

ORIGIN(S) OF THE DIPLOID HYBRID SPECIES *HELIANTHUS DESERTICOLA* (ASTERACEAE)¹

BRIANA L. GROSS,² ANDREA E. SCHWARZBACH,³ AND
LOREN H. RIESEBERG

Department of Biology, Indiana University, Jordan Hall 142, 1001 East Third Street, Bloomington, Indiana 47405 USA

Homoploid hybrid speciation has traditionally been considered a rare event, dependent on the establishment of both a novel, balanced genotype and reproductive isolating barriers between the new species and its progenitors. However, more recent studies have shown that synthetic hybrids converge toward the chromosomal structure of natural hybrids after only a few generations, suggesting that this phenomenon may be more frequent than previously assumed. Here, the possibility that the diploid hybrid species *Helianthus deserticola* arose from more than one hybrid speciation event was investigated using patterns of variation from cpDNA, 18 nuclear microsatellite loci, and population interfertility. *Helianthus deserticola* contains cpDNA haplotypes characteristic of both parental species, is polyphyletic with one parental species based on nine microsatellite loci, and has a high degree of interfertility among populations. The data are consistent with either a single origin followed by introgression with the parental species or multiple origins. Analysis of microsatellite variation places the origin of *H. deserticola* between 170 000 and 63 000 years before present, making it unlikely that anthropogenic disturbances influenced its origin. Finally, the hybrid species generally has lower levels of genetic diversity but higher levels of differentiation among populations than either parental species.

Key words: Asteraceae; *Helianthus*; hybrid speciation; hybridization; parallel speciation; phylogeography; sunflowers.

Hybridization plays an important creative role in plant evolution, contributing to both intraspecific variation through introgression (Anderson, 1948) and to the establishment of new lineages through homoploid or polyploid hybrid speciation (Stebbins, 1950; Grant, 1971). The formation of a sexually reproducing species via either of these two mechanisms is largely dependent on two events: the hybrid lineage must become reproductively isolated from the parents (in parapatry or sympatry), and it must achieve a stable, fertile genotype. When fulfilled through allopolyploidy, these requirements are met by genome duplication in hybrids. Homoploid hybrid speciation is evolutionarily more difficult, but can be achieved through ecological reproductive isolating barriers or chromosomal rearrangements in the hybrid and will result in a new species with the same ploidy as the parents (Smith and Daly, 1959; Grant, 1966a, b, 1971; Gallez and Gottlieb, 1982; Buerkle et al., 2000). Although both modes of speciation are known to have generated stable and reproductively isolated hybrid lineages in nature, more is presently known about the frequency and evolutionary significance of allopolyploidy (reviewed in Otto and Whitton, 2000) than of homoploid hybrid speciation.

Botanists have long recognized the importance of polyploid hybrid speciation in plant evolution (Stebbins, 1950), and recent models demonstrate the ease with which new allopolyploid species can arise and become established (Rodriguez, 1996). Allopolyploidy also represents the most rapid kind of speciation known, since reproductive isolation is an instantane-

ous by-product of genome doubling. In addition to reproductive isolation, genome doubling often is accompanied by a diverse array of morphological, life history, and physiological changes (Levin, 1983; Thompson and Lumaret, 1992). These phenotypic changes sometimes result in modified niche preferences, which increase the likelihood of polyploid establishment (Rodriguez, 1996). Allopolyploid species are often polyphyletic, with individual populations arising from independent hybridization events between different populations of the same two parental species (Soltis et al., 1995). The mechanism by which allopolyploid species arise is conducive to the formation of polyphyletic species; each F1 hybrid contains the same initial chromosomal complement consisting of one part from each parent, and chromosomal doubling will establish this genome permanently in all polyploids formed.

In contrast to the frequent documentation of allopolyploid speciation in nature, only a handful of diploid hybrid species have been thoroughly investigated and verified using molecular markers (reviewed in Rieseberg, 1997). The creation of a homoploid hybrid species from an F1 hybrid has traditionally been considered an unusual event (Grant, 1971). This view derives from simple models in which the origin of a fertile hybrid segregant was achieved via the stochastic sorting of chromosomal and genic sterility factors that differentiated the parental species. Even with modest levels of differentiation between the parental species, the probability of generating a fertile hybrid segregant is low. Inbreeding increases the likelihood of achieving a fertile, stable genome, so early workers predicted that homoploid hybrid speciation would be most likely in inbreeding systems (Stebbins, 1957). Given the stochastic nature of the sorting process, it was anticipated that each hybrid speciation event resulting from hybridization between the same two parental species would produce a different genome and that the newly formed microspecies would be perpetuated through inbreeding (Stebbins, 1957; Grant, 1958).

The prediction that multiple homoploid hybrid speciation events involving the same parental species can result in distinct hybrid lineages is verified in the genus *Helianthus*, where

¹ Manuscript received 1 April 2003; revision accepted 11 July 2003.

The authors thank Eva Allen, Jennifer Durphy, Keith Gardner, and Mark Welch for assistance with various aspects of the data collections and analyses. We also thank all members of the Rieseberg laboratory for helpful discussions and comments on the manuscript. Special thanks to the Willamette University faculty and greenhouse staff for their time and space. This work was supported by the NSF REU, IGERT, and Predoctoral Fellowships to B. L. G., and NIH grant GM59065 to L. H. R.

² E-mail: brgross@indiana.edu.

³ Present address: Department of Biological Sciences, Kent State University, Kent, Ohio 44242 USA.

the same parents (*H. annuus* and *H. petiolaris*) are the progenitors of three different diploid hybrid species: *H. anomalus*, *H. paradoxus*, and *H. deserticola* (Rieseberg et al., 1990, 1991; Rieseberg, 1991). However, more recent theory has emphasized the important role that selection for fertility and viability plays in establishing stable combinations of sterility factors (Templeton, 1981; McCarthy et al., 1995; Buerkle et al., 2000). Experiments by Rieseberg and colleagues (Rieseberg et al., 1996; Rieseberg, 2000) have shown that experimental hybrids converged toward chromosomal combinations similar to those found in natural hybrid species when selection was made for fertility alone, suggesting a more prominent role for selection than chance in the establishment of a new hybrid lineage. This indicates that multiple origins for a single homoploid hybrid species in the wild is a possibility; evidence consistent with multiple origins of a single diploid hybrid species has now been documented in several systems (Brochmann et al., 2000; Wang et al., 2001; Schwarzbach and Rieseberg, 2002). Recent models and demonstrated cases of homoploid hybrid speciation further suggest that this mode of speciation may occur frequently in outcrossing species rather than being restricted to inbreeding taxa (Gallez and Gottlieb, 1982; Arnold et al., 1990; Wang et al., 1990; Wang and Szmidi, 1994; McCarthy et al., 1995; Sang et al., 1995; Allan et al., 1997; Rieseberg, 1997; Wolfe et al., 1998a, b; Buerkle et al., 2000).

This study was designed to evaluate the possibility that the diploid hybrid species, *H. deserticola*, has arisen multiple times in the wild using evidence from variation in chloroplast DNA (cpDNA), nuclear microsatellite loci, and interpopulation crossability. *Helianthus deserticola* is a xerophytic species found in sandy soils and distributed in small populations located in western Nevada, west central Utah, and along the border of Utah and Arizona, USA (Fig. 1). It is a confirmed diploid hybrid species based on comparison of isozyme, nuclear ribosomal DNA, and cpDNA with its parental species, *H. annuus* and *H. petiolaris* (Rieseberg, 1991). *Helianthus annuus* is distributed throughout the central and western United States and typically inhabits heavy, clay-based soils. *Helianthus petiolaris*, the smaller of the two parental species, is distributed mainly through the central United States and inhabits sandier soils than *H. annuus*. The two parental species co-occur and often hybridize throughout their range. The species are all annual, outcrossing, and have a haploid chromosome number of 17 (Heiser, 1947; Heiser et al., 1969; Rogers et al., 1982).

Variation in chloroplast DNA haplotypes and nuclear microsatellite loci were surveyed for both parental species and populations of *H. deserticola* throughout its geographic range. The genetic composition of *H. deserticola* was expected to show one of two possible patterns. The first was uniformity in cpDNA haplotype throughout its geographical range and reciprocal monophyly with the parental species based on microsatellites. This would be characteristic of a single origin for the species, followed by range expansion or widespread dispersal. The second possibility was that the cpDNA haplotypes for *H. deserticola* would be diverse and that the species would be polyphyletic with respect to the parental species. This latter pattern has two possible explanations: one is that *H. deserticola* resulted from a single homoploid hybrid speciation event followed by range expansion and subsequent introgression with different parental populations; the other is that the species arose from multiple homoploid hybrid speciation events. It is also possible that patterns of genetic variation indicative of

multiple origins might be erased by the homogenizing effects of gene flow between populations of *H. deserticola*.

Finally, interpopulation crosses were made between six of the eight *H. deserticola* study populations, and pollen viability of the progeny was used to verify interfertility among the geographically separated populations. Uniform patterns of interfertility between geographically disparate populations would be a sign of a single origin and a common genetic composition across the species. Discontinuities in patterns of pollen viability would suggest marked genetic differentiation, attributable to extreme population subdivision and isolation or to intrinsic differences in genetic composition due to multiple origins.

