Pollination, genetic structure, and adaptation to climate across the geographic range of *Clarkia pulchella*

by

Megan Bontrager

B.Sc. Plant Sciences, University of California, Santa Cruz, 2011

B.Sc. Molecular, Cell, and Developmental Biology, University of California, Santa Cruz, 2011

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF

Doctor of Philosophy

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES
(Botany)

The University of British Columbia
(Vancouver)

August 2018

The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, the dissertation entitled:

Pollination, genetic structure, and local adaptation across the geographic range of *Clarkia pulchella*

submitted by Megan Bontrager in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Botany

Examining committee:

Dr. Amy Angert

Supervisor

Dr. Jeannette Whitton

Supervisory Committee Member

Dr. Sally Aitken

Supervisory Committee Member

Dr. Darren Irwin

University Examiner

Dr. John Richardson

University Examiner

Additional supervisory committee members:

Dr. Michael Whitlock

Supervisory Committee Member

Abstract

Every species experiences limits to its geographic distribution on the landscape. Sometimes the barriers that limit geographic ranges are obvious. For example, oceans and topographic features may prevent a species from colonizing the areas beyond them. However, species' distributions frequently end at places on the landscape where no obvious barrier or abrupt shift in the environment occurs, and this raises the question of what limits the range at these edges, both proximately and in evolutionary time. This thesis investigates the contributions of pollination, climate, and gene flow to limiting range edge populations of an annual wildflower, *Clarkia pulchella*.

Pollinators may be important at range edges because many of the proposed characteristics of edge populations (small, isolated, or low density) are also features that might make pollination less reliable and in some cases favour the evolution of self-pollination. I found that climate influences floral morphology and that the capacity of plants to set seed in the absence of pollinators was slightly higher in northern range edge populations. All populations benefit from the service of pollinators.

Another factor that may limit populations at geographic range edges is the influence of asymmetric gene flow from central populations, which could prevent local adaptation in range edge populations. Alternatively, edge populations might have low genetic variance and therefore might benefit from gene flow. I tested these competing predictions by simulating gene flow between populations from across the species' range in the greenhouse and planting the progeny into common gardens at the northern range edge. This experiment took place during an extremely warm year. As a result, gene flow from warmer provenances improved performance. I also found a small benefit of gene flow independent of climate.

Finally, I found no evidence that environmental differences contribute to genetic differentiation of populations, though geographic distance is a strong predictor of genetic differentiation. Contrary to expectations, genetic variation was higher near the northern range edge. Together, these chapters shed light on important drivers of reproductive success and local adaptation in this species and allow for insights into what processes are likely (or unlikely) to generate range limits.

Lay Summary

All species occupy a limited geographic area on the landscape. My work seeks to understand what prevents species from occurring beyond their observed distributions. I investigated how floral traits and reproduction with vs. without pollinators vary across the range of an annual plant species, *Clarkia pulchella*. I found that some floral traits are associated with climate and northern populations were somewhat less reliant on pollinators compared to other populations. I investigated whether populations that are in different climate environments are more genetically differentiated. I did not find support for an effect of climate differences on genetic structure. I also investigated how populations from different climate regimes performed in a common environment. Populations that were from historically warm places performed better than local populations, likely because it was a very warm year. These results indicate that adaptation to climate and the availability of pollinators may influence the geographic distribution of *Clarkia pulchella*.

Preface

Chapter 2 has been published as "Bontrager, Megan, and Amy L. Angert. Effects of range wide variation in climate and isolation on floral traits and reproductive output of *Clarkia pulchella*. American Journal of Botany 103.1 (2016): 10-21." I conceived of the project in consultation with Dr. Angert and she provided guidance throughout the project. I collected the data, generated the species distribution model, and performed data analyses. Dr. Angert and I conceived of the structure of the manuscript together. I wrote the initial draft of the paper and Dr. Angert contributed edits to this draft. I submitted the paper for publication and led the revision process with guidance from Dr. Angert.

A version of Chapter 3 is in preparation for publication with Chris D. Muir and Amy L. Angert and is available as a preprint on bioRxiv (doi: https://doi.org/10.1101/372375). I conceived of the project in consultation with Dr. Angert and she provided guidance throughout the project. I performed all of the field work and data collection. Dr. Muir and I collaborated on the statistical analyses with input from Dr. Angert. Dr. Muir diagnosed model fit issues. Dr. Angert and I conceived of the structure of the manuscript together. I generated all figures and tables and wrote the manuscript with feedback from Dr. Angert and Dr. Muir.

I conceived of Chapter 4 in consultation with Amy L. Angert, and she provided guidance throughout the project. I collected the plant material, performed all DNA extraction and library preparation, cleaned and analyzed the sequences, and performed all data analyses. I generated all figures and tables and wrote the manuscript with frequent conversation and comments from Dr. Angert. A version of this manuscript is in preparation for publication and is available as a preprint on bioRxiv (doi: https://doi.org/10.1101/374454).

I conceived of Chapter 5 in consultation with Amy L. Angert, and she provided guidance throughout the project. I collected the seeds and performed the hand pollinations in the greenhouse. I led the installation of transplant gardens and performed all monitoring and data collection. I performed all data analyses with helpful advice from Dr. Angert. I generated all figures and tables and wrote the manuscript with frequent conversation and comments from Dr. Angert.

Table of Contents

Abstract . . . . . **iii**

Lay Summary . . . . . **iv**

Preface . . . . . **v**

Table of Contents . . . . . **vi**

List of Tables . . . . . **x**

List of Figures . . . . . **xi**

Acknowledgments . . . . . **xiii**

1 Introduction **1**

1.1 Why range limits? 1

1.2 Equilibrial vs. disequilibrial range limits 1

1.3 Mechanisms generating equilibrial range limits: theory 2

1.3.1 Drift, limited genetic variance, and adaptive trade-offs 2

1.3.2 Swamping gene flow 3

1.3.3 Metapopulation models 3

1.4 Abiotic and biotic constraints on geographic ranges 4

1.5 Assumption of smooth environmental gradients and abundant centre distributions . 4

1.6 Range limits: empirical examples 5

1.6.1 *Mimulus cardinalis*: contrasting patterns across elevation vs. latitude 5

1.6.2 *Clarkia xantiana*: both biotic and abiotic gradients affect range edges 6

1.6.3 *Drosophila birchii*: both gene flow and strong selection constrain the range . 7

1.7 Investigating the effects of pollinators and gene flow across the range of *Clarkia pulchella* 8

2 Effects of range-wide variation in climate and isolation on floral traits and reproductive output of *Clarkia pulchella* **10**

2.1 Introduction 10

2.2 Methods . . . 12
2.2.1 Study system . . . 12
2.2.2 Specimen selection and measurements . . . 12
2.2.3 Estimating geographic isolation . . . 14
2.2.4 Locality-specific climate data . . . 15
2.2.5 Statistical analyses . . . 16
2.3 Results . . . 17
2.3.1 Climate and plant reproductive output . . . 17
2.3.2 Climate, isolation, and floral traits . . . 17
2.3.3 Variation in climate, isolation, and plant characteristics across the range . . . 17
2.4 Discussion . . . 18
2.4.1 Climate, range position, and reproductive fitness . . . 18
2.4.2 Climate, range position, and floral traits . . . 19
2.4.3 Isolation, range position, and self-pollination . . . 20
2.4.4 Metapopulation dynamics . . . 20
2.4.5 Use of herbarium specimens . . . 21
2.4.6 Conclusions and future directions . . . 21

3 Geographic and climatic drivers of reproductive assurance in *Clarkia pulchella* 32
3.1 Introduction . . . 32
3.2 Methods . . . 34
3.2.1 Study system . . . 34
3.2.2 Plot establishment and monitoring . . . 35
3.2.3 Climate variable selection . . . 36
3.2.4 Statistical analyses . . . 37
3.3 Results . . . 37
3.3.1 Variation in response to pollinator limitation across the range . . . 37
3.3.2 Response of patch density to seed production in the previous year . . . 38
3.3.3 Variation in fruit production across the range . . . 38
3.4 Discussion . . . 39
3.4.1 Reproductive assurance is driven by geography rather than climate . . . 39
3.4.2 Reallocation to flower and fruit production under pollen limitation . . . 40
3.4.3 Implications for responses to climate change . . . 41
3.4.4 Conclusions and future directions . . . 41

4 Genetic differentiation is determined by geographic distance in *Clarkia pulchella* 48
4.1 Introduction . . . 48
4.2 Methods . . . 50
4.2.1 Study species . . . 50
4.2.2 Population selection, climate characterization, and seed collection . . . 50

4.2.3 DNA Extraction . . . 51
4.2.4 Library preparation and sequencing . . . 51
4.2.5 Alignment and SNP calling . . . 52
4.2.6 Quantifying isolation by environment vs. isolation by distance . . . 53
4.2.7 Assessment of spatially continuous vs. discrete genetic differentiation . . . 54
4.2.8 Exploring spatial patterns in genetic diversity . . . 54
4.3 Results . . . 54
4.3.1 Isolation by environment vs. geographic distance . . . 54
4.3.2 Genetic structure of populations . . . 55
4.3.3 Geographic trends in genetic diversity . . . 55
4.4 Discussion . . . 55
4.4.1 Populations of *Clarkia pulchella* are isolated by distance . . . 56
4.4.2 Populations are admixtures of northern and southern genetic groups . . . 56
4.4.3 Genetic diversity increases with latitude . . . 57
4.4.4 Conclusions . . . 57

5 Gene flow disrupts local adaptation but improves performance at the northern range edge of *Clarkia pulchella* . . . 66
5.1 Introduction . . . 66
5.2 Methods . . . 69
5.2.1 Study system, seed collection, and site selection . . . 69
5.2.2 Greenhouse generation and crossing design . . . 69
5.2.3 Common garden design and installation . . . 70
5.2.4 Monitoring and measuring . . . 71
5.2.5 Climate data . . . 72
5.2.6 Population genetic data . . . 72
5.2.7 Statistical analyses . . . 73
5.3 Results . . . 75
5.3.1 Climate of origin explains performance in common gardens . . . 75
5.3.2 Gene flow may confer some benefits to edge populations . . . 76
5.3.3 Genetic differentiation between parental populations is positively correlated with fitness . . . 76
5.4 Discussion . . . 77
5.4.1 Climate of origin predicts performance . . . 78
5.4.2 Gene flow confers benefits independent of climate . . . 79
5.4.3 Limited inference about population persistence . . . 80
5.4.4 Conclusions . . . 81

6 Conclusions 92
6.1 Major findings 92
6.1.1 Chapter 2: Associations of climate and geography with herbarium specimen characteristics 92
6.1.2 Chapter 3: Exclusion of pollinators in natural populations of *Clarkia pulchella* 93
6.1.3 Chapter 4: Genetic structure across the geographic range of *Clarkia pulchella* 94
6.1.4 Chapter 5: Effects of gene flow on performance at the northern range edge . 95
6.2 What limits the range in *Clarkia pulchella*? Synthesis and future directions 96
6.3 Next steps in range limit research 98

Bibliography 101

A Supporting Materials 117

List of Tables

Table 2.1 Effects of climate on reproductive output 27
Table 2.2 Effects of climate on floral traits 28
Table 2.3 Effects of isolation on floral traits 29
Table 2.4 Effect of range position on spatial isolation 29
Table 2.5 Relationship between range position and climate 30
Table 2.6 Relationship between range position and reproductive output or herkogamy . . . 31

Table 3.1 Geographic data for experimental sites 46
Table 3.2 Effects of pollinator exclusion, region, and climate on seed set per fruit. 47
Table 3.3 Effects of pollinator exclusion, region, and climate on fruit number. 47

Table 4.1 Geographic information for populations included in population genetic analyses . 64
Table 4.2 Results of partial Mantel tests of pairwise differences 65
Table 4.3 Covariance contributions of each layer in conStruct models 65

Table 5.1 Geographic information for populations used in the transplant experiment 87
Table 5.2 Effect of local vs. foreign origin on performance of *Clarkia pulchella* 87
Table 5.3 Effects of absolute precipitation and temperature differences on component lifestages 88
Table 5.4 Effects of being a within-population cross vs. a between-population cross 89
Table 5.5 Effects of genetic differentiation and climate differences on performance 90
Table 5.6 Effects of genetic differentiation between parental populations on performance . . 91

Table A.1 Sensitivity analyses of tests involving isolation to distribution model decisions . . 118
Table A.2 Correlation among climate variables from pollinator exclusion sites 123

List of Figures

Figure 2.1 Conceptual map of analytical framework 22
Figure 2.2 Map of *Clarkia pulchella* localities across the species' range 23
Figure 2.3 Diagram showing the measurements made on flowers 24
Figure 2.4 Effects of climate, isolation, and range position on *Clarkia pulchella* 25
Figure 2.5 Range position, climate, and *Clarkia pulchella* characteristics 26

Figure 3.1 Experimental sites relative to the geographic range of *Clarkia pulchella* 42
Figure 3.2 Climate conditions in experimental sites in each region 43
Figure 3.3 Seed set in plots with and without pollinators by region 44
Figure 3.4 Effect of seed input in 2015 on the number of adult plants in 2016 44
Figure 3.5 Fruits per plant in plots with and without pollinators by region 45
Figure 3.6 Effects of climate and pollinator exclusion on per-plant fruit production 45
Figure 3.7 Reproductive assurance in each region as a proportion of control seed set 46

Figure 4.1 The geographic range of *Clarkia pulchella* and sampling locations 59
Figure 4.2 Relationships of climate and geography across the range of *Clarkia pulchella* . . 60
Figure 4.3 Pairwise genetic differentiation vs. geographic distance 61
Figure 4.4 Effect sizes of temperature and precipitation vs. geographic distance 62
Figure 4.5 Results of conStruct cross validation 62
Figure 4.6 Admixture proportions of each of 32 populations of *Clarkia pulchella* 63
Figure 4.7 Expected heterozygosity across the range of *Clarkia pulchella* 63

Figure 5.1 Geographic locations and climate averages of populations used in this experiment 82
Figure 5.2 Lifetime fitness from populations of *Clarkia pulchella* with foreign vs. local parents 83
Figure 5.3 Effects of temperature and precipitation differences on performance 84
Figure 5.4 Effects of gene flow on performance of *Clarkia pulchella* 85
Figure 5.5 Effects of genetic differentiation and climate differences on performance 86

Figure A.1 Distribution of per-locus F_{ST} 125
Figure A.2 Marginal posterior distributions of BEDASSLE analyses 126
Figure A.3 Pairwise differences and genetic differentiation among northern populations . . . 127
Figure A.4 Pairwise differences and genetic differentiation among central populations 128

Figure A.5 Crossing design for common garden experiment 129
Figure A.6 Distribution of climate of within- vs. between-population crosses 130

Acknowledgments

I would like to thank my supervisor, Dr. Amy Angert, for countless conversations that broadened my thinking on the subjects explored here and for her support and encouragement throughout my Ph.D. I feel incredibly fortunate to have worked with you and to have learned how to be a scientist from you. You helped me whenever I got stuck, and more importantly, you've taught me how to get myself unstuck. Thank you so much for your patience, generosity, and the seemingly limitless intellectual energy you brought to our work together. It would be nice to have all of my thesis chapters published already, but I am also very happy to have a reason to keep bothering you after I move on.

