Check for updates

Received: 26 November 2020 | Revised: 17 May 2021 | Accepted: 26 May 2021
DOI: 10.1111/mec.16008

SPECIAL ISSUE

Standing variation rather than recent adaptive introgression probably underlies differentiation of the *texanus* subspecies of *Helianthus annuus*

Gregory L. Owens[1] | Marco Todesco[2] | Natalia Bercovich[2] | Jean-Sébastien Légaré[2] | Nora Mitchell[3,4] | Kenneth D. Whitney[4] | Loren H. Rieseberg[2]

[1]Department of Biology, University of Victoria, Victoria, BC, Canada
[2]Department of Botany and Biodiversity Research Centre, University of British Columbia, Vancouver, BC, Canada
[3]Department of Biology, University of Wisconsin – Eau Claire, Eau Claire, WI, USA
[4]Department of Biology, University of New Mexico, Albuquerque, NM, USA

Correspondence
Gregory L. Owens, Department of Biology, University of Victoria, Victoria, BC, Canada.
Email: grego@uvic.ca

Funding information
Genome Canada, Grant/Award Number: LSARP2014-223SUN; National Science Foundation, Grant/Award Number: DEB 1257965

Abstract

The origins of geographic races in wide-ranging species are poorly understood. In Texas, the *texanus* subspecies of *Helianthus annuus* has long been thought to have acquired its defining phenotypic traits via introgression from a local congener, *H. debilis*, but previous tests of this hypothesis were inconclusive. Here, we explore the origins of *H. a. texanus* using whole genome sequencing data from across the entire range of *H. annuus* and possible donor species, as well as phenotypic data from a common garden study. We found that although it is morphologically convergent with *H. debilis*, *H. a. texanus* has conflicting signals of introgression. Genome wide tests (Patterson's *D* and TreeMix) only found evidence of introgression from *H. argophyllus* (sister species to *H. annuus* and also sympatric), but not *H. debilis*, with the exception of one individual of 109 analysed. We further scanned the genome for localized signals of introgression using PCAdmix and found minimal but nonzero introgression from *H. debilis* and significant introgression from *H. argophyllus* in some populations. Given the paucity of introgression from *H. debilis*, we argue that the morphological convergence observed in Texas is probably from standing genetic variation. We also found that genomic differentiation in *H. a. texanus* is mostly driven by large segregating inversions, several of which have signatures of natural selection based on haplotype frequencies.

KEYWORDS
adaptation, chromosomal inversion, ecological genetics, hybridization, population genetics – empirical, speciation

1 | INTRODUCTION

Patterns of genetic variation in wide-ranging species are a product of multiple processes, including local adaptation, demography, and migration. Since dispersal is limited in most plants, isolation by distance is common (Heywood, 1991; Wright, 1943). In wide-ranging species that span environmental gradients, migration between environmental niches is often reduced due to selection against immigrants (i.e., isolation by environment) (Sexton et al., 2014). This can result in a genome-wide signal of reduced migration, or be restricted to specific genomic regions depending on the strength of selection and the genomic architecture of adaptation. Reduced migration between populations can lead to the development of intraspecific forms that are variously called races, ecotypes, varieties and subspecies. Regardless of taxonomic classification, intraspecific geographic variation represents an important intermediate stage

 © 2021 John Wiley & Sons Ltd

between population and species, and understanding the underlying genetic architecture is of broad interest to students of adaptation and speciation.

Two other factors that influence how genetic diversity is distributed within a species are hybridization and genome structure. Hybridization and introgression transfers diversity between species, including potentially adaptive alleles. For example, *Arabidopsis arenosa* became adapted to drought and toxic levels of soil minerals in serpentine soil through introgression of alleles from *Arabidopsis lyrata* (Arnold et al., 2016). Similarly, brown winter coats in snowshoe hares from areas with little snow probably originated through introgression with black-tailed jackrabbits (Jones et al., 2018). It is likely, however, that most introgression is not adaptive, but rather neutral or even maladaptive (Martin & Jiggins, 2017). Genome structure, especially large chromosomal rearrangements such as inversions, shape the distribution of intraspecific diversity through their effects on the recombination of linked alleles. By suppressing recombination among locally adapted alleles, structural variants not only facilitate adaptation in the presence of gene flow, but they also create heterogenity in migration rates and patterns of genetic divergence across the genome (Kirkpatrick & Barton, 2006).

Wild sunflowers of the genus *Helianthus* are well-suited for studying processes that affect the formation and maintenance of intraspecific geographic variation. The genus contains approximately 50 species native to North America, which are found in diverse habitats ranging from the beaches of Florida and Baja California to active sand dunes in Texas and Colorado to hot springs in the Canadian Rocky Mountains (Heiser et al., 1969). There can be considerable geographic and ecological variation within species, even for narrow endemics, and many of the species contain multiple intraspecific taxa, including well-characterized ecotypes (Heiser et al., 1969; Kantar et al., 2015). Genetic studies have demonstrated that both isolation by distance (Baute et al., 2016; Moyers & Rieseberg, 2016) and isolation by environment (Andrew & Rieseberg, 2013) contribute to patterns of intraspecific geographic variation. Furthermore, chromosomal structural variants, especially inversions, have been shown to contribute importantly to ecotype formation within sunflower species (Huang et al., 2020; Todesco et al., 2020) and to differential introgression between them (Barb et al., 2014). Lastly, many of the species, but especially the 10–12 members of the annual clade, overlap in geographic distribution and form natural hybrids (Heiser, 1947, 1951a; Rieseberg et al., 1998). While hybrids typically are semi-sterile (Owens & Rieseberg, 2014), substantial interspecific gene flow has been demonstrated between several species pairs (e.g. Baute et al., 2016; Kane et al., 2009; Strasburg & Rieseberg, 2008; Zhang et al., 2019).

Here we attempt to identify and order the evolutionary factors contributing to geographic variation in the most widespread sunflower species, *Helianthus annuus*, which can be found across the United States, as well as southern Canada and northern Mexico. More recently, the species has spread into California, where it absorbed alleles from the native *Helianthus bolanderi* (Heiser, 1949; Owens et al., 2016). It has been suggested that the wide range of *H. annuus* is due to its ability to incorporate locally adaptive alleles from congeners (Heiser, 1951a). This idea is highlighted by studies of Texas populations, which are phenotypically convergent with a local congener, *Helianthus debilis* var. *cucumerifolius*, at several traits including smaller flower heads, earlier flowering and purple mottled stems (Heiser, 1951a). This, combined with observations that *H. annuus* was exclusive to human-disturbed locations in this area, and documented hybridization between the two species, led Heiser to argue that the expansion of *H. annuus* into eastern central Texas was facilitated by introgression from *H. debilis* (Heiser, 1951a,1951b). Since populations of *H. annuus* are interfertile across their range, these putatively introgressed populations in Texas were given subspecific status as *H. annuus* subsp. *texanus* (Heiser, 1954), while populations in the rest of the range are referred to as *H. annuus* subsp. *annuus*.

Molecular analyses have tested this hypothesis using rDNA, cpDNA, AFLPs and microsatellites, each finding some evidence that alleles diagnostic for *H. debilis* can be found in *H. a. texanus* at low frequency (Rieseberg et al., 1990, 2007; Scascitelli et al., 2010). Furthermore, field experiments found that local *H. a. texanus* had higher fitness than *H. a. annuus* when grown in Texas, and that fitness-related trait values were shifted towards *H. debilis* (Whitney et al., 2006, 2010). In those experiments, the pattern of selection favouring local trait values was also seen in BC1 hybrids between *H. a. annuus* and *H. debilis*, supporting the adaptive potential of introgression. Finally, an eight-year field experimental evolution study found that *H. a. annuus* × *H. debilis* hybrids (synthesized to mimic the putative early ancestors of *H. a. texanus*) rapidly evolved higher fitness than nonhybrid *H. a. annuus* controls when exposed to the central Texas environment (Mitchell et al., 2019). Together these results suggest that *H. a. texanus* has introgressed ancestry from *H. debilis* and that these alleles should control adaptive and convergent morphological traits.

However, a phylogenetic analysis of the annual sunflowers based on genotyping by sequencing (GBS) data failed to find a significant signal of introgression with *H. debilis* using the ABBA-BABA test, but suggested the occurrence of introgression with another local congener, *Helianthus argophyllus* (Baute et al., 2016). Although Heiser recognized that *H. argophyllus* was a possible donor due to its location and the presence of hybrids with *H. annuus*, the lack of morphological convergence between *H. argophyllus* and *H. a. texanus* made *H. debilis* seem a more likely choice (Heiser, 1951b). Finally, a recent study of the role of structural variation in local adaptation in wild sunflowers found several haploblocks (large regions of suppressed recombination, mostly associated with chromosomal inversions) that differed in frequency between Texas *H. annuus* and the rest of the range, suggestive of a possible role in the formation of *H. a. texanus* (Todesco et al., 2020).

In this study, we used common garden phenotypic measurements and genomic sequencing to understand what defines the locally adaptive subspecies, *H. a. texanus*. Our goal is three-fold, to definitely settle whether introgression has occurred in *H. a. texanus*, identify the donor(s) of any introgressions that are detected, and assess how local adaptation has evolved in *H. a. texanus*.

2 | MATERIALS AND METHODS

2.1 | Morphological analysis

Before evaluating the genetics of *H. a. texanus*, we first reassessed whether *H. annuus* samples from Texas were morphologically distinct from samples from the rest of the species range. The original subspecies designation by Heiser (1954) suggested that *H. a. texanus* was found in south-eastern Texas, and graded into other subspecies in northern and western Texas. Taking this observation into account, we divided our samples into "South Texas" samples below 30° latitude, "North Texas" samples above 30° latitude but within Texas, and "non-Texas", including all others. We assessed the morphology using data collected from a previously reported common garden of *H. annuus* (Todesco et al., 2020). Briefly, 614 *H. annuus* from 63 populations across the native range were grown in Vancouver, BC, Canada and measured for a range of morphological and phenological traits.

Based on the original description of *H. a. texanus*, we examined traits that were expected to differ between *H. a. annuus* and *H. a. texanus*, including stem colour, flower head size and seed size. We then used R to conduct a Mann-Whitney test for each variable asking if the southern and/or northern Texas populations differed from the rest of the species (R Core Team, 2013). To see if Texas samples are exceptional in general, we ran a principal component analysis (PCA) of all traits using FactoMineR, and missing data imputed with MissMDA (Josse & Husson, 2016; Lê et al., 2008). Based on this analysis, when directly comparing *H. a. annuus* and *H. a. texanus* as a whole, we included southern Texas samples (identifying them as *H. a. texanus*) and the non-Texas samples (identifying them as *H. a. annuus*) and excluded intermediate north Texas samples. When analysing individual samples, we included all Texas *H. annuus* to probe possible geographic patterns. Figure colours were chosen from the PNWColors palettes (Lawlor, 2020).

