Int. J. Plant Sci. 175(6):688–701. 2014.
Copyright is not claimed for this article.
DOI: 10.1086/676305

EFFECT OF A GEOGRAPHIC BARRIER ON ADAPTATION IN THE DWARF SUNFLOWER (*HELIANTHUS PUMILUS* NUTT.)

Patrick A. Reeves[1,*] and Christopher M. Richards*

*USDA-ARS, National Center for Genetic Resources Preservation, 1111 South Mason Street, Fort Collins, Colorado 80521, USA

Editor: Pamela Soltis

Premise of research. Geographic isolation promotes local adaptation but may also prevent the movement of advantageous gene complexes throughout a species distribution. We identify a geographic barrier that inhibits gene flow and adaptation across the disjunct species range of the Rocky Mountain endemic sunflower *Helianthus pumilus*.

Methodology. Habitat suitability modeling and least-cost path analysis were used to characterize environmental preferences across the species range and to identify barriers to gene flow. Simple sequence repeat data were used to assess genetic structure and estimate recent levels of gene flow across the distributional disjunction. A common garden was used to study phenotypic differentiation and local adaptation. Controlled crosses tested for intrinsic barriers to reproduction.

Pivotal results. Despite significant differences in available habitat, individuals in northern and southern regions occupy similar sites, suggesting limited adaptive differentiation in environmental preference. A region of low-suitability habitat separates the two regions. Northern and southern plants are genetically distinct. There has been little recent gene flow. Southern plants had higher fitness (measured as seed set) in the northern region than northern plants. A suite of morphological characters, present in half of southern plants, conveyed a twofold fitness advantage in the north, but this adaptive phenotype is virtually nonexistent in the northern range. Between-region crosses were equal to or more productive than within-region crosses.

Conclusions. The Pikes Peak massif, a high-altitude eastward projection of the Rocky Mountains, forms a geographic barrier to gene flow in *H. pumilus*, a lowland foothills species. No intrinsic barriers to gene exchange between regions have yet evolved. Therefore, the geographic barrier appears to impede the movement of the adaptive phenotype found in the southern region, resulting in a northern group that is less fit than it might be, contrary to the expectation of local adaptation.

Keywords: fitness, gene flow, niche model, population structure, reproductive isolation, common garden.

Online enhancement: appendix.

Introduction

Geographic isolation drives biological diversification. Restrictions in gene flow between populations caused by simple spatial disjunction are ultimately responsible for many of the species we observe today (Jordan 1905; Mayr 1963; Barraclough and Vogler 2000; Coyne and Orr 2004, pp. 83–124, 168–178; Fitzpatrick and Turelli 2006). Excluding polyploid speciation, a few cases of sympatric speciation have been documented in plants (Gottlieb 1973; Savolainen et al. 2006; Papadopulos et al. 2011), but most botanists would agree that allopatric speciation appears to be the dominant mode by which plants diversify.

Geographic isolation promotes diversification by functioning as a powerful, premating, extrinsic reproductive isolating mechanism (Wiens 2004). Given that dispersal distances for seeds and pollen are usually limited (by pollination vectors, dispersal agents, or physical processes), spatial disjunction reduces the flow of genes between populations in proportion to the magnitude of the distance between them. Geographic isolation, over time, leads to the serendipitous accumulation of other—intrinsic—isolating mechanisms, which evolve as a byproduct of adaptation to distinct environments or via genetic drift, thereby promoting the formation of new species (Dobzhansky 1951, pp. 206–211; Mayr 1959; Schluter 2001; Coyne and Orr 2004). While spatial discontinuity alone, independent of any intrinsic mechanism, might cause complete reproductive isolation, one must also consider the possibility that the spatial discontinuity itself is a consequence of genetic differences—adaptations—that preclude a continuous distribution (Stebbins 1950, pp. 197–199; Schemske 2000). Sobel et al. (2010) recognized this conundrum and proposed two definitions that echo Stebbins (1950, p. 196): "effective geographic isolation" describes the contribution of both extrinsic

[1] Author for correspondence; e-mail: pat.reeves@ars.usda.gov.

Manuscript received June 2013; revised manuscript received December 2013; electronically published May 27, 2014.

and intrinsic factors to reductions in gene flow between allopatric populations, whereas "ecogeographic isolation," on the other hand, describes the contribution of intrinsic factors alone, i.e., adaptations that promote geographic isolation. For convenience, we utilize a third term from Stebbins (1950), "spatial isolation," which refers to the reduction in gene flow attributable solely to the physical distance between groups, i.e., to dispersal limitation. Spatial isolation plus ecogeographic isolation equals effective geographic isolation.

The contribution of geographic isolation relative to other isolating mechanisms has rarely been quantified, but where it has it appears to be important. When pollinator isolation was suspected a priori to be the predominant mechanism maintaining species distinctness between sympatric species pairs, two studies found that geography (technically, effective geographic isolation) was, in fact, an equivalent or stronger isolating force, accounting for 35%–60% of total reproductive isolation (Ramsey et al. 2003; Kay 2006). In a diploid/tetraploid species pair, Husband and Sabara (2004) found that geographic isolation was more important than triploid block, a strong isolating mechanism caused by ploidy incompatibilities in the backcross generation (Ramsey and Schemske 1998). This finding is remarkable because polyploid speciation is often cited as the most plausible form of sympatric speciation in plants (Otto and Whitton 2000), yet geographic isolation may still function as the critical (and largely unappreciated) factor allowing divergence.

Effective geographic isolation can be estimated using dispersal distance and range overlap (Ramsey et al. 2003; Kay 2006). Best practices for teasing apart, in a quantitative sense, the contribution of spatial and ecogeographic components of effective geographic isolation have yet to be developed, but Sobel et al. (2010) present experimental recommendations to help determine which form predominates. First, reciprocal transplants can be used to estimate the relative fitness of two disjunct populations in sympatry. If local genotypes perform better than foreign genotypes, local adaptation has occurred and ecogeographic isolation is indicated. If local and foreign genotypes perform equivalently or if foreign genotypes outperform locals, ecogeographic isolation should be rejected. Second, habitat suitability modeling (Guisan and Zimmermann 2000; Soberón and Peterson 2005; Elith et al. 2006; Kozak et al. 2008) can be used to estimate the degree to which the habitat occupied by two populations differs. If the populations occupy similar or identical habitat, ecogeographic isolation is probably not strong, whereas if occupied habitat differs, intrinsic differences in ecological preference are possible. The latter conclusion is complicated by the fact that environmental characteristics, which define "habitat" in most modeling exercises, change in a correlated manner across space such that any two populations occupying different geographic regions will likely be inferred to exist in distinct "habitats" (Legendre 1993; Elith and Leathwick 2009). Glor and Warren (2011) devised a statistical test to control for this spatial autocorrelation in environmental variables. The background similarity test allows one to determine whether two populations occupy significantly different habitat, given that the habitat available to them in their respective regions differs.

Geographic barriers to gene flow would seem to lie at the root of most allopatric differentiation processes, so pinpointing their location and describing their consequences in populations, as well as farther along the speciation continuum, is critical for fully understanding plant diversification. For some well-collected species, the location of a potential barrier can be found using maps, as a geographic gap in the species distribution. If a geographic gap can be shown to contain suitable habitat, either by transplant experiment or by modeling, it may form a contemporary barrier due to dispersal limitations but is unlikely to remain as such over evolutionary time periods, as intervening suitable habitat is colonized and a conduit for gene flow is established. If, on the other hand, intervening habitat is unsuitable, the gap may restrict gene flow by preventing colonization. The latter type of gap is more likely to function as a barrier over evolutionary time because gene flow across it requires infrequent, long-distance dispersal events. Finally, if a gap can be associated with a restriction in gene flow, it may then be considered a barrier. Two conditions are necessary. First, genetic differentiation should be demonstrable across the barrier. This can be tested using molecular marker data (Pearse and Crandall 2004) or via traditional morphological studies. Second, the barrier itself must be responsible for the genetic differentiation, as opposed to migration from already-differentiated source populations (e.g., distinct glacial refugia). To satisfy this condition, one must demonstrate that habitat on both sides of the barrier has been continuously occupied and that the barrier region has been unoccupied, or unsuitable, for an evolutionarily relevant period of time. This will generally be difficult, although for a minority of species fossil or palynological data may be available to help prove the case. For most species, paleoclimatic habitat suitability modeling seems to offer a reasonable alternative (Hugall et al. 2002; Carstens and Richards 2007).

Geographic barriers to gene flow are generally thought of as a creative force in evolution, one that promotes adaptive differentiation (Slatkin 1987; Garant et al 2007) by disrupting maladaptive gene flow from other, ecologically distinct regions (Dobzhansky 1940; Kirkpatrick and Barton 1997; Lenormand 2002; Coyne and Orr 2004). Localized, adaptively differentiated populations result and are an important step toward new species formation. But barriers might also impede adaptation—species-wide—by restricting the movement of novel adaptive alleles or coadapted gene complexes between disjunct populations. Evolutionary innovations have to arise somewhere, and in a fragmented distribution they may not be able to permeate the species range prior to local extinction or the evolution of reproductive isolation. The latter phenomenon, where the lack of gene flow impedes adaptation, has been addressed in the conservation literature (Templeton et al. 2001) but is less frequently described in evolutionary studies, possibly because it might be uncommon: local adaptation is pervasive (Linhart and Grant 1996; Savolainen et al. 2007), and the spread of advantageous alleles between populations is highly efficient (Slatkin 1976; Morjan and Rieseberg 2004).

