Gene flow improves fitness at a range edge under climate change

Megan Bontrager[1,2,3] (iD) **and Amy L. Angert[4]** (iD)

[1]*Department of Botany, University of British Columbia, Vancouver British Columbia V6T 1Z4, Canada*

[2]*Department of Evolution and Ecology, University of California, Davis. Davis, California 95616, United States*

   [3]*E-mail: mgbontrager@gmail.com*

[4]*Departments of Botany and Zoology, University of British Columbia, Vancouver British Columbia V6T 1Z4, Canada*

Populations at the margins of a species' geographic range are often thought to be poorly adapted to their environment. According to theoretical predictions, gene flow can inhibit these range edge populations if it disrupts adaptation to local conditions. Alternatively, if range edge populations are small or isolated, gene flow can provide beneficial genetic variation and may facilitate adaptation to environmental change. We tested these competing predictions in the annual wildflower *Clarkia pulchella* using greenhouse crosses to simulate gene flow from sources across the geographic range into two populations at the northern range margin. We planted these between-population hybrids in common gardens at the range edge and evaluated how genetic differentiation and climatic differences between edge populations and gene flow sources affected lifetime fitness. During an anomalously warm study year, gene flow from populations occupying historically warm sites improved fitness at the range edge and plants with one or both parents from warm populations performed best. The effects of the temperature provenance of gene flow sources were most apparent at early life history stages, but precipitation provenance also affected reproduction. We also found benefits of gene flow that were independent of climate: after climate was controlled for, plants with parents from different populations performed better at later lifestages than those with parents from the same population, indicating that gene flow may improve fitness via relieving homozygosity. Further supporting this result, we found that increasing genetic differentiation of parental populations had positive effects on fitness of hybrid seeds. Gene flow from warmer populations, when it occurs, is likely to contribute adaptive genetic variation to populations at the northern range edge as the climate warms. On heterogeneous landscapes, climate of origin may be a better predictor of gene flow effects than geographic proximity.

KEY WORDS: *Clarkia pulchella*, climate change, common garden, gene flow, geographic range limits, local adaptation, swamping gene flow.

Impact Summary

What limits species' geographic ranges on the landscape? One process of interest when trying to answer this question is gene flow, which is the movement of genetic material between populations, as might occur in plants when seeds or pollen move across the landscape. One hypothesis that has been proposed is that gene flow from populations in other environments prevents populations at range edges from adapting to their local habitats. Alternatively, it has been suggested that these populations might benefit from gene flow, as it would provide more genetic material for natural selection to act upon.

We tested these predictions in an annual wildflower, *Clarkia pulchella*. We simulated gene flow by pollinating plants from the range edge with pollen from other populations. Then we planted the resulting seeds into common gardens in the home sites of the range edge

populations and recorded their germination, survival, and reproduction. The weather during our experiment was much warmer than historic averages in our garden sites, and perhaps because of this, we found that gene flow from warm locations improved the performance of range edge populations. This result highlights the potential role of gene flow and dispersal in aiding adaptation to warming climates. We also found some positive effects of gene flow that were independent of climate. Even after we statistically controlled for adaptation to temperature and precipitation, plants that were the result of gene flow pollinations produced more seeds and fruits than plants with both parents from the same population. Rather than preventing adaptation, in our experiment, gene flow generally had positive effects on fitness.

Species are limited in their geographic extents on the landscape. In many cases, the limits of species' geographic distributions are the result of niche limitation, rather than simply an inability to disperse to suitable areas beyond their current distribution (Lee-Yaw et al. 2016). This raises the question of what prevents populations on the range periphery from adapting to sites beyond the range edge (Antonovics 1976; Bridle and Vines 2007), particularly when boundaries are not co-incident with an abrupt shift in the abiotic environment. The putative causes of limits to adaptation at the range edge hinge upon demographic and genetic features of metapopulations (Sexton et al. 2009).

If range limits represent limits to adaptation, this could be the result of insufficient genetic variation in range edge populations. Limited genetic variation at range edges is predicted because range edge populations are often characterized as small, with more frequent or severe changes in population size and elevated rates of turnover relative to central populations (Vucetich and Waite 2003). Populations that are on the leading edge of range expansions may also exhibit patterns of low genetic variation as a result of successive founder events (Pujol and Pannell 2008). Significant declines in neutral genetic variation near range edges is a common (although not ubiquitous) pattern (Eckert et al. 2008; Pironon et al. 2017), indicating that some of these processes are likely to affect some range edges in some species. If the observed declines in neutral variation also reflect reduced adaptive genetic variation, this might result in marginal populations being less locally adapted when compared to central populations, as they have less capacity to respond to local selection pressures. Maladaptation is expected to lead to poor demographic performance, reducing colonization opportunities in sites beyond the range,

and potentially creating (or reinforcing) a range edge at equilibrium along an environmental gradient (Kirkpatrick and Barton 1997).

Swamping gene flow is another often-invoked hypothesis for how equilibrial range limits might form and persist (Lenormand 2002; Sexton et al. 2009). Under swamping gene flow, peripheral populations are unable to adapt to their local conditions because they experience maladaptive gene flow from central populations (Kirkpatrick and Barton 1997). This process is predicted to occur when populations are arranged along an environmental gradient where individuals are well adapted and abundant in the center of that gradient. Because of this asymmetry in abundance, net gene flow is asymmetric and brings alleles that are adaptive in central environments to edge populations, disrupting local adaptation to edge environments. This causes edge populations to become demographic sinks, where death rates exceed birth rates, and prevents further range expansion. According to this model, the fitness of edge populations will depend upon the rate of gene flow from center to edge as well as the steepness of the environmental gradient (i.e., the magnitude of environmental differences between the sources of the gene flow and the recipient populations).

Comprehensive empirical tests of the swamping gene flow hypothesis are difficult to conduct because they require demonstrating both the negative effects of gene flow on edge populations as well as the occurrence of asymmetric gene flow on the landscape. Evidence to date indicates that in some systems swamping gene flow might limit adaptation along geographic gradients (Paul et al. 2011) or between habitat types (Anderson and Geber 2010; Farkas et al. 2016) and may sometimes limit the geographic range (Fedorka et al. 2012; Holliday et al. 2012). However, in other systems, there are no detectable fitness costs of gene flow across environmental gradients (Emery 2009; Moore and Hendry 2009; Samis et al. 2016) and strong local adaptation persists despite gene flow (Yeaman and Jarvis 2006; Gould et al. 2014). Outbreeding depression may generate patterns similar to those resulting from the disruption of local adaptation, but the effects of genetic incompatibilities can be discerned from those of swamping by experimental designs that allow for decoupling of environmental and genetic differentiation.

Most theory about swamping gene flow at range edges has been developed with the assumption of smooth environmental gradients underlying the range, however, this assumption is unrealistic for most species. Topography, continentality, and other landscape features make transects from range centers to edges heterogenous with regards to climate. Other habitat variables, such as soil type or the biotic community (which may mediate responses to climate in addition to imposing selection on their own) are also likely to be spatially heterogeneous. This complicates predictions of the swamping gene flow hypothesis: range edge populations may experience gene flow from

environmentally divergent neighboring populations, or environmentally similar central populations, as well as combinations falling anywhere in between. In this case, geography cannot be used as a proxy for predictions about the effects of gene flow; rather, these predictions must be informed by the environmental differences between populations. Gene flow between populations in similar environments may be beneficial, even when populations are geographically disparate, because gene flow can allow for the spread of environment-specific beneficial alleles that arise in a single population (Sexton et al. 2011). Abundant-center distribution patterns and asymmetric gene flow have been documented in some species, but are not ubiquitous (Sagarin and Gaines 2002), perhaps at least in part as a result of complex environmental gradients.

