

High genetic diversity in a rare and endangered sunflower as compared to a common congener

J. R. ELLIS, C. H. PASHLEY, J. M. BURKE* and D. E. McCAULEY

Vanderbilt University, Department of Biological Sciences, VU Station B 35-1634, Nashville, TN 37232, USA

Abstract

Determining the genetic structure of isolated or fragmented species is of critical importance when planning a suitable conservation strategy. In this study, we use nuclear and chloroplast SSRs (simple sequence repeats) to investigate the population genetics of an extremely rare sunflower, *Helianthus verticillatus* Small, which is known from only three locations in North America. We investigated levels of genetic diversity and population structure compared to a more common congener, *Helianthus angustifolius* L., using both nuclear and chloroplast SSRs. We also investigated its proposed hybrid origin from *Helianthus grosseserratus* Martens and *H. angustifolius*. Twenty-two nuclear SSRs originating from the cultivated sunflower (*Helianthus annuus* L.) expressed sequence tag (EST) database, and known to be transferable to *H. verticillatus* and its putative parental taxa, were used in this study thereby allowing for statistical control of locus-specific effects in population genetic analyses. Despite its rarity, *H. verticillatus* possessed significantly higher levels of genetic diversity than *H. angustifolius* at nuclear loci and equivalent levels of chloroplast diversity. Significant levels of population subdivision were observed in *H. verticillatus* but of a magnitude comparable to that of *H. angustifolius*. Inspection of multilocus genotypes also revealed that clonal spread is highly localized. Finally, we conclude that *H. verticillatus* is not of hybrid origin as it does not exhibit a mixture of parental alleles at nuclear loci, and it does not share a chloroplast DNA haplotype with either of its putative parents.

Keywords: chloroplast DNA (cpDNA), conservation, EST-SSRs, genetic diversity, *Helianthus*, hybridization

Received 15 November 2005; revision received 27 January 2006; accepted 13 February 2006

Introduction

The genetics of endangered species have been of great interest to both evolutionary biologists and conservation managers for some time (Hedrick 2001, e.g. volumes by Falk & Holsinger 1991; Avise & Hamrick 1996; Young & Clarke 2000). Conservation biologists are interested in knowing if there are generalities that can be made with regard to rare species, such as whether or not they typically exhibit reduced genetic diversity or restricted gene flow between populations, as predicted by population genetic theory when populations are small and isolated. While there is indeed a trend for rare species to exhibit reduced genetic diversity, some exhibit equivalent levels of diver-

sity compared to their common congeners (Gitzendanner & Soltis 2000). In order to determine if a rare species does exhibit low diversity, we must have a measure or standard with which to compare. Many studies of rare plants can make comparisons against other plant species that share similar life histories by making use of compendiums of studies of genetic diversity that utilize allozyme (Hamrick & Godt 1989), RAPD (randomly amplified polymorphic DNA), or anonymous SSR (simple sequence repeat) markers (Nyblom 2004). However, when using novel types of markers for which no compendium exists, comparisons against a common congener provide a useful standard against which rare species can be evaluated. Such comparisons minimize the confounding effects of phylogeny and life history on population genetic parameters (Felsenstein 1985; Karron 1987, 1991; Baskauf *et al.* 1994).

In addition to considering species-wide levels of diversity, knowledge of the partitioning of genetic variation within and between populations, or population structure,

Correspondence: Jennifer R. Ellis, Fax: (615) 343-6707; E-mail: jennifer.ellis@vanderbilt.edu

Address as of August 1 2006: Department of Plant Biology, University of Georgia, Athens, GA 30602, USA.

is important when considering a conservation strategy for an endangered species, especially if not all populations can be protected. With a low level of population structure, the loss of a single population may have little impact on the species-wide genetic diversity. With a high level of structure, the loss of a single population might significantly reduce overall genetic variation. Further, a species whose distribution has been reduced to small and isolated populations may be at particular risk of extinction due to: (i) the fixation of deleterious alleles within populations as a result of inbreeding due to restricted gene flow, (ii) reduced genetic variation, and consequently, an inability to adapt to a changing environment (Barrett & Kohn 1991), and/or (iii) demographic or environmental stochasticity (Lande 1988, 1993). In fact, Lande (1988) argues that demographic factors may have a more immediate effect on population persistence than genetic factors. Despite this, Reed & Frankham (2003) found a significant positive correlation between heterozygosity and fitness in a meta-analysis of 34 plant and animal data sets, indicating that genetic variability is an important component to consider when formulating management plans.

Levels and patterns of genetic diversity can also be shaped by natural hybridization, wherein hybrids might exhibit elevated levels of genetic diversity resulting from the mixing of parental genomes (Rieseberg & Wendel 1993; Arnold 1997). Given this possibility, comparisons made with a common congener may not accurately reflect the overall effects of rarity on the level of genetic variation found within a rare hybrid derivative. Thus, for rare or endangered species in which a hybrid origin has been suggested, it is important to test for hybridity. Hybrid ancestry is most reliably established with molecular data (Rieseberg & Ellstrand 1993; Chapman & Abbott 2005), and conservation studies that use appropriate genetic markers may help to characterize genetic relationships between taxa when hybridization and introgression might have occurred (e.g. Bruneau *et al.* 2005). Furthermore, hybrid ancestry can impact the status of endangered species that might otherwise be eligible for listing under the US Endangered Species Act, as the listing of hybrids has sometimes been difficult (Allendorf *et al.* 2001).

Here, we report the results of a population genetic survey of a rare and endangered sunflower species, *Helianthus verticillatus* Small, using both EST (expressed sequence tag) nuclear and chloroplast SSRs. Our nuclear markers are derived from the cultivated sunflower (*Helianthus annuus* L) EST database (<http://cgpdb.ucdavis.edu>) and have been found to be highly transferable across species within the genus *Helianthus* (CH Pashley & JR Ellis, unpublished data). As such, they are particularly useful for our purposes. Similarly, the chloroplast markers that we employed have been used successfully in species from across the Compositae (Wills *et al.* 2005). Thus, we were

able to include in our survey a more widely distributed congener (*Helianthus angustifolius* L.) as a phylogenetic and life history control whilst statistically controlling for inherent differences in the level of genetic variation from one locus to another. Specifically, we compare the species with regard to the level of standing genetic variation found at these markers and the degree to which that variation is partitioned among populations. Further, since *H. verticillatus* is thought to be clonal, we used EST-SSRs to determine if closely spaced stalks were indeed a single genetic individual or perhaps represented several individuals. This is an important issue when evaluating genetic effective population size from census data. Finally, we looked for a genetic signature of hybridization in *H. verticillatus* through a comparison with its putative parents, *H. angustifolius* and *Helianthus grosseserratus* Martens. If *H. verticillatus* is a hybrid, then its genome should consist of a mixture of alleles from its parents.

