doi:10.1111/evo.13507

The relationship between postmating reproductive isolation and reinforcement in *Phlox*

Sevan S. Suni[1,2,3,4] and Robin Hopkins[1,2,5]

[1]*The Arnold Arboretum of Harvard University, 1300 Centre Street, Boston, MA 02131*

[2]*Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA*

[3]*Present Address: The University of San Francisco, 2130 Fulton Street, San Francisco, CA 94117*

[4]*E-mail: ssuni@usfca.edu*

[5]*E-mail: rhopkins@fas.harvard.edu*

Received October 23, 2017
Accepted April 14, 2018

The process of speciation involves the accumulation of reproductive isolation (RI) between diverging lineages. Selection can favor increased RI via the process of reinforcement, whereby costs to hybridization impose selection for increased prezygotic RI. Reinforcement results in phenotypic divergence within at least one taxon, as a result of costly hybridization between sympatric taxa. The strength of selection driving reinforcement is determined by the cost of hybridization and the frequency of hybridization. We investigated the cost of hybridization by quantifying postmating RI barriers among *Phlox* species that comprise one of the best-studied cases of reinforcement. We determined if the strength of RI differs among lineages that have and have not undergone reinforcement, how much variability there is within species in RI, and whether RI is associated with phylogenetic relatedness. We found high RI for the species that underwent phenotypic divergence due to reinforcement; however, RI was also high between other species pairs. We found extensive variability in RI among individuals within species, and no evidence that the strength of RI was associated with phylogenetic relatedness. We suggest that phenotypic divergence due to reinforcement is associated with the frequency of hybridization and introgression, and not the cost of hybridization in this clade.

KEY WORDS: **hybrid sterility, hybridization, *Phlox*, reinforcement, Reproductive isolation, speciation.**

The process of reinforcement occurs when selection favor traits that decrease hybridization because hybrids are sterile, inviable, or maladapted (Dobzhansky 1937, 1940). Many studies across a diversity of taxa have found evidence of reinforcement (reviewed in Howard 1993; Rundle and Schluter 1998; Servidio and Noor 2003; Hopkins 2013), suggesting that it can play an important role in the process of speciation. Although most research has focused on verifying that reinforcement has occurred, it is also important to understand why reinforcement occurred. Specifically, under which conditions does increased premating reproductive isolation (RI) evolve due to selection against hybridization?

The strength of reinforcing selection is determined by of the cost of hybridization and the frequency of hybridization (Liou and Price 1994; Kelly and Noor 1996). Quantifying the frequency of past hybridization that generated reinforcing selection is difficult if not impossible, but assessing the cost of hybridization is feasible. High costs to hybridization are common across taxa, and can result from a variety of postmating RI barriers, such as gametic incompatibilities, hybrid inviability, or hybrid sterility (reviewed in Coyne and Orr 2004). Investigating the strength and relative importance of these barriers to total RI can suggest the primary source of selection against hybridization driving reinforcement.

There are a number of hypotheses regarding how the strength and distribution of postmating reproductive barriers affect if and how divergence due to reinforcement will occur. Comparing sympatric regions that have and have not undergone reinforcement can provide a framework to begin to test some of these hypotheses. Theory predicts that reinforcement is more likely to succeed when

selection is strong (Liou and Price 1994; Kelly and Noor 1996), suggesting that sympatric zones with reinforcement may have stronger postmating RI than sympatric zones that have not undergone reinforcement. Although we know of no cases that explicitly investigated if the occurrence of reinforcement is associated with higher postmating RI in sympatry, support for this notion comes from work in *Drosophila* showing that premating and postmating RI are highly correlated in sympatry (Yukilevich 2012). All else being equal, if phenotypic divergence due to reinforcement occurs only in one of the two sympatric species, we may expect that the diverged species will have a greater cost to hybridization than the species that did not diverge. Consistent with this hypothesis, work in animals has found evidence that asymmetric phenotypic divergence is correlated with an asymmetric cost to hybridization (Jaenike et al. 2006; Cooley 2007; Yukilevich 2012). In plants, every known example of reinforcement involves asymmetric phenotype divergence (Hopkins 2013). Although asymmetries in postmating barriers to reproduction are common (Tiffin et al. 2001; Lowry et al. 2008; Widmer et al. 2009; Baack et al. 2015), to our knowledge the association between asymmetric phenotypic divergence in sympatry and asymmetry in the strength of postmating RI has not been investigated in plants.

Quantifying the cost to hybridization is the key to understanding if and why reinforcement occurs; potentially, determining the variability in the cost to hybridization may also provide insights into reinforcement. The frequency of alleles that cause RI within sympatric populations should influence the strength of reinforcing selection in those populations. Populations that are polymorphic for alleles at RI loci should experience lower reinforcing selection than populations that are fixed for these RI alleles. It follows, then, that species may have lower average costs to hybridization, and therefore not diverge due to reinforcement, because of polymorphisms in RI alleles. Therefore, we hypothesize that species with higher variability in RI may have lower RI and therefore not diverge. Most studies assume there is not genetic variation in postmating RI (Orr and Turelli 2001). However, variation in RI within pairs of taxa has been documented in several systems (Sweigart et al. 2007; Scopece et al. 2010; Kozlowska et al. 2011; Charron et al. 2014; Mandeville et al. 2015). To our knowledge, investigating the within-species variability in RI in the context of reinforcement has never been done.

Another factor that could influence the strength of postmating RI is the amount of time that has passed since two lineages diverged. Postmating RI is influenced by both postmating prezygotic barriers such as gametic incompatibilities, and also by postzygotic barriers such as zygote mortality, hybrid inviability, and hybrid sterility. Postzygotic RI tends to increase with phylogenetic divergence in animals (Coyne and Orr 1989, 1997; Sasa et al. 1998; Presgraves and Noor 2002; Price and Bouvier 2002; Bolnick and Near 2005; Singhal and Moritz 2013). In plants, several taxa show a positive association of RI and genetic distance, such as *Silene* (Moyle et al. 2004), *Coreopsis* (Archibald et al. 2005), food-deceptive orchids (Scopece et al. 2007), and *Helianthus* and Madiinae (Owens and Rieseberg 2014). Others, such as *Glycine* and *Streptanthus*, show no such association (Moyle et al. 2004). A positive association between the strength of postzygotic RI and phylogenetic divergence presumably occurs because lineages that diverged long ago have had more time to accumulate and fix mutations causing RI (Widmer et al. 2009). This could theoretically lead to a greater likelihood of reinforcement in lineages that diverged long ago, assuming that appropriate mutations occurred and that they were able to evolve in sympatry without constraint (Walsh and Blows 2009).

We studied postmating RI among three *Phlox* species that comprise one of the best-studied cases of reinforcement (Hopkins and Rausher 2012; Hopkins 2013). *Phlox drummondii*, *Phlox cuspidata*, and *Phlox roemeriana* are three annual herbs with overlapping ranges in Texas. The species diverged about two million years ago and, although there is extensive discordance across gene trees, *P. drummondii* and *P. roemeriana* are sister taxa while *P. cuspidata* is sister to that pair (Roda et al. 2017). All three species have light-blue flowers throughout most of their ranges. However, *P. drummondii* has evolved dark-red flowers where it exists in sympatry with *P. cuspidata*. The evolution of this novel dark-red flower color is due to reinforcement (Levin 1985), whereby it reduces hybridization between *P. drummondii* and *P. cuspidata* (Hopkins and Rausher 2012). Unlike *P. drummondii*, there is no evidence that *P. cuspidata* has diverged due to reinforcement in sympatry. Hybrids between *P. drummondii* and *P. cuspidata* are found in nature (Levin 1967; Ferguson et al. 1999; Ruane and Donohue 2008; Ruane 2009), and data from crosses suggest that the fitness of these hybrids is low (Levin 1967; Ruane and Donohue 2008). However, the maternal and paternal species of the hybrids used in Levin (1967), and the species to which hybrids were crossed in Ruane and Donohue (2008) were not specified.