In addition to contributing alleles that are adaptive or maladaptive in a given environment, gene flow may provide relief from homozygosity caused by drift or inbreeding. Gene flow is expected to increase heterozygosity and reintroduce variation that can allow for masking or purging of fixed deleterious alleles. As a result, gene flow can improve fitness in peripheral populations (Sexton et al. 2011). The extent to which gene flow causes heterosis depends upon the genetic divergence of populations (Ingvarsson and Whitlock 2000), but not explicitly on the magnitude of the environmental differences between the source and recipient of gene flow, although environmental differences are correlated with genetic differentiation in some species (Sexton et al. 2014).

Gene flow may also be beneficial when maladaptation arises due to disequilibrium between a populations' optimal conditions and the environment. This could occur when a species is undergoing a range expansion, or when the environmental landscape is moving out from under individuals, as is occurring under climate change (Aitken and Whitlock 2013). If a population is locally adapted to historic conditions in a site, and the environment changes rapidly, then gene flow from populations with historic conditions that are more similar to these new local conditions is expected to improve population performance.

To investigate how gene flow affects peripheral populations, we simulated gene flow among populations spanning the northern half of the range of an annual wildflower, *Clarkia pulchella*, and measured lifetime fitness of individuals in two common gardens at the species' northern range edge. We asked: (1) Are range edge populations of *C. pulchella* locally adapted? (2) What climatic factors predict fitness at the northern range edge? (3) Does gene flow positively or negatively affect edge populations? And (4) How does the effect of gene flow from other populations depend upon the genetic differentiation and climatic distances of these populations? Under conditions where the range edge is not at equilibrium with climate, we expect that gene flow from sites that are historically similar to the experimental conditions will improve

performance. Under conditions in which this species' range is at equilibrium with climate, and if this edge is limited by adaptation, we expect that gene flow from populations in similar climates will have a positive effect on fitness via heterosis or the contribution of adaptive alleles, but that gene flow from strongly contrasting climates will be detrimental. If populations have genetic incompatibilities (which need not be the result of divergent selection, but could simply be the result of drift under prolonged separation) then we expect the offspring of crosses between populations that are more genetically divergent to perform worse, regardless of the conditions of the test environment. However, if homozygosity is negatively related to fitness, we expect a greater benefit from gene flow between more genetically divergent populations.

Methods
STUDY SYSTEM, SEED COLLECTION, AND SITE SELECTION

Clarkia pulchella Pursh (Onagraceae) is a winter annual that grows on sparsely vegetated, south-facing slopes with low canopy cover throughout eastern Washington and Oregon, Idaho, and western Montana (United States) and southeastern British Columbia (Canada). This species germinates in fall and overwinters as a seedling before flowering in late May, June, and early July. It has no observed seed dormancy, but seeds will not germinate immediately upon dehiscing and require an after-ripening period of several weeks. It has showy pink flowers and is visited by a diverse array of pollinators, although it has some capacity to self-pollinate in the absence of pollinators or mates (MacSwain et al. 1973; Bontrager et al. 2018). Individual plants that survive to reproduction typically produce 1–35 flowers (mean 3.11, SD 3.07; M. Bontrager, unpublished data).

Seeds of *C. pulchella* were collected from 15 populations in July 2014 (Fig. 1A; Table S1). We used the two northwestern-most populations of the continuous distribution of *C. pulchella* as common garden sites (hereafter referred to as focal populations). Other populations (hereafter, donor populations) were selected with the goal of sampling representative variation in major climatic axes (temperature, precipitation, and seasonality of these variables; Fig. 1B) across the northern half of the species' range. In each of these populations, seeds were collected haphazardly from plants spaced >0.5 m apart.

GREENHOUSE GENERATION AND CROSSING DESIGN
We grew field-collected seeds in the greenhouse and implemented a controlled crossing design to generate seeds for the field transplant. We planted maternal families from each population into 22 randomized blocks, with each block containing one family from each population. Two types of crosses were performed:
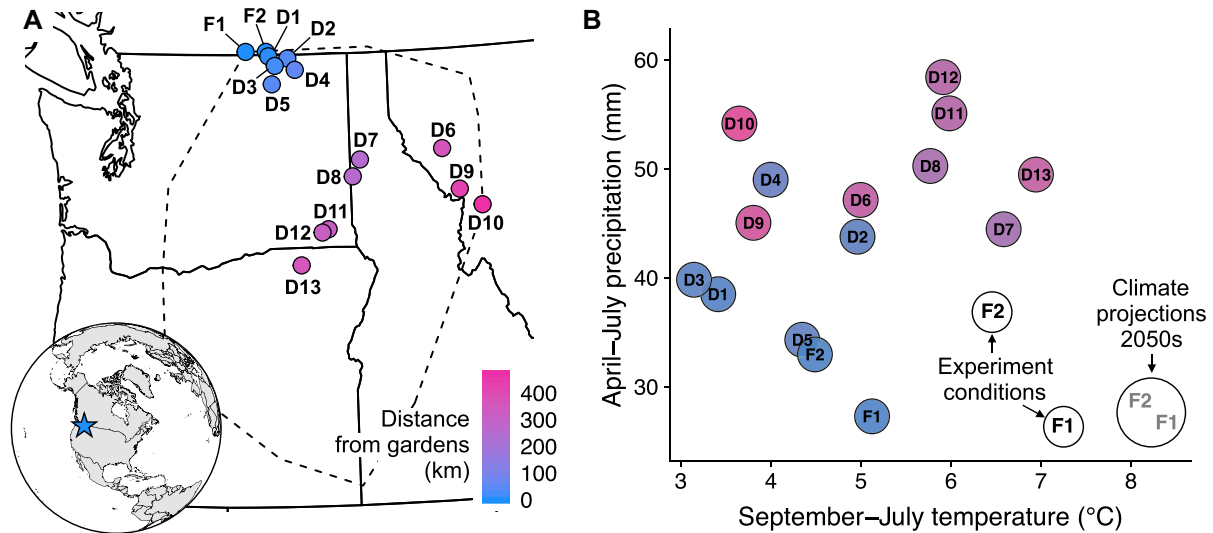
Figure 1. Geographic locations and climate averages of populations used in this experiment. (A) Populations span the northern half of the geographic range of *Clarkia pulchella* (indicated by the dashed line; the range is represented by the extent of localities in the Consortium of Pacific Northwest Herbaria database). Focal population sites (where common gardens were installed) are indicated by "F" and donor population sites are indicated by "D". Identifying codes for each population correspond to the Map ID column in Table S1. (B) Temperature and precipitation conditions of common gardens and source provenances (calculated from monthly PRISM data; PRISM Climate Group 2017). Colored circles represent the 1951-1980 average September-July temperature and summed April-July precipitation in the home site of each population. Open circles represent weather conditions during the experiment. Focal populations are historically intermediate relative to donor populations in average historic temperature (*x*-axis), but are from the driest sites of any population used in the experiment (*y*-axis). Conditions in common gardens during the experiment were hot relative to normal conditions at those sites, and hot and dry relative to average conditions of all populations in the experiment. Conditions during the experiment were not quite as warm as projected temperatures for 2055 under the HadGEM2-ES RCP 4.5 model (grey text; downscaled using ClimateNA, Wang et al. 2016).

"within-population" crosses and "between-population" crosses. For within-population crosses, dams were pollinated using pollen from the plant of the same population in the subsequent block (Fig. 2A). Each plant from each of the 15 populations was therefore used as both a sire and a dam with other plants from the same population. For between-population crosses, flowers on plants from the two focal populations within each block were pollinated using each of the donor plants in that block (Fig. 2B). These crosses simulate an early stage of gene flow: the progeny of a mating event that is the result of long distance pollen dispersal (or the progeny of a cross between a native individual and a recent immigrant). We collected the seeds from crosses as fruits ripened. Further details about our greenhouse conditions and crossing design are in Supporting Information Methods 1.

