Molecular Ecology (2002) **11**, 000–000

Likely multiple origins of a diploid hybrid sunflower species

A. E. SCHWARZBACH* and L. H. RIESEBERG†
*Department of Biological Sciences, Kent State University, Kent, OH 44242, USA; †Department of Biology, Indiana University, Jordan Hall 142, 1001 East Third Street, Bloomington, IN 47405, USA

Abstract

The recurrent origin of diploid hybrid species is theoretically improbable because of the enormous diversity of hybrid genotypes generated by recombination. Recent greenhouse experiments, however, indicate that the genomic composition of hybrid lineages is shaped in part by deterministic forces, and that recurrent diploid hybrid speciation may be more feasible than previously believed. Here we use patterns of variation from chloroplast DNA (cpDNA), nuclear microsatellite loci, cross-viability and chromosome structure to assess whether a well-characterized diploid hybrid sunflower species, *Helianthus anomalus*, was derived on multiple occasions from its parental species, *H. annuus* and *H. petiolaris*. Chloroplast DNA and crossability data were most consistent with a scenario in which *H. anomalus* arose three times: three different *H. anomalus* fertility groups were discovered, each with a unique cpDNA haplotype. In contrast, there was no clear signature of multiple, independent origins from the microsatellite loci. Given the age of *H. anomalus* (> 100 000 years bp), it may be that microsatellite evidence for recurrent speciation has been eroded by mutation and gene flow through pollen.

Keywords: *Helianthus*, hybridization, hybrid speciation, phylogeography, recurrent speciation, sunflowers

Received 24 April 2002; revision received 6 May 2002; accepted 6 May 2002

Introduction

Reports that species may sometimes be multiply derived have broad evolutionary implications (Levin 2000). Frequent recurrent speciation, for example, would support a larger role for determinism in evolution and might undermine the use of monophyly as a species criterion. Multiply derived species also provide natural experiments for studying the causes of speciation (Schluter 2000). Ecological speciation (i.e. speciation caused by divergent selection in contrasting environments) and alloplyploidy appear most conducive to recurrent origins. The most striking evidence for the former derives from threespine sticklebacks, in which sympatric limnetic and benthic species have evolved independently on multiple occasions in separate lakes (Schluter & Nagel 1995; Taylor & McPhail 2000). The case for recurrent allopolyploid speciation is even more compelling. For example, molecular marker evidence suggests two allotetraploid goatsbeard species (*Tragopogon mirus* and *T. miscellus*) may have arisen as many as 30 times in the past century alone (Soltis *et al*. 1995). Indeed, the multiple origin of allopolyploid species now appears to be the rule rather than the exception (Soltis & Soltis 1993).

Evidence for the recurrent origins of species derived through other means is less convincing, although it has been suggested for species originating through shifts in mating system (e.g. Strid 1970; Goodwillie 1999; Levin 2000) and for two putative diploid hybrid species, *Argyranthemum sundingii* (Brochmann *et al*. 2000) and *Pinus densata* (Wang *et al*. 2001). Unambiguous documentation of the multiple origins of diploid hybrid species is particularly difficult because evolutionary phenomena other than hybridization can account for the shared morphological and/or molecular features often used as evidence of hybrid origins (Avise 1994; Rieseberg 1997). Even when evidence for a hybrid ancestry is conclusive, introgression following hybrid origin may produce patterns that are almost indistinguishable from those produced by recurrent hybrid speciation.

Correspondence: Loren H. Rieseberg. Fax: +1 812 855 6705; E-mail: lriesebe@indiana.edu

The study of recurrent diploid hybrid speciation has also been hampered by the perception that it is improbable (Stebbins 1950; Grant 1958, 1981). Unlike allopolyploids, which contain complete copies of both parental species' genomes, diploid hybrid species are the products of many generations of recombination. As a result, hybrid lineages with similar genotypes should be rare. The segregation of genetic factors that cause sterility has received the most attention. For example, Stebbins (1950) has shown that the number of different classes of fertile F_2 offspring from a given cross is 2^n, where n is the number of independent pairs of complementary sterility factors that differentiate the parental lines. That is, four classes of fertile F_2 plants can theoretically be extracted from a species' pair that differs by two sets of complementary sterility factors, eight classes from species differentiated by three sets, 16 classes from species differentiated by four sets and so forth. Consequently, multiple origins of a diploid hybrid species was considered unlikely, particularly when hybrid sterility had a complex genetic basis (Stebbins 1950; Grant 1958).

Recent experimental evidence suggests, however, that diploid hybrid speciation may be more repeatable than previously suspected, even when numerous hybrid incompatibilities are involved (Rieseberg *et al*. 1996; Rieseberg 2000). For example, Rieseberg *et al*. (1996) demonstrated that three experimentally generated hybrid lineages of sunflower converged onto a combination of chromosomal blocks similar to that found in an ancient hybrid species (*Helianthus anomalus*) with the same parentage. Crossing studies verified the genomic composition data: the three synthetic lineages were cross-compatible with each other and with *H. anomalus* (Rieseberg 2000). Thus, it appears that fertility selection may shape the genomes of hybrid sunflower lineages. The cross-compatibility of the hybrid lineages contrasts with the near sterility of first-generation hybrids between the parental species, *H. annuus* and *H. petiolaris* (Ungerer *et al*. 1998), which is caused by a minimum of 16 genetic factors (Rieseberg *et al*. 1999).

The ease with which a genome similar to that of *H. anomalus* could be created in the greenhouse raises the possibility that this species might have been derived on multiple occasions in the wild. To address the question of multiple origins, we assayed natural populations of *H. anomalus* and its parental species for restriction site variation in the chloroplast genome, chromosomal structural variation in the nuclear genome, as well as length variation at nuclear microsatellite loci. If *H. anomalus* was multiply derived, populations of the species from different geographical areas might be expected to possess different combinations of chloroplast DNA (cpDNA), chromosomal and microsatellite markers. Alternatively, if *H. anomalus* had a single origin, then populations from across the range of the species should share a similar combination of markers. Of course, other evolutionary phenomena can produce patterns suggestive of multiple origins, or alternatively, mask evidence for recurrent speciation. For example, local introgression with one or both parental species following speciation could produce patterns of cpDNA and microsatellite variation indistinguishable from those caused by multiple hybrid speciation. Likewise, gene flow within species could erase footprints of recurrent speciation. Fortunately, chromosomal structural variants are resistant to introgression with the parental species or to homogenization by gene flow within species (Rieseberg 2001) and thus may offer a means for distinguishing between these hypotheses.

Materials and methods

Study system

Helianthus anomalus is probably the best documented example of a diploid hybrid species in either animals or plants. This diploid ($n = 17$), self-incompatible, annual, combines the molecular markers of both parental taxa (Rieseberg 1991; Rieseberg *et al*. 1996; Ungerer *et al*. 1998), and, as mentioned previously, is genomically similar and cross-compatible with synthetic hybrid lineages derived from its parental species, *H. annuus* and *H. petiolaris*. Although clearly of hybrid origin, *H. anomalus* has diverged considerably from its parents in both karyotype and ecological preference. Much of the karyotypic change can be attributed to the sorting of chromosomal rearrangements that differentiate the parental species, but *H. anomalus* also possesses several unique arrangements, possibly induced by recombination (Rieseberg *et al*. 1995b). As a result, *H. anomalus* is almost completely intersterile with its parental species: F_1 values average 2.6–2.9% viable pollen and 0.16–0.18% viable seed (Rieseberg 2000). Ecologically, *H. anomalus* is a sand dune endemic, restricted to a handful of dunes in central Utah and northern Arizona.

In contrast to the restricted distribution of *H. anomalus*, both parental species are widespread and occur commonly throughout the central and western USA. The two species prefer different habitats: *H. annuus* is found in mesic, clay-based soils and *H. petiolaris* in light, sandy soils. Hybridization occurs in areas where these two habitats are juxtaposed, but the two species retain their genetic integrity because of the synergistic action of several reproductive barriers (Rieseberg *et al*. 1995a, 1999; Schwarzbach *et al*. 2001).

