought to examine what the evolutionary relationship was between *H. dasycalyx* and its two sympatric species, *H. moscheutos* and *H. laevis.* This study provided that *H. dasycalyx* was able to produce fertile offspring when crossed with both sympatric species. It was also found through enzyme electrophoresis examining protein polymorphism in allozymes that the three species were all diploid and shared major alleles in all enzyme systems except three which possessed banding that was nearly identical in *H. dasycalyx* and *H. laevis* and absent in *H. moscheutos* (Klips, 1995). Klips (1995) was unable to warrant the endemic Texas rose mallow as a true species and suggested that it may be referred to as an ecotype or variety of the widespread *H. laevis*. Small (2004) sought to resolve if *Hibiscus* sect. *Muenchhusia* was monophyletic and where its phylogenetic placement was within the genus *Hibiscus* and tribe *Hibisceae,* and also to determine what the phylogenetic relationship was between the species of *Hibiscus* sect. *Muenchhusia*. This study provided that sect. *Muenchhusia* was a monophyletic group within a clade with other *Hibiscus* species, members of the tribe Malvavisceae, and other genera of Hibisceae, based on two chloroplast genes, *ndhF* and *rpL16* (Small 2004). The phylogenetic relationship between the species of *Hibiscus* sect. *Muenchhusia* was found to support previous studies, with the species falling in two main clades: one including *H. grandifloras* and *H. moscheutos* and the other including *H. coccineus*, *H. dasycalyx*, and *H. laevis*, with this work based on the nuclear *GBSSI* gene. There were also sequence polymorphisms found in one *H. dasycalyx* and *H. grandiflorus* sample that were inferred to be due to gene flow with *H. moscheutos*. Small (2004) was able to determine the monophly of sect. Muenchhusia, the relationship between the five species within this taxon, and provide some evidence of hybridization between *H. dasycalyx* and *H. grandifloras* with *H. moscheutos*.

Past studies have sought to understand the relationship of the endemic East Texas *H. dasycalyx* to its co-occurring, more widely distributed sister species *H. laevis* and *H. moscheutos*, and these results have not been able to conclusively identify *H. dasycalyx* as a separate, true species (Blanchard 1976; Klips 1995; Small 2004). Therefore, this is a crucial problem to understand because of the recent listing of *H. dasycalyx* as a threatened species. It important to understand whether this status is indeed warranted.

The purpose of this study is to gather a better understanding of the taxonomic standing of the threatened *H. dasycalyx* by: (1) producing genetic and genomic analyses of the three *Hibiscus* species that will aid in understanding their relationships to one another; (2) determining the specific habitat requirements and potential distribution of the endemic *H. dasycalyx* and its congeners using ecological niche modeling; and (3) comparing ecological niche models of *H. dasycalyx* and its two co-occurring congeners to understand if there are any ecological distinctions among the three *Hibiscus* species.

Objective

(1) Resolve the taxonomic relationships among *Hibiscus dasycalyx* and its congeners (*H. laevis* and *H. moscheutos*); (2) quantify the hybridization threat posed by *H. laevis* and *H. moscheutos* to *H.*

dasycalyx; and (3) create ground-truthed, geo-referenced maps of East Texas, showing the areas of suitable habitat for *H. dasycalyx* versus its congeners.

Location

Brazos, Trinity, Angelina, Neches, Sabine, Sulphur, and Caddo Lake watersheds in East Texas, USA. See Figures 1, 8, 9, and 10.

Methods

Task #1. *Intensive (non-destructive) leaf sampling of* H. dasycalyx *and its congeners in the field*

We collected tissue and voucher specimens from four previously identified populations of the *H. dasycalyx*. We have also collected specimens of two congeneric species, *H. moscheutos* and *H. leavis*. These species are much more widespread and we have collections covering a wider area of East Texas, in addition to specimens from the sites in sympatry with *H. dasycalyx*.

Task #2. *Phylogenetic and population genetic analysis of* H. dasycalyx *and its congeners using modern genetic methods*

DNA was extracted from young leaves of each plant using a DNeasy Plant Mini Kit (Qiagen). The DNA was sequenced in two different ways: using Sanger methodology (Sanger and Coulson 1975) to sequence the plants at the nuclear-encoded gene *GRANULE-BOUND STARTCH SYNTHASE I* (*GBSSI*); and using next-generation genome-wide sequencing methodology to sequence the plants genome-wide.

Sanger sequencing. PCR and sequencing primers are given in Table 1. PCR reactions were performed in 50 ml volumes with the following reaction components 35.75 µL RNase-free H_2O, 5 µL 10x *ExTaq* buffer (TaKaRa), 4 µL dNTPs, 2 µL $MgCl_2$, 2 µL BSA, 0.5 µL each 2- µmol primer, 0.25 µL *ExTaq* (TaKaRa), and 1 µL DNA (Small, 2004). The addition of bovine serum albumin (BSA) was used to help improve the amplification of difficult templates. PCR cycling conditions used for the amplification of the *GBSSI* nDNA were: 30 cycles each of denaturation at 94°C for 30 sec., primer annealing at 60°C for 30 sec, primer extension at 72°C for 2 min. A final extension step consisted of 5 min at 72°C (Small 2004). All PCR reactions were performed in Eppendorf Mastercycler personal thermal cyclers.

Verification of PCR product amplification was performed via gel electrophoresis. PCR products were purified prior to sequencing with illustra MicroSpin G-50 Columns (GE Healthcare). Purified PCR products were sent to Eurofins MWG Operon to be sequenced on an ABI 3730xl DNA sequencer. Sequencher 5.2.4 (Gene Codes Corporation, Ann Arbor, MI) was used to manually proofread and edit sequenced DNA. ClustalX (Thompson et al. 1997) was used to align all sequences before a final round of editing. Exon regions of all sequences were removed and intron regions were spliced together using Mesquite 3.01 (Maddison and Maddison 2014).

Genome-wide sequencing. DNA samples were sent off to Floragenex (Portland, OR) for restriction site-associated DNA sequencing (RAD-seq; Miller et al. 2007) and identification of single nucleotide polymorphisms (SNPs). The Florgenex protocols were as follows: The genome was first digested with a restriction endonuclease *PstI*, and then a series of sequencing adapters were ligated to the resulting DNA fragments. The DNA fragments were subjected to 1x100bp Seq on Illumina Hi Seq 2000 15-30x (Bentley et al. 2008).

Following Floragenex's standard bioinformatics pipeline, one sample (M7) was assembled *de novo* and used as the pseudoreference to call SNPs for the rest of the samples. Filters were applied at three levels of stringency: relaxed, standard, and stringent. Subsequent analyses used the SNPs called by

the standard criteria, specifically a cluster depth of 10 – 1000 and 2 – 4 variants per cluster. Sequence data has been archived under NCBI BioProject PRJNA382435.

Phylogenetic analysis. For the *GBSSI* sequence data, a maximum likelihood (ML) tree was generated using PhyML 3.1 (Guindon et al. 2010). The dataset was based on a nuclear gene, which contains heterozygosity that can complicate phylogenetic inference (Lischer et al. 2014). We followed Small (Small 2004) and leveraged the heterozygosity to our advantage by pseudo-phasing the dataset into haploid alleles that were analyzed as separate individuals. Pseudo-phasing was performed using the Excoffier-Laval-Balding (ELB) algorithm (Excoffier 2003) in Arlequin v. 3.5.2.2 (Excoffier and Lischer 2010).

To statistically support the ML phylogeny, a non-parametric bootstrap resampling using 1000 bootstrap replicates was performed (Felsenstein 1981). jModeltest 2.16 v20140903 was used to determine the substitution model to use by using the Bayesian Information Criterion (BIC) (Darriba et al. 2011). HKY was determined to be the best model of sequence evolution for the data according to the jModeltest. The default HKY85 model was used as the substitution model in PhyML. The ML tree was rooted using a sequence from *H. trionum* (Small 2004; GenBank accession No. AY341422) as the outgroup species.

For the genome-wide SNP data, phylogenetic trees were generated using two separate approaches: a ML tree and a Bayesian coalescence-based tree. The ML tree was generated using RAxML (Stamatakis 2014), a program for phylogenetic analysis of large datasets that implements a tree search algorithm returning trees with reliable likelihood scores. JModeltest 2.16 v20140903 identified a General Time Reversible (GTR) model as the best model of sequence evolution for the concatenated SNP alignment under the Akaike information criterion (AIC). A phylogeny was constructed in RAxML 3.1 using the rapid bootstrapping with subsequent ML search option under a GTR model of evolution with an ascertainment bias correction (ASC), given that only variant SNP sites were included in the alignment (as discussed in the RAxML manual). RAxML assessed support for the phylogeny using non-parametric bootstrap resampling of 100 replicates (Felsenstein 1981). *Hibiscus trionum* again served as the outgroup. It was initially obtained from a commercial source and provided by Dr. Edwige Moyroud at the University of Cambridge.

The Bayesian coalescence-based tree was generated using the program Bayesian Evolutionary Analysis by Sampling Trees (BEAST), with the add-on package SNP and AFLP Package for Phylogenetic analysis (SNAPP) (Bryant et al. 2012). This package is designed for inferring species trees and species demographics from independent (unlinked) biallelic markers such as well spaced SNPs (Bryant et al. 2012). This program implements a full coalescent model, but uses a novel algorithm to integrate over all possible gene trees, rather than sampling them explicitly. Following Yoder et al. (2013), we analyzed our SNP data using a multispecies coalescent approach in SNAPP version 1.3.0 within BEAST2 v2.3.2. The analysis utilized the sam GTR model of evolution and proceeded for 10,000,000 generations with 1,000,000 (10%) discarded as burnin. The full SNP data were converted to a 0, 1, 2 format for analysis, with 1 representing a heterozygous genotype. Once the program completed, the results were analyzed in Tracer (Drummond and Rambaut 2007) for performance and accuracy. As a primary analysis, we used all individuals of our focal in-group species and the single individual of *H. trionum* as an outgroup to facilitate rooting as with the maximum likelihood phylogeny.

F_{ST}. For the *GBSSI* sequence data, gene flow within each of the three species was assessed using the statistic F_{ST}. *F*-statistics (Wright 1965) are used to quantify genetic differentiation between different populations, and F_{ST} ranges from zero to one, with low values indicating a high amount of gene flow among the populations (panmixis), and high values indicating a low amount of gene flow among the populations (indicating they are genetically isolated from one another). If populations of a species are isolated, they are more inbred and therefore at greater risk of local extinction. F_{ST} was calculated by

performing a molecular analysis of variance (AMOVA) using Arlequin v. 3.5.2.2 (Excoffier and Lischer 2010).

For the genome-wide SNP data, we again calculated F_{ST}, but this time it was used to measure differentiation *among* the species rather than gene flow *within* them. This is because we did not have SNP data from multiple individuals within each population. The F_{ST} statistic in this case measures how genetically similar the species are to one another. Values range from zero to one, where zero indicates that the species are genetically indistinguishable from one another and one indicates that the species are completely diverged from one another at polymorphic loci. The program Arlequin was once again used for this analysis (Excoffier and Lischer 2010), and F_{ST} was calculated for (i) the entire data set (all three species) and (ii) *H. dasycalyx* and *H laevis* only.

Bayesian clustering analysis. The potential number of genetic clusters and the membership of each individual were estimated using STRUCTURE Version 2.3.4 (Pritchard et al. 2000). The software uses Markov chain Monte Carlo (MCMC) simulations to estimate those parameters, with the number of clusters to be tested (K) specified by the user (Blanco-Bercial and Bucklin 2016). The MCMC simulation was run for 300,000 iterations, after a burn-in period of 100,000 iterations. The traces were examined graphically to confirm chain convergence. The most likely *K* present in the data was inferred following Evanno et al. (2005). For each value of K (number of potential ancestral populations, which ranged from 1 to the number of presumed populations + 1), the genetic ancestry of each individual was estimated based on the admixture model without any prior population assignment. For the entire population set, K ranged from 1 to 10. The optimal K between the species in the 10 subsets was visualized and then chosen using the lowest log-likelihood (Rohlf and Sokal 1995). The same procedure described above was used for the *GBSSI* sequence data and the genome-wide data, except that, for the genome-wide data, we ran the Bayesian clustering analysis with and without *H. moscheutos*.

Task #3. *Creation of ecological niche models*

The Maxent method for niche modeling uses a general purpose machine learning method that estimates the probability of a species distribution by finding the probability of a distribution that is closest to uniform and then altering one environmental variable at a time repeatedly to maximize the likelihood of the occurrence dataset (Hernandez et al. 2006; Phillips et al. 2006). Maxent produces a heat map that visualizes a fitted cloglog link function relating the environmental data to the habitat suitability of every parcel of the landscape (at the grain size of the environmental data) (Phillips 2017). The habitat suitability scores range on a scale from 0 (most unsuitable) to 1 (most suitable).

Five continuous soil variables were incorporated into the models: calcium carbonate concentration (%), erodibility (K_f; ranges from 0.02 to 0.69 where higher values mean more susceptibility to rill erosion by rainfall; Oregon Department of Transportation 2005), liquid limit of the soil layer (% moisture by weight), slope of the map unit (%), and depth to the seasonally high water table (cm). Soil characteristics were obtained from the State Soil Geographic (STATSG0) Data Base (USDA, 1994), and the data processing steps used to make this dataset are described by Wolock (1997). All environmental layers and occurrence data were projected to NAD 1983 UTM Zone 15N (units: meters) using ArcMap 10.3. Environmental layers were converted to raster files and resampled to a common resolution of 100m x 100m. Then each raster was clipped to the extent of the study area and converted to ASCII files. The correlations among the layers were less than |0.65| (not shown).

To minimize spatial autocorrelation, we used the 'thin' function of the package spThin (Aiello-Lammens et al. 2015) in R version 3.3.3 (R Development Core Team 2017) to remove all but one occurrence points within 1 km of one another. For modeling, we used a "leave-one-out" or "*n*-1" cross-validation method, as previously described by Pearson et al. (2007), which is appropriate for small sample sizes. We set the number of folds for each species to equal the number of samples, so that each

fold contained *n* -1 observations, where *n* is the total sample size. This means that each fold only had a single test data point, and that each observation was the test data point, in turn, for a separate fold. Model statistics were then averaged across the *n* folds for each species.

Models were validated using the test AUC, or the area under the operator receiving curve. AUC measures the probability that a randomly chosen presence site will be ranked above a randomly chosen pseudoabsence site (Phillips and Dudik 2008). The test AUCs represent the average percentage of the pseudoabsence data with lower habitat suitability scores than single "test" presence locations left out of the model building process for each model fold. Ecological niche models with AUC values above 0.75 are considered useful (Elith 2002).

Task #4. *Refinement of the ecological niche models and analysis of niche separation among them*

Based on the ecological niche models for each of the three species, we went into the field and verified the ecological niche models. Specifically, we tested whether the species is found in the most suitable habitats and absent from the least suitable habitats. For each species, we picked locations where the target habitat is supposed to be favorable and locations where the target habitat is supposed to be unfavorable, as well as locations in between. All locations were on the banks of perennial rivers and tributaries. We then verified whether the species is found in the highly favorable habitats, not found in the unfavorable habitats, and found at some intermediate frequency in the intermediate habitats.

Additional locations were searched specifically for *H. dasycalyx* at Boggy Slough (http://www.conservationfund.org/projects/boggy-slough), a large conservation easement located in the Neches River floodplain on the border of Trinity and Angelina Counties. These searches were performed by the principal investigator on this project (Banta), as well as personnel from US Fish & Wildlife Service and the T. L. L. Temple Foundation. *H. dasycalyx* is already documented to occur at Boggy Slough. The goal was document more populations within this area and test whether they were predicted by the model. New locations of *H. dasycalyx* at Boggy Slough were added to the ground-truthing analyses (below).

We thinned the survey locations for each species to one location per 1 km radius using the 'thin' function of the package spThin (Aiello-Lammens et al. 2015) in R version 3.3.3 (R Development Core Team 2017). Then, for each species, we used logistic regression (Cox 1958) to test whether there was a significant relationship between the habitat suitability score of a location (independent variable) and the presence or absence of the species at that location (dependent variable). Logistic regression was performed using the 'lrm' function of the rms package (Harrell 2017) in R version 3.3.3 (R Development Core Team 2017). We assessed significance of the association between habitat suitability and the presence/absence of a species with a likelihood ratio test, which is the recommended procedure to assess the contribution of individual "predictors" to a given logistic regression model (Hosmer and Lemeshow 2000). We also tested whether the locations where the species was present had higher habitat suitabilities than the locations where the species was absent using Student's *t*-tests (Sokal and Rohlf 1995); habitat suitability was the dependent variable and presence/absence was the independent variable. *t*-tests were performed using the 'ttest' function in Microsoft® Excel for Mac version 15.36, using one-tailed tests with the assumption of homoscedasticity.

With this new ground-truthed data, we then updated the ecological niche models, thus improving the habitat maps. The new points were added to the previous points and then we re-performed ecological niche modeling as described above. Finally, we used the habitat maps to determine what environmental variables were separating the distributions of the different species. Specifically, for each species we graphed the average habitat suitability of each environmental variable at different levels of the variable. Using this approach, we picked out when one species was differentiated from another species, by looking for differences in how suitable a particular level of an

environmental variable is for the different species. This allowed us to make preliminary conclusions about how differentiated the three species are ecologically, and what environmental variables are most important for distinguishing their ranges.

3. Results

Task #1. *Intensive (non-destructive) leaf sampling of* H. dasycalyx *and its congeners in the field*

We collected 104 *H. dasycalyx* samples, 11 morphological hybrid samples, 61 *H. laevis* samples, and 67 *H. moscheutos* samples. The list of samples with their unique identification numbers and GPS coordinates are in Table 2. Herbarium specimens are currently being stored at the University of Texas at Tyler. The specimens will be submitted to the Botanical Research Institute of Texas (BRIT) herbarium in Ft. Worth, TX after peer-reviewed publications have been prepared. Herbarium specimens were prepared for all *H. dasycalyx* individuals as well as all morphological hybrids. Some morphologically pure *H. laevis* and *H. moscheutos* herbarium specimens were prepared as well.

Different plant samples were used for the Sanger sequencing and the genome-wide sequencing, as indicated in Table 2. Briefly, a subset of the *L. laevis* and *L. moscheutos* samples from the Sanger sequencing were also used for genome-wide sequencing, along with a different set of *H. dasycalyx* plants from Lovelady, TX.

