J. Amer. Soc. Hort. Sci. 147(1):62–69. 2022. https://doi.org/10.21273/JASHS05095-21

Genetic Diversity and Population Structure of *Chionanthus virginicus*

Phillip A. Wadl
U.S. Department of Agriculture, Agricultural Research Service, U.S. Vegetable Laboratory, Charleston, SC 29414

Timothy A. Rinehart
U.S. Department of Agriculture, Agricultural Research Service, National Program Leader - Specialty Crops, Beltsville, MD 20705

Richard T. Olsen
U.S. Department of Agriculture, Agricultural Research Service, Floral and Nursery Plants Research Unit, U.S. National Arboretum, Washington, DC 20002

Benjamin D. Waldo
U.S. Department of Agriculture, Agricultural Research Service, U.S. Vegetable Laboratory, Charleston, SC 29414

Joseph H. Kirkbride, Jr.
Department of Botany, National Museum Natural History, Smithsonian Institution, Washington, DC 20013

Additional index words. american fringetree, microsatellites, population genetics, tree genetics, woody ornamental

Abstract. **The genus *Chionanthus*, known as fringetrees, is a member of the olive family (Oleaceae). *Chionanthus virginicus* is an understory tree or shrub with a wide range in forests of the eastern United States and is used as an ornamental tree that is known to be free of insects and disease in the wild. The species is tolerant of a wide range of environmental conditions, and there is interest in developing new cultivars with improved horticultural traits, such as tree form or upright growth habit and superior flowering display that are widely adapted. To identify genepools in the native range of *C. virginicus* for use in breeding programs, the genetic diversity and population structure were assessed for 274 individuals from 12 locations in four states (Florida, Maryland, North Carolina, and Texas) using 26 simple sequence repeats (SSRs). An average of 12.54 alleles/locus were detected, allelic richness averaged 2.80. Genetic differentiation was 0.11, indicating moderate differentiation among subpopulations. Despite the high genetic diversity and low population differentiation, Bayesian clustering analysis identified six genetic groups that match the geographic distribution of collection sites. Analysis of molecular variance indicated that most (82%) of the variation is explained within individuals, and 11% and 7% of the variation is due to differences among individuals within populations and among populations. Analysis of isolation by distance across all samples showed a weak positive relationship between geographic distance and genetic distance. The *C. virginicus* samples analyzed in this study indicate there is sufficient diversity for germplasm collection for use in breeding programs. Given the relatively moderate genetic differentiation, there are not likely to be unique islands of genetic diversity that may be missed when gathering parental materials for a breeding program**

Received for publication 4 Aug. 2021. Accepted for publication 21 Oct. 2021. Published online 23 December 2021.
We are grateful for permission to collect samples in the Serpentine Barrens Conservation Park, Montgomery County, Maryland, Idylwild Wildlife Management Area, Maryland Department of Natural Resources, Maryland, EPA Research Triangle Park Campus, North Carolina, Joe Mountain Wildlife Sanctuary, Humane Society Wildlife Land Trust, North Carolina, and Sam. B. Hayter Estate, Nacogdoches, Texas, and the Florida Department of Environmental Protection, Division of Recreation and Parks (permit #03271110). We thank Drs. Carole Bergmann, Robert Gibbs, Peter Schubert, Christopher Frye, John Moulis, and David L. Creech for granting or assisting in obtaining those permissions, and William Phelan, John Moulis, Andrew Walker, Kevin Conrad, and David L, Creech for their support and assistance during the collections. Mention of trade names or commercial products in this article is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the U.S. Department of Agriculture (USDA). USDA is an equal opportunity employer.
P.A.W. is the corresponding author. E-mail: Phillip.Wadl@usda.gov.

The genus *Chionanthus* is a member of the olive family (Oleaceae) that belongs to the subtribe Oleinae (Wallander and Albert, 2000) and has more than 140 species (Gouvaerts and Green, 2021; Green, 2004) with at least five distinct lineages that are paraphyletic (Dupin et al., 2020; Olofsson et al., 2019). Three species are temperate, deciduous trees and shrubs, *Chionanthus pygmaeus* (pygmy fringetree) and *C. virginicus* (american fringetree) from southeastern North America are a distinct group and the only temperate lineage found in North America, and *C. retusus* from central and southeastern China (Fang et al., 2009; Hong-Wa, 2013; Hong-Wa and Besnard, 2013). The remaining species are tropical, evergreen trees and shrubs, which are pantropical. *C. virginicus* is an understory tree or shrub in forests of southeastern North America. Whereas *C. pygmaeus* is a small shrub endemic to the Lake Wales Ridge in central Florida and occurs in low scrub vegetation over white sand. The North American species can easily be distinguished

from the Chinese species by their inflorescence position. The inflorescences of *C. pygmaeus* and *C. virginicus* are lateral and develop from buds on the previous year's growth, and the inflorescences of *C. retusus* are terminal on new growth.

C. virginicus is dioecious and often used as a woody landscape or ornamental shrub or small tree (5 × 3 m) for its fast growth, fragrant floral display, and because it is known to be free of insects and disease in the wild (Dirr, 2009). Seed dispersal from the immediate area is facilitated by birds and rodents. Several cultivars are commercially available, such as White Knight, Emerald Knight, Spring Fleecing, and Prodigy (Dirr and Warren, 2019). Selected materials represent chance seedlings from the wild. The main goal for breeding *C. virginicus* is improved horticultural traits such as tree form or upright growth habit and superior flowering display. Given the wide range of this species (U.S. Department of Agriculture cold hardiness zones 3 to 9), it has high potential as a native tree to tolerate environmental conditions throughout much of the United States. However, recent reports suggest it may also be a potential host for emerald ash borer, which significantly detracts from its commercial appeal in the landscape and nursery industry (Cipollini, 2015).

