Phylogenomic patterns of divergence and gene flow detail the evolution of reinforcement and hybrid speciation in *Phlox* wildflowers

Austin G. Garner [1,2*], Benjamin E. Goulet-Scott [1,2], and Robin Hopkins [1,2*]

[1] *Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 021382, USA*

[2] *The Arnold Arboretum, Harvard University, Boston, MA 02131, USA*

[*] Authors to whom correspondence should be addressed: aggarner@g.harvard.edu; rhopkins@fas.harvard.edu

ABSTRACT

It is becoming increasingly clear that the evolutionary histories of many groups of organisms are riddled with reticulation. Some clades of organisms appear to be hotspots of speciation by hybridization and gene flow. For example, there are two cases of reinforcement and up to nine cases of hypothesized hybrid speciation in the eastern standing *Phlox* (*Polemoniaceae*) wildflowers. However, the relationships of these taxa and the hypotheses about their evolution have received limited attention with genomic data. We performed a comprehensive phylogenomic study of the eastern standing *Phlox* to resolve their evolutionary relationships and test long-standing hypotheses about if and how hybridization and gene flow acted as creative forces in their diversification. Our phylogenomic analyses on genome-wide markers provides resolution into the evolutionary relationships of thirty-two eastern standing *Phlox* taxa, including well-supported non-monophyletic relationships across species complexes. Explicit tests for introgression support gene flow occurring alongside the evolution of one of the two cases of reinforcement, and patterns of divergence and gene flow support one of the five hypothesized homoploid hybrid speciation events. Additionally, comparative read mapping allowed assignment of putative subgenome ancestries for four allotetraploids hybrid species. Our findings demonstrate the utility and importance of phylogenomics in confirming hypothesized evolutionary histories of non-model systems and add to the growing evidence that gene flow across species boundaries can play a role in the generation of novel biodiversity.

Keywords: phylogeny; RADseq; speciation; reinforcement; hybridization; gene flow; hybrid species; polyploid

INTRODUCTION

Hybridization and gene flow can play important and diverse roles during the formation and divergence of species (Taylor and Larson 2019). While these processes are often viewed as homogenizing forces that mix genetic variation across species boundaries, they can also generate novel variation. For example, hybridization and gene flow can create stable hybrid species and/or fuel the diversification of adaptive radiations (Abbott et al. 2013; Schumer et al. 2014; Marques et al. 2019). The act of hybridization itself can also create selective pressures favoring increased reproductive trait divergence to reduce mating between species, i.e. reinforcement (Howard 1993; Garner et al. 2018). The effect of hybridization and gene flow on speciation has been investigated across the tree of life, yet we still lack a broader understanding of the frequency and extent to which these forces impact speciation across clades of organisms (but see (Eaton et al. 2015; Pease et al. 2016)**)**. Characterizing the context and consequences of hybridization and gene flow across groups of species is necessary for discerning how these forces influence the generation and maintenance of biodiversity.

Much of the foundational work on hybridization as a creative force in speciation comes from historic research on non-model plant systems (Anderson 1949; Stebbins 1959; Grant 1981). Patterns of morphological, physiological, and ecological trait variation across and within plant species created taxonomic conflict, inspiring biologists to infer evolutionary histories of hybridization and gene flow between species (Goulet et al. 2017). In particular, populations with individuals exhibiting intermediate, recombinant, and novel trait combinations where closely related species come into geographic contact motivated hypotheses about the formation of polyploid and homoploid hybrid species (Gottlieb 1972; Soltis and Soltis 2009; Yakimowski and Rieseberg 2014; Barker et al. 2016).

Additionally, patterns of trait divergence in sympatry, where closely related species coexist, but not in allopatry, motivated hypotheses of reinforcement during plant speciation (Hopkins 2013). Reinforcement occurs when maladaptive hybridization between species generates selection for traits that increase reproductive isolation. In plants, reinforcement can result in divergence in reproductive traits such as flower color, flowering time, and cross-compatibility in sympatry (McNeilly and Antonovics 1968; Levin 1985; Fishman and Wyatt 1999). Although hybridization drives selection for reinforcement, gene flow and recombination can impede adaptive divergence by reinforcement (Felsenstein 1981). Therefore, much debate and theory has focused on if and how reinforcement can evolve if hybridization also causes interspecific gene flow (Servedio and Noor 2003). Yet, only a few empirical studies have evaluated the presence and extent of gene flow alongside the evolution of reinforcement in nature (Kulathinal et al. 2009; Lemmon and Juenger 2017; Roda et al. 2017; Turissini and Matute 2017; Dyer et al. 2018).

Most hypotheses of hybrid speciation and reinforcement are based on patterns of phenotypic variation and/or a few genetic markers. However, this evidence can be misleading (Noor 1999; Goulet-Scott et al. 2021; Poelstra et al. 2021). Therefore, it is important to return to known and hypothesized cases of reticulate evolution with genomic tests for gene flow to confirm these evolutionary processes and further understand the context of their evolution (Goulet et al. 2017). Hybridization and gene flow result in the mixing of genetic variation between species. Therefore, if these processes coincide with the formation of novel diverging lineages, admixture of parental species' genetic variation will be present in the novel diverging lineage's genome.

Traditional phylogenetic methods infer evolution along bifurcating trees. However, modern analyses can use discordance in allele patterns under these bifurcating trees (Green et al. 2010; Durand et al. 2011) or modeling of allele frequencies at tips (Gutenkunst et al. 2009; Pickrell and Pritchard 2012; Excoffier et al. 2013) to identify a history of gene flow between non-sister taxa and/or the formation of new lineages through hybridization (Kong and Kubatko 2021). Under a bifurcating phylogenetic framework, novel lineages evolving by reinforcement with gene flow or hybrid speciation should be observed sister to one of their parental lineages and have a significant signature of gene flow with these methods from their secondary parental lineage. In this way, evolutionary relationships inferred from genomic data alongside phylogenetic tests for gene flow can evaluate and confirm histories of genetic exchange between lineages (Hibbins and Hahn 2021). Contemporary genome-wide sequencing methods, such as restriction site-associated DNA (RAD) sequencing (Andrews et al. 2016) allow application of these methods to radically clarify the evolutionary histories of reticulate non-model systems (Eaton and Ree 2013; Eaton et al. 2015; Bombonato et al. 2020; Léveillé-Bourret et al. 2020; Guo et al. 2021; Suissa et al. 2022).

The eastern standing *Phlox* (*Polemoniaceae*) are an ideal system for exploring the role of hybridization and gene flow in speciation. They form a monophyletic group (Landis et al. 2018) and are distinct from the rest of *Phlox* by their upright growth habit and natural occurrence across mid to eastern North America. The natural ranges of these taxa intertwine and overlap extensively, generating numerous zones of species contact (Fig. 1). Reproductive barriers between these species are highly permissive (Levin 1966a) and a mix of morphological, ecological, and biochemical evidence suggests pervasive gene flow between some taxa

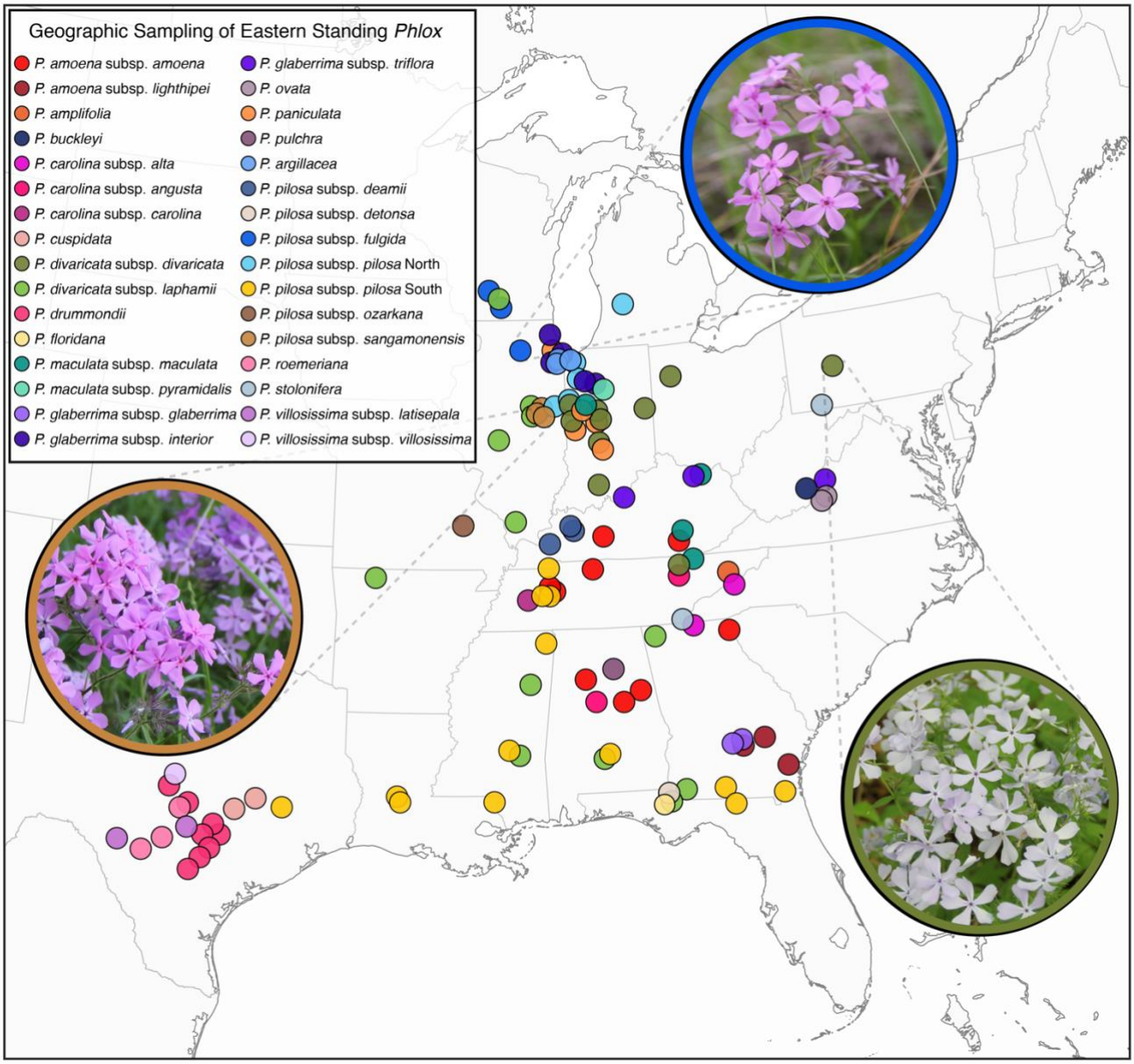


Figure 1. Map of individuals sampled across the interdigitating ranges of the thirty-two eastern standing *Phlox* taxa included for phylogenetic analyses. Each dot represents the locality of a single individual with the color corresponding to an individual's taxonomic identity. Three insets show *P. pilosa* subsp. *sangamonensis* (brown), *P. pilosa* subsp. *fulgida* (blue), and *P. divaricata* subsp. *divaricata* (green) in their natural habitat.

(Anderson and Gage 1952; Erbe and Turner 1962; Levin 1966b; Levin and Smith 1966; Levin and Kerster 1967; Levin and Schaal 1970a; Levin 1985). This clade has served as testing grounds of evolutionary theory on how reproductive isolation, hybridization, and gene flow can influence species divergence, including hybrid speciation and reinforcement. The scientific legacy of the eastern standing *Phlox* makes it a prime system to use modern genomic methods to study how species diverge with hybridization and gene flow.

Phlox has been a taxonomically difficult genus (Locklear 2011a). Traditional delineations of *Phlox* species relied on morphological, ecological, and geographic characteristics, resulting in a long history of describing, promoting, and demoting species, subspecies, and varieties (Whitehouse 1945; Wherry 1955, 1965; Erbe and Turner 1962; Turner 1998; Locklear 2011b, 2011a). Conspicuous variation in pistil style length has been used to divide the eastern standing *Phlox* into long-styled (10-26mm) and short-styled (1.5-4mm) species groups, with the style longer or shorter than the sepals respectively (Wherry 1930, 1931, 1932, 1955; Ferguson et al. 1999). However, other morphological traits vary widely within and between species, and species-level relationships within the group remain either poorly supported or incongruent across phylogenetic studies (Ferguson et al. 1999; Ferguson and Jansen 2002; Landis et al. 2018; Goulet-Scott et al. 2021). Resolving the phylogenetic relationship of the eastern standing *Phlox* will provide a powerful framework for testing evolutionary hypotheses of hybridization and gene flow during their diversification.

