Genetic diversity, reproductive success, and genetic differentiation from congeners in the narrow endemic Phlox pilosa ssp. sangamonensis1

Authors: Zinnen, Jack, Matthews, Jeffrey W., and Zaya, David N.

Source: The Journal of the Torrey Botanical Society, 151(2) : 138-163

Published By: Torrey Botanical Society

URL: https://doi.org/10.3159/TORREY-D-23-00025.1

Journal of the Torrey Botanical Society 151(2): 138–163, 2024.

Genetic diversity, reproductive success, and genetic differentiation from congeners in the narrow endemic *Phlox pilosa* ssp. *sangamonensis*[1]

Jack Zinnen,[2,5] Jeffrey W. Matthews,[2] and David N. Zaya[3,4]

[2] Department of Natural Resources and Environmental Sciences, University of Illinois Urbana-Champaign, 1102 S. Goodwin Avenue, Urbana, IL 61801

[3] Illinois Natural History Survey, Prairie Research Institute, University of Illinois Urbana-Champaign, 1816 S. Oak Street, Champaign, IL 61820

[4] Department of Biological Sciences, University of Illinois Chicago, 900 W. Taylor Street, Chicago, IL 60607

Abstract. Risk factors that make rare plant taxa particularly susceptible to population declines include a self-incompatible breeding system combined with small population size, small range, and isolated populations separated by unsuitable habitat. *Phlox pilosa* ssp. *sangamonensis* is an endangered, self-incompatible, and narrowly endemic taxon with isolated population fragments in east-central Illinois. Here we combined a field and genetic study of *Phlox pilosa* ssp. *sangamonensis* as a study of its taxonomic status and conservation, especially because some of its remaining populations are small (< 30 flowering individuals). First, we used six polymorphic nuclear microsatellite loci developed for *P. pilosa* ssp. *pilosa* to characterize the genetics of 212 individuals in all 10 known populations of *P. pilosa* ssp. *sangamonensis*. We tested the taxon's genetic differentiation from congeners in east-central Illinois: two populations of *Phlox pilosa* ssp. *fulgida*, two populations of *Phlox pilosa* ssp. *pilosa*, and three populations of *Phlox divaricata* ssp. *laphamii*. We also quantified reproduction in each population. We surveyed fruit set for three years and tested correlations with flowering population size and distance to other flowering individuals. For one year, we collected data on seed set from successfully formed capsules. *Phlox pilosa* ssp. *sangamonensis* was genetically distinct from, and had putatively lower genetic variation, than *Phlox divaricata* ssp. *laphamii*, *Phlox pilosa* ssp. *fulgida*, and *Phlox pilosa* ssp. *pilosa*. Overall genetic diversity in *P. pilosa* ssp. *sangamonensis* appeared low, especially in small populations. Fruit set was positively associated with *P. pilosa* ssp. *sangamonensis* population size, with larger population (> 100 flowering individuals) fruit set being more than twice as high compared to smaller (< 30 flowering individuals) ones. Across populations, individuals with a greater distance to flowering neighbors also showed reduced fruit set, although the relationship was not as strong as for population size. In the year we studied seed set, we also found a positive association between population size and seed set, with approximately 30% more seeds per capsule being produced in the larger populations relative to small ones. Our data suggest that *P. pilosa* ssp. *sangamonensis* is a genetically distinct taxon from nearby *Phlox* taxa, and its smaller populations could be at risk of further decline, possibly due to mate or pollen limitation, and/or low genetic diversity. Management for this endangered taxon should facilitate recruitment in small (< 100 flowering individuals) populations and maintain habitat quality for the large populations.

Key words: conservation genetics, genetic structure, habitat fragmentation, microsatellites, *Phlox pilosa* complex, rare plants

[1] This work would not have been possible without funding from the Elmhurst Garden Club and the Garden Club of Downers Grove. Two anonymous reviewers provided constructive feedback on the previous version of the manuscript. We thank M. Davis for access to the Collaborative Ecological Genetics Laboratory (Illinois Natural History Survey) and his management of the lab. We are grateful to the property owners for giving us permission to conduct the project. We specifically thank the Alexander, Halfar, Kasper, McCullough, Seten, Willard, and Wolf families. We thank L. M. Smith, C. Carroll-Cunningham, M. K. Solecki, and P. Marcum for assistance in locating the populations. The Champaign County Forest Preserve District and Grand Prairie Friends kindly provided access to populations of *P. pilosa* ssp. *sangamonensis* and *P. pilosa* conspecifics, respectively. Special thanks to S. Mason for assisting with fieldwork. We are appreciative to S. Fehlberg and C. Ferguson for providing sequences of additional unpublished microsatellite primers, and P. Zale and K. Robertson for providing tissue samples. Species records were obtained from the Illinois Department of Natural Resources' Natural Heritage Database. Data for this project can be obtained from the Illinois Data Bank: https://doi.org/10.13012/B2IDB-5376622_V1.

[5] Author for correspondence: Jack Zinnen (jzinnen2@illinois.edu)

doi: 10.3159/TORREY-D-23-00025.1

Received for publication August 21, 2023, and in revised form April 1, 2024; first published June 3, 2024.

Small population size can threaten the persistence of rare plant populations. Small populations are more sensitive to inbreeding depression, genetic erosion and drift, and mate limitation (Ellstrand and Elam 1993; Willi et al. 2022). The risks associated with small population sizes are particularly intense for rare plant taxa with few and fragmented extant populations, as well as outcrossing taxa susceptible to mate limitation (Leimu et al. 2006; Aguilar et al. 2019). Further complicating the conservation of such taxa, rare plants may have naturally lower genetic variation relative to widespread species (Gitzendanner and Soltis 2000), which could reduce fitness and environmental adaptability. Thus, habitat loss, fragmentation, and degradation present genetic and reproductive challenges for small plant populations.

In this study, we report on genetic and reproductive analyses of the rare *Phlox pilosa* L. ssp. *sangamonensis* Levin & D.M. Sm., representing the first work on the taxon intended to provide detailed information about its conservation and genetics. *Phlox pilosa* ssp. *sangamonensis* is an Illinois state-endangered taxon (Illinois Endangered Species Protection Board 2015) and is one of the only plant taxa that is endemic to Illinois (Robertson 2001). *Phlox pilosa* ssp. *sangamonensis* is a self-incompatible taxon native to open or semi-open plant communities, occurring in a small, band-like range in Piatt and Sangamon Counties (Levin and Smith 1965). Its putative conservation threats include habitat loss, fragmentation, and degradation (e.g., canopy closure) due to its natural range being dominated by agriculture and development. *Phlox pilosa* taxa have been previously shown to have limited capabilities for long-distance gene flow or dispersal (Levin and Kerster 1968), making continued habitat destruction and fragmentation in its small range threatening to population-level persistence. There is a risk the taxon is suffering from low reproduction due to mate limitation or low genetic diversity. However, detailed information about the taxon's historical attributes (e.g., geographic extent, population sizes and conservation risks, or reproductive output) is not available, presenting a limitation for the taxon's future conservation.

Studying the genetics and reproduction of *P. pilosa* ssp. *sangamonensis* can contribute to the understanding of its taxonomy and improve its management by revealing vulnerable populations. First, studying *P. pilosa* ssp. *sangamonensis* genetics has an initial benefit of providing insight into its relationship to other subspecies in the taxonomically complex *Phlox pilosa* group (Zale 2014; e.g., Fehlberg et al. 2014). *Phlox pilosa* ssp. *sangamonensis* is recognized as a distinct subspecies (Levin and Smith 1965; USDA NRCS 2022). Previous studies show that *P. pilosa* ssp. *sangamonensis* has weak to strong reproductive barriers to other *Phlox* taxa in this group (Levin and Smith 1965; Levin 1966) and may not be closely related to *Phlox pilosa* taxa found in the same region (Garner et al. 2022). Thus, we suspect that *P. pilosa* ssp. *sangamonensis* is indeed genetically distinct from other Illinoisan *Phlox pilosa* complex congeners, but additional data could reinforce the finding from previous studies. Second, a combination of genetic and reproductive information of the remaining *P. pilosa* ssp. *sangamonensis* populations can inform their management. Smaller population sizes and greater individual isolation can create reproductive challenges to outcrossing plant taxa (Wagenius 2006). For example, one study showed that *Phlox pilosa* populations in prairie remnants were found to have lower fruit set in smaller populations (Hendrix and Kyhl 2000). If this pattern applies to the state-endangered *P. pilosa* ssp. *sangamonensis*, targeted management may be needed to maintain or increase remaining populations. Surveying genetic diversity using microsatellites (Fehlberg et al. 2008) can suggest the best populations for sourcing suitable propagules or show those that are genetically depauperate needing reinforcement (DeMauro 1993; Tecic et al. 1998). Third, our study provides a comprehensive survey and multifaceted study of the remaining populations. A study that combines a rigorous study of the populations' sizes, genetic status, and reproductive success will provide novel insights into the biology and conservation of the taxon.

Our specific objectives were to 1) compare the genetics of *P. pilosa* ssp. *sangamonensis* to closely related congeners *P. divaricata* ssp. *laphamii* (Alph. Wood) Wherry, *P. pilosa* ssp. *fulgida* (Wherry) Wherry, and *P. pilosa* L. ssp. *pilosa*; 2) assess the genetic diversity and structure of remaining *P. pilosa* ssp. *sangamonensis* populations; and 3) survey population size and reproductive success in all known *P. pilosa* ssp. *sangamonensis* populations. We expected our findings to suggest that *P. pilosa* ssp. *sangamonensis* is genetically distinct from other Illinois *Phlox* taxa, including the two conspecifics. We hypothesized that *P. pilosa* ssp. *sangamonensis* genetic variation would be low, and that flowering

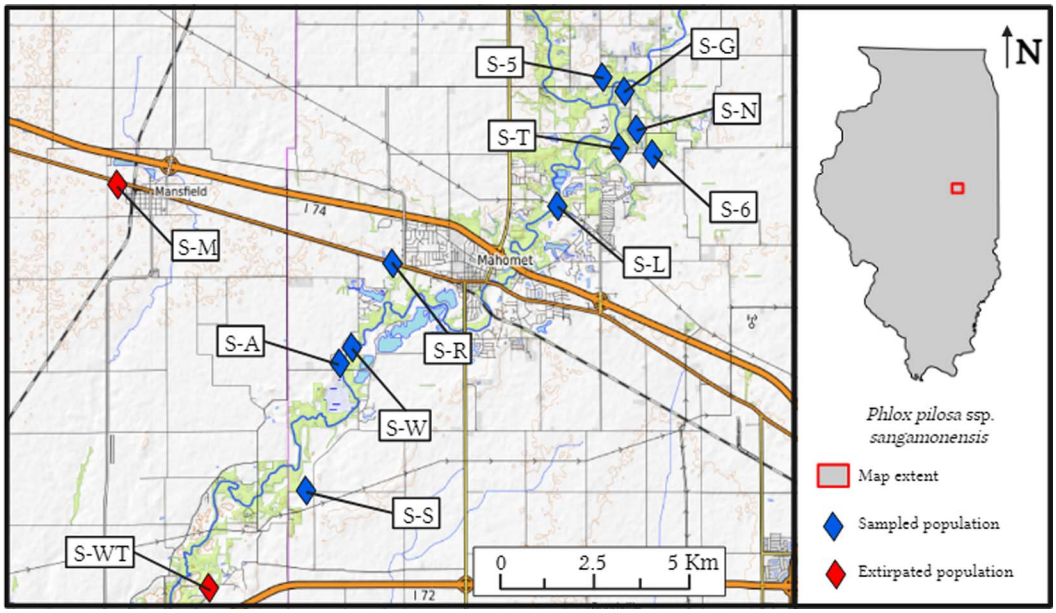

FIG. 1. Surveyed populations of *P. pilosa* ssp. *sangamonensis* with population names shown in Table 1. Shown in the right panel is the state of Illinois and the extent of the study area. Ten extant populations (blue diamonds) of *P. pilosa* ssp. *sangamonensis* were characterized using microsatellites and surveyed for their population sizes and reproductive statuses; two populations (red diamonds) were not found during visits in 2018-2019 and were presumed extirpated. Extant populations are distributed close to the Sangamon River and its surrounding woodlands (shown as green). OpenTopoMap was used for the base map (http://www.opentopomap.org; CC-BY-NC-SA 4.0).

population size would be positively associated with genetic diversity and reproductive success.