Task #2. *Phylogenetic and population genetic analysis of* H. dasycalyx *and its congeners using modern genetic methods*

Phylogenetic analysis. The plant samples used for genetic analysis are shown in Figure 1 and Table 2. 1,867 nucleotides of the 1,927 nucleotide *GBSSI* gene were sequenced and aligned from 10 *H. dasycalyx*, 14 *H. laevis*, and 14 *H. moscheutos* individuals, as well as the outgroup *H. trionum*. The intron-only alignment consisted of 1,089 nucleotides; 17 were variable and 6 were parsimoniously informative. The most important result is that we found *H. moscheutos* to be distinct from *H. dasycalyx* and *H. laevis*, with moderate bootstrap support (Figure 2). This corresponds to the Group I (*H. moscheutos*) and Group II (*H. dasycalyx* and *H. laevis*) clades that Wise and Menzel (1971) and Small (2004) found. Within Group II, there were also three clades with moderate bootstrap support: one with primarily *H. laevis*, one with *H. dasycalyx*, and another one with *H. laevis*. One of the *H. laevis* clades contained most of the *H. laevis* specimens (L4, L5, L6, L7, L10, L11, L12, L13, L40, and L41), and the main *H. dasycalyx* clade contained alleles from most of the *H. dasycalyx* specimens (D2, D24, D25, D37, D38, RD, and D49) and none of the *H. laevis* specimens. The main *H. laevis* clade and the main *H. dasycalyx* clade were sister groups to one another. The main *H. laevis* clade also contained alleles from *H. dasycalyx* specimens D1, D2, and D37. In fact, both of the D1 alleles were in the *H. laevis* clade. For both D2 and D37, however, one of their alleles was in the *H. dasycalyx* clade and the other one was in the main *H. laevis* clade. The fact that both of the D1 alleles were in the main *H. laevis* clade is consistent with D1 being a misidentified specimen of *H. laevis*. The fact that D2 and D37 both had alleles in both the H. dasycalyx and H. laevis clades is consistent with D2 and D37 being hybrids (whether first generation or advance-generation) between *H. dasycalyx* and *H. laevis*. Yet the herbarium specimens for both D1 and D2 appear to have all of the diagnostic characteristics associated with *H. dasycalyx*, including thin lobed leaves and a hairy calyx; the herbarium specimen for D37 was inconclusive. Finally, there was another *H. laevis* clade with moderate bootstrap support, containing specimens L15, L31, and L32, whose placement within the larger phylogeny was undetermined.

The RAD-Seq analysis yielded large amounts of genome-wide data for six *H. dasycalyx*, four *H. laevis*, and five *H. moscheutos* individuals, as well as the outgroup *H. trionum*. The number of quality

filtered RAD tags via the standard output of reads passing FASTQ quality filters were 14,354,883, and the number of failing reads was 480,151. The total number of contigs extracted from the provisional clusters were 44,054, and the total number of contigs in the final assembly were 71,194 with an average base pair length of 92. The total cluster length was 6,549,848 bp. Out of the 16 samples screened, the total number of candidate variants detected was 117,026, and the number of candidate variants filtered (due to missing or low-quality data) was 102,622. The number of candidate variants passing all filters was 14,062.

The rooted maximum likelihood tree based on the genome-wide data showed *H. dasycalyx* and *H. laevis* to be more closely related to each other than either are to *H. moscheutos* (Figure 3). This again corresponds to the Group I and Group II that were seen earlier. Furthermore, it showed *H. laevis* and *H. dasycalyx* to each be a separate monophyletic group, albeit closely related. The analysis separated the three species into two major clades: one that contained only *H. moscheutos*, and the other that contained both *H. dasycalyx* and *H. laevis*. Within the H. *dasycalyx-H.laevis* clade, the two species were reciprocally monophyletic sister taxa. Bootstrap support for all nodes was high, except for some internal nodes within the *H. dasycalyx* clade. The rooted Bayesian coalescent tree based on the genome-wide data also showed two major clades: one containing only *H. moscheutos* and one containing both *H. dasycalyx* and *H. laevis* (Figure 4). While this analysis still recovered the Group I and Group II clades as before, the difference was that *H. dasycalyx* was nested within *H. laevis*. The major clades, as well as the paraphyly of *H. dasycalyx*, had high posterior support.

F_{ST}. The list of plant samples used for analysis of molecular variance (AMOVA) of the gene *GBSSI* are listed in Table 3. We could only use plant samples from locations where multiple plants were collected, because the analysis required replication within populations. In this case, the statistic F_{ST} refers to differentiation among populations *within* species. The F_{ST} statistics for each species are listed in Table 4. The only F_{ST} statistic that was significantly different from zero was the one for *H. laevis* (F_{ST} = 0.69; *P*-value = 0.007), indicating that the *H. laevis* populations were significantly genetically differentiated from one another. The other two F_{ST} statistics were statistically indistinguishable from zero (*H. moscheutos*: F_{ST} = 0.14, *P*-value = 0.29; *H. dasycalyx*: F_{ST} = 0.28; *P*-value = 0.15, indicating that the populations within these species were not significantly genetically differentiated from one another.

For the genome-wide data, F_{ST} refers to genetic differentiation *among* (rather than within) species. For the genome-wide AMOVA including all three species, F_{ST} is 0.58 and the *P*-value is < 0.01, indicating that the three species are significantly genetically differentiated from one another. For the genome-wide AMOVA including only *H. dasycalyx* and *H. laevis*, F_{ST} is 0.22 and the *P*-value is < 0.01, indicating that *H. dasycalyx* and *H. laevis* are significantly genetically differentiated from one another (Table 4).

Bayesian clustering analysis. For the *GBSSI* sequence data, the most parsimonious number of inferred ancestral groups (*K*) was 5. *H. moscheutos* was inferred to have contributions from two ancestral groups, color-coded purple and green (Figure 5). *H. laevis* and *H. dasycalyx* were inferred to have contributions from three different ancestral groups, color-coded blue, yellow, and red. *H. dasycalyx* was primarily associated with the blue-colored inferred ancestral group, whereas *H. laevis* was primarily associated with the red and yellow inferred ancestral groups. There was no evidence of admixture between *H. moscheutos* and either *H. dasycalyx* or *H. laevis*, as there were no substantial inferred ancestral contributions that were in common among these groups. There was evidence of admixture or incomplete lineage sorting between *H. dasycalyx* and *H. laevis*, with several *H. dasycalyx* individuals having simultaneous contributions from the blue, yellow, and red inferred ancestral groups (D2, D37, D38, and D48; Figure 5) and one *H. laevis* individual having simultaneous contributions from the red and blue inferred ancestral groups (L14; Figure 5).

For the genome-wide Bayesian cluster analysis of all three species, the most parsimonious number of inferred ancestral groups was two. It showed that *H. moscheutos* clustered separately from

H. dasycalyx and *H. laevis*, but that *H. laevis* and *H. dasycalyx* did not cluster separately from one another. It also showed no evidence of admixture among *H. moscheutos* and the *H. dasycalyx*/*H. laevis* group (Figure 6). For the genome-wide Bayesian cluster analysis of just *H. dasycalyx* and *H. laevis*, the most parsimonious number of inferred ancestral groups was six. In this case, the analysis was able to detect more fine-scale differentiation between the two species. It revealed that *H. dasycalyx* clustered separately from *H. laevis*. Even though there were two different inferred ancestral groups within *H. dasycalyx*, these inferred ancestral groups were not shared by *H. laevis*. And while *H. laevis* showed evidence of genetic diversity, with multiple inferred ancestral contributions to its genome, these inferred ancestral contributions were not shared by *H. dasycalyx* (Figure 7). Thus, the two species were reciprocally differentiated from each other in this dataset with no evidence of admixture.

Task #3. *Creation of the initial ecological niche models*

The study area was restricted to East Texas including the watersheds of the Trinity, Neches, and Angelina rivers. This extent included all of East Texas, so as to incorporate locations for *H. laevis* and *H. moscheutos* that were included in the models (Figure 8). Species occurrence data was obtained for *H. dasycalyx*, *H. laevis*, and *H. moscheutos* via personal collections from the field, herbaria records and iNaturalist records (http://inaturalist.org) (Figure 9 and Table 4). The GPS coordinates of one *H. laevis* location are redacted at the landowner's request; it is approximately 14 km southwest of Groveton, TX.

The test AUC values for the three species are listed in Table 4. Briefly, the test AUC values for *H. laevis* and *H. dasycalyx* were well above the 0.75 threshold that deems the models useful, whereas the test AUC for *H. moscheutos* was at the cusp of the 0.75 threshold.

Task #4. *Refinement of the ecological niche models and analysis of niche separation among them*

Ground-truth locations for each species are given in Figure 10 and Tables 6 – 8. We documented three additional populations of *H. laevis* (Table 6), four additional populations of *H. moscheutos* (Table 7), and four additional populations of *H. dasycalyx* (Table 8). The ground-truthed areas where *H. laevis* and *H. moscheutos* were found had significantly higher habitat suitabilities than the areas where they were not found. The ground-truthed areas where *H. dasycalyx* was found had marginally significantly higher habitat suitabilities than the areas where it was not found (Figure 11). When the new confirmed locations of the three species were fed into the ecological niche models, the test AUC values of all three species increased. In fact, adding the updated locations increased the test AUC for *H. moscheutos* above the 0.75 threshold (Table 5).

The updated ecological niche models are presented in Figures 12 – 14. The most suitable habitats for *H. laevis* were along the upper and middle Neches River and its tributaries; the Angelina River and its tributaries; the tributaries (but not the main stem) of the middle Sabine River along the border with Louisiana; and a group of tributaries of the Trinity River (White Rock Creek, Tantabogue Creek, Little White Rock Creek, and Caney Creek) (Figure 12). The most suitable habitats for *H. moscheutos* were along the middle Neches River; the headwaters of the Angelina River and its tributaries (Mud Creek, Boules Creek, and Johnson Creek); the headwaters of the Neches River and its tributaries (Kickapoo Creek, Prairie Creek, and Indian Creek); the tributaries of the upper Sabine River (but not the main stem); the lower Sabine River on the border with Louisiana; and the tributaries of Caddo Lake (Black Cypress Bayou, Little Cypress Bayou, and Big Cypress Bayou) (Figure 13). The most suitabilt habitats for *H. dasycalyx* were along the middle Neches River and its tributaries (notably Piney Creek); a small set of tributaries of the Angelina River (Odell Creek and Linton Creek); and the Angelina River where it meets the Neches River (Figure 14).

Some differences were apparent in the habitat suitabilities of the different species. All three species had higher habitat suitabilities in areas with lower calcium carbonate concentrations, but *H. dasycalyx* was the most sensitive to increases calcium carbonate concentration, and *H. laevis* was the most tolerant; *H. moscheutos* was in the middle (Figure 15a). *H. dasycalyx* consistently preferred more erodible soil, whereas *H. laevis* and *H. moscheutos* preferred two different levels of soil erodibility (moderately erodible and highly erodible) (Figure 15b). *H. dasycalyx* habitat suitability decreased with increasing liquid limit of the soil layer, whereas *H. laevis* had a strong preference for soils with a liquid limit around 31% moisture, and *H. moscheutos* preferred soils with a liquid limit of either ~35% moisture or ~68% moisture (Figure 15c). All three species preferred soils with as little slope as possible, consistent with floodplain alluvium, but *H. moscheutos* was much more tolerant of non-flat areas (Figure 15d). *H. moscheutos* and *H. dasycalyx* preferred soils as close to the seasonally-high water table as possible, whereas *H. laevis* preferred soils that were 4 – 8 cm above the seasonally-high water table and was less tolerant of conditions outside of this range (15e).

Discussion

Overview

This is the most comprehensive study to date of the federally threatened Texas endemic plant *Hibiscus dasycalyx* and its widespread congeners, *H. laevis* and *H. moscheutos*. This study utilized both single-gene data and genome-wide data, along with ecological data, to create an integrative picture of the taxonomy and distribution of *H. dasycalyx* in the context of similar species.

The totality of evidence presented here is consistent with *H. dasycalyx* being a distinct taxon from *H. laevis* and *H. moscheutos*, and being more closely related to *H. laevis* than to *H. moscheutos*. The single-gene phylogenetic analysis suggests that there is a distinct group of *H. dasycalyx* plants, and the genome-wide maximum-likelihood phylogenetic analysis strongly supports this conclusion. The genome-wide F_{ST} statistics and the genome-wide Bayesian clustering analysis evidence provide further support. The genome-wide F_{ST} statistic measures how genetically similar the species are to one another, and this statistic suggests that *H. dasycalyx* has a distinct gene pool from both *H. laevis* and *H. moscheutos*. The genome-wide Bayesian clustering analysis also suggests that the gene pool of *H. dasycalyx* is distinct. Surprisingly, the genome-wide Bayesian coalescence phylogenetic analysis contradicts these conclusions, suggesting that *H. dasycalyx* is part of the genetic variation within *H. laevis*. Furthermore, the single-gene Bayesian clustering analysis is ambiguous as to whether *H. dasycalyx* is distinct. There is no simple way to reconcile these contradictions, except to conclude that the *overall* weight of evidence is consistent with *H. dasycalyx* being a separate taxon, despite some conflicting or ambiguous signals in the data. We believe it would be unjustified, based on these results, to change the taxonomic status of *H. dasycalyx* relative to *H. laevis* and *H. moscheutos*.

The single-gene results presented here (both from the phylogenetic analysis as well as the Bayesian clustering analysis) have uncovered evidence of hybridization between *H. dasycalyx* and *H. laevis*, but not between *H. moscheutos* and either *H. dasycalyx* or *H. laevis*. In fact, there is evidence of hybridization in 25% of our *H. dasycalyx* genetic samples, at two of the three sites from which we have genetic data (Boggy Slough and Mud Creek). This suggests that hybridization with *H. laevis* is a real and widespread threat to the *H. dasycalyx* gene pool. Hybridization can result in genetic swamping, where the rarer species is subsumed by the more common one by successive generations of inbreeding with, and backcrossing to, the more common species (Rhymer and Simberloff 1996). We recommend that hybridization mitigation efforts should be considered, such as removing *H. laevis* from *H. dasycalyx* populations. On the other hand, we have no evidence from our study that hybridization with *H. moscheutos* constitutes a serious threat to *H. dasycalyx*. An important caveat is that we did not

genetically or genomically analyze morphological hybrids. This means that hybridization with *H. laevis* could be an even greater threat to *H. dasycalyx* than our data reveals, and *H. moscheutos* could be hybridizing with *H. dasycalyx* even though its alleles are not evident in morphologically pure *H. dasycalyx* specimens.

The ecological data presented here suggests that *H. dasycalyx* has distinct habitat preferences from *H. laevis* and *H. moscheutos*. It appears to have a more restricted range (and range potential) than *H. laevis* or *H. moscheutos*, preferring a narrower range of environmental conditions than the other two species. This further supports the distinctiveness of *H. dasycalyx*.

Phylogenetic and population genetic analysis of H. dasycalyx *and its congeners using modern genetic methods*

Our study illustrates how single-gene sequence data and genome-wide data can be used in concert to better understand plants of conservation concern and how to protect them. Single-gene nuclear data can better accommodate hybridization in a phylogeny. Most phylogenetic software cannot process heterozygous loci per se, yet the exclusion of these loci can bias the results (Lischer et al. 2014). With single-gene nuclear data, however, it is possible to decompose sequences into separate alleles (by phasing the data, as described in the methods section) and phylogenetically analyze each allele separately (see Small 2004). The reason that we did not use haploid sequences from chloroplast or mitochondrial DNA, or nuclear loci from an internal transcribed spacer region (ITS), is that the species in our study are very closely related. Small (2004) was unable to resolve the relationships among these taxa with chloroplast or ITS sequences, but he found the nuclear gene *GBSSI* to have enough molecular genetic variation to be useful at this taxonomic level.

With the single-gene sequence data, we are able to identify a clade of *H. dasycalyx* that is separate from *H. laevis* and *H. moscheutos*. This suggests that *H. dasycalyx* is a distinct taxon. We also identify two potential hybrids between *H. dasycalyx* and *H. laevis*: D2 and D37. In both cases, one of their alleles is in the *H. dasycalyx* clade and the other one is in the main *H. laevis* clade. These putative hybrids are both from populations where *H. dasycalyx* and *H. laevis* co-occur (Banta, personal observation), making their hybrid ancestry plausible. Yet the herbarium specimen for D2 clearly matches *H. dasycalyx* morphologically (whereas the herbarium specimen for D37 is inconclusive due to missing plant parts). More surprisingly, the phylogenetic analysis suggests that D1 may be a misidentified pure specimen of *H. laevis*, rather than *H. dasycalyx*, even though the herbarium specimen matches *H. dasycalyx* morphologically. Both D1 and D2 are from the same population at Boggy Slough, so the fact that both of them have putative *H. laevis* ancestry (as opposed to just one of them) makes sense. This suggests that it may be easier to confuse *H. dasycalyx* and *H. laevis* in the field than is currently appreciated, especially in areas of sympatry.

The results from the single-gene Bayesian clustering analysis largely mirror the results from the single-gene phylogenetic analysis: based on their inferred ancestral contributions, D1 appears to be a misidentified specimen that is actually *H. laevis*, and D2 and D37 appear to be hybrids with *H. laevis*. The Bayesian clustering analysis is an entirely different computational approach, so the fact that the results largely mirror those of the phylogenetic analysis provides another piece of evidence in favor those conclusions. Furthermore, the Bayesian clustering approach explicitly incorporates heterozygous sites into the analysis, which is another benefit of including it in the study. Interestingly, the analysis infers that the same ancestral group contributes to *H. dasycalyx* as well as *H. laevis* specimen L14. Specimen L14 is from a population where *H. dasycalyx* is not known to occur, and the plants in this population match *H. laevis* morphologically. But this specimen is from Harrison County, TX, where a population of *H. dasycalyx* was known to occur in 1980 (albeit at a different location; USFWS 2013). Thus, it is possible that L14 represents hybridization with *H. dasycalyx*, followed by advance-generation introgression into

H. laevis. Interestingly, US Fish & Wildlife Service (2013) reports that the Harrison County specimen from 1980 was morphologically ambiguous, and was once considered to be *H. laevis* before being declared to be *H. dasycalyx*. Gene flow between *H. dasycalyx* and *H. laevis* may help to explain this ambiguity, in which case L14 could be evidence of this gene flow.

The single-gene F_{ST} statistics for the three species suggest that gene flow among populations is relatively open within *H. dasycalyx* and *H. moscheutos*. We found no significant evidence of inbreeding within either of those species. This is especially important for *H. dasycalyx*, which is of conservation concern. Our results do not show that the populations of *H. dasycalyx* are hampered by being so limited demographically as well as disjunct geographically. Interestingly, *H. laevis* did show evidence of restricted gene flow among its populations. This could be due to specimens L15, L31, and L32, which are from the same population. These specimens clustered separately from the rest of the *H. laevis* specimens in the single-gene phylogenetic analysis (although their exact position within the phylogeny was unclear), and they were inferred to have a different main ancestral contribution than the other *H. laevis* specimens. This raises the possibility that L15, L31, and L32 represent a different subspecies of *H. laevis*, (no subspecies are currently recognized). Incidentally, we emphasize that while *H. laevis* appears paraphyletic in the phylogeny, this should not be concluded; the actual placement of L15, L31, and L32 in the phylogeny is unclear, so their visualized placement is arbitrary. Similarly, it may appear that *H. dasycalyx* is paraphyletic as well, but the placement of D38-2, D48-2, and D49 are similarly unresolved and therefore also arbitrary.

An alternative explanation for many of the findings from the single-gene analyses (including the apparent hybridization) is that lineage sorting at *GBSSI* may be incomplete (Nichols 2001). In other words, some alleles could have been inherited from a common ancestor and still be present across the different species. This possibility cannot be ruled out, which is where the findings from the genome-wide data become important. Our genome-wide data integrates over thousands of loci at random places in the genome, so that the signal from incompletely sorted loci should be drowned out.