My committee—Dr. Jeannette Whitton, Dr. Mike Whitlock, and Dr. Sally Aitken—provided feedback that helped me think more clearly about my projects. All the components of this thesis were improved by their thoughtful input and I really appreciate the time that they put into reading and commenting on my work. I would also like to thank other faculty who have given me help and support over the years, including Dr. Jill Jankowski, Dr. Matthew Pennell, Dr. Jennifer Williams, and Dr. Rachel Germain. Colleagues and friends at the Biodiversity Research Centre have provided me with invaluable support, feedback, and inspiring conversations. Especially important among these people are Jasmine Ono, Anna Bazzicalupo, Alison Porter, Sean Naman, Ken Thompson, Nathaniel Sharp, Bill Harrower, Barb Gass, and most especially Matthew Osmond.

I would like to thank earlier mentors I had in my educational arc. I am very grateful to Dr. Ingrid Parker and Dr. Jenn Yost. Any advantages I had coming into graduate school are because of your mentorship and teaching and I have tried not to forget the good habits you taught me. Dr. Matt Ritter and Dr. Sara Grove were also great sources of support in my early days as a biologist. I would like to thank teachers I had over the years who encouraged my curiosity and helped me learn to communicate: Dr. David Sullivan, Marina Martin, Elaine Adams, and Maria Fahrner.

I have been really fortunate in the funding support I have received during my PhD. I would like to acknowledge the UBC Four-year Fellowship, the Vladimir J. Krajina research prize, and the Li Tze Fong affiliated fellowship. I would also like to thank Dr. Roy Turkington for his support of my research. The Botanical Society of America and the Washington Native Plant Society provided grants that supported my work as well.

I would like to thank my family—Laura, Keith, Anna and Juliana—for their endless support and kindness.

Acknowledgements by chapter

Chapter 2

I would like to thank C. Kopp, C. Muir, R. Sargent, S. Pironon, A. Hargreaves, the Angert lab group, and two anonymous reviewers for comments that improved this manuscript. M. Bayly also provided helpful comments and shared code. T. Edwards led a workshop that provided code and guidance for species distribution modelling. E. Fitz provided much appreciated assistance with georeferencing and specimen measurements. Thanks to L. Jennings for herbarium access at the University of British Columbia and to J. Smith for allowing us to use a sample image from the Snake River Plains herbarium at Boise State University.

Chapter 3

I would like to thank B. Harrower, R. Germain, and members of the Angert lab for their thoughtful comments on this project. E. Fitz assisted with fieldwork. Permission to work in field sites was granted by British Columbia Parks, Umatilla National Forest, Ochoco National Forest, and the Vale District Bureau of Land Management.

Chapter 4

I would like to thank C. Caseys and M. Todesco for their generous guidance and training during library preparation. G. Owens provided helpful advice on bioinformatics methods. E. Fitz assisted with locating populations of *C. pulchella* in the field and A. Wilkinson assisted with plant cultivation.

Chapter 5

I would like to thank M. Osmond, L. Bontrager, C. Leven, D. Gamble, A. Porter, A. Wilkinson, J. Chan, T. Mitchell, and P. Chen for their help in the field. A. Wilkinson and M. Zink Yi provided plant care and pest management in the greenhouse. B. Gass and D. Holtz also helped with plant care. E. Fitz and D. Gamble assisted with greenhouse pollinations and fruit collection. A. Porter, J. Ono, and J. Chan assisted with preparing the seeds for planting. British Columbia Provincial Parks permitted the garden installation and monitoring.

Chapter 1

Introduction

1.1 Why range limits?

Insightful observers of the natural world have frequently posed the question of what limits species' distributions on the landscape (Darwin, 1859; MacArthur, 1972). In some cases, the landscape features that limit species' distributions are obvious, such as when there is an abrupt shift in the environment at some point in space. However, species' distributions frequently end at seemingly arbitrary places on the landscape and this raises the fascinating question of what processes prevent the species from occurring beyond that limit (Antonovics, 1976). This question can be framed with regard to what restricts geographic ranges in the present, the extent to which these range limits are temporally stable, and what forces act to stabilize range limits in evolutionary time.

1.2 Equilibrial vs. disequilibrial range limits

When considering why a range limit exists, one of the first questions to consider is whether that range limit represents a niche limit or whether the species' range is limited by its ability to disperse. A species' niche is the intersectional space of many environmental axes within which individuals can survive and reproduce such that populations can persist (Hutchinson, 1957). In the case of a niche-limited range, the range edge occurs where one or more environmental variables have changed such that they no longer allow population persistence (Pulliam, 2000). Alternatively, a species' range may be limited not because the areas beyond it are unsuitable for persistence but because individuals of that species have not dispersed into them. Dispersal-limited ranges may result when species fail to track temporally changing environments across space, as is the case in some temperate species which have not expanded their ranges into areas that have become climatically suitable since the last glacial maximum (Svenning et al., 2008). While dispersal limitation can explain some range limits, the majority of experiments that transplant individuals to sites beyond the range edge detect declines in performance, as is predicted if range limits represent niche limits (Lee-Yaw et al., 2016). Interest in the mechanisms that prevent adaptation to conditions beyond the range edge have driven a rich body of theoretical and empirical work in the past twenty-five

years (reviewed in Hoffmann and Blows, 1994; Holt, 2003; Bridle and Vines, 2007; Sexton et al., 2009).

1.3 Mechanisms generating equilibrial range limits: theory

Theoretical and conceptual explorations of equilibrial range-limiting processes (i.e., processes other than dispersal limitation) can be roughly divided into three groups. The first are those that invoke limited genetic variance, higher genetic drift, or other mechanisms limiting adaptation in range edge populations without requiring an effect of central populations (Hoffmann and Blows, 1994). Next, there are models that posit that maladaptation at range edges may be maintained because of influence from central populations that are adapted to different conditions (Kirkpatrick and Barton, 1997; Polechová and Barton, 2015). Finally, some models show that range edges can result from gradients in metapopulation dynamics. Models in this last group usually do not explicitly address limits to adaptation, but are not mutually exclusive with adaptation-limited ranges (Holt and Keitt, 2000; Lennon et al., 1997). They lend an important perspective on range limits because they consider landscape-level processes such as colonization and dispersal, rather than solely focusing on dynamics within populations.

1.3.1 Drift, limited genetic variance, and adaptive trade-offs

Edge populations are often characterized as smaller or lower density than those in the centre of a species' range (but see Sagarin and Gaines (2002) and discussion in Section 1.5), based on the assumption that populations at the range centre occupy an optimal position on underlying environmental gradients, while populations at range periphery experience less favourable conditions (Brown, 1984). The idea that edge populations have smaller population sizes and occupy environments at the limits of the species' tolerance has led to a suite of hypotheses about what limits the ability of edge populations to adapt to their environments (reviewed in Antonovics, 1976; Hoffmann and Blows, 1994; Bridle and Vines, 2007). Among these hypotheses is the idea that if edge populations are smaller, they may experience stronger genetic drift and this could lead to the fixation of deleterious alleles and maladaptation. There is also less opportunity for beneficial mutations to arise in small populations. Lower genetic variation may also arise if the environments that edge populations occupy are at the limits of the species' physiological tolerance, because in this situation range edge populations may experience strong and persistent directional selection that reduces genetic variance in fitness-related traits. Finally, some environmental conditions may pose adaptive challenges if the optimal phenotype requires change in traits that have antagonistic genetic correlations. Similarly, if multiple phenotypic changes are required for adaptation, they may be unlikely to arise in the same individual. Any of these phenomena might limit or slow the adaptive potential of range edge populations when they occur.

1.3.2 Swamping gene flow

Swamping gene flow is an often-invoked process that could inhibit adaptation in range edge populations and suppress their population growth rates such that they are prevented from exerting propagule pressure on areas beyond the range. In their classic model of this scenario, Kirkpatrick and Barton (1997) consider populations connected by gene flow arranged along a smooth environmental gradient. The optimum phenotype changes along this environmental gradient. Given sufficient genetic variation and limited gene flow, the expectation is that in this spatially heterogeneous environment populations will evolve such that their phenotypes are optimal under the conditions they typically experience (Felsenstein, 1977), resulting in a pattern known as local adaptation. However, when this underlying environmental gradient is steep, gene flow from centre of the gradient (which is densely populated with well-adapted individuals) can inhibit local adaptation at the range edge (García-Ramos and Kirkpatrick, 1997). If this swamping gene flow is strong enough, it can turn populations at the edge of the range into demographic sinks, even if carrying capacity is high across the entire gradient (i.e., the intrinsic rate of increase at any point along the gradient would be positive if the population at that point were locally adapted; Kirkpatrick and Barton, 1997). The deleterious effects of gene flow can, under some circumstances, be counteracted by the benefits of an influx of genetic variance (Barton, 2001; Alleaume-Benharira et al., 2006).

1.3.3 Metapopulation models

A final set of mechanisms that may generate stable range limits arise from gradients in metapopulation dynamics across the species' range (Holt and Keitt, 2000; Lennon et al., 1997). Range limits may arise if there is a lower abundance of suitable patches near range edges, if extinction probabilities are higher in range edge habitats, or if the probability of colonization of empty but suitable patches is lower near range edges. While this result may seem intuitive, Holt and Keitt (2000) showed that range limits could arise at positions along a gradient where suitable habitat had not yet disappeared altogether, before extinction was certain, and before colonization was impossible. Their model is not specific about the causes of these gradients in metapopulation dynamics; one possibility is that they arise as a result of underlying environmental gradients. However, in contrast to other range limit hypotheses, in some configurations these gradients need not actually affect the selective environment that populations experience. Rather, range limits may occur due to gradients in the availability of habitable environments or in the degree to which the matrix surrounding habitable patches facilitates dispersal. Similarly, in a patchy landscape, Allee effects may limit the spread of a species; under some circumstances this can occur in the absence of any underlying environmental gradient (Keitt et al., 2001). While not explicitly incorporated into these models, the increased rates of population turnover and greater isolation of populations predicted at range edges are likely to impact population genetic characteristics and demography within patches in ways that might negatively affect persistence of populations at range edges. It is also possible that fitness declines due to the mechanisms considered in other models could act as drivers of increased extinction probability or decreased colonization probability in a metapopulation framework.

1.4 Abiotic and biotic constraints on geographic ranges

While the environmental gradients that generate range limits at equilibrium in the models described above are often conceptualized as abiotic (such as gradients in temperature or salinity), variation in the frequency or strength of interspecific interactions may also underlie geographic ranges. For example, competitive interactions along a gradient in resources may reinforce the range limits of each of the interacting species (Price and Kirkpatrick, 2009) and can relax the steepness of the environmental gradient that might otherwise be required for gene flow to limit adaptation (Case and Taper, 2000). Hybridization with closely related parapatric species also has the potential to shape range boundaries via character displacement (Goldberg and Lande, 2006). Most of the theoretical investigations of how species interactions might create or reinforce geographic range limits focus upon competitive or predatory interactions, rather than mutualisms. Mutualists may be critical to population persistence (Lennartsson, 2002), can play a role in determining the genetic structure of populations (Kramer et al., 2011), and have even been found to affect geographic distributions in some systems (Afkhami et al., 2014). However, mutualisms are not well incorporated into range limit theory. Like other proposed forces shaping ranges, spatial variation in mutualistic relationships may have more nuanced effects on geographic distributions than simply being necessary for persistence of one or both of the partners.

1.5 Assumption of smooth environmental gradients and abundant centre distributions

The classic theoretical predictions for evolutionarily stable geographic range limits generally assume that a species' range overlays a smooth environmental gradient (Kirkpatrick and Barton, 1997; Holt and Keitt, 2000). Perhaps for this reason, some of the systems in which we understand geographic range limits the best are those with relatively straightforward underlying environmental gradients, especially in climate (such as the precipitation and temperature gradients experienced by rainforest *Drosophila* spp.; Kellermann et al., 2009). The assumption of smooth underlying gradients results in predictions of geographic distributions in which the centre of the range has a higher density of individuals than the range edge, a characterization of the range consistent with other inferences in the literature (Brown, 1984). However, neither the pattern of smooth underlying environmental gradients nor the abundant centre distribution consistently describe populations on the landscape (Sagarin and Gaines, 2002). Global gradients in temperature and precipitation are often quite heterogeneous at the scale of species' ranges due to topographic features or continentality. Less than half of studies that quantify abundance across geographic ranges result in patterns consistent with the abundant centre distribution (Sagarin and Gaines, 2002). In light of this, a critical next step for both theoretical and empirical investigations of range limits is to try to understand how robust predictions from classic papers are to the partial decoupling of geography, abundance, and environment, and to develop new predictions that specifically consider more realistic landscapes and abundance patterns (Pironon et al., 2017).

1.6 Range limits: empirical examples

It has now been about two decades since many of the classic theoretical explorations of range limits were first published. Despite a large number of studies that have endeavoured to test these theories in natural systems, there are still relatively few systems where we have a clear understanding of what shapes and limits species' geographic distributions. Here I summarize findings from three key study systems, each of which have been the subject of over a decade of research on range limits. I discuss how the results of empirical work to-date sometimes support and sometimes contradict predictions of range limit theory.