In this small clade, there are five hypothesized cases of homoploid hybrid species (*P. argillacea*, *P. maculata* subsp. *pyramidalis*, *P. pilosa* subsp. *deamii*, *P. pilosa* subsp. *detonsa*, and *P. amoena* subsp. *lighthipei*) and at least four cases of allotetraploid hybrid species (*P. villosissima* subsp. *villosissima*, *P. villosissima* subsp. *latisepala*, *P. floridana*, *P. buckleyi*)

(Table 1). Support for these hypotheses were originally based on the observation that putative hybrid lineages had reproductive, morphological, and ecophysiological trait values that appear to be intermediate, recombined, or transgressive with putative parental taxa, making them potentially reproductively and ecologically distinct from their parental lineages (Levin 1963, 1969; Levin and Smith 1966; Hadley and Levin 1969)**.** Further analysis of karyology (Smith and Levin 1967; Levin 1968), cross compatibility (Levin 1966a), biochemical profiles (Levin and Schaal 1970a, 1972; Levy and Levin 1974, 1975), and microsatellite markers (Fehlberg et al. 2014) built support for these hypotheses. Finally, natural hybrid zones and synthetic crosses between pairs of putative parental taxa generated hybrids with phenotypes similar to the putative stable hybrid lineages in nature (Levin 1963, 1966b; Levin and Smith 1966). However, other treatments of these putative hybrids species have assumed them to be diverging varieties without hybridization defining their formation (Wherry 1956) or found no support for hybrid ancestry with genome-wide markers (Goulet-Scott et al. 2021).

In this same clade of *Phlox,* there are also two cases of reinforcement. During the speciation of both *P. drummondii* and *P. pilosa* subsp. *pilosa*, flower color evolution within sympatric populations decreases hybridization with closely related species (reviewed in (Hopkins 2013), Fig. 3). *P. drummondii* has evolved from being light-blue colored to dark-red to prevent hybridization with sympatric light-blue colored *P. cuspidata* (Levin 1985; Hopkins and Rausher 2012), and *P. pilosa* subsp. *pilosa* flowers have evolved from pink colored to white to prevent hybridization where it cooccurs in high frequency with pink colored *P. glaberrima* subsp. *interior* near Lake Michigan (Levin and Kerster 1967). Biochemical evidence suggests gene flow occurs between sympatric *P. pilosa* subsp. *pilosa* and *P. glaberrima* subsp. *interior* (Levin and Schaal 1972), and genomic analyses have inferred a history of gene flow between sympatric *P.*

Table 1. Hybrid speciation hypotheses in the eastern standing *Phlox*

Hypothesized Hybrid Species	Ploidy	Putative Parent Species 1	Putative Parent Species 2	Motivating Previous Evidence
P. argillacea	2N	*P. glaberrima* subsp. *interior*	*P. pilosa* subsp. *pilosa* North	morphology[A], physiology[A], seed protein polymorphism[B], experimental crosses[C]
P. maculata subsp. *pyramidalis*	2N	*P. glaberrima* subsp. *interior*	*P. maculata* subsp. *maculata*	morphology[D], physiology[E], seed protein polymorphism[F], experimental crosses[C]
P. amoena subsp. *lighthipei*	2N	*P. pilosa* subsp. *pilosa* South	*P. amoena* subsp. *amoena*	morphology[G], seed protein polymorphism[F], experimental crosses[CG]
P. pilosa subsp. *deamii*	2N	*P. pilosa* subsp. *pilosa* South	*P. amoena* subsp. *amoena*	morphology[G], experimental crosses[CG], microsatellites[H]
P. pilosa subsp. *detonsa*	2N	*P. pilosa* subsp. *pilosa* South	*P. carolina*	seed protein polymorphism[F], experimental crosses[C]
P. floridana	4N	*P. pilosa* subsp. *pilosa* South	*P. carolina* subsp. *angusta*	karyology[U], flavonoids[K], seed protein polymorphism[F], experimental crosses[C]
P. villosissima subsp. *villosissima*	4N	*P. pilosa* subsp. *pilosa* South	*P. drummondii*	karyology[J], flavonoids[L], seed protein polymorphism[F], experimental crosses[C]
P. villosissima subsp. *latisepala*	4N	*P. pilosa* subsp. *pilosa* South	*P. drummondii*	karyology[U], flavonoids[L], seed protein polymorphism[F], experimental crosses[C]
P. buckleyi	4N	unknown	unknown	karyology[J]

Supporting Literature: A (Levin 1969), B (Levin an Schaal 1972), C (Levin 1966b), D (Levin 1963), E (Hadley Levin 1969), F (Levin an Schaal 1970), G (Levin and Smith 1966), H (Fehlberg et al. 2014), I (Smith and Levin 1967), J (Levin 1968), K (Levy and Levin 1975), L (Levy and Levin 1974).

155 *drummondii* and *P. cuspidata* but analyses were limited to single individuals (Roda et al. 2017).
156 Reexamining these hypotheses of reinforcement with gene flow using geographically widespread
157 sampling and extensive genomic sequencing will provide a clearer understanding of the
158 evolution of reproductive trait divergence in sympatry.

159 In this study, we apply double-digest RAD (ddRAD) sequencing with phylogenomic
160 approaches to resolve the evolutionary relationships of the eastern standing *Phlox*. We then
161 leverage this phylogenetic framework to examine the contribution of hybridization and gene flow
162 in the diversification of this system. Specifically, we test if gene flow has occurred alongside the
163 evolution of two cases of reinforcement, and investigate patterns of genetic composition, and
164 gene flow in multiple hypothesized cases of homoploid and allotetraploid hybrid speciation.
165

166 **MATERIALS AND METHODS**

167 *Sampling, ddRAD Library Sequencing, and Data Processing*

168 We collected wild plants from across the native ranges of the eastern standing *Phlox* in
169 North America during the summers of 2017-2019. We acquired additional samples from natural
170 populations preserved in tissue culture at the Ohio State University Ornamental Plant Germplasm
171 Center (OPGC). Our sampling includes seventeen major recognized eastern standing *Phlox*
172 species, with representation from thirty-two described subspecies, and three outgroup taxa from
173 the mat-forming *Phlox* (Fig. 1; Supplementary Table S1). We erred on the side of over-splitting
174 taxa as described in Locklear 2011a, except for the Texas annuals which we only designated at
175 species level.

176 Plants were grown in the greenhouse under controlled conditions until flowering. We
177 extracted genomic DNA from bud tissue using a hybrid Omega Bio-Tek EZNA DNA kit and

CTAB/Chloroform extraction protocol. Genomic libraries were constructed with a double-digest restriction-site-associated digestion (ddRAD) sequencing protocol and paired-end sequenced as described in (Goulet-Scott et al. 2021).

Raw reads were demultiplexed, filtered, and *de novo* assembled as described in (Goulet-Scott et al. 2021). In brief, raw sequence data was demultiplexed and clone filtered using custom python scripts and the 'process radtags' and 'clone filtering' functions in STACKS v.2.4 (Catchen et al. 2013). Demultiplexed reads were further filtered for adapter sequences and trimmed for base quality score using Trimmomatic v.0.39 (Bolger et al. 2014) under default settings. Reads were *de novo* assembled into loci and variant called using iPyRAD v.0.9.50 (Eaton and Overcast 2020), with a consensus locus clustering threshold of 0.93. Samples with minimal informative shared loci were removed, resulting in a final dataset of 116 ingroup and 4 outgroup individuals.

For the 120 retained samples, raw reads were filtered to an average of 3,060,421 reads per individual. Filtered reads were used to construct four datasets with varying levels of missingness, requiring a locus to be shared among a minimum of N = 4, 10, 20, and 30 individuals, named MS4, MS10, MS20, and MS30 respectively. These assembled datasets produced matrices with a total of 165,549 to 5,3871 loci and 1,884,423 to 105,363 SNP sites with a locus missingness rate of 90%, 83%, 72%, and 60% respectively.

Phylogenetic Inferences

We inferred the phylogenetic relationships of our samples using concatenation and coalescent-based methods. First, whole ddRAD loci from the MS10, MS20 and MS30 datasets were individually concatenated into supermatrices. Maximum likelihood (ML) phylogenies were estimated from each supermatrix using IQ-TREE v.1.6.10 (Nguyen et al. 2015) with the GTR +

gamma nucleotide substitution model and 1000 bootstrap replicates using ultrafast bootstrap approximation (Minh et al. 2013; Hoang et al. 2018).

Second, we inferred a population-level species tree using SVDquartets (Chifman and Kubatko 2014), as implemented in PAUP v.4.0a166 (Swofford 2002). This method uses unlinked SNPs to infer relationships among quartets of taxa under a coalescent model and then estimates a species tree using a quartet assembly method. We conducted the analysis on our dataset with least missingness, MS30, with the mat-forming species removed and *P. stolonifera* set as the outgroup. To avoid genetic linkage, we sampled a single SNP-site with the least amount of missing data from each locus, totaling 5,306 variant sites. We used all 7,160,245 possible quartets and 100 non-parametric bootstrap replicates to assess topological support and to generate a majority-rule consensus species tree.

Support values between these two inference methods are derived from two different estimation methods and cannot be compared one to one. Standard bootstrap supports (SBS), as used with our SVDquartets inferred tree, are conservative and underestimate the likelihood of relationships; SBS supports of 80% have a 95% likelihood of being correct, so SBS >= 80 are considered trusted supports. However ultrafast bootstrap supports (UBS), as used for our concatenation inference method, are unbiased when >70, so a 95% support has a 0.95 probability of being correct (Minh et al. 2013).

Tests for Gene Flow Alongside Reinforcement

We tested for evidence of gene flow alongside the evolution of reinforcement in *P. drummondii* and *P. pilosa* subsp. *pilosa*. First, we computed Patterson's D-statistic (Green et al. 2010; Durand et al. 2011) to infer if gene flow occurred between two lineages in a given phylogeny. This test measures the asymmetry in the ratio of two discordant allele patterns,

ABBA and BABA, across the topology of a four-taxon pectinate tree ((P1, P2), P3), O). Under stochastic lineage divergence without gene flow between lineages P3 and P1 or P3 and P2, the proportion of discordant allele patterns from incomplete lineage sorting are unbiased (ABBA=BABA); however, when introgression has occurred between P3 and P1 or P3 and P2, one discordant allele pattern is more prevalent than the other (ABBA>BABA, BABA>ABBA).

We defined the P1, P2, P3, and O lineages as a taxa and a "test" as a unique combination of individuals from these taxa. We subset the complete MS4 data matrix to include the individuals from a set of four taxa that fit the pectinate topology under our ML phylogenetic trees. P1 is the taxon that diverged by reinforcement (*P. drummondii* and *P. pilosa* subsp. *pilosa*), P2 as a taxon out to P1 but not involved in the reinforcement hypothesis, P3 as the taxon sympatric with P1 driving divergence by reinforcement (*P. cuspidata* and *P. glaberrima* subsp. *interior*), and O as an individual from a proximal lineage out to all other lineages. Tests were iterated over all possible combinations of all individuals of the P1 and P3 lineages, up to three randomly sampled individuals of the P2 lineage not observed growing geographically near populations of P1 or P3, and two individuals of the O lineage. The white flowered *P. pilosa* subsp. *pilosa* lineage that evolved by reinforcement diverged from the lineage of northern *P. pilosa* subsp. *pilosa* populations (Fig. 2), therefore, we only used *P. pilosa* subsp. *pilosa* North individuals in these tests. We were also able to consider two different species as the P2 lineage in the hypothesis concerning *P. pilosa* subsp. *pilosa* North. For all possible tests, we implemented the D-statistic test for 1,000 bootstrap iterations with loci resampling with replacement, as described in (Eaton et al. 2015). Tests with |Z-score| >3 are conservatively statistically significant.