Materials and methods

Study species

The whorled sunflower, *Helianthus verticillatus*, is an extremely rare, diploid ($n = 17$), perennial restricted to only three locations in the southeast interior of the United States: one in western Tennessee (35.49N, -88.72W; Madison County), one in northeastern Alabama (34.13N, -85.44; Cherokee County), and one in northwestern Georgia (34.14N, -85.38 W; Floyd County). This species is a candidate for federal listing for the Endangered Species Act and is listed as endangered in each of the three states. First collected in western Tennessee in 1892, *H. verticillatus* was not found again in the field until 1994 in Georgia (Matthews *et al.* 2002). In 1996 and 1998, populations of *H. verticillatus* in Alabama and Tennessee were also discovered. The Alabama and Georgia populations are about 3.5 km from each other whereas the Tennessee population is about 350 km from the others. The soil type in the Alabama and Georgia habitats is deep, poorly drained soils formed in alluvium and residuum from limestone, and the Tennessee soil type is silt loam from alluvial deposits of Tertiary Porters Creek clay (Matthews *et al.* 2002). *H. verticillatus* is clonal with slender rhizomes, a glaucous stem, leaves mostly verticillate in threes or fours, prefers wet habitats, and flowers August to October. This species ranges in height from 0.6 to 4.2 m, and its clones occur in somewhat distinct clusters in nature.

There is no information available on the historical range of *H. verticillatus*; the species may represent a narrow endemic or a relictual species that was once more extensive throughout this region. Since this species was not collected during most of the 20th century, several authors studying the 1892 herbarium specimen speculated that it might be of

hybrid origin, having resulted from matings between either *Helianthus angustifolius* ($n = 17$) \times *Helianthus eggertii* Small ($n = 51$) (Beatley 1963) or *H. angustifolius* \times *Helianthus grosseserratus* ($n = 17$) (Heiser *et al.* 1969). The former hybrid combination seems unlikely since *H. eggertii* is hexaploid and *H. verticillatus* is diploid. However, *H. angustifolius* and *H. grosseserratus* are both diploid species and might reasonably represent the parents of *H. verticillatus*. Matthews *et al.* (2002) reported on the current status of the species and concluded on the basis of several morphological characters that *H. verticillatus* should be considered to be a distinct species. However, the USDA Plant Database (<http://plants.usda.gov/>) continues to list *H. verticillatus* as a hybrid between *H. angustifolius* and *H. grosseserratus*.

Helianthus angustifolius, a close relative of *H. verticillatus*, is commonly distributed over most of the eastern United States from New York to Florida and west to Texas. The species is perennial with slender or lacking rhizomes, leaves linear to narrowly lanceolate, usually found in moist, shady areas, and flowers from September to October (Heiser *et al.* 1969). In the areas where *H. verticillatus* is located, *H. angustifolius* is the most common sunflower species. *H. grosseserratus* is a perennial with short to medium rhizomes, leaves lanceolate to ovate, found in dry to moderately wet prairies, and flowers August to October. The species is also commonly distributed across the eastern United States from New England to South Dakota and south to Texas (Heiser *et al.* 1969). All three species are members of the section *Atrorubens* within the genus *Helianthus* (Seiler & Gulya 2004), have overlapping distribution ranges, and are outcrossers pollinated by generalists.

Collection of plant material and DNA extraction

Leaf material of *H. verticillatus* was collected from 22, 22, and 27 clusters of stalks found in the three known locations in Tennessee (TN), Georgia (GA), and Alabama (AL), respectively. The species grows in clusters of up to five or six stalks, separated from other clusters by at least 1 m. In order to determine if each cluster represented a single clone we collected leaves from two to three stalks per cluster for analysis. The total number of clusters varies among the populations, with about 70 in Tennessee and 30 in Georgia. In Alabama, the species is not found in well-defined clusters as in the other two populations — there are about 200–300 stalks. Individuals collected from GA were found in a single field, as was the case for the AL population. In contrast, the TN collection consisted of three subpatches separated by 100–200 m.

Helianthus angustifolius leaf material was collected from two locations: (i) a population located about 10 km from the *H. verticillatus* TN population, and (ii) a large continuous

population consisting of thousands of plants, which connects the GA and AL *H. verticillatus* populations. We collected 13 individuals from the TN population and 25 individuals from throughout the continuous GA/AL site (hereafter referred to as the AL population). While *H. grosseserratus* is known to occur in Tennessee, difficulties in making collections from TN populations required us to obtain seeds from the North Central Regional Plant Introduction Station (NCRPIS; Ames, IA). Seeds were nicked with a razor blade, germinated on moist filter paper, and grown in the Vanderbilt University Department of Biological Sciences greenhouse. When the resulting plants were large enough, a leaf was collected for DNA extraction. Sampled accessions were: South Dakota (NCRPIS accession Ames 2742), North Dakota (Ames 22739), Wisconsin (PI 547187), Illinois (PI 547205) and Iowa (PI 613793). In all species, total genomic DNA was isolated from ~200 mg of fresh leaf tissue using the Doyle & Doyle (1987) CTAB method. All DNA samples were quantified using a TKO-100 fluorometer (Hoefer Scientific Instruments).

Selection of loci, PCR conditions, and genotyping

Twenty-two EST-SSR loci developed for *Helianthus annuus* and proven cross-transferable to *H. verticillatus* were chosen as genetic markers for this study (CH Pashley & JR Ellis, unpublished data). Nineteen loci amplified in both *H. verticillatus* and *H. angustifolius*, and the remaining three amplified in *H. verticillatus* alone (Table 1). All but one of the 22 EST-SSRs amplified in *H. grosseserratus*. For the survey of cpDNA (chloroplast DNA) variation and hybrid origin, three polymorphic chloroplast SSRs [cpSSRs: N39 and N30 (Bryan *et al.* 1999) and C7 (Weising & Gardner 1999)] were analysed in the three species.

SSR genotyping was performed using a modified version of the fluorescent labelling protocol of Schuelke (2000), as detailed in Wills *et al.* (2005). Polymerase chain reaction (PCR) was performed in a total volume of 20 μ L containing 2 ng of template DNA for *H. verticillatus*, or 10 ng of DNA in the cases of both *H. angustifolius* and *H. grosseserratus*, 30 mM Tricine pH 8.4-KOH, 50 mM KCl, 2 mM $MgCl_2$, 125 μ M of each dNTP, 0.2 μ M M13 Forward (-29) sequencing primer labelled with either VIC, 6-FAM or TET, 0.2 μ M reverse primer, 0.02 μ M forward primer and 2 U of *Taq* polymerase. The PCR conditions were as follows: 3 min at 95 $^{\circ}$ C; 10 cycles of 30 s at 94 $^{\circ}$ C, 30 s at 65 $^{\circ}$ C and 45 s at 72 $^{\circ}$ C, annealing temperature decreasing to 55 $^{\circ}$ C by 1 $^{\circ}$ C per cycle, followed by 30 cycles of 30 s at 94 $^{\circ}$ C, 30 s at 55 $^{\circ}$ C, 45 s at 72 $^{\circ}$ C, followed by 20 min at 72 $^{\circ}$ C.