MATERIALS AND METHODS

Collections and DNA isolations—Leaves and/or achenes were collected from 12 natural populations of *H. annuus*, six of *H. petiolaris*, and eight of *H. deserticola* in Arizona, Utah, and Nevada, the three states from which *H. deserticola* has been reported (Table 1, Fig. 1). Total genomic DNA was isolated for 149 *H. annuus* individuals, 61 *H. petiolaris* individuals, and 97 *H. deserticola* individuals from frozen and silica-dried field material, achenes, or greenhouse-grown material. Isolations were made using the DNeasy Plant Mini Kit (QIAGEN, Valencia, California, USA), and DNA concentrations were quantified with a fluorometer (Hoefer Scientific Instruments, San Francisco, California, USA).

Chloroplast DNA polymerase chain reaction-restriction fragment length polymorphisms—The chloroplast DNA analysis represents an extension of a previous study of another diploid hybrid species (*H. anomalus*) and its parents, *H. annuus* and *H. petiolaris* (Schwarzbach and Rieseberg, 2002). Because both *H. anomalus* and *H. deserticola* have the same parental species and occur in the desert Southwest, many of the same parental species populations were informative for both studies. Thus, haplotype information has been previously reported for all six *H. petiolaris* populations and for six of the 12 *H. annuus* populations in Table 1 (Schwarzbach and Rieseberg, 2002). Chloroplast DNA PCR amplification and restriction fragment length polymorphism (cpDNA PCR-RFLP) analysis of the remaining six populations of *H. annuus* and of all eight populations of *H. deserticola* are reported here for the first time.

As described previously, primers *rpoC1*—195—*rpoC2*—1364 (Liston, 1992), *trnC*—*trnD* (Demesure et al., 1995), *trnF*—*trnV*, and *trnV*—*rbcL* (Dumolin-Lapegue et al., 1997) were employed to amplify cpDNA fragments for all individuals (Schwarzbach and Rieseberg, 2002; Welch and Rieseberg, 2002). All amplify large regions of cpDNA containing considerable variation (Schwarzbach and Rieseberg, 2002; Welch and Rieseberg, 2002). The PCR reactions were performed in volumes of 50 μ L containing 1 μ L of template DNA, 0.08 μ g of each primer, and 1 unit *Taq* DNA polymerase, at a final concentration of 30 mmol/L Tricine, 50 mmol/L KCl, 2 mmol/L MgCl₂, and 100 μ mol/L each dNTP. Reactions were performed in an MJ Research (Watertown, Massachusetts, USA) PTC-100 Thermal Cycler programmed for an initial denaturation step of 1 min at 94°C, 37 cycles of 45 s at 94°C, 45 s at 55°C, and 3 min at 72°C, with a final extension at 72°C for 7 min.

The four PCR products were digested with restriction endonucleases known to yield distinct restriction profiles for the parental species in previous studies (*Hha*I, *Acc*I, *Dde*I, *Bst*UI, *Hae*III, *Mse*I, *Taq*I) (Table 2). Digestions were performed in volumes of 15 or 20 μ L, containing 5 or 10 μ L of PCR product, respectively. Each digestion contained 2.5 units restriction enzyme, 1.5 μ L buffer (New England Biolabs, Beverly, Massachusetts, USA), 0.15 μ L of bovine serum albumin (BSA) if required and distilled water. After mixing, samples were incubated for a minimum of 3 h at either 37° or 60°C, as recommended by the supplier. Digests were run on 1.5% agarose gels, stained with ethidium bromide, and visualized under UV-light. Restriction profiles were scored as either 1 or 0 (Table 3), depending on the presence or absence of a restriction site. Restriction digests that yielded more than one site were scored separately for each site. A most parsimonious network was manually

TABLE 1. Population codes, locality, sample size, and cpDNA haplotypes. Sample size is shown for the cpDNA/microsatellite analysis. Parentheses denote numbers of individuals with each cpDNA haplotype per population.

Taxon/Population	Location ^a	N	cpDNA haplotypes	Code
<i>Helianthus annuus</i>				
Schwarzbach & Welch 1/99	Hwy 6, 1.6 km S of Price, Carbon Co., UT	12/14	1(12)	ANN1/99
Rieseberg 1281	Hwy 160, between Mm 418 and 419, Navajo Co., AZ	12/10	13(12)	ANN1281
Rieseberg 1286	Jct I-15, Hwy 6 exit 261, E side, Utah Co., UT	12/12	1(12)	ANN1286
Rieseberg 1295	0.40 km N of N city limits of Hanksville, Wayne Co., UT	12/15	1(12)	ANN1295
Schwarzbach & Welch 26/99	Hwy 125, 1.6 km W of jct. Hwy 125 and 50, Millard Co., UT	12/14	1(12)	ANN26/99
Rieseberg 1308	Mm 33 on Hwy 89, 51.5 km east of Kanab, Kane Co., UT	11/12	19(7), 10(4)	ANN3
Rieseberg 1309	15.3 km E of Zion National Park on State Hwy 9, Kane Co., UT	12/12	1(4), 10(3), 13(5)	ANN4
Schwarzbach & Welch 4/99	Hwy 95 between Natural Bridge National Monument and Blanding, San Juan Co., UT	12/12	1(12)	ANN4/99
Rieseberg 1310	4.8 km S of Toquerville exit of I-15 on Old Hwy 91, Washington Co., UT	12/12	1(5), 18(3), 19(1), 13(1), 20(2)	ANN5
Rieseberg 1311	2.4–3.2 km SW of Riverside Bridge, exit 112 off I-15, Clark Co., NE	11/12	6(11)	ANN6
Schwarzbach & Welch 9/99	Hwy 191, just S of Rock Point, Apache Co., AZ	12/12	1(12)	ANN9/99
Rieseberg Reno	48.3 km E of Reno, Washoe Co., NE	12/12	17(12)	ANNReno
<i>H. deserticola</i>				
Rieseberg 1236	7.6 km SW of Riverside Bridge, Clark Co., NE	11/11	13(11)	DES1236
Rieseberg 1265	Jericho Campground, Little Sahara Recreation Area, Juab Co., UT	13/13	1(13)	DES1265
Rieseberg 1270	6.4 km S of Toquerville exit off I-15 on Old Hwy 91, Washington Co., UT	15/15	13(10), 16(5)	DES1270
Rieseberg 1274	N edge of Glen Canyon National Recreation Area, Hwy 89, Kane Co., UT	10/10	13(5), 8(5)	DES1274
Rieseberg 1275	Horse corral, Snow Canyon State Park, Washington Co., UT	12/12	13(11)	DES1275
Rieseberg 1296	Jct. of road to Lone Rock and Hwy 89, Coconino Co., AZ	10/10	13(1), 8(9)	DES1296
Rieseberg DesB	S of road 4.3 km E-NE of Lee Hot Springs, Churchill Co., NE	14/14	13(14)	DESB
Rieseberg DesC	7.7 km N of Hwy 50 on Soda Lake Road, Churchill Co., NE	12/12	13(12)	DESC
<i>H. petiolaris</i>				
Rieseberg 1271	0.40 km W of exit 95, jct. I-5, Hwy 20, Iron Co., UT	12/10	1(1), 14(5), 10(6)	PET1271
Rieseberg 1277	16.1 km S of Page on Hwy 89, Coconino Co., AZ	12/9	8(12)	PET1277
Rieseberg 1279	20.0 km E of jct. Hwy 6 and 264, Hwy 264, Navajo Co., AZ	11/10	13(9), 7(2)	PET1279
Rieseberg 1283	8.5 km E of jct. Hwy 160 and 264, Hwy 264, Coconino Co., AZ	12/9	13(5), 8(5), 9(2)	PET1283
Rieseberg 1285	Kodachrome Basin, Chimney Rock, Kane Co., UT	12/11	12(8), 10(1), 13(3)	PET1285
Rieseberg 1287	0.40 km E of Zion National Park entrance on Hwy 9, Kane Co., UT	12/12	10(7), 6(4), 5(1)	PET1287

^a Abbreviations: AZ, Arizona; Co., County; E, east; Hwy, highway; I, interstate highway; Jct., junction; Mm, mile marker; N, north; NE, Nebraska; S, south; UT, Utah; W, west.

constructed based on restriction site polymorphisms and verified with PAUP* 4.0b8 (Swofford, 2002).

Microsatellites—Microsatellite markers were used to determine the relatedness of *H. deserticola* to populations of its parental species. Primer pairs developed for *H. annuus* (Tang et al., 2002) were screened for clean amplification and codominance in the taxa included in the study, and the following 18 loci were chosen for surveys of all populations and species: *ORS3**, *ORS4**, *ORS5**, *ORS7**, *ORS8**, *ORS10**, *ORS12**, *ORS59**, *ORS297*, *ORS299*, *ORS377*, *ORS437*, *ORS442*, *ORS484*, *ORS541*, *ORS613*, *ORS618*, *ORS733*. Primer sequences can be found at <http://www.css.orst.edu/knapp-lab/sunflower>. Also, note that the eight loci marked with an asterisk were previously assayed in six populations of each parental species (see Table 1) by Schwarzbach and Rieseberg (2002). Primers were synthesized, and the 5' end

of each forward primer was labeled with one of three fluorescent dyes (6FAM, HEX, or NED) by MWG Biotech (High Point, North Carolina, USA), Invitrogen (Grand Island, New York, USA), or PE Biosystems (Foster City, California, USA).