Collections and DNA isolations

Leaves or achenes (one-seeded fruits) of *H. anomalus* and its parental species were collected from populations in Utah and Arizona (Table 1). With one exception (*H. anomalus* 1280), achenes or plants were collected from a large number of individuals (> 20) and 12 achenes or plants

Table 1 Geographic location, population code and number of sampled individuals

Taxon/popn	Total no. individuals	Location	Population code
Helianthus anomalus			
Rieseberg 1244	12	Little Sahara recreation area, nr white sands camp site, Juab Co., UT	ANO1244
Rieseberg 1276	12	Jericho picnic area, Little Sahara recreation area, Juab Co., UT	ANO1276
Rieseberg 1273	12	26.5 miles S of Hanksville on Hwy. 95, Garfield Co., UT	ANO1273
Rieseberg 1260	12	15 miles N of Hanksville on Hwy 24, Emery Co., UT	ANO1260
Rieseberg 1288	12	Hwy. 276, ≈ 14 miles E of Hall's Crossing, San Juan Co., UT	ANO1288
Rieseberg 1280	12	Wash on E side of New Oraibe, Navajo Co., AZ	ANO1280
Rieseberg 1282	12	6 miles E of Mexican Water on Hwy 160, Apache Co., AZ	ANO1280
H. petiolaris			
Rieseberg 1287	12	0.25 miles E of Zion park entrance on Hwy 9, Kane Co., UT	PET1287
Rieseberg 1271	12	0.25 miles W of exit 95, jct. I-5, Hwy 20, Iron Co., UT	PET1271
Rieseberg 1285	12	Kodachrome Basin, Chimney Rock, Kane Co., UT	PET1285
Rieseberg 1292	4	Hanksville city limits, North side, Wayne Co., UT	PET1292
Rieseberg 1283	12	5.3 miles E of jct. 160 and 264 on 264, Coconino Co., AZ	PET1285
Rieseberg 1277	12	10 miles S of Paige on Hwy. 89, Coconino Co., AZ	PET1277
Rieseberg 1279	12	12.4 miles E of jct. 6/264 on Hwy 264, Navajo Co., AZ	PET1279
H. annuus			
Rieseberg 1286	12	Jct I-15, Hwy 6 behind Chevron station, Exit 261, E side, Utah Co., UT	ANN1286
Rieseberg 1295	12	0.25 miles N of Northern city limits of Hanksville, Wayne Co., UT	ANN1295
Rieseberg 1298	12	5.6 miles E of Zion Natl Park on Hwy 9, Kane Co., UT	ANN1298
Schwarzbach & Welch 4/99	12	Hwy 95 between Natural Bridge National Monument and Blanding, San Juan, UT	ANN4/99
Schwarzbach & Welch 1/99	12	Hwy. 6, 1 mile S of Price, Carbon Co., UT	ANN1/99
Schwarzbach & Welch 26/99	12	Hwy. 125, 1 mile W of jct. 125/50, close to Delta, Millard Co., UT	ANN26/99
Schwarzbach & Welch 9/99	12	Hwy. 191, just S of Rock Point, Apache Co., AZ	ANN9/99
Rieseberg 1281	12	on Hwy 160, between miles 418 and 419, Navajo Co., AZ	ANN1281

per population were chosen randomly for DNA isolations and subsequent analyses (below). Attempts were made to sample populations of each species from throughout the range of *H. anomalus*. However, the parental species are unevenly distributed, with *H. annuus* more prevalent in northern Utah and *H. petiolaris* more common in southern Utah and northern Arizona. These skewed distributions are reflected in our sampling scheme (Table 1; Fig. 1). DNA was extracted from silica gel-dried leaves, seeds or greenhouse-grown seedlings using the DNeasy plant extraction kit (QIAGEN Inc., Chatsworth, CA, USA).

Chloroplast DNA analysis

Four rapidly evolving regions in the chloroplast genome were amplified via the polymerase chain reaction (PCR) and subjected to restriction fragment length polymorphism (RFLP) analysis: *trn*C-*trn*D (Demesure *et al*. 1995), *trn*F-*trn*V (Dumolin-Lapegue *et al*. 1997), *trn*V-*rbc*L (Dumolin-Lapegue *et al*. 1997) and *rpo*C1-*rpo*C2 (Liston 1992). PCR reactions (50 μL) contained 5–10 ng of template DNA, 0.08 μg of each primer, 30 mM Tricine, 50 mM KCl, 2 mM $MgCl_2$, 100 μM each dNTP and 2 Units *Taq* DNA polymerase. The PCR profile consisted of 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

Initially, amplified fragments from several individuals of the parental species and *H. anomalus* were screened for RFLPs using the following 16 restriction enzymes: *Aci*I, *Alu*I, *Bfa*I, *Bsa*BI, *Bst*UI, *Dde*I, *Hae*III, *Hha*I, *Hin*fI, *Mbo*I, *Mse*I, *Msp*I, *Rsa*I, *Sau*96I, *Scr*fI and *Taq*I. Six of these (*Aci*I, *Bst*UI, *Dde*I, *Hae*III, *Hha*I and *Taq*I) yielded polymorphic profiles for one or more fragments and were used for RFLP assays across all individual/fragment combinations. Restriction digestions followed the manufacturer's recommendations (New England Biolabs, Beverley, MA, USA) and restriction fragments were separated by electrophoresis in 1.5% agarose gels and visualized by staining with ethidium bromide. Restriction profiles were interpreted in terms of the absence (0) or presence (1) of particular restriction sites (Tables 2 and 3). Because of the small number of polymorphisms detected, a most parsimonious haplotype network was initially constructed manually and then confirmed using PAUP* Version 4.0b8 (Swofford 2000).

Microsatellite analyses

DNAs from several individuals of *H. annuus*, *H. anomalus* and *H. petiolaris* were tested for ≈ 100 microsatellite

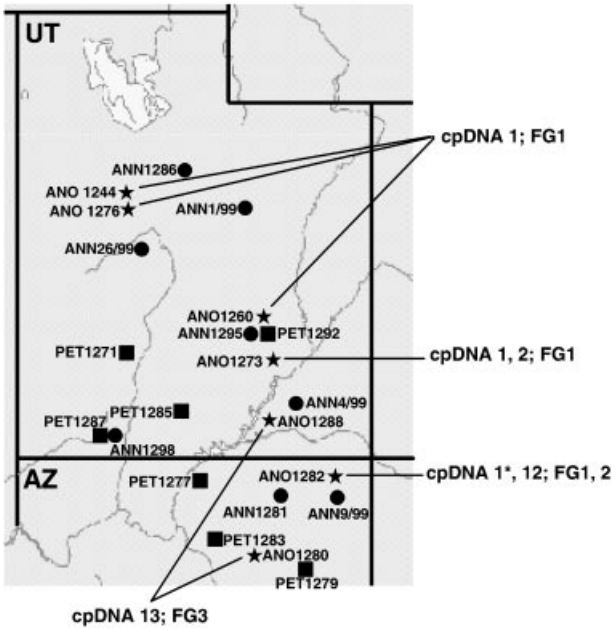

Fig. 1 Geographic distribution of *Helianthus* populations sampled, cpDNA haplotypes and fertility groups (FG). ★, populations of *H. anomalus*; ●, populations of *H. annuus*; ■, populations of *H. petiolaris*. *Note that cpDNA 1 reported for ANO1282 was detected in an earlier study (Rieseberg 1991) rather than our analysis.

primers developed by Tang *et al*. (2002). From this initial screen, 12 primer pairs were chosen for the full analysis that: (i) amplified in all three species; and (ii) produced single-locus, co-dominant banding patterns: ORS3, ORS4, ORS5, ORS7, ORS8, ORS10, ORS12, ORS13, ORS29, ORS59, ORS78, ORS297. (Primer sequences may be obtained from the Sunflower Genome Database: http://www.css.orst.edu/knapp-lab/sunflower) The forward primer from each pair was 5′-labelled with one of three fluorophores (6FAM, HEX or NED) and PCR reactions (10 μL) were performed with 20 ng of template DNA, 0.0016 μg of each primer, 30 mM Tricine, 50 mM KCl, 2 mM $MgCl_2$, 100 μM each dNTP and 0.5 Units of *Taq* DNA polymerase. Cycle parameters included an initial denaturation step of 2 min at 94 °C, 30 cycles of 30 s at 94 °C, 30 s at 50 °C and 1 min at 72 °C, with a final extension at 72 °C for 10 min.