PCR products were visualized on an MJ Research Base-Station automated DNA sequencer, and MapMarker 1000 ROX size standards (BioVentures) were run in each

Table 1 Locus descriptions and mean values for Nei's gene diversity (1987) for *Helianthus verticillatus* and *Helianthus angustifolius*

Locus	Type of repeat	Location	<i>H. verticillatus</i>	<i>H. angustifolius</i>
BL0001	tri	UTR	0.711	0.336
BL0002†‡	tetra	UTR	0.133	N/A
BL0003	tri	Coding	0.689	0.704
BL0004‡	tri	Coding	0.347	0.090
BL0005‡	tri	UTR	0.817	0.699
BL0006†‡	tetra	UTR	0.656	N/A
BL0007‡	tri	Coding	0.609	0
BL0008‡	tri	Coding	0.777	0.646
BL0010	tri	Coding	0.703	0.508
BL0011	tetra	UTR	0.586	0
BL0012‡	tetra	UTR	0.222	0
BL0013†	tetra	UTR	0.659	N/A
BL0014	tri	Coding	0	0.213
BL0017‡	tri	UTR	0.533	0.442
BL0018†§	tri	Coding	0.738	0.784
BL0020	tri	Coding	0.493	0.529
BL0022	tri	UTR	0	0
BL0023	tri	UTR	0.381	0.419
BL0025	tetra	Coding	0.651	0.400
BL0027	tri	Coding	0.820	0.767
BL0029	tetra	UTR	0	0
BL0030	tetra	UTR	0	0
Mean			0.478 ± 0.06*	0.344 ± 0.07*

*Mean values ± SE are significantly different from each other at the $P < 0.05$ level, two-factor ANOVA, see text for details.

†EST-SSRs that amplified in *H. verticillatus* only.

‡Nine EST-SSRs genotyped for analysis of clonal structure.

§EST-SSR that did not amplify in *H. grosseserratus*.

lane to allow for accurate determination of fragment size. CARTOGRAPHER version 1.2.6 (MJ Research) was used to infer individual genotypes according to the fragment sizes of the PCR products.

Analysis of clonal structure

Leaves collected from 13 putative clones of *H. verticillatus* were initially genotyped for nine arbitrarily selected polymorphic EST-SSR loci (Table 1 and see below). The probability that each cluster was a single genet and that identical genotypes were not simply obtained by chance, was calculated using a multilocus probability for codominant genotypes, $P_{\text{cgen}} = (\prod p_i)2^h$, where p_i is the frequency of each allele observed in the multilocus genotype and h is the number of heterozygous loci (Parks & Werth 1993; Sydes & Peakall 1998). The probability of obtaining $n - 1$ more copies of that genotype by chance is given by $(P_{\text{cgen}})^{n-1}$, where n is the number of times the genotype was observed.

Analyses of genetic diversity and population structure

Measures of genetic diversity, including mean number of alleles, observed and expected heterozygosities, and the inbreeding coefficient (f , Weir & Cockerham 1984) were calculated for each population of *H. verticillatus* and *H. angustifolius* using GDA version 1.0 (Lewis & Zaykin 2001). Unbiased gene diversities (Nei 1987) were calculated for each locus in the two species using FSTAT (Goudet 2001). Since *H. grosseserratus* seeds were obtained from NCRPIS and the collections were made without knowledge of specific location, this species was not included in the population genetic calculations but was used in the investigation of hybrid origin.

ANOVAS were performed using JMP version 4 (SAS Institute) to test for significant differences in measures of genetic diversity between *H. verticillatus* and *H. angustifolius*, without regard to population. We ran the models with the 19 markers that amplified in both species (Table 1). The main effects included in the ANOVA were species and locus, with the dependent variables being either the number of alleles at a locus or the expected heterozygosity. Use of the same genetic markers in both taxa resulted in increased statistical power because locus-to-locus variation was explicitly included in the model. Differences between the two species in population level measures of diversity were also tested for statistical significance. These comparisons were made for number of alleles, observed and expected heterozygosity, and the inbreeding coefficient. We calculated Spearman's nonparametric correlation to test whether there was a significant correlation among gene diversities across loci in *H. verticillatus* and *H. angustifolius*. The effects of species and repeat motif (i.e. tri- or tetranucleotide repeat) on gene diversity were investigated via nested ANOVA with loci nested within repeat motifs. All proportions were transformed with an angular transformation prior to analysis (Sokal & Rohlf 1995).

Population genetic structure was estimated in an analysis of molecular variance (AMOVA) framework (Weir & Cockerham 1984; Excoffier *et al.* 1992) using ARLEQUIN version 2.0 (Schneider *et al.* 2000). This hierarchical analysis of variance (ANOVA) partitions the total variance into that found within and among populations. The proportion of total diversity that was found among populations was reported as F_{ST} . F_{ST} estimates were analysed in a two-factor ANOVA with species and locus as main effects. Once again, this model allows locus effects to be included as a factor in the model instead of being ascribed to error, thereby increasing the power to detect differences between species' F_{ST} (McCauley *et al.* 1995; Sokal & Rohlf 1995). Principal coordinate analysis (PCO) was conducted on pairwise genetic distances among all three populations of *H. verticillatus* using the covariance standardized method

implemented in the program GENALEX (Peakall & Smouse 2002).

Analysis of genetic admixture

The Bayesian clustering program STRUCTURE (Pritchard *et al.* 2000) was used to test whether *H. verticillatus* represented a genetic mixture of its putative parents, *H. angustifolius* and *H. grosseserratus*, as might be expected if it is a hybrid derivative. We used the admixture model and correlated allele frequencies parameter. In this program, one assumes K populations contribute to the gene pool of the sample population. In this analysis, we set $K = 2$ and the data input consisted of multilocus genotypes from individuals of all three species. For our analysis, the admixture model considers *H. angustifolius* and *H. grosseserratus* as the two populations and determines what proportion of each 'parent' is present in each of the *H. verticillatus* individuals. The putative parents were used as prior population information by employing the USEPOPINFO feature. We assume that a proportion of a proposed hybrid's genotype is drawn from both population one and two (Beaumont *et al.* 2001; James & Abbott 2005). Results are reported as q , the estimated proportion of membership from a given cluster. We used a burn-in period of 50 000 with 10^6 MCMC iterations. GENALEX (Peakall & Smouse 2002) was again used to conduct PCO using the covariance standardized method on pairwise genetic distances among all three species in order to evaluate how distinct they are from one another.