We complemented the D-statistic results with TreeMix v1.13 (Pickrell and Pritchard 2012), a model-based method for inferring population splits and migration (gene flow) under a tree framework, as implemented in iPyRAD. We targeted a potential signal of introgression by modeling a subtree of our total phylogenetic sampling. We subset the complete MS4 data matrix to include all individuals from all non-hybrid lineages in the clades containing the P1 and P3 lineages, described above, all intermediate clades, and from a lineage out to this subtree. Individuals were grouped based on species identity with *P. drummondii* and *P. pilosa* subsp. *pilosa* North split based on flower color phenotype. We denote these subtrees as "*P. pilosa* subsp. *pilosa* North" and "*P. drummondii*" respectively. We required SNPs to be shared by at least 20% of individuals from each species and randomly sampled one unlinked SNP from each locus. The *P. drummondii* (4,638 unlinked SNPs) and *P. pilosa* subsp. *pilosa* North (3,383 unlinked SNPs) subtrees were modeled in TreeMix with m = 1-8 possible migration edges. The best supported number of migration edges was qualitatively chosen as the value at which the log-likelihood of the model began to increase only marginally.

Testing Origins of Homoploid Hybrid Species

To evaluate if the five putative homoploid hybrid species in the standing *Phlox* (Table 1) have genetic support we leveraged our phylogenetic inferences and tests for gene flow to evaluate four criteria: 1) the putative hybrid sits sister to or nested within one of its putative parental taxon in our bifurcating phylogenetic reconstructions, 2) the putative hybrid shows a signal of gene flow with its other parental taxon, 3) the signal of gene flow is exclusive to the putative hybrid, and 4) the signal of gene flow goes into the putative hybrid.

For putative hybrids sister to or within one of their putative parent lineages, we tested for the presence of gene flow with the second putative parental lineage using Patterson's D-statistic,

as described above. For each hybrid hypothesis, a four-taxon tree was generated by subsetting the MS4 dataset to all individuals of the putative hybrid species and putative second parent species, three individuals from a species lineage between the putative hybrid and putative parent, and two individuals of a closely related outgroup. The putative hybrid and the putative parent lineages served as either P1 or P3 based on the structure of our ML phylogenies. We also considered multiple species as the P2 lineage. For putative hybrid lineages with significant evidence of gene flow with their parental taxon, we evaluated if the signature of gene flow was localized to the hybrid lineage by testing for any gene flow between the two putative parental taxa. All hybrid hypotheses tested with D-statistics were also modeled with TreeMix v.1.13, as described above. The *P. maculata* subsp. *pyramidalis*, *P. amoena* subsp. *lighthipei*, and *P. pilosa* subsp. *detonsa* modeled subtrees used 4,490, 4,294, and 1,033 unlinked SNPs respectively. We used the *P. pilosa* subsp. *pilosa* North modeled subtree for the *P. argillacea* hypothesis given their identical sampling.

Inferring Origins of Allotetraploid Hybrid Species

We investigated the hypothesized ancestries of the four allotetraploid hybrid species (Table 1) using a consensus and comparative read alignment approach, similar to (Wang et al. 2021)**.** Sequenced representatives of these taxa were previously confirmed to be tetraploid via flow cytometry at the OGPC.

First, we created reference consensus sequences from diploid lineages. Whole ddRAD loci present in at least 60% of diploid taxa from the MS4 dataset were concatenated to generate taxon-level consensus sequences using the iPyRAD window extractor toolkit, retaining the most common allele in the consensus sequence for species with multiple samples. We then mapped the filtered sequence reads for all allotetraploids to the diploid consensus sequences using bwa-

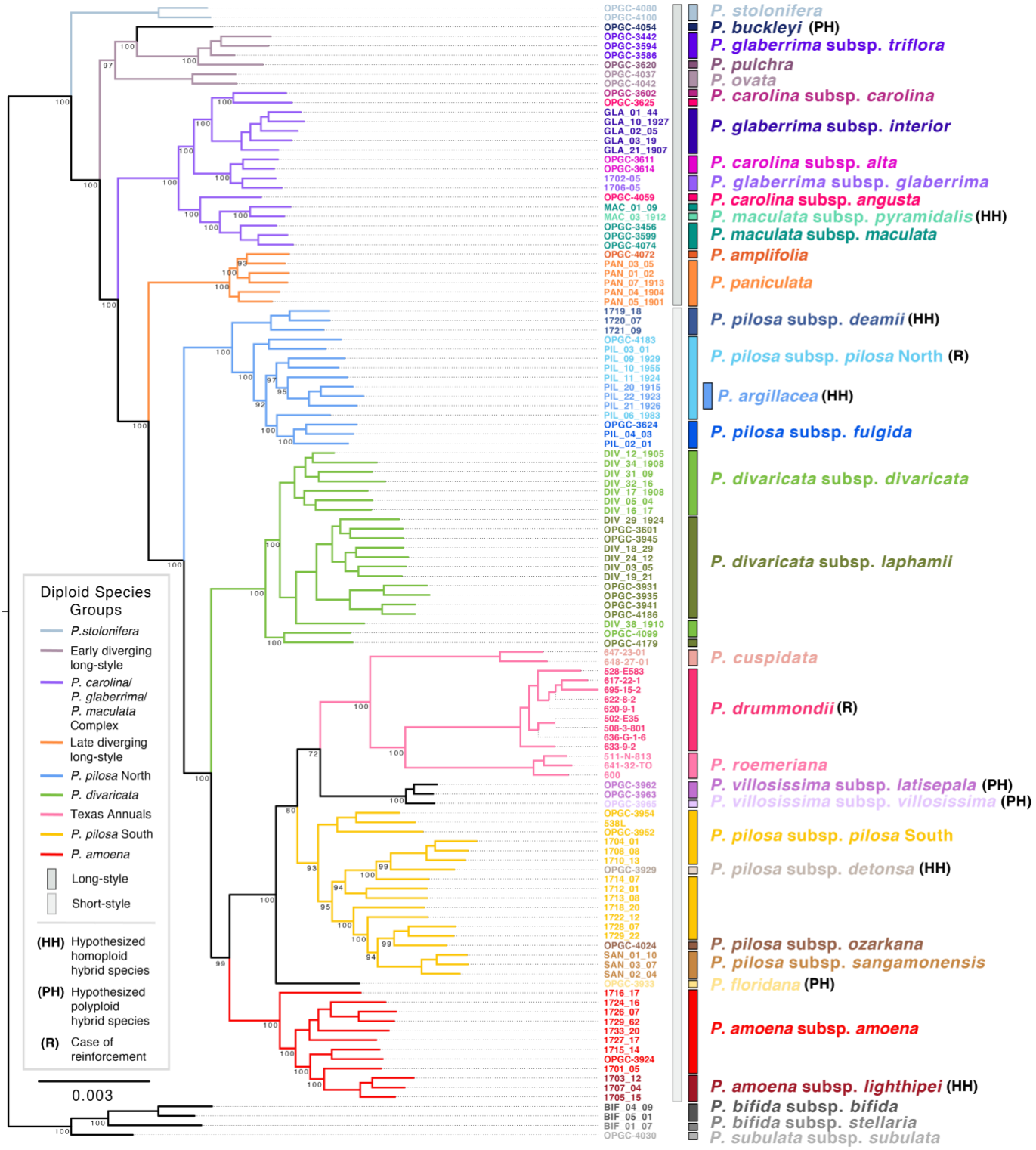
292 mem (Li 2013). We tallied the proportion of primary read alignments with a mapping quality ≥ 5
293 for the forward and reverse reads. We reasoned the diploid consensus receiving the highest
294 proportion of quality read alignments would be most closely related to the ancestor of the
295 mapped reads. We tested this assumption by also mapping one representative from each diploid
296 taxon to the consensus sequences and observing the highest rates of mapping to the same taxon
297 consensus with some carry over to closely related sister taxa. At a minimum, this method
298 allowed identification of the ancestral clade the reads originated from. Under this logic we
299 inferred the putative ancestry of the allopolyploid subgenomes to diploid clades that received the
300 highest rates of read mapping (Wang et al. 2021).
301
302 **RESULTS**
303 *Phylogenomic Inferences of Species Evolutionary Relationships*
304 Phylogenetic analyses of genome-wide nuclear ddRAD loci provided novel resolution of
305 species relationships and partly challenge the historic view of standing *Phlox* evolution. ML trees
306 inferred from each of three concatenated datasets were highly congruent, despite variation in
307 missing data (Fig. 2; Supplementary Fig. S1, S2). Therefore, we focused our results to trends
308 across phylogenetic inferences with discussion of notable differences in interspecific
309 relationships across datasets.
310 The base of the eastern standing *Phlox* is notably composed of long-styled taxa. We recover an
311 early split between *P. stolonifera* and all other taxa (UFB=100), followed by a split of a well-
312 supported clade (UFB=100) containing *P. ovata* out to *P. pulchra*, *P. glaberrima* subsp. *triflora*,
313 and the allotetraploid *P. buckleyi.* The remaining *P. glaberrima* subspecies (*P. glaberrima* subsp.
314 *glaberrima* and *P. glaberrima* subsp. *interior*) are found in a neighboring clade, with subspecies

of *P. carolina* and *P. maculata* (UFB=100). Across all inferences, the *P. glaberrima* and *P. carolina* subspecies are not monophyletic within their species delineation. The next branch subtends a well-supported small clade including the final long-styled species, *P. amplifolia* and *P. paniculata* (UFB=100), with the placement of *P. amplifolia* within or sister to *P. paniculata* varying across levels of missingness.

Across all datasets, our analyses inferred strong support for a single transition from long to short-styled species, with all short-styled taxa forming a monophyletic group (UFB=100). This short-styled group contains taxa from the *P. pilosa* complex, the *P. divaricata* and *P. amoena* subspecies, and the three Texas annual species (*P. drummondii*, *P. roemeriana*, and *P. cuspidata*). Half of our *P. pilosa* subspecies samples, including *P. pilosa* subsp. *fulgida*, *P. pilosa* subsp. *pilosa*, and *P. pilosa* subsp. *deamii*, form a well-supported clade out to the remaining short-style species (UFB=100). We refer to this group and the *P. pilosa* subsp. *pilosa* individuals within it as "North". Notably, the white flower morph of *P. pilosa* subsp. *pilosa* that has evolved by reinforcement, and sometimes described as the putative hybrid species *P. argillacea*, is found in this clade, and *P. pilosa* subsp. *fulgida* is found nested within *P. pilosa* subsp. *pilosa* North. Both *P. divaricata* and *P. amoena* species respectively form strongly supported monophyletic groups (UFB=100), although *P. divaricata* subspecies are interdigitated and do not form monophyletic groups.

Figure 2. Maximum likelihood phylogenetic inference of eastern standing *Phlox* from whole concatenated ddRAD loci. Loci were required to be shared by at least 30 individuals. Numbers on branches represent bootstrap supports from 100 bootstrap replicates using ultrafast bootstrap approximation (UFB). Bootstrap supports are shown for all nodes differentiating between taxa.

338



339

The remaining *P. pilosa* subspecies, Texas annuals, and the allotetraploids *P. floridana, P. villosissima* subsp. *villossisma* and *P. villosissima* subsp. *latisepala* form a large, well supported clade (UFB=100). Along the backbone of this clade we observed reduced branch supports; however, the three annual species, *P. cuspidata*, *P. drummondii*, and *P. roemeriana,* are monophyletic (UFB=100), with *P. roemeriana* sister to *P. drummondii,* consistent with previous genomic inferences (Roda et al. 2017)**.** Adjacent to the annuals lie the remaining *P. pilosa* subspecies. We refer to this *P. pilosa* group as "South" because of the relative southern origin of these samples to our other *P. pilosa* subspecies samples. This group contains *P. pilosa* subsp. *pliosa*, *P. pilosa* subsp. *detonsa*, and *P. pilosa* subsp. *ozarkana.* Notably we also observed *P. pilosa* subsp. *sangamonensis* is within this *P. pilosa* South group, despite being from well into the *P. pilosa* North group's range in Illinois.