The PCR reactions were performed in volumes of 10 μ L containing 20 ng DNA, 1 unit *Taq* DNA polymerase, and 0.0016 μ g of each primer at a final concentration of 30 mmol/L Tricine, 50 mmol/L KCl, 2 mmol/L MgCl₂, and 100 μ mol/L each dNTP. Fragments were amplified using a "touchdown" PCR protocol, developed to reduce nonspecific primer binding and fragment amplification (Don et al., 1991). An initial denaturing cycle of 3 min at 95°C was followed by 10 touchdown cycles (the program starts 10°C above the appropriate annealing temperature, and the temperature drops 1°C each cycle) of 30 s at 94°C, 30 s at the annealing temperature, and 45 s at 72°C. These 10 cycles were followed by 29 cycles of 30 s at 94°C, 30 s at the appropriate

TABLE 2. Description of cpDNA restriction fragment length polymorphisms for 26 populations of *Helianthus*. Fragment sizes listed in brackets were not observed, but inferred from the length of the complementary fragment.

Code	cpDNA fragment	Restriction enzyme	Digestion fragment sizes
A	trnC-D	<i>TaqI</i>	1600 → 1500 + [100]
B	trnF-trnV	<i>BstUI</i>	2550 → 2000 + 550
C	trnF-trnV	<i>HhaI</i>	680 → 500 + 180
D	trnF-trnV	<i>DdeI</i>	350 → 240 + 110
E	trnF-trnV	<i>AcII</i>	790 → 600 + 190 (site #1)
F	trnF-trnV	<i>AcII</i>	600 → 450 + 150 (site #2)
G	trnV-rbcL	<i>AcII</i>	850 → 600 + [250]
H	trnV-rbcL	<i>BstUI</i>	1150 → 1000 + 150 (site #1)
I	trnV-rbcL	<i>BstUI</i>	1150 → 960 + 190 (site #2)
J	trnV-rbcL	<i>DdeI</i>	1100 → 800 + 300
K	trnV-rbcL	<i>HaeIII</i>	2140 → 1800 + 340
L	trnV-rbcL	<i>MseI</i>	540 → 330 + 210
M	rpoC	<i>HhaI</i>	660 → 530 + [130] (site #1)
N	rpoC	<i>HhaI</i>	620 → 580 + [40] (site #2)
O	rpoC	<i>HhaI</i>	1950 → 1050 + 900 (site #3)

annealing temperature, and 45 s at 72°C, with a final elongation period at 72°C for 20 min. Annealing temperatures are as follows: 51°C for *ORS3* and *ORS377*, 52°C for *ORS7*, *ORS8*, *ORS10*, and *ORS12*, and 55°C for *ORS4*, *ORS5*, *ORS59*, *ORS297*, *ORS299*, *ORS437*, *ORS442*, *ORS484*, *ORS541*, *ORS613*, *ORS618*, and *ORS733*.

Microsatellite fragments were assayed via electrophoresis on an ABI 3700 capillary sequencer (Applied Biosystems, Foster City, California, USA). The PCR fragments of non-overlapping size and color were pooled and diluted 1 : 20 with ddH₂O. Finally, 1 μL of the diluted PCR product was added to a 10 μL mixture of 9.82 μL ddH₂O and 0.18 μL of the GenSize R500 ROX size standard (GenPak, St. James, New York, USA). Samples were denatured at 95°C and snap cooled on ice before loading onto the 3700. Chromatographs of sequencer data were generated using GENESCAN 3.5 (Applied Biosystems, Foster City, California, USA) and fragment lengths were scored using GENOTYPER 3.6 (Applied Biosystems, Foster City, California, USA).

Linkage disequilibrium between all pairs of microsatellite loci and number of migrants per generation (*N_m*) were calculated using GENEPOP (Raymond and Rousset, 1995; Hendrie et al., 1998). Estimates of *N_m* are based on the private alleles method (Barton and Slatkin, 1986). *F* statistics were calculated using ARLEQUIN 2.0 (Schneider et al., 2000), assuming an infinite alleles model according to the methods of Weir (1996). Average gene diversity (*H_e*) was also calculated using ARLEQUIN 2.0. A Mantel test was used to check for correlations between geographic distance and *F_{ST}*, as implemented by the

R PACKAGE 4.0 (Casgrain and Legendre, 2001). Significance was tested using 999 permutations of the data.

Phylogenetic analysis of microsatellite frequency data was performed using PHYLIP 3.6 (Felsenstein, 1993). Neighbor-joining trees (Saitou and Nei, 1987) were constructed based on Nei's genetic distances (Nei, 1987). Support for nodes was evaluated based on 1000 bootstrap replicates. Hybrid lineages, by definition, contain a mixture of parental loci, and it is acknowledged that reticulate evolution presents a special challenge to phylogeny reconstruction (McDade, 1990, 1992; Rieseberg and Ellstrand, 1993). It has also been demonstrated that independent hybrid lineages may become fixed for the same parental loci (Rieseberg et al., 1996). Thus, populations from two independently derived hybrid lineages may appear more similar to each other than to any populations of the parental species. One way to alleviate this problem is to construct trees based only on loci derived from one parental species (Schwarzbach and Rieseberg, 2002). Loci in *H. deserticola* were assigned to one parent or another using a maximum likelihood approach (Rieseberg et al., 1998). The value produced by this analysis is essentially a hybrid index based on allele frequencies; a hybrid index score was generated at every locus for the species as a whole and for each population individually. Two separate neighbor-joining trees were then constructed using the loci that could be assigned to a particular parent and containing populations of only that parental species and the hybrid species. The allele frequencies of the other parental species were pooled and used as an outgroup.

The divergence times among populations of *H. deserticola* were calculated as a proxy for time of origin of the hybrid species. Divergence times were calculated based on microsatellite allele frequency according to the methods of Zhivotovsky (2001), using Equation 1. Here, $T_D = D_1/2w - V_0/w$; T_D is the divergence time in generations, D_1 is the average over all loci of the squared difference in repeat numbers for pairs of alleles drawn from different populations (from equation 2 of Goldstein et al., 1995), w is the effective mutation rate, and V_0 is the average over all loci of the variance in repeat number in the ancestral population. V_0 , a property of the ancestral population, is unknown, so we followed Zhivotovsky's recommendation to set V_{0min} equal to 0, and V_{0max} equal to the average within-population variance of the two populations being compared. This allowed us to generate a lower and upper bound for the divergence times (T_{Dmin} and T_{Dmax} , respectively). The mutation rate, w , was set at 2×10^{-4} based on estimated rates of mutation for 20 microsatellites in soybean populations of known pedigree (Diwan and Cregan, 1997).

Crossing studies—Genetic relatedness of six of the eight *H. deserticola* study populations were evaluated using interpopulational crosses and pollen viability counts; no seeds were available from populations DESB and DESC, so they were excluded from the analysis. A minimum of 10 plants from each population was propagated from seed in an Indiana University greenhouse.

TABLE 3. Chloroplast DNA haplotypes characterized by 15 different restriction sites for 26 populations of *Helianthus*. Restriction sites are represented by letters A–O; see Table 2 for description of fragment lengths.

Haplotype number	Code														
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	0	0	0	1	1	0	0	0	0	0	0	1	1	0	0
5	0	0	0	1	1	0	1	1	0	0	1	0	0	0	0
6	0	0	0	1	1	0	1	1	0	1	1	0	0	0	0
7	0	0	0	1	1	0	1	0	0	1	1	0	0	0	0
8	0	0	0	1	0	0	1	0	0	1	1	0	0	0	0
9	0	0	0	1	1	1	1	0	0	1	1	0	0	0	0
10	0	0	0	0	1	1	1	0	0	1	1	0	0	0	0
12	0	1	1	1	1	1	1	0	0	1	1	0	0	0	0
13	1	0	0	1	1	1	1	0	0	1	1	0	0	1	0
14	1	0	0	1	1	1	1	0	1	1	1	0	0	1	0
16	0	0	0	1	1	1	1	0	0	1	1	0	0	1	1
17	0	0	0	1	1	0	0	0	0	0	0	1	0	1	0
18	0	0	0	1	1	0	1	0	0	0	0	1	1	0	0
19	1	0	0	1	1	1	1	0	0	1	1	0	0	0	0
20	1	0	0	1	1	1	0	0	0	1	1	0	0	1	0