PCR products from different loci were pooled on the basis of size or fluorophore. The resulting mixture was diluted 20-fold and 1 μL of the diluted PCR pool was combined with 9.8 μL ddH_2O and 0.20 μL of the GenSize R500 ROX size standard (GenPak, St. James, NY, USA). Samples were denatured for 5 min at 95 °C, quickly cooled on ice and then genotyped on an ABI 3700 using GENESCAN Version 3.5 and GENOTYPER Version 3.6 software (Applied Biosystems, Foster City, CA, USA).

Basic population genetic statistics, including allele frequencies, deviations from Hardy–Weinberg equilibrium (F_{IS}), linkage disequilibrium between all pairs of loci, genetic subdivision among populations (F_{ST}) and number of migrants (N_em) per generation were calculated using the computer program GENEPOP (Raymond & Rousset 1995). N_em values were based on the private allele model of Barton & Slatkin (1986) and corrected for sample size. Additional estimates of genetic variation and genetic structure were conducted using ARLEQUIN Version 1.1 (Schneider *et al*. 1997). A sequential Bonferroni correction (Rice 1989) was applied whenever multiple statistical tests were conducted.

Spatial autocorrelation analysis, as implemented by the 'R PACKAGE' (Casgrain & Legendre 2000), was used to

Table 2 Chloroplast DNA polymorphisms found in *Helianthus annuus*, *H. petiolaris* and *H. anomalus*

Polymorphism	cpDNA region	Enzyme	Recognition sequence	Mutation (fragment size in bp)	
				0	1
A	*trn*C-*trn*D	*Taq*I	TCGA	1600	1500 + (100)
B	*trn*F-*trn*V	*Bst*UI	CGCG	2550	2000 + 550
C	*trn*F-*trn*V	*Hha*I	GCGC	680	500 + 180
D	*trn*F-*trn*V	*Dde*I	CTNAG	350	240 + 110
E	*trn*F-*trn*V	*Aci*I site 1	CCGC	790	600 + 190
F	*trn*F-*trn*V	*Aci*I site 2	CCGC	600	450 + 150
G	*trn*V-*rbc*L	*Aci*I	CCGC	850	600 + (250)
H	*trn*V-*rbc*L	*Bst*UI site1	CGCG	1150	1000 + 150
I	*trn*V-*rbc*L	*Bst*UI site2	CGCG	1150	960 + 190
J	*trn*V-*rbc*L	*Dde*I	CTNAG	1100	800 + 300
K	*trn*V-*rbc*L	*Hae*III	GGCC	2140	1800 + 340
L	*trn*V-*rbc*L	*Mse*I	TTAA	540	210 + (330)
M	*rpo*C1-*rpo*C2	*Hha*I site 1	GCGC	660	530 + (130)
N	*rpo*C1-*rpo*C2	*Hha*I site 2	GCGC	620	580 + (40)

Table 3 Chloroplast DNA haplotypes found in *Helianthus annuus*, *H. petiolaris* and *H. anomalus*. For listing of polymorphisms see Table 2. Restriction site present: 1; restriction site absent: 0

	Polymorphisms													
Haplotype	A	B	C	D	E	F	G	H	I	J	K	L	M	N	Populations (Frequency)
1	0	0	0	1	1	0	0	0	0	0	0	1	1	0	ANN12861.0, ANN26/99(1.0), ANN1298(0.5), ANN1295(1.0), ANN1/99(1.0), ANN4/99(1.0), ANN9/99(1.0), ANO1244(1.0), ANO1276(1.0), ANO1260(1.0), ANO1273(1.00), PET1271(0.08)
2	0	0	0	1	1	0	0	0	0	0	0	1	0	0	ANO1273(0.08)
3	0	0	0	0	1	0	0	0	0	0	0	1	1	0	ANN1298(0.17)
4	0	0	0	1	1	0	0	0	0	1	1	1	1	0	ANN1298(0.08)
5	0	0	0	1	1	0	1	1	0	0	1	0	0	0	PET1287(0.08)
6	0	0	0	1	1	0	1	1	0	1	1	0	0	0	PET1287(0.33)
7	0	0	0	1	1	0	1	0	0	1	1	0	0	0	PET1279(0.18)
8	0	0	0	1	0	0	1	0	0	1	1	0	0	0	PET1277(1.0)
9	0	0	0	1	1	1	1	0	0	1	1	0	0	0	ANN1298(0.17), PET1283(0.17), PET1292(0.50)
10	0	0	0	0	1	1	1	0	0	1	1	0	0	0	PET1271(0.50), PET1287(0.58), PET1285(0.25)
11	0	0	0	0	1	1	1	0	0	0	1	1	0	0	ANN1298(0.08)
12	0	1	1	1	1	1	1	0	0	1	1	0	0	0	ANO1282(1.0), PET1285(0.75)
13	1	0	0	1	1	1	1	0	0	1	1	0	0	1	ANN1281(1.0), ANO1280(1.0), ANO1288(1.0), PET1279(0.82), PET1283(0.42), PET1285(0.08), PET1292(0.5)
14	1	0	0	1	1	1	1	0	1	1	1	0	0	1	PET1271(0.42)

estimate the relationship between genetic (F_{ST}) and geographical distances. The significance of the autocorrelations (i.e. Mantel's *r*) was tested using 1000 permutations. Dendrograms were generated from the F_{ST} matrix (F_{ST} distances do not differ substantially from other more standard distance measures) by the neighbour-joining method (Saitou & Nei 1987) using the PHYLIP package. An initial dendrogram was generated using all loci and all populations. However, because loci found in *H. anomalus* might be derived from just one or the other parental species, we worried that evidence for the independent origins of individual *H. anomalus* populations might be masked in a combined analysis. Thus, we used a maximum likelihood approach (Rieseberg *et al*. 1998) to estimate a hybrid index for each locus in *H. anomalus* (Table 4). A dendrogram was then constructed using loci derived from *H. petiolaris* only. No loci were derived exclusively from *H. annuus*. Reliability of all trees was tested by 1000 bootstrap replicates. In addition, genetic relationships among populations were graphically assessed by principal components analysis (PCA) of a weighted allele frequency matrix. Weighting was used to avoid overestimating the contribution of alleles from allelic rich loci and underestimating those from loci with few alleles. Weighting was accomplished by conducting a separate analysis of each locus and then weighting alleles in the combined analysis by 1/(the first eigenvalue derived from the separate analysis of that locus). All analyses were performed with ADE-4, a free software package from the University of Lyon, France.

The approximate timing of the origin(s) of *H. anomalus* was estimated from the microsatellite data using equation 1 of Zhivotovsky (2001): $T_D = D_1/2w - V_0/w$. T_D refers to 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 (equation 2 of Goldstein *et al*. 1995); *w* represents the effective mutation

Table 4 Hybrid index scores for 12 microsatellite loci in *Helianthus anomalus*

Marker	Lower support limit*	Hybrid Index†	Upper support limit*
ORS3‡	0	0.11	0.29
ORS4	0.28	0.44	0.60
ORS5‡	0.22	0.35	0.48
ORS7‡	0	0.08	0.30
ORS8	0.24	0.38	0.52
ORS10	0.26	0.48	0.72
ORS12	0.12	0.35	0.69
ORS13	0.32	0.55	0.83
ORS29‡	0	0	0.15
ORS59‡	0	0.13	0.31
ORS78‡	0	0	0.13
ORS297	0.42	0.58	0.72

*The upper and lower support limits represent two log-likelihood units.

†A hybrid index score of 0 means that *H. anomalus* individuals are genotypically similar to *H. petiolaris* at that locus, whereas a hybrid index score of 1 means they are genotypically similar to *H. annuus*.

‡Markers considered to be derived from *H. petiolaris* for phylogenetic purposes.

rate; and V_0 is the average over all loci of the variance in repeat number in the ancestral population. Simulations indicate that this method is robust to variation in population size and weak migration. Because V_0 is unknown, we followed the suggestion of Zhivotovsky (2001) and used minimum (V_{0min}) and maximum (V_{0max}) estimates of the variance in the ancestral population to generate an upper (T_{Dmax}) and lower bound (T_{Dmin}) for divergence times. V_{0min} was placed at 0 and V_{0max} was set to the average of the within-population variance of any two populations being compared (populations of *H. anomalus* in this study). The value employed for w of 2×10^{-4} was based on estimates of mutation rates for over 20 SSRs with di- and trinucleotide repeats in seven soybean pedigrees (Diwan & Cregan 1997).