The genome-wide maximum likelihood phylogeny, the genome-wide F_{ST} statistic, and the genome-wide Bayesian clustering analysis all support the distinctness of *H. dasycalyx*. The genome-wide maximum likelihood phylogeny strongly supports *H. dasycalyx* as distinct from *H. laevis*. The genome-wide F_{ST} statistic, which in this case measures differentiation among species, shows that the *H. dasycalyx* gene pool is distinct from *H. laevis*. Finally, the genome-wide Bayesian clustering analyses suggests that *H. dasycalyx* has distinct ancestral contributions from those of *H. laevis* or *H. moscheutos*. The genome-wide results do not show evidence of hybridization between *H. dasycalyx* with the other species, but the genome-wide data uses a smaller dataset. The population of *H. dasycalyx* included in the genome-wide study does not show evidence of hybridization in the single-gene study (the *Hibiscus* preserve at Lovelady, TX). Therefore, the fact that hybridization is not detected in the genome-wide study is not surprising. This is not inconsistent with the single-gene results, since evidence of hybridization was detected in different *H. dasycalyx* populations than the one used in the genome-wide analyses.

The Bayesian coalescence phylogenic results are anomalous as compared to the other genome-wide results. They show *H. dasycalyx* nested within *H. laevis*, as opposed to being distinct. But Wielstra (Wielstra et al. 2014) found that the Bayesian coalescence approach had difficulty resolving the relationships among closely related taxa. *H. dasycalyx* and its two congeners are very closely related, which adds caution to the anomalous Bayesian coalescence results we found here. Given that our overall findings suggest *H. dasycalyx* is distinct from *H. laevis* and *H. moscheutos*, we believe our overall conclusion should reflect those findings.

Ecological Niche Modeling and ground-truthing

The ground-truthing confirmed that the ecological niche modeling generated predictive models: the ground-truthed locations where the species are found have significantly (or marginally significantly) higher predicted habitat suitabilities than the ground-truthed locations where the species are not found. Furthermore, the addition of the ground-truthed points to the ecological niche models improves them: the area under the operator receiving curves (AUCs) for each model increases with the addition of the ground-truthed data. This suggests that the ecological niche models, which are provided with this report in raster format, can be used to identify more populations of *H. dasycalyx*. But because *H. dasycalyx* is so rare, caution should be exercised in applying the model. The suitable habitat for *H. dasycalyx* overlaps with the suitable habitat for *H. laevis* and *H. moscheutos*, although the range of *H. dasycalyx* is more restricted. We recommend searching for new *H. dasycalyx* populations within a radius of documented populations, using the ecological niche model to find the most suitable habitats to search within that radius. The ecological niche model can make searching for this very rare species more efficient. All of the new locations of *H. dasycalyx* that we found are within 15km of an already-documented population.

The predictive models generated by ecological niche modeling support the conclusion that *H. dasycalyx* is distinct. *H. dasycalyx* is distinguished by being highly intolerant of soils with calcium carbonate, by preferring soils that have as low a liquid limit as possible, by being intolerant of sloped areas that lie outside of the floodplain, and by preferring soils that are as close to the seasonally-high water table as possible. The habitat affinities/tolerances of *H. dasycalyx* as a function of these environmental variables can be used to evaluate locations for reintroductions of *H. dasycalyx* as well as habitat restoration projects. Our modeling suggests that *H. dasycalyx* is tightly associated with very flat floodplains that are easily flooded. This is in contrast to *H. moscheutos*, which is much more tolerant of areas with steeper slopes, and in contrast to *H. laevis*, which prefers soils higher above the water table and therefore less easily flooded. *H. dasycalyx* has a much clearer preference than the other two species for highly erodible soils, consistent with frequently recharged floodplain alluvium. In summary, *H. dasycalyx* is predicted to generally be closer to the banks of waterways than the other two species, and relegated to very flat, broad, frequently-flooded areas with highly erodible alluvial deposits.

Acknowledgements

Funding for this work came from a Section 6 Traditional Grant (TX E-161-R) to JB, JP, RS, LW, and MW from Texas Parks & Wildlife Department (TPWD) and US Fish & Wildlife Service (USFWS), as well as grants from Ellen Temple and the T. L. L. Temple Foundation to JB and JP. The following individuals helped with field work or performed additional field work used in this project: Amber Miller and Jeffrey Ried from the USFWS; Jackie Poole from TPWD; En Tze Chong, Melody Sain, Megan Seawright, Kayla Key, and Julia Norrell from the UT-Tyler; Ellen Temple and personnel from the T. L. L. Temple Foundation; and Minnette Marr from the Ladybird Johnson Wildflower Center. Access to private land was provided by Ellen Temple, the T. L. L. Temple Foundation, and a private landowner who wishes to remain anonymous. Melody Sain, Katherine Barthel, and Julia Norrell at UT-Tyler performed the DNA extractions with help from Justin Hendy at the University of Tennessee-Knoxville. Melody Sain, Julia Norrell, Megan Seawright, En Tze Chong, Alyssa Blanton, Rebbecca Nolen, Samuel Davis, Kandice Hays, and Breanna Herring at UT-Tyler prepared and catalogued the herbarium specimens. Samuel Davis, at UT-Tyler assisted with ground-truthing analyses as part of undergraduate honor's thesis. Julia Norrell and Melody Sain prepared large parts of this project for their Masters theses at UT-Tyler. Jackie Poole, Anna Strong, and Jason Singhurst, state botanists for TPWD, provided substantial advice, guidance, and other forms of support for this work.

Literature Cited

Aiello-Lammens, M. E., R. A. Boria, A. Radosavljevic, B. Vilela, and R. P. Anderson. 2014. spThin: functions for spatial thinning of species occurrence records for use in ecological models. https://CRAN.R-project.org/package=spThin. Accessed 28 June 2017.

Bentley, D. R., S. Balasubramanian, H. P. Swerdlow, G. P. Smith, J. Milton, C. G. Brown… and A. J. Smith. 2008. Accurate whole human genome sequencing using reversible terminator chemistry. Nature 456:53-59.

Blake, S. E. 1958. Two species of *Hibiscus* from Texas. Journal of the Washington Academy of Science 48:277-280.

Blanchard, O. J. 1976. A revision of species segregated from *Hibiscus* sect. *Trionum* (Medicus) de Candolle sensu lato (Malvaceae). Ph.D. dissertation, Cornell University, Ithaca, NY.

Blanco-Bercial, L. and A. Bucklin. 2016. New view of population genetics of zooplankton: RAD-seq analysis reveals population structure of the North Atlantic planktonic copepod Centropages typicus. Molecular ecology.

Bryant, D., R. Bouckaert, J. Felsenstein, N. A. Rosenberg, and A. RoyChoudhury. 2012. Inferring species trees directly from biallelic genetic markers: bypassing gene trees in a full coalescent analysis. Molecular Biology and Evolution 29:1917-1932.

Correll, D. S., and H. B. Correll. 1972. Aquatic and wetland plants of southwestern United States. https://nepis.epa.gov/Exe/ZyPURL.cgi?Dockey=9100UIEA.TXT. Accessed 18 August 2017. Environmental Protection Agency, Washington, DC, USA.

Cox, D. R. 1958. The regression analysis of binary sequences. Journal of the Royal Statistical Society Series B: Statistical Methodology 20:215-242.

Coyne, J. A. and H. A. Orr. 2004. Speciation. Sinauer Associates, Inc., Sunderland, Massachusetts.

Darriba, D., G. L. Taboada, R. Doallo, and D. Posada. 2012. jModelTest 2: more models, new heuristics and parallel computing. Nature Methods 9:772-772.

Drummond, A. J. and A. Rambaut. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evolutionary Biology 7.

Elith, J. 2002. Quantitative methods for modeling species habitat: comparative performance and an application to Australian plants. Quantitative methods for conservation biology (eds. Ferson S, Burgman M). Springer, New York.

Excoffier, L. 2003. Analysis of Population Subdivision. Pp. 713-750 *in* D. Balding, M. Bishop, and C. C., eds. Handbook of Statistical Genetics, 2nd Edition. John Wiley and Sons, Ltd., New York.

Excoffier, L. and H. E. L. Lischer. 2010. Arlequin suite ver 3.5.2.2: A new series of programs to perform population genetics analyses under Linux and Windows. Molecular Ecology Resources 10:564-567.

Evanno, G., S. Regnaut, and J. Goudet. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular Ecology 14:2611-2620.

Felsenstein, J. 1981. Evolutionary trees from DNA sequences: a maximum likelihood approach. Journal of Molecular Evolution 17:368-376.

Fryxell, P. A. 1988. Malvaceae of Mexico. Systematic Botany Monographs 25:1-522.

Guindon, S., J. F. Dufayard, V. Lefort, M. Anisimova, W. Hordijk, and O. Gascuel. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Systematic Biology 59:307-321.

Harrell Jr., F. E. 2017. rms: regression modeling strategies. R package version 5.1-1. https://cran.r-project.org/package=rms. Accessed 9 Aug. 2017.

Hernandez, P. A., C. H. Graham, L. L. Master, and D. L. Albert. 2006. The effect of sample size and species characteristics on performance on different species distribution modeling methods. Ecography 29:773-785.

Hörandl, E. and K. Emadzade. 2012. Evolutionary classification: A case study on the diverse plant genus *Ranunculus* L. (Ranunculaceae). Perspectives in Plant Ecology, Evolution, and Systematics 14:310-324.

Hosmer, D. W. and S. Lemeshow. 2000. Applied logistic regression (2nd ed.). John Wiley and Sons, Hoboken, NJ.

Klipps. R. A. 1995. Genetic affinity of the rare eastern Texas endemic. *Hibiscus dasycalyx*. (Malvaceae). American Journal of Botany 82:1463-1472.

Lischer, H. E., L. Excoffier, and G. Heckel. 2014. Ignoring heterozygous sites biases phylogenomic estimates of divergence times: implications for the evolutionary history of microtus voles. Molecular Biology and Evolution 31:817-831.

Maddison, W.P., and D. R. Maddison. 2014. Mesquite: a modular system for evolutionary analysis. Version 3.01 http://mesquiteproject.org.

Mendoza, E. A. 2004. Genetic diversity within *Hibiscus dasycalyx*, *Hibiscus laevis* and *Hibiscus moscheutos u*tilizing ISSR techniques. M.S. dissertation, Stephen F. Austin University, Nacogdoches, TX.

Miller, M. R., J. P. Dunham, A. Amores, W. A. Cresko, and E. A. Johnson. 2007. Rapid and cost-effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. Genome Research 17:240-248.

Nichols, R. 2001. Gene trees and species trees are not the same. Trends Ecol Evol 16:358-364.

Oregon Department of Transportation. 2005. ODOT erosion control manual: guidelines for developing and implementing erosion and sediment controls. Appendix B. Prepared by Harza Engineering Company and ODOT Geo/Environmental Section. http://www.deq.state.or.us/wq/stormwater/docs/escmanual/appxb.pdf. Accessed 1 June 2017.

Pearson, R. G., C. J. Raxworthy, M. Nakamura, and A. T. Peterson. 2007. Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. Journal of Biogeography 34:102-117.

Phillips, S. 2017. A brief tutorial on Maxent. AT&T Research. https://biodiversityinformatics.amnh.org/open_source/maxent/Maxent_tutorial2017.pdf. Accessed 18 August 2017.

Phillips, S. J., and M. Dudik. 2008. Modeling of species distributions with MaxEnt: new extensions and a comprehensive evaluation. Ecography 31:161-175.

Phillips, S. J., M. Dudík, and R. E. Schapire. Maxent software for modeling species niches and distributions (Version 3.4.1). 2017. http://biodiversityinformatics.amnh.org/open_source/maxent/. Accessed 27 June 2017.

R Development Core Team. 2017. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Raxworthy, C. J., C. M. Ingram, N. Rabibisoa, and R. G. Pearson. 2007. Applications of ecological niche modeling for species delimitation: A review and empirical evaluation using day geckos (*Phelsuma*) from Madagascar. Systematic Biology 56:907-923.

Rhymer, J. M. and D. Simberloff. 1996. Extinction by hybridization and introgression. Annual Review of Ecology and Systematics 27:83-109.

Sanger, F. and A. R. Coulson. 1975. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. Journal of Molecular Biology 94:441-448.

Scotland, R. W., R. G. Olmstead, and J. R. Bennett. 2003. Phylogeny reconstruction: the role of morphology. Systematic Biology 52:539-548.

Small, R. L. 2004. Phylogeny of *Hibiscus* sect. Muenchhusia (Malvaceae) based on chloroplast *rpl16* and *ndhf*, and nuclear ITS and GBSSI sequences. Systematic Botany 29:385-392.

Sokal, R. R. and F. J. Rohlf. 1995. Biometry: the principles and practice of statistics in biological research. W.H. Freeman and Company, New York.

Stamatakis, A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30:1312-1313.

Texas Parks & Wildlife Department. 2011. Neches River rose-mallow fact sheet. http://www.tpwd.state.tx.us/huntwild/wild/wildlife_diversity/texas_rare_species/petition/neches-riverrosemallow/media/account.pdf. Accessed 6 June 2017.

Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin, and D. G. Higgins. 1997. The ClustalX windows interface: flexible strategies for multiple alignment aided by quality analysis tools. Nucleic Acids Research 24:4876-4882.

Thompson, K., C. S. Baker, A. van Helden, S. Patel, C. Millar, and R. Constantine. 2012. The world's rarest whale. Current Biology 22:R905-R906.

US Department of Agriculture. 1994. State soil geographic (STATSG0) Data Base: Data use information. Miscellaneous Publication Number 1492. http://www.fsl.orst.edu/pnwerc/wrb/metadata/soils/statsgo.pdf. Accessed 27 June 2017.

US Fish & Wildlife Service. 2013. 50 CFR Part 17: Endangered and threatened wildlife and plants; designation of critical habitat for Texas Golden Gladecress and Neches River Rose-Mallow; final rule. http://www.gpo.gov/fdsys/pkg/FR-2013-09-11/pdf/2013-22083.pdf. Accessed September 8th, 2014.

Wielstra, B., J. W. Arntzen, K. J. van der Gaag, M. Pabijan, and W. Babik. 2014. Data concatenation, bayesian concordance and coalescent-based analyses of the species tree for the rapid radiation of *Triturus* newts. PLoS ONE 9:e111011.

Wise, D. A. and M. Y. Menzel. 1971. Genetic affinities of the North American species of *Hibiscus* sect. Trionum. Brittonia 23:425-437.

Wolock, D. M. 1997. STATSGO soil characteristics for the conterminous United States. Open-file report 97-656. https://water.usgs.gov/GIS/metadata/usgswrd/XML/muid.xml. Accessed 30 May 2017. US Geological Survey, Reston, VA.

Wright, S. 1965. The interpretation of population structure by F-statistics with special regard to systems of mating. Evolution 19:395-420.

Yoder, J. B., R. Briskine, J. Mudge, A. Farmer, T. Paape, K. Steele, G. D. Weiblen, A. K. Bharti, P. Zhou, G. D. May, N. D. Young, and P. Tiffin. 2013. Phylogenetic signal variation in the genomes of *Medicago* (Fabaceae). Systematic Biology 62:424-438.

Appendices

laevis.tif: a georeferenced raster map of the habitat suitability for *Hibiscus laevis* across East Texas, from the corresponding ecological niche model. It can be opened in geographic information software such as ArcGIS, GRASS GIS, or QGIS.

moscheutos.tif: a georeferenced raster map of the habitat suitability for *Hibiscus moscheutos* across East Texas, from the corresponding ecological niche model. It can be opened in geographic information software such as ArcGIS, GRASS GIS, or QGIS.

dasycalyx.tif: a georeferenced raster map of the habitat suitability for *Hibiscus dasycalyx* across East Texas, from the corresponding ecological niche model. It can be opened in geographic information software such as ArcGIS, GRASS GIS, or QGIS.

Table 1. *GBSSI* amplification (Amp) and sequencing (Seq) primers used for Sanger sequencing.

Primer	Sequence (5' to 3')	Amp/Seq	Reference
1F	CTG GTG GAC TCG GTG ATG TTC TTG	Amp	Evans et al. 2000
9R	CTC TTC TAG CCT GCC AAT GAA CC	Amp	Evans et al. 2000
3R	TCR AGG AAC AYR GGG TGA TC	Seq	Small 2004
3F	ACT GTY CGR TTC TTC CAC	Seq	Small 2004
6R	AGA GCA GTG TGC CAA TCA TTG	Seq	Small 2004
8R	TCA CCR GAW ACA AGC TCC TG	Seq	Small 2004
8F	CCT GTC AAG GGA AGG AAA AT	Seq	Small 2004

Table 2. Plants samples collected in this study. Herbarium specimen ID numbers are provided, when applicable, as well as the dates of collection, the general site of collection, the species found, and the precise latitude and longitude of collection (stored in decimal degrees using a WGS 1984 latitude-longitude pseudoprojection). Also provided are details on whether the plant was used for single-gene Sanger genetic sequencing or genome-wide next-generation genetic sequencing.

