

# The role of anthropogenic dispersal in shaping the distribution and genetic composition of a widespread North American tree species

Graham E. Wyatt | J. L. Hamrick | Dorset W. Trapnell 

Department of Plant Biology, University of Georgia, Athens, Georgia, USA

## Correspondence

Dorset W. Trapnell, Department of Plant Biology, University of Georgia, Athens, GA 30602, USA.

Email: dorset@uga.edu

## Funding information

The University of Georgia Office of the Vice President for Research award to DWT

## Abstract

Dispersal and colonization are among the most important ecological processes for species persistence as they allow species to track changing environmental conditions. During the last glacial maximum (LGM), many cold-intolerant Northern Hemisphere plants retreated to southern glacial refugia. During subsequent warming periods, these species expanded their ranges northward. Interestingly, some tree species with limited seed dispersal migrated considerable distances after the LGM ~19,000 years before present (YBP). It has been hypothesized that indigenous peoples may have dispersed valued species, in some cases beyond the southern limits of the Laurentide Ice Sheet. To investigate this question, we employed a molecular genetics approach on a widespread North American understory tree species whose fruit was valued by indigenous peoples. Twenty putative anthropogenic (near pre-Columbian habitations) and 62 wild populations of *Asimina triloba* (pawpaw), which produces the largest edible fruit of any North American tree, were genetically assayed with nine microsatellite loci. Putative anthropogenic populations were characterized by reduced genetic diversity and greater excess heterozygosity relative to wild populations. Anthropogenic populations in regions that were glaciated during the LGM had profiles consistent with founder effects and reduced gene flow, and shared rare alleles with wild populations hundreds of kilometers away (mean = 723 km). Some of the most compelling evidence for human-mediated dispersal is that putative anthropogenic and wild populations sharing rare alleles were separated by significantly greater distances (mean = 695 km) than wild populations sharing rare alleles (mean = 607 km;  $p = .014$ ). Collectively, the genetic data suggest that long-distance dispersal played an important role in the distribution of pawpaw and is consistent with the hypothesized role of indigenous peoples.

## KEYWORDS

*Asimina triloba*, clonality, genetic structure, pawpaw, population genetics, pre-Columbian peoples

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. *Ecology and Evolution* published by John Wiley & Sons Ltd.

## 1 | INTRODUCTION

Dispersal and colonization are critical ecological processes that affect the ability of organisms to undergo range shifts as they track a changing environment and thus to persist over evolutionary time scales. During ice age glacier advances, most cold-intolerant Northern Hemisphere plant species retreated to southern refugia and during subsequent warming periods there was extensive northward migration and often range expansion (Davis & Shaw, 2001; Hewitt, 2000; Lyford et al., 2003). Understanding historical patterns of plant dispersal and mechanisms responsible for movement during the warming climate that ensued after the last glacial maximum (LGM; Clark et al., 2009) allows inferences and predictions regarding their capacity for range modification in response to future climatic shifts.

Sessile plants disperse their seeds by abiotic means (i.e., barochory, hydrochory, or anemochory) and/or biotic vectors (epi- and endozoochory) with some species having more than one dispersal mechanism. Interestingly, some taxa with limited seed dispersal, such as large-seeded tree species without an obvious dispersal vector, appear to have migrated considerable distances after glacial retreat in the late Pleistocene (~19,000 years before present [YBP]), a phenomenon known as "Reid's Paradox" (Clark et al., 1998; MacDougall, 2003). In 1899, Clement Reid was the first of many to note the surprisingly rapid rate of northward migration from glacial refugia by large-seeded trees, such as oaks, into formerly glaciated regions of Great Britain (Clark et al., 1998). Some have hypothesized that humans may have mediated dispersal of valued tree species beyond the limits of their southern glacial refugia (Abrams & Nowacki, 2008; Delcourt & Delcourt, 2004; MacDougall, 2003).

There is considerable evidence that pre-Columbian peoples had a marked impact on the North American landscape prior to European arrival (Abrams & Nowacki, 2008; Denevan, 1992; Pyne, 1983; Stewart, 2002). However, little attention has been paid to the possible role of native peoples in the dispersal of useful plant species (but see Abrams & Nowacki, 2008; Cottrell et al., 2002; Delcourt & Delcourt, 2004; MacDougall, 2003; Petit et al., 2002) and some population genetic studies have intentionally avoided species valued or impacted by humans (Petit et al., 2003). One of the few exceptions involving a North American tree species used genetic markers and packrat middens (pollen and macrofossils) to investigate an isolated (>200 km) northern population of *Pinus edulis* (Colorado pinyon pine; Bentancourt et al., 1991). The authors concluded that this isolated population originated by long-distance dispersal by Native Americans who valued pine nuts as a food source. Additional phytogeographic evidence of human-mediated dispersal in North America comes from tobacco (*Nicotiana* sp.) and bottle gourd (*Lagenaria siceraria*; Asch, 1995). Both species are thought to have been absent from the area presently known as Illinois (U.S.A.) until their introduction by external contacts and subsequent cultivation as useful specialty-plant species. Wykoff

(1991) asserted that black walnut (*Juglans nigra*), some hickories (*Carya* spp.), oaks (*Quercus* spp.), and medicinal plant species were likely introduced into what is now recognized as New York state (U.S.A.) prior to European contact.

At least 15,000 YBP an ice-free corridor connected Alaska and the rest of the continental United States, permitting human migration into this region. Earlier arrival of humans in North America ~16,000 YBP could also have been possible along the North Pacific coast (Erlandson et al., 2015). It is generally agreed that indigenous peoples inhabiting North America manipulated their environment to manage resources soon after their arrival ~15,000 (Delcourt & Delcourt, 2004; Morse & Morse, 1983) to 17,000 YBP (Adovasio et al., 1978; Richland et al., 2007); however, debate remains about the uniformity of their influence over space and time (Muñoz et al., 2014).

While Denevan (2011) estimated two million indigenous people lived in North America in 1492, when Europeans commenced colonization, there is much controversy over the extent of anthropogenic impact on the present distribution of plant species. Some investigators maintain that impacts by indigenous peoples were ubiquitous, with few ecosystems unaffected (Abrams & Nowacki, 2008; Denevan, 1992; Kay, 2002; Krech, 1999). However, after the introduction of Old World diseases by Spanish explorers and the resulting decimation of indigenous populations (Lovell, 1992), North American ecosystems recovered somewhat, which may have obscured indigenous imprints on the landscape (Denevan, 1992). Others have argued that landscape alteration near densely settled areas showed greater impacts, while more remote, sparsely inhabited areas exhibited little anthropogenic alteration (Muñoz et al., 2014; Parker, 2002; Vale, 1998, 2002). Thus, plants used by indigenous people may show signs of manipulation in proximity to settlements but may exist in an un-manipulated state farther away (e.g., Parker et al., 2014). Muñoz et al. (2014) contended that agricultural and silvicultural impacts were greatest near settlements, as well as along riparian corridors and along trade routes. By approximately 2,000 YBP, indigenous societies throughout the river valleys of Eastern North America were interacting widely, erecting monuments, and producing sophisticated material cultures (Wright, 2017). Various land-use patterns exhibited by indigenous peoples include the following: (a) semi-nomadic hunter-gatherers with temporary encampments, (b) seasonal movements and periodic occupation of resource-specific sites with hunting grounds, nut trees, fish, and/or shellfish, and (c) stable villages associated with agricultural fields or aquatic food resources (Abrams & Nowacki, 2008). Highlands and Piedmont Lowlands were characterized by temporary seasonal encampments (Pagoulatos, 1992). More stable villages were located along major rivers and tributaries of Ridge and Valley areas and the Coastal Plain, while foraging destinations were scattered across the landscape. Hunter-gatherer groups were more typical in the north, while more agricultural-based subsistence was practiced from Massachusetts southward, especially along major river valleys (Patterson & Sassaman, 1988). By 2,000 YBP, large

agricultural fields associated with settlements were established in the floodplains of larger rivers in the Midwest and the south (Fritz, 1990; Scarry & Scarry, 2005). Some cultures relied heavily on nutritious nuts (i.e., balanocultures) of mast-producing species such as chestnut, hickory, and oaks and are thought to have planted orchards and manipulated the distribution of some tree species (Abrams & Nowacki, 2008). Acorns served as a particularly important dietary staple for many indigenous peoples (Bainbridge, 1985; Logan, 2005), and widespread acorn use almost certainly predates widespread corn use (Bainbridge, 1985). There were many other crops that were cultivated as well (e.g., Delcourt & Delcourt, 2004; Mueller, 2017; Schroeder, 1999). Inadvertent or intentional anthropogenic seed dispersal, particularly in heavily impacted portions of the landscape, may have contributed to postglacial maximum range expansion (e.g., Abrams & Nowacki, 2008; Keener & Kuhns, 1998; White, 1906; Wykoff, 1991) and help explain Reid's Paradox. It is important that this question be investigated in more depth in order to develop a better understanding of plant dispersal and factors that have shaped the distribution of plant taxa.

*Asimina triloba* (Annonaceae), whose common name is pawpaw or Indian banana, has the largest (5–16 cm long, 3–7 cm wide) edible fruits of any native North American tree. The flavorful fruits (described as a cross between banana, mango, and pineapple) were prized as a food source by native peoples (e.g., Brooks & Johannes, 1990; Keener & Kuhns, 1998; Peterson, 1991), and the fibrous bark was used to make rope and cloth (Peattie, 1950). It may also have been utilized medicinally (Krochmal & Krochmal, 1973; Peterson, 1991). In recent years, there has been considerable research on promising anti-cancer properties of *A. triloba* seeds and bark (e.g., Zhao et al., 1993).

Seeds or carbonized remains of *A. triloba* have been found along the Mississippi and Ohio Rivers and at numerous archaeological sites in Arkansas, Illinois, Indiana, Kansas, Kentucky, Mississippi, and Missouri (Blasing, 1986; Cutler & Blake, 1976; Gilmore, 1931; Jones, 1936; King, 1982; Waselkov, 1977; Wedel, 1943; Yarnell & Watson, 1966) and may have been dispersed by various cultures (Brooks & Johannes, 1990; Keener & Kuhns, 1998; Peterson, 1991; Wykoff, 1991). The first written record of *A. triloba* comes from Hernando de Soto's expedition in 1541 across Southeastern North America, which reported widespread planting of the tree by indigenous tribes of the southeast and mentions its flavor and fragrance (Pickering, 1879). Some authors have asserted that *A. triloba* was grown by pre-Columbian people (Cai et al., 2019; Hormaza, 2004; Peattie, 1950; Peterson, 1991), although there is no known evidence for the cultivation of *A. triloba* orchards by Native Americans beyond such written reports. There is however molecular evidence for pre-Columbian dispersal of *Annona cherimola* between Central and South America (Larranaga et al., 2017), a species that shares some attributes of *A. triloba* (i.e., a woody perennial that produces large, edible, and nutritious fruits).

It has been hypothesized that the remarkably wide distribution of pawpaw throughout much of eastern North America relative

to its seven North American congeners, which are restricted to the southeastern United States, is due in part to its spread by indigenous peoples who may have actively traded propagules, encouraged its growth (e.g., Abrams & Nowacki, 2008; Keener & Kuhns, 1998; MacDougall, 2003), and/or inadvertently dispersed its seeds (White, 1906). White (1906) for example observed Native Americans using pawpaw seeds as game pieces, contending that this game is centuries old and that Sac and Fox Indians may have unintentionally dispersed the seeds. That the fruit cannot easily be stored or transported suggests the advantages of planting pawpaw seeds along well-traveled routes and/or near settlements. Its occurrence north of the southern Pleistocene glacial boundary suggests that its post-LGM dispersal could have been human-mediated. While pawpaw was clearly used by indigenous peoples in North America, there has been debate regarding the relative importance of natural versus anthropogenic dispersal (Keener & Kuhns, 1998; Murphy, 2001) with more authors favoring an anthropogenic role (e.g., Abrams & Nowacki, 2008; Keener & Kuhns, 1998; MacDougall, 2003).

The overarching goal of our research is to employ a molecular genetics approach to test the hypothesis that indigenous pre-Columbian people contributed to the dispersal and range expansion of a valued North American understory tree. To address this question, we used genetic signatures contained within contemporary *A. triloba* populations of wild and putative anthropogenic origin to assess levels of genetic diversity and test whether patterns of genetic variation within and among populations across the species' range are consistent with anthropogenic dispersal. Specifically, we (a) estimate and compare levels of genetic diversity of wild populations and those of putative anthropogenic origin (i.e., populations occurring near known Native American settlements), and (b) test whether patterns of genetic structure among populations are consistent with the hypothesis of anthropogenic dispersal, particularly into regions glaciated during the LGM. Our a priori expectation is that if humans dispersed *A. triloba* that resulting populations would be characterized by lower genetic diversity because of founder effects and higher genetic structure with most other populations. We further predict that populations of anthropogenic origin share alleles with geographically distant wild populations.

## 2 | MATERIALS AND METHODS

### 2.1 | Study species

*Asimina triloba* (L.) Dunal (Annonaceae) is the most widespread member of the genus, occurring in 26 states in the eastern United States, ranging from Texas and Iowa to the eastern seaboard and north into southern Ontario, Canada (Little, 1977). This understory tree, with trunks  $\leq 20$  cm diameter at breast height (DBH; Wyatt, personal observation), generally occurs at elevations  $< 350$  m above sea level,

primarily in mesic, alluvial forests along rivers and creeks (Freeman & Hulbert, 1985; Murphy, 2001; Pomper et al., 2010). Vegetative spreading occurs via root-suckers (Keener & Kuhns, 1998), resulting in dense patches and genets that persist for a long time (Hosaka et al., 2005). While there are no reports of longevity of individual ramets or genets, we documented 32 annual growth rings in one stem from Georgia with a DBH of 11.4 cm.

