Reexamination of a Putative Diploid Hybrid Taxon Using Genetic Evidence: The Distinctiveness of *Phlox pilosa* subsp. *deamii* (Polemoniaceae)

Author(s): Shannon D. Fehlberg, Maureen C. Ty, and Carolyn J. Ferguson

Source: *International Journal of Plant Sciences*, Vol. 175, No. 7 (September 2014), pp. 781-793
Published by: The University of Chicago Press
Stable URL: http://www.jstor.org/stable/10.1086/677228
Accessed: 10/10/2014 14:37

Int. J. Plant Sci. 175(7):781–793. 2014.

1058-5893/2014/17507-0005$15.00 DOI: 10.1086/677228

REEXAMINATION OF A PUTATIVE DIPLOID HYBRID TAXON USING GENETIC EVIDENCE: THE DISTINCTIVENESS OF *PHLOX PILOSA* SUBSP. *DEAMII* (POLEMONIACEAE)

Shannon D. Fehlberg,* Maureen C. Ty,[1,†] and Carolyn J. Ferguson[2,†]

*Research, Conservation, and Collections, Desert Botanical Garden, 1201 North Galvin Parkway, Phoenix, Arizona 85008, USA; and †Herbarium and Division of Biology, Kansas State University, Manhattan, Kansas 66506, USA

Editor: Alan W. Meerow

Premise of research. A number of recent studies have shown that hybridization plays a key role in the evolution of plants and can lead to the development of genetic and taxonomic novelty, even without a change in ploidy level. Documenting patterns of population genetic diversity for taxa of proposed homoploid hybrid origin and their putative parents leads to an improved understanding of the role of hybridization in the generation of plant diversity.

Methodology. *Phlox pilosa* subsp. *deamii* is hypothesized to have arisen from diploid hybridization between *P. pilosa* subsp. *pilosa* and *Phlox amoena*. Here we analyzed 18 populations representing all three taxa using a population genetic approach, including microsatellite and chloroplast DNA sequence data. We addressed questions concerning genetic diversity and structure, hybrid origin, and ongoing evolutionary processes.

Pivotal results. Several different analyses revealed that genetic variation in *P. pilosa* subsp. *deamii* was a mixture of (or was intermediate to) the genetic variation found in *P. pilosa* subsp. *pilosa* and *P. amoena*. However, populations of *P. pilosa* subsp. *deamii* were also notably distinct genetically and maintained high levels of genetic diversity. Comparisons between nuclear and chloroplast genetic data provided some evidence for ongoing gene exchange within this system.

Conclusions. *Phlox pilosa* subsp. *deamii* is distinctive in its genetics, as well as its morphology and ecology. Combined with data from previous studies, population genetic data support a diploid hybrid origin for *P. pilosa* subsp. *deamii*, although with sufficient time since its origin for establishment of unique genetic variation. Furthermore, its formation may have come about through complex interactions among the closely related parental taxa, potentially with multiple hybrid generations, backcrossing, and introgression. Such a complex scenario of formation for this subspecific taxon raises questions with respect to the current distinction we make between homoploid hybrid origin of taxa, which is viewed as rare, and introgression, which is undoubtedly prevalent in plants.

Keywords: genetic diversity, hybrid origin, hybrid species, introgression, microsatellites, *Phlox*, reproductive isolation.

Online enhancement: appendix table.

Introduction

Hybridization is recognized as an important process in plant evolution, and the creative role that hybridization plays in the generation of genetic and taxonomic novelty (e.g., adaptive introgression, hybrid origin of taxa) is of particular interest (Abbott et al. 2013; Soltis 2013). While taxonomic novelty generated via allopolyploidy has been well documented (see Soltis and Soltis 2009; e.g., Kim et al. 2008; Brokaw and Hufford 2010; Clarkson et al. 2010; Weiss-Schneeweiss et al. 2012), there are few well-supported cases of taxa of homoploid hybrid origin (see Abbot et al. 2010; e.g., Arnold 1993; James and Abbott 2005; Rieseberg 2006; Pan et al. 2007; see also table 1 in Paun et al. 2009). This may be due to the rarity of homoploid hybrid origins or, perhaps more likely, to the difficulty of detecting and definitively identifying such origins (Mallet 2007; Kelly et al. 2010; Clay et al. 2012). To better understand the role of homoploid hybridization in generating biodiversity, additional examples and study systems are needed.

In contrast to allopolyploidization (with ploidy level differences rapidly effecting intrinsic postzygotic reproductive isolation), the establishment of a stable, fertile, and reproductively

[1] Current address: Division of Parasitology, Department of Microbiology, New York University School of Medicine, New York, New York 10016, USA.

[2] Author for correspondence; e-mail: ferg@ksu.edu.

Manuscript received December 2013; revised manuscript received April 2014; electronically published July 29, 2014.

isolated lineage resulting from hybridization between two distinct species without a change in ploidy level requires a combination of several different processes. Reproductive isolation in such cases may involve chromosomal rearrangements (intrinsic postzygotic), ecological divergence, and/or spatial isolation (ecological prezygotic and extrinsic postzygotic; Gross and Rieseberg 2005). The results of several theoretical models demonstrate that the colonization of a novel habitat by a homoploid hybrid taxon is critical to its establishment and persistence as an independent, stable entity (McCarthy et al. 1995; Buerkle et al. 2000, 2003). Furthermore, all confirmed cases of homoploid hybrid origins show evidence of ecological divergence from parental taxa (reviewed in Gross and Rieseberg 2005; Abbott and Rieseberg 2012).

Our understanding of the process of homoploid hybrid formation of taxa improves as more examples are documented. Genetic studies are foundational work yielding evidence relative to hypotheses of hybrid origins. Data from both phylogenetic studies with broad taxonomic sampling and focused population genetic studies have been employed to investigate homoploid hybrid origins. When phylogenetic data are available from multiple gene regions, incongruence among gene trees and/or divergent or recombinant copies of nuclear genes can provide support for hypotheses of hybrid origins of taxa (e.g., Kelly et al. 2010). Some well-known examples of taxa of hybrid origin have also been investigated using a population genetic approach with data from microsatellite loci (e.g., *Helianthus* [Gross et al. 2003], *Senecio* [Brennan et al. 2012]), and such data have been informative for other taxa as well (Sherman and Burke 2009; Yuan et al. 2010; Rentsch and Leebens-Mack 2012; Jiang et al. 2013). Although taxa of homoploid hybrid origin may show expected patterns of parental genome additivity, other patterns are also possible due to phenomena such as shared alleles between parental taxa, interaction among multiple hybrid generations, backcrossing, and multiple origins (Soltis and Soltis 2009). Taxa of hybrid origin may also have had enough time since their formation to accumulate novel alleles through processes such as mutation and recombination (e.g., Golding and Strobeck 1983; Bradley et al. 1993; Hoffman and Brown 1995) and may be notably genetically distinct from parental taxa (Brennan et al. 2012; Gross 2012). Continued investigation of population genetic patterns in taxa of suspected hybrid origin increases our knowledge of the expected range of parental genome representation and levels of diversity and divergence in hybrid derivatives under varying circumstances.

The genus *Phlox* (Polemoniaceae) is known for both diploid and polyploid hybridization (e.g., Levin 1967, 1968; Smith and Levin 1967; Levin and Schaal 1970*a*, 1970*b*; Levy and Levin 1974, 1975), and there is extensive background information on the study system, including work on taxonomy, evolutionary process, and phylogeny (e.g., Wherry 1955; Levin 1975, 1978; Levin and Kerster 1968; Ferguson et al. 1999; Ferguson and Jansen 2002; Fehlberg and Ferguson 2012; Hopkins et al. 2012). One particular taxon, *Phlox pilosa* L. subsp. *deamii* D. Levin, is strongly supported as being of homoploid hybrid origin on the basis of previous work and presents an excellent case for focused population genetic study. *Phlox pilosa* subsp. *deamii* is a diploid (Smith and Levin 1967) and is thought to be the hybrid derivative of *P. pilosa* subsp. *pilosa* and *Phlox amoena* Sims (Levin 1966; Levin and Smith 1966; Levin and Schaal 1970*b*). *Phlox pilosa* subsp. *deamii* occurs uncommonly in open woods from a portion of northwestern Tennessee through western Kentucky and north into southern Indiana (Levin 1966). The putative parental taxa are each wide ranging and their general ranges overlap or abut those of *P. pilosa* subsp. *deamii*, but fine-scale distributions are principally allopatric. *Phlox pilosa* subsp. *pilosa* occurs on moderately rich soils of grasslands and forest edges from the eastern United States through the midwestern and southern states to the eastern edge of the Great Plains. Chromosome counts and flow cytometry data indicate that this taxon is diploid across most of its range, although tetraploid populations occur in some areas on the western edge of the range in Oklahoma and Texas, and possibly in Arkansas (Smith and Levin 1967; Worcester et al. 2012). *Phlox amoena*, which is diploid (e.g., Flory 1934; Smith and Levin 1967), occurs on more xeric, often sandy neutral-acidic soils of the southeastern coastal plain and into the adjacent piedmont and interior plateau. It reaches the northern and western limits of its range in south-central Kentucky and eastern Alabama, respectively. Where *P. pilosa* subsp. *pilosa* and *P. amoena* occur in local sympatry, they are usually spatially separated due to different habitat preferences (although rare hybrid zones are reported; Levin 1966; Levin and Smith 1966).