P. drummondii and *P. roemeriana* also share a region of broad sympatry or parapatry and there is no evidence that either species has diverged due to reinforcement in this region of overlap. Controlled crosses revealed that matings between *P. drummondii* as the maternal seed parent and *P. roemeriana* as the paternal pollen donor can result in seed production (Erbe and Turner 1962). There are currently no data about other postmating barriers between *P. drummondii* and *P. roemeriana*, or about the extent of postmating RI between *P. roemeriana* and *P. cuspidata*.

We tested the following four hypotheses in this system. First, we hypothesize that the strength of postmating RI is higher between sympatric species that undergo reinforcement than between those that do not. Second, asymmetric divergence due to reinforcement is associated with asymmetric postmating RI. Third, RI is stronger between more distantly related lineages. Fourth,

variation in postmating RI among individuals within a species is lower for lineages that undergo phenotypic divergence due to reinforcement. To investigate these hypotheses, we assessed RI within and among species pairs at three postmating barriers via crosses among the species. We took a focal lineage perspective and characterized RI for each species from each other species (six estimates of RI), and then compared estimates among the species combinations. We also developed a statistical framework for characterizing whether a particular stage in mating and reproduction significantly reduces fitness in heterospecific versus conspecific crosses.

Materials and Methods

COLLECTIONS AND PLANT REARING

In May of 2014, we collected seeds from 23 natural populations of *Phlox* throughout Texas, including five *P. cuspidata*, five *P. roemeriana*, and 13 *P. drummondii* sites (Table S1). We brought seeds back to the Arnold Arboretum of Harvard University, planted them in Farfard germination soil, and grew them in growth chambers (Conviron MTPC144) set to long days (16 h of light) at 27°C. We transferred seedlings to Promix high porosity soil with mycorrhizae, watered as needed, and fertilized them regularly with Dyna-Gro Grow™ (Richmond, CA) and Dyna-Gro Bloom™. Germination and survival of seedlings was variable among species and maternal half-sibling families within species. Our final sample of 98 plants included seeds from 1–9 maternal half-sibling families per population, 85 maternal half-sibling families, and 1–2 half siblings per maternal half-sibling family (Table S1).

EXPERIMENTAL OVERVIEW

We evaluated the extent to which postmating barriers to reproduction, including gametic incompatibilities or zygote mortality, hybrid inviability, and hybrid sterility, contribute to RI among Texas *Phlox* species. For each species, we tested if success at each stage of mating and reproduction was significantly decreased by heterospecific versus conspecific matings (see 'Modeling variation in mating and germination success,' below). We then calculated an index of RI for all stages that showed significantly lower heterospecific reproductive success (see 'Quantifying RI,' below). We used this RI index to compare RI estimates among species pairs. To explore variability within each species in its RI from each other species, we compared the variation in mating success between conspecific crosses and heterospecific/hybrid crosses (see 'Within-species variability in RI,' below).

EXPERIMENTAL METHODS

To assess RI caused by gametic incompatibility or zygotic mortality we compared seed production from heterospecific crosses and conspecific crosses. We performed reciprocal crosses among 26 *P. cuspidata*, 43 *P. drummondii*, and 17 *P. roemeriana* individuals. These parents are the F_0 generation plants. Hereafter, we refer to crosses among these individuals by their cross-type with maternal species listed first and paternal species second. For example, a drum-cusp cross-type is between a *P. drummondii* maternal plant and a *P. cuspidata* pollen donor. For each cross-type, we haphazardly choose paternal plants from the available plants, so crosses were within and among different populations. As a control, we also performed conspecific crosses with these same plants. The number of crosses we performed averaged 5.6 crosses per cross-type per plant (range 1–38 crosses depending on the number of flowers produced) for a total of 1297 crosses. Table 1 shows the number of maternal plants used and the number of heterospecific and conspecific crosses performed for each cross-type. We emasculated all flowers three days prior to crossing to prevent self-fertilization. We used forceps to transfer pollen for crosses, and washed them with 70% isopropyl alcohol between crosses. We bagged crossed inflorescences for seed collection and counted seeds once they were mature. *P. drummondii* and *P. cuspidata* have three ovules per flower and *P. roemeriana* has three to five ovules (Erbe and Turner 1962).

To evaluate RI caused by seed inviability, we planted F_1 seeds from the crosses described above. We compared germination rates of the hybrid seeds relative to seeds resulting from conspecific crosses. From each F_0 maternal plant, we planted an average of 8.6 seeds per cross-type (range 1–38) using similar plant-care protocols described above. We used seeds from 83 maternal F_0 plants, for a total of 1446 seeds (Table 2).

To quantify hybrid female sterility and incompatibility, we performed backcrosses in which the hybrid plants received pollen from both their maternal and paternal species (Fig. 1). To assess RI caused by hybrid male sterility and incompatibility, we performed backcrosses with the hybrid plants as pollen donors to both their maternal and paternal species. As controls, we performed conspecific crosses between the pure-species plants used in hybrid crosses. These pure-species F_1 individuals came from conspecific crosses among the F_0 generation plants.

In our evaluations of both female and male sterility, we used at least 10 plants of each hybrid type and species when there were enough germinants to do so, and for each cross-type we performed at least five replicate crosses per individual (Table 3). We did not perform hybrid backcrosses for plants that were mothered by *P. roemeriana* and fathered by *P. cuspidata* (roem-cusp cross-type) because the flowers of three of the four plants were deformed. This deformity could be due to hybrid incompatibility but our low sample sizes make this barrier uncertain. We removed crosses with one *P. cuspidata* individual, one *P. drummondii* individual, and two *P. roemeriana* individuals from analyses because they did not set seed or produce viable seeds when used as pollen donors

Table 1. Differences in seed set between heterospecific and conspecific crosses used to assess effects of gametic incompatibilities or zygote mortality (GI & ZM; above dashed line) and backcrosses used to asses hybrid sterility (below dashed line). Differences among cross types were assessed using likelihood ratio tests on nested GLMMs that included maternal half sibling family as a random factor.

Conspecific				Heterospecific				Model output		
Species	*N*	Crosses	Mean ± SE	Cross	*N*	Crosses	Mean ± SE	Estimate	*z*-value	*P*-value
cusp	25	156	1.5 ± 0.23	cusp <- drum	22	210	1.5 ± 0.23	0.18	1.8	0.07
				cusp <- roem	21	31	0.97 ± 0.21	0.07	0.72	0.47
drum	43	276	1.63 ± 0.14	drum <- cusp	38	132	0.67 ± 0.14	–0.76	–7.6	<0.001
				drum <- roem	38	35	0.59 ± 0.11	–0.56	–6.3	<0.001
roem	11	46	1.65 ± 0.56	roem <- cusp	15	151	0.60 ± 0.25	–0.77	–3	0.003
				roem <- drum	12	258	1.38 ± 0.43	–0.41	–1.8	0.08
cusp	10	93	1.14 ± 0.22	c-d <- cusp	13	87	0.17 ± 0.07	–2.3	–3.2	<0.001
				cusp <- c-d	10	57	0.04 ± 0.04	–2.9	–4.8	<0.001
				c-r <- cusp	10	54	0.18 ± 0.11	2.5	–3.1	0.002
				cusp <- c-r	10	72	0.19 ± 0.08	–1.7	–5.1	<0.001
				d-c <- cusp	8	50	0.10 ± 0.10	–3.2	–3.3	<0.001
				cusp <- d-c	7	36	0.28 ± 0.18	–0.78	–2.5	0.01
drum	15	133	1.62 ± 0.19	c-d <- drum	8	46	0.13 ± 0.07	–2.3	–5.9	<0.001
				drum <- c-d	13	95	0 ± 0	~	~	~
				d-c <- drum	8	50	0.52 ± 0.32	–1	–4.9	<0.001
				drum <- d-c	8	60	0.16 ± 0.16	–2.4	–6.7	<0.001
				d-r <- drum	4	20	0.5 ± 0.44	–1.1	–3.4	<0.001
				drum <- d-r	11	71	0.16 ± 0.12	–2.3	–7.3	<0.001
				r-d <- drum	2	11	0 ± 0	~	~	~
				drum <- r-d	2	23	0.6 ± 0.6	–1.1	–3.5	<0.001
roem	4	41	2.65 ± 0.81	c-r <- roem	10	56	0.67 ± 0.14	–1.8	–3.1	0.002
				roem <- c-r	7	38	0.32 ± 0.22	–2.9	–5.6	<0.001
				d-r <- roem	11	61	0.13 ± 0.11	–4.1	–5.1	<0.001
				roem <- d-r	11	52	0.05 ± 0.03	–4.9	–7.3	<0.001
				r-d <- roem	2	13	0.07 ± 0.07	–4.2	–3	0.002
				roem <- r-d	2	10	1.4 ± 1.4	–2.6	-4.3	<0.001