Methods. DESCRIPTION AND ECOLOGY OF *PHLOX PILOSA* SSP. *SANGAMONENSIS*. *Phlox pilosa* ssp. *sangamonensis* (Polemoniaceae) is an herbaceous, perennial forb. It has similar morphological characteristics to the more widespread *P. pilosa* ssp. *pilosa*, but is distinguished by its glabrous leaves, stems, and a less hirsute calyx (Levin and Smith 1965). *Phlox pilosa* ssp. *sangamonensis* is ~0.2–0.5 m tall, with dull to hot pink flowers that bloom from mid-May through early July. *Phlox pilosa* taxa are self-incompatible and are primarily pollinated by large butterflies (Robertson 1928; Levin 1966). *Phlox. pilosa* ssp. *sangamonensis* and other *Phlox pilosa* taxa have seed that is dispersed by ballistic dispersal (Levin and Kerster 1968; Zale 2014), meaning dehisced capsules physically erupt to move the seed away from the parent. Individuals can be long-lived, possibly up to 30 years (Levin 1984). *Phlox pilosa* taxa are known to spread vegetatively via rhizomes.

Phlox pilosa ssp. *sangamonensis* is narrowly endemic to a small area in east-central Illinois, USA, in Champaign and Piatt Counties. The populations are only found within the upper Sangamon River watershed area, forming a linear, band-like range (Levin and Smith 1965). The causes of the strict endemism of *P. pilosa* ssp. *sangamonensis* are unclear because it has a relatively wide habitat breadth and no apparent edaphic specialization; it is typically found in partly shaded edge habitats (viz., savannas and wooded roadsides), and was occasionally present in prairies, old fields, and forest bluffs (Levin and Smith 1965; Iverson et al. 1997).

As a narrow endemic, the conservation status of *P. pilosa* ssp. *sangamonensis* is unclear and warrants further investigation. In the 1970s, *P. pilosa* ssp. *sangamonensis* had at least 4,000 individuals across all populations (Levin 1984), and the taxon was described as having "large populations" by Levin and Smith (1965). Habitat and population decline over the past decades has led to its state-endangered status (Illinois Endangered Species Protection Board 2015). Its range has putatively decreased since its description (cf. Levin and Smith 1965 to Fig. 1). However, besides sporadic, sparse surveys recorded by the Illinois Department

of Natural Resources for the past thirty years, we lacked detailed data on its population locations, sizes, or causes for putative declines in range and population sizes. *Phlox pilosa* ssp. *sangamonensis* has poor tolerance to modern anthropogenic disturbances (Taft et al. 1997), and requires historical ecological regimes (e.g., occasional fires) to maintain its habitat. All but one of the remaining populations are located on unprotected land, mostly on privately owned properties. Risks to its populations include habitat destruction through herbicide, woody encroachment, mowing and brush hogging, as well as locally intense herbivory from *Odocoileus virginianus* (pers. observation).

Taxonomic Status and Origin of *Phlox pilosa* ssp. *sangamonensis*. Our first objective was to compare the genetic composition of *P. pilosa* ssp. *sangamonensis* to other *Phlox* taxa in Illinois. We studied this due to the uncertainty and complexity of its systematic relationships. The *Phlox pilosa* complex is a taxonomically vexing group of *Phlox* species in eastern North America with semipermeable reproductive barriers. The clade is taxonomically difficult due to variable morphology, interspecific gene flow, and intermediate taxa of possibly stable hybrid origin (Levin 1966). Below, we provide an abridged description of previous work on its taxonomy and possible origins.

Phlox pilosa ssp. *sangamonensis* was first formally studied and described by Levin and Smith (1965). These authors tested the hypothesis that it was of hybrid origin between some combination of *P. pilosa* ssp. *pilosa*, *P. divaricata*, or *P. glaberrima* L. ssp. *interior* (Wherry) Wherry. However, the authors found that *P. pilosa* ssp. *sangamonensis* was not of hybrid origin. Furthermore, they showed that *P. pilosa* ssp. *sangamonensis* was reproductively incompatible with *P. pilosa* ssp. *pilosa*.

Morphological, chromosomal, and biochemical characteristics suggested that *P. pilosa* ssp. *sangamonensis* is a disjunct relict of long-distance dispersal event of *P. pilosa* ssp. *detonsa* (A. Gray) Wherry, a subspecies found primarily in the southeastern USA (Wherry 1955; Levin and Smith 1965; Levin 1984). Reproductive barriers between *P. pilosa* ssp. *detonsa* and *P. pilosa* ssp. *sangamonensis* were also weaker compared to crosses of *P. pilosa* ssp. *sangamonensis* and *P. pilosa* ssp. *pilosa* (Levin and Smith 1965). Furthermore, the two taxa had similar chromosomal abnormalities relative to other *P. pilosa* subspecies (Smith and Levin 1967). Levy (1983) found that *P. pilosa* ssp. *sangamonensis* flavone profiles were identical to those of *P. pilosa* ssp. *detonsa* individuals found near the Kentucky-Tennessee border; therefore, Levy (1983) suggested that the origin of *P. pilosa* ssp. *sangamonensis* was a restricted disjunction of *P. pilosa* ssp. *detonsa*. Likewise, Levin (1984) showed that allozymes of *P. pilosa* ssp. *sangamonensis* were a smaller and more homogenous subset of allozymes found in *P. pilosa* ssp. *detonsa*, suggesting a past genetic bottleneck. Some evidence has suggested that *P. pilosa* ssp. *detonsa* was itself a stable hybrid derivative of *P. pilosa* and *P. carolina* L. Seed protein profiles of *P. pilosa* ssp. *sangamonensis* and *P. pilosa* ssp. *detonsa* were more like *P. carolina* than *P. pilosa* ssp. *pilosa* (Levin and Schaal 1970).

Recent phylogenetic techniques have provided more evidence supporting the *P. pilosa* ssp. *detonsa* hypothesis and the taxonomic distinctiveness of *P. pilosa* ssp. *sangamonensis*. Recently, Garner et al. (2022) used RADseq analyses of nuclear DNA to clarify the phylogeny of the *Phlox pilosa* complex. They found strong evidence that some *P. pilosa* taxa are polyphyletic. These authors found that *P. pilosa* ssp. *sangamonensis* is part of a distinct clade that includes *P. pilosa* ssp. *detonsa*, *P. pilosa* ssp. *ozarkana* (Wherry) Wherry, and southern collections of *P. pilosa* ssp. *pilosa*. This study suggested that *P. pilosa* ssp. *sangamonensis* is not closely related to other northern *Phlox*, taxa, including conspecifics. Despite these findings, more evidence could help clarify or reinforce the taxonomic status of *P. pilosa* ssp. *sangamonensis* relative to other members of the *Phlox pilosa* complex.

Identifying Populations of *Phlox pilosa* ssp. *sangamonensis*. We first sought to relocate all known populations of *P. pilosa* ssp. *sangamonensis*. In May 2018, we surveyed seven elements of occurrence records (EORs). EORs are descriptive records of populations of endangered and threatened species which are managed by the Illinois Department of Natural Resources. From these seven EORs, we located a total of 10 distinct populations of *P. pilosa* ssp. *sangamonensis* (Table 1). These 10 populations were identified as distinct populations because they were separated from other populations by at least 400 m of dissimilar habitat or anthropogenic land uses, and likewise had separate site conditions and conservation

Table 1. Summary statistics of 17 populations of the four *Phlox* taxa with genetic data from six microsatellites.*

Taxon	Population name	N	$\bar{x}$	A	A_P	A_R	H_o	H_e	F_{IS}
P. pilosa ssp. *sangamonensis*	S-5	22	29	23	2	2.59	0.447	0.489	0.087
	S-6	30	333	29	1	3.11	0.583	0.604	0.023
	S-A	22	122	20	0	2.48	0.424	0.476	0.138
	S-G	7	5	16	0	2.37	0.500	0.423	−0.090
	S-L	28	432	26	0	2.71	0.440	0.487	0.114
	S-N	23	253	23	0	2.78	0.565	0.557	0.007
	S-R	19	17	19	2	2.40	0.533	0.463	−0.145
	S-S	29	423	27	0	2.77	0.511	0.511	0.005
	S-T	28	62	22	0	2.63	0.512	0.527	0.022
	S-W	4	3	16	0	2.41	0.417	0.438	0.220
	Mean:			22.1	0.5	2.62	0.493	0.497	0.038
	sd:			4.4	0.8	0.23	0.058	0.055	0.107
P. divaricata ssp. *laphamii*	D-L	17	NA	47	5	3.96	0.666	0.687	0.058
	D-N	13	NA	39	5	3.82	0.620	0.669	0.083
	D-S	7	NA	29	3	3.41	0.607	0.581	0.035
	Mean:		NA	38.3	4.3	3.73	0.631	0.646	0.0599
	sd:		NA	9.0	1.2	0.29	0.031	0.057	0.024
P. pilosa ssp. *fulgida*	PF-B	11	NA	27	3	3.14	0.606	0.558	−0.052
	PF-W	8	NA	32	2	3.75	0.583	0.660	0.162
	Mean:		NA	29.5	2.5	3.45	0.595	0.609	0.055
	sd:		NA	3.5	0.7	0.43	0.016	0.072	0.151
P. pilosa ssp. *pilosa*	PP-L	11	NA	44	11	4.12	0.668	0.656	0.022
	PP-P	12	NA	46	6	4.18	0.717	0.718	0.047
	Mean:		NA	45.0	8.5	4.15	0.693	0.687	0.034
	sd:		NA	1.4	3.5	0.04	0.035	0.044	0.018

* Sample sizes are shown after for removing 18 clones or probable clones for *P. pilosa* ssp. *sangamonensis*. N = sample size, A = total number of alleles observed, A_P = number of observed private alleles across all populations, A_R = mean rarified allelic richness, H_o = mean observed heterozygosity, H_e = mean expected heterozygosity across loci, F_{IS} = mean inbreeding coefficient per population. $\bar{x}$ indicates the mean number of flowering individuals in each population across three years (2018–2020) for *P. pilosa* ssp. *sangamonensis*; the number of flowering individuals for the other three taxa were not counted.

risks. Average flowering population sizes ranged from 3–432 for these 10 populations. Because two populations recorded in EORs were not found in 2018 or 2019, and had not been seen in nearly a decade, they were presumed extirpated (Fig. 1).

MICROSATELLITE CHARACTERIZATION OF *PHLOX PILOSA* SSP. *SANGAMONENSIS* AND OTHER *PHLOX* TAXA. During 2019, we haphazardly collected leaf tissue from a total of 212 unique individuals of *P. pilosa* ssp. *sangamonensis* from the 10 extant populations for genetic analysis (Table 1). Eighteen *P. pilosa* ssp. *sangamonensis* individuals that were initially sampled were clones or probable clones (i.e., repeat samples); those samples were removed from all analyses (see *Statistical analysis*).

To compare the genetic similarity and diversity of *P. pilosa* ssp. *sangamonensis* to other closely related taxa, we also sampled leaf tissues from nearby populations of three other taxa in the *Phlox pilosa* complex. The three taxa were *P. pilosa* ssp. *fulgida*, *P. pilosa* ssp. *pilosa*, and *P. divaricata* ssp. *laphamii*—all of which are native to east-central Illinois. *Phlox pilosa* ssp. *pilosa* is the most widespread subspecies of *P. pilosa* (Zale 2014); populations of *P. pilosa* ssp. *fulgida* and *P. pilosa* ssp. *pilosa* occupy nearby prairie remnants throughout east-central Illinois and can intergrade (Levin and Levy 1971). We sampled two populations of each *P. pilosa* subspecies, all of which were found in prairie remnants that occurred 35–85 km from the *P. pilosa* ssp. *sangamonensis* range. We sampled 19 *P. pilosa* ssp. *fulgida* and 23 *P. pilosa* ssp. *pilosa* individuals from two populations each. We sampled these two taxa because they are conspecific to *P. pilosa* ssp. *sangamonensis* with similar morphologies, and likewise the taxa may be their close relatives (Ferguson and Jansen 2002). The last taxon in the *Phlox pilosa* complex we sampled, *Phlox divaricata* ssp. *laphamii*, is sympatric with multiple *P. pilosa* ssp. *sangamonensis* populations (e.g., flowering individuals

separated by < 2 m distance; Appendix Fig. S1). We sampled 37 individuals from three *P. divaricata* ssp. *laphamii* populations. Since *P. divaricata* ssp. *laphamii* and *P. pilosa* ssp. *sangamonensis* have been shown to hybridize in the wild (Zinnen et al. 2024), we were interested in characterizing possible hybridization or introgression between the taxa because it may represent a genetic risk to the *P. pilosa* ssp. *sangamonensis*. Across all four *Phlox* taxa, leaf tissue sample sizes varied per population and are presented in Table 1.