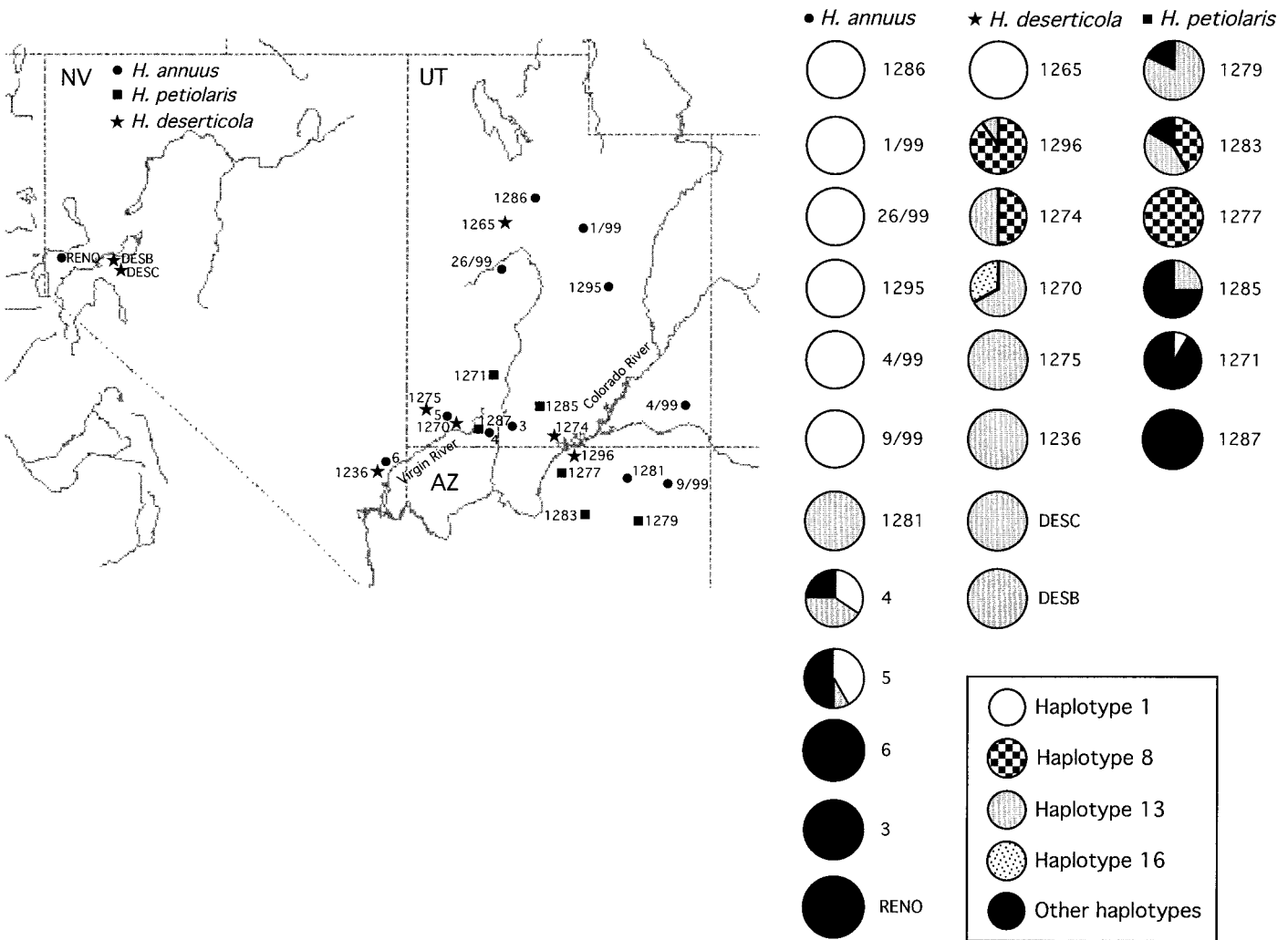


Fig. 1. Range map showing locations of *Helianthus* populations from Utah, Arizona, and Nevada, USA, used in this study. Haplotype composition of each population is shown to the right of the map. For simplicity, only the four haplotypes found in *H. deserticola* are depicted, and the other 11 haplotypes are represented in black.

Reciprocal crosses were made between the six populations and among individuals within each population, resulting in 36 sets of seeds representing every inter- and intrapopulation parental combination.

Seeds were germinated at the Willamette University greenhouses, and pollen viabilities of progeny resulting from the crosses were estimated by staining with a solution of 30% sucrose and 0.1% MTT (Chandler et al., 1986). Pollen viability was scored for two plants resulting from each parental combination, and ~500 pollen grains were scored per flowering head. Means and standard errors were calculated for pollen viabilities for each cross combination. Correlations between pollen viability and either geographic distance or F_{ST} were tested using the Mantel test, implemented by the R PACKAGE 4.0 (Casgrain and Legendre, 2001). Significance was tested using 999 permutations of the data.

RESULTS

Chloroplast DNA—In total, 15 different restriction sites were scored for each individual and 15 cpDNA haplotypes were discernable from the restriction digests (Tables 2 and 3). Ten of the haplotypes were previously reported by Schwarz-

bach and Rieseberg (2002), and the same numerical designations are employed in both studies. Haplotypes unique to this study were assigned new numbers. Haplotypes varied considerably within and among populations of all three species, with eight haplotypes in *H. annuus*, (1, 6, 10, 13, 17, 18, 19, and 20), 10 in *H. petiolaris* (1, 5, 6, 7, 8, 9, 10, 12, 13, and 14), and four in *H. deserticola*: 1, 8, 13, and 16 (Table 1; Figs. 1 and 2). Of the four haplotypes present in *H. deserticola*, one was characteristic of *H. petiolaris* (8), one was characteristic for *H. annuus* (1), one was found in equal frequencies in both parental species (13), and one was unique to *H. deserticola* (16), but was more similar to *H. petiolaris* types than *H. annuus* (Fig. 2). All three species were occasionally polymorphic at the population level, with up to five different cpDNA haplotypes occurring in a single population (Table 1).

The diversity of *H. annuus* haplotypes was distributed across the geographic range of the species, although the northern and eastern populations tended to be fixed for the common *H. annuus* haplotype (haplotype 1, populations ANN1286,

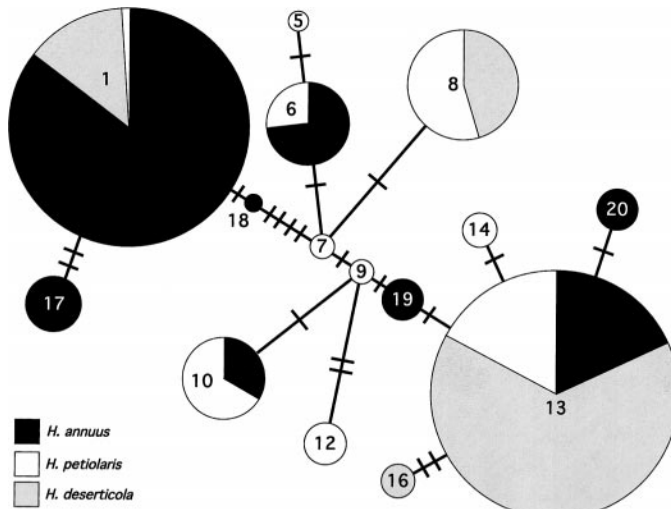


Fig. 2. Most parsimonious network of cpDNA haplotypes of *Helianthus*, with haplotype numbers given within each circle. Sizes of the circle roughly correspond to the number of individuals with that haplotype, and the three species are represented by shades within each circle. Mutational steps are indicated by bars between haplotypes.

ANN1/99, ANN26/99, ANN1295, ANN4/99, and ANN9/99). In *H. petiolaris*, haplotypes 7, 8, 9, 12, and 13 occurred only in the four populations farthest east along the Utah–Arizona border, all located in the Colorado River drainage basin (PET1285, PET1283, PET1279, and PET1277). Types 1, 5, 6, and 14 were localized to two populations in the Virgin River drainage basin of Utah’s southwest corner (PET1287 and PET1271). Type 10 occurred in populations of *H. petiolaris* in both of these areas, but was found predominantly in the two Virgin River populations and represented by only a single individual in the Colorado River drainage basin.

Of the four different haplotypes present in *H. deserticola*, haplotype 13 occurred in seven of the eight populations surveyed. The two *H. deserticola* populations farthest east along the Utah–Arizona border in the Colorado River drainage basin also contained some individuals with haplotype 8 (DES1274 and DES1296). Type 16, unique to *H. deserticola*, occurred in only one population near the southwest corner of Utah and made up part of a population that was predominantly type 13 (DES1270). The single *H. deserticola* population with haplotype 1 (the typically *H. annuus* haplotype) occurred in the Little Sahara National Recreation Area located in west central Utah (DES1265).

Microsatellites—Analysis of linkage disequilibrium revealed that none of the 18 loci were significantly linked to any others. All loci were included in subsequent analyses. According to estimates of F_{ST} , *H. deserticola* showed the most extreme partitioning of genetic variation among populations (F_{ST} of 0.249), followed by *H. petiolaris* and *H. annuus* (F_{ST} of 0.194 and 0.159, respectively). In all cases, the majority of the genetic variation exists within populations, as compared to among populations or species. There was no correlation, however, between geographic and genetic distances within *H. deserticola*. N_m values are estimated as 0.70 for *H. deserticola*, 1.0 for *H. annuus*, and 0.73 for *H. petiolaris*. Average and median variation in microsatellite fragment size, average and median number of alleles per microsatellite locus, and average

gene diversity (H_e) were calculated to assess genetic diversity of the three species (Table 4). *Helianthus deserticola* was generally the least variable according to all estimates, the only exception being that the mean allele size variance was slightly higher for *H. deserticola* than for *H. annuus*, but still much lower than that for *H. petiolaris*.