Note: Abbreviations are "cusp" referring to *P. cuspidata*, "drum" referring to *P. drummondii*, and "roem" referring to *P. roemeriana*. Hybrids are abbreviated with two letter codes, for example "c-r" represents a hybrid with maternal species *P. cuspidata* and paternal species *P. roemeriana*. "N" refers to the number of maternal individuals crossed to assess GI & ZM or the number of hybrid plants used in backcrosses. Tilda refers to levels of the independent variable for which coefficients were not estimated because seeds did not develop from any cross of that type.

in any conspecific crosses. In total, we performed 1229 crosses among 223 pairs of F_1 plants.

Differences in seed production from crosses in which hybrid pollen was used may be explained by genetic incompatibilities between pollen and pistils or low pollen viability in hybrids. Therefore, we also directly assessed pollen viability. We applied a modified Alexander (1969) stain to pollen that dyes viable pollen orange and inviable pollen green. The stain contained 10 ml of 95% ethanol, 1 ml of malachite green (1% in 95% Ethanol), 5 ml fuchsin acid (1% in water), 0.5 ml orange G (1% in water), 5 g phenol, 2 ml glacial acetic acid, 25 ml glycerol, and 50 ml distilled water. We soaked one anther from each of three open flowers in 500 μl of stain for 24 h at room temperature in the dark. We then placed 2 μl of the solution onto a slide and counted the number of viable and inviable pollen grains using a light microscope (average number of pollen grains per plant assessed = 128 ± 5.2 SE).

MODELING VARIATION IN MATING AND GERMINATION SUCCESS

We statistically evaluated if each potential RI barrier significantly decreased reproductive success in heterospecific versus conspecific matings for all possible species pairs. We used generalized linear mixed effect models (GLMMs) implemented in R (R core team 2013) to test whether variation in seed set, germination, or number of viable pollen grains was due to maternal and paternal

Table 2. Differences in germination between seeds from conspecific and heterospecific crosses, where the conspecific germinants were the same species as the maternal plant of each hybrid, as evaluated by GLMMs.

Conspecific				Heterospecific				Model output		
Species	Planted	Germ	Mean prop.	Hybrid	Planted	Germ	Mean prop.	Estimate	*z*-value	*P*-value
cusp	176 (17)	46	0.35 ± 0.06	c-d	161 (16)	81	0.43 ± 0.06	0.65	3.4	<0.001
				c-r	185 (13)	71	0.31 ± 0.06	0.35	1.8	0.07
drum	296 (29)	139	0.5 ± 0.05	d-c	114 (19)	36	0.31 ± 0.06	–0.43	–1.7	0.1
				d-r	196 (125)	64	0.3 ± 0.07	–0.31	–1.4	0.18
roem	71 (5)	11	0.24 ± 0.17	r-c	27 (4)	4	0.36 ± 0.24	0.39	0.58	0.56
				r-d	25 (5)	2	0.25 ± 0.02	–0.64	–0.64	0.52

Note: Number of pure-species and hybrid seeds planted is indicated and number of maternal families is in parentheses. Germ indicates number that germinated and Mean prop. is mean proportion that germinated, with standard error. Abbreviations are "cusp" referring to *P. cuspidata*, "drum" referring to *P. drummondii*, and "roem" referring to *P. roemeriana*. Hybrids are abbreviated with two letter codes, for example "c-r" represents a hybrid with maternal species *P. cuspidata* and paternal species *P. roemeriana*.

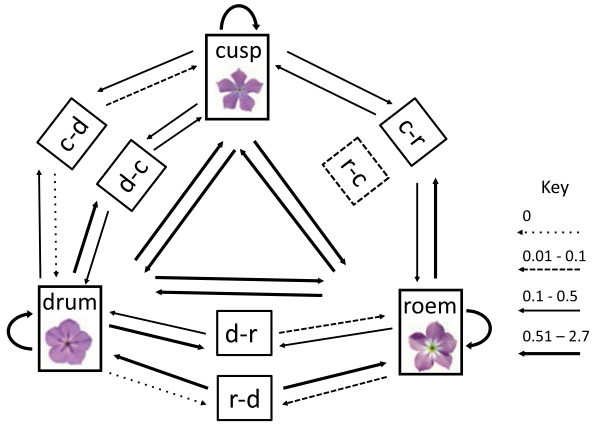

Figure 1. Seed set from controlled crosses used to evaluate reproductive isolation among Texas *Phlox* species (see Table 1 for means and standard errors for each cross-type). Here, drum represents *P. drummondii*, cusp represents *P. cuspidata*, roem represents *P. roemeriana*, and "c-r" represents a hybrid in which the maternal species was *P. cuspidata* and the paternal species was *P. roemeriana*. Arrows in the center represent heterospecific crosses conducted to assess gametic incompatibilities and zygote mortality (GI & ZM). Peripheral arrows represent crosses conducted to assess hybrid maternal and paternal sterility/incompatibility. Dashed box represents hybrids that were not crossed due to hybrid flower deformity. Arrow style/thickness reflects number of seeds set per cross, as specified in the key.

species or hybrid identity while controlling for individual-based variation.

To evaluate if there were differences in seed set between heterospecific and conspecific crosses among the F_0 plants, we ran three GLMMs (one for each species) with a Poisson error structure and log link function with the package lme4 (Bates et al. 2015). In these models, our dependent variable was number of seeds and paternal species was the independent variable. We included the number of crosses performed per individual as an offset and maternal plant identity as a random effect. We evaluated whether there were significant differences between conspecific and heterospecific crosses using R's summary function, which implements Wald tests to test for differences in means among levels of fixed factors.

To evaluate whether there were differences in germination between pure species and hybrid seeds, we ran three GLMMs (one for each species) with a negative binomial error structure. In these models, our dependent variable was the number of seeds that germinated and our independent variable was the paternal species of the seed. We included the number of seeds planted as an offset and the maternal plant identity as a random effect. As above, we evaluated whether seeds fathered by heterospecific or conspecific individuals differed in germination using Wald tests.

To evaluate whether there were differences in seed set between conspecific crosses and backcrosses, we ran three GLMMs (one model for each species) with a Poisson error structure and log link function. In these models, backcross-type was our independent variable and the number of crosses per individual was included as an offset. Backcross-type was a combination of pure species identity and hybrid-type used in the cross, along with a specification of whether hybrids were used as seed parents or pollen donors (see Table 1 for backcross-types). We included maternal plant identity as a random effect. We did not run models for roem-cusp hybrids due to hybrid flower deformity.