Levin first put forth the hypothesis of hybrid origin of *P. pilosa* subsp. *deamii* when describing the taxon (Levin 1966) and, over time (with several coworkers), presented supporting data from field and morphological observations, crossing studies, and electrophoretic study of seed proteins. This putative hybrid derivative strikingly combines morphological attributes of *P. pilosa* subsp. *pilosa* and *P. amoena* (Levin 1966; Levin and Smith 1966; see also Wherry 1955). For example, it exhibits intermediacy with respect to the extent of vegetative growth, leaf shape, and compactness of the inflorescence. It exhibits the long, eglandular calyx vestiture characteristic of *P. amoena* but the externally pubescent corolla tube typically exhibited by *P. pilosa* subsp. *pilosa* (see Levin 1966; Levin and Smith 1966). Greenhouse crossing studies of eastern *Phlox* demonstrated that crosses could be effected readily among diploid accessions of the *P. pilosa* complex (Levin 1966), and percentage seed set for crosses involving *P. pilosa* subsp. *pilosa*, *P. pilosa* subsp. *deamii*, and *P. amoena* ranged from 28% to 42% (Levin and Smith 1966), highlighting the potential for current (and, by inference, historical) hybridization in the group. Artificial F_1 hybrids between the putative parental taxa exhibited general morphological intermediacy (Levin and Smith 1966). Moreover, two natural hybrid zones between *P. pilosa* subsp. *pilosa* and *P. amoena* were discovered (in Winston County, AL, and Decatur County, TN, both outside the range of *P. pilosa* subsp. *deamii*; Levin and Smith 1966). Through morphological study of plants in these hybrid zones, Levin and Smith (1966) detected several putative F_1's (morphologically similar to the artificial hybrids) and putative backcrosses, with one individual completely matching *P. pilosa* subsp. *deamii* for the morphological characters investigated. The natural occurrence of individuals approximating *P. pilosa* subsp. *deamii* was discussed as a "resynthesis" of the putative hybrid derivative from its putative parental taxa.

Finally, Levin and Schaal (1970*b*) included *P. pilosa* subsp.

deamii in a seed protein study of several eastern North American *Phlox*. Seed proteins were separated by gel electrophoresis, and the resulting patterns of bands were considered "protein profiles" for the taxa. They found that *P. pilosa* subsp. *pilosa* and *P. amoena* had distinctive profiles differing in four bands (one band specific to the former, based on 12 populations from Illinois and Indiana, and three bands specific to the latter, based on two populations from Tennessee; Levin and Schaal 1970*b*). As part of their study, they generated artificial hybrids (among various species) and confirmed that the F_1 hybrids exhibited additive protein profiles relative to their parents (with no new bands generated), supporting the assumption that the protein markers are inherited in a codominant fashion and are useful for inferring origins. Samples of *P. pilosa* subsp. *deamii* from two Kentucky populations exhibited a protein profile completely additive to those of the putative parental taxa (and also matching that of the artificial hybrids; Levin and Schaal 1970*b*). The protein data thus provided further evidence for the hybrid origin of *P. pilosa* subsp. *deamii*.

More recently, *P. pilosa* subsp. *deamii* was included in phylogenetic studies focused on eastern North American members of the genus (using the internal transcribed spacer [ITS] region of the nrDNA and cpDNA restriction site and sequence data; Ferguson et al. 1999, 2008; Ferguson and Jansen 2002). Broad sampling included all eastern species of *Phlox* and numerous samples of members of the *P. pilosa* species complex (see Ferguson 1998 for a discussion of taxonomy), including one sample of *P. pilosa* subsp. *deamii*. These phylogenetic findings held taxonomic implications for *Phlox*, but the most striking result from these studies was complicated patterns of relationships for groups of eastern North American *Phlox* in each phylogeny coupled with strong incongruence between the ITS and cpDNA trees (see Ferguson and Jansen 2002). *Phlox pilosa* subsp. *deamii* always resolved with some samples of *P. pilosa* subsp. *pilosa* (as well as with other taxa, including *P. amoena* in the cpDNA tree); however, in each case it was a member of a clade exhibiting some incongruence with the alternate phylogeny (Ferguson and Jansen 2002). With respect to prior hypotheses of hybrid origins for particular *Phlox* taxa, including *P. pilosa* subsp. *deamii*, the phylogenies did not provide simple, straightforward support (e.g., a putative hybrid derivative grouping with one putative parent in one phylogeny and with the other putative parent in another; see Ferguson and Jansen 2002). Moreover, the broad sampling revealed the presence of noteworthy genetic variation, particularly within *P. pilosa* across its range. For example, samples of *P. pilosa* subsp. *pilosa* (which included geographical representation from 11 states; Ferguson et al. 1999) were placed among multiple distinct, well-supported clades in each phylogeny. The authors concluded that the *P. pilosa* complex has a complicated evolutionary history, consistent with a history of hybridization and introgression (also postulated by Levin [1966]), and that further, focused studies of putative hybrid derivatives at the population level are needed to advance our understanding of the role of hybridization in the evolution of this group.

In this study, we use a population genetic approach to investigate diversity in the putative homoploid hybrid derivative *P. pilosa* subsp. *deamii*. Individuals from populations spanning the range of *P. pilosa* subsp. *deamii* and individuals of the putative parental taxa from geographically proximate populations are assessed for five *Phlox*-specific microsatellite loci. In addition, three chloroplast DNA regions are sequenced and analyzed for a subset of individuals. We seek to address the following questions: (1) What is the genetic diversity and structure found within and among populations of *P. pilosa* subsp. *deamii*? (2) How does this diversity and structure compare to that found in geographically proximate populations of the putative parental taxa, *P. pilosa* subsp. *pilosa* and *P. amoena*? (3) Do population-level molecular data support the hypothesis of a homoploid hybrid origin for *P. pilosa* subsp. *deamii*, as previously proposed on the basis of ecological, morphological, crossing, and protein electrophoretic data? (4) How do patterns of nuclear microsatellite diversity compare to general patterns of chloroplast haplotype diversity among populations? (5) What possible evolutionary processes are evidenced by the observed patterns of genetic diversity, admixture, and/or additivity in *P. pilosa* subsp. *deamii*?

Material and Methods

Plant Material and DNA Extraction

Leaf material was collected from a total of 136 individuals from six populations of *Phlox pilosa* subsp. *deamii* (including all populations that could be located during the sampling season and spanning the general range of this rare taxon), seven populations of *P. pilosa* subsp. *pilosa*, and five populations of *Phlox amoena* (fig. 1; app. A). The collection sites ranged from southern Indiana and Illinois through Kentucky and Tennessee to northern Alabama (fig. 1). The putative parental taxa were not found co-occurring with *P. pilosa* subsp. *deamii*. However, *P. pilosa* subsp. *pilosa* and *P. amoena* were found together at one site in south-central Tennessee: in Decatur County, in the same general area (though not the precise locality) of a hybrid zone mentioned by Levin and Smith (1966). Individuals of *P. pilosa* subsp. *pilosa* and *P. amoena* co-occurring in that locality were included in the study (see populations P3 and A2; fig. 1; app. A). Voucher specimens were collected for each population and deposited in the Kansas State University Herbarium (KSC; app. A). Leaf material was flash-frozen in liquid nitrogen in the field and stored at −80°C until DNA extraction. Total genomic DNA was extracted using the small-scale CTAB extraction protocol of Loockerman and Jansen (1996), modified from Doyle and Doyle (1987). Diploid ploidy levels for all populations were confirmed through flow cytometry analysis of frozen leaf material following the same methods as Worcester et al. (2012; based on one individual per population; data not shown).

Nuclear Microsatellite Genotyping

Population genetic variation was assessed using five microsatellite loci developed in our laboratory: PHL84, PHL98, PHL113, PHL121, and PHL137 (Fehlberg et al. 2008). Amplification procedures followed those described by Fehlberg et al. (2008). Amplification products were screened by gel electrophoresis to determine reaction success, diluted according to band intensity, and genotyped on an ABI 3130xl genetic analyzer at the Kansas State University DNA sequencing and genotyping facility. ROX 500 or LIZ 600 (Life Technologies, Grand Island, NY) was used as an internal standard in each lane, depending

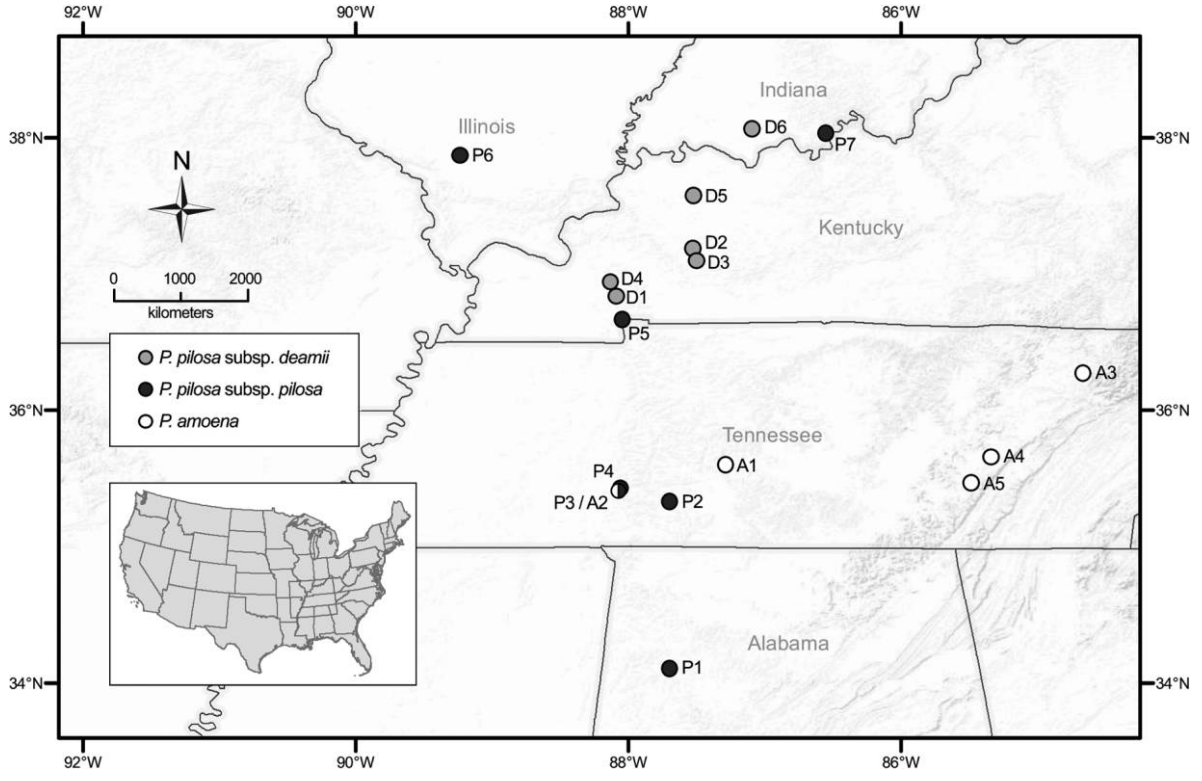

Fig. 1 Geographic localities of the 18 study populations of *Phlox pilosa* subsp. *deamii*, *P. pilosa* subsp. *pilosa*, and *Phlox amoena* in the midwestern United States (population name details are provided in app. A).

on the size of the fragment. The raw output from the genetic analyzer was read using the program GENEMARKER (ver. 1.8; SoftGenetics, State College, PA), and allele sizes were checked manually and recorded.