1.6.1 *Mimulus cardinalis*: contrasting patterns across elevation vs. latitude

Mimulus cardinalis is a riparian perennial plant growing in coastal and montane regions along the west coast of North America. Angert and Schemske (2005) transplanted individuals to sites beyond the species' high elevation margin and found that this range limit represents a niche limit: individuals moved beyond the high elevation edge generally failed to reproduce and had very low fitness. Small differences in fitness between populations in common gardens arranged across an elevation transect indicated limited local adaptation to elevation. This suggests that these populations are limited in their ability to adapt to the local climate conditions imposed by elevation. Optimal phenotypes differ at different elevations (Angert et al., 2008), but adaptive differentiation of populations could be inhibited if gene flow or dispersal frequently expose phenotypes that are favourable in one environment to selection in other environments. Investigation of the climatic drivers of a latitudinal cline in physiological traits indicates that gene flow among populations of *M. cardinalis* may result in adaptation to the average climatic conditions in a region, rather than to the precise conditions in a site (Muir and Angert, 2017). Taken together, these results suggest that the high elevation range edge of *M. cardinalis* may be a result of gene flow and trade-offs across elevation gradients limiting adaptation to conditions beyond the current elevational range.

While the high elevation range edge in this species may represent limits to adaptation, populations near this edge do not appear to be especially maladapted; population growth rates do not decline towards high elevation range limits (Angert, 2006, 2009). This contrasts with predictions of Kirkpatrick and Barton (1997), in which populations near range margins are expected to be more phenotypically mismatched to their environments. It is possible that this discrepancy could be the result of low adaptive differentiation across elevations without the strong asymmetry in gene flow that is central to the model of Kirkpatrick and Barton (1997).

In contrast to populations along an elevation gradient, populations of *M. cardinalis* are phenotypically differentiated across latitude (Muir and Angert, 2017). Angert and colleagues have employed several methods to investigate processes that may limit the latitudinal range of this species. Paul et al. (2011) found that there is asymmetry in gene flow among populations of *M. cardinalis*, with more immigration from central populations into those at high and low latitudes. They also found that gene flow from other latitudes was correlated with mismatch between the average phenotype of a population and the optimum phenotype in its home site. However, the

populations with high amounts of phenotypic mismatch were not necessarily those at high and low latitudes. This perhaps corroborates their findings along elevation gradients: gene flow may limit adaptive differentiation, but on complex landscapes this disruption may not be concentrated at range edges. Under these scenarios gene flow may be a force that maintains niche stability (Morjan and Rieseberg, 2004), but geographic patterns of local adaptation may diverge from model predictions.

Finally, the northern geographic range limit of *M. cardinalis* is likely dispersal limited. Bayly (2015) transplanted *M. cardinalis* to sites just beyond the species' northern range edge. Population growth rates estimated in these sites were as high as those estimated in experimental populations inside the range, indicating that these sites are suitable for *M. cardinalis*, should the species disperse to them. This is consistent with findings that the occupancy of suitable sites declines towards the northern range edge (Angert et al., 2018). At the northern range limit dispersal limitation may be the result of postglacial disequilibrium, but could also reflect a failure to track habitat that has more recently become suitable as a result of climate change. Across both latitudinal and elevational transects, high population growth rates towards cool edges and low rates at warm edges indicate that the warming climate is affecting population dynamics (Angert, 2006; Sheth and Angert, 2018).

1.6.2 *Clarkia xantiana*: both biotic and abiotic gradients affect range edges

Clarkia xantiana ssp. *xantiana* and ssp. *parviflora* are winter annual plants that are endemic to the foothills of the Sierra Nevada in California. These two subspecies have a small zone of overlap and present a particularly interesting system for range limit research because of the potential for comparison of sister taxa with contrasting reproductive strategies. The majority of the research on range limits in this system has focused on the mechanisms limiting *C. xantiana* ssp. *xantiana* at the subspecies' eastern range limit (where it meets the range of ssp. *parviflora*) though some studies have also investigated the coincident western range limit of ssp. *parviflora*. A variety of environmental gradients affecting population growth rate underlie the range: some run in parallel and some in contrasting directions (Eckhart et al., 2011).

Population growth rates of *C. xantiana* ssp. *xantiana* decline from the centre to the eastern edge of its geographic range (Eckhart et al., 2011), and fitness is low in experimental populations beyond the range edge (Geber and Eckhart, 2005). This is consistent with the hypothesis that underlying environmental gradients generate a range limit where there is a niche limit. However, population genetic analyses, quantification of adaptive differentiation between populations, and estimates of the heritability of fitness-related traits do not lend support to the leading theoretical predictions of what limits adaptation at range edges (Moeller et al., 2011; Gould et al., 2014). Heritabilities of fitness-related traits are high in range edge populations and edge populations do not appear to be lacking genetic variance in these traits (Gould et al., 2014). While there may be some asymmetric gene flow from centre to edge (Moeller et al., 2011), trait-environment correlations are just as strong among edge populations as they are among central populations; this indicates that countergradient gene flow is not disrupting local adaptation in a suite of fitness-related traits (Gould et al.,

2014). Range edge populations do not have genetic signatures of recent founder events or frequent population turnover, indicating that gradients in metapopulation dynamics are unlikely to play a large role at this range edge (Moeller et al., 2011).

Simultaneous work in this system has focused on the role of biotic interactions in limiting fitness within the range and possibly preventing range expansion. Investigations of pollinator activity in populations across the species' range have highlighted the potential for abiotic gradients to affect critical mutualistic relationships, as pollinator abundances in populations of the outcrossing *C. xantiana* ssp. *xantiana* decline along a gradient in precipitation (Moeller et al., 2012). Pollinators may limit reproduction at the range edge, but traits facilitating self-pollination have not evolved in response. Recent experiments have shown that herbivory has strong effects on fitness beyond the eastern range margin (Benning et al., 2018). Herbivory not only impacts experimental plants beyond the range edge, but also suppresses population growth rates within edge populations, which may limit propagule pressure beyond the edge. This makes herbivory a compelling proximate driver of the eastern range limit of *C. xantiana* ssp. *xantiana*. Herbivory has especially strong negative effects on late-flowering plants, and it might exert selection for earlier flowering time at this range limit. Populations within the range are quite differentiated in flowering time, and among these populations flowering time can be predicted by abiotic variables (Gould et al., 2014). These results suggest the possibility that abiotic and biotic features of the landscape may exert conflicting selection pressures on traits such as flowering time; if this were the case, it would make adaptation to conditions beyond the range edge difficult.

1.6.3 *Drosophila birchii*: both gene flow and strong selection constrain the range

Drosophila birchii, a fruit fly that is endemic to the rainforests on the eastern coast of Australia, has been the focus of research on processes that limit adaptive potential in range edge populations. This system is an ideal one for these types of questions because populations can be sampled along both elevational and latitudinal transects, each of which have strong underlying environmental gradients. Like other *Drosophila* spp., *D. birchii* can be easily reared in the lab, making measurements of the genetic variation and heritability of quantitative traits feasible (Hoffmann et al., 2003).

As in *M. cardinalis*, the causes of elevational and latitudinal range limits appear to differ for *D. birchii*. In parts of the range where there are steep climatic gradients (caused by steep elevational gradients), it appears that gene flow among populations in different environments prevents adaptive divergence along the gradient. This is in contrast to populations occurring along shallower elevation gradients, which show clinal divergence in climate-associated traits, such as cold tolerance (Bridle et al., 2009). Along these shallow gradients, populations with greater genetic variation in cold tolerance were also more closely matched to the predicted trait optimum at their site, suggesting that local adaptation may be facilitated when moderate levels of gene flow increase genetic variation. This work sheds light on what generates range limits along very steep gradients in this species, but further work will be necessary to understand what prevents further clinal differentiation and range

expansion along shallow elevation gradients. It is possible that trade-offs among traits (Jenkins and Hoffmann, 1999) or reductions in genetic variation due to strong directional selection (van Heerwaarden et al., 2009) could affect these range limits, as has been documented in another member of the genus or at other range limits of *D. birchii*.

At the dry end of the latitudinal range of *D. birchii*, heritability of desiccation tolerance is low, and this may prevent adaptation to increasingly dry conditions beyond this edge (Hoffmann et al., 2003). This limit to adaptation can likely be attributed to strong directional selection depleting genetic variation, as estimates of gene flow, neutral genetic diversity, and divergence among populations do not support other causes of low genetic variation (van Heerwaarden et al., 2009). The absence of asymmetric gene flow across latitudes, combined with trait differentiation among populations, indicates that maladaptation due to swamping gene flow is unlikely at this range edge. Similarly, genetic markers indicated that population sizes were not smaller towards the species' latitudinal range limit, so adaptive potential is unlikely constrained by demographic factors alone (van Heerwaarden et al., 2009).

Work that has focused on this single species of *Drosophila* is complemented by other studies that contrast traits, geographic distributions, and genetic variation across the phylogeny. Among species, upper thermal limits (Kellermann et al., 2012) and desiccation tolerance (Kellermann et al., 2012) show strong phylogenetic constraints, and widespread species have both higher rates of stress tolerance and greater genetic variation of stress resistance traits (Kellermann et al., 2009). As the processes that limit genetic variation and adaptive responses are characterized in other *Drosophila* spp., it may allow for generalizations about where and when different processes (including swamping gene flow, strong directional selection, or small population size) are most likely to limit adaptation at range margins.

1.7 Investigating the effects of pollinators and gene flow across the range of *Clarkia pulchella*

As evidenced by the examples above, understanding the range dynamics of a species may be a goal more appropriate for a career than a dissertation. However, in the work presented here I try to address some key gaps in our current understanding of range limiting processes. In particular, I investigate the effects of gene flow across a climatically heterogeneous landscape, the genetic structure of populations on the landscape, and the effects of climate and geography on reproduction and reproductive traits. My work focuses on the winter annual plant *Clarkia pulchella*, which occupies a climatically complex landscape in the interior Pacific Northwest. The climatic conditions underlying the range of this species are spatially heterogeneous, so in each chapter I consider the effects of climate independent of geography, but interpret these effects in a geographic context.

In Chapters 2 and 3, I explore variation in floral traits, reproduction, and the effects of pollinator exclusion across the geographic range of *C. pulchella*. Plant mating systems and geographic range limits are conceptually linked by shared underlying drivers, such as heterogeneity in climate and in the abundance of the plant species, but potential feedbacks between mating system variation

and range limiting processes are under-explored. I use herbarium specimens to examine spatial variation in floral morphology and reproductive output (Chapter 2) and perform field manipulations to measure the extent to which plants rely upon pollinators for reproduction (Chapter 3) with an interest in understanding whether the abiotic environment and pollinator availability interact to limit reproduction of *C. pulchella* in parts of the species' geographic range. In both of these chapters, I not only describe spatial patterns but also investigate abiotic variables that may be generating these patterns. In Chapter 4 I test whether climatic differences between populations play a role in determining population genetic structure. I also examine whether genetic variance declines towards range edges in this species, as might be expected if the range edge is limited by adaptation. Finally, in Chapter 5 I examine the effects of gene flow on edge populations of *C. pulchella* in a common garden experiment at the northern range edge. I assess whether gene flow has positive effects on edge populations, as might be expected if edge populations have experienced strong drift or lack adaptive genetic variance. As an alternative, I test for the potential of gene flow to swamp local adaptation at these edges. Taken together, these four chapters shed light upon important drivers of local adaptation in this species and emphasize the importance of explicitly considering how environmental variables vary across space when testing range limit hypotheses.

Chapter 2

Effects of range-wide variation in climate and isolation on floral traits and reproductive output of *Clarkia pulchella*

2.1 Introduction

The ecological and evolutionary factors shaping and maintaining mating system variation are of fundamental interest to plant biologists because of their potential impact on genetic and demographic processes. Many of the factors that affect mating system variation within species are also implicated in setting the boundaries of species' distributions (Hargreaves and Eckert, 2014). Studies of geographic distributions focus on variation in environment across space and the associated variation in species' abundance patterns. If environmental gradients underlying a species' range cause gradients in mate limitation (which could result from either lower densities of conspecific plants or lower service by pollinators), then mating system also may vary according to the position of a population in the species' range. Self-pollination may evolve in mate-limited populations (Moeller and Geber, 2005a; Fishman and Willis, 2008), overcoming the demographic consequences of mate limitation, but introducing potential genetic consequences. Self-pollination may reduce fitness via effects of increased homozygosity, and may also lead to smaller effective population sizes (Schoen and Brown, 1991) and lower genetic diversity within populations (Takebayashi and Morrell, 2001; Glémin et al., 2006), limiting a population's ability to adapt to novel conditions. Small population sizes and low genetic diversity in marginal populations have the potential to maintain evolutionarily stable geographic distributions by limiting response to selection (Hoffmann and Blows, 1994). Therefore, a gradient that causes mate limitation may also act to limit the species' range via genetic processes.

Species' ranges are often conceptualized as the geographic area within which environmental

variables are suitable for population growth or maintenance (Gaston, 2003; Sexton et al., 2009). Often, the centre of a species' range is expected to have the most optimal conditions for a species, because the underlying environmental gradients approach extremes or exhibit greater temporal variance at the range edges (Brown et al., 1996). This spatial pattern in environmental conditions may result in higher fitness of individuals near the centre of the range relative to those near the edge. Additionally, optimality could be reflected by a greater density of suitable patches or by greater carrying capacity for populations. These potential patterns lead to the expectation of an "abundant centre" distribution, in which the centre of the range has the highest density, while the margins of a species' distribution are predicted to have a sparser distribution of populations (Brown et al., 1996) and smaller population sizes (Holt and Keitt, 2000).

Environmental gradients may represent gradients in mate limitation because of their potential effects on the density of reproductive individuals, either in time or in space. In addition to reducing the availability of mates, low local abundance can limit mating opportunities if it leads to reduced pollinator services because of the low density of floral rewards. In some species, Allee effects have been documented, where low local abundance or density may prevent populations from maintaining or attracting pollinators (Groom, 1998; Knight, 2003; Moeller and Geber, 2005a). Such Allee effects have been shown to reinforce range limits in theoretical models (Keitt et al., 2001). Environmental extremes may also impose selection on mating system via their direct effects on individuals; for example, they may reduce resources available for allocation to pollinator attraction (Jorgensen and Arathi, 2013). All of these mechanisms of mate limitation may give a fitness advantage to individuals that can produce offspring autonomously and may select for traits that promote self-pollination at range limits.