In plants, the extent to which a geographic barrier acts as a creative or constraining force can be examined empirically. Outperformance of local over foreign genotypes in a common garden—the signature of local adaptation—implies a creative influence of a barrier on the species: gene flow has been cut off, and adaptive differentiation has proceeded. The opposite result, where foreign genotypes perform better than locals (i.e.,

local maladaptation), suggests a negative effect. In this case, the barrier acts as a constraint by blocking access to adaptive alleles or gene complexes, thereby limiting the evolutionary potential of isolated populations. Since biologically based reproductive isolating mechanisms may accumulate between geographically isolated groups over time, it is necessary to establish that any restriction in gene flow is due to the barrier alone and not other factors. This can be accomplished by examining the contribution of common isolating mechanisms (e.g., temporal isolation, pollinator isolation, and hybrid inviability) to reductions in gene flow between disjunct groups (following Ramsey et al. [2003], Husband and Sabara [2004], and Kay [2006]).

In this study, we identify and characterize a geographic barrier to gene flow in *Helianthus pumilus*, the dwarf sunflower, using habitat suitability modeling. We test whether genetic differentiation is evident between plants from disjunct portions of the range using neutral molecular markers and morphological data. We examine evidence for local adaptive differentiation and the evolution of intrinsic reproductive isolation by measuring fitness components in a common garden and conducting experimental crosses. We discuss the nature of geographic isolation in this species and consider whether it impedes or promotes adaptation.

Material and Methods

Study System

Helianthus pumilus Nutt. is a bisexual, outcrossing, diploid perennial sunflower species common in the Front Range foothills and Laramie Mountains at the eastern edge of the Rocky Mountains, within the US states of Colorado and Wyoming. Preferred habitat is arid and includes sandy or rocky slopes and disturbed areas. Beginning on jumbled granite and gneiss hillsides in the Sangre de Cristo Mountains at the southern end of the range, the species distribution gradually declines in altitude from south to north to include sedimentary hogbacks fronting the foothills in the central range and finally, at the northern range limits, to isolated mineral outcrops and intermittent stream beds in the abutting Great Plains.

Taxonomically, *H. pumilus* belongs to *Helianthus* series *Pumilii*, a clade of three perennial species from the western United States that occupies a position in the *Helianthus* phylogeny as sister to section *Helianthus*, a well-studied group of annual sunflowers containing the domesticated species *H. annuus* (Schilling and Heiser 1981; Timme et al. 2007). Improvement of the cultivated sunflower relies on introgression of valuable traits from wild, related species (Dorrell and Whelan 1978). *Helianthus pumilis* is of agricultural interest because of its pathogen resistance and its capacity to produce seeds with high linoleic fatty acid content under conditions that, in agricultural terms, would be considered extreme drought (Seiler 1994; Nikolova et al. 2004). Barriers to hybridization with the cultivar are strong but incomplete, and fertile offspring, though not documented in the wild, have been produced under controlled conditions (Nikolova et al. 2004). *Helianthus pumilus* is a convenient system to study the effect of barriers to gene flow on adaptation because of an obvious distributional disjunction, easy access to the entire species range, free availability of seeds in the US National Plant Germplasm System from comprehensive sampling expeditions in the wild, and transferable molecular markers from the cultivar.

Population Structure

DNA samples were collected in 2009–2010 from five to eight individuals in 29 populations distributed across the species range (fig. 1; table A1). Genotypic data were obtained from 16 expressed sequence tag–based simple sequence repeat (SSR) loci (Pashley et al. 2006; Chapman et al. 2008) and one anonymous SSR locus (Tang et al. 2003). A three-primer PCR labeling procedure was used, followed by visualization on a LICOR 2400 DNA Sequencer (Oetting et al. 1995) or an Applied Biosystems 3100 Genetic Analyzer (Schuelke 2000). Scoring used SAGA software (ver. 3.3; LI-COR) or P2EMA (http://sourceforge.net/projects/p2ema/); binning used Flexibin (Amos et al. 2007). Genotyping details are provided in table B1 (tables B1–B8 are available online). Descriptive statistics, including heterozygosity, null allele frequency, and deviations from Hardy-Weinberg equilibrium were calculated using GDA

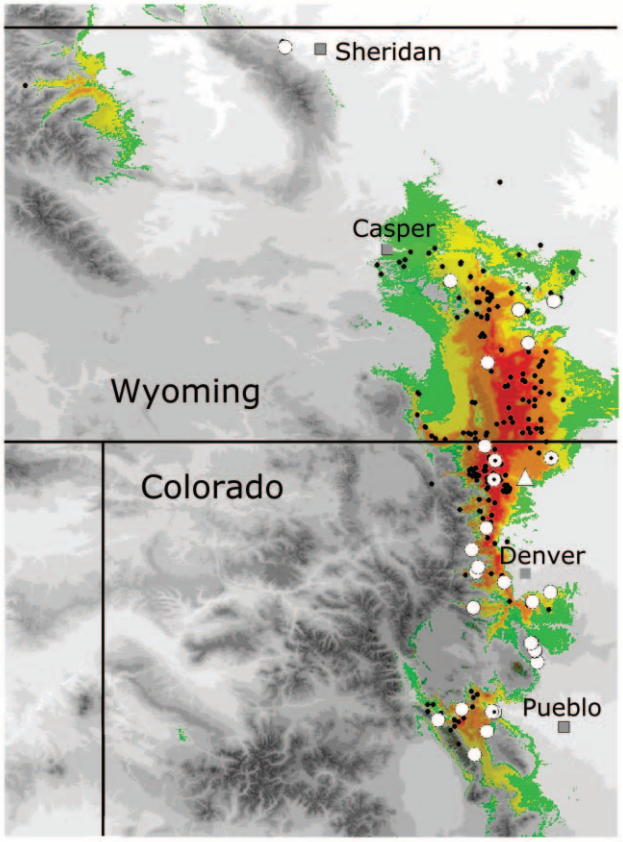

Fig. 1 Distribution of *Helianthus pumilus* in the western United States, with habitat suitability model overlaid. Warmer colors indicate more suitable habitat (values <0.17 are transparent; increment = 0.17). Small black dots mark known *H. pumilus* populations. Large open circles indicate populations sampled for DNA and grown in a common garden (with bull's-eye, absent from common garden). Triangle = common-garden site.

(Lewis and Zaykin 2001) and MICRO-CHECKER (Van Oosterhout et al. 2004).

The number of genetically distinct clusters was estimated using the software STRUCTURAMA (Huelsenbeck and Andolfatto 2007). STRUCTURAMA uses the approach of Pritchard et al. (2000) to calculate the posterior probability of assignment of individual genotypes to clusters. It also allows K, the number of populations, to be treated as a random variable; hence, STRUCTURAMA can estimate the number of discrete populations present in a data set. One hundred replicate analyses were performed. Admixture was not permitted. The concentration parameter prior (α) followed a gamma distribution with shape and scale parameters assigned at random from 0 to 10. A single Markov chain was run for 1×10^6 generations, with samples taken every 500. The initial 1000 samples were discarded. The most likely number of populations was inferred from the mean probability of each K value calculated across replicate runs.

STRUCTURE (Pritchard et al. 2000) was used to estimate admixture among populations. Twenty replicate runs were performed using the value for K determined by STRUCTURAMA. After 1×10^5 generations of burn-in, 1×10^5 generations were sampled. Admixture was permitted, and allele frequencies were assumed to be correlated. Average membership coefficients were calculated across replicate runs using CLUMPP (Jakobsson and Rosenberg 2007).

BA3 (Wilson and Rannala 2003) was used to estimate migration rates between northern and southern regions. Two independent analyses were run for 2×10^7 generations, with the first 2×10^6 generations discarded as burn-in. Tuning parameters for allele frequencies, inbreeding coefficients, and migration rates were set to 0.3, 0.3, and 0.08, respectively, to ensure proper mixing of the chain. Acceptance rates ranged from 20% to 35% using these settings. The chains were sampled every 5000 generations after burn-in. Convergence was assessed by comparing allele frequency estimates between the two independent runs. Parameters of interest included population migrant proportion and individual immigrant ancestry. The former indicates the level of migration between regions, and the latter estimates the probability that sampled individuals are first- or second-generation migrants. Together these parameter estimates were used to establish the frequency and direction of migration in the recent past.

Habitat Suitability Modeling

Habitat suitability across the study area was modeled using Maxent (ver. 3.3.3a; Phillips et al. 2006; Phillips and Dudík 2008). A total of 239 localities spanning the entire geographic range were used. Localities were obtained from herbarium records (specimens held at CS and RM) and the USDA Germplasm Resources Information Network (http://www.ars-grin.gov/npgs/). Nineteen biologically relevant climatic layers derived from current temperature and moisture data (the BioClim data layers; WorldClim database, http://www.worldclim.org), plus altitude, were provided to Maxent as 30″ grids cropped to the species range. Default settings were accepted for all parameters except that 20 replicate models were generated using the "Subsample" option, with 25% of samples randomly set aside as test localities. A habitat suitability model was computed as the average across replicates. Model validation included an examination of omission rates on test localities and application of a binomial test to determine whether the model predicted presence at test sites significantly better than random (Anderson et al. 2002; Phillips et al. 2006).