2.2 | Population genetics

We used genomic data to determine how genetically differentiated *H. a. texanus* is from *H. a. annuus*. From a previously published whole genome resequencing data sets (~6.34× mean coverage), we created two sets of samples. One *annuus*-specific set included all *H. annuus* samples spanning the entire range. The second, multispecies set included all *H. annuus* samples from Texas and one random sample per population for the remaining wild *H. annuus*, as well as from the annual species *H. argophyllus*, *H. debilis*, *Helianthus petiolaris* subsp. *fallax*, *H. petiolaris* subsp. *petiolaris* and *Helianthus niveus* subsp. *canescens* (Figure 1a,b; Todesco et al., 2020). For the multispecies set we also selected four outgroup perennial samples, one from each of *Helianthus divaricatus*, *Helianthus giganteus*, *Helianthus decapetalus* and *Helianthus grosseserratus*. Samples from each set were variant-called together using the same pipeline as Todesco et al. (2020). While it would be preferable to variant-call all samples together, the level of sequence variation and repetitive sequence caused issues with GATK at high sample number, leading us to use two separate data sets. All sample information, including SRA accession numbers for raw sequence data, are collated in Supporting Information files S1 and S2. Briefly, samples were trimmed using Trimmomatic, aligned to the *H. annuus* XRQv1 reference genome using NextGenMap, and variant-called using GATK. Variants were filtered using the GATK VQSR using a set of 67 cultivated *H. annuus* as the truth set, and the 90% tranche was retained (Badouin et al., 2017; Bolger et al., 2014; Sedlazeck et al., 2013; Van der Auwera et al., 2013). The data sets were further filtered to retain only biallelic SNPs with minor allele frequency ≥1% and a genotyping rate ≥90%. Variants were then remapped to the *H. annuus* Ha412HOv2 using BWA, as this has been shown to dramatically improve SNP ordering (Todesco et al., 2020). Finally, the data set was phased and imputed using Beagle (Browning & Browning, 2007). The data set was subset for further analyses using bcftools and converted to specific program input formats using plink and custom perl scripts (Li, 2011; Purcell et al., 2007). We used haploblock genotype calls from Todesco et al. (2020) based on diagnostic markers for each haploblock. All scripts are available at https://github.com/owensgl/texanus_ancestry.

We sought to understand the genomic landscape of differentiation between *H. annuus* species and their Texas congeners. Using the multispecies data set, we calculated genome-wide Weir and Cockerham F_{ST} between *H. a. texanus*/*H. a. annuus* and *H. argophyllus*/*H. debilis* using a custom perl script (Weir & Cockerham, 1984), requiring a minor allele frequency ≥1% for each locus in tested samples. For the multispecies data set, our sampling strategy was unbalanced and included fewer samples but more populations for *H. a. annuus* compared to *H. a. texanus* (59 samples in 59 populations vs. 109 samples in 11 populations), and this combined with minor allele frequency filters may bias results (i.e., if rare population-specific alleles are being preferentially retained in one set). To combat this, when comparing interspecies F_{ST} scores between *H. a. annuus* and *H. a. texanus*, we filtered loci so that both comparisons had an F_{ST} value (that is, the site is variable and above MAF cutoffs in both *H. annuus* groups). Additionally, we subsampled down to 11 samples from different populations for each *H. annuus* group, required a minor allele frequency ≥5% and repeated F_{ST} calculations. We then visualized these results in 1 Mbp nonoverlapping sliding windows, with at least 10 loci, by summing the numerator and denominator of F_{ST} within the window, using the tidyverse library in R (Wickham et al., 2019). We tested if there was a significant difference in window F_{ST} between the *H. annuus* groups and *H. argophyllus* or *H. debilis* using a paired *t* test. Additionally, we imputed the recombination rate for each F_{ST} window using a consensus genetic map built from five domestic *H. annuus* genetic maps (Todesco et al., 2020). We visualized F_{ST} in five quantiles divided by ranked recombination rate.

The previous analyses were done using the multispecies set to include comparisons with interspecies comparisons. We also calculated F_{ST} between *H. a. texanus* and *H. a. annuus*, using the annuus-specific set which includes all samples. This allowed us to more accurately identify regions of genetic differentiation between the species. This

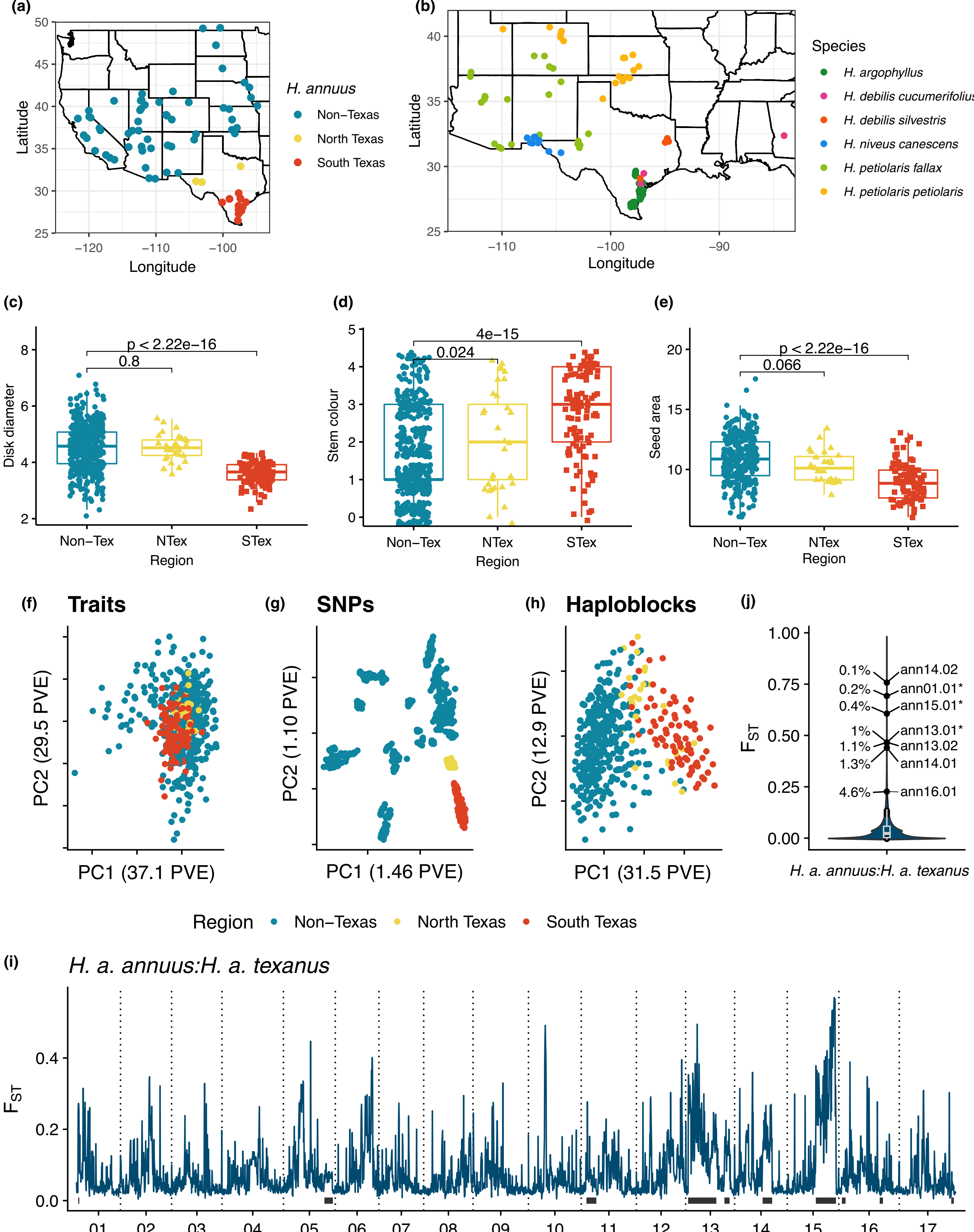

(a)
Latitude
Longitude
H. annuus
Non−Texas
North Texas
South Texas
(b)
Species
H. argophyllus
H. debilis cucumerifolius
H. debilis silvestris
H. niveus canescens
H. petiolaris fallax
H. petiolaris petiolaris
(c)
Disk diameter
p < 2.22e−16
0.8
Non−Tex
NTex
STex
Region
(d)
Stem colour
4e−15
0.024
(e)
Seed area
p < 2.22e−16
0.066
(f) Traits
PC2 (29.5 PVE)
PC1 (37.1 PVE)
(g) SNPs
PC2 (1.10 PVE)
PC1 (1.46 PVE)
(h) Haploblocks
PC2 (12.9 PVE)
PC1 (31.5 PVE)
(j)
0.1% ann14.02
0.2% ann01.01*
0.4% ann15.01*
1% ann13.01*
1.1% ann13.02
1.3% ann14.01
4.6% ann16.01
H. a. annuus:H. a. texanus
Region Non−Texas North Texas South Texas
(i) H. a. annuus:H. a. texanus
Chromosome

FIGURE 1 Sampling, morphology and population structure in *H. annuus*. (a) *H. annuus* samples from across South, central and West USA. (b) Samples from other annual *Helianthus* species. (c–e) Diagnostic traits for *H. a. texanus* differ between regions. *p*-values are from Mann-Whitney tests. (f–h) PCA analysis of traits, SNPs and haploblocks. (i) The genome-wide pattern of genetic differentiation between *H. annuus* subspecies, in 1MBp windows. Haploblock locations are highlighted in black. J. Haploblock F_{ST} in comparison to SNPs outside of haploblocks. * indicates lowest BayeScan q value [Colour figure can be viewed at wileyonlinelibrary.com]

was done for both SNPs and for the 11 *H. annuus* haploblocks. When comparing SNPs to haploblocks, we excluded all SNPs within known haploblocks because they would represent the same genomic region. Since haploblocks had high relatively F_{ST}, we ran BayeScan to see if this is consistent with selection. We filtered our haploblocks-region free variant set for minor allele frequency >5% to match haploblocks allele frequencies and randomly subsampled down to 5% of remaining SNPs (76,185 sites). We did this because the full set was too large to run on BayeScan and because we only intend to see if haploblocks are exceptional, rather than identify selected SNPs.

To understand population structure across the range of *H. annuus*, we conducted a principal component analysis on our *annuus*-specific data set using SNPRelate (Zheng et al., 2012). We pruned variants for LD (r^2 < .2) in a sliding window using snpgdsLDpruning. We then replicated this analysis using the 11 *H. annuus* haploblocks as loci and to understand if the distribution of haploblock alleles matched the genome wide pattern. To explore whether haploblocks are associated with the phenotypic and environmental adaptations of *H. a. texanus*, we used genome wide association values and genome environment association values from Todesco et al. (2020). These values were calculated using EMMAX and BayENV (Günther & Coop, 2013; Zhou & Stephens, 2012).

2.3 | Genome wide introgression analysis

To further understand the genetic differences between *H. annuus* subspecies, we looked for introgression into *H. a. texanus* using multiple methods. We first calculated Patterson's *D* using Admixr, an R wrapper for Admixtools (Patterson et al., 2012; Petr et al., 2019). *D*, also known as the ABBA-BABA test, asks if there is a greater number of shared derived alleles between taxa than expected under incomplete lineage sorting. In this case, we asked if *H. a. texanus* shares more derived alleles with a donor species (e.g., *H. debilis*) than *H. a. annuus*, which is allopatric and should not be affected by gene flow (although see below). We used three possible donor species: *H. debilis*, the hypothesized donor, *H. argophyllus*, and *H. petiolaris fallax*.

One possible confounding feature of *D* is that when there are multiple hybridization events in different parts of the tree, the signal of specific introgression events can be masked. Previous studies have suggested ongoing gene flow between *H. annuus* and *H. petiolaris*, and we expect that this most recently occurs in the non-Texas populations of *H. a. annuus* because *H. petiolaris* has a limited range in Texas. Since *H. debilis* is closely related to *H. petiolaris*, introgression between *H. a. annuus* and *H. petiolaris* may mask introgression from *H. a. texanus* and *H. debilis* (see Section 4). To explore this possible issue, we also used *H. petiolaris fallax* as a possible donor species.

