THE GENETICS OF SUNFLOWER DOMESTICATION

by

DAVID M. WILLS

(Under the Direction of John M. Burke)

ABSTRACT

The chapters of this dissertation investigate the early stages of the evolution of cultivated sunflower. Previous archeological and genetic investigations of domesticated sunflower have argued for either: (1) a single origin of domestication in the east central United States, or (2) one origin in the United States and a second origin in Mexico, potentially represented by the primitive Hopi and Havasupai landraces. The first chapter reports an analysis of chloroplast haplotype diversity in wild and domesticated sunflower, the results of which support a single origin of domestication of the extant cultivated lineages, most likely somewhere outside of Mexico. These results, combined with previous analyses based on nuclear markers, confirmed that the Hopi and Havasupai landraces represent the most basal domesticated sunflower lineages available. Having confirmed the Hopi sunflower as a primitive domesticate, the second chapter utilizes a wild × Hopi mapping population to investigate the minimum genetic changes required to transform wild sunflower into a useful crop plant. The genetic architecture of sunflower domestication was found to involve a large number of loci, most of which had small to moderate phenotypic effects. This is a unique genetic architecture, which previously has been undocumented in other crop species. Despite being one of the two most basal

domesticated lineages, the Hopi sunflower exhibits several unique traits – primarily achene shape, anthocyanin pigmentation, and exceptionally late flowering time – which appear to be byproducts of post-domestication selection. The third chapter thus utilizes the same wild × Hopi mapping population investigate the genetic architecture of these traits. Unlike typical domestication-related traits, these traits exhibited a relatively simple genetic basis, with two genomic regions being largely responsible for the divergence of the Hopi sunflower from other cultivated sunflower lineages. The simple genetic control of these traits may be the product of constraint imparted by the complex genetic architecture of domestication.

INDEX WORDS: Sunflower, *Helianthus annuus*, domestication, quantitative trait locus, chloroplast haplotype, origin, population bottleneck, post-domestication diversification, Hopi, genetic architecture, domestication syndrome

THE GENETICS OF SUNFLOWER DOMESTICATION

by

DAVID M. WILLS

B.S., University of Tennessee, 2000

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2008

THE GENETICS OF SUNFLOWER DOMESTICATION

by

DAVID M. WILLS

Major Professor: John M. Burke

Committee: James L. Hamrick
James H. Leebens-Mack
Steven J. Knapp
Michael L. Arnold

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
May 2008

DEDICATION

This work is dedicated to Kathleen Wills, Michael Wills, and Daniel Roberts for giving me my start as a scientist.

ACKNOWLEDGMENTS

I would like to thank my advisor John Burke for his excellent guidance and showing me how amazing crops are. Additional thanks to my committee members. This work was started at Vanderbilt University with David McCauley, Patrick Abbot, and Scott Williams and then completed at the University of Georgia with James Hamrick, Steven Knapp, James Leebens-Mack, and Michael Arnold.

A special thanks to Kathleen Wills, Michael Wills, Kevin Wills, Stephen Wills, Erica Hall, and Taressa Fenus for the love and support that enables all that I do.

TABLE OF CONTENTS

Page

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

The domestication of plant lineages involves selection on a variety of traits to make products more useful for man. Common domestication traits include: increased seed or fruit size, more determinate growth and flowering, suppression of natural seed dispersal, and loss of self-incompatibility. Many of these changes make a plant less fit in the wild, and thus a plant is fully domesticated when its survival is completely dependant on humans (Harlan 1992). Although a great variety of plants have been domesticated, traits that are subjected to selection are typically the same and are collectively referred to as the "domestication syndrome" (Hammer 1984; Harlan 1992). Because plant domestication typically involves a rapid response to intense selection, which produces large changes in quantitative traits, it can provide insight into the origins of new taxa, the genetic consequences of selective sweeps, the response of populations to long-term directional selection, and the limitations imposed on selective response by genetic architecture (e.g., Stuber *et al.* 1980; Wang *et al.* 1999; Bost *et al.* 2001).

Derived from the wild, common sunflower (*Helianthus annuus*), cultivated sunflower (also *H. annuus*) has developed a number of traits that are consistent with the domestication syndrome. These adaptations to human cultivation include a dramatic increase in apical dominance, an increase in seed size, the loss of natural seed dispersal

and seed dormancy, and the loss of self-incompatibility. Yet, this transition in sunflower has been shown previously to have occurred in a singularly unique manner amongst crops (Burke *et al.* 2002; 2005). The goals of my dissertation research were to follow up on this earlier finding by identifying the number of origins of sunflower domestication, assessing the genetic architecture of sunflower domestication, and to investigate the genetic architecture of a unique instance of post-domestication diversification.

In the time since its domestication, sunflower has become one of the world's most important oilseed crops, an important source of confectionery seeds, and a major source of cut flowers (Putt 1997). The weedy, common sunflower thrives in disturbed sites, and is widely distributed in the western United States and the Great Plains area where it was adapted for dispersal by buffalo herds (Berthoud 1892; Heiser 1954). Despite being morphologically distinct, cultivated and common sunflower are considered members of the same species, and are fully reproductively compatible. Results of previous research indicate that, whenever they come into contact and flower coincidentally, common and cultivated sunflower hybridize readily, sometimes over surprisingly long distances (Arias and Rieseberg 1994; Whitton *et al.* 1997; Linder *et al.* 1998). Wild sunflower's adaptation to disturbance may have predisposed it for human use. Heiser postulated that sunflower first occurred as a camp following weed, which was eventually used for food, and then ultimately domesticated somewhere in the east-central United States (Heiser 1954; Heiser *et al.* 1969).

As with all crop plants, sunflower is presumed to have experienced a population bottleneck concomitant with domestication. Despite the occurrence of this bottleneck, however, domesticated sunflower has been shown to harbor a great deal of genetic

diversity (e.g., Tang and Knapp 2003; Liu and Burke 2006). Such high levels of genetic diversity may be the result of a relatively mild population bottleneck and/or a rapid recovery to large population size, or it may be the result of multiple domestication events, each of which may have infused the cultivated gene pool with novel genetic variants. Although sunflower was long thought to be the product of a single domestication event, both the archaeological and genetic investigations of the origin of sunflower have failed to conclusively confirm a single origin, leading multiple researchers to consider the possibility of a second origin in Mesoamerica (e.g., Heiser 1985; Lentz *et al.* 2001; Tang and Knapp 2003).

Archaeological finds have indicated the presence of domesticated sunflowers in what is now the central United States and southern Mexico by about 4000 years before present (BP). Carbonized seeds from the Hayes Site in middle Tennessee provide the earliest record of domesticated sunflower (ca. 4,300 years before present; Brewer 1973, Ford 1985; Crites 1993). Lentz *et al.* (2001) reported finding what appeared to be domesticated sunflower seeds in southern Mexico. These seeds were approximately the same age (~4,000 BP) as those found in the United States, and no older seeds have been recovered. Based on this finding, Lentz *et al.* (2001) determined that sunflower had in fact been first domesticated in Mexico and spread northward. Recently, Heiser (2008) has concluded his verification of the seeds found by Lenz *et al.* (2001) as sunflower seeds to be incorrect, suggesting that they are, in fact, bottle gourd seeds. Like the archaeological evidence, genetic studies have not conclusively ruled out multiple origins of domesticated sunflower. Rieseberg and Seiler (1990) identified four chloroplast haplotypes present in wild sunflower using RFLPs. Consistent with a single origin, only one haplotype was

fixed in the domesticates. The relatively high frequency (27%) of the cultivar haplotype in the wild, however, allows for the possibility of a second domestication event that may have been obscured if it received the same haplotype. Investigations using nuclear markers have found contradictory results, with domesticated sunflower being found to be either monophyletic or paraphyletic (Tang and Knapp 2003; Harter *et al.* 2004). Tang and Knapp (2003) speculated that Hopi and Havasupai landraces, which are separated from the balance of the domesticates by a substantial genetic distance, may represent the remnants of an independent origin of domestication. It is thus noteworthy that these landraces are geographically the most proximate landraces to the putative Mexican origin of domestication.

The aim of the first chapter of my dissertation is to determine if domesticated sunflower is singly or multiple derived, and to search for the signature of a possible origin of domestication in Mexico. Chloroplast microsatellites were employed because they have two advantages over previous approaches. First, because the chloroplast genome is maternally inherited, chloroplast data will be much more resistant to the confounding effects of hybridization than are nuclear markers such as those employed by Tang and Knapp (2003) and Harter *et al.* (2004). Second, microsatellites are highly polymorphic, providing a much greater level of resolution than is typically available for other types of chloroplast markers, such as the RFLPs employed by Rieseberg and Seiler (1990). Indeed, high levels of chloroplast microsatellite variation have already been documented in a number of species (Provan *et al.* 2001), and such markers have been used to successfully investigate the origin of domestication in soybean (Xu *et al.* 2002). These primers were used to survey chloroplast haplotype diversity in 26 wild populations

selected throughout the native range and 15 domesticated lineages focused primarily on the primitive domesticates. The frequency and distribution of chloroplast haplotypes in wild and domesticated lineages were then used to infer the number and location(s) of sunflower domestication events.

Whether singly or multiply derived, the phenotypic changes that sunflower has undergone during domestication largely follow the domestication syndrome. To assess the genetic architecture of these phenotypic changes, quantitative trait loci (QTL) can be identified to estimate of the number of genomic regions and their actions and interactions which control these phenotypes. QTL analysis of the domestication syndrome in other crops has shown that domestication traits are typically under very simple genetic control (e.g., Doebley *et al.* 1990; Koinange *et al.* 1996). Maize provides the most striking example, in which five genomic regions control the majority of the phenotypic difference between teosinte and maize. Similar patterns have since been documented in tomato, common beans, rice, and soybean (Paterson *et al.* 1991; Koinange *et al.* 1996; Xiong *et al.* 1999; Liu *et al.* 2007).

Substantial progress has already been made in understanding the manner in which sunflower was domesticated. Based on earlier work, sunflower does not appear to follow the pattern of relatively simple genetic control of domestication related traits. In fact, a previous QTL analysis of domestication in sunflower resulted in the identification of 78 QTL with only four QTL explaining more than 25% of the segregating phenotypic variance (Burke *et al.* 2002). In addition, over one-third of the QTL identified had effects in the opposite direction as would be expected based on the parental phenotypes. This result was first interpreted as evidence that the variation necessary for domestication

might have already existed in wild sunflower populations, such that domestication required only a selective force (i.e., man) to proceed. An alternative explanation would be that this pattern is an artifact of the cross that the authors employed. More specifically, this work was based on a cross between wild sunflower and a highly improved oilseed cultivar. It could therefore be that the complex breeding history of modern cultivated sunflower resulted in the hitchhiking of apparently maladaptive (i.e., wild-like) alleles when other traits (e.g., disease resistance, seed oil characters, etc.) were introgressed into the cultivar gene pool. A follow-up analysis provides evidence for this having occurred. When oil content and composition traits were analyzed in that same population, a new cluster of QTL of major effect determining seed oil composition was identified on linkage group 6 (LG 6) (Burke *et al.* 2005), which was previously show to harbor a cluster of QTL that largely acted in the wrong direction.

In view of the possibility that the previous investigations had confounded genetic changes that occurred during domestication with those that occurred during the subsequent period of improvement, the purpose of the second chapter of my dissertation is to map QTL underlying domestication-related traits in an F_2 population derived from a cross between common sunflower and a primitive Native American domesticate, the Hopi sunflower. Such an analysis provides a much better picture of the minimum number of genetic changes required to produce a domesticated sunflower. The Hopi landrace was chosen because it and Havasupai appear to represent the most basal of the extant cultivated sunflower lineages (see Chapter 1; Harter *et al.* 2004; Tang and Knapp 2003; Wills and Burke 2006). The use of an unimproved domesticate was intended to protect against the effects of improvement-related alleles that may have produced artifactual

results in the analysis of Burke *et al.* (2002), thereby providing a better understanding of the changes which transformed wild into domesticated sunflower.

As alluded to above, when trying to assess the genetic architecture of domestication, QTL responsible for post-domestication diversification should be ignored. However, these sorts of QTL are themselves interesting for other reasons. For example, as a crop disseminates away from its initial origin of domestication, adaptation to local growing conditions will almost inevitably occur. In addition, certain lineages will inevitably be selected to improve particular traits whereas others may not. Finally, as new cultures acquire a crop, certain cultivated lineages may be adapted for a new use. This complex history of domestication and improvement is preserved to some degree in extant germplasm collections. For example, as noted above, the Hopi sunflower is one of the two most basal extant cultivated sunflower lineages, yet it has a set of unique characteristics that have arisen post-domestication which make it distinct from other domesticated lineages. Specifically, the Hopi sunflower exhibits exceptionally late flowering, which is presumably an adaptation to the long growing season in the desert southwest (Heiser 1946), and the achenes have a unique morphology featuring a greater length-to-width ratio, higher kernel-to-pericarp weight ratio, and an extremely high concentration of anthocyanin in the achene pericarp (i.e., the 'shell' that surrounds the sunflower seed; Heiser 1945). The high anthocyanin content of the Hopi sunflower seeds made them an ideal source of source of purple dye for use in basketry and textiles, as well as in body paint (Whiting 1939; Heiser 1951).

The final chapter of my dissertation involves an investigation of the genetic architecture of this unique suite of traits in the Hopi landrace. The same mapping

population produced for the identification of domestication QTL was phenotyped for the seed characteristics that differentiate Hopi from the other domesticates. The identification of QTL for these post-domestication traits provides insight into the issue of whether or not the genetic architecture of post-domestication specialization is similar to that of the phenotypic changes that occurred during domestication.

CHAPTER 2

CHLOROPLAST DNA VARIATION CONFIRMS A SINGLE ORIGIN OF DOMESTICATED SUNFLOWER (*HELIANTHUS ANNUUS* L.)[1]

[1] WILLS, D.M. and J.M. BURKE. Chloroplast DNA variation confirms a single origin of domestication sunflower (*Helianthus annuus* L.). *Journal of Heredity*.2006. 97 (4):403-408. Reprinted here with permission of the American Genetic Association.

ABSTRACT

Although sunflower was long thought to be the product of a single domestication in what is now the east-central United States, recent archaeological and genetic evidence have suggested the possibility of an independent origin of domestication, perhaps in Mexico. We therefore used hypervariable chloroplast simple-sequence repeat (cpSSR) markers to search for evidence of a possible Mexican origin of domestication. This work resulted in the identification of 45 chloroplast haplotypes from 26 populations across the range of wild sunflower as well as 3 haplotypes from 15 domesticated lines, representing both primitive and improved cultivars. The 3 domesticated haplotypes were characterized by one primary haplotype (found at a frequency of 6.7% in the wild), as well as two rare haplotypes, which are most likely the products of mutation or introgression. One of these rare haplotypes was not observed in the wild, bringing the total number of haplotypes identified to 46. A principal coordinate analysis revealed the presence of three major haplotype clusters, one of which contained the primary domesticated haplotype, the two rare domesticated variants, as well as haplotypes found across much of the range of wild sunflower. The Mexican haplotypes, on the other hand, fell well outside of this cluster. While our data do not provide insight into the specific location of sunflower domestication, the relative rarity of the primary domesticated haplotype in the wild, combined with the dissimilarity between this haplotype and those found in the Mexican populations surveyed, provides further evidence that the extant domesticated sunflowers are the product of a single domestication event somewhere outside of Mexico.

INTRODUCTION

The majority of crop plants were domesticated between 4,000–10,000 years ago (Hancock 2004) and, in most cases, the wild progenitors of these crops have been satisfactorily identified. We are, however, continually gaining insight into the details surrounding the domestication of these plants. For example, it is clear in some cases (such as barley, maize, and potato) that the crop form arose just once (Badr *et al.* 2000, Matsuoka *et al.* 2002, and Spooner *et al.* 2005). Thus, the current ranges of cultivation of these crops reflect post-domestication diffusion from their centers of origin. In other cases, such as rice, cotton, and soybean, the crop appears to be the product of multiple origins of domestication (Second 1982, Wendel *et al.* 1995, and Xu *et al.* 2002), sometimes in geographically disparate locales. Here we report the results of an investigation into the origin of domesticated sunflower based on patterns of chloroplast DNA variation.

Domesticated sunflower (*Helianthus annuus*) is one of the world's most important oilseed crops, and is also a major source of confectionery seeds (Putt 1997). Derived from the common sunflower (also *H. annuus*), domesticated sunflower was initially thought to have arisen just once in what is now the east-central United States (Heiser 1954, 1978). In fact, Heiser (1954) first hypothesized that the use of wild sunflowers by Native Americans as a food source resulted in the production of a camp-following weed that eventually spread eastward, and that this weed ultimately served as the progenitor of domesticated sunflower. However, Heiser (1985) later discussed the possibility of an additional origin of domestication, perhaps in Mexico. Until recently (see below), the

available archaeological evidence (Brewer 1973, Ford 1985; Crites 1993) was most consistent with the single origin hypothesis, with carbonized achenes (i.e., single-seeded fruits) from the Hayes Site in Middle Tennessee providing the earliest record of domesticated sunflower (ca. 4,300 years before present; Crites 1993).

In terms of genetic data, Rieseberg and Seiler (1990) surveyed a broad collection of wild and domesticated sunflower lines and found that the domesticates exhibited reduced allozyme variability, and that they were all characterized by a single chloroplast DNA (cpDNA) RFLP haplotype. While this result is consistent with a single origin of domestication, these data are far from conclusive, as the domesticated cpDNA haplotype was geographically widespread and present at relatively high frequency (27%) in the wild. It is thus conceivable that independently derived lines could share the same chloroplast haplotype by chance. In a subsequent survey of allozyme polymorphism, however, Cronn *et al.* (1997) reported that the domesticates form a "genetically coherent group" (p. 532), a result that was once again consistent with the hypothesis of a single origin. The possibility of a second origin of domestication was thus eventually dismissed based on the total weight of the archaeological and genetic evidence available at the time (Seiler and Rieseberg 1997).

The debate over the origin of domesticated sunflower was, however, revived when Lentz and colleagues (2001) reported the discovery of carbonized achenes of domesticated sunflower in southern Mexico, beyond the current range of wild sunflower. These achenes dated to roughly the same time period as those recovered at the Hayes site (ca. 4,000 YBP), and no older achenes have been recovered since that time. Shortly after this discovery, Tang and Knapp (2003a) used a suite of 122 nuclear simple-sequence

repeats (SSRs) to examine patterns of genetic diversity in both wild and domesticated sunflower. Based on their results, these authors concluded that "the single ancestor hypothesis … seems improbable" (p. 999). Rather, they suggested that the Hopi and Havasupai landraces, which are separated from the balance of the domesticates by a substantial genetic distance, might represent the descendants of the hypothesized 'other' origin of domestication. Adding to this is the fact that the Hopi and Havasupai lines are native to the desert southwest, making them the geographically most proximate landraces to the previously hypothesized Mexican origin of domestication. In the most comprehensive molecular analysis to date, however, Harter *et al.* (2004) argued that the eight extant Native American landraces, from which the modern cultivars are presumably derived, can all be reliably assigned to a single population genetic cluster based on patterns of nuclear SSR diversity. This result led them to conclude that these lines do, in fact, trace back to a single origin of domestication, most likely somewhere in central North America. Under their interpretation, the Hopi and Havasupai landraces represent the most primitive of the extant domesticates.

Here we reconsider the issue of single vs. multiple origins of sunflower domestication based on patterns of cpDNA variation in wild and domesticated sunflower. More specifically, we investigate the question of whether or not the Hopi and Havasupai landraces represent the descendants of an independent origin of sunflower domestication in Mexico. In order to answer this question, we used a suite of highly variable chloroplast SSRs (cpSSRs), which provided us with far greater levels of population genetic resolution than were available at the time of Rieseberg and Seiler's (1990) original RFLP-based survey of cpDNA diversity.

MATERIALS AND METHODS

Plant Materials and DNA Extractions

Wild and domesticated sunflower accessions were obtained from the North Central Regional Plant Introduction Station (NCRPIS; Ames, IA). Twenty-six wild accessions were selected to represent the range of common sunflower across North America, whereas fifteen domesticated lines were selected to represent the Native American landraces as well as improved lines (Table 2.1). Seeds were sown in flats and allowed to germinate in the greenhouse. Following seedling emergence, 200 mg of leaf tissue were collected from each of 4–6 individuals per accession. Total genomic DNA was then extracted from each sample using the Qiagen DNeasy Plant Mini Kit (Qiagen, Valencia, CA).

cpSSR Genotyping and Analysis

Conserved primer pairs that flank cpSSRs have been identified from a number of angiosperm species (Bryan *et al.* 1999; Powell *et al.* 1995; Weising and Gardner 1999). We used six such primer pairs, which have previously been shown to reveal polymorphisms within *H. annuus* (ccmp 2, ccmp 7, NTCP 9, NTCP 30, NTCP 40, and NTCP 18; Wills *et al.* 2005), to genotype each individual. We used a modification of the PCR methodology presented by Schuelke (2000), wherein we added an arbitrarily selected sequence (the M13 Forward [-29] sequencing primer, 5'-CAC GAC GTT GTA AAA CGA C-3') to the 5′ end of the forward primer. In order to allow for the visualization of multiple loci per lane on an automated DNA sequencer, PCR products were labeled by including a fluorescently-tagged M13 Forward (-29) primer (carrying

either HEX, FAM, or TET) in the reaction mixture. Reactions were performed in 10 µl total volume containing 10 ng of template DNA, 30 mM Tricine pH 8.4-KOH, 50 mM KCl, 2mM $MgCl_2$, 100 µM of each dNTP, 0.02 µM forward primer, 0.1 µM of both the reverse primer and the fluorescently labeled M13 primer, and 2 units of *Taq* polymerase. Cycling conditions were as follows: initial denaturation at 95°C for 3 min, followed by ten cycles of 30 s at 94°C, 30 s at 58°C (annealing temperature was reduced by one degree per cycle), 45 s at 72°C, followed by 30 cycles of 30 s at 94°C, 30 s at 48°C, 45 s at 72°C, and a final extension time of 20 min at 72°C.

Amplification products were visualized on an MJ Research BaseStation automated DNA sequencer (South San Francisco, CA) with MapMarker® 1000 ROX size standards (BioVentures Inc., Murfreesboro, TN) included in each lane to allow for accurate determination of fragment size. Alleles were called using the software package Cartographer (MJ Research), and the resulting data were analyzed using Arlequin (Excoffier *et al.* 2005) to generate summary statistics and GenAlEx 6 (Peakall and Smouse 2006) to perform a principle coordinate (PCO) analysis.

RESULTS AND DISCUSSION

Chloroplast Diversity

We determined the chloroplast haplotype of four to six individuals from each of the twenty-six wild populations and fifteen domesticated lineages using the six cpSSRs described above. Individual cpSSR loci harbored an average of 6.0 ± 1.1 alleles per locus (mean ± standard error), resulting in an average gene diversity of 0.59 ± 0.34 (Table 2.2).

