Title: Phylogenomic analyses re-examine the evolution of reinforcement and hypothesized hybrid speciation in *Phlox* wildflowers

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

Affiliations:

[1] *Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138, USA*

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

[*] Authors to whom correspondence should be addressed:

Robin Hopkins; email: rhopkins@fas.harvard.edu, phone: 617-496-9613

Word Count:

Summary: 200

Main Text: Total: 6,498

Introduction: 1,290 (20%)

Methods: 1,739 (27%)

Results: 1,570 (24%)

Discussion: 1,897 (29%)

In-Text Figures: N=5, Fig.1-Fig.5 are in color.

In-Text Tables: N=3

Supporting Information: Figures S1-S10; Table S1-S4

SUMMARY

- The tree of life is riddled with reticulate evolutionary histories, and some clades, such as the eastern standing *Phlox*, appear to be hotspots of hybridization. In this group there are two cases of reinforcement and nine hypothesized hybrid species. Given their historical importance in our understanding of plant speciation, the relationships of these taxa and the role of hybridization in their diversification require genomic validation.
- Using phylogenomic analyses, we resolve the evolutionary relationships of the eastern standing *Phlox* and interrogate hypotheses about if and how hybridization and gene flow played a role in their diversification.
- Our results provide novel resolution of the phylogenetic relationships in this group, including paraphyly across some taxa. We identify gene flow during one case of reinforcement, and find genomic support for a hybrid lineage underlying one of the five hypothesized homoploid hybrid speciation events. Additionally, we estimate the ancestries of four allotetraploid hybrid species.
- Our results are consistent with hybridization contributing to diverse evolutionary outcomes within this group; although, not as extensively as previously hypothesized. This study demonstrates the importance of phylogenomics in evaluating hypothesized evolutionary histories of non-model systems and adds to the growing support of interspecific genetic exchange in the generation of biodiversity.

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

INTRODUCTION

Hybridization and gene flow can play important and diverse roles during the formation and divergence of species (Taylor & Larson, 2019). While these processes are often viewed as homogenizing forces, they can also generate novel variation. For example, gene flow can create stable hybrid species and/or fuel 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 the evolution of reproductive trait divergence to reduce mating between species, i.e. reinforcement (Howard, 1993; Garner *et al.*, 2018). The effects of hybridization and gene flow on speciation have been investigated across the tree of life, yet we lack a broader understanding of the frequency and extent to which these forces impact speciation across clades (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 propose some species arose from interspecific mating (Goulet *et al.*, 2017). In particular, observations of populations with combinations of intermediate, recombinant, and novel traits to those of nearby geographically overlapping species inspired hypotheses of polyploid and homoploid hybrid speciation (Gottlieb, 1972; Soltis & Soltis, 2009; Yakimowski & Rieseberg, 2014; Barker *et al.*, 2016). It is important to return to these hypothesized cases of hybrid speciation with genomic tests for gene flow to evaluate their hypothesized origins and further understand the context and dynamics of their evolution (Goulet *et al.*, 2017).

Similarly, patterns of trait divergence in sympatry, where species coexist, but not in allopatry, motivated hypotheses of hybridization causing reproductive trait divergence by reinforcement (Hopkins, 2013). Yet, only a few empirical studies have evaluated the presence and extent of gene flow alongside the reinforcement's evolution (Kulathinal *et al.*, 2009; Roda *et al.*, 2017; Turissini & Matute, 2017; Lemmon & Juenger, 2017; Dyer *et al.*, 2018)**.** Maladaptive hybridization can drive selection for reinforcement and the divergence of plant reproductive traits, such as flower color and cross-compatibility (McNeilly & Antonovics, 1968; Levin, 1985;

Fishman & Wyatt, 1999). However, gene flow and recombination can also impede divergence by reinforcement (Felsenstein, 1981). Much debate and theory has focused on if and how reinforcement can evolve if hybridization also causes interspecific gene flow (Servedio & Noor, 2003), but further genomic study of cases of reinforcement are needed to understand if and how this reproductive trait divergence has evolved with gene flow.

Hybridization and gene flow result in genetic mixing between species. If these processes coincide with the formation of novel lineages, admixture of parental species' genetic variation will be present in the new lineage. Traditional phylogenetic methods infer evolution along bifurcating trees. However, modern phylogenomic analyses can use discordance in allele patterns under these trees (Green *et al.*, 2010; Durand *et al.*, 2011) or model allele frequencies at the tips (Gutenkunst *et al.*, 2009; Pickrell & Pritchard, 2012; Excoffier *et al.*, 2013) to identify a history of gene flow between taxa (reviewed in Hibbins & Hahn, 2021). Contemporary genome-wide sequencing methods, such as restriction site-associated DNA (RAD) sequencing (Andrews *et al.*, 2016), allow cost-efficient application of these methods to infer the evolutionary histories of genetic exchange in non-model systems (Eaton & Ree, 2013; Eaton *et al.*, 2015; Léveillé-Bourret *et al.*, 2020; Bombonato *et al.*, 2020; Guo *et al.*, 2021; Suissa *et al.*, 2022).

The eastern standing *Phlox* (Polemoniaceae) wildflowers are an ideal system for exploring the role of hybridization and gene flow in speciation. They are a monophyletic group (Landis *et al.*, 2018) distinct from the rest of *Phlox* by their upright growth habit and natural occurrence across North America. Their natural ranges overlap and intertwine, generating numerous zones of species contact (Fig. 1), and reproductive barriers between these taxa are permissive (Levin, 1966b). Decades of morphological, ecological, and biochemical evidence have suggested hybridization and gene flow between taxa, leading to hypotheses of their diversification by hybrid speciation and reinforcement (Anderson & Gage, 1952; Erbe & Turner, 1962; Levin, 1966a; Levin & Smith, 1966; Levin & Kerster, 1967; Levin & Schaal, 1970b; Levin, 1985).

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) within the eastern standing *Phlox*. These hypotheses were originally based on observations 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 & Smith, 1966; Hadley & Levin, 1969). Further analysis of karyology (Smith & Levin, 1967; Levin, 1968), cross-compatibility (Levin, 1966b), biochemical profiles (Levin & Schaal, 1970b, 1972; Levy & 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 hybrid species (Levin, 1963, 1966a; Levin & Smith, 1966). However, other treatments of these putative hybrid 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).

This clade of *Phlox* also has two cases of reinforcement. During the speciation of both *P. drummondii* and *P. pilosa* subsp. *pilosa*, flower color divergence within sympatric populations decreases matings with closely related species (reviewed in Hopkins, 2013; Fig. 3). *P. drummondii* has evolved from having light-blue to dark-red flower color to prevent hybridization with sympatric light-blue colored *P. cuspidata* (Levin, 1985; Hopkins & Rausher, 2012), and *P. pilosa* subsp. *pilosa* have evolved from pink to white flower color to prevent hybridization where it co-occurs with pink colored *P. glaberrima* subsp. *interior* (Levin & Kerster, 1967). Biochemical evidence suggests gene flow occurred between sympatric *P. pilosa* subsp. *pilosa* and *P. glaberrima* subsp. *interior* (Levin & Schaal, 1972), and transcriptomic analyses have inferred a history of gene flow between sympatric *P. drummondii* and *P. cuspidata* but analyses were limited to single individuals (Roda *et al.*, 2017).

Phlox has been a taxonomically and phylogenetically difficult genus. Traditional delineations of *Phlox* taxa relied on morphological and eco-geographic characteristics, resulting in decades of describing, promoting, and demoting species, subspecies, and varieties (Whitehouse, 1945; Wherry, 1955, 1965; Erbe & Turner, 1962; Turner, 1998; Locklear, 2011a,b). 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 (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 & Jansen, 2002; Landis *et al.*, 2018; Goulet-Scott *et al.*, 2021). Resolving the phylogenetic relationship of the eastern standing *Phlox* with genomic methods will provide a framework for returning to evolutionary hypotheses about their speciation and understanding the dynamics of their diversification.

Here, we apply double-digest RAD (ddRAD) sequencing with phylogenomic approaches to resolve the evolutionary relationships of the eastern standing *Phlox*. We leverage this phylogenetic framework to test if gene flow has occurred alongside the evolution of two cases of reinforcement and investigate evidence of genomic composition and gene flow in multiple hypothesized cases of homoploid and allotetraploid hybrid speciation.

MATERIALS AND METHODS

Sampling, sequencing, and data processing

We collected wild plants from across the native ranges of the eastern standing *Phlox* from 2017-2019, and we acquired additional samples from natural populations preserved in tissue culture at the Ohio State University Ornamental Plant Germplasm Center (OPGC). Our sampling includes thirty-two described eastern standing *Phlox* taxa, and three taxa from the outgroup mat-forming *Phlox* (Fig. 1; Supporting Information Table S1). We split taxa following Locklear, 2011b, except for the Texas annuals for which we only designated at species level.

Plants were grown in the greenhouse under controlled conditions until flowering. We extracted genomic DNA from bud tissue, processed this DNA into double-digest restriction-site-associated digestion (ddRAD) sequencing libraries, and paired-end sequenced the libraries as described in Goulet-Scott *et al.*, 2021. Raw sequence reads were demultiplexed, filtered, and *de novo* assembled as described in Goulet-Scott *et al.*, 2021, resulting in a final dataset of 116 ingroup and 4 outgroup individuals.

For all 120 samples, filtered reads (average of 3,060,421 reads per individual) were processed by iPyRAD v.0.9.50 (Eaton & Overcast, 2020) to construct four data matrices 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 data matrices contain 165,549 to 53,871 loci and 1,884,423 to 105,363 SNP sites with a locus missingness rate of 90%, 83%, 72%, and 60% respectively.

169

170 *Phylogenetic inferences*

171 Phylogenetic relationships of our samples were inferred using concatenation and
172 coalescent-based methods. First, ddRAD loci from the MS10, MS20 and MS30 datasets were
173 individually concatenated into supermatrices. Maximum likelihood (ML) phylogenies were
174 estimated from each supermatrix using IQ-TREE v.1.6.10 (Nguyen *et al.*, 2015) with the GTR +
175 gamma nucleotide substitution model and 1,000 bootstrap replicates using ultrafast bootstrap
176 approximation (Minh *et al.*, 2013; Hoang *et al.*, 2018). For the MS30 dataset, we estimated site
177 concordance factors, or the level of concordance between the inferred tree and quartets at
178 individual sites, as implemented in IQ-Tree (Minh *et al.*, 2020).

179 Second, we inferred a species-level phylogeny using a coalescent model in the program
180 SVDquartets (Chifman & Kubatko, 2014), as implemented by PAUP v.4.0a166 (Swofford,
181 2002). Analyses were performed using the MS30 dataset, with all individuals grouped by taxa
182 and omitting the mat-forming outgroup. For the SNP-site with the least amount of missing data
183 from each locus, totaling 5,306 variant sites, we sampled all possible quartets with 100 non-
184 parametric bootstrap replicates to generate a majority-rule consensus species tree. To consider
185 the effects of introgression on the inferred relationships, this analysis was repeated, dropping all
186 allopolyploid hybrids and individuals identified to have a signal of gene flow using D-statistics
187 in analyses described below.

188 Support values between our concatenation and coalescent-based inference methods are
189 derived from different estimation methods, standard bootstrap support (SBS) and ultrafast
190 bootstrap support (UBS) respectively, and cannot be directly compared (Minh *et al.*, 2013;
191 Chifman & Kubatko, 2014). SBS underestimates the likelihood of relationships (SBS ≥ 70% has
192 a 95% probability of being correct), while UBS are unbiased when >70% (UBS = 95% has a
193 0.95% probability of being correct)(Hillis & Bull, 1993; Minh *et al.*, 2013).

194

195 *Tests for gene flow alongside reinforcement*

196 We tested for evidence of gene flow alongside the evolution of reinforcement in *P.*
197 *drummondii* and *P. pilosa* subsp. *pilosa*. First, we computed Patterson's D-statistic (Durand *et*
198 *al.*, 2011) to infer if gene flow occurred between two lineages in a given phylogeny. This test
199 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.

We defined the P1, P2, P3, and O lineages as 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 also considered two different species as the P2 lineage in the hypothesis concerning *P. pilosa* subsp. *pilosa* North. In each test, only the loci present across all four taxa were used to quantify ABBA and BABA patterns. 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 statistically significant.

To complement our D-statistic tests, a history of gene flow between the focal taxa in each reinforcement scenario was also evaluated by modeling subtrees of our phylogenetic sampling using TreeMix v1.13 (Pickrell & Pritchard, 2012), as implemented in iPyRAD. We first generated two subsets of the MS4 data matrix for each reinforcement scenario, referred to as the "*P. pilosa*" and "*P. drummondii*" submatrices respectively. These submatrices include all individuals from all non-hybrid lineages in the clades containing the focal taxa, *P. drummondii* and *P. cuspidata* or *P. pilsoa* subsp. *pilosa* North and *P. glaberrima* subsp. *interior*, all intermediate clades, and from an outgroup to these lineages. In iPyRAD, individuals were grouped based on their taxon identity, with individuals of *P. drummondii* and *P. pilosa* subsp.

pilosa North split based on flower color phenotype, and one unlinked SNP-site present in at least 20% of individuals within a species was sampled from each locus in the submatix. The *P. drummondii* (4,638 unlinked SNPs) and *P. Pilosa* submatrices (3,383 unlinked SNPs) were analyzed in TreeMix to estimate subtree models 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 assess if there is genetic support for the five homoploid hybrid speciation hypotheses in this group (Table 1), we leveraged our phylogenetic inferences and evidence of gene flow to evaluate four expectations of genomic variation under a scenario of hybrid speciation: 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 that meet the first expectation, 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. For these analyses, the "*P. maculata* subsp. *pyramidalis"*, "*P. amoena* subsp. *lighthipei"*, and "*P. pilosa* subsp. *detonsa"* submatrices analyzed in TreeMix included 4,490, 4,294, and 1,033 unlinked SNPs respectively. We used the "*P. pilosa"* submatrix 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, previously demonstrated in birch trees (Wang *et al.*, 2021). Sequenced representatives of these taxa were previously confirmed to be tetraploid via flow cytometry at the OPGC.