Chloroplast DNA Sequencing

To augment population genetic data for these particular populations, phylogenetic data were obtained. Three regions were amplified and sequenced from the chloroplast genome for one individual per study population: the *trnL-trnF* region (*trnL* intron–3′–*trnL-trnF*, using primers "c" and "f" of Taberlet et al. 1991), *trnH-psbA* (Shaw et al. 2005), and *trnD-trnT* (Demesure et al. 1995). PCR was performed in 50-μL volumes containing 1 × PCR buffer (Promega, Madison, WI), 2.5 mM $MgCl_2$, 0.2 mM each dNTP (Promega), 0.2 μM each primer, 1.25 U of *taq* polymerase (Promega), and ~50 ng of template DNA. Thermocycling conditions (MJ Research PTC-200) for PCR were as follows: 94°C for 5 min; 27–33 cycles of 94°C for 1 min, 54°C for 1 min, and 72°C for 1 min; and 72°C for 5 min. For *trnH-psbA* and other select cases, the *taq* polymerase was added following the initial 5-min denaturation (a "hot start"), as this greatly increased product yield.

PCR products were purified with a QIAquick PCR Purification Kit (Qiagen, Valencia, CA) or by adding 2.0 μL of ExoSap-IT enzyme (Affymetrix, Santa Clara, CA) to 5.0 μL of PCR product and incubating at 37°C for 15 min followed by 80°C for 15 min. Sequences were obtained using 1/4 or 1/16 BigDye Terminator (ver. 3.1; Life Technologies) cycle sequencing reactions. Thermocycling conditions were 94°C for 1 min and then 25 cycles of 94°C for 10 s, 50°C for 5 s, and 60°C for 30 s. Sequencing products were purified and sequenced on an ABI 3130xl or 3730xl genetic analyzer at the Arizona State University or the Iowa State University DNA sequencing facility. Both forward and reverse sequences were obtained for each region using the amplification primers stated above; an internal primer was additionally used for the *trnL-trnF* region ("d" of Taberlet et al. 1991). Sequence data were edited using SEQUENCHER (ver. 5.1; GeneCodes, Ann Arbor, MI) or GENEIOUS (ver. 6.1; Biomatters, Auckland, New Zealand; available at http://www.geneious.com). Each region was aligned individually using SE-AL (ver. 2.0a11; Rambaut 1996) or GENEIOUS, and then regions were concatenated into a single matrix. All sequences were deposited in GenBank, and accession numbers are provided in appendix A.

Nuclear Microsatellite Data Analyses

Descriptive population genetic statistics were calculated for each population and taxon using GENALEX (ver. 6.501; Peakall and Smouse 2006, 2012). Statistics included percentage of polymorphic loci, total number of alleles, mean number of alleles per locus, total number of private alleles, mean number of private alleles per locus, gene diversity (calculated as Nei's [1978] unbiased expected heterozygosity), fixation index, deviation from Hardy-Weinberg equilibrium, genetic differentiation (Φ_{PT}; Peakall and Smouse 2006; significance based on 1000 permutations), number of migrants, and genetic distance

(Peakall et al. 1995). The apportionment of genetic variation within and among populations and taxa was calculated using an analysis of molecular variance (AMOVA) based on genetic distances. The relationship between genetic distances and the natural log of geographic distances between populations was evaluated using a Mantel test. Relationships among individuals and populations were visualized using principal coordinate analysis (PCoA) of the covariance matrix with data standardization of genetic distances. Neighbor-joining trees were constructed on the basis of Cavalli-Sforza and Edwards's (1967) chord distances (calculated from allele frequency data) in POWERMARKER (ver. 3.25; Liu and Muse 2005). Nodal support for neighbor-joining trees was calculated with 1000 bootstrap replicates generated in POWERMARKER using the CONSENSE program in PHYLIP (ver. 3.67; distributed by J. Felsenstein, Department of Genome Sciences, University of Washington, Seattle, WA; http://evolution.gs.washington.edu/phylip.html).

Bayesian clustering analysis was performed in the program STRUCTURE (ver. 2.3; Pritchard et al. 2000; Falush et al. 2007), first to test for distinctiveness of taxa and then to evaluate the ancestry of *P. pilosa* subsp. *deamii* individuals. To determine the appropriate number of clusters given the data, all individuals were included in the analysis without a priori population designations. The likelihood of K, where K is the number of distinct genetic clusters, was calculated for $K = 1$ to 18 (18 being the total number of populations sampled) using a model of admixture and correlated allele frequencies. Each value of K was evaluated with 20 independent runs of 500,000 iterations preceded by a burn-in of 100,000 iterations. To determine the most likely value of K, log probabilities and the change in log probabilities (Pritchard et al. 2000; Evanno et al. 2005) were examined using the program STRUCTURE HARVESTER (Earl and vonHoldt 2012). Clusters were aligned and averaged using the program CLUMPP (ver. 1.1) with the Greedy algorithm and 1000 permutations of randomized input order (Jakobsson and Rosenberg 2007). Resulting assignments were visualized using the program DISTRUCT (ver. 1.1; Rosenberg 2004).

After all taxa were found to be distinct in the initial STRUCTURE analysis, the proportion of ancestry for *P. pilosa* subsp. *deamii* individuals was evaluated by assuming $K = 2$, providing a priori information about the population origin of all individuals, activating the USEPOPINFO feature for *P. pilosa* subsp. *pilosa* and *P. amoena*, and treating *P. pilosa* subsp. *deamii* as having unknown ancestry (following the methods of James and Abbott 2005). Proportion of ancestry was estimated with 20 independent runs of 500,000 iterations preceded by a burn-in of 100,000 iterations using a model of admixture and correlated allele frequencies. Clusters were aligned, averaged, and visualized using CLUMPP and DISTRUCT, as described above.

Chloroplast DNA Sequence Data Analyses

Phylogenetic relationships were reconstructed on the basis of the concatenated chloroplast sequence matrix using maximum likelihood (ML) and Bayesian inference (BI). *Phlox stolonifera* Sims was selected as an appropriate outgroup on the basis of broader phylogenetic studies of *Phlox* (see Ferguson and Jansen 2002). jModelTest (ver. 2.0; Guindon and Gascuel 2003; Darriba et al. 2012) was used to calculate likelihood scores of different substitution models for the sequence matrix, and models were ranked according to the Akaike Information Criterion (AIC; Akaike 1974). The model TPM1uf selected by AIC was used for ML analysis; however, because that model was not available in the analysis software, the more parameterized GTR model was used for BI analysis. ML analysis was conducted using PHYML (Guindon et al. 2010), and support was estimated with 1000 bootstrap replicates. Bayesian inference was conducted using MRBAYES (Huelsenbeck and Ronquist 2001; Ronquist and Huelsenbeck 2003). The Markov chain Monte Carlo algorithm was run for 1,100,000 generations with one cold and three heated chains, starting from random trees and sampling one out of every 200 generations. The first 10% of trees were discarded as burn-in. The remaining trees were used to construct 50% majority-rule consensus trees. For ML and BI analyses, indels (including gaps resulting from mononucleotide length variation) were treated as missing. PHYML and MRBAYES analyses were implemented in GENEIOUS. Additional analyses performed with indels recoded and/or under a model of parsimony resulted in congruent relationships among populations and taxa to those described in the results from ML and BI analyses (results not shown).

Results

Nuclear Microsatellite Data

A total of 155 alleles were found across all loci and populations, with 94 found in *Phlox pilosa* subsp. *deamii*, 100 in *P. pilosa* subsp. *pilosa*, and 66 in *Phlox amoena* (table 1). Of the 155 alleles, 23 were shared by all three taxa. *Phlox pilosa* subsp. *deamii* shared some alleles with each putative parent exclusively; 33 alleles were shared only with *P. pilosa* subsp. *pilosa*, and 12 were shared only with *P. amoena*. Taxa were genetically distinct from each other, with 26 alleles unique (private) to *P. pilosa* subsp. *deamii*, 30 unique to *P. pilosa* subsp. *pilosa*, and 17 unique to *P. amoena*. All three taxa exhibited polymorphism across all loci and populations (table 1), with the exception of one population of *P. pilosa* subsp. *pilosa*, which was missing data at one locus for five of the eight samples. On average, diversity measures were similar for all three taxa, although *P. pilosa* subsp. *deamii* showed the highest values for number of alleles per locus and unbiased expected heterozygosity, and *P. pilosa* subsp. *pilosa* showed the lowest values (table 1). Few private alleles were found within populations, and *P. amoena* populations typically had fewer private alleles than did *P. pilosa* subsp. *deamii* and *P. pilosa* subsp. *pilosa* populations (table 1). Mean fixation indices were generally small or close to zero across all populations and taxa, and none of the loci deviated significantly from Hardy-Weinberg equilibrium (after Bonferroni correction; data not shown).