Crossing studies and meiotic analyses

Two plants from each of the seven populations of *H. anomalus* were grown in Indiana University greenhouses and used as both maternal and paternal parents in experimental crosses. Each wild population was crossed with all other wild populations (21 combinations) by rubbing flowering heads together. In addition, crosses were made between the two individuals within each of the seven populations (seven combinations). All cross-combinations were replicated with two plants per accession (four combinations) and each plant was used as both a pollen donor and recipient (two directions) for a total of 224 crosses: (21 + 7) cross-combinations × 4 replicate crosses × 2 directions.

One plant from each of the 224 crosses was grown to reproductive maturity and fresh pollen was harvested from the first flowering head (seed set was not measured because some flowering heads were damaged by heat stress). The pollen was treated with a solution of 30% sucrose and 0.1% MTT (Chandler *et al*. 1986), and 300 pollen grains were scored for the staining indicative of enzyme activity. Fully darkened grains were considered viable. Because pollen was sometimes not fully mature, all counts were replicated and counts that were made too early were discarded. Means and standard errors were calculated for pollen viabilities for each cross-combination. Three distinct fertility groups were discovered (see Results) and this information was used to guide the meiotic analyses.

To estimate the contribution of chromosomal rearrangements to fertility reduction, immature flowering heads were collected from at least two crosses between each fertility group and two crosses within fertility groups 1 and 3 (fertility group 2 was represented by a single plant). Meiotic analyses followed Rieseberg (2000). Briefly, floral buds were fixed in Carnoy's solution for 24 h and then stored at 4 °C in 70% EtOH. Anther squashes were stained in 2% propiono-carmine and observed under a light microscope. Interpretations were based on analyses of 10–20 diakinesis and metaphase I pollen mother cells from each cross.

Results

Chloroplast DNA analyses

Restriction site analyses of the four chloroplast fragments revealed 14 polymorphisms, all of which could be interpreted in terms of the gain or loss of a restriction site (Table 2). The majority of polymorphisms were found in the fragments *trn*F-*trn*V (five) and *trn*V-*rbc*L (six), with only a single variant detected for *trn*C-*trn*D and two variants for *rpo*C1-*rpo*C2. The 14 polymorphisms were distributed among 14 different cpDNA haplotypes (Table 3). A most parsimonious network connecting the haplotypes (Fig. 2) required 19 steps.

Two clusters of haplotypes were detected that roughly correspond to the parental species, *Helianthus annuus* and *H. petiolaris* (Fig. 2). Haplotypes 1–4 are mostly found in *H. annuus* (or the hybrid, *H. anomalus*) and haplotypes 5–14 are mostly restricted to populations of *H. petiolaris* (or *H. anomalus*). There were exceptions to this. For example, ANN1298, a population near Zion National Park, contained a wide diversity of cpDNA haplotypes (1, 3, 4, 9 and 11) including two that occurred within the *H. petiolaris* cluster (Fig. 2; Table 3). Recent field work indicates, however, that this population occurs at one end of a hybrid zone between *H. annuus* and *H. petiolaris*. Thus, the presence of *petiolaris*-like haplotypes in the population is not surprising. Another population, ANN1281, is fixed for the cpDNA type (haplotype 13) characteristic of abundant

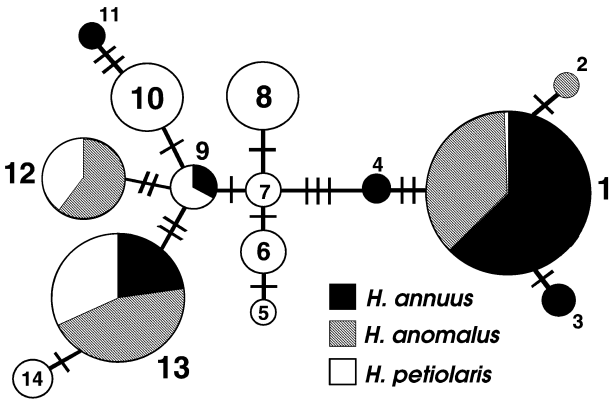

Fig. 2 Most parsimonious network of cpDNA haplotypes (Table 3). Haplotype numbers are given beside each circle. The sizes of circles roughly correspond to the frequency of a given haplotype across all three species, and the proportional representation from each species is indicated by the pie charts within each circle. Mutational steps are indicated by bars between haplotypes.

nearby populations of *H. petiolaris*, and most likely represents a recent example of cytoplasmic introgression, which has been reported previously for these two species (Dorado *et al*. 1992). One possible example of cytoplasmic introgression in the opposite direction was also observed, in which a single individual of *H. petiolaris* from PET1271 possessed the common *H. annuus* cpDNA type (haplotype 1). These patterns might also be interpreted as the product of phylogenetic sorting, but this explanation seems unlikely because mismatches between cpDNA haplotype and species are restricted to parapatric population samples and the common local haplotype is almost always involved.

Four different cpDNA haplotypes were detected in the hybrid species, *H. anomalus* (Figs 1 and 2). The four most northerly populations were largely monomorphic for haplotype 1 of *H. annuus* (Figs 1 and 2), although one individual had a unique cpDNA type (haplotype 2) that differed by a single mutation from haplotype 1. In contrast, the southern populations were fixed for haplotypes of *H. petiolaris* (Figs 1 and 2): ANO1280 and ANO1288 for haplotype 13 and ANO1282 for haplotype 12. Although these data might initially seem most consistent with three separate origins for *H. anomalus*, one in the north and two in the south, it is not possible to rule out the alternative hypothesis of cytoplasmic introgression with these data alone. This is particularly problematic in sunflowers, as cpDNA introgression appears to be frequent (Rieseberg *et al*. 1990, 1991; Dorado *et al*. 1992).

Microsatellite analyses

Analyses of linkage disequilibrium among all pairs of loci revealed significant associations between *ORS10* and *ORS13* in all populations. These two loci are most likely tightly linked, but this has not yet been confirmed by genetic mapping studies. No other locus pairs were significantly associated. Despite the strong linkage disequilibrium observed between *ORS10* and *ORS13*, both were retained in subsequent analyses because they did not have identical patterns of allelic variability and information would have been lost by eliminating either locus.

Estimates of genetic variability revealed that *H. anomalus* populations were as diverse genetically as populations of either parental species (Table 5). These data suggest that either *H. anomalus* did not undergo a population bottleneck during its formation or that it originated long enough ago so that the effects of a bottleneck are no longer apparent. The latter hypothesis is supported by divergence time estimates (below) that indicate an ancient origin for *H. anomalus*. Nonetheless, the majority of genetic variability still resides within populations (70.3%) or among populations within species (17.4%). Only 12.3% of the variance occurs among species. Indirect estimates of gene flow also were similar among the three species, with N_em values of 0.88, 0.91 and 0.90 in *H. anomalus*, *H. annuus* and *H. petiolaris*, respectively.

Table 5 Genetic variation across 12 microsatellite loci in three *Helianthus* species

Species/popn	Percentage polymorphic loci	Average H_E (SE)	Average number alleles/locus
ANN1281	91.7	0.47 (0.26)	3.58
ANN1286	100	0.61 (0.33)	4.83
ANN1295	100	0.62 (0.32)	4.92
ANN1298	90.9	0.51 (0.28)	3.75
ANN1/99	91.7	0.62 (0.34)	4.33
ANN26/99	91.7	0.50 (0.28)	3.67
ANN4/99	91.7	0.60 (0.33)	3.75
ANN9/99	100	0.53 (0.34)	4.75
ANO1244	100	0.61 (0.33)	4.25
ANO1260	100	0.67 (0.37)	5.17
ANO1273	100	0.50 (0.28)	4.0
ANO1276	100	0.54 (0.30)	4.08
ANO1280	90.0	0.38 (0.22)	2.42
ANO1282	100	0.68 (0.35)	5.08
ANO1288	100	0.61 (0.34)	4.58
PET1271	100	0.49 (0.27)	4.25
PET1277	100	0.66 (0.36)	5.25
PET1279	75.0	0.54 (0.35)	4.5
PET1283	100	0.71 (0.38)	5.75
PET1285	100	0.60 (0.33)	4.33
PET1287	90.9	0.32 (0.19)	3.33
PET1292	100	0.72 (0.43)	3.5

There was a significant correlation between geographical and genetic distances in *H. anomalus* (Mantel's $r = 2.50$; $P = 0.006$). Th33e correlation was primarily due to the shortest distance class (0–100 km; Mantel's $r = -0.62$; $P = 0.003$). Note that positive autocorrelation produces negative values of r in the low distance classes (Casgrain & Legendre 2000). The finding of a correlation between geographical and genetic distance in *H. anomalus* might be due to recurrent speciation or isolation by distance or some combination of both.