For each *H. a. texanus* sample, we calculated *D* using the arrangement [a single *H. a. texanus* sample, all *H. a. annuus*, potential donor, all perennials]. We ran this independently using *H. argophyllus*, *H. debilis or H. petiolaris fallax* as the potential donor in position 3. We used the default (1 Mbp) window size for block bootstrapping. This allowed us to look for variation in introgression between individuals to determine if introgression is present in all *H. a. texanus* samples. From this, we found that *D* scores using *H. debilis* or *H. argophyllus* as a donor appeared correlated, so we calculated the correlation between *D* scores using these two species. We also ran the following arrangements to test for *H. argophyllus* – *H. debilis* introgression using the following sets: (*H. debilis*, *H. petiolaris fallax*, *H. argophyllus*, perennials), (*H. a. texanus*, *H. argophyllus*, *H. debilis*, perennials), (*H. debilis*, *H. petiolaris fallax*, *H. a. texanus*, perennials) and (*H. debilis*, *H. petiolaris fallax*, *H. a. annuus*, perennials). For tests comparing whole species, rather than populations, we repeated these tests using *H. petiolaris petiolaris* instead of *H. petiolaris fallax* to check for subspecies effects.

Genomic admixture can also be detected using programs for identifying population structure, such as ADMIXTURE (Alexander & Lange, 2011). To remove linkage, we used SNPrelate to prune for markers with high linkage (r^2 ≥ .2 within 500 Kbp) and selected all *H. annuus*, *H. argophyllus* and *H. debilis* samples from our multi-species dataset (Zheng et al., 2012). We ran ADMIXTURE with *K* from 1 to 10 with 200 bootstrap replicates, and cross validation to determine the best fit *K*. We then used CLUMPAK to synchronize groupings between *K* values (Kopelman et al., 2012). We focus on results at *K* = 5, which was the minimum *K* value that best separated known species. To confirm ADMIXTURE results, we ran STRUCTURE with *K* = 5 for 20 replicates each randomly subset to 5% of the total sites due to computational limitations (Pritchard et al., 2000). These results were also synchronized with CLUMPAK and we report the two major output patterns.

Another way of looking at gene flow between species is through the covariance of allele frequencies, which can be produced through shared drift or gene flow. TreeMix creates a tree of populations based on shared genetic drift and adds migration events to explain exceptional shared drift (Pickrell & Pritchard, 2012). This method requires unlinked markers, so we used SNPrelate to prune out markers with high linkage as above. For this analysis, we included all species in our data set and divided *H. annuus* into three groups: non-Texas (*H. a. annuus*), South Texas (*H. a. texanus*), and North Texas (intermediate). We also included *H. niveus canescens* to better represent the full annual clade. Each tree was rooted with the four perennial species which were grouped into a single population for simplicity. We plotted the residual fit from the maximum likelihood tree by dividing the residual covariance between pairs of populations by the average standard error across all pairs. We selected an appropriate number of

migration edges based on the decay in likelihood improvements with successive numbers.

Since the initial TreeMix analyses did not include a migration branch from *H. debilis* into south Texas *H. annuus*, we specifically tested the likelihood of migration from *H. debilis* or from *H. argophyllus*. For each potential donor, we manually added a single migration event, with 0.001–0.4 admixture proportion, and recorded the likelihood. We plot the likelihood relative to the highest likelihood proportion.

Lastly, we calculated intercross ancestry by selecting fixed differences between *H. a. annuus* and *H. argophyllus* or *H. debilis*. From these sites we measured ancestry (i.e., the proportion of alleles from each species) and heterozygosity for all *H. a. texanus* samples (e.g., Fitzpatrick, 2012).

2.4 | Introgression candidates

After testing for full genome signals of introgression, we next looked for localized genomic patterns of introgression. To do this we first measured D and f_d using a custom perl script (Martin et al., 2015). This approach uses allele frequencies instead of genotypes from individual samples, and allows us to extract window values. We visualized this in 100 SNP nonoverlapping windows and selected the top 1% of windows based on f_d as outliers and potential introgressed regions. We used f_d instead of D because it is more robust to low diversity regions (Martin et al., 2015). We also ran this analysis for a single *H. a. texanus* sample (ANN1363) that was identified as having extra *H. debilis* ancestry in admixture and genome wide D.

To see if differences in F_{ST} between *H. annuus* subspecies correspond to f_d scores, we compared F_{ST} to f_d in 100 kb windows. For this, we calculated the correlation between f_d with *H. debilis* as donor and the difference between F_{ST} (*H. a. annuus – H. debilis*) and F_{ST} (*H. a. texanus – H. debilis*). We repeated this with *H. argophyllus* as possible donor and comparator in F_{ST}.

Another way of detecting introgression is detecting genomic regions where samples do not genetically cluster with others of their species. We applied this approach using PCAdmix (Brisbin et al., 2012). In this analysis, we included all samples from *H. a. annuus*, *H. argophyllus* and *H. debilis* as parental populations and all *H. a. texanus* samples as potentially admixed. Genetic map positions were imputed from a 1 Mbp resolution genetic map for Ha412HOv2 (Todesco et al., 2020). In regions with zero recombination across 1 Mbp, we smoothed the cM position across the nearest neighbouring positions with a nonzero cM/Mbp rate. We filtered for linkage and used 100 SNP windows (–r^2 .8 –w 100). We selected windows with *H. debilis* ancestry in >40% of chromosomes as being candidate introgression regions. As a control, we reran the analysis pulling out a single allopatric *H. a. annuus* sample and testing it as an admixed sample, repeating for each *H. a. annuus* sample.

For both f_d and PCAdmix, we explored the role of recombination by plotting the f_d scores or introgression frequency in windows grouped by recombination rate quantile.

3 | RESULTS

3.1 | Morphology

A comparison of samples from across the range of *H. annuus* (Figure 1a) identified patterns of phenotypic variation consistent with the existing subspecies description (Heiser, 1954). *H. annuus* samples from south Texas tend to begin branching closer to the ground, have smaller inflorescences (measured as the diameter of the central disk), more purplish stem coloration, and have smaller seeds (Figure 1c–e, Figure S1), although their values are within the range found across the range of non-Texas *H. annuus*. This is seen in the trait PCA, which locates Texas samples along a restricted portion of the first PC axis, while non-Texas samples span the entire axis (Figure 1f). North Texas samples are often morphologically intermediate and the exact geographic range of *H. a. texanus* is not defined (Figure 1c–f, Figure S1), so moving forward we excluded these samples when directly comparing *H. a. texanus* to *H. a. annuus* as a whole. This ensures that tests comparing *H. annuus* subspecies only include unambiguous samples. These samples are included in analysis that do not make this direct comparison (i.e., TreeMix, admixture and structure).

3.2 | Population genetics

Using genome-wide SNPs, we found geographic population structure across the range of *H. annuus*. South Texas populations could be differentiated based on the first two PCs but were not dramatically different from other populations (Figure 1g). In contrast, when we use genotype information at 11 recently described putative segregating structural variants (haploblocks; Todesco et al., 2020), the major axis of variation divides South Texas and non-Texas populations, with North Texas populations falling intermediate (Figure 1h). Consistent with this, we found that *H. a. texanus* was only mildly differentiated from *H. a. annuus* (F_{ST} = 0.0799; 0.0043–0.57 in 1 Mbp windows, Figure 1i), with regions of high differentiation often colocalizing with haploblocks. As previously reported, when haploblocks were treated as loci, seven out of 11 fall within the top 5% of highest F_{ST} values for syntenic SNPs and three have the lowest possible BayeScan q-value (0) along with 1.1% of SNPs (Figure 1j). For these haploblocks, North Texas populations often had intermediate frequency between South Texas and non-Texas populations (Figure S2).

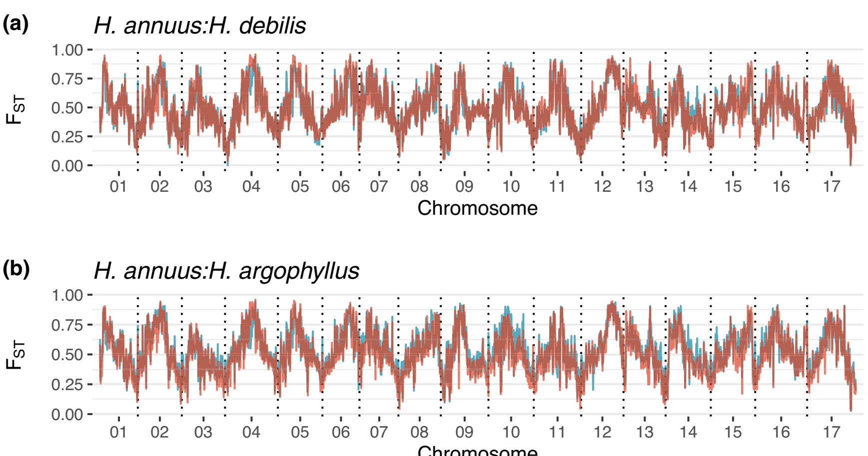
For both subspecies of *H. annuus*, there was high differentiation from *H. argophyllus* and *H. debilis*. F_{ST} was significantly lower for *H. a. texanus* samples than for *H. a. annuus* samples when both were compared to *H. argophyllus* (paired $t[3093] = 30.923$, $p \le 2.2e{-}16$), but only marginally lower when compared to *H. debilis* (paired $t[3082] = 1.7514$, $p = .08$]) (Figure 2a–c). When tested using a subset of samples, so both *H. annuus* subspecies had equal numbers of populations and samples, we obtained qualitatively similar results, although differences between *H. a. annuus* and *H. a. texanus* were even less pronounced in the *H. debilis* comparison; *H. argophyllus* (paired

$t[3075] = 18.871$, $p \le 2.2e{-}16$) and *H. debilis* (paired $t[3074] = -0.5312$, $p = .5$]). We further found that F_{ST} was lowest in regions of high recombination in both comparisons (Figure 2d,e).

3.3 | Introgression detection via ABBA-BABA tests

We calculated Patterson's *D* for each Texas *H. annuus* sample, looking for introgression from *H. argophyllus*, *H. debilis* and *H. petiolaris fallax*. We included *H. petiolaris fallax* because it is a relative of *H. debilis* that is known to hybridize and exchange genes with *H. a. annuus*, which may affect *D* tests of other species (see Section 4) (Kane et al., 2009). A majority of Texas *H. annuus* samples had a significant signal of introgression from *H. argophyllus*, and particularly high values were found in three coastal populations: ANN_55, ANN_56 and ANN_57 (Figure 3a,b). In contrast, a majority of samples had negative scores for *H. debilis* introgression (suggesting, surprisingly, greater *H. debilis* allele sharing with allopatric *H. a. annuus* than with *H. a. texanus*). Higher *H. debilis D* values occurred in samples with high *H. argophyllus D* scores; *H. argophyllus D* significantly explained *H. debilis D* ($R^2 = .6585$, $p < e^{-15}$), suggesting that *H. argophyllus* is acting as a bridge for *H. debilis* ancestry or that *H. argophyllus* ancestry is being mistaken as *H. debilis* (Figure 3c). Lastly, almost all *H. a. texanus* samples had a negative *H. petiolaris fallax D*, suggesting greater allele sharing between the parapatric *H. a. annuus* and *H. petiolaris fallax* than between the latter and *H. a. texanus*. It is possible that this introgression is masking a possible signal of introgression between *H. a. texanus* and *H. debilis* (See Section 4. Figure 3d,e). We did find one *H. a. texanus* sample, ANN1363, with positive *H. debilis D* not accompanied by higher *H. argophyllus D* (Figure 3a,c). We calculated *D* in genomic windows using just this sample and found localized evidence of introgression on a large portion of chromosome 3, suggesting this is an advanced generation backcross individual (Figure S3).