We used a GLMM with a Poisson error structure and log link function to test how pollen sterility varied between pure-species and hybrid individuals. Our model included the number of viable pollen grains as the dependent variable, cross-type of the F_0 species used to generate the hybrid as the independent variable, the number of pollen grains assessed as an offset, and the identity of the plant as a random factor. We used Wald tests

Table 3. RI and associated 95% confidence intervals at post mating barriers, total post-mating RI, and RI due to pollen sterility.

Focal species	RI from this species	GI & ZM	Hybrid inviability	Hybrid female sterility /incompatibility		Hybrid male sterility /incompatibility		Avg hybrid sterility	Total Post-mating RI	Pollen sterility
				H_m	H_p	H_m	H_p			
cusp	drum	ns	–0.32 (–0.14, –0.49)	0.76 (0.58, 0.92)	0.85 (0.57, 1)	0.92 (0.75, 1)	0.55 (0.08, 1)	0.77	0.59 (0.33, 0.87)	0.43 (0.23, 0.61)
cusp	roem	ns	ns	0.76 (0.50, 1)	NA	0.78 (0.61, 0.94)	NA	0.77	0.77 (0.62, 0.92)	0.21 (0.08, 0.38)
drum	cusp	0.34 (0.17, 0.53)	ns	0.47 (0.06, 093)	0.82 (0.64,0.98)	0.84 (0.53, 0.85)	1 (1, 1)	0.78	0.88 (0.87, 0.98)	0.43 (0.27, 0.61)
drum	roem	0.26 (0.10, 0.45)	ns	0.50 (0.03, 1)	1 (NA, NA)	0.81 (0.57, 0.98)	0.52 (NA, NA)	0.71	0.81 (0.78, 1)	0.53 (0.38, 0.69)
roem	cusp	0.51 (0.18, 0.81)	ns	NA	0.59 (0.32, 0.77)	NA	0.62 (0.21, 0.99)	0.61	0.86 (0.65, 0.97)	NA
roem	drum	ns	ns	0.95 (0.86, 1)	0.92 (0.75, 1)	0.39 (–0.19, 1)	0.95 (0.89, 0.99)	0.8	0.89 (0.74, 0.99)	0.53 (0.37, 0.68)

Note: drum represents *P. drummondii*, cusp represents *P. cuspidata*, and roem represents *P. roemeriana*. GI & ZM is RI from gametic incompatibilities and zygote mortality. Hybrid sterility can occur in hybrids generated from crosses in which the focal species is the maternal (H_m) or paternal (H_p) parent of the hybrid. Here, "ns" represents "not significant", based on results from generalized linear mixed models, and "NA" represents not assessed due to low sample sizes or deformed flowers. RI due to pollen sterility was not included in the total RI measure because it is already reflected in the hybrid male sterility/incompatibility estimates.

to compare mean number of viable pollen grains among cross-types. We confirmed that pollen viability predicts seed set in our system using a linear mixed model (LMM) with log transformed number of seeds produced as the dependent variable, and the percent viable pollen of the father in the cross as the independent variable. This LMM included maternal species as a covariate and the cross-type, the identity of the maternal plant, and the identity of the paternal plant as random factors.

QUANTIFYING RI

We calculated an RI index for each barrier at which our above models found statistical support that heterospecific/hybrid matings decreased reproductive success when compared to conspecific matings. We calculated RI as:

$$\mathrm{RI} = 1 - 2\,(\mathrm{H}/\,(\mathrm{H} + \mathrm{C}))\,, \tag{1}$$

where H is heterospecific success and C is conspecific success (Sobel and Chen 2014). This RI index ranges from –1 to 1, with –1 representing complete disassortative mating, 0 representing no RI, and 1 representing complete RI. The quantity H/H+C represents the probability of heterospecific gene flow. For each of the postmating RI barriers, we preformed multiple crosses on multiple individuals. Our goal was to get an overall assessment of the strength of the barriers to reproduction, so we calculated an average H and C for each trait weighted by the number of crosses performed per individual or pollen grains assessed. We used the following formulas to calculate weighted averages of heterospecific success across all individuals:

$$\overline{H} = \frac{\sum_i H_i}{\sum_i n_i}, \tag{2}$$

where the average heterospecific success is $\bar{H}$, heterospecific seeds produced or seeds germinated for a given individual i is H_i, and the number of crosses done with that individual or seeds planted from that individual is n_i. When calculating $\bar{H}$ due to pollen sterility, we included hybrids that were both mothered and fathered by the focal species. We calculated the average conspecific success as

$$\overline{C} = \frac{\sum_i C_i}{\sum_i n_i}, \tag{3}$$

where individual conspecific success is C_i and the number of crosses performed is n_i.

We used these weighted averages of heterospecific and conspecific success to estimate total RI for each species. We calculated total RI following equation 4S3 in Sobel and Chen (2014), from the RI estimates for each barrier, as:

$$1 - 2\left(\frac{\prod_{i=1}^{n} G_i}{\prod_{i=1}^{n} G_i + \prod_{i=1}^{n}(1 - G_i)}\right),$$

where G_i is the probability of heterospecific gene flow ($\bar{H}/(\bar{H}+\bar{C})$) for barrier i and 1- G_i is the probability of conspecific gene flow for barrier i. We did not include pollen sterility in our calculation of total RI because it was redundant with hybrid male sterility/viability. The success of seed set sired by hybrid pollen reflects pollen sterility, hybrid pollen-pistil incompatibility, and early zygote mortality. Therefore, adding pollen sterility to the total RI estimate would be including the same factor twice.

We calculated 95% confidence intervals around our estimates of RI using bootstrap resampling implemented in R. Specifically, for each measure of incompatibility, we resampled at the individual level with replacement and recalculated $\bar{H}$ and $\bar{C}$ based on the sample. Each resampled pool contained the same number of individuals as our observed pool, and our 95% confidence intervals were determined from resampling 10,000 times.

To test if species pairs differed in hybrid sterility and total RI, we calculated 95% confidence intervals around the differences of total RI and hybrid sterility using bootstrap resampling. Specifically, we resampled individual results from each cross and calculated the difference in RI from these sampled estimates from two species pairs. We repeated this resampling and difference calculation 10,000 times to generate confidence intervals in our estimate. If these confidence intervals overlapped with zero we concluded the two RI measurements were not statistically different.

WITHIN-SPECIES VARIABILITY IN RI

We tested for variability of RI across individuals by calculating a standardized measure of dispersion for heterospecific and conspecific crosses. We did this only for RI caused by gametic incompatibilities or zygote mortality because GLMMs revealed that there is no RI caused by hybrid inviability among these species (see results). Specifically, we determined if variability was greater for heterospecific mating success than conspecific mating success. Such a pattern would suggest genetic variation in a postmating barrier to reproduction. Variability in seed set and pollen viability could also be caused by general plant health, which we expect to be random with regard to cross-type.

In additional to general plant health, variability in conspecific matings could be due to within-species polymorphism in incompatibilities among individuals of the same species (Scopece et al. 2010). We therefore compared seed set between conspecific crosses conducted with plants from the same and different populations. We used the lme4 package in R to run GLMMs for each cross type with the number of seeds produced as the dependent variable. The independent variable was whether the population from which the father came was the same or different from which the mother came. We added the number of crosses as an offset, and the random factors were maternal and paternal plant identity, and maternal half sibling family.

We used the coefficient of variation (CV) as the estimate of mating success variability. CV is a unit-free measure of the dispersion of a probability distribution, defined as the SD divided by the mean (Gulhar 2012). We calculated the CV and associated 95% confidence intervals using the modified McKay method (Vangel 1996) and implemented calculations using R. We calculated the CV for each cross type using a dataset that contained the seeds per cross produced by each individual.