Results

Clonal structure in *Helianthus verticillatus*

Ample levels of genetic diversity were found at the nine nuclear loci used to detect clonal structure in *Helianthus verticillatus* (Table 1 and see below). Our investigation of clonal structure revealed that all stalks from the same cluster yielded identical multilocus genotypes, consistent with the hypothesis that they are members of the same genet. In fact, the probabilities that the same EST-SSR multilocus genotype would be encountered a second time in an obligate outcrosser purely by chance ranged from 9.67×10^{-8} to 4.01×10^{-11} , and the probabilities that the same genotype would be encountered n more times (where n is the number of ramets that we surveyed) ranged from 2.05×10^{-9} to 2.62×10^{-20} . Thus, it is highly unlikely that these genotypes are the result of sexual reproduction. Rather, all stalks from each of the observed clusters most likely represent the same genet, and we never found the same multilocus genotype in disjunct clusters. This result suggests that genets can be identified in the

field based solely on the clustering of stalks and that the genetic population size is much smaller than the number of stalks.

Genetic diversity

In *H. verticillatus*, 18 of 22 EST-SSRs were polymorphic, 13 of 19 were polymorphic in *Helianthus angustifolius*, and 19 of 21 markers were polymorphic in *Helianthus grosseserratus*. Average gene diversities calculated without regard to the population from which samples were drawn for each locus in *H. verticillatus* and *H. angustifolius* are shown in Table 1. Gene diversity ranged from 0 to 0.82 (0.48 ± 0.06 , mean \pm SE) in *H. verticillatus* (based on 22 loci) and from 0 to 0.78 (0.34 ± 0.07) in *H. angustifolius* (based on 19 loci). Gene diversity was significantly positively correlated across shared loci ($r_s = 0.74$; $P = 0.0006$). A nested two-factor ANOVA on average gene diversities with species and repeat motif as main effects, and loci nested within motifs, yielded significant results for all effects (Table 2), and revealed that *H. verticillatus* had significantly higher average gene diversity than *H. angustifolius* ($F_{1,17} = 6.95$, $P = 0.017$). For *H. verticillatus*, the mean number of alleles per polymorphic locus was 7.7 ± 0.96 (6.3 ± 0.83 , all loci), and for *H. angustifolius*, the mean number of alleles per polymorphic locus was 4.9 ± 0.72 (3.3 ± 0.64 , all loci). These differences were significant (one-way ANOVA, polymorphic loci $F_{1,29} = 5.01$, $P = 0.033$; all loci $F_{1,29} = 5.74$, $P = 0.022$). Calculations of number of alleles, observed and expected heterozygosity, and the inbreeding coefficient were also made with regard to populations (Table 3). Mean expected heterozygosity within populations was significantly higher in *H. verticillatus* than in *H. angustifolius* (one-way ANOVA H_E , $F_{1,3} = 30.94$, $P = 0.012$). The other measures of diversity (number of alleles, observed heterozygosity, inbreeding coefficient) did not differ significantly between the two species.

Twelve unique chloroplast haplotypes were found in the three populations of *H. verticillatus*, whereas eight were found in the two populations of *H. angustifolius*. Among all the *H. grosseserratus* individuals, six unique chloroplast haplotypes were found. None of these haplotypes

Table 2 Two-factor nested ANOVA of Nei's gene diversity (1987)

Source	Degrees of freedom	Sum of squares	F ratio	P
Species	1	0.360	6.95	0.017
Repeat motif	1	1.42	27.8	< 0.0001
Locus (repeat motif)	17	3.08	4.42	0.002
Species \times repeat	1	0.079	1.54	0.231
Error	17	0.870		

Species	Population (N)	A	H_O	H_E	f
<i>H. verticillatus</i>	TN (22)	4.3 (0.54)	0.35 (0.05)	0.46 (0.31)	0.23 (0.14)
	GA (22)	4.5 (0.57)	0.40 (0.08)	0.51 (0.31)	0.22 (0.04)
	AL (27)	3.5 (0.43)	0.32 (0.01)	0.46 (0.29)	0.32 (0.10)
<i>H. angustifolius</i>	TN (13)	2.8 (0.45)	0.23 (0.05)	0.34 (0.31)	0.39 (0.08)
	AL (25)	3.6 (0.43)	0.28 (0.04)	0.36 (0.30)	0.14 (0.06)
Mean <i>H. verticillatus</i>		4.1 (0.31)	0.36 (0.02)	0.48 (0.02)*	0.26 (0.03)
Mean <i>H. angustifolius</i>		3.2 (0.40)	0.26 (0.03)	0.35 (0.01)*	0.27 (0.13)

Values are averaged over all loci in each population: A, mean number \pm SE of alleles per locus; H_O , mean observed heterozygosity \pm SE; H_E , mean expected heterozygosity \pm SE; f, within population coefficient of inbreeding \pm SE. *Mean values significantly different from one another, $P < 0.05$; two-factor ANOVA, species and locus effects (see text for details).

Table 4 Chloroplast simple sequence repeat (SSR) diversity in three populations of *Helianthus verticillatus* and two populations of *Helianthus angustifolius*

Species	Population (N)	Haplotypes	H_E
<i>H. verticillatus</i>	TN (22)	4 (A, B, C, D)	0.29
	GA (22)	6 (E, F, G, H, I, J)	0.55
	AL (27)	2 (K, L)	0.05
<i>H. angustifolius</i>	TN (13)	5 (M, N, O, P, Q)	0.31
	AL (25)	3 (R, S, T)	0.35
Mean <i>H. verticillatus</i>		4 (1.2)	0.30 (0.14)
Mean <i>H. angustifolius</i>		4 (1.0)	0.33 (0.02)

Mean \pm SE values in the two species were not significantly different from each other.

were shared between populations or species, and measures of chloroplast genetic diversity did not differ between *H. verticillatus* and *H. angustifolius* (Table 4).

Population structure

The *H. verticillatus* populations were moderately differentiated in terms of both nuclear and chloroplast diversity ($F_{STnuc} = 0.118$, $P < 0.0001$; $F_{STcp} = 0.620$, $P < 0.0001$). The two *H. angustifolius* populations were somewhat more differentiated with $F_{STnuc} = 0.207$ ($P < 0.0001$) and $F_{STcp} = 0.700$ ($P < 0.0001$). These nuclear and chloroplast measures of population differentiation are similar to values reported in other studies of plant populations (Petit *et al.* 2005). The two-factor ANOVA conducted on 12 common, polymorphic loci revealed variation among loci, but not among species for F_{STnuc} (locus effect $F_{11,11} = 2.84$, $P = 0.049$; species effect $F_{1,11} = 0.23$, $P = 0.63$). Pairwise values, all of which were significantly different from zero ($P < 0.0001$), were as follows: Georgia and Alabama ($F_{STnuc} = 0.083$

Table 3 Expressed sequence tag–simple sequence repeat (EST–SSR) diversity in three populations of *Helianthus verticillatus* and two populations of *Helianthus angustifolius*

and $F_{STcp} = 0.589$), Tennessee and Georgia ($F_{STnuc} = 0.146$ and $F_{STcp} = 0.389$), Tennessee and Alabama ($F_{STnuc} = 0.128$ and $F_{STcp} = 0.814$). Subdividing the three patches of the TN population revealed slight but significant differentiation in nuclear markers among these patches ($F_{STnuc} = 0.048$, $P < 0.0323$) and greater differentiation for chloroplast markers ($F_{STcp} = 0.432$, $P < 0.0001$). The PCO carried out on *H. verticillatus* revealed some overlap between individuals from GA and AL, whereas the TN population formed a distinct cluster, separated along PCO1 (PCO 1: 10.0%, PCO 2: 7.2%; Fig. 1a).