Despite the appealing simplicity and theoretical support for the abundant centre hypothesis, the frequency and scale at which this pattern occurs in nature is unclear, and it is equivocally supported by empirical studies (Sagarin and Gaines, 2002). An exciting forefront for studies of both range limits and geographic variation in mating systems is to consider how deviations from simple abundant centre patterns affect predictions for phenotypic evolution. One mechanism that might lead to distributions that do not fit the abundant centre pattern is the decoupling of space (range position) from environmental variables that influence fitness and abundance. Environmental gradients underlying species' ranges may not be smooth due to topography, vegetation structure, and numerous other landscape features. Thus, when investigating range-wide patterns, one should not assume that spatial position and environmental suitability are correlated (Sagarin et al., 2006; Dixon et al., 2013). Another important consideration is that not all edges are structured by the same limiting variables. For example, the environmental variables that influence abundance and fitness might differ for northern vs. southern edges. Recent studies (Lira-Noriega and Manthey, 2014; Wang and Bradburd, 2014) have advocated for a focus on geographic variation in variables for which range position has often been used as a proxy. Rather than relying on range position alone as a predictor, studies should examine spatial changes in the mean and temporal variance of critical environmental variables that influence fitness and the spatial distribution of abundance.

In this study, we investigated the relationships between climate, climatic variability, spatial isolation, and potential for self-pollination across the geographic range of a mixed-mating annual, *Clarkia pulchella*. First, we examined which climatic variables (including deviations from average climatic conditions) drive variation in reproductive output (Figure 2.1A). Second, we identified which climatic variables best predict two floral traits, petal size and herkogamy, which we use as indicators of propensity for self-pollination (Figure 2.1B). We focused on precipitation and temperature variables that are likely to influence reproduction via direct effects on plant growth and indirect effects due to length of growing season. Though our study focused on reproductive characters, we also examined climate during germination and vegetative growth periods because this is likely to affect plant size and thus reproduction. Third, we examined whether spatially isolated populations tend to have floral trait values consistent with a greater propensity for self-pollination (Figure 2.1C). Next, we examined how drivers of reproductive output and floral traits vary spatially across the range (Figure 2.1DE). Finally, we examined whether reproductive output or floral traits are correlated with distance from the centre of the range, ignoring intermediate climatic predictors (Figure 2.1FG).

2.2 Methods

2.2.1 Study system

Clarkia pulchella Pursh (Onagraceae) is a winter annual that grows east of the Cascade Mountains in southern British Columbia, Canada, and in Washington, Oregon, Idaho and Montana in the United States (Figure 2.2). This species grows on dry, rocky slopes in forest gaps. It is self-compatible; however, as in other members of the genus, temporal and spatial separation of male and female functions promote outcrossing (Lewis, 1953). Flowers are pink and four-petaled and are pollinated by a diverse group of pollinators, including solitary bees (Palladini and Maron, 2013). The seeds of *C. pulchella* exhibit very little dormancy and have no specific dispersal mechanism (Newman and Pilson, 1997). Germination occurs in fall, most flowering occurs in June and July, and by August most plants have dried out and fruits are mature and dehiscing. Lewis (1953) noted that populations of species in the genus *Clarkia* seem to be more temporally stable than other annual wildflowers.

2.2.2 Specimen selection and measurements

Herbarium specimens were selected for measurements based on the availability of high-resolution images and the precision of associated locality data. Additionally, at least one flower on the specimen sheet had to meet the criteria described below. Images of 308 herbarium specimens were downloaded from the Consortium of Pacific Northwest Herbaria website in September of 2014 (www.pnwherbaria.org). An additional 15 specimens were photographed at the University of British Columbia herbarium. When multiple specimens were associated with the same location

in the same year, we used just one, chosen haphazardly. Records with coordinates provided were checked in Google Earth and assigned an error distance based on the specificity of the coordinates relative to the collector's description. Error distances were assigned by estimating the radius of the area that a specimen could have been collected in, given the specificity of the description. When the coordinates provided with a record did not match the locality description, they were edited manually, given an informative enough description. For example, a town name was not considered adequate to assign a precise locality, but a distance and direction from a distinct landmark was typically adequate. Records without any coordinates provided that had adequate descriptions were also georeferenced and assigned error distances. We then excluded all records with an error distance greater than 1 km. In all, we obtained 120 specimens with adequate locality data and specimen quality: 105 from the consortium and 15 from the UBC herbarium. These specimens cover the range of *C. pulchella* (Figure 2.2) and were collected between 1897 and 2013.

On each herbarium sheet, one flower was haphazardly selected from among those in good condition (petals, stigma, and at least one anther intact, visible, and well pressed). Additionally, the stigma had to be open and the anthers dehiscent. The amount of spatial separation between the stigma and anthers, or herkogamy, is positively correlated with outcrossing rates in many taxa (Karron et al., 1997; Takebayashi and Delph, 2000; Herlihy and Eckert, 2007; Luo and Widmer, 2013), including other *Clarkia* species (Lewis, 1953; Holtsford and Ellstrand, 1992; Moeller, 2006), and a pilot pollinator-exclusion study performed in the greenhouse found a significant relationship between low herkogamy and autonomous seed set in *C. pulchella* (M. Bontrager, unpublished data). In many species, herkogamy is a continuously varying, heritable trait, and low herkogamy contributes to a plant's ability to self-pollinate autonomously as well as to the probability that a pollinator will facilitate transfer of self-pollen (Carr and Fenster, 1994). The anthers of *C. pulchella* curl as they dehisce, so we measured the stamen in two ways: we measured the path length and also the height of the stamen from the base of the filament to the farthest point of the anther (this was not typically the anther tip, but instead the most distant point on the curled anther; Figure 2.3). Style length was measured from the base of the style to the centre of the stigma lobes. We calculated herkogamy as the difference between the path length of the stamen and length of the style. We chose to use the path length of the stamen rather than the height of the stamen because stamen height changes with floral age as the anther curls. Path length can be compared among flowers of different ages; therefore, it is a more useful representation of herkogamy for this study. Realized herkogamy is likely slightly greater than estimated here, because the anther dehisces once it has begun to curl back toward the filament. We used the ratio of the two stamen measurements as an indicator of flower age. We used our floral age metric to ensure that our metric of herkogamy did not change with age of the flower across the specimens we measured ($R^2 = 0.014$, $P = 0.196$). Self-pollination is often associated with a reduction in overall flower size (Goodwillie et al., 2010; Button et al., 2012; Dart et al., 2011). Therefore, we also measured petal characteristics: the length of one petal as well as its width, which we measured as the distance between the tips of the two lateral lobes of the petal. Petal length and petal width are correlated ($r = 0.81$), so we

use petal length only as a proxy for flower size in our analyses. To ensure that pressing did not dramatically alter floral measurements of *C. pulchella*, we measured herkogamy and petal length on fresh greenhouse-grown flowers and then measured them again after several weeks in a plant press. For both traits, correlations between fresh and pressed measurements were high (r = 0.88).

We counted all buds, flowers, and fruits on each plant, and summed them to obtain a metric of reproductive output. Although the fruits vary in the number of seeds set and herbarium specimens may not represent the exact reproductive output of these plants (i.e., not all buds may develop into fruits, or plants may have been collected before developing their full count of reproductive structures), this metric is a coarse proxy for reproduction, which is likely an important fitness component in these annual plants. Because these plants are small and often multiple specimens of various sizes were pressed on a single sheet, there should be little bias introduced from collectors preferring plants that fit in their presses. On 115 specimens, the roots were collected with the plant, so total counts were obtained. On the remaining five specimens, we could not confirm that the entire plant was collected so we measured floral traits only. All measurements were made to the nearest 0.1 mm using the segmented line tool in ImageJ (Rasband, 2012).

2.2.3 Estimating geographic isolation

To estimate each specimen's potential geographic isolation from other populations on the landscape, we used a distribution modelling approach to project habitat suitability across the landscape and then estimated average suitability within 1, 5, and 10 km radii of each specimen occurrence. Although it would be ideal to determine the relevant radius for isolation based on known distances for pollinator movement and seed dispersal, in the absence of such information for our study species we used a range of areas. This proxy for spatial isolation assumes that specimens surrounded by habitat of higher average suitability are less likely to be isolated from other populations than specimens surrounded by habitat of lower average suitability. This assumption is likely to be most valid if occupancy of suitable areas across the range is even and if temporal changes in suitability are low. We used MaxEnt (Phillips et al., 2004, 2006) to build a model of habitat suitability across the species range. MaxEnt modelling and associated spatial analyses were performed in R (R Core Team, 2013) using the packages 'dismo' (Hijmans et al., 2016), 'raster' (Hijmans and van Etten, 2014), 'rgdal' (Bivand et al., 2014), 'rgeos' (Bivand and Rundel, 2013), and 'sp' (Pebesma and Bivand, 2005). All occurrence records available with and without coordinates were downloaded from the Consortium of Pacific Northwest Herbaria in September of 2014. This resulted in 815 records (including those of specimens on which we measured plant characteristics). Additional localities were added from specimens at the University of British Columbia herbarium that had not been added to the consortium database (eight records). The geographic coordinates of each occurrence record were checked as described above and manually georeferenced as needed. Additional occurrences were added from field surveys (50 records). After removing duplicate records (those that fell in the same 0.0083 by 0.0083 degree grid cell) and records with inadequate locality information, we had 310 records with locality error distances of 1 km or less. An additional 31 localities were

spatial duplicates, but were collected in unique years; these were used in later analyses but were not used to build the distribution model. Our final set of localities covers the continuous range of *C. pulchella* (Figure 2.2). Although this species is mentioned to have occurred in northern California and South Dakota (Lewis, 1955), no records could be found based on queries of herbarium databases and online floras of these states.

We defined the background extent for the distribution model as the polygon created by the union of 100-km-radius buffers around each locality point. From this extent, we randomly sampled 3100 background points. We selected climatic predictor variables from the full Bioclim variable set (Hijmans et al., 2005) based on correlation among predictors across 2000 background points (avoiding including multiple predictors with $r > 0.9$) and the performance of each variable in distinguishing between presence and background in univariate GLM models. Ultimately, we used annual mean temperature (bio1), temperature seasonality (bio4), maximum temperature of the warmest month (bio5), minimum temperature of the coldest month (bio6), temperature annual range (bio7), mean temperature of the wettest quarter (bio8), precipitation of the wettest month (bio13), and precipitation seasonality (bio15). Additionally, we included a forest canopy cover layer (Geospatial Information Authority of Japan, Chiba University, and collaborating organizations, 2008) and a total green vegetation layer (Broxton et al., 2014) in our model because the occurrence of *C. pulchella* was associated with canopy gaps in field surveys (M. Bontrager, unpublished data). Our choices of a fairly high correlation threshold, the inclusion of a relatively large number of variables, and a high ratio of background points to presence points reflect our intention to use the model as a predictor of current occurrence, rather than for interpretation of the relative importance of the variables and their ecological effects or for extrapolation (Merow et al., 2014). We ran the model with MaxEnt default features. Model performance was evaluated by calculating the area under the receiver operating characteristic curve (AUC) across five replicate model runs using a 5-fold cross validation procedure, in which a model was built using subsets of the locality data and the performance of the model was tested on the unused data; this process was repeated with different data partitions. For details about the sensitivity of model performance to changes in background extent, number of background points, and choice of features see Table A.1. Suitability scores produced by MaxEnt are bounded by 0 and 1, with scores near 1 representing high suitability. The scores used for calculating isolation were at a resolution of 0.0083 by 0.0083 degrees. Our MaxEnt model performed reasonably well, with an average AUC score of 0.805 from five cross-validation runs; therefore we proceeded with our calculations of population isolation. Our isolation metrics were calculated as 1 - (average suitability of all cells in a 1, 5, or 10 km radius of each point).

2.2.4 Locality-specific climate data

We chose to use climate data from ClimateWNA (Wang et al., 2012) for our analyses because this program provides annual data and because ClimateWNA uses elevation and partial derivative functions to downscale climate data to precise localities rather than averaging across a grid cell. Site-specific data associated with each locality was downloaded across all years of data availability

(1902-2012). We then pulled out year-specific values for each record as well as averages of the 30 preceding years. These data were compiled for two sets of localities: the set of specimens we measured (plant characteristics data set, n = 120) and all available *C. pulchella* localities including the specimen localities (spatial analyses data set, n = 287; we did not include field observations because field surveys were concentrated in the northern half of the range). For specimens collected before 1933 (n = 18 in spatial analyses data set, n = 2 in plant characteristics data set), we did not have 30 years of data to average, so the averages for these specimens represent the data available. Specimens collected before 1902 or after 2012 (n = 7 in spatial analyses data set, n = 2 in plant characteristics data set) were not used in the climate analyses, but were included in spatial isolation analyses. For each specimen, we calculated the difference in each climatic variable between the year of collection and the 30-year average. We maintained directionality when calculating deviation in precipitation and the beginning of the frost-free period; a negative precipitation deviation represents less precipitation than average in the year of collection, and a negative beginning of the frost-free period represents an earlier beginning than average. Because we hypothesized that both hot and cold deviations in temperature would negatively affect reproductive output, we used the absolute deviation for temperature and degree-days variables. We did not include predictors that were correlated above r = 0.75 within each temporal category (year of collection, 30-year average, and deviation from average) in these analyses. This resulted in the exclusion of degree-days above 5°C (correlated with all temperature measures) and the beginning of the frost-free period (correlated with spring temperatures) from the year of collection and 30-year average analyses. Some climate variables had to be transformed to obtain normality: year of collection fall precipitation, year of collection spring precipitation, 30-year average fall precipitation, and 30-year average spring precipitation values were log-transformed; year of collection summer precipitation, the deviation from average degree days above 5°C, and the deviation from average temperature in each season were square-root-transformed; the deviation from average spring and summer precipitation was translated so that the minimum value was 1 and then square-root-transformed; and 30-year average spring precipitation was log and square-root-transformed.