Approximately 20 mg of leaf tissue per individual was used for genomic DNA extraction using the DNeasy Plant Mini Kit (Qiagen). DNA concentrations and purities were assessed using a spectrophotometer (Implen NanoPhotometer, Munich, Germany®). We used primers for six nuclear microsatellite loci which had been previously developed for North American *Phlox* species (Appendix Table S1; Fehlberg et al. 2008; Fehlberg and Ferguson 2012); an additional six primer pairs were screened but not used in the study (Appendix Table S2). Our PCR protocol followed Fehlberg et al. (2008) with some modifications: each 15-uL reaction contained ~25 ng of genomic DNA, 1 unit of EU-Taq DNA Polymerase (eEnzyme), 1× Taq Buffer (eEnzyme), 2.5-mM $MgCl_2$, 0.2 mM dNTPs (eEnzyme), 1.5-μg bovine serum albumin, 0.025-μM forward M13(–18) tagged primer, 0.25-μM reverse primer, and 0.25-μM fluorescent dye. Thermocycling was conducted at 94°C for 2 min, followed by 30 cycles of 94°C for 45 s, 51°C for 1 min, 72°C for 2 min, and a final extension at 72°C for 5 min. We initially visualized and screened PCR products using 2% agarose gel in TBE buffer. Fragment analyses were conducted using an Applied Biosystems (ABI) GeneAnalyzer 3730xl at the Roy J. Carver Biotechnology Center, University of Illinois. Each sample analysis included 600 LIZ dye size standard (GeneScan). Allele sizes for all loci were manually scored using Geneious Prime® v. 2021.1.1 (Biomatters, Auckland, New Zealand).

Field Surveys: Reproductive Monitoring of *Phlox pilosa* ssp. *sangamonensis*. In 2018–2020, we monitored reproductive success of the 10 remaining *Phlox pilosa* ssp. *sangamonensis* populations. We haphazardly selected 10–20 individuals (if available) during the peak blooming in late-May. We flagged individuals and counted the total number of flower buds. To quantify physical isolation of individuals, we averaged the distance to its two nearest flowering neighbors (a better predictor of reproductive success than the distance to the nearest or second nearest neighbor, data not shown). In each season, we revisited the populations during the first week of July to assess fruit set (i.e., fruit-flower ratio). We counted flowers which had turned into capsules; capsules were either directly observed or were inferred due to retracted sepals "star-pattern" that indicated ballistic eruption (Appendix Fig. S2).

In 2019, we also estimated seed set for seven of the 10 extant populations (S-5, S-6, S-A, S-L, S-N, S-S, S-T), and used this as an additional characterization of reproductive success. We collected mature but undehisced capsules during July 2019 from haphazardly selected individuals (range: 6–19 individuals); capsules were isolated and labeled per unique individual, and if available, multiple capsules were collected per individual. Sampled individuals were not necessarily those used for the concurrent capsule formation study described above. Because *Phlox* taxa have a maximum of three seeds formed per capsule (Hendrix and Kyhl 2000; Zale 2014), we counted the number of resulting seeds, and then calculated the seed-ovule ratio (i.e., capsules × 3 seeds/capsule).

Data Analyses. We conducted all the analyses described below, apart from the Bayesian clustering (Structure) analyses, using R v. 4.0.2 (R Core Team 2020).

Prior to our genetic analyses described below, we first checked and removed repeat genotypes. This was because *Phlox* taxa can form dense neighborhoods, complicating the sampling of distinct genets in the field. We used the P_{sib} statistic calculated by the *allelematch* package (Galpern et al. 2012) to identify and remove possible clones. P_{sib} is the probability of two individuals having identical multilocus genotypes due to being siblings rather than clones (Woods et al. 1999). We used an *alleleMismatch* parameter of $\hat{m} = 1$ to account for possible errors when scoring allele sizes, which allowed two genotypes to differ at one allele and still be considered identical. Across all *P. pilosa* ssp. *sangamonensis* populations, we removed five complete matches; we also removed an additional 13 samples of *P. pilosa* ssp. *sangamonensis* samples that were possibly clones, using the conservative estimate of $P_{sib} < 0.1$. We found no evidence of clones for *P. divaricata* ssp. *laphamii*, *P. pilosa* ssp. *fulgida*, or *P. pilosa* ssp. *pilosa*.

We used the package *adegenet* (Jombart 2008) to summarize the microsatellite results. We checked for the presence of null alleles in each taxon using the package *PopGenReport* (Adamack and Gruber 2014). This package uses the method developed by Brookfield (1996) to estimate the frequency of null alleles per locus. We inferred significant evidence of null allele presence if the 95% CI of estimated null allele frequency did not include 0. We recorded private alleles, the number of alleles that are unique to a population, using the package *poppr* (Kamvar et al. 2014). We used different subsets of the microsatellite dataset to find private alleles. First, we determined private allelic richness among all 17 populations for the four taxa. Second, we determined the number of unique allele sizes within each taxon by pooling all individuals of *P. pilosa* ssp. *sangamonensis*, *P. pilosa* ssp. *fulgida*, *P. pilosa* ssp. *pilosa*, and *P. divaricata* ssp. *laphamii*. Once these basic genetic characteristics were recorded, we tested for deviations from Hardy–Weinberg equilibrium (HWE) in each population-locus combination using the package *pegas* (Paradis 2010).

To assess the possible genetic differentiation among the four taxa, we used STRUCTURE v. 2.3.4 (Pritchard et al. 2000). We did not include taxon or population identity in our analyses with STRUCTURE; the program uses Bayesian clustering and Markov chain and Monte Carlo simulations to estimate an individual's identity using posterior probabilities. The results of STRUCTURE give an estimated proportion of an individual's genome that belongs to a particular cluster. The user defines the number of clusters, *K*, to be tested. We tested *K* from 1 to 17 for an analysis of all four taxa to gauge their genetic differences. We then tested *K* from 1 to 10 for an analysis that only included *P. pilosa* ssp. *sangamonensis* samples because we wanted to assess their genetic structure among populations. For both analyses, we chose the best-supported number of clusters following Evanno et al. (2005) and using STRUCTURE HARVESTER (Earl and vonHoldt 2012). We used the admixture model assuming correlated allele frequencies. We averaged results from 10 runs, each with 200,000 iterations after an initial burn-in of 50,000 iterations.

To compare genetic diversity among the taxa, we compared rarefied allelic richness (A_R) and observed heterozygosity (H_o) using Kruskal–Wallis rank sum tests, followed by post hoc comparisons using Bonferroni correction. Allelic richness is an informative, basic proxy of genetic variation (Kalinowski 2004) that was used to compare genetic diversity within the populations of *P. pilosa* ssp. *sangamonensis* and to compare it to the other taxa. Allelic richness was rarefied due to differences in sample sizes and was conducted with *PopGenReport* (Adamack and Gruber 2014). We also calculated the inbreeding coefficient (F_{IS}) for the four taxa using *adegenet*. We partitioned genetic diversity within and between populations of each taxon using an analysis of molecular variance (AMOVA) using the *poppr* package (Kamvar et al. 2014). We combined *P. pilosa* ssp. *fulgida* and *P. pilosa* ssp. *pilosa* into a single taxon for comparing genetic diversity (A_R and H_o) and AMOVA. We combined the subspecies because we did not find evidence of genetic differentiation using STRUCTURE (see Results), and because only two populations were available for each taxon.

To determine the possible relationship between genetic diversity and population size for *P. pilosa* ssp. *sangamonensis*, we used linear regression to compare mean population sizes (mean across the growing seasons of 2018–2020) to A_R and H_o. We log-transformed population sizes to meet statistical assumptions and moderate the effect of the largest populations on the analyses. We used a Mantel test (Mantel 1967) with 1,000,000 permutations to determine if there was significant isolation by distance among *P. pilosa* ssp. *sangamonensis* populations, that is, if geographic distance was associated with genetic distance. We calculated pairwise geographic distances between populations using the *geosphere* package (Hijmans 2021) and calculated pairwise genetic distances as Jost's *D* (Jost 2008) using the *mmod* package (Winter 2012).

We used generalized linear mixed-effects models with a binomial response to determine predictors of reproductive success. The primary measure of reproductive success, which served as the response variable in our models, was fruit set (i.e., fruit-flower ratio) per recorded individual. All models were constructed using the *lme4* package (Bates et al. 2015). Model selection was conducted based on AIC_C using the *MuMIn* package (Barton 2020). The explanatory variables tested were estimated reproductive population size (number of flowering individuals in each population in each year), mean distance to the two nearest neighbors, and rarefied allelic richness (A_R) at the population-level as a measure of genetic

diversity. We constructed separate models for each explanatory variable. First, we constructed a null model that included year as the only fixed effect (categorical) with population as the random effect (intercept only). Then, we compared this null model to models that incorporated one of three variables (population size, mean neighbor distance, or A_R) as a fixed effect. For the models including population size and mean neighbor distance as fixed effects, we made alternative random effects structures: one included the random population intercept only and the second included a random slope and intercept for population; both models had year as a fixed effect. For the model including A_R, we only constructed a model with a random population intercept because only one estimate of genetic diversity was available for each population across years. For each of the three variables, we compared the null model to the models that included the variable as a fixed effect (two additional models for population size and mean neighbor distance, one additional model for genetic diversity). We used the Akaike information criterion corrected for small sample sizes (AIC_C) for model comparison and selected the best-supported model for further consideration. Once the optimal model was selected for each variable, we used AIC_C to compare models across explanatory variables to identify the strongest explanatory variable. To meet statistical assumptions regarding the distribution of residuals, we used log-transformations of the nearest-neighbor distances and population size.

For our second measure of reproductive success, we constructed two linear models to determine if there were differences in seed set (seed-ovule ratio) in each population; seed set per individual was the response variable in both cases. The first model was a null model that only incorporated the global average, whereas the second model included population. We used AIC_C to identify the strongest model. The proportion was logit-transformed to meet assumptions for the distribution of model residuals. We then used the *emmeans* package (Lenth 2022) as a post hoc test to estimate differences among the populations with respect to seed formation during 2019 and used linear regressions (Pearson's product-moment correlation) to correlate the estimated means for each population to log-transformed flowering population size or A_R.

Results. GENETIC ANALYSES. We scored six polymorphic microsatellites for 212 individuals of *P. pilosa* ssp. *sangamonensis*, 37 of *P. divaricata* ssp. *laphamii*, 19 of *P. pilosa* ssp. *fulgida*, and 23 of *P. pilosa* ssp. *pilosa*. Due to homozygote excesses in some loci, we found evidence of null alleles in four taxon-locus combinations. In *P. pilosa* ssp. *sangamonensis*, we found significant evidence of null alleles in two loci (Phl-84 and Phl-113; Appendix Table S3). For *P. pilosa* ssp. *fulgida*, we found significant evidence of null alleles at two loci, Phl-28 and Phl-137. We found no significant evidence of null alleles in the *P. pilosa* ssp. *pilosa* and *P. divaricata* ssp. *laphamii* samples ($P > 0.05$).

Across all loci, we observed a total of 44 alleles in *P. pilosa* ssp. *sangamonensis*, 68 in *P. divaricata* ssp. *laphamii*, 51 in *P. pilosa* ssp. *fulgida*, and 69 in *P. pilosa* ssp. *pilosa*. *Phlox divaricata* ssp. *laphamii* had the greatest number (n = 23) of alleles not observed in any other taxon, followed by *P. pilosa* ssp. *pilosa* (n = 21). *Phlox pilosa* ssp. *sangamonensis* had fewer alleles (n = 12) that were not observed in any other taxon; *P. pilosa ssp. fulgida* (n = 6) had the fewest. The Phl-113 (A = 37), Phl-84 (A = 33), and Phl-33 (A = 27) loci were the most variable in our dataset (Appendix Table S4). In contrast, the Phl-28 (A = 10) and Phl-115 loci (A = 12) were the least variable. The six loci varied in observed heterozygosity across the 17 populations (Appendix Table S4). Because of the higher diversity and evenness of alleles, the Phl-113 and Phl-84 loci were highly polymorphic; when averaged across the populations, $H_o > 0.70$. The other 4 loci were less heterozygous on average, with $H_o < 0.50$. Mean H_o across loci was 0.553 ± 0.090 (standard deviation); the average H_e across loci was 0.559 ± 0.093. Although there were deviations from HWE, no locus consistently (>12%) violated HWE across the populations. We observed seven locus-population combinations that were significantly ($P < 0.05$) in violation of HWE assumptions. However, assuming the threshold of $\alpha = 0.05$ for violating HWE, the number of violations of HWE we observed across our dataset was not statistically different from the number (5.1) expected under null conditions (one-tailed binomial test $P = 0.249$); this finding was also robust when the taxa were separated (one-tailed binomial test $P > 0.118$).