Because the chloroplast genome is thought to be a non-recombining unit, we tested for linkage disequilibrium among loci using Slatkin's (1994) extension of Fisher's exact test. As expected, all loci were found to be in strong disequilibrium with one another (all $P < 0.001$).

In total, we identified forty-five unique wild sunflower haplotypes, with the most common haplotype occurring at a frequency of 10.7% (Figure 2.1). In contrast, 17 wild individuals carried unique haplotypes. With the exception of the Hidatsa and Maíz de Tejas landraces, all domesticated individuals shared a single haplotype (hereafter referred to as the 'primary' domesticated haplotype) which was the second most common haplotype found in the wild, occurring at a frequency of 6.7%. All six of the Hidatsa individuals that were initially surveyed shared a unique haplotype that differed from the primary domesticated haplotype at two of the six loci (ccmp 7 and NTCP 18) and was not found in the other domesticated lineages, nor in any of the wild populations. To confirm this finding, we requested a second Hidatsa accession from the NCRPIS and genotyped six additional individuals as described above. All six of these individuals contained the same haplotype that was found in the first Hidatsa accession. In the case of Maíz de Tejas, three of the four individuals surveyed exhibited the primary domesticated haplotype, whereas the fourth contained a haplotype that differed by a single base pair at one locus (NTCP 18).

Principal coordinates analysis was performed using the 46 cpSSR haplotypes that we identified (43 of which were found in wild sunflower only, 2 of which were shared between the wild and domesticated accessions, and 1 of which was unique to the Hidatsa landrace). Using the 'distance-not standardized' setting in GenAlEx 6 (Peakall and

Smouse 2005), wherein each of the 3 to 9 alleles per locus (mean = 6.0) were considered to be a single mutational step from all others, the first two coordinates explained 44.8% of the total variance. Inspection of Figure 2.2 reveals that the haplotypes appear to form three main clusters with three outlying haplotypes, and all three of the domesticated haplotypes occurring in the largest cluster. This cluster contains 30 of the 46 haplotypes, including the five most common haplotypes across the range of wild sunflower. Note that the two wild accessions from Mexico (#6 and #7) fell outside of this cluster, one as an outlier and the other within cluster #3.

Insights into the Origin of Domesticated Sunflower

Our data revealed the presence of three haplotypes within the primitive domesticates but only one in all other domesticates. While this finding is superficially consistent with the occurrence of independent origins, the two rare domesticated haplotypes seem more likely to be the result of other processes. For example, although the occurrence of a unique haplotype in the Hidatsa lineage could have resulted from a separate origin of domestication, previous researchers have found no evidence (based on nuclear markers) to suggest that this landrace arose independently of the others, and these plants come from North Dakota, which is geographically distant from the hypothesized 'other' origin of domestication. Thus, it seems most likely that this haplotype is the result of mutation and subsequent fixation within the Hidatsa landrace, or possibly introgression/chloroplast capture. While this haplotype has not been found in the wild, it falls within PCO cluster #1 (Figure 2.2). In the case of the Maíz de Téjas individual that differs from the primary domesticated haplotype by a single base pair at one of the six loci, the most likely

explanation seems to be that the haplotype carried by this individual arose as a result of a unique mutational event.

With regard to the hypothesis that the Hopi and Havasupai landraces trace to a second origin of domestication in Mexico, the relative rarity of the primary domesticated haplotype in the wild (ca. 6.7%) makes a second origin rather unlikely. Moreover, this haplotype was not found outside of the United States, and both of the Mexican haplotypes that we identified fell well outside of the PCO cluster that contains all of the domesticated lines that we surveyed. Thus, while we cannot rule out the possibility of a second origin of domestication in Mexico, the descendants of which ultimately went extinct, our data point to a single origin of the extant domesticated sunflowers somewhere outside of Mexico.

ACKNOWLEDGEMENTS

We thank Preston Aldrich, Mark Chapman, Aizhong Liu, Natasha Sherman, Jessica Wenzler and one anonymous reviewer for comments on an earlier version of the manuscript. Tamara Berthel, Chris Buckner, Jonathan Ertelt, Natasha Sherman, and Jessica Wenzler provided assistance in the greenhouse and/or lab. This work was supported by grants to JMB from the NSF Plant Genome Research Program (DBI-0332411) and the National Research Initiative of the USDA Cooperative State Research, Education and Extension Service (#03-35300-13104 and 03-39210-13958).

LITERATURE CITED

Badr A, Muller K, Schafer-Pregl R, *et al.* (2000) On the origin and domestication history of barley (*Hordeum vulgare*). *Molecular Biology and Evolution,* **17**, 499-510.

Brewer AJ (1973) Analysis of floral remains from the Higgs site (40L045). In: *Excavation of the Higgs and Doughty Site. I-75 Salvage Archaeology Tennessee Archaeological Society, Miscellaneous Papers No. 12* (eds. McCullough MCR, Faulknor CJ), pp. 141-147. University of Tennessee, Knoxville, TN.

Bryan GJ, McNicoll J, Ramsay G, Meyer RC, DeJong WS (1999) Polymorphic simple sequence repeat markers in chloroplast genomes of Solanaceous plants. *Theoretical and Applied Genetics,* **99**, 859-867.

Crites GD (1993) Domesticated sunflower in 5th millennium B.P. temporal context: new evidence from middle Tennessee. *American Antiquity,* **58**, 146-148.

Cronn R, Brothers M, Klier K, Bretting PK, Wendel JF (1997) Allozyme variation in domesticated annual sunflower and its wild relatives. *Theoretical and Applied Genetics,* **95**, 532-545.

Excoffier L, Laval G, Schneider S (2005) Arlequin Ver. 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online,* **1**, 47-50.

Ford RI (1985) Some botanical considerations of the early domesticated plants north of Mexico. In: *Prehistoric Food Production in North America, Museum of Anthropology, Anthropological Paper No. 75,* (ed. Ford RI), pp. 1-18. University of Michigan, Ann Arbor, MI.

Hancock JF (2004) *Plant Evolution and the Origin of Crop Species.* CABI Publishing, Cambridge, MA.

Harter AV, Gardner KA, Falush D*, et al.* (2004) Origin of extant domesticated sunflowers in eastern North America. *Nature,* **430**, 201-205.

Heiser CB (1954) Variation and subspeciation in the common sunflower, *Helianthus annuus*. *American Midland Naturalist,* **51**, 287-305.

Heiser CB, Jr. (1978) Taxonomy of *Helianthus* and origin of domesticated sunflower. In: *Sunflower Science and Technology* (ed. Carter JF), pp. 31-53. American Society of Agronomy, Madison, WI.

Heiser CB, Jr. (1985) Some botanical considerations of the early domesticated plants north of Mexico. In: *Prehistoric Food Production in North America, Museum of Anthropology, Anthropological Paper No. 75,* (ed. Ford RI), pp. 57-72. University of Michigan, Ann Arbor, MI.

Lentz DL, Pohl MED, Pope KO, Wyatt AR (2001) Prehistoric sunflower (*Helianthus annuus* L.) domestication in Mexico. *Economic Botany,* **55**, 370-376.

Matsuoka Y, Vigouroux Y, Goodman MM*, et al.* (2002) A single domestication for maize shown by multilocus microsatellite genotyping. *Proceedings of the National Academy of Sciences, USA,* **99**, 6080-6084.

Peakall R, Smouse PE (2006) Genalex 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes,* 6, 288-295.

Powell W, Morgante M, McDevitt R, Vendramin GG, Rafalski JA (1995) Polymorphic simple sequence repeat regions in chloroplast genomes: applications to the

population genetics of pines. *Proceedings of the National Academy of Sciences, USA,* **92**, 7759-7763.

Putt ED (1997) Early history of sunflower. In: *Sunflower Production and Technology* (ed. Scheiter AA), pp. 1-19. American Society of Agronomy, Madison, WI.

Rieseberg LH, Seiler GJ (1990) Molecular evidence and the origin and development of the domesticated sunflower (*Helianthus annuus*, Asteraceae). *Economic Botany* 44 (Supplement 3), 79-91.

Schuelke M (2000) An economic method for the fluorescent labeling of PCR fragments. *Nature Biotechnology,* **18**, 233-234.

Second G (1982) Origin of the genic diversity of cultivated Rice (*Oryza*-Spp): Study of the polymorphism scored at 40 isoenzyme loci. *Japanese Journal of Genetics,* **57**, 25-57.

Seiler GJ, Rieseberg LH (1997) Systematics, origin, and germplasm resources of the wild and domesticated sunflower. In: *Sunflower Production and Technology* (ed. A. SA), pp. 21-65. American Society of Agronomy, Madison, WI.

Slatkin M (1994) Linkage disequilibrium in growing and stable populations. *Genetics,* **137**, 331-336.

Spooner DM, McLean K, Ramsay G, Waugh R, Bryan GJ (2005) A single domestication for potato based on multilocus amplified fragment length polymorphism genotyping. *Proceedings of the National Academy of Sciences, USA,* **102**, 14694-14699.

Tang SX, Knapp SJ (2003) Microsatellites uncover extraordinary diversity in Native American land races and wild populations of cultivated sunflower. *Theoretical and Applied Genetics,* **106**, 990-1003.

Weising K, Gardner RC (1999) A set of conserved PCR primers for the analysis of simple sequence repeat polymorphisms in chloroplast genomes of dicotyledonous angiosperms. *Genome,* **42**, 9-19.

Wendel JF, Schnabel A, Seelanan T (1995) Bidirectional interlocus concerted evolution following allopolyploid speciation in cotton (*Gossypium*). *Proceedings of the National Academy of Sciences, USA,* **92**, 280-284.

Wills DM, Hester ML, Liu AZ, Burke JM (2005) Chloroplast SSR polymorphisms in the Compositae and the mode of organellar inheritance in *Helianthus annuus*. *Theoretical and Applied Genetics,* **110**, 941-947.

Xu DH, Abe J, Gai JY, Shimamoto Y (2002) Diversity of chloroplast DNA SSRs in wild and cultivated soybeans: Evidence for multiple origins of cultivated soybean. *Theoretical and Applied Genetics,* **105**, 645-653.

Table 2.1 Accession numbers and improvement status of wild populations and domesticated lines surveyed. Each haplotype was assigned a number representing its rank from most to least frequent, and the identity of the haplotypes found in each accession are reported. In addition, the PCO cluster to which each haplotype was assigned is included. See text for details.

Sample Location	Accession ID	Improvement Status	Sample Size	Haplotype ID	PCO Cluster
Arizona, USA	Ames 14400	Wild	6	2, 22	1, 3
Arkansas, USA	PI 613727	Wild	6	5	1,
California, USA	PI 613732	Wild	6	12, 29, 30	1, 2
Colorado, USA	PI 586840	Wild	6	3, 15, 25, 31, 32	1, 2
Iowa, USA	PI 597895	Wild	5	10	1
Illinois, USA	PI 547168	Wild	5	1, 23	1
Kansas, USA	PI 413027	Wild	6	13, 24	1
Minnesota, USA	PI 613745	Wild	6	2	1
Missouri, USA	PI 413011	Wild	5	11	1
Montana, USA	PI 531032	Wild	6	3, 21, 34, 35	1, 3

Table 2.1 Contiuned.

Sample Location	Accession ID	Improvement Status	Sample Size	Haplotype ID	PCO Cluster
North Dakota, USA	PI 596910	Wild	6	18, 19	1, 3
Nebraska, USA	PI586865	Wild	6	20, 26, 36	1, 2
Ohio, USA	Ames 23238	Wild	4	14	1
Oklahoma, USA	PI 435619	Wild	6	1	1
Oregon, USA	PI 531015	Wild	6	4	1
South Dakota, USA	Ames 23940	Wild	6	1, 5, 27, 43	1
Tennessee, USA	PI 435552	Wild	6	1	1
Texas, USA	Ames 7442	Wild	6	8	1
Utah, USA	PI 531009	Wild	6	9	1
Washington, USA	PI 531016	Wild	5	4, 15	1
Wyoming, USA	PI 586822	Wild	6	21, 28, 44, 45	1, other
Alberta, Canada	PI 592308	Wild	6	3, 16	1, 3

Table 2.1 Contiuned.

Sample Location	Accession ID	Improvement Status	Sample Size	Haplotype ID	PCO Cluster
Manitoba, Canada	PI 592327	Wild	5	17, 25, 33	1
Saskatchewan, Canada	PI 592317	Wild	6	37, 38, 39, 40, 41, 42	1, 2, 3, other
España, Mexico	PI 413067	Wild	6	7	other
Mayo, Mexico	PI 413123	Wild	6	6	3
Hopi	PI 432504	Domesticated	6	2	1
Havasupai	PI 369358	Domesticated	6	2	1
Seneca	PI 369360	Domesticated	4	2	1
Mandan	PI 600717	Domesticated	4	2	1
Hidatsa	PI 600721	Domesticated	4	46	1
Hidatsa[1]	PI 600720	Domesticated	6	46	1
Arikara	PI 369357	Domesticated	4	2	1
Maíz de Tejas	Ames 6859	Domesticated	4	2, 35	1

Table 2.1 Contiuned.

Sample Location	Accession ID	Improvement Status	Sample Size	Haplotype ID	PCO Cluster
Maíz Negro	Ames 19070	Domesticated	4	2	1
Mennonite	Ames 7574	Improved	4	2	1
Jupiter	PI 296289	Improved	4	2	1
Tchernianka Select W-13	PI 343794	Improved	4	2	1
Sunrise	PI 162454	Improved	5	2	1
Mammoth	PI 476853	Improved	5	2	1
Damaya	PI 496263	Improved	5	2	1

[1] Denotes the second Hidatsa accession that was surveyed to confirm the occurrence of a unique haplotype within this landrace.

Table 2.2 Results of our survey of cpSSR polymorphisms across twenty-six wild *Helianthus annuus* populations. Allele sizes are reported in basepairs, and reflect the inclusion of the 19 bp extension on the 5' end of the forward primer (see Materials and Methods for additional details).

Locus Name	**Primer Sequences (5'–3')**	**Allele Size Range**	**Number of Alleles**	**Gene Diversity**
ccmp 2	F: GATCCCGGACGTAATCCTG R: ATCGTACCGAGGGTTCGAAT	228-230	3	0.452
ccmp 7	F: CAACATATACCACTGTCAAG R: ACATCATTATTGTATACTCTTTC	139-149	9	0.771
NTCP 9	F: CTTCCAAGCTAACGATGC R: CTGTCCTATCCATTAAGACAATG	279-284	6	0.758
NTCP 18	F: CTGTTCTTTCCATGACCCCTC R: CCACCTAGCCAAGCCAGA	207-218	9	0.607
NTCP 30	F: GATGGCTCCGTTGCTTTAT R: TGCCGGAGAGTTCTTAACAATA	176-181	6	0.606
NTCP 40	F: TAATTTGATTCTTCGTCGC R: GATGTAGCCAAGTGGATCA	277-278	3	0.346
Mean		NA	6.0	0.590

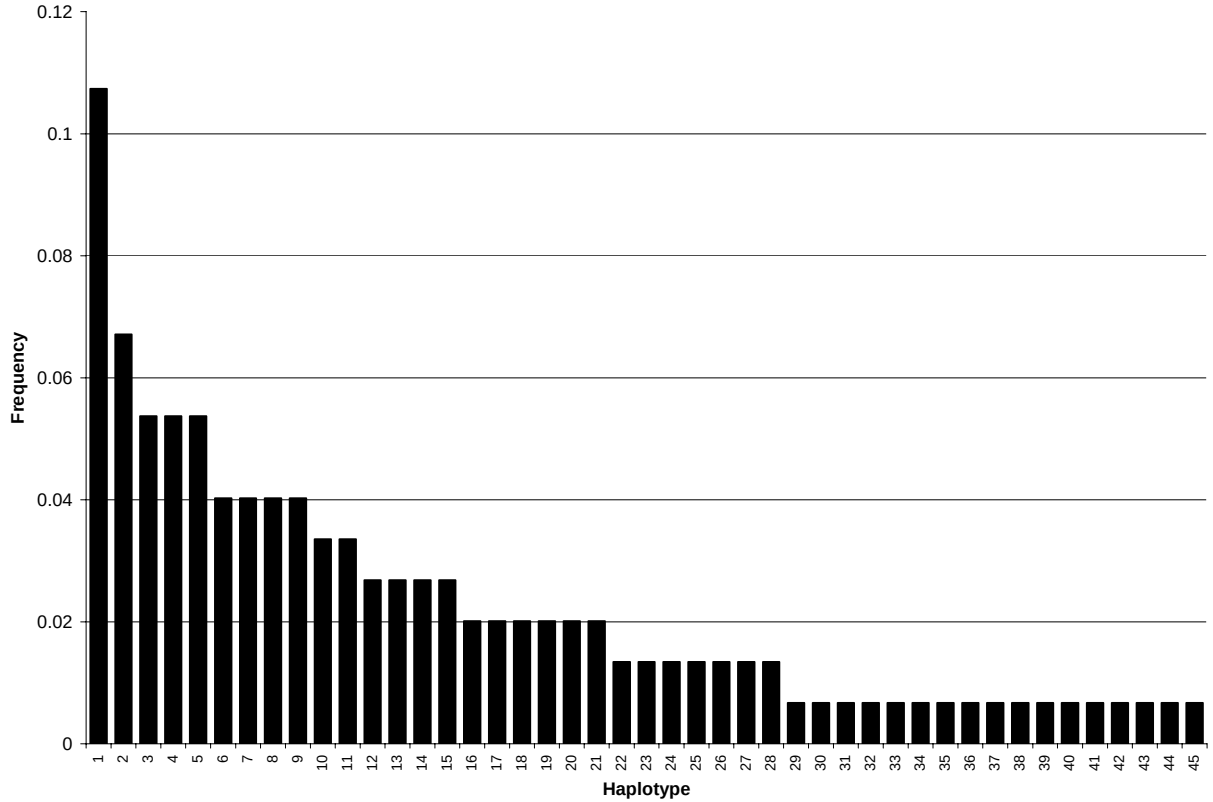

Figure 2.1 Frequency distribution of the forty-five unique cp SSR haplotypes identified in wild *Helianthus annuus*.

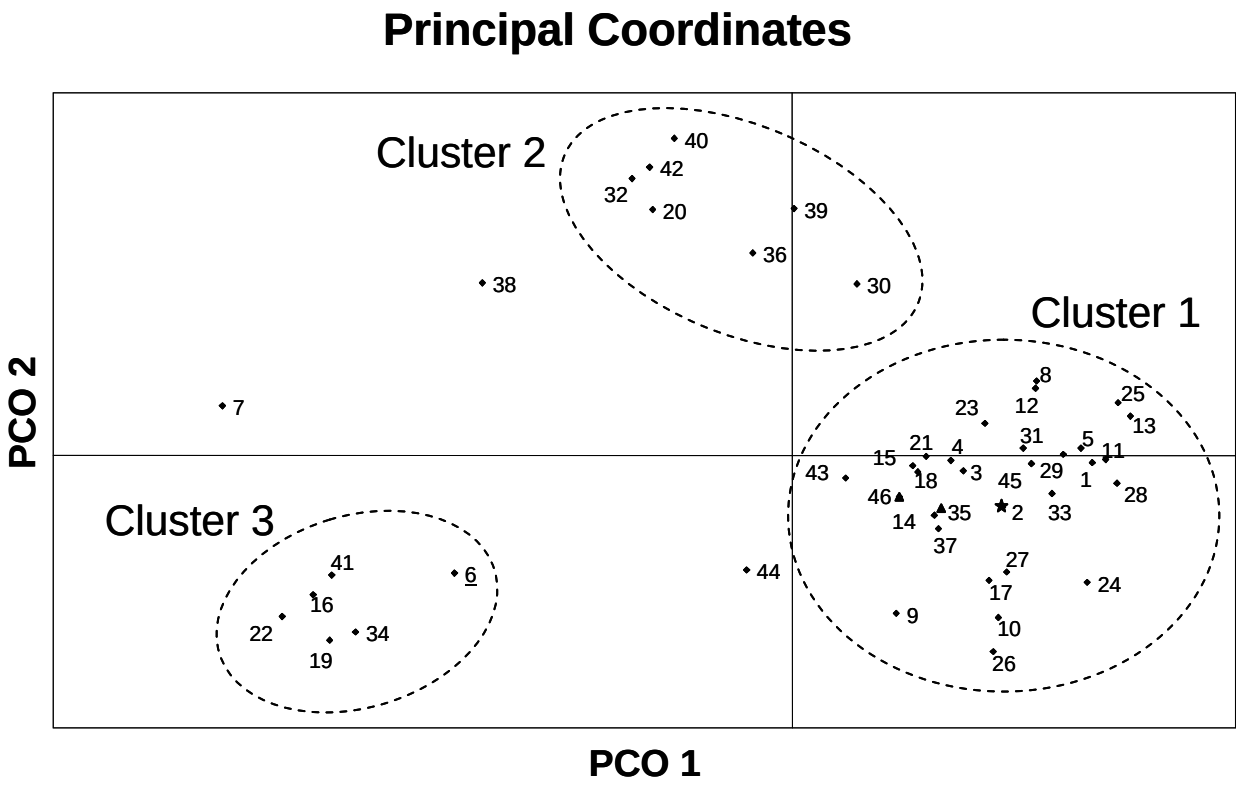

Figure 2.2 Principal coordinate (PCO) analysis of the cpSSR haplotypes observed in *Helianthus annuus*.

CHAPTER 3

QTL ANALYSIS OF THE EARLY DOMESTICATION OF SUNFLOWER[1]

[1] WILLS, D.M. and J.M. BURKE. QTL analysis of the early domestication of sunflower. *Genetics*. 2007. 176:2589-2599. Reprinted here with permission of the Genetics Society of America.

ABSTRACT

Genetic analyses of the domestication syndrome have revealed that domestication-related traits typically have a very similar genetic architecture across most crops, being conditioned by a small number of quantitative trait loci (QTL), each with a relatively large effect on the phenotype. To date, the domestication of sunflower (*Helianthus annuus* L.) stands as the only counterexample to this pattern. In previous work involving a cross between wild sunflower (also *H. annuus*) and a highly improved oilseed cultivar, we found that domestication-related traits in sunflower are controlled by numerous QTL, typically of small effect. To provide insight into the minimum genetic changes required to transform the weedy common sunflower into a useful crop plant, we mapped QTL underlying domestication-related traits in a cross between a wild sunflower and a primitive Native American landrace that has not been the target of modern breeding programs. Consistent with the results of the previous study, our data indicate that the domestication of sunflower was driven by selection on a large number of loci, most of which had small to moderate phenotypic effects. Unlike the results of the previous study, however, nearly all of the QTL identified herein had phenotypic effects in the expected direction, with the domesticated allele producing a more crop-like phenotype and the wild allele producing a more wild-like phenotype. Taken together, these results are consistent with the hypothesis that selection during the post-domestication era has resulted in the introduction of apparently maladaptive alleles into the modern sunflower gene pool.