2.2.5 Statistical analyses

We hypothesized that precipitation and average temperature during germination and seedling establishment (September-November), vegetative growth (March-May), and reproduction (June-July) would affect reproductive output (Figure 2.1A). We also included the date of the beginning of the frost-free period and the degree-days above 5°C in our analyses. We did not examine winter variables since we thought these were likely to affect survival only. Winter (December-February) climate averages were also strongly correlated with fall (September-November) averages (temperature, r = 0.86, precipitation, r = 0.96). We regressed log-transformed reproductive output on the year of collection values for each variable. We also regressed log-transformed reproductive output on deviation from average for each variable (in this case, including degree-days above 5°C and the beginning of the frost-free period) to test whether deviation from normal conditions affected

reproductive output.

We hypothesized that drought stress due to low precipitation and high average temperature in spring and summer would increase propensity for self-pollination and that this could occur due to both via plastic effects within year and longer-term selection (Figure 2.1B). To test this, we regressed both petal length and herkogamy on climate in the year of collection and the 30-year average of each climate variable. We also predicted that spatial isolation would be related to propensity for self-pollination (Figure 2.1C), so we regressed petal length and herkogamy on the suitability-based spatial isolation metric calculated over 1, 5, and 10 km buffers.

Finally, we performed a linear regression of reproductive output and floral traits (herkogamy and petal length; here we use the plant characteristics data set) with distance from the centre of the range (Figure 2.1FG). Additionally, we regressed reproductive output and floral traits on distance from the centre of the range broken down by geographic quadrant. We only tested these relationships in quadrants where significant climatic predictors of a given plant characteristic were also significantly related to distance from the range centre.

2.3 Results

2.3.1 Climate and plant reproductive output

An overview of significant results is provided in Figure 2.4. Specimens from sites with high summer precipitation in the year of collection had higher reproductive output (Table 2.1, Figure 2.5E). Similarly, specimens collected in years with higher positive deviations from average summer precipitation in their collection sites had higher reproductive output (Table 2.1). Year of collection and deviation from average values for other climatic variables were not related to reproductive output (Table 2.1).

2.3.2 Climate, isolation, and floral traits

Plants from sites with warmer temperatures in spring (both in the year of collection and on average) and summer (average only) had reduced herkogamy (Table 2.2, Figure 2.5B). Precipitation variables did not predict herkogamy (Table 2.2). Petal length was not related to any of the year of collection climatic variables or 30-year averages of climatic variables that we examined (Table 2.2). Isolation was not related to either floral trait on any spatial scale (Table 2.3).

2.3.3 Variation in climate, isolation, and plant characteristics across the range

Isolation increased with increasing distance from the centre of the range when calculated across a 10 km area around populations (Table 2.4). When broken down by geographic quadrant, isolation increased toward the southern and western range edges at all spatial scales, but not toward northern and eastern range edges (Table 2.4).

Significant predictors of reproductive output included year of collection summer precipitation

and the deviation from average summer precipitation. The coefficient of variation in summer precipitation decreased with distance from the range centre toward the northern range margin and the eastern range margin and increased toward the southern range margin and the western range margin (Table 2.5). Year of collection summer precipitation decreased with distance from centre toward the western range edge only (Table 2.5, Figure 2.5).

Significant predictors of floral traits include 30-year averages of spring and summer temperature and year of collection spring temperature. Spring temperatures of both timespans decreased toward range edges across all points (Table 2.5). When broken down by geographic quadrant, spring temperatures decreased toward northern and eastern range margins (Table 2.5). Summer temperatures increased toward the southern range margin (Table 2.5, Figure 2.5A).

Petal length, herkogamy, and reproductive output were not related to distance from the centre of the range (petal length: $F_{1,118} = 0.0292$, $P = 0.86$; herkogamy: $F_{1,118} = 0.0460$, $P = 0.0830$; reproductive output: $F_{1,113} = 2.35$, $P = 0.128$). When broken down by quadrant, only the western quadrant showed significant declines in reproductive output with increasing distance from the centre of the range (Table 2.6, Figure 2.5F).

2.4 Discussion

In this study, we examined the relationship between climate and reproductive output as well as the relationship between climate, spatial isolation, and mating-system-related floral traits of the annual herb, *Clarkia pulchella*. Once we had determined the significant predictors of plant characteristics, we examined which of these predictors varied predictably across the range of the species and then tested whether the characteristics of interest changed in space along with their climatic predictors. We found that low summer precipitation was related to low reproductive output toward western (and possibly southern) range edges, while high spring and summer temperatures may increase propensity for self-pollination at the southern range margin. On the whole, this suggests that underlying climatic drivers cause spatial patterns in mating-system-related floral traits and reproductive output, but that these patterns may only occur at some range edges. Below we discuss these results in more detail and their implications for understanding feedbacks between range geography, climate, and mating systems.

2.4.1 Climate, range position, and reproductive fitness

Of the variables we considered, the one with the strongest relationship with reproductive output of *Clarkia pulchella* is summer precipitation. This influence is reflected by the positive effects of both precipitation in the year of collection (Figure 2.5E) and positive deviations from average precipitation, which are correlated with each other. Summer precipitation in sites occupied by *C. pulchella* tends to decrease toward the species' western range margin and may be an important factor limiting reproductive output on that edge (Figure 2.5DF). Similarly, populations toward both the southern and western range edges experience greater interannual variation in precipitation, which may contribute to variance in reproductive output and hence population declines. In

contrast, precipitation is unlikely to limit reproductive output at the northern and eastern edges, because populations near those edges do not experience declines in summer precipitation and show significant reductions in interannual variation in summer precipitation when compared with populations near the range centre. Our results support the inference that different edges are likely limited by different climatic factors.

Although summer precipitation is a significant predictor of reproductive output, and summer precipitation changes with range position, the proportion of variation in reproductive output explained by each of these analyses is low. This unexplained variation may be why, with the exception of the west range quadrant, we failed to detect relationships between range position and reproductive output. This result highlights the fact that when conducting studies of geographic variation across ranges, it is critical to consider intermediate mechanisms, such as climate, in addition to spatial position. Otherwise, important patterns may be obscured by landscape heterogeneity.

2.4.2 Climate, range position, and floral traits

Greater potential for self-pollination (as suggested by reduced herkogamy) is positively related to temperatures in spring and summer (Figure 2.5B). High temperatures may increase drought stress, which may shorten plant lifespans or accelerate flower senescence, making self-pollination adaptive (Mazer et al., 2010). Summer temperature increases toward the southern range margin, however, herkogamy did not decline towards the southern range edge (Figure 2.5C). The climatic predictors of reduced herkogamy were not correlated with low numbers of reproductive structures, indicating that self-pollination is not likely a result of the inability of individuals to allocate resources to pollinator attraction. A relationship between climate and mating system may not be caused by direct effects of climate on plants, but may be mediated by changes in pollinator abundance along climatic gradients (Moeller, 2006). In another member of the genus, *Clarkia xantiana* ssp. *xantiana*, absence of pollinators contributes to one range edge, and beyond this range edge, a self-pollinating sister species occurs (Moeller et al., 2012). Though floral size is indicative of mating system within and among other species of *Clarkia* (Mosquin, 1964; Gottlieb and Ford, 1988; Runions and Geber, 2000), petal length did not show the same patterns as herkogamy in our analyses. Overall, we may have had greater statistical power to detect relationships with reproductive output than with floral traits due to the latter's lower range of variation relative to measurement precision.

Increasing prevalence of climatic conditions that correlate with self-pollination-related traits near the southern range margin may have genetic repercussions for these populations. Experimental populations of *C. pulchella* showed that low genetic effective population sizes can reduce fitness and increase population extinction probability (Newman and Pilson, 1997). It is possible that feedback between the demographic benefits of self-pollination and the genetic effects of self-pollination could maintain a stable range boundary at this edge. However, as in our analyses of reproductive output, there is still a large amount of unexplained variation in the relationship between temperature and herkogamy (Table 2.2) and between range position and temperature (Table 2.5). Perhaps because of this unexplained variation, we did not detect a significant relationship between range position

and herkogamy in the southern quadrant, although the slope of the nonsignificant trend is in the anticipated direction (reduced herkogamy toward range margins; Figure 2.5C).

2.4.3 Isolation, range position, and self-pollination

Isolation, as we have quantified it in this study, increases toward southern and western range margins, consistent with the abundant-centre hypothesis. However, isolation is not correlated with floral traits. We hypothesized that isolation would promote self-pollination due to limited mate availability. However, on heterogeneous landscapes with high gene flow, self-pollination may prevent genetic swamping of local adaptation by gene flow from other populations. In that case, self-pollination would be expected to be advantageous in areas with high spatial environmental heterogeneity and high potential for maladaptive gene flow, which may be areas of high population density. If this occurs, isolation is likely to have complex effects on mating system that differ from our predictions.

The scale at which isolation affects mate availability is an important consideration. Our metric, calculated at a 1-10 km scale, is a proxy for the density of populations or patches on the landscape. It is possible that for many species, including *C. pulchella*, population size and local density within a patch at the scale of meters is important for attracting pollinators and achieving successful pollen transfer. If so, our metric is not likely to have captured the relevant scale for selection on mating system. Another potentially important factor not considered here is the community context of pollination. Competition for pollinators may reduce visitation rates in a plant population (Mitchell et al., 2009) and increase selection for self-pollination (Fishman and Wyatt, 1999). The presence of exotic neighbouring plants can reduce pollinator visitation to *C. pulchella* (Palladini and Maron, 2013). Alternatively, proximity to other plant species that share pollinators may increase the potential for a plant community to support pollinators, and this could help overcome Allee affects that a plant population might face in the absence of that neighbouring plant community (Johnson et al., 2003; Moeller and Geber, 2005a). A final consideration with regard to isolation is its temporal scale. Our isolation metric is based on recent climate normals, but if isolation has changed over longer timescales, then the effects of historic isolation on present-day mating systems would not be captured by our analyses.

2.4.4 Metapopulation dynamics

Alternative predictions for geographic patterns of mating system variation have been derived from metapopulation models. Metapopulation models of geographic distributions are built on underlying gradients of extinction rates, colonization rates, or habitat availability. Some models indicate that range edges may have greater rates of population turnover (Lennon et al., 1997; Holt and Keitt, 2000). Baker (1955) predicted that self-compatible individuals are more likely to establish populations after dispersal. If these two predictions are considered together, it is expected that populations on the periphery of a species' range are likely to be founded by self-compatible individuals with floral traits that facilitate self-pollination (Pannell and Barrett, 1998; Brys et al.,

2013). If true, these predictions could yield range-wide patterns in mating system similar to those predicted along climatic gradients or gradients of increasing isolation. In the case of species' ranges that are not in equilibrium, self-pollination may be prevalent on expanding edges, since populations are likely to be founded by self-compatible (or autonomously self-pollinating) individuals (Baker, 1955; Van Kleunen et al., 2007).

2.4.5 Use of herbarium specimens

This study highlights the potential utility of herbarium specimens for studies of within-species variation. Herbarium specimens may offer a greater temporal and spatial range of sampling than field logistics will typically allow. Efforts to add specimen information and images to public databases are very important for improving the efficiency and comprehensiveness of research that relies on herbarium data. There are, of course, limitations to the utility of these specimens. They do not allow for the analysis of within-population variation or population means, which are both statistically and biologically important. This limitation likely contributed to the unexplained variance in our analyses. Additionally, the geographic coordinates associated with specimens vary in their reliability and availability. Further, if some populations experience shorter seasons than others, they may have less opportunity to be collected; therefore, specimens from localities with climatic conditions that shorten the flowering season may be underrepresented in herbarium collections. Finally, geographic sampling is likely to be biased toward roads and areas frequented by collectors. Nonrandom sampling of the geographic range may lead to distribution model predictions that model sampling effort rather than suitability.

2.4.6 Conclusions and future directions

The results of this study suggest that some aspects of climate contribute to variation in reproductive output and herkogamy in *Clarkia pulchella* and that spatial variation in these plant characteristics is suggestive of climatically driven range-limitation at some edges. Field studies that consider plant population size and pollinator communities will tell us more about how climate affects plant fitness and mating system, and such studies may be particularly appropriate at the southern and western range edges of *C. pulchella*. These should be complemented by studies of the mechanism by which temperature affects mating-system-related traits and by studies testing the link between floral traits, environmental conditions, and realized rates of self-pollination. Understanding the effects of abundance on mating-system-related traits requires further consideration of the relevant scale of population isolation, the role of population size and density in shaping selection on mating-system-related traits, and the geographic distribution of population sizes and densities. Future work should also consider the effects of climatic conditions and co-flowering species on pollinator visitation. Reproductive output at the northern and eastern range margins does not appear to be limited by the climatic variables tested here. Future work should consider other factors that may limit the range at these edges, including the effects of environmental conditions on life stages other than reproduction, the role of swamping gene flow, and dispersal limitation.

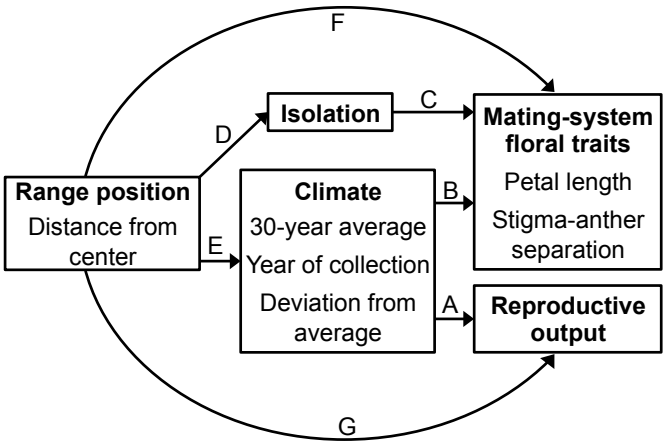

Figure 2.1: Conceptual map of analytical framework for assessing the effects of climate, isolation, and range position on floral traits and reproductive output of *Clarkia pulchella*. Each arrow represents a tested relationship between range position, climate, isolation, and plant characteristics. Letters are referenced in text. For details of analyses, see methods.