Woody plants are perennial and generally outcrossing species, resulting in higher genetic diversity and less differentiation between populations than self-pollinated species such as annuals (Hamrick and Godt, 1996). The mode of pollination, wind vs. insect, and the mode of seed dispersal also contribute to gene flow within and among populations. Given the extensive distribution of *C. virginicus*, presumably due to the migration of environmentally tolerant plants during glacial retreat, we expect to find historical evidence for a large contiguous population with substantial gene flow by migration (Elfers, 1989). Alternatively, high levels of genetic differentiation may indicate inbreeding due to founder effects, isolating barriers such as rivers, or recent habitat fragmentation that could restrict distributions of local seed and/or pollen. We are particularly interested in the relationship between geographic distance and genetic diversity because this species is relatively long lived and has overlapping generations.

SSRs, or microsatellites, are powerful tools for characterizing genetic diversity and population structure, in paternity analyses and hybrid detection, and in genetic mapping (Goncalves-Vidigal and Rubiano, 2011; Merritt et al., 2015). SSRs are based on a highly specific polymerase chain reaction (PCR) that is used to detect variations of repetitive motifs at specific locations within the genome. The power of these markers lies in their high levels of polymorphism, codominant inheritance, multiallelism, and random distribution throughout the genome. In addition, SSRs have been found to be highly transferable between species and the highest in species with long generation times and mixed or outcrossing breeding systems (Barbará et al., 2007). This high level of cross-species transfer of SSRs has been consistent in Oleaceae genera (De La Rosa et al., 2002; Dervishi et al., 2018; Gorrilliot et al., 2021; Lefort et al., 1999; Noakes et al., 2014; Rallo et al., 2003). In this study, we used SSRs from *C. retusus* (Arias et al., 2011) with demonstrated cross-species transfer to *C. virginicus* to characterize the genetic diversity and population structure for the species to identify genepools in the native range of *C. virginicus* for use in breeding programs.

Materials and Methods

DNA collection and isolation. Twelve collection sites were selected based on the number of individual plants in each population and their locations (Table 1, Fig. 1). Maryland populations were used because they are in the northeastern corner of the native range of *C. virginicus*, the Texas population because it

Table 1. Sampling locality information for *Chionanthus virginicus* populations analyzed with 26 simple sequence repeats.

Population	Location	Code	Plants (no.)	Elevation range (m)	Center coordinates	Approximate area (ha)
Serpentine Barrens Conservation Park	Potomac, MD	SERP	38	92–139	lat. 39°4′10.3″N long. 77°14′3.5″W	83.0
Idylwild Wildlife Management Area	Federalsburg, MD	IDYL	33	2–10	lat. 38°44′1.7″N long. 75°45′7.4″W	72.5
EPA Research Triangle Park	Research Triangle, NC	EPA	50	91–102	lat. 35°52′6.9″N long. 78°52′9.8″W	1.3
Joe Mountain	Joe Mountain, NC	JOE	44	435–609	lat. 36°0′58.4″N long. 81°9′24.9″W	3.9
Sam B. Hayter Estate	Nacogdoches, TX	SAM	49	127–137	lat. 31°38′28.8″N long. 94°45′15.6″W	3.8
Beltsville Agricultural Research Center	Beltsville, MD	BARC	12	26	lat. 39°1′25.9″N long. 76°54′5.7″W	2.9
O'Leno State Park	High Springs, FL	OL	6	13–14	lat. 29°55′18.2″N long. 82°34′7.8″W	0.3
Lafayette Blue Springs State Park	Mayo, FL	LAF	8	19–24	lat. 30°7′41.9″N long. 83°13′36.3″W	1.6
San Felasco Hammock Preserve State Park	Gainesville, FL	SF	5	54–57	lat. 29°42′50.5″N long.82°27′38.2″W	0.2
Lake Griffin State Park	Fruitland Park, FL	LG	8	22–24	lat. 28°51′35.2″N long. 81°54′5.4″W	0.1
Colt Creek State Park	Lakeland, FL	CC	11	27	lat. 28°17′44.7″N long. 82°1′46.6″W	0.9
Hillsborough River State Park	Thonotosassa, FL	HILL	10	13–14	lat. 28°8′52.2″N long. 82°13′27.2″W	0.4

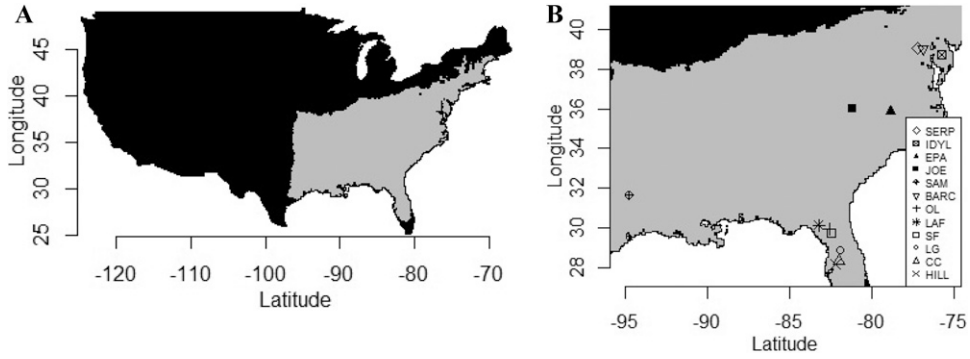

Fig. 1. Range distribution (**A**) and sampling localities (**B**) of *Chionanthus virginicus*. Complete information regarding sampling locations can be found in Table 1.

is in the southwestern corner of the range, and the Florida populations because they are in the southeastern extent of the range. Thus, the extreme limits of the species were sampled. The North Carolina populations were used because they are in the middle of the eastern side of the species distribution and closer to Maryland than to Texas. Two populations were sampled in Maryland and in North Carolina to examine the effect of proximity. Maps were generated in R (version 4.0.4; R Foundation for Statistical Computing, Vienna, Austria) with the "BIEN" package (Maitner, 2020), "maps" package (Becker and Wilks, 2006), and the "sp" package (Pebesma and Bivand, 2005).