Several loci in each of the *H. verticillatus* and *H. angustifolius* populations were found to be significantly out of Hardy–Weinberg equilibrium. Among the three populations of *H. verticillatus* five loci were consistently out of Hardy–Weinberg: BL0001, BL0008, BL0010, BL0018, and BL0027. Loci significantly out of Hardy–Weinberg equilibrium for *H. angustifolius* were BL0018 and BL0025. We used ARLEQUIN to calculate F_{ST} across loci with and without the assumption of Hardy–Weinberg equilibrium. Values of F_{ST} were only slightly different in both cases (i.e. *H. verticillatus* $F_{STnuc} = 0.118$ vs. $F_{STnuc} = 0.113$, respectively), and in no instances did the level of significance change. A test for linkage disequilibrium was not carried out on the data as a rejection of the linkage test could be due to departures from Hardy–Weinberg equilibrium (Excoffier & Slatkin 1998). Furthermore, with the large number of alleles per locus and large number of loci, a likelihood ratio test of linkage disequilibrium may not be valid due to a small number of expected individuals per genotypic class (Schneider *et al.* 2000). However, because the number of loci exceeds the number of chromosomes (17), undoubtedly some markers occur on the same linkage group.

Genetic admixture in *Helianthus verticillatus*

The three species under consideration shared equivalent numbers of nuclear SSR alleles with each other. The mean

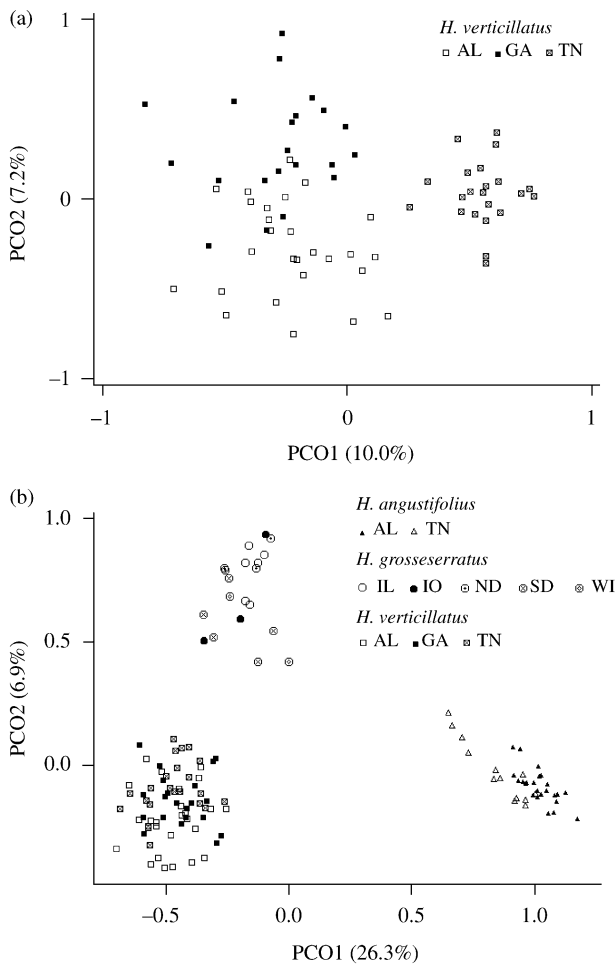


Fig. 1 (a) Principal coordinate analysis (PCO) plot based on 22 loci among three *Helianthus verticillatus* populations. (b) PCO plot based on 18 common loci among *H. verticillatus* (squares), *Helianthus grosseserratus* (circles), and *Helianthus angustifolius* (triangles).

number of shared alleles per locus was 1.79 ± 0.31 between *H. verticillatus* and *H. angustifolius*, 2.09 ± 0.20 between *H. verticillatus* and *H. grosseserratus* and 1.50 ± 0.31 between *H. angustifolius* and *H. grosseserratus*. For one locus, BL0022, *H. verticillatus* and *H. angustifolius* were fixed for different alleles. On the other hand, *H. grosseserratus* was polymorphic for this locus exhibiting four alleles, including the *H. verticillatus* allele. As noted above, each population of the three species possessed unique sets of chloroplast haplotypes.

Genetic admixture analysis indicated that *H. verticillatus* is not a hybrid derivative of *H. angustifolius* and *H. grosseserratus*. The proportion of population membership of each species assigned by STRUCTURE using $K = 2$ was 99.8% of the *H. angustifolius* individuals in one population, and 99.4% and 99.8% of *H. grosseserratus* and *H. verticillatus*

individuals in the other. Therefore, using a model with two groups corresponding to the two putative parents, all *H. verticillatus* individuals were assigned into the cluster with *H. grosseserratus*, suggesting that *H. verticillatus* is more closely related to *H. grosseserratus* than to *H. angustifolius*. It is important to note that, when using $K = 3$, *H. verticillatus* no longer groups with *H. grosseserratus*; rather, it formed a distinct cluster, indicating that *H. verticillatus* is genetically distinct from both *H. grosseserratus* and *H. angustifolius*. A PCO of all individuals from the three species also demonstrated these species are genetically distinct, with no overlap of *H. verticillatus* individuals with either of the putative parental clusters (PCO 1: 26.3%, PCO 2: 6.9%; Fig. 1b).

Discussion

Levels and patterns of genetic diversity

Despite the general expectation of reduced genetic variation in a rare species, *Helianthus verticillatus* does not exhibit a reduction in genetic diversity at either the population or the species level relative to its more common congener, *Helianthus angustifolius*. In fact, for nuclear EST-SSRs, *H. verticillatus* has significantly higher levels of gene diversity than does *H. angustifolius*. While this is not a common result, Gitzendanner & Soltis (2000) demonstrated that endangered species sometimes exhibit levels of diversity as high as, or higher than, a common congener [e.g. in the genera *Agastache* (Vogelmann & Gastony 1987), *Adenophorus* (Ranker 1994) and *Daviesia* (Young & Brown 1996)].

Ellstrand & Elam (1993) proposed that high levels of genetic diversity might be expected in rare species assuming they consist of relatively large populations. While the accurate determination of population sizes can be difficult in clonal species in general, the large number of available EST-SSR markers made the determination of clonal identity straightforward in TN and GA. We genotyped 2–3 stalks from 13 putative genets and found that all stalks within an observed cluster exhibited the same nine-locus genotype indicating that clonal identity can be reliably assessed by eye when the species is found in clusters. The number of distinct clusters in TN is about 70, whereas the GA population contains around 30. Because the clusters are less well-defined in AL, an estimate of the number of genets is more difficult to make; this population is, however, not likely to be large, as there are only a few hundred stalks, and some fraction of these are likely to be ramets of the same genet. Thus, large population size does not seem to be a likely explanation of the relatively high level of diversity present in this species.