Low support for partitioning of the *P. pilosa* South group from the Texas annuals across our datasets may be due to the presence of individuals from the two allotetraploid subspecies *P. villosissima* subsp. *latisepala* and *P. villosissima* subsp. *villosissima.* Across datasets, these taxa are inferred either out to the annuals or out to the *P. pilosa* South group with variable support (UFB=41-100). Variation in their inferred placement between these groups is consistent with their theorized ancestry as hybrids between populations of *P. pilosa* subsp. *pilosa* South and *P. drummondii* (Table 1). Similar uncertainty can also be observed for the allotetraploid *P. floridana* across datasets, observed either out or nested within the *P. pilosa* South group. This may also be a result of its putative hybrid ancestry between populations of *P. pilosa* subsp. *pilosa* South and *P. carolina* subsp. *angusta* (Table 1).

Coalescent-based species tree inference using SVDquartets recovered a species tree with nearly identical relationships to our ML inferred trees (Supplementary Fig. S3); however, we did

observe differences in branch supports at some interspecific positions. We inferred strong support for most nodes along the backbone of the tree (SBS > 80) and splits between major species groupings, except for early splits differentiating relative placement within the long-style species. *P. pulchra*, *P. glaberrima* subsp. *triflora*, and *P. buckleyi* remained monophyletic (SBS=84) and separate from the remaining monophyletic *P. glaberrima*/*P.maculata*/*P.carolina* subspecies (SBS=100). We also recovered the previously observed polyphyletic relationships of the *P. pilosa* North and *P. pilosa* South groups. Unlike our ML analyses, our coalescent inferences showed support for a cleaner partitioning of a monophyletic clade of *P. pilosa* South and the allotetraploids *P. floridana*, *P. villosissima* subsp. *villosissima* and *P. villosissima* subsp. *latisepala*, sister to the Texas Annuals (SBS=89).

Evidence of Gene Flow Alongside Reinforcement

We used D-statistics and TreeMix v.1.13 to test for a history of gene flow in two cases of reinforcement in *Phlox*. Specifically, we leveraged our sampling to compare if a significant signature of gene flow exists between the focal sympatric species in each case of reinforcement. All D-statistic tests were structured so a significant negative D supports a history of introgression between these focal species.

In *P. drummondii*, reinforcement against hybridization with *P. cuspidata* has favored the divergence of flower color from light-blue to dark-red in sympatry (Fig. 3a). With D-statistic tests for gene flow between *P. drummondii* and *P. cuspidata*, we found no evidence of introgression between light-blue allopatric populations (D values were negative but not significant); However, tests including dark-red *P. drummondii* individuals sympatric with *P. cuspidata* (N=48), all tests had a negative D value and 24 out of 48 tests showed a significant excess of BABA patterns with a |Z-score| > 3 (Table 2). 18 of these significant tests used

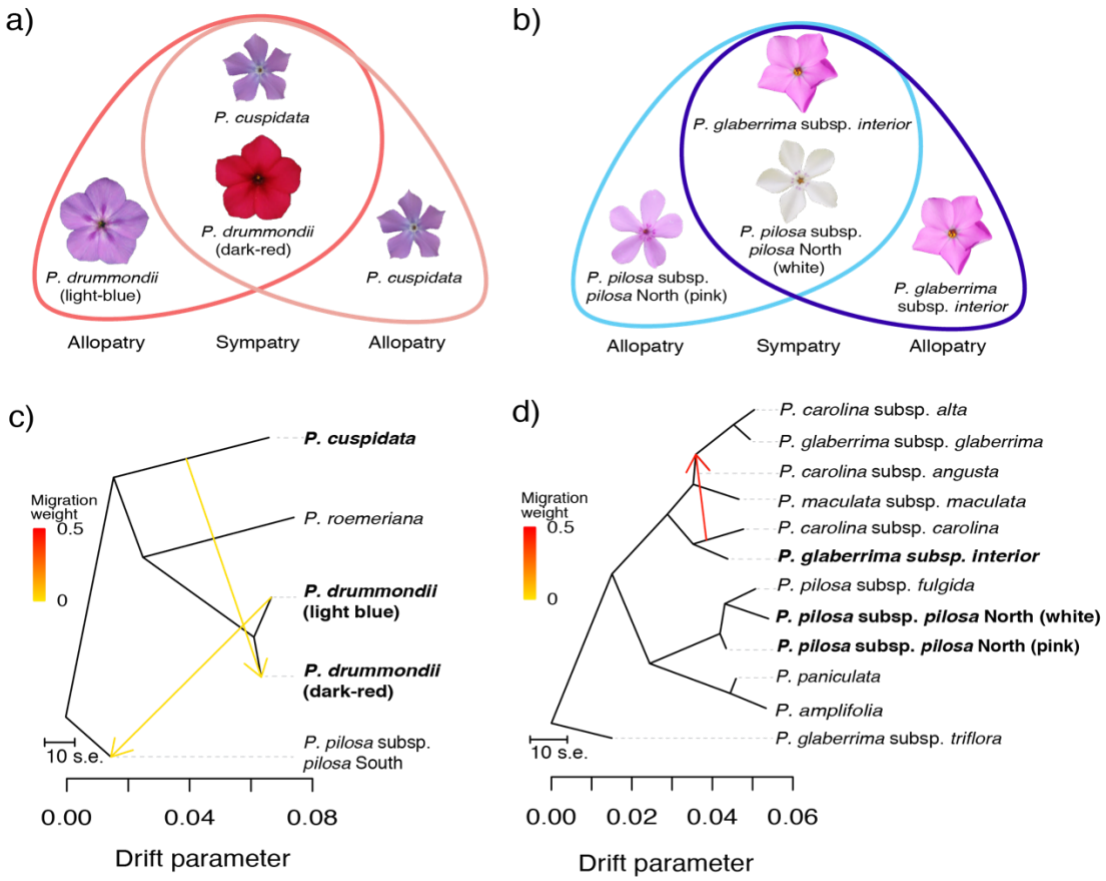


Figure 3. (a and b) Schematics of flower color divergence due to reinforcement in two species of *Phlox* to prevent maladaptive hybridization with another species in sympatry. (a) *P. drummondii* (red) has similar light-blue colored flowers to *P. cuspidata* (pink) in allopatry but *P. drummondii* has evolved dark-red flowers in sympatry to reduce hybridizing with *P. cuspidata*. (b) In allopatry *P. pilosa* subsp. *pilosa* North (light blue) has similar pink colored flowers to *P. glaberrima* subsp. *interior* (dark blue), but where they co-occur in high frequency in sympatry, *P. pilosa* subsp. *pilosa* North has evolved white flowers to reduce hybridizing with *P. glaberrima* subsp. *interior*. (c and d) Best fit models from TreeMix inferred gene flow between *P. cuspidata* and dark-red *P. drummondii* (c) (best fit m=2) but did not support evidence of gene flow between *P. glaberrima* subsp. *interior* and either color morph of *P. pilosa* subsp. *pilosa* North (d) (best fit m=1).

Table 2. Summary results for four taxon D-statistic tests for introgression in two cases of reinforcement

Reinforcement Scenario	P1	P2	P3	Range D	Range Z†	nSig/N‡ -D	nSig/N‡ +D	Test ID
P. drummmondii	*P. drummmondii* (light-blue)	*P. roemariana*	*P. cuspidata*	(-0.12. -0.01)	(0.187, 2.479)	0/24	0/24	T2:1 - T2:24
	P. drummmondii (dark-red)	*P. roemariana*	*P. cuspidata*	(-0.16. -0.01)	(0.341, **4.692**)	**24**/48	0/48	T2:25 - T2:72
P. pilosa subsp. *pilosa* North								
	P. pilosa subsp. *pilosa* North (pink)	*P. divaricata* subsp. *divaricata*	*P. glaberrima* subsp. *interior*	(-0.10. 0.08)	(0.008, 2.771)	0/180	0/180	T2:73 - T2:252
	P. pilosa subsp. *pilosa* North (white)	*P. divaricata* subsp. *divaricata*	*P. glaberrima* subsp. *interior*	(-0.08. 0.09)	(0.016, 2.418)	0/90	0/90	T2:253 - T2:342
	P. pilosa subsp. *pilosa* North (pink)	*P. paniculata*	*P. glaberrima* subsp. *interior*	(-0.09. 0.05)	(0.025, 2.707)	0/180	0/180	T2:343 - T2:522
	P. pilosa subsp. *pilosa* North (white)	*P. paniculata*	*P. glaberrima* subsp. *interior*	(-0.09. 0.10)	(0.026, 2.566)	0/90	0/90	T2:523 - T2:612

P1, P2, P3 refer to the three lineages used for the ABBA-BABA D statistic test (See Methods). Outgroups not shown. All tests can be found in Table S2.

† Z-score to assess significance of whether D deviates from zero. Bold indicates significance at α=0.01.

‡ Number of significant tests over all possible sampled individuals. Significant -D identifes gene flow between P1 and P3, while significant +D identifies gene flow between P2 and P3.

topologies with a *P. cuspidata* individual sampled from sympatry. Our best fit model (m=2) of the *P. drummondii* subtree in TreeMix recovered *P. cuspidata* out to *P. drummondii* and *P. roemeriana* with a migration event going from *P. cuspidata* into dark-red sympatric *P. drummondii* (Fig. 3c). This migration edge remained present for all models m ≥2 (Supplementary Fig. S4, S9). These results support the hypothesis of gene flow occurring between sympatric *P. drummondii* and *P. cuspidata* but do not show evidence of gene flow between allopatric *P. drummondii* and *P. cuspidata*.

In *P. pilosa* subsp. *pilosa*, reinforcement against hybridization with *P. glaberrima* subsp. *interior* has favored the divergence of *P. pilosa* subsp. *pilosa* flower color from pink to white (Fig. 3b). Our phylogenetic analyses demonstrate the white *P. pilosa* subsp. *pilosa* populations are monophyletic and nested within a group of pink *P. pilosa* subsp. *pilosa* in the *P. pilosa* North group and polyphyletic from the rest of *P. pilosa* subsp. *pilosa* South (Fig. 2). This discovery changes the context of this evolutionary scenario, with *P. pilosa* subsp. *pilosa* North more closely related with *P. glaberrima* subsp. *interior* than previously thought.

Unlike reinforcement in *P. drummondii*, the ranges of *P. pilosa* subsp. *pilosa* North and *P. glaberrima* subsp. *interior* are mosaic and cannot be distinctly divided as geographically allopatric or sympatric (Levin and Kerster 1967). Therefore, we grouped our tests for introgression by flower color phenotype of the *P. pilosa* subsp. *pilosa* North individuals, the ancestral pink flower color and the derived white flower color that evolved by reinforcement, instead of by geography. Our analyses show both an inconsistency in the direction D deviated from zero and a lack of significant signature of gene flow between *P. glaberrima interior* with pink flowered (N=360) and white flowered (N=180) colored *P. pilosa* subsp. *pilosa* North (Table 2). All TreeMix models of the *P. pilosa* subsp. *pilosa* North subtree recovered similar topologies

to our ML phylogenetic inferences but no migration events involving pink or white *P. pilosa* subsp. *pilosa* North (Fig. 3d; Supplementary Fig. S5, S9).

Evidence of Divergence and Gene Flow Underlying Hypotheses of Putative Homoploid Hybrid Species

The eastern standing *Phlox* have five hypothesized cases of homoploid hybrid speciation (Table 1). We leveraged our phylogenetic reconstructions and tests for gene flow to ask if these hypotheses are supported with phylogenetic evidence; specifically, if putative hybrid species sit nested within or sister to one of its hypothesized parental species and demonstrating a signature of gene flow, and therefore hybridization, with its other putative parental species.