Our Bayesian clustering analyses robustly separated *P. pilosa* ssp. *sangamonensis* from the other three *Phlox* taxa, despite STRUCTURE not including *a priori* information about the taxon or population identity of the individuals (Fig. 2A). We found the greatest evidence for $K = 2$ when all four taxa

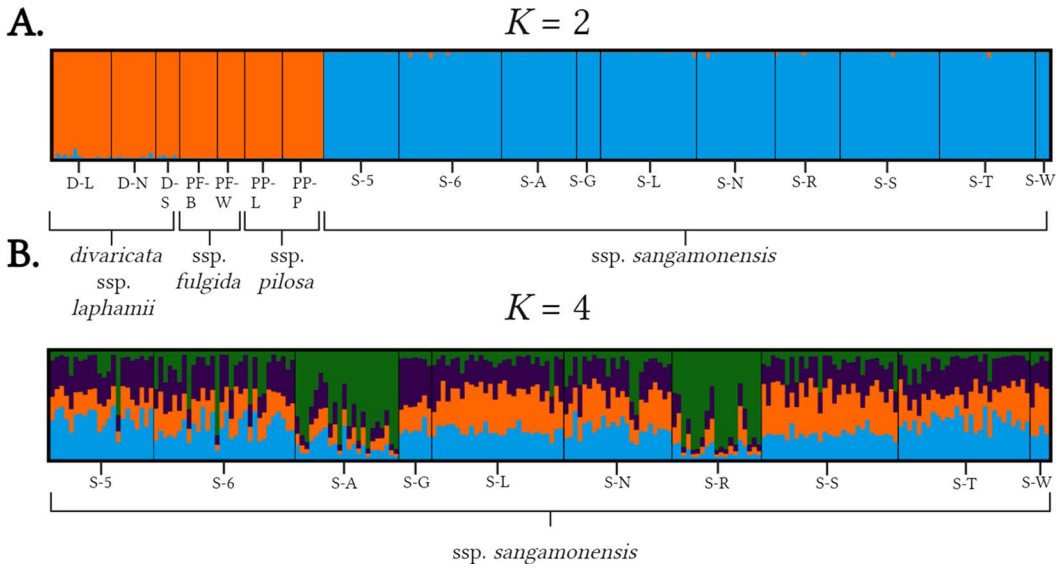

FIG. 2. Bayesian clustering with STRUCTURE for *P. pilosa* ssp. *sangamonensis*, *P. divaricata* ssp. *laphamii*, *P. pilosa* ssp. *fulgida*, and *P. pilosa* ssp. *pilosa*. Analyses were conducted with genotypes from six nuclear microsatellite DNA loci. (A) $K = 2$ was identified as the optimal number of clusters and differentiated *P. pilosa* ssp. *sangamonensis* (blue) from *P. divaricata* ssp. *laphamii, P. pilosa* ssp. *fulgida*, and *P. pilosa* ssp. *pilosa* (orange). (B) When *P. pilosa* ssp. *sangamonensis* was analyzed separately, $K = 4$ clusters was identified as the optimal number of clusters, which particularly differentiated populations S-A and S-R (green) from the other eight populations (blue, orange, and purple). This figure was created using STRUCTURESELECTOR (Li and Liu 2018) and CLUMPAK (Kopelman et al. 2015).

were included in STRUCTURE (Appendix Fig. S3A). With $K = 2$, populations of *P. pilosa* ssp. *sangamonensis* were grouped, whereas *P. divaricata* ssp. *laphamii*, *P. pilosa* ssp. *fulgida*, and *P. pilosa* ssp. *pilosa* were grouped into a single cluster (Fig. 2A). We did not find evidence of hybridization or introgression in *P. pilosa* ssp. *sangamonensis*; at $K = 2$, posterior probabilities for every individual did not meet typical thresholds of being considered of hybrid origin (e.g., Abraham et al. 2011). When individuals of *P. pilosa* ssp. *sangamonensis* were analyzed separately, $K = 4$ had the greatest support (Appendix Fig. S3B). Generally, there was evidence of admixture among *P. pilosa* ssp. *sangamonensis* populations, though two populations (S-A and S-R) showed some signs of differentiation with increasing K (Fig. 2B).

When averaged across the six loci, genetic diversity was relatively high among samples of *P. divaricata* ssp. *laphamii*, noticeably lower in *P. pilosa* ssp. *sangamonensis*, and high for the other two *P. pilosa* conspecifics (Table 1). Population-level A_R of *P. pilosa* ssp. *sangamonensis* was significantly lower than the other taxa (Appendix Fig. S4A; $\chi^2 = 11.7$; df = 2; $P < 0.01$). *Phlox pilosa* ssp. *sangamonensis* had significantly lower H_o than both *P. divaricata* ssp. *laphamii* and the two pooled *P. pilosa* conspecifics (Appendix Fig. S4B; $\chi^2 = 11.4$; df = 2; $P < 0.01$). Despite much greater sampling in *P. pilosa* ssp. *sangamonensis*, it had the fewest number of alleles and the second fewest private alleles at the taxon level. Furthermore, *P. pilosa* ssp. *sangamonensis* often had narrower allele size ranges per locus (Appendix Table S5).

As is typical for outcrossing species, most genetic diversity was distributed within rather than between populations (Appendix Table S6). Conspecifics *P. pilosa* ssp. *fulgida* and *P. pilosa* ssp. *pilosa*, both prairie-specialists, had relatively high (~15%) genetic diversity partitioned between populations. This was likely an indication of their fragmented remnant habitat. The taxon with the least variance between populations was *P. divaricata* ssp. *laphamii* (~4%), which was unsurprising considering its relatively intact woodland habitat and small range of our sampling (~13 km). *Phlox pilosa* ssp. *sangamonensis* had an intermediate level of genetic diversity partitioned between populations (~8%) compared to the other sampled congeners.

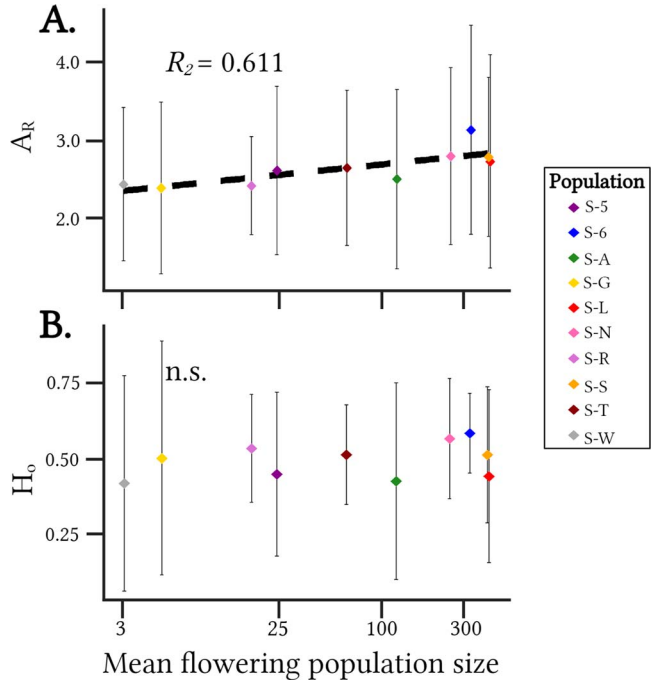

FIG. 3. The association between population sizes of *P. pilosa* ssp. *sangamonensis* and two measures of genetic diversity, A_R (subpanel A) and observed heterozygosity (H_o; subpanel B). There was a strong, positive correlation between log-transformed population size and A_R (A; $P = 0.008$; $R^2 = 0.611$). In both panels, the population sizes are displayed logarithmically to improve data interpretation, and 95% confidence intervals are shown for estimates of A_R and H_o.

Of the 10 *P. pilosa* ssp. *sangamonensis* populations, the H_o ranged from 0.417 in S-W to 0.583 in S-6; A_R ranged from 2.37 in S-G to 3.11 in S-6 (Table 1). Populations S-6 and S-N appeared to be somewhat more genetically diverse across the six loci ($A_R > 2.75$; $H_o > 0.55$; Table 1). Several smaller populations, namely S-W, S-G, and S-A, were the least genetically diverse ($A_R < 2.50$; $H_o < 0.51$; Table 1). There was a strong positive relationship between mean *P. pilosa* ssp. *sangamonensis* flowering population size and A_R ($P = 0.003$; $R^2 = 0.611$; Fig. 3A). However, there was no evidence that H_o and population size were related ($P = 0.318$; $R^2 = 0.124$; Fig. 3B). We also did not find evidence that population size was related to F_{IS} ($P = 0.949$; $R^2 < 0.001$). We found weak, statistically nonsignificant evidence of isolation by distance among the *P. pilosa* ssp. *sangamonensis* populations (Appendix Fig. S5): the Mantel test was marginally significant ($P = 0.059$). The southernmost population, S-S (Fig. 1), was peculiar in this analysis. The S-S was the most geographically distant population on average, yet it had low genetic distances to the northernmost populations (Appendix Fig. S5).

ASSOCIATIONS WITH REPRODUCTIVE SUCCESS. For our field surveys, we found nine populations of *P. pilosa* ssp. *sangamonensis* that were previously recorded in EORs and found an additional, small, previously unknown population (S-G). Between 2018-2020, we counted a mean of nearly 1,700 flowering plants across all populations. Populations S-W ($\bar{x} = 3$) and S-G ($\bar{x} = 5$) had the smallest average number of flowering individuals, whereas S-L had the largest number of mean flowering individuals ($\bar{x} = 432$), followed closely by S-S ($\bar{x} = 423$).

Initially, to understand potential drivers of fruit-flower ratio, we incorporated the three variables as fixed effects in a global model. However, a preliminary analysis revealed that these variables were highly correlated; thus, associations with these variables were analyzed separately, and then assessed to find which was the strongest explanatory variable. Specifically, we found a strong inverse relationship between population size and neighbor distances ($R^2 = 0.667$). In other words, smaller populations generally had low densities of individuals and thus high average neighbor distances,

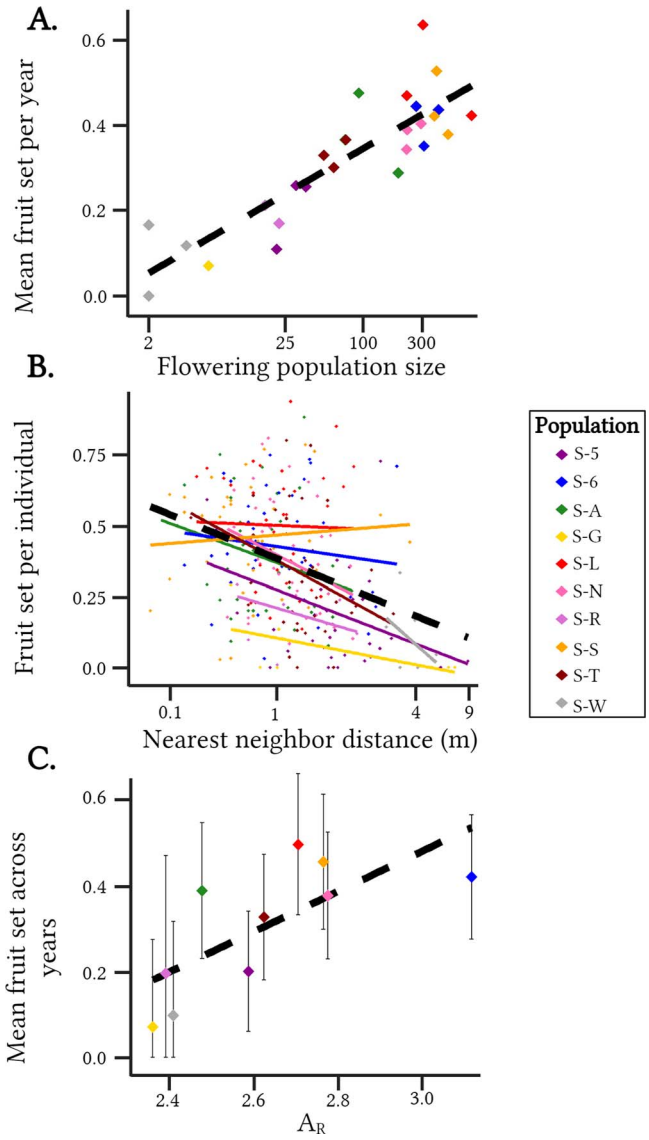

FIG. 4. Fruit set predicted by measures of population size, neighbor distance, and genetic diversity. The proportion of capsules that developed is positively associated with log-transformed flowering population size (A), negatively associated with neighbor distance (B; calculated as mean of the nearest two neighbors), and positively associated with rarefied allelic richness, A_R (C). Neighbor distance (m) and population sizes were analyzed after log-transformation to meet model assumptions and improve interpretation. Note that the larger datapoints in subpanels A and C are population-level metrics (per year for population size); smaller points in subpanel B represent values for individuals with the specific mean neighbor distances shown, with color-coded solid lines showing the random slope for each population. For subpanel C, 95% confidence intervals are shown for estimates of the A_R.

whereas larger populations consisted of lower nearest neighbor distances due having denser clusters of individuals. Furthermore, A_R was strongly correlated to population size ($R^2 = 0.611$).