Overall, eight of the loci present in *H. deserticola* are derived from *H. annuus*, while nine are derived from *H. petiolaris* according to the hybrid index analysis. One locus (ORS618) could not be attributed to either parental species, as none of the alleles present in the hybrid species were present in the parental species. Within 12 of the 17 loci amenable to the hybrid index analysis, individual populations of *H. deserticola* revealed different patterns of parental origin for the locus in question as compared to the overall assignment (Table 5). That is, the origin of a particular locus might be assigned to *H. annuus* when *H. deserticola* is considered as a single population, but parental assignments might vary for individual populations.

Helianthus deserticola forms a single, well-supported clade (86.9%) when an unrooted neighbor-joining tree is constructed based on all 18 microsatellite loci (Fig. 3). In a tree constructed using the eight loci largely inherited from *H. annuus*, *H. deserticola* is polyphyletic (Fig. 4). Four of the populations of *H. deserticola* were more closely related to each other than to any populations of *H. annuus*, although the clade had very weak support. Two populations occurred in basal positions, and one of these appeared to be more closely related to the *H. petiolaris* outgroup than to any other populations. The two remaining populations of *H. deserticola* (DESB and DESC) grouped most closely with two populations of *H. annuus* (ANNReno and ANN5). Although this clade had only weak support, it is of interest in that ANNReno is the population of *H. annuus* most geographically proximal to these populations of *H. deserticola*. *Helianthus deserticola* is monophyletic when a tree is constructed based on the nine microsatellite loci largely derived from *H. petiolaris*, although the clade again has weak support (40%) (Fig. 5). Note that incongruence between the two trees was not unexpected, because both introgression and multiple origins can have different effects on unlinked loci.

Divergence times among populations of *H. deserticola* were calculated independently for each locus, then averaged across loci. Comparisons were made among populations of the hybrid species rather than between the hybrid species and its progenitors because the latter comparison might be inflated by differences that had accumulated between the parental species. The average maximum age of the species (that is, when V_0 was set to 0) is $170\,000 \pm 12\,000$ generations. The average minimum age of the species (that is, when V_0 was considered to be the average within-population variance of the two populations being compared) is $63\,000 \pm 11\,000$ generations. Because all species involved in the study are annual, these values can be translated roughly into years, giving a date of origin between 170 000 and 63 000 years before present. The divergence times are not substantially affected when the calculation is made using only the five geographically proximal population in southern Utah/northern Arizona (DES1236, DES1270, DES1275, DES1274, and DES1296), which all appear to have originated from the same hybridization event. In this case, the maximum age of the species is $190\,000 \pm 21\,000$ years before present, and the minimum age of the species is $78\,000 \pm$

TABLE 4. Estimates of genetic diversity within populations and species, including variance in microsatellite allele size, number of alleles per locus, and the average gene diversity (H_e). Diversity rank reflects average gene diversity (H_e).

Code	Mean size variance	Median size variance	Mean no. alleles per locus	Median no. alleles per locus	Gene diversity H_e (SE)	Diversity rank
ANN1/99	88.0	38.1	6.1	6.0	0.61 (0.32)	11
ANN1281	83.2	18.3	3.6	4.0	0.45 (0.24)	21
ANN1286	52.1	32.5	5.1	5.0	0.64 (0.33)	7
ANN1295	41.6	37.1	5.4	5.5	0.53 (0.28)	15
ANN26/99	46.4	18.2	4.5	4.0	0.40 (0.22)	22
ANN3	103.8	19.5	4.8	4.0	0.66 (0.34)	5
ANN4	65.4	34.1	5.9	6.5	0.67 (0.35)	4
ANN4/99	65.2	12.0	3.8	4.0	0.34 (0.19)	25
ANN5	66.5	16.5	6.7	6.0	0.72 (0.38)	1
ANN6	72.1	25.9	6.1	6.0	0.69 (0.36)	2
ANN9/99	49.5	27.9	4.7	5.0	0.49 (0.26)	17
ANNReno	39.6	26.8	4.6	4.5	0.62 (0.32)	10
Total	64.4	24.3	5.1	5.0	0.67 (0.33)	
DES1236	82.3	13.0	3.3	3.0	0.48 (0.26)	19
DES1265	129.4	17.1	4.4	4.5	0.57 (0.30)	13
DES1270	40.2	5.3	4.1	4.0	0.49 (0.26)	18
DES1274	103.1	11.5	5.6	5.0	0.64 (0.34)	6
DES1275	34.5	2.3	2.8	2.5	0.46 (0.25)	20
DES1296	97.6	11.7	5.4	5.0	0.62 (0.33)	8
DESB	47.6	4.1	3.3	3.0	0.39 (0.21)	23
DESC	33.6	3.5	2.9	3.0	0.30 (0.17)	26
Total	71.1	7.7	4.0	4.0	0.62 (0.31)	
PET1271	105.8	13.3	4.1	4.5	0.52 (0.28)	16
PET1277	58.8	31.3	5.0	5.0	0.59 (0.32)	12
PET1279	142.5	41.8	5.0	5.5	0.56 (0.30)	14
PET1283	121.4	21.5	5.6	5.5	0.67 (0.36)	3
PET1285	111.3	30.4	5.0	5.0	0.62 (0.33)	9
PET1287	133.4	9.9	3.3	3.0	0.38 (0.21)	24
Total	112.2	23.9	4.7	5.0	0.66 (0.33)	
Grand total	77.5	18.6	4.7	4.0	0.73 (0.36)	

22000 years before present. It should be noted that the presence of a seed bank in these populations may have the effect of extending the generation time of the species, making these estimates conservative.

Crossing studies—Average pollen viability for any population as maternal or paternal parents ranged from 85 to 94%. This degree of interfertility is much greater than that achieved in crosses between *H. deserticola* and either parental species

TABLE 5. Parental origin of 17 loci, based on hybrid index scores; A = *Helianthus annuus*, P = *H. petiolaris*. The hybrid index scores ranged from 0 (*H. petiolaris*) to 1 (*H. annuus*), and each score was assigned an upper and lower support limit, representing two log-likelihood units. Loci for which the upper and lower support limits are both within the range for a single parent (0 to 0.5 for *H. petiolaris*, 0.51 to 1.0 for *H. annuus*) are in boldface type. Parental origin of *ORS442*, *DES1275* could not be assigned because no parental alleles existed in the population.

Locus	Population								All Pops
	DES1236	DES1265	DES1270	DES1274	DES1275	DES1296	DESB	DESC	
<i>ORS8</i>	A	P	P	P	P	P	P	A	P
<i>ORS377</i>	A	P	P	P	P	A	P	P	P
<i>ORS3</i>	P	P	P	P	A	A	A	A	P
<i>ORS484</i>	P	P	P	P	P	P	A	A	P
<i>ORS437</i>	P	P	P	P	P	A	P	P	P
<i>ORS7</i>	P	P	P	P	A	P	P	P	P
<i>ORS12</i>	P	P	P	P	P	P	P	P	P
<i>ORS733</i>	P	P	P	P	P	P	P	P	P
<i>ORS299</i>	P	P	P	P	P	P	P	P	P
<i>ORS541</i>	A	A	P	P	A	P	A	A	A
<i>ORS59</i>	A	A	A	P	A	P	P	P	A
<i>ORS442</i>	A	P	P	A	?	A	P	P	A
<i>ORS5</i>	P	A	P	A	A	A	A	A	A
<i>ORS297</i>	P	A	P	A	P	P	A	A	A
<i>ORS613</i>	P	P	A	A	A	P	A	A	A
<i>ORS10</i>	A	A	A	A	A	A	A	A	A
<i>ORS4</i>	A	A	A	A	A	A	A	A	A