Our ML and coalescent based phylogenetic inferences show support for four of the five putative hybrids sitting nested or within one of their putative parents (Fig. 2, Supplementary Fig. S1, S2, S3); *P. amoena* subsp. *lighthipei* within *P. amoena* subsp. *amoena*, *P. maculata* subsp. *pyramidalis* within *P. maculata* subsp. *maculata*, *P. argillacea* within *P. pilosa* subsp. *pilosa* North, and *P. pilosa* subsp. *detonsa* in *P. pilosa* subsp. *pilosa* South. *P. pilosa* subsp. *deamii* was not found out or within *P. pilosa* subsp. *pilosa* South or *P. amoena* subsp. *amoena*, consistent with (Goulet-Scott et al. 2021)**.**

D-statistic tests for gene flow did not find support for gene flow between three of the putative hybrid species and their putative second parental species, *P. amoena* subsp. *lighthipei* with *P. pilosa* subsp. *pilosa* South (N=624), *P. argillacea* with *P. glaberrima* subsp. *interior* (N=270), and *P. pilosa* subsp. *detonsa* with *P. carolina* subspecies (N=90*)* (Table 3). Our TreeMix analyses generally recapitulated similar topologies of the ML phylogenetic inferences but found no support for gene flow underlying these hypotheses (Supplementary Fig. S5, S6, S7, S9).

Table 3. Summary results for four taxon D-statistic tests for introgression in four hypothesized cases of hybrid speciation

Hypothesized Hybrid	P1	P2	P3	Range D	Range Z[†]	nSig/N[‡] -D	nSig/N[‡] +D	Test ID
P. argillacea								
	P. argillacea	*P. pilosa* subsp. *pilosa* North	*P. glaberrima* subsp. *interior*	(-0.09, 0.08)	(0.023, 2.115)	0/90	0/90	T3:1 - T3:90
	P. argillacea	*P. divaricata* subsp. *divaricata*	*P. glaberrima* subsp. *interior*	(-0.08. 0.09)	(0.016, 2.418)	0/90	0/90	T3:91 - T3:180
	P. argillacea	*P. paniculata*	*P. glaberrima* subsp. *interior*	(-0.09. 0.10)	(0.026, 2.566)	0/90	0/90	T3:181 - T3:270
P. pilosa subsp. *detonsa*								
	P. pilosa subsp. *detonsa*	*P. pilosa* subsp. *pilosa* South	*P. carolina* subsp. *angusta*	(-0.14. 0.01)	(0.033, 2.383)	0/12	0/12	T3:271 - T3:282
	P. pilosa subsp. *detonsa*	*P. pilosa* subsp. *pilosa* South	*P. carolina* subsp. *alta*	(-0.09. 0.02)	(0.042, 1.520)	0/12	0/12	T3:283 - T3:294
	P. pilosa subsp. *detonsa*	*P. pilosa* subsp. *pilosa* South	*P. carolina* subsp. *carolina*	(-0.13. 0.06)	(0.207, 1.794)	0/6	0/6	T3:295 - T3:300
	P. pilosa subsp. *detonsa*	*P. drummondii*	*P. carolina* subsp. *angusta*	(-0.08. -0.01)	(0.223, 1.880)	0/12	0/12	T3:301 - T3:312
	P. pilosa subsp. *detonsa*	*P. drummondii*	*P. carolina* subsp. *alta*	(-0.05. 0.05)	(0.164, 1.238)	0/12	0/12	T3:313 - T3:324
	P. pilosa subsp. *detonsa*	*P. drummondii*	*P. carolina* subsp. *carolina*	(-0.04. 0.06)	(0.024, 1.147)	0/6	0/6	T3:325 - T3:330
	P. pilosa subsp. *detonsa*	*P. roemeriana*	*P. carolina* subsp. *angusta*	(-0.05, 0.06)	(0.055, 1.266)	0/12	0/12	T3:331 - T3:342
	P. pilosa subsp. *detonsa*	*P. roemeriana*	*P. carolina* subsp. *alta*	(-0.05, 0.13)	(0.024, 2.696)	0/12	0/12	T3:343 - T3:354
	P. pilosa subsp. *detonsa*	*P. roemeriana*	*P. carolina* subsp. *carolina*	(-0.05, 0.02)	(0.188, 1.102)	0/6	0/6	T3:355 - T3:360
P. amoena subsp. *lighthipei*								
	P. amoena subsp. *lighthipei*	*P. amoena amoena*	*P. pilosa* subsp. *pilosa* South	(-0.14. 0.08)	(0.005, 2.887)	0/234	0/234	T3:361 - T3:594
	P. pilosa subsp. *pilosa* South	*P. drummondii*	*P. amoena* subsp. *lighthipei*	(-0.13, 0.06)	(0.007, 2.631)	0/234	0/234	T3:595 - T3:828
	P. pilosa subsp. *pilosa* South	*P. cuspidata*	*P. amoena* subsp. *lighthipei*	(-0.10, 0.13)	(0.007, 2.559)	0/156	0/156	T3:829 - T3:984
P. maculata subsp. *pyramidalis*								
	P. maculata subsp. *pyramidalis*	*P. maculata* subsp. *maculata*	*P. glaberrima* subsp. *interior*	(-0.50, 0.08)	(1.869, **20.718**)	**27**/30	0/30	T3:985 - T3:1014
	P. glaberrima subsp. *interior*	*P. carolina* subsp. *alta*	*P. maculata* subsp. *pyramidalis*	(-0.34, -0.06)	(2.002, **13.515**)	**17**/20	0/20	T3:1015 - T3:1034
	P. glaberrima subsp. *interior*	*P. glaberrima* subsp. *glaberrima*	*P. maculata* subsp. *pyramidalis*	(-0.31, -0.06)	(2.167, **12.277**)	**19**/20	0/20	T3:1035 - T3:1054
	P. glaberrima subsp. *interior*	*P. carolina* subsp. *alta*	*P. maculata* subsp. *maculata*	(0.01, 0.20)	(0.299, **6.323**)	0/80	**62**/80	T3:1055 - T3:1134
	P. glaberrima subsp. *interior*	*P. glaberrima* subsp. *glaberrima*	*P. maculata* subsp. *maculata*	(-0.02, 0.19)	(0.090, **6.313**)	0/80	**59**/80	T3:1135 - T3:1214

P1, P2, P3 refer to the three lineages used for the ABBA-BABA D statistic test (See Methods). Outgroups not shown. All tests can be referenced by Test ID in Table S3.

[†] Z-score to assess significance of whether D deviates from zero. Bold indicates significance at α=0.01.

[‡] Number of significant tests over all possible sampled individuals. Significant -D identifes gene flow between P1 and P3, while significant +D identifies gene flow between P2 and

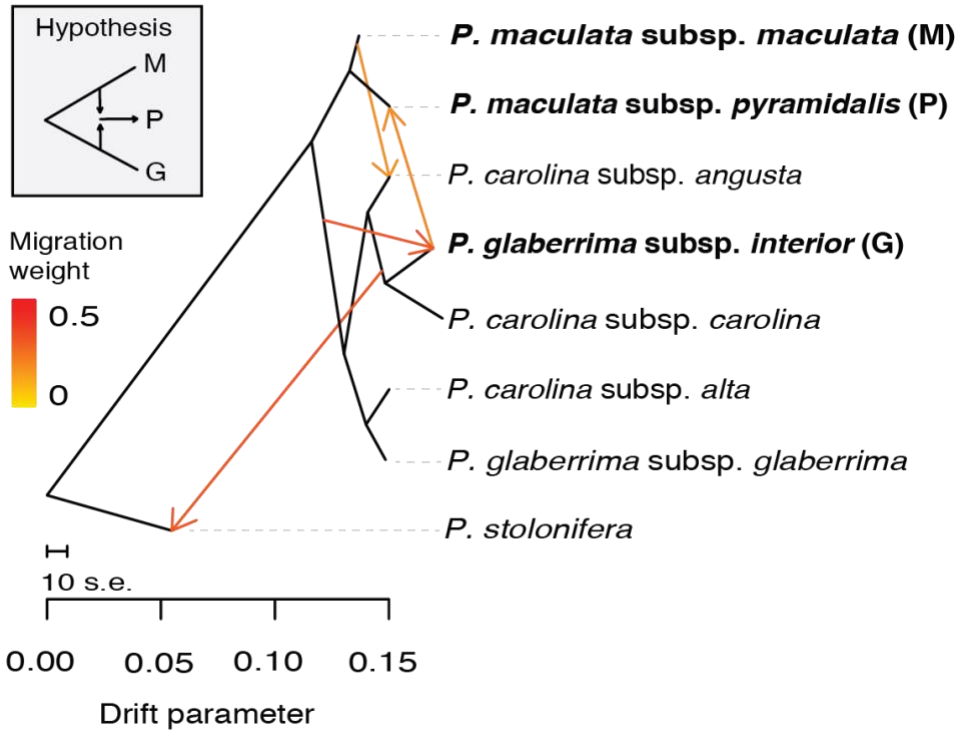


Figure 4. Results from TreeMix support the hybrid speciation hypothesis for *P. maculata pyramidalis* in a pruned subtree (best fit m=4). Modeling of 4,490 unlinked SNPs for individuals clustered by taxon identity inferred *P. maculata* subsp. *pyramidalis* sister to *P. maculata* subsp. *maculata* and receiving a migration event (gene flow) from *P. glaberrima* subsp. *interior*.

However, our D-statistic tests did identify strong support for gene flow between *P. glaberrima* subsp. *interior* and *P. maculata* subsp. *pyramidalis*, with significantly negative values in 63 of 70 tests. We found no evidence of gene flow between *P. glaberrima* subsp. *interior* and *P. maculta* subsp. *maculta*, N=160 (Table 3). All models of the pyramidalis subtree in TreeMix (best fit m=4) confirmed this result, with a migration edge from *P. glaberrima* subsp. *interior* into *P. maculta* subsp. *pyramidalis* (Fig. 4, Supplementary Fig. S8, S9). These results are

consistent with *P. maculata* subsp. *pyramidalis* as a hybrid lineage arising from *P. maculata* subsp. *maculata* and *P. glaberrima* subsp. *interior* parental lineages.

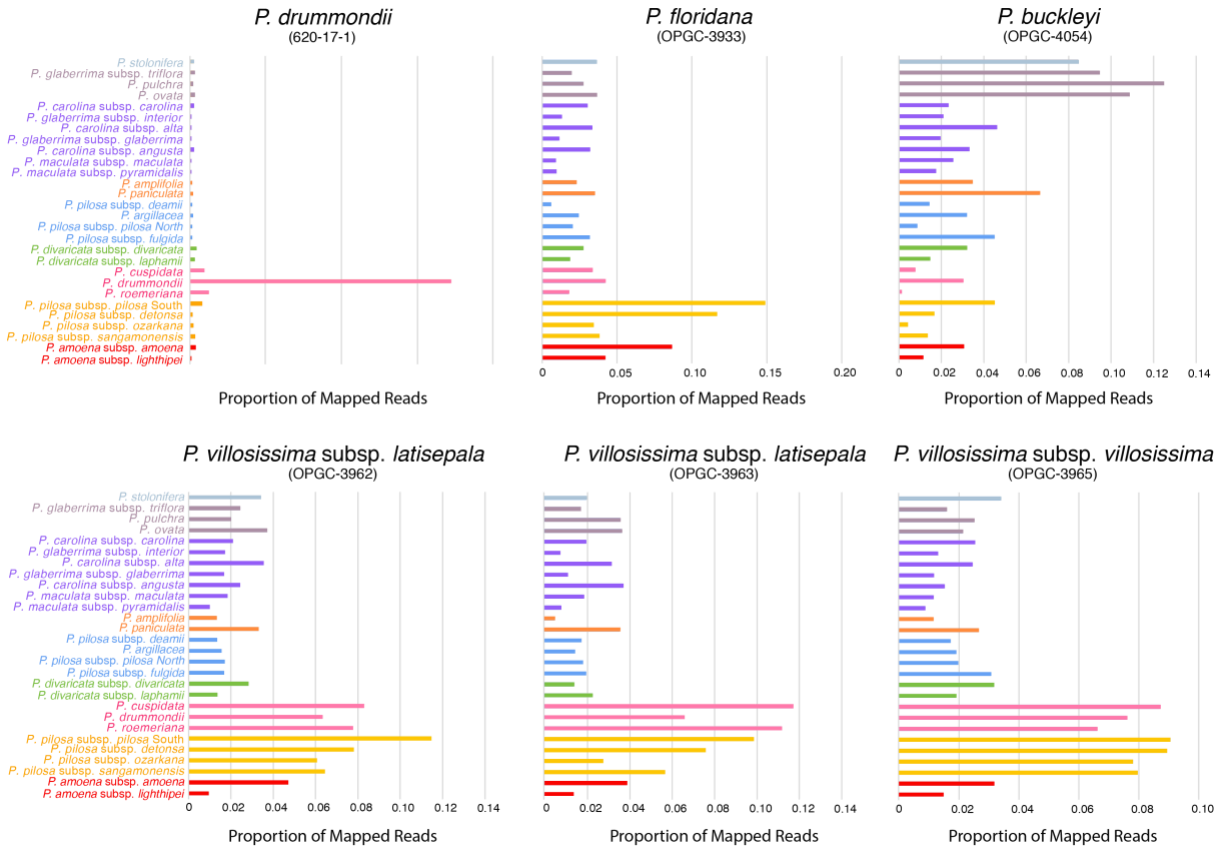
Inferring Origins of Allotetraploid Hybrid Species

Hybridization of divergent diploid genomes formed four, reproductively isolated, tetraploid *Phlox* species (Table 1) (Levin 1966a, 1968; Smith and Levin 1967). *P. villosissima* subsp. *villosissima* and *P. villosissima* subsp. *latisepala* are thought to be independent polyploid hybridization events between *P. pilosa* subsp. *pilosa* South and *P. drummondii*. *P. floridana* is hypothesized to be the product of *P. pilosa* subsp. *pilosa* South and *P. carolina* subsp. *angusta***,** and finally, *P. buckleyi* contains two discordant diploid karyotypes not observed elsewhere and the ancestors are hypothesized to be extinct (Wherry 1955)**.** We leveraged comparative read mapping to infer subgenome ancestry of these taxa and evaluate their hypothesized origins.