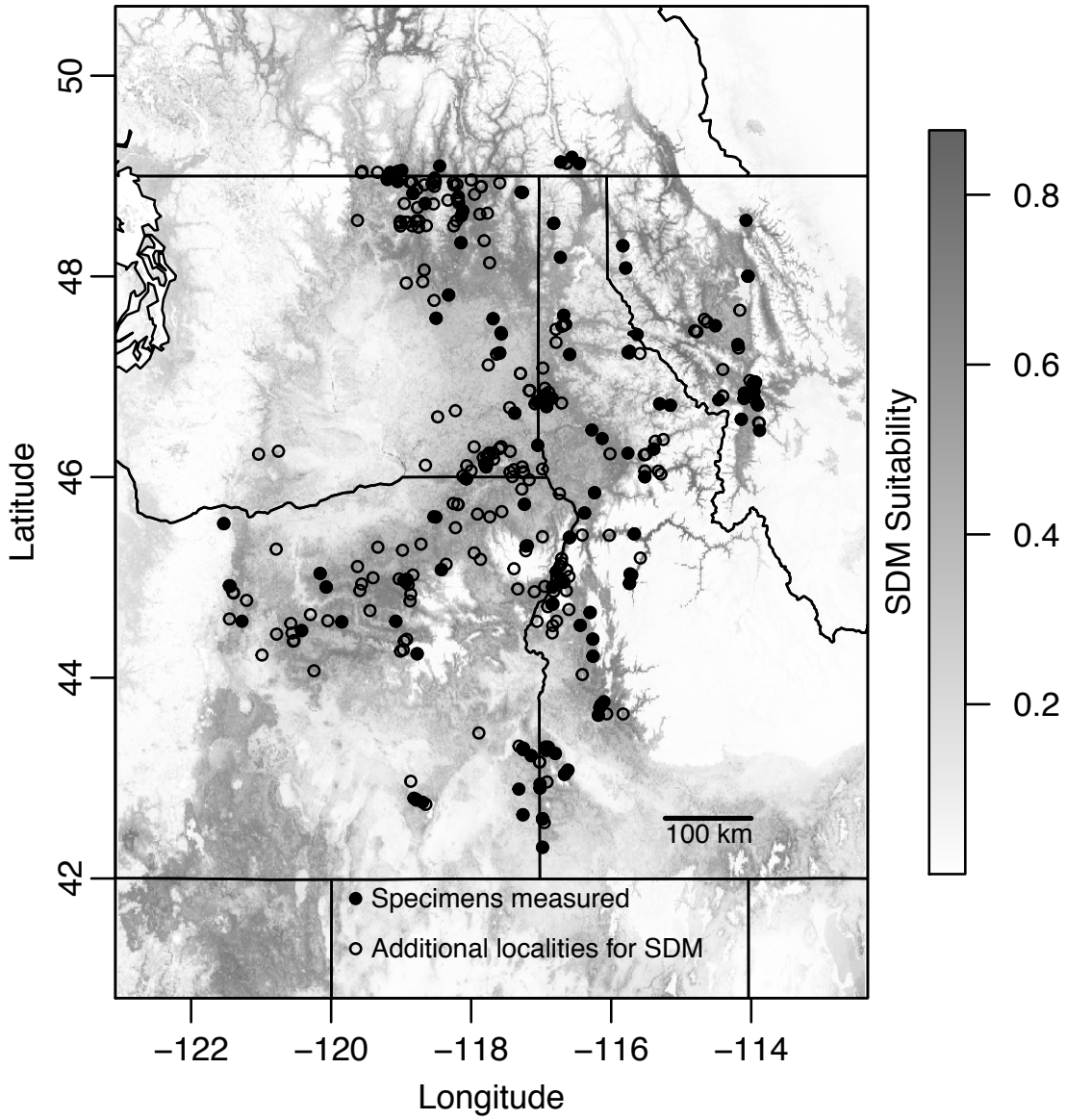

Figure 2.2: Map of *Clarkia pulchella* localities across the species' range in the Pacific Northwest. Filled circles represent herbarium specimens measured for analyses of mating system traits or reproductive output. Open circles represent additional localities used to build a species distribution model. Background shading shows predictions of the species distribution model, where darker shades indicate higher suitability.

Figure 2.3: Diagram showing the measurements made on flowers of *Clarkia pulchella* herbarium specimens.

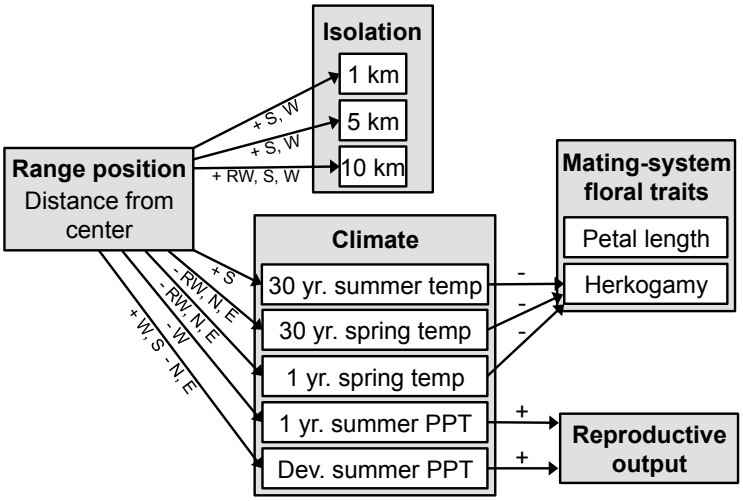

Figure 2.4: Summary of results of analyses of the effects of climate, isolation, and range position on floral traits and reproductive output of *Clarkia pulchella*. Each arrow represents a significant relationship from linear regressions. Positive (+) and negative (-) relationships are indicated along each arrow, as well as whether the relationships are significant range-wide (RW) or within certain range quadrants (N, W, E, or S), when applicable.

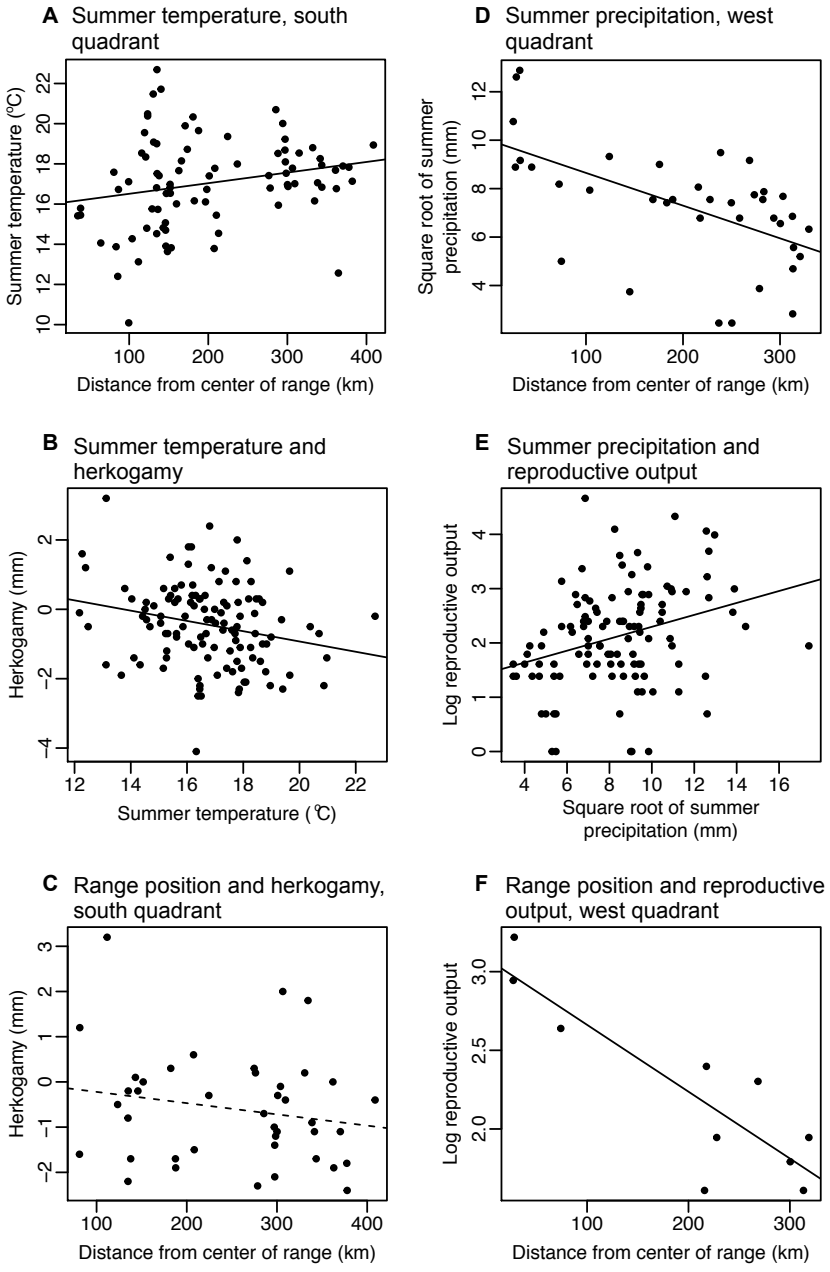

Figure 2.5: Relationships between range position, climate, and *Clarkia pulchella* characteristics. Solid lines represent fits of significant linear models; the dashed line represents a nonsignificant trend. **(A)** The 30-year average summer temperatures increase with distance of localities from the range centre in the southern range quadrant. **(B)** Herkogamy declines with increasing summer temperatures for specimens collected across the range. **(C)** There is no significant effect of increasing distance from the centre of the range on herkogamy in the southern range quadrant of *C. pulchella*. **(D)** Summer precipitation in the year of collection decreases with increasing distance of specimens from the range centre in the western range quadrant. **(E)** Reproductive output is positively correlated with summer precipitation for all specimens across the range of *C. pulchella*. **(F)** Reproductive output declines with increasing distance from the range centre in the western quadrant of the range.

Table 2.1: Effects of climate on reproductive output of *Clarkia pulchella* across the species' range. Year of collection variables are climatic conditions for each specimen in the year that it was collected. Deviations from average are calculated as the difference between the value in the year of collection and the average of the 30 years preceding the year of collection. For temperature and degree-day variables, all deviations are absolute; however, for precipitation and the beginning of the frost-free period, directionality of deviation was maintained. Log-transformed reproductive output was regressed on each climatic variable. n = 113 for all tests. Bold text indicates significant tests.

Climate variable	Slope	Slope SE	$F_{1,111}$	P	R^2
Year of collection					
Fall precipitation (Oct-Dec)[a]	0.096	0.149	0.41	0.521	0.004
Spring precipitation (Mar-May)[a]	-0.014	0.160	0.01	0.930	0.000
Summer precipitation (Jun-Jul)[b]	**0.110**	**0.033**	**11.33**	**0.001**	**0.093**
Fall temperature	-0.088	0.048	3.34	0.070	0.029
Spring temperature	-0.039	0.049	0.64	0.426	0.006
Summer temperature	-0.059	0.041	2.11	0.149	0.019
Deviation from average					
Fall precipitation (Oct-Dec)[a]	0.002	0.001	1.73	0.191	0.015
Spring precipitation (Mar-May)[b]	-0.022	0.033	0.43	0.512	0.004
Summer precipitation (Jun-Jul)[b]	**0.143**	**0.041**	**12.44**	**0.001**	**0.101**
Beginning of the frost-free period	0.006	0.008	0.67	0.415	0.006
Degree days > 5°C[b]	-0.015	0.022	0.48	0.490	0.004
Fall temperature[b]	-0.136	0.261	0.27	0.604	0.002
Spring temperature[b]	0.103	0.234	0.19	0.662	0.002
Summer temperature[b]	-0.153	0.253	0.36	0.547	0.003

[a] Log-transformed before analysis.
[b] Square-root-transformed before analysis.

Table 2.2: Effects of climate on floral traits of *Clarkia pulchella* across the species' range. Year of collection variables are climatic conditions of each specimen in the year that it was collected, and 30-year averages are the average of the 30 years preceding the year of collection. Two floral traits, petal length and herkogamy, were regressed on each climatic variable. Petal length was square-root-transformed. n = 118 for all tests. Bold text indicates significant tests.

Climate variable and floral measure	Slope	Slope SE	$F_{1,116}$	P	R^2
30-year averages					
Spring PPT (Mar-May)[a]					
Petal length	0.172	0.505	0.12	0.735	0.001
Herkogamy	0.830	1.201	0.48	0.491	0.004
Summer PPT (Jun-Jul)[b]					
Petal length	0.014	0.118	0.01	0.908	0.000
Herkogamy	0.342	0.280	1.49	0.225	0.013
Spring temperature					
Petal length	0.002	0.029	0.00	0.947	0.000
Herkogamy	**-0.145**	**0.067**	**4.68**	**0.033**	**0.039**
Summer temperature					
Petal length	-0.009	0.025	0.13	0.715	0.001
Herkogamy	**-0.148**	**0.057**	**6.66**	**0.011**	**0.054**
Year of collection					
Spring PPT (Mar-May)[b]					
Petal length	0.032	0.084	0.14	0.708	0.001
Herkogamy	0.219	0.200	1.20	0.275	0.010
Summer PPT (Jun-Jul)[c]					
Petal length	0.014	0.018	0.64	0.425	0.006
Herkogamy	0.039	0.042	0.84	0.361	0.007
Spring temperature					
Petal length	0.013	0.026	0.24	0.627	0.002
Herkogamy	**-0.126**	**0.061**	**4.30**	**0.040**	**0.036**
Summer temperature					
Petal length	-0.002	0.022	0.01	0.943	0.000
Herkogamy	-0.082	0.051	2.53	0.114	0.021

[a] Square-root- and log-transformed before analysis.
[b] Log-transformed before analysis.
[c] Square-root-transformed before analysis.

Table 2.3: Effects of isolation on floral traits of *Clarkia pulchella* across the species' range. Isolation was calculated at three spatial scales: 1, 5, and 10 km. Two floral traits, petal length and herkogamy, were then regressed on isolation. Petal length was square-root-transformed. n = 120 for all tests.