There was a strong positive correlation between population size and fruit set (Fig. 4A). The two smallest populations (S-G and S-W) had approximately 10% fruit set across years, whereas larger populations (those with mean > 200 flowering individuals per year) had upwards of 40% fruit set. Also, there was a negative association between neighbor distance and reproductive success, though

Table 2. Comparisons of the optimal generalized linear mixed effects models to identify the best explanatory variable for fruit set.*

Fixed effects	Random effects	Intercept	Slope (fixed effect)	df	ΔAIC_C	Weight
Y + N	Intercept: P	–0.49	0.112	7	0.0	> 0.999
Y + D	Intercept: P, Slope: D	0.09	–0.360	7	13.9	0.0
$Y + A_R$	Intercept: P	–7.92	2.587	5	185.5	0.0
Y	Intercept: P	–1.14	NA	4	192.1	0.0

* Model terms: Y = year (categorical variable), N = log-transformed flowering population size per year, P = population random intercept, D = log-transformed mean of distance to two nearest neighbors, A_R = rarefied allelic richness.

the strength of the trend varied across populations (Fig. 4B). There was strong statistical support for the relationships between fruit set and both population size and neighbor distance (Appendix Table S7). During model selection, the two best-supported models for population size and neighbor distance included both random slopes and intercepts (Appendix Table S7; weight > 0.999 for both). We also found a weaker, but still strong, positive relationship between fruit set and A_R (Fig. 4C; $\Delta AIC_C = 6.6$; weight = 0.96). Of the three explanatory variables (flowering population size, neighbor distance, and A_R) used to predict fruit set (i.e., the fruit-flower ratio or capsule formation), population size was the strongest explanatory variable (Table 2; $\Delta AIC_C > 13$; weight > 0.999).

For the 2019 seed formation data set, the mean percent seed set per population ranged from 60.0% to 79.3% across seven populations. We found statistical support for a positive association between population identity and seed set ($\Delta AIC_C = 3.7$; weight = 0.862). After we estimated the marginal means of seed set per population, we found a positive correlation between flowering population size and seed set (Fig. 5A; $R^2 = 0.87$); there was also a positive correlation between A_R and seed set, but the correlation was somewhat weaker (Fig. 5B; $R^2 = 0.61$).

Discussion. General Findings. We found evidence supporting the taxonomic distinctiveness of *P. pilosa* ssp. *sangamonensis* compared to other *Phlox pilosa* complex taxa found nearby in Illinois. Our genetic findings were also consistent with findings of other studies suggesting that *P. pilosa* ssp. *sangamonensis* is dissimilar to nearby *P. pilosa* conspecifics and may have lower genetic variation compared to other *Phlox* taxa. Small population sizes were associated with poorer reproductive output in *P. pilosa* ssp. *sangamonensis*.

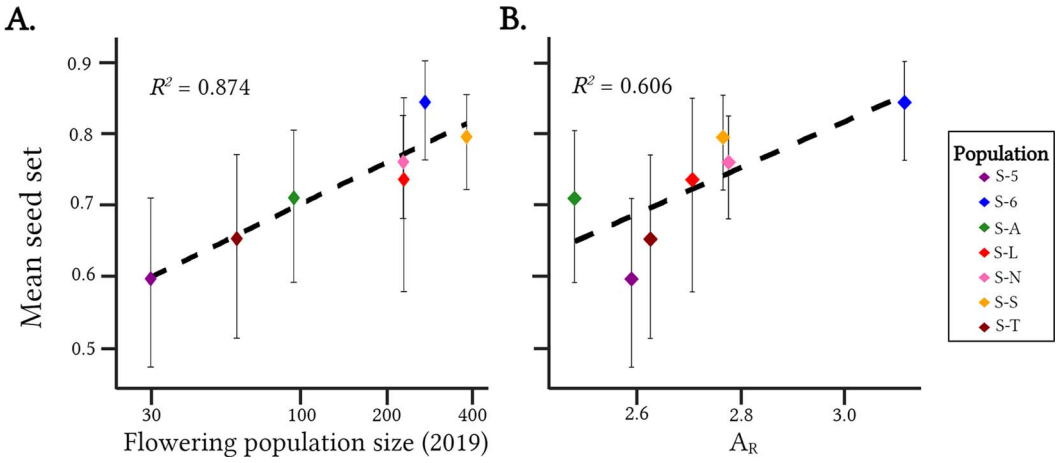

Fig. 5. The association between seed set and log-transformed flowering population sizes (A) and A_R (B) of seven *P. pilosa* ssp. *sangamonensis* populations in 2019. Seed set is defined as the proportion of ovules that develop into seeds. There were strong, positive correlations between the proportion of developed seeds and both log-transformed flowering population size ($R^2 = 0.874$) and A_R ($R^2 = 0.606$). Note that statistical tests were conducted using logit-transformed proportion data, but back-transformed proportions and 95% confidence intervals are displayed on the y-axis of this figure.

This may have been due to pollen limitation in this insect-pollinated outcrossing taxon, and perhaps the lower genetic diversity of smaller populations. Furthermore, smaller population sizes coincided with lower density and greater distances between possible mates, which could have posed an additional barrier to reproduction.

COMPARISON OF *PHLOX PILOSA* SSP. *SANGAMONENSIS* TO CONGENERS. Our results suggest that *P. pilosa* ssp. *sangamonensis* is genetically differentiated from the other three *Phlox* taxa in this study, particularly the more prevalent and widespread *P. pilosa* conspecifics in Illinois. Experimental and observational data of *P. pilosa* ssp. *sangamonensis* has suggested it is not closely related to *P. pilosa* ssp. *fulgida* or *P. pilosa* ssp. *pilosa* populations in Illinois (Levin and Smith 1965; Levy 1983). Most notably though, our results are supported by Garner et al. (2022), who used RADseq across thousands of DNA markers to provide a phylogeny of the *Phlox pilosa* complex. Garner et al. (2022) found evidence that *P. pilosa* ssp. *pilosa* is polyphyletic, with populations in the southern USA forming a distinct clade that is not closely related to the northern USA *P. pilosa* ssp. *pilosa* clade. In Garner et al. (2022), the southern *P. pilosa* ssp. *pilosa* clade included *P. pilosa* ssp. *sangamonensis*, *P. pilosa* ssp. *detonsa*, and *P. pilosa* ssp. *ozarkana*. These authors found that *P. pilosa* ssp. *fulgida* was nested in a clade with northern *P. pilosa* ssp. *pilosa* populations. Thus, our results were consistent with these other studies that showed the taxonomic distinctiveness of *P. pilosa* ssp. *sangamonensis*. While we cannot directly comment on the taxonomy of *P. pilosa* ssp. *sangamonensis*, these studies have identified the southeastern United States as the likely origin of the taxon. Moreover, we showed that adjacent *P. divaricata* ssp. *laphamii* populations are not threatening the genetic integrity of *P. pilosa* ssp. *sangamonensis* (Fig. 2). Although the taxa can hybridize (Zinnen et al. 2024), our results suggest hybridization or introgression is rare.

Our data on genetic diversity also provided indirect support for the hypothesized origin of *P. pilosa* ssp. *sangamonensis*. If *P. pilosa* ssp. *sangamonensis* was indeed the result of a long-distance dispersal event from *P. pilosa* ssp. *detonsa* populations (Levy 1983; Levin 1984), or perhaps some other group associated with southern *P. pilosa* ssp. *pilosa* (see Garner et al. 2022), our results were consistent with the expected genetic consequences of a founder effect. Founder or bottleneck events are known to reduce genetic diversity in plant populations (Young et al. 1996). Putative genetic variation in *P. pilosa* ssp. *sangamonensis* was low despite high intensity sampling across all its known populations. Similar to our results, Levin (1984) studied allozymes and found that, compared to its putative ancestor *P. pilosa* ssp. *detonsa*, *P. pilosa* ssp. *sangamonensis* had relatively few alleles at polymorphic loci and had no unique alleles.

PHLOX PILOSA SSP. *SANGAMONENSIS* GENETIC DIVERSITY AND STRUCTURE. Larger populations of *P. pilosa* ssp. *sangamonensis* had higher allelic richness (Fig. 3A). Population fragmentation and erosion have been widely shown to reduce the number of alleles among plants (Aguilar et al. 2008). Yet, we did not find evidence that smaller populations were more homozygous—or inbred—on average. Small populations may have had comparable genetic measures to larger populations due to occasional gene flow, recent population declines combined with long-lived remnant individuals, or importantly for this taxon, barriers to inbreeding imposed by self-incompatibility (Ellstrand and Elam 1993; Luijten et al. 2000).

Phlox pilosa ssp. *sangamonensis* had low genetic variation compared to congeners *P. divaricata* ssp. *laphamii*, *P. pilosa* ssp. *fulgida*, and *P. pilosa* ssp. *pilosa*, and we were surprised at how genetically homogenous our samples were. Despite much greater sampling in *P. pilosa* ssp. *sangamonensis*, it had the fewest number of alleles and the second fewest private alleles at the taxon level. Moreover, the total number of alleles of just 11 individuals of *P. pilosa* ssp. *pilosa* in the P-L population equaled the total number of alleles found in 212 *P. pilosa* ssp. *sangamonensis* individuals (Table 1). We also note that *P. pilosa* ssp. *sangamonensis* had relatively low genetic variation when compared to findings from previous studies for most other *Phlox* taxa using some of the microsatellite loci we used. *Phlox pilosa* ssp. *sangamonensis* had lower H_o and H_e values than have been reported for most populations of *P. pilosa* ssp. *pilosa* (Fehlberg et al. 2008; Zale 2014) and *P. pilosa* L. ssp. *deamii* Levin (Fehlberg et al. 2014). However, there was comparable genetic variation between *P. pilosa* ssp. *sangamonensis* and that of *P. pilosa* ssp. *ozarkana* found by Zale (2014). Furthermore, *P. pilosa* ssp. *sangamonensis*

was more genetically variable than *Phlox pungens* Dorn, a species narrowly endemic to Wyoming (Waselkov et al. 2020). Although the methods of these other studies were not totally congruent to our methods, these other studies' findings suggest that *P. pilosa* ssp. *sangamonensis* may have relatively low genetic variation compared to other *Phlox* taxa.

Generally, most populations of *P. pilosa* ssp. *sangamonensis* seemed genetically similar (Fig. 2B). But a surprising finding was that there may have been weak genetic structure among *P. pilosa* ssp. *sangamonensis* populations. There was marginal evidence for isolation by distance in the Mantel test, and populations S-A and S-R were identified as somewhat divergent populations in STRUCTURE for *P. pilosa* ssp. *sangamonensis*. In fact, S-R had the greatest number of private alleles within *P. pilosa* ssp. *sangamonensis* (data not shown). Although it is possible these populations harbor distinct genetic diversity, we still caution against treating these populations as unique. We used a limited number of microsatellite loci—which represented a small portion of the genome—and some of which were affected by null alleles. This finding may have been due to the small sizes of those populations, which increases the likelihood of divergence due to bottleneck/founder effects and genetic drift.

