Molecular Evidence and the Evolutionary History of the Domesticated Sunflower

The domestication of plants and animals by prehistoric humans was perhaps the most far-reaching cultural development in human history. Not only were domesticated organisms crucial to the rise of modern civilization, but their widespread use has dramatically altered the ecology and evolutionary history of numerous other species (Diamond, 2002). As a consequence, there is great interest in determining the geographic origins and timing of domestication (Sauer, 1952; Harlan, 1971). Although seemingly straightforward, this task is complicated by poor preservation of plant remains, particularly in tropical regions, and by the difficulty of discriminating between wholly independent origins of domestication and the secondary introduction of crop plants from a core region (Cowan and Watson, 1992; Denham et al., 2003; Neumann, 2003).

In the New World, these complications have led to conflicting interpretations of archaeological and paleobotanical evidence regarding the relationship between Mesoamerica and other regions where evidence of food production is found. One interpretation holds that Mesoamerica served as a primary center of domestication from which domesticated plant lineages and food production practices spread to areas of secondary innovation (Harlan, 1971; Lentz et al., 2001). In this view, the midlatitude woodland region of eastern North America is considered to be one of these secondary areas, and the
domestication of indigenous North American plant species is hypothesized to have been triggered by the introduction of major crops from Mesoamerica (Lentz et al., 2001). The alternative and more widely accepted interpretation is that agriculture in eastern North America arose wholly independently (Smith, 1989; Cowan and Watson, 1992, Neumann, 2003).

Evidence of an independent origin of agriculture in eastern North America derives primarily from the archaeobotanical record of four indigenous crops: thick-walled cucurbit or squash (\textit{Cucurbita pepo} ssp. \textit{ovifera}), sumpweed (\textit{Iva annua}), goosefoot (\textit{Chenopodium berlandieri}), and sunflower (\textit{Helianthus annuus}). All exhibit morphological changes in reproductive propagules that are associated with domestication (Asch and Asch, 1985; Smith, 1989). The transition to fully domesticated forms occurred between 4000 and 3000 years BP (Smith, 1989), which substantially predates the introduction of maize \textit{circa} 1800 years BP (Chapman and Crites, 1987); note that maize is thought to be the first tropical crop to be introduced into eastern North America (Smith, 1989). In addition, knotweed (\textit{Polygonum erectum}), maygrass (\textit{Phalaris caroliniana}), and little barley (\textit{Hordeum pusillum}) were used as minor seed crops before the introduction of maize (Cowan, 1978; Asch and Asch, 1985), but there is insufficient evidence to establish strong cases for their domesticated status.

Despite strong archaeobotanical support, the eastern North American origin of three of the four main indigenous domesticates (thick-walled cucurbit, goosefoot, and sunflower) has been questioned. For example, a recent mitochondrial DNA study (Sanjur et al., 2002) was consistent with an origin for \textit{C. pepo} ssp. \textit{ovifera} from wild gourds in either northeastern Mexico (\textit{C. pepo} ssp. \textit{fraternalia}) or eastern North America (\textit{C. pepo} ssp. \textit{ovifera} var. \textit{ozarkana}). However, a possible progenitor role for \textit{C. pepo} ssp. \textit{fraternalia} was quickly ruled out by random amplified polymorphic DNA (RAPD) data (Decker-Walters et al., 2002), which places the domesticate with \textit{C. pepo} ssp. \textit{ovifera} var. \textit{ozarkana} as originally proposed (Decker-Walters et al., 1993). Likewise, Wilson (1990) postulates that goosefoot might have a Mexican origin because of its close resemblance to the Mexican cultivar \textit{Chenopodium berlandieri} ssp. \textit{nutalliae}.

The most serious challenge to the eastern North American domestication hypothesis derives from the discovery of a sunflower achene and seed at the San Andrés site in Tabasco, Mexico, that date to 4130 ± 40 years BP and 4085 ± 50 years BP, respectively (accelerator mass spectrometry [AMS] determined) (Lentz et al., 2001; Pope et al., 2001). The achene and seed clearly represent the domesticated form, and their age rivals that of the earliest domesticated achenes from eastern North America, which are from the Hayes site in Tennessee and date to 4265 ± 60 years BP (AMS determined; Crites, 1993). However, Lentz et al. (2001) questions the shrinkage factors used to correct carbonized achene sizes at sites from eastern North America (Yarnell, 1978) and argues that the achenes from the Hayes site and other early finds actually represent wild material (but see Smith, 2003). If the Lentz et al. arguments were valid, then the earliest domesticated sunflower remains in eastern North America would derive from the Higgs site in eastern Tennessee (2850 ± 85 years BP, AMS determined; Brewer, 1973) and the Marble Bluff Rockshelter in northwest Arkansas (2842 ± 44 years BP, AMS determined; Fritz, 1997).