INTRODUCTION

Plant domestication typically involves intense directional selection which produces large changes in quantitative traits, often accompanied by some degree of reproductive isolation between wild and domesticated taxa. Crop evolution thus allows for the investigation of basic evolutionary phenomena such as the phenotypic response of populations to long-term directional selection, the genetic consequences of recent selective sweeps, and the limitations imposed on selection response by genetic architecture (e.g., Stuber *et al.* 1980; Wang *et al.* 1999; Bost *et al.* 2001). Unlike researchers studying most wild systems, students of domestication often enjoy historical insights into the likely timing of selection, as well as the types of traits that have been subjected to selection. Common domestication traits include: increased seed or fruit size, more determinate growth and flowering, suppression of natural seed dispersal, and loss of self-incompatibility. Termed the "domestication syndrome" (Harlan 1992), these traits make crop plants easier to cultivate and result in more valuable products for human use.

Genetic analyses of the domestication syndrome have revealed that these traits have a similar genetic architecture across most crops (e.g., Doebley *et al.* 1990; Doebley and Stec 1991, 1993; Paterson *et al.* 1991; Koinange *et al.* 1996; Xiong *et al.* 1999). More specifically, crop-related traits are typically conditioned by a small number of quantitative trait loci (QTL), each with a relatively large effect on the phenotype (reviewed in Ross-Ibarra, 2005). Perhaps the most well known example of this is maize, wherein just five genomic regions account for the majority of the phenotypic differentiation between teosinte and maize (Doebley and Stec 1991, 1993). Doebley and

Stec (1991, p. 294) argued that if "evolution is opportunistic, one would predict that major shifts in the morphological traits of plants could be controlled by the full range of genetic mechanisms from few genes with large effects to many genes with small effects." They further argued that "The relative importance in plant evolution of these contrasting modes of inheritance remains to be determined."

Although wild populations have been shown to respond to selection in a variety of ways (e.g., Bradshaw *et al.* 1998; Fishman *et al.* 2002), the pattern of few QTL of large effect is nearly universal in crop plants. In fact, there is just one counterexample – the evolution of domesticated sunflower, *Helianthus annuus* L. (Burke *et al.* 2002). In terms of the phenotypic response to cultivation, sunflower is a very typical crop. Human-mediated selection has resulted in a dramatic increase in apical dominance relative to its wild progenitor (common sunflower; also *H. annuus*), an increase in seed size, and the loss of natural seed dispersal, seed dormancy, and self-incompatibility. However, when these traits were investigated at a genetic level in a cross between common sunflower and an elite oilseed cultivar, they were found to be under the control of a large number of QTL of predominantly minor effect, with only 5% of all QTL detected accounting for ≥ 25% of the segregating phenotypic variation. Traits of obvious importance for domestication, such as seed weight, branching, and shattering, all lacked QTL of major effect, and seed dormancy was later shown to be under similarly complex genetic control in a different crop × wild mapping population (Gandhi *et al.* 2005).

The foregoing results suggest that sunflower may indeed be an exception to the rule, thereby supporting the notion that evolution under domestication is an opportunistic process, making use of whatever genetic variation happens to be available (Doebley and

Stec 1991). However, a subsequent study of seed oil content and composition revealed that these original findings may have been influenced by the complex post-domestication breeding history of sunflower (Burke *et al.* 2005). In fact, it now seems clear that certain portions of the cultivated sunflower genome experienced post-domestication selective sweeps, meaning that the use of a modern inbred line as the cultivar parent in the original study likely confounded the effects of selection during domestication with the effects of selection during the subsequent period of breeding and improvement.

Here we report the results of an investigation of the genetic architecture of sunflower domestication utilizing a cross between common sunflower and a primitive Native American domesticate. This work is thus designed to provide insight into the genetic changes that were necessary for the initial transformation of the weedy, common sunflower into a useful crop plant. We have focused primarily on a suite of domestication-related traits that have previously been analyzed, and are thus able to make a direct comparison to the results of earlier research.

MATERIALS AND METHODS

Mapping Population

The mapping population described in this study was derived from a common × domesticated sunflower cross. The wild parent used in this cross was drawn from the same population (Ann1238) in Keith County, Nebraska that served as the source of the wild parent in the previous QTL analysis of sunflower domestication (Burke *et al.* 2002). This population is located within the same general range of the common sunflower that is

thought to have given rise to domesticated sunflower (Harter *et al.* 2004). The domesticated parent was the Hopi sunflower (USDA PI 432504), which was selected for analysis because it represents one of the two most primitive extant cultivated sunflower lineages (Harter *et al.* 2004; Tang and Knapp 2003; Wills and Burke 2006). A single, self-compatible F_1 individual from the initial wild × domesticated cross was self-pollinated to produce the F_2 generation. F_2 seeds were nicked with a razor blade and allowed to germinate on moist filter paper prior to being sown in flats. Seedlings were then transplanted into pots and grown under 16-hr days in the greenhouse. The final mapping population consisted of 378 F_2 individuals. Fifteen individuals each from the Hopi landrace and the Ann1238 population were grown along with the mapping population to estimate the phenotypic means of the parental lines when grown under these conditions.

Phenotypic Trait Measurements

Thirteen domestication-related traits that have been shown to differ between wild and domesticated sunflower were measured in all 378 F_2 plants as well as in the fifteen individuals from each parental line (or their selfed progeny in the case of seed traits; Table 3.1). The number of days to flowering was recorded for each individual. At flowering, the number of rays, disc diameter of the primary head, stem length, length and width of largest leaf, and stem diameter 3 cm above the soil were recorded for each individual. Leaf size was calculated as length × width. The primary head on each individual was bagged to prevent pollination from neighboring plants, and rubbed to ensure self-pollination until florets ceased to emerge (ca. nine days). Plants were maintained in the greenhouse until their seeds were mature, at which time the primary

head was harvested and the number of heads and branches were recorded for each individual. Primary heads were then dried for three days at 40º. To quantify shattering of the capitulum, the dried heads were dropped three times from a height of 12 cm. The total number of seeds released from the capitulum was then recorded, the heads were threshed, and the total seed output was recorded. Shattering was scored as the percentage of seeds released and 100-seed weight was estimated for each line. Seeds were then stored at 4º for three months. To quantify seed dormancy, twenty seeds from each selfed F_2 individual with sufficient seed output were sown in pots at a soil depth of 2 cm and allowed to germinate in a growth chamber under 16 hour days with constant bottom watering. The number of germinated seeds was recorded each day, and pots were monitored for 100 days. Seeds that failed to germinate during the course of this trial were scored as having germinated on the 100th day and the mean number of days until germination was calculated for each F_2 line.

Genotyping

Total genomic DNA was extracted from a sample of leaf tissue from each F_2 individual using the Qiagen DNeasy Plant Mini Kit (Qiagen, Valencia, CA). Genotyping for the genetic map was then carried out for 111 variable, codominant loci, including 108 simple-sequence repeats (SSRs) that were previously mapped in sunflower (Yu *et al.* 2003; Tang *et al.* 2003; Tang *et al.* 2002; Lai *et al.* 2005). The SSRs were fluorescently labeled with 6FAM, TET, HEX, or VIC either by direct labeling of the 5' end of the forward primer, or using a modification of the three-primer PCR methodology presented by Schuelke (2000), previously adapted for sunflower by Wills *et al.* (2005). This technique involves incorporation of an arbitrarily selected sequence (the M13 Forward [-

29] sequencing primer, 5'-CAC GAC GTT GTA AAA CGA C-3') to the 5′ end of the forward primer. PCR products are then labeled by including a fluorescently-tagged (6FAM, TET, or HEX) M13 Forward (-29) primer in the reaction mixture. All reactions were performed in 10 µl total volume containing 10 ng of template DNA, 30 mM Tricine pH 8.4-KOH, 50 mM KCl, 2mM $MgCl_2$, 100 µM of each dNTP, 0.02 µM forward primer, 0.1 µM of both the reverse primer and the fluorescently labeled M13F primer, and 2 units of *Taq* polymerase. When PCR was carried out with directly-labeled primer pairs, the M13F primer was left out and the forward primer was increased to 0.1 µM. Cycling conditions were as follows: initial denaturation at 95° for 3 min, followed by ten cycles of 30 s at 94°, 30 s at 58° (annealing temperature was reduced by one degree per cycle), 45 s at 72°, followed by 30 cycles of 30 s at 94°, 30 s at 48°, 45 s at 72°, and a final extension time of 20 min at 72°.

Amplification products were visualized on either an MJ Research BaseStation automated DNA sequencer (South San Francisco, CA) or an Applied Biosystems 3730xl DNA Analyzer (Foster City, CA). MapMarker® 1000 ROX size standard (BioVentures Inc., Murfreesboro, TN) was included in each lane to allow for accurate determination of fragment size. Alleles were called using the software package Cartographer (MJ Research) for the BaseStation runs or GeneMarker (SoftGenetics, State College, PA) for the 3730 data. The final map included three additional, previously unpublished markers: HT39, HT135, and HT1490 (S. Tang and S.J. Knapp, unpublished data). HT39 amplicons were visualized via SSCP gel electrophoresis using 0.5X MDE™ gels that were run for 14 hours at 4 watts (Slabaugh *et al.* 1997) followed by silver staining (Sanguinetti *et al.* 1994). HT135 exhibited a length polymorphism in this cross and was scored in the same

manner as the SSRs described above. HT1490 was not length polymorphic, and could not be reliably scored via SSCP analysis. Thus, this locus was sequenced and a restriction polymorphism corresponding to an *Fnu*4HI restriction site was found to be segregating within the mapping population. The forward primer was therefore 5' end-labeled with 6FAM and each individual was amplified as described above. All PCR amplicons were then digested at 37° overnight with one unit of *Fnu*4HI (New England Biolabs, Inc., Ipswich, MA). The PCR-RFLP products were then run on an Applied Biosystems 3730xl and scored using GeneMarker.

Map Construction

The linkage map was constructed using MAPMAKER 3.0/EXP (Lander *et al.* 1987; Lincoln and Lander 1992). Initial linkage groups were identified using the 'group' command with LOD > 5.0 and $\theta < 0.2$. Preliminary map orders within groups were then set based on the results from previous sunflower mapping studies (Burke *et al.* 2002; Yu *et al.* 2003; Tang *et al.* 2003; Tang *et al.* 2002; Lai *et al.* 2005). Final map orders were then confirmed using the 'ripple' and 'compare' commands such that the map orders presented herein reflect the statistically most likely order based on the data at hand.

QTL Analysis

The initial QTL analysis followed the same general approach as outlined by Burke *et al.* (2002). Because shattering was scored as a proportion, the data for this trait were arcsine-square root transformed prior to analysis using JMP 4 (SAS Institute Cary, NC). Composite interval mapping (CIM; Zeng 1993, 1994) was then performed as implemented by the program Zmapqtl (model 6) of the software package QTL Cartographer version 1.17 (Basten *et al.* 1994, 2004). CIM was run with a 10-cM window

and five background cofactors. Tests were performed at 2-cM intervals, and cofactors were selected via forward-backward stepwise regression using the program SRmapqtl. Genome-wide threshold values ($\alpha = 0.05$) for declaring the presence of QTL were estimated from 1000 permutations for each trait (Churchill and Doerge 1994; Doerge and Churchill 1996). A likelihood-ratio decline of 9.21 (equivalent to a LOD decline of 2.0) between adjacent peaks on a linkage group was taken as evidence of multiple, linked QTL and one-LOD support limits for the position of each QTL were calculated from the CIM results. The degree of dominance of the Hopi allele at each locus was calculated as the dominance effect divided by the additive effect (*d*/*a)*, and the following arbitrary thresholds were used to classify the mode of gene action for each QTL: underdominant $\leq$ -1.25 < recessive $\leq$ -0.75 < partially recessive $\leq$ -0.25 < additive $\leq$ 0.25 < partially dominant $\leq$ 0.75 < dominant < 1.25 $\leq$ overdominant. Finally, to allow a direct comparison to the results of Burke *et al.* (2002), we used arbitrary PVE thresholds of 10% and 25% to classify QTL as having "minor," "intermediate," or "major" effects.

Multiple interval mapping (MIM; Kao and Zeng 1997; Kao *et al.* 1999) was then used to search for epistatic interactions amongst the QTL identified via CIM. The CIM results were used as the initial model for the MImapqtl module in QTL Cartographer (Basten *et al.* 1994, 2004), and the maximum number of allowable pairwise interactions was set to 19. Only those interactions that significantly improved the fit of the model were retained. As recommended by the authors, significance was determined based on the information criterion IC(k) = -2(log(L) - k c(n) / 2) with c(n) = log(n) as the penalty function and a threshold of 0.0.

RESULTS

Linkage Analysis

The map coalesced into the expected seventeen linkage groups, and covered a total of 906.4 cM with an average intermarker distance of 8.2 cM. As has been previously observed for common × cultivated sunflower crosses (e.g., Burke *et al.* 2002), this map showed evidence of suppressed recombination, with common marker intervals exhibiting nearly 20% compression when compared against the sunflower reference map, which is based on a cross between two elite inbred lines (RHA280 × RHA801; Tang *et al.* 2002; S. Tang and S.J. Knapp, unpublished data). Based on a comparison of shared markers, coverage of the map described herein is equal to or exceeds that of the map constructed by Burke *et al.* (2002) for sixteen of the seventeen linkage groups (LGs). The one exception was a portion of the top of LG 17, which we were unable to cover due to a lack of polymorphism. However, no QTL have been detected previously in this region, so this small gap in coverage is unlikely to influence our overall findings.

QTL analysis

For the thirteen domestication-related traits that we analyzed, CIM detected 61 QTL (Table 3.2; Figure 3.1). The number of QTL per trait ranged from 2 to 8 (mean = 4.7; Figure 3.2A) with shattering and the number of heads produced being the only traits with multiple QTL on a single linkage group. QTL were found on all linkage groups with the exception of LG 11, and the one-LOD support intervals, which provide an approximate confidence interval for the true location of a given QTL, ranged from 1.1 to 30.8 cM (mean = 13.0 cM). Multiple, overlapping QTL were observed on most linkage groups;

the exceptions were LG 2, LG 3, and LG 4, as well as LG 11 which (as noted above) was devoid of QTL.

As previously documented, there was a paucity of QTL of major effect. Individual QTL explained 2.5% – 59.0% of the observed phenotypic variance for a particular trait, but only four QTL had PVE ≥ 25% (Figure 3.2B). Three of these major-effect QTL co-occur on the bottom of LG 15 and influence days to flower, plant height, and number of leaves along the main stem. The only other QTL of major effect is located on LG 10 and influences the number of heads produced. Contrary to previous findings, the majority of domestication-related QTL identified here (56 of 61) had effects in the expected direction (Figure 3.2C). That is, the Hopi allele produced a more crop-like phenotype, and the wild allele produced a more wild-like phenotype. The exceptions were QTL for shattering on LG 4, number of heads produced on LG 6, days to flower and number of leaves along the main stem on LG 7, and leaf size on LG 16. All five QTL with effects in 'wrong' direction were minor, explaining from 2.5 to 6.4 % of the phenotypic variance. In terms of the mode of gene action, the Hopi allele at each locus exhibited a dominance ratio (d/a) ranging from -2.19 to 11.47 with a mean of 0.38 (Table 3.2; Figure 3.2D).

Results of the search for epistasis amongst significant QTL using MIM are presented in Table 3.3. There were 30 significant pairwise interactions across the thirteen traits. For the most part, these interactions were minor, and their phenotypic effects were mixed. Indeed, 28 of 30 interactions had an effect of less than 5%, and half were in the expected direction, whereas the other half were in the 'wrong' direction. The two exceptions were interactions for branching and seed germination. In the former case, there was a dominant x additive interaction with an effect of 9.3% between QTL on LG

13 and LG 17, whereas the latter case had an additive x additive interaction with an effect of 11.5%. In both cases, the interactions acted in the expected direction, with Hopi/Hopi genotypes producing an even more Hopi-like phenotype.

DISCUSSION

QTL Numbers and Magnitudes of Effect

In general terms, the results of this study confirm that sunflower is an exception to the rule in that its domestication involved changes at a large number of loci, each of relatively small effect (Figures 3.2A and 3.2C). Indeed, we analyzed 13 traits and identified a total of 61 QTL, only 4 of which had PVE > 25%. In contrast, the domestication of crops such as maize (Doebley and Stec 1991, 1993), rice (Xiao *et al.* 1998; Xiong *et al.* 1999), and beans (Koinange *et al.* 1996) were all driven by relatively major changes at a much smaller number of loci. This observed lack of major QTL suggests that the transition from wild to domesticated sunflower was relatively smooth with few major phenotypic leaps.

Gene Action and Interaction

In terms of the predominant mode of gene action, our results mirror those of Burke *et al.* (2002), and stand in stark contrast to the view that domestication is generally driven by recessive genetic changes (e.g., Ladizinsky 1985; Lester 1989). In fact, inspection of Figure 3.2B reveals a preponderance of non-recessive QTL, suggesting that selection during domestication likely resulted in a rapid phenotypic response, as most of these QTL would have been at least partially visible to selection, even when rare. These findings are

in accord with previous QTL results from other taxa, including tomato (Paterson *et al.* 1991) and maize (Doebley *et al.* 1994).

With regard to the role of gene interaction in domestication, our results provide somewhat limited evidence of epistasis. When combined with the overall lack of epistasis documented by Burke *et al.* (2002), these results suggest that neither the initial domestication of sunflower nor its subsequent improvement relied heavily upon the fixation of favorably interacting gene complexes. While MIM detected significant QTL × QTL interactions for ten of the thirteen traits, the vast majority of these interactions had effects of 5% or less (Table 3.3). The exceptions to this were an interaction between two branching QTL located on LG 13 and LG 17, and an interaction between the two seed dormancy QTL on LG 12 and LG 15. This latter case, which involves a synergistic, additive × additive interaction between two QTL of intermediate effect (PVE = 17.3% and 17.8%, respectively), is particularly noteworthy because seed dormancy was not previously analyzed by Burke *et al.* (2002). Gandhi *et al.* (2005) did, however, map QTL related to seed dormancy in a different elite × wild cross and identified a QTL of intermediate effect in the same region of LG 15, as well as two other QTL that were not recovered here; no significant epistatic interactions were detected among those QTL.

QTL Concordance

Despite the foregoing similarities between our results and those of Burke *et al.* (2002), a direct comparison of QTL locations reveals a relatively low level of concordance. Indeed, comparing the 59 QTL identified in this study (ignoring the two seed dormancy QTL because this trait was not previously analyzed) to the 56 QTL previously identified for this same suite of traits reveals only fifteen cases in which QTL for the same trait mapped

to the same linkage group in both studies. Twelve of these cases involved QTL with overlapping 1-LOD support intervals, suggesting that the same QTL was detected in both crosses, two showed clear evidence of non-overlap, and in one case the degree of overlap could not be determined due to a paucity of shared markers (LG 5). This relatively low rate of correspondence between studies is likely due to a combination of factors, including differences between the parents used in each cross, QTL × environment interactions, and difficulties associated with reliably detecting QTL of small effect (Beavis 1994).

Because sunflower is an annual plant, it was impossible to use the same wild individual in both the present and previous crosses. Moreover, wild sunflower is an obligate outcrosser which exhibits high levels of heterozygosity (e.g., Ivanov 1975; Fernandez-Martinez and Knowles 1978; Tang and Knapp 2003; Harter *et al.* 2004). Thus, in an attempt to minimize problems with intra-taxon polymorphism and maintain continuity with previous work, the wild parent for the present cross was drawn from the same population that Burke *et al.* (2002) utilized. Although variation is evident in the wild for all of the traits in question, the phenotypic differences between wild and cultivated sunflower are largely consistent across environments. Despite this, it is still possible that some of the differences between the two studies resulted from allelic variation between the wild parents used in the two studies.

A more likely explanation is that a sizable fraction of the differences result from the cultivar parents used in the two studies having very different evolutionary histories; in fact, this was the primary motivation of the present study. The cultivar parent utilized by Burke *et al.* (2002) was a highly improved, elite oilseed line that has subsequently been

found to bear the signature of post-domestication selective sweeps, presumably due to selection on oil-related characters (Burke *et al.* 2005). In contrast, the cultivar parent used in the present study is a primitive Native American landrace. The results presented herein should, therefore, provide a much more accurate picture of the genetic changes necessary to transform wild into domesticated sunflower, as they are largely free from the confounding effects of improvement subsequent to the initial domestication event. In this context, it is worth noting that LG 6 has previously been shown to a harbor a large cluster of QTL that mostly have effects in the 'wrong' direction (Burke *et al.* 2002). Subsequent work has suggested that at least some of these QTL arose as a byproduct of selection during sunflower improvement (Burke *et al.* 2005), and our results are fully consistent with this hypothesis. Indeed, only a subset of the QTL that were initially identified were recovered in the present analysis, and all but one of the QTL on this linkage group now have effects in the expected direction.

Another key difference between the cultivars utilized in these two studies is that they are adapted to relatively different habitats. As such, some of the QTL that have been identified in just one population are likely to reflect differences in local adaptation. Most notable in this context is flowering time (and associated traits, such as height and number of stem leaves) in the Hopi × wild cross analyzed here. The Hopi landrace exhibits late flowering, presumably as an adaptation to the extremely long growing season of the desert southwest. Our results indicate that this flowering time difference is conditioned by a major QTL at the bottom of LG 15 which, as one might expect, was not present in the previous cross.

As noted above, other factors that could account for the relatively low level of QTL concordance are QTL x environment interactions and the difficulties associated with reliably detecting QTL of small effect. With regard to QTL x environment interactions, it has previously been shown that individual QTL can vary in their degree of environmental sensitivity, with some QTL being robust across environments while others can only be detected under certain conditions (e.g., Paterson *et al.* 1991; Paterson *et al.* 2003). Thus, even though both populations were greenhouse grown, and the traits of interest are reasonably robust across environments, it is conceivable that some fraction of the QTL were detected in one study but not the other because of differences in growing conditions.