Flowers, which are strongly protogynous (Lagrange & Tramer, 1985), emerge in the spring after the last frost, and flowering on a given stem continues for approximately a month (Wyatt, personal observation). Flowering is also highly synchronous (Lagrange & Tramer, 1985). *Asimina triloba* pollinators are weak flying insects such as flies (including *Drosophila* spp.), beetles, and small Lepidoptera (Goodrich et al., 2006; Robertson, 1928; Willson & Schemske, 1980). Fruit set tends to be low (Lagrange & Tramer, 1985; Willson & Schemske, 1980), with only 0.4% of flowers and 15% of stems producing fruit in a central Illinois population (Willson & Schemske, 1980). Willson and Schemske (1980) found that fruit set was subject to pollen-limitation as well as resource limitation related to the size of the fruits. *Asimina triloba* produces the largest edible fruits ( $\leq 1$  kg; Darrow, 1975) of any native North American tree. Lagrange and Tramer (1985) also report fruit set relative to the number of flowers ranging from 0.5% to 3.4% in Tennessee and 0% in Toledo, Ohio.

Natural dispersal of *A. triloba* is likely by hydrochory and zoochory. Fruits float and are potentially dispersed along waterways (Bowden & Miller, 1951; Peterson, 1991; Wyatt, obs. pers.). The fruit is consumed by various animals with seeds reported in the scat of raccoons (Murphy, 2001; Willson, 1993; Willson & Schemske, 1980; Yeager & Elder, 1945), opossums, red fox (Murphy, 2001; Willson, 1993), and coyotes (Cypher & Cypher, 1999; Wyatt, personal observation) suggesting that these animals may mediate seed dispersal. There are reports of black bear in Virginia and North Carolina eating the fruit but seeds have not been documented in scat (Willson, 1993). Turkeys also consume the fruits (Murphy, 2001) although the seeds are digested and destroyed in the process (Wyatt, personal observation). Deer are unlikely vectors as there is no evidence that they consume the fruit other than in captivity (Murphy, 2001; Wyatt, personal observation). It is believed that historically the fruit was primarily consumed by megafauna that are now extinct or extirpated from their former range (Janzen & Martin, 1982; Poor, 1984) and that these megafauna may have mediated long-distance dispersal of pawpaw. Mammoths (*Mammuthus columbi*), tapirs (*Tapirus veroensis*), bison (*Bison bison*), horses (*Equus* sp.), mastodons (*Mammuth americanum*), and ground sloths (*Paramylodon harlani*, two species of *Eremotherium*, and three *Megalonyx* spp.) were present in Southeastern North America (Fields et al., 2012; Janzen & Martin, 1982; Poor, 1984) as recently as 12,700 YBP (Perrotti, 2018). Evidence suggests that horses and mastodons, for example, underwent annual migrations in Southeastern North America in excess of 150 km (Hoppe & Koch, 2007). Bison also migrated extensively before their numbers were decimated (Kauffman et al., 2021).

## 2.2 | Sampling and genotyping

Leaf tissue samples were collected from 82 mapped populations of *A. triloba* distributed throughout the species' range (Tables 1 and 2; Table S1; Figure 1). Populations were separated from one another by a mean of 681.3 km (range = 0.2 to 1,845.8 km). Twenty populations were classified as of putative anthropogenic origin based on their location  $< 1$  km from documented pre-Columbian habitations (i.e., villages or mound sites). We selected this objective criterion based on Muñoz et al. (2014) who indicate that impacts on plants by indigenous North Americans were localized around settlements as well as along travel corridors. Proximity to pre-Columbian settlements increases the likelihood that populations are anthropogenic, but anthropogenic origins become less probable with increasing distance from former settlements. Six of the 20 putative anthropogenic populations (LM1, LM2, LM3, MC1, MC2, and MC3) are located near archeological sites where seeds or carbonized remains of *A. triloba* have been found (Keener & Kuhns, 1998; Yarnell & Watson, 1966). Sixty-two populations were designated as "wild" as there are no known indigenous sites or trade routes located within 1 km of these populations. We compared the elevation of putative anthropogenic and wild populations and tested for significance with a two-sample *t* test assuming equal variances.

Leaf samples were collected from a mean of 28.1 individuals (range = 5–50) from each anthropogenic population and a mean of 29.6 individuals (range = 10–62) from each wild population (Table 2). Samples were taken from mature ( $\geq 3$  m in height) stems separated by  $\geq 10$  m when possible to avoid sampling multiple ramets belonging to the same genet. Samples were snap-frozen in liquid nitrogen and transported to the University of Georgia for genetic analyses.

Genomic DNA was extracted from frozen leaf tissue ( $\sim 0.05$  g) using a modified CTAB protocol (Doyle & Doyle, 1987), and DNA quality and quantity were evaluated using an ND-1000 NanoDrop® spectrophotometer. Nine nuclear microsatellite loci (Lu et al., 2011; Table S2) were PCR amplified using a 3-primer protocol whereby a CAG-tag sequence (Hauswaldt & Glenn, 2003) was added to the 5' end of one primer and a third fluorescently labeled (FAM, HEX, or NED) primer identical to the CAG-tag was included. The 12.5  $\mu$ l reaction volumes contained 3.25  $\mu$ l molecular grade ddH<sub>2</sub>O, 2.5  $\mu$ l 5 $\times$  One Taq® standard reaction buffer (New England Biolabs, Ipswich, MA), 0.35  $\mu$ l 25 mM MgCl solution (Sigma-Aldrich), 1.25  $\mu$ l 10 $\times$  (1 mg/ml) bovine serum albumin (Thermo-Fisher), 1  $\mu$ l 2.5 mM dNTP mix (New England Biolabs, Ipswich, MA), 1  $\mu$ l primer mix (0.5  $\mu$ M CAG-labeled primer and 5  $\mu$ M unlabeled primer), 0.45  $\mu$ l 10  $\mu$ M universal dye-labeled primer, 0.2  $\mu$ l One Taq® Hot Start DNA polymerase (New England Biolabs, Ipswich, MA), and 2.5  $\mu$ l of diluted template DNA (20 ng/ $\mu$ l).

Presence of null alleles was evaluated in Genepop v. 4.2 using Brookfield's method (Brookfield, 1996; Raymond & Rousset, 1995). Because *A. triloba* propagates vegetatively, we tested whether duplicate multilocus genotypes (MLGs) within populations represented clones. To estimate the likelihood of two individuals within a population having identical MLGs by chance, we calculated the probability of identity (PI) for each population which reflects the number of

**TABLE 1** Locations and elevation of twenty *Asimina triloba* populations of putative anthropogenic origin, with the pre-Columbian culture/period of nearby archeological evidence indicated. Populations are arranged in order of ascending latitude. Mean elevation is 194.6 m ( $SD = 83.6$ )

Pop	Location	Latitude	Longitude	Elevation (m)	Culture or period	Reference
FCP	Florida Caverns State Park, FL	30.810	-85.227	135	Fort Walton	Gardner (1966)
TUN	Tunica Hills WMA, LA	30.930	-91.514	91	Tunica	Perrault et al. (2008)
OSF	Old Stone Fort State Park, TN	35.485	-86.103	314	Middle Woodland	Yerka (2010)
PET	Pettigrew State Park, NC	35.793	-76.413	11	Late Woodland	Phelps (1983)
SEL	Sellars Farm Archeological Site, TN	36.169	-86.236	191	Mississippian	Butler (1981)
MC3	Mammoth Cave, KY	37.148	-86.093	248	Woodland	Yarnell and Watson (1966)
MC2	Mammoth Cave, KY	37.178	-86.108	164	Woodland	Yarnell and Watson (1966)
MC1	Mammoth Cave, KY	37.204	-86.139	245	Woodland	Yarnell and Watson (1966)
POC	Pocahontas State Park, VA	37.381	-77.583	87	Powhatan	Higgins et al. (1995); Kandle et al. (1996)
FHM	Fort Hill Memorial, OH	39.123	-83.401	288	Hopewell	Prufer (1964)
SCI	Scioto Trail State Park, OH	39.216	-82.974	309	Middle Woodland	Carr (2008)
LM3	Little Miami River, OH	39.409	-84.100	203	Fort Ancient	Connolly (2004)
LM2	Little Miami River, OH	39.410	-84.101	204	Fort Ancient	Connolly (2004)
LM1	Little Miami River, OH	39.410	-84.099	206	Fort Ancient	Connolly (2004)
SU1	Susquehanna State Park, MD	39.610	-76.147	74	Late Woodland	MacNamara (1982)
CFP	Cunningham Falls State Park, MD	39.633	-77.456	313	Late Woodland	Bedell et al. (2011)
CON	Conneaut, OH	41.933	-80.611	221	Woodland	Wallace (1952); Brose (1976)
LCN	Love Creek Nature Center, MI	41.951	-86.306	225	Late Woodland	Hinsdale (1931); Lovis (1974)
CHA	Chautauqua Creek, NY	42.339	-79.604	186	Late Woodland	Wallace (1952); Bennett (1986)
WEN	Wendt County Beach Park, NY	42.679	-79.052	176	Late Woodland	Bennett (1986)

loci, allele frequencies, and sample size. The multilocus probability of identity ( $PI_m$ ) was calculated as  $PI_m = \prod_s (PI_s)$ , where

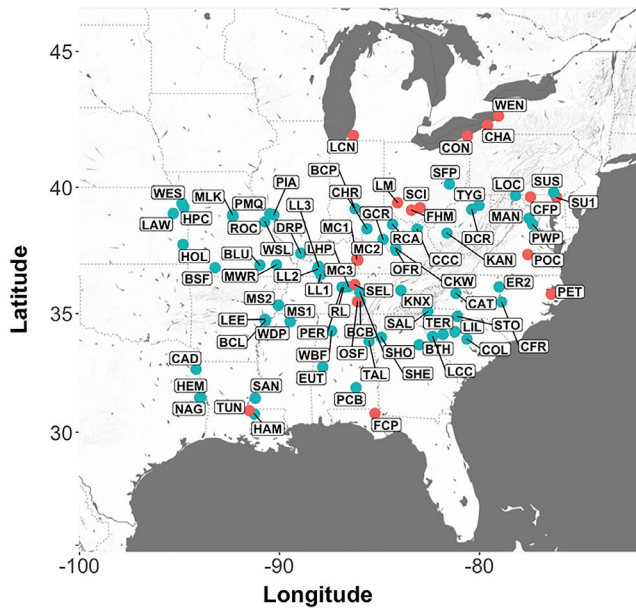
$$PI_s = \sum_i p_i^4 + \sum_i \sum_{j>1} (2p_i p_j)^2$$

and  $p_i$  and  $p_j$  are the frequencies of the  $i$ th and  $j$ th alleles at a locus, respectively (Gonzales et al., 2008; Paetkau & Strobeck, 1994). The PI was adjusted for sample size ( $N$ ) by  $PI = 1 - [(1 - PI_m)^N]$  (Parker et al., 1998). Duplicate MLGs that represented clones were removed, leaving one individual per genet for subsequent analyses.

The number of MLGs and the number of MLGs adjusted for population sample size (sMLG; i.e., rarefaction) were calculated using FSTAT (Goudet, 1995, 2001). For the rarefaction analysis, all populations were treated as consisting of 10 individuals to reflect the number of individuals in the smallest population after  $SEL_A$  ( $N = 5$ ),  $LM2_A$  ( $N = 6$ ), and  $STO_W$  ( $N = 9$ ) were omitted. Genetic diversity statistics were estimated separately for putative anthropogenic populations and wild populations using GenAEx v. 6.51b2 (Peakall & Smouse, 2006, 2012). Measures of genetic diversity were percent of polymorphic loci ( $P$ ), total number of alleles ( $A_T$ ), effective number of alleles ( $A_E$ ), number of private alleles ( $P_A$ ; an allele found in a

single population), rare alleles ( $A_{RARE}$ ; alleles in <10% of all populations), observed heterozygosity ( $H_O$ ), expected heterozygosity ( $H_E$ ), and the inbreeding coefficient ( $F_{IS}$ ). For all genetic diversity statistics, we used a two-sample  $t$  test assuming equal variances to assess whether there were significant differences between anthropogenic and wild populations. Population values were averaged across populations as well as pooled. Observed heterozygosity ( $H_O$ ) and Hardy-Weinberg expected heterozygosity ( $H_E$ ) were compared for each locus in each population, using Wright's inbreeding coefficient ( $F_{IS}$ ; Wright, 1922, 1951). Deviations from Hardy-Weinberg expectations were tested for significance using chi-square (Li & Horvitz, 1953). A Bonferroni correction for multiple comparisons was applied using FSTAT (Goudet, 1995, 2001). Correlations between MLGs and  $F_{IS}$  values, as well as between latitude and  $F_{IS}$  values, were tested for significance using linear regression analyses.

Genetic structure among populations was estimated using Nei's (1973)  $G_{ST}$ . Pairwise  $G_{ST}$  values were calculated for all possible pairs of populations in GenAEx.v. 6.5.1b2 (Peakall & Smouse, 2006, 2012). A two-sample  $t$  test assuming equal variances was used to test for significant differences between mean pairwise  $G_{ST}$  values and mean pairwise geographic distances among anthropogenic populations versus wild populations.