Date	Herbarium Specimen ID	Name for genetics	Genetic analysis used in	Site	Species found	Latitude	Longitude
9/5/14	D-090514-2	D1	Single-gene	Boggy Slough	*H. dasycalyx*	31.287688	-94.908549
4/1/16	D-040116-73	D11	Genome-wide	Lovelady Preserve	*H. dasycalyx*	31.100892	-95.476486
4/1/16	D-040116-60	D1c	Genome-wide	Lovelady Preserve	*H. dasycalyx*		
9/5/14	D-090514-3	D2	Single-gene	Boggy Slough	*H. dasycalyx*	31.287688	-94.908549
9/6/14	D-090614-25	D24	Single-gene	Boggy Slough	*H. dasycalyx*	31.28261375	-94.90192413
9/6/14	D-090614-26	D25	Single-gene	Boggy Slough	*H. dasycalyx*	31.28261375	-94.90192413
9/12/14	D-091214-37	D37	Single-gene	Mud Creek/Hwy 204 ROW	*H. dasycalyx*	31.901167	-95.0995
9/12/14	D-091214-38	D38	Single-gene	Mud Creek/Hwy 204 ROW	*H. dasycalyx*	31.901167	-95.0995
10/24/14	D-102414-48	D48	Single-gene	Lovelady Preserve	*H. dasycalyx*	31.101368	-95.476899
10/24/14	D-102414-49	D49	Single-gene	Lovelady Preserve	*H. dasycalyx*	31.101368	-95.476899
4/1/16	D-040116-59	D5a	Genome-wide	Lovelady Preserve	*H. dasycalyx*	31.101061	-95.476922
4/1/16	D-040116-64	D6b	Genome-wide	Lovelady Preserve	*H. dasycalyx*	31.101397	-95.477181
4/1/16	D-040116-67	D8	Genome-wide	Lovelady Preserve	*H. dasycalyx*	31.101367	-95.476806
4/1/16	D-040116-65	D9a	Genome-wide	Lovelady Preserve	*H. dasycalyx*	31.101333	-95.476861
6/29/14	L-062914-109	L10	Single-gene	Dallas Trinity River	*H. laevis*	32.703228	-96.7043896
2014		L10	Single-gene		*H. laevis*	32.703228	-96.7043896
2014		L11	Single-gene and genome-wide		*H. laevis*	33.32034211	-95.80344937
7/11/14	L-071114-112	L13	Single-gene	HWY 294 Cherokee & Anderson County line Neches River	*H. laevis*	31.629037	-95.284583
2014		L14	Single-gene		*H. laevis*	32.63586	-94.67286
2014		L15	Single-gene and genome-wide		*H. laevis*	32.63586	-94.67286
2014		L31	Single-gene and genome-wide		*H. laevis*	32.67161	-94.42331
2014		L32	Single-gene		*H. laevis*	32.67161	-94.42331
2014		L4	Single-gene		*H. laevis*	32.703079	-96.704306
10/22/14	L-102214-139	L40	Single-gene	Near Boggy Slough under bridge on Neches River	*H. laevis*	31.286128	-94.8912

10/22/14	L-102214-140	L41	Single-gene and Genome-wide	Near Boggy Slough under bridge on Neches River	*H. laevis*	31.286128	-94.8912
2014		L5	Single-gene		*H. laevis*	32.703079	-96.704306
2014		L6	Single-gene		*H. laevis*	32.70403465	-96.7043604
2014		L7	Single-gene		*H. laevis*	32.70403465	-96.7043604
2015		L59	Single-gene		*H. laevis*	Redacted	Redacted
2015		L60	Single-gene		*H. laevis*	Redacted	Redacted
2014		M10	Single-gene and genome-wide		*H. moscheutos*	32.62743	-94.51598
2014		M11	Single-gene and genome-wide		*H. moscheutos*	32.62743	-94.51598
2014		M32	Single-gene		*H. moscheutos*	32.678896	-94.502723
2014		M33	Single-gene		*H. moscheutos*	32.678896	-94.502723
2014		M37	Single-gene and genome-wide		*H. moscheutos*	32.615227	-94.580391
2014		M38	Single-gene and genome-wide		*H. moscheutos*	35.58216476	-89.4261932
2014		M39	Single-gene		*H. moscheutos*	35.58216476	-89.4261932
2014		M48	Singe-gene		*H. moscheutos*	35.56882095	-89.4823608
2014		M49	Singe-gene		*H. moscheutos*	35.56882095	-89.4823608
2014		M5	Single-gene		*H. moscheutos*	32.313542	-95.46005
2014		M6	Single-gene		*H. moscheutos*	32.312607	-95.460385
6/30/14	M-063014-163	M7	Single-gene and genome-wide	Hwy 69 outside of Mineola towards Lindale	*H. moscheutos*	32.58368287	-95.458304
2014		M8	Single-gene		*H. moscheutos*	32.14039887	-95.3112737
2014		M9	Single-gene		*H. moscheutos*	32.14039887	-95.3112737
4/1/16	D-040116-61			Lovelady Preserve	*H. dasycalyx*		
4/1/16	D-040116-62			Lovelady Preserve	*H. dasycalyx*		
4/1/16	D-040116-63			Lovelady Preserve	*H. dasycalyx*		
4/1/16	D-040116-66			Lovelady Preserve	*H. dasycalyx*		
4/1/16	D-040116-68			Lovelady Preserve	*H. dasycalyx*	31.101356	-95.476761
4/1/16	D-040116-69			Lovelady Preserve	*H. dasycalyx*		
4/1/16	D-040116-70			Lovelady Preserve	*H. dasycalyx*	31.101292	-95.476608
4/1/16	D-040116-71			Lovelady Preserve	*H. dasycalyx*		
4/1/16	D-040116-72			Lovelady Preserve	*H. dasycalyx*		
7/6/16	D-070616-74			Lovelady Near Preserve	*H. dasycalyx*		

7/6/16	D-070616-75	Lovelady Near Preserve	*H. dasycalyx*		
8/30/16	D-083016-1	Boggy Slough	*H. dasycalyx*		
8/30/16	D-083016-76	Boggy Slough	*H. dasycalyx*		
8/30/16	D-083016-77	Boggy Slough	*H. dasycalyx*		
8/30/16	D-083016-78	Boggy Slough	*H. dasycalyx*		
8/30/16	D-083016-79	Boggy Slough	*H. dasycalyx*		
8/30/16	D-083016-80	Boggy Slough	*H. dasycalyx*		
8/30/16	D-083016-81	Boggy Slough	*H. dasycalyx*		
8/30/16	D-083016-82	Boggy Slough	*H. dasycalyx*		
8/30/16	D-083016-83	Boggy Slough	*H. dasycalyx*		
8/30/16	D-083016-84	Boggy Slough	*H. dasycalyx*		
8/30/16	D-083016-85	Boggy Slough	*H. dasycalyx*		
8/30/16	D-083016-86	Boggy Slough	*H. dasycalyx*		
8/30/16	D-083016-87	Boggy Slough	*H. dasycalyx*	31.287511	-94.912378
8/30/16	D-083016-88	Boggy Slough	*H. dasycalyx*	31.285278	-94.914231
8/30/16	D-083016-89	Boggy Slough	*H. dasycalyx*	31.284147	-94.914422
8/30/16	D-083016-90	Boggy Slough	*H. dasycalyx*	31.281808	-94.911383
8/30/16	D-083016-91	Boggy Slough	*H. dasycalyx*	31.287642	-94.926803
8/30/16	D-083016-92	Boggy Slough	*H. dasycalyx*	31.283794	-94.901628
8/30/16	D-083016-94	Boggy Slough	*H. dasycalyx*	31.287692	-94.908394
8/30/16	D-083016-95	Boggy Slough	*H. dasycalyx*	31.287858	-94.908608
8/30/16	D-083016-96	Boggy Slough	*H. dasycalyx*	31.324097	-94.913794
8/30/16	D-083016-97	Boggy Slough	*H. dasycalyx*	31.283183	-94.900658
9/5/14	D-090514-10	Boggy Slough	*H. dasycalyx*	31.287688	-94.908549
9/5/14	D-090514-11	Boggy Slough	*H. dasycalyx*	31.287688	-94.908549
9/5/14	D-090514-12	Boggy Slough	*H. dasycalyx*	31.287688	-94.908549
9/5/14	D-090514-13	Boggy Slough	*H. dasycalyx*	31.287688	-94.908549
9/5/14	D-090514-14	Boggy Slough	*H. dasycalyx*	31.287688	-94.908549
9/5/14	D-090514-15	Boggy Slough	*H. dasycalyx*	31.287688	-94.908549
9/5/14	D-090514-16	Boggy Slough	*H. dasycalyx*	31.287688	-94.908549
9/5/14	D-090514-17	Boggy Slough	*H. dasycalyx*	31.287688	-94.908549
9/5/14	D-090514-18	Boggy Slough	*H. dasycalyx*	31.287688	-94.908549
9/5/14	D-090514-19	Boggy Slough	*H. dasycalyx*	31.287688	-94.908549
9/5/14	D-090514-20	Boggy Slough	*H. dasycalyx*	31.287688	-94.908549
9/5/14	D-090514-21	Boggy Slough	*H. dasycalyx*	31.287688	-94.908549
9/5/14	D-090514-22	Boggy Slough	*H. dasycalyx*	31.287688	-94.908549

9/5/14	D-090514-23	Boggy Slough	*H. dasycalyx*	31.287688	-94.908549
9/5/14	D-090514-24	Boggy Slough	*H. dasycalyx*	31.287688	-94.908549
9/5/14	D-090514-4	Boggy Slough	*H. dasycalyx*	31.287688	-94.908549
;l....	D-090514-5	Boggy Slough	*H. dasycalyx*	31.287688	-94.908549
9/5/14	D-090514-6	Boggy Slough	*H. dasycalyx*	31.287688	-94.908549
9/5/14	D-090514-7	Boggy Slough	*H. dasycalyx*	31.287688	-94.908549
9/5/14	D-090514-8	Boggy Slough	*H. dasycalyx*	31.287688	-94.908549
9/5/14	D-090514-9	Boggy Slough	*H. dasycalyx*	31.287688	-94.908549
9/6/14	D-090614-27	Boggy Slough	*H. dasycalyx*	31.28261375	-94.90192413
9/6/14	D-090614-28	Boggy Slough	*H. dasycalyx*	31.28261375	-94.90192413
9/6/14	D-090614-29	Boggy Slough	*H. dasycalyx*	31.28261375	-94.90192413
9/6/14	D-090614-30	Boggy Slough	*H. dasycalyx*	31.28261375	-94.90192413
9/6/14	D-090614-31	Boggy Slough	*H. dasycalyx*	31.28261375	-94.90192413
9/6/14	D-090614-32	Boggy Slough	*H. dasycalyx*	31.28261375	-94.90192413
9/6/14	D-090614-33	Boggy Slough	*H. dasycalyx*	31.28261375	-94.90192413
9/6/14	D-090614-34	Boggy Slough	*H. dasycalyx*	31.28261375	-94.90192413
9/6/14	D-090614-35	Boggy Slough	*H. dasycalyx*	31.28261375	-94.90192413
9/6/14	D-090614-36	Boggy Slough	*H. dasycalyx*	31.28261375	-94.90192413
9/12/14	D-091214-39	Mud Creek/Hwy 204 ROW	*H. dasycalyx*	31.901167	-95.0995
9/12/14	D-091214-40	Mud Creek/Hwy 204 ROW	*H. dasycalyx*	31.901167	-95.0995
9/12/14	D-091214-41	Mud Creek/Hwy 204 ROW	*H. dasycalyx*	31.901167	-95.0995
9/12/14	D-091214-42	Mud Creek/Hwy 204 ROW	*H. dasycalyx*	31.901167	-95.0995
9/12/14	D-091214-43	Mud Creek/Hwy 204 ROW	*H. dasycalyx*	31.901167	-95.0995
9/12/14	D-091214-44	Mud Creek/Hwy 204 ROW	*H. dasycalyx*	31.901167	-95.0995
9/12/14	D-091214-45	Mud Creek/Hwy 204 ROW	*H. dasycalyx*	31.901167	-95.0995
9/12/14	D-091214-46	Mud Creek/Hwy 204 ROW	*H. dasycalyx*	31.901167	-95.0995
9/12/14	D-091214-47	Mud Creek/Hwy 204 ROW	*H. dasycalyx*	31.901167	-95.0995
9/12/16	D-091216-100	Mud Creek	*H. dasycalyx*	31.901042	-95.099389
9/12/16	D-091216-98	Mud Creek	*H. dasycalyx*	31.901042	-95.099464
9/12/16	D-091216-99	Mud Creek	*H. dasycalyx*	31.900628	-95.101417
9/28/16	D-092816-233	Boggy Slough	*H. dasycalyx*	31.2807	70' south of -94.90569
10/4/16	D-100416-235	Boggy Slough	*H. dasycalyx*	31.282321	-94.9116

10/4/16	D-100416-236	Boggy Slough	*H. dasycalyx*	31.282361	-94.91161
10/4/16	D-100416-237	Boggy Slough	*H. dasycalyx*	31.28241	-94.91086
10/4/16	D-100416-238	Boggy Slough	*H. dasycalyx*	31.28237	-94.91088
10/24/14	D-102414-50	Lovelady	*H. dasycalyx*	31.101368	-95.476899
10/24/14	D-102414-51	Lovelady	*H. dasycalyx*	31.101368	-95.476899
10/24/14	D-102414-52	Lovelady	*H. dasycalyx*	31.101368	-95.476899
10/24/14	D-102414-53	Lovelady	*H. dasycalyx*	31.101368	-95.476899
10/24/14	D-102414-54	Lovelady	*H. dasycalyx*	31.101368	-95.476899
10/24/14	D-102414-55	Lovelady	*H. dasycalyx*	31.101368	-95.476899
10/24/14	D-102414-56	Lovelady	*H. dasycalyx*	31.101368	-95.476899
10/24/14	D-102414-57	Lovelady	*H. dasycalyx*	31.101368	-95.476899
10/24/14	D-102414-58	Lovelady	*H. dasycalyx*	31.101368	-95.476899
8/30/16	DL-082216-227	Boggy Slough	*H. dasycalyx/H. laevis putative hybrid*	31.282853	-94.901244
9/6/14	DL-090614-226	Boggy ROW	*H. dasycalyx/H. laevis putative hybrid*	31.28261375	-94.90192413
9/12/16	DL-091216-228	Mudd Creek	*H. dasycalyx/H. laevis putative hybrid*	31.901056	-95.100478
9/12/16	DL-091216-229	Mud Creek	*H. dasycalyx/H. laevis putative hybrid*	31.900878	-95.100931
9/12/16	DL-091216-230	Mud Creek	*H. dasycalyx/H. laevis putative hybrid*	31.901186	-95.101097
9/12/16	DL-091216-231	Mud Creek	*H. dasycalyx/H. laevis putative hybrid*	31.901006	-95.101036
9/12/16	DL-091216-232	Mud Creek	*H. dasycalyx/H. laevis putative hybrid*	31.901133	-95.100128
10/4/16	DL-100416-239	Boggy Slough	*H. dasycalyx/H. laevis putative hybrid*	31.282691	-94.91125
10/4/16	DL-100416-240	Boggy Slough	*H. dasycalyx/H. laevis putative hybrid*	31.282361	-94.91126
10/4/16	DL-100416-241	Boggy Slough	*H. dasycalyx/H. laevis putative hybrid*	31.2876881	-94.908549

8/30/16	DM-083016-93	Boggy Slough	*H. dasycalyx/H. moscheutos putative hybrid*	31.285564	-94.903625
6/29/14	L-062914-103	Dallas Audubon Center	*H. laevis*	32.703079	-96.704346
6/29/14	L-062914-104	dallas Audubon Center	*H. laevis*	32.703079	-96.704346
6/29/14	L-062914-105	dallas Audubon Center	*H. laevis*	32.70403465	-96.7043604
6/29/14	L-062914-106	dallas Audubon Center	*H. laevis*	32.70403465	-96.7043604
6/29/14	L-062914-107	Dallas Audubon Center	*H. laevis*	32.70403465	-96.7043604
6/29/14	L-062914-108	Dallas Audubon Center	*H. laevis*	32.704029	-96.705193
6/30/14	L-063014-110	HWY 24 S of Cooper Jernigan Creek	*H. laevis*	33.32034211	-95.8034493
6/30/14	L-063014-111	HWY 24 S of Cooper Jernigan Creek	*H. laevis*	33.32034211	-95.8034493
7/29/14	L-072914-113		*H. laevis*	32.63586	
7/29/14	L-072914-114		*H. laevis*	32.63586	-94.67286
7/29/14	L-072914-115		*H. laevis*	32.63586	-94.67286
7/29/14	L-072914-116		*H. laevis*	32.63586	-94.67286
7/29/14	L-072914-117		*H. laevis*	32.63586	-94.67286
7/29/14	L-072914-118		*H. laevis*	32.63586	-94.67286
7/29/14	L-072914-119		*H. laevis*	32.63586	-94.67286
7/29/14	L-072914-120		*H. laevis*	32.63586	-94.67286
7/29/14	L-072914-121		*H. laevis*	32.63586	-94.67286
7/29/14	L-072914-122		*H. laevis*	32.63586	-94.67286
7/29/14	L-072914-123		*H. laevis*	32.63586	-94.67286
7/29/14	L-072914-124		*H. laevis*	32.63586	-94.67286
7/29/14	L-072914-125		*H. laevis*	32.63586	-94.67286
7/29/14	L-072914-126		*H. laevis*	32.63586	-94.67286
7/29/14	L-072914-127		*H. laevis*	32.63586	-94.67286
7/29/14	L-072914-128		*H. laevis*	32.63586	-94.67286
7/29/14	L-072914-129		*H. laevis*	32.63586	-94.67286
7/29/14	L-072914-130		*H. laevis*	32.67161	-94.42331
7/29/14	L-072914-131		*H. laevis*	32.67161	-94.42331
7/29/14	L-072914-132		*H. laevis*	32.67161	-94.42331
7/29/14	L-072914-133		*H. laevis*	32.67161	-94.42331
7/29/14	L-072914-134		*H. laevis*	32.67161	-94.42331
7/29/14	L-072914-135		*H. laevis*	32.67161	-94.42331
7/29/14	L-072914-136		*H. laevis*	32.67161	-94.42331
7/29/14	L-072914-137		*H. laevis*	32.67161	-94.42331

7/29/14	L-072914-138		*H. laevis*	32.67161	-94.42331
8/12/16	L-081216-152	Cypress Creek	*H. laevis*	32.635642	-94.672592
8/12/16	L-081216-153	Cypress Creek	*H. laevis*	32.635642	-94.672447
8/12/16	L-081216-154	Cypress Creek	*H. laevis*	32.635117	-94.672975
8/12/16	L-081216-155	Cypress Creek	*H. laevis*	32.635114	-94.672997
8/12/16	L-081216-156	Cypress Creek	*H. laevis*	32.635053	-94.673003
8/12/16	L-081216-157	Cypress Creek	*H. laevis*	32.634936	-94.672858
8/12/16	L-081216-158	Cypress Creek	*H. laevis*	32.636397	-94.673033
8/15/13	L-081513-101	HWY 204, Mud Creek Site 1	*H. laevis*	31.901167	-95.0995
8/15/13	L-081513-102	Side of road	*H. laevis*		
8/22/16	L-082216-144	Striker Creek	*H. laevis*	32.003478	-94.990433
8/22/16	L-082216-145	Striker Creek	*H. laevis*	32.003522	-94.990433
8/22/16	L-082216-146	Striker Creek	*H. laevis*	32.003161	-94.992019
8/22/16	L-082216-147	Mud Creek	*H. laevis*	31.855847	-94.997917
8/22/16	L-082216-148	Mud Creek	*H. laevis*	31.855628	-94.997908
8/22/16	L-082216-149	Mud Creek& Hwy 79	*H. laevis*	31.976274	-95.166628
8/22/16	L-082216-150	Striker Creek	*H. laevis*	32.003535	-94.991994
8/22/16	L-082216-151	Mud Creek	*H. laevis*	31.856039	-94.99665
10/4/16	L-100416-234	Boggy Slough	*H. laevis*	31.282291	-94.9148
10/22/14	L-102214-141	near boggy under bridge on neches	*H. laevis*	31.286128	-94.8912
10/22/14	L-102214-142	near boggy under bridge on neches	*H. laevis*	31.286128	-94.8912
10/22/14	L-102214-143	near boggy under bridge on neches	*H. laevis*	31.2867	-94.888842
6/28/14	M-062814-161	River Oaks Chandler TX	*H. moscheutos*	32.313542	-95.46005
6/28/14	M-062814-162	side of road Chandler TX	*H. moscheutos*	32.312607	-95.460385
7/22/14	M-072214-164	Hwy 69S out of Tyler	*H. moscheutos*	32.14039887	-95.3112737
7/22/14	M-072214-165	Hwy 69S out of Tyler	*H. moscheutos*	32.14039887	-95.3112737
7/29/14	M-072914-166	Lammy Site Harrison Cnty	*H. moscheutos*	32.62743	-94.51598
7/29/14	M-072914-167	Lammy Site Harrison Cnty	*H. moscheutos*	32.62743	-94.51598
7/29/14	M-072914-168	Lammy Site Harrison Cnty	*H. moscheutos*	32.62743	-94.51598