So far, molecular evidence has had little impact on the debate over the geographic origins of the domesticated sunflower, although it has been interpreted as supporting both sides of the debate (Heiser, 2001; Lentz et al., 2001). Given disagreements regarding the interpretation of earlier molecular studies and the recent completion of a comprehensive microsatellite survey of sunflower origins (Harter et al., 2004), it seemed worthwhile to provide a critical review of molecular data relating to sunflower domestication. We will show that although sunflower appears to be easily domesticated, molecular evidence indicates that all extant domesticated sunflowers had a single origin in eastern North America.

**Systematics and Biogeography of \textit{H. annuus}**

\textit{Helianthus} comprises approximately 50 species of sunflower, all of which are native to North America (Schilling and Heiser, 1981; Seiler and Rieseberg, 1997). The genus is monophyletic (Schilling et al., 1994) and includes diploids (n = 17), tetraploids, and hexaploids. Although most species are perennial, section \textit{Helianthus} (formerly section \textit{Annuus}) includes 11 or 12 species, most of which are self-incompatible, diploid annuals. Molecular phylogenetic studies indicate that the section is monophyletic and consistently place \textit{H. annuus} in a clade with three other species: \textit{H. argophyllus}, \textit{H. bolanderi}, and \textit{H. exilis} (Rieseberg, 1991; Rieseberg et al., 1991; Schilling, 1997; Schilling et al., 1998). In all trees, \textit{H. argophyllus}, a silver-leaved sunflower from southern Texas, is sister to \textit{H. annuus}. The two species do hybridize in areas of contact in southern Texas but retain their distinctive morphology and karyotype, presumably because of divergent ecological selection and a fairly strong chromosomal sterility barrier (Heiser, 1951a).
The domesticated sunflower is clearly derived from the wild form of *H. annuus*, or common sunflower (Heiser 1951b, 1954). Hybrids between wild and domesticated *H. annuus* are fully fertile (Heiser, 1954), and molecular studies all confirm the predicted progenitor-derivative relationship (e.g., Rieseberg and Seiler, 1990; Cronn et al., 1997; Tang and Knapp, 2003). Heiser (1954) gave formal recognition to four different forms of the common sunflower: *H. annuus* ssp. *lenticularis* (the western North American subspecies), *H. annuus* ssp. *texasus* (a form of *H. annuus* from Texas that has converged toward a local species, *H. debilis*, with which it hybridizes), *H. annuus* ssp. *annuus* (the midwestern and more weedy form of the species), and *H. annuus* ssp. *annuus* var. *macrocarpus* (the domesticated sunflower). Heiser later recognized the inadequacy of this classification because of extensive intergradation between forms, so he adopted a less formal treatment in his monograph of the genus (Heiser et al., 1969). However, in later discussions, Heiser (1976, 1978) once again used subspecific nomenclature but restricted the definition of ssp. *annuus* to the urban weed form of *H. annuus*. Molecular evidence indicates that there is significant structuring among populations of *H. annuus*, but it more closely tracks geography (i.e., isolation by distance) than subspecific categories (Harter et al., 2004).

Wild *H. annuus* currently occurs throughout the continental United States, southern Canada, and northern Mexico (Heiser et al., 1969; González-Elizondo and Gómez-Sánchez, 1992), but its prehistoric distribution is poorly understood. Heiser (1951b) speculated that the species was restricted to the southwestern United States before the arrival of *Homo sapiens* into the Americas. Native Americans used wild *H. annuus* for food, so Heiser (1951b) proposed that it became a camp-following weed and was thereby introduced into the central and eastern United States, where it was domesticated. However, it seems more likely that buffalo was the primary dispersal agent (Asch, 1993) and that wild *H. annuus* was widely distributed throughout the Great Plains, western United States, and northern Mexico before the colonization of North America by humans.

Previous Molecular Studies

The first comprehensive molecular analysis of sunflower domestication assayed chloroplast DNA (cpDNA) and allozyme variation in 5 Native American varieties, 3 modern cultivars, 15 old landraces, and 12 wild populations from throughout the continental United States (Rieseberg and Seiler, 1990). All 23 cultivars had the same chloroplast DNA haplotype, implying a single origin for extant domesticated sunflowers. This haplotype was also found in wild populations from Missouri, New Mexico, and California, so no conclusions could be made regarding the geographic origin of the domesticates.