**FIGURE 1** Map of 20 putative anthropogenic populations (red circles) and 62 wild populations (blue circles) of *Asimina triloba* in the eastern United States

For the nine putative anthropogenic populations located at, or within, the boundary of the Wisconsin glaciation we used a two-sample *t* test assuming equal variances to assess significance between (a) pairwise  $G_{ST}$  values among the nine populations versus pairwise  $G_{ST}$  values between these populations and the wild populations with which they were most genetically similar, and (b) pairwise geographic distances of putative anthropogenic populations located at or within the boundary of the Wisconsin glaciation and the wild populations with which they shared their lowest  $G_{ST}$  values.

Isolation by distance (IBD; Wright, 1943) was estimated using linear regression analysis of  $G_{ST} / (1 - G_{ST})$  against log geographic distance (Rousset, 1997) for all 82 populations as well as separately for anthropogenic and wild populations. Significance was tested using a Mantel test in GenAlEx.v. 6.5.1b2 (Peakall & Smouse, 2006, 2012).

STRUCTURE v. 2.3.4 (Pritchard et al., 2000), a Bayesian clustering approach, was used to estimate levels of genetic admixture among all 82 populations and the number of genetically distinct clusters ( $K$ ). Levels of admixture and  $K$  were also estimated separately for putative anthropogenic populations as a part of the present study and wild populations as reported by Wyatt et al. (2021). Ten independent simulations at each  $K$ -value from 1 to 20 were run, using a burn-in of 500,000 repetitions and a run length of 1,000,000 Markov chain Monte Carlo (MCMC) iterations. The admixture model was chosen to infer alpha ( $\alpha$ ). We also employed a model based on correlated allele frequencies with no a priori assumptions regarding sampling locations. The Evanno et al. (2005) method was used to determine the optimal number of genetic clusters ( $K$ ) in STRUCTURE HARVESTER (Earl & vonHoldt, 2012).

The occurrence of rare alleles in <10% of all wild and putative anthropogenic populations (i.e., alleles present in 2–8 populations) was used to more closely investigate relationships between putative anthropogenic and wild populations. A two-sample *t* test assuming

equal variances was used to assess whether the number of rare alleles in anthropogenic versus wild populations differed significantly. We also used linear regression analyses to test whether there was a significant relationship between latitude<sub>RA</sub> and the number of rare alleles in anthropogenic and wild populations.

We were particularly interested in examining rare alleles shared by putative anthropogenic and wild populations. We estimated pairwise  $G_{ST}$  values and pairwise geographic distances of all wild populations ( $W$ ), pairs of wild populations sharing rare alleles ( $W_{RA}$ ), wild and anthropogenic populations sharing rare alleles ( $W - A$ )<sub>RA</sub>, anthropogenic populations in areas not formerly glaciated and wild populations sharing rare alleles ( $A_{NG} - W$ )<sub>RA</sub>, as well as anthropogenic populations in formerly glaciated areas and wild populations sharing rare alleles ( $A_G - W$ )<sub>RA</sub>. We used a two-sample *t* test assuming equal variances to test for significance between pairwise  $G_{ST}$  values and geographic distances of (a) wild populations ( $W$ ) versus wild populations that share rare alleles ( $W_{RA}$ ), (b) wild populations ( $W$ ) versus wild and anthropogenic populations that share rare alleles ( $(W - A)_{RA}$ ), (c) wild populations that share rare alleles ( $W_{RA}$ ) versus wild and anthropogenic populations that share rare alleles ( $(W - A)_{RA}$ ), and (d) ( $A_{NG} - W$ )<sub>RA</sub> versus ( $A_G - W$ )<sub>RA</sub>.

### 2.3 | Fruit set

To better evaluate whether populations of putative anthropogenic origin and populations at the edge of the species range exist beyond or at the edge of their ecological limit, we collected fruit set data. For a species to persist in a particular landscape, individuals must be able to survive and reproduce. Fruit set was recorded for each population between May and August, although primarily in July and August, after flowering had ended at which point rapidly enlarging ovaries are visually conspicuous. Because of the large number of populations and their widespread distribution, data were collected over the course of three summers with each population surveyed once. Developing fruits occurring on all stems sampled for genotyping were counted. The stage of fruit maturation was not recorded as it was impossible to visit all populations within a sufficiently small window of time to render comparable data.

For each population, the number of fruits/stem and fruits/MLG were calculated and a two-sample *t* test assuming equal variances was used to assess whether anthropogenic and wild populations differed significantly. We used linear regression analysis separately for anthropogenic and wild populations to test correlations between (a) the number of MLGs and the number of fruits produced, (b) latitude and the number of fruits/stem, and (c) latitude and the number of fruits/MLG.

## 3 | RESULTS

Ten (12%) of the 82 populations are in landscapes within, or at, the boundary of the Wisconsin glaciation maximum. Of these, nine are anthropogenic ( $WEN_A$ ,  $CHA_A$ ,  $LCN_A$ ,  $CON_A$ ,  $LM1_A$ ,  $LM2_A$ ,  $LM3_A$ ,

SCI<sub>A</sub>, and FHM<sub>A</sub>) and one (BCP<sub>W</sub>) is wild. Thus, 45% of anthropogenic populations and 2% of wild populations occur in landscapes that were colonized after retreat of the LGM. The mean elevation of putative anthropogenic populations ( $194.6 \pm 83.6$  m) exceeded the mean for wild populations ( $178.8 \pm 94.3$  m), but the difference was nonsignificant ( $p = .253$ ; Tables 1; Table S1).

### 3.1 | Genetic diversity

Null allele frequencies were nonsignificant for all loci and populations ( $p = .05$ ). Mean probability of identity (PI) for anthropogenic and wild populations was  $1.46 \times 10^{-4}$  and  $1.15 \times 10^{-3}$ , respectively

(Table S3), indicating a high likelihood that identical MLGs within populations represent clones. Despite efforts to avoid sampling multiple ramets per genet, identical MLGs were present in all but three wild populations (BTH<sub>W</sub>, LIL<sub>W</sub>, and LL3<sub>W</sub>). There were no MLGs shared by individuals in different populations. Although mean sample sizes for anthropogenic and wild populations were 27.3 and 29.4, respectively, six anthropogenic (30%) and eight wild (13%) populations had a single MLG.

The mean number of MLGs was significantly lower for anthropogenic than wild populations (5.3 and 10.1, respectively;  $p = .011$ ; Table 2; Figure S1). When adjusted for sample size, the mean number of sMLGs was also significantly lower for anthropogenic populations (4.6 and 7.0, respectively;  $p = .029$ ; Table 2).

**TABLE 2** Summary statistics of *A. triloba* nuclear genetic diversity in 20 putative anthropogenic and 62 wild populations arranged in order of ascending latitude. Data for wild populations are from Wyatt et al. (2021)

Population	N	MLG	sMLG	P	A <sub>T</sub>	A <sub>E</sub>	P <sub>A</sub>	A <sub>RARE</sub>	H <sub>O</sub>	H <sub>E</sub>	F <sub>IS</sub>
Anthropogenic											
FCP	35	3	3.0	44	17	1.8	2	4	0.185	0.290	0.348 <sup>¥</sup>
TUN	21	7	6.1	100	45	3.9	2	7	0.651	0.668	0.044
OSF	35	3	3.0	100	30	2.9	—	4	0.630	0.537	-0.179
PET	24	22	15.1	89	33	2.3	—	2	0.479	0.495	0.051
SEL	5	2	—	78	22	2.3	—	2	0.722	0.472	-0.562
MC3	12	1	1.0	44	13	1.4	—	1	0.444	0.222	-1.000 <sup>¥</sup>
MC2	18	10	9.6	100	36	2.5	—	1	0.658	0.541	-0.194
MC1	16	7	6.9	100	33	2.7	—	4	0.501	0.532	0.098
POC	48	6	3.7	67	21	2.0	—	—	0.537	0.403	-0.400 <sup>¥</sup>
FHM	46	6	3.7	100	29	2.4	2	—	0.491	0.545	0.058
SCI	24	1	1.0	67	15	1.7	1	1	0.667	0.333	-1.000 <sup>¥</sup>
LM3	14	3	3.4	78	20	2.1	—	—	0.593	0.444	-0.383
LM1	17	14	12.0	100	43	3.3	—	2	0.583	0.632	0.125
LM2	6	1	—	67	15	1.7	—	2	0.667	0.333	-1.000 <sup>¥</sup>
SU1	33	11	6.7	89	39	3.1	—	3	0.606	0.559	-0.111
CFP	48	1	1.0	56	13	1.4	—	1	0.556	0.278	-1.000 <sup>¥</sup>
CON	30	2	1.9	89	24	2.4	—	1	0.778	0.528	-0.467 <sup>¥</sup>
LCN	48	4	2.5	89	18	1.8	—	3	0.778	0.448	-0.690 <sup>¥</sup>
CHA	38	1	1.0	33	11	1.2	—	—	0.333	0.167	-1.000 <sup>¥</sup>
WEN	10	1	1.0	67	15	1.7	—	—	0.667	0.333	-1.000 <sup>¥</sup>
Mean	27.3	5.3	4.59	77.9	24.6	2.24	0.35	1.9	0.576	0.438	-0.413
SD	13.6	5.5	4.11	21.4	10.6	0.70	0.50	1.7	0.145	0.138	0.467
Pooled	546	106	—	100	98	5.69	7	—	0.591	0.709	0.170
Wild											
HAM	47	2	1.9	100	28	2.2	—	3	0.444	0.528	0.170
SAN	21	1	1.0	78	16	1.8	1	2	0.778	0.389	-1.000 <sup>¥</sup>
NAG	29	3	2.6	78	20	3.1	—	1	0.667	0.586	-0.158
HEM	40	5	2.6	89	26	2.8	—	3	0.578	0.544	-0.016
PCB	48	14	8.0	89	29	1.9	—	6	0.397	0.391	0.013
CAD	36	21	8.0	100	31	3.3	1	2	0.633	0.657	0.085

(Continues)

TABLE 2 (Continued)

Population	N	MLG	sMLG	P	A <sub>T</sub>	A <sub>E</sub>	P <sub>A</sub>	A <sub>RARE</sub>	H <sub>O</sub>	H <sub>E</sub>	F <sub>IS</sub>
EUT	33	32	16.5	100	57	3.9	—	13	0.463	0.650	0.231
BTH	36	36	17.3	100	54	3.4	2	8	0.620	0.665	0.066 <sup>Y</sup>
TAL	24	3	2.7	89	25	2.2	—	1	0.611	0.480	-0.266
COL	48	18	8.2	78	29	1.9	—	2	0.394	0.397	-0.015
SHE	24	5	4.2	89	26	2.3	—	3	0.511	0.479	-0.045
LCC	24	23	15.8	89	38	2.8	—	4	0.476	0.555	0.147
TER	33	27	14.7	89	41	2.9	—	5	0.529	0.579	0.081
LIL	24	24	14.9	89	44	2.7	—	4	0.452	0.562	0.194
WBF	43	22	13.1	89	53	3.5	—	13	0.578	0.623	0.082
WDP	27	3	2.6	89	26	2.5	1	4	0.444	0.531	0.192
BCL	48	12	10.6	100	69	2.7	1	5	0.503	0.573	0.125
LEE	24	12	9.8	100	56	4.1	—	8	0.643	0.703	0.096
STO	9	5	—	89	25	2.3	—	1	0.454	0.498	0.129
SAL	30	14	9.9	100	43	3.0	—	1	0.563	0.612	0.030
MS1	24	4	3.7	78	18	3.5	—	—	0.611	0.674	0.072
MS2	23	1	1.0	67	15	0.8	1	—	0.111	0.056	-1.000 <sup>Y</sup>
CFR	20	5	4.1	89	30	2.5	1	1	0.689	0.550	-0.238
BCB	25	8	6.4	78	21	1.9	—	3	0.597	0.404	-0.367*
CAT	16	1	1.0	67	14	1.7	—	—	0.667	0.333	-1.000 <sup>Y</sup>
SHO	24	16	12.2	89	49	3.4	—	6	0.596	0.566	-0.062
KNX	25	3	2.7	100	31	2.8	2	2	0.537	0.600	0.133
RL	27	10	8.7	100	44	3.6	2	4	0.681	0.639	-0.064
PER	46	4	3.3	100	37	3.4	—	—	0.583	0.642	0.044
ER2	10	4	4.0	67	17	1.7	—	1	0.500	0.326	-0.527
LHP	20	4	3.5	89	22	1.9	—	2	0.611	0.406	-0.442*
LL1	25	1	1.0	56	14	1.6	—	1	0.556	0.278	-1.000 <sup>Y</sup>
LL2	25	7	6.2	78	38	3.4	—	2	0.613	0.580	-0.055
BSF	23	21	14.9	89	43	3.1	—	5	0.546	0.543	-0.010
LL3	24	24	16.0	78	51	3.7	1	8	0.495	0.577	0.141
BLU	25	9	7.32	78	28	2.5	1	1	0.460	0.506	0.062
MWR	25	14	10.3	78	38	2.3	1	4	0.524	0.442	-0.092
DRP	23	13	10.3	100	41	3.3	2	7	0.562	0.599	0.070
OFR	33	7	5.7	89	36	2.9	—	5	0.550	0.512	-0.082
CKW	22	13	10.0	100	40	2.5	4	2	0.523	0.526	-0.024
HOL	10	9	9.0	89	34	2.9	—	2	0.545	0.495	-0.098
GCR	48	5	3.5	78	26	2.4	1	1	0.356	0.471	0.316 <sup>S</sup>
KAN	43	10	7.7	89	43	3.4	—	1	0.559	0.599	0.090
CCC	26	12	9.5	89	43	3.2	—	5	0.477	0.614	0.222
CHR	30	19	12.5	89	45	3.2	—	4	0.580	0.598	0.030
PWP	48	27	14.3	100	56	3.6	—	6	0.540	0.670	0.169
RCA	24	16	11.6	100	49	3.8	1	2	0.490	0.645	0.273
WSL	11	9	9.0	100	40	3.5	—	6	0.600	0.647	0.078
MAN	48	1	1.0	33	10	1.1	—	—	0.333	0.167	-1.000 <sup>Y</sup>
ROC	25	1	1.0	56	14	1.6	—	—	0.556	0.278	-1.000 <sup>Y</sup>
MLK	24	2	1.9	67	18	1.8	1	—	0.611	0.375	-0.622 <sup>Y</sup>