In each population, a series of parallel transects were searched for individuals to sample. The distance between transects varied depending on the size of the populations and the area covered. In the Maryland populations, transects were ≈5 m apart, and in the Florida, North Carolina, and Texas populations, which were smaller and much denser, transects were 1 to 2 m apart. Individuals were sampled for DNA extraction by collecting four or five leaves from branch apices, immediately placing them into plastic bags on ice, and shipping them overnight to the laboratory of the second author. Also, two branch ends were collected and prepared as herbarium vouchers for each sample. One of the vouchers for each sample is deposited in the herbarium of the U.S. National Arboretum in Washington, DC. For selected collections, the latitude and longitude were determined using a handheld global positioning satellite unit. The elevation of these data points and the size of each population sampled were determined by drawing a polygon around the collection points (Google Earth Pro; Google, Mountain View, CA) and calculating the area, in hectares, enclosed in the polygon. Sex of the individuals sampled was not recorded.

Simple sequence repeat development and sample processing. Twenty-six SSR loci that were previously indicated as polymorphic in *C. virginicus* were selected based on Arias et al. (2011). DNA was extracted from 1-cm^2 pieces of fresh leaf tissue using a Plant Mini Kit (Qiagen, Valencia, CA), quantified using a spectrophotometer (NanoDrop; Thermo Fisher Scientific, Waltham, MA), and diluted to a final concentration of 5 ng·μL^{-1}. SSR amplification was performed using a three-primer protocol modified from (Waldbieser et al., 2003). Fluorescence-labeled PCR fragments were visualized by automated capillary gel electrophoresis (ABI3730xl, Thermo Fisher Scientific) using ROX-500 size standard (Applied Biosystems, Foster City, CA). GeneMapper version 3.7 (Applied Biosystems) was used to recognize and size peaks.

Data analysis. Allele size data from 26 SSR loci were compiled and GenAlEx 6.5 (Peakall and Smouse, 2012) and Arlequin version 3.5 (Excoffier and Lischer, 2010) were used to estimate genetic diversity and calculate diversity indices, including the mean number of alleles (N_A), effective number of alleles (N_E), observed (H_O) and expected (H_E) heterozygosity, and F statistics across all populations for each locus. The program HP-Rare (Kalinowski, 2005) was used to calculate allelic richness (A_R). Population structure was assessed using a Bayesian analysis [STRUCTURE version 2.3.4 (Pritchard et al., 2000)] using the admixture model, which infers whether the individual i has inherited a portion of its genetic material from ancestors in population K. For measuring different values of K, 20 independent replicates were made for each K value between 1 and 8. A burn-in period of 250,000 iterations and 250,000 Markov Chain Monte Carlo repetitions were used in all analyses. The most likely number of clusters (K) was evaluated considering the plateau criterion proposed by Pritchard et al. (2000) using the nonparametric Wilcoxon test and the Delta K method using STRUCTURE HARVESTER (Earl and von Holdt, 2012).

Differentiation among populations was quantified using a hierarchical analysis of molecular variance (AMOVA) using Arlequin version 3.5. Two different analyses were conducted, the first one included all sites as a single hierarchical group and the second one accounted for populations grouped according to regions identified by the program STRUCTURE. The significances of variance components for each hierarchical comparison (among populations, among individuals, and among individuals within populations) were tested using 99,999 permutations. GenAlEX 6.5 was used for pairwise calculations of genetic differentiation (FST) and gene flow estimates between populations. To determine the occurrence of isolation by distance (IBD), a Mantel test between the genetic and geographic distances was evaluated using GenAlEx 6.5 with 9999 permutations.

Results

Collection sites varied in size and elevation (Table 1). The largest sites were 83 and 72.5 ha for SERP and IDYL, respectively. However, the remaining sites were all less than 4.0 ha each. Both SERP and IDYL are in Maryland and represent populations that are near each other, similar to the close proximity of JOE and EPA that are both located in North Carolina (Table 2). Texas samples (SAM) are most distant from all other sites and represent the extreme southwestern edge of the *C. virginicus* native range (Table 2, Fig. 1). Maryland samples, SERP and IDYL, represent the northeastern extreme. Samples from state parks in Florida represent the most southeastern populations in the native range of *C. virginicus* (Fig. 1). North Carolina samples, JOE and EPA, are intermediate along the eastern edge of the native distribution but are closer to Maryland than to Texas (Fig. 1, Table 2). Table 1 also includes samples not used in the initial analyses, such as BARC, which is in-between the

Table 2. Geographic distances in kilometers (above diagonal) and pairwise population differentiation (FST) values (below diagonal) for *Chionanthus virginicus* populations (see Table 1 for location codes).

	BARC	SERP	SAM	IDYL	EPA	JOE	OL	LAF	SF	LG	CC	HILL
BARC	—	29	1814	105	391	503	1135	1144	1152	1218	1281	1303
SERP	0.11	—	1799	133	385	485	1126	1133	1143	1211	1275	1296
SAM	0.04	0.06	—	1984	1555	1348	1181	1114	1197	1274	1282	1270
IDYL	0.09	0.09	0.03	—	418	564	1161	1175	1175	1234	1295	1318
EPA	0.07	0.07	0.05	0.08	—	207	745	756	761	828	891	913
JOE	0.08	0.10	0.06	0.10	0.02	—	689	681	710	797	860	878
OL	0.08	0.08	0.05	0.09	0.02	0.01	—	68	25	134	188	199
LAF	0.07	0.09	0.05	0.08	0.03	0.02	0.01	—	87	191	234	240
SF	0.09	0.13	0.07	0.11	0.04	0.03	0.03	0.02	—	109	163	175
LG	0.11	0.16	0.10	0.14	0.08	0.07	0.07	0.06	0.05	—	64	85
CC	0.09	0.03	0.06	0.08	0.05	0.07	0.07	0.08	0.09	0.12	—	25
HILL	0.14	0.05	0.08	0.11	0.09	0.10	0.11	0.13	0.16	0.19	0.02	—

two populations in Maryland (SERP and IDYL), and samples from six state parks in Florida.

All SSR loci analyzed were polymorphic, and a total of 312 alleles were detected in 12 *C. virginicus* populations (Table 3). Fifteen of 26 SSR loci did not deviate significantly from Hardy-Weinberg equilibrium (HWE) for the 12 populations (Table 3). The N_A varied considerably between loci, from 3 (seq967 and seq1006) to 32 (seq411), and the average number of alleles/loci detected was 12.54. A_R ranged from 1.46 to 4.39 and averaged 2.80. H_O and H_E also varied considerably between loci, but the average for all loci was 0.49 and 0.55, respectively. This corresponds with F statistics on inbreeding coefficient (FIS) and allele fixation (FIT), which ranged from −0.25 to 0.85, −0.26 to 0.88, and 0.05 to 0.30 for FIS, FIT, and FST, respectively (Table 3). Mean FIS was 0.09 and indicated a slight excess of homozygotes within individuals, and

Table 3. Genetic diversity estimates and F statistics for 26 simple sequence repeats in 12 *Chionanthus virginicus* populations.