7/29/14	M-072914-169	Lammy Site Harrison Cnty	*H. moscheutos*	32.62743	-94.51598
7/29/14	M-072914-170	Lammy Site Harrison Cnty	*H. moscheutos*	32.62743	-94.51598
7/29/14	M-072914-171	Lammy Site Harrison Cnty	*H. moscheutos*	32.62743	-94.51598
7/29/14	M-072914-172	Lammy Site Harrison Cnty	*H. moscheutos*	32.62743	-94.51598
7/29/14	M-072914-173	Lammy Site Harrison Cnty	*H. moscheutos*	32.62743	-94.51598
7/29/14	M-072914-174	Lammy Site Harrison Cnty	*H. moscheutos*	32.62743	-94.51598
7/29/14	M-072914-175	Lammy Site Harrison Cnty	*H. moscheutos*	32.62743	-94.51598
7/29/14	M-072914-176	Lammy Site Harrison Cnty	*H. moscheutos*	32.62743	-94.51598
7/29/14	M-072914-177	Lammy Site Harrison Cnty	*H. moscheutos*	32.62743	-94.51598
7/29/14	M-072914-178	Lammy Site Harrison Cnty	*H. moscheutos*	32.62743	-94.51598
7/29/14	M-072914-179	Lammy Site Harrison Cnty	*H. moscheutos*	32.62743	-94.51598
7/29/14	M-072914-180	Lammy Site Harrison Cnty	*H. moscheutos*	32.62743	-94.51598
7/29/14	M-072914-181	Lammy Site Harrison Cnty	*H. moscheutos*	32.62743	-94.51598
7/29/14	M-072914-182	Lammy Site Harrison Cnty	*H. moscheutos*	32.62743	-94.51598
7/29/14	M-072914-183	Lammy Site Harrison Cnty	*H. moscheutos*	32.62743	-94.51598
7/29/14	M-072914-184	Lammy Site Harrison Cnty	*H. moscheutos*	32.62743	-94.51598
7/29/14	M-072914-185	Lammy Site Harrison Cnty	*H. moscheutos*	32.62743	-94.51598
7/29/14	M-072914-186	Lammy Site Harrison Cnty	*H. moscheutos*	32.62743	-94.51598
7/29/14	M-072914-187	Lammy Site Harrison Cnty	*H. moscheutos*	32.62743	-94.51598

7/29/14	M-072914-188	Lammy Site Harrison Cnty	*H. moscheutos*	32.678896	-94.502723
7/29/14	M-072914-189	Lammy Site Harrison Cnty	*H. moscheutos*	32.678896	-94.502723
7/29/14	M-072914-190	Lammy Site Harrison Cnty	*H. moscheutos*	32.678896	-94.502723
7/29/14	M-072914-191	Lammy Site Harrison Cnty	*H. moscheutos*	32.678896	-94.502723
7/29/14	M-072914-192	Lammy Site Harrison Cnty	*H. moscheutos*	32.678896	-94.502723
7/29/14	M-072914-193	ROW Harrison Cnty	*H. moscheutos*	32.615227	-94.580391
8/16/14	M-081614-194	around pond TN Hwy 54	*H. moscheutos*	35.58216476	-89.42619323
8/16/14	M-081614-195	around pond TN Hwy 54	*H. moscheutos*	35.58216476	-89.42619323
8/16/14	M-081614-196	around pond TN Hwy 54	*H. moscheutos*	35.58216476	-89.42619323
8/16/14	M-081614-197	around pond TN Hwy 54	*H. moscheutos*	35.58216476	-89.42619323
8/16/14	M-081614-198	around pond TN Hwy 54	*H. moscheutos*	35.58216476	-89.42619323
8/16/14	M-081614-199	around pond TN Hwy 54	*H. moscheutos*	35.58216476	-89.42619323
8/16/14	M-081614-200	around pond TN Hwy 54	*H. moscheutos*	35.58216476	-89.42619323
8/16/14	M-081614-201	around pond TN Hwy 54	*H. moscheutos*	35.58216476	-89.42619323
8/16/14	M-081614-202	around pond TN Hwy 54	*H. moscheutos*	35.58216476	-89.42619323
8/16/14	M-081614-203	around pond TN Hwy 54	*H. moscheutos*	35.58216476	-89.42619323
8/16/14	M-081614-204	ROW TN Hwy 54	*H. moscheutos*	35.56882095	-89.48236083
8/16/14	M-081614-205	ROW TN Hwy 54	*H. moscheutos*	35.56882095	-89.48236083
8/16/14	M-081614-206	ROW TN Hwy 54	*H. moscheutos*	35.56882095	-89.48236083

8/16/14	M-081614-207	ROW TN Hwy 54	*H. moscheutos*	35.56882095	-89.4823608
8/16/14	M-081614-208	ROW TN Hwy 54	*H. moscheutos*	35.56882095	-89.4823608
8/16/14	M-081614-209	ROW TN Hwy 54	*H. moscheutos*	35.56882095	-89.4823608
8/16/14	M-081614-210	ROW TN Hwy 54	*H. moscheutos*	35.56882095	-89.4823608
8/16/14	M-081614-211	ROW TN Hwy 54	*H. moscheutos*	35.56882095	-89.4823608
8/16/14	M-081614-212	ROW TN Hwy 54	*H. moscheutos*	35.56882095	-89.4823608
8/16/14	M-081614-213	ROW TN Hwy 54	*H. moscheutos*	35.56882095	-89.4823608
8/16/14	M-081614-214	ROW TN Hwy 54	*H. moscheutos*	35.56882095	-89.4823608
8/16/14	M-081614-215	ROW TN Hwy 54	*H. moscheutos*	35.56882095	-89.4823608
8/16/14	M-081614-216	ROW TN Hwy 54	*H. moscheutos*	35.56882095	-89.4823608
8/16/14	M-081614-217	ROW TN Hwy 54	*H. moscheutos*	35.56882095	-89.4823608
8/16/14	M-081614-218	ROW TN Hwy 54	*H. moscheutos*	35.56882095	-89.4823608
8/16/14	M-081614-219	ROW TN Hwy 54	*H. moscheutos*	35.56882095	-89.4823608
8/16/14	M-081614-220	ROW TN Hwy 54	*H. moscheutos*	35.56882095	-89.4823608
8/16/14	M-081614-221	ROW TN Hwy 54	*H. moscheutos*	35.56882095	-89.4823608
8/16/14	M-081614-222	ROW TN Hwy 54	*H. moscheutos*	35.56882095	-89.4823608
8/30/16	M-083016-223	Boggy Slough	*H. moscheutos*	31.319381	-94.927008
9/12/16	M-091216-225	Mud Creek/Hwy 204 ROW	*H. moscheutos*	31.901056	-95.100478
9/17/13	M-091713-159	Boggy Slough	*H. moscheutos*	31.282667	-94.902

9/17/13	M-091713-160	Boggy Slough	*H. moscheutos*	31.282667	-94.902
9/27/16	M-092716-224	Mud Creek Hwy 110	*H. moscheutos*	32.162522	-95.171586

Table 3. The plant samples used for analysis of molecular variance (AMOVA) of *GBSSI* to derive F_{ST} values. For each plant sample (rows), the designations of which population they belonged to, as well as their genetic identification numbers (columns), are provided.

Species	Population #	Identification #
H. dasycalyx	1	D48
H. dasycalyx	1	D49
H. dasycalyx	2	D2
H. dasycalyx	2	D24
H. dasycalyx	2	D25
H. dasycalyx	3	D37
H. dasycalyx	3	D38
H. laevis	4	L11
H. laevis	4	L12
H. laevis	5	L10
H. laevis	5	L6
H. laevis	5	L7
H. laevis	5	L4
H. laevis	5	L5
H. laevis	6	L59
H. laevis	6	L60
H. laevis	7	L1
H. laevis	7	L40
H. laevis	7	L41
H. laevis	8	L14
H. laevis	8	L15
H. laevis	8	L31
H. laevis	8	L32
H. moscheutos	9	M5
H. moscheutos	9	M6
H. moscheutos	10	M8
H. moscheutos	10	M9
H. moscheutos	11	M10
H. moscheutos	11	M11
H. moscheutos	11	M32
H. moscheutos	11	M33
H. moscheutos	11	M37
H. moscheutos	12	M38
H. moscheutos	12	M39
H. moscheutos	12	M48
H. moscheutos	12	M49

Table 4. Occurrence data used for the initial ecological niche modeling of *H. dasycalyx*, *H. laevis*, and *H. moscheutos* (highlighted in red) as selected from the total list of possible locations. The species at each location is noted, as well as the precise latitude and longitude of collection (stored in decimal degrees using a WGS 1984 latitude-longitude pseudoprojection). The coordinates of one location are redacted at the landowner's request; it is approximately 14km southwest of Groveton, TX. The modeling locations were selected so that they are at least 1km away from each other. This was done to minimize spatial autocorrelation and pseudoreplication (see Methods).

Species	Longitude	Latitude
H. dasycalyx	-95.0406	31.34355
H. dasycalyx	-94.9085	31.28769
H. dasycalyx	-94.9019	31.28261
H. dasycalyx	-95.0995	31.90117
H. dasycalyx	-95.4769	31.10137
H. dasycalyx	-95.0411	31.34297
H. dasycalyx	-94.9053	31.283
H. dasycalyx	-94.9085	31.28765
H. dasycalyx	-94.9086	31.28765
H. dasycalyx	-94.9086	31.28753
H. dasycalyx	-94.9087	31.28757
H. dasycalyx	-94.9088	31.28762
H. dasycalyx	-94.9088	31.2877
H. dasycalyx	-94.9088	31.28775
H. dasycalyx	-94.9089	31.28775
H. dasycalyx	-94.9086	31.28743
H. dasycalyx	-94.9086	31.28745
H. dasycalyx	-94.9087	31.28738
H. dasycalyx	-94.9087	31.28742
H. dasycalyx	-94.9086	31.28737
H. dasycalyx	-94.9086	31.28728
H. dasycalyx	-94.9086	31.28723
H. dasycalyx	-94.9085	31.2872
H. dasycalyx	-94.9086	31.28722
H. dasycalyx	-94.9086	31.2873
H. dasycalyx	-94.9086	31.28742
H. dasycalyx	-94.9084	31.28753
H. dasycalyx	-94.9084	31.28755
H. dasycalyx	-94.9085	31.2876
H. dasycalyx	-94.9084	31.28763
H. dasycalyx	-94.9084	31.28763
H. dasycalyx	-94.9083	31.28763
H. dasycalyx	-94.9086	31.28768
H. dasycalyx	-94.9085	31.28763
H. dasycalyx	-94.9085	31.28763
H. dasycalyx	-94.9083	31.28759
H. dasycalyx	-94.9083	31.28756
H. dasycalyx	-94.9083	31.28754

H. dasycalyx	-94.9083	31.28758
H. dasycalyx	-94.9084	31.28761
H. dasycalyx	-94.9084	31.28753
H. dasycalyx	-94.9084	31.28755
H. dasycalyx	-94.9085	31.28714
H. dasycalyx	-94.9084	31.28701
H. dasycalyx	-94.9087	31.28741
H. dasycalyx	-94.9087	31.28741
H. dasycalyx	-94.9087	31.28743
H. dasycalyx	-94.9087	31.28747
H. dasycalyx	-94.9087	31.28747
H. dasycalyx	-94.9086	31.28766
H. dasycalyx	-94.9083	31.2878
H. dasycalyx	-94.9054	31.28303
H. dasycalyx	-94.9017	31.28453
H. dasycalyx	-94.9022	31.28498
H. dasycalyx	-94.9022	31.285
H. dasycalyx	-94.9024	31.285
H. dasycalyx	-94.9027	31.28504
H. dasycalyx	-94.9027	31.28516
H. dasycalyx	-94.9027	31.28505
H. dasycalyx	-94.9027	31.28501
H. dasycalyx	-94.9027	31.28501
H. dasycalyx	-94.9026	31.28525
H. laevis	Redacted	Redacted
H. laevis	-95.0995	31.90117
H. laevis	-96.7043	32.70308
H. laevis	-96.7044	32.70435
H. laevis	-96.7052	32.70403
H. laevis	-96.7044	32.70323
H. laevis	-95.8034	33.32034
H. laevis	-95.2846	31.62904
H. laevis	-94.6729	32.63586
H. laevis	-94.4233	32.67161
H. laevis	-95.57	33.65111
H. laevis	-95.5287	30.92031
H. laevis	-94.08	30.5
H. laevis	-94.751	32.67313
H. laevis	-95.6211	29.36615
H. laevis	-94.8888	31.2867
H. laevis	-95.1895	32.54451
H. laevis	-94.8912	31.28613
H. laevis	-95.364182	31.137274
H. laevis	-94.901721	31.282694
H. laevis	-95.102199	31.900913
H. moscheutos	-94.902	31.28267
H. moscheutos	-95.4601	32.31354

H. moscheutos	-95.4604	32.31261
H. moscheutos	-95.4583	32.58368
H. moscheutos	-95.3113	32.1404
H. moscheutos	-94.516	32.62743
H. moscheutos	-94.5027	32.6789
H. moscheutos	-94.5804	32.61523
H. moscheutos	-94.8986	31.2825
H. moscheutos	-95.7781	29.1225
H. moscheutos	-93.9019	30.26028
H. moscheutos	-94.3772	29.625
H. moscheutos	-94.3737	29.67281
H. moscheutos	-93.7946	30.55082
H. moscheutos	-95.3174	32.20589
H. moscheutos	-94.5156	32.62784
H. moscheutos	-94.5952	32.60126
H. moscheutos	-94.7935	32.46132
H. moscheutos	-94.3193	33.03356
H. moscheutos	-94.1879	32.69605
H. moscheutos	-94.9304	31.31666

Table 5. Area under the operator receiving curve based upon the set-aside test data (test AUC) for each species. The test AUC values for the initial ecological niche models and the updated models (columns) are shown for the three species (rows).

	Test AUC	
Species	Initial model	Updated model
H. laevis	0.851	0.907
H. moscheutos	0.747	0.787
H. dasycalyx	0.952	0.973

Table 6. Ground-truth locations for *H. laevis* (rows). In columns, the GPS coordinates of each location are provided (stored in decimal degrees using a WGS 1984 latitude-longitude pseudoprojection), as well as the habitat suitability that was predicted for that location (Suitability) and whether the species was found there or not (Present?).

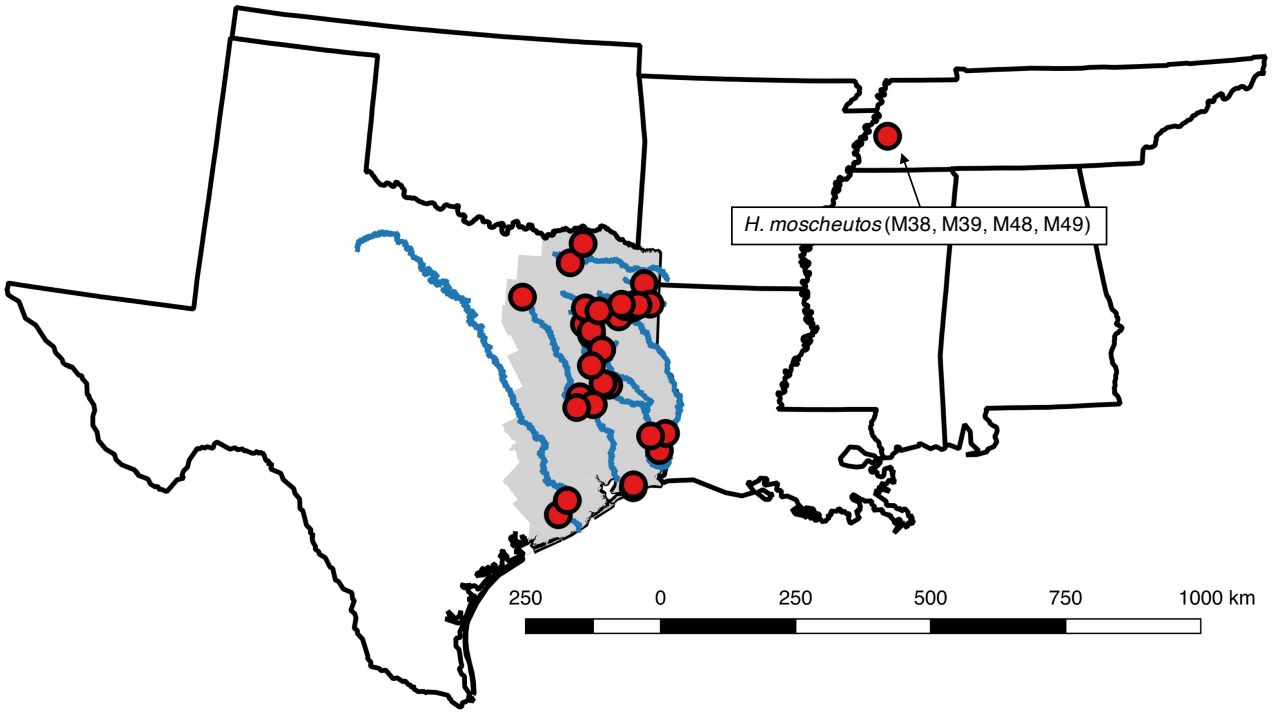
Longitude	Latitude	Suitability	Present?
-94.990433	32.003478	0.785971	Yes
-94.997917	31.855847	0.785971	Yes
-95.166628	31.976274	0.785971	Yes
-95.163078	32.158806	0.822833	No
-95.106658	31.885711	0.00424675	No
-95.16185	31.897114	0.00424675	No
-95.129081	31.911231	0.785971	No
-94.910179	32.197902	0.00832754	No
-96.256592	32.733667	0.156597	No
-94.9635	32.52625	0.391538	No
-95.485231	32.611242	0.574521	No

Table 7. Ground-truth locations for *H. moscheutos* (rows). In columns, the GPS coordinates of each location are provided (stored in decimal degrees using a WGS 1984 latitude-longitude pseudoprojection), as well as the habitat suitability that was predicted for that location (Suitability) and whether the species was found there or not (Present?).

Longitude	Latitude	Suitability	Present?
-94.997917	31.855847	0.476522	No
-95.106658	31.885711	0.175389	No
-95.16185	31.897114	0.175389	No
-95.129081	31.911231	0.476522	No
-95.166628	31.976274	0.476522	No
-94.990433	32.003478	0.476522	No
-95.485231	32.611242	0.364906	No
-96.256592	32.733667	0.10381	No
-95.100478	31.901056	0.476522	Yes
-95.171586	32.162522	0.995693	Yes
-94.9635	32.52625	0.725407	Yes
-94.910179	32.197902	0.503615	Yes

Table 8. Ground-truth locations for *H. dasycalyx* (rows). In columns, the GPS coordinates of each location are provided (stored in decimal degrees using a WGS 1984 latitude-longitude pseudoprojection), as well as the habitat suitability that was predicted for that location (Suitability) and whether the species was found there or not (Present?).