Populations were moderately to well differentiated, with average Φ_{PT} equal to 0.277 across all populations, 0.165 across *P. pilosa* subsp. *deamii* populations, 0.435 across *P. pilosa* subsp. *pilosa* populations, and 0.324 across *P. amoena* populations (individual pairwise values are shown in app. B, which

Table 1

Genetic Diversity Statistics Based on Microsatellite Variation

Population name	N_{ind}	% polymorphic loci	Total N_a	N_a/locus Range	N_a/locus Mean	Total P_a	P_a/locus Range	P_a/locus Mean	H_e	F_{IS}
Phlox pilosa subsp. *deamii*:										
D1	6	100	28	4–7	5.6	4	0–2	.8	.785	.319
D2	7	100	28	5–6	5.6	2	0–1	.4	.827	.069
D3	8	100	33	4–10	6.6	3	0–2	.6	.845	.235
D4	7	100	26	3–7	5.2	4	0–4	.8	.749	.097
D5	6	100	20	2–6	4	2	0–1	.4	.716	.229
D6	8	100	35	5–9	7	4	0–2	.8	.854	.145
Mean	7	100	28.3	2–10	5.7	3.2	0–4	.6	.796	.182
All	42	100	94	10–26	18.6	26	0–12	5.2	.901	.336
P. pilosa subsp. *pilosa*:										
P1	8	100	23	4–6	4.6	3	0–1	.6	.700	.242
P2	8	80	19	1–7	3.8	1	0–1	.2	.530	−.074
P3	8	100	25	3–9	5	2	0–2	.4	.704	.188
P4	8	100	25	2–9	5	3	0–1	.6	.671	.221
P5	8	100	21	2–6	4.2	4	0–4	.8	.649	.275
P6	7	100	29	4–8	5.8	8	0–4	1.6	.828	.314
P7	7	100	21	3–5	4.2	2	0–2	.4	.812	.146
Mean	7.7	97	23.3	1–9	4.7	3.3	0–4	.7	.699	.187
All	54	100	100	11–30	20	30	2–13	6	.881	.412
Phlox amoena:										
A1	8	100	30	3–9	6	4	0–4	.8	.750	−.107
A2	8	100	23	2–9	4.6	3	0–1	.6	.716	.117
A3	8	100	23	2–8	4.6	1	0–1	.2	.652	−.072
A4	8	100	28	4–10	5.6	0	...	.0	.735	.109
A5	8	100	27	2–7	5.4	2	0–1	.4	.761	.277
Mean	8	100	26.2	2–10	5.2	2	0–4	.4	.723	.065
All	40	100	66	7–21	13.2	18	1–10	3.6	.815	.189

Note. N_{ind} = number of individuals sampled for analysis, N_a = number of alleles, P_a = number of private alleles, H_e = unbiased expected heterozygosity, F_{IS} = fixation index as calculated in GENALEX.

is available online). Number of migrants (N_m) calculated from Φ_{PT} values generally indicated very low gene flow among populations, with average N_m equal to 0.854 across all populations, 1.879 across *P. pilosa* subsp. *deamii* populations, 0.677 across *P. pilosa* subsp. *pilosa* populations, and 1.409 across *P. amoena* populations (individual values are shown in app. B). Almost all N_m values calculated between populations of different taxa were less than 1.00, with the exception of two *P. pilosa* subsp. *pilosa* populations (populations P6 and P7) compared with *P. pilosa* subsp. *deamii* (app. B). Results from the AMOVA indicated that a considerable amount of the observed variation was due to differences among populations (21%) and among taxa (9%), although most of the observed variation was due to differences within populations (70%; $P = 0.001$). There was a significant correlation between genetic and geographic distances when all populations were tested together ($r_{xy} = 0.354$, $R^2 = 0.126$, $P = 0.001$) and when populations of each taxon were tested separately (*P. pilosa* subsp. *deamii*, $r_{xy} = 0.346$, $R^2 = 0.120$; *P. pilosa* subsp. *pilosa*, $r_{xy} = 0.611$, $R^2 = 0.373$; *P. amoena*, $r_{xy} = 0.340$, $R^2 = 0.116$; all values significant at $P = 0.001$).

Results were similar for all three types of analyses examining overall genetic structure. First, PCoA based on genetic distances among populations revealed that each taxon was generally distinct, with the exception of a single *P. pilosa* subsp. *pilosa* population (P6) that grouped with *P. pilosa* subsp. *deamii* (fig. 2*A*). *Phlox pilosa* subsp. *deamii* formed a cluster intermediate to the clusters of *P. pilosa* subsp. *pilosa* and *P. amoena* (fig. 2*A*). PCoA based on genetic distances among individuals also showed general distinction among taxa and the intermediacy of *P. pilosa* subsp. *deamii* (fig. 2*B*). However, some individuals of *P. pilosa* subsp. *pilosa* and *P. pilosa* subsp. *deamii* were intermingled (fig. 2*B*), especially those from *P. pilosa* subsp. *pilosa* populations P6 and P7.

Second, similar to the PCoA results, three generally distinct groups corresponding to each of the three taxa were evident from the neighbor-joining analysis of Cavalli-Sforza and Edward's chord distances, with moderate support for some groups (*P. amoena* with 70% bootstrap support, *P. pilosa* subsp. *deamii* with 58% bootstrap support; fig. 3). However, the placement of *P. pilosa* subsp. *deamii* population D5 was uncertain, and *P. pilosa* subsp. *pilosa* populations P6 and P7 grouped with *P. pilosa* subsp. *deamii* (fig. 3).

Finally, examination of the log probabilities and the change in log probabilities from Bayesian clustering analysis revealed $K = 3$ as the optimal number of clusters (fig. 4*A*). In all $K = 3$ models from 20 independent runs, each cluster corresponded to one of the three taxa, with the exception of individuals from *P. pilosa* subsp. *pilosa* populations P6 and P7, which were assigned to the *P. pilosa* subsp. *deamii* cluster. Some admixture was also evident in *P. pilosa* subsp. *deamii* population D5. When the data were analyzed with $K = 2$ and a priori pop-

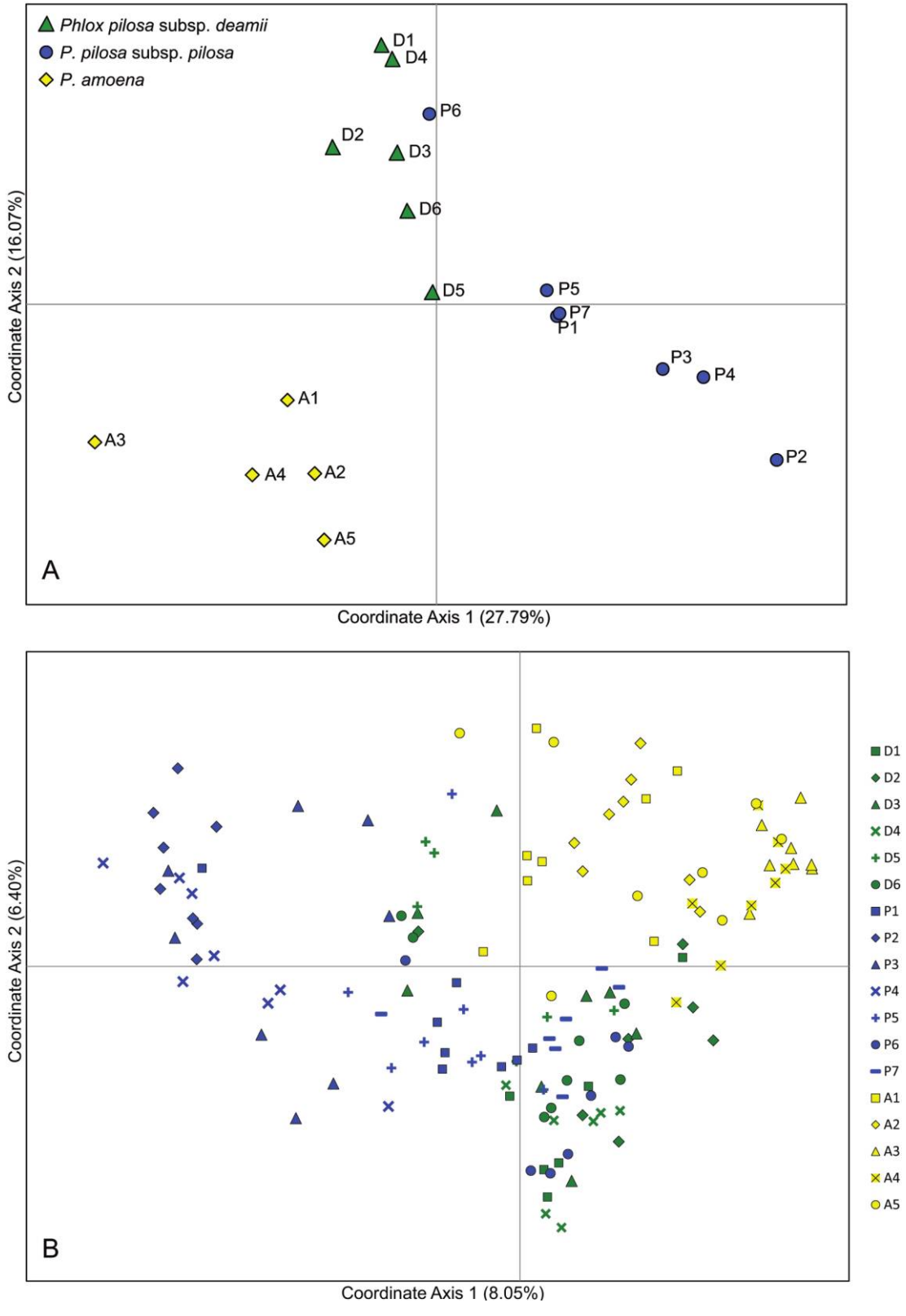

Fig. 2 Plot of first and second principle coordinate scores resulting from an analysis of pairwise genetic distances between populations (*A*) and individuals (*B*) of *Phlox pilosa* subsp. *deamii*, *P. pilosa* subsp. *pilosa*, and *Phlox amoena*. Axes are labeled with the percentage of variance explained.