We also tested for introgression among Texas annual sunflower species outside of the focal *H. a. texanus* system (Table 1). We found that there was excess allele sharing between *H. argophyllus* and *H. debilis*, consistent with the admixture signals seen in ADMIXTURE (see below). Despite this, there is a greater signal of excess allele



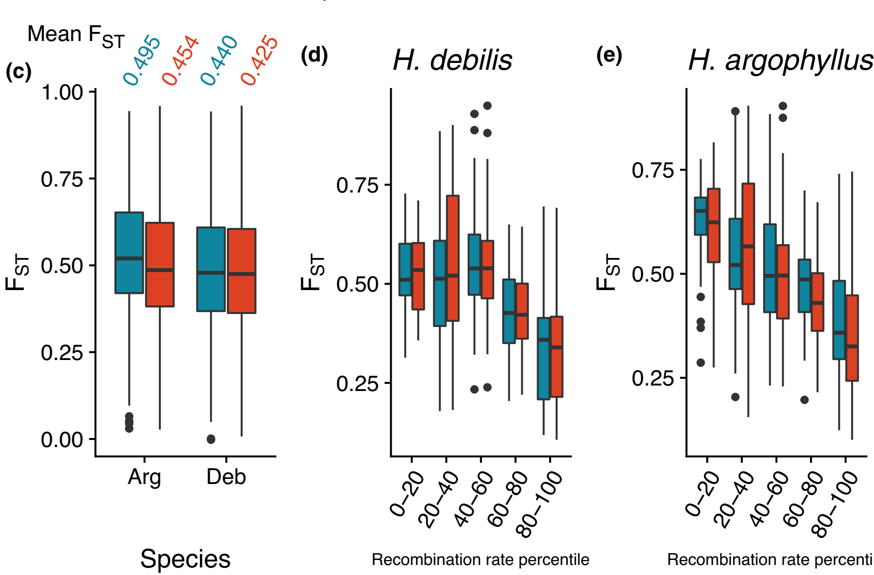
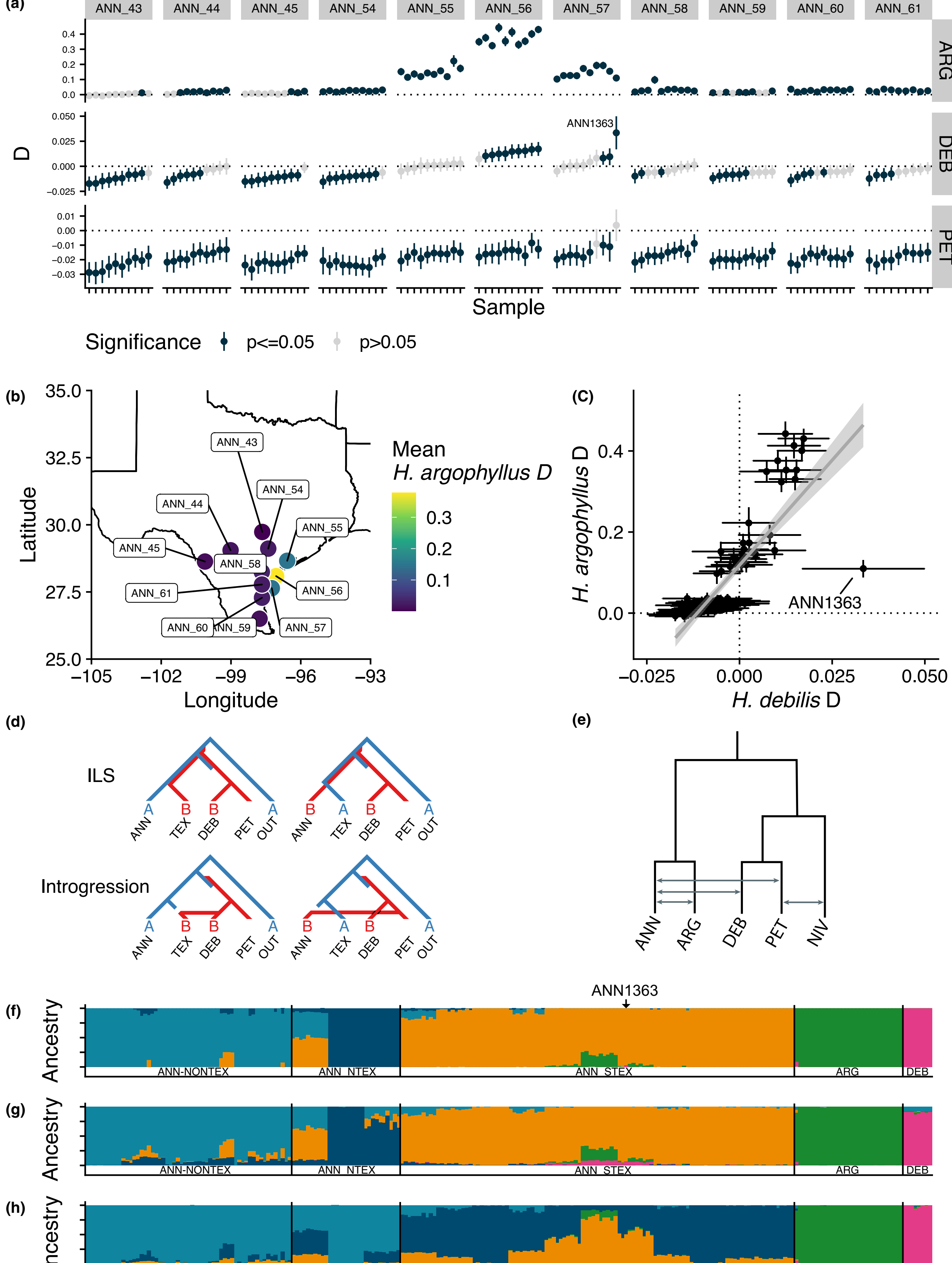


FIGURE 2 F_{ST} between *H. a. annuus*, *H. a. texanus* and congeners. (a) F_{ST} between *H. annuus* and *H. debilis* in 1 Mbp windows. (b) F_{ST} between *H. annuus* and *H. argophyllus* in 1 Mbp windows. (c) Boxplot of F_{ST} in 1 Mbp windows, with genome-wide F_{ST}. (d–e) Comparison of F_{ST} with recombination rate based on a *H. annuus* genetic map. F_{ST} windows are binned into 20% percentile groups based on recombination rate [Colour figure can be viewed at wileyonlinelibrary.com]



(a)
ANN_43
ANN_44
ANN_45
ANN_54
ANN_55
ANN_56
ANN_57
ANN_58
ANN_59
ANN_60
ANN_61
ARG
DEB
PET
D
ANN1363
Sample
Significance p<=0.05 p>0.05
(b)
Latitude
Longitude
Mean
H. argophyllus D
(C)
H. argophyllus D
H. debilis D
ANN1363
(d)
ILS
Introgression
A B B A
ANN TEX DEB PET OUT
(e)
ANN ARG DEB PET NIV
(f)
(g)
(h)
Ancestry
ANN1363
ANN-NONTEX
ANN_NTEX
ANN_STEX
ARG
DEB

FIGURE 3 Measures of admixture in *H. a. texanus*. (a) Patterson's *D* with 95% bootstrap confidence intervals for *H. a. texanus* samples with three possible donor species. ARG: *H. argophyllus*, DEB: *H. debilis*, PET: *H. petiolaris fallax*. (b) Mean Patterson's *D* for *H. a. texanus* populations with *H. argophyllus* as the possible donor. (c) The relationship between Patterson's *D* for *H. a. texanus* samples using *H. argophyllus* and *H. debilis* as donors. Error bars represent 95% bootstrap confidence interval. (d) Phylogenetic representation of *D* test for showing how *H. petiolaris* – *H. a. annuus* introgression could affect tests for introgression from *H. debilis*. (e) The phylogenetic relationship between annual sunflowers with documented introgression events highlighted. (f) ADMIXTURE results for *H. annuus*, *H. argophyllus* and *H. debilis* samples with $K = 5$. (g) STRUCTURE results for 11/20 replicates. (h) STRUCTURE results for 9/20 replicates[Colour figure can be viewed at wileyonlinelibrary.com]

sharing between the allopatric *H. petiolaris* and *H. argophyllus* than between *H. debilis* and *H. argophyllus*. This is also true when using *H. a. annuus* or *H. a. texanus* instead of *H. argophyllus*, suggesting that this may reflect known introgression between *H. annuus* and *H. petiolaris* (Yatabe et al., 2007; see Section 4).

3.4 | Admixture detection via ADMIXTURE, STRUCTURE, and diagnostic markers

ADMIXTURE analysis using the multispecies set best matched species designations with five groups, one primarily in each of non-Texas *H. annuus*, north Texas *H. annuus*, south Texas *H. annuus*, *H. argophyllus* and *H. debilis*. At this number of groups we saw *H. argophyllus* ancestry in three populations of *H. a. texanus* previously identified using *D*, as well as *H. debilis* ancestry (~2%) in ANN1363 (Figure 3f). We did not see any additional ancestry from *H. debilis* in *H. a. texanus*, at this or any other *K* value (Figure S4). In contrast to ADMIXTURE, STRUCTURE produced two nearly equally supported ancestry scenarios. In one scenario, *H. debilis* was largely its own cluster but contained ancestry from non-Texas *H. annuus*, and *H. debilis* ancestry was found in three *H. a. texanus* populations that also contain *H. argophyllus* ancestry. In the other scenario, there is no *H. debilis* ancestry in *H. a. texanus* except for the previously identified ANN1363, which has ~2% (Figure 3g,h).

Ancestry in *H. annuus* samples were partitioned different amongst three groups in our ADMIXTURE and two supported STRUCTURE results. A PCA on the full *H. annuus* data set suggests that there is continuous geographic population structure, so we interpreted the ancestry estimates as reflecting the difficulty of partitioning ancestry when variation is mostly continuous and not discrete.

Based on diagnostic markers fixed between *H. a. annuus* and *H. argophyllus*, we found between 0.6% and 23% *H. argophyllus* ancestry in *H. a. texanus* samples. Samples with higher *H. argophyllus* ancestry also had higher diagnostic marker heterozygosity suggestive that the admixture is relatively young and has not had time for *H. argophyllus* ancestry to fix (Figure S5). When using *H. debilis* as a possible parent, we found between 0.6 and 4% *H. debilis* ancestry for *H. a. texanus* samples (Figure S5).

3.5 | Introgression detection via TREEMIX

We explored gene flow between multiple species using TREEMIX, which calculates a phylogeny of populations based on shared drift and tests whether migration edges (i.e., introgression) improve the model fit. It found that likelihood improvements declined after five migration edges so we present that value here (Figure 4a; Figure S6). At five migration edges, TREEMIX found evidence for gene flow from *H. petiolaris* into *H. debilis*, *H. petiolaris fallax* into *H. niveus canescens*, *H. debilis* into *H. argophyllus*, *H. argophyllus* into *H. a. texanus* and an ancestral node into *H. petiolaris fallax* (Figure 4a). The overall tree is largely consistent with previously described phylogeny, except *H. debilis* is thought to be sister to *H. petiolaris*, probably explaining the strongest migration edge bringing those two together (Stephens et al., 2015). Even at 10 migration edges, TREEMIX does not suggest migration from *H. debilis* into *H. a. texanus* (Figure S6). TREEMIX selects migration edges based on additional positive residual covariance after the initial tree but we see that the residual covariance with *H. debilis* and *H. a. texanus* is actually negative and less in *H. a. texanus* than in other *H. annuus* subspecies (Figure 4b). When specifically tested, the most likely migration edge between *H. debilis* and *H. a. texanus* is 0. In contrast, the migration edge between *H. argophyllus* and *H. a. texanus* has the highest support at 0.19 (Figure 4c).

3.6 | Localized introgression detection via *D*, f_d and PCADMIX

Across the genome we see much more variation and more positive values of *D* when *H. argophyllus* is used as a donor instead of *H. debilis* (Figure S3). For *H. debilis*, there are two large regions with negative *D* scores which correspond to haploblocks on chromosomes 5 and 13 (ann05.01, and ann13.02; Figure S3) (Todesco et al., 2020).