Models were also computed for sites partitioned into two groups according to the geographic disjunction. In this case, 213 sites belonged to a northern group, and 26 sites belonged to a southern group. The background similarity test and two range-breaking statistical tests were applied to explore the potential for local adaptive differentiation between groups (Warren et al. 2008, 2010; Glor and Warren 2011). Background mask regions (from which random sites are sampled) were defined using the 10 percentile training presence threshold as a cutoff, to ensure that only highly suitable habitat was considered when generating null distributions. Both blob and linear range-breaking tests were conducted. All unique bipartitions of the data were considered. This was accomplished by generating 10,000 random samples using the blob and linear models of data bisection. A filtered set containing only unique bipartitions was used to generate null distributions (linear range break test, 142 unique data sets; blob, 36).

Paleoclimatic modeling utilized global 2.5′ environmental layers reconstructed under the Community Climate System Model to approximate conditions at ~21,000 yr before present, available from WorldClim. The current-conditions Maxent model was projected onto the 19 paleolayers to predict habitat suitability at the last glacial maximum. Regions of poor prediction that arise due to novel climate conditions in the projection data were identified following Elith et al. (2010). This included inspection of Maxent output grids describing regions affected by clamping (a modeling feature that may introduce out-of-range errors), as well as exploration of the multivariate environmental similarity surface, which highlights regions of environmental novelty in the reconstructed layers ("no-analog climates") relative to current conditions.

Barriers to Migration

To model dispersal, we determined least-cost paths between sampled populations across the habitat suitability model (Wang et al. 2008, 2013). SAGA-GIS (ver. 2.0.5) was used for grid manipulations (available at http://www.saga-gis.org). A cost surface ("resistance matrix") was calculated as 1 minus the mean habitat suitability grid from Maxent. To determine the least-cost path between two populations, one site was set as the destination, and an accumulated cost grid was calculated using the module "Accumulated Cost (Isotropic)." The accumulated cost grid describes, for every cell, the cumulative cost to move from that cell to the destination point as measured along the cost surface. The least-cost path was then calculated by defining an origin and using a valley-seeking algorithm to "drain" to the destination point. The specific path chosen is that which has the lowest possible accumulated cost at each step between origin and destination. Least-cost paths were calculated for all pairwise combinations of sampled populations. The total cost of migration between pairs was calculated as the sum of values along the least-cost path across the cost surface. This value measures the cost due to geographic distance (longer routes have more steps) as well as difficulty, in

terms of habitat suitability, of traversing the path. It is an estimate of effective geographic isolation, hereafter referred to as "effective geographic distance" (EGD).

Barriers to gene flow were visualized by plotting the value present at each step along the least-cost path in the cost surface. Sites with low-habitat suitability have high values on the cost surface; thus, peaks in the trace can be used to identify regions with low probability of occupation but that must be crossed for gene flow to occur. A matrix of EGDs was assembled for all pairwise combinations of sampled sites. Mantel tests (Bonnet and Van de Peer 2002) were performed to examine whether EGD distances better predicted genetic divergence among populations (Nei 1978) than straight-line distances.

Common Garden

Seeds from 24 accessions from the US National Plant Germplasm System were used to produce seedlings for a common-garden experiment (accessions, table A1; germination conditions, table B2). Seeds were collected in 2005 (with one exception from 1991) by sampling seed heads from 10–100 individuals per population and then pooling achenes into a single bulked sample intended to represent the diversity present in the wild population (Seiler et al. 2012). In 2010, a common garden was established at the periphery of *H. pumilus* native range in suitable habitat (fig. 1; Maxent logistic suitability = 0.314; suitability of occupied sites in wild = 0.195–0.841). The site was located within an agricultural field brought into production in the 1940s; thus, some (although not extensive) modification of natural soil properties should be expected. Soil type was Otero sandy loam; grade was 0%–3%. Three randomized blocks, each containing one row per accession with a maximum of 10 plants per row spaced on 60-cm centers with 80 cm between rows, were constructed. Immediately following transplantation, plants were irrigated once to promote establishment, then left to grow under prevailing environmental conditions without supplemental water or fertilizer. Of 599 seedlings, 454 survived transplantation (76%). In 2011, plants were thinned to 120-cm centers; 235 individuals remained (per population sample sizes, table A1). Plots were hand-weeded throughout the growing season to eliminate nonuniform competition across the experiment. Plants were open-pollinated.

Flowering time and seed set were measured in 2010 and 2011. Flowering phenology was studied by making weekly counts of the number of active inflorescences on all individuals. Active inflorescences were defined as containing both ray flowers with unsenesced ligules plus disk flowers with functional stamens and/or receptive pistils. Total seed set, an indicator of potential fitness, was calculated as the mean number of seeds per head (four heads counted per individual) multiplied by the total number of heads produced (counted at the end of the growing season). A set of morphological traits was measured on all individuals (2010 only), including disk diameter, inflorescence diameter (radius between ligule tips), and ray flower number (all expressed as the average of three inflorescences); leaf length, width, and thickness (average of three largest leaves); and branchiness (number of lateral branches longer than 5 cm) and plant height, both scored at the R5 stage of development, which marks the commencement of flowering (Schneiter and Miller 1981).

Crossing Study

Two crossing studies were undertaken to test for intrinsic reproductive isolation between northern and southern groups. The first study (2011) used hand-pollination in the common garden. Three classes of crosses were performed: self-fertilization, within region, and between region. For between-region crosses, 14 individuals in the common garden were used as parents, seven from each geographic region. Mating pairs were assigned randomly; southern plants acted as pollen donor to a northern recipient. To perform within-region crosses, additional individuals were chosen randomly and paired by region with the individuals already chosen for between-region crosses. Individuals acting as pollen recipients were also used for self-fertilization experiments. Four to 7 d before hand-pollinations, immature, unopened inflorescences at the R3 stage were isolated from pollinators using glassine envelopes. Pollen was collected at daybreak from bagged inflorescences on pollen donors and then applied to receptive stigmas on the recipient plants by means of cotton swabs. Four inflorescences per plant were pollinated and then resealed in labeled envelopes. Self-pollinations proceeded similarly, except that pollen from the same individual was used. All pollinations were completed within 3 h of sunrise. After ripening and drying, inflorescences were collected and seeds were counted.

To better reflect natural pollination processes, the second crossing study was conducted in isolation plots established along the Front Range within predicted suitable habitat. Plants were allowed to open-pollinate throughout the growing season using local pollinators. The minimum distance between sites was 1.7 km, and the shortest distance to any indigenous *H. pumilus* was 0.6 km. Plots testing self-fertilization contained one plant and were used to establish a baseline level of seed production in the wild due to selfing plus, potentially, pollen flow from distant populations. This measurement was necessary to confirm that the majority of seed produced in crosses was derived from the intended pollen donor. Four self-fertilization experiments contained individuals from the south; five contained northern individuals. Within-region crosses consisted of three plants grown together at a single site. Four such plots with southern individuals were used. Plots containing northern individuals failed. Between-region crosses contained two individuals from one region (pollen donors) and a single individual from the alternate region (pollen recipient). On the pollen recipient, all seeds except those due to selfing should be the result of a between-region cross. Three between-region crosses targeted a northern recipient, and two targeted a southern recipient. Plants for each treatment were randomly selected from the common garden and transplanted in the spring of 2012. The mean number of seeds per inflorescence (eight seed heads counted) was used to measure fertility for each treatment class.

Results

Population Structure

Substantial genetic variation was present across the range of *Helianthus pumilus*. Allele counts at 17 SSR loci varied from three to 32 per locus (mean = 9.2). Heterozygote deficit was observed for four loci, attributable to null alleles or population