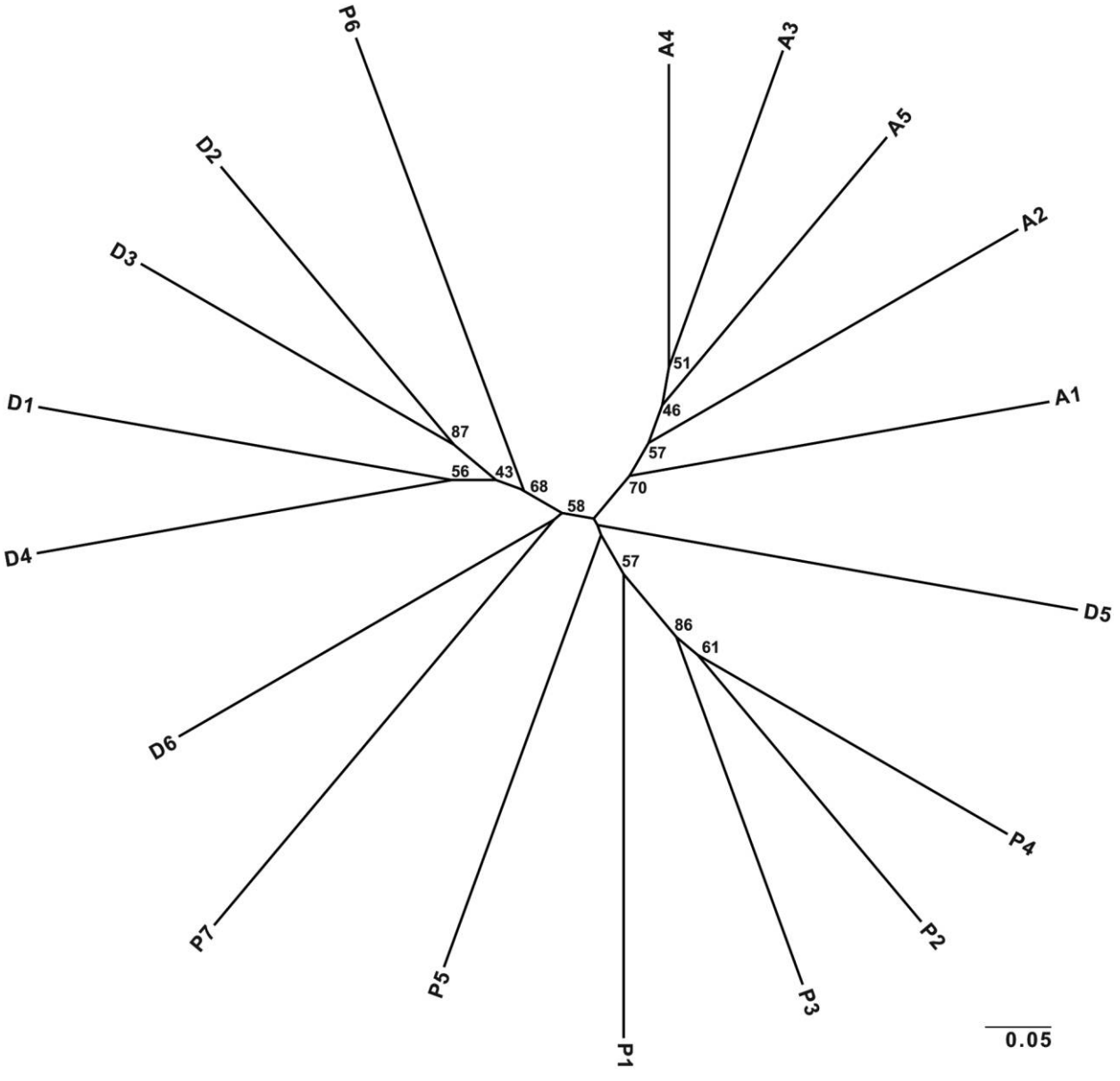

Fig. 3 Neighbor-joining dendrogram showing genetic relationships among populations of *Phlox pilosa* subsp. *deamii*, *P. pilosa* subsp. *pilosa*, and *Phlox amoena* based on microsatellite data. The dendrogram was constructed from Cavalli-Sforza and Edwards's (1967) chord distances calculated from allele frequency data. Nodal support shown above branches is based on 1000 bootstrap replicates.

ulation designations for *P. pilosa* subsp. *pilosa* and *P. amoena*, results indicated that *P. pilosa* subsp. *deamii* individuals exhibited on average 65% membership in the *P. pilosa* subsp. *pilosa* cluster (range: 18%–89%) and 35% membership in the *P. amoena* cluster (range: 11%–82%; fig. 4*B*). The unexpected placement of individuals from *P. pilosa* subsp. *pilosa* population P7 in the *P. amoena* cluster in this analysis may be indicative of introgression or, more likely, the general dissimilarity of population P7 from all other *P. pilosa* subsp. *pilosa* populations (fig. 4*B*).

Chloroplast DNA Sequence Data

Three chloroplast regions were sequenced for six individuals of *P. pilosa* subsp. *deamii*, seven individuals of *P. pilosa* subsp. *pilosa*, and five individuals of *P. amoena*, each from a different study population (app. A). The combined length of the three regions after alignment was 2584 bp (outgroup included) or 2581 bp (outgroup excluded). Variable sites among the ingroup sequences included 18 substitutions and 11 indel regions (including indels 1–34 bp in length and several areas of mono-

nucleotide length variation). When all variable sites were considered, 16 ingroup haplotypes were identified.

Phylogenetic analyses revealed low resolution for relationships among populations and taxa, and ML and BI recovered congruent topologies and similar levels of branch support (fig. 5). Samples of taxa were not resolved as monophyletic in the cpDNA tree. Relationships among most individuals of *P. pilosa* subsp. *pilosa* and *P. amoena* were unresolved, while all individuals of *P. pilosa* subsp. *deamii* grouped together in a moderately supported derived position along with three individuals of *P. pilosa* subsp. *pilosa* and one individual of *P. amoena* (66% ML bootstrap, 0.9972 BI posterior probability). Only one haplotype was shared among individuals of the putative parental taxa *P. pilosa* subsp. *pilosa* and *P. amoena*, and it occurred in a geographic area where the taxa co-occur (including samples from Decatur County, TN: P3, P4, and A2; fig. 1). All of the ingroup haplotypes detected in this study are placed within a large clade in the broader cpDNA phylogeny for *Phlox* that includes previously sampled *P. pilosa* subsp. *deamii*, the majority of samples of *P. pilosa* subsp. *pilosa*, and *P. amoena*, as well as numerous other eastern North American taxa (Ferguson et al. 2008; corresponding to clade "d" of Ferguson and Jansen [2002]; data not shown).

Discussion

Results presented here are consistent with the hypothesis of a homoploid hybrid origin for *Phlox pilosa* subsp. *deamii*, as previously proposed on the basis of ecological, morphological, crossing, and protein electrophoretic data (Levin 1966; Levin and Smith 1966; Levin and Schaal 1970*b*). A number of analyses reveal the mosaic nature of the *P. pilosa* subsp. *deamii* genome relative to its putative parental taxa, including shared alleles between *P. pilosa* subsp. *deamii* and each of its parents to the exclusion of the other parent, intermediacy of overall *P. pilosa* subsp. *deamii* genetic diversity relative to its parents in PCoA and neighbor-joining analyses, ancestry from both parents indicated by STRUCTURE analysis, and chloroplast haplotypes shared with both parents. If *P. pilosa* subsp. *deamii* has arisen via homoploid hybridization between *P. pilosa* subsp. *pilosa* and *Phlox amoena*, it is likely that ecology has played a role in its divergence. Gross and Rieseberg (2005) found that most, if not all, homoploid hybrid taxa show evidence of ecological divergence from parental taxa. *Phlox pilosa* subsp. *deamii* and *P. amoena* are not known to overlap in distribution. *Phlox pilosa* subsp. *deamii* and *P. pilosa* subsp. *pilosa* occur—or historically occurred—in close geographical proximity in at least two areas. Levin (1966) noted range overlap of the taxa in northwestern Tennessee but found the taxa to be consistently spatially isolated, with populations of the hybrid derivative occurring locally in more xeric habitats (see also Levin and Smith 1966). Both of these taxa are also known from Perry County, Indiana, although ecological affinities in that area have not been explored. Furthermore, life history characteristics of these taxa could have augmented the effects of spatial isolation. Gene flow in *Phlox* is restricted by predominantly local pollen movement by insect pollinators and seed dispersal by capsule dehiscence in the local vicinity of the maternal plant (Levin and Smith 1966; see Grant 1959; Levin and Kerster 1968). In line with these characteristics, genetic results reveal low levels of gene flow, low numbers of migrants, and genetic variation that is highly structured geographically (strong isolation by distance). This system warrants further studies addressing the fine-scale distribution patterns and specific ecological and environmental habitat characteristics of these taxa.