(Continues)



TABLE 2 (Continued)

Population	N	MLG	sMLG	P	A <sub>T</sub>	A <sub>E</sub>	P <sub>A</sub>	A <sub>RARE</sub>	H <sub>O</sub>	H <sub>E</sub>	F <sub>IS</sub>
PIA	14	1	1.0	44	15	1.4	1	1	0.444	0.222	-1.000 <sup>¥</sup>
PMQ	59	4	3.5	89	26	2.2	—	1	0.583	0.441	-0.313 <sup>§</sup>
LAW	30	12	7.8	89	29	1.9	—	2	0.492	0.392	-0.185
DCR	15	1	1.0	44	13	1.4	—	—	0.444	0.222	-1.000 <sup>¥</sup>
BCP	48	2	2.0	67	17	1.7	1	—	0.444	0.292	-0.547 <sup>¥</sup>
HPC	25	15	10.6	89	31	2.1	—	—	0.467	0.469	-0.020
TYG	48	2	2.0	89	20	2.0	—	1	0.722	0.444	-0.600 <sup>¥</sup>
WES	22	9	7.3	100	45	3.1	—	5	0.537	0.590	0.078
LOC	38	3	2.2	56	16	1.6	—	—	0.278	0.264	-0.134
SUS	30	8	6.7	100	37	3.0	—	2	0.602	0.603	0.088
SFP	25	7	6.0	100	27	2.1	1	—	0.456	0.482	0.014
Mean	29.4	10.1	7.01	84.7	32.5	2.59	0.44	2.9	0.530	0.496	-0.153
SD	11.5	8.6	4.83	15.7	13.5	0.77	0.75	2.9	0.108	0.141	0.387
Pooled	1824	626	—	100	153	6.52	27	—	0.553	0.765	0.273
Species mean	28.9	8.9	6.46	83.0	30.6	2.50	0.41	2.7	0.541	0.482	0.216
Species pooled	2,370	732	—	100	160	6.60	34	54	0.559	0.760	0.260

Note: N = sample size, MLG = number of multilocus genotypes, sMLG = number of multilocus genotypes adjusted for sample sizes after the three smallest populations were omitted (LM2<sub>A</sub>, SEL<sub>A</sub>, and STO<sub>W</sub>), P = percent of polymorphic loci, A<sub>T</sub> = total number of alleles, A<sub>E</sub> = effective number of alleles, P<sub>A</sub> = number of private alleles, A<sub>RARE</sub> = number of rare alleles (<10% of all anthropogenic and wild populations), H<sub>O</sub> = observed heterozygosity, H<sub>E</sub> = expected heterozygosity, F<sub>IS</sub> = inbreeding coefficient.

Significant F<sub>IS</sub> values are indicated with \* ( $p < .05$ ), § ( $p < .01$ ) and ¥ ( $p < .005$ ).

A total of 160 alleles occurred across all 82 populations (Table 2); however, the mean was significantly lower in anthropogenic (24.6) than wild (32.5) populations ( $p = .010$ ). A total of 34 private alleles occurred in four anthropogenic (20%) and 20 wild populations (32%). Anthropogenic populations of *A. triloba* had less genetic diversity than wild populations for all population genetic statistics except observed heterozygosity, although the differences were significant only for the number of MLGs (5.3 vs. 10.1;  $p = .011$ ), sMLGs (4.6 vs. 7.0;  $p = .029$ ), and A<sub>T</sub> (24.6 vs. 32.5;  $p = .010$ ). Other diversity statistics for anthropogenic versus wild populations were P (77.9 vs. 84.7;  $p = .198$ ), A<sub>E</sub> (2.2 vs. 2.6;  $p = .056$ ), P<sub>A</sub> (0.35 vs. 0.44;  $p = .948$ ), A<sub>RARE</sub> (1.9 vs. 2.9;  $p = .071$ ), and H<sub>E</sub> (0.438 vs. 0.496;  $p = .114$ ). Anthropogenic populations had a higher mean H<sub>O</sub> (0.576) than wild populations (0.530;  $p = .199$ ), although the difference was not significant, and a significantly lower mean inbreeding coefficient (F<sub>IS</sub> = -0.413 vs. -0.153;  $p = .032$ ), indicative of an excess of heterozygosity relative to Hardy-Weinberg expectations (HWE; Table 2). F<sub>IS</sub> values were highly significant for 50% of anthropogenic populations ( $p < .005$ ) and significant for 19% wild populations ( $p < .05$ ; Table 2). For nearly all populations with an inbreeding coefficient that differed significantly from HWE, there was excess heterozygosity. There was a significant positive correlation between MLGs and F<sub>IS</sub> values in anthropogenic ( $r = 0.640$ ;  $p = .001$ ) and wild populations ( $r = 0.570$ ;  $p = 6.50 \times 10^{-7}$ ). F<sub>IS</sub> values also decreased significantly with increasing latitude in anthropogenic populations ( $r = -0.599$ ;

$p = .005$ ) but the relationship was nonsignificant in wild populations ( $r = -0.192$ ;  $p = .134$ ).

### 3.2 | Genetic structure

Genetic structure ( $G_{ST}$ ) among anthropogenic and wild populations was 0.388 and 0.345, respectively. There was significant IBD among all 82 populations ( $r = 0.143$ ; (P(rxy-rand) > = rxydata) = 0.010) and among the 62 wild populations ( $r = 0.158$ ; (P(rxy-rand) > = rxydata) = 0.020), but among anthropogenic populations there was an absence of significant IBD ( $r = 0.068$ ; (P(rxy-rand) > = rxydata) = 0.230). Although the mean geographic distance separating pairs of anthropogenic populations was less than between wild populations (629.3 vs. 660.3 km;  $p = .125$ ), the mean pairwise  $G_{ST}$  value was significantly higher for anthropogenic populations (0.265 vs. 0.228;  $p = 1.12 \times 10^{-7}$ ).

The nine northernmost populations of putative anthropogenic origin occupying landscapes at, or within, the boundary of the Wisconsin glaciation maximum (CHA<sub>A</sub>, CON<sub>A</sub>, FHM<sub>A</sub>, LCN<sub>A</sub>, LM1<sub>A</sub>, LM2<sub>A</sub>, LM3<sub>A</sub>, SCI<sub>A</sub>, and WEN<sub>A</sub>) had a mean pairwise  $G_{ST}$  = 0.289 (range = 0.071 to 0.523) and were separated from one another by a mean of 317 km (range = 0.2 to 602 km). These populations instead shared their lowest pairwise  $G_{ST}$  values (mean = 0.137; range = 0.037 to 0.268) with wild populations that were a mean of 865 km away

(range = 579 to 1,342 km). Not only did these anthropogenic populations and their genetically closest wild populations have significantly lower pairwise  $G_{ST}$  values ( $p = 1.29 \times 10^{-4}$ ) than these nine anthropogenic populations had with each other, but they were also significantly further apart ( $p = 1.69 \times 10^{-8}$ ). These anthropogenic populations and their genetically closest wild populations also had significantly lower pairwise  $G_{ST}$  values ( $p = .001$ ) and were significantly further apart ( $p = .040$ ) than all wild populations. The one wild population at the boundary of the Wisconsin glaciation ( $BCP_W$ ) is 350 km from the population with which it was most genetically similar ( $G_{ST} = 0.161$ ).

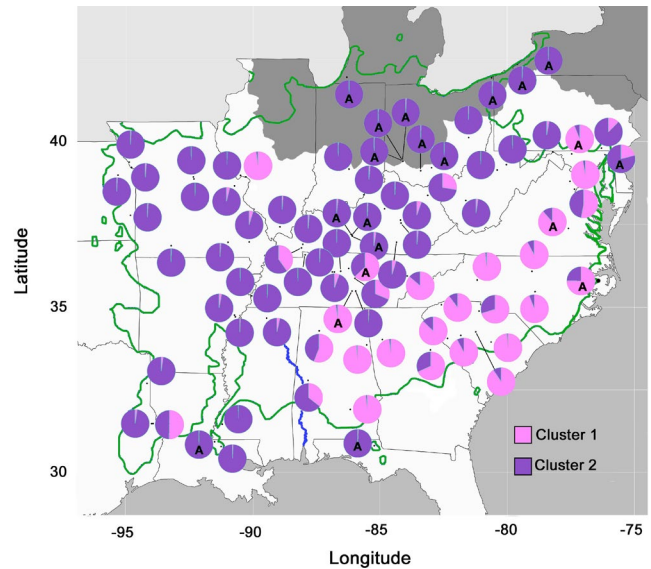
Simulations in STRUCTURE yielded an optimal  $K = 2$  for all 82 populations, with mean  $\ln P(D)$  of  $-67,119.0$  and  $\Delta K$  of 20.6 (Table 3; Table S4; Figure 2; Figure S2). The next most likely number of clusters was  $K = 4$  with mean  $\ln P(D)$  of  $-63,063.9$  and  $\Delta K$  of 8.5 (Table 3; Figures S2 and S3). Caution is required in interpretation of STRUCTURE results for all 82 populations as well as the 62 wild populations as the presence of significant IBD violates one of the underlying assumptions of STRUCTURE (Pritchard et al., 2000). Simulations for the 20 putative anthropogenic populations also found  $K = 2$  to be optimal, but the clusters are not geographically distinct: populations with a higher genetic assignment to one cluster are often geographically embedded in the other cluster (Figure 3).

### 3.3 | Rare alleles

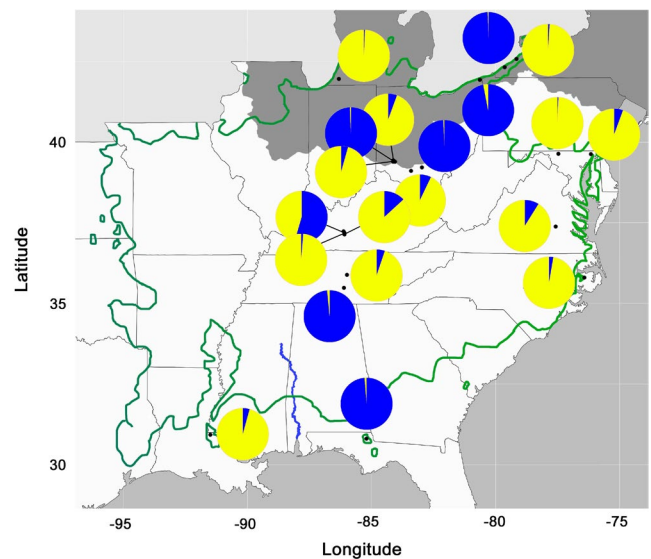
A total of 54 rare alleles occurred in 75% (15) of anthropogenic and 81% (50) of wild populations. Only 24 rare alleles were shared by both anthropogenic and wild populations, and no rare alleles occurred exclusively in anthropogenic populations (Figure S4). There were also fewer rare alleles in anthropogenic populations (mean = 1.9 vs. 2.9;  $p = .071$ ). The number of rare alleles per population decreased significantly with increasing latitude for both anthropogenic ( $r = -0.679$ ;  $p = .001$ ) and wild populations ( $r = -0.302$ ;  $p = .017$ ; Figure 4).

As predicted, the pairwise  $G_{ST}$  values between wild populations that share rare alleles ( $W_{RA}$ ) were significantly smaller than for all

wild populations ( $W$ ; 0.166 and 0.228, respectively;  $p = 1.37 \times 10^{-23}$ ) and pairwise geographic distances for  $W_{RA}$  were significantly smaller (607.1 and 660.3 km, respectively;  $p = .014$ ; Table 4). Wild and anthropogenic populations that shared rare alleles ( $W - A$ ) $_{RA}$  had a significantly higher pairwise  $G_{ST}$  values than  $W_{RA}$  (0.185 vs. 0.166;  $p = .012$ ) but were also separated by significantly larger pairwise geographic distances (695.3 vs. 607.1 km;  $p = .014$ ; Table 4). Anthropogenic populations in formerly glaciated landscapes and the



**FIGURE 2** All sampled populations ( $N = 82$ ) showing proportion of each population assigned to the two clusters identified by STRUCTURE v 2.3.4. The Tombigbee River in Alabama is indicated in blue. The reported range of *A. triloba* from Little (1977) is outlined in green, and the maximum extent of the Wisconsin glaciation is indicated in dark gray. A = putative anthropogenic populations. Wild populations have no designation



**FIGURE 3** Putative anthropogenic populations of *A. triloba* showing the proportion of assignment to the two genetic clusters identified by STRUCTURE v 2.3.4

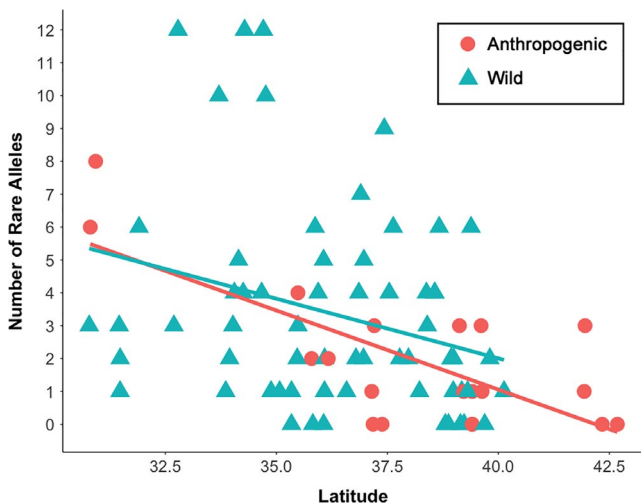
**TABLE 3** STRUCTURE output for all 82 populations of *A. triloba* for  $K$  values 2 – 10, without population location data included in the models

$K$	Mean $\ln P(K)$	Delta $K$
1	-70,698.6	–
2	-67,119.0	20.5702
3	-65,873.1	0.7354
4	-63,063.9	8.4830
5	-61,303.8	3.7374
6	-60,298.4	0.8305
7	-58,730.3	1.1389
8	-57,735.0	0.4290
9	-56,464.2	0.1425
10	-55,300.5	–

wild populations with which they shared rare alleles were separated by a larger mean pairwise distance (723.1 vs. 687.7 km;  $p = 0.345$ ) and had a significantly higher mean  $G_{ST}$  (0.216 vs. 0.176;  $p = 0.018$ ) (Table 4) than anthropogenic populations in areas not formerly glaciated and wild populations that share rare alleles.