DRIVERS OF REPRODUCTIVE SUCCESS OF *PHLOX PILOSA* SSP. *SANGAMONENSIS*. We found that reproductive success was associated with population size, neighbor distance, and genetic diversity. In general, smaller *P. pilosa* ssp. *sangamonensis* populations had lower reproductive success than larger ones. We showed a negative association between distance to neighbors and capsule formation. Our results were similar to Hendrix and Kyhl (2000), who showed that larger populations of *P. pilosa* ssp. *pilosa* in isolated prairie remnants had greater fruit set. In one year of seed set data, we also showed that smaller populations were producing fewer seeds per capsule compared to larger ones (Fig. 5). Our findings suggest that smaller populations face a "double tax" to their reproductive success: individuals in smaller populations may face lower formation of fruit from flowers, and they may then have lower seed for developed fruits (e.g., Hendrix and Kyhl 2000; Brys et al. 2004).

However, because of the limitations of our observational approach, we cannot be certain of the root causes of lower reproductive success in smaller *P. pilosa* ssp. *sangamonensis* populations. The three explanatory variables we tested for capsule formation were all correlated with one another and potentially confounding. For example, on average, smaller populations had substantially greater distances between flowering neighbors compared to larger ones (Appendix Fig. S6). Thus, our measured variables could have been indirectly capturing flowering density, which can be an even more important driver of reproductive success than population size (Bernhardt et al. 2008). Other unmeasured factors could have influenced these results, including pollen limitation (Waites and Ågren 2004), sympatric flowering plant species, and differential visitation by pollinators at each population (Seifan et al. 2014). For self-incompatible species, smaller populations can lose genetic diversity at mate compatibility loci, which can further reduce the number of suitable mates in the population (i.e., mate limitation; Wagenius et al. 2007). Future studies could investigate these factors and disentangle their interplaying effects to reproductive success.

Despite the difficulty of interpreting the causes of the results, they provide some insights into the reproductive success of *P. pilosa* ssp. *sangamonensis*. Flowering population size was a better predictor than nearest neighbor distances and genetic diversity. This suggests that flowering population size is important for its reproductive success, or that some other factors that corresponding to population size were important reproductive drivers. In addition to more suitable mates, additional benefits to having larger populations include being more visible or attractive to pollinators (Groom 1998; Cheptou and Avendaño 2006) or having more individuals that are in flowering synchrony (Ison et al. 2014). Other underlying factors could also lead to both greater population sizes and reproductive success, such as general habitat quality (e.g., Vergeer et al. 2003). Regardless of the specific causes, these results suggest reproductive success might be improved by increasing the sizes of the smaller populations (< 100 flowering individuals).

THE CONSERVATION STATUS AND MANAGEMENT OF *PHLOX PILOSA* SSP. *SANGAMONENSIS*. A byproduct of this study was the clarification of the current conservation status of *P. pilosa* ssp. *sangamonensis*. EOR data is available for the taxon, but is non-comprehensive, and only general comments are available by Levin and Smith (1965) and Levin (1984) about its historical population locations

and sizes. Broadly, our resurveys of *P. pilosa* ssp. *sangamonensis* populations suggest a decline during the past 40 years. We note that the extent of *P. pilosa* ssp. *sangamonensis* has likely declined since its description (c.f. Levin and Smith 1965; Fig. 1). Although detailed comparisons of the EOR data were not an objective of this study, those documents suggest severe localized declines for some populations. For example, our counts suggest that the S-R population declined by nearly 90% since the mid-1990s. Levin (1984) provided an important but unclear descriptor of historical population sizes, commenting that the "actual size of the system is about 4,000 plants." It was unclear if this referred to the "three large populations" of *P. pilosa* ssp. *sangamonensis* sampled in the study or all the populations known at the time. Regardless of the specific meaning, we counted fewer than half of the individuals counted by Levin (1984).

With the decline of *P. pilosa* ssp. *sangamonensis* in mind, our results offer potentially helpful information about habitat and genetic management of the populations. Future management activities include translocations to increase genetic diversity or establish new populations in ecological restorations to safeguard against localized population decline or habitat degradation. Because there appears to be relatively low genetic differentiation among most of the populations, we suggest the largest populations (S-6, S-L, and S-S) could be sensible propagule sources. These populations may also have the additional benefit of having relatively high genetic diversity (A_R). Because we found possible evidence of genetic variation in populations S-A and S-R, and these populations are small and relatively vulnerable, we suggest translocating individuals from those populations.

Conclusion. We have demonstrated the value of genetic analysis combined with field studies for the understanding the biology of rare, endemic taxa. Our results match other findings of other studies that have examined small, fragmented populations and plant reproduction. Specifically, we found that small population sizes were associated with lower genetic variation and poorer reproductive success. Moreover, distance to other mates was associated with poorer individual fruit set. Our study highlights the utility of combining genetic and field data to assist the conservation of this susceptible plant taxon.

Literature Cited

ABRAHAM, S. T., D. N. ZAYA, W. D. KOENIG, AND M. V. ASHLEY. 2011. Interspecific and intraspecific pollination patterns of valley oak, *Quercus lobata*, in a mixed stand in coastal central California. International Journal of Plant Sciences 172: 691–699.

ADAMACK, A. T. AND B. GRUBER. 2014. POPGENREPORT: simplifying basic population genetic analyses in R. Methods in Ecology and Evolution. 5: 384–387.

AGUILAR, R., M. QUESADA, L. ASHWORTH, Y. HERRERIAS-DIEGO, AND J. LOBO. 2008. Genetic consequences of habitat fragmentation in plant populations: Susceptible signals in plant traits and methodological approaches. Molecular Ecology 17: 5177–5188.

AGUILAR, R., E. J. CRISTÓBAL-PÉREZ, F. J. BALVINO-OLVERA, M. DE JESÚS AGUILAR-AGUILAR, N. AGUIRRE-ACOSTA, L. ASHWORTH, J. A. LOBO, S. MARTÉN-RODRÍGUEZ, E. J. FUCHS, G. SANCHEZ-MONTOYA, G. BERNARDELLO, AND M. QUESADA. 2019. Habitat fragmentation reduces plant progeny quality: A global synthesis. Ecology Letters 22, 1163–1173.

BARTON, K. 2020. MuMIn: Multi-model inference. R package version 1.43.17. https://CRAN.R-project.org/package=MuMIn.

BATES, D., M. MAECHLER, B. BOLKER, AND S. WALKER. 2015. Fitting linear mixed-effects models using lme4. Journal of Statistical Software 67: 1–48.

BERNHARDT, C. E., R. J. MITCHELL, AND H. J. MICHAELS. 2008. Effects of population size and density on pollinator visitation, pollinator behavior, and pollen tube abundance in *Lupinus perennis*. International Journal of Plant Sciences 169: 944–953.

BROOKFIELD, J. F.Y. 1996. A simple new method for estimating null allele frequency from heterozygote deficiency. Molecular Ecology 5: 453–455.

BRYS, R., H. JACQUEMYN, P. ENDELS, F. VAN ROSSUM, M. HERMY, L. TRIEST, L. DE BRUYN, AND G. D. E. BLUST. 2004. Reduced reproductive success in small populations of the self-incompatible *Primula vulgaris*. Journal of Ecology 92: 5–14.

CHEPTOU, P. O. AND L. G. AVENDAÑO. 2006. Pollination processes and the Allee effect in highly fragmented populations: Consequences for the mating system in urban environments. New Phytologist 172: 774–783.

DEMAURO, M. M. 1993. Relationship of breeding system to rarity in the lakeside daisy (*Hymenoxys acaulis* var. *glabra*). Conservation Biology 7: 542–550.

EARL, D. A. AND B. M. VONHOLDT. 2012. STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genetics Resources 4: 359–361.

ELLSTRAND, N. C. AND D. R. ELAM. 1993. Population genetic consequences of small population size: Implications for plant conservation. Annual Review of Ecology and Systematics 24: 217–242.

EVANNO, G., S. REGNAUT AND J. GOUDET. 2005. Detecting the number of clusters of individuals using the software structure: A simulation study. Molecular Ecology 14: 2611–2620.

FEHLBERG, S. D. AND C. J. FERGUSON. 2012. Intraspecific cytotypic variation and complicated genetic structure in the *Phlox amabilis–P. woodhousei* (Polemoniaceae) complex. American Journal of Botany 99: 865–874.

Fehlberg, S. D., M. C. Ty, and C. J. Ferguson. 2014. Reexamination of a putative diploid hybrid taxon using genetic evidence: The distinctiveness of *Phlox pilosa* subsp. *deamii* (Polemoniaceae). International Journal of Plant Sciences 175: 781–793.

Fehlberg, S. D., K. A. Ford, M. C. Ungerer, and C. J. Ferguson. 2008. Development, characterization and transferability of microsatellite markers for the plant genus *Phlox* (Polemoniaceae). Molecular Ecology Resources 8: 116–118.

Ferguson, C. J. and R. K. Jansen. 2002. A chloroplast DNA phylogeny of eastern *Phlox* (Polemoniaceae): Implications of congruence and incongruence with the ITS phylogeny. American Journal of Botany 89: 1324–1335.

Galpern, P., M. Manseau, P. Hettinga, K. Smith, and P. Wilson. 2012. allelematch: An R package for identifying unique multilocus genotypes where genotyping error and missing data may be present. Molecular Ecology Resources 12: 771–778.

Garner, A. G., B. E. Goulet-Scott, and R. Hopkins. 2022. Phylogenomic patterns of divergence and gene flow detail the evolution of reinforcement and hybrid speciation in *Phlox* wildflowers. BioRxiv [preprint]. https://doi.org/10.1101/2022.04.15.488502.

Gitzendanner, M. A. and P. S. Soltis. 2000. Patterns of genetic variation in rare and widespread plant congeners. American Journal of Botany 87: 783–792.

Groom, M. J. 1998. Allee effects limit population viability of an annual plant. The American Naturalist 151: 487–496.

Hendrix, S. D. and J. F. Kyhl. 2000. Population size and reproduction in *Phlox pilosa*. Conservation Biology 14: 304–313.

Hijmans, R. J. 2021. *geosphere*: Spherical trigonometry. https://cran.r-project.org/package=geosphere.

Illinois Endangered Species Protection Board. 2015. Checklist of Illinois endangered and threatened animals and plants. Springfield, IL.

Ison, J. L., S. Wagenius, D. Reitz, and M. V. Ashley. 2014. Mating between *Echinacea angustifolia* (Asteraceae) individuals increases with their flowering synchrony and spatial proximity. American Journal of Botany 101: 180–189.

Iverson, L. R., A. M. Prasad, and D. M. Ketzner. 1997. A summary of the Illinois flora based on the Illinois plant information network. Transactions of the Illinois State Academy of Science 90: 41–64.

Jombart, T. 2008. *adegenet*: An R package for the multivariate analysis of genetic markers. Bioinformatics 24: 1403–1405.

Jost, L. 2008. G_{st} and its relatives do not measure differentiation. Molecular Ecology 17: 4015–4026.

Kalinowski, S. T. 2004. Counting alleles with rarefaction: Private alleles and hierarchical sampling designs. Conservation Genetics 5: 539–543.

Kamvar, Z. N., J. F. Tabima, and N. J. Grünwald. 2014. *Poppr*: An R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. PeerJ 2: e281.

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

Leimu, R., P. Mutikainen, J. Koricheva, and M. Fischer. 2006. How general are positive relationships between plant population size, fitness and genetic variation? Journal of Ecology 94: 942–952.

Lenth, R. 2022. *emmeans*: Estimated marginal means, aka least-squares means. R package version 1.8.0. https://CRAN.R-project.org/package=emmeans.

Levin, D. A. 1966. The *Phlox pilosa* complex: Crossing and chromosome relationships. Brittonia 18: 142–162.

Levin, D. A. 1984. Genetic variation and divergence in a disjunct *Phlox*. Evolution 38: 223–225.

Levin, D. A. and H. W. Kerster. 1968. Local gene dispersal in *Phlox*. Evolution 22: 130–139.

Levin, D. A. and M. Levy. 1971. Secondary intergradation and genome incompatibility in *Phlox pilosa* (Polemoniaceae). Brittonia 23: 246–265.

Levin, D. A. and B. A. Schaal. 1970. Reticulate evolution in *Phlox* as seen through protein electrophoresis. American Journal of Botany 57: 977–987.

Levin, D. A. and D. M. Smith. 1965. An enigmatic *Phlox* from Illinois. Brittonia 17: 254–266.

Levy, M. 1983. Flavone variation and subspecific divergence in *Phlox pilosa* (Polemoniaceae). Systematic Botany 8: 118–126.