The most striking finding from our analyses is the general genetic distinctiveness of *P. pilosa* subsp. *deamii*. This taxon is distinguished overall from *P. pilosa* subsp. *pilosa* and *P. amoena* by all analyses of population genetic variation. Bayesian and neighbor-joining clustering analyses assign these taxa to three different groups, and each taxon is separated in PCoA analysis (although, in various analyses, one or two populations of *P. pilosa* subsp. *pilosa* show affinity with *P. pilosa* subsp.

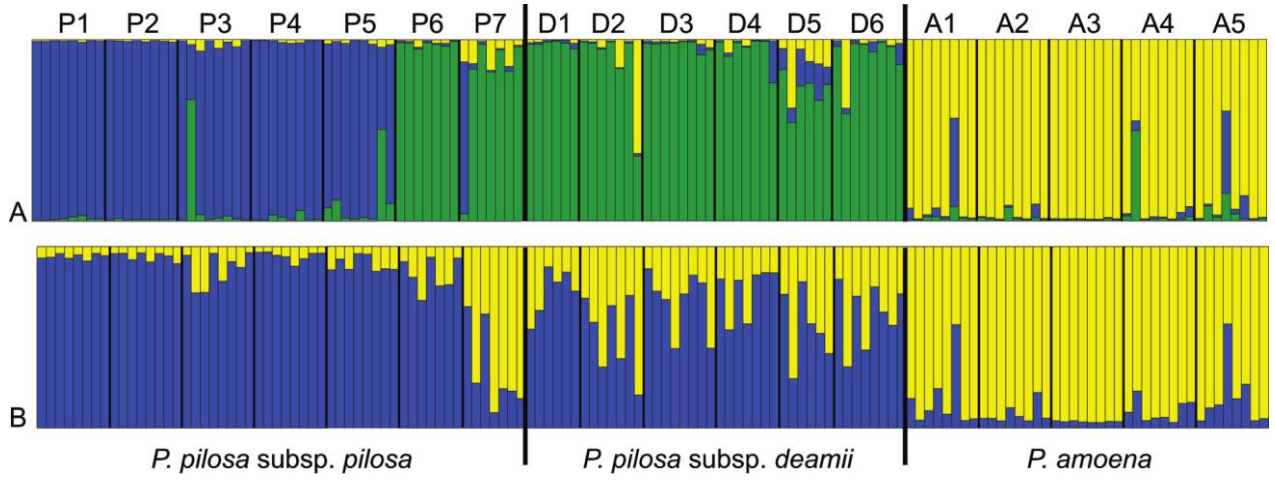

Fig. 4 STRUCTURE analysis of *Phlox pilosa* subsp. *deamii*, *P. pilosa* subsp. *pilosa*, and *Phlox amoena* based on microsatellite data. *A*, $K = 3$, the optimal number of clusters computed without a priori population information. *B*, Estimated proportion of ancestry in *P. pilosa* subsp. *deamii* individuals assuming $K = 2$ and providing a priori population information for *P. pilosa* subsp. *pilosa* and *P. amoena*.

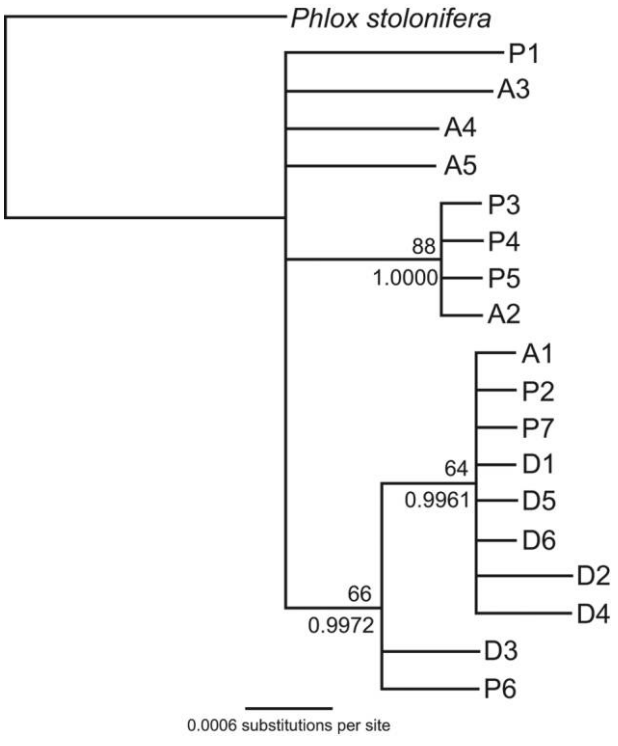

Fig. 5 Phylogeny for study populations of *Phlox pilosa* subsp. *deamii*, *P. pilosa* subsp. *pilosa*, and *Phlox amoena* based on maximum likelihood and Bayesian inference analyses of chloroplast DNA sequence data (the tree was generated using GENEIOUS, which adds a terminal branch to each operational taxonomic unit). Numbers above and below branches represent bootstrap values >50% and posterior probability values, respectively.

deamii). Furthermore, the percentage of genetic variation attributable to differences among taxa is slightly higher than that found for two species of western *Phlox* for which population genetic data are available, *P. amabilis* Brand and *P. woodhousei* (Torr. ex A. Gray) E. E. Nelson (9% here compared with 7% in the study of western species; Fehlberg and Ferguson 2012). Genetic differentiation among taxa in the present study is similar to that found in the western *Phlox* species ($\Phi_{PT} = 0.277$ here compared with $\Phi_{PT} = 0.305$; Fehlberg and Ferguson 2012) and to perennials and outcrossing species more generally ($\Phi_{PT} = 0.25$ and 0.27, respectively; Nybom 2004). The amount of genetic differentiation expected between a hybrid taxon and its parental taxa is an intriguing question: in a brief survey of available data, Gross (2012) found that genetic distance between a hybrid species and its parental taxa could equal or exceed the genetic distance found between the parental taxa. Here, the genetic differentiation and distance (Φ_{PT} and D, respectively) found between *P. pilosa* subsp. *deamii* and *P. pilosa* subsp. *pilosa* ($\Phi_{PT} = 0.272$, $D = 10.74$) or *P. pilosa* subsp. *deamii* and *P. amoena* ($\Phi_{PT} = 0.274$, $D = 10.25$) is similar to the genetic differentiation and distance found between *P. pilosa* subsp. *pilosa* and *P. amoena* ($\Phi_{PT} = 0.341$, $D = 12.07$). Certainly, knowledge of the extent of genetic divergence of hybrid taxa from their putative parental taxa combined with knowledge of their phenotypic divergence will lead to improved insights regarding the role of hybridization in speciation and the generation of biodiversity (Gross 2012), and the present study adds to our broader understanding.

Populations of *P. pilosa* subsp. *deamii* maintain high levels of genetic diversity, as indicated by heterozygosity and the number of polymorphic loci, alleles, and private alleles. These high levels of diversity are important from a conservation standpoint (Booy et al. 2000; Amos and Balmford 2001), but they are also significant from an evolutionary standpoint in that *P. pilosa* subsp. *deamii* does not have reduced genetic diversity relative to its putative parental taxa. This lack of reduced genetic diversity could be attributed to several factors, such as the time since formation of the hybrid taxon (with older taxa harboring more unique alleles; Clay et al. 2012; Rentsch and Leebens-Mack 2012), processes generating diversity in hybrid taxa (higher mutation rates, recombination of divergent genomes; Golding and Strobeck 1983; Bradley et al. 1993; Hoffman and Brown 1995), selection, multiple origins, and/or continued genetic exchange with parental taxa. Although it is possible that the large number of private alleles observed in *P. pilosa* subsp. *deamii* is due to unsampled alleles in the more broadly ranging parental taxa, it may also be that enough time has passed for genome stabilization and the subsequent generation of genetic diversity in this hybrid taxon. While *P. pilosa* subsp. *deamii* may have once been more widespread, it seems unlikely that it ranged beyond eastern North America, and cpDNA haplotypes do not suggest genetic affinities with *P. pilosa* subsp. *pilosa* from the western part of its range (from which some populations are known to harbor divergent haplotypes similar to those in the Texas annual clade, for instance; e.g., Ferguson and Jansen 2002).

Although the *P. pilosa* subsp. *deamii* genome appears to have genetic contributions from both parental taxa, there is disproportionate representation of the *P. pilosa* subsp. *pilosa* genome, as evidenced by a greater number of nuclear alleles and chloroplast haplotypes shared between *P. pilosa* subsp. *deamii* and *P. pilosa* subsp. *pilosa*, patterns of overlap of populations and individuals in PCoA, and patterns of admixture and ancestry in STRUCTURE. Such a departure from strict additivity could be expected as a result of regular segregation and independent assortment of alleles in subsequent generations of the hybrid taxon following formation, producing an array of genotypes that are more or less like either parental taxon (Soltis and Soltis 2009). In addition, particularly in light of their current distributions, it is possible that *P. pilosa* subsp. *deamii* and *P. pilosa* subsp. *pilosa* currently have and/or have had in the past a greater opportunity for backcrossing and introgression (see the discussion of hybridization below). Furthermore, the initial formation of a hybrid taxon could result from a much more complex scenario than the straightforward hybridization of genetically distinct parental taxa followed by rapid, strong reproductive isolation of the hybrid derivative. Soltis and Soltis (2009) describe one possible complex scenario wherein a hybrid species originates from a hybrid swarm between parental taxa that are heterozygous and share alleles at some loci, with F_1's, F_2's, later generations, and even backcrosses all contributing to the genome of the eventually stabilized hybrid derivative.