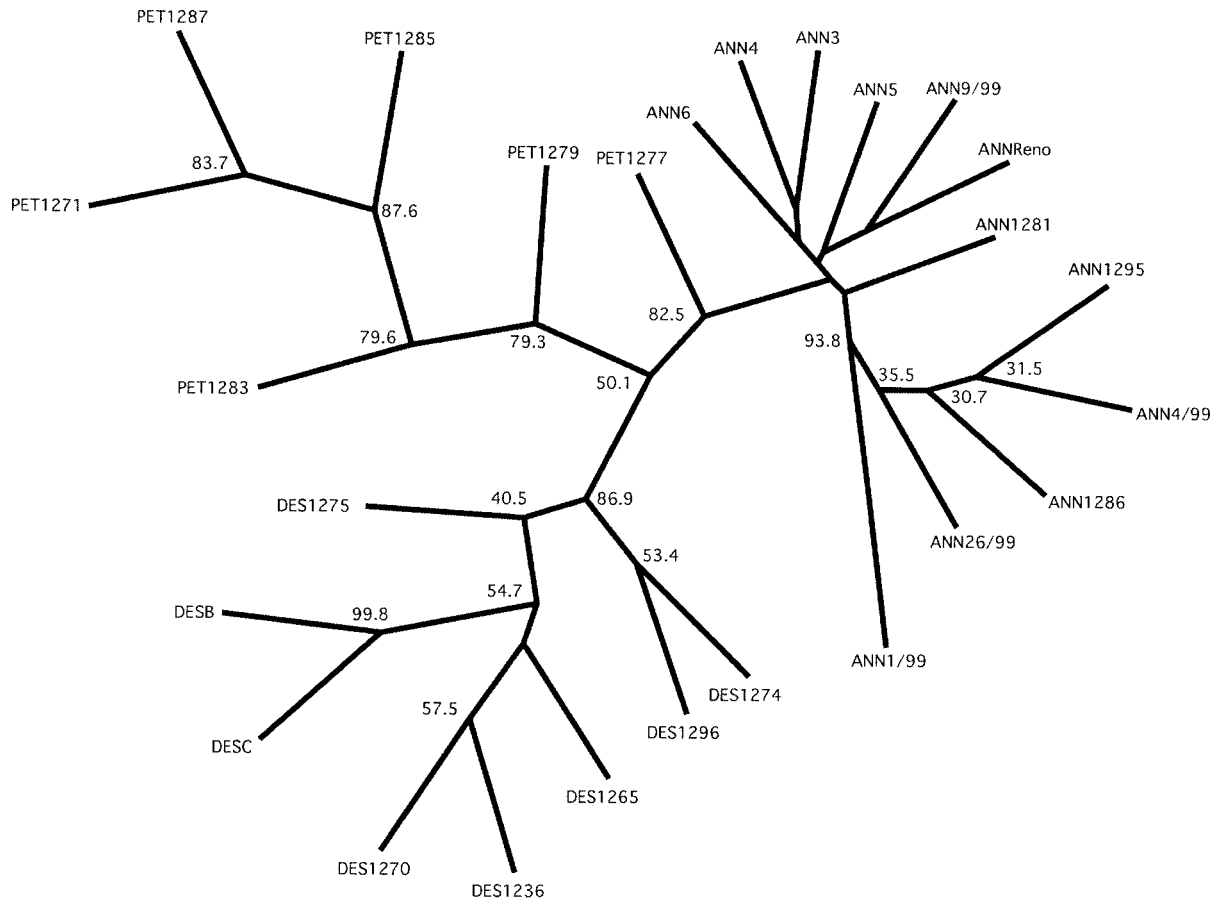


Fig. 3. Unrooted neighbor-joining dendrogram constructed using all 18 microsatellite loci and containing all populations of *Helianthus*. Numbers represent bootstrap support after 1000 replicates. Only bootstrap values over 30 are shown.

and also suggests that there is potential for gene flow between all populations of the hybrid species (Fig. 6). There were no obvious discontinuities in the degree of interfertility, nor were there any statistically significant associations between interfertility and F_{ST} or geographical location.

DISCUSSION

We had predicted two possible patterns for the cpDNA and microsatellite variation in *H. deserticola*. One potential pattern was a uniform cpDNA haplotype across populations and also monophyly for the species based on microsatellite frequencies. The second possible pattern was that *H. deserticola* would possess a variety of cpDNA haplotypes and that the lineage would be polyphyletic with respect to populations of the parental species. The results followed the second pattern, and the two possible explanations for this variation are considered next.

Single origins with subsequent introgression—The cpDNA haplotypes from both parents are present in *H. deserticola*. A single origin is only in accordance with this evidence if the possibility of cytoplasmic introgression is considered. Given the current distribution of haplotypes and the constraint of a single origin, all populations of *H. deserticola* would likely derive from a single diploid hybrid speciation event in which the maternal population of *H. petiolaris* contained haplotypes 8 and 13, probably a population in the Colorado River basin

(e.g., PET1283). The predominance of haplotype 13 in the southwest corner of Utah and in Nevada could be attributed to genetic drift or to founder events, while the polymorphism in eastern populations may be due to retention of ancestral variation (provided sufficient population size) or continued gene flow with neighboring populations of *H. petiolaris*. Population DES1265 in northern Utah could have gained haplotype 1 through introgression with *H. annuus*. This explanation is plausible, as haplotypes 1, 6, 10, and 13 are found in both of the parental species (e.g., PET1271 contains one individual with the *H. annuus* haplotype 1, and ANN1281 is fixed for an *H. petiolaris* haplotype 13), and cytoplasmic introgression is not an uncommon event in taxa capable of interspecific hybridization (Rieseberg, 1995).

A striking case of polyphyly occurs when the tree based on the eight microsatellite loci largely derived from *H. annuus* is considered. Here, the two populations of *H. deserticola* from western Nevada (DESB and DESC) group more closely with a local population of *H. annuus* than with conspecifics. A single origin would require hybridization and nuclear introgression between these or ancestral populations, an event that seems reasonable based on the patterns of cpDNA variation noted earlier. The inclusion of ANN5 in this group is curious, given that the population is not geographically proximal; one explanation for its occurrence may be long-distance dispersal.

The case for a single origin of *H. deserticola* followed by dispersal or range expansion is also supported by the fact that

speciation event involving a maternal population of *H. petiolaris* in the Colorado River basin. The single *H. deserticola* population with an *H. annuus* haplotype (haplotype 1) in northern Utah may have originated in a distinct hybridization event where *H. annuus* served as the maternal parent. These two hypothetical hybrid speciation events occur in localities that are geographically isolated via distance; this supports the possibility of multiple origins rather than the possibility that *H. deserticola* has spread to these different locations after a single hybrid speciation event. If this population of *H. deserticola* were, in fact, locally derived, then we might expect it to be closely related to one of the proximal populations of *H. annuus* based on patterns of microsatellite variation. Instead, this population appears to be more closely related to geographically distant populations of *H. deserticola* than to any nearby populations of the parental species according to both the neighbor-joining trees and F_{ST} values, implying that the population of *H. annuus* that served as the maternal parent in this speciation event (or hybridized with *H. deserticola*) was not sampled in our survey. The high interfertility and moderate values of N_m among populations of *H. deserticola* (0.70) are also relevant to this issue, however, because gene flow among populations of the species may have eroded initial differences resulting from multiple origins.

The patterns of microsatellite variation provide stronger support for an alternate hybrid speciation event; this one would have occurred in western Nevada, where there are currently two isolated populations of *H. deserticola*, DESB and DESC. The potential for a unique speciation event is suggested by the close phylogenetic relationship between the populations of *H. deserticola* and the ANNReno population based on the alleles derived from *H. annuus*; this suggests that ANNReno (or an ancestral population) may have served as the pollen parent in the original hybrid speciation event (Fig. 4). These populations of *H. deserticola* have no nearby *H. petiolaris* sample populations, so it is not possible to say if the type 13 cpDNA present in those populations is locally derived or the result of the proposed hybrid speciation event in the Colorado River basin. These populations are geographically distant from other populations of *H. deserticola*, and *H. petiolaris* does occur in this area historically (Rogers et al., 1982), so an independent hybrid speciation event is a possibility.

Finally, it is interesting to consider some of the broad patterns relevant to this question. First, no two populations seem to share an identical genetic makeup. When loci were assigned an origin in either of the two parents, 12 of the 17 loci amenable to this analysis had varying patterns at the population level (Table 5). If the species was derived from a single speciation event, one might predict a greater uniformity at the genome level. The fact that populations of the hybrid species have retained loci from different parents suggests that the species is either a product of a diverse founder population or that there were multiple origins of the species, resulting in the fixation of different loci in many cases. The five loci that do not show this pattern may show the effects of selection or certain genetic constraints on the formation of *H. deserticola*; i.e., they may be vital to survival in the desert environment, or they may be the only possible combination of parental loci that do not contribute to negative epistasis in the hybrid. Second, there is a considerable variation in the degree of interfertility among populations of *H. deserticola*, ranging from 67.7% (± 3.1) to 96.8% (± 1.2). While this variation is not significantly correlated with geographic or genetic variation, it

exceeds the level of variation normally found among populations of the parental species (Rieseberg, 2000) and suggests that the populations are not genetically uniform.

Diploid hybrid species and their origins—The possibility of multiple origins of a diploid hybrid species has only recently been explored in the literature (Brochmann et al., 2000; Wang et al., 2001; Schwarzbach and Rieseberg, 2002; Welch and Rieseberg, 2002). In contrast, researchers have already shown that multiple origins of allopolyploids are almost as common as single origins (Soltis and Soltis, 1993). It is possible that the potential for multiple origins of diploid hybrid species was not investigated earlier because theory has held it to be unlikely. Whatever the reason, research into this scenario is less clear-cut than is research into the origins of allopolyploids. The main difficulty lies in the potential for continued gene flow between the hybrid species and its progenitors; the cytoplasmic and nuclear introgression that result from these events can leave patterns that are essentially indistinguishable from those caused by multiple origins. The strong reproductive isolating barrier between allopolyploids and their diploid parents makes introgression less likely (although see Husband, 2000) and the answers less equivocal.

In this study, we have shown that *H. deserticola* is a genetically diverse species, containing cpDNA haplotypes from both parents and showing patterns of polyphyly at 18 microsatellite loci. The occurrence of multiple cpDNA haplotypes is not unexpected because the same phenomenon is seen in other diploid hybrid species, such as *Pinus densata*, *Argyranthemum sundingii*, and *H. anomalus* (Wang and Szmidi, 1994; Brochmann et al., 2000; Schwarzbach and Rieseberg, 2002). Para- or polyphyly for nuclear markers in a diploid hybrid lineage has also been reported previously; populations of *P. densata* appear to be paraphyletic, based on evidence from allozyme loci (Wang et al., 2001). However, this is the first report of polyphyly of a hybrid species based on nuclear microsatellite loci.