Li, Y. L. and J. X. Liu. 2018. StructureSelector: a web-based software to select and visualize the optimal number of clusters using multiple methods. Molecular Ecology Resources 18: 176–177.

Luijten, S. H., A. Dierick, J. Gerard, B. Oostermeijer, L. E.L. Raijmann, and H. C.M. Den Nijs. 2000. Population size, genetic variation, and reproductive success in a rapidly declining, self-incompatible perennial (*Arnica montana*) in the Netherlands. Conservation Biology 14: 1776–1787.

Mantel, N. 1967. The detection of disease clustering and a generalized regression approach. Cancer Research 27: 209–220.

Paradis, E. 2010. pegas: An R package for population genetics with an integrated–modular approach. Bioinformatics 26: 419–420.

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

R Core Team. 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Robertson, C. 1928. Flowers and insects. Lists of visitors of four hundred and fifty-three flowers. Science Press Printing Company, Lancaster, PA.

Robertson, K. R. 2001. Planning with plants in Illinois. pp. 28–45 *In* J. B. Phipps, and P. M. Catling, eds. Bioconservation and Systematics. Canadian Botanical Association, London, Ontario, Canada.

Seifan, M., E. M. Hoch, S. Hanoteaux, and K. Tielbörger. 2014. The outcome of shared pollination services is affected by the density and spatial pattern of an attractive neighbour. Journal of Ecology 102: 953–962.

Smith, D. M. and D. A. Levin. 1967. Karyotypes of eastern North American *Phlox*. American Journal of Botany 54: 324–334.

TAFT, J. B., G. S. WILHELM, D. M. LADD, AND L. A. MASTERS. 1997. Floristic Quality Assessment for vegetation in Illinois, a method for assessing vegetation integrity. Erigenia 15: 3–95.

TECIC, D. L., J. L. MCBRIDE, M. L. BOWLES, AND D. L. NICKRENT. 1998. Genetic variability in the federal threatened Mead's milkweed, *Asclepias meadii* Torrey (Asclepiadaceae), as determined by allozyme electrophoresis. Annals of the Missouri Botanical Garden 85: 97–109.

USDA NRCS. 2022. The PLANTS database. http://plants.usda.gov/.

VERGEER, P., R. RENGELINK, A. COPAL, AND N. J. OUBORG. 2003. The interacting effects of genetic variation, habitat quality and population size on performance of *Succisa pratensis*. Journal of Ecology 91: 18–26.

WAGENIUS, S. 2006. Scale dependence of reproductive failure in fragmented *Echinacea* populations. Ecology 87: 931–941.

WAGENIUS, S., E. LONSDORF, AND C. NEUHAUSER. 2007. Patch aging and the S-Allee effect: Breeding system effects on the demographic response of plants to habitat fragmentation. The American Naturalist 169: 383–397.

WAITES, A. R. AND J. ÅGREN. 2004. Pollinator visitation, stigmatic pollen loads and among-population variation in seed set in *Lythrum salicaria*. Journal of Ecology 92: 512–526.

WASELKOV, K., M. SANTIAGO, B. HEIDEL, M. H. MAYFIELD, AND C. J. FERGUSON. 2020. Population genetics of the Wyoming endemic *Phlox pungens* Dorn (Polemoniaceae). Western North American Naturalist 80: 369–380.

WHERRY, E. T. 1955. The genus *Phlox*. Associates of the Morris Arboretum, Philadelphia, PA.

WILLI, Y., T. N. KRISTENSEN, C. M. SGRÒ, A. R. WEEKS, M. ØRSTED, AND A. A. HOFFMANN. 2022. Conservation genetics as a management tool: The five best-supported paradigms to assist the management of threatened species. Proceedings of the National Academy of Sciences 119: e2105076119.

WINTER, D. J. 2012. mmod: An R library for the calculation of population differentiation statistics. Molecular Ecology Resources 12: 1158–1160.

WOODS, J. G., D. PAETKAU, D. LEWIS, B. N. MCLELLAN, M. PROCTOR, AND C. STROBECK. 1999. Genetic tagging of free-ranging black and brown bears. Wildlife Society Bulletin 27: 616–627.

YOUNG, A., T. BOYLE AND T. BROWN. 1996. The population and genetic consequences of habitat fragmentation for plants. Trends in Ecology and Evolution 11: 413–418.

ZALE, P. J. 2014. Germplasm collection, characterization, and enhancement of eastern *Phlox* species. Ph.D. thesis. The Ohio State University, Columbus, OH. 372 pp.

ZINNEN, J., J. W. MATTHEWS, AND D. N. ZAYA. 2024. Noteworthy collections: First record of a natural hybrid between *Phlox divaricata* ssp. *laphamii* (Alph. Wood) Wherry and *Phlox pilosa* ssp. *sangamonensis* (Levin & D.M. Sm.). Castanea 88: 184–190.

Appendix

FIG. S1. An example of sympatry of *P. pilosa* ssp. *sangamonensis* (upper right) and *P. divaricata* ssp. *laphamii* (upper and lower left). This photograph was taken at the S-S population.

FIG. S2. An example of the appearance of successfully dehisced capsules (viz., ballistic dispersal) for *Phlox pilosa* ssp. *pilosa*. Capsule appearances are almost identical to that of *P. pilosa* ssp. *sangamonensis*, though the *P. pilosa* ssp. *sangamonensis* is more glabrous than pictured here. This image was taken by Ken Robertson and can be found at http://phytoimages.siu.edu.

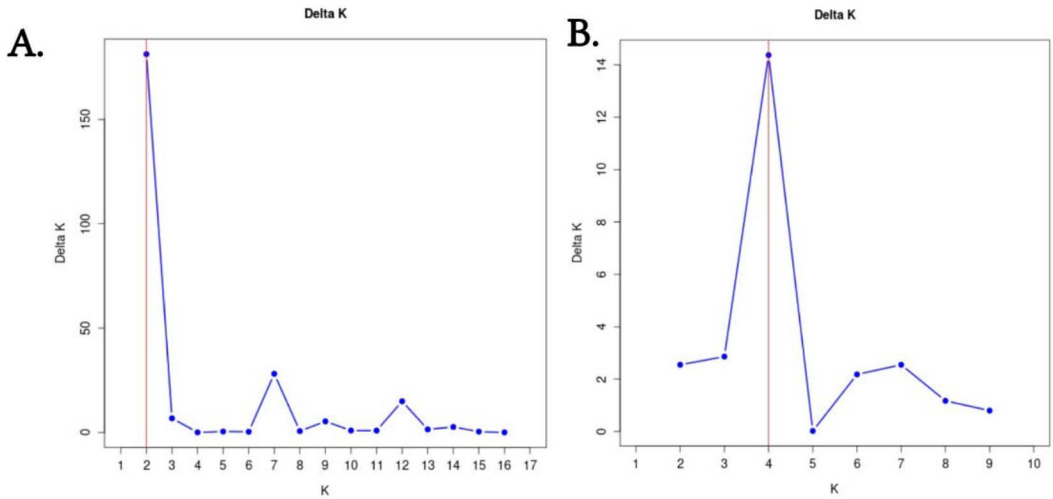

FIG. S3. Selecting the optimal numbers of K in STRUCTURE for all four *Phlox* taxa in the study (A), and the STRUCTURE that only considered the 10 populations of *P. pilosa* ssp. *sangamonensis* (B). Both selections represent the method described in Evanno et al. (2005).

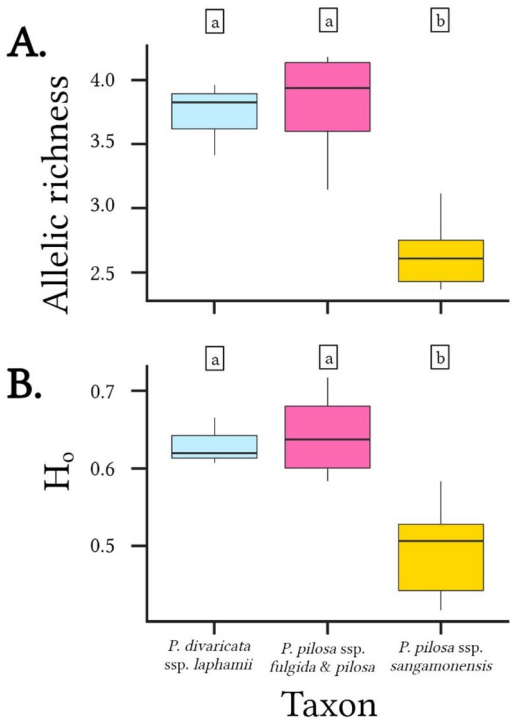

FIG. S4. Comparisons of population-level genetic diversity statistics among *Phlox* taxa. Two populations each of *P. pilosa* ssp. *fulgida* and *P. pilosa* ssp. *pilosa* were combined due to a lack of differentiation and low sample sizes. There was a significant difference of A_R among the taxa (A; $\chi^2 = 11.7$; df = 2; $P = 0.003$): values for *P. pilosa* ssp. *sangamonensis* were significantly lower than either *P. divaricata* ssp. *laphamii* or its *P. pilosa* conspecifics. Likewise, there was also a significant (B; $\chi^2 = 11.4$; df = 2; $P = 0.003$) difference in observed heterozygosity (H_o) among the taxa; *P. pilosa* ssp. *sangamonensis* had significantly lower H_o. Uncapitalized letters indicate the results of post hoc tests for differences after Bonferroni correction.

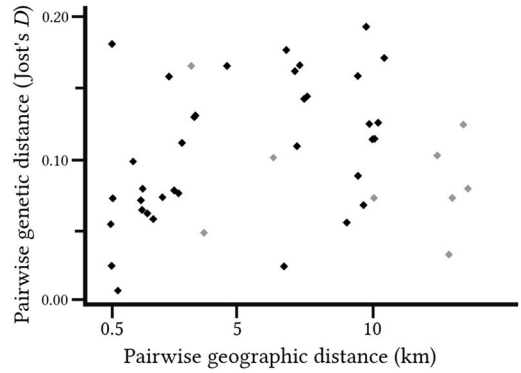

FIG. S5. Relationship between genetic and geographic distance in *Phlox pilosa* ssp. *sangamonensis*. There was marginal statistical support for isolation by distance (Mantel test, P = 0.059) for isolation by distance. Genetic distances were calculated using Jost's *D* (Jost 2008). Pairwise differences for the S-S population are shown as gray points; the isolation by distance relationship may have been obscured by S-S having an unusually small genetic distance to geographically distant populations.

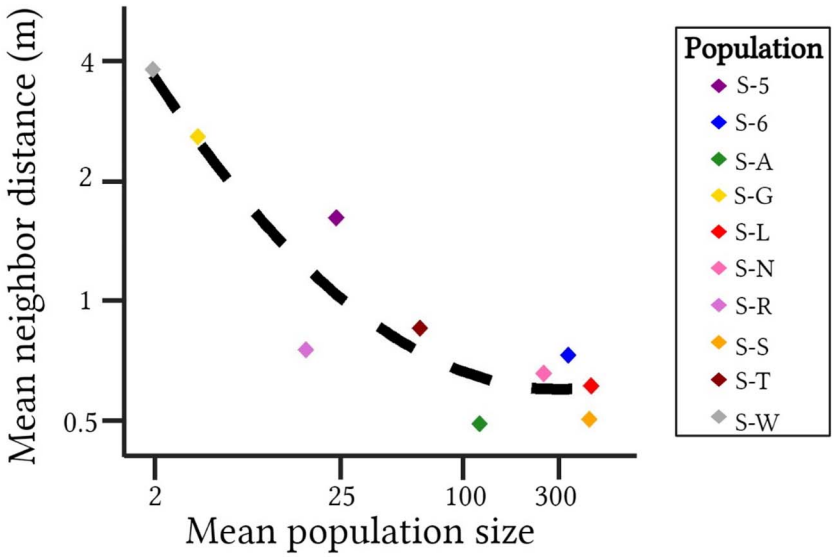

FIG. S6. Larger population size is associated with reduced mean neighbor distances for individuals of *P. pilosa* ssp. *sangamonensis*. Each data point shows the mean across the population for year, as well as the mean across individuals for mean neighbor distance. Both axes are displayed on the log-scale to improve data interpretation. The trendline represents the quadratic relationship between mean population size and mean neighbor distance.

Table S1. Basic descriptions of the six polymorphic microsatellite loci used for the study.