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-mem (Li, 2013). We tallied the proportion of primary read alignments with a mapping quality ≥ 5 for the forward and reverse reads. We reasoned the diploid consensus receiving the highest proportion of quality read alignments would be most closely related to the ancestor of the mapped reads (Wang *et al.*, 2021).

We tested this assumption by mapping one representative from each diploid taxon to the consensus sequences. We observed the consensus receiving the highest relative proportion of mapped reads corresponding to the identity of the diploid individual the read came from (Supporting Information Table S4). This method allows identification of small subclades of diploid taxa to which the reads are most similar, and we inferred the ancestry of the allopolyploid subgenomes as the diploid subclades receiving the highest rates of read mapping (Wang *et al.*, 2021).

Finally, reticulate evolutionary relationships resulting from allopolyploidy are not well represented by a bifurcating tree. Therefore, we performed a NeighborNet analysis in SplitsTree4 to reconstruct the possible network-like evolutionary relationships among taxa across the whole group (Huson & Bryant, 2006). Our NeighbourNet analysis was performed on all informative SNPs from the MS30 dataset, with 1,000 replicate bootstrap iterations.

RESULTS

Phylogenomic inferences of evolutionary relationships

Maximum likelihood (ML) trees inferred from our concatenated datasets were highly congruent, despite varying missingness (Fig. 2; Supporting Information Fig. S1, S2). Therefore, we focus our results to trends in interspecific relationships observed across ML inferences.

The base of the eastern standing *Phlox* is notably composed of long-styled taxa. We recover an early split between *P. stolonifera* and all other taxa (UFB=100), followed by a split of a well-supported clade (UFB=100) containing *P. ovata* out to *P. pulchra, P. glaberrima* subsp. *triflora,* and the allotetraploid *P. buckleyi* (UFB=100). The remaining *P. glaberrima* subspecies 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 clade including *P. amplifolia* and *P. paniculata* (UFB=100).

Across all datasets, we inferred all short-styled taxa form a monophyletic group evolving from the long-styled taxa (UFB=100). This short-styled group contains the *P. pilosa*, *P. divaricata, P. villosissima*, and *P. amoena* subspecies, *P. floridana*, 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*, *P. argillacea*, and *P. pilosa* subsp. *deamii*, form a clade out to the remaining short-style species (UFB=100). Given the relative northern origin of these samples, we call this the *P. pilosa* "North" group. This group also contains the white flower morph of *P. pilosa* subsp. *pilosa* that evolved by reinforcement, sometimes described as *P. argillacea*. Both *P. divaricata* and *P. amoena* species form monophyletic groups (UFB=100), although *P. divaricata* subspecies are interdigitated.

The remaining *P. pilosa* subspecies, Texas annuals, and the allotetraploids *P. floridana, P. villosissima* subsp. *villosissma* and *P. villosissima* subsp. *latisepala* form a large clade (UFB=100). We observe reduced branch supports along the backbone of this clade; however, the Texas annual species are monophyletic (UFB=100), with *P. roemeriana* sister to *P. drummondii.* Adjacent to the annuals lie the remaining *P. pilosa* subspecies. We refer to this *P. pilosa* group as "South", given the more southern origin of these taxa relative to the *P. pilosa* North group.

Our low support partitioning the *P. pilosa* South group from the Texas annuals may be due to the presence of the allotetraploid *P. villosissima* subspecies. These taxa were inferred either out to the annuals or out to the *P. pilosa* South group with variable support (UFB=41-100). Variation in their placement is consistent with their hypothesized hybrid ancestry between *P.*

322 *pilosa* subsp. *pilosa* South and *P. drummondii* (Table 1). Similar phylogenetic uncertainty is
323 observed for the allotetraploid *P. floridana* across analyses. This may also be a result of its
324 polyploid hybrid ancestry (Table 1). Notably, the majority site patterns from our site
325 concordance factor analysis are concordant with the MS30 ML tree relationships leading up to
326 clades, with exception at nodes flanking splits involving allopolyploid lineages between the *P.*
327 *amoena* and the Texas annuals groups (Fig 2).
328 Coalescent-based species tree inference using SVDquartets recovered species trees with
329 generally similar relationships between the major species groupings observed in our ML inferred
330 trees (Supporting Information Fig. S3). Our "full data" and "no hybrids" inferred trees differ
331 primarily in their ability to resolve the relationships between the early diverging long-style
332 species and the *P. glaberrima*/*P.maculata/P.carolina* subspecies. While the "full data" tree
333 provides weak support that these two groups are distinct (SBS=75), as seen in the ML analyses,
334 the no hybrids inference weakens this confidence (SBS=68). However, all datasets reiterate
335 support for *P. amplifolia* and *P. paniculata* forming a monophyletic branch (SBS=100)
336 subtending a monophyletic group containing all short-styled taxa (SBS=93-94). We recovered
337 support for the *P. pilosa* North and *P. pilosa* South groups as being polyphyletic, and we
338 clarified the relationship of the Texas annuals and *P. pilosa* South groups as sister, when
339 omitting hybrids (SBS=100).
340
341 *Evidence of gene flow alongside reinforcement*
342 We applied D-statistics and TreeMix v.1.13 to infer if gene flow occurred between
343 sympatric populations of focal taxa in two cases of reinforcement. All D-statistic tests were
344 structured so a negative D supports a history of introgression between these focal taxa.
345 In *P. drummondii*, reinforcement against hybridization with *P. cuspidata* has favored the
346 divergence of light-blue to dark-red flower color in sympatry (Fig. 3a)(Levin, 1985; Hopkins &
347 Rausher, 2012). D-statistic tests between allopatric *P. drummondii* and *P. cuspidata* show no
348 evidence of gene flow; however, all 48 tests between sympatric dark-red *P. drummondii* and
349 every *P. cuspidata* were negative, with 24 being significant (Table 2). Our best fit *P.*
350 *drummondii* subtree modeled in TreeMix (m=2) recovered *P. cuspidata* out to *P. drummondii*
351 and *P. roemeriana,* with a low-weight migration event (0.11) going from *P. cuspidata* into dark-
352 red sympatric *P. drummondii* (Fig. 3c). This migration edge indicates 11% of alleles in the

353 sympatric *P. drummondii* population are from the *P. cuspidata* lineage. This migration edge
354 remained present for all models m ≥2 (Supporting Information Fig. S4, S9).
355 In *P. pilosa* subsp. *pilosa* North, reinforcement against hybridization with *P. glaberrima*
356 subsp. *interior* has favored the divergence of flower color from pink to white (Fig. 3b)(Levin &
357 Kerster, 1967). Unlike reinforcement in *P. drummondii*, the ranges of *P. pilosa* subsp. *pilosa*
358 North and *P. glaberrima* subsp. *interior* are mosaic and are not distinctly divisible as
359 geographically allopatric or sympatric (Levin & Kerster, 1967). Therefore, we grouped our tests
360 for gene flow based on the flower color of the *P. pilosa* subsp. *pilosa* North individual used, pink
361 or white, instead of by geography. Our tests do not support gene flow between pink (N=360) nor
362 white (N=180) flowered *P. pilosa* subsp. *pilosa* North and *P. glaberrima interior* (Table 2).
363 TreeMix modeled *P. pilosa* subtrees produced similar topologies to our ML phylogenetic
364 inferences but did not infer migration events involving *P. pilosa* subsp. *pilosa* North (Fig. 3d;
365 Supporting Information Fig. S5, S9).
366
367 *Evidence of divergence and gene flow underlying hypotheses of putative homoploid hybrid*
368 *species*
369 We leveraged our inferred phylogenies to evaluate if evidence of gene flow is consistent
370 with the formation of a hybrid lineage under the five hypothesized cases of homoploid hybrid
371 speciation in the eastern standing *Phlox* (Table 1). First, we asked if a putative hybrid species sits
372 nested within or sister to one of its hypothesized parental species and we find support for four of
373 the five putative hybrids showing this relationship (Fig. 2, Supporting Information Fig. S1, S2,
374 S3); *P. amoena* subsp. *lighthipei* within *P. amoena* subsp. *amoena*, *P. maculata* subsp.
375 *pyramidalis* within *P. maculata* subsp. *maculata*, *P. argillacea* within *P. pilosa* subsp. *pilosa*
376 North, and *P. pilosa* subsp. *detonsa* in *P. pilosa* subsp. *pilosa* South. *P. pilosa* subsp. *deamii* was
377 not found out or within *P. pilosa* subsp. *pilosa* South or *P. amoena* subsp. *amoena*, consistent
378 with (Goulet-Scott *et al.*, 2021).
379 Second, we ask if there is evidence of gene flow with the other, non-sister, lineage. For *P.*
380 *amoena* subsp. *lighthipei, P. argillacea, and P. pilosa* subsp. *detonsa,* D-statistic tests and
381 TreeMix analyses did not find support for gene flow between the putative hybrids and their
382 second hypothesized parental species (Table 3; Supporting Information Fig. S5, S6, S7,
383 S9). However, D-statistic tests do support gene flow between *P. glaberrima* subsp. *interior* and

384 *P. maculata* subsp. *pyramidalis,* with 63 of 70 tests being significantly negative. Importantly, no
385 evidence of gene flow was found between *P. glaberrima* subsp. *interior* and *P. maculta* subsp.
386 *maculata,* N=160 (Table 3). All *P. maculata* subsp. *pyramidalis* subtrees modeled in TreeMix
387 confirmed this result, with a low-weight migration edge from *P. glaberrima* subsp. *interior* into
388 *P. maculta* subsp. *pyramidalis* (Fig. 4, Supporting Information Fig. S8, S9). In our best-fit model
389 (m=4), this migration edge has a weight of 0.189, indicating 18.9% of alleles in the *P. maculata*
390 subsp. *pyramidalis* lineage are from *P. glaberrima* subsp. *interior.*
391
392 *Inferring origins of allotetraploid hybrid species*
393 We leveraged comparative read mapping to infer the subgenome ancestries of four
394 allotetraploid *Phlox* species (Table 1) (Levin, 1966b, 1968; Smith & Levin, 1967). Concatenated
395 ddRAD loci were used to generate 439,271 bp long reduced consensus sequences for each of the
396 28 diploid species, with a locus missingness rate of 25% among them. Mapping of reads to this
397 consensus dataset resulted in differential assignment of sequence ancestry for each tetraploid
398 sample to subclades of diploid taxa (Fig. 5).
399 Reads from *P. villosissima* subsp. *villosissima* and *P. villosissima* subsp. *latisepala*
400 mapped with high frequency to the *P. pilosa* South group and the Texas annuals, each receiving
401 30-34% and 22-26% respectively. Within these two groups the top recipients of mapped reads
402 were *P. pilosa* subsp. *pilosa* South and *P. cuspidata* for all three samples. Our NeighbourNet
403 analysis mirrors this result (Supporting Information Fig. S10), placing the polyploid lineages
404 between the ancestors of the *P. pilosa* South clade and the Texas annuals.
405 *P. floridana* showed a high mapping rate to the *P. pilosa* South group (34%), with most
406 reads mapping to *P. pilosa* subsp. *pilosa* South and second most to *P. pilosa* subsp. *detonsa*.
407 Contrary to the hypothesized origins of this species, we did not observe elevated read mapping to
408 the *P. carolina subspecies* consensus sequences. Instead, we observed elevated read mapping to
409 *P. amoena* subsp. *amoena*. Again, the NeighbourNet analysis also reflects our mapping results,
410 showing *P. floridana* intermediate partitioning between the *P. pilosa* South and *P. amoena*
411 lineages.
412 Finally, *P. buckleyi* reads mapped with the highest rate to the *P. ovata*, *P. pulchra*, *P.*
413 *glaberrima* subsp. *triflora* clade (38%), with the NeighborNet analysis showing *P. buckleyi*
414 being of intermediate placement to *P. ovata* and *P. glaberrima* subsp. *triflora*.

DISCUSSION

With genomic data and modern phylogenomic analyses, evolutionary biology can understand the presence, extent, and consequences of hybridization and gene flow across large clades of organisms. Here, we present well-resolved phylogenetic relationships of the eastern standing *Phlox*, demonstrate clear support for most described species relationships, and reveal novel non-monophyletic relationships of subspecies in historically taxonomically difficult species complexes. Using this phylogenetic framework, our analyses identify evidence of a hybrid lineage consistent with one case of hypothesized homoploid hybrid speciation, identify putative ancestries of multiple polyploid species, and find evidence of gene flow in one case of reinforcement. However, we also find no support for many hypothesized cases of gene flow in the formation of new lineages. 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 hybridization and gene flow across species boundaries does play diverse roles in generating novel biodiversity.

Evolutionary relationships of the eastern standing Phlox

Previous phylogenetic inference of the eastern standing *Phlox* either relied on a handful of genetic markers or limited taxonomic sampling, leaving most of the evolutionary relationships of this group unclear (Ferguson *et al.*, 1999; Ferguson & Jansen, 2002; Roda *et al.*, 2017; Landis *et al.*, 2018; Goulet-Scott *et al.*, 2021). Our phylogenomic inferences on genome-wide ddRAD markers from a broad taxonomic sampling clarify the evolutionary relationships of the eastern standing *Phlox* and provide novel insight into their diversification (Fig. 2; Supporting Information Fig. S1-S3).