### 3.4 | Fruit Set

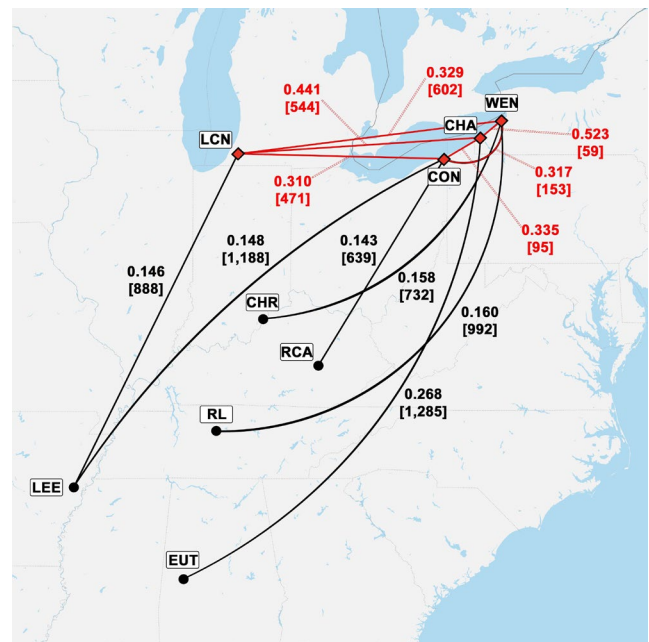
Fruit set occurred in fewer anthropogenic than wild populations (35% vs. 40%). Anthropogenic populations also produced fewer fruits per stem (mean = 0.10 vs. 0.29;  $p = .175$ ) and fewer fruits per MLG than wild populations (mean = 0.40 vs. 0.47;  $p = .386$ ; Table S5). Of the nine anthropogenic populations in formerly glaciated habitat, only two produced fruit ( $SCI_A = 1$  and  $LCN_A = 1$ ; Table S5). In wild populations, fruit set was significantly positively correlated with the number of MLGs ( $r = 0.498$ ;  $p = 3.77 \times 10^{-5}$ ), but the relationship was not significant for anthropogenic populations ( $r = 0.219$ ;  $p = .354$ ). In both anthropogenic and wild populations, fruit set per stem and per MLG decreased nonsignificantly ( $p > .05$ ) with increasing latitude.



**FIGURE 4** Correlation between latitude and the number of rare alleles (i.e., present in <10% of all 82 populations) for twenty anthropogenic ( $r = -0.679$ ;  $p = .001$ ) and 62 wild populations ( $r = -0.301$ ;  $p = .017$ ) of *Asimina triloba*

## 4 | DISCUSSION

Consistent with our a priori expectations, all measures of genetic diversity indicate that putative anthropogenic populations possess less genetic variation than wild populations, with four measures differing significantly. Only observed heterozygosity was higher in anthropogenic populations, consistent with their significantly lower  $F_{IS}$  values. The low number of MLGs and sMLGs that characterize anthropogenic populations appear to be the most important distinguishing characteristics. Reduced genetic diversity in populations near sites historically occupied by humans has also been reported in *Agave* (Parker et al., 2007), *Spondias* (Miller & Schaal, 2006), *Phaseolus*



**FIGURE 5** Map showing pairwise  $G_{ST}$  values and geographic distances in km (shown in brackets) among the four northernmost *A. triloba* populations (LCN, CON, CHA, and WEN), of putative anthropogenic origin, as well as between these four populations and the wild populations with which they share the lowest  $G_{ST}$  values. Red diamonds designate anthropogenic populations, and black circles designate wild populations. Red lines connect the anthropogenic populations while black lines connect anthropogenic and wild populations

**TABLE 4** Summary of pairwise geographic distances (km) and  $G_{ST}$  values between all wild populations of *A. triloba* (W), wild populations that share rare alleles ( $W_{RA}$ ), wild and anthropogenic populations that share rare alleles ( $(W - A)_{RA}$ ), anthropogenic populations in areas not formerly glaciated and wild populations that share rare alleles ( $(A_{NG} - W)_{RA}$ ), and anthropogenic populations in formerly glaciated areas and wild populations that share rare alleles ( $(A_G - W)_{RA}$ ). Statistically significant comparisons are bolded

	W	$W_{RA}$	W	$(W - A)_{RA}$	$W_{RA}$	$(W - A)_{RA}$	$(A_{NG} - W)_{RA}$	$(A_G - W)_{RA}$
Mean km	660.3	607.1	660.3	695.3	607.1	695.3	687.7	723.1
SD	351.4	331.8	351.4	392.3	331.8	392.2	410.5	331.7
$p$ -value	<b>.014</b>		.150		<b>.014</b>		.345	
Mean $G_{ST}$	0.228	0.166	0.228	0.185	0.166	0.185	0.176	0.216
SD	0.091	0.064	0.091	0.083	0.064	0.083	0.080	0.090
$p$ -value	<b><math>1.37 \times 10^{-23}</math></b>		<b><math>2.71 \times 10^{-7}</math></b>		<b>.012</b>		<b>.018</b>	

(Bitocchi et al., 2012), *Theobroma* (Motamayor et al., 2002), and *Annona* (Larranaga et al., 2017; see also Levis et al., 2017). Reduced genetic variation is consistent with founder effects where few seeds or ramets served as colonists, introducing only a subset of the genetic variation present in source populations. It also suggests that putative anthropogenic populations of *A. triloba* have experienced little gene flow subsequent to their initiation.

The absence of more significant differences may relate to the unavoidably imperfect characterization of study populations. Some populations that we designated as wild may in fact be anthropogenic in origin. Particularly, suspect are populations with genetic profiles (e.g., few MLGs and sMLGs, low  $A_T$  and high  $F_{IS}$ ) similar to those of populations associated with documented pre-Columbian settlements or mound sites. It is also possible that some populations that we designated as anthropogenic are actually wild. Incomplete documentation regarding pre-Columbian population sites could lead to erroneous designations, and the dating of occupation of some sites as far back as 15,000–17,000 YBP (Adovasio et al., 1978; Delcourt & Delcourt, 2004; Morse & Morse, 1983; Williams et al., 2018) could mean the erosion of genetic signatures of anthropogenic origins.

Regarding the inverse relationship between genetic diversity and excess heterozygosity, wild populations of *Fagus* in Europe reveal a similar pattern (Comps et al., 2001). Excess heterozygosity appears to be correlated with asexual reproduction (Ballou et al., 2003; Birky, 1996; Bryzski & Culley, 2011; Meloni et al., 2013) which agrees with our finding of significantly higher clonality and excess heterozygosity in anthropogenic populations of *A. triloba* (Table 2). Excess heterozygosity may result from dominance of well-adapted, highly heterozygous genotypes or the introduction of a few seeds from multiple genetically divergent source populations. If humans intentionally transported the fruits or seeds from wild populations, it is likely that propagules were selected on the basis of favorable traits (i.e., fruit production, fruit size, flavor, and/or fiber quality; Peattie, 1950; Peterson, 1991) possibly associated with higher heterozygosity.

The significant decline in  $F_{IS}$  values with increasing latitude in anthropogenic populations is also consistent with studies showing that populations at the edge of a species' range, where they are at their ecological limit, tend to have more asexual reproduction (Billingham et al., 2003; Dorken & Eckert, 2001; Eckert, 2002; Garcia et al., 2000; Silvertown, 2008). That this inverse relationship was highly significant in anthropogenic populations but not in wild populations supports the hypothesis that the more northern populations resulted from anthropogenic dispersal into habitat at *A. triloba*'s ecological limit. The low fruit set documented in these northern populations is consistent with low rates of gene flow and possible marginal ecological conditions for *A. triloba*. White (1906) also reported that flower and fruit production in *A. triloba* decreased with increasing latitude north of the Ohio River mouth (Cairo, Illinois) and that its vegetative growth at the northern edge of its contemporary range is beyond its fruiting limit. Murphy (2001) further mentions the reduced size of *A. triloba* in northern populations, although he

attributes this to possible overgrazing by a deer population that has been unnaturally large since 1912 (Bartlett, 1958).

Compelling evidence that putative anthropogenic and wild populations were shaped by different processes comes from the finding of significantly higher mean pairwise  $G_{ST}$  values among anthropogenic populations despite their smaller pairwise geographic distances relative to wild populations. Some of the most interesting data come from the nine northernmost anthropogenic populations that occur in formerly glaciated landscapes. Although separated from one another by a mean of 317 km, they shared their lowest pairwise  $G_{ST}$  values with wild populations at a mean distance of 865 km. This suggests that the wild populations with which they shared their lowest  $G_{ST}$  values may have served as seed sources and that subsequent to long-distance dispersal and colonization, there has been little gene flow among these northern populations (Hewitt, 1999). The four northernmost anthropogenic populations and their most genetically similar wild populations are illustrated in Figure 5.

The finding of significant IBD in all 82 populations and all wild populations is consistent with findings for about half of angiosperms surveyed ( $N = 71$ ; Perez et al., 2018). The absence of significant IBD in anthropogenic populations however is more noteworthy and is consistent with their hypothesized initiation by propagules from different and geographically more distant wild populations, as well as limited gene flow.

The two genetic clusters identified by STRUCTURE analysis of all 82 populations are consistent with results for wild populations only (Wyatt et al., 2021), with clusters primarily located on either side of the Appalachian Mountains and separated by an east-west phylogeographic discontinuity along the Tombigbee drainage in Alabama (Figure 2). However, some anthropogenic populations (e.g., FCP<sub>A</sub>, OSF<sub>A</sub>, SEL<sub>A</sub>, and SU1<sub>A</sub>) are located outside the geographic boundaries of their inferred genetic cluster, which is consistent with long-distance human-mediated dispersal.

Alleles that occur in only a small proportion of populations can be useful for inferring patterns of dispersal. A rare allele in an anthropogenic population possessing low levels of genetic variation and located a considerable distance from the nearest wild population possessing that allele suggests long-distance dispersal. Cases where the anthropogenic population occurs in habitat that was glaciated during the Late Wisconsin Glaciation suggest that dispersal may have been mediated by human activity as the megafauna potentially capable of such long-distance dispersal of *A. triloba* went extinct ~12,000 YBP (Janzen & Martin, 1982; Perrotti, 2018; Poor, 1984). Our data show a pattern of putative anthropogenic populations occurring well north of the wild populations with which they share rare alleles. While alleles identical in size and occurring in such widely separated populations may reflect homoplasy, the low  $G_{ST}$  values between geographically distant anthropogenic and wild populations as well as reduced genetic diversity of putative anthropogenic populations lend support to the hypothesis that human-mediated long-distance dispersal was responsible for the founding of these anthropogenic populations.

Some of the most compelling evidence comes from the pairwise  $G_{ST}$  values and geographic distances between populations sharing rare alleles. As predicted, wild populations that share rare alleles ( $W_{RA}$ ) had significantly lower pairwise  $G_{ST}$  values and are separated by significantly smaller distances than all wild populations (W; Table 4). However, wild and anthropogenic populations that shared rare alleles,  $(W - A)_{RA}$ , are separated by significantly larger distances and had significantly higher pairwise  $G_{ST}$  values than  $W_{RA}$  (Table 4). Anthropogenic populations in formerly glaciated landscapes and the wild populations with which they share rare alleles,  $(A_G - W)_{RA}$  are separated by the largest mean geographic distance investigated (Table 4).