Isolation scale and floral measure	Slope	Slope SE	$F_{1,118}$	P	R^2
Isolation, 1 km[a]					
Petal length	-0.043	0.152	0.08	0.778	0.001
Herkogamy	0.103	0.362	0.08	0.777	0.001
Isolation, 5 km					
Petal length	-0.151	0.344	0.19	0.662	0.002
Herkogamy	0.382	0.821	0.22	0.643	0.002
Isolation, 10 km					
Petal length	-0.185	0.376	0.24	0.623	0.002
Herkogamy	0.157	0.899	0.03	0.862	0.000

[a] Log-transformed before analysis.

Table 2.4: Effect of range position on spatial isolation of populations of *Clarkia pulchella*. Range position was measured as the distance between a specimen's latitude and longitude coordinates and the coordinates of the range centroid. Isolation (at three spatial scales) was then regressed on distance from the centre. Each test was performed on all localities across the range, and separately on localities occurring in each of four geographic quadrants, as designated by NW-SE and NE-SW diagonals through the range centroid. Isolation variables were all log-transformed before analysis. Bold text indicates significant tests.

Isolation scale and region	n	Slope	Slope SE	F	df	P	R^2
1 km							
All	260	0.0003	0.0002	2.90	1,258	0.090	0.011
North	81	-0.0004	0.0003	2.47	1,79	0.120	0.032
South	**84**	**0.0014**	**0.0003**	**20.64**	**1,82**	**<0.001**	**0.201**
West	**37**	**0.0009**	**0.0004**	**4.77**	**1,35**	**0.036**	**0.120**
East	58	-0.0004	0.0004	1.36	1,56	0.249	0.024
5 km							
All	260	0.0003	0.0001	3.51	1,258	0.062	0.014
North	81	-0.0002	0.0003	0.86	1,79	0.356	0.011
South	**84**	**0.0011**	**0.0003**	**16.11**	**1,82**	**<0.001**	**0.164**
West	**37**	**0.0009**	**0.0003**	**7.08**	**1,35**	**0.012**	**0.168**
East	58	-0.0001	0.0003	0.03	1,56	0.681	0.003
10 km							
All	**260**	**0.0003**	**0.0001**	**4.69**	**1,258**	**0.031**	**0.018**
North	81	0.0000	0.0002	0.03	1,79	0.854	0.000
South	**84**	**0.0009**	**0.0002**	**14.42**	**1,82**	**<0.001**	**0.150**
West	**37**	**0.0009**	**0.0003**	**7.06**	**1,35**	**0.012**	**0.168**
East	58	0.0000	0.0003	0.03	1,56	0.873	0.000

Table 2.5: Relationship between range position and climate. Climatic variables used in these analyses only include significant drivers of reproductive output (coefficient of variation of summer precipitation and year of collection summer precipitation) and herkogamy (spring and summer temperature). These variables were then regressed on the distance of specimens from the centre of the range. Each test was performed on all localities across the range, and separately on localities occurring in each of four geographic quadrants. Bold text indicates significant tests.

Climate variable and region	Slope	Slope SE	n	F	df	P	R^2
Coefficient of variation of summer precipitation							
North	**-0.0001**	**0.0000**	**92**	**20.79**	**1,90**	**<0.001**	**0.188**
South	**0.0006**	**0.0000**	**88**	**298.78**	**1,86**	**<0.001**	**0.776**
West[a]	**0.0005**	**0.0001**	**37**	**26.30**	**1,35**	**<0.001**	**0.429**
East	**-0.0002**	**0.0000**	**60**	**40.55**	**1,58**	**<0.001**	**0.411**
Year of collection summer precipitation (Jun-Jul)							
All[b]	-0.0016	0.0015	278	1.11	1,276	0.294	0.004
North[b]	-0.0011	0.0021	90	0.29	1,88	0.593	0.003
South[b]	-0.0034	0.0024	90	2.05	1,88	0.156	0.023
West[b]	**-0.0108**	**0.0035**	**37**	**9.71**	**1,35**	**0.004**	**0.217**
East[b]	0.0050	0.0032	61	2.48	1,59	0.121	0.040
Year of collection spring temperature (Mar-May)							
All	**-0.0037**	**0.0011**	**278**	**11.41**	**1,276**	**0.001**	**0.040**
North	**-0.0061**	**0.0018**	**90**	**11.84**	**1,88**	**0.001**	**0.119**
South	0.0007	0.0021	90	0.10	1,88	0.749	0.001
West[a]	0.0001	0.0004	37	0.06	1,35	0.808	0.002
East	**-0.0059**	**0.0025**	**61**	**5.32**	**1,59**	**0.025**	**0.083**
30-year average of spring temperature (Mar-May)							
All	**-0.0037**	**0.0009**	**283**	**16.42**	**1,281**	**<0.001**	**0.055**
North	**-0.0062**	**0.0015**	**90**	**17.10**	**1,88**	**<0.001**	**0.163**
South	0.0003	0.0016	95	0.03	1,93	0.862	0.000
West[a]	-0.0027	0.0026	37	1.13	1,35	0.295	0.031
East	**-0.0050**	**0.0021**	**61**	**5.57**	**1,59**	**0.022**	**0.086**
30-year average of summer temperature (Jun-Jul)							
All	-0.0007	0.0011	283	0.44	1,281	0.507	0.002
North	-0.0031	0.0017	90	3.07	1,88	0.083	0.034
South	**0.0007**	**0.0003**	**95**	**5.90**	**1,93**	**0.017**	**0.060**
West[b]	-0.0003	0.0003	37	0.66	1,35	0.421	0.019
East	-0.0029	0.0022	61	1.73	1,59	0.194	0.028

[a] Log-transformed prior to analysis.
[b] Square-root-transformed prior to analysis.

Table 2.6: Relationship between range position and reproductive output or herkogamy of *Clarkia pulchella* by quadrant. Tests were only performed using data from quadrants where results of prior analyses indicated that reproductive output or herkogamy might be associated with range position. Either reproductive output or herkogamy was regressed on distance from the range centre. Reproductive output was log-transformed before analysis. Bold text indicates significant tests.

Plant measure, range quadrant	Slope	Slope SE	n	F	df	*P*	R^2
Reproductive output							
North	0.0017	0.0015	37	1.38	1,35	0.247	0.038
East	0.0004	0.0022	30	0.03	1,28	0.853	0.001
South	0.0000	0.0018	38	0.00	1,36	0.991	0.000
West	**-0.0043**	**0.0009**	**10**	**25.28**	**1,8**	**0.001**	**0.760**
Herkogamy							
North	0.0011	0.0020	38	0.29	1,36	0.600	0.008
East	0.0023	0.0020	31	1.41	1,29	0.245	0.046
South	0.0025	0.0020	41	1.49	1,39	0.229	0.037

Chapter 3

Geographic and climatic drivers of reproductive assurance in *Clarkia pulchella*

3.1 Introduction

Climate change can affect population dynamics directly by altering the survival and reproduction of individuals (McGraw et al., 2015). In addition to these direct effects, climate change can indirectly affect species by altering their interactions with mutualists, predators, or competitors (Miller-Struttmann et al., 2015). To make informed predictions about species' responses to climate change, we must understand both direct and indirect effects. For plant species, pollinators are likely to be an important medium for these indirect effects, as the reproductive success of primarily outcrossing taxa is often highly dependent on the actions of these mutualists (Burd, 1994; Ashman et al., 2004). Changing environmental conditions can disrupt the reliability of pollination (Kudo et al., 2004). For example, changes in phenological cues might lead to mismatch between plants and pollinators (Kudo and Ida, 2013), pollinator populations may decline if they are maladapted to changing conditions (Williams et al., 2007), and the presence of invasive species can reduce visitation to native plants (Bjerknes et al., 2007; Bruckman and Campbell, 2016).

In the face of sustained mate or resource limitation, reliance on outcross pollen can limit seed production, and selection might favour individuals with floral traits that facilitate reproductive assurance via self-pollination (Bodbyl Roels and Kelly, 2011), including traits that allow for delayed self-pollination when outcross pollen has not been delivered. Reproductive assurance is the ability to self-pollinate, either autonomously or with the assistance of a pollinator, in order to offset deficits in pollen delivery. Limited resources, including limited water availability, can increase the cost of producing and maintaining attractive floral displays (Galen et al., 1999). This could lead to selection for individuals that can achieve high reproductive success without incurring the costs of showy displays. Similarly, short flowering seasons may increase the risks of waiting for pollinator

service. Some habitat characteristics, such as limited numbers of suitable growing sites, may lead to sparser or smaller populations and in turn, mate limitation. Mate limitation can also occur even when conspecific individuals are abundant if pollinators are low in abundance or prefer to visit co-occurring species (Knight et al., 2005). When temporal variability in environmental conditions is high, selection might alternatively favour plasticity that allows for increased self-pollination in response to environmental cues associated with pollen limitation (Kay and Picklum, 2013).

Mate and resource limitation can co-vary with climatic conditions. Therefore, patterns in mating system traits may be correlated with the climatic gradients that underlie a species' geographic distribution. While climatic conditions can exert selection on mating system and, as a result, indirectly affect demographic (Lennartsson, 2002; Moeller and Geber, 2005b) and genetic processes (Eckert et al., 2010; Kramer et al., 2011), climate can also directly affect demographic components. Climatic gradients may shape variation in life history or in the sensitivity of population growth rate to a specific demographic stage, leading to measurable correlations between some fitness components and climate variables across space (Doak and Morris, 2010). Inter-annual variability in climate may also be correlated with temporal variation in vital rates within a single population (Coulson et al., 2001). Our understanding of how climate affects population dynamics will benefit from examining the relationships of multiple variables (fitness components or strengths of biotic interactions) to variation in climate.

Biogeographic processes also shape mating system variation on the landscape. During range expansions, individuals capable of reproduction in the absence of mates or pollinators are expected to be more likely to found new populations (Baker, 1955; Pannell et al., 2015), creating a geographic cline in mating system variation, with a greater degree of self-compatibility or capacity for self-pollination near expanding or recently expanded range edges. Similar patterns might arise in regions where populations turn over frequently, where the ability to reproduce autonomously may be an important trait for individuals that are colonizing empty patches. Geographic variation in mating system can also be attributed to range overlap with pollinator taxa or with plant taxa that share pollinators. In mixed-mating plants, parts of the range that overlap with a reliable pollinator community might experience little selection for self-pollination. Overlap with a competing plant species may reduce pollination success and lead to selection for self-pollination.

Empirical examinations of mating systems are infrequently carried out at the scale of geographic ranges (with exceptions including Busch 2005; Herlihy and Eckert 2005; Moeller and Geber 2005b; Dart et al. 2011; Mimura and Aitken 2007) and investigations of geographic variation in vital rates rarely consider mating system variation. The interplay of vital rates and mating systems across geographic and climatic space may be relevant not only to population dynamics within the range, but also to the dynamics that limit geographic distributions. Across environmental gradients, mating system variation might interact with other genetic and demographic processes to influence population persistence and adaptive response. For example, while highly selfing individuals might be expected to be good colonizers, they also might have limited genetic variation for adaptation to novel environments beyond the range edge (Wright et al., 2013). Investigating range-wide variation

in reproduction may shed light on climate variables that limit range expansion.

In this study, we investigate the relationships among climate, pollinator exclusion, and reproductive fitness components of a winter annual wildflower, *Clarkia pulchella*. In a previous study, we used herbarium specimens to examine relationships between climate, mating system, and reproductive characteristics of this species. We found that summer precipitation was positively correlated with reproductive output and that warm temperatures were correlated with traits indicative of self-pollination (Chapter 2; Bontrager and Angert, 2016). Here, we employed field manipulations across the range of *C. pulchella* to examine whether reproductive assurance co-varies with geographic range position and/or climate. We were specifically interested in the autonomous component of reproductive assurance, that is, the ability to transfer self-pollen in the absence of a pollinator (rather than the degree to which pollinators transfer self-pollen). *C. pulchella* grows in sites that are very dry during the flowering season, particularly at the northern and southern range edges, so we expected that plants in these regions might have greater capacity to self-pollinate as a means of ensuring reproduction before drought-induced mortality. We therefore predicted that range edge populations would have greater capacity to self-pollinate in the absence of pollinators, and that this geographic pattern would be attributable to climate, in particular, drought stress during the flowering season (summer precipitation and temperature). We also sought to determine whether short-term drought relief produced consistent mating system responses across the range of *C. pulchella*. We hypothesized that drought would induce a plastic increase in self-pollination, and that as a result we would see reduced reproductive assurance when drought relief was combined with pollinator limitation. Finally, we examined how variation in pollinator availability and climate affect different components of reproduction. We anticipated that drought relief would have opposing effects on reproductive assurance and fruit production: while drought may prompt plastic increases in reproductive assurance, higher water availability likely increases plant longevity and productivity during the flowering season.

3.2 Methods

3.2.1 Study system

Clarkia pulchella Pursh (Onagraceae) is a mixed-mating winter annual that grows east of the Cascade Mountains in the interior Pacific Northwest of North America (Figure 3.1). The species is found in populations ranging in size from hundreds to thousands of individuals on dry, open slopes in coniferous forest and sagebrush scrub. It is primarily outcrossed by solitary bees (Palladini and Maron, 2013) with a diverse array of other pollinators (MacSwain et al., 1973), but selfing can be facilitated by spatial and temporal proximity of fertile anthers and stigma within flowers. As the anthers dehisce, pollen is often suspended from the anthers on viscin threads, and may come into contact with the stigma. A large portion of the range of *C. pulchella* is in the Okanagan Valley, which is expected to experience warmer temperatures and redistributed rainfall in the coming decades (Figure 3.2). Temperature increases are expected to be especially prominent in the

summer months (Wang et al., 2012; Meyer et al., 2014). Anticipated changes in precipitation are variable and uncertain across the range of our focal species, with many sites expected to experience decreases in summer precipitation, but central sites projected to experience slight increases in annual precipitation (Wang et al., 2012; Meyer et al., 2014).