Results

VARIATION IN HETEROSPECIFIC MATING SUCCESS

We found variation among species pairs in whether heterospecific crosses resulted in reduced seed set relative to conspecific crosses (Table 1). For *P. drummondii* (the species that underwent phenotypic divergence due to reinforcement in sympatry with *P. cuspidata*), seed set was lower from crosses with both *P. cuspidata* and *P. roemeriana* than from conspecific crosses. For *P. roemeriana* (the sister species to *P. drummondii*), seed set was lower from crosses with *P. cuspidata* than from crosses with conspecifics, but not lower from crosses with *P. drummondii*. For *P. cuspidata*, seed set did not differ between conspecific and heterospecific crosses with either species.

For no species pair was germination of hybrid seeds lower than germination of conspecific crosses. In fact, germination success was higher for hybrid seeds mothered by *P. cuspidata* and fathered by *P. drummondii* (c-d hybrids) than for seeds that developed from cusp-cusp crosses (Table 2).

From crosses using the germinated hybrids, seed set was lower for all hybrid backcrosses than for conspecific crosses (Table 1). For example, crosses with pollen from c-d hybrids produced no or few seeds with either parent. Crosses with pollen from r-d hybrids (mothered by *P. roemeriana* and fathered by *P. drummondii*) produced some seeds with both parents but still about half as many seeds as when crossed with conspecific pollen. Some hybrids, such as c-d, had low seed set when crossed with either parent. Other hybrids, such as d-r, showed moderate seed set with at least one parent's pollen.

COMPARISONS OF RI AMONG SPECIES PAIRS

We calculated an RI index for each barrier showing a significant reduction in heterospecific versus conspecific success (Table 3), and we evaluated differences among species pairs in their RI (Table 4). RI caused by gametic incompatibilities or zygote mortality ranged from 0.26 to 0.51. RI caused by hybrid inviability for *P. cuspidata* from *P. drummondii* was significantly less than zero, indicating hybrid vigor. RI caused by hybrid male sterility and incompatibility ranged from 0.39 to 1, and hybrid female sterility

Table 4. Differences among species pairs in total post-mating RI (below diagonal) and average hybrid sterility (above diagonal), with associated 95% confidence intervals.

	cusp drum	cusp roem	drum cusp	drum roem	roem cusp	roem drum
cusp drum	–	0 [–0.18, 0.20]	–0.01 [–0.22, 0.20]	0.06 [–0.23, 0.32]	0.16 [–0.13, 0.45]	0.03 [–0.28, 0.20]
cusp roem	0.18 [–0.36, 0.56]	–	–0.01 [–0.22, 0.19]	0.06 [–0.23, 0.31]	0.16 [–0.14, 0.45]	0.03 [–0.28, 0.19]
drum cusp	**0.29 [0.06, 0.62]**	0.11 [–0.09, 0.57]	–	0.07 [–0.21, 0.32]	0.17 [–0.11, 0.45]	0.02 [–0.26, 0.20]
drum roem	**0.22 [0.01, 0.60]**	0.04 [–0.15, 0.56]	–0.07 [–0.17, 0.09]	–	0.1 [–0.23, 0.44]	0.09 [–0.40, 0.22]
roem cusp	0.27 [–0.07, 0.57]	0.09 [–0.23, 0.53]	–0.02 [–0.29, 0.06]	–0.05 [0.28, 0.13]	–	0.19 [–0.50, 0.12]
roem drum	0.3 [–0.01, 0.59]	0.12 [–0.17, 0.55]	0.01 [–0.20, 0.09]	0.08 [–0.19, 0.16]	0.03 [–0.16, 0.27]	–

Note: Bolded entries correspond to differences for which the confidence intervals do not overlap with zero. The abbreviations drum, cusp, and roem represent *P. drummondii*, *P. cuspidata*, and *P. roemeriana*, respectively. The species on the left in each species pair is the species *for* which we characterized RI *from* the species on the right. For example, cusp drum represents RI for *P. cuspidata* from *P. drummondii*.

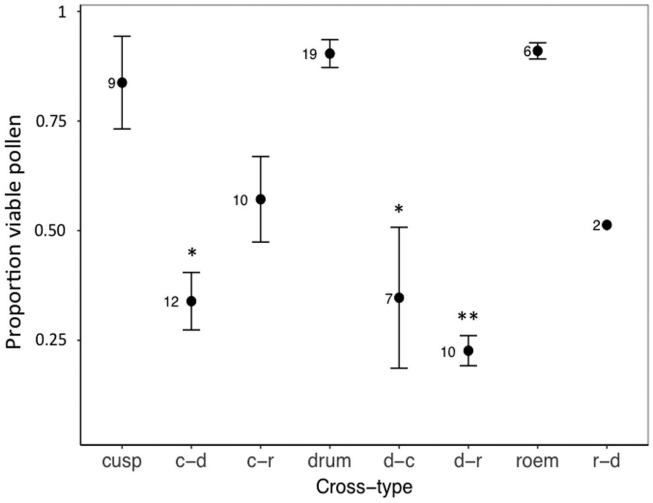

Figure 2. Proportion of viable pollen for pure species and hybrids generated from crosses among F_0 generation plants. Asterisks represent significant differences between that hybrid and its maternal species, with one asterisk representing $P < 0.01$, and two representing $P < 0.001$. Sample sizes of plants assessed are to the left of the points. For the hybrids, the letter preceding the dash represents the hybrids maternal species and the letter following the dash represents the hybrids paternal species. Confidence intervals were not generated for r-d hybrids due to low sample sizes.

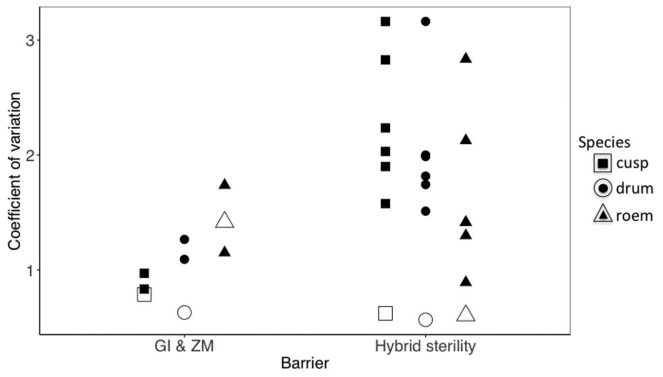

Figure 3. Variability in seed set within cross-type as evaluated by coefficient of variation (CV). On the left is the CV among conspecific (open shapes) and heterospecific (solid shapes) crosses used to assess RI caused by gametic incompatilities and zygote mortality (GI & ZM) on the left. On the right is the CV for seed set from conspecific (open shapes) and hybrid (solid shapes) crosses used to assess hybrid sterility. Table S2 shows 95% confidence intervals for each CV estimate.

and incompatibility ranged from 0.47 to 1. Total postmating RI was high for *P. drummondii* from *P. cuspidata*, and this was significantly higher than total postmating RI for *P. cuspidata* from *P. drummondii* (Tables 3 and 4). RI caused by pollen sterility ranged from 0.21 to 0.53 (Table 3; Fig. 2) and was lower than RI estimated from hybrid male sterility/inviability as measured by seed set sired by hybrid pollen.

WITHIN-SPECIES VARIATION IN RI

P. drummondii showed significantly more variability in seed set from heterospecific crosses with both *P. cuspidata* and *P. roemeriana* than from conspecific crosses as measured by coefficient of variation (Fig. 3; Table S2). For *P. cuspidata* and *P. roemeriana* variability in seed set did not differ between conspecific and heterospecific crosses (Fig. 3; Table S2). For *P. drummondii* and *P. roemeriana*, there was no differences in seed set between crosses in which conspecific pollen donors were from the same or different populations (For *P. drummondii*: GLMM est = –0.09; $P = 0.5$; *P. roemeriana*: est. = 0.98, $P = 0.23$). For *P. cuspidata*, being from the same population reduced seed set (est. = –0.49; $P = 0.027$), likely due to inbreeding depression. Greater variability in seed set within species makes our assessments of RI

variability conservative. Point estimates of variability were higher for all hybrid backcrosses than for conspecific crosses (Fig. 3). Estimates of variability were higher for all hybrid backcrosses than for conspecific crosses (Fig. 3), but the 95% confidence intervals overlapped for these crosses (Table S2).