Levels of genetic diversity within the hybrid species *H. deserticola* are similar to those present in *H. paradoxus*, a species with a single origin, i.e., *H. deserticola* is genetically depauperate when compared to its parents, suggesting a population bottleneck at its origin. This finding confirms reports based on allozyme variability, where *H. deserticola* and *H. paradoxus* generally contained a lower percentage of polymorphic loci, mean number of alleles per locus, and mean heterozygosity than the parental species. Particularly, *H. anomalus*, the hybrid species most likely to have originated from multiple hybridization events, yielded higher levels of genetic diversity than the other two hybrid species based allozyme data (Rieseberg et al., 1991), and data from microsatellite loci revealed no significant difference between the levels of diversity in the hybrid compared to the parental species (Schwarzbach and Rieseberg, 2002). Two other diploid hybrid species, *Iris nelsonii* and *Stephanomeria diegensis*, are characterized by levels of diversity that are roughly equivalent or slightly lower than those found in either parent according to allozyme markers (Gallez and Gottlieb, 1982; Arnold et al., 1990). *Pinus densata* differs from *H. deserticola* in that the hybrid species has a higher level of genetic diversity than either parent based on allozyme markers (Wang et al., 1990, 2001).

Remarkably, despite the variety of cpDNA types and polyphyletic nature of *H. deserticola*, the geographically disparate populations have a high degree of interfertility, especially

compared to interspecific crosses, indicating that *H. deserticola* is a “good” species (Fig. 6). The species likely originated between $170\,000 \pm 12\,000$ and $63\,000 \pm 11\,000$ years before present, and thus its inception likely preceded human disturbance and was the result of hybridization events in the wild (Dixon, 2001). Unfortunately, it is impossible to say which of the diploid hybrid sunflowers came first, as the other studies yield similar dates; the estimated origin of *H. anomalus* is between 144 000 and 116 000 years before present, while *H. paradoxus* likely originated between 208 000 and 78 000 years ago (Schwarzach and Rieseberg, 2002; Welch and Rieseberg, 2002). The estimated maximum and minimum age of *H. deserticola* differ greatly, and the dates should be viewed with reservation because it was necessary to assume that the markers evolved according to the multistep mutational model employed by Zhivotovsky (2001) and that they evolved at the same rate as SSRs in soybean (Diwan and Cregan, 1997). However, note that the estimated dates place the origin of all three hybrid species after the colonization of North America by bison approximately 200 000 years BP (Dary, 1974). Bison are considered to be the primary dispersal agent for sunflower and may have brought the parental species into contact and created the habitat disturbances that appear to facilitate hybridization between them (Asch, 1993).

Conclusions and implications—The patterns of genetic variation within *H. deserticola*, when compared with the parental species, have two potential explanations, neither of which appears to be superior. It is possible that there was a single origin for the species, followed by cytoplasmic and nuclear introgression with different populations of the parental species. It is also possible that *H. deserticola* originated from up to three unique hybrid speciation events and that the present day populations are the result of parallel selection pressure and a coordinate response based on a common genetic starting point.

The recent interest in the potential for multiple origins of diploid hybrid species comes on the heels of convincing evidence for parallel speciation due to ecological factors from several laboratories (Schluter and Nagel, 1995; Rundle et al., 2000; Levin, 2001). In light of these findings, it is important to consider that a diploid hybrid speciation event depends on both ecological and genomic factors to enforce reproductive isolation between the nascent hybrid and its parental species. This represents a contrast to cases of allopolyploid hybrid speciation where ploidy differences alleviate the requirement for immediate ecological divergence (although not the eventual necessity). We propose that despite their apparent rarity, homoploid hybrid species represent potentially powerful study systems for the investigation of principles basic to more common forms of speciation. For example, homoploid hybrid speciation represents a kind of speciation with gene flow. Likewise, this mode of speciation is facilitated by rapid ecological divergence, where hybrid species typically inhabit extreme environments in comparison with the parental species. This dependence on strong ecological selection in the face of ongoing gene flow is clearly pertinent to speciation as a general process, particularly given that the ancestral genotype of diploid hybrid species can be reproduced in a laboratory setting (Rieseberg et al., 1996; Rieseberg, 2000). Thus, diploid hybrid species represent tractable systems in which to investigate ecologically relevant traits that are of importance early in the speciation process and that diverge in the presence of gene flow (Lexer et al., 2003, in press).

LITERATURE CITED

- ALLAN, G. J., C. CLARK, AND L. H. RIESEBERG. 1997. Distribution of parental DNA markers in *Encelia virginensis* (Asteraceae: Heliantheae), a diploid species of putative hybrid origin. *Plant Systematics and Evolution* 205: 205–221.
- ANDERSON, E. 1948. Introgressive hybridization. Wiley, New York, New York, USA.
- ARNOLD, M. L., J. L. HAMRICK, AND B. D. BENNETT. 1990. Allozyme variation in Louisiana irises: a test for introgression and hybrid speciation. *Heredity* 65: 297–306.
- ASCH, D. L. 1993. Common sunflower (*Helianthus annuus* L.): the pathway toward its domestication. Proceedings of the 58th Annual Meeting for the Society for American Archaeology, St. Louis, Missouri, USA: 1–29.
- BARTON, N. H., AND M. SLATKIN. 1986. A quasi-equilibrium theory of the distribution of rare alleles in a subdivided population. *Heredity* 56: 409–415.
- BROCHMANN, C., L. BORGEN, AND O. E. STABBETORP. 2000. Multiple diploid hybrid speciation of the Canary Island endemic *Argyranthemum sundin-gii* (Asteraceae). *Plant Systematics and Evolution* 220: 77–92.
- BUERKLE, C. A., R. J. MORRIS, M. A. ASMUSSEN, AND L. H. RIESEBERG. 2000. The likelihood of homoploid hybrid speciation. *Heredity* 84: 441–451.
- CASGRAIN, P., AND P. LEGENDRE. 2001. The R Package for multivariate and spatial analysis, version 4.0 d4, user's manual. Department de sciences biologiques, Universite de Montreal, Montreal, Quebec, Canada.
- CHANDLER, J., C. JAN, AND B. BEARD. 1986. Chromosomal differentiation among the annual *Helianthus* species. *Systematic Botany* 11: 353–371.
- DARY, D. 1974. The buffalo book; the full saga of the American animal. Sage Books, Chicago, Illinois, USA.
- DEMASURE, B., N. SODZI, AND R. PETIT. 1995. A set of universal primers for amplification of polymorphic noncoding regions of mitochondrial and chloroplast DNA in plants. *Molecular Ecology* 4: 129–131.
- DIWAN, N., AND P. B. CREGAN. 1997. Automated sizing of fluorescent-labeled simple sequence repeat (SSR) markers to assay genetic variation in soybean. *Theoretical and Applied Genetics* 95: 723–733.
- DIXON, E. 2001. Human colonization of the Americas: timing, technology and process. *Quaternary Science Review* 20: 277–299.
- DON, R., P. COX, B. WAINWRIGHT, K. BAKER, AND J. MATTICK. 1991. Touch-down PCR to circumvent spurious priming during gene amplification. *Nucleic Acids Research* 19: 4008–4008.
- DUMOLIN-LAPEGUE, S., M.-H. PEMONGE, AND R. PETIT. 1997. An enlarged set of consensus primers for the study of organelle DNA in plants. *Molecular Ecology* 6: 393–397.
- FELSENSTEIN, J. 1993. PHYLIP (Phylogeny Inference Package) version 3.5c. Distributed by the author, Department of Genetics, University of Washington, Seattle, Washington, USA.
- GALLEZ, G. P., AND L. D. GOTTLIEB. 1982. Genetic evidence for the hybrid origin of the diploid plant *Stephanomeria diegensis*. *Evolution* 36: 1158–1167.
- GOLDSTEIN, D. B., A. R. LINARES, L. L. CAVALLI-SFORZA, AND M. W. FELDMAN. 1995. An evaluation of genetic distances for use with microsatellite loci. *Genetics* 139: 463–471.
- GRANT, V. 1958. The regulation of recombination in plants. *Cold Spring Harbor Symposia on Quantitative Biology* 23: 337–363.
- GRANT, V. 1966a. The origin of a new species of *Gilia* in a hybridization experiment. *Genetics* 54: 1189–1199.
- GRANT, V. 1966b. Selection for vigor and fertility in the progeny of a highly sterile species hybrid in *Gilia*. *Genetics* 53: 757–775.
- GRANT, V. 1971. Plant speciation. Columbia University Press, New York, New York, USA.
- HEISER, C. 1947. Hybridization between the sunflower species *Helianthus annuus* and *H. petiolaris*. *Evolution* 1: 249–262.
- HEISER, C., D. SMITH, S. CLEVINGER, AND W. MARTIN. 1969. The North American sunflowers (*Helianthus*). *Memoirs of the Torrey Botanical Club* 22: 1–218.
- HENDRIE, R., J. ABERCROMBIE-HIGGINS, T. CHANG, R. MADDOCKS, B. TRUONG, E. MORGAN, AND J. WETHERALL. 1998. Genepop on the Web. Website: <http://wbiomed.curtin.edu.au/genepop/>.
- HUSBAND, B. C. 2000. Constraints on polyploid evolution: a test of the minority cytotype exclusion principle. *Proceedings of the Royal Society of London Series B-Biological Sciences* 267: 217–223.