Such a scenario is also interesting with regard to taxonomic recognition. When a recognized species is evidenced to be of

hybrid origin, hybrid speciation is invoked. *Phlox pilosa* subsp. *deamii* is recognized taxonomically as a subspecies within *P. pilosa*; the evidenced homoploid hybrid origin of this taxon presents a scenario similar to that of hybrid speciation but with weaker expected reproductive isolation from *P. pilosa* (i.e., greater expected potential for gene flow; see below). In fact, it is conceivable that many species—particularly wide-ranging, variable species that have received limited taxonomic scrutiny—harbor entities that could be recognized as infraspecific taxa (subspecies or varieties) that have been derived through interspecific hybridization. Where infraspecific taxa (with morphogeographical distinctness and increased reproductive isolation) are not recognized, such cases of hybrid origin of taxa would be considered cases of introgression (with gene flow unrestricted by taxonomic boundaries). In other words, the distinction between homoploid hybrid origin of taxa (which we view to be rare) and introgression (which is undoubtedly prevalent in plants) may be less pronounced than is generally thought.

Comparisons between nuclear and chloroplast genetic data do provide some support for ongoing gene exchange in this study system. First, there are two instances of incongruent patterns between nuclear and chloroplast data. Populations of *P. pilosa* subsp. *pilosa* and *P. amoena* from the same site (P3 and A2; fig. 1; app. A) and from sites in close geographic proximity (P2 and A1; fig. 1; app. A) share chloroplast haplotypes but are clearly separated according to taxonomy in analyses of nuclear variation. These could reflect cases of chloroplast capture as a result of hybridization and introgression, although it is also possible that these patterns are the result of shared ancestral variation. Second, populations of *P. pilosa* subsp. *deamii* and *P. pilosa* subsp. *pilosa* from sites in close geographic proximity (D6 and P7; fig. 1; app. A) share both nuclear and chloroplast variation. This observed pattern may well be attributable to hybridization and introgression among populations of both taxa in this particular region: the sample of *P. pilosa* subsp. *pilosa* (P7) was collected at a site in Perry County, Indiana, that is within 10 mi of a historical locality for *P. pilosa* subsp. *deamii* (in fact, the type locality for the hybrid derivative taxon, although extant populations of *P. pilosa* subsp. *deamii* could not be located in Perry County during the course of fieldwork for the present study). Our findings may be indicative of recent gene flow (and perhaps this area historically harbored populations exhibiting the "limited intergradation" between the two subspecies noted by Levin [1966]), although the possibility of shared ancestral genetic variation cannot be ruled out. Some other cases of shared nuclear and chloroplast variation are more difficult to explain. Detailed phylogeographic study of populations, ideally using phylogenetic data from the nuclear genome in addition to the chloroplast genome, as well as expanded population genetic work including additional populations (perhaps also sampling additional eastern *Phlox* taxa) would contribute to a greater understanding of ongoing hybridization among these three taxa.

This population genetic study of *P. pilosa* subsp. *deamii* highlights the genetic uniqueness of the taxon relative to its putative parental taxa, *P. pilosa* subsp. *pilosa* and *P. amoena*, and provides support for its homoploid hybrid origin. The genetic findings lend strong support to the current recognition of *P. pilosa* subsp. *deamii* as a taxonomic entity; furthermore, coupled with previously documented morphological distinctness and geographic integrity of the taxon, our findings could be taken as support for recognition of the hybrid derivative at the species level. A more thorough assessment of this possibility, ideally along with detailed morphometric work and ecological study by regional botanists, is warranted. From a broader perspective, such cases of homoploid hybrid origins—and the taxonomic questions they raise—may be more common than currently recognized, and focused population genetic studies will continue to advance our understanding of the consequences of hybridization for biodiversity.

Acknowledgments

We thank Bethany Wright, Theresa Melhem, and Paul Carman for valuable laboratory technical assistance; Veronica Nixon for assistance in producing the map figure; and Kate Waselkov, Mark Mayfield, the Ferguson lab group, and two anonymous reviewers for comments on the manuscript. Original collection of plant material was accomplished with support from the NSF (Carolyn J. Ferguson's DDIG award, DEB-9623386 to Bob Jansen), and we thank Steve Olson and Mike Homoya for assistance with collecting and/or discussion of collection localities during that early fieldwork, as well as the US Forest Service and TVA's Land between the Lakes for collecting permits. Research was supported by the KSU Center for the Understanding of Origins, the KSU Ecological Genomics Institute, and HHMI (supporting Maureen C. Ty's research through the KSU Undergraduate Research Scholars Program). This is publication 14-182-J of the Kansas Agricultural Experiment Station.

Appendix A

Study Sample Information and GenBank Accession Numbers

The following information is provided for the study samples of *Phlox pilosa* subsp. *deamii*, *P. pilosa* subsp. *pilosa*, and *Phlox amoena*, as well as the outgroup for the cpDNA phylogenetic work, *Phlox stolonifera*: population name; voucher (all vouchers are housed at KSC); general locality; and GenBank accession numbers for the *trnL-trnF* region, *trnH-psbA*, and *trnD-trnT*.

***Phlox pilosa* subsp. *deamii*: D1**; CJF 383; Trigg Co., KY; KJ686534, KJ686552, KJ686571. **D2**; CJF 384; Hopkins Co., KY; KJ686535, KJ686553, KJ686572. **D3**; CJF 385; Christian Co., KY; KJ686536, KJ686554, KJ686573. **D4**; CJF 386; Lyon Co., KY; KJ686537, KJ686555, KJ686574. **D5**; CJF 390; Webster Co., KY; KJ686538, KJ686556, KJ686575. **D6**; CJF 391; Spencer Co., IN; KJ686539, KJ686557, KJ686576. ***Phlox pilosa* subsp. *pilosa*: P1**; CJF 362; Marion Co., AL; KJ686540, KJ686558, KJ686577. **P2**; CJF 371; Wayne Co., TN; KJ686541, KJ686559, KJ686578. **P3**; CJF 372; Decatur Co., TN; KJ686542, KJ686560,

KJ686579. **P4**; CJF 375; Decatur Co., TN; KJ686543, KJ686561, KJ686580. **P5**; CJF 381; Stewart Co., TN; KJ686544, KJ686562, KJ686581. **P6**; CJF 389; Jackson Co., IL; KJ686545, KJ686563, KJ686582. **P7**; CJF 393; Perry Co., IN; KJ686546, KJ686564, KJ686583. ***Phlox amoena***: **A1**; CJF 370; Maury Co., TN; KJ686547, KJ686565, KJ686584. **A2**; CJF 374; Decatur Co., TN; KJ686548, KJ686566, KJ686585. **A3**; CJF 399; Morgan Co., TN; KJ686549, KJ686567, KJ686586. **A4**; CJF 403; Van Buren Co., TN; KJ686550, KJ686568, KJ686587. **A5**; CJF 404; Sequatchie Co., TN; KJ686551, KJ686569, KJ686588. ***Phlox stolonifera***: G. R. and C. Mayfield sn, 26 April 1995; Yancey Co., NC; EF433269, KJ686570, EF433198.

Literature Cited

Abbott RJ, D Albach, S Ansell, JW Arntzen, SJE Baird, N Bierne, J Boughman, et al 2013 Hybridization and speciation. J Evol Biol 26:229–246.

Abbott RJ, MJ Hegarty, SJ Hiscock, AC Brennan 2010 Homopoloid hybrid speciation in action. Taxon 59:1375–1386.

Abbott RJ, LH Rieseberg 2012 Hybrid speciation. *In* eLS. Wiley, Chichester. doi:10.1002/9780470015902.a0001753.pub2.

Akaike H 1974 A new look at the statistical model identification. IEEE Trans Automat Control 19:716–723.

Amos W, A Balmford 2001 When does conservation genetics matter? Heredity 87:257–265.

Arnold ML 1993 *Iris nelsonii* (Iridaceae): origin and genetic composition of a homoploid hybrid species. Am J Bot 80:577–583.

Booy G, RJJ Hendriks, MJM Smulders, JM Van Groenendael, B Vosman 2000 Genetic diversity and the survival of populations. Plant Biol 2:379–395.

Bradley RD, JJ Bull, AD Johnson, DM Hillis 1993 Origin of a novel allele in a mammalian hybrid zone. Proc Natl Acad Sci USA 90: 8939–8941.

Brennan AC, D Barker, SJ Hiscock, RJ Abbott 2012 Molecular genetic and quantitative trait divergence associated with recent homoploid hybrid speciation: a study of *Senecio squalidus*. Heredity 108: 87–95.

Brokaw JM, L Hufford 2010 Origins and introgression of polyploid species in *Mentzelia* section *Trachyphytum* (Loasaceae). Am J Bot 97:1457–1473.

Buerkle CA, RJ Morris, MA Asmussen, LH Rieseberg 2000 The likelihood of homoploid hybrid speciation. Heredity 84:441–451.

Buerkle CA, DE Wolf, LH Rieseberg 2003 The origin and extinction of species through hybridization. Pages 117–141 *in* CA Brigham, MW Schwartz, eds. Population viability in plants. Springer, Berlin.

Cavalli-Sforza LL, AWF Edwards 1967 Phylogenetic analysis: models and estimation procedures. Am J Hum Genet 19:233–257.

Clarkson JJ, LJ Kelly, AR Leitch, S Knapp, MW Chase 2010 Nuclear glutamine synthetase evolution in *Nicotiana*: phylogenetics and the origins of allotetraploid and homoploid (diploid) hybrids. Mol Phylogenet Evol 55:99–112.

Clay DL, SJ Novak, MD Serpe, DC Tank, JF Smith 2012 Homoploid hybrid speciation in a rare endemic *Castilleja* from Idaho (*Castilleja christii*, Orobanchaceae). Am J Bot 99:1976–1990.

Darriba D, GL Taboada, R Doallo, D Posada 2012 jModelTest 2: more models, new heuristics and parallel computing. Nat Meth 9: 772.