COMMON GARDEN DESIGN AND INSTALLATION

For the transplant, we used seeds from 15 of our greenhouse blocks, and substituted seeds from the same type of cross from other greenhouse blocks when they were unavailable from our primary 15. Seeds were glued to toothpicks to expedite planting and monitoring in the field. Two seeds were glued to each toothpick

with a tiny dab of water-soluble glue (when seeds were limited, just one seed was glued to each toothpick). At each of the two sites, toothpicks were planted into 10 fully randomized plots. Each plot contained two toothpicks from each cross type from each of the 15 replicates. We only planted between-population crosses with local dams at each of the two focal sites (i.e., Blue Lake plots only contained between-population crosses performed on Blue Lake plants, and Johnstone Creek plots only contained between-population crosses performed on Johnstone Creek plants). Within-population crosses from all populations were planted out at both sites. Therefore, each plot contained two replicates of each of 15 crosses from 29 cross types (14 between-population groups and 15 within-population groups). For some cross types, less than 15 families had sufficient seeds for the full design, therefore, each plot contained 832 toothpicks at Johnstone Creek and 836 toothpicks at Blue Lake. In total, our design included 16,680 toothpicks and 32,755 seeds. Seeds were planted in September 2015 and watered once in late October, although at that time most seeds that were checked had already begun germinating. Plots were protected with wire cages; further details of transplant installation are in Supporting Information Methods 2.
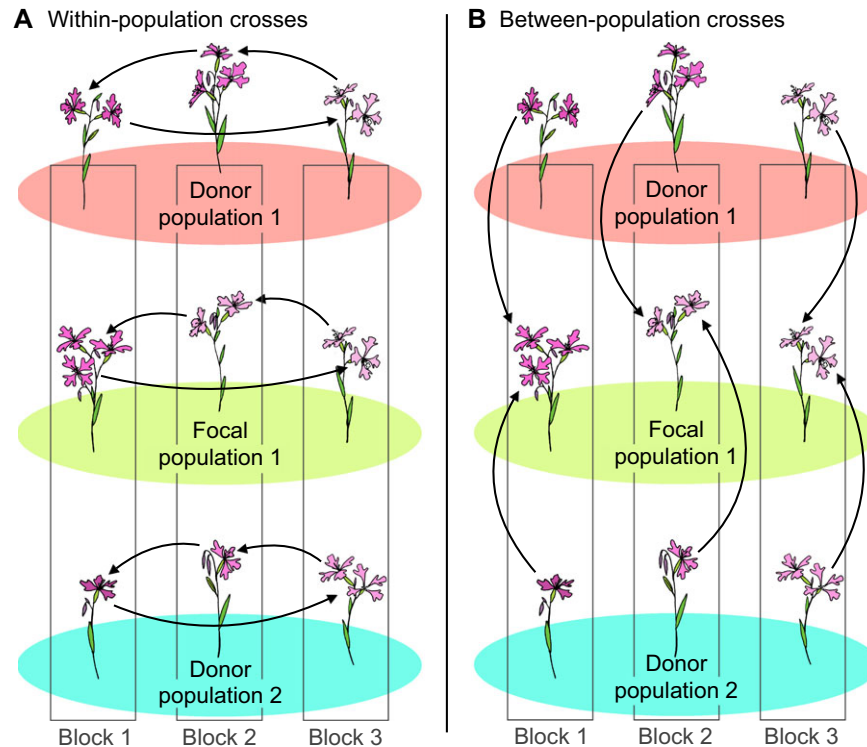
Figure 2. Schematic of the greenhouse crossing design to generate seeds for our common gardens. Three blocks are diagrammed here, but crosses were performed in 22 blocks and seeds from 15 blocks were transplanted into the field. (A) Within-population crosses: for each seed family from each of 15 populations (three are shown), plants were crossed in a "daisy chain" design, in which each plant was hand-pollinated using pollen from another individual of the same population from the subsequent block. (B) Between-population crosses: we used pollen from each of 13 donor populations (two are shown) to pollinate flowers on plants from each of the two focal populations (one is shown). Each focal plant served as a dam for multiple between-population crosses, that is, each focal seed family had one flower pollinated by a plant from each of 13 donor populations and from the other focal population.

MONITORING AND MEASURING

Germination was censused 16–20 November 2015. We documented the emergence of either zero, one, or two germinants at each toothpick. If two germinants were present, these were randomly thinned so that just one remained. Overwinter survival was assessed 17–21 March 2016. During both germination and overwinter survival censuses, plant size was measured as the distance in millimeters from the tip of one leaf to the tip of the other (most plants had just one pair of leaves). Further details of monitoring and measuring are provided in Supporting Information Methods 3. In early June 2016, we censused the spring survival of all plants and began measuring reproduction every 2–3 days. As a proxy for seed production, we measured the ovary length of each flower produced on each plant. Pollinators were excluded from our plots to prevent the escape of nonnative alleles so we calibrated a conversion from ovary length to seed production using hand pollinations (linear regression of seed production on ovary length: $R_0^2 = 0.87$, $P < 0.0001$; Supporting Information Methods 4). We continued reproductive censuses until all plants had senesced, but reduced census frequency when flowering slowed in July.

CLIMATIC AND GENETIC DIFFERENTIATION

We compiled monthly temperature and precipitation data from 1951 to 1980 for all seed sources, as well as the gardens during the months of the experiment (September 2015 to July 2016) from PRISM (PRISM Climate Group 2017). We calculated historic (prewarming, 1951–1980) climate averages for each site, which we compared to conditions experienced by plants during the transplant in our analyses. Our decision to use normals from this time window, rather than a more recent time window, reflects an assumption that populations have been limited in their capacity to adapt to changing conditions, such that historic normals best represent each population's climate optima. However, we note that contemporary normals (1984–2013) are highly correlated with those that we used (September–July temperature: $r = 0.99$, $P < 0.0001$; April–July precipitation: $r = 0.96$, $P < 0.0001$, Figs. S1 and S2), so our choice of time period is unlikely to affect our results. Inter-annual climate variability was very similar in each of the seed collection sites (data not shown), so we do not further consider variability in climate, and focus on averages only. Pairwise population genetic differentiation

(F_{ST}) was calculated from 2982 SNPs that were genotyped in up to 12 parental individuals from each population (genotyping and alignment methods are described in Bontrager and Angert 2018). F_{ST} was calculated using the implementation of Weir and Cockerham (1984) in the R package hierfstat (Goudet and Jombart 2015).

STATISTICAL ANALYSES

Did local populations outperform foreign populations?

We tested whether local populations were, on average, superior to foreign populations using within-population crosses only. We compared lifetime fitness of local versus foreign individuals in a generalized linear mixed model (GLMM) with a zero-inflated negative binomial distribution using the package glmmTMB (Brooks et al. 2017). These zero-inflated models allow specification of fixed effects for both the zero-inflation part of the model (the probability of a non-zero value) as well as the conditional part of the model (the effect on the response after accounting for zero-inflation). Generally, we consider the zero-inflation part of the model to reflect early lifestages, as the majority of plants that produced zero seeds did so as a result of failing to germinate or survive winter. The conditional part of the model may reflect both differences during reproduction as well as differences among individuals accumulated across all lifestages. In addition to testing for local adaptation represented by lifetime fitness, we tested whether local populations performed better than foreign populations at any component lifestage: germination, size after germination, overwinter survival, size after winter, fruit count, and estimated seed production. Plant size was modeled with a Gaussian response distribution, germination and survival with binomial response distributions and logit link functions, and fruit counts and seed production with zero-inflated negative binomial response distributions and log link functions. For component lifestage analyses, we included only individuals that had survived to the preceding census and always included plant size at the previous census to account for differences that had accumulated at earlier lifestages. For all of these models, we initially included a random effect structure of block within site, dam within dam population, and sire within sire population. However, models of later lifestages and lifetime fitness frequently failed to converge with this parameterization. When convergence failed, we reduced random effects to only block within site and sire population. Predictions and confidence intervals were visualized with ggeffects (Lüdecke 2018).