An alternative explanation of the relatively high diversity in *H. verticillatus* is that the widespread *H. angustifolius*

exhibits unexpectedly low diversity. However, gene diversity in *H. verticillatus* does not differ significantly from the extremely widespread common sunflower, *Helianthus annuus*. In this latter species, average gene diversity for the same 19 EST-SSRs, based on data from 13 populations (52 individuals in total), was 0.57 ± 0.02 (range 0.45–0.70; C.H. Pashley & J.M. Burke, unpublished), as compared to 0.48 in *H. verticillatus* (two-factor ANOVA with species and locus as effects; $F_{1,18} = 2.65$, $P = 0.12$). While we did not do a formal statistical comparison with *Helianthus grosseserratus* because of differences in the way the samples were obtained, it is worth noting that gene diversity in *H. verticillatus* did not differ markedly from that of *H. grosseserratus* at the 19 loci (0.48 vs. 0.44, respectively). Because we have no information concerning the historical distribution or population size of *H. verticillatus*, we have no way of knowing how long it has been since the species became rare. When combined with the presently small population sizes, the relatively high levels of genetic diversity in *H. verticillatus* may indicate that this species has not been rare for a long time, especially when we consider that it is a clonal perennial (see below).

Another explanation for the relatively high levels of genetic variation exhibited by *H. verticillatus* is that it is of hybrid origin, as higher levels of genetic diversity may result from a mixing of parental alleles. Hybridization often plays a significant role in the evolution and speciation of plants (Arnold 1997; Rieseberg 1997), and its role in the evolution of the annual *Helianthus* species has been studied extensively (e.g. Heiser 1947; Rieseberg 1991; Rieseberg *et al.* 1995, 1996). Some *Helianthus* species freely hybridize in the wild, resulting in hybrid swarms, and three homoploid hybrid sunflower species have been reported (Rieseberg 1991). Based on our SSR data, however, *H. verticillatus* does not appear to be the product of hybridization between *H. angustifolius* and *H. grosseserratus*, as was proposed by Heiser *et al.* (1969). The STRUCTURE analysis with $K = 2$ places *H. angustifolius* into one population, whereas *H. grosseserratus* and *H. verticillatus* correspond to another. This result clearly indicates that *H. verticillatus* does not exhibit mixed ancestry as would be expected in the case of a hybrid swarm, as individuals of hybrid origin would likely consist of a mixture of the genomes of each of its parents. When combined with the findings of Matthews *et al.* (2002), who concluded from morphological evidence that *H. verticillatus* should be considered a distinct species, our data suggest that *H. verticillatus* represents a good taxonomic species of nonhybrid origin. While introgressive hybridization could account for the high levels of heterozygosity observed in this study, our analysis likewise failed to provide any evidence of introgression.

A final possibility is that the unexpectedly high levels of genetic variation in *H. verticillatus* result from the fact that relatively few generations have passed since it became

rare. *H. verticillatus* is a clonal perennial and, because of these life history attributes, populations may have not experienced extensive loss of variation due to the effects of small populations such as genetic drift and inbreeding. While nothing is known of the prior history of the species, F_{STnuc} may give insight into the historical distribution. The number of populations of *H. verticillatus* is very low (only three are known), and one of these is quite disjunct from the other two. Hence, gene flow between TN and GA/AL is probably rare. Despite this, there is only modest genetic differentiation among populations ($F_{STnuc} = 0.118$). It is therefore reasonable that a larger number of populations existed in a more continuous range in the past, but the species has experienced severe reduction in population numbers due to the removal of suitable habitat. In fact, other plant species associated with *H. verticillatus* populations are considered to have strong prairie affinities (e.g. *Hypericum sphaerocarpum* Michaux, *Silphium terebinthinaceum* Jacq., *Andropogon gerardii* Vitman), and fire suppression and conversion of large tracts of land to farmland during European settlement may have significantly reduced the prairie habitat that was once present in this region (Allison 1995). Another possible explanation is that *H. verticillatus* represents a historically narrow endemic, and one of the complexes (TN or GA/AL) has recently been derived from the other. However, we did not see evidence for a recent bottleneck in the form of loss of diversity associated with a founder event. In any case, genetic divergence is not as high as might be expected for disjunct populations with a long history of isolation, especially when compared to estimates of population structure in *H. angustifolius* ($F_{STnuc} = 0.207$).

On the utility of EST-SSRs in evolutionary genetics

This study is unique in that it involves a population genetic survey of an endangered plant species based on a large number of EST-SSRs. Compared to traditional methods of SSR development which are laborious and expensive (Zane *et al.* 2002; Squirrell *et al.* 2003), the transfer of SSRs from a species with an existing EST database to an endangered species is far less time-consuming and costly. In addition, since EST-SSRs are more transferable across taxonomic boundaries than are anonymous SSRs (Varshney *et al.* 2005; CH Pashley & JR Ellis, unpublished data), one can survey two or more taxa with a common set of genetic markers, thereby allowing for the statistical control of locus-specific effects when comparing estimates of genetic diversity and/or population structure. Indeed, if we had conducted a one-way ANOVA (accounting only for the effect of species identity) on genetic diversity in *H. verticillatus* and *H. angustifolius*, the means would not have been statistically different from one another ($F_{1,36} = 1.60$, $P = 0.21$). Thus, the higher statistical power afforded by the use of

common markers across taxa allowed us to detect real differences between these species that would have otherwise gone undocumented.

A point to consider when using EST-SSRs is that selection on these loci could affect population genetic parameters. However, Woodhead *et al.* (2005) found that population differentiation does not seem to be affected by selection as F_{ST} values based on EST-SSRs were similar to those based on anonymous SSRs and AFLPs (amplified fragment length polymorphisms). Given that EST-SSRs appear, on average, to be neutral, they can be used to study the effects of demography on the standing level of genetic variation, a common goal in conservation genetics. Such insights are a prerequisite for understanding the potential influence of rarity and/or fragmentation on adaptive variation, which is likely to play a role in population persistence.

Another possible concern with SSRs in general has to do with the occurrence of null alleles (i.e. alleles that fail to amplify because of mutations in the primer sites flanking the SSR repeat; Callen *et al.* 1993). Null alleles could account, at least in part, for the significant heterozygote deficits that were observed at some loci. Other explanations for the deviations may be selfing or biparental inbreeding due to spatial structuring. Sunflowers typically exhibit sporophytic self-incompatibility with rare selfing seen in the annuals but none encountered in the perennials (C.B. Heiser, personal communication). Furthermore, greenhouse work with all three of these species has resulted in no seed set from selfed flower heads (J.R. Ellis, unpublished). It seems more likely that this pattern results from biparental inbreeding due to spatial structuring and nonrandom mating within populations. Recall that in the Tennessee population, a small but significant amount of micropopulation structure was seen.