Wild and domesticated sunflowers were very similar at allozyme loci as well. Twenty-nine of 30 alleles found in the domesticates also occurred in wild populations, with an average genetic identity (I) between wild and domesticated populations of 0.93, a value only slightly lower than that for comparisons between wild populations (I = 0.96). Because of these very similar high genetic identities, the question of geographic origins could not be addressed. Nonetheless, high levels of allozyme variability in wild plants and virtual monomorphism in cultivated lines reinforced the cpDNA results: Extant domesticated sunflowers had a single origin from a very limited gene pool (Rieseberg and Seiler, 1990).

Shortly after this initial study, Arias and Rieseberg (1995) attempted to locate the geographic center of domestication for sunflower using RAPD markers. However, the high RAPD identity between wild populations and domesticated *H. annuus* (I = 0.976 to I = 0.997) once again precluded determination of geographic origin. In fact, Arias and Rieseberg were skeptical that molecular evidence could ever solve this problem, suggesting that the weedy, human-dispersed nature of wild *H. annuus* populations probably had erased evidence of geographic structure. Fortunately, as will be discussed later in this chapter, we were unnecessarily pessimistic.

In 1997, another attempt was made to use allozyme variation to ascertain the geographic origin of the domesticated sunflower (Cronn et al., 1997). This study differed from that of Rieseberg and Seiler (1990) in its inclusion of four additional allozyme loci, increased sampling of both wild and cultivated accessions, the use of clusters of related populations as operational taxonomic units in genetic distance trees, and the inclusion of related wild species for rooting the trees. This improved method led to the discovery of limited geographic structure among wild populations. More significantly, they found that the domesticated sunflower was slightly more similar genetically to wild populations from the Great Plains than from the Southwest or California. However, support for this relationship was very weak.

Recently, the development of microsatellite loci for sunflower has greatly enhanced our ability to analyze genetic relationships between domesticated and wild accessions (Whitton et al., 1997; Tang et al., 2002). In previous work, cpDNA haplotypes and RAPD and allozyme allele frequencies were
not sufficiently differentiated between geographic locations to determine likely source populations for domesticated sunflower. However, microsatellites have proved superior to these markers for the study of domestication because there is more intraspecific genetic variation at these loci, making it feasible to dissect relationships between recently divergent populations.

In sunflower, microsatellites were first used for this purpose by Tang and Knapp (2003). With the exception of a wild accession from North Dakota, which appears to be the product of crop–wild hybridization, their study provided the first strong statistical support for the genetic separation of cultivated from wild material. Unfortunately, there was insufficient resolution between the wild populations and inadequate geographic coverage to determine the geographic origin of the domesticated sunflower. However, it is noteworthy that a wild population from the Great Plains (Oklahoma) clustered most closely with the domesticates, and the single wild population from Mexico was most distant.

The most intriguing result of Tang and Knapp (2003) was the large genetic distances observed between two of the Native American varieties (Hopi and Havusupai) and other domesticated sunflowers (0.714 to 0.798). Tang and Knapp interpreted the large distances as evidence that the domesticated sunflower might have multiple origins. This interpretation was consistent with earlier observations by Heiser (1976) on the morphological distinctness of the Hopi and Havusupai varieties, the discovery of domesticated sunflower remains at archaeological sites in both Mexico (Lenz et al., 2001) and eastern North America (Yarnell, 1978), and quantitative trait locus studies of domestication traits (Burke et al., 2002) indicating that sunflower was easily domesticated (domestication entailed few major genetic changes, and wild populations contain numerous alleles with effects in the direction of the cultivar).

There are two weaknesses with the multiple-origin hypothesis. First, the genetic distances reported by Tang and Knapp (2003) are exaggerated because only a single sample was analyzed per accession. Second, all sampled domesticated sunflowers appear to form a monophyletic lineage that derives from within the pool of wild variation. Note that this is not immediately apparent in figure 6 of Tang and Knapp (2003) because the consensus unweighted pair group method with arithmetic mean tree was rooted with a highly derived cultivar lineage rather than a primitive wild form, and they included the hybrid North Dakota population in the tree. If there were multiple origins of the domesticates, we would expect independently derived cultivar lineages to be placed sister to the wild progenitor populations from which they were derived, and this is not the case. On the other hand, given the lack of extensive sampling from Mexico, it was perfectly reasonable for Tang and Knapp to assume that probable progenitor populations for at least one of the origins had not been sampled.