Does climate of origin explain performance in common gardens?

We built GLMMs using the methods described above to evaluate the effects of mismatch between the experiment conditions and populations' historic climates on lifetime fitness and all fitness components, again using only within-population crosses. We calculated the absolute difference between the garden conditions and the historic temperature and precipitation of each source population. We use absolute differences because we expect that mismatch in either direction along a climate axis will negatively impact fitness. Very few source populations were from sites drier or hotter than conditions during the experiment, so absolute differences mostly result from source populations being historically cooler or wetter than the experiment. Our lifetime fitness model included absolute differences in temperature (for the experiment duration, September–July) in both the conditional and zero-inflation parts of the model, as well as absolute differences in spring and summer precipitation (April–July) in the conditional part of the model only. We isolated the specific lifestages affected by each of these climatic predictors using the methods described above. In lifestage-specific analyses, we calculated climate differences using only the months in each census window. In all analyses, all continuous predictors were scaled and centered.

Does gene flow help or hurt edge populations?

Based on the results of the analyses above, we expected that gene flow from some populations was likely to confer benefits by contributing adaptive genetic variation to focal populations experiencing an anomalous climate. To evaluate whether there were benefits of gene flow that were independent of these climate effects, we calculated the midparent historic temperature average for all individuals (i.e., the average temperature of dam and sire sites) and then calculated the absolute difference between this temperature and the experimental temperature (Fig. S3). We calculated a metric of absolute midparent precipitation difference (summed over April–July) in the same manner. We used GLMMs as described above to test for an effect of gene flow in addition to an anticipated effect of midparent climate differences on lifetime fitness and each component lifestage. In these models, gene flow was included as a categorical fixed effect (within-population cross vs. between-population cross) along with midparent temperature and precipitation differences. We included gene flow and temperature differences in both the conditional and zero-inflation parts of the lifetime fitness model, and precipitation differences in only the conditional part. We had difficulty disentangling effects of precipitation differences and gene flow (see Results) so we ran these models with and without precipitation differences.

Do the effects of gene flow depend upon the genetic differentiation between focal and donor populations?

We examined whether the genetic differentiation (F_{ST}) between the two parental populations of the between-population crosses positively or negatively affected offspring fitness. We could only

estimate genetic differentiation between parental populations for individuals with parents from different populations, so in these analyses we use between-population crosses only. We built zero-inflated GLMMs as described above using lifetime fitness and included predictors of absolute midparent temperature and precipitation differences as well as F_{ST}. We also tested the effects of these parameters on each component lifestage. Our ability to detect significant effects of climate in full models was limited, likely due to the narrow range of midparent climatic variability across between-population crosses, so we also built separate models of each of our three predictors on lifetime fitness and each lifestage. All statistical analyses were implemented in R version 3.4.3 (R Core Team 2017).

Results

CLIMATE OF ORIGIN EXPLAINS PERFORMANCE IN COMMON GARDENS

Local populations were not superior to the average foreign population in their cumulative fitness across all lifestages, or in any component lifestage (Fig. S4, Table S2). Populations that were best matched to experimental temperatures performed best in our gardens; lifetime fitness declined with increasing absolute temperature differences between the source and the experimental conditions (Fig. 3A). This occurred via effects on both the probability of producing any seeds (the zero-inflation part of the model; $\beta = -0.341$, SE $= 0.033$, $P < 0.001$; Fig. 3B), and the number of seeds produced (the conditional part of the model; $\beta = -0.114$, SE $= 0.050$, $P = 0.022$; Fig. 3C). Local populations, which are historically intermediate in temperature (Fig. 1B), were mismatched from the experiment conditions and performed worse than populations from warmer sites that were more climatically similar to the garden conditions (Supporting Information Analysis 1).

Analyses of component lifestages support these inferences (Fig. 3D, Table S3): being poorly matched to experimental temperatures had negative effects on germination proportion, overwinter survival, and the size of plants after winter. While precipitation differences were not significant in the model of lifetime fitness ($\beta = -0.067$, SE $= 0.049$, $P = 0.177$, Fig. 3C), they did have a negative effect on seed production among plants surviving the winter (Fig. 3D, Table S3).

GENE FLOW MAY CONFER SOME BENEFITS TO EDGE POPULATIONS

As in the analyses of within-population plants only, both midparent temperature differences and midparent precipitation differences had negative effects on lifetime fitness in our common gardens (Table S4A; Fig. 4AB). Gene flow (i.e., being a between-population rather than a within-population cross) did not have a

significant effect in the lifetime fitness model that also included climate differences.

It is difficult to disentangle the effects of precipitation differences from the effects of gene flow in these analyses. This is because our focal populations are already among the driest provenances in our experiment. Therefore, the average between-population plant is better matched to the experimental conditions than the average within-population plant, because the midparent precipitation of between-population plants is always calculated with at least one very dry focal parent (Fig. S3A). This was not an issue with temperature differences, because our focal populations are intermediate to other provenances in terms of temperature (Fig. S3B). When lifetime fitness was analyzed without precipitation differences in the model, we found that gene flow (being a between-population cross, rather than a within-population cross), had a positive effect on lifetime fitness in addition to effects of temperature (Table S4B).

A small positive effect of gene flow, independent of climatic differences, is supported by analyses of some lifestage components (Fig. 4C; Table S4C). Negative effects of precipitation and temperature differences were similar to those found in the analyses of climatic drivers of performance, while gene flow (i.e., being from a between-population vs. a within-population cross) had a positive effect on fruit production and a marginal positive effect on seed production.

GENETIC DIFFERENTIATION BETWEEN PARENTAL POPULATIONS POSITIVELY AFFECTS FITNESS

Both midparent temperature difference from the garden conditions and genetic differentiation between parental populations had significant effects on fitness (Fig. 5, Tables S5AB and S6A). Genetic differentiation between parental populations had a positive effect on lifetime fitness via increasing the probability of producing seeds (the zero-inflation part of the model). The effects of genetic differentiation on lifetime fitness were mirrored in the analyses of single lifestages: F_{ST} had a positive effect on germination, size after winter, fruit count, and seed production (Fig. 5C, Tables S5A and S6A).