At a total genome level, PCADMIX was consistent with *D* and showed high levels of *H. argophyllus* admixture in three populations of *H. a. texanus* (Figure 4d). The inferred *H. debilis* introgression was much lower (~0.5%), although these values are higher than most, but not all, of the "control" non-Texas *H. a. annuus* samples. Across the genome, inferred admixture was more common in the higher recombination regions on the ends of chromosomes, strongly for *H. argophyllus* and very slightly for *H. debilis* (Figure 4e). We also found that f_d scores were higher for high recombination regions when *H. argophyllus* was the potential donor, but not when *H. debilis* was the donor (Figure 4f).

We found that there was a positive correlation between the difference in F_{ST} and f_d when testing with *H. argophyllus* but not *H. debilis* ($r = .32$, $p < e^{-4}$; $r = -.13$, $p = .08$). The positive correlation with *H. argophyllus* means that regions with lower F_{ST} between *H. a. texanus* and a *H. argophyllus* also tended to share more derived alleles.

Based on our methods, we also attempted to identify candidate introgressed regions. From f_d measures testing for *H. debilis* gene

TABLE 1 *D*-statistic among Texas annual sunflowers

P1	P2	P3	P4	*D*	*p*
H. argophyllus	*H. a. texanus*	*H. debilis*	Perennials	−0.049	$<e^{-21}$
H. p. fallax	*H. debils*	*H. argophyllus*	Perennials	−0.062	$<e^{-21}$
H. p. petiolaris	*H. debils*	*H. argophyllus*	Perennials	−0.092	$<e^{-39}$
H. p. fallax	*H. debils*	*H. a. annuus*	Perennials	−0.090	$<e^{-50}$
H. p. petiolaris	*H. debils*	*H. a. annuus*	Perennials	−0.122	$<e^{-72}$
H. p. fallax	*H. debils*	*H. a. texanus*	Perennials	−0.081	$<e^{-43}$
H. p. petiolaris	*H. debils*	*H. a. texanus*	Perennials	−0.112	$<e^{-64}$

Note: Negative *D* values indicate greater allele sharing between P1 and P3.

flow into *H. a. texanus*, we selected the top 1% of windows, representing 101 regions spanning 49.3 Mbp. For PCAdmix, we required >40% *H. debilis* ancestry across all *H. a. texanus* samples, resulting in only 13 regions spanning 1.0 Mbp being selected (Figure S7, Supporting Information S3, S4). Between these two measures for detecting *H. debilis* introgression, we found two regions of overlap on chromosomes 13 and 15. The only gene found in these regions is Ha412HOChr13g0625071 on chromosome 13. This gene shows similarity to anthocyanidin 3-O-glucosyltransferases, enzymes which are known to play a major role in the accumulation of anthocyanin pigments (Saito et al., 2013). However, when we tested the association between *H. debilis* ancestry from PCAdmix at this region and stem colour in *H. a. texanus* samples using a linear model, we found an insignificant relationship (β = 0.21, *p* = 0.09, adjusted R^2 = 0.016) suggesting that introgression in this region is unlikely to play a major role in the phenotypic differences between the subspecies.

4 | DISCUSSION

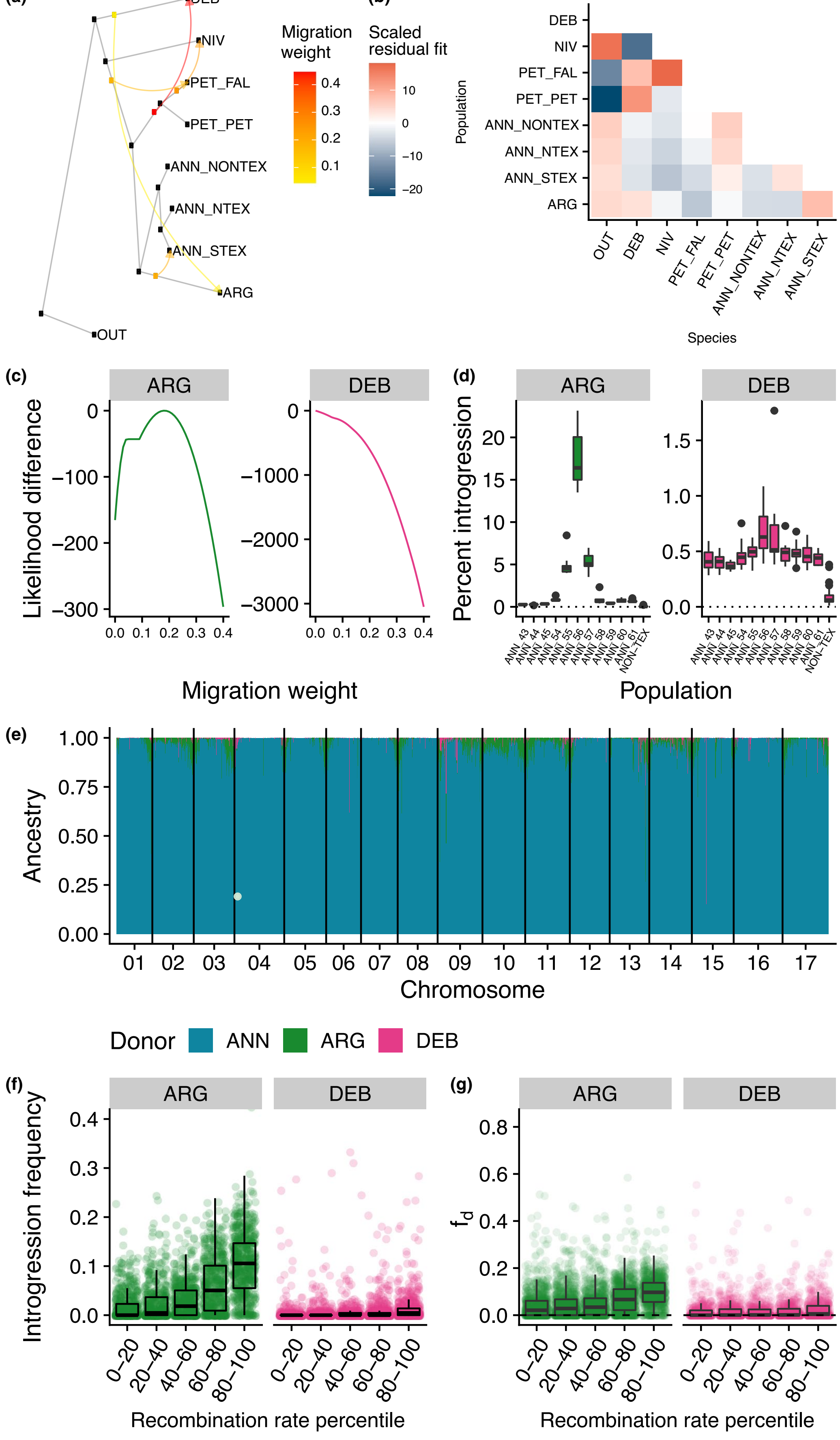
4.1 | Subspecies differentiation in *H. annuus*

Heiser (1951a) hypothesized that the morphologically divergent southeastern Texas form of *H. annuus* was the result of adaptive introgression from the narrowly distributed Texas sunflower *H. debilis*, with which it is morphologically convergent (Heiser, 1951a; Whitney et al., 2006, 2010). QTL mapping of *H. annuus* × *H. debilis* hybrids identified numerous candidate loci for which *H. debilis* alleles shift the phenotype towards *H. a. texanus*, as well as three candidate regions where *H. debilis* alleles are associated with increased fitness (Whitney et al., 2015). Further, field experimental evolution trials have shown that hybridization with *H. debilis* can allow populations to more quickly adapt over generations (Mitchell et al., 2019). This evidence has until now made *H. a. texanus* one of the better-supported candidates for adaptive introgression in plants (Suarez-Gonzalez et al., 2018), although see Baute et al. (2016).

We confirm that *H. a. texanus* populations have different mean trait values than *H. a. annuus* for traits that are considered to be distinctive of this subspecies, as previously described. This phenotypic variation, however, is not unique to *H. a. texanus*, and similar values for these traits are found elsewhere in the larger population of *H. a. annuus* (Figure 1c, Figure S1). Based on aggregate genome-wide SNP variation, *H. annuus* has geographic population structure which does separate Texas populations, but not more-so than one would expect based on the geography. Perhaps more importantly, four different approaches found either very little (admixture, Patterson's *D*, PCAdmix) or no (TreeMix) introgression from *H. debilis* into *H. a. texanus*. These same approaches did detect introgression from *H. argophyllus* into *H. a. texanus*, which was not previously thought to be a significant contributor to *H. a. texanus* ancestry. These introgressed regions appear to play only a minor role in separating *H. a. texanus* from *H. a. annuus*; the strongest contributors to genetic differentiation between the two subspecies are instead large segregating haploblocks (i.e., inversions), which are unrelated to *H. debilis* or *H. argophyllus*.

The original designation of *H. a. texanus* was based on mainly floral characteristics (Heiser, 1954). Heiser recognized that there was continuous variation in morphology across the range and emphasized that the subspecies designation was descriptive and for facilitating discussion. Our results agree with this looser definition of subspecies in the system, meaning that *H. a. texanus* represents one end of a continuous spectrum. Interestingly, this is very different from the subspecies of a related species (*H. petiolaris*), which have a deep phylogenetic division on a similar scale to that of different *Helianthus* species (Todesco et al., 2020). This highlights the utility of genomic data for understanding the relatedness of subspecies and populations, which can vary dramatically.

FIGURE 4 TreeMix and PCAdmix output testing for introgression. (a) TreeMix output with five migration edges. (b) Scaled residual fit after tree but without migration edges from TreeMix. (c) Relative likelihood values for different weights of migration into *H. a. texanus*. (d) Total percent introgression (>0.9 confidence) for *H. annuus* populations from PCAdmix. Non-Texas are *H. a. annuus* samples tested using a "leave one out" strategy. (e) Total percent introgressed ancestry for all *H. a. texanus* across the genome, scaled by basepair. (f) Proportion introgressed ancestry for *H. argophyllus* (ARG) and *H. debilis* (DEB) in windows separated by recombination rate percentile. (g) f_d in 100 SNP windows separated by recombination rate percentile for the donor species *H. argophyllus* (ARG) and *H. debilis* (DEB)[Colour figure can be viewed at wileyonlinelibrary.com]


(a)
DEB
NIV
PET_FAL
PET_PET
ANN_NONTEX
ANN_NTEX
ANN_STEX
ARG
OUT
Migration weight
0.4
0.3
0.2
0.1
(b)
Scaled residual fit
10
0
−10
−20
Population
DEB
NIV
PET_FAL
PET_PET
ANN_NONTEX
ANN_NTEX
ANN_STEX
ARG
OUT
Species
(c)
ARG
DEB
Likelihood difference
Migration weight
(d)
Percent introgression
Population
(e)
Ancestry
Chromosome
Donor
ANN
ARG
DEB
(f)
Introgression frequency
Recombination rate percentile
(g)
f_d
Recombination rate percentile

4.2 | Adaptation from standing genetic variation in *H. a. texanus*

Despite our findings that *H. a. texanus* showed no consistent (across all populations) pattern of introgression from either *H. debilis* or *H. argophyllus*, it is still a locally adapted subpopulation of *H. annuus*, as evidenced by previous common garden experiments in Texas showing increased fitness relative to nonlocal *H. a. annuus* (Whitney et al., 2006, 2010). One possible reason why adaptation was not achieved through introgression is due to the strong post-zygotic reproductive isolation between *H. annuus* and *H. debilis* (~94% F_1 pollen sterility) (Chandler et al., 1986; Heiser, 1954), although even stronger barriers have not eliminated introgression between *H. annuus* and other species (e.g., Sambatti et al., 2012). Given that *H. a. texanus* is morphologically within the range of variation of *H. a. annuus*, it is more likely that adaptation could have occurred purely from selection on standing variation, rather than required an outside source. Under this model, the convergence between *H. a. texanus* and *H. debilis* is due to the common selective environment found in Texas.