Taxonomic treatments of the eastern standing *Phlox* have grouped these taxa by differences in style length (Wherry, 1930, 1931, 1932, 1955). Unlike prior phylogenetic studies, our inferences support a large monophyletic clade of short-styled taxa subtended by multiple paraphyletic clades of long-styled taxa. Divergence in style-length can generate reproductive isolation (Kay, 2006; Brothers & Delph, 2017), and future work should investigate if and how the transition in style length is involved in the diversification of the short-styled taxa.

Our analyses also reveal unprecedented resolution into relationships within species groups. Our inferences support *P. drummondii* and *P. roemeriana* are sister taxa within the Texas annuals (Fig. 2). This topology is supported in Roda *et al.*, 2017 but not supported in studies with few genetic loci (Ferguson *et al.*, 1999; Ferguson & Jansen, 2002; Landis *et al.*, 2018). In the *P. glaberrima*/*P. carolina*/*P. maculata* complex, we find support for non-monophyletic relationships among subspecies (Fig. 2). Additionally, the subspecies of the *P. pilosa* complex form distinct polyphyletic groups coinciding with their relative northern or southern location of origin, as suggested in Goulet-Scott *et al.*, 2021 (Fig. 2). The one exception is *P. pilosa* subsp. *sangamonensis* which sits in the southern clade but is found in the north (Fig. 1), a discovery consistent with the hypothesis of *P. pilosa* subsp. *sangamonensis* arising by long-distance dispersal from southern populations of *P. pilosa* subsp. *pilosa* (Levin & Smith, 1965; Levin, 1984). The non-monophyletic relationship among the northern and southern *P. pilosa* lineages and within *P. glaberrima* were suggested in previous phylogenetic studies, but remained uncertain due to low phylogenetic supports and reliance on few genetic loci (Ferguson *et al.*, 1999; Ferguson & Jansen, 2002; Landis *et al.*, 2018). Only with phylogenomics are we able to definitively show support for these relationships. Classic taxonomic delineations within *Phlox* are largely based on morphological characteristics (Wherry, 1955; Locklear, 2011b), yet our phylogenetic inferences demonstrate current taxonomy may not best reflect the true evolutionary relationships of some taxa within this group.

Evolution of reinforcement with and without gene flow

Limited empirical study has evaluated if interspecific hybridization generating selection for reinforcement also results in reinforcement evolving in the face of gene flow (Garner *et al.*, 2018). We investigated if gene flow accompanied two cases of divergence by reinforcement in *Phlox* (Fig. 3).

P. drummondii has diverged from light-blue to dark-red flower color to prevent hybridization with *P. cuspidata* (Levin, 1985; Hopkins & Rausher, 2012). Our sampling revealed evidence of gene flow between *P. cuspidata* and multiple sympatric dark-red *P. drummondii* individuals but not between *P. cuspidata* and any allopatric light-blue *P. drummondii,* consistent with previous observations (Roda *et al.*, 2017). A history of gene flow from hybridization in sympatry but not in allopatry mirrors the model of phenotypic divergence by reinforcement in

476 sympatry but not in allopatry (Garner *et al.*, 2018). This pattern suggests the signal of gene flow
477 is the product of contemporary hybridization in sympatry and not due to admixture between the
478 distant ancestors of *P. cuspidata* and *P. drummondii*. Although F1 hybrids have high sterility
479 (Suni & Hopkins, 2018), our observation of a low proportion of *P. cuspidata* alleles (11%) in the
480 sympatric dark-red *P. drummondii* lineage indicate hybrids do backcross in nature, resulting in
481 interspecific gene flow. Therefore, reinforcement in *P. drummondii* evolved despite gene flow in
482 sympatry.

483 Conversely, our analyses do not support a history of gene flow coinciding with
484 reinforcement in *P. pilosa* subsp. *pilosa* North. *P. pilosa* subsp. *pilosa* North has evolved from
485 pink to white colored flowers to reduce hybridization where it co-occurs with *P. glaberrima*
486 subsp. *interior* in high frequency (Fig. 2) (Levin, 1966b; Levin & Kerster, 1967; Levin & Schaal,
487 1970a,b, 1972). Reproductive isolation between these taxa is high but field experiments and
488 experimental crosses demonstrate these species do generate hybrid seed (Levin, 1966b; Levin &
489 Kerster, 1967; Levin & Schaal, 1970a,b, 1972)(Fig. 2). However, we do not detect gene flow
490 between these species (Fig. 3; Table 2). This result suggests hybrid offspring between these taxa
491 may be completely inviable, lethally maladapted, or have exceedingly high sterility and that
492 reinforcement may have occurred after gene flow had ceased and postzygotic RI was complete
493 between the lineages.

494 The presence/absence of gene flow in these two cases of reinforcement presents a
495 powerful opportunity to investigate the dynamics by which reproductive trait divergence, and
496 flower color specifically, evolves by reinforcing selection. In both cases the *Phlox* responded to
497 reinforcing selection through the evolution of the same trait – flower color – while in one case
498 the trait evolved in the presence of gene flow and in the other case there is no gene flow between
499 sympatric species. For reproductive trait divergence to evolve by reinforcement, alleles
500 conferring increased assortative mating within a species must remain associated with alleles
501 causing costly hybridization between species (Servedio & Kirkpatrick, 1997; Kirkpatrick, 2000).
502 When gene flow between species is present, recombination can disassociate these alleles,
503 impeding successful divergence. In the two *Phlox* cases, we see similar responses to selection but
504 different potentials for gene flow to result in recombination. These contrasting scenarios could
505 allow future research to compare if and how other factors influencing the probability of
506 successful reinforcement may also differ between these two examples and contribute to our

empirical understanding of when reinforcement is likely to evolve. For example, are there differing genetic architectures and genetic linkage underlying the cost of hybridization and flower color between these cases, varying strengths of reproductive isolating barriers between the species pairs upon their secondary contact, and/or differences in the functional cost of hybridization (i.e. offspring or gamete loss)? Further study of the strength and genetic architecture of reproductive isolating barriers between these *Phlox* species will better inform how reinforcement evolves.

Varying support for homoploid and allotetraploid hybrid species hypotheses

Within the eastern standing *Phlox* there are five homoploid and four allopolyploid hybrid speciation hypotheses (Table 1). Of the homoploid hybrid speciation hypotheses, we only observed genomic evidence for one lineage, *P. maculata* subsp. *pyramidalis,* arising from the distinct lineages of its hypothesized parental taxa, *P. maculata* subsp. *maculata* and *P. glaberrima* subsp. *interior*. *P. maculata* subsp. *pyramidalis* is found phylogenetically closely related to multiple *P. maculata* subsp. *maculata* individuals (Fig. 2), with a strong signal of gene flow from *P. glaberrima* subsp. *interior* (Fig. 4; Table 3.). Our analyses also identify *P. maculata* subsp. *pyramidalis* is an advanced generation hybrid, with 18.9% of its alleles derived from *P. glaberrima* subsp. *interior.* This evidence identifies *P. maculata* subsp. *pyramidalis* resulted from hybridization, and is consistent with the long-standing homoploid hybrid speciation hypothesis of this lineage. However, while evidence of historical hybridization is necessary to infer homoploid hybrid speciation, it is not sufficient to demonstrate a stable, reciprocally reproductively isolated lineage has evolved by homoploid hybrid speciation. Future work is necessary to determine if the hybrid origin of *P. maculata* subsp. *pyramidalis* gave rise to a stable reproductive isolated lineage from its putative parental species, demonstrating it evolved by homoploid hybrid speciation (Schumer *et al.*, 2018).

Additionally, a classic interpretation of the homoploid hybrid speciation hypothesis requires the stable reproductively isolated homoploid hybrid species to arise by hybridization between two distinct species. This interpretation is rooted in the biological species concept and may bring to question whether a subspecies of one of the parental taxa, like P. maculate subsp. pyramidalis, can arise via homoploid hybrid speciation. We caution how this interpretation is applied and required across systems where the taxonomic delineation between species and

subspecies is controversial, in flux, and not based on the levels of reproductive isolation between the taxa. Quantifications of reproductive isolation between taxa involved in the *P. maculata subsp. pyramidalis* hypothesis will aid in evaluating the extent to which these taxa are distinct biological species, and fit this classic definition of homoploid hybrid speciation, despite their current taxonomy

The remaining four hypothesized cases of homoploid hybrid speciation were not supported by our genomic investigation, despite extensive phenotypic and biomolecular evidence. Three taxa are closely related to one of their hypothesized parental species, and may just be divergent population from this lineage, while the fourth taxon is distantly related to both hypothesized parents and may possess phenotypic traits resembling these species through convergence or incomplete lineage sorting.

Polyploid hybrid lineages can be significantly easier to identify than homoploid hybrids through chromosome structure and number. Despite this advantage, inferring the evolutionary origin of these lineages remains challenging (Rothfels, 2021). Although the signals from our comparative read mapping and SplitsTree analyses are coarse, our approaches suggest the identity of the progenitor clades that gave rise to these polyploid hybrid species. Our evidence supports *P. villosissima* subsp. *latisepala* and *P. villosissima* subsp. *villosissima* are derived from *P. pilosa* (*P. pilosa* subsp. *pilosa* South) and a Texas annual species. *P. drummondii* was previously hypothesized as the parent but our analyses cannot resolve which of the Texas annual species or their ancestor is the parent. As previously hypothesized, we find support for one parent of *P. floridana* to be *P. pilosa* subsp. *pilosa* South, yet we find evidence that *P. amoena* subsp. *amoena* may be the second parent instead of the *P. carolina subspecies*. We also observe *P. buckleyi* is genetically similar to the early-diverging long-styled *Phlox*, but we cannot confirm 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 been a potential and realized source 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 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. *P. pilosa*'s developmental biology may make it prone to evolutionary innovation, with many populations having high rates of unreduced gametes and whole genome duplications across their range (Worcester, Mayfield, & Ferguson, 2012). 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

Much of the data supporting the role of hybridization and gene flow in the speciation of the eastern standing *Phlox* stemmed from comparisons of trait variation between lineages found in sympatry. 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. However, our findings also indicate a lack of interspecific genetic exchange for many hypothesized cases of hybrid speciation. Further study of these confirmed cases may inform how specific traits and even genes move between species and contribute to reproductive isolation between lineages. While our lack of support in some cases suggests forces other than hybridization may underlie patterns of trait variation observed across taxa in this group, this finding also motivates a strong need to reevaluate our previous evidence for the existence of gene flow, and to rethink our expectations of how and why traits evolve within and between closely related species. Future investigation into the phenotypic variation of these taxa may provide new insights into how differences in ecological selective pressures across overlapping species ranges, effective population size, and incomplete lineage sorting contribute to phenotypic divergence within and between 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 S.V. Edwards, E.M. Kramer, J. Wakeley, and J. Mallet for thoughtful discussions that improved this study, and to C.J. Ferguson, G.A. Burgin, and A.E. Berardi 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.

AUTHOR 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 JW**, **Baird SJE**, **Bierne N**, **Boughman J**, **Brelsford A**, **Buerkle CA**, **Buggs R**, *et al.* **2013**. Hybridization and speciation. *Journal of evolutionary biology* **26**: 229–246.

Anderson E. **1949**. Introgressive hybridization. *Introgressive hybridization.*

Anderson E, **Gage A**. **1952**. Introgressive hybridization in phlox bifida. *American journal of botany* **39**: 399–404.

Andrews KR, **Good JM**, **Miller MR**, **Luikart G**, **Hohenlohe PA**. **2016**. Harnessing the power of RADseq for ecological and evolutionary genomics. *Nature reviews genetics* **17**: 81–92.

Barker MS, **Arrigo N**, **Baniaga AE**, **Li Z**, **Levin DA**. **2016**. On the relative abundance of autopolyploids and allopolyploids. *The new phytologist* **210**: 391–398.

Bombonato JR, **do Amaral DT**, **Silva GAR**, **Khan G**, **Moraes EM**, **da Silva Andrade SC**, **Eaton DAR**, **Alonso DP**, **Ribolla PEM**, **Taylor N**, *et al.* **2020**. The potential of genome-wide RAD sequences for resolving rapid radiations: a case study in Cactaceae. *Molecular phylogenetics and evolution* **151**: 106896.

Brothers AN, **Delph LF**. **2017**. Divergence in style length and pollen size leads to a postmating-prezygotic reproductive barrier among populations of Silene latifolia. *Evolution; international journal of organic evolution* **71**: 1532–1540.

Chifman J, **Kubatko L**. **2014**. Quartet inference from SNP data under the coalescent model. *Bioinformatics (Oxford, England)* **30**: 3317–3324.

Durand EY, **Patterson N**, **Reich D**, **Slatkin M**. **2011**. Testing for ancient admixture between closely related populations. *Molecular biology and evolution* **28**: 2239–2252.

Dyer KA, **Bewick ER**, **White BE**, **Bray MJ**, **Humphreys DP**. **2018**. Fine-scale geographic patterns of gene flow and reproductive character displacement in Drosophila subquinaria and Drosophila recens. *Molecular ecology*.

Eaton DAR, **Hipp AL**, **González-Rodríguez A**, **Cavender-Bares J**. **2015**. Historical introgression among the American live oaks and the comparative nature of tests for introgression. *Evolution; international journal of organic evolution* **69**: 2587–2601.

Eaton DAR, **Overcast I**. **2020**. ipyrad: Interactive assembly and analysis of RADseq datasets. *Bioinformatics (Oxford, England)* **36**: 2592–2594.

Eaton DAR, **Ree RH**. **2013**. Inferring phylogeny and introgression using RADseq data: an example from flowering plants (Pedicularis: Orobanchaceae). *Systematic biology* **62**: 689–706.

Erbe L, **Turner BL**. **1962**. A biosystematic study of the phlox cuspidata-phlox drummondii complex. *The American midland naturalist* **67**: 257.

Excoffier L, **Dupanloup I**, **Huerta-Sánchez E**, **Sousa VC**, **Foll M**. **2013**. Robust demographic inference from genomic and SNP data. *PLoS genetics* **9**: e1003905.