Subsequent to extinction of North American megafauna and the LGM, the primary nonanthropogenic candidates for dispersal are waterways, raccoons, opossums, turkeys, red foxes, and coyotes. However, hydrochory cannot explain the northward migration of *A. triloba*. Many of the anthropogenic populations occur west of the Appalachians along the Mississippi/Ohio river drainages where the directionality of water flow is from the northeast to the southwest. While this makes hydrochory an unlikely mechanism of northward dispersal, the distribution of anthropogenic populations is consistent with known pre-Columbian trade routes (e.g., Forsberg, 2003; Muñoz et al., 2014). Major rivers east of the Appalachians flow from west to east or northwest to southeast, also precluding the possibility of northward dispersal by hydrochory. While raccoons consume pawpaw, they are unlikely to mediate long-distance dispersal due to their generally small home ranges of 1–3 km (maximum = 6.4 km; Stuewer, 1943; Butterfield, 1944; Feldhamer et al., 1982). Opossums are also unlikely long-distance seed dispersers with average home ranges of 64.4 ha and 141.6 ha for females and males, respectively, and average nightly movements of 1,465 m (females) and 1,835 m (males; Ryser, 1995). Wild turkeys have annual home ranges of 150 to 550 ha but they are nonmigratory (National Wild Turkey Federation, 2021). Furthermore, turkeys that consume pawpaw appear to destroy the seeds (Wyatt, personal observation). Red foxes (*Vulpes vulpes*) have larger home ranges; however, native red foxes are a boreal species whose distribution prior to European settlement overlapped very little with the postglacial distribution of pawpaw (Kamler & Ballard, 2002). Non-native red foxes from Europe were introduced into the eastern United States and expanded westwards in the early 1900s (Kamler & Ballard, 2002). Neither the distribution of native red foxes nor the migration of introduced red foxes are consistent with northward range expansion of *A. triloba*. Coyotes are the only other animals known to consume pawpaw that are capable of long-distance movement; however, their distribution during the Pleistocene, early Holocene, pre-Columbian period, and the early 1600s was confined to southern Mexico and Central America (Hidalgo-Mihart et al., 2004). It is only in the last century that coyotes have expanded their range to overlap with the distribution of pawpaw (Hody & Kays, 2018). Thus, while hydrochory and zoochory have certainly contributed to *A. triloba* migration and range shifts, our data strongly suggest that humans played a role in long-distance dispersal.

In summary, this study represents the first investigation into human-mediated dispersal of a useful understory tree species in Eastern North America using genetic markers. Perhaps the most salient finding was that nearly all genetic diversity statistics are higher, although not significantly so, in wild populations than putative anthropogenic populations, except  $H_D$  which is significantly lower in wild populations. Lower genetic diversity in anthropogenic populations of *A. triloba* is likely associated with founder events and dispersal of seeds or ramets from wild populations to areas isolated and/or less ecologically suitable. The high levels of genetic structure observed in northern anthropogenic populations suggest independent dispersal and colonization events from southern source populations and low levels of subsequent gene flow between these populations, which is consistent with the expectation of human-mediated dispersal into northern habitats.

As we learn more about movement and trade practices of indigenous people in Eastern North America, it has become increasingly evident that they were exchanging goods over distances of 1,000 km as early as the Archaic Period (~8,000–3,000 YBP; Sanger et al., 2018). The preponderance of anthropogenic populations along major river valleys is consistent with known trade corridors and continental-scale exchange along large mid-continental rivers including the Mississippi and Ohio (e.g., Fowler, 1975; Smith, 2010). Given the active and widespread trade practices of indigenous peoples of Eastern North America as well as the duration of such practices, human-mediated dispersal is a viable explanation for the patterns of genetic variation we detected.

Populations of *A. triloba* become very patchy and infrequent north of 40°N (Keener & Kuhns, 1998; Wyatt, personal observation). Regardless of the dispersal vector, the high  $G_{ST}$  values among the northernmost *A. triloba* populations are indicative of multiple separate introductions from genetically distinct source populations followed by historically low rates of gene flow. The low fruit set documented in these northern populations is consistent with little sexual reproduction, which may be related to the extent of clonality in these populations and self-incompatibility (Wyatt et al., personal observation) as well as low rates of gene flow. Low gene flow in turn may result from the isolation of populations. It is also possible that these populations exist at or beyond their ecological limit, and/or that of their pollinators, such that pollination is limited and populations persist primarily through asexual reproduction. Excellent examples of this phenomenon come from *Agave delamateri* and *A. murpheyi* (Parker et al., 2007), which are believed to have been introduced into Arizona by pre-Columbian peoples because of their value as a reliable source of food, fiber, and medicine. Both species are associated with indigenous settlements that were abandoned centuries ago due to a drying climate and have persisted through clonal spreading with no evidence of sexual reproduction (Parker et al., 2007). It is thought that Arizona is outside the normal range of these species, precluding sexual reproduction. While the low pawpaw fruit set we documented represents a snapshot of one reproductive season, it is consistent with the complete absence of *A. triloba* seeds in refuse middens from Archaic sites around Lake Erie

and northward (Keener & Kuhns, 1998). Such absence is consistent with the arrival of *A. triloba* at these locations more recently than 3,000 YBP and/or with the persistence of these populations primarily through asexual reproduction. This work elucidates the possible origins of populations of a widespread Eastern North American tree species at the northern limit of its range. Furthermore, closer examination of populations of *A. triloba* that are designated as wild for genetic profiles more similar to anthropogenic populations may reveal hitherto undetected locations of pre-Columbian settlements and trade routes, thus furthering our understanding of indigenous cultures (e.g., Parker et al., 2014).

## ACKNOWLEDGMENTS

This research was supported by an award to DWT by the University of Georgia Office of the Vice President for Research.

## CONFLICT OF INTEREST

The authors declare no conflict of interest.

## AUTHOR CONTRIBUTIONS

**Graham E. Wyatt:** Conceptualization (supporting); Data curation (lead); Formal analysis (equal); Investigation (lead); Methodology (equal); Visualization (lead); Writing-original draft (equal). **J. L. Hamrick:** Conceptualization (supporting); Methodology (equal); Supervision (supporting); Writing-original draft (supporting); Writing-review & editing (equal). **Dorset W. Trapnell:** Conceptualization (lead); Data curation (supporting); Formal analysis (equal); Funding acquisition (lead); Methodology (lead); Project administration (lead); Resources (lead); Supervision (lead); Writing-original draft (equal); Writing-review & editing (lead).

## DATA AVAILABILITY STATEMENT

Genetic data are accessible at Dryad (<https://doi.org/10.5061/dryad.5x69p8d3g>).

## ORCID

Dorset W. Trapnell  <https://orcid.org/0000-0001-7184-4700>

## REFERENCES

- Abrams, M. D., & Nowacki, G. J. (2008). Native Americans as active and passive promoters of mast and fruit trees in the eastern USA. *The Holocene*, 18, 1123–1137. <https://doi.org/10.1177/0959683608095581>
- Adovasio, J. M., Gunn, J. D., Donahue, J., & Stuckenrath, R. (1978). Meadowcroft Rockshelter, 1977: An overview. *American Antiquity*, 43, 632–651. <https://doi.org/10.2307/279496>
- Asch, D. L. (1995). *Aboriginal specialty-plant propagation: Illinois prehistory and an eastern North American post-contact perspective*, Unpublished doctoral dissertation. University of Michigan.
- Bainbridge, D. A. (1985). The rise of agriculture: A new perspective. *Ambio*, 14, 148–151.
- Ballou, F., Lehmann, L., & de Meeùs, T. (2003). The population genetics of clonal and partially clonal diploids. *Genetics*, 164, 1635–1644. <https://doi.org/10.1093/genetics/164.4.1635>
- Bartlett, C. O. (1958). *A study of some deer and forest relationships in Rondeau Provincial Park*. Ontario Department of Lands and Forests.
- Bedell, J., Katz, G., Shellenhamer, J., Kraus, L., & Groesbeck, S. (2011). *The people of the mountain: archeological overview, assessment, identification, and evaluation study of Catocin Mountain Park, Maryland Vol 1. Final Report*. Prepared by: Louis Berger Group Inc. Washington, DC. Prepared for: National Park Service, Washington, DC.
- Bennett, B. A. (1986). *Geology of the Mallory Site, Fredonia, New York*, Master's thesis, Department of Anthropology and Sociology, SUNY College, Fredonia, NY.
- Bentancourt, J. L., Schuster, W. S., Mitton, J. B., & Anderson, R. S. (1991). Fossil and genetic history of a pinyon pine (*Pinus edulis*) isolate. *Ecology*, 72, 1685–1697. <https://doi.org/10.2307/1940968>
- Billingham, M. R., Reusch, T. B. H., Alberto, F., & Serrao, E. A. (2003). Is asexual reproduction more important at geographical limits? A genetic study of the seagrass *Zostera marina* in the Ria Formosa, Portugal. *Marine Ecology Progress Series*, 265, 77–83. <https://doi.org/10.3354/meps265077>
- Birky, C. W. Jr (1996). Heterozygosity, heteromorphy, and phylogenetic trees in asexual eukaryotes. *Genetics*, 144, 427–437. <https://doi.org/10.1093/genetics/144.1.427>
- Bitocchi, E., Bellucci, E., Giardini, A., Rau, D., Rodriguez, M., Biagetti, E., Santilocchi, R., Spagnoletti, Z., Gioia, T., Logozzo, G., Attene, G., Nanni, L., & Papa, R. (2012). Molecular analysis of the parallel domestication of the common bean (*Phaseolus vulgaris*) in Mesoamerica and the Andes. *New Phytologist*, 197, 300–313.
- Blasing, R. (1986). *Archaeological survey of the Upper Deep Creek Drainage, Kansas*. Department of Anthropology Report, Wichita State University, Wichita, KS.
- Bowden, W. M., & Miller, B. (1951). Distribution of the Pawpaw, *Asimina triloba* (L.) Dunal, in southern Ontario. *The Canadian Field Naturalist*, 65, 27–31.
- Brookfield, J. F. Y. (1996). A simple new method for estimating null allele frequency from heterozygote deficiency. *Molecular Ecology*, 5, 453–455. <https://doi.org/10.1111/j.1365-294X.1996.tb00336.x>
- Brooks, R. R., & Johannes, D. (1990). *Historical, ethno-, and economic botany series*, Vol. 3. Dioscorides Press.
- Brose, D. A. (1976). Conneaut Fort, a prehistoric Whittlesey Focus Village in Ashtabula County, OH. *Pennsylvania Archaeologist*, 46, 29–77.
- Bryzski, J. R., & Culley, T. M. (2011). Genetic variation and clonal structure of the rare, riparian shrub *Spiraea virginiana* (Rosaceae). *Conservation Genetics*, 12, 1323–1332. <https://doi.org/10.1007/s10592-011-0233-x>
- Butler, B. M. (1981). Sellars: A small mound center in the Hinterlands. *Tennessee Anthropologist*, 6, 37–60.
- Butterfield, R. T. (1944). Populations, hunting pressure, and movement of Ohio raccoons. *Transactions, 9th North American Wildlife Conference*, 9, 337–344.
- Cai, Z., Gold, M., & Brannan, R. (2019). An exploratory analysis of US consumer preferences for North American pawpaw. *Agroforest Systems*, 93, 1673–1685. <https://doi.org/10.1007/s10457-018-0296-5>
- Carr, C. (2008). Chapter 2: Environmental setting, natural symbols, and subsistence. In D. T. Case, & C. Carr (Eds.), *The Scioto Hopewell and their neighbors: Bioarchaeological documentation and cultural understanding* (pp. 122–128). Springer.
- Clark, J. S., Fastie, C., Hurr, G., Jackson, S. T., Johnson, C., King, G. A., Lewis, M., Lynch, J., Pacala, S., Prentice, C., Schupp, E. W., Webb, T. III, & Wyckoff, P. (1998). Reid's Paradox of rapid plant migration. *BioScience*, 48, 13–24. <https://doi.org/10.2307/1313224>
- Clark, P. U., Dyke, A. S., Shakun, J. D., Carlson, A. E., Clark, J., Wohlfarth, B., Mitrovica, J. X., Hostetler, S. W., & McCabe, A. M. (2009). The last glacial maximum. *Science*, 325, 710–714. <https://doi.org/10.1126/science.1172873>
- Comps, B., Gömöry, D., Letouzey, J., Thiébaud, B., & Petit, R. J. (2001). Diverging trends between heterozygosity and allelic richness during postglacial colonization in the European Beech. *Genetics*, 157, 389–397. <https://doi.org/10.1093/genetics/157.1.389>