The negative effect of midparent temperature differences in between-population crosses is generally consistent with our analyses of climatic drivers of performance in within-population crosses. Between-population plants with donor parents that were well-matched to temperatures during the experiment were more likely to produce seeds, as indicated by the significant negative effect of midparent temperature differences in the zero-inflation part of that model (Table S5B and S6A). Midparent temperature differences did not significantly affect any single lifestage, but had marginal negative effects on size after germination, size after winter, and fruit number (Fig. 5C, Table S5B). Midparent
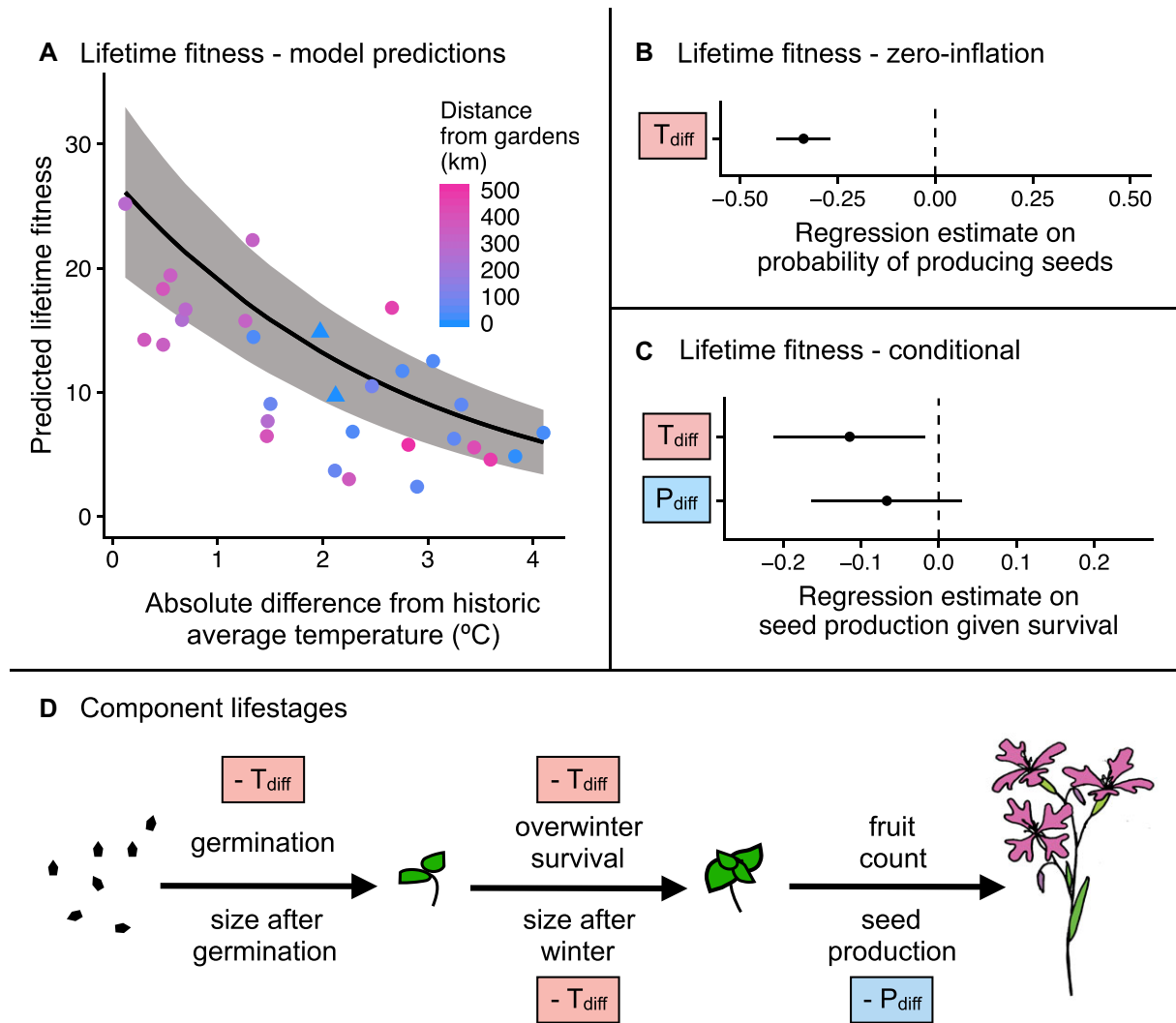
Figure 3. Effects of absolute temperature difference (September–July; T_{diff}) and absolute precipitation difference (April–July; P_{diff}) on performance of *Clarkia pulchella* in common gardens. These analyses include within-population plants only (no gene flow). (A) Lifetime fitness (number of seeds produced by each seed that was planted) declines with increasing differences in temperature between the historic average of the source population and the experimental conditions in the transplant gardens. The regression line and 95% confidence interval incorporate both the conditional and zero-inflation model components; the confidence interval is conditioned on fixed effects only. Though these temperature differences are expressed as absolute, almost all populations were from sources that are historically cooler than the transplant sites were during the experiment. Points are raw averages for each source population in each garden, colored by the distance between the source population and the transplant garden. Triangles are focal populations in their home sites. (B) Regression estimates and standard errors from the zero-inflation part of a model of lifetime fitness. (C) Regression estimates and standard errors from the conditional part of a model of lifetime fitness. (D) Schematic of effects of absolute temperature differences (T_{diff}) and absolute precipitation difference (P_{diff}) on component lifestages of *Clarkia pulchella*. Directionality of effects is illustrated with "-"; in these analyses all significant effects were negative. Predictors in boxes are significant ($P < 0.05$). Size in the previous lifestage is not shown here, but has a significant positive effect on overwinter and reproductive lifestages. This summarizes the significant results of separate models for each lifestage; full statistical results of these tests are in Table S3.

precipitation differences did not significantly affect lifetime fitness or component lifestages (Table S5C).

It is important to note that genetic differentiation is not especially correlated with temperature differences ($r = -0.25$, $P = 0.193$), although genetic differentiation and precipitation differences are correlated ($r = 0.64$, $P = 0.0003$); plants whose parents are more genetically differentiated also have having larger differences between their historic midparent precipitation and conditions during the experiment. We think it is unlikely that the apparent positive effects of genetic differentiation are actually driven by precipitation differences, because we would expect high precipitation differences to negatively affect fitness. The
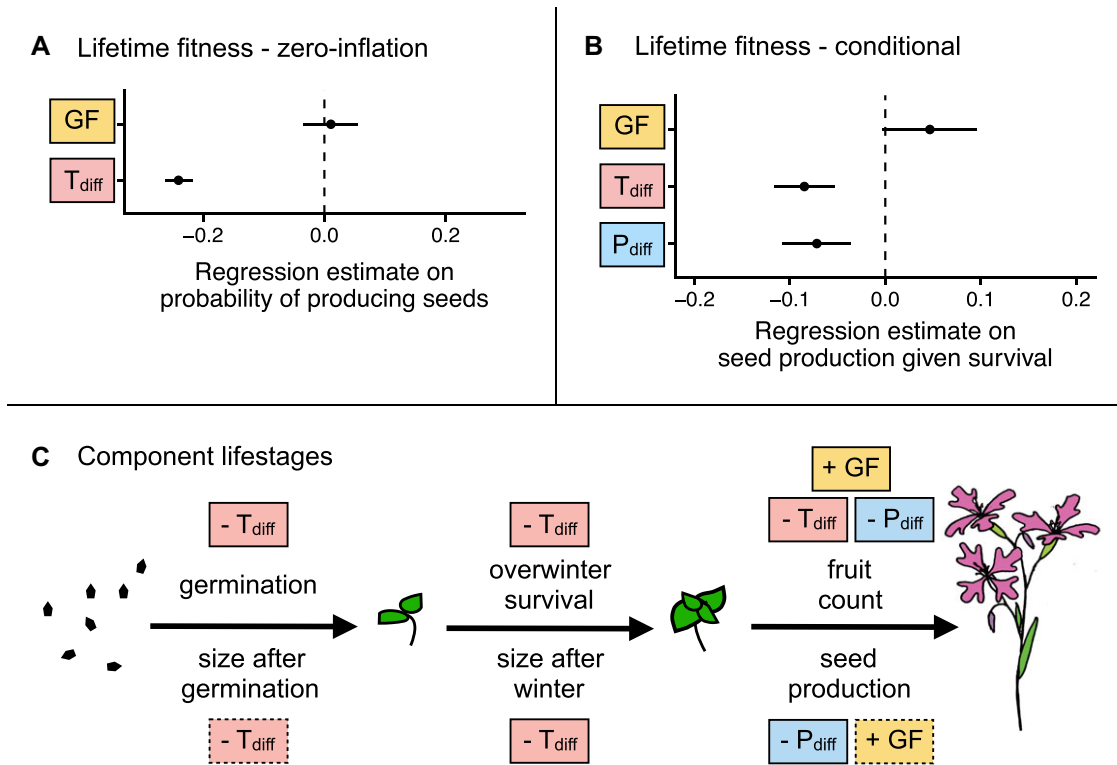
Figure 4. Effects of gene flow (differences between between-population and within-population crosses) on performance of *Clarkia pulchella*, accounting for midparent temperature and precipitation. Lifetime fitness (the number of seeds produced by each seed that was planted) incorporates germination, survival, and reproduction. (A) Regression estimates and standard errors from the zero-inflation part of a model of lifetime fitness. (B) Regression estimates and standard errors from the conditional part of a model of lifetime fitness. (C) Effects of gene flow (GF), absolute midparent temperature differences (T_{diff}), and absolute midparent precipitation differences (P_{diff}) on component lifestages of *Clarkia pulchella*. Directionality of effects is illustrated with "+" and "-". Not quite significant parameters (0.05 < *P* < 0.10) are shown in boxes with dashed margins, predictors in solid boxes are significant (*P* < 0.05). Size in the previous lifestage is not shown here, but has a significant positive effect on overwinter and reproductive lifestages. This summarizes the significant results of separate models for each lifestage; full statistical results of these tests are in Table S4.

overall weak or absent effects of temperature and precipitation differences in these models may be due to a narrower range of variation in midparent climate for the between-population crosses relative to the within-population crosses (Fig. S3).