One intriguing possibility is that local adaptation of *H. a. texanus* is mediated through selection on large segregating haploblocks (mainly inversions) that differentiate it from *H. a. annuus*; we found that seven of 11 haploblocks identified in *H annus* were in the top 5% of most differentiated loci, and three had the maximum possible signal of adaptation from BayeScan. To follow up, we evaluated genome-environment associations for the seven *H. a. texanus* specific haploblocks previously identified in Todesco et al. (2020). We found strong associations with continentality (i.e., the range of temperature between summer and winter) for ann13.01 and ann15.01 along with several other temperature-related associations for ann14.02 and ann15.01. These associations, based on populations across the full range of *H. annuus*, support a role for haploblocks being a source of abiotic local adaptation in *H. a. texanus*. We also evaluated the possibility that haploblocks were involved in the specific phenotypic traits that define *H. a. texanus*. To evaluate this question, we looked at genome-wide association values for haploblocks for traits that were expected to, and did, phenotypically differ between *H. a. annuus* and *H. a. texanus* (Todesco et al., 2020). For these 17 traits and 11 haploblocks, we found only eight significant ($p < .05$) associations, and only two of which were in the predicted direction (i.e., the haplotype common in *H. a. texanus* caused a shift in the phenotype towards *H. a. texanus*); ann13.01 to disk diameter, and ann14.01 to stem colour, although neither were significant after correction for multiple testing (Figure S8). Taken together, these data offer little evidence that phenotypes that identify *H. a. texanus* are caused by haploblocks, although some of the phenotypic traits that are associated to this subspecies, like resistance to herbivory, were not tested. However, they do bring new prominence to the role of inversions in environmental adaptation, as has been seen in several other systems (Hager et al., 2021; Wellenreuther & Bernatchez, 2018). Follow-up experiments should directly test phenotypic and fitness effects of Texas specific haploblocks in their home environment.

4.3 | Potential confounders for tests of introgression

Despite the relatively unambiguous result of a lack of widespread introgression into *H. a. texanus*, we have reasons to be cautious in our interpretation. In addition to the morphological convergence that prompted the initial hypothesis, there are strong reasons to expect that *H. debilis* ancestry is present in *H. a. texanus*. Both species hybridize in the wild, although with high F_1 sterility, and *H. debilis* is the source of agronomically important traits in cultivated *H. annuus*, brought in through breeding (Heiser, 1951a; Jan & Chandler, 1985). Artificial introduction of *H. debilis* genetic material to *H. a. annuus*, followed by natural selection in the wild, can result in rapid fitness increases relative to control (non-introgressed) stock (Mitchell et al., 2019), suggesting that at least some introgressed *H. debilis* regions would be favoured and retained in wild *H. a. texanus*. Thus, it is somewhat surprising that we detect so little admixture; below, we consider the evidence provided by each of the analyses we performed, and explore possible confounding factors masking the presence of *H. debilis* introgressions.

Patterson's *D* is the most direct and model-free method for detecting introgression that we used, but can be confounded by introgression in other branches of the tree. A previous study has shown that *H. a. annuus* has gene flow with *H. petiolaris* (Figure 3e, Kane et al., 2009), and our *D* scores using *H. petiolaris fallax* as the donor confirms that there is greater *H. petiolaris* introgression into the parapatric *H. a. annuus* than the allopatric *H. a. texanus*. Both *H. petiolaris* and *H. debilis* are within the same clade, so introgression from one species could be interpreted as introgression from the other (Figure 3d). Thus, a test of Patterson's *D* using *H. a. annuus*, *H. a. texanus* and *H. debilis* is, to some extent, actually asking if there is greater gene flow between *H. a. texanus* and *H. debilis* versus between *H. a. annuus* and *H. petiolaris*. The degree that this would affect the test is hard to predict and depends on the divergence between *H. petiolaris* and *H. debilis*. To properly control for this, we would need a population of *H. a. annuus* without introgression; but given the broad range of *H. petiolaris* and probably continuous gene flow, we do not know of such a population.

The signal of *H. debilis* introgression into *H. a. texanus* is strongly correlated with *H. argophyllus*' stronger introgression signal (Figure 3c). This could be caused by *H. argophyllus* acting as a bridge for *H. debillis* alleles into *H. a. texanus*. Alternatively, if *H. argophyllus* ancestry is in *H. debils*, as suggested by *D* and some TreeMix analyses, then the signal may be reflecting that both *H. debilis* and *H. a. texanus* share *H. argophyllus* ancestry. We believe it is likely that both of these hypotheses are true to some extent, but both suggest that the modest *H. debilis* allele sharing with some *H. a. texanus* populations is mediated by *H. argophyllus*. In contrast, the *H. a. texanus* sample ANN1363 does not follow the same pattern of ancestry correlations, and has a clear localized signal of introgression, suggesting that when introgression is recent, we can detect it.

This potential confounding factor of multiple introgression events is less relevant for TreeMix, which considers all species

together, so if both gene flow events are occurring, TreeMix should detect them. TreeMix suggested several introgression edges that agreed with what is known in the system. For example, *H. petiolaris fallax* and *H. niveus canescens* overlap in range, intergrade together morphologically, and have high levels of introgression according to genotyping-by-sequencing analyses (Heiser et al., 1969; Zhang et al., 2019). The introgression from *H. argophyllus* into *H. a. texanus* and from *H. debilis* into *H. argophyllus* that we observed in the TreeMix results are both also supported by our admixture and *D* results. Despite specifically testing for *H. debilis* into *H. a. texanus* introgression, TreeMix does not support any migration edge.

admixture is similarly clear that *H. debilis* is not a part of *H. a. texanus* ancestry, with the exception of ANN1363. At $K = 4$, admixture does place *H. debilis* in the same group with ANN_56, an *H. a. texanus* population that shows the highest proportion of *H. argophyllus* introgressions in ABBA-BABA tests, but this grouping is not consistent at other *K* values, or supported by other tests (Figure S4). One of two equally-probable results from structure is consistent with admixture, while the other supports some *H. debilis* admixture in the same populations that have *H. argophyllus* ancestry. In contrast with previous programs but partially consistent with one result from structure, PCAdmix found that *H. a. texanus* samples had low but higher than baseline (allopatric *H. a. annuus*) levels of *H. debilis* and *H. argophyllus* ancestry (Figure 4d). There are a few possible reasons why PCAdmix detected introgression not seen in other methods. PCAdmix is designed for recently admixed populations, so if *H. debilis* introgression is old, expectations of recombination in the PCAdmix model will probably be violated. Furthermore, PCAdmix is not robust to incomplete lineage sorting, which occurs at the high effective population sizes seen in sunflowers (Yuan et al., 2017). We attempted to use the allopatric *H. a. annuus* as a control but this is not perfect. *H. a. texanus* is genetically different from *H. a. annuus* (Figure 1i), and this difference may drive *H. a. texanus* samples into intermediate positions in the two-dimensional PCA, more so than a single non-Texas *H. a. annuus* sample, even in the absence of introgression. Our intercross ancestry estimates were consistent with values from PCAdmix, but probably had similar issues with population structure between *H. a. annuus* and *H. a. texanus*.

Taken together, it is unlikely that *H. debilis* contributed significantly to *H. a. texanus*' ancestry, but that does not necessarily mean it did not contribute at all. In both f_d and PCAdmix, we found small regions that are consistent with *H. debilis* introgression into *H. a. texanus*, although we do not know if they are due to incomplete lineage sorting (ILS) or introgression. Between PCAdmix and f_d we found two regions of overlap, which is perhaps not surprising considering both methods are looking for similar patterns. The *Arabidopsis* homologue of the single gene in those regions is involved in anthocyanidin production, although ancestry at this region does not significantly predict stem colour. Future studies to QTL map herbivore resistance in *H. a. texanus*, a proposed adaptively introgressed trait, could cross reference with this list of regions to identify candidates, or rule out introgression as the source of the trait. It is also possible that more complicated structural variation, for example, copy number variation, is poorly captured in our SNP data and is introgressed but not detected in our scans.

4.4 | Reconciling previous signals of introgression

The first study to show *H. debilis* introgression into *H. a. texanus* used cpDNA and rDNA restriction digests (Rieseberg, Beckstrom-Sternberg, et al., 1990). They found a single cpDNA pattern in non-Texas *H. annuus*, while *H. debilis* had roughly a split between a unique cpDNA pattern and the non-Texas *H. annuus* pattern. In *H. a. texanus*, 6% of the samples had the *H. debilis* pattern, while the rest had matched the *H. a. annuus* pattern. A recent study using next-generation sequencing (NGS) found two main cpDNA lineages across the annual sunflowers (Lee-Yaw et al., 2019), which appear to correspond to the different cpDNA patterns reported previously. Both cpDNA clades are found in *H. debilis* from Texas. *Helianthus annuus* primarily contains cpDNA from Clade II, but a handful of samples have Clade I cpDNA. Simulations suggested that this pattern was best explained by cytoplasmic introgression, although most probably involving different donor species. The examination of nuclear markers (ribosomal DNA, rDNA) found some *H. debilis* markers in 10% of *texanus* samples (Rieseberg, Beckstrom-Sternberg, et al., 1990). While this may represent introgression, it could also be a product of ILS. Although this study used markers that were fixed for alternate alleles in reference panels, those panels were of limited size (30 and 16) so it is possible that the markers were not actually fixed differences, but polymorphic within *H. annuus*. This chance increases in the presence of population structure, which we now know exists between *H. a. annuus* and *H. a. texanus*. To test this, we subsampled 30 random *H. a. annuus* samples and selected all sites with fixed differences between this subset and all eight sequenced *H. debilis*. We then measured the allele frequency in *H. a. texanus*. We found that for these markers that are ostensibly fixed differences, 36% of loci were at >1% minor allele frequency in *H. a. texanus*, while 4% were at >10% minor allele frequency. Thus, isolated introgression-like signals can be found in a data set with no overall pattern of introgression.

A follow-up study found higher levels of cpDNA discordance, but slightly lower levels of rDNA discordance (7%) (Rieseberg et al., 2007). At the time, ILS was dismissed because *H. annuus* and *H. debilis* are in separate, divergent clades. In our current filtered data set, non-Texas *H. annuus* and *H. debilis* share polymorphism at 18% of loci, suggesting much higher amounts of shared variation than previously appreciated.

A more recent study (Scascitelli et al., 2010) used 88 microsatellite loci and found, again, very low levels of possible introgression. Using structure, the authors found 3/150 *H. a. texanus* samples were early generation admixed with *H. debilis*, while 1/90 *H. a. annuus* samples were similarly admixed. No other *H. annuus* samples had admixture levels significantly above zero. Although this study estimated a nonzero migration rate from *H. debilis* into *H. a. texanus*, it estimated a similar rate into *H. a. annuus*, which is geographically

implausible, suggesting it did not accurately capture recent migration rates from *H. debilis* or was detecting introgression from a third species.

Altogether, previous studies found low and inconsistent *H. debilis* ancestry in *H. a. texanus*. Although this has been interpreted as evidence that *H. a. texanus* is a hybrid lineage, with more complete genomic data it seems more likely to have resulted from ILS or possibly introgression with other species, especially *H. argophyllus*.