We concatenated ddRAD loci to generate reduced consensus sequences for each of the 28 diploid species. Each consensus was 439,271 bp long with a locus missingness rate of 25%. For one test representative of each diploid species, an average of 37,948 reads mapped with high quality to the consensus multi-sequence FASTA file. The distribution of mapping rates across consensus sequences demonstrated a clear ability to assign ancestry of diploid genomes, with the consensus receiving the highest relative proportion of mapped reads corresponding to the known identity of the diploid individual (Supplementary Table S4).

We repeated this mapping for each allotetraploid, with an average of 17,737 high-quality mapped reads per polyploid individual. We observed a more diffuse read mapping rate across consensus sequences relative to the test diploid taxa, yet clear patterns of elevated read mapping constrained to specific subclades (Fig. 5). Reads of *P. villosissima* subsp. *vollosissima* and *P.*

485 *villosissima* subsp. *latisepala* mapped primarily to the *P. pilosa* South group and the Texas
486 annuals, with each receiving 30-34% and 22-26% of the mapped reads respectively. Within these
487 two groups the top recipients of mapped reads were *P. pilosa* subsp. *pilosa* South and *P.*
488 *cuspidata* for all three allotetraploid individuals. *P. floridana* showed a similar high mapping rate
489 to the *P. pilosa* South group (34%), with most reads mapping to *P. pilosa* subsp. *pilosa* South
490 and second most to *P. pilosa* subsp. *detonsa*. Contrary to the hypothesized origins of this species,



491

492 **Figure 5.** Proportion of reads mapped to diploid species-level consensus sequences for a diploid
493 representative of *P. drummondii* and all allotetraploid hybrid individuals. Bar colors correspond
494 to diploid species groups in Figure 2.

495

we did not observe elevated read mapping to the *P. carolina subspecies* consensus sequence. Instead, we observed elevated read mapping to *P. amoena* subsp. *amoena*. Finally, *P. buckleyi* showed high rates of read mapping to taxa within the deeper branch of the eastern standing *Phlox*, with highest rate of mapping to the *P. ovata*, *P. pulchra*, *P. glaberrima* subsp. *triflora* clade (38%), with highest recipients being *P. pulchra* and *P. ovata.*

DISCUSSION

Phylogenomic inferences of species relationships and explicit tests for interlineage gene flow provide a systematic way for inferring the presence, extent, and context of hybridization and gene flow across a phylogeny. Here we present well-resolved phylogenetic relationships of the eastern standing *Phlox*. We demonstrate clear support for most described species relationships and reveal novel non-monophyletic relationships of subspecies in the historically taxonomically difficult *P. pilosa* and *P. glaberrima*/ *P. carolina*/ *P. maculata* species complexes. Using this phylogenetic framework, we find support for one case of hypothesized homoploid hybrid speciation event, identify the putative ancestries of multiple polyploid species, and find evidence of gene flow in one of two cases of reinforcement. Our findings demonstrate the utility and importance of phylogenomics in confirming hypothesized evolutionary histories of non-model systems and add to the growing evidence that gene flow across species boundaries can play a role in the generation of novel biodiversity.

Evolutionary Relationships of the Eastern Standing Phlox

For nearly one hundred years, the eastern standing *Phlox* have been a foundational system for understanding plant evolution, yet despite previous phylogenetic study on this clade (Ferguson et al. 1999; Ferguson and Jansen 2002; Roda et al. 2017; Landis et al. 2018; Goulet-

519 Scott et al. 2021), the evolutionary relationships of these taxa have remained unclear. Our
520 phylogenomic inferences on genome-wide ddRAD sequence data clarifies the evolutionary
521 relationships of the taxa in this group and provides insight into their diversification.

522 Historic treatments of the eastern standing *Phlox* have grouped the species by
523 conspicuous differences in style length (Wherry 1930, 1931, 1932, 1955). Unlike prior
524 phylogenetic studies, we observe support for a large monophyletic clade of short-styled taxa
525 subtended by multiple paraphyletic clades of long-styled taxa (Fig. 2). Divergence in style-length
526 can generate reproductive isolation (Kay 2006; Brothers and Delph 2017), and this transition
527 from long to short style-length may have reduced competition and stimulated the diversification
528 of the short-styled clade.

529 Our phylogenetic analyses recapitulate previously inferred relationship of the Texas
530 Annuals species, as observed (Roda et al. 2017) but not in other studies (Ferguson et al. 1999;
531 Ferguson and Jansen 2002; Landis et al. 2018). We also find clear monophyletic support for the
532 *P. amoena* and *P. divaricata* species groups.

533 In the taxonomically more difficult *P. glaberrima*/ *P. carolina*/ *P. maculata* complex, we
534 find support for non-monophyletic relationships among subspecies (Fig. 2). Notably, *P.*
535 *glaberrima triflora* is phylogenetically distinct from the rest of the interdigitated taxa of this
536 species complex. Additionally, the subspecies of the *P. pilosa* complex form two distinct
537 polyphyletic groups roughly coinciding with their relative northern or southern location of
538 origin, as suggested in (Goulet-Scott et al. 2021). The one exception is *P. pilosa* subsp.
539 *sangamonensis* which genetically clusters with the southern clade but is found in the north (Fig.
540 1); A discovery consistent with the allopatric speciation hypothesis of *P. pilosa* subsp.
541 *sangamonensis* arising by long-distance dispersal from southern populations of *P. pilosa* subsp.

pilosa (Levin and Smith 1965; Levin 1984). The non-monophyletic relationship among the northern and southern *P. pilosa* taxa and within *P. glaberrima* were suggested in previous phylogenetic studies, but remained uncertain due to low phylogenetic supports and reliance on single genetic loci (Ferguson et al. 1999; Ferguson and Jansen 2002; Landis et al. 2018).

Taxonomic delineations within *Phlox* are largely based on morphological characteristics (Wherry 1955; Locklear 2011a), yet our phylogenetic inferences demonstrate current taxonomy may not best reflect the true evolutionary relationships of some taxa within the eastern standing *Phlox* group. Discovery of these new phylogenetic relationships also reframe the phylogenetic context of the historic discussions on hybridization, gene flow, reinforcement, and hybrid speciation within this system, as discussed below.

Evolution of Reinforcement With and Without Gene Flow

Interspecific hybridization can play an important role in speciation by generating selection for reinforcement. However, limited empirical study has evaluated if this hybridization also resulted in reinforcement evolving in the face of gene flow (Garner et al. 2018). We applied D-statistics and TreeMix to assess if a history of gene flow coincides with the evolution of flower color divergence by reinforcement in *P. drummondii* and *P. pilosa* subsp. *pilosa* North (Fig. 3).

Divergence from light-blue to dark-red colored flowers in *P. drummondii* has evolved to prevent hybridization with *P. cuspidata* (Levin 1985; Hopkins and Rausher 2012)**.** Our geographically widespread sampling of these species revealed evidence of introgression between *P. cuspidata* and multiple dark-red *P. drummondii* individuals from across the range of sympatry but not between *P. cuspidata* and any allopatric light-blue *P. drummondii.* The significant signature of gene flow in sympatry but not allopatry mirrors the classic evolutionary model of

565 divergence due to reinforcement to reduce hybridization in sympatry but not in allopatry (Garner
566 et al. 2018). This pattern also suggests the signal of gene flow is the product of hybridization in
567 sympatry and not due to ancient introgression between the *P. cuspidata* and *P. drummondii*
568 lineages. Our results corroborate previous inferences from transcriptome sequencing of single
569 individuals of each species (Roda et al. 2017). Although F1 hybrids are highly sterile and have
570 reduced ability to backcross in the lab (Suni and Hopkins 2018), our results demonstrate hybrids
571 do backcross in nature, resulting in gene flow between the species. Reinforcement for increased
572 reproductive isolation in *P. drummondii* evolved despite gene flow in sympatry.

573 Conversely, our analyses do not support a history of gene flow coinciding with
574 hybridization and reinforcement between *P. pilosa* subsp. *pilosa* North and *P. glaberrima* subsp.
575 *interior*. These two species overlap for most of their ranges but co-flower in higher frequency at
576 the lower tip of lake Michigan. In this zone of elevated contact, *P. pilosa* subsp. *pilosa* North has
577 evolved from pink to white colored flowers, reducing the formation of hybrid seed by half in
578 sympatry, making it more adapted to coexist with *P. glaberrima* subsp. *interior* (Levin 1966a;
579 Levin and Kerster 1967; Levin and Schaal 1970b, 1970a, 1972). Phylogenetic distance (Fig. 2)
580 and barriers to reproduction between these taxa are higher than between *P. cuspidata* and *P.*
581 *drummondii*, but experimental crosses demonstrate these species can generate hybrid seed, with
582 *P. pilosa* subsp. *pilosa* North as the maternal parent (Levin 1966a). Our results do not detect
583 gene flow between *P. pilosa* subsp. *pilosa* North and *P. glaberrima* subsp. *interior* regardless of
584 flower color (Fig. 3; Table 2). This lack of gene flow suggests hybrid offspring between these
585 taxa may be completely inviable, lethally maladapted, or have exceedingly high sterility.

586 The absence and presence of gene flow in these two case studies of reinforcement yield a
587 powerful comparative opportunity to investigate the dynamics by which reproductive trait

588 divergence, and flower color specifically, evolves by selection. For reproductive trait divergence
589 to evolve by reinforcement, alleles conferring increased assortative mating within a species must
590 remain associated with alleles causing costly hybridization between species (Servedio and
591 Kirkpatrick 1997; Kirkpatrick 2000). When gene flow between species is present, recombination
592 can disassociate these alleles, impeding successful divergence. Yet, we find the same trait
593 evolving due to similar types of selection in two *Phlox* species despite gene flow occurring
594 alongside one case but not the other. This result may be the consequence of differing genetic
595 architectures and genetic linkage underlying the cost of hybridization and flower color between
596 these cases. For example, strong genetic linkage between loci causing postzygotic
597 incompatibility and flower color divergence can reduce recombination between these traits and
598 allow buildup of linkage disequilibrium and the successful evolution of stronger reproductive
599 isolation in sympatry. Further investigation of the genetic architecture underlying reproductive
600 isolating barriers across these *Phlox* species will better inform the role of trait architecture in the
601 evolution of reinforcement.
602
603 *Varying Support for Hypothesized Homoploid and Allotetraploid Hybrid Species*
604 Within the eastern standing *Phlox* there are five hypothesized homoploid hybrid species
605 and four hypothesized allopolyploid species (Table 1). We only observed genomic support for *P.*
606 *maculata* subsp. *pyramidalis* being a homoploid hybrid species between *P. maculata* subsp.
607 *maculata* and *P. glaberrima* subsp. *interior*. *P. maculata* subsp. *pyramidalis* is found
608 phylogenetically closely related to multiple *P. maculata* subsp. *maculata* individuals (Fig. 2)
609 with a strong signal of receiving gene flow from *P. glaberrima* subsp. *interior* (Fig. 4; Table
610 3.). We can conclude our sampled *P. maculata* subsp. *pyramidalis* is of hybrid origin, however

611 future work will be necessary to discern if the hybrid origin of *P. maculata* subsp. *pyramidalis*
612 actually gave rise to reproductive isolation of its lineage from its putative parental species
613 (Schumer et al. 2018).