Longitude	Latitude	Suitability	Present?
-94.990433	32.003522	0.826178	No
-94.997917	31.855847	0.826178	No
-95.166628	31.976274	0.826178	No
-95.163078	32.158806	0.772046	No
-95.106658	31.885711	0.00026805	No
-95.16185	31.897114	0.00026805	No
-95.129081	31.911231	0.826178	No
-94.910179	32.197902	0.00038363	No
-96.256592	32.733667	1.55E-07	No
-94.9635	32.52625	0.164213	No
-95.485231	32.611242	0.224171	No
-94.914422	31.284147	0.659131	Yes
-94.926803	31.287642	0.659131	Yes
-94.902608	31.2852469	0.857051	Yes
-94.913794	31.324097	0.857051	Yes


H. moscheutos (M38, M39, M48, M49)
250 0 250 500 750 1000 km

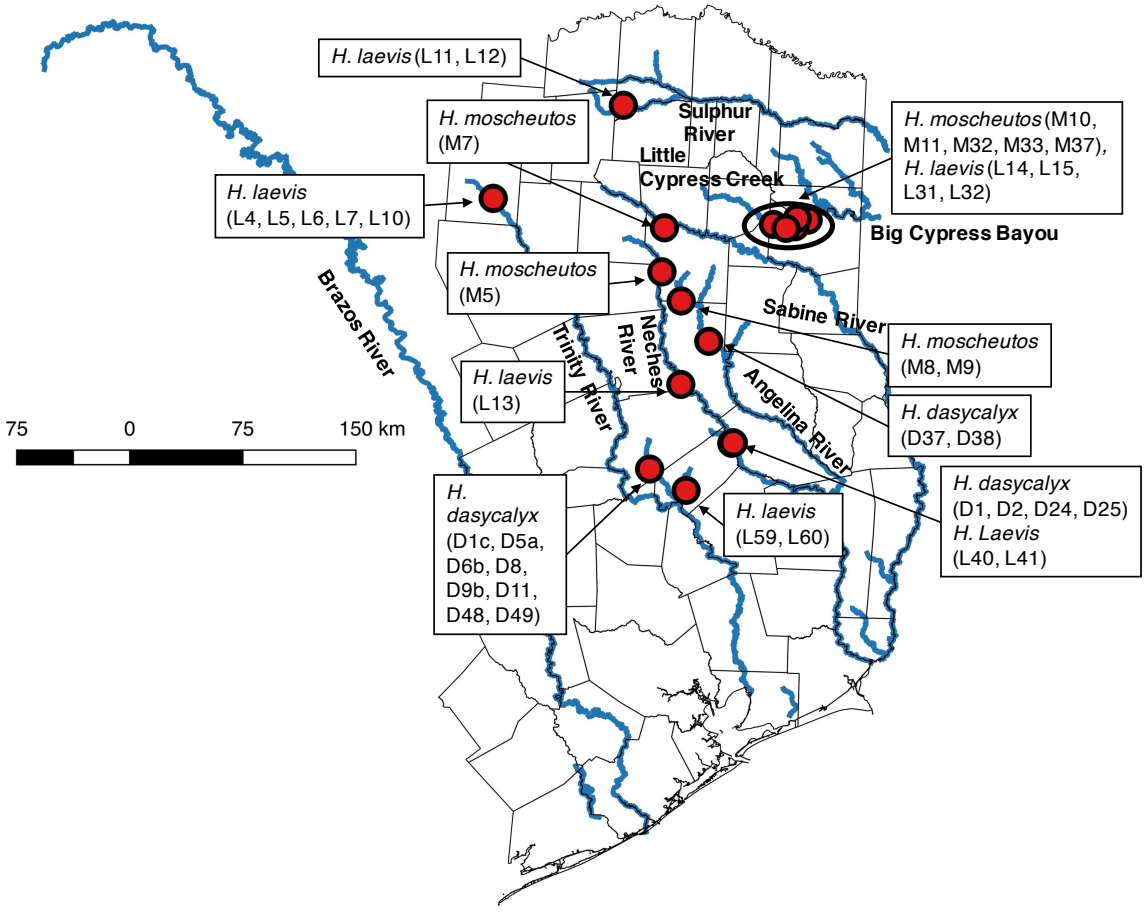

H. laevis (L11, L12)
H. moscheutos (M7)
Sulphur River
Little Cypress Creek
H. moscheutos (M10, M11, M32, M33, M37), H. laevis (L14, L15, L31, L32)
Big Cypress Bayou
H. laevis (L4, L5, L6, L7, L10)
H. moscheutos (M5)
Brazos River
Trinity River
Neches River
Sabine River
H. moscheutos (M8, M9)
H. laevis (L13)
Angelina River
H. dasycalyx (D37, D38)
H. dasycalyx (D1, D2, D24, D25) H. Laevis (L40, L41)
H. dasycalyx (D1c, D5a, D6b, D8, D9b, D11, D48, D49)
H. laevis (L59, L60)
75 0 75 150 km

Figure 1. Locations of the populations used in the genetic analyses (red circles). The identification numbers in parentheses refer to specific plants that were used for genetic analysis (see Table 1). The major rivers in East Texas are indicated in blue. The names of the East Texas rivers, as well as the East Texas county boundaries (black lines), are provided.

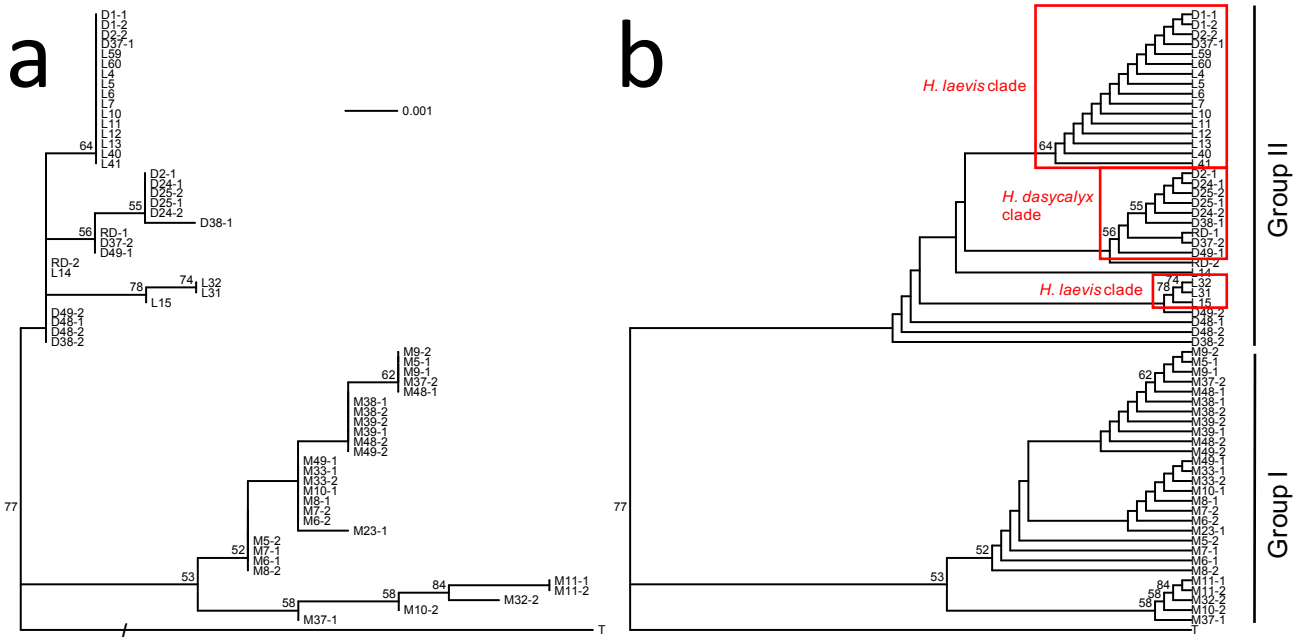


Figure 2. (a) Rooted maximum likelihood tree showing phylogenetic relationships of *H. dasycalyx*, *H. laevis*, and *H. moscheutos* inferred from *GBSSI* intron sequences, with proportional branch lengths. When heterozygotes were present, each phased allele was analyzed separately as if it were a separate individual. Alleles are indicated by "-1" and "-2" suffixes. Bootstrap values greater than 50% are shown at each node. Each individual (tip) is labeled by species. D represents *H. dasycalyx*, L represents *H. laevis*, M represents *H. moscheutos*, RD represents an *H. dasycalyx* individual from Small (2004), and T represents the outgroup *H. trionum*. The slash on the outgroup branch indicates that the length has been shortened for visualization purposes. (b) The same phylogeny, but without proportional branch lengths so that clades are more easily visible.

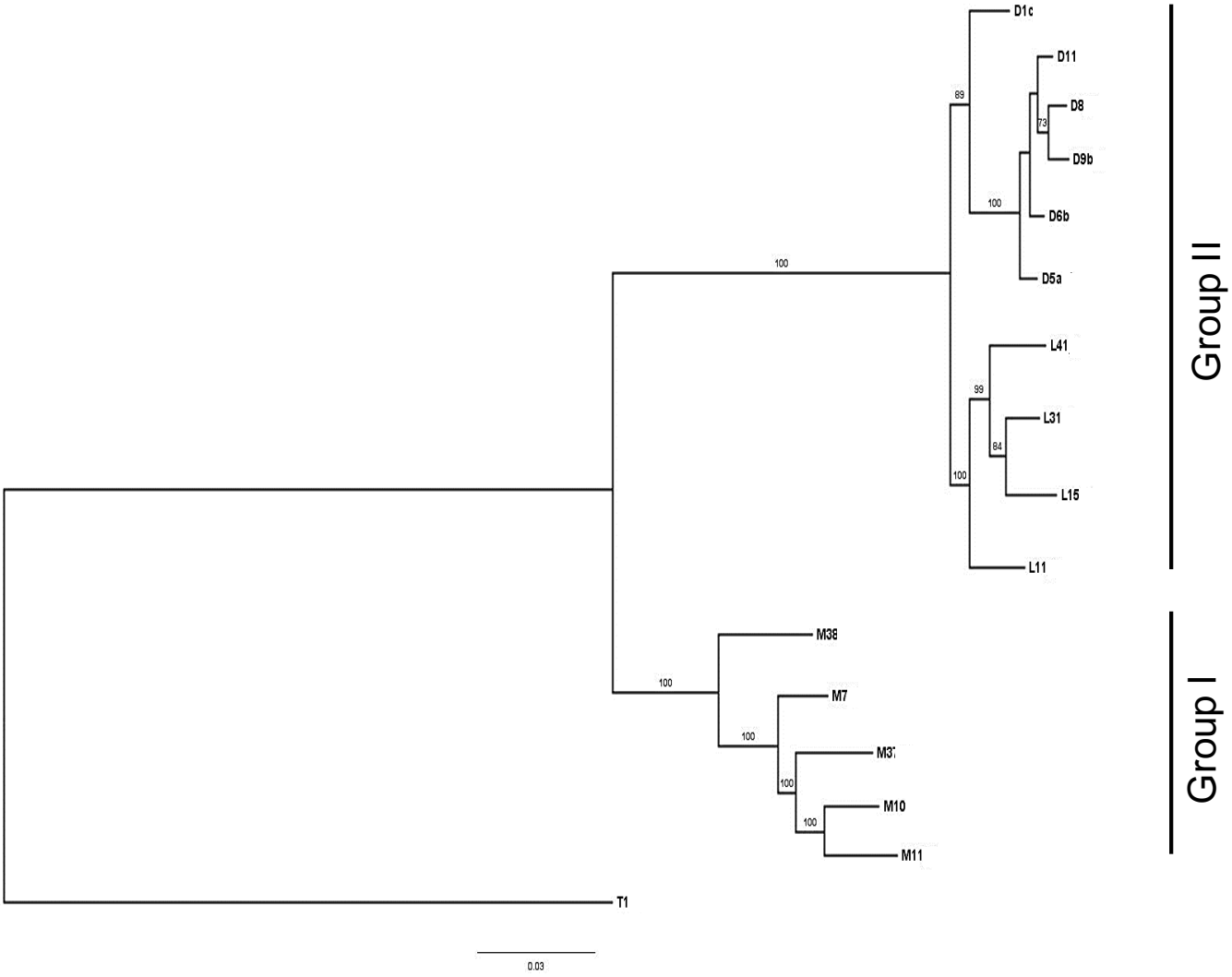


Figure 3. Rooted maximum likelihood tree showing phylogenetic relationships of *H. dasycalyx*, *H. laevis*, and *H. moscheutos* inferred from the genome-wide data. Bootstrap values greater than 60% are shown on each branch. Each accession is labeled by species D represent *H. dasycalyx*, L represents *H. laevis*, M represents *H. moscheutos*, and T1 represents the outgroup *H. trionum*.

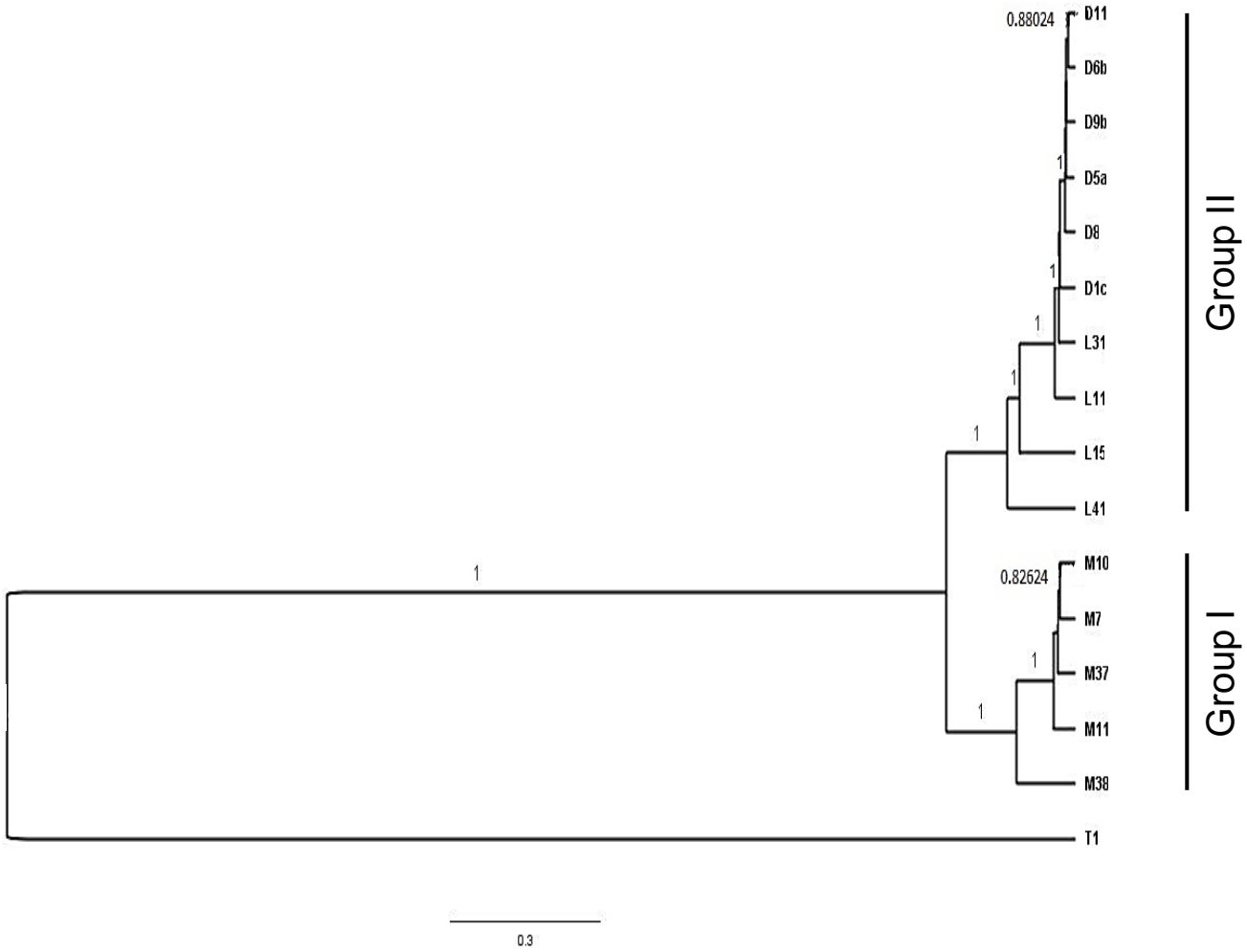


Figure 4. Rooted Bayesian coalescence tree showing phylogenetic relationships of *H. dasycalyx*, *H. laevis*, and *H. moscheutos* inferred from the genome-wide data. Posterior support values greater than 0.7 are shown above each branch. Each accession is labeled by species. D represents *H. dasycalyx*, L represents *H. laevis*, M represents *H. moscheutos*, and T1 represents the outgroup *H. trionum*.

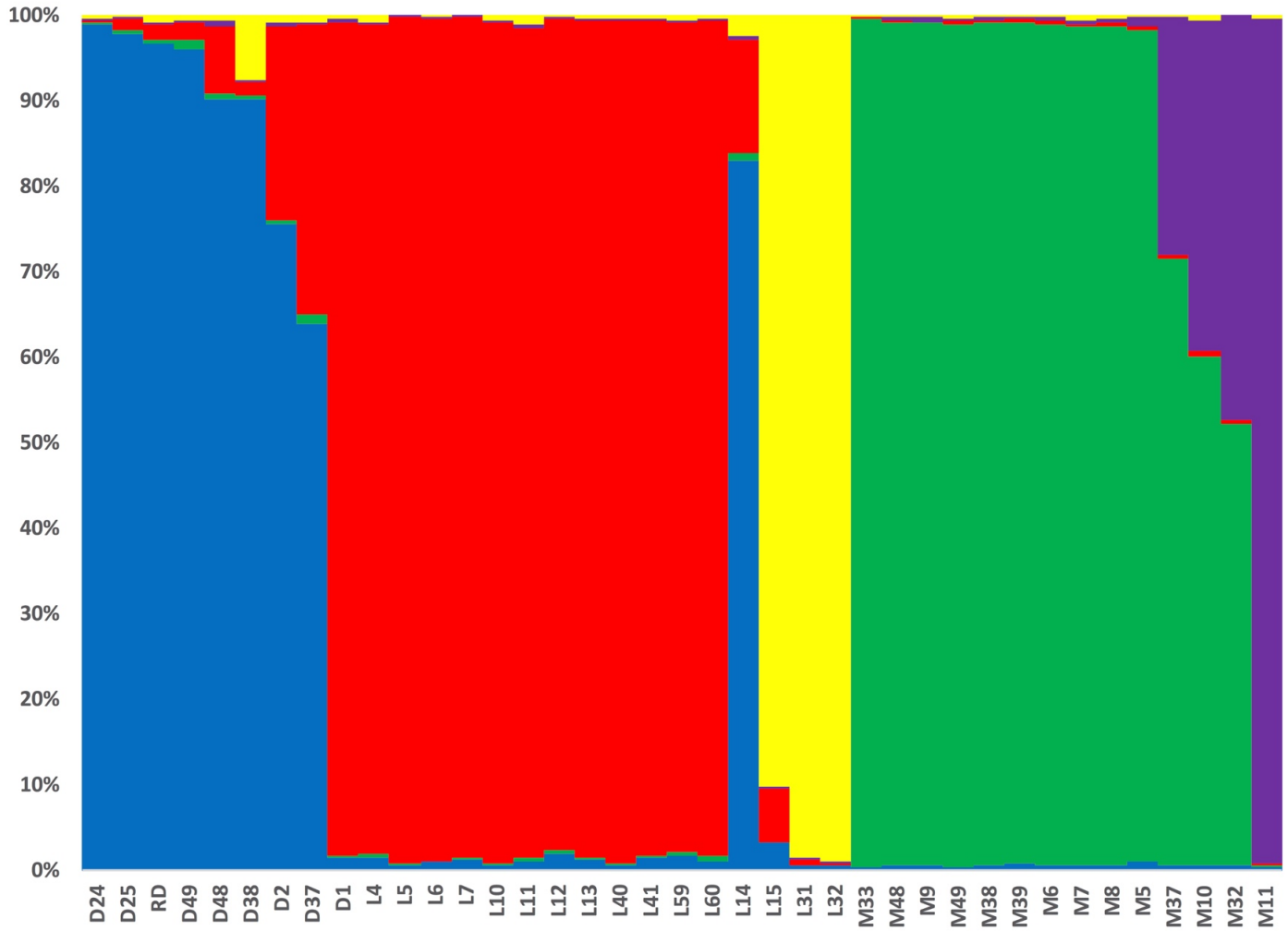


Figure 5. A graphical representation of the Bayesian cluster analysis of the *GBSSI* sequence data, showing the percentages of different inferred ancestries (*y*-axis) comprising each individual (*x*-axis). The different inferred ancestral groups are color-coded blue, red, yellow, green, and purple. D represents *H. dasycalyx*, L represents *H. laevis*, M represents *H. moscheutos*, and RD represents an *H. dasycalyx* individual from Small (2004).