An additional finding of this study was that gene diversity in *H. verticillatus* and *H. angustifolius* was dependent upon the repeat motif of the locus in question. In the two-factor nested ANOVA, repeat motif had a significant effect on genetic diversity, with trinucleotide repeats exhibiting higher genetic diversity than tetranucleotide repeats (0.74 vs. 0.30, $F_{1,17} = 27.77$, $P < 0.0001$). This point is particularly important in the context of cross-species comparisons, as the use of different loci in different taxa could easily bias estimates of genetic diversity. This possibility highlights the value of being able to use the same genetic markers across related species as different loci are likely to have quite different evolutionary histories.

Conclusions and conservation implications

It is clear that the expectation of reduced genetic diversity in rare species is not always borne out, and low genetic diversity does not appear to be an immediate concern for

H. verticillatus at this time. Furthermore, populations of *H. verticillatus* exhibited moderate levels of population differentiation using presumably neutral markers. Since these populations are geographically distinct and vary somewhat in ecological conditions, they are likely to be at least as differentiated at adaptive loci, if not more. Therefore, to preserve maximum species diversity, all three populations of *H. verticillatus* should be protected. Habitat loss is probably the cause of rarity in *H. verticillatus* as the species seems to be adapted to prairie habitats which have declined since European settlement (Matthews *et al.* 2002). Thus, habitat protection is of great concern and is probably the most immediate action to take at this time to preserve the species. Finally, the species is a candidate for federal listing on the US Endangered Species Act, and our results have important implications for the species' listing. More specifically, there has been little resolution concerning the issue of whether or not hybrids should warrant legal protection under the Endangered Species Act (Allendorf *et al.* 2001). However, *H. verticillatus* does not appear to be a hybrid between the two proposed taxa; as such the hybrid issue should not inhibit its listing.

Acknowledgements

We thank Temple-Inland Corporation, Sam Breyfogle, Jim Allison, Natasha Sherman, and the Tennessee Natural Heritage Program (especially Claude Bailey and David Lincome) for assistance in locating and collecting samples. We also thank Mark Welch, Mark Chapman, and David Wills for assistance in the laboratory. This work was supported by funds provided by USDA Grant 2002-02213 to D.E.M. This work was also supported by grants from the NSF Plant Genome Research Program (DBI-0332411 to J.M.B.) and the USDA Plant Genome program (#03-39210-13958 to JMB). *Helianthus annuus* EST data was obtained from The Compositae Genome Project database (<http://compgenomics.ucdavis.edu>), which was supported by a grant from the USDA IFAFS program.

References

- Allendorf FW, Leary RF, Spruell P, Wenburg JK (2001) The problem with hybrids: setting conservation guidelines. *Trends in Ecology & Evolution*, **16**, 613–622.
- Allison JR (1995) Prairies ... in Georgia!. *Tipularia*, **10**, 2–8.
- Arnold ML (1997) *Natural Hybridization and Evolution*. Oxford University Press, New York.
- Awise JC, Hamrick JL, eds. (1996) *Conservation Genetics: Case Histories from Nature*. Chapman & Hall, New York.
- Barrett SCH, Kohn JR (1991) Genetic and evolutionary consequences of small population size in plants: implications for conservation. In: *Genetics and Conservation of Rare Plants* (eds Falk DA, Holsinger KE), pp. 3–30. Oxford University Press, New York.
- Baskauf CJ, McCauley DE, Eickmeier WG (1994) Genetic analysis of a rare and a widespread species of *Echinacea* (Asteraceae). *Evolution*, **48**, 180–188.