Regarding the issue of detectability, it is well known that QTL of minor effect suffer a higher false-negative rate as compared to QTL of major effect (Beavis 1994). Indeed, Doebley and Stec (1993) found much higher agreement in QTL locations in a comparison between two teosinte × maize populations for QTL of major effect (81% concordance for QTL with $r^2 \geq 20\%$) as compared to QTL of intermediate or minor effect (55% and 28% concordance for QTL with $10\% \leq r^2 \leq 20\%$ and QTL with $r^2 < 10\%$, respectively). Given the typically small effect sizes associated with QTL identified in both the present and previous analyses, the relatively low QTL concordance is therefore not surprising. Consistent with this idea is the fact that the handful of major QTL identified by Burke *et al.* (2002), including QTL for flowering time and the number of stem leaves on LG 6 and the number of selfed seeds on LG 17, were all recovered in the present study. The key difference is that the estimated effect sizes for all of these QTL were much lower in the present study. In the case of flowering time and the number of

stem leaves, the reduced PVE in the Hopi × wild cross is likely due to an overall increase in phenotypic variance for these traits within the mapping population (data not shown) due to the adaptation of the Hopi sunflower to the long growing seasons of the desert southwest (see above). In the case of selfed seed production, the prior identification of two QTL at the bottom of LG 17 (Burke *et al.* 2002) has subsequently been shown to be an artifact of inconsistent locus ordering; this region is now believed to harbor the S-locus (Gandhi *et al.* 2005). Values reported in Figure 3.2 from the earlier study have been adjusted to account for the reordering of these markers. The low PVE associated with this locus in the present analysis is potentially an artifact of extreme segregation distortion in this region (all four markers on this linkage group deviate significantly from the expected segregation ratios, with all $P < 0.001$).

Conversely, the four QTL of major effect identified in the present study had not been previously identified. This result is not surprising for the QTL related to flowering time that are located near the bottom of LG 15, as they are likely a byproduct of adaptation of the Hopi landrace to local growing conditions. In the one remaining case (number of heads produced; LG 10), however, this is somewhat unexpected. Indeed, this QTL explains 28% of the segregating phenotypic variance and co-localizes with the *B* locus, which is known to influence apical branching (Tang *et al.* 2006). In fact, the region harboring the *B* locus is known to have manifold effects, influencing not only plant architecture but also achene/seed morphology. In the present analysis, this QTL is embedded within a larger cluster of loci that influence apical dominance as well as leaf and disc morphology, seed weight, and shattering. In fact, Burke *et al.* (2002) also found QTL related to seed size in this vicinity, but none related to branching. One possible

explanation for this is that, despite their simple inheritance in crosses between cultivars, branching-related traits in wild × cultivar crosses are thought to be genetically complex (Burke *et al.* 2002), and may well be influenced by genetic background.

QTL Directionality and Evidence of Selection

Perhaps the greatest departure between our results and those of Burke *et al.* (2002) relates to the directionality of QTL effects (Fig. 2D). As previously noted, nearly one-third of all QTL identified by Burke *et al.* (2002) had effects in the 'wrong' direction, with the wild allele producing a more crop-like phenotype and vice versa. As noted above, however, a subsequent analysis of improvement-related traits (i.e., seed oil content and composition) combined with a population genetic scan for selection suggested that this result was due to post-domestication selection and breeding (Burke *et al.* 2005). The low frequency of 'wrong-way' QTL in the present study (just 5 of 61 QTL had such effects) is fully consistent with this hypothesis. While data on QTL directionality can be used to statistically test for past directional selection (Orr 1998), the power of this approach is limited by QTL numbers. Thus, following the methods of Rieseberg *et al.* (2002), we pooled our data across traits and tested for selection on the domestication syndrome as a whole. In this case, the results were highly significant ($P < 0.001$), providing clear evidence that sunflower domestication was driven by consistent directional selection on a wide variety of traits.

Summary and Conclusions

Our results confirm that the domestication of sunflower was driven by selection on a large number of loci, most of which had small to moderate phenotypic effects. The underlying cause of this departure from the 'typical' genetic architecture of domestication,

however, remains a mystery. For example, while sunflower is an ancient polyploid (Adams and Wendel 2005; Sossey-Alaoui *et al.* 1998), and thus potentially exhibits high levels of genetic redundancy across the genome, this sort of redundancy alone cannot be the explanation. Indeed, virtually all major crops have experienced large-scale genome duplication at some point in their evolutionary history. Another possibility is that the population bottleneck leading to domesticated sunflower was less severe than that which occurred during the evolution of other crop lineages, resulting in a relatively large effective population size during domestication. This, in turn, would allow for selection to target mutations of minor effect more efficiently in sunflower than in other crop lineages. While the available data indicate that sunflower suffered a similar loss in genetic diversity during domestication as compared to other crop plants (e.g., Liu and Burke 2006) – a fact which argues against the idea that sunflower experienced a relatively mild bottleneck – a better understanding of the dynamics of the sunflower domestication bottleneck awaits more rigorous analysis. Ultimately, it may be that Doebley and Stec (1991) had it right; evolution under domestication may simply be an opportunistic process that makes use of whatever genetic variation happens to be available. Whether or not this lack of suitable mutations of major effect in sunflower reflects some sort of fundamental genomic constraint remains an open question.

ACKNOWLEDGEMENTS

We thank Tamara Berthel, Chris Buckner, Mark Chapman, Michael Foster, Meghan Jennings, Natasha Sherman, Ben Stephens, and Jessica Wenzler for assistance in the

greenhouse and/or lab, and Dörthe Dräger, Adam Heesacker, Steve Knapp, and Shunxue Tang for technical assistance with genotyping and access to unpublished marker data. We also thank Jim Birchler, Jutta Burger, Mark Chapman, Aizhong Liu, Natasha Sherman, and an anonymous reviewer for comments on a previous version of this manuscript. This work was supported by grants from the United States Department of Agriculture Cooperative State Research, Education and Extension Service National Research Initiative (CSREES-NRI) Plant Genome Program (Award #03-35300-13104 to JMB) and the National Science Foundation (Award #DBI-0332411 to JMB).

LITERATURE CITED

Adams KL, Wendel JF (2005) Polyploidy and genome evolution in plants. *Current Opinion in Plant Biology,* **8**, 135-141.

Basten CJ, Weir BS, Zeng Z-B (1994) Zmap-a QTL cartographer. In: *Proceedings of the 5th World Congress on Genetics Applied to Livestock Production: Computing Strategies and Software* (eds. Smith C, Gavora JS, Benkel B*, et al.*), pp. 65-66. Organizing Committee, 5th World Congress on Genetics Applied to Livestock Production, Guelph, Ontario, Canada.

Basten CJ, Weir BS, Zeng Z-B (2004) *QTL Cartographer, Version 1.17.* Department of Statistics, North Carolina State University, Raleigh, NC.

Beavis WD (1994) The power and deceit of QTL experiments: lesson from comparative QTL studies. *In: Proceedings of the 49th Annual Corn and Sorghum Research Conference*, pp. 250-266. American Seed Trade Association.

Bost B, de Vienne D, Hospital F, Moreau L, Dillmann C (2001) Genetic and nongenetic bases for the L-shaped distribution of quantitative trait loci effects. *Genetics,* **157**, 1773-1787.

Bradshaw HD, Otto KG, Frewen BE, McKay JK, Schemske DW (1998) Quantitative trait loci affecting differences in floral morphology between two species of monkeyflower (*Mimulus*). *Genetics,* **149**, 367-382.

Burke JM, Knapp SJ, Rieseberg LH (2005) Genetic consequences of selection during the evolution of cultivated sunflower. *Genetics,* **171**, 1933-1940.

Burke JM, Tang S, Knapp SJ, Rieseberg LH (2002) Genetic analysis of sunflower domestication. *Genetics,* **161**, 1257-1267.

Churchill GA, Doerge RW (1994) Empirical threshold values for quantitative trait mapping. *Genetics,* **138**, 963-971.

Doebley J, Bacigalupo A, Stec A (1994) Inheritance of kernel weight in 2 maize-teosinte hybrid populations: implications for crop evolution. *Journal of Heredity,* **85**, 191-195.

Doebley J, Stec A (1991) Genetic analysis of the morphological differences between maize and teosinte. *Genetics,* **129**, 285-295.

Doebley J, Stec A (1993) Inheritance of the morphological differences between maize and teosinte: comparison of results for 2 F_2 populations. *Genetics,* **134**, 559-570.

Doebley J, Stec A, Wendel J, Edwards M (1990) Genetic and morphological analysis of a maize-teosinte F_2 population: Implications for the origin of maize. *Proceedings of the National Academy of Sciences, USA,* **87**, 9888-9892.

Doerge RW, Churchill GA (1996) Permutation tests for multiple loci affecting a quantitative character. *Genetics* **142**, 285-294.

Fernandez-Martinez J, Knowles PF (1978) Inheritance of self-incompatibility in wild sunflower. In: *8th International Sunflower Conference*, pp. 484-489. International Sunflower Association, Paris.

Fishman L, Kelly AJ, Willis JH (2002) Minor quantitative trait loci underlie floral traits associated with mating system divergence in *Mimulus*. *Evolution,* **56**, 2138-2155.

Gandhi SD, Heesacker AF, Freeman CA, *et al.* (2005) The self-incompatibility locus (*S*) and quantitative trait loci for self-pollination and seed dormancy in sunflower. *Theoretical and Applied Genetics*, **111**, 619-629.

Harlan JR (1992) *Crops & Man*, 2nd edn. American Society of Agronomy, Madison, WI.

Harter AV, Gardner KA, Falush D, *et al.* (2004) Origin of extant domesticated sunflowers in eastern North America. *Nature,* **430**, 201-205.

Ivanov IG (1975) Study on compatibility and incompatibility display in crossing selfed sunflower lines. *Rastenievud Nauki,* **12**, 36-40.

Kao CH, Zeng ZB (1997) General formulas for obtaining the MLEs and the asymptotic variance-covariance matrix in mapping quantitative trait loci when using the EM algorithm. *Biometrics,* **53**, 653-665.

Kao CH, Zeng ZB, Teasdale RD (1999) Multiple interval mapping for quantitative trait loci. *Genetics,* **152**, 1203-1216.

Koinange EMK, Singh SP, Gepts P (1996) Genetic control of the domestication syndrome in common bean. *Crop Science,* **36**, 1037-1045.

Ladizinsky G (1985) Founder effect in crop plant evolution. *Economic Botany,* **39**, 191-199.

Lai Z, Livingstone K, Zou Y, *et al.* (2005) Identification and mapping of SNPs from ESTs in sunflower. *Theoretical and Applied Genetics,* **111**, 1532-1544.

Lander ES, Green P, Abrahamson J, *et al.* (1987) MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. *Genomics,* **1**, 174-181.

Lester RN (1989) Evolution under domestication involving disturbance of genic balance. *Euphytica,* **44**, 125-132.

Lincoln SE, Lander ES (1992) Systematic detection of errors in genetic linkage data. *Genomics,* **14**, 604-610.

Liu AZ, Burke JM (2006) Patterns of nucleotide diversity in wild and cultivated sunflower. *Genetics* **173**, 321-330.

Orr HA (1998) Testing natural selection vs. genetic drift in phenotypic evolution using quantitative trait locus data. *Genetics,* **149**, 2099-2104.

Paterson AH, Damon S, Hewitt JD*, et al.* (1991) Mendelian factors underlying quantitative traits in tomato: Comparison across species, generations, and environments. *Genetics,* **127**, 181-197.

Paterson AH, Saranga Y, Menz M, Jiang CX, Wright RJ (2003) QTL analysis of genotype x environment interactions affecting cotton fiber quality. *Theoretical and Applied Genetics,* **106**, 384-396.

Rieseberg LH, Widmer A, Arntz AM, Burke JM (2002) Directional selection is the primary cause of phenotypic diversification. *Proceedings of the National Academy of Sciences, USA,* **99**, 12242-12245.

Ross-Ibarra J (2005) Quantitative trait loci and the study of plant domestication. *Genetica* **123**, 197-204.

Sanguinetti CJ, Neto ED, Simpson AJG (1994) Rapid silver staining and recovery of PCR products separated on polyacrylamide gels. *Biotechniques,* **17**, 914-921.

Schuelke M (2000) An economic method for the fluorescent labeling of PCR fragments. *Nature Biotechnology,* **18**, 233-234.

Slabaugh MB, Huestis GM, Leonard J, *et al.* (1997) Sequence based genetic markers for genes and gene families: Single strand conformational polymorphisms for the fatty acid synthesis genes of Cuphea. *Theoretical and Applied Genetics,* **94**, 400-408.

Sossey-Alaoui K, Serieys H, Tersac M, *et al.* (1998) Evidence for several genomes in *Helianthus*. *Theoretical and Applied Genetics,* **97**, 422-430.

Stuber CW, Moll RH, Goodman MM, Schaffer HE, Weir BS (1980) Allozyme frequency changes associated with selection for increased grain yield in maize (*Zea Mays* L). *Genetics,* **95**, 225-236.

Tang S, Yu JK, Slabaugh MB, Shintani DK, Knapp SJ (2002) Simple sequence repeat map of the sunflower genome. *Theoretical and Applied Genetics,* **105**, 1124-1136.

Tang SX, Kishore VK, Knapp SJ (2003) PCR-multiplexes for a genome-wide framework of simple sequence repeat marker loci in cultivated sunflower. *Theoretical and Applied Genetics,* **107**, 6-19.

Tang SX, Knapp SJ (2003) Microsatellites uncover extraordinary diversity in Native American land races and wild populations of cultivated sunflower. *Theoretical and Applied Genetics,* **106**, 990-1003.

Tang SX, Leon A, Bridges WC, Knapp SJ (2006) Quantitative trait loci for genetically correlated seed traits are tightly linked to branching and pericarp pigment loci in sunflower. *Crop Science,* **46**, 721-734.

Wang R-L, Stec A, Hey J, Lukens L, Doebley J (1999) The limits of selection during maize domestication. *Nature,* **398**, 236-239.

Wills DM, Burke JM (2006) Chloroplast DNA variation confirms a single origin of domesticated sunflower (*Helianthus annuus* L.). *Journal of Heredity,* **97**, 403-408.

Wills DM, Hester ML, Liu AZ, Burke JM (2005) Chloroplast SSR polymorphisms in the Compositae and the mode of organellar inheritance in *Helianthus annuus*. *Theoretical and Applied Genetics,* **110**, 941-947.

Xiao JH, Li JM, Grandillo S, *et al.* (1998) Identification of trait-improving quantitative trait loci alleles from a wild rice relative, *Oryza rufipogon*. *Genetics,* **150**, 899-909.

Xiong LX, Liu KD, Dai XK, Xu CG, Zhang QF (1999) Identification of genetic factors controlling domestication-related traits of rice using an F_2 population of a cross between *Oryza sativa* and *O. rufipogon*. *Theoretical and Applied Genetics,* **98**, 243-251.

Yu JK, Tang S, Slabaugh MB, *et al.* (2003) Towards a saturated molecular genetic linkage map for cultivated sunflower. *Crop Science,* **43**, 367-387.

Zeng ZB (1993) Theoretical basis for separation of multiple linked gene effects in mapping quantitative trait loci. *Proceedings of the National Academy of Sciences, USA,* **90**, 10972-10976.

Zeng ZB (1994) Precision mapping of quantitative trait loci. *Genetics,* **136**, 1457-1468.

Table 3.1 Comparison of 14 traits between a primitive sunflower domesticate (the Hopi landrace) and its wild progenitor (*Helianthus annuus* var. *annuus*). All values are expressed as mean ± standard error.

Trait	**Hopi Landrace**	**Common sunflower**
Days to Flower	100.0 ± 4.4	84.8 ± 4.4
Stem Diameter (mm)	21.7± 0.7	10.8 ± 0.7
Height (cm)	358 ± 14.0	171 ± 16.9
No. Main Stem Leaves	47.5 ± 1.0	21.7 ± 1.9
Leaf Size (cm^2)	687 ± 23.8	335 ± 27.7
No. Branches	0.4 ± 0.2	9.4 ± 1.6
No. Heads	1 ± 0.0	4.2 ± 0.6
Disk Diameter (mm)	75.3 ± 4.3	23.3 ± 1.9
No. Ray Flowers	42.5 ± 3.8	23.7 ± 1.6
Self-Compatibility	Yes	No
Achene Weight (g per 100)	2.9 ± 1.1	0.6 ± 0.1
Shattering	No	Yes
Seed Dormancy	No	Yes

Table 3.2 Putative QTL positions, effect magnitudes, and modes of gene action for 13 traits using composite interval mapping in an F_2 population derived from a cross between a primitive sunflower domesticate (the Hopi landrace) and its wild progenitor (*H. annuus* var. *annuus*).

Trait	Linkage Group[a]	Position[b]	Nearest Marker	1-LOD Interval[c]	Additive Effect[d]	Dominance Ratio[e]	PVE[f]	Previously Identified?[g]
Days to Flower	6	57.6	ORS483	53.6-57.7	4.6	-0.23	7.6	Yes
	7	1.0	ORS1041	0-5.3	-2.6	-0.11	2.5	Yes
	15	57.1	ORS687	57-58.2	10.4	-0.49	46.9	No
Stem Diameter	1	7.0	HT1018	4.6-10.4	1.2	0.72	10.0	Yes
	2	1.7	ORS925	0-15.0	0.7	-0.03	3.0	No
	3	3.4	ORS665	0-9.9	1.1	-0.33	6.5	Yes
	8	43.8	HT668	37.8-46.8	1.3	0.84	8.0	No
	15	56.4	ORS1141	52.4-58.2	1.7	0.16	15.7	No

Table 3.2 Continued.

Trait	Linkage Group[a]	Position[b]	Nearest Marker	1-LOD Interval[c]	Additive Effect[d]	Dominance Ratio[e]	PVE[f]	Previously Identified?[g]
Height	1	8.0	ORS716	4.6-10.0	29.6	0.45	11.9	No
	6	57.6	ORS483	47.6-57.7	23.1	0.01	6.4	Yes
	9	10.0	ORS1265	2.0-19.0	15.7	0.56	3.0	No
	14	16.1	HT319	10.1-18.0	11.5	1.19	3.0	No
	15	57.1	ORS687	57-58.2	53.2	-0.31	39.4	No
No. Main Stem Leaves	6	57.6	ORS483	55.6-57.7	2.7	0.09	4.9	Yes
	7	1.0	ORS1041	0-7.3	-1.8	0.31	2.7	Yes
	9	19.0	HT294	13.0-39.8	2.6	-0.23	5.3	Yes
	15	57.1	ORS687	57.1-58.2	8.0	-0.49	57.0	No

Table 3.2 Continued.

Trait	Linkage Group[a]	Position[b]	Nearest Marker	1-LOD Interval[c]	Additive Effect[d]	Dominance Ratio[e]	PVE[f]	Previously Identified?[g]
Leaf Size	5	31.6	ORS852	21.6-44.5	64.8	-0.71	9.1	Unknown
	8	35.6	ORS1161	32.4-35.8	56.4	-0.14	5.6	No
	10	15.8	ORS437	7.8-18.9	46.7	0.64	4.4	No
	14	10.1	ORS307	3.1-18.0	36.4	1.34	4.9	No
	15	57	ORS1141	50.5-58.2	42.2	0.97	3.7	No
	16	45.4	ORS407	37.4-60.1	-31.6	-1.96	5.1	No
No. Branches	10	17.1	ORS437	9.8-24.8	-1.3	0.26	4.6	No
	13	0	HT848	0-17.6	-1.4	-0.1	5.2	No
	16	30.1	ORS899	22.0-36.1	-1.3	-0.87	7.0	No
	17	22.0	ORS735	16.0-30.1	-0.2	11.47	8.4	No

Table 3.2 Continued.

Trait	Linkage Group[a]	Position[b]	Nearest Marker	1-LOD Interval[c]	Additive Effect[d]	Dominance Ratio[e]	PVE[f]	Previously Identified?[g]
No. Heads	6	41.1	ORS1229	22.8-53.6	0.8	-0.43	3.1	No
	8	29.5	ORS147	19.5-35.2	-1.3	-0.06	5.9	No
	10	15.8	ORS437	11.8-18.9	-2.4	0.59	28.1	No
	13a	0	HT848	0-2.0	-1.2	-0.2	5.3	No
	13b	15.6	ORS317	5.6-27.6	-1.3	0.03	6.5	No
	16	34.1	ORS993	28.1-45.4	-0.5	-2.19	3.4	No
	17	24.8	ORS735	14.0-32.8	-0.4	2.31	3.4	Yes
Disc Diameter	1	14.4	HT39	0-18.3	2.7	0.67	4.4	No
	6	53.6	ORS381	45.6-57.7	2.3	-1.09	4.9	No
	8	35.2	ORS456	32.4-45.8	4.1	0.23	9.0	No

Table 3.2 Continued.

Trait	Linkage Group[a]	Position[b]	Nearest Marker	1-LOD Interval[c]	Additive Effect[d]	Dominance Ratio[e]	PVE[f]	Previously Identified?[g]
Disc Diameter	9	17.0	ORS1265	8.0-22.5	3.8	0.16	7.7	No
	10	13.8	ORS437	7.8-17.1	4.4	0.74	13	No
	14	12.1	ORS307	0-18.0	2.8	0.7	5.6	No
	15	50.5	ORS7	35.1-57.1	2.0	1.61	4.3	No
	17	4.0	ORS565	0-10.0	3.8	-0.14	5.5	No
No. Ray Flowers	5	19.6	ORS505	8.5-29.6	1.9	-0.2	6.7	No
	8	32.4	ORS147	23.5-41.8	0.8	1.77	3.1	No
	10	13.8	ORS534	4.0-22.9	1.4	0.04	3.6	No
	12	72.3	HT466	65.7-72.8	1.9	0.28	4.3	No
	15	57.1	ORS687	48.5-58.2	2.8	0.23	13.1	No

Table 3.2 Continued.

Trait	Linkage Group[a]	Position[b]	Nearest Marker	1-LOD Interval[c]	Additive Effect[d]	Dominance Ratio[e]	PVE[f]	Previously Identified?[g]
Disc Diameter	9	17.0	ORS1265	8.0-22.5	3.8	0.16	7.7	No
	10	13.8	ORS437	7.8-17.1	4.4	0.74	13	No
	14	12.1	ORS307	0-18.0	2.8	0.7	5.6	No
	15	50.5	ORS7	35.1-57.1	2.0	1.61	4.3	No
	17	4.0	ORS565	0-10.0	3.8	-0.14	5.5	No
No. Ray Flowers	5	19.6	ORS505	8.5-29.6	1.9	-0.2	6.7	No
	8	32.4	ORS147	23.5-41.8	0.8	1.77	3.1	No
	10	13.8	ORS534	4.0-22.9	1.4	0.04	3.6	No
	12	72.3	HT466	65.7-72.8	1.9	0.28	4.3	No
	15	57.1	ORS687	48.5-58.2	2.8	0.23	13.1	No

Table 3.2 Continued.