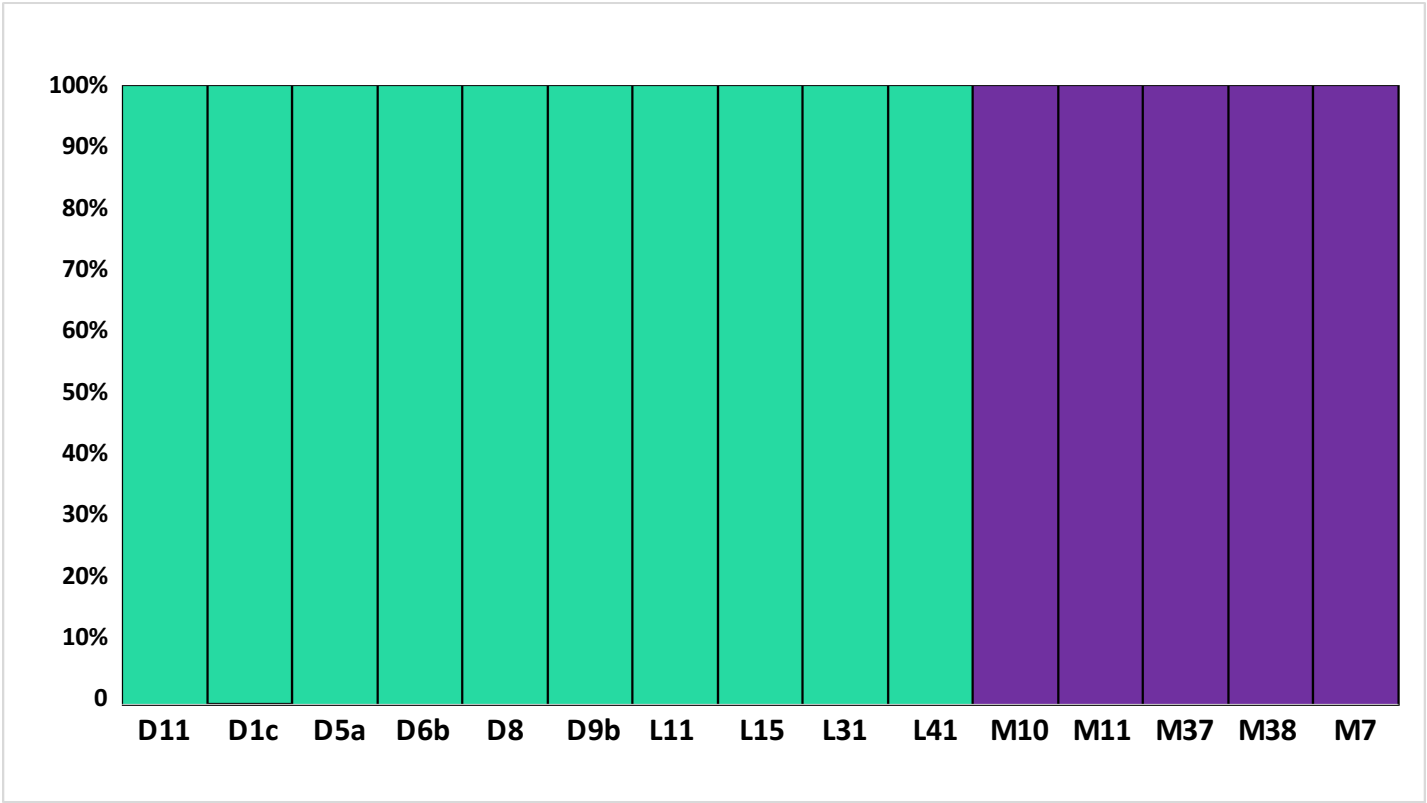


Figure 6. A graphical representation of the Bayesian cluster analysis of the genome-wide data, showing the percentages of different inferred ancestries (*y*-axis) comprising each individual (*x*-axis). The different inferred ancestral groups are color-coded green and purple. D represents *H. dasycalyx*, L represents *H. laevis*, and M represents *H. moscheutos*.

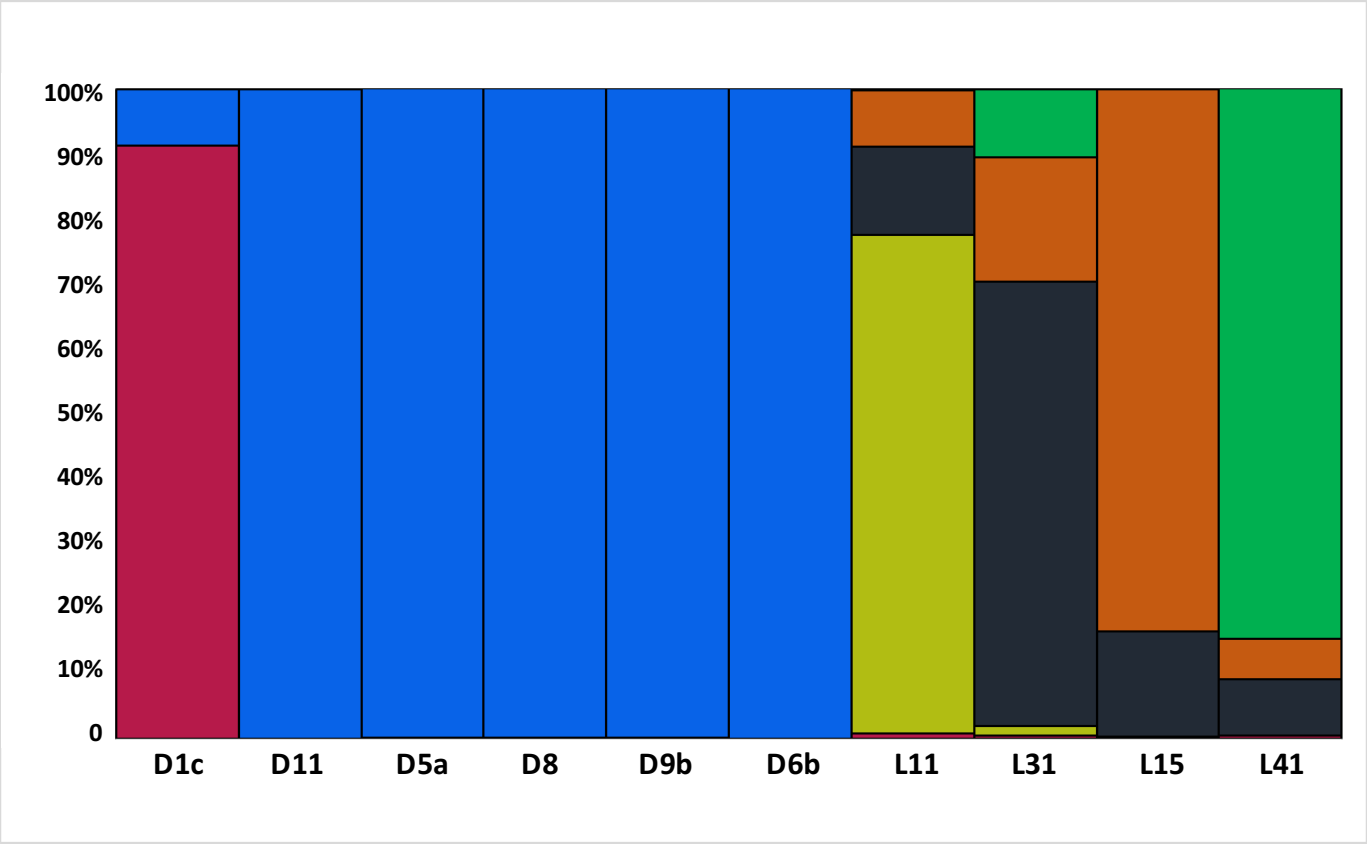


Figure 7. A graphical representation of the Bayesian cluster analysis of the genome-wide data, showing the percentages of different inferred ancestries (*y*-axis) comprising each individual (*x*-axis). The different inferred ancestral groups are color-coded red, blue, pea green, black, orange, and green. D represents *H. dasycalyx* and L represents *H. laevis*. *H. moscheutos* is excluded.

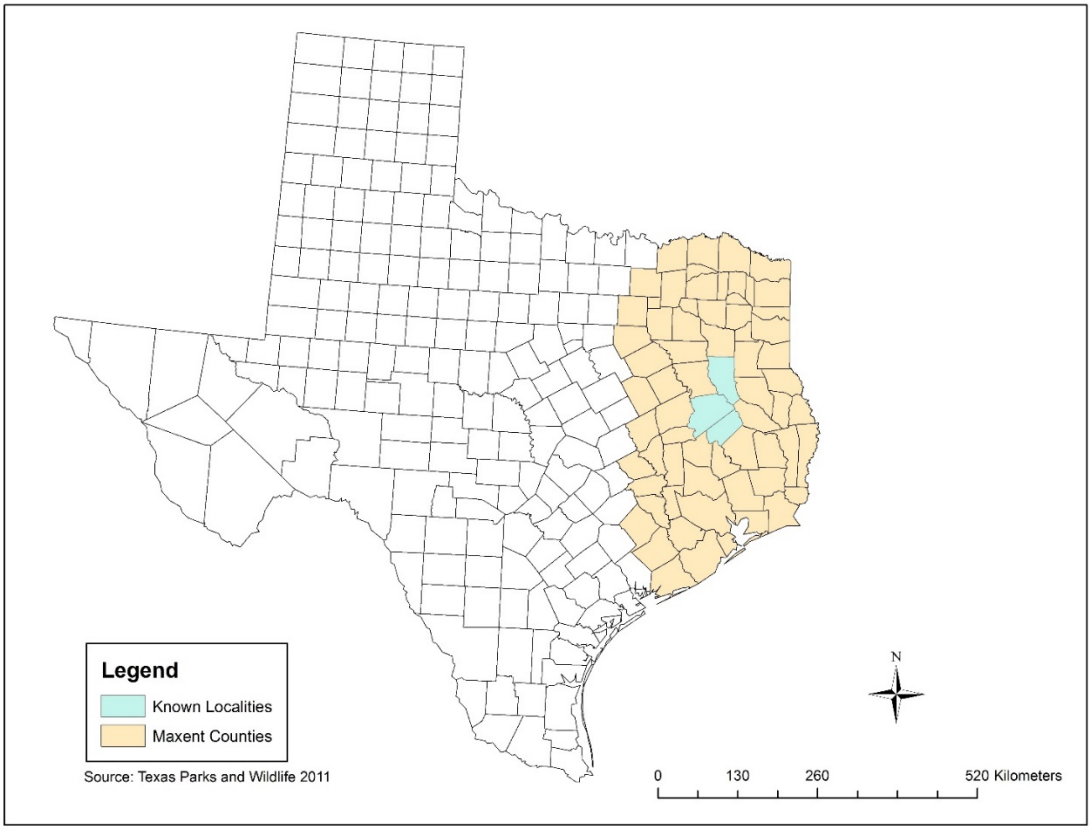
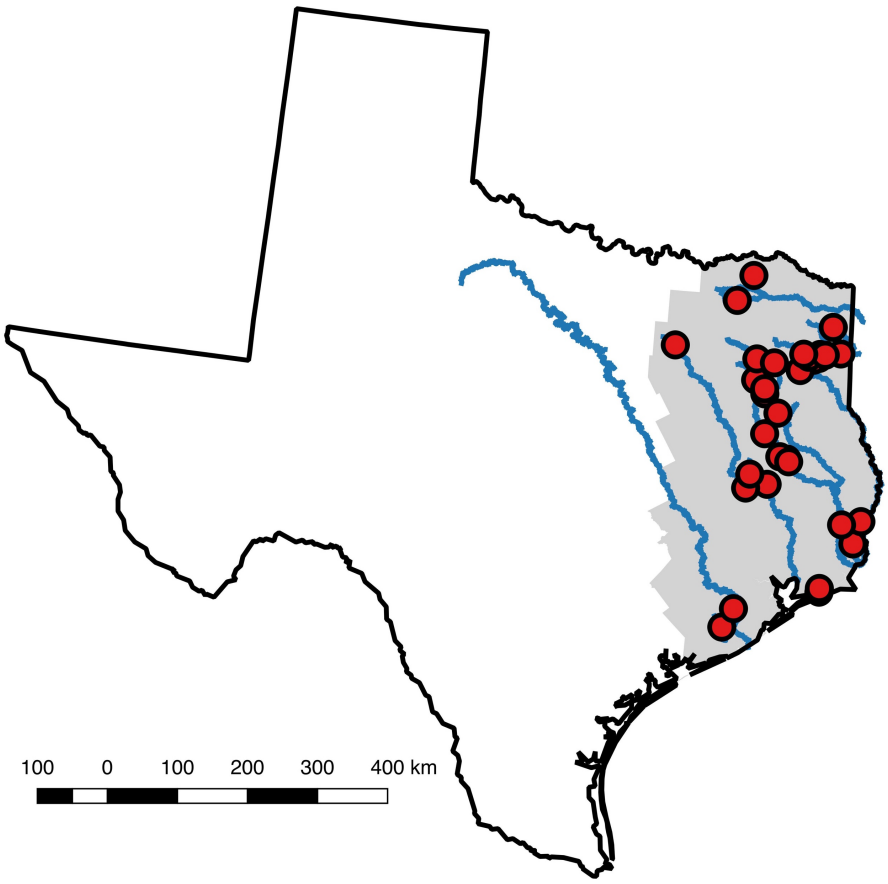


Figure 8. In blue: counties with existing documented native populations of *H. dasycalyx* in East Texas: Cherokee, Houston, and Trinity. In beige: counties included in the ecological niche modeling analyses.


100 0 100 200 300 400 km

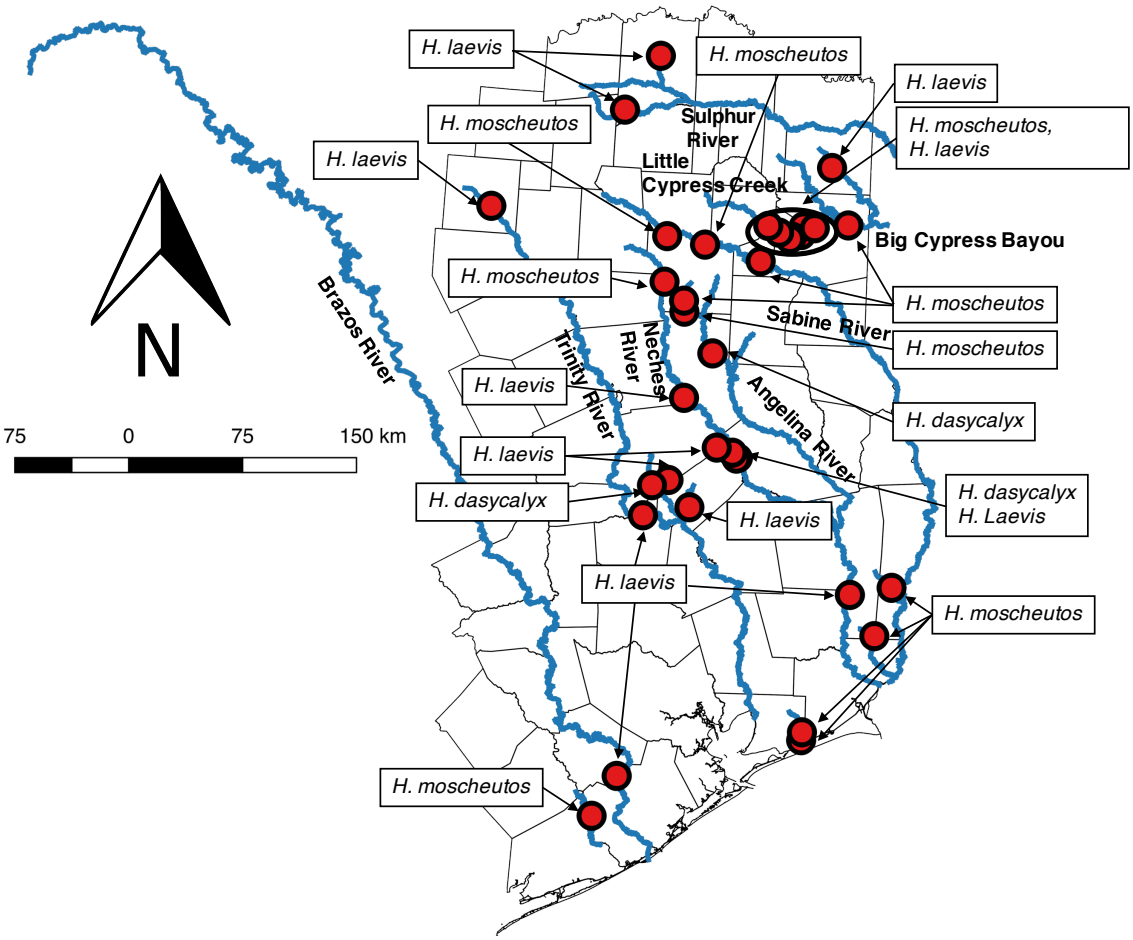

N
75 0 75 150 km
H. laevis
H. moscheutos
H. laevis
H. moscheutos
Sulphur River
H. moscheutos, H. laevis
H. laevis
Little Cypress Creek
Big Cypress Bayou
H. moscheutos
H. moscheutos
Sabine River
H. moscheutos
Brazos River
Trinity River
Neches River
Angelina River
H. laevis
H. dasycalyx
H. laevis
H. dasycalyx
H. dasycalyx H. Laevis
H. laevis
H. laevis
H. moscheutos
H. moscheutos

Figure 9. Locations of the populations used in the initial ecological niche modeling of the three species (red circles). The major rivers in East Texas are indicated in blue. The names of the East Texas rivers, as well as the East Texas county boundaries (black lines), are provided.