Analysis of genetic relationships among populations failed to confirm the hypothesis of multiple origins. In both the neighbour-joining tree and the PCA, all populations of *H. anomalus* clustered into a single group (Figs 3 and 4). The same relationship was found whether the analyses were based on the entire data set or on the subset of loci derived from *H. petiolaris* (Table 4). *H. annuus* also formed a discrete cluster in both graphical representations, but populations of *H. petiolaris* were genetically more heterogeneous. In fact, one population (PET1277) actually clustered more closely with *H. anomalus* than with other populations of *H. petiolaris*. There is no evidence that this population has hybridized with *H. annuus* or *H. anomalus*

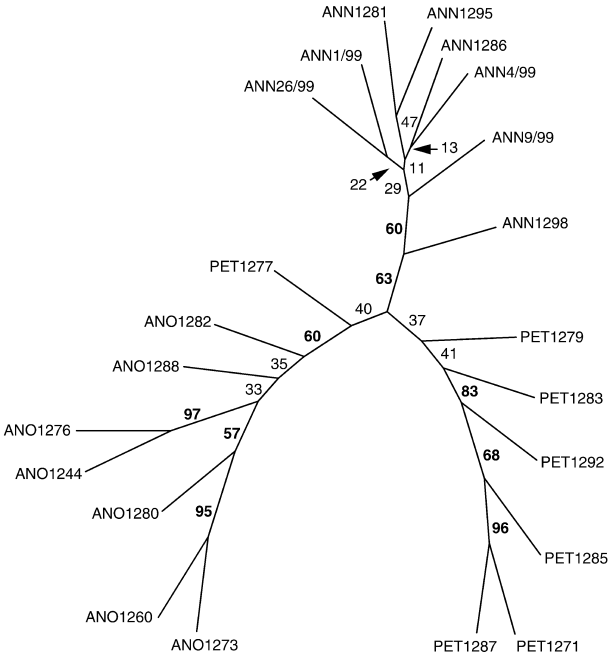

Fig. 3 Neighbour-joining tree for *Helianthus* populations (Table 1) based on F_{ST} genetic distances from all 12 microsatellite loci. Numbers along branches represent bootstrap support (1000 replicates); those ≥ 50% are in bold.

in the past, so its apparent genetic affinity is difficult to explain.

Another puzzle from the microsatellite data was the bias of the hybrid index scores for individual loci toward *H. petiolaris* (Table 4). Two of the loci seem to be derived exclusively from *H. petiolaris* and four additional loci have hybrid index scores that deviate significantly in the direction of *H. petiolaris*. In contrast, no loci deviate significantly in frequencies toward *H. annuus*. These results are at odds with previously published genetic mapping data based on 330 species-specific RAPD and AFLP markers, which indicate that the genomic contributions of the parental species to the *H. anomalus* genome fit a 1:1 ratio (Ungerer *et al*. 1998). However, additional marker screening in several individuals of the three species (L. H. Rieseberg unpublished) indicates that the bias toward *H. petiolaris* for microsatellite loci analysed in this study results from sampling error.

Analyses of the distribution of population- or group-specific markers also failed to provide compelling evidence either for or against the hypothesis of multiple origins. Two alleles (*ORS4-140* and *ORS3-261*) that were restricted to a single population of one of the parental species were found in all populations of *H. anomalus*. This is the kind of pattern that would be expected if *H. anomalus* had a single

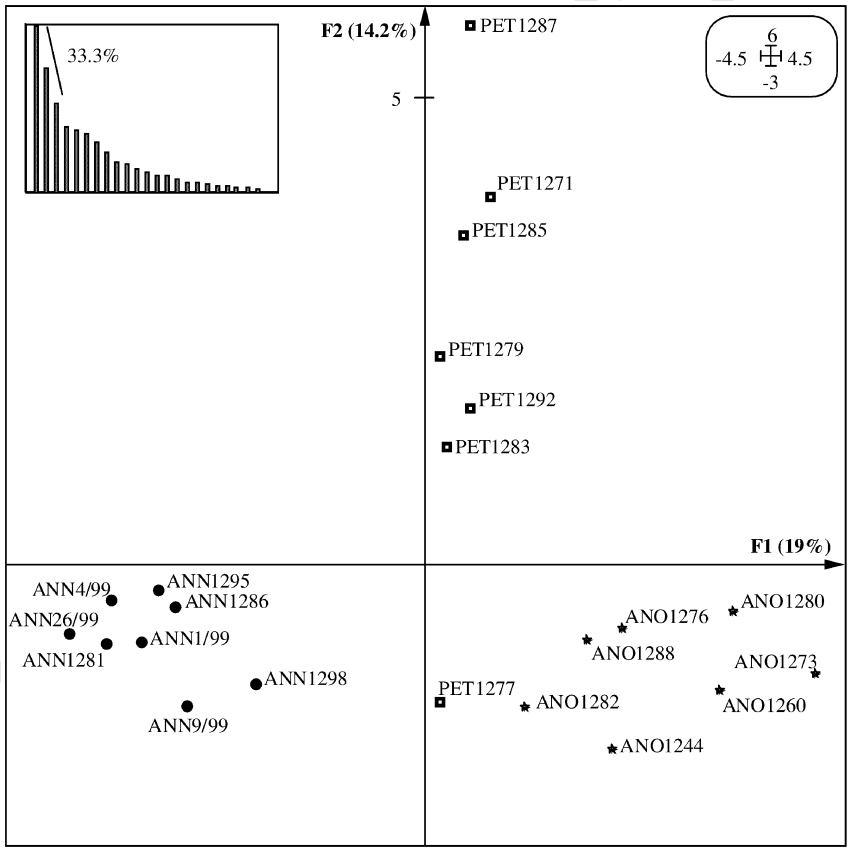

Fig. 4 Principal components analysis of weighted microsatellite allele frequencies for three *Helianthus* species (Factor 1·Factor 2). The box at the upper left hand corner of the figure shows the proportion of phenotypic variation explained by each factor, whereas the rounded rectangle at the upper right hand corner of the figure gives the minimum and maximum ordinates for factors 1 and 2. Population abbreviations follow Table 1.

origin. However, we also found alleles that were exclusive to population groupings based on cpDNA haplotype (Fig. 2) — an observation most consistent with recurrent speciation. For example, the northern populations of *H. anomalus*, which are almost completely fixed for the common *H. annuus* cpDNA type, also share the following exclusive alleles: *ORS59-162*, *ORS3-258* and *ORS29-130*.

Minimum and maximum times of divergence were estimated for all pairs of populations of *H. anomalus* using the methods described previously. Minimum divergence times averaged 116 000 ± 17 000 (mean ± SE) generations, whereas maximum divergence times averaged 144 000 ± 17 000 generations. Because *H. anomalus* is an annual, these values can be viewed as a conservative estimate of when *H. anomalus* arose, if it had a single origin. Speculating that *H. anomalus* was multiply derived has little impact of divergence times. For example, minimum divergence times for the four northern populations with cpDNA haplotype 1 average 117 000 ± 29 000 generations; maximum divergence times for this group average 148 000 ± 28 000 generations. The divergence times for the two populations that share cpDNA haplotype 13 are 141 000–160 000 generations. Of course, all of these dates should be viewed with considerable scepticism because we do not know whether microsatellites in sunflower evolve according to the multistep mutational model employed by Zhivotovsky (2001) or whether mutation rates in sunflower approximate the values used in this study, which were derived from soybean. Nonetheless, they do suggest fairly ancient origin(s) for this species.

Crossing studies and meiotic analyses

Pollen viability of progeny from interpopulational crosses within *H. anomalus* ranged from 47.9 to 96.0% (Fig. 5). These values are considerably higher than those reported for interspecific crosses between *H. anomalus* and its parental species, which average < 3% (Rieseberg 2000). Thus, crossability data accord well with earlier morphological and ecological data (Heiser *et al.* 1969; Nabhan & Reichhardt 1983) that classify populations of *H. anomalus* from throughout Utah and northern Arizona as belonging to a single species. Pollen viability was not significantly affected by the direction of the cross, indicative of nuclear rather than cytoplasmic control of pollen fertility.