Discussion

The strength of selection driving reinforcement is determined by the cost of hybridization (the strength of postmating RI barriers) and hybridization frequency (Liou and Price 1994; Kelly and Noor 1996). Here, we use a classic case of reinforcement in plants to explore if the occurrence of reinforcement is associated with higher postmating RI. Using thousands of crosses and a rigorous statistical framework, we demonstrate that species pairs that did and did not undergo reinforcement have similarly high levels of post-zygotic RI. Furthermore, within the species pair that underwent reinforcement, we find that RI is likely similar between the lineage that underwent phenotypic divergence and the lineage that did not. We discuss the roles that post-reinforcement divergence, hybridization frequency, and genetic constraints, in addition to the strength of RI barriers, may play in reinforcement.

ASSOCIATION BETWEEN RI AND THE OCCURRENCE OF REINFORCEMENT

Theory predicts that strong selection increases the likelihood of successful divergence due to reinforcement. The strength of reinforcing selection is correlated with the cost of hybridization (Servidio and Kirkpatrick 1997; Kirkpatrick and Servidio 1999). We therefore tested the hypothesis that post-mating RI is stronger between sympatric lineages that underwent reinforcement than sympatric lineages that did not. In the Texas *Phlox*, reinforcement occurred between *P. drummondii* and *P. cuspidata,* whereby *P. drummondii* evolved a novel flower color in sympatry that decreases hybridization with *P. cuspidata* (Hopkins and Rausher 2012). We did find strong RI between these two species, largely due to hybrid sterility. However, we also found strong RI between *P. drummondii* and its sister species *P. roemeriana*. Given the fitness reductions due to hybridization, why is there no evidence of phenotypic divergence occurring in the sympatric region of *P. drummondii* and *P. roemeriana*? These species have overlapping ranges, grow within a few kilometers of one another, and are pollinated by the same suite of pollinator species (Erbe and Turner 1962). Although not extensively studied, there are no known records of hybrids in the field. Therefore, despite the high postmating RI that we document between these taxa, we hypothesize that a low frequency of hybridization may result in weak reinforcing selection in nature. A detailed study of the degree of eco-geographic RI between *P. drummondii* and *P. roemeriana* would be useful to evaluate this hypothesis. Alternatively, the availability of genetic variation may be limiting the evolution of reinforcement in this sympatric zone. We hypothesize this is less likely because an assessment of population genetic structure in *P. drummondii* suggested ample opportunity for the spread of the alleles causing reinforcement (flower color variation) from the sympatric region with *P. cuspidata* into the sympatric region with *P. roemeriana* (Hopkins et al. 2012). Nevertheless, further research should investigate this alternative hypothesis.

DRIVERS OF ASYMMETRIC PHENOTYPIC DIVERGENCE

A limited number of studies in animals have found that asymmetric divergence due to reinforcement is associated with asymmetric cost to hybridization (Jaenike et al. 2006; Cooley 2007). In line with this hypothesis, we expected the strength of RI to be higher for *P. drummondii* from *P. cuspidata* than for *P. cuspidata* from *P. drummondii*. We found that hybrid sterility was equivalent for these two parental species, but that total postmating RI was higher for *P. drummondii* from *P. cuspidata*. This asymmetry was driven by stronger gametophytic incompatibilities or zygote mortality for crosses with *P. drummondii* as the maternal species. More work is needed to determine if this barrier is due to pollen–pistil incompatibility, which would actually decrease hybridization, or due to zygote mortality, which would increase the cost of hybridization. Preliminary evidence suggests that pollen–pistil incompatibility is responsible for this barrier to gene flow (Roda and Hopkins 2017) and therefore, it is likely that *P. drummondii* and *P. cuspidata* suffer similar costs to hybridization.

Given that hybrid sterility is high for both *P. drummondii* and *P. cuspidata*, what explains the asymmetric phenotypic divergence? We outline three plausible scenarios. First, historical costs to hybridization could have been similarly high for both *P. drummondii* and *P. cuspidata*, and either standing genetic variation or a novel mutation that happened to be present in *P. drummondii* and not in *P. cuspidata* could be responsible for the asymmetric divergence. Second, it is possible that asymmetries in hybridization costs were present historically, but that post-reinforcement divergence led to the similarly high hybridization costs we see for both species. In addition to intrinsic post-zygotic RI, extrinsic postzygotic isolation, for example, based on the higher tolerance of *P. drummondii* to high soil calcium concentrations (Ruane and Donohue 2007) could cause variation in the strength of selection. Third, it is possible that a higher frequency of hybridization in *P. drummondii* relative to *P. cuspidata* drove the asymmetric phenotypic divergence.

Genome wide patterns of genetic variation indicate there is asymmetric gene flow from *P. cuspidata* into *P. drummondii*, which could result in stronger reinforcing selection for *P. drummondii* than *P. cuspidata* (Roda et al. 2017). While our RI data suggest it should be easier for more hybrids to be mothered by

P. cuspidata than by *P. drummondii*, in the field most hybrids are mothered by *P. drummondii* (Ferguson 1999), and backcrossing likely occurs more frequently into *P. drummondii* than into *P. cuspidata.* This may be because *P. cuspidata* has high rates of self-fertilization that limit the formation of hybrids mothered by *P. cuspidata* (Levin 1978). It could also be because pollinators are more likely to switch from *P. cuspidata* to *P. drummondii* than from *P. drummondii* to *P. cuspidata* when foraging, due to the higher nectar reward of *P. drummondii* (average greenhouse-assessed nectar volume (microliter) of dark-red *P. drummondii*: 0.98 ± 0.19 (SE; $N = 6$); nectar volume of *P. cuspidata*: 0.19 ± 0.09 (SE; $N = 7$)). Our finding that neither drum-cusp hybrids nor cusp-drum hybrids are completely sterile suggests plausible paths for introgression from *P. cuspidata* into *P. drummondii*. Taken together, the genomic evidence of asymmetric gene flow (Roda et al. 2017) and our evidence of strong reproductive barriers to gene flow into both species, support the hypothesis that the higher frequency of hybrid introgression experienced by *P. drummondii* relative to *P. cuspidata* can explain why phenotypic divergence evolved in *P. drummondii*.

ASSOCIATION OF RI AND PHYLOGENETIC RELATEDNESS

We found no evidence that the strength of RI was associated with phylogenetic relatedness. We predicted that the two sister species, *P. drummondii* and *P. roemeriana*, would have lower post-zygotic RI than either of these species with *P. cuspidata*. Instead we found similar levels of high postzygotic RI for all species pairs. This contrasts with the general pattern that is emerging in animals, in which postzygotic RI seems to evolve in a clock-like manner and increase with increasing genetic distance (Coyne and Orr 1989, 1997; Sasa et al. 1998; Presgraves and Noor 2002; Price and Bouvier 2002; Bolnick and Near 2005; Singhal and Moritz 2013). In plants, there seems to be more variability in the relationship between genetic distance and post-zygotic RI (Moyle et al. 2004; Baack et al. 2015). The gradual increase in RI with genetic distance in some taxa may suggest that many genes of small effect underlie RI (Edmands 2002). The genetic basis of RI in *Phlox* is currently unknown, but the lack of an association between RI and phylogenetic distance may suggest that a few genes of large effect, or changes in chromosomal structure (Edmands 2002) underlie RI, or that incomplete lineage sorting or introgression within this clade (Roda et al. 2017) is obscuring the relationship between phylogenetic distance and RI. It is worth noting that we used three species pairs in this study, and studies incorporating many more species pairs would be useful to test the general hypothesis that stronger postmating RI is associated with a higher occurrence of reinforcement across taxa.