GenBank accession	n	N_A	H_O	H_E	HWE	A_R	FIS	FST	FIT
GQ118094	12	3	0.47	0.45	NS	2.09	−0.04	0.08	0.05
GQ118114	12	4	0.31	0.31	NS	1.71	0.01	0.14	0.15
GQ117386	12	7	0.62	0.58	NS	2.69	−0.06	0.15	0.10
GQ117428	12	14	0.50	0.54	NS	2.64	0.06	0.09	0.15
GQ117431	12	20	0.52	0.65	**	3.02	0.19	0.11	0.28
GQ117475	12	19	0.71	0.71	NS	3.55	0.01	0.08	0.09
GQ117545	12	10	0.71	0.70	NS	3.36	−0.02	0.06	0.04
GQ117553	12	29	0.84	0.79	NS	4.02	−0.07	0.09	0.03
GQ117554	12	13	0.57	0.69	***	3.35	0.18	0.10	0.26
GQ117600	12	20	0.59	0.74	NS	3.67	0.20	0.14	0.31
GQ117632	12	32	0.53	0.81	ND	4.39	0.34	0.13	0.43
GQ117667	12	16	0.49	0.73	***	3.57	0.33	0.10	0.40
GQ117694	12	31	0.50	0.78	***	4.01	0.36	0.10	0.42
GQ117701	12	5	0.65	0.52	NS	2.48	−0.25	0.06	−0.18
GQ117761	12	6	0.26	0.24	NS	1.64	−0.06	0.15	0.10
GQ117762	12	15	0.65	0.65	NS	3.28	0.00	0.15	0.14
GQ117328	12	14	0.68	0.74	NS	3.64	0.08	0.11	0.18
GQ117875	12	5	0.18	0.18	ND	1.46	0.01	0.09	0.10
GQ117341	12	8	0.60	0.41	***	2.11	−0.45	0.13	−0.26
GQ117927	12	8	0.08	0.50	ND	2.44	0.85	0.23	0.88
GQ117932	12	9	0.50	0.66	***	3.02	0.24	0.05	0.28
GQ117981	12	6	0.33	0.50	***	2.52	0.33	0.30	0.53
GQ117295	12	5	0.24	0.28	NS	1.80	0.15	0.06	0.20
GQ117358	12	5	0.12	0.11	ND	1.30	−0.09	0.07	−0.02
GQ118064	12	3	0.61	0.58	NS	2.58	−0.06	0.13	0.07
GQ118087	12	5	0.49	0.53	NS	2.36	0.06	0.08	0.13
Calculated means		12.54	0.49	0.55		2.80	0.09	0.11	0.19

n = number of populations; N_A = number of alleles; H_O = observed heterozygosity; H_E = expected heterozygosity; HWE = significance of deviation from Hardy-Weinberg equilibrium (NS = not significant, * = significant at $P \leq 0.05$, ** = significant at $P \leq 0.01$, *** = significant at $P \leq 0.001$, ND = not done); A_R = allelic richness; FIS = inbreeding coefficient of individuals relative to the population; FST = variance among subpopulations relative to the total variance; FIT = variance in the total population.

FST was 0.11, indicating moderate differentiation among subpopulations.

Despite substantial differences in sample and N_A for each population, A_R, H_O, and H_E are comparable between sites (Table 4). N_A was highest in the EPA (195) population and lowest in the SF (81) population. A_R ranged from 2.55 (HILL) to 3.08 (EPA) and averaged 2.80. Total H_O (0.49) was similar to H_E (0.55) and ranged from 0.38 (SERP) to 0.55 (SF), whereas H_E ranged from 0.49 (HILL, LG, and SF) to 0.63 (EPA). FIS ranged from −0.12 (SF) to 0.24 (SERP). Seventy private alleles were detected in the 12 sampling sites and ranged from 1 (HILL, JOE, and SERP) to 18 (LG). Private alleles were found at 19 loci, with seq326 (10) and seq486 (10) having the most, and seq9, seq711, seq793, and seq1032 having one each. Fifty-one private alleles occurred at a frequency of <0.05, 13 at a frequency between 0.05 and 0.09, and eight at a frequency greater than 0.10.

The lowest FST (0.037) was found between EPA and JOE populations, which are located 207 km apart in North Carolina (Table 2). The only two sites closer to each other are IDYL and SERP in Maryland, which are 133 km apart (Table 2) and have slightly higher FST at 0.055, which is greater than most Maryland and North Carolina sites (Table 2). The highest FST value was found between SAM and EPA populations, which are separated by 1555 km, but this value was less than 10% divergence using Nei's minimum genetic distance (Table 2). Pairwise genetic divergence between SAM and the other four sites is greater than values for comparisons between the other sites, but all values indicate less than 10% divergence (Table 2).

STRUCTURE analysis found evidence for six distinct groups as indicated by the maximum for Delta K at K = 6 (Fig. 2). The mean log-likelihood curve attained a maximum value ≈K = 6, beyond which the mean log-likelihood values reached a plateau and the standard deviations associated with the estimates increased. Cluster 1 contained individuals from two Maryland locations, BARC and SERP (Fig. 3). Cluster 2 consisted of the individuals from Texas. Cluster 3 is composed of individuals sampled at IDYL in Maryland. The samples from the EPA in North Carolina were assigned to cluster 4, and the individuals from JOE in North Carolina were assigned to cluster 5. All individuals from the five Florida sampling locations were assigned to cluster 6.

Table 4. Genetic variability estimates for 12 *Chionanthus virginicus* populations based on genotyping with 26 simple sequence repeats (see Table 1 for location codes).