Recent Work

A second microsatellite survey was recently completed by Harter et al. (2004). This study differed from that of Tang and Knapp (2003) in that there was complete geographic coverage of the prehistoric range of sunflower, including Mexico. Also, all wild populations were collected by the authors and attempts were made to choose large populations from natural sites that were far from cultivated fields to minimize the potential for crop–wild gene flow. Finally, in addition to standard tree-building methods, sophisticated model-based clustering approaches were used that are more appropriate and powerful for assigning domesticates to wild populations and for reconstructing the pattern of genetic drift between wild populations and domesticated strains arising from the domestication process.

Individuals from 21 geographically diverse populations of wild H. annuus from North America and Mexico and 10 domesticated lineages including 2 commercial lines and 8 Native American–developed landraces (figure 2.1) were genotyped for 18 microsatellite loci (Harter et al., 2004). The resultant data set was analyzed in three ways. Pairwise genetic distances between populations were calculated and used to construct a neighbor-joining (NJ) tree (figure 2.2). Second, a model-based clustering approach was implemented with the software program STRUCTURE (Pritchard et al., 2000; Falush et al., 2003) to infer population structure in wild H. annuus and then to assign the domesticates to inferred populations. Third, the STRUCTURE program (Pritchard et al., 2000; Falush et al., 2003) was used to make inferences about ancestral allele frequencies in the common ancestor of wild and domesticated sunflower and the degree of drift away from the ancestral genomic composition in each population.

Neighbor-Joining Tree

The topology of an NJ tree based on pairwise genetic distances between populations closely follows their geographic distribution, although some nodes are not well supported (figure 2.2). The Mexico plus Arizona grouping is supported by a high bootstrap value of 90% and includes
FIGURE 2.1 Map of sampling locations used by Harter et al. (2004), archaeological sites and Native American groups. Shaded areas = centers of domestication, with eastern North America to the north and Mesoamerica to the south; numbers = sampling locations of wild populations, where 1 = Sinaloa, 2 = Sonora5, 3 = Sonora3, 4 = Sonora6, 5 = Tamaulipas, 6 = Zacatecas, 7 = Nuevo León, 8 = Chihuahua, 9 = Arizona, 10 = Texas, 11 = Oklahoma2, 12 = Kansas, 13 = Colorado, 14 = Montana1, 15 = Montana2, 16 = North Dakota, 17 = South Dakota, 18 = Iowa, 19 = Missouri, 20 = Oklahoma1, 21 = Tennessee; names = historical locations of Native American groups; and letters = archaeological sites with oldest remains of domesticated sunflower, where A = San Andres, Tabasco, MX (4130 ± 40 bp), B = Higgs, TN, USA (2850 ± 85 bp), C = Hayes, TN, USA (4265 ± 60 bp) and D = Marble Bluff, AR, USA (2843 ± 44 bp). Identities of indigenous groups associated with Maíz de Tejas and Maíz Negro are unknown. USDA and Mammoth are modern cultivars derived from Russian stock. Therefore these strains do not appear on the map.

FIGURE 2.2 Majority rule consensus neighbor-joining tree summarized the genetic distances, $D_A$ (Nei et al., 1983) between groups. West Mexico populations are underlined, east-central Mexico populations are in boxes, U.S. Great Plains populations are in italics, east-central U.S. populations are in bold, and cultivars are in bold and italics. Numbers in parentheses correspond to sampling locations of wild H. annuus populations, as shown in figure 2.1. Numbers along branches are mean drift values; the value for each domesticated strain is the average across all comparisons with wild populations, and the value for each wild population is the average across all comparisons with domesticated strains. Numbers at nodes indicate bootstrap values greater than 50% (1000 replicates). Because of space considerations, the 74% bootstrap value for the node subtending Colorado, the 54% bootstrap for the node subtending Seneca, and the 62% bootstrap for the node subtending Mammoth–Maíz de Tejas do not appear on the tree.
two clusters that correspond to the western coastal plain (Sinaloa, Sonora4, Sonora6, Sonora5) and northeastern Mexico (Tamaulipas, Nuevo León, Zacatecas), plus more interior populations (Arizona and Chihuahua) basal to them. The U.S. cluster has lower bootstrap values, but the Great Plains populations (Montana2, Montana1, North Dakota, South Dakota, Colorado, Kansas, Texas, and Oklahoma2) form a discrete group within which the branching order reflects geographic relationships. Populations to the east of the Great Plains (Tennessee, Missouri, Iowa, and Oklahoma1) do not form a distinct group. Instead, each is sister to the Great Plains group.