Fehlberg SD, **Ty MC**, **Ferguson CJ**. **2014**. Reexamination of a putative diploid hybrid taxon using genetic evidence: The distinctiveness ofPhlox pilosasubsp.deamii(polemoniaceae). *International journal of plant sciences* **175**: 781–793.

Felsenstein J. **1981**. Skepticism Towards Santa Rosalia, or Why are There so Few Kinds of Animals? *Evolution; international journal of organic evolution* **35**: 124.

Ferguson CJ, **Jansen RK**. **2002**. A chloroplast DNA phylogeny of eastern Phlox (Polemoniaceae): implications of congruence and incongruence with the ITS phylogeny. *American journal of botany* **89**: 1324–1335.

Ferguson CJ, **Kramer F**, **Jansen RK**. **1999**. Relationships of eastern north American phlox (polemoniaceae) based on ITS sequence data. *Systematic botany* **24**: 616.

Fishman L, **Wyatt R**. **1999**. Pollinator-mediated competition, reproductive character displacement, and the evolution of selfing in Arenaria uniflora (Caryophyllaceae). *Evolution; international journal of organic evolution* **53**: 1723–1733.

Garner AG, **Goulet BE**, **Farnitano MC**, **Molina-Henao YF**, **Hopkins R**. **2018**. Genomic Signatures of Reinforcement. *Genes* **9**.

Gottlieb LD. **1972**. Levels of confidence in the analysis of hybridization in plants. *Annals of the Missouri Botanical Garden. Missouri Botanical Garden* **59**: 435.

Goulet BE, **Roda F**, **Hopkins R**. **2017**. Hybridization in plants: Old ideas, new techniques. *Plant physiology* **173**: 65–78.

Goulet-Scott BE, **Garner AG**, **Hopkins R**. **2021**. Genomic analyses overturn two long-standing homoploid hybrid speciation hypotheses. *Evolution; international journal of organic evolution* **75**: 1699–1710.

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

Green RE, **Krause J**, **Briggs AW**, **Maricic T**, **Stenzel U**, **Kircher M**, **Patterson N**, **Li H**, **Zhai W**, **Fritz MH-Y**, *et al.* **2010**. A Draft Sequence of the Neandertal Genome. *Science* **328**: 710–722.

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

Gutenkunst RN, **Hernandez RD**, **Williamson SH**, **Bustamante CD**. **2009**. Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data. *PLoS genetics* **5**: e1000695.

Hadley EB, **Levin DA**. **1969**. Physiological evidence of hybridization and reticulate evolution in phlox maculata. *American journal of botany* **56**: 561–570.

Hibbins M, **Hahn MW**. **2022**. Phylogenomic approaches to detecting and characterizing introgression. *Genetics* **220**: iyab173.

Hillis DM, **Bull JJ**. **1993**. An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. *Systematic biology* **42**: 182–192.

Hoang DT, **Chernomor O**, **von Haeseler A**, **Minh BQ**, **Vinh LS**. **2018**. UFBoot2: Improving the ultrafast bootstrap approximation. *Molecular biology and evolution* **35**: 518–522.

Hopkins R. **2013**. Reinforcement in plants. *The new phytologist* **197**: 1095–1103.

Hopkins R, **Rausher MD**. **2012**. Pollinator-mediated selection on flower color allele drives reinforcement. *Science (New York, N.Y.)* **335**: 1090–1092.

Howard DJ. **1993**. Reinforcement: origin, dynamics, and fate of an evolutionary hypothesis. *Hybrid zones and the evolutionary process*: 46–69.

Huson DH, **Bryant D**. **2006**. Application of phylogenetic networks in evolutionary studies. *Molecular biology and evolution* **23**: 254–267.

Kay KM. **2006**. Reproductive isolation between two closely related hummingbird-pollinated neotropical gingers. *Evolution; international journal of organic evolution* **60**: 538–552.

Kirkpatrick M. **2000**. Reinforcement and divergence under assortative mating. *Proceedings. Biological sciences* **267**: 1649–1655.

Kulathinal RJ, **Stevison LS**, **Noor MAF**. **2009**. The genomics of speciation in Drosophila: diversity, divergence, and introgression estimated using low-coverage genome sequencing. *PLoS genetics* **5**: e1000550.

Landis JB, **Bell CD**, **Hernandez M**, **Zenil-Ferguson R**, **McCarthy EW**, **Soltis DE**, **Soltis PS**. **2018**. Evolution of floral traits and impact of reproductive mode on diversification in the phlox family (Polemoniaceae). *Molecular phylogenetics and evolution* **127**: 878–890.

Lemmon EM, **Juenger TE**. **2017**. Geographic variation in hybridization across a reinforcement contact zone of chorus frogs (Pseudacris). *Ecology and evolution* **7**: 9485–9502.

Léveillé-Bourret É, **Chen B-H**, **Garon-Labrecque M-È**, **Ford BA**, **Starr JR**. **2020**. RAD sequencing resolves the phylogeny, taxonomy and biogeography of Trichophoreae despite a recent rapid radiation (Cyperaceae). *Molecular phylogenetics and evolution* **145**: 106727.

Levin DA. **1963**. Natural Hybridization Between Phlox maculata and Phlox glaberrima and its Evolutionary Significance. *American journal of botany* **50**: 714.

Levin DA. **1966a**. Chromatographic evidence of hybridization and evolution in phlox maculata. *American journal of botany* **53**: 238–245.

Levin DA. **1966b**. The phlox pilosa complex: Crossing and chromosome relationships. *Brittonia* **18**: 142.

Levin DA. **1968**. The genome constitutions of eastern north American phlox amphiploids. *Evolution; international journal of organic evolution* **22**: 612.

Levin DA. **1969**. The challenge from a related species: A stimulus for saltational speciation. *The american naturalist* **103**: 316–322.

Levin DA. **1984**. Genetic variation and divergence in a disjunct phlox. *Evolution; international journal of organic evolution* **38**: 223–225.

Levin DA. **1985**. Reproductive Character Displacement in Phlox. *Evolution; international journal of organic evolution* **39**: 1275.

Levin DA, **Kerster HW**. **1967**. Natural selection for reproductive isolation in phlox. *Evolution; international journal of organic evolution* **21**: 679–687.

Levin DA, **Schaal BA**. **1970a**. Corolla color as an inhibitor of interspecific hybridization in phlox. *The american naturalist* **104**: 273–283.

Levin DA, **Schaal BA**. **1970b**. Reticulate evolution in phlox as seen through protein electrophoresis. *American journal of botany* **57**: 977–987.

Levin DA, **Schaal BA**. **1972**. Seed Protein Polymorphism in Phlox pilosa (Polemoniaceae). *Brittonia* **24**: 46.

Levin DA, **Smith DM**. **1965**. An enigmatic phlox from Illinois. *Brittonia* **17**: 254.

Levin DA, **Smith DM**. **1966**. Hybridization and Evolution in the Phlox pilosa Complex. *The american naturalist* **100**: 289–302.

Levy M, **Levin DA**. **1974**. Novel flavonoids and reticulate evolution in the phlox pilosa–p. Drummondii complex. *American journal of botany* **61**: 156–167.

Levy M, **Levin DA**. **1975**. The novel flavonoid chemistry and phylogenetic origin of Phlox floridana. *Evolution; international journal of organic evolution* **29**: 487–499.

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

Lindsey W, Mayfield MH, Ferguson CJ. **2012**. Cytotypic Variation in Phlox pilosa ssp. pilosa (Polemoniaceae) at the western edge of its range in the central United States. *Journal of the botanical research institute of texas* **6**: 443–451.

Locklear JH. **2011a**. Phlox ovata L. (Polemoniaceae): Clarification of the Nomenclature of the Allegheny Phlox. *Castanea* **76**: 116–117.

Locklear JH. **2011b**. *Phlox a natural history and gardener's guide*. Portland, OR: Timber Press.

Marques DA, **Meier JI**, **Seehausen O**. **2019**. A combinatorial view on speciation and adaptive radiation. *Trends in ecology & evolution* **34**: 531–544.

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

Minh BQ, **Hahn MW**, **Lanfear R**. **2020**. New methods to calculate concordance factors for phylogenomic datasets. *Molecular biology and evolution* **37**: 2727–2733.

Minh BQ, **Nguyen MAT**, **von Haeseler A**. **2013**. Ultrafast approximation for phylogenetic bootstrap. *Molecular biology and evolution* **30**: 1188–1195.

Nguyen L-T, **Schmidt HA**, **von Haeseler A**, **Minh BQ**. **2015**. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular biology and evolution* **32**: 268–274.

Pease JB, **Haak DC**, **Hahn MW**, **Moyle LC**. **2016**. Phylogenomics Reveals Three Sources of Adaptive Variation during a Rapid Radiation. *PLoS biology* **14**: e1002379.

Pickrell JK, **Pritchard JK**. **2012**. Inference of population splits and mixtures from genome-wide allele frequency data. *PLoS genetics* **8**: e1002967.

Roda F, **Mendes FK**, **Hahn MW**, **Hopkins R**. **2017**. Genomic evidence of gene flow during reinforcement in Texas Phlox. *Molecular ecology* **26**: 2317–2330.

Rothfels CJ. **2021**. Polyploid phylogenetics. *The new phytologist* **230**: 66–72.

Schumer M, **Rosenthal GG**, **Andolfatto P**. **2014**. How common is homoploid hybrid speciation? *Evolution; international journal of organic evolution* **68**: 1553–1560.

Schumer M, **Rosenthal GG**, **Andolfatto P**. **2018**. What do we mean when we talk about hybrid speciation? *Heredity* **120**: 379–382.

Servedio MR, **Kirkpatrick M**. **1997**. The effects of gene flow on reinforcement. *Evolution; international journal of organic evolution* **51**: 1764.

Servedio MR, **Noor MAF**. **2003**. The role of reinforcement in speciation: Theory and data. *Annual review of ecology, evolution, and systematics* **34**: 339–364.

Smith DM, **Levin DA**. **1967**. Karyotypes of eastern north American phlox. *American journal of botany* **54**: 324.

Soltis PS, **Soltis DE**. **2009**. The role of hybridization in plant speciation. *Annual review of plant biology* **60**: 561–588.

Stebbins GL. **1959**. The Role of Hybridization in Evolution. *Proceedings of the american philosophical society* **103**: 231–251.

Suissa JS, **Kinosian SP**, **Schafran PW**, **Bolin JF**, **Taylor WC**, **Zimmer EA**. **2022**. Homoploid hybrids, allopolyploids, and high ploidy levels characterize the evolutionary history of a western North American quillwort (Isoëtes) complex. *Molecular phylogenetics and evolution* **166**: 107332.

Suni SS, **Hopkins R**. **2018**. The relationship between post-mating reproductive isolation and reinforcement in Phlox. *Evolution; international journal of organic evolution*.

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

Taylor SA, **Larson EL**. **2019**. Insights from genomes into the evolutionary importance and prevalence of hybridization in nature. *Nature ecology & evolution* **3**: 170–177.

Turissini DA, **Matute DR**. **2017**. Fine scale mapping of genomic introgressions within the Drosophila yakuba clade. *PLoS genetics* **13**: e1006971.

Turner BL. **1998**. Atlas of the Texas species of Phlox (Polemoniaceae). *Phytologia* **85**: 309–326.

Wang, **Kelly LJ**, **McAllister HA**, **Zohren J**, **Buggs RJA**. **2021**. Resolving phylogeny and polyploid parentage using genus-wide genome-wide sequence data from birch trees. *Molecular phylogenetics and evolution* **160**: 107126.

Wherry ET. **1930**. The Eastern Short-styled Phloxes. *Bartonia; proceedings of the hiladelphia botanical club ...*: 24–53.

Wherry ET. **1931**. The Eastern Long-styled Phloxes, part i. *Bartonia; proceedings of the Philadelphia botanical club ...*: 18–37.

Wherry ET. **1932**. The Eastern Long-styled Phloxes, part 2. *Bartonia; proceedings of the Philadelphia botanical club ...*: 14–26.

Wherry ET. **1955**. The genus Phlox, Monograph III. *Morris Arbor, Philadelphia, USA*.

Wherry ET. **1956**. The Genus Phlox. *The American midland naturalist* **56**: 508.

Wherry ET. **1965**. The Genus Phlox, Ten Years After. *Bartonia; proceedings of the Philadelphia botanical club ...*: 13–16.

Whitehouse E. **1945**. Annual Phlox Species. *The American midland naturalist* **34**: 388.

Yakimowski SB, **Rieseberg LH**. **2014**. The role of homoploid hybridization in evolution: a century of studies synthesizing genetics and ecology. *American journal of botany* **101**: 1247–1258.

SUPPLEMENTARY INFORMATION:

Figure S1. Maximum likelihood phylogenetic inference from whole concatenated loci, with a locus shared among at least 10 individuals.

Figure S2. Maximum likelihood phylogenetic inference from whole concatenated loci, with a locus shared among at least 20 individuals.

Figure S3. Coalescent-based phylogenetic inference using SVDquartets, with a locus shared among at least 30 individuals.

Figure S4. TreeMix models for the *P. drummondii* subtree.

Figure S5. TreeMix models for the *P. pilosa subsp. pilosa* North subtree.

Figure S6. TreeMix models for the *P. pilosa subsp. detonsa* subtree.

Figure S7. TreeMix models for the *P. amoena subsp. lighthipei* subtree.

Figure S8. TreeMix models for the *P. maculata subsp. pyramidalis* subtree.

Figure S9. TreeMix model best fit estimation.

Figure S10. NeighborNet whole phylogeny splitsgraph.

Table S1. Species sampling.

Table S2. All D-statistic test results for introgression in two cases of reinforcement.

Table S3. All D-statistic test results for introgression in four hypothesized cases of hybrid speciation.