Code	Plants (no.)	A	N_A	A_R	H_O	H_E	FIS	P_A
BARC	12	121	4.65	2.95	0.53	0.59	0.09	3
SERP	38	144	5.54	2.60	0.38	0.51	0.24	1
SAM	49	184	7.08	2.91	0.49	0.59	0.14	7
IDYL	33	147	5.65	2.73	0.52	0.57	0.07	6
EPA	51	195	7.50	3.08	0.53	0.63	0.13	2
JOE	44	193	7.42	2.99	0.50	0.60	0.14	1
OL	6	97	3.73	2.93	0.53	0.56	0.04	2
LAF	8	104	4.00	2.86	0.53	0.57	0.05	8
SF	5	81	3.12	2.62	0.55	0.49	−0.12	6
LG	8	92	3.54	2.58	0.46	0.49	0.06	18
CC	11	108	4.15	2.74	0.44	0.55	0.15	15
HILL	10	95	3.65	2.55	0.43	0.49	0.08	1
Mean	22.91	130.08	5.00	2.80	0.49	0.55	0.09	5.83

A = number of alleles; N_A = mean number of alleles; A_R = allelic richness; H_O = observed heterozygosity; H_E = expected heterozygosity; FIS = inbreeding coefficient; P_A = private alleles.

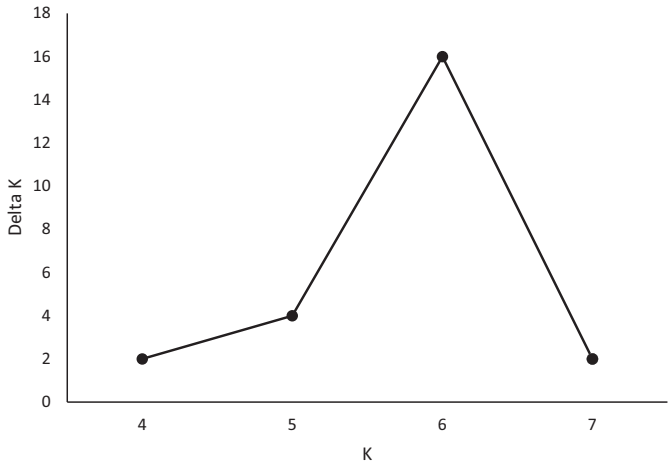

Fig. 2. Plot of Delta K vs. K produced from STRUCTURE HARVESTER (Earl and von Holdt, 2012) using STRUCTURE (Pritchard et al., 2000) results. Delta K indicates that the most likely K is where the largest change in magnitude of the second-order rate of change in ln Pr(X|K) against successive K values occurs. K is the assumed number of populations or genetic groups.

AMOVA grouping of all populations together indicated that most (82%) of the variation is explained within individuals, and 11% and 7% of the variation is due to differences among individuals within populations and among populations, respectively (Table 5). Additional AMOVA was performed by grouping populations by the results of the STRUCTURE analysis and found similar results to the population structure when all populations were analyzed as a single hierarchical group (Table 5). Regardless of AMOVA hierarchical grouping, similar F statistic values were estimated (Table 5). Regarding pairwise FST, the estimated values obtained ranged from 0.01 (OL and JOE; OL and LAF) to 0.19 (HILL and LG). The genetic divergence among subpopulations is reflected in the population pairwise FST values (Table 2). Individual-based analysis of IBD across all samples showed a positive relationship between geographic distance (kilometers) and genetic distance (Fig. 4).

Discussion

Genetic diversity within and among *C. virginicus* was assessed with 26 SSR loci, providing a robust analysis for the species. This is the first report of SSR diversity from wild-collected *C. virginicus* in the United States and the first estimates of population structure of this native species. The SSR markers we used were highly polymorphic and repeat type was previously described by Arias et al. (2011). Data generated had clean profiles with a low incidence of null alleles, providing high confidence in identifying private alleles. The distribution of SSR loci within the genome is not known, but we used a relatively large number of SSRs, 26 markers, for more robust estimates of gene diversity.

Alleles per locus values were high (average of 12.54, with a range of 3 to 32), suggesting adequate support for genetic diversity estimates. Private alleles, alleles not observed in other samples, were most frequent (over 70%) in the Florida sampling locations, whereas only 4.3%, 10.0%, and 14.3% were identified in the North Carolina, Maryland, and Texas locations, respectively.