WITHIN-SPECIES VARIABILITY IN RI

Our extensive replication of parental and hybrid crosses allowed us to investigate variability in RI within each species, and determine if variability in RI is lower in species that diverged due to reinforcement. Our rationale for testing this hypothesis is that genetic polymorphisms in alleles causing RI could cause populations to experience lower costs to hybridization, relative to populations that are fixed for RI alleles. Variability in RI suggests that some hybrid matings are less costly than others and thus the average strength of selection is lower than if these RI barriers were fixed. We therefore predicted that *P. drummondii* would express less variability in RI and have higher RI, whereas the other species would have lower costs to hybridization and more variability in alleles responsible for these costs. To evaluate variability in RI we looked for a higher coefficient of variation in seed set from heterospecific or hybrid crosses relative to conspecific crosses. Suprisingly, we found that *P. drummondii* showed significantly higher variability in seed set from heterospecific crosses than conspecific crosses, suggesting genetic variation in gametic incompatibility between species or hybrid seed mortality. This was not the case for the other two species, which showed similar levels of variation in seed set in heterospecific and conspecific crosses. All three species showed similar patterns of higher variability in seed set from backcrosses with hybrids than crosses with conspecifics. These results indicate that there is genetic variation for hybrid sterility and incompatibility, and that reinforcement can cause phenotypic divergence despite variation within species at loci that cause post-zygotic RI.

Genetic variation in loci contributing to RI could result in some studies finding strong or complete barriers to gene flow while others do not. Our evidence of significant variability in post-mating RI is consistent with the conflicting reports of the strength of RI caused by hybrid sterility of these *Phlox* species. We found that many of our hybrid individuals showed complete sterility (no seed set or pollination success). But, for only two hybrid types did we find that every individual produced zero seeds from a particular type of cross. The proportion of hybrid individuals that produced some seeds or no seeds from each cross-type is shown in Figure 4. Our sterility results differ slightly from previous reports. For example, we found that no seeds developed when *P. drummondii* was pollinated by cusp-drum hybrid pollen. In contrast, Ruane and Donohue (2008) report that seeds did develop from this type of cross. Furthermore, Ruane and Donohue (2008) reported that hybrids sired by *P. cuspidata* are seed-sterile; although this report does not specify which pure species' pollen was transferred to the hybrids. In contrast, we find that 38% of the drum-cusp hybrid individuals produce seeds with *P. drummondii* pollen, and 13% produced seeds with *P. cuspidata* pollen (Fig. 4). In addition, Levin (1967) reported that no seeds were

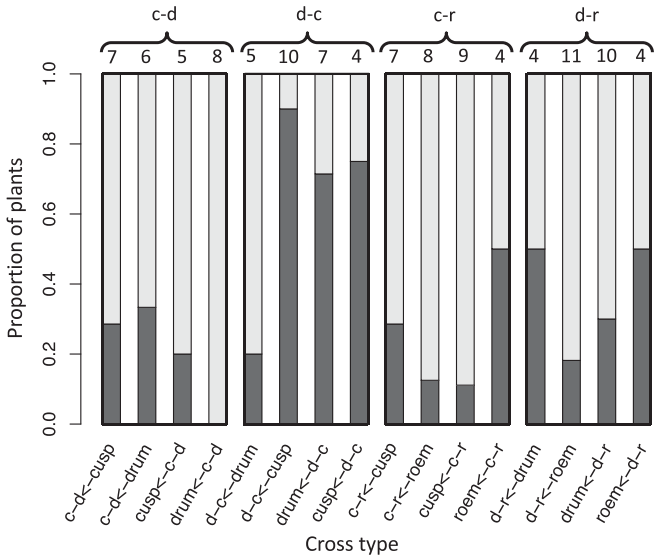

Figure 4. Within cross-types, hybrid individuals varied in whether they produced seeds. For each hybrid-type, the proportion of plants that produced at least one seed (dark grey) and no seeds (light grey) from each cross type. Arrows represent the direction of the cross, such that "cusp <- d-c" represents a cross in which pollen from a hybrid mothered by *P. drummondii* and fathered by *P. cuspidata* was transferred onto *P. cuspidata*. Hybrid-types used in crosses are above curly brackets. The number above each bar is the number of unique plant combinations used for that cross type. See Table 1 for sample sizes of hybrids and pure-species individuals used in crosses.

produced in hybrid backcrosses to *P. cuspidata*. The hybrid types used in these backcrosses were unspecified. We found that 13–50% of hybrid individuals could successfully pollinate *P. cuspidata*. The variability in the strength of post-zygotic RI within species demonstrates that hybrid incompatibilities are continuing to evolve through standing genetic variation within species. Understanding the extent of variation and the genetic basis of variation in hybrid sterility will provide new insight into how loci causing RI evolve within species.

In conclusion, we found little evidence that the cost of hybridization, as measured by postmating RI between species pairs, predicts which species will diverge in sympatry due to reinforcement. Additionally, our assessment of postmating RI provides strong evidence for genetic variation in RI within species. The strength of our study is the statistical framework we used to test if heterospecific matings significantly decrease seed set or germination, and our subsequent calculation of RI for only those barriers that showed a significant decrease. Other factors that could influence the occurrence of reinforcement include the frequency of hybridization, or genetic constraints to phenotypic divergence. In our system, there is evidence suggesting the frequency of hybridization may better explain the occurrence of reinforcement. To our knowledge, the role that genetic constraints may play in preventing lineages from undergoing phenotypic divergence due to reinforcement is unexplored and would provide valuable insight into the evolutionary dynamics of cases in which there is a lack of phenotypic divergence but high postmating RI.

AUTHOR CONTRIBUTIONS

S. Suni and R. Hopkins designed the experiments, S. Suni conducted the experiments, analyzed the data, and wrote the manuscript, and R. Hopkins contributed to data analysis, interpretation of results, and manuscript editing.

ACKNOWLEDGMENTS

We thank Federico Roda and Emma Dillon for their help performing crosses, Callin Switzer and Daniel Flynn for statistical advice, and two anonymous reviewers and Editor Jeremiah Busch for helpful comments on the manuscript.

DATA ARCHIVING

Data is archived at Dryad digital archive. DOI: https://doi.org/10.5061/dryad.31m1462.

LITERATURE CITED

Archibald, J.K., M. E. Mort, D. J. Crawford, J. K. Kelly. 2005. Life history affects the evolution of reproductive isolation among species of Coreopsis (Asteraceae). Evolution 59:2362–2369.

Baack, E., M. Clara Melo, L. H. Riesberg, and D. Ortiz-Barrientos. 2015. The origins of reproductive isolation in plants. New Phytol. 207:968–984.

Bates, D., M. Maechler, B. Bolker, and S. Walker. 2015. Fitting Linear Mixed-Effects Models Using lme4. J. Statist. Soft. 67:1–48.

Bolnick, D.I., and T. J. Near. 2005. Tempo of hybrid inviability in centrarchid fishes (Teleostei: Centrarchidae). Evolution 59:1754–1767.

Charron, G., J. Leducq, and C. R. Landry. 2014. Chromosomal variation segregates within incipient species and correlates with reproductive isolation. Mol. Ecol. 23:4362–4372.

Cooley, J. R. 2007. Decoding asymmetries in reproductive character displacement. Proc. Acad. Nat. Sci. Phil. 156:89–96.

Coyne, J. A., and H. A. Orr. 1989. Patterns of speciation in Drosophila. Evolution 43:362–381.