Table S4. Summary of comparative read mapping of diploid and allotetraploid species to diploid consensus sequences.

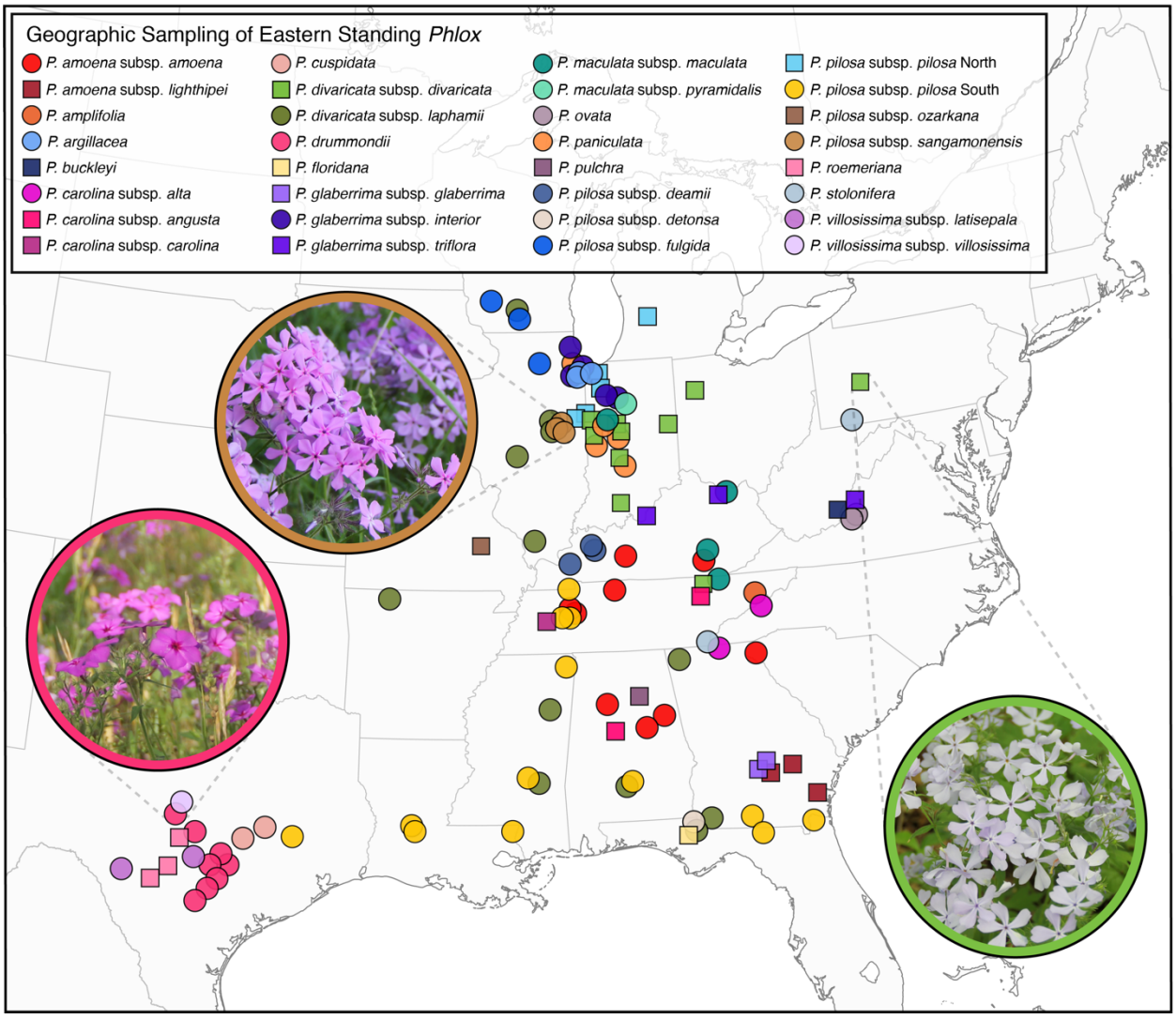


Figure 1. Map of individuals sampled across the interdigitating ranges of the thirty-two eastern standing *Phlox* taxa included for phylogenetic analyses. Each point represents the locality of a single individual with the color and shape corresponding to an individual's taxonomic identity. Three insets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

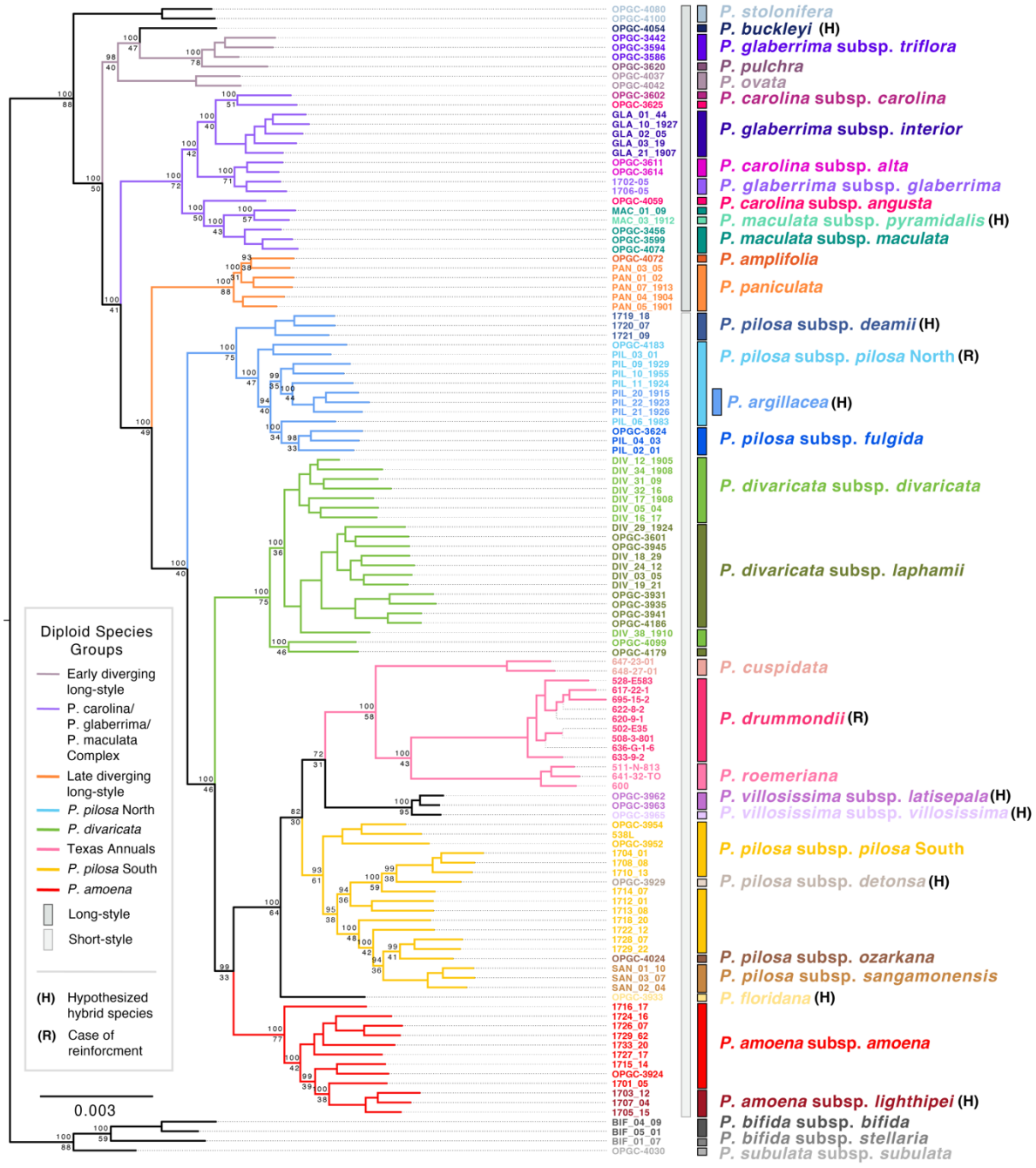


Figure 2. Maximum likelihood phylogenetic inference of eastern standing *Phlox* from whole concatenated ddRAD loci that were shared by at least 30 individuals. Nodes differentiating taxa are labeled with ultrafast bootstrap approximation support (above) and site concordance factor values (below).

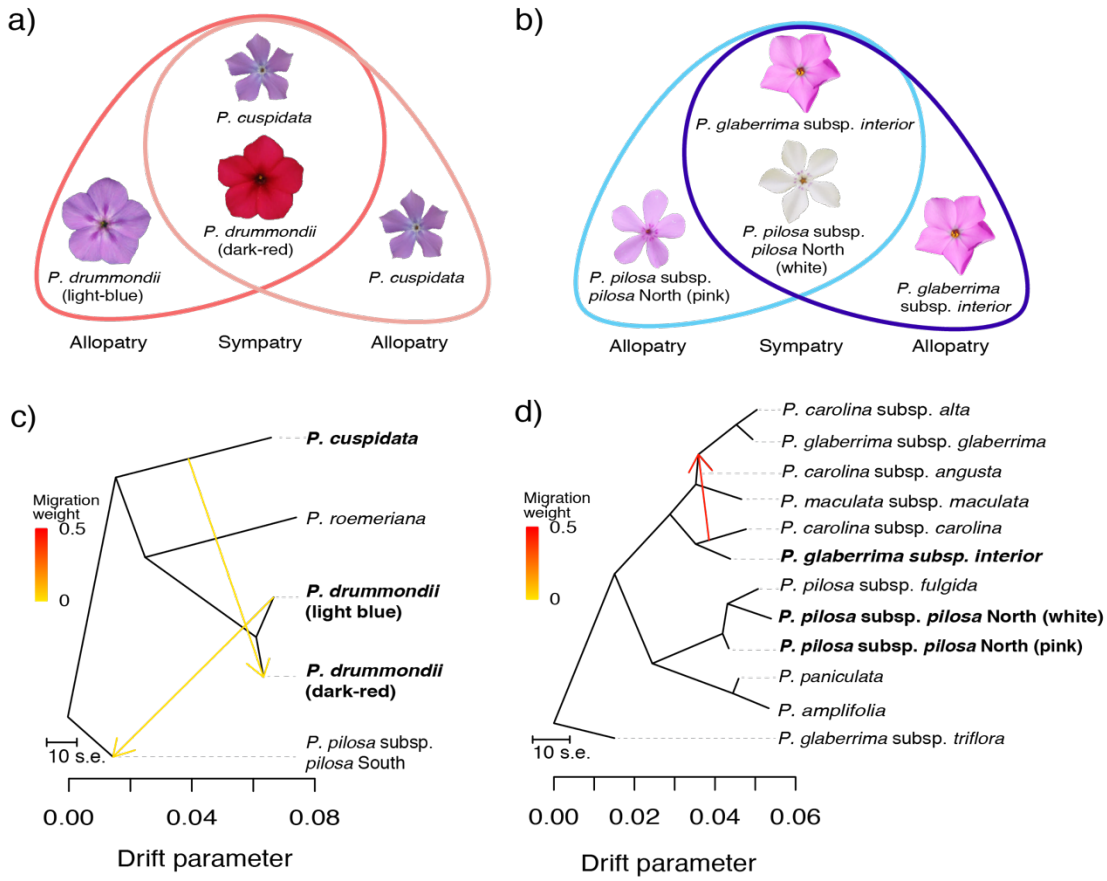
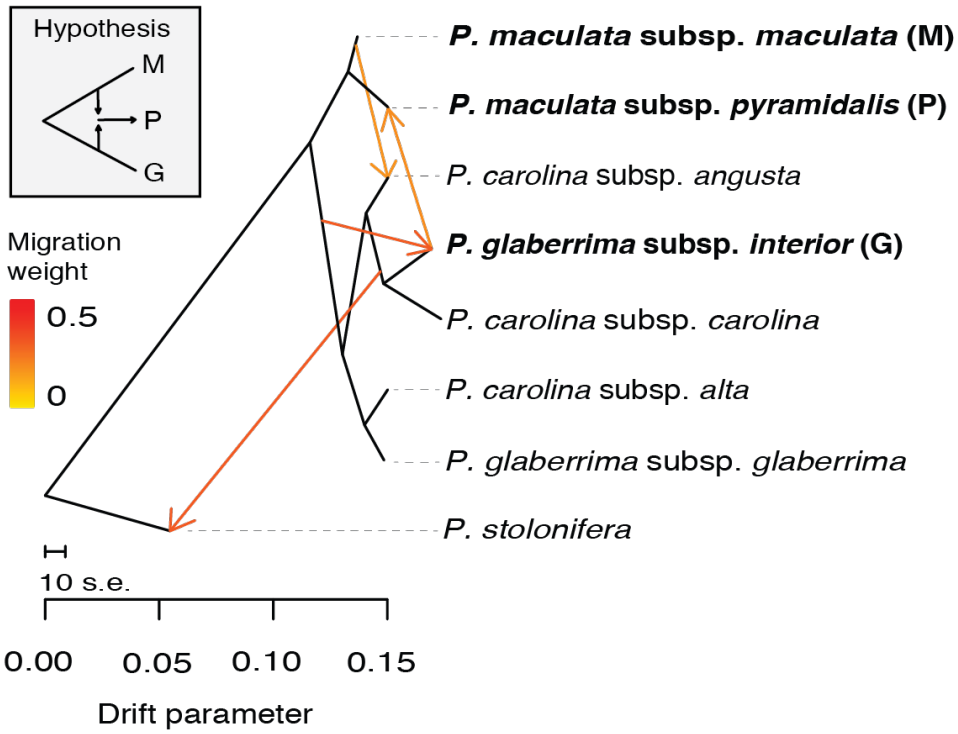
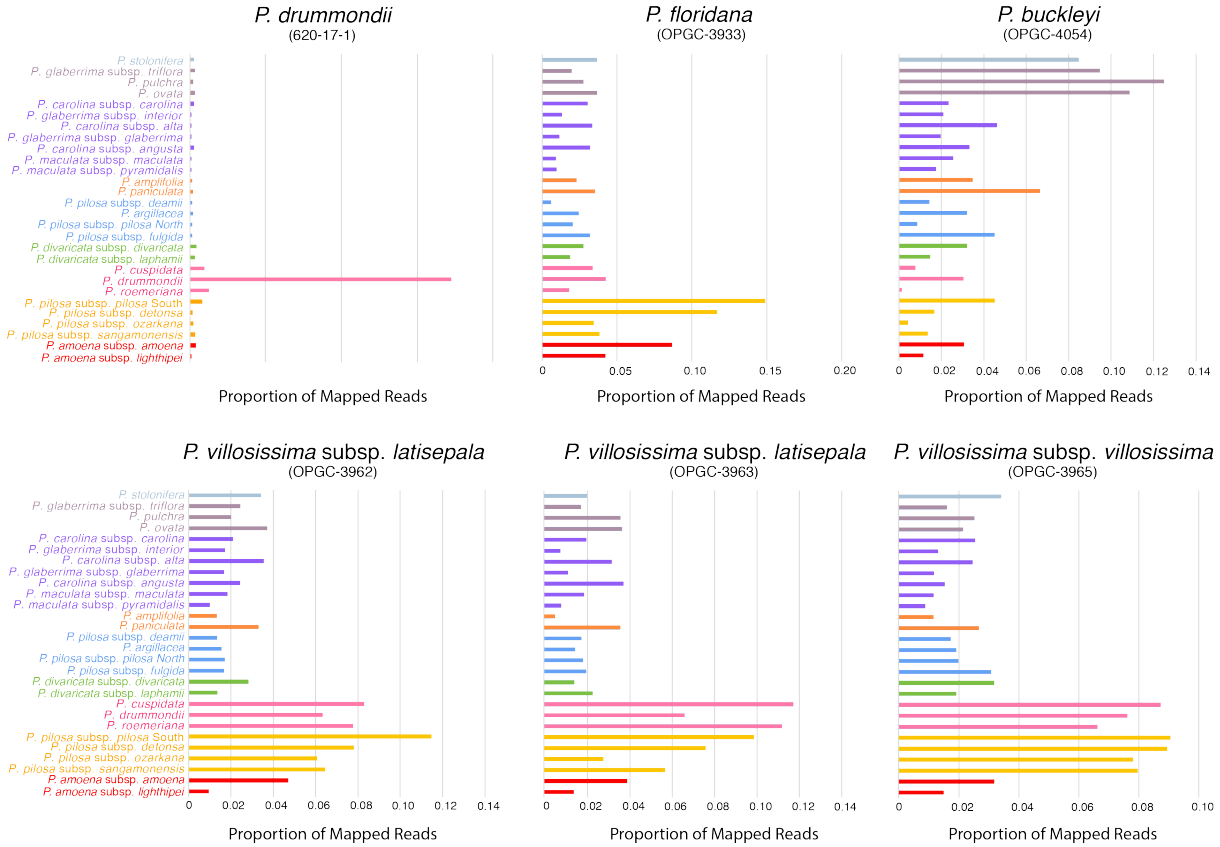