Trait	Linkage Group[a]	Position[b]	Nearest Marker	1-LOD Interval[c]	Additive Effect[d]	Dominance Ratio[e]	PVE[f]	Previously Identified?[g]
No. Ray Flowers	17	26.8	ORS735	10-33.7	1.2	0.92	4.6	No
No. Selfed Seeds	1	14.4	HT39	8.0-30.3	48.7	-0.5	6.6	No
	8	15.5	ZVG34	6.2-29.5	44.2	0.69	5.4	No
	12	72.3	HT466	65.7-72.8	60.4	0.06	6.8	No
	17	18.0	ORS735	10.0-33.7	37.3	1.08	7.2	No
Achene Weight	1	6.6	HT1018	2.0-18.3	1.6	0.70	8.6	No
	8	35.2	ORS456	19.5-35.8	1.7	0	7.6	No
	9	19.0	HT294	6.0-35.8	1.3	0.37	4.2	Yes
	10	15.8	ORS437	9.8-18.9	2.6	0.36	19.0	Yes

Table 3.2 Continued.

Trait	Linkage Group[a]	Position[b]	Nearest Marker	1-LOD Interval[c]	Additive Effect[d]	Dominance Ratio[e]	PVE[f]	Previously Identified?[g]
Shattering[g]	4a	0	HT298	0-4.0	-0.1	0.26	10.7	n/a
	4b	33.4	ORS674	32.6-41.4	0.1	0.04	6.4	n/a
	10	15.8	ORS437	7.8-20.1	-0.1	0.38	9.0	n/a
Seed Germination	12	72.3	HT466	71.4-72.8	-10.0	-1.31	17.3	n/a
	15	57.1	ORS687	48.5-58.2	-20.3	0.16	17.8	n/a

Table 3.2 Continued.

[a] When multiple QTL for a single trait occured on the same linkage group, a letter was used to uniquely identify them.

[b] Absolute position from the top of the linkage group (in cM).

[c] Refers to the region flanking each QTL peak in which the LOD score declines by one.

[d] Refers to the additive effect (*a*) of the Hopi allele. Underlined values indicate instances in which the allelic effects are in the 'wrong' direction. See text for details.

[e] Refers to the dominance ratio (*d*/*a*) of the Hopi allele.

[f] Percentage of phenotypic variation explained by each QTL using CIM. PVE values for QTL with effects in the direction of the wild phenotype are underlined.

[g] Indicates whether or not a given QTL was detected in the previous cultivated × wild sunflower QTL analysis (Burke *et al.* 2002). The determination of overlap between studies was based on the one-LOD confidence intervals. Note that the previous analysis of shattering was based on a different (indirect) measure, and that seed germination has not been previously analyzed.

Table 3.3 Summary of significant interactions amongst individually significant QTL.

Trait	Linkage Groups	Type of Interaction[a]	Phenotypic Effect[b]	Effect (%)
Height	6 × 15	A × A	12.5	2.2
	9 × 15	D × A	17.8	1.3
	" "	D × D	-22.9	1.0
No. Main Stem Leaves	6 × 15	A × A	2.4	3.8
	7 × 9	A × A	-1.2	1.3
Leaf Size	5 × 8	A × A	-46.8	1.6
	5 × 14	D × D	100.2	2.2
	10 × 14	D × A	70.5	1.2
No. Branches	10 × 17	D × A	-1.5	0.8
	13 × 17	D × A	-5.0	9.3
	" "	D × D	3.0	-1.1
	16 × 17	A × A	1.4	3.7
No. Heads	8 × 16	A × A	0.8	0.8
	8 × 17	D × A	-1.4	2.2
	13a × 17	D × A	-1.2	1.1
Disk Diameter	1 × 8	A × A	-1.9	0.4
	1 × 14	A × A	-2.3	1.0

Table 3.3 Continued.

Trait	Linkage Groups	Type of Interaction[a]	Phenotypic Effect[b]	Effect (%)
Disk Diameter	1 × 17	A × A	-2.9	2.1
	6 × 14	D × A	2.7	1.1
	6 × 15	A × A	2.3	1.8
	6 × 17	A × A	2.3	1.3
	8 × 14	A × A	-2.7	1.8
	15 × 17	A × A	-2.2	0.3
No. Ray Flowers	8 × 15	D × A	2.1	1.8
	10 × 12	D × A	-2.0	1.9
	10 × 15	D × A	-1.9	1.5
No. Selfed Seeds	1 × 17	A × A	38.5	0.3
	8 × 17	A × A	42.2	1.0
Achene Weight	1 × 8	A × A	-2.3	2.0
Seed Germination	12 × 15	A × A	-14.3	11.5

[a] A × A = Additive × Additive; A × D = Additive × Dominant; D × A = Dominant × Additive; D × D = Dominant × Dominant.

[b] Underlined values indicate an interaction with effects in the 'wrong' direction.

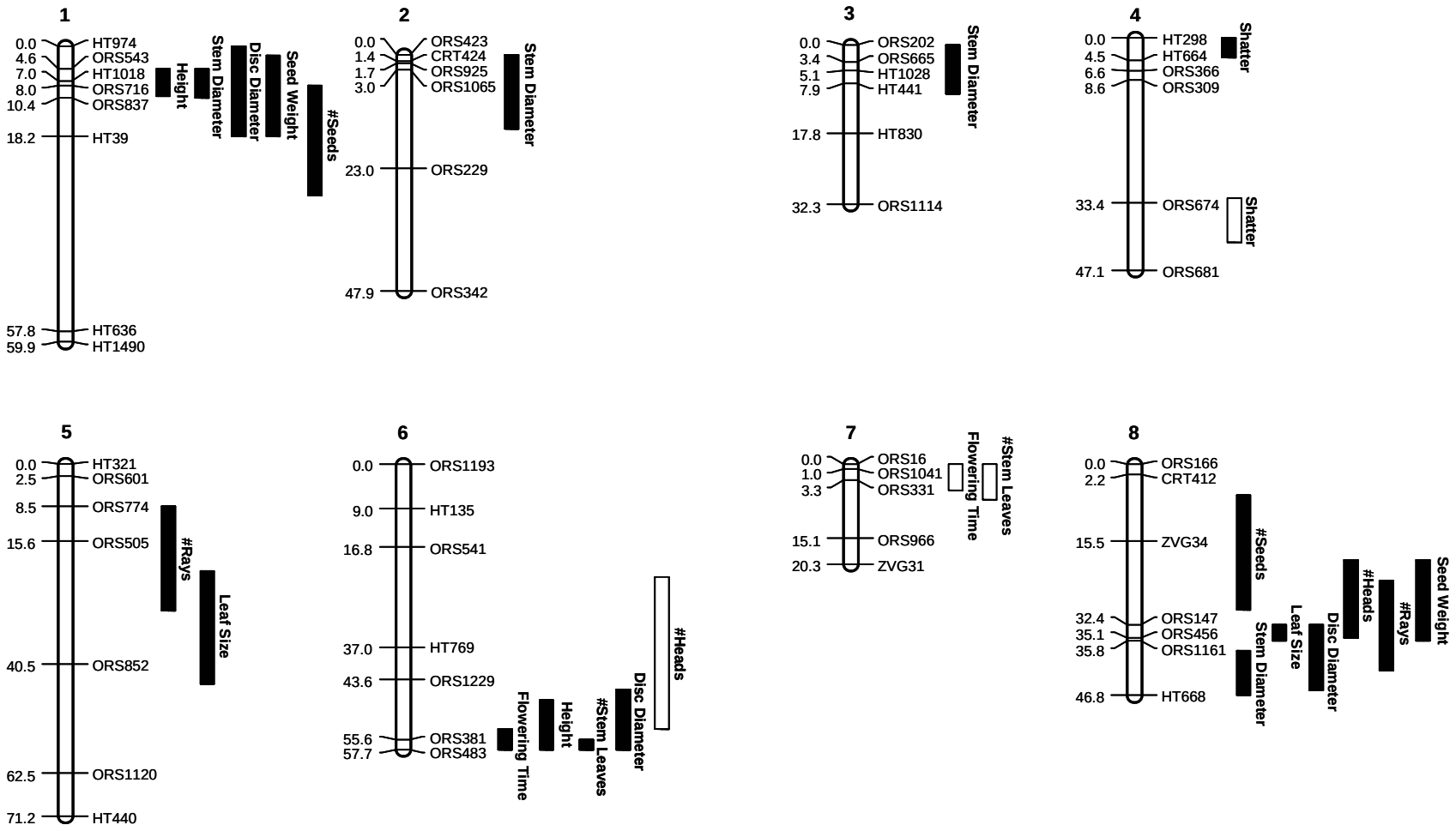
1
0.0 HT974
4.6 ORS543
7.0 HT1018
8.0 ORS716
10.4 ORS837
18.2 HT39
57.8 HT636
59.9 HT1490
Height
Stem Diameter
Disc Diameter
Seed Weight
#Seeds
2
0.0 ORS423
1.4 CRT424
1.7 ORS925
3.0 ORS1065
23.0 ORS229
47.9 ORS342
Stem Diameter
3
0.0 ORS202
3.4 ORS665
5.1 HT1028
7.9 HT441
17.8 HT830
32.3 ORS1114
Stem Diameter
4
0.0 HT298
4.5 HT664
6.6 ORS366
8.6 ORS309
33.4 ORS674
47.1 ORS681
Shatter
Shatter
5
0.0 HT321
2.5 ORS601
8.5 ORS774
15.6 ORS505
40.5 ORS852
62.5 ORS1120
71.2 HT440
#Rays
Leaf Size
6
0.0 ORS1193
9.0 HT135
16.8 ORS541
37.0 HT769
43.6 ORS1229
55.6 ORS381
57.7 ORS483
Flowering Time
Height
#Stem Leaves
Disc Diameter
#Heads
7
0.0 ORS16
1.0 ORS1041
3.3 ORS331
15.1 ORS966
20.3 ZVG31
Flowering Time
#Stem Leaves
8
0.0 ORS166
2.2 CRT412
15.5 ZVG34
32.4 ORS147
35.1 ORS456
35.8 ORS1161
46.8 HT668
#Seeds
Stem Diameter
Leaf Size
Disc Diameter
#Heads
#Rays
Seed Weight

Figure 3.1

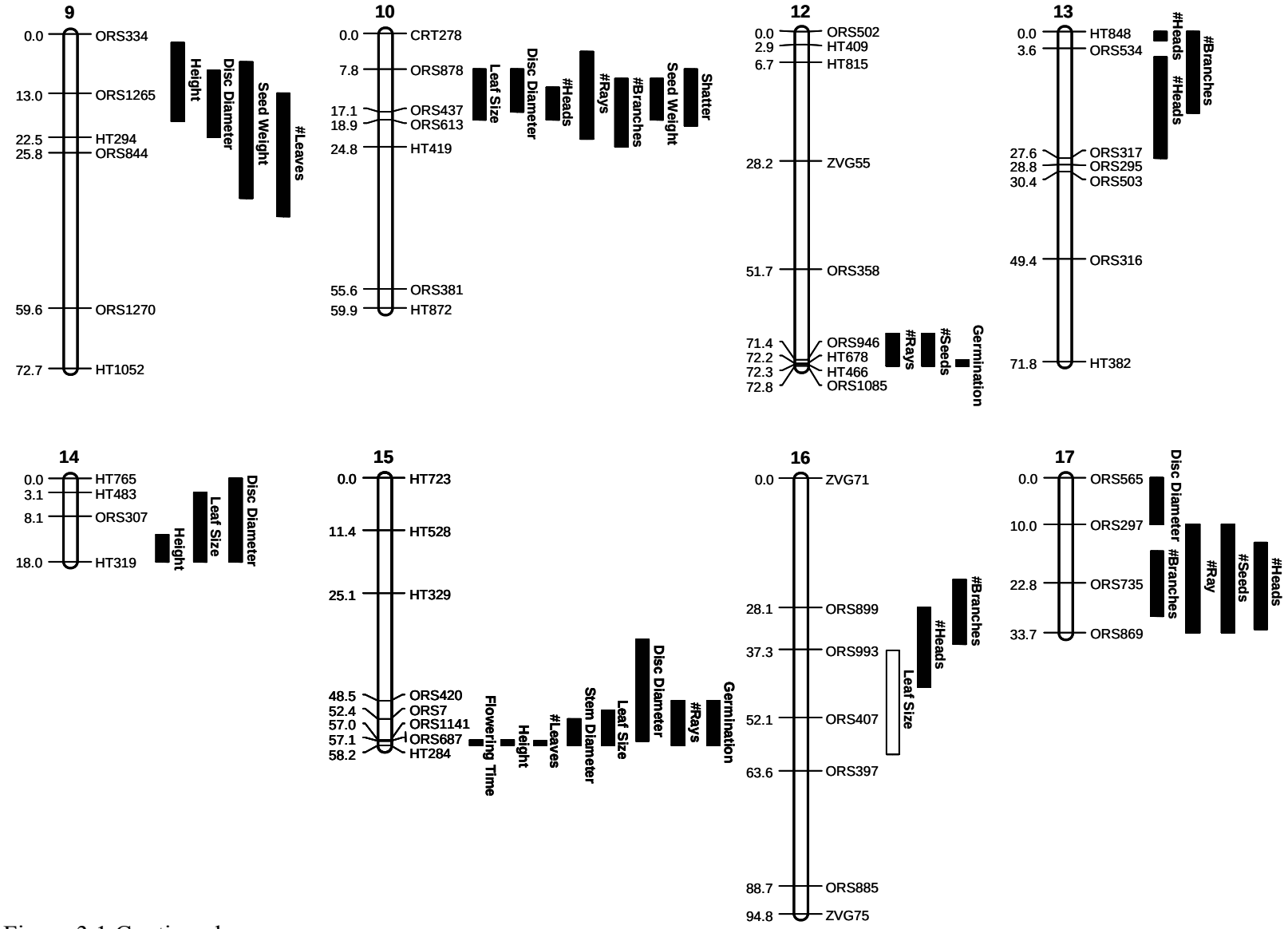

Figure 3.1 Continued.

Figure 3.1 Results of the CIM analysis for the sixteen linkage groups on which QTL were detected. QTL positions are indicated by bars alongside each linkage group. The length of each bar is equal to the one-LOD support interval for that QTL. Loci at which the crop allele had the expected effect are indicated by a filled bar, whereas those at which the crop allele conferred a wild-like phenotype are represented by unfilled bars.

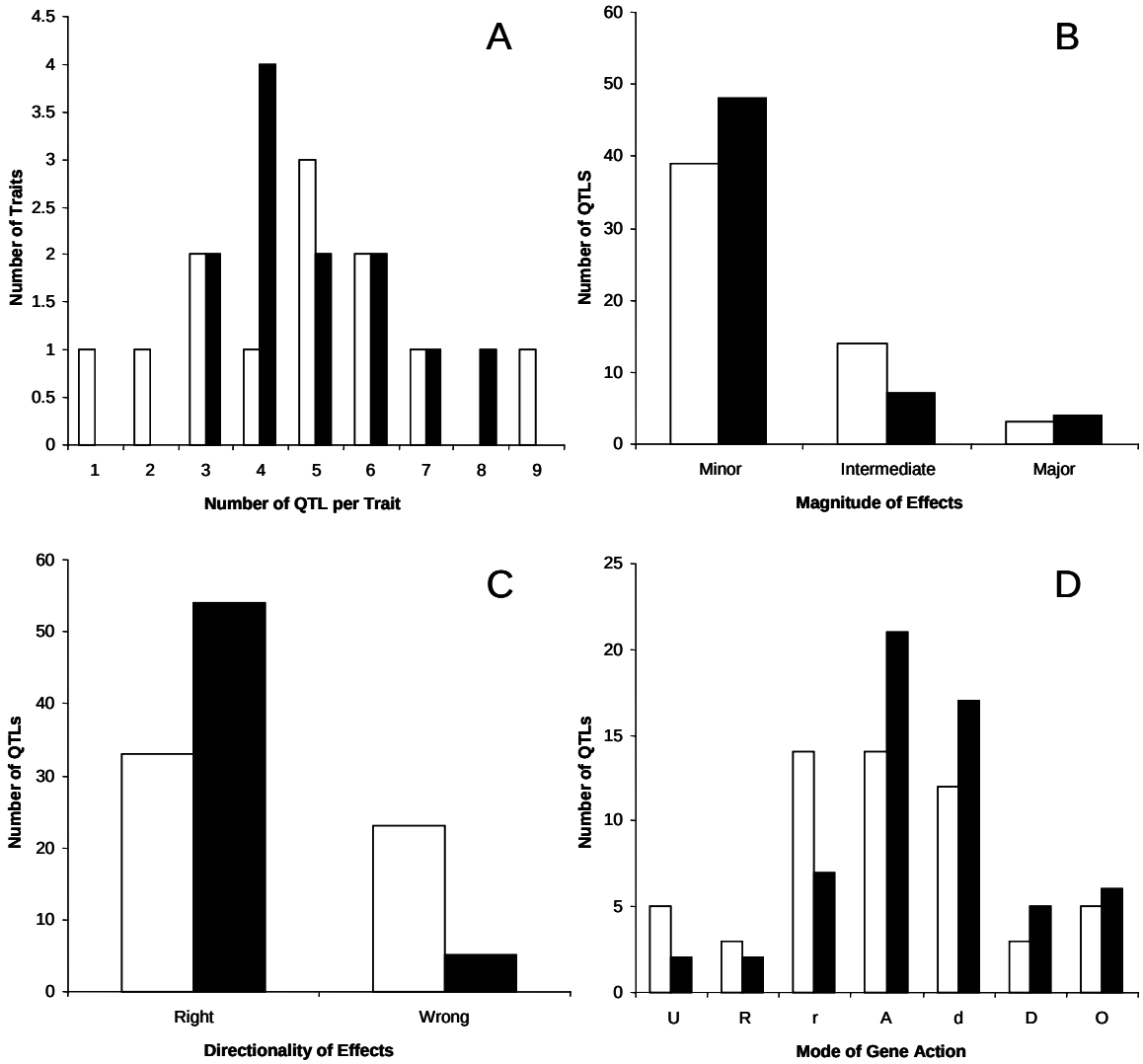
A
Number of Traits
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
1
2
3
4
5
6
7
8
9
Number of QTL per Trait
B
Number of QTLS
60
50
40
30
20
10
0
Minor
Intermediate
Major
Magnitude of Effects
C
Number of QTLs
60
50
40
30
20
10
0
Right
Wrong
Directionality of Effects
D
Number of QTLs
25
20
15
10
5
0
U
R
r
A
d
D
O
Mode of Gene Action

Figure 3.2

Figure 3.2 Comparison of the number of QTL per trait (A), magnitude of effects (B), directionality of effects (C), and mode of gene action (D) between a previous study based on an elite × wild sunflower mapping population (open bars; Burke *et al.* 2002) and the present study, which was based on a primitive × wild mapping population (closed bars). The following thresholds were used to classify the mode of gene action for each QTL: underdominant $\leq -1.25 <$ recessive $\leq -0.75 <$ partially recessive $\leq -0.25 <$ additive $\leq 0.25 <$ partially dominant $\leq 0.75 <$ dominant $< 1.25 \leq$ overdominant.

CHAPTER 4

GENETIC ARCHITECTURE OF POST-DOMESTICATION TRAITS IN SUNFLOWER[1]

[1] WILLS, D.M., H. ABDEL-HALEEM, S.J.KNAPP, and J.M. BURKE. To be submitted to *New Phytologist*.

ABSTRACT

Following the initial domestication of a new crop plant, different lineages are expected to undergo diversification to disparate habitats, or to fill novel agricultural roles. From specialized landraces to highly derived cultivars, this process has resulted in the numerous distinct varieties found in modern-day crop germplasm collections. In sunflower, the genetic architecture of domestication has previously been shown to be complex, involving numerous quantitative trait loci, typically of small to moderate effect. In the present study, we mapped QTL underlying a number of novel post domestication traits in the Hopi sunflower, a primitive, Native American domesticate. These traits included the extremely late flowering time and the achene characters (primarily shape and anthocyanin pigmentation) that are characteristic of the Hopi sunflower. Unlike typical domestication-related traits, these traits exhibited a relatively simple genetic basis, with two genomic regions being largely responsible for the divergence of the Hopi sunflower from other cultivated sunflower lineages. Whether the simple genetic control of these traits is a result of the complex genetic architecture of domestication or an inherent quality of the underlying developmental pathways of the traits themselves remains an open question.

INTRODUCTION

The evolution of crop plants can be viewed as occurring in somewhat distinct phases (e.g., Yamasaki *et al.* 2005; Burke *et al.* 2005, 2007). During the initial phase of domestication

from its wild progenitor, the proto-domesticate typically experiences changes in a suite of traits collectively known as the "domestication syndrome" (Hammer 1984; Harlan 1992). Following this initial period of domestication, crop lineages often experience selection for adaptation to disparate environments, and/or to fill different agricultural niches. Beyond this, crop plants are generally subjected to selection on traits such as yield, quality, and disease resistance. This complex evolutionary history is often preserved, at least in part, in germplasm collections. Early landraces which have undergone minimal post-domestication selection are thus available for many crops, creating an opportunity to genetically dissect the process of crop evolution.

Previous genetic analyses of phenotypic evolution under domestication have revealed that domestication-related traits often have a relatively simple genetic architecture, with a small number of quantitative trait loci (QTL), each of relatively large effect, accounting for much of the phenotypic transformation (e.g., Doebley *et al.* 1990; Doebley and Stec 1991, 1993; Paterson *et al.* 1991; Koinange *et al.* 1996; Xiong *et al.* 1999; Liu *et al.* 2007). The domestication of sunflower (*Helianthus annuus* L.) is perhaps the clearest counterexample to this pattern. Sunflower domestication appears to have been driven by selection on a large number of loci, most of which had small to moderate phenotypic effects (Burke *et al.* 2002; Wills and Burke 2007).

While we now have a relatively well-developed understanding of the genetics of crop domestication, much less is known about the genetic changes that occur during the early stages of post-domestication crop diversification. Sunflower is thought to have been initially domesticated as a source of edible seeds and was ultimately developed into an important oilseed crop (Heiser 1945, 1954; Putt 1997), though certain early domesticates

have subsequently become adapted for growth in novel environments and/or for new agricultural uses. For example, the Hopi sunflower, which is native to the desert southwest of the United States, exhibits exceptionally late flowering and ultimately came to be used as an important source of dye for basketry, textiles, and body paint in the Hopi culture (Whiting 1939; Heiser 1950, 1951). The pigments used for these purposes were water-soluble anthocyanins extracted from the achene (single-seeded fruit) pericarps—that is, the 'shell' surrounding the sunflower kernel. It is, however, unlikely that anthocyanin content was selected during the initial domestication of sunflower. Although domesticated sunflower appears to have been singly derived (Harter *et al.* 2004; Wills and Burke 2006), this trait varies among the primitive domesticated lineages. Genetic evidence identifies the Hopi and Havasupai landraces as the most basal extant domesticated sunflower lineages (Heiser 1951; Tang and Knapp 2003; Harter *et al.* 2004; Wills and Burke 2006); unlike Hopi, however, the Havasupai landrace is quite variable when it comes to pericarp anthocyanin levels (Heiser 1945, 1951). The available evidence is thus consistent with the high anthocyanin phenotype of the Hopi landrace being selected for post-domestication.