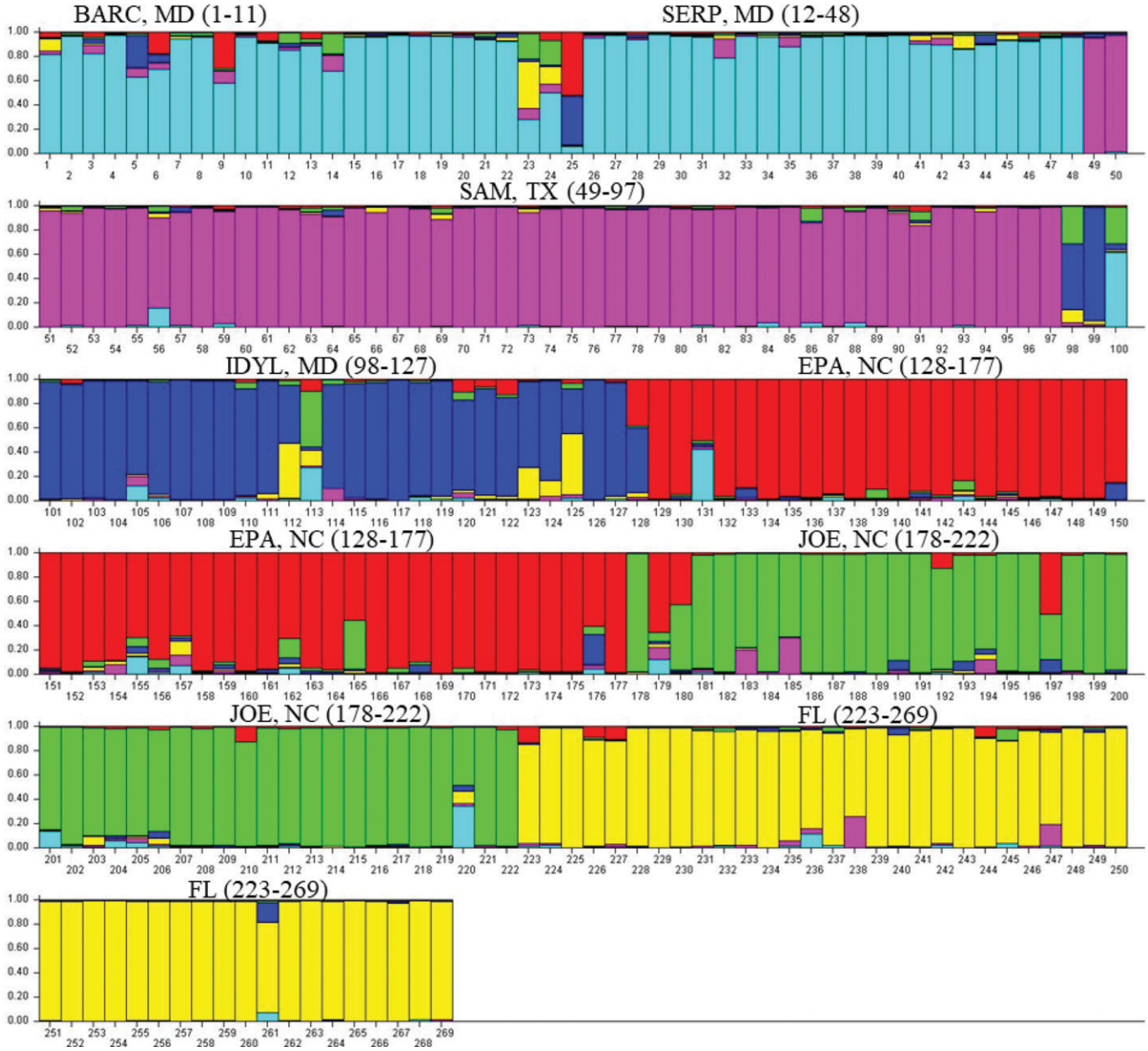

Fig. 3. Bar plot of individual Bayesian assignment probabilities for *Chionanthus virginicus* populations using the program STRUCTURE (Pritchard et al., 2000) for six genetic clusters. Each vertical line represents an individual's probability of belonging to one of six clusters [represented by different colors (K = number of populations or genetic groups)] or a combination if ancestry is mixed. Complete information regarding sampling locations can be found in Table 1.

H_O ranged from 0.12 to 0.84, suggesting high levels of population genetic diversity, and H_O values were only slightly less than the H_E values. This moderate indication of inbreeding was supported by FIS values, which were relatively low but significant at eight loci. The high genetic diversity reported here could be maintained by large population sizes, environmental heterogeneity across the native distribution, and life history traits such as overlapping generations.

Population differentiation was low to moderate (most pairwise FST values were less than 0.1), indicating weak differentiation among the geographic populations. High genetic diversity and low differentiation are often associated with outcrossing woody plants (Hamrick and Godt, 1996). AMOVA indicated that most of the genetic variation (82%) is within individuals rather than among the other two hierarchies, viz., among groups and individuals within groups. This also could be related to outcrossing and the dispersal of pollen and seed (Dirr, 2009; Nicholson, 1990; Ueda, 1996). Sufficient gene flow between populations would limit divergence of populations. Bee-mediated pollination is known to travel long distances and robust floral display make *C. virginicus* an ideal plant for widespread bee-mediated pollination. Similarly, fruits are attractive to birds and persistent on the tree, making birds and rodents ideal vectors for long-range seed dispersal (Stiles, 1980). Migration, whether ongoing or historical, likely plays a strong role in keeping population differentiation moderate and genetic diversity high. Forest succession may also help retain genetic variability with fast-growing *C. virginicus* shrubs occupying middle to late successional stages, especially on granite outcrops.

Table 5. Analysis of molecular variance for 12 populations of *Chionanthus virginicus* using Arlequin 3.5 (Excoffier and Lischer, 2010).

Variance partition	df	Sum of squares	Variance component	Variation (% of total)	*P*
A. One hierarchical group					
Among groups	11	266.548	0.41228	6.72	<0.0001
Among individuals within groups	262	1681.414	0.69603	11.35	<0.0001
Within individuals	274	1377.00	5.02555	81.93	<0.0001
Total	547	3324.962	6.13385	100	
Fixation indices: FIS = 0.12, FST = 0.07, FIT = 0.18					
B. Six hierarchical groups					
Among groups	5	165.893	0.29165	4.76	<0.0001
Among individuals within groups	268	1782.068	0.81198	13.25	<0.0001
Within individuals	274	1377.000	5.02555	81.99	<0.0001
Total	547	1972.394	4.438	100	
Fixation indices: FIS = 0.14, FST = 0.05, FIT = 0.18					

A = the first analysis included all sampling sites as one hierarchical group; B = the second hierarchical analysis accounted for sampling locations grouped according to the clusters identified by STRUCTURE; FIS = inbreeding coefficient of individuals relative to the population; FST = variance among subpopulations relative to the total variance; FIT = variance in the total population.

All locations of *C. virginicus* sampled here, including the BARC and FL samples, could be considered a single genetically homogeneous population based solely on the FST values. Although not sampled as extensively as other sites, the Florida and BARC samples provide additional sites for testing hypotheses about geographic relationships. The inclusion of Florida samples provides information on the most southeastern range of *C. virginicus*, and these accessions appear more related to Texas than more northern populations. BARC samples lie between the two more extensively sampled Maryland sites (SERP and IDYL), providing extra data to test the effects of proximity, because preliminary analyses of SERP and IDYL did not show clustering.