Figure 3. (a and b) Schematics of flower color divergence due to reinforcement in two species of *Phlox* to prevent maladaptive hybridization with another species in sympatry. (a) *P. drummondii* (red line) has similar light-blue colored flowers to *P. cuspidata* (pink line) in allopatry but *P. drummondii* has evolved dark-red flowers in sympatry to reduce hybridizing with *P. cuspidata*. (b) In allopatry *P. pilosa* subsp. *pilosa* North (light blue line) has similar pink colored flowers to *P. glaberrima* subsp. *interior* (dark blue line), 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).

859



860 **Figure 4.** Results from TreeMix support the hybrid speciation hypothesis for *P. maculata*
861 *pyramidalis* in a pruned subtree (best fit m=4). Modeling of 4,490 unlinked SNPs for individuals
862 clustered by taxon identity inferred *P. maculata* subsp. *pyramidalis* sister to *P. maculata* subsp.
863 *maculata* and receiving a migration event (gene flow) from *P. glaberrima* subsp. *interior*.
864



865

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

869

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[IJ], 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[IJ], flavonoids[L], seed protein polymorphism[F], experimental crosses[C]
P. buckleyi	4N	unknown	unknown	karyology[I]

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).

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.

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

New Phytologist

***New Phytologist* Supporting Information**

Article title: Phylogenomic analyses re-examine the evolution of reinforcement and hypothesized hybrid speciation in *Phlox* wildflowers

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

Article acceptance date: 22 March 2024

Fig. S1: Maximum likelihood phylogenetic inference of eastern standing *Phlox* relationships from whole concatenated ddRAD loci. Loci were required to be shared by at least 10 individuals. Numbers on branches represent bootstrap supports using ultrafast bootstrap approximation (UFB). Bootstrap supports are shown for all nodes differentiating between taxa.

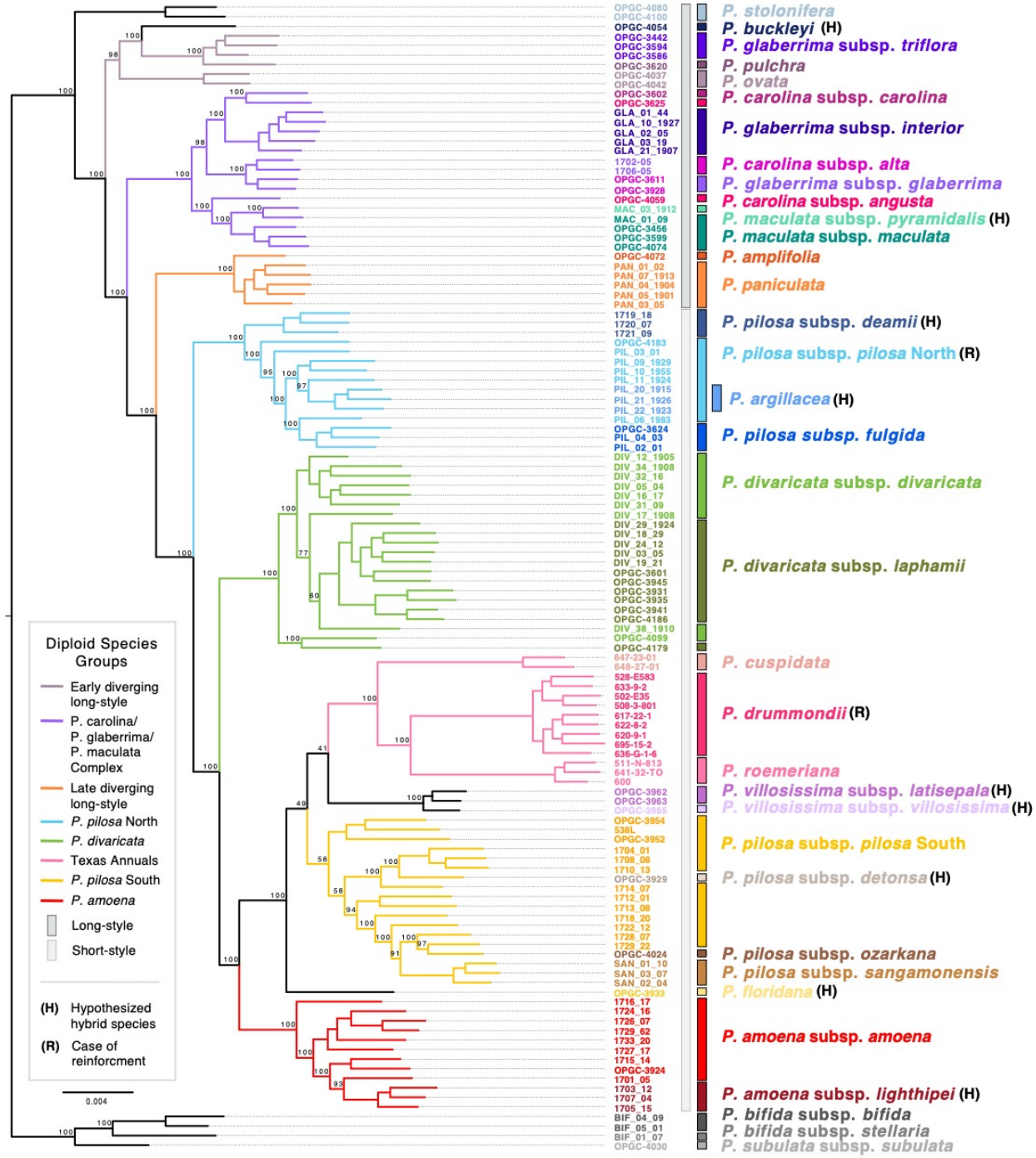


Fig. S2: Maximum likelihood phylogenetic inference of eastern standing *Phlox* relationships from whole concatenated ddRAD loci. Loci were required to be shared by at least 20 individuals. Numbers on branches represent bootstrap supports using ultrafast bootstrap approximation (UFB). Bootstrap supports are shown for all nodes differentiating between taxa.

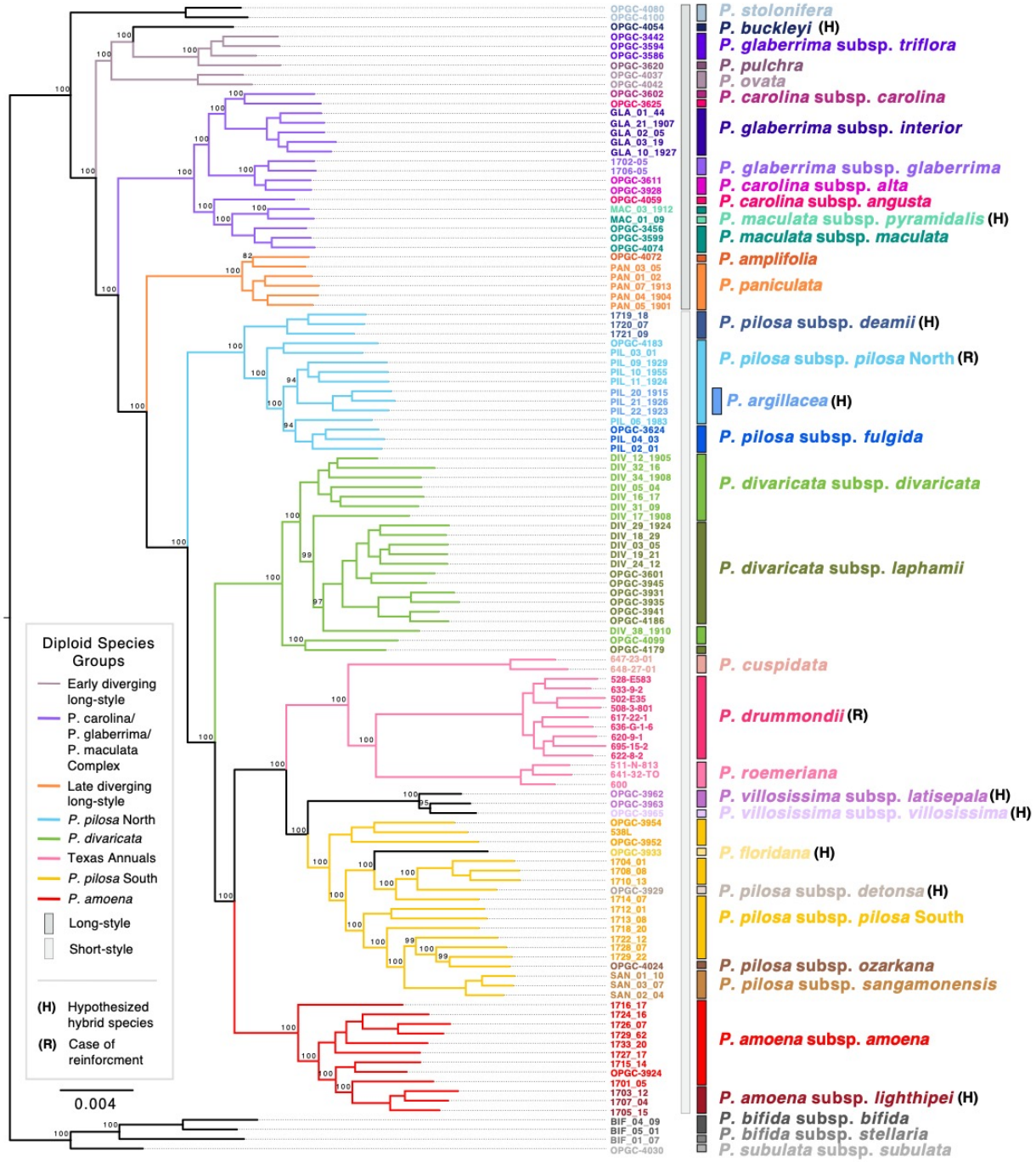


Fig. S3: Coalescent-based phylogenetic inferences of eastern standing *Phlox* relationships from 5,306 unlinked SNPs using SVDquartets, both A) with and B) without identified hybrid taxa. Loci were required to be shared by at least 30 individuals. Numbers on branches represent bootstrap supports from 100 nonparametric bootstrap replicates (SBS). Unlabeled branches had supports of SBS=100.