In addition to late flowering and high anthocyanin concentration, the Hopi sunflower possesses achenes which exhibit a unique overall morphology, with an elevated kernel-to-pericarp weight ratio and a high achene length-to-width ratio resulting in cylindrical shape. Furthermore, seed oil content in the Hopi sunflower is somewhat elevated relative to that of wild sunflower, though there is no evidence that this trait was directly selected by Native Americans. In this article, we report on the genetic architecture of the unique trait transitions observed in the Hopi sunflower, including

delayed flowering, increased seed anthocyanin and oil content, and elongated seed shape. We leveraged an existing Hopi × wild sunflower mapping population that was previously used to map QTL underlying domestication-related traits; as such, we are able to directly compare the genetic basis of these changes to the types of changes that occurred during the initial domestication of sunflower.

MATERIALS AND METHODS

Mapping Population and Genotyping

The mapping population used for this study was previously described by Wills and Burke (2007). This population consists of 378 F_2 individuals derived from a cross between an individual from a wild population in Keith County, Nebraska (Ann1238) and the Hopi sunflower (USDA PI 432504). The linkage map, which consists of 111 codominant loci, was constructed using MAPMAKER 3.0/EXP (Lander *et al.* 1987; Lincoln and Lander 1992). The initial linkage groups were identified using the 'group' command with LOD > 5.0 and $\theta < 0.2$ and the final orders were then confirmed using the 'ripple' and 'compare' commands such that the map orders presented herein reflect the statistically most likely order based on the data at hand.

Phenotypic Analysis

The phenotypes examined in this study include nine aspects of seed morphology, pericarp anthocyanin concentration, and percent seed oil content for a total of 11 traits. Only 300 of the 378 F_2 individuals set selfed seeds and were thus available for analysis. Parental

phenotypic values are based on averages from eight wild (Ann1238) and seven Hopi individuals grown concurrently with the mapping population in the greenhouse.

Flowering time and achene weight data are the same values that were reported previously (Wills and Burke 2007). The number of days from germination to flowering was recorded for all 378 individuals and 100 achene weight is based on the 300 F_2 individuals that set selfed seeds. Pericarp weights were only available for the 150 families assayed for anthocyanin concentration (see below). For those 150 lines, kernel weight was estimated as average achene weight minus average pericarp weight, and kernel-to-pericarp weight ratio was estimated from those measures. The remaining seed size and shape phenotypes were measured using Tomato Analyzer 2.1.0.0 (Brewer *et al.* 2006), which is a software package capable of measuring a variety of morphological characteristics of fruits or seeds from a scanned image. Fifteen achenes from all lines with sufficient seed, 230 in total, were scanned at 1200 dpi and saved as JPEG files. The achene images were rotated to a common orientation, and the perimeter, area, length, width, and achene shape were estimated for each. Using the standard nomenclature of Brewer *et al.* (2006), these measures correspond to perimeter, area, height at midpoint, width at midpoint, and fruit shape index (fs II, which is essentially a length-to-width ratio), respectively. These values were then averaged across seeds for each line.

Due to the destructive nature of the anthocyanin assay, the 150 lines with the largest number of seeds were selected for analysis. To quantify pericarp anthocyanin content, we extracted anthocyanin from achenes of each line using an adaptation of the protocol described by Bullard *et al.* (1989). The pericarps from 5 achenes from each of seven Hopi individuals were removed, weighed, and placed in a 2 mL microcentrifuge

tube containing 250 mg of sand, one 4.83 mm steel bead, and 1.2 mL of 0.5% HCl in methanol. Seeds were ground for 2 minutes at 30 hertz in a Qiagen TissueLyser (Qiagen, Valencia, CA). Samples were then centrifuged for 10 minutes at 16,100 g and 1 mL of supernatant was removed. A SmartSpec 3000 spectrophotometer (Bio-Rad Laboratories, Inc., Hercules, CA) was used to find the absorbance of the supernatant from 400 nm to 800 nm in 1 nm intervals. Peak heights were divided by the mass of pericarp material extracted and the value at each wavelength was averaged across all seven individuals. The highest average peak was 525 nm, which corresponds closely to the expected sunflower anthocyanin value reported by Mazza and Gao (1993). Phenotyping of the mapping population involved the characterization of pericarps from five selfed seeds from each of 150 F_2 individuals. Extractions were performed as above, and absorbance at 525 nm was measured. Trait values were recorded as OD525 per gram pericarp weight and Box-Cox transformed in JMP 4.0 (SAS Institute, Cary, NC) prior to QTL analysis.

Seed oil content was estimated following the methods outlined by Burke *et al.* (2005) and Tang *et al.* (2006). Briefly, achene samples weighing between 0.5–1.0 g were drawn from the selfed offspring of each of 150 selfed F_2 individuals and placed in flat-bottomed sample tubes. Percentage oil content was measured by pulsed nuclear magnetic resonance (NMR) analysis on a Bruker MQ20 Minispec NMR Analyzer (The Woodlands, TX). The data were proportions of total mass, so trait values were arcsine square-root transformed for data analysis (Sokal and Rohlf 1995).

QTL Analysis

The QTL analysis followed the same general approach outlined by Wills and Burke (2007) and Burke *et al.* (2002) using QTL Cartographer version 1.17 (Basten *et al.* 1994,

2004). Composite interval mapping (CIM; Zeng 1993, 1994) was performed with a 10-cM window and five background cofactors identified using both forward and backward regression, with tests performed at 2-cM intervals. 1000 permutations for each trait were performed to set threshold values ($\alpha = 0.05$) to identify the presence of QTL (Churchill and Doerge 1994; Doerge and Churchill 1996). A LOD decline of 2.0 between adjacent peaks was taken as evidence of multiple, linked QTL on a linkage group. To search for epistatic interactions between the identified QTL, multiple interval mapping was employed (MIM; Kao and Zeng 1997; Kao *et al.* 1999). The maximum number of allowable pairwise interactions was set to 19 and significance was determined based on the information criterion $IC(k) = -2(\log(L) - k\, c(n) / 2)$ with $c(n) = \log(n)$ as the penalty function and a threshold of 0.0, as recommended by the authors. The degree of dominance of the Hopi allele at each locus was calculated as the dominance effect divided by the additive effect (d/a), such that a value of -1 reflects recessive gene action, a value of 0 reflects additivity, and a value of +1 reflects dominance. Arbitrary percentage of variance explained (PVE) thresholds of 10 and 25% were used to classify QTL as having "minor," "intermediate," and "major" effects. These were the same thresholds used to describe previously identified QTL in sunflower (Burke *et al.* 2002; Wills and Burke 2007).

RESULTS

The linkage map used for the present study has been described elsewhere (Wills and Burke 2007). The map was 906.4 cM long, with the expected 17 linkage groups (LGs),

and had an average intermarker distance of 8.4 cM. CIM identified 42 QTL for the 12 traits (including flowering time and achene weight from the previous analysis). QTL were found on ten of the 17 linkage groups (Figure 4.1 and Table 4.2). One-LOD support interval estimates for QTL locations ranged from 0.1 to 29.8 centimorgans (cM) (mean = 9.3 cM). The number of QTL per trait ranged from two to six and the average QTL explained 14.7% of the variance (range = 2.5 to 46.9%). Achene shape, pericarp weight, and kernel-to-pericarp weight ratio were the only traits with multiple QTL on the same linkage group. Each of these traits had two non-overlapping QTL on LG 16 (Figure 4.1). The QTL for achene perimeter on LG 4 was the only QTL in which the 1-LOD interval did not overlap with previously identified domestication QTL.

In our previous investigation of sunflower domestication, three QTL were identified for flowering time on LGs 6, 7, and 15 (Wills and Burke 2007) and four QTL for achene weight were identified on LGs 1, 8, 9, and 10 (Wills and Burke 2007). Each of the achene weight QTL co-localized with a QTL for kernel weight or pericarp weight in the present analysis, but not always both. Furthermore, QTL for kernel and pericarp weight were identified on LGs 7, 15, and 16 without a significant effect on total seed weight. Above that location on LG 16, is a second QTL for pericarp weight which co-localizes with a QTL of large effect for kernel-to-pericarp weight ratio (40.8% PVE). A final QTL for kernel-to-pericarp ratio, which mapped to a third genomic region on LG 16, above the other two, explains an additional 11.6% of the variance and once again had no significant effect on total seed weight. Most of the QTL (17 of 19) for achene area, perimeter, length, width, and achene shape co-localized with QTL for achene, kernel, or pericarp weight identified in the present or previous analyses of this cross (Wills and

Burke 2007). The only two exceptions were QTL for achene length and perimeter on LG 4. Only two QTL were detected for seed oil content in the Hopi × wild population, with the QTL on LG 16 controlling 39.5% of the phenotypic variance. Similarly, the concentration of anthocyanin in the pericarp was controlled by two QTL in this cross, one each on LGs 15 and 16. These QTL explained 28.7% and 17.7% of the trait variance, respectively.

For those traits with two or more QTL, we searched for epistatic interactions using MIM (Table 4.3). Ten significant two-way interactions accounting for 0.1-8.1% of the phenotypic variability were identified for 13 traits. The contribution of epistasis to the observed phenotypes was minor, as was previously observed for domestication-related traits. Only a single dominant × additive interaction for achene shape had an effect greater than 5%. Pericarp weight was the only other trait for which the total effect for all epistatic interactions summed to greater than 5%. No significant epistatic interactions were detected for oil content or anthocyanin content, though power to detect an interaction for these traits was reduced by the reduced sample size of 150 individuals. Nonetheless, 150 observations were sufficient to identify two minor epistatic interactions for pericarp weight.

DISCUSSION

As one of the most basal domesticated sunflower lineages, the Hopi landrace has provided important insights into the genetic architecture of the early stages of the evolution of cultivated sunflower (Wills and Burke 2007). Despite its primitive nature,

the Hopi sunflower exhibits a number of unique traits, at least some which appear to be byproducts of post-domestication selection for adaptation to local growing conditions as well as for cultural purposes. As noted above, these traits include exceptionally late flowering time, as well as a suite of unique seed phenotypes.

Flowering Time

On average, the Hopi individuals included in our study initiated flowering after 100 days, which was over two weeks (15.2 days) later than individuals from the wild parental population (Table 4.1). This result accords well with the previous findings of Heiser (1951), who observed a mean flowering time in the field of 102 days post-germination. This finding led him to note that "Such late maturing forms obviously can only be grown where there is an extremely long growing season" (Heiser 1951). Our results indicate that this phenological shift is largely due to a single QTL on LG 15 with an additive effect of 10.4 days with two other QTL playing a lesser role (Table 4.2; Figure 4.1).

Seed Phenotypes

Any investigation of the achene characters present in the Hopi landrace must first account for the tremendous increase in seed size that likely drove sunflower domestication (Burke *et al.* 2002). As noted above, 17 of 19 QTL conditioning increases in achene length, width, perimeter, and area, or altering achene shape, co-localized with previously mapped QTL controlling some aspect of achene mass on LGs 1, 8, 9, and 10 (Burke *et al.* 2002; Tang *et al.* 2006; Wills and Burke 2007), and are thus likely a byproduct of prior selection for increased seed size during domestication. When the change in seed morphology is measured in terms of achene shape (i.e., fs II), two QTL are detected on LG 16 which collectively account for 38.1 % of the observed phenotypic variance. These

two QTL co-localize with two QTL underlying kernel-to-pericarp weight ratio on that same linkage group (Figure 4.1). Thus, LG 16 appears to be a key determinant of the unique achene morphology of the Hopi sunflower, independent of the increase in seed mass during domestication. The one QTL of major effect (PVE = 39.5%) for seed oil content is also present on this linkage group. This QTL has been detected in multiple crosses and environments, and has been found in other studies to co-localize with *Hyp,* a phenotypically-defined locus causing a white pigment to be deposited in the hypodermis of the pericarp (Leon *et al.* 1995, 1996; Tang *et al.* 2006).

Both wild and domesticated sunflowers produce anthocyanin in multiple tissues including the hypocotyl, margins of young leaves, upper part of the petioles, basal leaf veins, stems, top of the calyx, stigmas, corolla of the disk florets, as well as in the pericarp of the achene (Miller and Fick 1997, Fick and Miller 1997). Anthocyanin pigmentation in sunflower has previously been shown to be under multilocus control, and pleiotropic effects have been observed for some of the pigment loci. The presence of anthocyanin in any tissue is controlled by *T*, a locus with epistatic effects (Stoenescu 1974; Joshi *et al.* 1994), which has been mapped in close proximity to a nuclear male sterility locus on LG 11 (Burke *et al*. 2002; Pérez-Vich *et al.* 2005). Prior to our work, this was the only locus controlling anthocyanin pigmentation that had been mapped in sunflower. Although the interval containing *T* was well-covered in our map, QTL for anthocyanin concentration were not found on LG 11. Presumably, this locus was not segregating in our mapping population, as all offspring were able to produce anthocyanin. Mosjidis (1982) found evidence for three loci controlling the presence of anthocyanin in sunflower. A dominant allele at one locus (*C*, which likely corresponds to *T*) was

necessary, but not sufficient, for anthocyanin production. The *P* locus governs anthocyanin pigmentation in the hypodermis of the pericarp as well as the corolla (note that the *P* locus in this instance is distinct from the *P* locus controlling phytomelanin pigmentation; Stoenescu 1974; Joshi *et al*. 1994). Additionally, the *Y* locus was responsible for producing a "diluted" purple phenotype in the hypodermis. The anthocyanin pigment QTL detected in the present study may correspond to *Y* and *P*, though we did not observe a diluted purple phenotype in any individuals. It is conceivable that the phenotype observed by Mosjidis (1982) was the result of the joint effects of the QTL that we mapped to LG 16 and the *Hyp* locus (Leon *et al.* 1996), whose location falls in an adjacent, non-overlapping interval. Together, these loci might control the deposition of both anthocyanin and the white pigment in the pericarp, resulting in a diluted purple phenotype.

Insights into the Origin of the Hopi Sunflower

Two genomic regions harbor clusters of QTL influencing seed morphology: the interval containing the *B* locus on LG 10 and the middle portion of LG 16. QTL on LG 10 are predominantly related to seed size, and are thus likely a byproduct of domestication. In contrast, QTL on LG 16 influence most of the achene characters that are unique to the Hopi landrace, including achene shape, increased kernel-to-pericarp weight ratio, and pericarp anthocyanin pigmentation. Thus, the distinguishing characteristics of the Hopi landrace,– i.e., the unique achene shape and anthocyanin pigmentation, as well as the exceptionally late flowering and resulting tremendous plant size – are largely controlled by QTL on linkage groups 15 and 16.

The foregoing results suggest that the unique traits of the Hopi sunflower have a relatively simple genetic basis; yet, the source of the necessary alleles remains unknown. In principle, there are three possible sources of such allelic variation: mutation, introgression, or standing variation within the domesticated sunflower gene pool. While novel mutations are a possibility, this explanation would require the occurrence of mutations at two loci shortly after domestication, and therefore seems improbable. If the necessary variation arose via introgression, one might expect to see a pattern similar to that documented by Burke *et al.* (2005) in an analysis of oilseed sunflower evolution, wherein a putatively introgressed region harbored QTL conditioning desirable traits, as well as multiple QTL with maladaptive (wild-like) effects. However, all of the QTL found at the bottom of LG 15 acted in the expected direction, and just one QTL underlying a domestication-related trait on LG 16 had an effect in the wrong direction. In the case of standing variation, the genetic bottleneck imposed by domestication will determine the amount of residual variation present for selection. For domestication-related traits, very little (if any) genetic variation is expected to make it through the bottleneck (Hanson *et al.* 1996; Tenaillon *et al.* 2004), but for other traits the severity of the bottleneck will ultimately determine the amount of available variation (Eyre-Walker *et al.* 1998). In the case of sunflower, a substantial amount of molecular variation is thought to have passed through the domestication bottleneck (Liu and Burke 2006; Kolkman *et al*. 2007), suggesting that there may have been a large amount of functional variation segregating in the early domesticates. For the flowering time and anthocyanin QTL on LGs 15 and 16, the Hopi alleles were all partially or completely recessive, which would have allowed their persistence until they became desirable.

Whatever the source of the variation necessary for producing the specialized traits that distinguish the Hopi lineage, these traits appear to be largely conditioned by the QTL-rich regions on LGs 15 and 16. This relatively simple genetic architecture may be a byproduct of constraints due to the genetically complex nature of domestication-related traits in sunflower. That is, the large number of domestication-related QTL may limit the fraction of the genome that is available for subsequent diversification. The strong genetic correlation between late flowering time and high anthocyanin pigmentation on the bottom of LG 15 suggests two possible scenarios. First, local adaptation to the long growing season of the desert southwest could have initially brought the 'Hopi' allele in this genomic region to higher frequency, at which time culturally-mediated selection for high anthocyanin content took over. Alternatively, the dissemination of domesticated sunflower to a locale with an exceptionally long growing season, such as the desert southwest, may have allowed for the evolution of highly pigmented seeds in the absence of natural selection against late flowering. Whether the simple genetic control of these traits is a result of the complex genetic architecture of domestication or an inherent quality of the underlying developmental pathways of the traits themselves remains an open question.

ACKNOWLEDGEMENTS

The authors wish to thank Jenny Wood, Ben Craddock, Daniel Hill, John Hvala, Chris Buckner, Jonathan Ertelt, Ben Stephens, and Jessica Wenzler for their assistance in the greenhouse and/or lab. We also thank James Hamrick, James Leebens-Mack, Mark Chapman, and Natasha Sherman for comments on a previous version of this manuscript. This work was supported by grants to JMB from the United States Department of Agriculture (#03-39210-13958 and #03-35300-13104) and the National Science Foundation (DBI-0332411).

LITERATURE CITED

Basten CJ, Weir BS, Zeng Z-B (1994) Zmap-a QTL cartographer. In: *Proceedings of the 5th World Congress on Genetics Applied to Livestock Production: Computing Strategies and Software* (eds. Smith C, Gavora JS, Benkel B, *et al.*), pp. 65-66. Organizing Committee, 5th World Congress on Genetics Applied to Livestock Production, Guelph, Ontario, Canada.

Basten CJ, Weir BS, Zeng Z-B (2004) *QTL Cartographer, Version 1.17.* Department of Statistics, North Carolina State University, Raleigh, NC.

Beavis WD (1994) The power and deceit of QTL experiments: lesson from comparative QTL studies. *In: Proceedings of the 49th Annual Corn and Sorghum Research Conference*, pp. 250-266. American Seed Trade Association.

Brewer MT, Lang LX, Fujimura K, *et al.* (2006) Development of a controlled vocabulary and software application to analyze fruit shape variation in tomato and other plant species. *Plant Physiology,* 141, 15-25.

Bullard RW, Woronecki PP, Dolbeer RA, Mason JR (1989) Biochemical and morphological characteristics in maturing achenes from purple-hulled and oilseed sunflower cultivars. *Journal of Agricultural and Food Chemistry,* **37**, 886-890.

Burke JM, Burger JC, Chapman MA (2007) Crop evolution: from genetics to genomics. *Current Opinion in Genetics and Development,* **17**, 525-532.

Burke JM, Knapp SJ, Rieseberg LH (2005) Genetic consequences of selection during the evolution of cultivated sunflower. *Genetics,* **171**, 1933-1940.

Burke JM, Tang S, Knapp SJ, Rieseberg LH (2002) Genetic analysis of sunflower domestication. *Genetics,* **161**, 1257-1267.

Churchill GA, Doerge RW (1994) Empirical threshold values for quantitative trait mapping. *Genetics,* **138**, 963-971.

Doebley J, Stec A (1991) Genetic analysis of the morphological differences between maize and teosinte. *Genetics,* **129**, 285-295.

Doebley J, Stec A, Wendel J, Edwards M (1990) Genetic and morphological analysis of a maize-teosinte F_2 population: Implications for the origin of maize. *Proceedings of the National Academy of Sciences, USA,* **87**, 9888-9892.

Doerge RW, Churchill GA (1996) Permutation tests for multiple loci affecting a quantitative character. *Genetics* **142**, 285-294.

Eyre-Walker A, Gaut RL, Hilton H, Feldman DL, Gaut BS (1998) Investigation of the bottleneck leading to the domestication of maize. *Proceedings of the National Academy of Sciences, USA,* **95**, 4441-4446.

Fick GN, Miller JF (1997) Sunflower breeding. In: *Sunflower Production and Technology* (ed. Scheiter AA), pp. 395-440. American Society of Agronomy, Madison, WI.

Hammer K (1984) Das Domestikationssyndrom. *Kulturpflanze,* 32, 11-34.

Hanson MA, Gaut BS, Stec AO*, et al.* (1996) Evolution of anthocyanin biosynthesis in maize kernels: The role of regulatory and enzymatic loci. *Genetics,* **143**, 1395-1407.

Harlan JR (1992) *Crops & Man*, 2nd edn. American Society of Agronomy, Madison, WI.

Harter AV, Gardner KA, Falush D*, et al.* (2004) Origin of extant domesticated sunflowers in eastern North America. *Nature,* **430**, 201-205.

Heiser CB (1945) The Hopi sunflower. *Missouri Botanical Garden Bulletin* **33**, 163-167.

Heiser CB (1946) A "new" cultivated sunflower from Mexico. *Madrono,* **8**, 226-229.

Heiser CB (1950) The sunflower among the North American Indians. *American Journal of Botany,* **37**, 662-662.

Heiser CB (1951) The sunflower among the North American Indians. *Proceedings of the American Philosophical Society,* **95**, 432-448.

Joshi SS, Basavalingappa, Giriraj K (1994) Pleiotropy in sunflower (*Helianthus annuus* L.). *Helia,* **20**, 1-6.

Kao CH, Zeng ZB (1997) General formulas for obtaining the MLEs and the asymptotic variance-covariance matrix in mapping quantitative trait loci when using the EM algorithm. *Biometrics,* **53**, 653-665.

Kao CH, Zeng ZB, Teasdale RD (1999) Multiple interval mapping for quantitative trait loci. *Genetics,* **152**, 1203-1216.

Koinange EMK, Singh SP, Gepts P (1996) Genetic control of the domestication syndrome in common bean. *Crop Science,* **36**, 1037-1045.

Kolkman JM, Berry ST, Leon AJ*, et al.* (2007) Single nucleotide polymorphisms and linkage disequilibrium in sunflower. *Genetics,* **177**, 457-468.

Lander ES, Green P, Abrahamson J*, et al.* (1987) MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. *Genomics,* **1**, 174-181.

Leon AJ, Lee M, Rufener GK, Berry ST, Mowers RP (1995) Use of RFLP markers for genetic linkage analysis of oil percentage in sunflower seed. *Crop Science,* **35**, 558-564.