Although the crossability data do support the current taxonomy, there is considerable subdivision within *H. anomalus* with respect to fertility (Figs 1 and 5). The four northern populations of *H. anomalus* (ANO1273, ANO1244, ANO1276 and ANO1260), plus one individual from

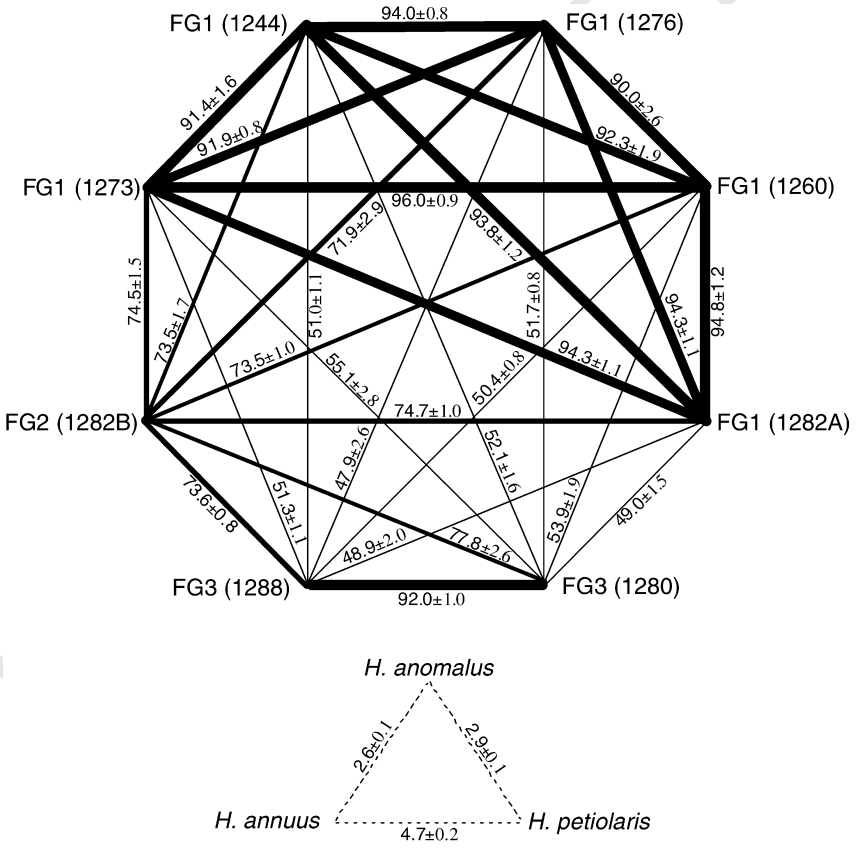

Fig. 5 Crossability relationships among populations of *Helianthus anomalus*. Line thickness is proportional to cross-compatibility. Numbers along lines are the percentages of viable pollen with standard errors. The small figure below the *H. anomalus* crossability diagram shows the percentages of pollen viability obtained from crosses between the three *Helianthus* species (Rieseberg 2000).

Mexican Water, Arizona (ANO1282A), are completely interfertile when crossed, with pollen viability of progeny averaging > 90% for all crosses. We labelled this cluster of interfertile populations as fertility group 1 (FG1; Figs 1 and 5). The second individual from the Mexican Water population (ANO1282B) was uniformly less compatible in crosses with other populations or with ANO1282A, with pollen viability of progeny ranging from 71.9 to 77.8%. We have dubbed this individual as the sole sampled representative of fertility group 2 (FG2; Figs 1 and 5). Finally, populations ANO1288 and ANO1280, from northern Arizona and southern Utah, respectively, were highly interfertile (pollen viabilities averaged 92.0%) and had pollen viabilities ranging from 47.9 to 55.1% when crossed with individuals from fertility group 1. These populations were classified as members of fertility group 3 (FG3; Figs 1 and 5). Note that only two individuals were sampled per population, however, and it may be that a larger number of fertility groups will be discovered with increased sampling.

The discrete nature of the fertility groups suggested to us that the variation in fertility might be caused by large chromosomal rearrangements, which are common in annual sunflowers (Heiser 1947; Chandler *et al*. 1986; Rieseberg *et al*. 1995b) and have previously been shown to cause fertility reductions (Quillet *et al*. 1995; Rieseberg *et al*. 1999). Thus, we examined meiotic configurations in progeny derived from crosses within and among the fertility groups. The data were unambiguous. Meiotic cells from crosses within the fertility groups had perfect bivalent pairing for all 17 chromosomes. In contrast, 9 of 21 cells examined in crosses between FG1 and FG2 had a single quadrivalent, suggesting that the two fertility groups differ by a single interchromosomal translocation. A single quadrivalent was also observed in close to 60% of cells in progeny from FG2 × FG3, suggestive of an interchromosomal translocation. The situation was more complex in meioses from progeny of FG1 × FG3. All but 1 of 40 cells examined exhibited meiotic abnormalities: a single quadrivalent was observed in 16 cells, and two quadrivalents were detected in the remaining 23 cells. These observations suggest that FG1 and FG3 likely differ by two interchromosomal translocations.

One puzzle from the meiotic work is that the translocations detected in this study do not cause the 50% reduction in viable pollen predicted by models (Chandler *et al*. 1986). Similar observations were made by Rieseberg (2000), who suggested that the translocations might be small and/or that recombination is suppressed, leading to lower than expected reductions in pollen viability.

The fertility groups detected correspond closely to groupings defined on the basis of cpDNA haplotypes (Fig. 1). The only exception was ANO1282, which had a single cpDNA type (haplotype 12) in this study (Table 3), but was polymorphic for an interchromosomal translocation. However, a previous molecular phylogenetic study (Rieseberg 1991) detected two different chloroplast DNA haplotypes from the same locality, with one derived from *H. petiolaris* (presumably haplotype 12) and the other from *H. annuus* (haplotype 1).

Discussion

Single vs. multiple origins of Heliathus anomalus

The primary goal of this study was to assess whether a well-characterized diploid hybrid species, *Helianthus anomalus*, was derived on multiple occasions, and if so, to estimate the number of times it arose. Patterns of variation in cpDNA, nuclear microsatellite loci, cross-compatibility and chromosome structure were used to address this question and the results suggest three possible scenarios for the origin and evolution of *H. anomalus*.

Scenario 1. H. anomalus may have had a single origin as suggested by the close genetic relationships among *H. anomalus* populations revealed by analyses of microsatellite loci. To account for the diversity of cpDNA haplotypes in *H. anomalus*, hybrid speciation and subsequent range expansion would have to be followed by a minimum of two independent episodes of cytoplasmic introgression, in which populations of *H. anomalus* captured different cpDNA haplotypes from one or the other parental species. In addition, one or two interchromosomal translocations would have to arise in the same populations that had acquired the new cytoplasmic variants.

There are two difficulties with this hypothesis. First, unlike the examples of recent cytoplasmic introgression between the parental species, populations of *H. anomalus* sometimes do not have the common haplotype of the local parental species' populations. For example, ANO1282 from north-eastern Arizona has a haplotype (cpDNA 12) found only in PET1285 from south-central Utah rather than in geographically local populations from Arizona. To further verify this assertion, we examined the cpDNA haplotype of four individuals from *H. petiolaris* that were within 100 m of the ANO1282 population. None of them had haplotype 12. A second difficulty with this scenario is the correlation observed between cpDNA haplotype and fertility group (Fig. 1). Because interchromosomal translocations are under-dominant in sunflowers (Chandler *et al*. 1986; Rieseberg *et al*. 1999), they are unlikely to have introduced through recent introgression with local parental populations. Thus, we are left with the unsatisfying hypothesis that this correlation arose entirely by chance.