Discussion

We conducted a common garden experiment at the northern range margin of *Clarkia pulchella* to examine how the effects of gene flow on peripheral populations vary with climatic and genetic differentiation between focal and source populations. We examined predictors of fitness of within-population crosses, in which both parents originated from the same source population, as well as between-population crosses, in which one parent was local to the common gardens and the other was from another population from across the northern half of the range of *C. pulchella*. In our experiment, provenances of *C. pulchella* from climates that most closely matched the conditions during the experiment performed

best. Populations of *C. pulchella* at the northern range margin benefited from gene flow from warm source locations during the warm year of our experiment. Gene flow also conferred some benefits independent of climate, as evidenced by the positive effect of gene flow on later lifestages when controlling for temperature differences and the positive effect of increasing genetic differentiation between the parental populations.

CLIMATE OF ORIGIN PREDICTS PERFORMANCE

Populations of *Clarkia pulchella* are adapted to their historic climate regimes, a pattern consistent with findings in many other species (Anderson et al. 2015; Wilczek et al. 2014). When grown in common sites, the performance of individuals was determined by the degree to which conditions during the experiment deviated from historic temperature and precipitation averages of each provenance (Fig. 3). Because of this local adaptation to climate, gene flow from sites that deviate from local conditions (in our experiment, sites that are cooler than the focal populations) had
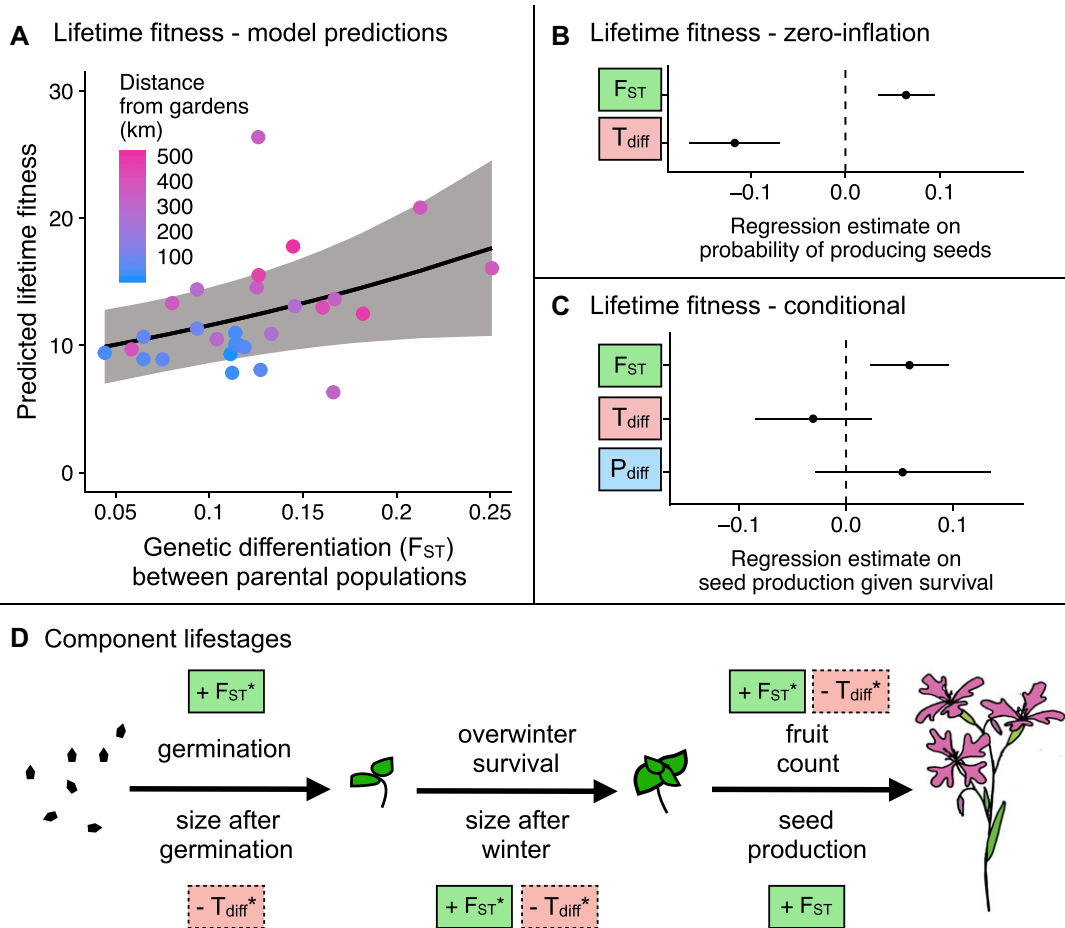
Figure 5. Effects of genetic differentiation between parental populations, as well as midparent temperature and precipitation on performance of *Clarkia pulchella*. **(A)** Among between-population crosses, increased genetic divergence between parental populations had a positive effect on lifetime fitness (number of seeds produced by each seed that was planted). The regression line and 95% confidence interval show the effects of F_{ST} when T_{diff} and P_{diff} are held constant. These incorporate both the conditional and zero-inflation model components; the confidence interval is conditioned on fixed effects only. Points are raw averages for each gene flow source in each garden, colored by the distance between the source population and the garden (the home of the recipient population). **(B)** Regression estimates and standard errors of genetic differentiation (F_{ST}) and absolute midparent temperature differences (T_{diff}) on the probability of producing seeds. **(C)** Effects of genetic differentiation, temperature differences, and absolute precipitation differences (P_{diff}) on conditional seed production. **(D)** Effects of T_{diff} and F_{ST} on component lifestages of *Clarkia pulchella*. Precipitation differences were not significant when tested for component lifestages. Directionality of effects is illustrated with "+" and "-". Not quite significant parameters ($0.05 < P < 0.10$) are shown in boxes with dashed margins, predictors in solid boxes are significant ($P < 0.05$). Size in the previous lifestage is not shown here, but has a significant positive effect on overwinter and reproductive lifestages. * indicates predictors that are only significant in separate models, not in full models with all predictors. Complete statistical results of these tests are in Table S5 and Table S6.

the potential to disrupt local adaptation, as indicated by the somewhat negative effects of midparent temperature differences on between-population plants (Fig. 5, Table S5). Although simulated gene flow from divergent environments had the potential to reduce fitness in this experiment, it is unlikely that gene flow occurs at a high enough rate among natural populations to swamp local adaptation. However, our results highlight that gene flow and dispersal need not be from populations that are geographically distant (or from the center of the range) to be climatically divergent from

historic or current conditions. Rather, two of the most temperature-mismatched populations used in the experiment are from sites nearest to our common gardens (populations D1 and D3; Fig. 1).