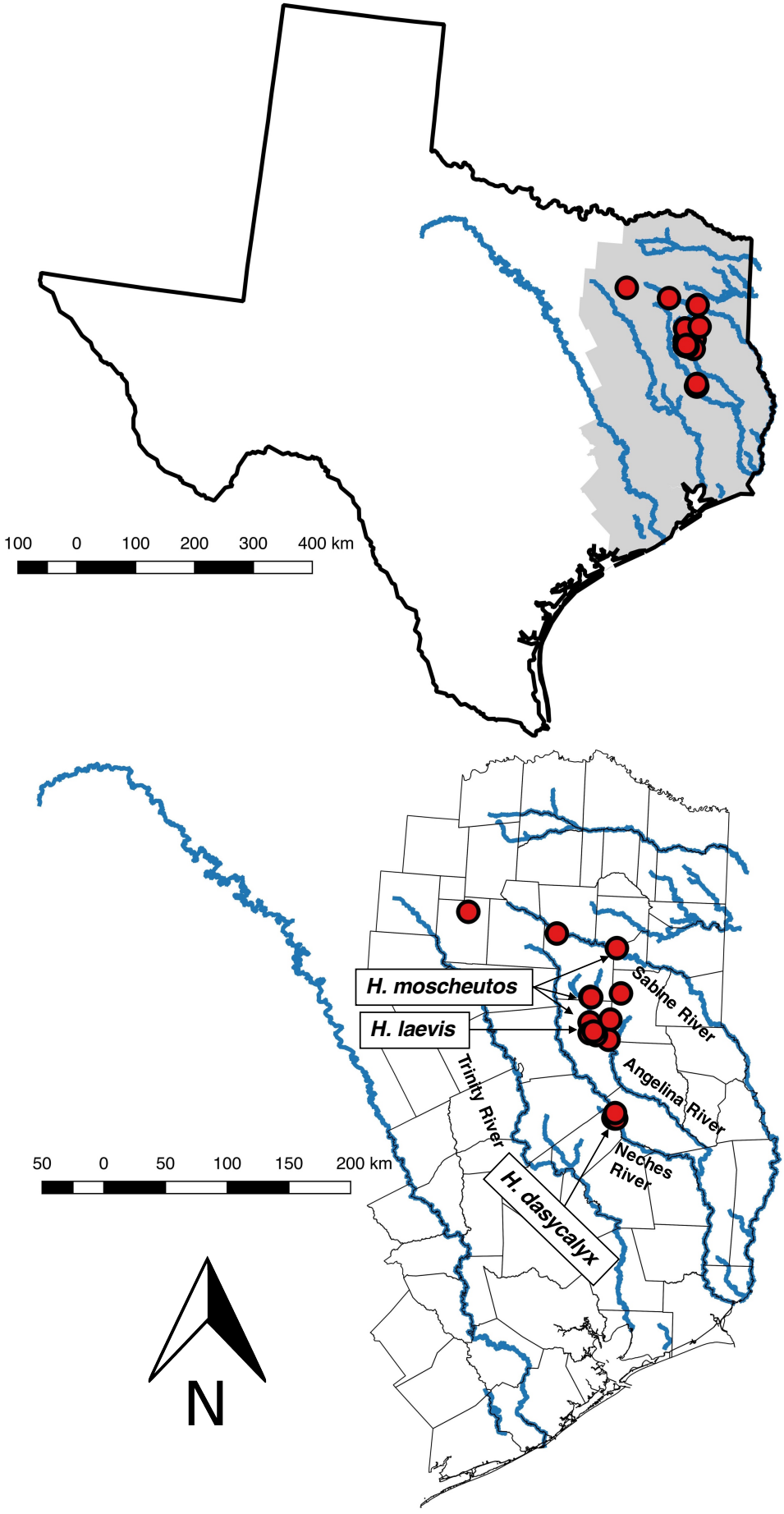

100 0 100 200 300 400 km
50 0 50 100 150 200 km
H. moscheutos
H. laevis
H. dasycalyx
Sabine River
Angelina River
Trinity River
Neches River
N

Figure 10. Ground-truth locations that were searched for the three species (red circles). The locations where plants were found are indicated with arrows and labels. The major rivers in East Texas are indicated in blue. The names of the East Texas rivers, as well as the East Texas county boundaries (black lines), are provided.

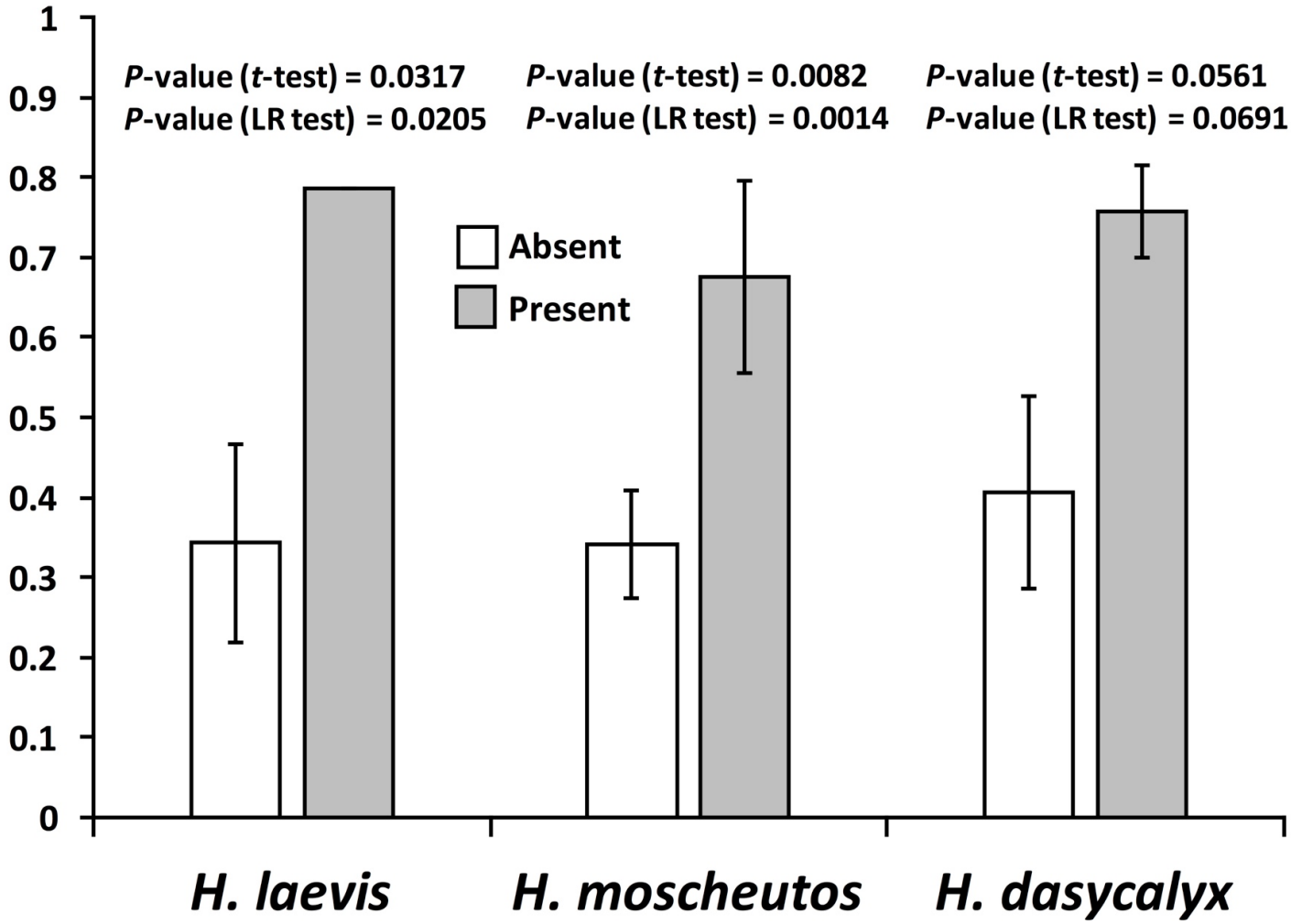


Figure 11. Ground-truthing of the initial ecological niche models for each of the three species. The white columns represent the average habitat suitability (from the ecological niche model) of the locations where the species was not found. The grey columns represent the average habitat suitability of the locations where the species was found. The bars represent ± 1 SE. For each species, the *P*-values from the Student's *t*-test and the likelihood ratio (LR) test are presented, indicating whether or not the difference in habitat suitability between the absence and presence locations was significant.

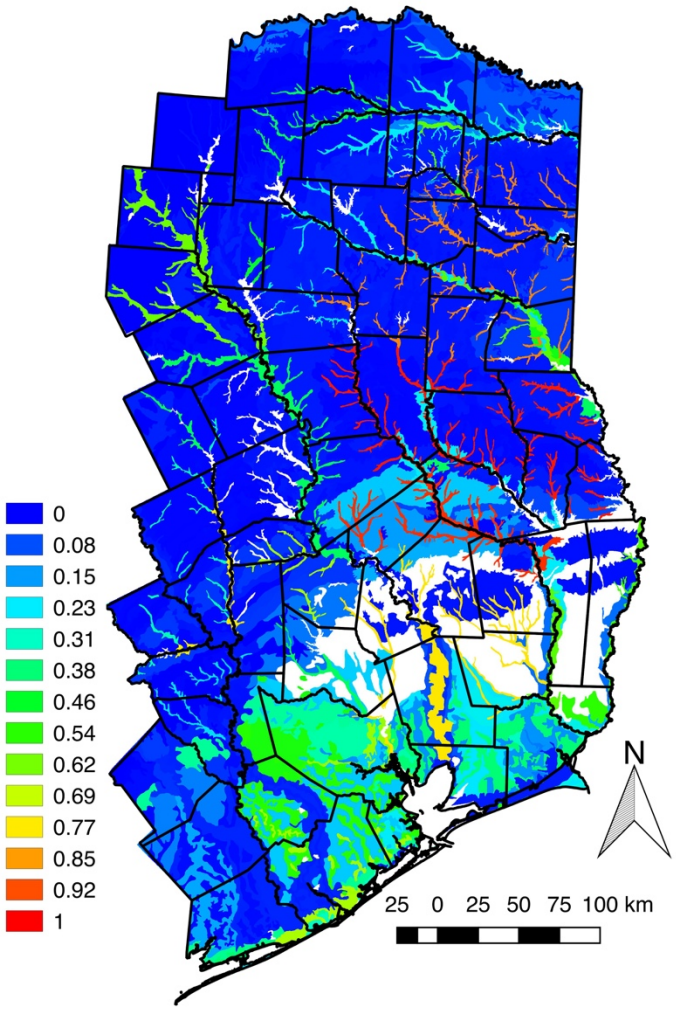


Figure 12. Habitat suitability for *H. laevis* across East Texas, as forecast by the updated ecological niche model that includes the ground-truthing data. The colors on the landscape represent the habitat suitability at that location at a gran size of 100m x 100m. Habitat suitabilities are color coded from zero to one, with zero being the least suitable habitat and one being the most suitable habitat. The East Texas county boundaries are outlined in black for context. White spaces indicate missing data.

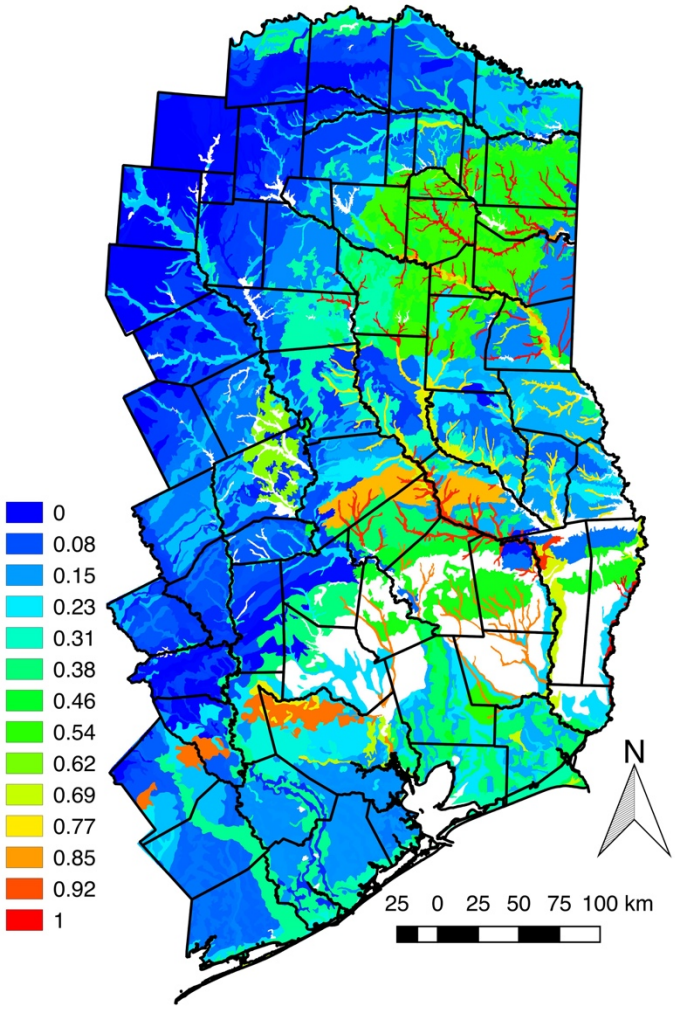


Figure 13. Habitat suitability for *H. moscheutos* across East Texas, as forecast by the updated ecological niche model that includes the ground-truthing data. The colors on the landscape represent the habitat suitability at that location at a gran size of 100m x 100m. Habitat suitabilities are color coded from zero to one, with zero being the least suitable habitat and one being the most suitable habitat. The East Texas county boundaries are outlined in black for context. White spaces indicate missing data.

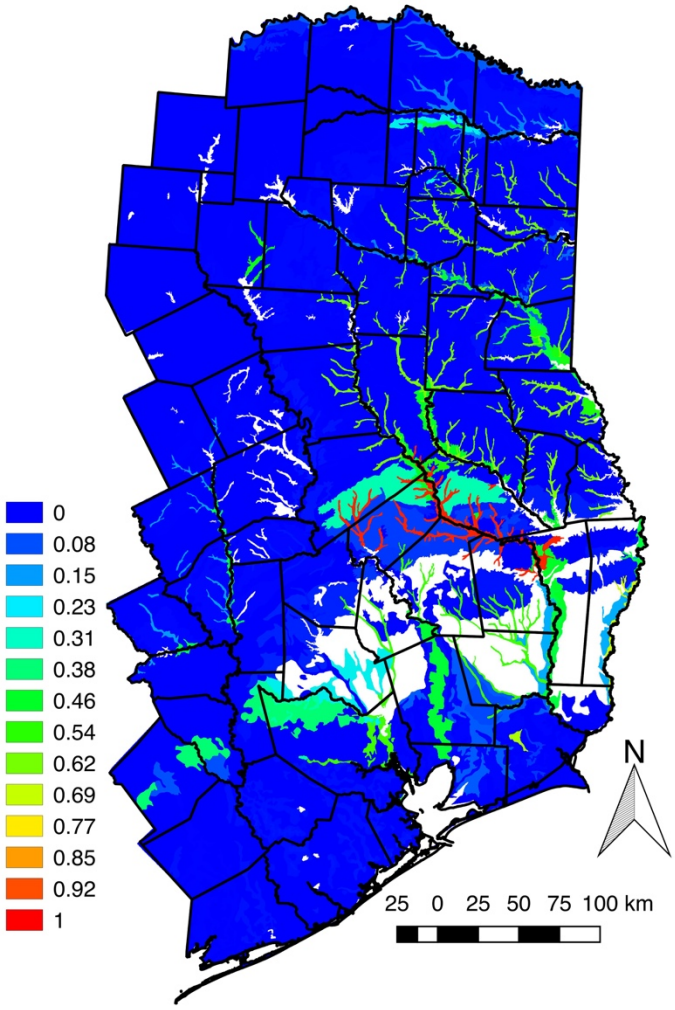


Figure 14. Habitat suitability for *H. dasycalyx* across East Texas, as forecast by the updated ecological niche model that includes the ground-truthing data. The colors on the landscape represent the habitat suitability at that location at a gran size of 100m x 100m. Habitat suitabilities are color coded from zero to one, with zero being the least suitable habitat and one being the most suitable habitat. The East Texas county boundaries are outlined in black for context. White spaces indicate missing data.

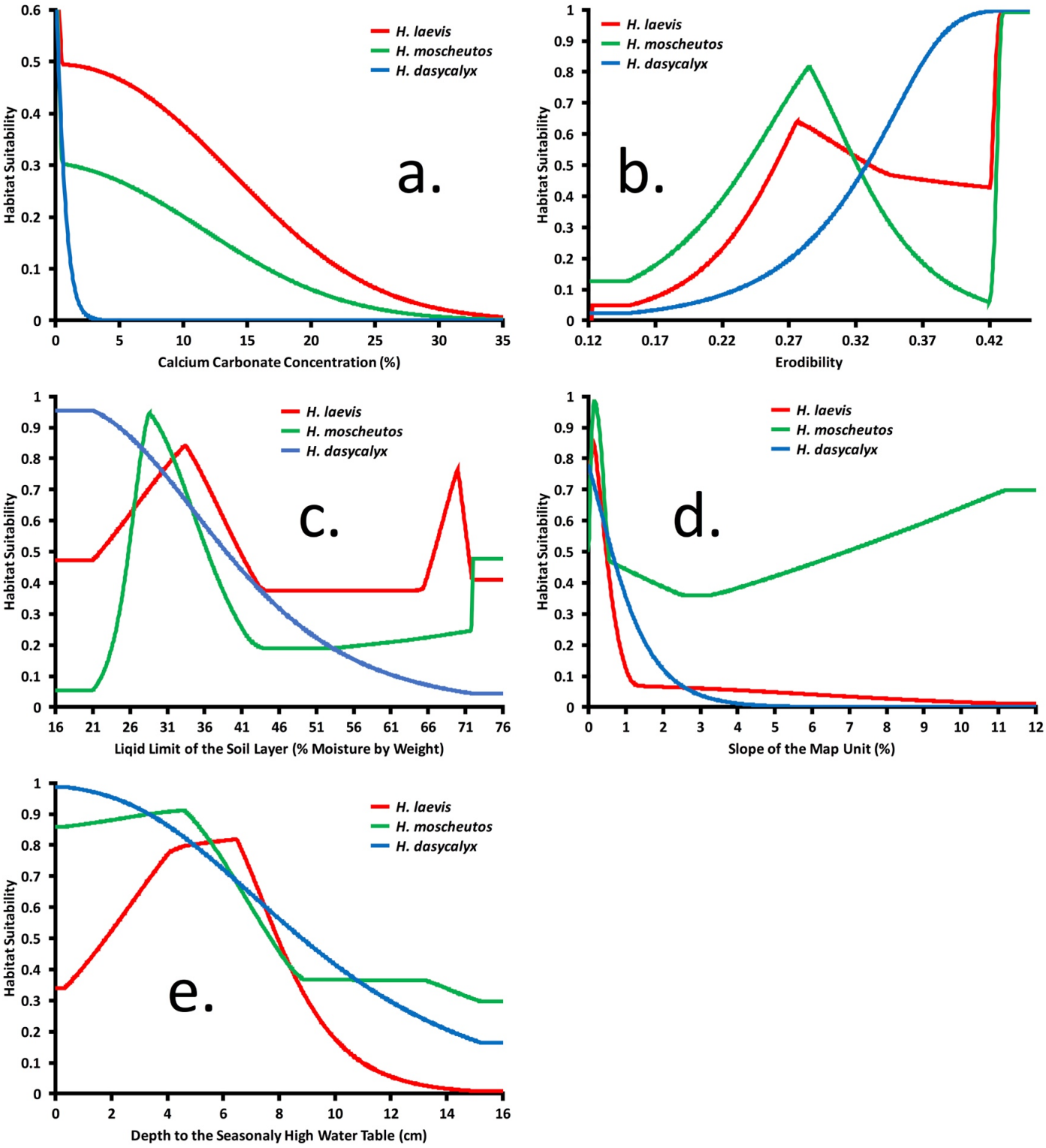


Figure 15. Average habitat suitability the three species in terms of (a) calcium carbonate concentration, (b) erodibility, (c) liquid limit of the soil layer, (d) slope of the map unit, and (e) depth to the seasonally-high water table. The red line is *H. laevis*, the green line is *H. moscheutos*, and the red line is *H. dasycalyx*.