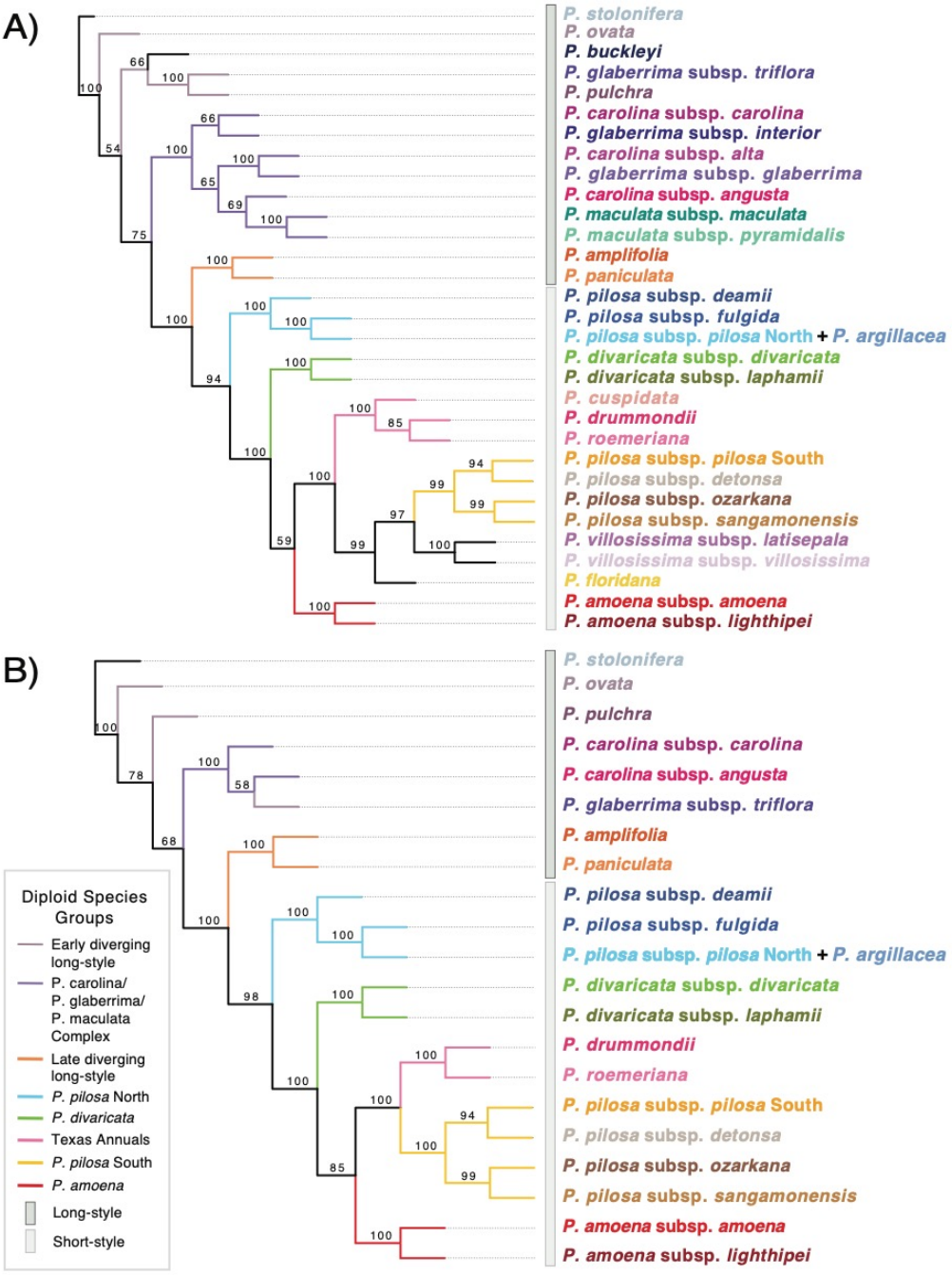


Fig. S4: Qualitative approximation of best fit TreeMix subtree model by the decay in increase in log likelihood across subtrees modeled with TreeMix with increasing number of migration events m = 0-8.

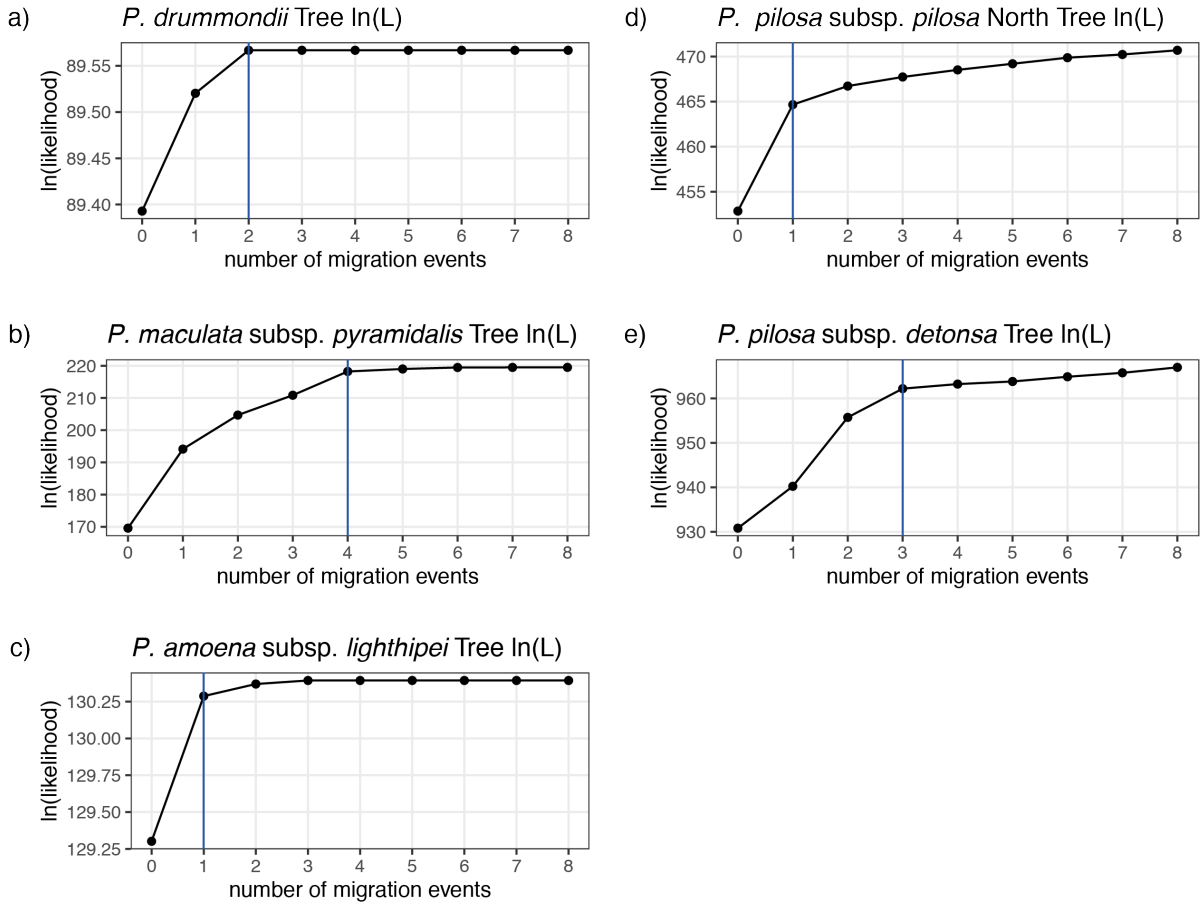


Fig. S5: TreeMix *P. drummondii* subtree models with m=1-8 migration events. Best fit model (m=2) identified by the decay in log likelihood across models (Supporting Information Fig. S4).

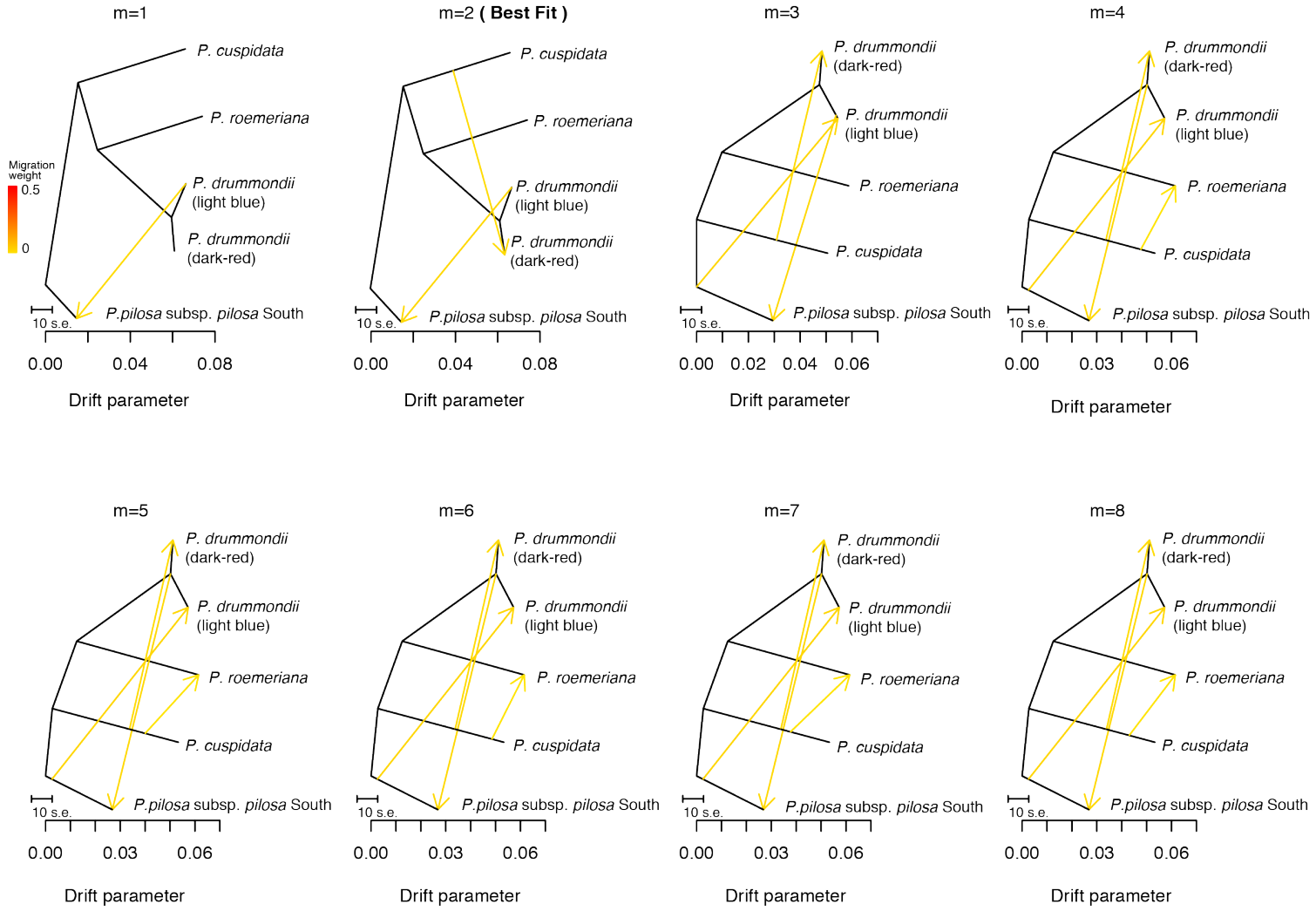


Fig. S6: TreeMix *P. pilosa* subtree models with m=1-8 migration events. Best fit model (m=1) identified by the decay in log likelihood across models (Supporting Information Fig. S4).

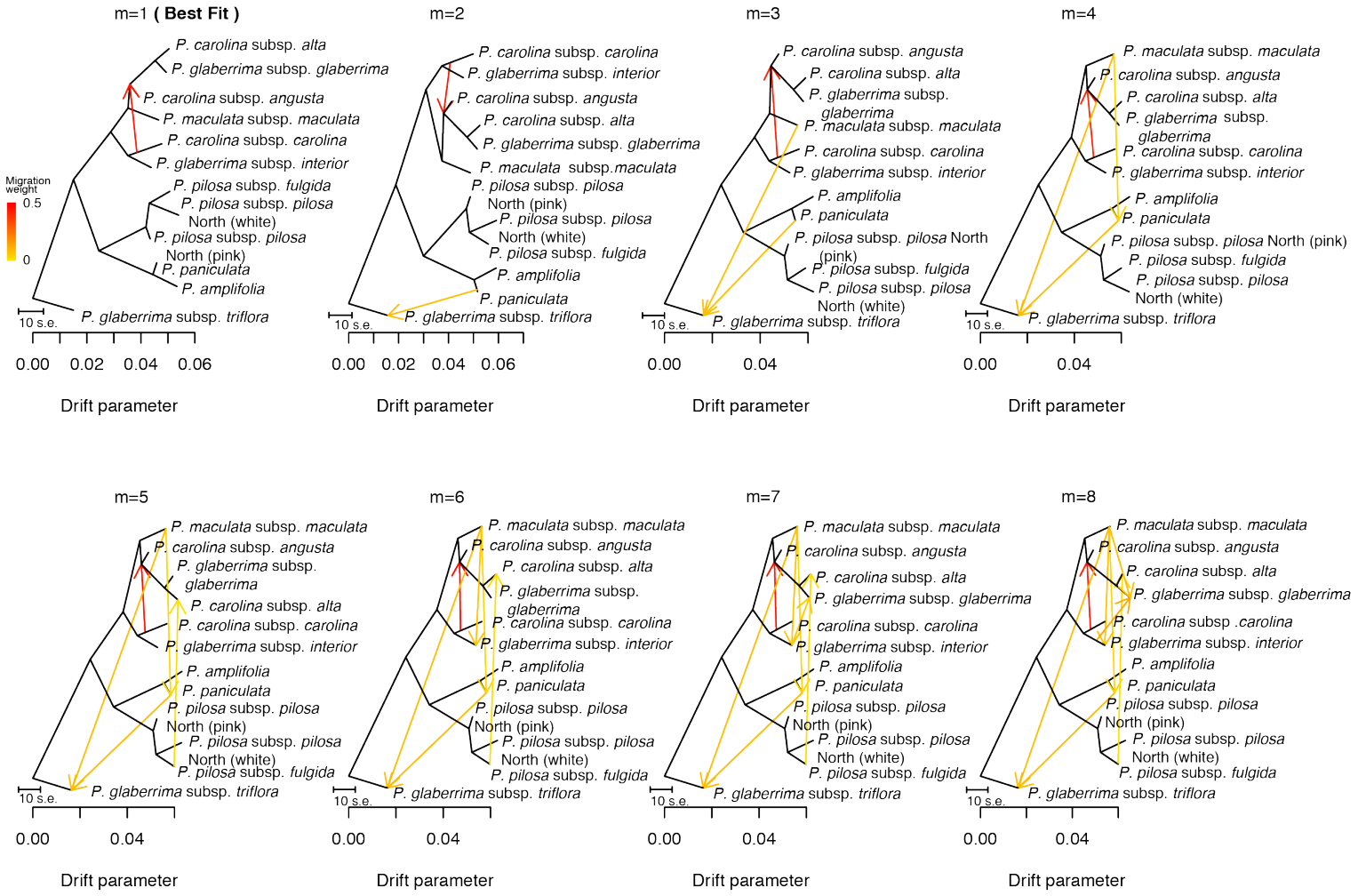


Fig. S7: TreeMix *P. pilosa* subsp. *detonsa* subtree models with m=1-8 migration events. Best fit model (m=3) identified by the decay in log likelihood across models (Supporting Information Fig. S4).