Under climate change, local adaptation to historic climate regimes may generate local maladaptation in field trials. We see this in our results: populations from warmer locations performed best in our gardens (Fig. 3, Table S3) and outperformed local populations (Supporting Information Analysis 1 and Fig. S5), and

gene flow from warmer locations had positive effects on some lifestages (Fig. 5, Table S5). This lagging adaptation to climate has been documented in other recent common garden studies. In a reciprocal transplant experiment of a long-lived sedge, McGraw et al. (2015) found that populations were displaced 140 km south of their optimal climate conditions. Wilczek et al. (2014) found that local genotypes of *Arabidopsis thaliana* from across Europe were consistently outperformed in common gardens by accessions from historically warmer locations. These results indicate that dispersal and gene flow are important processes promoting range stasis as climate warms, as they allow alleles that are beneficial in warm environments to spread from historically warm populations to recently warming sites. However, climate is multivariate, and as the climate changes it may generate combinations of conditions that no population has historically experienced (Williams and Jackson 2007; Mahony et al. 2017). The particular combination of hot and dry conditions in our common gardens was unlike any of our populations' historic temperature and precipitation combinations and approached future climate projections for these sites (Fig. 1), although they are similar to normals from some populations not included in our experiment, primarily from the southern half of the species range (data not shown). While precipitation conditions were similar to those historically experienced by the focal populations, temperature conditions favored another set of populations. Whether the optimal traits for different climatic axes are antagonistic and whether segregation and recombination will allow adaptation to novel climates are important considerations in predicting climate change responses.

GENE FLOW CONFERS BENEFITS INDEPENDENT OF CLIMATE

We saw some additional positive effects of gene flow once the effects of climate are controlled for (Fig. 4, Table S4). These positive effects may be the result of reduced homozygosity when parental plants come from two different populations; this inference is supported by the positive effect of genetic differentiation between parental populations on performance (Fig. 5, Table S5). These results are also generally consistent with previous work in which experimental populations of *Clarkia pulchella* with higher genetic effective population sizes had lower extinction probabilities (Newman and Pilson 1997). An interesting direction for future models of swamping gene flow along environmental gradients might be to explore whether incorporating heterosis-dependent increases in the effective migration rate (Ingvarsson and Whitlock 2000) alters predictions (this could be done with various dispersal distances, under scenarios of various magnitudes of isolation-by-distance). However, an important question is whether the benefits of reduced homozygosity (or increased heterozygosity) are transient effects among F1s, how long they would persist in

future generations if our between-population plants backcrossed into the focal populations, and whether these benefits may be counteracted by outbreeding depression as recombination disrupts co-adapted gene complexes. The answers to these questions are likely to depend on many factors, including the genetic architecture of local adaptation and population size (Willi et al. 2007), but fitness declines in subsequent generations are not uncommon after between-population crosses (Fenster and Galloway 2000; Johansen-Morris and Latta 2006). Novel environments may alter the costs and benefits of outbreeding: increases in variation among individuals might help populations adapt, despite temporary decreases in mean fitness due to outbreeding depression.

During this study, the effects of being well matched to the experimental conditions seemed to dominate over potential benefits of being from a local population (e.g., the benefits of being adapted to local soil conditions or herbivores). This inference is supported by the fact that lifetime fitness of local populations did not differ from that of foreign populations even once climate differences were controlled for (Supporting Information Analysis 2, Table S7), although our experiment was not especially well suited to test this because we have only two local populations.

LIMITED INFERENCE ABOUT POPULATION PERSISTENCE

Our ability to make inferences from our results about the longer-term effects of gene flow on the persistence and adaptive potential of range edge populations is limited. It seems clear that gene flow from warm sites is likely to accelerate adaptation to warming conditions, but we do not know whether these populations were historically limited by adaptation and whether the additional genetic variation introduced by gene flow would permit better adaptation to local conditions and range expansion on an evolutionary time scale. These types of questions are difficult to test in field systems (but see Etterson and Shaw 2001), but inferences can be made by examining genetic variance of wild populations in the laboratory (Kellermann et al. 2006; Hoffmann et al. 2003). The development of experimental evolution systems to test equilibrial range dynamics is an exciting avenue for future work—this would be a natural extension of recent studies of range expansion dynamics using experimental evolution in the laboratory (Ochocki and Miller 2017; Williams et al. 2016).

It is also important to note that all populations in our experiment had reproductive rates that were well above replacement (one seed produced per seed planted, see *y*- axes on Figs. 3A and 5A), so we have no evidence that gene flow has the potential to drive populations extinct or to rescue them from extinction. The high lifetime fitness we observed during our experiment could be due to several factors. First, perhaps warm conditions over the entire season are favorable for all sources, but are more favorable for warm-adapted

populations. Alternatively, we could have increased fitness by limiting antagonistic biotic interactions. Additionally, our conversion from ovary length to seed production estimates assumes that pollinator service is at least as effective as our hand pollinations and that seed production is not resource-limited. If preventing pollinator access increased fruit production due to within-plant resource reallocation, our estimates of fitness from ovary length could be overestimates. Finally, and perhaps most plausibly, our plot placement may have upwardly biased our fitness estimates. We placed plots in patches that appeared favorable to *C. pulchella*, but naturally dispersing seeds are likely to land in a mix of favorable and unfavorable patches. We do not know whether any of these factors interact with provenance, in which case they might change the relative performances of populations in our experiment.

CONCLUSIONS AND FUTURE DIRECTIONS

On heterogeneous landscapes, the effects of gene flow are likely to depend strongly on climate of origin, which may be decoupled from geography. Although gene flow from divergent climates may result in reduced fitness, gene flow can also confer benefits to range edge populations via both preadapted alleles and reduced homozygosity. As a result, moderate rates of gene flow likely represent a net benefit to edge populations, particularly as the environment is changing. This study highlights the challenges of testing hypotheses about equilibrial range limits in the field, where climate change is a persistent reality. Even if populations were once locally adapted, they are likely no longer at equilibrium with climate. The signal of climate anomalies disrupting local adaptation can be detected in published literature to date (Bontrager et al., in prep.). In light of this, future studies of local adaptation at range edges should be designed in such a way that the results will be informative even in non-equilibrial conditions.

AUTHOR CONTRIBUTIONS
M.B. and A.L.A. conceived of the project. M.B. performed greenhouse crosses, garden installation, and data collection. M.B. analyzed the data in consultation with A.L.A. M.B. wrote the manuscript with advice and edits from A.L.A.

DATA ARCHIVING
All data and code required to recreate these analyses is publicly available at github.com/meganbontrager/clarkia-pulchella-transplant. All data and code are archived on Dryad (https://doi.org/10.5061/dryad.1b90t69).

LITERATURE CITED

Aitken, S. N., and M. C. Whitlock. 2013. Assisted gene flow to facilitate local adaptation to climate change. Annu. Rev. Ecol. Evol. System. 44:367–388.

Anderson, J. T., and M. A. Geber. 2010. Demographic source-sink dynamics restrict local adaptation in elliott's blueberry (*Vaccinium elliottii*). Evolution 64:370–384.

Anderson, J. T., N. Perera, B. Chowdhury, and T. Mitchell-Olds. 2015. Microgeographic patterns of genetic divergence and adaptation across environmental gradients in *Boechera stricta* (Brassicaceae). Am. Nat. 186:S60–S73.

Antonovics, J. 1976. The nature of limits to natural selection. Ann. Mo. Bot. Gard. 63:224–247.

Bontrager, M., and A. L. Angert. 2018. Genetic differentiation is determined by geographic distance in clarkia pulchella. bioRxiv P. 374454.

Bontrager, M., C. D. Muir, and A. Angert. 2018. Geographic and climatic drivers of reproductive assurance in *Clarkia pulchella*. bioRxiv P. 372375.

Bridle, J. R., and T. H. Vines. 2007. Limits to evolution at range margins: when and why does adaptation fail? Trends Ecol. Evol. 22:140–147.

Brooks, M. E., K. Kristensen, K. J. van Benthem, A. Magnusson, C. W. Berg, A. Nielsen, H. J. Skaug, M. Machler, and B. M. Bolker. 2017. glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. R J. 9:378–400.

Eckert, C., K. Samis, and S. Lougheed. 2008. Genetic variation across species' geographical ranges: the central–marginal hypothesis and beyond. Mol. Ecol. 17:1170–1188.