All cultivars belong to a single, strongly supported group (bootstrap = 100%) in the NJ tree. Although Hopi and Havasupai form a distinct and well-supported clade within the cultivar group, genetic distances (0.436 to 0.696) are not as large as those reported by Tang et al. (2003). Wild populations from the east-central United States, especially Tennessee, Missouri, and Iowa, which represent the eastern wild form (H. annuus ssp. annuus), have the closest genetic relationship with all the domesticated accessions. More broadly, the Great Plains populations, as a whole, cluster more closely with the domesticates (bootstrap = 90%) than do populations from Mexico. These results suggest a single origin of extant domesticated sunflowers from the east-central United States as originally hypothesized by Heiser (1951b). Note that this result is not inconsistent with genetic data suggesting that sunflower is readily domesticated (Burke et al., 2002) because domestication of even the most amenable wild taxon is a long and arduous process when compared with the spread of an already domesticated form.

Model-Based Clustering

The admixture model included in the STRUCTURE program was used to define genetic populations or clusters in wild H. annuus based on allele frequencies and then to assign domesticated genotypes probabilistically to these defined clusters. Genetic populations were defined at both a regional and a local scale. Note that the admixture model allows individuals to originate from more than one source population.

At the regional scale, two genetic populations or clusters of wild H. annuus were consistently found by the STRUCTURE program. One cluster comprised all Mexican populations plus Arizona, whereas all central U.S. populations (i.e., populations from the Great Plains and east-central United States) formed a second cluster (figure 2.3a). Assignment of domesticated individuals to these two clusters revealed that all extant domesticates had central U.S. ancestry (figure 2.4). Indeed, the average estimated ancestry for each domesticated strain was at least 0.997!

The regional clusters (figure 2.3a) were subjected to further independent analyses to identify local genetic populations. Tests for population structure on the Mexican subsample identified two clusters that correspond to distinct geographic regions: west Mexico and east-central Mexico (figure 2.3b). Likewise, the North America subsample could be subdivided genetically into a U.S. Great Plains and east-central U.S. cluster (figure 2.3b). Assignment of domesticated lineages to these local clusters revealed that, as predicted by the NJ tree, all domesticates were assigned to the east-central United States, with average estimated membership of at least 0.994 for all domesticates (figure 2.4). Thus both regional and local clustering analyses indicate that domesticated sunflowers are most similar to wild H. annuus from the central United States, particularly the easternmost populations.

Patterns of Genetic Drift

All previous studies of genetic variation in wild and domesticated sunflowers have reported much lower levels of variability in domesticated than in wild sunflowers (Rieseberg and Seiler, 1990; Cronn et al., 1997; Tang and Knapp, 2003), as would be predicted if there were a genetic bottleneck associated with domestication. Using the F model of the STRUCTURE program, Harter et al. (2004) investigated the pattern of genetic drift between wild and domesticated sunflowers in order to determine whether this pattern was consistent with domesticates arising via genetic drift from wild U.S. populations or from wild Mexican populations. The F model assumes that populations have independently drifted from the allele frequencies found in their common ancestor and uses a Bayesian approach to make inferences about ancestral allele frequencies and the rate of drift away from the ancestor. Wild populations that are most similar in allele frequency to the common ancestor of wild and domesticated H. annuus should exhibit little evidence of drift (i.e., have low drift values). Likewise, if domestication is associated with a strong selective bottleneck, domesticated lines should have much larger drift values than wild populations.
FIGURE 2.3 Map of genetic populations or clusters of wild *H. annuus*. Numbers = sampling locations of wild populations, where 1 = Sinaloa, 2 = Sonora5, 3 = Sonora4, 4 = Sonora6, 5 = Tamaulipas, 6 = Zacatecas, 7 = Nuevo León, 8 = Chihuahua, 9 = Arizona, 10 = Texas, 11 = Oklahoma2, 12 = Kansas, 13 = Colorado, 14 = Montana1, 15 = Montana2, 16 = North Dakota, 17 = South Dakota, 18 = Iowa, 19 = Missouri, 20 = Oklahoma1, 21 = Tennessee. (A) Regional clusters of wild *H. annuus*: Mexico plus Arizona and central United States. (B) Local clusters of wild *H. annuus*: west Mexico, east-central Mexico, Great Plains, and east-central United States.