- LEVIN, D. A. 1983. Polyploidy and novelty in flowering plants. *American Naturalist* 122: 1–25.
- LEVIN, D. A. 2001. The recurrent origin of plant races and species. *Systematic Botany* 26: 197–204.
- LEXER, C., M. E. WELCH, J. L. DURPHY, AND L. H. RIESEBERG. 2003. Natural selection for salt tolerance quantitative trait loci (QTLs) in wild sunflower hybrids: implications for the origin of *Helianthus paradoxus*, a diploid hybrid species. *Molecular Ecology* 12: 1225–1235.
- LEXER, C., M. E. WELCH, O. RAYMOND, AND L. H. RIESEBERG. In press. The origin of ecological divergence in *Helianthus paradoxus* (Asteraceae): selection on transgressive characters in a novel hybrid habitat. *Evolution*.
- LISTON, A. 1992. Variation in the chloroplast genes *rpoC1* and *rpoC2* of the genus *Astragalus* (Fabaceae)—evidence from restriction site mapping of a PCR-amplified fragment. *American Journal of Botany* 79: 953–961.
- MCCARTHY, E. M., M. A. ASMUSSEN, AND W. W. ANDERSON. 1995. A theoretical assessment of recombinational speciation. *Heredity* 74: 502–509.
- MCDADE, L. 1990. Hybrids and phylogenetic systematics I. Patterns of character expression in hybrids and their implications for cladistic analysis. *Evolution* 44: 1685–1700.
- MCDADE, L. 1992. Hybrids and phylogenetic systematics II. The impact of hybrids on cladistic analysis. *Evolution* 46: 1329–1346.
- NEI, M. 1987. Molecular evolutionary genetics. Columbia University Press, New York, New York, USA.
- OTTO, S. P., AND J. WHITTON. 2000. Polyploid incidence and evolution. *Annual Review of Genetics* 34: 401–437.
- RAYMOND, M., AND R. ROUSSET. 1995. GENEPOP (version 1.2): population genetics software for exact test and ecumenicism. *Journal of Heredity* 86: 248–249.
- RIESEBERG, L. H. 1991. Homoploid reticulate evolution in *Helianthus* (Asteraceae): evidence from ribosomal genes. *American Journal of Botany* 78: 1218–1237.
- RIESEBERG, L. H. 1995. The role of hybridization in evolution: old wine in new skins. *American Journal of Botany* 82: 944–953.
- RIESEBERG, L. H. 1997. Hybrid origins of plant species. *Annual Review of Ecology and Systematics* 28: 359–389.
- RIESEBERG, L. H. 2000. Crossing relationships among ancient and experimental sunflower hybrid lineages. *Evolution* 54: 859–865.
- RIESEBERG, L. H., S. J. E. BAIRD, AND A. M. DESROCHERS. 1998. Patterns of mating in wild sunflower hybrid zones. *Evolution* 52: 713–726.
- RIESEBERG, L. H., S. M. BECKSTROM-STERNBERG, A. LISTON, AND D. M. ARIAS. 1991. Phylogenetic and systematic inferences from chloroplast DNA and isozyme variation in *Helianthus* sect. *Helianthus* (Asteraceae). *Systematic Botany* 16: 50–76.
- RIESEBERG, L. H., R. CARTER, AND S. ZONA. 1990. Molecular tests of the hypothesized hybrid origin of two diploid *Helianthus* species (Asteraceae). *Evolution* 44: 1498–1511.
- RIESEBERG, L. H., AND N. C. ELLSTRAND. 1993. What can morphological and molecular markers tell us about plant hybridization? *Critical Reviews in Plant Sciences* 12: 213–214.
- RIESEBERG, L. H., B. SINERVO, C. R. LINDER, M. C. UNGERER, AND D. M. ARIAS. 1996. Role of gene interaction in hybrid speciation: evidence from ancient and experimental hybrids. *Science* 272: 741–745.
- RODRIGUEZ, D. J. 1996. A model for the establishment of polyploidy in plants: viable but infertile hybrids, iteroparity, and demographic stochasticity. *Journal of Theoretical Biology* 180: 189–196.
- ROGERS, C., T. THOMPSON, AND G. J. SEILER. 1982. Sunflower species of the United States. National Sunflower Association, Bismarck, North Dakota, USA.
- RUNDLE, H. D., L. M. NAGEL, J. W. BOUGHMAN, AND D. SCHLUTER. 2000. Natural selection and parallel speciation in sympatric sticklebacks. *Science* 287: 306–308.
- SAITOU, N., AND M. NEI. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* 4: 406–426.
- SANG, T., D. J. CRAWFORD, AND T. F. STUESSY. 1995. Documentation of reticulate evolution in peonies (*Paeonia*) using internal transcribed spacer sequences of nuclear ribosomal DNA: implications for biogeography and concerted evolution. *Proceedings of the National Academy of Science* 92: 6813–6817.
- SCHLUTER, D., AND L. M. NAGEL. 1995. Parallel speciation by natural selection. *American Naturalist* 146: 292–301.
- SCHNEIDER, S., D. ROESSLI, AND L. EXCOFFIER. 2000. Arlequin ver. 2.000: a software for population genetics data analysis. Genetics and Biometry Laboratory, University of Geneva, Geneva, Switzerland.
- SCHWARZBACH, A. E., AND L. H. RIESEBERG. 2002. Likely multiple origins of a diploid hybrid sunflower species. *Molecular Ecology* 11: 1703–1715.
- SMITH, H., AND K. DALY. 1959. Discrete populations derived by interspecific hybridization and selection. *Evolution* 13: 476–487.
- SOLTIS, D. E., AND P. S. SOLTIS. 1993. Molecular data and the dynamic nature of polyploidy. *Critical Reviews in Plant Sciences* 12: 243–273.
- SOLTIS, P. S., G. M. PLUNKETT, S. J. NOVAK, AND D. E. SOLTIS. 1995. Genetic variation in *Tragopogon* species: additional origins of the allotetraploids *T. mirus* and *T. miscellus* (Compositae). *American Journal of Botany* 82: 1329–1341.
- STEBBINS, G. L. 1950. Variation and evolution in plants. Columbia University Press, New York, New York, USA.
- STEBBINS, G. L. 1957. The hybrid origin of microspecies in the *Elymus glaucus* complex. *Cytologia Supplemental* 36: 336–340.
- SWOFFORD, D. 2002. PAUP* 4.0b8 Phylogenetic analysis using parsimony (*and other methods). Sinauer Associates, Sunderland, Massachusetts, USA.
- TANG, S., J.-K. YU, M. SLABAUGH, D. SHINTANI, AND S. J. KNAPP. 2002. Simple sequence repeat map of the sunflower genome. *Theoretical and Applied Genetics*, Online DOI 10.1007/s00122-002-0989-y.
- TEMPLETON, A. R. 1981. Mechanisms of speciation—a population genetic approach. *Annual Review of Ecology and Systematics* 12: 23–48.
- THOMPSON, J. D., AND R. LUMARET. 1992. The evolutionary dynamics of polyploid plants: origins, establishment and persistence. *Trends in Ecology and Evolution* 7: 302–307.
- WANG, X.-R., AND A. E. SZMIDT. 1994. Hybridization and chloroplast DNA variation in a *Pinus* complex from Asia. *Evolution* 48: 1020–1031.
- WANG, X.-R., A. E. SZMIDT, A. LEWANDOWSKI, AND Z.-R. WANG. 1990. Evolutionary analysis of *Pinus densata* (Masters) a putative Tertiary hybrid. 1. Allozyme variation. *Theoretical and Applied Genetics* 80: 635–640.
- WANG, X.-R., A. E. SZMIDT, AND O. SAVOLAINEN. 2001. Genetic composition and diploid hybrid speciation of a high mountain pine, *Pinus densata*, native to the Tibetan Plateau. *Genetics* 159: 337–346.
- WEIR, B. S. 1996. Genetic data analysis II: methods for discrete population genetic data. Sinauer Associates, Sunderland, Massachusetts, USA.
- WELCH, M. E., AND L. H. RIESEBERG. 2002. Patterns of genetic variation suggest a single, ancient origin for the diploid hybrid species *Helianthus paradoxus*. *Evolution* 56: 2126–2137.
- WOLFE, A. D., Q.-Y. XIANG, AND S. K. KEPHART. 1998a. Diploid hybrid speciation in *Penstemon* (Scrophulariaceae). *Proceedings of the National Academy of Science* 95: 5112–5115.
- WOLFE, A. D., Q.-Y. XIANG, AND S. K. KEPHART. 1998b. Assessing hybridization in natural populations of *Penstemon* (Scrophulariaceae) using hypervariable intersimple sequence repeat (ISSR) bands. *Molecular Ecology* 7: 1107–1125.
- ZHIVOTOVSKY, L. A. 2001. Estimating divergence time with the use of microsatellite genetic distances: impacts of population growth and gene flow. *Molecular Biology and Evolution* 18: 700–709.