Demesure B, N Sodzi, RJ Petit 1995 A set of universal primers for amplification of polymorphic non-coding regions of mitochondrial and chloroplast DNA in plants. Mol Ecol 7:129–131.

Doyle JJ, JL Doyle 1987 A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull 19:11–15.

Earl DA, BM vonHoldt 2012 STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv Genet Resour 4:359–361.

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

Falush D, M Stephens, JK Pritchard 2007 Inference of population structure using multilocus genotype data: dominant markers and null alleles. Mol Ecol Notes 7:574–578.

Fehlberg SD, CJ Ferguson 2012 Intraspecific cytotypic variation and complicated genetic structure in the *Phlox amabilis–P. woodhousei* (Polemoniaceae) complex. Am J Bot 99:865–874.

Fehlberg SD, KA Ford, MC Ungerer, CJ Ferguson 2008 Development, characterization and transferability of microsatellite markers for the plant genus *Phlox* (Polemoniaceae). Mol Ecol Resour 8:116–118.

Ferguson CJ 1998 Molecular systematics of eastern *Phlox* L. (Polemoniaceae). PhD diss. University of Texas, Austin.

Ferguson CJ, SD Fehlberg, KA Ford, SC Strakosh 2008 Phylogenetic relationships of the genus *Phlox* (Polemoniaceae): comparing and combining data sets from the chloroplast genome. Pages 153–172 *in* AK Sharma, A Sharma, eds. Plant genome: biodiversity and evolution. Science, Enfield.

Ferguson CJ, RK Jansen 2002 A chloroplast DNA restriction site study of eastern North American *Phlox* (Polemoniaceae): implications of congruence and incongruence with the ITS phylogeny. Am J Bot 89:1324–1335.

Ferguson CJ, F Krämer, RK Jansen 1999 Relationships of eastern North American *Phlox* (Polemoniaceae) based on ITS sequence data. Syst Bot 24:616–631.

Flory WS 1934 A cytological study on the genus *Phlox*. Cytologia 6: 1–18.

Golding GB, C Strobeck 1983 Increased number of alleles found in hybrid populations due to intragenic recombination. Evolution 37: 17–29.

Grant V 1959 Natural history of the phlox family: systematic botany. Martinus Nijhoff, The Hague.

Gross BL 2012 Genetic and phenotypic divergence of homoploid hybrid species from parental species. Heredity 108:157–158.

Gross BL, LH Rieseberg 2005 The ecological genetics of homoploid hybrid speciation. J Hered 96:241–252.

Gross BL, AE Schwarzbach, LH Rieseberg 2003 Origin(s) of the diploid hybrid species *Helianthus deserticola* (Asteraceae). Am J Bot 90:1708–1719.

Guindon S, JF Dufayard, V Lefort, M Anisimova, W Hordijk, O Gascuel 2010 New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PHYML 3.0. Syst Biol 59:307–321.

Guindon S, O Gascuel 2003 A simple, fast and accurate method to estimate large phylogenies by maximum-likelihood. Syst Biol 52: 696–704.

Hoffman SMG, WM Brown 1995 The molecular mechanism underlying the "rare allele phenomenon" in a subspecific hybrid zone of the California field mouse, *Peromyscus californicus*. J Mol Evol 41: 1165–1169.

Hopkins R, DA Levin, MD Rausher 2012 Molecular signatures of selection on reproductive character displacement of flower color in *Phlox drummondii*. Evolution 66:469–485.

Huelsenbeck JP, F Ronquist 2001 MRBAYES: Bayesian inference of phylogeny. Bioinformatics 17:754–755.

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

James JK, RJ Abbott 2005 Recent, allopatric, homoploid hybrid speciation: the origin of *Senecio squalidus* (Asteraceae), in the British Isles from a hybrid zone on Mount Etna, Sicily. Evolution 59:2533–2547.

Jiang J, M Zhu, X Ai, L Xiao, G Deng, Z Yi 2013 Molecular evidence for a natural diploid hybrid between *Miscanthus sinensis* (Poaceae) and *M. sacchariflorus*. Plant Syst Evol 299:1367–1377.

Kelly LJ, AR Leitch, JJ Clarkson, RB Hunter, S Knapp, MW Chase 2010 Intragenic recombination events and evidence for hybrid speciation in *Nicotiana* (Solanaceae). Mol Biol Evol 27:781–799.

Kim ST, SE Sultan, MJ Donoghue 2008 Allopolyploid speciation in *Persicaria* (Polygonaceae): insights from a low-copy nuclear region. Proc Natl Acad Sci USA 105:12370–12375.

Levin DA 1966 The *Phlox pilosa* complex: crossing and chromosome relationships. Brittonia 18:142–162.

——— 1967 Hybridization between annual species of *Phlox*: population structure. Am J Bot 54:1122–1130.

——— 1968 The genome constitutions of eastern North American *Phlox* amphiploids. Evolution 22:612–632.

——— 1975 Minority cytotype exclusion in local plant populations. Taxon 24:35–43.

——— 1978 Genetic variation in annual *Phlox*: self-compatible versus self-incompatible species. Evolution 32:245–263.

Levin DA, HW Kerster 1968 Local gene dispersal in *Phlox*. Evolution 22:130–139.

Levin DA, BA Schaal 1970*a* Corolla color as an inhibitor of interspecific hybridization in *Phlox*. Am Nat 104:273–283.

——— 1970*b* Reticulate evolution in *Phlox* as seen through protein electorphoresis. Am J Bot 57:977–987.

Levin DA, DM Smith 1966 Hybridization and evolution in the *Phlox pilosa* complex. Am Nat 100:289–302.

Levy M, DA Levin 1974 Novel flavonoids and reticulate evolution in *Phlox pilosa–P. drummondii* complex. Am J Bot 61:156–167.

——— 1975 The novel flavonoid chemistry and phylogenetic origin of *Phlox floridana*. Evolution 29:487–499.

Liu K, SV Muse 2005 POWERMARKER: an integrated analysis environment for genetic marker analysis. Bioinformatics 21:2128–2129.

Loockerman D, RK Jansen 1996 The use of herbarium material for molecular systematic studies. Pages 205–220 *in* TF Stuessy, S Sohmer, eds. Sampling the green world. Columbia University Press, New York.

Mallet J 2007 Hybrid speciation. Nature 446:279–283.

McCarthy EM, MA Asmussen, WW Anderson 1995 A theoretical assessment of recombinational speciation. Heredity 74:502–509.

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

Nybom H 2004 Comparison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants. Mol Ecol 13: 1143–1155.

Pan J, D Zhang, T Sang 2007 Molecular phylogenetic evidence for the origin of a diploid hybrid of *Paeonia* (Paeoniaceae). Am J Bot 94:400–408.

Paun O, F Forest, MF Fay, MW Chase 2009 Hybrid speciation in angiosperms: parental divergence drives ploidy. New Phytol 182: 507–518.

Peakall R, PE Smouse 2006 GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol Ecol Notes 6:288–295.

——— 2012 GENALEX 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. Bioinformatics 28:2537–2539.

Peakall R, PE Smouse, DR Huff 1995 Evolutionary implications of allozyme and RAPD variation in diploid populations of dioecious buffalograss *Buchloe dactyloides*. Mol Ecol 4:135–147.

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

Rambaut A 1996 SE-AL: sequence alignment editor v.2.0a11, available at http://tree.bio.ed.ac.uk/software/seal/.

Rentsch JD, J Leebens-Mack 2012 Homoploid hybrid origin of *Yucca gloriosa*: intersectional hybrid speciation in *Yucca* (Agavoideae, Asparagaceae). Ecol Evol 2:2213–2222.

Rieseberg LH 2006 Hybrid speciation in wild sunflowers. Ann Mo Bot Gard 93:34–48.

Ronquist F, JP Huelsenbeck 2003 MRBAYES 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572–1574.

Rosenberg NA 2004 DISTRUCT: a program for the graphical display of population structure. Mol Ecol Notes 4:137–138.

Shaw J, EB Lickey, JT Beck, SB Farmer, W Liu, J Miller, KC Siripun, CT Winder, EE Schilling, RL Small 2005 The tortoise and the hare II: relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. Am J Bot 92:142–166.

Sherman NA, JM Burke 2009 Population genetic analysis reveals a homoploid hybrid origin of *Stephanomeria diegensis* (Asteraceae). Mol Ecol 18:4049–4060.

Smith DM, DA Levin 1967 Karyotypes of eastern North American *Phlox*. Am J Bot 54:324–334.

Soltis PS 2013 Hybridization, speciation and novelty. J Evol Biol 26: 291–293.

Soltis PS, DE Soltis 2009 The role of hybridization in plant speciation. Annu Rev Plant Biol 60:561–588.

Taberlet P, L Gielly, G Pautou, J Bouvet 1991 Universal primers for amplification of three non-coding regions of chloroplast DNA. Plant Mol Biol 17:1105–1109.

Weiss-Schneeweiss H, C Blöch, B Turner, JL Villaseñor, TF Stuessy, GM Schneeweiss 2012 The promiscuous and the chaste: frequent allopolyploid speciation and its genomic consequences in American daisies (*Melampodium* sect. *Melampodium*; Asteraceae). Evolution 66:211–228.

Wherry ET 1955 The genus *Phlox*. Morton Arboretum Monographs 3, Philadelphia.

Worcester L, MH Mayfield, CJ Ferguson 2012 Cytotypic variation in *Phlox pilosa* ssp. *pilosa* (Polemoniaceae) at the edge of its range in the central United States. J Bot Res Inst Tex 6:443–451.

Yuan JH, FY Cheng, S Zhou 2010 Hybrid origin of *Paeonia* × *yananensis* revealed by microsatellite markers, chloroplast gene sequences, and morphological characteristics. Int J Plant Sci 171:409–420.