614 The remaining four hypothesized cases of homoploid hybrid speciation were not
615 supported by our genomic investigation, despite extensive phenotypic and biomolecular
616 evidence. In the cases of *P. amoena* subsp. *lighthipei*, *P. argillacea*, and *P. pilosa* subsp. *detonsa*
617 the lineages thought to be hybrid in origin may simply be divergent geographically isolated
618 populations from one of their hypothesized parental lineages. Yet in the case of *P. pilosa* subsp.
619 *deamii*, the lineage is distantly related to both hypothesized parents and may possesses
620 phenotypic traits resembling these species through convergence or incomplete lineage sorting.

621 Polyploid hybrid lineages can be significantly easier to identify than homoploid hybrids
622 through chromosome structure and counts. Despite this advantage, inferring the exact
623 evolutionary origin of these lineages remains challenging (Rothfels 2021). Although the signals
624 from our comparative read mapping are coarse, our approach suggests the identity of the
625 progenitor clades that gave rise to the polyploid hybrid species in this group. We provide further
626 support that *P. villosissima* subsp. *latisepala* and *P. villosisima* subsp. *villosissima* are derived
627 from *P. pilosa* (*P. pilosa* subsp. *pilosa* South) and a Texas annual species. *P. drummondii* was
628 previously hypothesized as the parent but our analyses cannot resolve which of the Texas annual
629 species or ancestor is the parent. As previously hypothesized, we find support for one parent of
630 *P. floridana* to be *P. pilosa* subsp. *pilosa* South, yet we find evidence that *P. amoena* subsp.
631 *amoena* may be the second parent instead of the *P. carolina subspecies*. We also observe *P.*
632 *buckleyi* is genetically similar to the early-diverging long-styled *Phlox*, but we cannot confirm
633 the specific parental lineages from this group. Future work using longer haplotype phase may

help resolve the evolutionary origins of these species and better inform the dynamics and evolutionary trajectory of these allopolyploid species.

Across both the hypothesized homoploid and polyploid hybrid lineages, the *P. pilosa* group has stood out as a potential and realized hotspot of hybridization activity. Four of the five hypothesized homoploid hybrid lineages and three of the four polyploid lineages were thought to include *P. pilosa* subsp. *pilosa* as a parent. While our analyses indicate that none of the hypothesized homoploid hybrid lineages have hybrid origin with *P. pilosa* subspecies, three of the four hypothesized polyploid hybrid lineages are derived from *P. pilosa* subsp. *pilosa* South *subspecies* ancestry.

What has made *P. pilosa* prime for a role in these hypotheses? Notably the high occurrence of *P. pilosa* subspecies across their composite broad range has not only led to them being found in sympatry to many species, but potentially also as indicators of locally adapted phenotypes across these environments. *P. pilosa*'s developmental biology may also make it prone to evolutionary innovation, with many populations observed having high rates of unreduced gametes and whole genome duplications across their range. Future work may help address what factors underly this developmental instability and if it has an outsized role in the generation of novel evolutionary lineages relative to other *Phlox* species.

Implications on Evolutionary History of Trait Variation

Now that we have a strong phylogenetic backbone for the evolutionary relationships of taxa in the eastern standing *Phlox,* we can begin to untangle what forces have contributed to patterns of phenotypic divergence throughout the clade. Much of the data supporting the role of hybridization in speciation in this group stemmed from trait variation between distantly related lineages. Our findings support hybridization has likely contributed to some of the phenotypic

variation observed across taxa in these cases, both as a source of genetic variation and force for selective divergence. Further study of these confirmed cases may inform how specific traits and even genes move between species and contribute to reproductive isolation between lineages.

Our findings also indicate that forces other than hybridization may be responsible for convergent and recombinant patterns of trait variation across *Phlox.* Our lack of support for many hypothesized cases of hybridization motivates future investigation of what did give rise to the patterns of trait variation across and within these lineages. Further research into these taxa and their phenotypic variation may provide new insights into the importance of convergent evolution and/or incomplete lineage sorting to local adaptation within a species.

CONCLUSION

Hybridization between species can be an important force shaping the patterns of evolution across clades. For example, decades of research have suggested that hybridization and gene flow are extensive across the eastern standing *Phlox* and have directly played a role in their speciation. With the accessibility of genomic data and improved analyses, we are now able to test these hypotheses. We have built a robust phylogenetic framework using genome-wide markers to better understand the relationship between these well-studied species and document the occurrence of gene flow. We have determined hybridization has served as an important creative force in the generation of diversity in this clade, however it may not be as prevalent as previously thought. We do find evidence suggesting hybrid speciation and gene flow during reinforcement, but we also find a lack of support for many hypothesized cases of gene flow in the formation of evolutionary novel lineages. These lineages may just be the product of trait divergence within a species separate from interspecific genetic exchange. Our work suggests a

strong need to reevaluate our previous evidence for the existence of gene flow, and to rethink our expectations of how traits diverge within and between closely related species.

FUNDING

This study was supported by the National Science Foundation (DEB-1844906 to R.H., and Graduate Research Fellowship to A.G.G.); the Harvard Graduate School of Arts and Sciences' Herchel Smith Graduate Research Fellowship in Science (to A.G.G. and B.E.G.S.); and the Harvard University Herbaria Expedition Grant (to A.G.G.).

ACKNOWLEDGEMENTS

We are grateful to C.J. Ferguson, S.V. Edwards, J.S. Suissa, and the Hopkins Lab for their helpful comments on this manuscript. We are also thankful to P.S. Jourdan and the OPGC staff for sharing their *Phlox* collections with us.

CONTRIBUTIONS

A.G.G. and R.H. conceptualized the project. A.G.G., B.E.G.S., and R.H. collected and analyzed the data. A.G.G. and R.H. wrote the manuscript.

DATA ACCESSIBILITY

Supporting data underlying this article is found on Dryad Digital Repository: [insert rep ID}

REFERENCES

Abbott R., Albach D., Ansell S., Arntzen J.W., Baird S.J.E., Bierne N., Boughman J., Brelsford A., Buerkle C.A., Buggs R., Butlin R.K., Dieckmann U., Eroukhmanoff F., Grill A., Cahan S.H., Hermansen J.S., Hewitt G., Hudson A.G., Jiggins C., Jones J., Keller B., Marczewski T., Mallet J., Martinez-Rodriguez P., Möst M., Mullen S., Nichols R., Nolte A.W., Parisod C., Pfennig K., Rice A.M., Ritchie M.G., Seifert B., Smadja C.M., Stelkens R., Szymura J.M., Väinölä R., Wolf J.B.W., Zinner D. 2013. Hybridization and speciation. J. Evol. Biol. 26:229–246.

Anderson E. 1949. Introgressive hybridization. Introgressive hybridization.

Anderson E., Gage A. 1952. Introgressive hybridization in phlox bifida. Am. J. Bot. 39:399–404.

Andrews K.R., Good J.M., Miller M.R., Luikart G., Hohenlohe P.A. 2016. Harnessing the power of RADseq for ecological and evolutionary genomics. Nat. Rev. Genet. 17:81–92.

Barker M.S., Arrigo N., Baniaga A.E., Li Z., Levin D.A. 2016. On the relative abundance of autopolyploids and allopolyploids. New Phytol. 210:391–398.

Bolger A.M., Lohse M., Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 30:2114–2120.

Bombonato J.R., do Amaral D.T., Silva G.A.R., Khan G., Moraes E.M., da Silva Andrade S.C., Eaton D.A.R., Alonso D.P., Ribolla P.E.M., Taylor N., Zappi D., Franco F.F. 2020. The potential of genome-wide RAD sequences for resolving rapid radiations: a case study in Cactaceae. Mol. Phylogenet. Evol. 151:106896.

724 Brothers A.N., Delph L.F. 2017. Divergence in style length and pollen size leads to a
725 postmating-prezygotic reproductive barrier among populations of Silene latifolia.
726 Evolution. 71:1532–1540.

727 Catchen J., Hohenlohe P.A., Bassham S., Amores A., Cresko W.A. 2013. Stacks: an analysis tool
728 set for population genomics. Mol. Ecol. 22:3124–3140.

729 Chifman J., Kubatko L. 2014. Quartet inference from SNP data under the coalescent model.
730 Bioinformatics. 30:3317–3324.

731 Durand E.Y., Patterson N., Reich D., Slatkin M. 2011. Testing for ancient admixture between
732 closely related populations. Mol. Biol. Evol. 28:2239–2252.

733 Dyer K.A., Bewick E.R., White B.E., Bray M.J., Humphreys D.P. 2018. Fine-scale geographic
734 patterns of gene flow and reproductive character displacement in Drosophila subquinaria
735 and Drosophila recens. Mol. Ecol.

736 Eaton D.A.R., Hipp A.L., González-Rodríguez A., Cavender-Bares J. 2015. Historical
737 introgression among the American live oaks and the comparative nature of tests for
738 introgression. Evolution. 69:2587–2601.

739 Eaton D.A.R., Overcast I. 2020. ipyrad: Interactive assembly and analysis of RADseq datasets.
740 Bioinformatics. 36:2592–2594.

741 Eaton D.A.R., Ree R.H. 2013. Inferring phylogeny and introgression using RADseq data: an
742 example from flowering plants (Pedicularis: Orobanchaceae). Syst. Biol. 62:689–706.

743 Erbe L., Turner B.L. 1962. A biosystematic study of the phlox cuspidata-phlox drummondii
744 complex. Am. Midl. Nat. 67:257.

745 Excoffier L., Dupanloup I., Huerta-Sánchez E., Sousa V.C., Foll M. 2013. Robust demographic
746 inference from genomic and SNP data. PLoS Genet. 9:e1003905.

747 Fehlberg S.D., Ty M.C., Ferguson C.J. 2014. Reexamination of a putative diploid hybrid taxon
748 using genetic evidence: The distinctiveness ofPhlox pilosasubsp.Deamii(polemoniaceae).
749 Int. J. Plant Sci. 175:781–793.

750 Felsenstein J. 1981. Skepticism Towards Santa Rosalia, or Why are There so Few Kinds of
751 Animals? Evolution. 35:124.

752 Ferguson C.J., Jansen R.K. 2002. A chloroplast DNA phylogeny of eastern Phlox
753 (Polemoniaceae): implications of congruence and incongruence with the ITS phylogeny.
754 Am. J. Bot. 89:1324–1335.

755 Ferguson C.J., Kramer F., Jansen R.K. 1999. Relationships of eastern north American phlox
756 (polemoniaceae) based on ITS sequence data. Syst. Bot. 24:616.

757 Fishman L., Wyatt R. 1999. Pollinator-mediated competition, reproductive character
758 displacement, and the evolution of selfing in Arenaria uniflora (Caryophyllaceae).
759 Evolution. 53:1723–1733.

760 Garner A.G., Goulet B.E., Farnitano M.C., Molina-Henao Y.F., Hopkins R. 2018. Genomic
761 Signatures of Reinforcement. Genes . 9.

762 Gottlieb L.D. 1972. Levels of confidence in the analysis of hybridization in plants. Ann. Mo.
763 Bot. Gard. 59:435.

764 Goulet B.E., Roda F., Hopkins R. 2017. Hybridization in plants: Old ideas, new techniques.
765 Plant Physiol. 173:65–78.

766 Goulet-Scott B.E., Garner A.G., Hopkins R. 2021. Genomic analyses overturn two long-standing
767 homoploid hybrid speciation hypotheses. Evolution. 75:1699–1710.