- Beasley JC (1963) The sunflowers (genus *Helianthus*) in Tennessee. *Journal of the Tennessee Academy of Science*, **38**, 135–154.
- Beaumont M, Barratt EM, Gottelli D *et al.* (2001) Genetic diversity and introgression in the Scottish wildcat. *Molecular Ecology*, **10**, 319–336.
- Bruneau A, Joly S, Starr JR, Drouin J (2005) Molecular markers indicate that the narrow Quebec endemics *Rosa roousseauiorum* and *Rosa williamsii* are synonymous with the widespread *Rosa blanda*. *Canadian Journal of Botany*, **83**, 386–398.
- Bryan GJ, McNicoll J, Ramsay G, Meyer RC (1999) Polymorphic simple sequence repeat markers in chloroplast genomes of Solanaceous plants. *Theoretical and Applied Genetics*, **99**, 859–867.
- Callen DF, Thompson AD, Shen Y *et al.* (1993) Incidence and origin of null alleles in the (AC)_n microsatellite markers. *American Journal of Human Genetics*, **52**, 922–927.
- Chapman MA, Abbott RJ (2005) The origin of a novel form of *Senecio* (Asteraceae) restricted to sand dunes in southern Sicily. *New Phytologist*, **166**, 1051–1062.
- Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small amounts of fresh leaf tissue. *Phytochemical Bulletin*, **19**, 11–15.
- Ellstrand NC, Elam DR (1993) Population genetic consequences of small population size: implications for plant conservation. *Annual Review of Ecology and Systematics*, **24**, 217–242.
- Excoffier L, Slatkin M (1998) Incorporating genotypes of relatives into a test of linkage disequilibrium. *American Journal of Human Genetics*, **62**, 171–180.
- Excoffier L, Smouse P, Quattro J (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*, **131**, 343–359.
- Falk DA, Holsinger KE, eds. (1991) *Genetics and Conservation of Rare Plants*. Oxford University Press, New York.
- Felsenstein J (1985) Phylogenies and the comparative method. *American Naturalist*, **125**, 1–15.
- Gitzendanner MA, Soltis PS (2000) Patterns of genetic variation in rare and widespread plant congeners. *American Journal of Botany*, **87**, 783–792.
- Goudet J (2001) *FSTAT, a Program to Estimate and Test Gene Diversities and Fixation Indices*. Version 2.9.3. Available from <http://www.unil.ch/popgen/softwares/fstat.html>
- Hamrick JL, Godt M (1989) Allozyme diversity in plant species. In: *Plant Population Genetics, Breeding and Genetic Resources* (eds Brown A, Clegg M, Kahler A, Weir B), pp. 43–63. Sinauer Associates, Sunderland, Massachusetts.
- Hedrick PW (2001) Conservation genetics: where are we now? *Trends in Ecology & Evolution*, **16**, 629–636.
- Heiser CB (1947) Hybridization between sunflower species *Helianthus annuus* and *H. petiolaris*. *Evolution*, **1**, 249–262.
- Heiser CB, Smith DM, Clevenger S, Martin WC (1969) The North American sunflowers (*Helianthus*). *Memoirs of the Torrey Botanical Club*, **22**, 1–218.
- James JK, Abbott RJ (2005) Recent, allopatric, homoploid hybrid speciation: the origin of *Senecio squalidus* (Asteraceae) in the British Isles from a hybrid zone on Mount Etna Sicily. *Evolution*, **59**, 2533–2547.
- Karron JD (1987) A comparison of levels of genetic polymorphism and self-compatibility in geographically restricted and widespread congeners. *Evolutionary Ecology*, **1**, 47–58.
- Karron JD (1991) Patterns of genetic variation and breeding systems in rare plant species. In: *Genetics and Conservation of Rare Plants* (eds Falk DA, Holsinger KE), pp. 87–98. Oxford University Press, New York.
- Lande R (1988) Genetics and demography in biological conservation. *Science*, **241**, 1455–1460.
- Lande R (1993) Risks of population extinction from demographic and environmental stochasticity and random catastrophes. *American Naturalist*, **142**, 911–927.
- Lewis PO, Zaykin D (2001) *GDA version 1.0 (D16c): Computer Program for the Analysis of Allelic Data*. Available from: <http://hydrodictyon.eeb.uconn.edu/people/plewis/software.php>.
- Matthews JF, Allison JR, Ware RT Sr, Nordman C (2002) *Helianthus verticillatus* Small (Asteraceae) rediscovered and redescribed. *Castanea*, **67**, 13–24.
- McCauley DE, Raveill J, Antonovics J (1995) Local founding events as determinants of genetic structure in a plant metapopulation. *Heredity*, **75**, 630–636.
- Nei M (1987) *Molecular Evolutionary Genetics*. Columbia University Press, New York.
- Nybohm H (2004) Comparison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants. *Molecular Ecology*, **13**, 1143–1155.
- Parks JC, Werth CR (1993) A study of spatial features of clones in a population of bracken fern, *Pteridium aquilinum* (Dennstaedtiaceae). *American Journal of Botany*, **80**, 537–544.
- Peakall R, Smouse PE (2002) *GENALEX version 5.04: genetic analysis in Excel. Population Genetic Software for Teaching and Research*. Australian National University, Canberra, Australia. Available from: <http://www.anu.edu.au/BoZo/GenALEX/>.
- Petit RJ, Duminil J, Fineschi S *et al.* (2005) Comparative organizations of chloroplast, mitochondrial, and nuclear diversity in plant populations. *Molecular Ecology*, **14**, 689–701.
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.
- Ranker TA (1994) Evolution of high genetic variability in the rare Hawaiian fern *Adenophorus periens* and implications for conservation management. *Biological Conservation*, **70**, 19–24.
- Reed DH, Frankham R (2003) Correlation between fitness and genetic diversity. *Conservation Biology*, **17**, 230–237.
- Rieseberg LH (1991) Homoploid reticulate evolution in *Helianthus* (Asteraceae): evidence from ribosomal genes. *American Journal of Botany*, **78**, 1218–1237.
- Rieseberg LH (1997) Hybrid origins of plant species. *Annual Review of Ecology and Systematics*, **28**, 359–389.
- Rieseberg LH, Ellstrand NC (1993) What can morphological and molecular markers tell us about plant hybridization? *Critical Reviews in Plant Sciences*, **12**, 213–241.
- Rieseberg LH, Sinervo B, Linder CR, Ungerer MC, Arias DM (1996) Role of gene interactions in hybrid speciation: evidence from ancient and experimental hybrids. *Science*, **272**, 741–745.
- Rieseberg LH, Van Fossen C, Desrochers AM (1995) Hybrid speciation accompanied by genomic reorganization in wild sunflowers. *Nature*, **375**, 313–316.
- Rieseberg LH, Wendel J (1993) Introgression and its consequences in plants. In: *Hybrid Zones and the Evolutionary Process* (ed. Harrison RG), pp. 70–109. Oxford University Press, New York.
- Schneider S, Roessli D, Excoffier L (2000) *ARLEQUIN version 2.000: A software for population genetics data analysis*. Genetics and Biometry Laboratory, Department of Anthropology, University of Geneva, Geneva, Switzerland.

- Schuelke M (2000) An economic method for the fluorescent labeling of PCR fragments. *Nature Biotechnology*, **18**, 233–234.
- Seiler GJ, Gulya TJ (2004) Exploration for wild *Helianthus* species in North America: challenges and opportunities in the search for global treasures. In: *Proceedings of the 16th International Sunflower Conference* (ed. Seiler GJ), pp. 43–68. Fargo, North Dakota International Sunflower Association, Paris, France.
- Sokal RR, Rohlf FJ (1995) *Biometry: The Principles and Practice of Statistics in Biological Research*, 3rd edn. WH Freeman, New York.
- Squirrell J, Hollingsworth PM, Woodhead M *et al.* (2003) How much effort is required to isolate nuclear microsatellites from plants? *Molecular Ecology*, **12**, 1339–1348.
- Sydes MA, Peakall R (1998) Extensive clonality in the endangered shrub *Haloragodendron lucasii* (Haloragaceae) revealed by allozymes and RAPDs. *Molecular Ecology*, **7**, 87–93.
- Varshney RK, Sigmund R, Borner A *et al.* (2005) Interspecific transferability and comparative mapping of barley EST-SSR markers in wheat, rye and rice. *Plant Science*, **168**, 195–202.
- Vogelmann JE, Gastony GJ (1987) Electrophoretic enzyme analysis of North American and eastern Asian populations of *Agastache* sect. *Agastache* (Labiatae). *American Journal of Botany*, **74**, 385–393.
- Weir BS, Cockerham CC (1984) Estimating *F*-statistics for the analysis of population structure. *Evolution*, **38**, 1358–1370.
- 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.
- Wills DM, Hester ML, Liu AZ, Burke JM (2005) Chloroplast SSR polymorphisms in the Compositae and the mode of organelle inheritance in *Helianthus annuus*. *Theoretical and Applied Genetics*, **110**, 941–947.
- Woodhead M, Russell J, Squirrell J *et al.* (2005) Comparative analysis of population genetic structure in *Athyrium distentifolium* (Pteridophyta) using AFLPs and SSRs from anonymous and transcribed gene regions. *Molecular Ecology*, **14**, 1681–1695.
- Young AB, Brown AHD (1996) Comparative population genetic structure of the rare woodland shrub *Daviesia suaveolens* and its common congener *D. mimosoides*. *Conservation Biology*, **10**, 1220–1228.
- Young AG, Clarke GM, eds. (2000) *Genetics, Demography and Viability of Fragmented Populations*. Cambridge University Press, Cambridge, UK.
- Zane L, Bargelloni L, Patarnello T (2002) Strategies for microsatellite isolation: a review. *Molecular Ecology*, **11**, 1–16.

JR Ellis is a graduate student interested in the conservation genetics and population biology of *Helianthus verticillatus*. This manuscript will form part of JR Ellis's doctoral thesis at Vanderbilt University. CH Pashley was a postdoctoral researcher developing and utilising EST-SSRs to search for signals of selection during the domestication of *Helianthus annuus*. She is now working as a research associate in the Department of Infection, Immunity and Inflammation at the University of Leicester. JM Burke is an evolutionary geneticist whose research primarily focuses on crop domestication and the evolution of reproductive isolation in plants. DE McCauley is a Professor of Biological Sciences whose research focuses on various aspects of plant population biology.