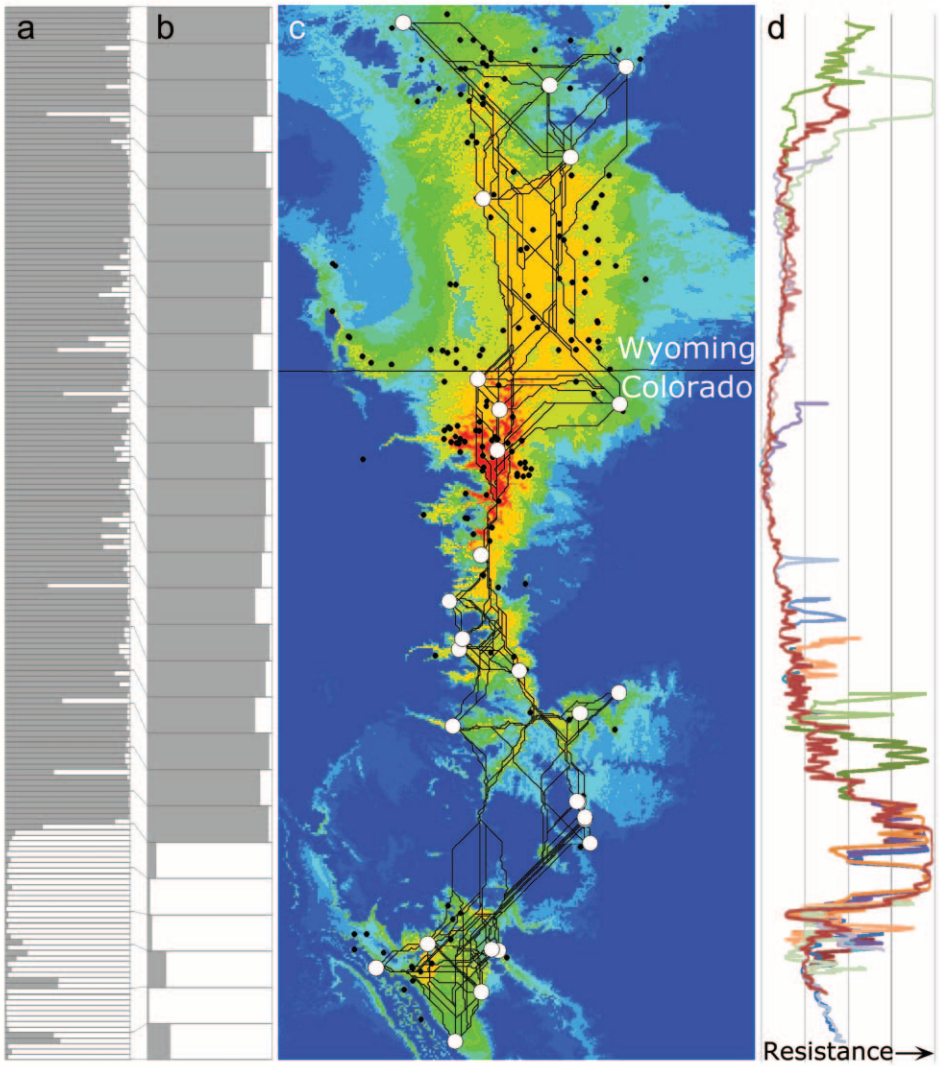

Fig. 2 Fractional assignment of *Helianthus pumilus* individuals (*a*) and populations (*b*) to one of two genotypic clusters. Samples from low to high latitude are arranged from bottom to top, in rough correspondence with their geographic position in *c*. *c*, Habitat suitability model for *Helianthus pumilus* based on Maxent modeling. Lines in map mark least-cost paths inferred between sampled populations (white circles) across the habitat suitability surface. *d*, Stepwise cost of migration across *c*; colors signify least-cost paths between representative population pairs. A genetic discontinuity, visualized in STRUCTURE results at left, is coincident with a region of unoccupied habitat across which resistance to migration is high.

structure (table B3). Only two of 493 population-locus pairings deviated significantly from Hardy-Weinberg equilibrium (table B4). The data deviated only in minor ways, typical of most SSR data sets, from assumptions of downstream analyses and were therefore considered suitable for estimation of gene flow rates and population structure.

The most probable number of populations, inferred using STRUCTURAMA, was two (mean posterior probability = 0.99948). The only other value returned, $K = 3$, was sampled very rarely (posterior probability = 0.00052). The Evanno et al. (2005) method for inferring K was not useful for verifying the STRUCTURAMA result because the method, which relies on the change in likelihood between consecutive K values, is ineffective for $K = 2$ (a peak in ΔK cannot be found at $K = 2$ because ΔK cannot be calculated for $K = 1$). The method was nevertheless applied; ΔK was highest at $K = 2$ and declined rapidly for larger K values, consistent with the STRUCTURAMA result.

Using STRUCTURE, populations could be assigned to one of two clusters with little admixture (fig. 2*b*). The lowest observed assignment probability of a population to its best fit cluster was 0.816. The same held true for individuals (fig. 2*a*). Only 18 individuals, from 13 populations scattered throughout the range, were assigned to one or the other cluster with a level of inferred ancestry less than 0.8. With few exceptions, individuals (and populations) from the southern geographic

region were assigned to one genetic cluster, while those from the north were assigned to the other.

Migration rates, inferred using BA3, indicated very little recent migration between northern and southern regions. The two independent chains converged to the same posterior distribution based on strong correlation in allele frequency estimates (Pearson's $r = 0.9999$). Results from a single chain with a post-burn-in effective sample size of 1600 are reported. The expected proportion of migrants in the southern region with origins in the north was 0.0167 ± 0.0135 (mean $\pm$ SD). The inferred migration rate from south to north was similarly low, 0.0036 ± 0.0033. The 95% credible interval (approximated as mean $\pm$ 1.96 $\times$ SD) thus included 0 for both directions of migration ($0 \leq m_{[n \leftarrow s]} \leq 0.010$; $0 \leq m_{[s \leftarrow n]} \leq 0.0432$). Only seven of 192 individuals had a posterior probability $\geq$0.05 of being either a first- or a second-generation immigrant. The remaining individuals were identified as nonimmigrants, with a posterior probability $\geq$0.95.

Habitat Suitability Modeling

The mean prediction of habitat suitability across the range is shown in figure 1. Models containing all sites, as well as the partitioned models, were robust. Omission rates on test sites were similar to predicted values except for the southern group, where few samples in a small geographic area caused spatial autocorrelation between test and training sites, resulting in lower than expected omission rates (in this case, below a cumulative threshold of 60%). Mean values for the area under the receiver operating characteristic curve were high (all sites, 0.954 ± 0.007; northern group, 0.959 ± 0.007; southern group, 0.986 ± 0.035), indicating a good fit between models and test data. Models predicted presence at test sites significantly better than random using 11 common thresholds (all sites, $P < 2.7 \times 10^{-21}$; north, $P < 1.1 \times 10^{-25}$; south, $P < 0.04$ for all but one nonsignificant test).

The background similarity test found that occupied sites in the north and south were more similar to each other than expected, given differences in the habitat available in the disjunct regions (I and D overlap metrics used; two-tailed single-sample Z-test, $P < 0.01$). This implies that plants occupy or "seek" similar habitat in the two regions. Range-breaking tests were also significant (linear, $P < 0.05$; blob, $P < 0.03$), indicating an important environmental breakpoint between the two defined groups. The null distributions for range-breaking tests were not normal. However, I and D were lower for the observed north/south partition than for any random partitions found in the null distributions, which were sampled exhaustively. Thus, one may still conclude that an important break in habitat-type availability exists between northern and southern regions, even though the Z-test may be invalid.

Inspection of Maxent output grids meant to identify regions of poor predictive power indicated that paleoclimatic models were robust. Regional artifacts due to novel climate conditions in the projection layers were minor. Clamping had little effect, and the multivariate environmental similarity surface suggested little environmental novelty relative to current-period data layers in the study area.

Barriers to Migration

A web of most-likely routes for gene flow between sampled sites is displayed in figure 2*c*. Visualization of the "cost per step" along these least-cost paths suggests a substantial barrier to migration between northern and southern regions, coincident with an unoccupied area of low-quality habitat. Simple Mantel tests suggest isolation by standard, straight-line distance ($P = 0.0225$), as well as by EGD ($P = 0.0015$). When straight line distance is partialed out, the correlation between EGD and genetic distance remains statistically significant ($P = 0.0029$). A negative correlation when controlling for EGD suggests that the correlation between straight line and genetic distance is a subset of that explained by EGD, as expected given the method of calculation (table B5). EGD was a better predictor of genetic divergence between populations ($R^2 = 0.1278$) than straight-line distance ($R^2 = 0.0675$).

Common Garden

When grown in a common garden, the average date of peak flowering (measured in 2011) in southern individuals was 9.3 d later than individuals from the north (Welch's t-test, $P = 0.0006$; fig. 3*a*). Plants measured in 2010 exhibited the same pattern (southern plants 4.48 d later; $P < 0.0001$). Plants from the south produced significantly more seeds than northern plants in 2010 and 2011 ($P = 0.0006$ and $P < 0.0001$, respectively; table B6). In 2011, southern individuals produced, on average, $37{,}018 \pm 20{,}976$ seeds per plant (mass = 126.8 ± 61.7 g), whereas those from the north produced $26{,}433 \pm 14{,}000$ (100.9 ± 56.3 g), a 40% difference. In 2010, southern plants produced 47% more seeds. Southern plants had fewer seeds per head (30 ± 10 vs. 37 ± 12; $P < 0.0001$); thus, their higher seed production is entirely attributable to having 70% more seed heads per plant (1227 ± 483 vs. 720 ± 347; $P < 0.0001$).

Of eight morphological characters measured, only plant height at first flowering differed significantly between regions (table B7). Southern plants were taller than northern plants (40.6 ± 12.8 vs. 32.8 ± 9.3 cm; $P < 0.0001$). A significant linear relationship between two characters (mean disk diameter and mean leaf length) and latitude was observed (simple linear regression, $P < 0.0001$ and $P = 0.026$, respectively). Plants from the south tend to have smaller disks (hence, smaller seed heads and lower seed capacity) and longer, more lanceolate leaves than northern plants.