Emery, N. C. 2009. Ecological limits and fitness consequences of cross-gradient pollen movement in *Lasthenia fremontii*. Am. Nat. 174:221–235.

Etterson, J. R., and R. G. Shaw. 2001. Constraint to adaptive evolution in response to global warming. Science 294:151–154.

Farkas, T. E., T. Mononen, A. A. Comeault, and P. Nosil. 2016. Observational evidence that maladaptive gene flow reduces patch occupancy in a wild insect metapopulation. Evolution 70:2879–2888.

Fedorka, K., W. Winterhalter, K. Shaw, W. Brogan, and T. Mousseau. 2012. The role of gene flow asymmetry along an environmental gradient in constraining local adaptation and range expansion. J. Evol. Biol. 25:1676–1685.

Fenster, C. B., and L. F. Galloway. 2000. Population differentiation in an annual legume: genetic architecture. Evolution 54:1157–1172.

Goudet, J., and T. Jombart. 2015. hierfstat: Estimation and tests of hierarchical F-statistics. URL https://CRAN.R-project.org/package=hierfstat. R package version 0.04-22.

Gould, B., D. A. Moeller, V. M. Eckhart, P. Tiffin, E. Fabio, and M. A. Geber, 2014. Local adaptation and range boundary formation in response to complex environmental gradients across the geographical range of *Clarkia xantiana* ssp. *xantiana*. J. Ecol. 102:95–107.

Hoffmann, A., R. Hallas, J. Dean, and M. Schiffer. 2003. Low potential for climatic stress adaptation in a rainforest *Drosophila* species. Science 301:100–102.

Holliday, J. A., H. Suren, and S. N. Aitken. 2012. Divergent selection and heterogeneous migration rates across the range of Sitka spruce (*Picea sitchensis*). Proc. R S. Lond. B Biol. Sci. 279:1675–1683.

Ingvarsson, P. K., and M. C. Whitlock. 2000. Heterosis increases the effective migration rate. Proc. R. Soc. Lond. B Biol. Sci. 267:1321–1326.

Johansen-Morris, A., and R. G. Latta. 2006. Fitness consequences of hybridization between ecotypes of *Avena barbata*: hybrid breakdown, hybrid vigor, and transgressive segregation. Evolution 60:1585–1595.

Kellermann, V. M., B. van Heerwaarden, A. A. Hoffmann, and C. M. Sgrò, 2006. Very low additive genetic variance and evolutionary potential in multiple populations of two rainforest *Drosophila* species. Evolution 60:1104–1108.

Kirkpatrick, M., and N. H. Barton. 1997. Evolution of a species' range. Am. Nat. 150:1–23.

Lee-Yaw, J. A., H. M. Kharouba, M. Bontrager, C. Mahony, A. M. Csergő, A. M. Noreen, Q. Li, R. Schuster, and A. L. Angert. 2016. A synthesis of transplant experiments and ecological niche models suggests that range limits are often niche limits. Ecol. Lett. 19:710–722.

Lenormand, T. 2002. Gene flow and the limits to natural selection. Trends Ecol. Evol. 17:183–189.

Lüdecke, D. 2018. ggeffects: Tidy data frames of marginal effects from regression models. J. Open Source Softw. 3:772.

MacSwain, J. W., P. H. Raven, and R. W. Thorp. 1973. Comparative behavior of bees and Onagraceae: IV. *Clarkia* bees of the Western United States. Univ. Calif. Publ. Entomol. 70:1–80 .

Mahony, C. R., A. J. Cannon, T. Wang, and S. N. Aitken. 2017. A closer look at novel climates: new methods and insights at continental to landscape scales. Global Change Biol. 23:3934–3955.

McGraw, J. B., J. B. Turner, S. Souther, C. C. Bennington, M. C. Vavrek, G. R. Shaver, and N. Fetcher. 2015. Northward displacement of optimal climate conditions for ecotypes of *Eriophorum vaginatum* L. across a latitudinal gradient in Alaska. Global Change Biol. 21:3827–3835.

Moore, J.-S., and A. P. Hendry. 2009. Can gene flow have negative demographic consequences? Mixed evidence from stream threespine stickleback. Philos. Trans. R. Soc. B Biol. Sci. 364:1533–1542.

Newman, D., and D. Pilson. 1997. Increased probability of extinction due to decreased genetic effective population size: experimental populations of *Clarkia pulchella*. Evolution 51:354–362.

Ochocki, B. M., and T. E. Miller. 2017. Rapid evolution of dispersal ability makes biological invasions faster and more variable. Nat. Commun. 8:14315.

Paul, J. R., S. N. Sheth, and A. L. Angert. 2011. Quantifying the impact of gene flow on phenotype-environment mismatch: a demonstration with the scarlet monkeyflower *Mimulus cardinalis*. Am. Nat. 178: S62–S79.

Pironon, S., G. Papuga, J. Villellas, A. L. Angert, M. B. García, and J. D. Thompson. 2017. Geographic variation in genetic and demographic performance: new insights from an old biogeographical paradigm. Biol. Rev. 92:1877–1909.

PRISM Climate Group. 2017. Oregon State University. URL http://prism.oregonstate.edu. Accessed: 2017-09-30.

Pujol, B., and J. R. Pannell. 2008. Reduced responses to selection after species range expansion. Science 321:96–96.

R Core Team. 2017. R: a language and environment for statistical computing, version 3.4.4. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Sagarin, R. D., and S. D. Gaines. 2002. The 'abundant centre' distribution: to what extent is it a biogeographical rule? Ecol. Lett. 5:137–147.

Samis, K. E., A. López-Villalobos, and C. G. Eckert. 2016. Strong genetic differentiation but not local adaptation toward the range limit of a coastal dune plant. Evolution 70:2520–2536.

Sexton, J. P., S. B. Hangartner, and A. A. Hoffmann. 2014. Genetic isolation by environment or distance: which pattern of gene flow is most common? Evolution 68:1–15.

Sexton, J. P., P. J. McIntyre, A. L. Angert, and K. J. Rice. 2009. Evolution and ecology of species range limits. Annu. Rev. Ecol. Evol. System. 40:415–436.

Sexton, J. P., S. Y. Strauss, and K. J. Rice. 2011. Gene flow increases fitness at the warm edge of a species range. Proc. Natl. Acad. Sci. USA 108:11704–11709.

Vucetich, J. A., and T. A. Waite. 2003. Spatial patterns of demography and genetic processes across the species' range: null hypotheses for landscape conservation genetics. Conserv. Genetics 4:639–645.

Wang, T., A. Hamann, D. Spittlehouse, and C. Carroll. 2016. Locally downscaled and spatially customizable climate data for historical and future periods for north america. PLoS One 11:e0156720.

Weir, B. S., and C. C. Cockerham. 1984. Estimating F-statistics for the analysis of population structure. Evolution 38:1358–1370.

Wilczek, A. M., M. D. Cooper, T. M. Korves, and J. Schmitt. 2014. Lagging adaptation to warming climate in *Arabidopsis thaliana*. Proc. Natl. Acad. Sci. USA 111:7906–7913.

Willi, Y., M. Van Kleunen, S. Dietrich, and M. Fischer. 2007. Genetic rescue persists beyond first-generation outbreeding in small populations of a rare plant. Proc. R. Soc. Lond. B Biol. Sci. 274:2357–2364.

Williams, J. L., B. E. Kendall, and J. M. Levine. 2016. Rapid evolution accelerates plant population spread in fragmented experimental landscapes. Science 353:482–485.

Williams, J. W., and S. T. Jackson. 2007. Novel climates, no-analog communities, and ecological surprises. Front. Ecol. Environ. 5:475–482.

Yeaman, S., and A. Jarvis. 2006. Regional heterogeneity and gene flow maintain variance in a quantitative trait within populations of lodgepole pine. Proc. R. Soc. Lond. B Biol. Sci. 273:1587–1593.

Associate Editor: S. Wright

Supporting Information

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