Leon AJ, Lee M, Rufener GK, Berry ST, Mowers RP (1996) Genetic mapping of a locus (*hyp*) affecting seed hypodermis color in sunflower. *Crop Science,* **36**, 1666-1668.

Lincoln SE, Lander ES (1992) Systematic detection of errors in genetic linkage data. *Genomics,* **14**, 604-610.

Liu AZ, Burke JM (2006) Patterns of nucleotide diversity in wild and cultivated sunflower. *Genetics* **173**, 321-330.

Liu B, Fujita T, Yan ZH, *et al.* (2007) QTL mapping of domestication-related traits in soybean (*Glycine max*). *Annals of Botany,* **100**, 1027-1038.

Mazza G, Gao L (1994) Malonylated anthocyanins in purple sunflower seeds. *Phytochemistry,* **35**, 237-239.

Miller JF, Fick GN (1997) The genetics of sunflower. In: *Sunflower Production and Technology* (ed. Scheiter AA), pp. 441-496. American Society of Agronomy, Madison, WI.

Mosjidis JA (1982) Inheritance of color in the pericarp and corolla of the disk florets in sunflower. *Journal of Heredity,* **73**, 461-464.

Paterson AH, Damon S, Hewitt JD, *et al.* (1991) Mendelian factors underlying quantitative traits in tomato: Comparison across species, generations, and environments. *Genetics,* **127**, 181-197.

Perez-Vich B, Berry ST, Velasco L, *et al.* (2005) Molecular mapping of nuclear male sterility genes in sunflower. *Crop Science,* **45**, 1851-1857.

Putt ED (1997) Early history of sunflower. In: *Sunflower Production and Technology* (ed. Scheiter AA), pp. 1-19. American Society of Agronomy, Madison, WI.

Sokal RR, Rohlf FJ (1995) *Biometry*, 3rd edn. W.H. Freeman and Company, New York.

Stoenescu F (1974) Genetics. In: *Floarea-soarelui* (ed. Vranceau AV), pp. 92-125. academiei Republicii Socialiste, Bucuresti, Romania.

Tang SX, Knapp SJ (2003) Microsatellites uncover extraordinary diversity in Native American land races and wild populations of cultivated sunflower. *Theoretical and Applied Genetics,* **106**, 990-1003.

Tang SX, Leon A, Bridges WC, Knapp SJ (2006) Quantitative trait loci for genetically correlated seed traits are tightly linked to branching and pericarp pigment loci in sunflower. *Crop Science,* **46**, 721-734.

Tenaillon MI, U'Ren J, Tenaillon O, Gaut BS (2004) Selection versus demography: A multilocus investigation of the domestication process in maize. *Molecular Biology and Evolution,* **21**, 1214-1225.

Whiting AF (1939) Ethnobotany of the Hopi. *Museum of Northern Arizona Bulletin,* **15**, 1-120.

Wills DM, Burke JM (2006) Chloroplast DNA variation confirms a single origin of domesticated sunflower (*Helianthus annuus* L.). *Journal of Heredity,* **97**, 403-408.

Wills DM, Burke JM (2007) Quantitative trait locus analysis of the early domestication of sunflower. *Genetics,* **176**, 2589-2599.

Wills DM, Hester ML, Liu AZ, Burke JM (2005) Chloroplast SSR polymorphisms in the Compositae and the mode of organellar inheritance in *Helianthus annuus*. *Theoretical and Applied Genetics,* **110**, 941-947.

Xiong LX, Liu KD, Dai XK, Xu CG, Zhang QF (1999) Identification of genetic factors controlling domestication-related traits of rice using an F_2 population of a cross between *Oryza sativa* and *O. rufipogon*. *Theoretical and Applied Genetics,* **98**, 243-251.

Yamasaki M, Tenaillon MI, Bi IV, *et al.* (2005) A large scale screen for artificial selection in maize identifies candidate agronomic loci for domestication and crop improvement. *Plant Cell,* **17**, 2859-2872.

Zeng ZB (1993) Theoretical basis for separation of multiple linked gene effects in mapping quantitative trait loci. *Proceedings of the National Academy of Sciences, USA,* **90**, 10972-10976.

Zeng ZB (1994) Precision mapping of quantitative trait loci. *Genetics,* **136**, 1457-1468.

Table 4.1 Comparison of parental population means for the 12 seed traits analyzed. All values are expressed as mean ± standard error.

Trait	Hopi Landrace	Common sunflower
Days to Flower	100.0 ± 4.4	84.4 ± 4.4
Achene Weight (g per 100)	5.4 ± 0.67	0.8 ± 0.06
Pericarp Weight (g per100)	1.3 ± 0.31	0.3 ± 0.04
Kernel Weight (g per 100)	4.0 ± 0.47	0.4 ± 0.04
Kernel-to-Pericarp Weight Ratio	3.9 ± 0.48	1.2 ± 0.12
Achene Length (mm)	10.8 ± 0.62	4.8 ± 0.09
Achene Width (mm)	3.9 ± 0.39	2.2 ± 0.07
Achene Area (mm^2)	35.0 ± 4.82	8.7 ± 0.40
Achene Perimeter (mm)	28.8 ± 1.67	14.3 ± 0.97
Achene Shape	2.9 ± 0.17	2.2 ± 0.06
Oil Content	29.6 ± 1.17	24.5 ± 0.80
Anthocyanin Concentration (OD525/gpericarp)	252.6 ± 59.16	5.0 ± 0.56

Table 4.2 Putative QTL positions, effect magnitudes, and modes of gene action for 12 seed traits using composite interval mapping in an F_2 population derived from a cross between the Hopi landrace and common sunflower.

Trait	Linkage Group[a]	Position[b]	Nearest Marker	1-LOD Interval[c]	Additive Effect[d]	Dominance Ratio[e]	PVE[f]
Days to Flower	6	57.6	ORS483	53.6-57.7	4.6	-0.23	7.6
	7	1.0	ORS1041	0-5.3	-2.6	-0.11	2.5
	15	57.1	ORS687	57-58.2	10.4	-0.49	46.9
Achene Weight	1	6.6	HT1018	2.0-18.3	1.6	0.70	8.6
	8	35.2	ORS456	19.5-35.8	1.7	0	7.6
	9	19.0	HT294	6.0-35.8	1.3	0.37	4.2
	10	15.8	ORS437	9.8-18.9	2.6	0.36	19.0
Pericarp Weight	9	21.0	HT294	15.0-24.5	0.74	0.15	7.7
	10	18.9	ORS613	15.8-24.8	0.21	0.41	16.1
	16a	43.4	ORS399	30.1-49.4	-0.03	-8.43	10.6
	16b	60.1	ORS397	52.1-67.6	-0.16	-1.09	14.8
Kernel Weight	1	7.0	HT1018	2.0-14.4	0.18	0.92	8.0
	7	9.3	ORS966	3.0-19.1	0.19	-0.4	8.0
	8	34.4	ORS456	29.5-35.2	0.17	-0.42	6.7

Table 4.2 Continued.

Trait	Linkage Group[a]	Position[b]	Nearest Marker	1-LOD Interval[c]	Additive Effect[d]	Dominance Ratio[e]	PVE[f]
Kernel Weight	9	22.5	HT294	17.0-24.5	0.21	0.16	9.0
	10	20.9	ORS613	13.8-22.9	0.34	-0.08	22.2
	15	48.5	ORS420	43.1-54.4	0.11	2.46	8.3
Kernel Weight/ Pericarp Weight	16a	26.0	ORS899	18.0-32.1	0.17	-2.57	11.6
	16b	49.4	ORS407	47.4-51.4	0.56	-0.93	41.1
Achene Length	4	4.5	HT664	2.0-24.6	0.19	0.64	5.0
	8	35.2	ORS456	29.5-35.2	0.30	0.48	8.7
	10	13.8	ORS437	11.8-17.1	0.44	0.56	25.9
	16	32.1	ORS899	30.1-36.1	0.41	-0.21	18.3
Achene Width	1	7.0	HT1018	6.6-10.4	0.20	0.50	5.8
	8	35.2	ORS456	34.4-35.2	0.18	-0.03	4.9
	9	17.0	HT294	6.0-24.5	0.20	0.23	5.7
	10	15.8	ORS437	13.8-17.1	0.43	0.41	33.3
Achene Area	8	35.2	ORS456	34.4-35.2	1.94	-0.04	7.8
	9	22.5	HT294	15.0-24.5	1.82	0.57	6.4
	10	15.8	ORS437	13.8-17.1	3.78	0.36	35.4

Table 4.2 Continued.

Trait	Linkage Group[a]	Position[b]	Nearest Marker	1-LOD Interval[c]	Additive Effect[d]	Dominance Ratio[e]	PVE[f]
Achene Perimeter	4	14.6	ORS309	8.6-26.7	0.73	0.73	6.0
	8	35.2	ORS456	27.5-35.2	0.85	0.45	5.5
	10	15.8	ORS437	13.8-17.1	1.62	0.45	26.0
	16	37.4	ORS399	30.1-43.4	0.79	0.78	6.7
Achene Shape	1	7.0	HT1018	6.6-8.0	-0.15	0.47	11.5
	10	15.8	ORS437	11.8-22.9	-0.13	0.48	10.3
	16a	22.0	ORS899	16.0-26.0	0.11	-1.84	16.8
	16b	47.4	ORS407	43.4-51.4	0.16	-0.95	21.3
%Oil Content	3	0.0	ORS202	0.0-2.0	1.49	-0.69	9.9
	16	58.1	ORS397	56.1-67.6	2.87	-0.82	39.5
Anthocyanin	15	58.2	HT284	54.4-58.2	3.52	-0.26	28.7
	16	47.4	ORS407	43.4-49.4	1.89	-1.24	17.7

Table 4.2 Continued.

[a] When multiple QTL for a single trait occurred on the same linkage group, a letter was used to uniquely identify them.

[b] Absolute position from the top of the linkage group (in cM).

[c] Refers to the region flanking each QTL peak in which the LOD score declines by one.

[d] Refers to the additive effect (a) of the Hopi allele. Underlined values indicate instances in which the allelic effects are in the 'wrong' direction. See text for details. Units are the same as Table 4.1 except for oil content and anthocyanin content which were arcsin-square root and Box-Cox transformed, respectively.

[e] Refers to the dominance ratio (d/a) of the Hopi allele.

[f] Percentage of phenotypic variation explained by each QTL using CIM. PVE values for QTL with effects in the direction of the wild phenotype are underlined.

Table 4.3. Summary of significant interactions amongst individually significant QTL.

Trait	Linkage Groups	Type of Interaction[a]	Phenotypic Effect[b]	Effect (%)
Achene Weight	1 × 8	A × A	-0.2	2.0
Pericarp Weight	10 × 16a	D × D	0.3	4.9
	9 × 16b	A × A	-0.1	2.7
Kernel Weight	7 × 10	D × A	-0.2	1.5
	7 × 10	A × A	0.2	0.1
KernelWeight/ Pericarp Weight	16a × 16b	D × D	0.6	3.9
Achene Length	None			
Achene Width	1 × 12	D × A	-0.2	0.7
Achene Area	9 × 10	A × A	1.3	2.7
Achene Perimeter	None			
Achene Shape	16a × 16b	A × A	-0.3	-2.6
	" "	D × A	-0.3	8.1
%Oil Content	None			
Anthocyanin	None			

[a] A × A = Additive × Additive; A × D = Additive × Dominant; D × A = Dominant × Additive; D × D = Dominant × Dominant.

[b] Underlined values indicate an interaction with effects in the 'wrong' direction.

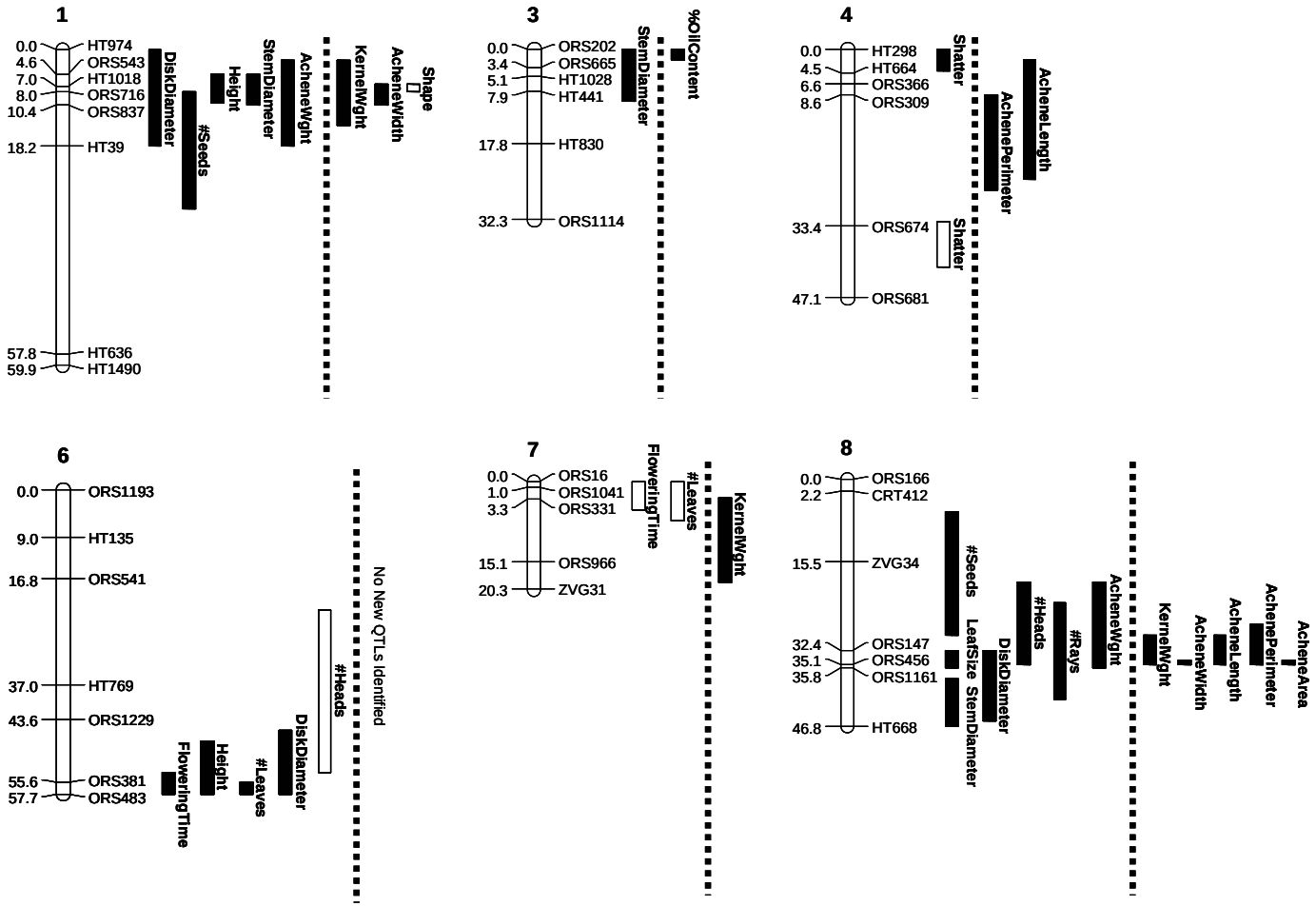
1
0.0 HT974
4.6 ORS543
7.0 HT1018
8.0 ORS716
10.4 ORS837
18.2 HT39
57.8 HT636
59.9 HT1490
DiskDiameter
#Seeds
Height
StemDiameter
AcheneWght
KerneWght
AcheneWidth
Shape
3
0.0 ORS202
3.4 ORS665
5.1 HT1028
7.9 HT441
17.8 HT830
32.3 ORS1114
StemDiameter
%OilContent
4
0.0 HT298
4.5 HT664
6.6 ORS366
8.6 ORS309
33.4 ORS674
47.1 ORS681
Shatter
AchenePerimeter
AcheneLength
Shatter
6
0.0 ORS1193
9.0 HT135
16.8 ORS541
37.0 HT769
43.6 ORS1229
55.6 ORS381
57.7 ORS483
FloweringTime
Height
#Leaves
DiskDiameter
#Heads
No New QTLs identified
7
0.0 ORS16
1.0 ORS1041
3.3 ORS331
15.1 ORS966
20.3 ZVG31
FloweringTime
#Leaves
KerneWght
8
0.0 ORS166
2.2 CRT412
15.5 ZVG34
32.4 ORS147
35.1 ORS456
35.8 ORS1161
46.8 HT668
#Seeds
LeafSize
StemDiameter
DiskDiameter
#Heads
#Rays
AcheneWght
KerneWght
AcheneWidth
AcheneLength
AchenePerimeter
AcheneArea

Figure 4.1

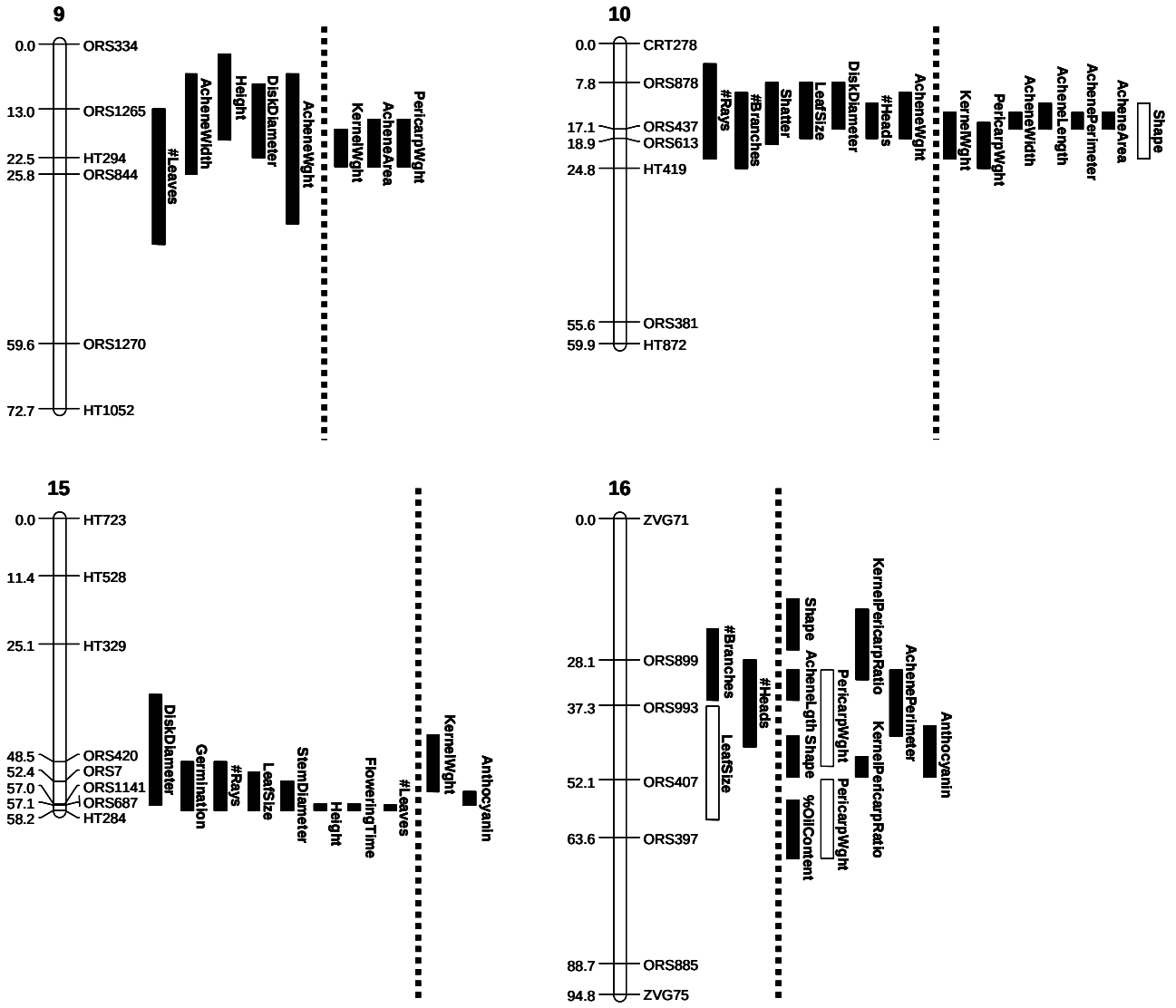

Figure 4.1 Continued.

Figure 4.1 Results of the CIM analysis for the ten linkage groups carrying QTL for the traits of interest. QTL positions are indicated by bars alongside each linkage group. The length of each bar is equal to the 1-LOD support interval for that QTL. Loci at which the crop allele had the expected effect are indicated by a filled bar, whereas those at which the crop allele conferred a wild-like phenotype are represented by unfilled bars. To the left of the dashed lines are QTL identified by Wills and Burke (2007). QTL to the right of the dashed line influence the post-domestication traits phenotyped in this study. Note that LG 6 is included because it harbors a flowering time QTL, though no new QTL were identified on this linkage group.

CHAPTER 5

CONCLUSIONS

Taken together, the results of my research provide detailed insight into the early evolution of cultivated sunflower. The results of my first study are consistent with a single domestication of extant cultivated sunflower in the east-central United States, and thus draw into question the findings of Lentz *et al.* (2001), who argued for a possible origin of domestication in southern Mexico. While a range-wide survey of common sunflower resulted in the identification of 45 unique chloroplast haplotypes, only three chloroplast haplotypes (one common, and two unique) were found in the cultivated lines that were sampled. The primary domesticated haplotype is actually present in all of the Native American landraces except Hidatsa, and has been carried forward into the modern cultivated lines. The Hidatsa haplotype differed from the common domesticated haplotype at two microsatellite loci and was not observed in any wild population. The presence of this haplotype in Hidatsa warrants further investigation. Sharing of the common domesticated haplotype suggests that the Hopi and Havasupai sunflowers are, in fact, basal domesticates, as opposed to being the descendants of a second origin of domestication. A principal coordinate analysis of these data revealed the presence of three major haplotype clusters, one of which contained the primary domesticated haplotype, the two rare domesticated variants, and a variety of haplotypes found across

much of the range of wild sunflower. The wild Mexican haplotypes, on the other hand, fell well outside of the cluster containing the primary domesticated haplotype. While my data do not provide insight into the specific location of sunflower domestication, the relative rarity of the primary domesticated haplotype in the wild, combined with the dissimilarity between this haplotype and those found in the Mexican populations surveyed, suggests that the extant domesticated sunflowers are the product of a single domestication event somewhere outside of Mexico. This conclusion is further corroborated by a recent paper in which Heiser (2008) argued that the seeds recovered by Lentz *et al.* (2001) were really bottle gourd seeds that had been misidentified as sunflower achenes.