Crossing Study

Fertility, measured as seed set per inflorescence, was lower in hand-pollinations than in isolation plots. This was expected because only a subset of the disk flowers were receptive on the pollination date for hand-pollinations, whereas in the isolation plots all flowers in an inflorescence had the opportunity to be pollinated. Nevertheless, the relative success rates were remarkably consistent (fig. 4; table B8). Selfing rates were low in both experiments and did not differ significantly between regions (for hand-pollination, $P = 0.5634$; for isolation plots, $P = 0.9858$). Between-region pairings showed the greatest fertility and were approximately sevenfold more productive than selfing, regardless of whether the measurement was made in isolation plots ($P = 0.0428$) or as a result of hand-polli-

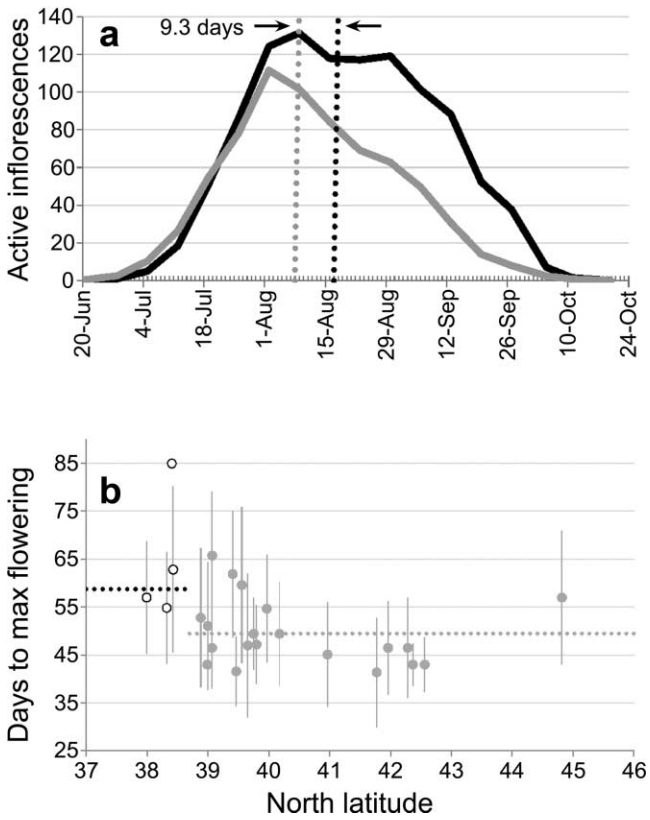

Fig. 3 *a*, Flowering phenology of two geographically disjunct groups of *Helianthus pumilus* measured in a common garden. *Y*-axis is mean number of active inflorescences per plant. Black line = southern group, gray line = northern group. Dotted lines indicate average date of peak flowering. Peak flowering occurred ~9 d later in plants from the southern region compared with those from the north. *b*, Mean peak flowering time (±1 SD) for wild populations in common garden (filled gray points = northern region, open circles = southern region). Dotted lines mark group means (black = south, gray = north). Under reproductive character displacement or reinforcement, populations close to the geographic barrier should differ more in peak flowering date than populations farther away. The four northern populations closest to the barrier did not differ significantly from the remainder of the northern populations (*t*-test, $P = 0.685$). Therefore, neither reproductive character displacement nor reinforcement are indicated for this trait.

nations ($P = 0.3552$). Within-region crosses were approximately four times as productive as self-fertilization.

Discussion

A Geographic Barrier to Gene Flow Splits the Helianthus pumilus *Range*

Helianthus pumilus is a showy perennial, inhabits areas that are easily accessed, and grows in a restricted geographic region that is relatively well populated by humans—precisely the kind of species that is frequently collected and for which an accurate species distribution can be mapped using locality information from herbarium specimens. A prominent gap in occupied habitat where no specimens have been collected is apparent around 38.7°N latitude. The gap in collections coincides with a region containing little suitable habitat, through which least-cost path analysis suggested strong resistance to migration (fig. 2*c*, 2*d*). Two genetically differentiated groups exist on opposite sides of the gap, and there has been little recent gene flow between them.

Whether genetic differentiation across the gap is the result of a barrier effect or recent migration from differentiated source populations can be considered using paleoclimatic information. While the paleodistribution of the species is ultimately unknowable, little of the extant range was glaciated at the last glacial maximum. Pinedale ice sheets extended less than a thousand meters below existing alpine areas in both regions (Madole 1986; Young et al. 2011), well above all but ~2% of current-day occupied sites. Paleoclimatic modeling suggests that currently occupied regions would have contained suitable habitat at the last glacial maximum, ~21,000 yr ago, and that a region of low-suitability habitat likely separated northern and southern regions then, as it does now (fig. 5). These results are consistent with a barrier effect.

We conclude that an ~50-km-wide barrier to gene flow is present within the *H. pumilus* range. The barrier is the Pikes Peak massif, a steep-sided group of mountains that includes Pikes Peak (4302 m), the easternmost high alpine peak in the Rocky Mountains. This geographic feature precludes the existence of much foothill habitat, which is favored by this species, in that particular portion of the range.

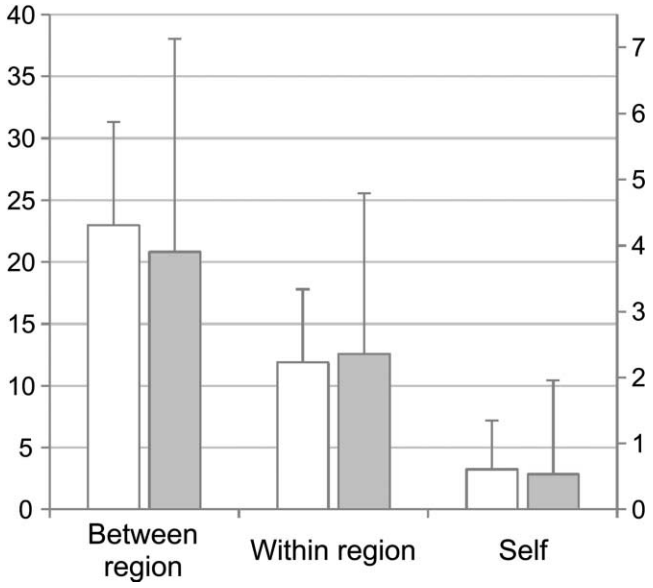

Fig. 4 Fertility rates for controlled crosses conducted in isolation plots and in the common garden. White bars show mean number of seeds (+1 SD) produced per inflorescence in isolation plots, and gray bars show results from hand-pollinations, plotted against the right *Y*-axis, which has been scaled to emphasize proportional similarities between experiments. *Helianthus pumilus* is largely self-incompatible. Higher between-region than within-region seed set indicates an absence of postpollination intrinsic reproductive isolation between disjunct groups.

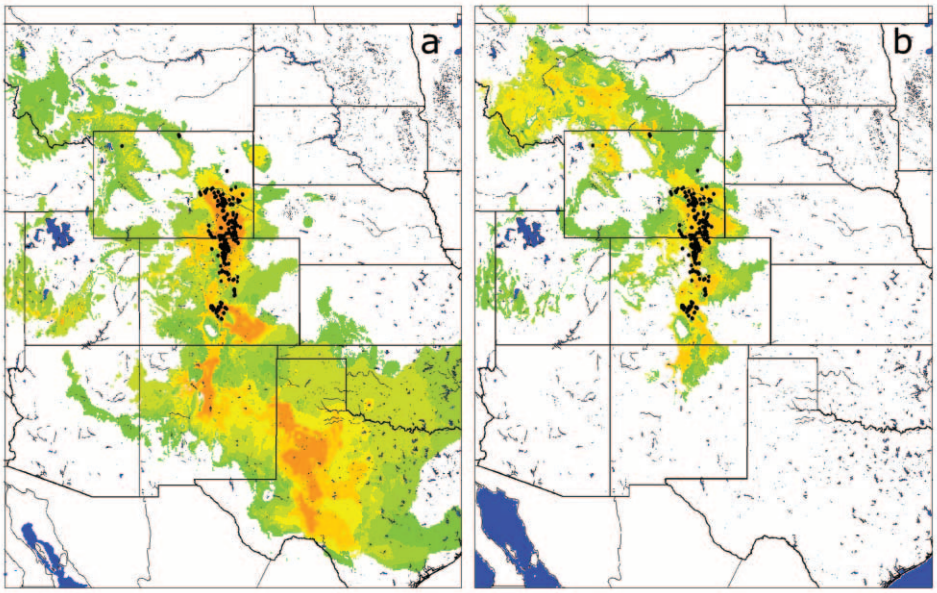

Fig. 5 Paleoclimatic modeling predicts that suitable habitat for *Helianthus pumilus* existed during the last glacial maximum in regions that are inhabited today. Warmer colors indicate more suitable habitat (values <0.1 are transparent; increment = 0.1). Points are current-day occupied sites. *a*, Habitat suitability ~21,000 yr before present. *b*, Habitat suitability under current conditions.

Local Adaptation and the Origin of an Adaptive Phenotype

A long history of common-garden experiments has demonstrated the prevalence of local adaptation in plants (Linhart and Grant 1996; Savolainen et al. 2007; cf. Leimu and Fischer 2008). In our common garden, individuals from the south produced 40% more seeds than individuals from the north (table B6), despite location of the common garden within the northern range. If we accept seed set as a reasonable correlate of fitness, then southern plants are better adapted to the northern range than are their local counterparts. Greater seed production on southern plants was caused by the occurrence of more inflorescences. Increased inflorescence number is related to the existence of a distinct vegetative morphology present in the southern region (fig. 6). Many (but not all) plants from the south have noticeably longer, more numerous, and more languid basal branches, which support more inflorescences without crowding. Plants from the north, on the other hand, have a stocky and robust habit and support development of fewer inflorescences.