- Connolly, R. (2004). Time, space, and function at Fort Ancient. In R. Connolly, & B. T. Lepper (Eds.), *The Fort Ancient earthworks: Prehistoric lifeways of the Hopewell culture in Southwestern Ohio*. Ohio Historical Society.
- Cottrell, J. E., Munro, R. C., Tabbener, H. E., Gillies, A. C. M., Forrest, G. I., Deans, J. D., & Lowe, A. J. (2002). Distribution of chloroplast DNA variation in British oaks (*Quercus robur* and *Q. petraea*): The influence of postglacial colonization and human management. *Forest Ecology and Management*, *156*, 181–195.
- Cutler, H. C., & Blake, L. W. (1976). *Plants from archaeological sites east of the Rockies*. Missouri Archaeological Society, 15. University of Missouri, Columbia, Missouri. Printed in microfiche L.C. No. 76 - 49422 (Corrected and updated version of the 1973 edition).
- Cypher, B. L., & Cypher, E. A. (1999). Germination rates of tree seeds ingested by coyotes and raccoons. *American Midland Naturalist*, *142*, 71–76.
- Darrow, G. M. (1975). Minor temperate fruits. In J. Janick, & J. N. Moore (Eds.), *Advances in fruit breeding* (pp. 276–277). Purdue University Press.
- Davis, M. B., & Shaw, R. G. (2001). Range shifts and adaptive responses to Quaternary climate change. *Science*, *292*, 673–679. <https://doi.org/10.1126/science.292.5517.673>
- Delcourt, P. A., & Delcourt, H. R. (2004). *Prehistoric native Americans and ecological change: Human ecosystems in Eastern North America since the Pleistocene* (pp. 42, 76, 167–168). Cambridge University Press.
- Denevan, W. M. (1992). The pristine myth: The landscape of the Americas in 1492. *Annals of the Association of American Geographers*, *82*, 369–385. <https://doi.org/10.1111/j.1467-8306.1992.tb01965.x>
- Denevan, W. M. (2011). The “pristine myth” revisited. *The Geographical Review*, *101*, 576–591. <https://doi.org/10.1111/j.1931-0846.2011.00118.x>
- Dorken, M. E., & Eckert, C. G. (2001). Severely reduced sexual reproduction in northern populations of a clonal plant, *Decodon verticillatus* (Lythraceae). *Journal of Ecology*, *89*, 339–350.
- Doyle, J. J., & Doyle, J. L. (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin*, *19*, 11–15.
- Earl, D. A., & vonHoldt, B. M. (2012). Structure Harvester: A website and program for visualizing structure output and implementing the Evanno method. *Conservation Genetics Resources*, *2*, 359–361. <https://doi.org/10.1007/s12686-011-9548-7>
- Eckert, C. G. (2002). The loss of sex in clonal plants. *Evolutionary Ecology*, *15*, 501–520. <https://doi.org/10.1023/A:1016005519651>
- Erlandson, J. M., Braje, T. J., Gill, K. M., & Graham, M. H. (2015). Ecology of the kelp highway: Did marine resources facilitate human dispersal from Northeast Asia to the Americas? *Journal of Island and Coastal Archaeology*, *10*, 392–411. <https://doi.org/10.1080/15564894.2014.1001923>
- Evanno, G. S., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Molecular Ecology*, *15*, 2611–2620. <https://doi.org/10.1111/j.1365-294X.2005.02553.x>
- Feldhamer, G. A., Thompson, B. C., & Chapman, J. A. (1982). *Wild mammals of North America*. The Johns Hopkins University Press.
- Fields, S. E., McDonald, H. G., Knight, J. L., & Sanders, A. E. (2012). The ground sloths (*Pilosa*) of South Carolina. *PalArch's Journal of Vertebrate Palaeontology*, *9*, 1–19.
- Forsberg, D. M. (2003). Native American life in the MNRRA corridor. In J. O. Anfanson (Ed.), *River of history – A historic resources study of the Mississippi National River and recreation area*. National Park Service and U.S. Army Corps of Engineers, St. Paul District.
- Fowler, M. L. (1975). A pre-Columbian urban center on the Mississippi. *Scientific American*, *233*, 92–101. <https://doi.org/10.1038/scientificamerican0875-92>
- Freeman, C. C., & Hulbert, L. C. (1985). An annotated list of the vascular flora of Konza Prairie Research Natural Area, Kansas. *Transactions of the Kansas Academy of Science*, *88*, 84–115. <https://doi.org/10.2307/3627880>
- Fritz, G. J. (1990). Multiple pathways to farming in precontact Eastern North America. *Journal of World Prehistory*, *4*, 387–435. <https://doi.org/10.1007/BF00974813>
- Garcia, D., Zamora, R., Gomez, J. M., Jordano, P., & Hódar, J. A. (2000). Geographical variation in seed production, predation and abortion in *Juniperus communis* throughout its range in Europe. *Journal of Ecology*, *88*, 436–466. <https://doi.org/10.1046/j.1365-2745.2000.00459.x>
- Gardner, W. M. (1966). The Waddell's Mill Pond Site. *Florida Anthropologist*, *19*, 43–64.
- Gilmore, M. R. (1931). Vegetal remains of the Ozark bluff-dweller culture. *Papers of the Michigan Academy of Science, Arts, and Letters*, *14*, 83–102.
- Gonzales, E., Hamrick, J. L., & Smouse, P. E. (2008). Comparison of clonal diversity in mountain and piedmont populations of *Trillium cuneatum* (Melanthiaceae-Trilliaceae), a forest understory species. *American Journal of Botany*, *95*, 1254–1261.
- Goodrich, K. R., Zjhra, M., Ley, C. A., & Raguso, R. A. (2006). When flowers smell fermented: The chemistry and ontogeny of yeasty floral scent in pawpaw (*Asimina triloba*: Annonaceae). *International Journal of Plant Sciences*, *167*, 33–46.
- Goudet, J. (1995). FSTAT (Version 1.2): A computer program to calculate F-statistics. *Journal of Heredity*, *86*, 485–486. <https://doi.org/10.1093/oxfordjournals.jhered.a111627>
- Goudet, J. (2001). *FSTAT, a program to estimate and test gene diversities and fixation indices, version 2.9.3*. Retrieved from <http://www2.unil.ch/popgen/softwares/fstat/htm>
- Hauswaldt, S. J., & Glenn, T. C. (2003). Microsatellite DNA loci from the diamondback terrapin (*Malaclemys terrapin*). *Molecular Ecology Notes*, *3*, 174–176. <https://doi.org/10.1046/j.1471-8286.2003.00388.x>
- Hewitt, G. M. (1999). Post-glacial recolonization of European biota. *Biological Journal of the Linnean Society*, *68*, 87–112. <https://doi.org/10.1111/j.1095-8312.1999.tb01160.x>
- Hewitt, G. M. (2000). The genetic legacy of the Quaternary ice ages. *Nature*, *405*, 907–913. <https://doi.org/10.1038/35016000>
- Hidalgo-Mihart, M. G., Cantú-Salazar, L., González-Romero, A., & López-González, C. A. (2004). Historical and present distribution of coyote (*Canis latrans*) in Mexico and Central America. *Journal of Biogeography*, *31*, 2025–2038. <https://doi.org/10.1111/j.1365-2699.2004.01163.x>
- Higgins, T. F. III, Downing, C. M., Linebaugh, D. W., Opperman, A. F., & Turner, R. E. III (1995). *Archaeological investigations of site 44CF7, Falling Creek ironworks, and vicinity Chesterfield County, Virginia*. William and Mary Center for Archaeological Research, Department of Anthropology, College of William and Mary, Williamsburg, Virginia.
- Hinsdale, W. B. (1931). *Archaeological Atlas of Michigan*. University of Michigan Press.
- Hody, J. W., & Kays, R. (2018). Mapping the expansion of coyotes (*Canis latrans*) across North and Central America. *ZooKeys*, *759*, 81–97. <https://doi.org/10.3897/zookeys.759.15149>
- Hoppe, K., & Koch, P. L. (2007). Reconstructing the migration patterns of late Pleistocene mammals from northern Florida, USA. *Quaternary Research*, *68*, 347–352. <https://doi.org/10.1016/j.yqres.2007.08.001>
- Hormaza, J. I. (2004). The pawpaw, a forgotten North American fruit tree. *Arnoldia*, *72*, 13–23.
- Hosaka, N., Gómez, S., Kachi, N., Stuefer, J. F., & Whigham, D. F. (2005). The ecological significance of clonal growth in the understory tree, Pawpaw (*Asimina triloba*). *Northeastern Naturalist*, *12*, 11–22.
- Janzen, D. H., & Martin, P. S. (1982). Neotropical anachronisms: The fruits the Gomphotheres ate. *Science*, *215*, 19–27. <https://doi.org/10.1126/science.215.4528.19>
- Jones, V. H. (1936). The vegetal remains of Newt Kash Hollow Shelter. In W. S. Webb, & W. D. Funkhouser (Eds.), *Rockshelters in Menifee County, Kentucky* (pp. 147–165). Reports in Anthropology and Archaeology 4, University of Kentucky.

- Kamler, J. F., & Ballard, W. B. (2002). A review of native and nonnative red foxes in North America. *Wildlife Society Bulletin*, 30(2), 370–379.
- Kandle, P. L., Downing, C. M., Linebaugh, D. W., & Deitrick, V. (1996). *Archaeological investigations at site 44CF102 Osbornes/Gatesville Town Chesterfield County, Virginia*. William and Mary Center for Archaeological Research, Department of Anthropology, College of William and Mary, Williamsburg, Virginia.
- Kauffman, M. J., Cagnacci, F., Chamailé-Jammes, S., Hebblewhite, M., Hopcraft, J. G. C., Merkle, J. A., Mueller, T., Mysterud, A., Peters, W., Roettger, C., Steingisser, A., Meacham, J. E., Abera, K., Adamczewski, J., Aikens, E. O., Bartlam-Brooks, H., Bennitt, E., Berger, J., Boyd, C., ... Zuther, S. (2021). Mapping out a future for ungulate migrations. *Science*, 372, 566–569. <https://doi.org/10.1126/science.abf0998>
- Kay, C. E. (2002). Are ecosystems structured from the top-down or bottom-up? In C. E. Kay, & R. T. Simmons (Eds.), *Wilderness and political ecology: Aboriginal influences and the original state of nature* (pp. 484–498). University of Utah Press.
- Keener, C., & Kuhns, E. (1998). The impact of Iroquoian populations on the northern distribution of pawpaws in the Northeast. *North American Archaeologist*, 18, 327–342. <https://doi.org/10.2190/ATOW-VEDT-OEOP-W21V>
- King, F. B. (1982). Archaeobotanical remains from Rodgers Shelter. In M. Kay (Ed.), *Holocene adaptations within the Lower Pomme de Terre River Valley, Missouri, Volume 1* (pp. 701–728). Illinois State Museum, Springfield, submitted to the U.S. Army Corps of Engineers, Kansas City District.
- Krech, S. (1999). *The ecological Indian: Myth and history*. W.W. Norton & Company.
- Krochmal, A., & Krochmal, C. (1973). *A guide to the medicinal plants of the U.S.* New York Times Books Company.
- Lagrange, R. L., & Tramer, E. J. (1985). Geographic variation in size and reproductive success in the paw paw (*Asimina triloba*). *The Ohio Journal of Science*, 85, 40–45.
- Larranaga, N., Albertazzi, F. J., Fontecha, G., Palmieri, M., Rainer, H., van Zonneveld, M., & Hormaza, J. I. (2017). A Mesoamerican origin of cherimoya (*Annona cherimola* Mill.): Implications for the conservation of plant genetic resources. *Molecular Ecology*, 26, 4116–4130.
- Levis, C., Costa, F. R. C., Bongers, F., Peña-Claros, M., Clement, C. R., Junqueira, A. B., Neves, E. G., Tamañá, E. K., Figueiredo, F. O. G., Salomão, R. P., Castilho, C. V., Magnusson, W. E., Phillips, O. L., Guevara, J. E., Sabatier, D., Molino, J.-F., López, D. C., Mendoza, A. M., Pitman, N. C. A., ... ter Steege, H. (2017). Persistent effects of pre-Columbian plant domestication on Amazonian forest composition. *Science*, 355, 925–931. <https://doi.org/10.1126/science.aal0157>
- Li, C. C., & Horvitz, D. G. (1953). Some methods of estimating the inbreeding coefficient. *American Journal of Human Genetics*, 5, 107–117.
- Little, E. L. Jr. (1977). *Atlas of the United States trees, Volume 4: minor eastern hardwoods*. Pub. 1342. US Department of Agriculture. 17 p 230 maps.
- Logan, W. B. (2005). *Oak: The frame of civilization*. W.W. Norton & Company.
- Lovell, W. G. (1992). Heavy shadows and black night: Disease and depopulation in colonial Spanish America. *Annals of the Association of American Geographers*, 82, 426–443. <https://doi.org/10.1111/j.1467-8306.1992.tb01968.x>
- Lovis, W. A. (1974). The Moccasin Bluff Site and the woodland cultures of Southwestern Michigan. *American Anthropologist*, 76, 463–464.
- Lu, L., Pomper, K. W., Lowe, J. D., & Crabtree, S. B. (2011). Genetic variation in pawpaw cultivars using microsatellite analysis. *Journal of the American Society of Horticultural Science*, 136, 415–521. <https://doi.org/10.21273/JASHS.136.6.415>
- Lyford, M. E., Jackson, S. T., Betancourt, J. L., & Gray, S. T. (2003). Influence of landscape structure and climate variability on a late Holocene plant migration. *Ecological Monographs*, 73, 567–583. <https://doi.org/10.1890/03-4011>
- MacDougall, A. (2003). Did Native Americans influence the northward migration of plants during the Holocene? *Journal of Biogeography*, 30, 633–647. <https://doi.org/10.1046/j.1365-2699.2003.00842.x>
- MacNamara, J. M. (1982). *Summary of the 1981 excavations at the Conowingo site (18CE14)*. Maryland Geological Survey, Division of Archaeology File Report 172. Baltimore, MD.
- Meloni, M., Reid, A., Caujapé-Castells, J., Marrero, Á., Fernández-Palacios, J. M., Mesa-Coelo, R. A., & Conti, E. (2013). Effects of clonality on the genetic variability of rare, insular species: The case of *Ruta microcarpa* from the Canary Islands. *Ecology and Evolution*, 6, 1569–1579.
- Miller, A. J., & Schaal, B. A. (2006). Domestication and the distribution of genetic variation in wild and cultivated populations of the Mesoamerican fruit tree *Spondias purpurea* L. (Anacardiaceae). *Molecular Ecology*, 15, 1467–1480. <https://doi.org/10.1111/j.1365-294X.2006.02834.x>
- Morse, D. E., & Morse, P. A. (1983). *Archaeology of the Central Mississippi Valley*. Academic Press.
- Motamayor, J. C., Risterucci, A. M., Lopez, P. A., Ortiz, C. F., Moreno, A., & Lanaud, C. (2002). Cacao domestication I: The origin of the cacao cultivated by the Mayas. *Heredity*, 89, 380–386. <https://doi.org/10.1038/sj.hdy.6800156>
- Mueller, N. G. (2017). Documenting domestication in a lost crop (*Polygonum erectum* L.): Evolutionary bet-hedgers under cultivation. *Vegetation History and Archaeobotany*, 26, 313–327. <https://doi.org/10.1007/s00334-016-0592-9>
- Muñoz, S. E., Mladenoff, D. J., Schroeder, S., & Williams, J. W. (2014). Defining the spatial patterns of historical land use associated with the indigenous societies of eastern North America. *Journal of Biogeography*, 41, 2195–2210. <https://doi.org/10.1111/jbi.12386>
- Murphy, J. L. (2001). Pawpaws, persimmons, and 'possums: On the natural distribution of pawpaws in the northeast. *North American Archaeologist*, 22, 93–115. <https://doi.org/10.2190/TT5Y-YAXJ-OKAU-GUFG>
- National Wild Turkey Federation. (2021). *Wild Turkey behavior*. Retrieved from <https://www.nwtf.org/hunt/wild-turkey-basics/behavior>
- Nei, M. (1973). Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences of the United States of America*, 70, 3321–3323. <https://doi.org/10.1073/pnas.70.12.3321>
- Paetkau, D., & Strobeck, C. (1994). Microsatellite analysis of genetic variation in black bear populations. *Molecular Ecology*, 3, 489–495. <https://doi.org/10.1111/j.1365-294X.1994.tb00127.x>
- Pagoulatos, P. (1992). Native American land-use patterns of New Jersey: Some testable hypotheses. *Journal of Middle Atlantic Archaeology*, 8, 57–77.
- Parker, K. C. (2002). Fire in the pre-European lowlands of the American Southwest. In T. R. Vale (Ed.), *Fire, native peoples, and the natural landscape* (pp. 101–142). Island Press.
- Parker, K. C., Hamrick, J. L., Hodgson, W. C., Trapnell, D. W., Parker, A. J., & Kuzoff, R. K. (2007). Genetic consequences of pre-Columbian cultivation for *Agave murpheyi* and *A. delamateri* (Agavaceae). *American Journal of Botany*, 94, 1479–1490.
- Parker, K. C., Trapnell, D. W., Hamrick, J. L., & Hodgson, W. C. (2014). Genetic and morphological contrasts between wild and anthropogenic populations of *Agave parryi* var. *huachucensis* in south-eastern Arizona. *Annals of Botany*, 113, 939–952. <https://doi.org/10.1093/aob/mcu016>
- Parker, P. G., Snow, A. A., Schug, M. D., Booton, G. C., & Fuerst, P. A. (1998). What molecules can tell us about populations: Choosing and using a molecular marker. *Ecology*, 79, 361–382. <https://doi.org/10.2307/176939>
- Patterson, W. A. III, & Sassaman, K. E. (1988). Indian fires in the prehistory of New England. In G. P. Nicholas (Ed.), *Holocene human ecology in northeastern North America* (pp. 107–135). Plenum Publishing.