4.5 | Introgression across the genus

In contrast with the minimal signal of introgression from *H. debilis*, there is unambiguous introgression from *H. argophyllus* into some *H. a. texanus* populations. We can see this in all except the most northern and western populations, which do not overlap with *H. argophyllus*' distribution (Figure 3b). The highest signals of introgression — up to 20% according to PCAdmix — are found in coastal populations of *H. a. texanus*. Interestingly, *H. argophyllus* presents two very distinct ecotypes (Moyers & Rieseberg, 2016). A late-flowering ecotype, which grows primarily inland, flowers later than most *H. annuus* populations, while an early-flowering ecotype's flowering time overlaps with *H. annuus*. The early-flowering ecotype occurs on the coast, where signals of introgression are strongest. This suggests that flowering time is an important reproductive barrier between these species.

The introgression from *H. argophyllus* probably contains deleterious incompatibility loci that are being selected against. We see this in the positive relationship between introgression and recombination rate (Figure 4f,g). This pattern can arise because higher recombination allows for deleterious incompatibility loci to be decoupled from the introgressed haplotype more rapidly (Nachman & Payseur, 2012; Schumer et al., 2018). That being said, we also see a negative relationship between F_{ST} and recombination rate, even when populations are allopatric (Figure 2e,f) which supports a role for linked selection (Burri, 2017). Given that sunflowers have large regions of extremely low recombination, polymorphic inversions and rampant structural changes between species, it will be important, but challenging, to account the effects of recombination when comparing genomes between species (Ostevik et al., 2020; Todesco et al., 2020).

We focused our attention on introgression from *H. argophyllus*, but it is also possible that there is introgression from *H. a. texanus* into *H. argophyllus*. A previous study found introgression of a haploblock on chromosome 6 from *H. annuus* into *H. argophyllus* is responsible for early flowering time (Todesco et al., 2020). In our data set, this was not seen in ADMIXTURE or STRUCTURE, probably because the data set was thinned for linkage and the introgression tract is in a region of low recombination, so was thinned to a small number of markers. This highlights how methods designed for unlinked markers can underrepresent signal from regions of low recombination.

In our TreeMix analysis, we failed to find a signal of *H. annuus* – *H. petiolaris* introgression, despite a history in the literature showing both hybridization and gene flow (Kane et al., 2009). Our *D* statistic analysis found much higher derived allele sharing between *H. annuus* and *H. petiolaris*, than *H. annuus* and *H. debilis*, but similar patterns when *H. argophyllus* is substituted for *H. annuus*. This suggests ancient gene flow between *H. petiolaris* and *H. annuus* that predate the *H. argophyllus* – *H. annuus* split. It is possible that TreeMix interprets this as a closer phylogenetic relationship, which could explain why *H. debilis* not placed as sister to *H. petiolaris*. We also do not fully understand how haploblocks, which can have dramatically different phylogenetic patterns, are affecting these phylogenetic and gene flow estimates. Regardless, a more detailed phylogenomic analysis of annual sunflowers is warranted to disentangle the complex hybridization patterns.

4.6 | Conclusion

The power of next-generation sequence to generate loci across the genome allows us to robustly test hypotheses originally generated using fewer markers. It is increasingly clear that different regions of the genome can have very different patterns of genetic diversity, for example due to differences in recombination rate or structural variation. Additionally, more markers allow the disentangling of processes such as incomplete lineage sorting and introgression. Here, we have used whole genome sequencing to show that, in contrast to previous expectations, *H. a. texanus* probably did not achieve local adaptation through introgression from *H. debilis*, but rather through standing variation and large segregating haploblocks. This brings new focus to the role of selection in the morphological convergence of *H. a. texanus* and *H. debilis*. Future studies should follow up on the adaptive potential of haploblocks in *H. a. texanus* by QTL mapping of phenotypic or fitness traits associated with segregating haploblocks and examining gene sequences within haploblocks to detect genes under selection.

ACKNOWLEDGEMENTS
Funding was provided by Genome Canada and Genome BC (LSARP2014-223SUN), the National Science Foundation (NSF DEB 1257965) and a Banting postdoctoral fellowship to G.L.O. We thank the editor and three anonymous reviewers for their time and efforts improving the manuscript.

CONFLICTS OF INTEREST
The authors declare no conflicts of interest.

AUTHOR CONTRIBUTIONS
Gregory L. Owens, Nora Mitchell, Kenneth D. Whitney and Loren H. Rieseberg conceived of the project. Marco Todesco and Natalia Bercovich collected genomic and phenotypic data. Gregory L. Owens and Jean-Sébastien Légaré processed sequence data. Gregory L. Owens conducted all analyses. Gregory L. Owens wrote the manuscript with contributions from all authors.

DATA AVAILABILITY STATEMENT

All raw sequence data has been made available on the Sequence Read Archive (SRA) in projectPRJNA532579. Identification numbers are listed in Supporting Information files 1 and 2.

ORCID

Gregory L. Owens https://orcid.org/0000-0002-4019-5215
Marco Todesco https://orcid.org/0000-0002-6227-4096
Nora Mitchell https://orcid.org/0000-0002-4745-438X
Kenneth D. Whitney https://orcid.org/0000-0002-2523-5469
Loren H. Rieseberg https://orcid.org/0000-0002-2712-2417

REFERENCES

Alexander, D. H., & Lange, K. (2011). Enhancements to the ADMIXTURE algorithm for individual ancestry estimation. *BMC Bioinformatics*, *12*(1), 246. https://doi.org/10.1186/1471-2105-12-246

Andrew, R. L., & Rieseberg, L. H. (2013). Divergence is focused on few genomic regions early in speciation: Incipient speciation of sunflower ecotypes. *Evolution*, *67*(9), 2468–2482. https://doi.org/10.1111/evo.12106

Arnold, B. J., Lahner, B., DaCosta, J. M., Weisman, C. M., Hollister, J. D., Salt, D. E., Bomblies, K., & Yant, L. (2016). Borrowed alleles and convergence in serpentine adaptation. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(29), 8320–8325. https://doi.org/10.1073/pnas.1600405113

Badouin, H., Gouzy, J., Grassa, C. J., Murat, F., Staton, S. E., Cottret, L., Lelandais-Brière, C., Owens, G. L., Carrère, S., Mayjonade, B., & Legrand, L. (2017). The sunflower genome provides insights into oil metabolism, flowering and Asterid evolution. *Nature*, *546*(7656), 148–152.

Barb, J. G., Bowers, J. E., Renaut, S., Rey, J. I., Knapp, S. J., Rieseberg, L. H., & Burke, J. M. (2014). Chromosomal evolution and patterns of introgression in Helianthus. *Genetics*, *197*(3), 969–979.

Baute, G. J., Owens, G. L., Bock, D. G., & Rieseberg, L. H. (2016). Genome-wide genotyping-by-sequencing data provide a high-resolution view of wild Helianthus diversity, genetic structure, and interspecies gene flow. *American Journal of Botany*, *103*(12), 2170–2177.

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Brisbin, A., Bryc, K., Byrnes, J., Zakharia, F., Omberg, L., Degenhardt, J., Reynolds, A., Ostrer, H., Mezey, J. G., & Bustamante, C. D. (2012). PCAdmix: Principal components-based assignment of ancestry along each chromosome in individuals with admixed ancestry from two or more populations. *Human Biology*, *84*(4), 343–364. https://doi.org/10.3378/027.084.0401

Browning, S. R., & Browning, B. L. (2007). Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *The American Journal of Human Genetics*, *81*(5), 1084–1097. https://doi.org/10.1086/521987

Burri, R. (2017). Interpreting differentiation landscapes in the light of long-term linked selection. *Evolution Letters*, *1*(3), 118–131. https://doi.org/10.1002/evl3.14

Chandler, J. M., Jan, C. C., & Beard, B. H. (1986). Chromosomal differentiation among the annual *Helianthus* species. *Systematic Botany*, *1*, 354–371. https://doi.org/10.2307/2419126

Fitzpatrick, B. M. (2012). Estimating ancestry and heterozygosity of hybrids using molecular markers. *BMC Evolutionary Biology*, *12*(1), 1–4. https://doi.org/10.1186/1471-2148-12-131

Günther, T., & Coop, G. (2013). Robust identification of local adaptation from allele frequencies. *Genetics*, *195*(1), 205–220. https://doi.org/10.1534/genetics.113.152462

Hager, E. R., Harringmeyer, O. S., Wooldridge, T. B., Theingi, S., Gable, J. T., McFadden, S., Neugeboren, B., Turner, K. M., & Hoekstra, H. E. (2021). A chromosomal inversion drives evolution of multiple adaptive traits in deer mice. *bioRxiv*. https://doi.org/10.1101/2021.01.21.427490

Heiser, C. B. Jr (1947). Hybridization between the sunflower species *Helianthus annuus* and *H. petiolaris*. *Evolution*, *1*(4), 249–262.

Heiser, C. B. Jr (1949). Study in the evolution of the sunflower species *Helianthus annuus* and *H. bolanderi*. *Botany-Botanique*, *23*, 157–208.

Heiser, C. B. Jr (1951). Hybridization in the annual sunflowers: *Helianthus annuus* × *H. debilis var. cucumerifolius*. *Evolution*, *1*, 42–51.

Heiser, C. B. Jr (1951). Hybridization in the annual sunflowers: *Helianthus annuus* × *H. argophyllus*. *The American Naturalist*, *85*(820), 65–72. https://doi.org/10.1086/281651

Heiser, C. B. Jr (1954). Variation and subspeciation in the common sunflower, *Helianthus annuus*. *American Midland Naturalist*, *1*, 287–305. https://doi.org/10.2307/2422222

Heiser, C. B., Smith, D. M., Clevenger, S. B., & Martin, W. C. (1969). The north American sunflowers (*Helianthus*). *Memoirs of the Torrey Botanical Club*, *22*(3), 1–218.

Heywood, J. S. (1991). Spatial analysis of genetic variation in plant populations. *Annual Review of Ecology and Systematics*, *22*(1), 335–355. https://doi.org/10.1146/annurev.es.22.110191.002003

Huang, K., Andrew, R. L., Owens, G. L., Ostevik, K. L., & Rieseberg, L. H. (2020). Multiple chromosomal inversions contribute to adaptive divergence of a dune sunflower ecotype. *Molecular Ecology*, *29*(14), 2535–2549. https://doi.org/10.1111/mec.15428

Jan, C. C., & Chandler, J. M. (1985). Transfer of powdery mildew resistance from *Helianthus debilis* Nutt. To cultivated sunflower 1. *Crop Science*, *25*(4), 664–666.

Jones, M. R., Mills, L. S., Alves, P. C., Callahan, C. M., Alves, J. M., Lafferty, D. J., Jiggins, F. M., Jensen, J. D., Melo-Ferreira, J., & Good, J. M. (2018). Adaptive introgression underlies polymorphic seasonal camouflage in snowshoe hares. *Science*, *360*(6395), 1355–1358.

Josse, J., & Husson, F. (2016). missMDA: A package for handling missing values in multivariate data analysis. *Journal of Statistical Software*, *70*(1), 1–31.

Kane, N. C., King, M. G., Barker, M. S., Raduski, A., Karrenberg, S., Yatabe, Y., Knapp, S. J., & Rieseberg, L. H. (2009). Comparative genomic and population genetic analyses indicate highly porous genomes and high levels of gene flow between divergent *Helianthus* species. *Evolution: International Journal of Organic Evolution*, *63*(8), 2061–2075.

Kantar, M. B., Sosa, C. C., Khoury, C. K., Castañeda-Álvarez, N. P., Achicanoy, H. A., Bernau, V., Kane, N. C., Marek, L., Seiler, G., & Rieseberg, L. H. (2015). Ecogeography and utility to plant breeding of the crop wild relatives of sunflower (*Helianthus annuus* L.). *Frontiers in Plant Science*, *6*, 841. https://doi.org/10.3389/fpls.2015.00841

Kirkpatrick, M., & Barton, N. (2006). Chromosome inversions, local adaptation and speciation. *Genetics*, *173*(1), 419–434. https://doi.org/10.1534/genetics.105.047985

Kopelman, N. M., Mayzel, J., Jakobsson, M., Rosenberg, N. A., & Mayrose, I. (2015). Clumpak: A program for identifying clustering modes and packaging population structure inferences across K. *Molecular Ecology Resources*, *15*(5), 1179–1191.