In extensive field observations and in the common garden, this southern phenotypic gestalt was not noted to occur with any frequency in northern populations. This observation was borne out by canonical discriminant analysis on four traits chosen to encapsulate the distinction between what we perceived as distinct northern and southern types (flowering time, number of seed heads, plant height, disk diameter). The discriminant function assigned 45% of southern individuals to a "southern type" (17 of 38 southern individuals phenotyped) and 55% to a "northern type" (overall Wilks's λ $F < 0.0001$), consistent with rough field observations. In contrast, only 4% of northern individuals were assigned as southern type (8 of 186 individuals, from four of 12 populations in the southern third of the northern range). Of these, only one was unambiguously assigned (probability >0.9) as southern type. In the south, 18% were unambiguously northern type. Thus, while we might expect to encounter the characteristic northern type frequently in the southern range, the southern type is essentially absent from the northern range.

When grown in the north, southern-type individuals produced nearly twice as many seeds as northern-type individuals (94% more). When southern types are excluded, there is no significant difference between north and south in terms of seed set ($P = 0.8558$). Thus, asymmetric geographic distribution of the southern type, rather than geographic region per se, is a more accurate explanation for the differences in seed set measured in the common garden.

Counter to the expectation of local adaptation, southern-type individuals appear to be better adapted to the northern range than northern types. This unexpected result can be explained in a number of ways (following Kawecki and Ebert 2004). First, the southern type might be an accidental preadaptation to conditions specific to the common garden that are not representative of the remainder of the northern range. The establishment of replicated common gardens in the northern range could address this issue definitively. An alternative hypothesis, that the southern type is a "generalist" phenotype, could be tested by establishment of a common garden in the southern region. A generalist would be expected to have relatively high fitness across a broad range of conditions, including those encountered in the northern range, but lower fitness than locally adapted individuals in its home range. This alternative might explain why the southern type is not fixed in the south.

Fig. 6 *Helianthus pumilus* plants from the southern disjunct portion of the range (*left*) are often tall and gracile, and they produce more inflorescences than the short, sturdy plants that grow in the north (*right*). These morphological differences result in a fitness advantage for southern plants relative to local genotypes when grown in the northern range. The genes responsible for this adaptive morphology appear to be blocked from entering the northern range by the Pikes Peak massif, which offers little suitable habitat for the species.

Second, the finding may be spurious because seed set might not correlate well with lifetime fitness in *H. pumilus*. Fecundity is commonly used to measure fitness in *Helianthus* (Burke and Rieseberg 2003; Mercer et al. 2006; Whitney et al. 2006; Gutierrez et al. 2011). However, fecundity may only rarely be the best predictor of population growth rates (Crone 2001). Estimating fitness using seed set alone may be particularly misleading for perennial species, where survivorship appears to be more important (Silvertown et al. 1993). In this study, survivorship would not have been a useful metric of relative fitness because it remained very high after establishment (≈99% per year) and mortality occurred with equal frequency among northern and southern individuals. Seed germination rates would be difficult to assess in nature because of considerable seed dormancy. Thus, seed set was chosen as a tractable fitness correlate for a short-term study.

Third, more recently established populations might be less well adapted to their environment than older populations. Given the expected pattern of northward migration of temperate species from southwestern glacial refugia in central North America since the last glacial maximum (Pielou 1991, pp. 232–244; Provan and Bennet 2008), the northern region might have been colonized more recently and sufficient time may not yet have passed for local adaptation to occur. This hypothesis, however, predicts lower genetic diversity in the north as well, which we did not observe (fig. B1; figs. B1 and B2 are available online). Moreover, paleoclimatic modeling shows that the current range offered suitable habitat and a similarly positioned barrier to gene flow at the last glacial maximum (fig. 5), suggesting continuous occupation of the current range. This strengthens the case against the idea that there has been insufficient time for adaptation in the north.

While not all alternative explanations can be refuted, the data may correctly indicate local maladaptation. The expectation of universal local adaptation has recently been challenged. In a meta-analysis of reciprocal transplant experiments, Leimu and Fischer (2008) found that one population performed better in both tested environments about as often as

local genotypes prevailed in their home environment. This implies that local adaptation, while common, is not ubiquitous and that results like ours, where foreign genotypes outperform locals, might be expected about half of the time.

Our best guess, given the current data, is that the advantageous southern type evolved in relative isolation, from standing genetic variation present in the southern region. The phenotype was not fixed in any populations examined from the south, suggesting that it may not be advantageous throughout the region, polymorphism perhaps being maintained by conflicting selection pressures and gene flow. Indeed, the southern region is more environmentally heterogeneous than the northern region (fig. B2), with *H. pumilus* populations existing in cool, moist, near-alpine habitats to seasonally hot and extremely dry lowland rain-shadow deserts within 25 km of one another.

The Barrier Impedes Species-Wide Adaptation

It is curious that the southern type has not spread more uniformly throughout the species range given its apparent fitness advantage. An obvious explanation is that the barrier between regions hampers its movement. Demonstration that geography is important in controlling the distribution of adaptive phenotypes requires the elimination of other potential isolating mechanisms that might similarly act to restrict gene flow between populations. We consider two prepollination isolation mechanisms, one intrinsic (temporal isolation), the other extrinsic (pollinator isolation). Evidence for ecogeographic isolation and postpollination intrinsic isolation (fertility rates in controlled crosses) is also discussed.

The potential exists for temporal isolation. In the common garden, peak flowering occurred 9 d later in southern individuals (fig. 3*a*). Differences in flowering phenology are thought to evolve quickly in allopatry, in response to differing environmental conditions, and have been mentioned as possible causal mechanisms in a few proposed cases of sympatric speciation, such is the strength with which they can drive divergence (McNeilly and Antonovics 1968; Levin 1971; Silvertown et al. 2005; Savolainen et al. 2006). While allochrony is sometimes leveraged as a reinforcing mechanism, preventing maladaptive gene flow and forcing the speciation process to completion, we see no indication of reinforcement or reproductive character displacement here (fig. 3*b*). Given substantial overlap in the overall flowering period (fig. 3*a*), we expect that the difference in peak flowering time would not act as a strong barrier to gene flow in nature. In this species, flowering is sustained for a long enough period that ample opportunity for cross-pollination would be available in sympatry.

We did not see any obvious indication of pollinator isolation between groups; plants in the common garden were frequented by the expected pollinator (native solitary bees) as well as ants, beetles, and grasshoppers. It was only examined indirectly here (under pollinator isolation we expect lower seed set in southern plants grown at the northern site, which we did not see, even when southern-type individuals were excluded), but pollinator isolation might generally require floral morphological differentiation (Grant 1949). Excepting disk diameter, this was not observed. Mortality in the common garden was low and evenly distributed between northern and southern individuals over a 2-yr period, and geospatial statistical tests (Warren et al. 2008; Glor and Warren 2011), though a crude tool, suggest that plants in the northern and southern range preferentially occupy similar sites even though the habitat available to them differs. Limited evidence for differentiation in environmental preference suggests that ecogeographic isolation should be weak.

With respect to postpollination intrinsic reproductive isolation, three features emerged from the crossing study. First, selfing rates were low (fig. 4; table B8); thus, there is substantial self-incompatibility in *H. pumilus*, as noted by others (Nikolova et al. 2004). This validates the experimental approach used to estimate fertility rates in the isolation plots by implying a low rate of seed set due to pollen flow from outside the plots. Second, between-region crosses were more productive than within-region crosses. Third, between-region fertility in isolation plots was not significantly different from fertility measured for the same individuals under open-pollinated conditions in the common garden (2010, $P = 0.4618$; 2011, $P = 0.152$). This shows that the isolation plots, even though they contained only three plants, were not pollen limited and adequately reflected fertility levels that might be achieved in a larger population. Overall, the results demonstrate that plants from different regions are perfectly capable of crossing and producing seeds and can do so at levels equal to or greater than plants from the same region. Thus, we did not see any evidence suggesting postpollination intrinsic reproductive isolation between regions.

Because none of the aforementioned isolating mechanisms seem likely to limit reproduction between individuals from the south and north, we propose that the geographic barrier alone is primarily responsible for preventing the movement of the southern adaptive phenotype. By blocking the northward movement of useful genes, the barrier appears to prevent global (i.e., species-wide) adaptation, resulting in a northern group that is less well adapted than it might be given greater gene flow. An alternative explanation is that, in cases where adaptive phenotypes are polygenic in nature, the movement of those phenotypes across a species range will be checked by recombination, even if the restrictions on gene flow that promote the evolution of the phenotype are lifted (Gavrilets 1996; Coyne et al. 1997). Thus, even without a barrier to migration we might not expect the southern type to be able to invade an occupied northern region. Cycles of local extinction followed by recolonization of unoccupied habitat might eventually overcome any such recombinational barrier (Lande 1979, 1985), allowing the southern type to move northward, although a major force supporting this dynamic in the Northern Hemisphere, periodic glaciation, may not pertain in this species (fig. 5).