FIGURE 2.4 Results of the domesticated *H. annuus* genotypic cluster assignment. Each domesticated individual's genome is represented by a thin vertical line that is partitioned into colored segments in proportion to the estimated membership in each of the wild source clusters. Cultivars are separated with black lines, with names below and sample sizes above. (Full-color version of this figure follows page 230.)

As predicted, domesticated populations had much higher drift values than wild populations (figure 2.2). The lowest mean drift value in a domesticate was more than 200 times that of the lowest mean value in a wild population, indicative of a strong genetic bottleneck associated with domestication (Harter et al., 2004).

Consistent with the cluster analyses, drift values in the wild populations place the ancestry of the domesticated sunflower in the central United States but fail to differentiate between a Great Plains or an east-central origin (figure 2.2). The nine lowest drift values are from central U.S. populations, and several populations actually have 90% credibility regions around the estimated drift value.
that includes 0.000 drift away from ancestral allele frequencies: Kansas, South Dakota, Oklahoma2 (all Great Plains), and Iowa (east-central United States). This is a remarkable result indicating that these contemporary wild populations are essentially identical in allele frequency to the wild ancestor of domesticated sunflowers. Several other populations from the Great Plains (North Dakota) or east-central United States (Oklahoma1 and Tennessee) have intermediate or high drift values, however, indicating that allele frequencies in these populations have drifted substantially from those found in the ancestral population. These populations cluster genetically with other wild populations from these areas (figure 2.2), a result consistent with strong localized drift events (Harter et al., 2004), perhaps because of recent founding events. This explanation is particularly likely for Tennessee, which is a roadside population far from the native range of *Helianthus annuus*.

The combined results from the genetic distance tree, model-based clustering, and drift analyses indicate that the progenitor of the domesticated sunflower was genetically most similar to wild populations in the central plains of the United States. A more precise geographic location is difficult (and perhaps nonsensical) to infer because of high levels of gene flow between populations in this area. Nonetheless, of the sampled populations, Iowa is probably most similar to the ancestor in that the 90% credibility region around the estimated drift value for this population includes 0.000, it is placed very close to the cultivars in the NJ tree, and it belongs to the east-central U.S. genetic population to which the domesticates were assigned in the model-based cluster analysis.

**Conclusions**

Several inferences can be made from molecular genetic studies of the domesticated sunflower. First, all studies agree that domestication was associated with a strong genetic bottleneck. As a consequence, allele frequencies have changed more than 50 times faster in domesticated lineages than in wild populations since their divergence from a common ancestor. Second, molecular evidence is most consistent with a single origin of all extant domesticated sunflowers. All domesticates share the same chloroplast DNA haplotype, and in the most recent and convincing molecular studies, all extant domesticates are placed in a monophyletic group that is well separated from all wild populations. Third, genetic distance trees, model-based clustering, and drift analyses of microsatellite data all indicate that the domesticated sunflower was derived from

wild sunflowers in the central plains of the United States. Because of extensive gene flow between populations in this area, it is not possible to assign the domesticated sunflower to a single wild population or local geographic area. However, of the sampled localities, a population from Iowa was most similar to the wild ancestor of the domesticated sunflower.

Based on these data, a likely scenario for domestication is that wild sunflowers from the central plains colonized adjacent regions to the east (i.e., Tennessee, Kentucky, Illinois, Missouri, Arkansas, and Ohio), perhaps because of human activities in the middle Holocene (Heiser, 1951b). The wild sunflowers were subsequently brought under cultivation and were domesticated over a period from approximately 4000 BP to 3000 BP (Smith, 1989). More generally, molecular evidence of a U.S. ancestry of extant domesticated sunflowers supports an origin in eastern North America independent of Mesoamerican domestication. However, the provenance of the domesticated achenes from the San Andrés site in Mexico remains a mystery. Possibly, there was an earlier and independent domestication in Mexico, but it does not appear to have influenced domestication in eastern North America. Alternatively, achenes may have been carried to San Andrés from the north. However, as far as we are aware, there is no evidence of long-distance trade at this time. Additional archaeobotanical work in Mexico is needed to establish the authenticity of the Mexican find by estimating the timing and duration of the Mexican domestication event (if it existed) and determining the date of extinction and its cause.

**References**