768 Grant V. 1981. Plant Speciation. New York Chichester, West Sussex: Columbia University
769 Press.

770 Green R.E., Krause J., Briggs A.W., Maricic T., Stenzel U., Kircher M., Patterson N., Li H., Zhai
771 W., Fritz M.H.-Y., Hansen N.F., Durand E.Y., Malaspinas A.-S., Jensen J.D., Marques-
772 Bonet T., Alkan C., Prüfer K., Meyer M., Burbano H.A., Good J.M., Schultz R., Aximu-
773 Petri A., Butthof A., Höber B., Höffner B., Siegemund M., Weihmann A., Nusbaum C.,
774 Lander E.S., Russ C., Novod N., Affourtit J., Egholm M., Verna C., Rudan P., Brajkovic
775 D., Kucan Ž., Gušic I., Doronichev V.B., Golovanova L.V., Lalueza-Fox C., la Rasilla
776 M. de, Fortea J., Rosas A., Schmitz R.W., Johnson P.L.F., Eichler E.E., Falush D., Birney
777 E., Mullikin J.C., Slatkin M., Nielsen R., Kelso J., Lachmann M., Reich D., Pääbo S.
778 2010. A Draft Sequence of the Neandertal Genome. Science. 328:710–722.

779 Guo C., Ma P.-F., Yang G.-Q., Ye X.-Y., Guo Y., Liu J.-X., Liu Y.-L., Eaton D.A.R., Guo Z.-H.,
780 Li D.-Z. 2021. Parallel ddRAD and genome skimming analyses reveal a radiative and
781 reticulate evolutionary history of the temperate bamboos. Syst. Biol. 70:756–773.

782 Gutenkunst R.N., Hernandez R.D., Williamson S.H., Bustamante C.D. 2009. Inferring the joint
783 demographic history of multiple populations from multidimensional SNP frequency data.
784 PLoS Genet. 5:e1000695.

785 Hadley E.B., Levin D.A. 1969. Physiological evidence of hybridization and reticulate evolution
786 in phlox maculata. Am. J. Bot. 56:561–570.

787 Hibbins M., Hahn M.W. 2021. Phylogenomic approaches to detecting and characterizing
788 introgression. EcoEvoRxiv.

789 Hoang D.T., Chernomor O., von Haeseler A., Minh B.Q., Vinh L.S. 2018. UFBoot2: Improving
790 the ultrafast bootstrap approximation. Mol. Biol. Evol. 35:518–522.

791 Hopkins R. 2013. Reinforcement in plants. New Phytol. 197:1095–1103.

792 Hopkins R., Rausher M.D. 2012. Pollinator-mediated selection on flower color allele drives
793 reinforcement. Science. 335:1090–1092.

794 Howard D.J. 1993. Reinforcement: origin, dynamics, and fate of an evolutionary hypothesis.
795 Hybrid zones and the evolutionary process.:46–69.

796 Kay K.M. 2006. Reproductive isolation between two closely related hummingbird-pollinated
797 neotropical gingers. Evolution. 60:538–552.

798 Kirkpatrick M. 2000. Reinforcement and divergence under assortative mating. Proc. Biol. Sci.
799 267:1649–1655.

800 Kong S., Kubatko L.S. 2021. Comparative Performance of Popular Methods for Hybrid
801 Detection using Genomic Data. Syst. Biol. 70:891–907.

802 Kulathinal R.J., Stevison L.S., Noor M.A.F. 2009. The genomics of speciation in Drosophila:
803 diversity, divergence, and introgression estimated using low-coverage genome
804 sequencing. PLoS Genet. 5:e1000550.

805 Landis J.B., Bell C.D., Hernandez M., Zenil-Ferguson R., McCarthy E.W., Soltis D.E., Soltis
806 P.S. 2018. Evolution of floral traits and impact of reproductive mode on diversification in
807 the phlox family (Polemoniaceae). Mol. Phylogenet. Evol. 127:878–890.

808 Lemmon E.M., Juenger T.E. 2017. Geographic variation in hybridization across a reinforcement
809 contact zone of chorus frogs (Pseudacris). Ecol. Evol. 7:9485–9502.

810 Léveillé-Bourret É., Chen B.-H., Garon-Labrecque M.-È., Ford B.A., Starr J.R. 2020. RAD
811 sequencing resolves the phylogeny, taxonomy and biogeography of Trichophoreae
812 despite a recent rapid radiation (Cyperaceae). Mol. Phylogenet. Evol. 145:106727.

813 Levin D.A. 1963. Natural Hybridization Between Phlox maculata and Phlox glaberrima and its
814 Evolutionary Significance. Am. J. Bot. 50:714.

815 Levin D.A. 1966a. The phlox pilosa complex: Crossing and chromosome relationships. Brittonia.
816 18:142.

817 Levin D.A. 1966b. Chromatographic evidence of hybridization and evolution in phlox maculata.
818 Am. J. Bot. 53:238–245.

819 Levin D.A. 1968. The genome constitutions of eastern north American phlox amphiploids.
820 Evolution. 22:612.

821 Levin D.A. 1969. The challenge from a related species: A stimulus for saltational speciation.
822 Am. Nat. 103:316–322.

823 Levin D.A. 1984. Genetic variation and divergence in a disjunct phlox. Evolution. 38:223–225.

824 Levin D.A. 1985. Reproductive Character Displacement in Phlox. Evolution. 39:1275.

825 Levin D.A., Kerster H.W. 1967. Natural selection for reproductive isolation in phlox. Evolution.
826 21:679–687.

827 Levin D.A., Schaal B.A. 1970a. Reticulate evolution in phlox as seen through protein
828 electrophoresis. Am. J. Bot. 57:977–987.

829 Levin D.A., Schaal B.A. 1970b. Corolla color as an inhibitor of interspecific hybridization in
830 phlox. Am. Nat. 104:273–283.

831 Levin D.A., Schaal B.A. 1972. Seed Protein Polymorphism in Phlox pilosa (Polemoniaceae).
832 Brittonia. 24:46.

833 Levin D.A., Smith D.M. 1965. An enigmatic phlox from Illinois. Brittonia. 17:254.

834 Levin D.A., Smith D.M. 1966. Hybridization and Evolution in the Phlox pilosa Complex. Am.
835 Nat. 100:289–302.

836 Levy M., Levin D.A. 1974. Novel flavonoids and reticulate evolution in the phlox pilosa–p.
837 Drummondii complex. Am. J. Bot. 61:156–167.

838 Levy M., Levin D.A. 1975. THE NOVEL FLAVONOID CHEMISTRY AND
839 PHYLOGENETIC ORIGIN OF PHLOX FLORIDANA. Evolution. 29:487–499.

840 Li H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM.
841 arXiv [q-bio.GN].

842 Locklear J.H. 2011a. Phlox a natural history and gardener's guide. Portland, OR: Timber Press.

843 Locklear J.H. 2011b. Phlox ovata L. (Polemoniaceae): Clarification of the Nomenclature of the
844 Allegheny Phlox. Castanea. 76:116–117.

845 Marques D.A., Meier J.I., Seehausen O. 2019. A combinatorial view on speciation and adaptive
846 radiation. Trends Ecol. Evol. 34:531–544.

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

849 Minh B.Q., Nguyen M.A.T., von Haeseler A. 2013. Ultrafast approximation for phylogenetic
850 bootstrap. Mol. Biol. Evol. 30:1188–1195.

851 Nguyen L.-T., Schmidt H.A., von Haeseler A., Minh B.Q. 2015. IQ-TREE: a fast and effective
852 stochastic algorithm for estimating maximum-likelihood phylogenies. Mol. Biol. Evol.
853 32:268–274.

854 Noor M.A.F. 1999. Reinforcement and other consequences of sympatry. Heredity (Edinb.).
855 83:503–508.

856 Pease J.B., Haak D.C., Hahn M.W., Moyle L.C. 2016. Phylogenomics Reveals Three Sources of
857 Adaptive Variation during a Rapid Radiation. PLoS Biol. 14:e1002379.

858 Pickrell J.K., Pritchard J.K. 2012. Inference of population splits and mixtures from genome-wide
859 allele frequency data. PLoS Genet. 8:e1002967.

860 Poelstra J., Montero B.K., Lüdemann J., Yang Z., Rakotondranary S.J., Hohenlohe P., Stetter N.,
861 Ganzhorn J.U., Yoder A.D. 2021. RADseq data reveal a lack of admixture in a mouse
862 lemur contact zone contrary to previous microsatellite results. bioRxiv.

863 Roda F., Mendes F.K., Hahn M.W., Hopkins R. 2017. Genomic evidence of gene flow during
864 reinforcement in Texas Phlox. Mol. Ecol. 26:2317–2330.

865 Rothfels C.J. 2021. Polyploid phylogenetics. New Phytol. 230:66–72.

866 Schumer M., Rosenthal G.G., Andolfatto P. 2014. How common is homoploid hybrid
867 speciation? Evolution. 68:1553–1560.

868 Schumer M., Rosenthal G.G., Andolfatto P. 2018. What do we mean when we talk about hybrid
869 speciation? Heredity (Edinb.). 120:379–382.

870 Servedio M.R., Kirkpatrick M. 1997. The effects of gene flow on reinforcement. Evolution.
871 51:1764.

872 Servedio M.R., Noor M.A.F. 2003. The role of reinforcement in speciation: Theory and data.
873 Annu. Rev. Ecol. Evol. Syst. 34:339–364.

874 Smith D.M., Levin D.A. 1967. Karyotypes of eastern north American phlox. Am. J. Bot. 54:324.

875 Soltis P.S., Soltis D.E. 2009. The role of hybridization in plant speciation. Annu. Rev. Plant Biol.
876 60:561–588.

877 Stebbins G.L. 1959. The Role of Hybridization in Evolution. Proc. Am. Philos. Soc. 103:231–
878 251.

879 Suissa J.S., Kinosian S.P., Schafran P.W., Bolin J.F., Taylor W.C., Zimmer E.A. 2022.
880 Homoploid hybrids, allopolyploids, and high ploidy levels characterize the evolutionary
881 history of a western North American quillwort (Isoëtes) complex. Mol. Phylogenet. Evol.
882 166:107332.

883 Suni S.S., Hopkins R. 2018. The relationship between post-mating reproductive isolation and
884 reinforcement in Phlox. Evolution.

885 Swofford D.L. 2002. PAUP*. Phylogenetic analysis using parsimony (* and other methods).
886 Vol. Sinauer Associates, Sunderland, MA.

887 Taylor S.A., Larson E.L. 2019. Insights from genomes into the evolutionary importance and
888 prevalence of hybridization in nature. Nat. Ecol. Evol. 3:170–177.

889 Turissini D.A., Matute D.R. 2017. Fine scale mapping of genomic introgressions within the
890 Drosophila yakuba clade. PLoS Genet. 13:e1006971.

891 Turner B.L. 1998. Atlas of the Texas species of Phlox (Polemoniaceae). Phytologia. 85:309–326.

892 Wang N., Kelly L.J., McAllister H.A., Zohren J., Buggs R.J.A. 2021. Resolving phylogeny and
893 polyploid parentage using genus-wide genome-wide sequence data from birch trees. Mol.
894 Phylogenet. Evol. 160:107126.

895 Wherry E.T. 1930. The Eastern Short-styled Phloxes. Bartonia.:24–53.

896 Wherry E.T. 1931. The Eastern Long-styled Phloxes, part i. Bartonia.:18–37.

897 Wherry E.T. 1932. The Eastern Long-styled Phloxes, part 2. Bartonia.:14–26.

898 Wherry E.T. 1955. The genus Phlox, Monograph III. Morris Arbor, Philadelphia, USA.

899 Wherry E.T. 1956. The Genus Phlox. Am. Midl. Nat. 56:508.

900 Wherry E.T. 1965. The Genus Phlox, Ten Years After. Bartonia.:13–16.

901 Whitehouse E. 1945. Annual Phlox Species. Am. Midl. Nat. 34:388.

902 Yakimowski S.B., Rieseberg L.H. 2014. The role of homoploid hybridization in evolution: a

903 century of studies synthesizing genetics and ecology. Am. J. Bot. 101:1247–1258.