Concluding Remarks

When geographic barriers to gene flow arise, a tension is set up between processes that favor local adaptation and lineage splitting and those that favor global adaptation and species cohesion. The outcome—speciation or cohesion—might be predicted in certain cases by understanding the permeability of the barrier, the selection differential of the phenotype in question, mutation rates, and demographic conditions in the affected populations. At a more basic level, it may be worth-

while to consider space—or spatial relationships—as a form of evolutionary constraint, similar in consequence to constraints arising from fundamental forces (gravity, electromagnetism), genomic organization (linkage, pleiotropy, allelic variation), or developmental integration, all of which lower the probability of certain phenotypic outcomes (see Gould and Lewontin 1979; Schwenk 1995). Distinct from the other forms of constraint, spatial constraints do not limit the production of variants on which selection may act. Instead, they prevent the spread of well-tested adaptations to regions isolated from the place of origin. In so doing, spatial constraints may promote new species that are, arguably, worse adapted to their surroundings than they might have been had a greater period of time for gene flow passed before intrinsic barriers arose to complete the process of allopatric speciation. In essence, a spatial constraint might be viewed as a "constraint of lost opportunity."

This conceptual perspective has applied value in the field of agricultural improvement, where spatial constraints that bound evolutionary trajectories in disjunct populations of wild-crop relatives can be overcome by human manipulation (Hawkes 1977; Tanksley and McCouch 1997). If valid, the perspective leads to a prediction: crop species whose wild relatives exist in highly subdivided distributions have more potential for improvement via artificial introgression of wild genes than those whose relatives exist in uniform distributions.

Acknowledgments

We thank our field assistants (Remi Bonnart, Jennifer Crane, Celeste Falcon, Briana Gross, Lisa Hill, Annette Miller, Ann Reilley, Valderês Sousa, Imke Thormann, Travis Vagher, and Randa Vagher) and a group of citizen scientists (Lesley Aaeng, Mary Eldridge, Ingrid Fischer, Marilyn Fogerty, Linda Gruidell, Maureen McCarthy, Chris McDonald, Millie Miller, Tracie Roth, Vickie Traxler, and Maria Zellar) who made the crossing experiments possible. We also thank Laura Marek and Gerald Seiler for their comprehensive collection efforts on behalf of the US National Plant Germplasm System, which enabled us to quickly identify suitable field sites, and Briana Gross, John R. Paul, and Seema Sheth for many helpful comments on the manuscript.

Appendix A

Table A1

Sources for *Helianthus pumilus* Seed and DNA

Population name	Latitude	Longitude	No. individuals, genotyping study	NPGS accession	No. individuals, common garden (2010/2011)
Westcliffe	37.9936	−105.4750	5	PI650041	26/13
Dead Mule Gulch	38.2155	−105.3578	8	PI650040	0
Hayden Pass	38.3233	−105.8283	8	PI650043	27/13
Cañon City	38.4000	−105.2819	6	Ames27895	0
Temple Canyon	38.4067	−105.3114	8	Ames27896	4/2
Texas Creek Gulch	38.4297	−105.5964	8	PI650042	23/12
Colorado Springs	38.8831	−104.8697	6	PI650039	11/5
Academy South	38.9897	−104.8933	8	PI650038	17/8
Academy North	38.9986	−104.8931	8	PI650037	24/13
Monument	39.0683	−104.9231	6	PI650036	13/6
Mount Herman	39.0692	−104.9300	8	PI650077	14/8
Bailey	39.4072	−105.4861	5	PI650044	19/10
Castle Pines	39.4667	−104.9167	6	Ames18009	21/10
Parker	39.5575	−104.7428	8	PI650078	15/8
Morrison	39.6547	−105.1867	6	Ames27938	12/7
Central City	39.7525	−105.4581	6	PI650045	24/12
Smith Hill	39.8000	−105.4422	6	PI650046	21/10
Nederland	39.9669	−105.5006	6	PI650047	21/12
South Saint Vrain	40.1761	−105.3589	6	PI650048	25/12
Rist Canyon	40.6456	−105.2853	8	PI650049	0
Cherokee Park	40.8275	−105.2758	7	PI650054	0
Pawnee	40.8528	−104.7372	7	PI650076	0
Virginia Dale	40.9653	−105.3736	6	PI650055	20/10
North Sybille Creek	41.7742	−105.3486	6	PI650066	25/13
Wheatland	41.9608	−104.9567	6	PI650067	25/14
El Rancho	42.2833	−105.0508	6	PI650074	18/12
Webb Canyon	42.3661	−104.7106	6	PI650068	14/6
LaPrele Creek	42.5633	−105.7064	6	PI650071	19/10
Dayton	44.8178	−107.3008	6	PI650072	16/9

Note. NPGS = National Plant Germplasm System.

Literature Cited

Amos W, JI Hoffmann, A Frodsham, L Zhang, S Best, AVS Hill 2007 Automated binning of microsatellite alleles: problems and solutions. Mol Ecol Notes 7:10–14.

Anderson RP, M Gómez-Laverde, AT Peterson 2002 Geographical distributions of spiny pocket mice in South America: insights from predictive models. Glob Ecol Biogeogr 11:131–141.

Barraclough T, A Vogler 2000 Detecting the geographical pattern of speciation from species-level phylogenies. Am Nat 155:419–434.

Bonnet E, Y Van de Peer 2002 Zt: a software tool for simple and partial Mantel tests. J Stat Softw 7:1–12.

Burke JM, LH Rieseberg 2003 Fitness effects of transgenic disease resistance in sunflowers. Science 300:1250.

Carstens BC, CL Richards 2007 Integrating coalescent and ecological niche modeling in comparative phylogeography. Evolution 61:1439–1454.

Chapman MA, CH Pashley, J Wenzler, J Hvala, S Tang, SJ Knapp, JM Burke 2008 A genomic scan for selection reveals candidates for genes involved in the evolution of cultivated sunflower (*Helianthus annuus*). Plant Cell 20:2931–2945.

Coyne JA, NH Barton, M Turelli 1997 A critique of Sewall Wright's shifting balance theory of evolution. Evolution 51:643–671.

Coyne JA, HA Orr 2004 Speciation. Sinauer, Sunderland, MA.

Crone EE 2001 Is survivorship a better fitness surrogate than fecundity? Evolution 55:2611–2614.

Dobzhansky T 1940 Speciation as a stage in evolutionary divergence. Am Nat 74:312–321.

——— 1951 Genetics and the origin of species. Columbia University Press, New York.

Dorrell DG, EDP Whelan 1978 Chemical and morphological characteristics of seeds of some sunflower species. Crop Sci 18:969–971.

Elith J, CH Graham, RP Anderson, M Dudík, S Ferrier, A Guisan, RJ Hijmans, et al 2006 Novel methods improve prediction of species' distributions from occurrence data. Ecography 29:129–151.

Elith J, M Kearney, S Phillips 2010 The art of modelling range-shifting species. Method Ecol Evol 1:330–342.

Elith J, JR Leathwick 2009 Species distribution models: ecological explanation and prediction across space and time. Annu Rev Ecol Evol Syst 40:677–697.

Evanno G, S Regnaut, J Goudet 2005 Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 14:2611–2620.

Fitzpatrick BM, M Turelli 2006 The geography of mammalian speciation: mixed signals from phylogenies and range maps. Evolution 60:601–615.

Garant D, SE Forde, AP Hendry 2007 The multifarious effects of dispersal and gene flow on contemporary adaptation. Funct Ecol 21:434–443.

Gavrilets S 1996 On phase three of the shifting-balance theory. Evolution 50:1034–1041.

Glor RE, D Warren 2011 Testing ecological explanations for biogeographic boundaries. Evolution 65:673–683.

Gottlieb LD 1973 Genetic differentiation, sympatric speciation, and the origin of a diploid species of *Stephanomeria*. Am J Bot 60:545–553.

Gould SJ, RC Lewontin 1979 The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme. Proc R Soc B 205:581–598.

Grant V 1949 Pollination systems as isolating mechanisms in angiosperms. Evolution 3:82–97.

Guisan A, NE Zimmermann 2000 Predictive habitat distribution models in ecology. Ecol Model 135:147–186.

Gutierrez A, M Cantamutto, M Poverene 2011 Persistence of sunflower crop traits and fitness in *Helianthus petiolaris* populations. Plant Biol 13:821–830.

Hawkes JG 1977 The importance of wild germplasm in plant breeding. Euphytica 26:615–621.

Huelsenbeck JP, P Andolfatto 2007 Inference of population structure under a Dirichlet process model. Genetics 175:1787–1802.

Hugall A, C Moritz, A Moussalli, J Stanisic 2002 Reconciling paleodistribution models and comparative phylogeography in the Wet Tropics rainforest land snail *Gnarosophia bellendenkerensis* (Brazier 1875). Proc Natl Acad Sci USA 99:6112–6117.

Husband BC, HA Sabara 2004 Reproductive isolation between autotetraploids and their diploid progenitors in fireweed, *Chamerion angustifolium* (Onagraceae). New Phytol 161:703–713.

Jakobsson M, NA Rosenberg 2007 CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics 23:1801–1806.

Jordan DS 1905 The origin of species through isolation. Science 22:545–562.