Scenario 2. A second possible scenario also posits a single origin for *H. anomalus*, but assumes that the ancestral hybrid population was highly polymorphic with respect to

cpDNA and chromosome structure, and that current variation patterns in the species are explained by the sorting of cpDNA and chromosomal variants. ANO1282, which is polymorphic for chromosome structure and cpDNA haplotype (Fig. 1), represents a candidate ancestral population. Again, there are two difficulties with this hypothesis. First, the ancestral population would not only have to be highly polymorphic, but it would also have to include cpDNA haplotypes from *H. petiolaris* that currently occur in geographically distant populations. Second, the perfect correlation between cpDNA haplotype and fertility group is unlikely to arise by random assortment.

Scenario 3. This scenario posits three independent origins for *H. anomalus* — one in the north and two in the south. This would account best for the correlation between cpDNA haplotype and fertility group. It also is noteworthy that three synthetic hybrid lineages generated between *H. annuus* and *H. petiolaris* (Rieseberg *et al*. 1996) differed by one or two translocations from each other and from *H. anomalus* (Rieseberg 2000) — the same number of translocations found among populations of *H. anomalus*. This scenario would further imply that the polymorphic population, ANO1282, likely represents a zone of secondary contact between the northern origin of *H. anomalus* and one of the two southern origins rather than the 'ancestral population.' Given the high mutation rate of microsatellite loci, mutation and gene flow through pollen could easily account for the lack of a clear signature of independent origins from the microsatellite data. The apparently ancient origins of *H. anomalus* (> 100 000 years bp) and modest N_em values (0.88) accord well with this explanation.

In conclusion, a recurrent origin of *H. anomalus* on three separate occasions provides, in our view, the most parsimonious explanation for the complex geographical patterning of cpDNA, nuclear microsatellites and crossability/chromosome variation. We are aware, of course, of the difficulties associated with the reconstruction of past evolutionary events and our conclusions should be viewed as tentative and worthy of further testing. However, these conclusions are consistent with earlier experimental work (Rieseberg *et al*. 1996; Rieseberg 2000) that demonstrated the ease with which the genome of this taxon could be generated in the greenhouse.

Dating the origins of H. anomalus

In a recent book review, Schemske (2000) argued that hybridization was unlikely to have played a significant role in the evolution of *Helianthus* because hybridization in the group is often associated with disturbance by humans. However, the divergence times reported here for populations of *H. anomalus* (116 000–160 000 years bp) imply that hybridization was an important evolutionary process in this group well before humans colonized North America. This assertion is further supported by divergence times for another hybrid *Helianthus* species, *H. paradoxus*, which also predates the colonization of North America by humans (Welch & Rieseberg unpublished). This molecular evidence for ancient hybridization between *H. annuus* and *H. petiolaris* accords well with field observations. Although many present-day hybrid swarms are associated with human disturbance, hybrid swarms often occur in naturally disturbed habitats such as flood plains or buffalo wallows (L. H. Rieseberg unpublished). Moreover, bison, which appear to have been the primary dispersal agent for annual sunflowers (Asch 1993), likely created ample disturbed habit for hybrid formation long before humans colonized North America.

Repeatability of diploid hybrid speciation

To date, the possibility of recurrent diploid hybrid speciation has been tested in only three species: *Argyranthemum sundingii* (Brochmann *et al*. 2000), *Pinus densata* (Wang *et al*. 2001) and *Helianthus anomalus*. All three examples are consistent with recurrent speciation, although uncertainties remain. For example, introgression with the parental species following hybrid speciation cannot be ruled out in either the *Argyranthemum* or *Pinus* examples. Nonetheless, the results from these studies, combined with experimental data (Rieseberg *et al*. 1996; Rieseberg 2000), suggest that diploid hybrid speciation should be included with ecological speciation and allopolyploidy as modes conducive to recurrent speciation.

There are differences, of course, in overall repeatability among these modes of speciation. Allopolyploid speciation is most repeatable because it involves a single macromutation (genome doubling) that arises frequently (Ramsey & Schemske 1998) and generates essentially identical products. Thus, it is perhaps not surprising that recurrent allopolyploid speciation appears to be the rule rather than the exception (Soltis & Soltis 1993). Recurrent ecological speciation is considerably rarer and more difficult to interpret. Not only are similar selection pressures required to produce ecological species recurrently, but also at least some fraction of the changes must have a similar genotypic basis. Otherwise, uniform selection will lead to divergent rather than parallel speciation (Wade & Goodnight 1998; Levin 2000).

Diploid hybrid speciation offers an interesting parallel to both allopolyploidy and ecological speciation. Notwithstanding the three putative examples of recurrent diploid hybrid speciation, it is clear that unlike allopolyploidy, multiple, different diploid hybrid species can be derived from the same cross. In *Helianthus*, for example, three different diploid hybrid species (*H. anomalus*, *H. deserticola* and *H. paradoxus*) are derived from the same parental species, *H. annuus* and *H. petiolaris*. The three hybrid

2

sunflower species occur in highly divergent habitats: *H. anomalus*, sand dunes; *H. deserticola*, the desert floor; and *H. paradoxus*, brackish salt marshes (Rogers *et al*. 1982). Thus, as with ecological speciation, the recurrent origin of hybrid sunflower species depends in part on uniform ecological selection. That is, *H. anomalus* is likely to arise only if a hybrid founder population happens to become established on or adjacent to sand dune habitats. If the hybrids happen to colonize desert floor or brackish marsh habitats, a different hybrid species may arise. Thus, both determinism and contingency interact to mediate the outcome of diploid hybrid speciation.

Acknowledgements

We are grateful to Keith Gardner, Briana Gross, Olivier Raymond and Mark Welch for assistance with various aspects of the data analyses. We also thank John Burke, Briana Gross, Kevin Livingstone and Mark Welch for helpful discussions and for comments on an earlier version of the manuscript. This work was supported by grants from the National Science Foundation (DEB9806290) and National Institutes of Health (GM59065) to L.H.R.

References

Asch DL (1993) Common sunflower (*Helianthus annuus* L.): the pathway toward domestication. *Proceedings of the 58th Annual Meeting of the Society for American Archaeology*, pp. 1–15. Society for American Archaeology, St. Louis, MO.

Avise JC (1994) *Molecular Markers, Natural History, and Evolution*. Chapman & Hall, New York.

Barton NH, Slatkin M (1986) A quasi-equilibrium theory of the distribution of rare alleles in a subdivided population. *Heredity*, **56**, 409–415.

Brochmann C, Borgen L, Stabbetorp OE (2000) Multiple diploid hybrid speciation of the Canary Island endemic *Argyranthemum sundingii* (Asteraceae). *Plant Systematics and Evolution*, **220**, 77–92.

Casgrain P, Legendre P (2000) *The R Package for Multivariate and Spatial Analysis, Version 4.0. (Development Release 2) User's Manual*. Université of Montréal, Quebec.

Chandler JM, Jan C, Beard BH (1986) Chromosomal differentiation among the annual *Helianthus* species. *Systematic Botany*, **11**, 353–371.

Demesure B, Sozi N, Petit RJ (1995) A set of universal primers for amplification of polymorphic non-coding regions of mitochondrial and chloroplast DNA in plants. *Molecular Ecology*, **4**, 129–131.

Diwan N, Cregan PB (1997) Automated sizing of fluorescent-labeled simple sequence repeat (SSR) markers to assay genetic variation in soybean. *Theoretical and Applied Genetics*, **95**, 723–733.

Dorado O, Rieseberg LH, Arias D (1992) Chloroplast DNA introgression in southern California sunflowers. *Evolution*, **46**, 566–572.

Dumolin-Lapegue S, Pemonge M-H, Petit RJ (1997) An enlarged set of consensus primers for the study of organelle DNA in plants. *Molecular Ecology*, **6**, 393–397.

Goldstein DB, Linares AR, Cavalli-Sforza LL, Feldman MW (1995) An evaluation of genetic distances for use with microsatellite loci. *Genetics*, **139**, 463–471.

Goodwillie C (1999) Multiple origins of self-compatibility in *Linanthus* section *Leptosiphon* (Polemoniaceae): phylogenetic evidence from the internal-transcribed spacer sequence data. *Evolution*, **53**, 1387–1395.

Grant V (1958) The regulation of recombination in plants. *Cold Spring Harbor Symposium in Quantitative Biology*, **23**, 337–363.

Grant V (1981) *Plant Speciation*. Columbia University Press, New York.

Heiser CB (1947) Hybridization between the 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.

Levin DA (2000) *The Origin, Expansion, and Demise of Plant Species*. Oxford University Press, New York.