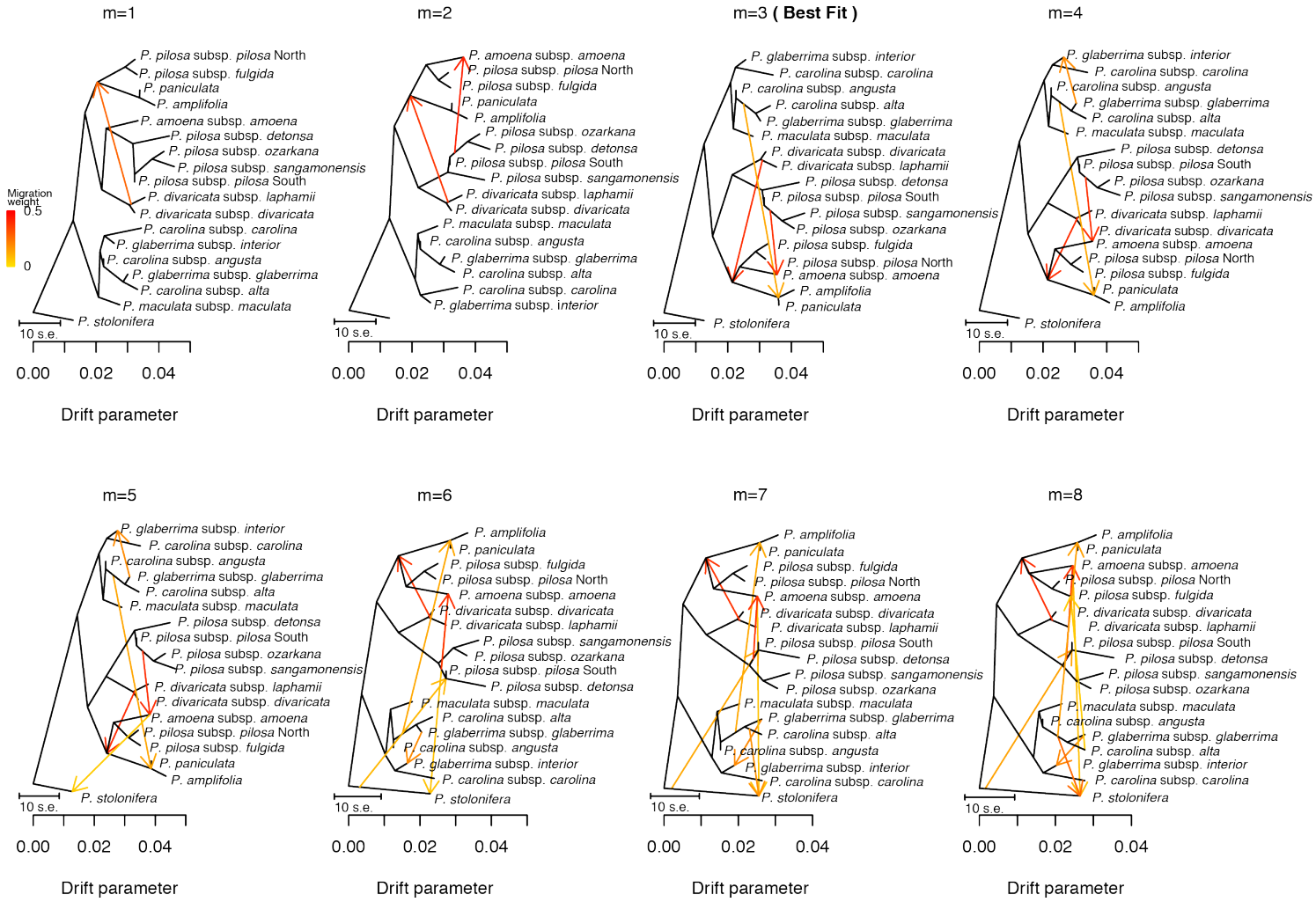


Fig. S8: TreeMix *P. amoena* subsp. *lighthipei* subtree models with m=1-8 migration events. Best fit model (m=1) identified by the decay in log likelihood across models (Supporting Information Fig. S4).

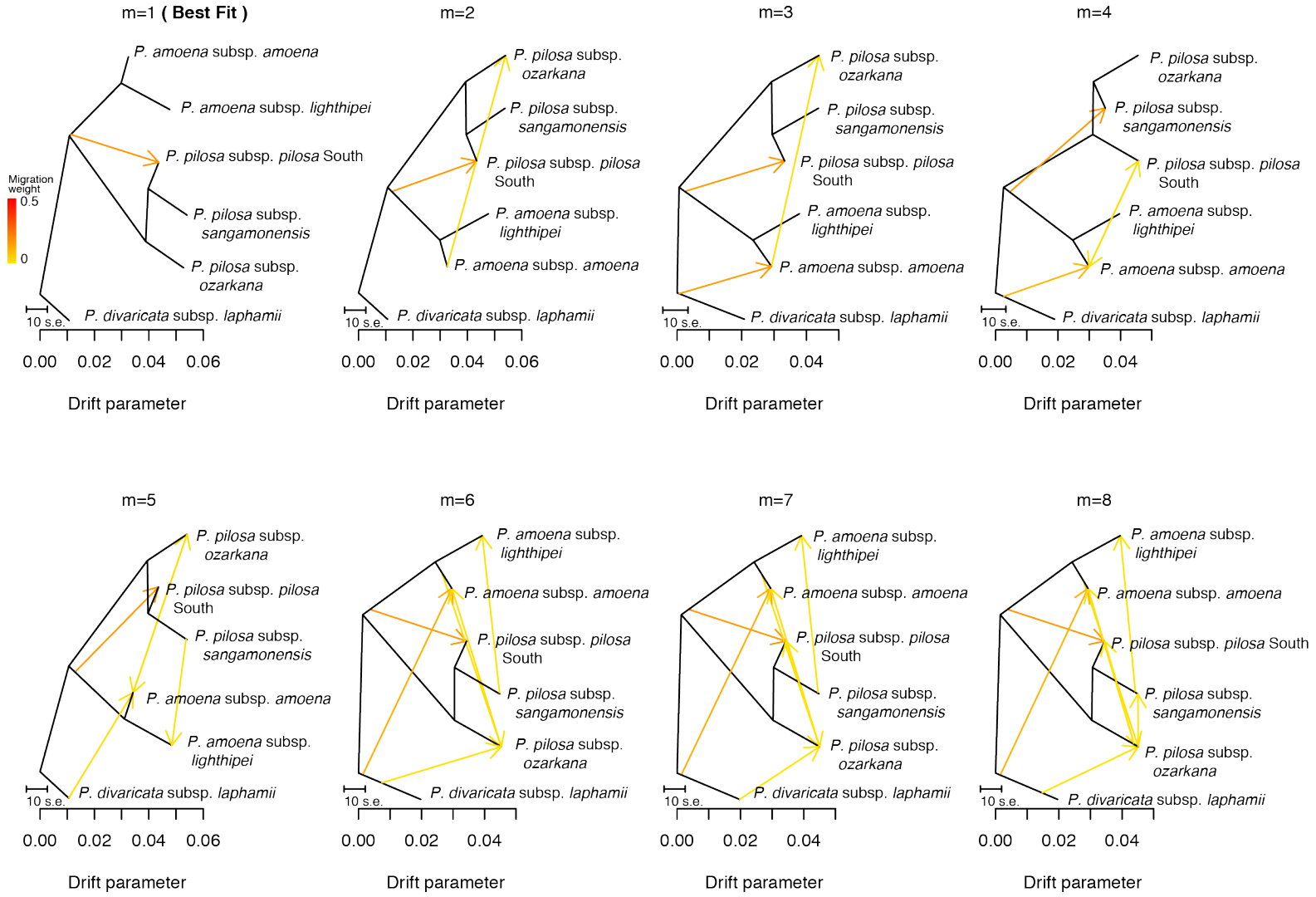


Fig. S9: TreeMix *P. maculata* subsp. *pyramidalis* subtree models with m=1-8 migration events. Best fit model (m=4) identified by the decay in log likelihood across models (Supporting Information Fig. S4).

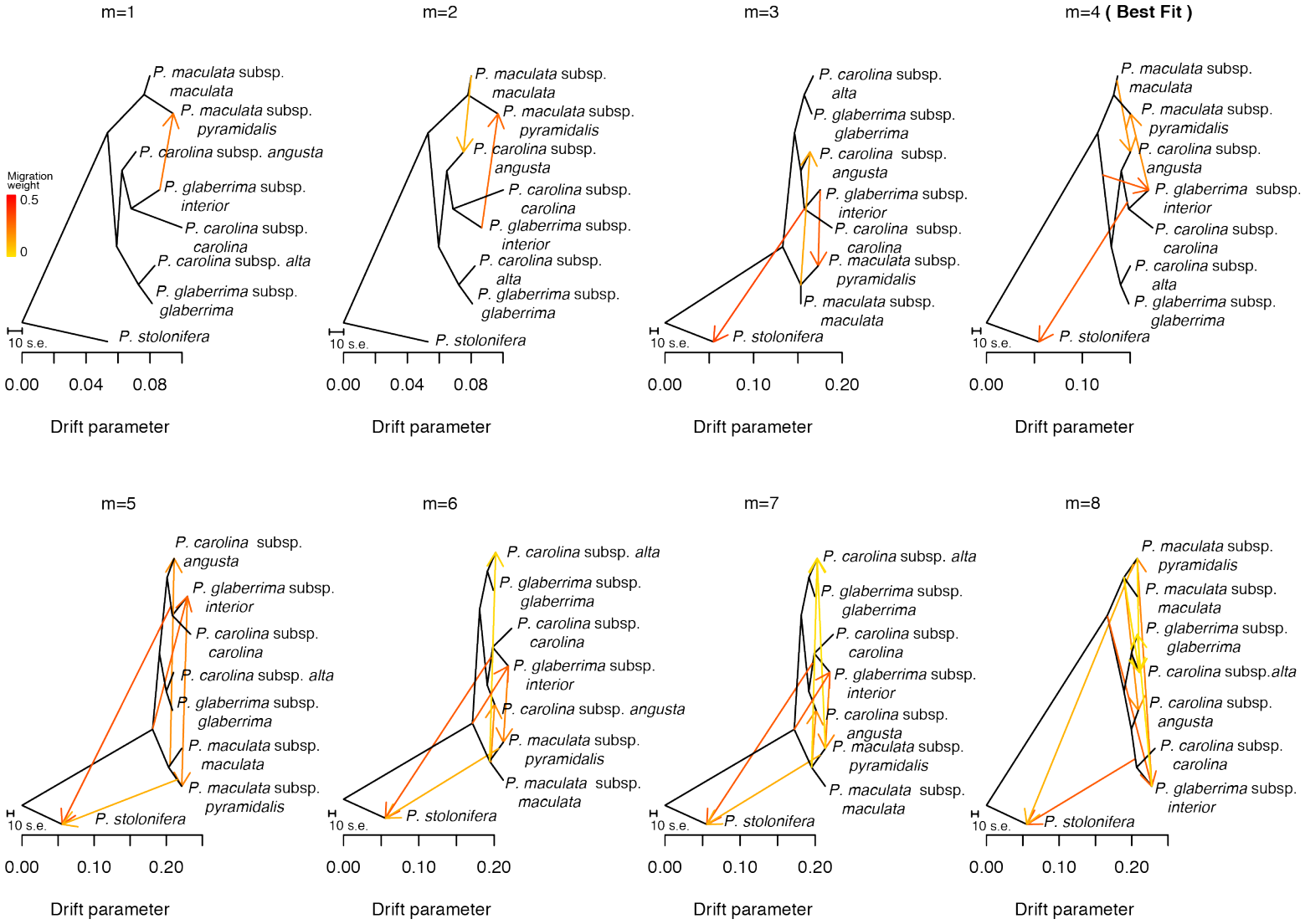


Fig. S10: SplitsTree NeighborNet inference of eastern standing *Phlox* relationships from informative SNPs in the MS30 dataset. Numbers on branches represent bootstrap supports from 1000 bootstrap replicates. Colors denote the major diploid species groups resolved in our phylogenetic inferences (Fig. 2, 5); while numbers denote the positions of the allopolyploid hybrid species; 1) *P. floridana* 2) *P. villosissima* subsp. *villosissima* 3) *P. villosissima* subsp. *latisepala* 4) *P. buckelyi*.