Locus	Forward and reverse primer	Repeat sequence	Data source
Phl-28	F: 5′-GTTGCCACCTCACAGATTCC-3′ R: 5′-AATTGGGCGGTAAAAATGAA-3′	$(CT)_{10}$	Fehlberg and Ferguson (2012)
Phl-33	F: 5′-GTCCCCCAGCTGATACTGGT-3′ R: 5′-CTTGCTTTGGCTTCTCCAAC-3′	$(GA\text{-}GA)_{22}$	Fehlberg et al. (2008)
Phl-84	F: 5′-TCAGACTAGGGGAGGCAGTG-3′ R: 5′-TTCTTTACCGCTGGCTGAAT-3′	$(CA)_{10}$	Fehlberg et al. (2008)
Phl-113	F: 5′-TGTCCACATGGGCTTGACTA-3′ R: 5′-ACGTACACGCCCAACTAAGG-3′	$(AG)_{10}$	Fehlberg et al. (2008)
Phl-115	F: 5′-CGCGCACTAGATCACTACCA-3′ R: 5′-GGGCGTCACCAACTTGTATT-3′	$(AG\text{-}AG)_{10}$	Fehlberg et al. (2008)
Phl-137	F: 5′-TCGGGCACCAGATTTTATTC-3′ R: 5′-TTCGACCCCCAGATAGTCAG-3′	$(GA\text{-}GA)_{20}$	Fehlberg et al. (2008)

Table S2. Six microsatellite loci we attempted to use but abandoned due to challenges with our *P. pilosa* ssp. *sangamonensis* samples.

Locus	Forward and reverse primer	Repeat sequence	Data source	Reason for abandonment and comments
Phl-68	F: 5′-ACGCAACAACCAAACTCCAT-3′ R: 5′-GATGCAGCCACACGAGTTTA-3′	$(CA\text{-}CA)_{17}$	Fehlberg et al. (2008)	There was a universal monomorphic fragment in *P. pilosa* ssp. *sangamonensis* (337bp).
Phl-77	F: 5′-AATAAGTTCAGCCGGGAAGG-3′ R: 5′-CCTACTGAGGACCCACCAAA-3′	$(AC)_{15}$	Fehlberg et al. (2008)	Poor amplification of fragments in the expected allele size range for *P. pilosa* ssp. *sangamonensis*; target amplification products possibly eclipsed by universal nontarget product with a size of ~318bp.
Phl-121	F: 5′-CTACTCCACGCTGCCCTTAC-3′ R: 5′-TCCTAGCTCGTAAAGCTCCA-3′	$(AC\text{-}AC)_{21}$	Fehlberg et al. (2008)	Substantial nonspecific amplification throughout the plausible microsatellite range made scoring infeasible for *P. pilosa* ssp. *sangamonensis*.
M-4	F: 5′-GTTCGCCGGAGATAGTTACG-3′ R: 5′-GGTAAACCCACGGGAGAACT-3′	$(GA)_7$	Fehlberg (unpublished data)	Monomorphic in *P. pilosa* ssp. *sangamonensis* (181bp) and extremely low variation (mostly 183bp) and/or no amplification in *P. pilosa* ssp. *pilosa*. Poor amplification and extremely low variation in *P. divaricata* ssp. *laphamii*.
M-187	F: 5′-CGAATCAATGAGGTAACTGCTG-3′ R: 5′-TCCTCCACCATTGTTGAAGTT-3′	$(ACC)_6$	Fehlberg (unpublished data)	Putatively monomorphic (445bp in *P. pilosa* ssp. *sangamonensis;* ~439bp in *P. divaricata* ssp. *laphamii*) and substantial nonspecific amplification.
M-225	F: 5′-TGCCGAACATAGAAACAGAAGA-3′ R: 5′-ATTTTCTCAACCCAGATGTGCT-3′	$(GA)_{13}$	Fehlberg (unpublished data)	Substantial nonspecific amplification throughout the plausible microsatellite range made scoring infeasible for *P. pilosa* ssp. *sangamonensis*.

Table S3. Estimates of null allele frequencies for six microsatellite loci in four *Phlox* taxa. For each observed frequency, we provide the 95% CI of the null allele frequency in parentheses. Superscripting parentheses, "ns" denotes no significant evidence ($p > 0.05$) of null alleles in that locus for the taxon, whereas "*" indicates significant ($p < 0.05$) evidence of null alleles in that locus for the taxon.

Locus	*P. pilosa* ssp. *sangamonensis*	*P. divaricata* ssp. *laphamii*	*P. pilosa* ssp. *fulgida*	*P. pilosa* ssp. *pilosa*
Phl-28	0.035 (−0.005–0.080)^{ns}	0.046 (−0.049–0.143)^{ns}	0.187 (0.053–0.314)*	0.003 (−0.098–0.139)^{ns}
Phl-33	0.042 (−0.005–0.086)^{ns}	0.044 (−0.042–0.146)^{ns}	0.073 (−0.048–0.200)^{ns}	0.020 (−0.049–0.086)^{ns}
Phl-84	0.101 (0.059–0.141)*	−0.004 (−0.070–0.073)^{ns}	0.017 (−0.063–0.088)^{ns}	−0.022 (−0.062–0.012)^{ns}
Phl-113	0.038 (0.008–0.071)*	0.059 (−0.012–0.134)^{ns}	−0.004 (−0.107–0.104)^{ns}	0.037 (−0.057–0.129)^{ns}
Phl-115	−0.009 (−0.035–0.021)^{ns}	−0.001 (−0.079–0.090)^{ns}	0.057 (−0.066–0.209)^{ns}	0.038 (−0.049–0.157)^{ns}
Phl-137	0.034 (−0.005–0.076)^{ns}	0.100 (−0.013–0.221)^{ns}	0.177 (0.037–0.311)*	0.060 (−0.040–0.169)^{ns}

Table S4. Observed (H_o) and expected (H_e) heterozygosity data organized by the six loci across all 17 populations in the study. Bolded H_o values indicate that population significantly deviated from Hardy–Weinberg equilibrium (HWE): * $P < 0.05$ and *** $P < 0.001$.

		Locus Phl-28		Locus Phl-33		Locus Phl-84		Locus Phl-113		Locus Phl-115		Locus Phl-137	
Taxon	Pop.	H_o	H_e	H_o	H_e	H_o	H_e	H_o	H_e	H_o	H_e	H_o	H_e
P. pilosa ssp. *sangamonensis*	S-5	0.455	0.573	0.318	0.491	0.591	0.711	0.818	0.742	0.455	0.376	0.045	0.044
	S-6	0.533	0.540	0.500	0.516	0.667	0.805	0.800	0.841	0.467	0.365	0.533	0.556
	S-A	0.409	0.503	**0.059*****	0.438	0.667	0.582	0.773	0.795	0.591	0.491	0.045	0.044
	S-G	0.286	0.439	0.429	0.337	0.857	0.663	1.000	0.765	0.000	0.000	0.429	0.337
	S-L	0.536	0.590	0.321	0.499	0.714	0.669	0.714	0.809	0.000	0.000	0.357	0.354
	S-N	0.478	0.474	0.565	0.509	0.696	0.754	0.870	0.798	0.435	0.386	0.348	0.423
	S-R	0.474	0.389	0.632	0.432	**0.526***	0.611	0.778	0.631	0.526	0.483	0.263	0.229
	S-S	0.414	0.476	0.483	0.450	0.586	0.674	0.862	0.788	0.207	0.185	0.517	0.494
	S-T	0.500	0.526	0.571	0.497	0.643	0.709	**0.679***	0.754	0.250	0.219	0.429	0.459
	S-W	0.750	0.531	0.000	0.375	0.750	0.656	0.500	0.563	0.000	0.000	0.500	0.500
P. divaricata ssp. *laphamii*	D-L	0.667	0.680	0.529	0.609	0.625	0.512	0.941	0.865	0.765	0.780	**0.467***	0.676
	D-N	0.692	0.799	0.538	0.601	0.692	0.704	**0.615*****	0.882	0.846	0.740	0.333	0.288
	D-S	0.500	0.403	0.429	0.367	0.857	0.796	0.857	0.745	0.714	0.724	0.286	0.449
P. pilosa ssp. *fulgida*	PF-B	0.182	0.165	0.727	0.674	0.818	0.736	0.545	0.397	0.455	0.541	0.909	0.835
	PF-W	0.625	0.461	**0.625***	0.859	0.875	0.859	0.750	0.773	0.500	0.469	**0.125***	0.539
P. pilosa ssp. *pilosa*	PP-L	0.545	0.492	0.818	0.860	1.000	0.905	0.727	0.773	0.100	0.095	0.818	0.814
	PP-P	0.500	0.490	0.917	0.878	0.917	0.868	0.833	0.830	0.500	0.497	0.636	0.744

Table S5. Locus-specific information on allelic richness and fragment size in the four taxa used in the study. For the "Taxon" column, *S* = *P. pilosa* ssp. *sangamonensis*; *D* = *P. divaricata* ssp. *laphamii*; *PF* = *P. pilosa ssp. fulgida*; *PP* = *P. pilosa* ssp. *pilosa*.

Locus	Taxon	A	Size range (bp)
Phl-28	*S*	3	260–262
	D	8	258–268
	PF	5	261–265
	PP	5	262–266
		$Total_{unique}$: 10	$Total_{range}$: 258–268
Phl-33	*S*	4	177–195
	D	8	139–187
	PF	11	152–223
	PP	16	150–185
		$Total_{unique}$: 27	$Total_{range}$: 139–223
Phl-84	*S*	16	183–236
	D	13	179–230
	PF	13	188–228
	PP	18	188–224
		$Total_{unique}$: 33	$Total_{range}$: 179–236
Phl-113	*S*	14	491–521
	D	22	485–535
	PF	7	495–573
	PP	15	484–573
		$Total_{unique}$: 37	$Total_{range}$: 484–573
Phl-115	*S*	4	391–401
	D	9	394–404
	PF	5	397–403
	PP	4	396–404
		$Total_{unique}$: 12	$Total_{range}$: 391–404
Phl-137	*S*	3	226–230
	D	8	222–233
	PF	10	224–236
	PP	11	222–237
		$Total_{unique}$: 15	$Total_{range}$: 222–237

Table S6. Results from analysis of molecular variance (AMOVA) within each *Phlox* taxa. AMOVA partitions genetic diversity among and between populations. AMOVA analyses were performed on the different taxa separately. Note that for this table, *P. pilosa* ssp. *fulgida* and *P. pilosa* ssp. *pilosa* were pooled due to the results of the Bayesian clustering analysis with STRUCTURE (Fig. 2A).

Taxon		df	Sum of squares	% variation
P. pilosa ssp. *sangamonensis*	Among populations	9	134.8	8.0
	Among samples within population	202	676.3	4.5
	Within samples	212	643.8	87.5
	Total	423	1454.8	100.00
P. divaricata ssp. *laphamii*	Between populations	2	11.9	4.0
	Between samples within population	34	104.8	4.3
	Within samples	37	104.3	91.7
	Total	73	221.0	100.00
P. pilosa ssp. *fulgida* and *P. pilosa* ssp. *pilosa*	Between populations	3	57.5	14.7
	Between samples within population	38	164.3	4.2
	Within samples	42	164.6	81.1
	Total	83	386.33	100.00

Table S7. Model selection results for three explanatory variables used to predict reproductive success. Model terms: S = reproductive success (proportional response variable), Y = year, N = log-transformed flowering population size per season, (1|P) = population random effect, D = log-transformed mean of distance to two nearest neighbors, A = rarified allelic richness.

	Model	Intercept	Effect term	df	AICc	ΔAICc	Weight
Population	$S \sim Y + N + (N\|P)$	−0.49	0.112	7	4423.5	0.000	> 0.999
	$S \sim Y + N + (1\|P)$	−0.13	−0.237	5	4605.9	182.4	0.0
	$S \sim Y + (1\|P)$	−1.14	NA	4	4615.6	192.1	0.0
Distance	$S \sim Y + D + (D\|P)$	0.087	−0.360	7	4437.4	0.000	> 0.999
	$S \sim Y + D + (1\|P)$	−0.250	−0.252	5	4483.1	45.7	0.0
	$S \sim Y + (1\|P)$	−1.135	NA	4	4615.6	178.2	0.0
Rarified allelic richness	$S \sim Y + A + (1\|A)$	−7.923	2.587	5	4609.0	0.0	0.96
	$S \sim Y + (1\|P)$	−1.135	NA	4	4615.6	6.59	0.04