Despite the high genetic diversity and moderate population differentiation, Bayesian clustering identified six genetic groups that match the geographic distribution of collection sites. Florida samples cluster with each other, as did BARC and SERP. All other collection sites clustered separately. Given the high heterozygosity, clustering is likely influenced by relatively rare alleles that are unique to each population. If we assume that the lack of differentiation is due to historically large population sizes, gene migration, and range expansion during glacial retreat, then the clustering of geographic sites and minor inbreeding observed here are probably due to recent fragmentation. For example, the Mississippi River could be a barrier to gene flow into and out of Texas, resulting in the geographic cluster with slightly lower genetic diversity. Presumably the relatedness between Florida and Texas is due to an ancestral, contiguous southern population that migrated north. Land development, loss of seed dispersal vectors such as birds and rodents, and even a decline in pollinators may play a role in recent isolation of geographic regions that are identified here. To determine the role of these factors, plastid markers (e.g., chloroplast SSRs) could be used to contrast patterns of genetic diversity with biparentally or maternally inherited markers to assess the relative contribution of seeds and pollen in gene flow across the species range.

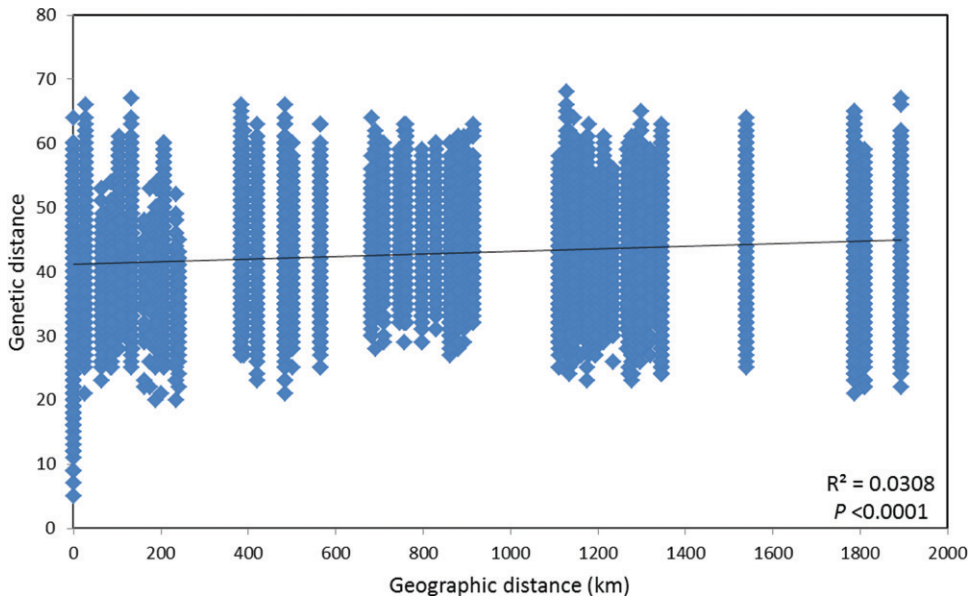

Fig. 4. Isolation by distance of *Chionanthus virginicus* populations. Correlation between pairwise genetic distance values and geographic distance (kilometers) as calculated using a Mantel test with 9999 permutations in GenAlEx 6.5 (Peakall and Smouse, 2012).

Although the large historical population of *C. virginicus* may be fragmented now, adequate gene flow appears to be intact, and accessions retain most of the genetic variability. There was a positive correlation between genetic differentiation and geographic distance (Fig. 4). If we invoke the IBD model, which would help explain the lower genetic diversity in Texas and Florida, germplasm collection efforts should be based on relatively few collections from across the broadest range to capture the most alleles possible. Ample genetic diversity is a requirement for plant breeders, and our study is an important first step in assessing the genetic diversity and population structure for *C. virginicus*. There is limited germplasm available for breeding purposes as there are only eight cultivars (Dirr and Warren, 2019) and eight accessions in the U.S. National Plant Germplasm System (U.S. Department of Agriculture, Agricultural Research

Service, 2021). To conclude, the *C. virginicus* populations characterized in this study contain sufficient genetic diversity to provide germplasm sources for ornamental breeding purposes. Given the moderate genetic differentiation, there are not likely to be unique islands of genetic diversity that may be missed when gathering parental materials for a breeding program. High genetic diversity and obligate outcrossing suggest parent selection should be based exclusively on ornamental traits that are currently lacking in existing cultivars. Improvement of current forms should be straightforward and occur within a few generations under rigorous selection. Last, any plant materials that are collected from the respective genetic clusters should first be phenotyped for the desired trait before being incorporated into a breeding program for cultivar development.

Literature Cited

Arias, R., N. Techen, T.A. Rinehart, R. Olsen, J. Kirkbride, and B. Scheffler. 2011. Development of SSR markers for *Chionanthus retusus* (Oleaceae) and cross amplification of closely related taxa. HortScience 46:23–29, https://doi.org/10.21273/HORTSCI.46.1.23.

Barbará, T., C. Palma-Silva, G.M. Paggi, F. Bered, M.F. Fay, and C. Lexer. 2007. Cross-species transfer of nuclear microsatellite markers: Potential and limitations. Mol. Ecol. 16:3759–3767, https://doi.org/10.1111/j.1365-294X.2007.03439.x.

Becker, R.A. and A.R. Wilks. 2016. Maps: Draw geographical maps. 21 July 2021. <https://cran.r-project.org/web/packages/maps/index.html>.

Cipollini, D. 2015. White fringetree as a novel larval host for emerald ash borer. J. Econ. Entomol. 108:370–375, https://doi.org/10.1093/jee/tou026.

De La Rosa, R., C.M. James, and K.R. Tobutt. 2002. Isolation and characterization of polymorphic microsatellites in olive (*Olea europaea* L.) and their transferability to other genera in the Oleaceae. Mol. Ecol. Notes 2:265–267, https://doi.org/10.1046/j.1471-8286.2002.00217.x.

Dervishi, A., J. Jakše, H. Ismaili, B. Javornik, and N. Štajner. 2018. Comparative assessment of genetic diversity in albanian olive (*Olea europaea* L.) using SSRs from anonymous and transcribed genomic regions. Tree Genet. Genomes 14:53, https://doi.org/10.1007/s11295-018-1269-6.

Dirr, M.A. 2009. Manual of woody landscape plants: Their identification, ornamental characteristics, culture propagation, and uses. Stipes Publ., Champaign, IL.

Dirr, M.A. and K.S. Warren. 2019. The tree book: Superior selections for landscapes, streetscapes, and gardens. Timber Press, Portland, OR.