Coyne, J. A., and H. A. Orr 1997. "Patterns of speciation in Drosophila" revisited. Evolution 51:295–303.

Coyne, J.A., and H. Orr. 2004. Speciation. Sinauer Associates, Sunderland, MA.

Dobzhansky, T.H. 1937. Genetics and the origin of species. Columbia Univ. Press, New York, NY.

Dobzhansky, T.H. 1940. Speciation as a stage in evolutionary divergence. Am. Nat. 74:312–321.

Edmands, S. 2002. Does parental divergence predict reproductive compatibility? Trends Ecol. Evol. 17:520–527.

Erbe, L., and B. L. Turner. 1962. A biosystematic study of the *Phlox cuspidata-Phlox drummondii* complex. Am. Mid. Nat. 67:257–281.

Ferguson, C., D. A. Levin, and R. K. Jansen. 1999. Natural hybridization between an outcrossing and a selfing*Phlox (Polemoniaceae)*: The maternal species of F_1 hybrids. Plant. Syst. Evol. 218:153–158.

Gulhar, M., B. M. G. Kibria, A. N. Albatineh, and N. U. Ahmed. 2012. A comparison of some confidence intervals for estimating the population coefficient of variation: a simulation study. Sort 36:45–68.

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

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

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

Howard, D. 1993. Reinforcement: origin, dynamics, and fate of an evolutionary hypothesis. Pp. 46–69 in R. G. Harrison, ed. Hybrid zones and the evolutionary process. Oxford Univ. Press, Oxford, U.K.

Jaenike, J., K. A Dyer, K. A. C. Cornish, and M. S. Minhas. 2006. Asymmetrical reinforcement and wolbachia infection in Drosophila. PLOS Biol. 4:e325.

Kelly, J. K, and M.A.F. Noor. 1996. Speciation by reinforcement: a model derived from studies of Drosophila. Evolution 143:1485–1497.

Kirkpatrick, M., and M. R. Servedio. 1999. The reinforcement of mating preferences on an island. Genetics 151:865–884.

Kozlowska, J.L., A. R. Ahmad, E. Jahesh, and A. D. Cutter. 2011. Genetic variation for postzygotic reproductive isolation between *Caenorhabditis briggsae* and *Caenorhabditis* Sp. 9. Evolution 66:1180–1195.

Levin, D. A. 1967. Hybridization between annual species of Phlox: Population structure. Am. J. Bot. 54:1122–1130.

Levin, D. A. 1978. The origin of isolating mechanisms in flowering plants. In: Hecht M.K., Steere W.C., Wallace B. (eds). Evolutionary Biology, vol 11. Springer, Boston, MA.

Levin, D. A. 1985. Reproductive character displacement in Phlox. Evolution 39:1275–1281.

Liou, L. W., and T. D. Price. 1994. Speciation by reinforcement of premating isolation. Evolution 48:1451–1459.

Lowry D. B., J. L. Modliszewski, K. M. Wright, C. A. Wu, and J. H. Wilis. 2008. The strength and genetic basis of reproductive isolating barriers in flowering plants. Phil. Trans. Roy. Soc. B. 363:3009–3021.

Mandeville, E.G., T. L. Parchman, D. B. McDonald, and A. C. Buerkle. 2015. Highly variable reproductive isolation among pairs of *Catostomus* species. Mol. Ecol. 24:1856–1872.

Moyle, L., M. S. Olson, and P. Tiffin. 2004. Patterns of reproductive isolation in three angiosperm genera. Evolution 58:1195–1208.

Orr, H.A., M. Turelli. 2001. The evolution of postzygotic isolation: accumulating Dobzhansky-Muller incompatibilities. Evolution 55:1058–1094.

Owens, G.L., and L. H. Rieseberg. 2014. Hybrid incompatibility is acquired faster in annual than in perennial species of sunflower and tarweed. Evolution 68:893–900.

Presgraves, D.C., and M. A. F. Noor. 2002. Patterns of postzygotic isolation in Lepidoptera. Evolution 56:1168–1183.

Price, T.D., and M. M. Bouvier. 2002. The evolution of F_1 postzygotic incompatibilities in birds. Evolution 56:2083–2089.

Ruane, L.G., and K. Donohue. 2007. Environmental effects on pollen-pistil compatibility between Phlox cuspidata and P. drummondii (Polemoniaceae): implications for hybridization dynamics. American Journal of Botany 94: 219–227.

Ruane, L.G., and K. Donohue. 2008. Pollen competition and environmental effects on hybridization dynamics between *Phlox drummondii* and *Phlox cuspidata*. Evol. Ecol. 22:229–241.

Ruane, L.G. 2009. Mating system and the hybridization between self-compatible Phlox cuspidata and self-incompatible Phlox drummondii. Evol. Ecol. 23:791–805.

Rundle, H. D., and D. Schluter. 1998. Reinforcement of stickleback mate preferences: sympatry breeds contempt. Evolution 52:200–208.

Roda, F., F. K. Mendes, M. W. Hahn, and R. Hopkins. 2017. Genomic evidence of gene flow during reinforcement in Texas *Phlox*. Mol. Ecol. https://doi.org/10.1111/mec.14041.

Roda, F., and R. Hopkins. 2017. Correlated evolution of self and interspecific incompatibility across the range of a Texas wildflower. bioRxiv. https://doi.org/10.1101/155986.

Sasa, M.M., P. T. Chippindale, and N. A. Johnson. 1998. Patterns of postzygotic isolation in frogs. Evolution 52:1811–1820.

Scopece, G., A. Musacchio, A. Widmer, and S. Cozzolino. 2007. Patterns of reproductive isolation in Mediterranean deceptive orchids. Evolution 6:2623–2642.

Scopece, G., C. Lexer, A. Widmer, and S. Cozzolino. 2010. Polymorphism of postmating reproductive isolation within plant species. Taxon 59:1367–1374.

Servidio M. R., and M. Kirkpatrick. 1997. The effects of gene flow on reinforcement. Evolution 51:1764–1772.

Servidio, M.R., and M. A. F. Noor. 2003. The role of reinforcement in speciation: theory and data. Ann. Rev. Ecol. Syst. 34:339–364.

Singhal, S., and C. Morits. 2013. Reproductive isolation between phylogeographic lineages scales with divergence. Proc. R. Soc. B 280:2013–2246.

Sobel J. M., and G. F. Chen. 2014. Unification of methods for estimating the strength of reproductive isolation. Evolution 68:1511–1522.

Sweigart, A.L., A. R. Mason, and J. Willis. 2007. Natural variation for a hybrid incompatibility between two species of *Mimulus*. Evolution 6:141–151.

Tiffin, P. 2001. Asymmetrical crossing barriers in angiosperms. Proc. Biol. Sci. B. 268:861–867.

Vangel, M. G. 1996. Confidence intervals for a normal coefficient of variation. Am. Stat. 15:21–26.

Walsh, B. and M. W. Blows. 2009. Abundant genetic variation + strong selection = multivariate genetic constraints: a geometric view of adaptation. Annual Review of Ecology, Evolution, and Systematics 40:41–59.

Widmer, A., C. Lexer, and S. Cozzolion. 2009. Evolution of RI in plants. Heredity 102:31–38.

Yukilevich, R. 2012. Asymmetrical patterns of speciation uniquely support reinforcement in Drosophila. Evolution 66:1430–1446.

Associate Editor: J. Busch
Handling Editor: P. Tiffin

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Table S1. For each population, the species, GPS location, the number of maternal plants from which seeds were collected and planted in the lab (number of half sibling families used in F1 crosses), and the total number of plants grown from seeds from each population.

Table S2. Assessment of RI variability including focal species, RI barrier, cross-type, and coefficient of variation and associated 95% confidence intervals.