Liston A (1992) Variation in the chloroplast genes *rpo*C1 and *rpo*C2 of the genus *Astragalus* (Fabaceae) — evidence from restriction site mapping of a PCR-amplified fragment. *American Journal of Botany*, **79**, 953–961.

Nabhan GP, Reichhardt KL (1983) Hopi protection of *Helianthus anomalus*, a rare sunflower. *Southwestern Naturalist*, **28**, 231–235.

Quillet MC, Madjidian N, Griveau T *et al.* (1995) Mapping genetic factors controlling pollen viability in an interspecific cross in *Helianthus* section *Helianthus*. *Theoretical and Applied Genetics*, **91**, 1195–1202.

Ramsey J, Schemske DW (1998) Pathways, mechanisms, and rates of polyploid formation. *Annual Review of Ecology and Systematics*, **19**, 467–502.

Raymond M, Rosset F (1995) GENEPOP Version 1.2: population genetics software for exact test and ecumenicism. *Journal of Heredity*, **86**, 248–249.

Rice WR (1989) Analyzing tables of statistical tests. *Evolution*, **43**, 223–225.

Rieseberg LH (1991) Homoploid reticulate evolution in *Helianthus*: 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 (2000) Crossing relationships among ancient and experimental sunflower hybrid lineages. *Evolution*, **54**, 859–865.

Rieseberg LH (2001) Chromosomal rearrangements and speciation. *Trends in Ecology and Evolution*, **16**, 351–358.

Rieseberg LH, Baird S, Desrochers A (1998) Patterns of mating in wild sunflower hybrid zones. *Evolution*, **52**, 713–726.

Rieseberg LH, Beckstrom-Sternberg SM, Doan K (1990) *Helianthus annuus* ssp. *texanus* has chloroplast DNA and nuclear ribosomal RNA genes of *Helianthus debilis* spp. *cucumerifolius*. *Proceedings of the National Academy of Sciences of the USA*, **87**, 593–597.

Rieseberg LH, Beckstrom-Sternberg SM, Liston A, Arias DM (1991) Phylogenetic and systematic inferences from chloroplast DNA and isozyme variation in *Helianthus* sect. *Helianthus* (Asteraceae). *Systematic Botany*, **16**, 50–76.

Rieseberg LH, Desrochers A, Youn SJ (1995a) Interspecific pollen competition as a reproductive barrier between sympatric species of *Helianthus* (Asteraceae). *American Journal of Botany*, **82**, 515–519.

Rieseberg LH, Sinervo B, Linder CR, Ungerer M, 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 A (1995b) Hybrid

speciation accompanied by genomic reorganization in wild sunflowers. *Nature*, **375**, 313–316.
Rieseberg LH, Whitton J, Gardner K (1999) Hybrid zones and the genetic architecture of a barrier to gene flow between two wild sunflower species. *Genetics*, **132**, 713–727.
Rogers CE, Thompson TE, Seiler GJ (1982) *Sunflower Species of the United States*. The National Sunflower Association, Fargo, ND.
Rundle HD, Nagel L, Boughman JW, Schluter D (2000) Natural selection and parallel speciation in sympatric sticklebacks. *Science*, **287**, 306–308.
Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*, **4**, 406–426.
Schemske DW (2000) Understanding the origin of species. *Evolution*, **54**, 1069–1073.
Schluter D (2000) Ecology and the origin of species. *Trends in Ecology and Evolution*, **16**, 372–380.
Schluter D, Nagel LM (1995) Parallel speciation by natural selection. *American Naturalist*, **146**, 292–301.
Schneider S, Kueffer JM, Roessli D, Excoffier L (1997) *ARLEQUIN 1.1*. University of Geneva, Switzerland.
Schwarzbach AE, Donovan LA, Rieseberg LH (2001) Transgressive character expression in a hybrid sunflower species. *American Journal of Botany*, **88**, 270–277.
Soltis PS, Plunkett GM, Novak SJ, Soltis DE (1995) Genetic variation in *Tragopogon* species: additional origins of the allotetraploids *T. mirus* and *T. miscellus* (Compositae). *American Journal of Botany*, **82**, 1329–1341.
Soltis DE, Soltis PS (1993) Molecular data and the dynamic nature of polyploidy. *Critical Reviews in Plant Sciences*, **12**, 243–275.
Stebbins GL (1950) *Variation and Evolution in Plants*. Columbia University Press, New York.
Strid A (1970) Studies in the Aegean flora. XVI. Biosystematics of the *Nigella arvensis* complex with special reference to the problem of non-adaptive radiation. *Opera Botanica*, **33**, 1–118.
Swofford D (2000) *PAUP**. *Phylogenetic Analysis Using Parsimony, Version 4.0*. Sinauer Associates, Sunderland, MA.
Tang S, Yu J-K, Slabaugh MB, Shintani DK, Knapp SJ (2002) Simple sequence repeat map of the sunflower genome. *Theoretical and Applied Genetics*, in press.
Taylor EB, McPhail JD (2000) Historical contingency and ecological determinism interact to prime speciation in sticklebacks *Gasterosteus*. *Proceedings of the Royal Society of London, Series B — Biological Sciences*, **267**, 2375–2384.
Ungerer MC, Baird S, Pan J, Rieseberg LH (1998) Rapid hybrid speciation in wild sunflowers. *Proceedings of the National Academy of Sciences of the USA*, **95**, 11757–11762.
Wade MJ, Goodnight CJ (1998) The theories of Fisher and Wright in the context of metapopulations: when nature does many small experiments. *Evolution*, **52**, 1537–1553.
Wang X-R, Szmidt AE, Savolainen O (2001) Genetic composition and diploid hybrid speciation of a high mountain pine, *Pinus densata*, native to the Tibetan plateau. *Genetics*, **159**, 337–346.
Zhivotovsky LA (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.

The laboratory of Loren Rieseberg studies the genetics of speciation in plants. This study, which was part of Andrea Schwarzbach's postdoctoral research, contributes to an ongoing research programme on the evolutionary outcomes of hybridization in the North American sunflowers.

Author Query Form

Journal: Molecular Ecology

Article: 1557

Dear Author,

During the copy-editing of your paper, the following queries arose. Please respond to these by marking up your proofs with the necessary changes/additions. Please write your answers on the query sheet if there is insufficient space on the page proofs. Please write clearly and follow the conventions shown on the attached corrections sheet. If returning the proof by fax do not write too close to the paper's edge. Please remember that illegible mark-ups may delay publication.

Many thanks for your assistance.

Query Refs.	Query	Remarks
1	**Schluter 2001** has been changed to **Schluter 2000** so that this citation matches the list, please check the date.	
2	**Rogers et al. 1981** has been changed to **Rogers et al. 1982** so that this citation matches the list. Please check the date	
3	**Rundle et al. 2000** has not been found in the text. Please add a citation or delete from this list.	

MARKED PROOF

Please correct and return this set

Any errors in this proof which have been noticed by the printer's reader have been marked in green. If you see any more printer's errors, please mark them in red: there is no charge for correcting these mistakes. For your own alterations, please use black or blue or any colour other than green or red. Please use the proof correction marks shown below for all alterations and corrections.

Instruction to printer	*Textual mark*	*Marginal mark*
Leave unchanged	··· under matter to remain	Stet
Insert in text the matter indicated in the margin		New matter followed by 
Delete	 through matter to be deleted	
Delete and close up	 through matter to be deleted	
Substitute character or substitute part of one or more word(s)	/ through letter or  through word	New letter or new word
Change to italics	— under matter to be changed	
Change to capitals	 under matter to be changed	
Change to small capitals	under matter to be changed	
Change to bold type	under matter to be changed	
Change to bold italic	under matter to be changed	
Change to lower case	Encircle matter to be changed	
Change italic to upright type	(As above)	
Insert 'superior' character	/ through character or  where required	 under character e.g. 
Insert 'inferior' character	(As above)	 over character e.g. 
Insert full stop	(As above)	
Insert comma	(As above)	
Insert single quotation marks	(As above)	 and/or
Insert double quotation marks	(As above)	and/or 
Insert hyphen	(As above)	
Start new paragraph		
No new paragraph		
Transpose		
Close up	linking ⌒ letters	
Insert space between letters	 between letters affected	#
Insert space between words	 between words affected	#
Reduce space between letters	 between letters affected	
Reduce space between words	between words affected	