Dupin, J., P. Raimondeau, C. Hong-Wa, S. Manzi, M. Gaudeul, and G. Besnard. 2020. Resolving the phylogeny of the olive family (Oleaceae): Confronting information from organellar and nuclear genomes. Genes (Basel) 11:1508, https://doi.org/10.3390/genes11121508.

Earl, D.A. and B.M. von Holdt. 2012. STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv. Genet. Resour. 4:359–361, https://doi.org/10.1007/s12686-011-9548-7.

Elfers, S.C. 1989. A biosystematic study of *Chionanthus* in the southeastern United States. Univ. Central Florida, Orlando, MS Thesis.

Excoffier, L. and H.E. Lischer. 2010. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. Mol. Ecol. Resour. 10:564–567, https://doi.org/10.1111/j.1755-0998.2010.02847.x.

Fang, J., Z. Wang, and Z. Tang. 2009. Atlas of woody plants in China: Distribution and climate. Vol. 1. Higher Educ. Press, Beijing, China.

Goncalves-Vidigal, M.C. and L.B. Rubiano. 2011. Development and application of microsatellites in plant breeding. Crop Breed. Appl. Biotechnol. 11:66–72.

Gorrilliot, O., C. Hong-Wa, F. Rakotonasolo, and G. Besnard. 2021. Microsatellite-assisted identification and comparative population genetics of malagasy olive species (*Noronhia* spp., Oleaceae). Bot. Lett., https://doi.org/10.1080/23818107.2021.1912636.

Gouvaerts, R. and P.S. Green. 2021. World checklist of Oleaceae. 8 Apr. 2021. <http://wcsp.science.kew.org/>.

Green, P.S. 2004. Oleaceae, p. 296–306. In: K. Kubitzki (ed.). The families and genera of vascular plants. Vol. 7. Springer-Verlag, Berlin, Germany.

Hamrick, J.L. and M.J.W. Godt. 1996. Effects of life history on genetic diversity in plant species. Phil. Trans. Biol. Sci. 351:1291–1298, https://doi.org/10.1098/rstb.1996.0112.

Hong-Wa, C. 2013. Diversification and coexistence in the madagascar olive (*Noronhia*, Oleaceae). Univ. Missouri, St. Louis, PhD Diss. 320.

Hong-Wa, C. and G. Besnard. 2013. Intricate patterns of phylogentic relationships in the olive family as inferred from multi-locus plastid and nuclear DNA sequences analyses: A close-up on *Chionanthus* and *Noronhia* (Oleaceae). Mol. Phylogenet. Evol. 67:367–378, https://doi.org/10.1016/j.ympev.2013.02.003.

Kalinowski, S.T. 2005. HP-RARE 1.0: A computer program for performance rarefaction on measures of allelic richness. Mol. Ecol. Notes 5:187–198, https://doi.org/10.1111/j.1471-8286.2004.00845.x.

Lefort, F., S. Brachet, N. Franscaria-Lacoste, K.J. Edwards, and G.C. Douglas. 1999. Identification and characterization of microsatellite loci in ash (*Fraxinus excelsior* L.) and their conservation in the olive family (Oleaceae). Mol. Ecol. 8:1088–1089, https://doi.org/10.1046/j.1365-294X.1999.00655_8.x.

Merritt, B.J., T.M. Culley, A. Avanesyan, R. Stokes, and J. Brzyski. 2015. An empirical review: Characteristics of plant microsatellite markers that confer higher levels of genetic variation. Appl. Plant Sci. 3:1500025, https://doi.org/10.3732/apps.1500025.

Maitner, B. 2020. BIEN: Tools for accessing the Botanical Information and Ecology Network database. R package version 1.2.4. 21 July 2021. <https://CRAN.R-project.org/package=BIEN>.

Nicholson, R.G. 1990. The fringe tree and its far-flung cousins. Arnoldia 50:24–31.

Noakes, A.G., T. Best, M.E. Staton, J. Koch, and J. Romero-Severson. 2014. Cross amplification of 15 EST-SSR markers in the genus *Fraxinus*. Conserv. Genet. Resour. 6:969–970, https://doi.org/10.1007/s12686-014-0260-2.

Olofsson, J.K., I. Cantera, C. Van de Paer, C. Hong-Wa, L. Zedane, L.T. Dunning, A. Alberti, P.-A. Christin, and G. Besnard. 2019. Phylogenomics using low-depth whole genome sequencing: A case study with the olive tribe. Mol. Ecol. Resour. 19:877–892, https://doi.org/10.1111/1755-0998.13016.

Peakall, R. and P.E. Smouse. 2012. GenAlEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research – An update. Bioinformatics 28:2537–2539, https://doi.org/10.1093/bioinformatics/bts460.

Pebesma, E.J. and R.S. Bivand. 2005. Classes and methods for spatial data in R. R News 5:9–13, https://cran.r-project.org/doc/Rnews/.

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

Rallo, P., I. Tenzer, C. Gessler, L. Baldoni, G. Dorado, and A. Martin. 2003. Transferability of olive microsatellite loci across the genus *Olea*. Theor. Appl. Genet. 107:940–946, https://doi.org/10.1007/s00122-003-1332y.

Stiles, E.W. 1980. Patterns of fruit presentation and seed dispersal in bird-disseminated woody plants in the eastern deciduous forest. Amer. Nat. 116:670–688.

Ueda, K. 1996. Androdioecism in *Chionanthus retusus* (Oleaceae). J. Phytotax. Taxon. 44:91–92, https://doi.org/10.2307/122684.

U.S. Department of Agriculture, Agricultural Research Service. 2021. U.S. National Plant Germplasm System, *Chionanthus virginicus*. 21 Oct. 2021. <https://npgsweb.ars-grin.gov/gringlobal/search?t=>.

Waldbieser, G.C., A.L. Bilodeau, and D.J. Nonneman. 2003. Complete sequence and characterization of the channel catfish mitochondrial genome. DNA Seq. 14:265–277, https://doi.org/10.1080/10425170310001149057.

Wallander, E. and V.A. Albert. 2000. Phylogeny and classification of Oleaceae based on rps16 and trnL-F sequence data. Amer. J. Bot. 87:1827–1841, https://doi.org/10.2307/2656836.