Lawlor, J. (2020). PNWColors: Color Palettes Inspired by Nature in the US Pacific Northwest. R package version 0.1.0. https://CRAN.R-project.org/package=PNWColors

Lê, S., Josse, J., & Husson, F. (2008). FactoMineR: An R package for multivariate analysis. *Journal of Statistical Software*, *25*(1), 1–8.

Lee-Yaw, J. A., Grassa, C. J., Joly, S., Andrew, R. L., & Rieseberg, L. H. (2019). An evaluation of alternative explanations for widespread cytonuclear discordance in annual sunflowers (Helianthus). *New Phytologist*, *221*(1), 515–526.

Li, H. (2011). A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation

from sequencing data. *Bioinformatics*, *27*(21), 2987–2993. https://doi.org/10.1093/bioinformatics/btr509

Martin, S. H., Davey, J. W., & Jiggins, C. D. (2015). Evaluating the use of ABBA–BABA statistics to locate introgressed loci. *Molecular Biology and Evolution*, *32*(1), 244–257. https://doi.org/10.1093/molbev/msu269

Martin, S. H., & Jiggins, C. D. (2017). Interpreting the genomic landscape of introgression. *Current Opinion in Genetics & Development*, *47*, 69–74. https://doi.org/10.1016/j.gde.2017.08.007

Mitchell, N., Owens, G. L., Hovick, S. M., Rieseberg, L. H., & Whitney, K. D. (2019). Hybridization speeds adaptive evolution in an eight-year field experiment. *Scientific Reports*, *9*(1), 1–2. https://doi.org/10.1038/s41598-019-43119-4

Moyers, B. T., & Rieseberg, L. H. (2016). Remarkable life history polymorphism may be evolving under divergent selection in the silverleaf sunflower. *Molecular Ecology*, *25*(16), 3817–3830. https://doi.org/10.1111/mec.13723

Nachman, M. W., & Payseur, B. A. (2012). Recombination rate variation and speciation: Theoretical predictions and empirical results from rabbits and mice. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1587), 409–421.

Ostevik, K. L., Samuk, K., & Rieseberg, L. H. (2020). Ancestral reconstruction of karyotypes reveals an exceptional rate of nonrandom chromosomal evolution in sunflower. *Genetics*, *214*(4), 1031–1145. https://doi.org/10.1534/genetics.120.303026

Owens, G. L., Baute, G. J., & Rieseberg, L. H. (2016). Revisiting a classic case of introgression: Hybridization and gene flow in Californian sunflowers. *Molecular Ecology*, *25*(11), 2630–2643.

Owens, G. L., & Rieseberg, L. H. (2014). Hybrid incompatibility is acquired faster in annual than in perennial species of sunflower and tarweed. *Evolution*, *68*(3), 893–900. https://doi.org/10.1111/evo.12297

Patterson, N., Moorjani, P., Luo, Y., Mallick, S., Rohland, N., Zhan, Y., Genschoreck, T., Webster, T., & Reich, D. (2012). Ancient admixture in human history. *Genetics*, *192*(3), 1065–1093. https://doi.org/10.1534/genetics.112.145037

Petr, M., Vernot, B., & Kelso, J. (2019). admixr—R package for reproducible analyses using ADMIXTOOLS. *Bioinformatics*, *35*(17), 3194–3195.

Pickrell, J., & Pritchard, J. (2012). Inference of population splits and mixtures from genome-wide allele frequency data. *PLOS Genetics*, *8*(11), e1002967. https://doi.org/10.1371/journal.pgen.1002967

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*(2), 945–959. https://doi.org/10.1093/genetics/155.2.945

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., Maller, J., Sklar, P., de Bakker, P. I. W., Daly, M. J., & Sham, P. C. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*, *81*(3), 559–575. https://doi.org/10.1086/519795

R Core Team (2013). *R: A language and environment for statistical computing*. R Core Team.

Rieseberg, L. H., Baird, S. J., & Desrochers, A. M. (1998). Patterns of mating in wild sunflower hybrid zones. *Evolution*, *52*(3), 713–726. https://doi.org/10.1111/j.1558-5646.1998.tb03696.x

Rieseberg, L. H., Beckstrom-Sternberg, S., & Doan, K. (1990). *Helianthus annuus* ssp. *texanus* has chloroplast DNA and nuclear ribosomal RNA genes of *Helianthus debilis* ssp. *cucumerifolius*. *Proceedings of the National Academy of Sciences of the United States of America*, *87*(2), 593–597.

Rieseberg, L. H., Kim, S. C., Randell, R. A., Whitney, K. D., Gross, B. L., Lexer, C., & Clay, K. (2007). Hybridization and the colonization of novel habitats by annual sunflowers. *Genetica*, *129*(2), 149–165. https://doi.org/10.1007/s10709-006-9011-y

Saito, K., Yonekura-Sakakibara, K., Nakabayashi, R., Higashi, Y., Yamazaki, M., Tohge, T., & Fernie, A. R. (2013). The flavonoid biosynthetic pathway in Arabidopsis: Structural and genetic diversity. *Plant Physiology and Biochemistry*, *72*, 21–34. https://doi.org/10.1016/j.plaphy.2013.02.001

Sambatti, J. B., Strasburg, J. L., Ortiz-Barrientos, D., Baack, E. J., & Rieseberg, L. H. (2012). Reconciling extremely strong barriers with high levels of gene exchange in annual sunflowers. *Evolution: International Journal of Organic Evolution*, *66*(5), 1459–1473.

Scascitelli, M., Whitney, K. D., Randell, R. A., King, M., Buerkle, C. A., & Rieseberg, L. H. (2010). Genome scan of hybridizing sunflowers from Texas (*Helianthus annuus* and *H. debilis*) reveals asymmetric patterns of introgression and small islands of genomic differentiation. *Molecular Ecology*, *19*(3), 521–541.

Schumer, M., Xu, C., Powell, D. L., Durvasula, A., Skov, L., Holland, C., Blazier, J. C., Sankararaman, S., Andolfatto, P., Rosenthal, G. G., & Przeworski, M. (2018). Natural selection interacts with recombination to shape the evolution of hybrid genomes. *Science*, *360*(6389), 656–660.

Sedlazeck, F. J., Rescheneder, P., & Von Haeseler, A. (2013). NextGenMap: Fast and accurate read mapping in highly polymorphic genomes. *Bioinformatics*, *29*(21), 2790–2791. https://doi.org/10.1093/bioinformatics/btt468

Sexton, J. P., Hangartner, S. B., & Hoffmann, A. A. (2014). Genetic isolation by environment or distance: Which pattern of gene flow is most common? *Evolution*, *68*(1), 1–5. https://doi.org/10.1111/evo.12258

Stephens, J. D., Rogers, W. L., Mason, C. M., Donovan, L. A., & Malmberg, R. L. (2015). Species tree estimation of diploid *Helianthus* (Asteraceae) using target enrichment. *American Journal of Botany*, *102*(6), 910–920.

Strasburg, J. L., & Rieseberg, L. H. (2008). Molecular demographic history of the annual sunflowers *Helianthus annuus* and *H. petiolaris*—Large effective population sizes and rates of long-term gene flow. *Evolution: International Journal of Organic Evolution*, *62*(8), 1936–1950.

Suarez-Gonzalez, A., Lexer, C., & Cronk, Q. C. (2018). Adaptive introgression: A plant perspective. *Biology Letters*, *14*(3), 20170688. https://doi.org/10.1098/rsbl.2017.0688

Todesco, M., Owens, G. L., Bercovich, N., Légaré, J. S., Soudi, S., Burge, D. O., Huang, K., Ostevik, K. L., Drummond, E. B., Imerovski, I., Lande, K., Pascual-Robles, M. A., Nanavati, M., Jahani, M., Cheung, W., Staton, S. E., Muños, S., Nielsen, R., Donovan, L. A., … Rieseberg, L. H. (2020). Massive haplotypes underlie ecotypic differentiation in sunflowers. *Nature*, *584*(7822), 602–607.

Van der Auwera, G. A., Carneiro, M. O., Hartl, C., Poplin, R., Del Angel, G., Levy-Moonshine, A., Jordan, T., Shakir, K., Roazen, D., Thibault, J., & Banks, E. (2013). From FastQ data to high-confidence variant calls: The genome analysis toolkit best practices pipeline. *Current Protocols in Bioinformatics*, *43*(1), 11.

Weir, B. S., & Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution*, *38*(6), 1358–1370.

Wellenreuther, M., & Bernatchez, L. (2018). Eco-evolutionary genomics of chromosomal inversions. *Trends in Ecology & Evolution*, *33*(6), 427–440.

Whitney, K. D., Broman, K. W., Kane, N. C., Hovick, S. M., Randell, R. A., & Rieseberg, L. H. (2015). QTL mapping identifies candidate alleles involved in adaptive introgression and range expansion in a wild sunflower. *Molecular Ecology*, *24*, 2194–2211.

Whitney, K. D., Randell, R. A., & Rieseberg, L. H. (2006). Adaptive introgression of herbivore resistance traits in the weedy sunflower *Helianthus annuus*. *The American Naturalist*, *167*(6), 794–807.

Whitney, K. D., Randell, R. A., & Rieseberg, L. H. (2010). Adaptive introgression of abiotic tolerance traits in the sunflower *Helianthus annuus*. *New Phytologist*, *187*(1), 230–239.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T., Miller, E., Bache, S., Müller, K., Ooms, J., Robinson, D., Seidel, D., Spinu, V., … Yutani, H. (2019). Welcome to the Tidyverse. *Journal*

of Open Source Software, *4*(43), 1686. https://doi.org/10.21105/joss.01686.
Wright, S. (1943). Isolation by distance. *Genetics*, *28*(2), 114–138. https://doi.org/10.1093/genetics/28.2.114
Yatabe, Y., Kane, N. C., Scotti-Saintagne, C., & Rieseberg, L. H. (2007). Rampant gene exchange across a strong reproductive barrier between the annual sunflowers, *Helianthus annuus* and *H. petiolaris*. *Genetics*, *175*(4), 1883–1893.
Yuan, K., Zhou, Y., Ni, X., Wang, Y., Liu, C., & Xu, S. (2017). Models, methods and tools for ancestry inference and admixture analysis. *Quantitative Biology*, *5*(3), 236–250. https://doi.org/10.1007/s40484-017-0117-2
Zhang, J. Q., Imerovski, I., Borkowski, K., Huang, K., Burge, D., & Rieseberg, L. H. (2019). Intraspecific genetic divergence within *Helianthus niveus* and the status of two new morphotypes from Mexico. *American Journal of Botany*, *106*(9), 1229–1239.
Zheng, X., Levine, D., Shen, J., Gogarten, S. M., Laurie, C., & Weir, B. S. (2012). A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics*, *28*(24), 3326–3328. https://doi.org/10.1093/bioinformatics/bts606
Zhou, X., & Stephens, M. (2012). Genome-wide efficient mixed-model analysis for association studies. *Nature Genetics*, *44*(7), 821–824. https://doi.org/10.1038/ng.2310

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Owens, G. L., Todesco, M., Bercovich, N., Légaré, J.-S., Mitchell, N., Whitney, K. D., & Rieseberg, L. H. (2021). Standing variation rather than recent adaptive introgression probably underlies differentiation of the *texanus* subspecies of *Helianthus annuus*. *Molecular Ecology*, 30, 6229–6245. https://doi.org/10.1111/mec.16008