- Peakall, R., & Smouse, P. E. (2006). GenAlEx 6: Genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, 6, 288–295.
- Peakall, R., & Smouse, P. E. (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>
- Peattie, D. C. (1950). *A natural history of trees of eastern and central North America*. Houghton Mifflin Co.
- Perez, M. F., Franco, F. F., Bombonato, J. R., Bonatelli, I. A. S., Khan, G., Romeiro-Brito, M., Fegies, A. C., Ribeiro, P. M., Silva, G. A. R., & Moraes, E. M. (2018). Assessing population structure in the face of isolation by distance: Are we neglecting the problem? *Diversity and Distributions*, 24, 1883–1889.
- Perrault, S. L., Saucier, R. T., Heinrich, P., Lee, D., Ryan, J., & Hahn, T. H. G. III (2008). *Phase I cultural resource survey of the Lake Killarney ecosystem restoration project, West Feliciana Parish, Louisiana*. Final report prepared by Coastal Environments, Inc. for U.S. Army Corps of Engineers, New Orleans District.
- Perrotti, A. G. (2018). Pollen and *Sporomiella* evidence for the terminal Pleistocene vegetation change and megafaunal extinction at Page-Ladson, Florida. *Quaternary International*, 466, 256–268.
- Peterson, R. N. (1991). Pawpaw (*Asimina*). *Acta Horticulturae*, 290, 567–600. <https://doi.org/10.17660/ActaHortic.1991.290.13>
- Petit, R. J., Aguinagalde, I., de Beaulieu, J., Bittaku, C., Brewer, S., Cheddadi, R., & Vendramin, G. G. (2003). Glacial refugia: Hotspots but not melting pots of genetic diversity. *Science*, 300, 1563–1565. <https://doi.org/10.1126/science.1083264>
- Petit, R. J., Brewer, S., Bordács, S., Burg, K., Cheddadi, R., Coart, E., Cottrell, J., Csai, U. M., van Dam, B., Deans, J. D., Espinel, S., Fineschi, S., Finkeldey, R., Glaz, I., Goicoechea, P. G., Jensen, J. S., König, A. O., Lowe, A. J., Madsen, S. F., ... Kremer, A. (2002). Identification of refugia and post-glacial colonisation routes of European white oaks based on chloroplast DNA and fossil pollen evidence. *Forest Ecology and Management*, 156, 49–74. [https://doi.org/10.1016/S0378-1127\(01\)00634-X](https://doi.org/10.1016/S0378-1127(01)00634-X)
- Phelps, D. S. (1983). The archaeology of the North Carolina coast and coastal plain: Problems and hypotheses. In M. A. Mathis, & J. J. Crow (Eds.), *The prehistory of North Carolina: An archaeological symposium* (pp. 1–51). North Carolina Division of Archives and History.
- Pickering, C. (1879). *Chronological history of plants* (Vol. 1). Little, Brown and Company.
- Pomper, K. W., Lowe, J. D., Lu, L., Crabtree, S. B., Dutta, S., Schneider, K., & Tidwell, J. (2010). Characterization and identification of pawpaw cultivars and advanced selections by simple sequence repeat markers. *Journal of the American Society for Horticulture Science*, 135, 143–149. <https://doi.org/10.21273/JASHS.135.2.143>
- Poor, J. M. (1984). *Plants that merit attention, Vol. 1: Trees*. Timber Press.
- Pritchard, J. K., Stephens, J. K., & Donnelly, M. (2000). Inference of population structure using multilocus genotype data. *Genetics*, 155, 945–959. <https://doi.org/10.1093/genetics/155.2.945>
- Prufer, O. H. (1964). New data on Hopewell hilltop enclosures in Southern Ohio. In W. Dancey, & P. J. Pacheco (Eds.), *Ohio hopewell community organization* (pp. 311–326). Kent State University Press.
- Pyne, S. J. (1983). Indian fires. *Natural History*, 2, 6–11.
- Raymond, M., & Rousset, F. (1995). GENEPOP (version 1.2): Population genetics software for exact tests and ecumenicism. *Journal of Heredity*, 86, 248–249. <https://doi.org/10.1093/oxfordjournals.jhered.a111573>
- Richland, D. M., Rundel, P. W., Jackson, S. T., Teskey, R. O., Aronson, J., Bytnerowicz, A., Wingfield, M. J., & Proches, S. (2007). Human impacts in pine forests: Past, present, and future. *Annual Review of Ecology, Evolution, and Systematics*, 38, 275–297.
- Robertson, C. (1928). *Flowers and insects: Lists of visitors of four hundred and fifty-three flowers*. Self-published, IL.
- Rousset, F. (1997). Genetic differentiation and estimation of gene flow from *F*-statistics under isolation by distance. *Genetics*, 145, 1219–1228.
- Ryser, J. (1995). Activity, movement and home range of Virginia opossum (*Didelphis virginiana*) in Florida. *Bulletin of the Florida Museum of Natural History*, 38, 177–194.
- Sanger, M. C., Hill, M. A., Lattanzi, G. D., Padgett, B. D., Larsen, C. S., Culleton, B. J., Kennett, D. J., Dussubieux, L., Napolitano, M. F., Lacombe, S., & Thomas, D. H. (2018). Early metal use and crematory practices in the American Southeast. *Proceedings of the National Academy of Sciences*, 115, 7672–7679. <https://doi.org/10.1073/pnas.1808819115>
- Scarry, M. C., & Scarry, J. F. (2005). Native American 'garden agriculture' in southeastern North America. *World Archaeology*, 37, 259–274. <https://doi.org/10.1080/00438243500095199>
- Schroeder, S. (1999). Maize productivity in the eastern woodlands and Great Plains of North America. *American Antiquity*, 64, 499–516. <https://doi.org/10.2307/2694148>
- Silvertown, J. (2008). The evolutionary maintenance of sexual reproduction: Evidence from the ecological distribution of asexual reproduction in clonal plants. *International Journal of Plant Science*, 169, 157–168. <https://doi.org/10.1086/523357>
- Smith, M. E. (2010). Trading patterns, ancient American. In W. H. McNeill, & J. H. Bentley (Eds.), *The Berkshire encyclopedia of world history* (2nd ed., pp. 2534–2535).
- Stewart, O. C. (2002). The effects of burning of grasslands and forests by aborigines the world over. In H. T. Lewis, & M. K. Anderson (Eds.), *Forgotten fires: Native Americans and the transient wilderness* (pp. 67–338). University of Oklahoma Press.
- Stuewer, F. W. (1943). Raccoons: Their habits and management in Michigan. *Ecological Monographs*, 13, 203–257. <https://doi.org/10.2307/1943528>
- Vale, T. R. (1998). The myth of the humanized landscape: An example from Yosemite National Park. *Natural Areas Journal*, 18, 231–236.
- Vale, T. R. (2002). The pre-European landscape of the United States: Pristine or humanized? In T. R. Vale (Ed.), *Fire, native peoples, and the natural landscape* (pp. 1–40). Island Press.
- Wallace, P. A. (1952). Historic Indian Paths of Pennsylvania. *The Pennsylvania Magazine of History and Biography*, 76, 411–439.
- Wasselkov, G. (1977). Prehistoric agriculture in the Central Mississippi Valley. *Agricultural History*, 51, 513–519.
- Wedel, W. R. (1943). *Archaeological investigations in Platte and Clay Counties, Missouri*. United States National Museum, Bulletin 183, pp. 26. Washington, DC.
- White, C. A. (1906). The northern limit of the Pawpaw tree. *Science*, 23, 749–751. <https://doi.org/10.1126/science.23.593.749>
- Williams, T. J., Collins, M. B., Rodrigues, K., Rink, W. J., Velchhoff, N., Keen-Zebert, A., Gilmer, A., Frederick, C. D., Ayala, S. J., & Prewitt, E. R. (2018). Evidence of an early projectile point technology in North America at the Gault Site, Texas, USA. *Science Advances*, 4, 1–7. <https://doi.org/10.1126/sciadv.aar5954>
- Willson, M. F. (1993). Mammals as seed-dispersal mutualists in North America. *Oikos*, 67, 159–176. <https://doi.org/10.2307/3545106>
- Willson, M. F., & Schemske, D. W. (1980). Pollinator limitation, fruit production, and floral display in Pawpaw (*Asimina triloba*). *Bulletin of the Torrey Botanical Club*, 107, 401–408. <https://doi.org/10.2307/2484160>
- Wright, A. P. (2017). Local and "global" perspectives on the Middle Woodland Southeast. *Journal of Archaeological Research*, 25, 37–83. <https://doi.org/10.1007/s10814-016-9096-5>
- Wright, S. (1922). Coefficients of inbreeding and relationship. *The American Naturalist*, 56, 330–338. <https://doi.org/10.1086/279872>
- Wright, S. (1943). Isolation by distance. *Genetics*, 28, 114–138. <https://doi.org/10.1093/genetics/28.2.114>

- Wright, S. (1951). The genetical structure of populations. *Annals of Eugenics*, 15, 323–354. <https://doi.org/10.1111/j.1469-1809.1949.tb02451.x>
- Wyatt, G. E., Hamrick, J. L., & Trapnell, D. W. (2021). Ecological niche modeling and phylogeography reveal range shifts of pawpaw, a North American understory tree. *Journal of Biogeography*, 48, 974–989. <https://doi.org/10.1111/jbi.14054>
- Wykoff, M. W. (1991). Black walnut on Iroquoian landscapes. *Northeast Indian Quarterly*, 8, 4–17.
- Yarnell, R. A., & Watson, P. J. (1966). Archaeological and paleoethnobotanical investigations in Salts Cave, Mammoth Cave National Park, Kentucky. *American Antiquity*, 31, 842–849. <https://doi.org/10.2307/2694457>
- Yeager, L. E., & Elder, W. H. (1945). Pre- and post-hunting season foods of raccoons on an Illinois goose refuge. *Journal of Wildlife Management*, 9, 48–53. <https://doi.org/10.2307/3795946>
- Yerka, S. J. (2010). *Geophysical study at Old Stone Fort State Archaeological Park, Manchester, Tennessee*. Master's thesis, Department of Anthropology, University of Tennessee, Knoxville, TN.
- Zhao, G.-X., Rieser, M. J., Hui, Y.-H., Miesbauer, L. R., Smith, D. L., & McLaughlin, J. L. (1993). Biologically active acetogenins from stem bark of *Asimina triloba*. *Phytochemistry*, 33, 1065–1073. [https://doi.org/10.1016/0031-9422\(93\)85024-L](https://doi.org/10.1016/0031-9422(93)85024-L)

#### SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

**How to cite this article:** Wyatt, G. E., Hamrick, J. L., & Trapnell, D. W. (2021). The role of anthropogenic dispersal in shaping the distribution and genetic composition of a widespread North American tree species. *Ecology and Evolution*, 00, 1–18. <https://doi.org/10.1002/ece3.7944>