Kawecki TJ, D Ebert 2004 Conceptual issues in local adaptation. Ecol Lett 7:1225–1241.

Kay KM 2006 Reproductive isolation between two closely related hummingbird-pollinated Neotropical gingers. Evolution 60:538–552.

Kirkpatrick M, NH Barton 1997 Evolution of a species' range. Am Nat 150:1–23.

Kozak KH, CH Graham, JJ Wiens 2008 Integrating GIS-based environmental data into evolutionary biology. Trends Ecol Evol 23:141–148.

Lande R 1979 Effective deme sizes during long-term evolution estimated from rates of chromosomal rearrangement. Evolution 33:234–251.

——— 1985 The fixation of chromosomal rearrangements in a subdivided population with local extinction and colonization. Heredity 54:323–332.

Legendre P 1993 Spatial autocorrelation: trouble or new paradigm? Ecology 74:1659–1673.

Leimu R, M Fischer 2008 A meta-analysis of local adaptation in plants. PLoS ONE 3:e4010.

Lenormand T 2002 Gene flow and the limits to natural selection. Trends Ecol Evol 17:183–189.

Levin DA 1971 The origin of reproductive isolating mechanisms in flowering plants. Taxon 20:91–113.

Lewis PO, D Zaykin 2001 Genetic data analysis: computer program for the analysis of allelic data. Version 1.0 (d16c). http://www.eeb.uconn.edu/people/plewis/software.php. Accessed May 2, 2013.

Linhart YB, MC Grant 1996 Evolutionary significance of local genetic differentiation in plants. Annu Rev Ecol Syst 27:237–277.

Madole RF 1986 Lake Devlin and Pinedale glacial history, Front Range, Colorado. Quat Res 25:43–54.

Mayr E 1959 Isolation as an evolutionary factor. Proc Am Philos Soc 103:221–230.

——— 1963 Animal species and evolution. Belknap, Cambridge, MA.

McNeilly T, J Antonovics 1968 Evolution in closely adjacent plant populations. IV. Barriers to gene flow. Heredity 23:205–218.

Mercer KL, DL Wyse, RG Shaw 2006 Effects of competition on the fitness of wild and crop-wild hybrid sunflower from a diversity of wild populations and crop lines. Evolution 60:2044–2055.

Morjan CL, LH Rieseberg 2004 How species evolve collectively: implications of gene flow and selection for the spread of advantageous alleles. Mol Ecol 13:1341–1356.

Nei M 1978 Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics 89:583–590.

Nikolova L, M Christov, G Seiler 2004 Interspecific hybridization between *H. pumilus* Nutt. and *H. annuus* L. and their potential for cultivated sunflower improvement. Helia 27:151–162.

Oetting WS, HK Lee, DJ Flanders, GL Wiesner, TA Sellers, RA King 1995 Linkage analysis with multiplexed short tandem repeat polymorphisms using infrared fluorescence and M13 tailed primers. Genomics 30:450–458.

Otto SP, J Whitton 2000 Polyploid incidence and evolution. Annu Rev Genet 34:401–437.

Papadopulos AST, WJ Baker, D Crayn, RK Butlin, RG Kynast, I Hutton, V Savolainen 2011 Speciation with gene flow on Lord Howe Island. Proc Natl Acad Sci USA 108:13188–13193.

Pashley CH, JR Ellis, DE McCauley, JM Burke 2006 EST databases as a source for molecular markers: lessons from *Helianthus*. J Hered 97:381–388.

Pearse DE, KA Crandall 2004 Beyond F_{ST}: analysis of population genetic data for conservation. Conserv Genet 5:585–602.

Phillips SJ, RP Anderson, RE Schapire 2006 Maximum entropy modeling of species geographic distributions. Ecol Model 190:231–259.

Phillips SJ, M Dudík 2008 Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. Ecography 31:161–175.

Pielou EC 1991 After the Ice Age: the return of life to glaciated North America. University of Chicago Press, Chicago.

Pritchard JK, M Stephens, P Donnelly 2000 Inference of population structure using multilocus genotype data. Genetics 155:945–959.

Provan J, KD Bennett 2008 Phylogeographic insights into cryptic glacial refugia. Trends Ecol Evol 23:564–571.

Ramsey J, HD Bradshaw, DW Schemske 2003 Components of reproductive isolation between the monkeyflowers *Mimulus lewisii* and *M. cardinalis* (Phrymaceae). Evolution 57:1520–1534.

Ramsey J, DW Schemske 1998 Pathways, mechanisms, and rates of polyploid formation in plants. Annu Rev Ecol Syst 29:467–501.

Savolainen O, T Pyhäjärvi, T Knürr 2007 Gene flow and local adaptation in trees. Annu Rev Ecol Evol Syst 38:595–619.

Savolainen V, M-C Anstett, C Lexer, I Hutton, JJ Clarkson, MV Norup, MP Powell, D Springate, N Salamin, WJ Baker 2006 Sympatric speciation in palms on an oceanic island. Nature 441: 210–213.

Schemske DW 2000 Understanding the origin of species. Evolution 54:1069–1073.

Schilling EE, CB Heiser 1981 Infrageneric classification of *Helianthus* (Compositae). Taxon 30:393–403.

Schluter D 2001 Ecology and the origin of species. Trends Ecol Evol 16:372–380.

Schneiter AA, JF Miller 1981 Description of sunflower growth stages. Crop Sci 21:901–903.

Schuelke M 2000 An economic method for the fluorescent labeling of PCR fragments. Nat Biotechnol 18:233–234.

Schwenk K 1995 A utilitarian approach to evolutionary constraint. Zoology 98:251–262.

Seiler G 1994 Oil concentration and fatty acid composition of achenes of North American *Helianthus* (Asteraceae) species. Econ Bot 48:271–279.

Seiler GJ, TJ Gulya Jr, L Fredrick-Marek 2012 Collection and evaluation of wild perennial *Helianthus pumilus* achenes for oil concentration and fatty acid composition [abstract]. Page 175 *in* 18th International Sunflower Conference program and abstracts. Mar del Plata and Barcarce, Argentina, February 27–March 1, 2012.

Silvertown J, M Franco, I Pisanty, A Mendoza 1993 Comparative plant demography—relative importance of lifecycle components to the finite rate of increase in woody and herbaceous perennials. J Ecol 81:465–476.

Silvertown J, C Servaes, P Biss, D Macleod 2005 Reinforcement of reproductive isolation between adjacent populations in the Park Grass Experiment. Heredity 95:198–205.

Slatkin M 1976 The rate of spread of an advantageous allele in a subdivided population. Pages 767–780 *in* S Karlin, E Nevo, eds. Population genetics and ecology. Academic Press, New York.

——— 1987 Gene flow and the geographic structure of natural populations. Science 236:787–792.

Sobel JM, GF Chen, LR Watt, DW Schemske 2010 The biology of speciation. Evolution 64:295–315.

Soberón J, AT Peterson 2005 Interpretation of models of fundamental ecological niches and species' distributional areas. Biodivers Inform 2:1–10.

Stebbins GL 1950 Variation and evolution in plants. Columbia University Press, New York.

Tang S, VK Kishore, SJ Knapp 2003 PCR-multiplexes for a genome-wide framework of simple sequence repeat marker loci in cultivated sunflower. Theor Appl Genet 107:6–19.

Tanksley SD, SR McCouch 1997 Seed banks and molecular maps: unlocking genetic potential from the wild. Science 277:1063–1066.

Templeton SR, RJ Robertson, J Brisson, J Strasburg 2001 Disrupting evolutionary processes: the effect of habitat fragmentation on collared lizards in the Missouri Ozarks. Proc Natl Acad Sci USA 98: 5426–5432.

Timme RE, BB Simpson, CR Linder 2007 High-resolution phylogeny for *Helianthus* (Asteraceae) using the 18S-26S ribosomal DNA external transcribed spacer. Am J Bot 94:1837–1852.

Van Oosterhout C, CWF Hutchinson, DPM Wills, P Shipley 2004 MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Mol Ecol Notes 4:535–538.

Wang IJ, RE Glor, JB Losos 2013 Quantifying the roles of ecology and geography in spatial genetic divergence. Ecol Lett 16:175–182.

Wang YH, KC Yang, CL Bridgman, LK Lin 2008 Habitat suitability modelling to correlate gene flow with landscape connectivity. Landsc Ecol 23:989–1000.

Warren DL, RE Glor, M Turelli 2008 Environmental niche equivalency versus conservatism: quantitative approaches to niche evolution. Evolution 62:2868–2883.

——— 2010 ENMTools: a toolbox for comparative studies of environmental niche models. Ecography 33:607–611.

Whitney KD, RA Randell, LH Rieseberg 2006 Adaptive introgression of herbivore resistance traits in the weedy sunflower *Helianthus annuus*. Am Nat 167:794–807.

Wiens JJ 2004 What is speciation and how should we study it? Am Nat 163:914–923.

Wilson GA, B Rannala 2003 Bayesian inference of recent migration rates using multilocus genotypes. Genetics 163:1177–1191.

Young NE, JP Briner, EM Leonard, JM Licciardi, K Lee 2011 Assessing climatic and nonclimatic forcing of Pinedale glaciation and deglaciation in the western United States. Geology 39:171–174.