Having confirmed that the Hopi sunflower is one of the most primitive domesticated lineages, a wild × Hopi sunflower mapping population was used to investigate the minimum genetic changes required to transform wild sunflower into a useful crop plant. The use of a primitive domesticated lineage successfully removed the effects of improvement which confounded the results of the previous QTL analysis of sunflower domestication (Burke *et al* 2002, 2005; Wills and Burke 2007). Interestingly, this approach resulted in a vast reduction in the number of apparently 'wrong-way' QTL that were identified as compared to earlier findings. Thus, it appears that modern crop improvement efforts have introduced a number of apparently maladaptive alleles into the cultivated sunflower gene pool. As was found in previous research (Burke *et al.* 2002), epistatic interactions were found to play only a minor role in determining the phenotype of cultivated sunflower. The only trait in which epistasis played a prominent role was seed dormancy, which was not investigated in the previous analysis of sunflower

domestication (Burke *et al.* 2002). In this case, a novel QTL (on LG 12) was identified that interacted in an additive-by-additive manner with a QTL on LG 15 to reduce dormancy; this latter QTL co-localizes with one that was previously identified by Ghandi *et al.* (2005). Most importantly, this study confirms that the unique nature of the genetic architecture of sunflower domestication, with a large number of loci, most of which had small to moderate phenotypic effects, contributing to the observed phenotypic transition.

Following the apparent single origin of domestication, sunflower experienced a substantial amount of diversification and improvement resulting in the wide variety of forms that are present in modern germplasm collections. The final component of my research involved using the same wild × Hopi mapping population described above to investigate the genetic architecture of several unique traits that arose within the Hopi lineage, apparently in response to selection following domestication. These traits include a suite of unique achene phenotypes, including an increased length-to-width ratio and a major increase in anthocyanin pigmentation in the achene pericarp, as well as exceptionally late flowering time. The gross morphology of the achenes in the mapping population was found to be largely influenced by two genomic regions (LG 10 and LG 16). The region on LG 10 corresponds to the classically-defined *B* locus, which is known to influence apical branching as well as a variety of other traits. In the wild × Hopi cross, this QTL region acted primarily to increase the seed size in all dimensions. The unique achene shape of the Hopi sunflower, with its much greater length-to-width ratio, is primarily the result of multiple QTL on LG 16. The increased anthocyanin content in the achenes was largely controlled by two genomic regions, including QTL on both LG 15 and LG 16. The QTL on LG 15 co-localized with a QTL that is responsible for the

majority of the variance in flowering time. Thus, unlike typical domestication-related traits, these traits exhibited a relatively simple genetic basis, with two genomic regions being largely responsible for the divergence of the Hopi sunflower from other cultivated sunflower lineages. Whether the simple genetic control of these traits is a result of the complex genetic architecture of domestication or an inherent quality of the underlying developmental pathways of the traits themselves remains an open question.

LITERATURE CITED

Adams KL, Wendel JF (2005) Polyploidy and genome evolution in plants. *Current Opinion in Plant Biology,* **8**, 135-141.

Arias DM, Rieseberg LH (1994) Gene flow between cultivated and wild sunflowers. *Theoretical and Applied Genetics,* **89**, 655-660.

Badr A, Muller K, Schafer-Pregl R, *et al.* (2000) On the origin and domestication history of barley (*Hordeum vulgare*). *Molecular Biology and Evolution,* **17**, 499-510.

Basten CJ, Weir BS, Zeng Z-B (1994) Zmap-a QTL cartographer. In: *Proceedings of the 5th World Congress on Genetics Applied to Livestock Production: Computing Strategies and Software* (eds. Smith C, Gavora JS, Benkel B, *et al.*), pp. 65-66. Organizing Committee, 5th World Congress on Genetics Applied to Livestock Production, Guelph, Ontario, Canada.

Basten CJ, Weir BS, Zeng Z-B (2004) *QTL Cartographer, Version 1.17.* Department of Statistics, North Carolina State University, Raleigh, NC.

Beavis WD (1994) The power and deceit of QTL experiments: lesson from comparative QTL studies. *In: Proceedings of the 49th Annual Corn and Sorghum Research Conference*, pp. 250-266. American Seed Trade Association.

Berthoud EL (1892) A peculiar case of plant dissemination. *The Botanical Gazette,* **17**, 321-326.

Bost B, de Vienne D, Hospital F, Moreau L, Dillmann C (2001) Genetic and nongenetic bases for the L-shaped distribution of quantitative trait loci effects. *Genetics,* **157**, 1773-1787.

Bradshaw HD, Otto KG, Frewen BE, McKay JK, Schemske DW (1998) Quantitative trait loci affecting differences in floral morphology between two species of monkeyflower (*Mimulus*). *Genetics,* **149**, 367-382.

Brewer AJ (1973) Analysis of floral remains from the Higgs site (40L045). In: *Excavation of the Higgs and Doughty Site. I-75 Salvage Archaeology Tennessee Archaeological Society, Miscellaneous Papers No. 12* (eds. McCullough MCR, Faulknor CJ), pp. 141-147. University of Tennessee, Knoxville, TN.

Brewer MT, Lang LX, Fujimura K, *et al.* (2006) Development of a controlled vocabulary and software application to analyze fruit shape variation in tomato and other plant species. *Plant Physiology,* 141, 15-25.

Bryan GJ, McNicoll J, Ramsay G, Meyer RC, DeJong WS (1999) Polymorphic simple sequence repeat markers in chloroplast genomes of Solanaceous plants. *Theoretical and Applied Genetics,* **99**, 859-867.

Bullard RW, Woronecki PP, Dolbeer RA, Mason JR (1989) Biochemical and morphological characteristics in maturing achenes from purple-hulled and oilseed sunflower cultivars. *Journal of Agricultural and Food Chemistry,* **37**, 886-890.

Burke JM, Burger JC, Chapman MA (2007) Crop evolution: from genetics to genomics. *Current Opinion in Genetics and Development,* **17**, 525-532.

Burke JM, Knapp SJ, Rieseberg LH (2005) Genetic consequences of selection during the evolution of cultivated sunflower. *Genetics,* **171**, 1933-1940.

Burke JM, Tang S, Knapp SJ, Rieseberg LH (2002) Genetic analysis of sunflower domestication. *Genetics,* **161**, 1257-1267.

Churchill GA, Doerge RW (1994) Empirical threshold values for quantitative trait mapping. *Genetics,* **138**, 963-971.

Crites GD (1993) Domesticated sunflower in 5th millennium B.P. temporal context: new evidence from middle Tennessee. *American Antiquity,* **58**, 146-148.

Cronn R, Brothers M, Klier K, Bretting PK, Wendel JF (1997) Allozyme variation in domesticated annual sunflower and its wild relatives. *Theoretical and Applied Genetics,* **95**, 532-545.

Doebley J, Bacigalupo A, Stec A (1994) Inheritance of kernel weight in 2 maize-teosinte hybrid populations: implications for crop evolution. *Journal of Heredity,* **85**, 191-195.

Doebley J, Stec A (1991) Genetic analysis of the morphological differences between maize and teosinte. *Genetics,* **129**, 285-295.

Doebley J, Stec A (1993) Inheritance of the morphological differences between maize and teosinte: comparison of results for 2 F_2 populations. *Genetics,* **134**, 559-570.

Doebley J, Stec A, Wendel J, Edwards M (1990) Genetic and morphological analysis of a maize-teosinte F_2 population: implications for the origin of maize. *Proceedings of the National Academy of Sciences, USA,* **87**, 9888-9892.

Doerge RW, Churchill GA (1996) Permutation tests for multiple loci affecting a quantitative character. *Genetics* **142**, 285-294.

Excoffier L, Laval G, Schneider S (2005) Arlequin Ver. 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online,* **1**, 47-50.

Eyre-Walker A, Gaut RL, Hilton H, Feldman DL, Gaut BS (1998) Investigation of the bottleneck leading to the domestication of maize. *Proceedings of the National Academy of Sciences, USA,* **95**, 4441-4446.

Fernandez-Martinez J, Knowles PF (1978) Inheritance of self-incompatibility in wild sunflower. In: *8th International Sunflower Conference*, pp. 484-489. International Sunflower Association, Paris.

Fick GN, Miller JF (1997) Sunflower breeding. In: *Sunflower Production and Technology* (ed. Scheiter AA), pp. 395-440. American Society of Agronomy, Madison, WI.

Fishman L, Kelly AJ, Willis JH (2002) Minor quantitative trait loci underlie floral traits associated with mating system divergence in *Mimulus*. *Evolution,* **56**, 2138-2155.

Ford RI (1985) Some botanical considerations of the early domesticated plants north of Mexico. In: *Prehistoric Food Production in North America, Museum of Anthropology, Anthropological Paper No. 75,* (ed. Ford RI), pp. 1-18. University of Michigan, Ann Arbor, MI.

Gandhi SD, Heesacker AF, Freeman CA, *et al.* (2005) The self-incompatibility locus (*S*) and quantitative trait loci for self-pollination and seed dormancy in sunflower. *Theoretical and Applied Genetics*, **111**, 619-629.

Hammer K (1984) Das Domestikationssyndrom. *Kulturpflanze,* 32, 11-34.

Hancock JF (2004) *Plant Evolution and the Origin of Crop Species.* CABI Publishing, Cambridge, MA.

Hanson MA, Gaut BS, Stec AO*, et al.* (1996) Evolution of anthocyanin biosynthesis in maize kernels: The role of regulatory and enzymatic loci. *Genetics,* **143**, 1395-1407.

Harlan JR (1992) *Crops & Man*, 2nd edn. American Society of Agronomy, Madison, WI.

Harter AV, Gardner KA, Falush D*, et al.* (2004) Origin of extant domesticated sunflowers in eastern North America. *Nature,* **430**, 201-205.

Heiser CB (1945) The Hopi sunflower. *Missouri Botanical Garden Bulletin* **33**, 163-167.

Heiser CB (1946) A "new" cultivated sunflower from Mexico. *Madrono,* **8**, 226-229.

Heiser CB (1950) The sunflower among the North American Indians. *American Journal of Botany,* **37**, 662-662.

Heiser CB (1951) The sunflower among the North American Indians. *Proceedings of the American Philosophical Society,* **95**, 432-448.

Heiser CB (1954) Variation and subspeciation in the common sunflower, *Helianthus annuus*. *American Midland Naturalist,* **51**, 287-305.

Heiser CB, Jr. (1978) Taxonomy of *Helianthus* and origin of domesticated sunflower. In: *Sunflower Science and Technology* (ed. Carter JF), pp. 31-53. American Society of Agronomy, Madison, WI.

Heiser CB, Jr. (1985) Some botanical considerations of the early domesticated plants north of Mexico. In: *Prehistoric Food Production in North America, Museum of Anthropology, Anthropological Paper No. 75,* (ed. Ford RI), pp. 57-72. University of Michigan, Ann Arbor, MI.

Heiser CB (2008) The sunflower (*Helianthus annuus*) in Mexico: further evidence for a North American domestication. *Genetic Resources and Crop Evolution,* **55**, 9-13.

Heiser CB, Jr., Smith DM, Clevenger SB, Martin WC, Jr. (1969) The North American sunflowers (*Helianthus*). *Memoirs of the Torrey Botanical Club,* **22**, 1-213.

Ivanov IG (1975) Study on compatibility and incompatibility display in crossing selfed sunflower lines. *Rastenievud Nauki,* **12**, 36-40.

Joshi SS, Basavalingappa, Giriraj K (1994) Pleiotropy in sunflower (*Helianthus annuus* L.). *Helia,* **20**, 1-6.

Kao CH, Zeng ZB (1997) General formulas for obtaining the MLEs and the asymptotic variance-covariance matrix in mapping quantitative trait loci when using the EM algorithm. *Biometrics,* **53**, 653-665.

Kao CH, Zeng ZB, Teasdale RD (1999) Multiple interval mapping for quantitative trait loci. *Genetics,* **152**, 1203-1216.

Koinange EMK, Singh SP, Gepts P (1996) Genetic control of the domestication syndrome in common bean. *Crop Science,* **36**, 1037-1045.

Kolkman JM, Berry ST, Leon AJ, *et al.* (2007) Single nucleotide polymorphisms and linkage disequilibrium in sunflower. *Genetics,* **177**, 457-468.

Ladizinsky G (1985) Founder effect in crop plant evolution. *Economic Botany,* **39**, 191-199.

Lai Z, Livingstone K, Zou Y, *et al.* (2005) Identification and mapping of SNPs from ESTs in sunflower. *Theoretical and Applied Genetics,* **111**, 1532-1544.

Lander ES, Green P, Abrahamson J, *et al.* (1987) MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. *Genomics,* **1**, 174-181.

Lentz DL, Pohl MED, Pope KO, Wyatt AR (2001) Prehistoric sunflower (*Helianthus annuus* L.) domestication in Mexico. *Economic Botany,* **55**, 370-376.

Leon AJ, Lee M, Rufener GK, Berry ST, Mowers RP (1995) Use of RFLP markers for genetic linkage analysis of oil percentage in sunflower seed. *Crop Science,* **35**, 558-564.

Leon AJ, Lee M, Rufener GK, Berry ST, Mowers RP (1996) Genetic mapping of a locus (*hyp*) affecting seed hypodermis color in sunflower. *Crop Science,* **36**, 1666-1668.

Lester RN (1989) Evolution under domestication involving disturbance of genic balance. *Euphytica,* **44**, 125-132.

Lincoln SE, Lander ES (1992) Systematic detection of errors in genetic linkage data. *Genomics,* **14**, 604-610.

Linder CR, Taha I, Seiler GJ, Snow AA, Rieseberg LH (1998) Long term introgression of crop genes into wild sunflower populations. *Theoretical and Applied Genetics,* **96**, 339-347.

Liu AZ, Burke JM (2006) Patterns of nucleotide diversity in wild and cultivated sunflower. *Genetics* **173**, 321-330.

Liu B, Fujita T, Yan ZH, *et al.* (2007) QTL mapping of domestication-related traits in soybean (*Glycine max*). *Annals of Botany,* **100**, 1027-1038.

Matsuoka Y, Vigouroux Y, Goodman MM, *et al.* (2002) A single domestication for maize shown by multilocus microsatellite genotyping. *Proceedings of the National Academy of Sciences, USA,* **99**, 6080-6084.

Mazza G, Gao L (1994) Malonylated anthocyanins in purple sunflower seeds. *Phytochemistry,* **35**, 237-239.

Miller JF, Fick GN (1997) The genetics of sunflower. In: *Sunflower Production and Technology* (ed. Scheiter AA), pp. 441-496. American Society of Agronomy, Madison, WI.

Mosjidis JA (1982) Inheritance of color in the pericarp and corolla of the disk florets in sunflower. *Journal of Heredity,* **73**, 461-464.

Orr HA (1998) Testing natural selection vs. genetic drift in phenotypic evolution using quantitative trait locus data. *Genetics,* **149**, 2099-2104.

Paterson AH, Damon S, Hewitt JD, *et al.* (1991) Mendelian factors underlying quantitative traits in tomato: Comparison across species, generations, and environments. *Genetics,* **127**, 181-197.

Paterson AH, Saranga Y, Menz M, Jiang CX, Wright RJ (2003) QTL analysis of genotype x environment interactions affecting cotton fiber quality. *Theoretical and Applied Genetics,* **106**, 384-396.

Peakall R, Smouse PE (2006) Genalex 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes,* 6, 288-295.

Perez-Vich B, Berry ST, Velasco L, *et al.* (2005) Molecular mapping of nuclear male sterility genes in sunflower. *Crop Science,* **45**, 1851-1857.

Powell W, Morgante M, McDevitt R, Vendramin GG, Rafalski JA (1995) Polymorphic simple sequence repeat regions in chloroplast genomes: applications to the population genetics of pines. *Proceedings of the National Academy of Sciences, USA,* **92**, 7759-7763.

Provan J, Powell W, Hollingsworth PM (2001) Chloroplast microsatellites: new tools for studies in plant ecology and evolution. *Trends in Ecology and Evolution,* **16**, 142-147.

Putt ED (1997) Early history of sunflower. In: *Sunflower Production and Technology* (ed. Scheiter AA), pp. 1-19. American Society of Agronomy, Madison, WI.

Rieseberg LH, Seiler GJ (1990) Molecular evidence and the origin and development of the domesticated sunflower (*Helianthus annuus*, Asteraceae). *Economic Botany* **44** (Supplement 3), 79-91.

Rieseberg LH, Widmer A, Arntz AM, Burke JM (2002) Directional selection is the primary cause of phenotypic diversification. *Proceedings of the National Academy of Sciences, USA,* **99**, 12242-12245.

Ross-Ibarra J (2005) Quantitative trait loci and the study of plant domestication. *Genetica* **123**, 197-204.

Sanguinetti CJ, Neto ED, Simpson AJG (1994) Rapid silver staining and recovery of PCR products separated on polyacrylamide gels. *Biotechniques,* **17**, 914-921.

Schuelke M (2000) An economic method for the fluorescent labeling of PCR fragments. *Nature Biotechnology,* **18**, 233-234.

Second G (1982) Origin of the genic diversity of cultivated Rice (*Oryza*-Spp): Study of the polymorphism scored at 40 isoenzyme loci. *Japanese Journal of Genetics,* **57**, 25-57.

Seiler GJ, Rieseberg LH (1997) Systematics, origin, and germplasm resources of the wild and domesticated sunflower. In: *Sunflower Production and Technology* (ed. A. SA), pp. 21-65. American Society of Agronomy, Madison, WI.

Slabaugh MB, Huestis GM, Leonard J*, et al.* (1997) Sequence based genetic markers for genes and gene families: Single strand conformational polymorphisms for the fatty acid synthesis genes of Cuphea. *Theoretical and Applied Genetics,* **94**, 400-408.

Slatkin M (1994) Linkage disequilibrium in growing and stable populations. *Genetics,* **137**, 331-336.

Sokal RR, Rohlf FJ (1995) *Biometry*, 3rd edn. W.H. Freeman and Company, New York.

Sossey-Alaoui K, Serieys H, Tersac M*, et al.* (1998) Evidence for several genomes in *Helianthus*. *Theoretical and Applied Genetics,* **97**, 422-430.

Spooner DM, McLean K, Ramsay G, Waugh R, Bryan GJ (2005) A single domestication for potato based on multilocus amplified fragment length polymorphism genotyping. *Proceedings of the National Academy of Sciences, USA,* **102**, 14694-14699.

Stoenescu F (1974) Genetics. In: *Floarea-soarelui* (ed. Vranceau AV), pp. 92-125. academiei Republicii Socialiste, Bucuresti, Romania.

Stuber CW, Moll RH, Goodman MM, Schaffer HE, Weir BS (1980) Allozyme frequency changes associated with selection for increased grain yield in maize (*Zea Mays* L). *Genetics,* **95**, 225-236.

Tang S, Yu JK, Slabaugh MB, Shintani DK, Knapp SJ (2002) Simple sequence repeat map of the sunflower genome. *Theoretical and Applied Genetics,* **105**, 1124-1136.

Tang SX, Kishore VK, Knapp SJ (2003) PCR-multiplexes for a genome-wide framework of simple sequence repeat marker loci in cultivated sunflower. *Theoretical and Applied Genetics,* **107**, 6-19.

Tang SX, Knapp SJ (2003) Microsatellites uncover extraordinary diversity in Native American land races and wild populations of cultivated sunflower. *Theoretical and Applied Genetics,* **106**, 990-1003.

Tang SX, Leon A, Bridges WC, Knapp SJ (2006) Quantitative trait loci for genetically correlated seed traits are tightly linked to branching and pericarp pigment loci in sunflower. *Crop Science,* **46**, 721-734.

Tenaillon MI, U'Ren J, Tenaillon O, Gaut BS (2004) Selection versus demography: A multilocus investigation of the domestication process in maize. *Molecular Biology and Evolution,* **21**, 1214-1225.

Wang R-L, Stec A, Hey J, Lukens L, Doebley J (1999) The limits of selection during maize domestication. *Nature,* **398**, 236-239.

Weising K, Gardner RC (1999) A set of conserved PCR primers for the analysis of simple sequence repeat polymorphisms in chloroplast genomes of dicotyledonous angiosperms. *Genome,* **42**, 9-19.

Wendel JF, Schnabel A, Seelanan T (1995) Bidirectional interlocus concerted evolution following allopolyploid speciation in cotton (*Gossypium*). *Proceedings of the National Academy of Sciences, USA,* **92**, 280-284.

Whiting AF (1939) Ethnobotany of the Hopi. *Museum of Northern Arizona Bulletin,* **15**, 1-120.

Whitton J, Wolf DE, Arias DM, Snow AA, Rieseberg LH (1997) The persistence of cultivar alleles in wild populations of sunflowers five generations after hybridization. *Theoretical and Applied Genetics,* **95**, 33-40.

Wills DM, Burke JM (2006) Chloroplast DNA variation confirms a single origin of domesticated sunflower (*Helianthus annuus* L.). *Journal of Heredity,* **97**, 403-408.

Wills DM, Burke JM (2007) Quantitative trait locus analysis of the early domestication of sunflower. *Genetics,* **176**, 2589-2599.

Wills DM, Hester ML, Liu AZ, Burke JM (2005) Chloroplast SSR polymorphisms in the Compositae and the mode of organellar inheritance in *Helianthus annuus*. *Theoretical and Applied Genetics,* **110**, 941-947.

Xiao JH, Li JM, Grandillo S*, et al.* (1998) Identification of trait-improving quantitative trait loci alleles from a wild rice relative, *Oryza rufipogon*. *Genetics,* **150**, 899-909.

Xiong LX, Liu KD, Dai XK, Xu CG, Zhang QF (1999) Identification of genetic factors controlling domestication-related traits of rice using an F_2 population of a cross between *Oryza sativa* and *O. rufipogon*. *Theoretical and Applied Genetics,* **98**, 243-251.

Xu DH, Abe J, Gai JY, Shimamoto Y (2002) Diversity of chloroplast DNA SSRs in wild and cultivated soybeans: Evidence for multiple origins of cultivated soybean. *Theoretical and Applied Genetics,* **105**, 645-653.

Yamasaki M, Tenaillon MI, Bi IV, *et al.* (2005) A large scale screen for artificial selection in maize identifies candidate agronomic loci for domestication and crop improvement. *Plant Cell,* **17**, 2859-2872.

Yu JK, Tang S, Slabaugh MB, *et al.* (2003) Towards a saturated molecular genetic linkage map for cultivated sunflower. *Crop Science,* **43**, 367-387.

Zeng ZB (1993) Theoretical basis for separation of multiple linked gene effects in mapping quantitative trait loci. *Proceedings of the National Academy of Sciences, USA,* **90**, 10972-10976.

Zeng ZB (1994) Precision mapping of quantitative trait loci. *Genetics,* **136**, 1457-1468.