3.2.2 Plot establishment and monitoring

Experimental plots were established in eight populations on 4-9 June 2015. These sites were located in three regions across the latitudinal range of *Clarkia pulchella*, with two at the species' northern edge in southern British Columbia (Canada), three in the range centre in southeastern Washington (USA), and three in the southwestern portion of the species range, in Oregon (USA; Figure 3.1, Table 3.1). Our original intention was to treat the southern and western edges of the range separately and establish three sites at each edge. However, due to difficulty finding populations of sufficient size in sites where we could also obtain permits, we used just two populations in the west and one in the south. Because the climatic similarity among these sites is nearly comparable to that among sites in other regions (Figure 3.2), we decided to treat them as a single region, the southwest. At each site, 5-8 blocks containing four plots each were marked with 6-inch steel nails, this resulted in a total of 50 blocks and 200 plots in the experiment. Each plot consisted of a 0.8 m^2 area. Plots were intentionally placed with the goal of obtaining 5-20 individuals per plot, therefore the density in plots was typically higher than the overall site density. Plots were placed closer to other plots in their block than to those in other blocks (with exceptions in two circumstances where low plant density meant very few suitable plot locations were available). Blocks were placed to capture variation in microhabitat characteristics across the site, and their spacing varied depending on the population size and density. Each plot was randomly assigned to one of four factorial treatment groups: control, water addition, pollinator exclusion, or both water addition and pollinator exclusion. Plots receiving water additions were at least 0.5 m away from unwatered plots, except when they were downslope from unwatered plots, in which case they were sometimes closer. Plots receiving pollinator exclusion treatments were tented in bridal-veil mesh with bamboo stakes in each corner and nails tacking the mesh to the ground. Some pollinator exclusion plots had their nets partially removed by wind or cows during the flowering season (n = 13 out of 100 total tented plots), so all analyses were performed without these plots.

The majority of the summer precipitation in these sites falls in summer storms. Plots receiving supplemental water were watered 1-2 times during the summer (when plants were flowering) to simulate additional rainfall events. During each watering event, 15 mm of water was added to each plot (9.6 L per plot). This approximated the typical precipitation of a summer rainfall event based on data from Wang et al. (2012), and in an average year, would have increased the total summer precipitation in these plots by 30-70%. However, our experiment was conducted during a drought year (Figure 3.2), therefore, in the central sites, plots receiving water additions still fell short of average summer precipitation levels. In southwestern and northern sites, the water addition likely raised the summer precipitation amount slightly above the historic average. In all sites, we consider

the water additions to represent a drought relief treatment, because unwatered plots were already experiencing natural drought. The first watering was performed when the experiment was set up. The second watering was performed 22-25 June 2015, except at two sites (SW3, C1), which had completed flowering and fruiting at that time. Efficacy of the water addition treatment was checked by measuring the soil water content with a probe (Hydrosense, Campbell Scientific Inc.) before and after water additions. Prior to water additions, there were no significant differences between plots receiving a water addition treatment and those not receiving this treatment (linear mixed effects model with a random effect of site and a fixed effect of water addition treatment; first watering: $P = 0.839$ (7 of 8 sites were measured); second watering: $P = 0.277$ (5 of 8 sites were measured)). Shortly after watering (within one hour), plots receiving a water addition treatment had higher soil moisture than those not receiving treatment (first watering: $P < 0.0001$, average soil moisture of unwatered plots = 11.0% , watered plots 22.2%; second watering: $P = 0.0001$, average soil moisture of unwatered plots 3.7%, watered plots 11.5%).

When flowering and fruiting were complete, we counted the number of plants in each plot and the number of fruits on each plant, and estimated the average number of seeds per fruit. The number of plants in each plot ranged from 1-43 (mean = 7.9, median = 7). We counted the number of fruits per plant on every plant in each plot, as a proxy for the number of flowers per plant (aborted fruits were rare overall). Plants that had died before producing any flowers were not included in our analyses. Some plants (n = 14, 0.7% of all plants counted) had experienced major damage prior to our final census making fruit counting impossible, so they were assigned the average number of fruits per plant in that plot type at that site for estimation of plot-level seed input, but we excluded them from analyses of fruit counts. Other plants (n = 25, 1.4% of all plants counted) still had flowers at the time of the final census. It was assumed that these flowers would ripen into fruits, so they were included in the fruit counts. When possible, up to four fruits per plot (average number of fruits per plot = 3.67) were collected for seed counting. After counting, seeds were returned to the plots that they were collected from by sprinkling them haphazardly over the plot from a 10 cm height. In 3 of 200 plots, no intact fruits were available for seed counting (all had dehisced), so these plots were excluded from analyses of seed set and plot-level seed input, but included in analyses of fruit counts. To assess the subsequent effects of pollinator limitation on populations in the following year, we revisited plots on 21-24 June and 29-31 July 2016 and counted the number of mature plants present in each. Some plot markers were missing, but we were able to relocate 182 of our 200 plots.

3.2.3 Climate variable selection

We expect long-term climatic conditions, particularly those that might contribute to drought stress, to influence selection for autonomous selfing. Concurrent work with *C. pulchella* (Chapter 5) has indicated that fall, winter, and spring growing conditions play a large role in overall plant growth and reproductive output, therefore we considered not only flowering season (June-July) climate variables but also annual temperature and precipitation for inclusion as predictors. We obtained

50-year climate normals (1963-2012) from ClimateWNA (Wang et al., 2012) and climate data during the study from PRISM (PRISM Climate Group, Oregon State University, prism.oregonstate.edu, downloaded 10 October 2016). Our selected set of climatic variables included annual temperature normals (MAT), annual precipitation normals (MAP), summer temperature during the experiment, and summer precipitation during the experiment. Among these, MAT and precipitation during the experiment were correlated (r = -0.84). A full set of annual and seasonal variable correlations is presented in Table A.2.

3.2.4 Statistical analyses

We used generalized linear mixed effects models (GLMMs) to evaluate the effects of pollinator exclusion, region, and each of the selected climate variables on reproductive assurance and fruits per plant. Initial data exploration indicated that our watering treatment did not have a strong or consistent biological effect, so we omitted this factor from our analyses to keep models simple and facilitate interpretation of interactions between the other factors. For each predictor variable of interest (the four climate variables and region), we built a model with a two-way interaction between this variable and pollinator exclusion on both seed counts and fruit counts. We used negative binomial GLMMs for both seeds and fruits, and we included a zero-inflation parameter when modelling seed counts. In all models we included random effects of blocks nested within sites. Because our data do not contain true zero fruit counts (i.e., we did not include plants that did not survive to produce fruits, so all plants in our dataset produced at least one fruit), we subtracted one from all counts of fruits per plant prior to analysis in order to better conform to the assumptions of the negative binomial model. All climate predictors were scaled prior to analyses by subtracting their mean and dividing by their standard deviation. We evaluated the relationship between total plot-level seed production in 2015 and the number of plants present in each plot in summer of 2016 using a GLMM with a negative binomial distribution and random effects of block nested within site. All models were built in R (R Core Team, 2017) using the package glmmTMB (Brooks et al., 2017) and predictions, averaged across random effects, were visualized using the package ggeffects (Lüdecke, 2018).

3.3 Results

3.3.1 Variation in response to pollinator limitation across the range

In all regions, *Clarkia pulchella* produced fewer seeds in the absence of pollinators (Table 3.2). We define reproductive assurance as the number of seeds produced in the absence of pollinators. Climatic or geographic drivers of variation in reproductive assurance were indicated by our models of seeds per fruit when there was a significant interaction between pollinator exclusion and region or pollinator exclusion and a given climate variable. We found that reproductive assurance varied by region, with greater rates of reproductive assurance in northern populations (Figure 3.3, Ta-

ble 3.2). We did not find any strong effects of climate on seed production or reproductive assurance (Table 3.2). However, there was a marginally significant interaction between mean annual precipitation (MAP) and pollinator exclusion: populations in historically wetter sites tended to be more negatively affected by pollinator exclusion (i.e., populations in drier sites had slightly higher rates of reproductive assurance) (Table 3.2). This could be a causal relationship, or the correlation could have been driven by the high degree of reproductive assurance in the northern part of the range, which has low MAP. If low MAP was really a driver of reproductive assurance, we might expect to have seen a greater degree of reproductive assurance in the southwestern sites, which also have low MAP. However, this was not the case in our data.

3.3.2 Response of patch density to seed production in the previous year

Across sites, there was a positive relationship between the number of seeds produced in a plot in 2015 and the number of adult plants present in 2016 ($P < 0.0001$, $\beta = 0.00044$, SE $= 0.000061$; Figure 3.4). This is not simply a result of plots with large numbers of plants in 2015 being similarly dense in 2016, because seed input was decoupled from plant density in 2015 by the pollinator exclusion treatments. The effect of seed input remained significant ($P < 0.0001$) when the number of plants in 2015 was included in the model as a covariate (results not shown).

3.3.3 Variation in fruit production across the range

Plants in the north produced more fruits (on average 4.0, compared to 1.5 and 1.7 in the centre and southwest, respectively; Table 3.3, Figure 3.5). This regional trend could be due to the relatively lower normal annual temperatures in the northern sites (Figure 3.2), the effects of which are discussed below. Pollinator exclusion tended to result in a slight increase in fruit production, possibly due to reallocation of resources within a plant in order to produce more flowers when ovules are left unfertilized (Table 3.3). This effect was small—plants in plots without pollinators produced an additional 0.4 fruits, on average.

We found that the effects of pollinator exclusion on fruit production depended upon the amount of summer precipitation during the experiment (Table 3.3, Figure 3.6). Fruit production was higher in wetter sites, and pollinator-excluded plants that were in the wettest sites showed a greater positive effect of pollinator exclusion on fruit production (Table 3.3). However, it should be noted that while both the main effect of climate and its interaction with pollinator exclusion were significant, the difference between plots with and without pollinators in wetter sites did not appear to be particularly strong, and when visualized the confidence intervals were largely overlapping (Figure 3.6A). We also found a main effect of mean annual temperature (MAT) on fruit production (Table 3.3). Fruit production was higher in cooler sites (Figure 3.6B). Disentangling these two climatic drivers of increased fruit production is not possible with this dataset, however, because summer precipitation during the experiment was negatively correlated with normal MAT. Therefore, it could have been either higher water resources during flowering or cooler temperatures over the growing season that resulted in increased fruit production. It is worth noting, however, that summer temperature

during the experiment was not correlated with either of these variables, so if temperature was the driver of this pattern, it was likely because of temperature effects on earlier life-history stages.

3.4 Discussion

Pollinator exclusion in eight populations of *Clarkia pulchella* revealed increased autonomous reproductive assurance in populations in the northern part of the species' range, as compared to the centre or southwest. Plants in the northern part of the species' range also produced more fruits. Fruit production was higher in sites that are cooler or that received higher amounts of precipitation during the experiment. Plants also produced slightly more fruits in response to pollinator exclusion, however, this reallocation was not, in general, large enough to offset the reduction in seed production caused by pollen limitation.

3.4.1 Reproductive assurance is driven by geography rather than climate

Pollinator limitation reduced reproduction across the range of *C. pulchella*. Contrary to our prediction, we did not observe plastic responses of decreased reproductive assurance in response to our water addition treatment, or in sites with high summer precipitation during the experiment. There is some indication that plants in sites with lower average precipitation may have adapted to have greater reproductive assurance (Table 3.2), perhaps due to shorter season lengths or because gradients in pollinator abundance may be driven by water availability. However, increased reproductive assurance is only apparent at the northern range edge (Figure 3.3) despite the fact that mean annual precipitation is lower at both the northern and southwestern range edges. This general pattern persists even after accounting for regional differences in seed set in control plots, i.e., when reproductive assurance is represented as a proportion of the average seed set in control plots (Figure 3.7). In light of this, we suggest that for this species, reproductive assurance is better explained by the latitudinal position of populations relative to the range than by any single climate variable. The locations of our northern populations were covered by the Cordilleran ice sheet during the last glacial maximum; the patterns we see could be the result of a post-glacial range expansion, in which the founders of these northern populations were individuals who had a greater capacity for autonomous reproduction. It is possible that during colonization there is a low probability of pollinators foraging on a novel plant species and moving conspecific pollen between sparse individuals. Reproductive assurance has evolved in other species when populations have experienced historic bottlenecks (Busch, 2005), and contrasts of species' range sizes indicate that species capable of autonomous self-pollination have a greater ability to colonize new sites (Randle et al., 2009). While latitude is not a strong predictor of among-species variation in mating system (Moeller et al., 2017), within-species variation may be more closely tied to postglacial colonization routes.

An alternative possibility is that our northern sites are distinct because they differ in community composition from sites in other parts of the range. These community differences could be in the regional suite of pollinators. A survey of *Clarkia* pollinators in western North America (MacSwain

et al., 1973) notes that visitors to *C. pulchella* differ from the characteristic groups that visit more southern members of the genus, and it is possible that a similar gradient in pollinator communities exists within the geographic range of *C. pulchella*. Similarly, co-occurring plant species can influence pollinator availability and deposition of conspecific pollen on a focal species (Palladini and Maron, 2013), and it is possible that populations in the northern portion of the range have adapted to a different pollination environment caused by overlap with different plant species.

Across the range, adult plant density was positively correlated with seed production in the previous year. Because our pollinator exclusion treatment led to plot-level seed input being decoupled from the number of plants in 2015 (data not shown), we can attribute differences in 2016 plant density to seed input, rather than to patch quality. Seed production is important enough to have an effect on subsequent density despite differences between plots in the availability of germination sites or the probability of survival to flowering. This, in combination with the consistent negative reproductive response to pollinator exclusion, indicates that populations would likely be negatively impacted by disruption of pollinator service.

3.4.2 Reallocation to flower and fruit production under pollen limitation

Either cool temperatures during the growing season, high summer precipitation, or a combination of the two increase overall fruit production. Germination of *C. pulchella* is inhibited under warm temperatures (Lewis, 1955), so plants in sites with cooler fall temperatures could have earlier germination timing and develop larger root systems, giving them access to more resources during the flowering season. *Clarkia pulchella* individuals appear to be capable of reallocating some resources to flower production when pollen is limited (Table 3.3). Our finding of a modest amount of reallocation under pollinator exclusion contrasts with work in another *Clarkia* species, *C. xantiana* ssp. *parviflora*, which found that individuals do not reallocate resources based on the quantity of pollen received (Briscoe Runquist and Moeller, 2013). These contrasting results can potentially be explained by two factors. First, the focal species of our study produces buds continuously over the flowering season, while *C. xantiana* ssp. *parviflora* produces nearly all of its buds at the beginning of the flowering season, leaving individuals little opportunity to respond to the pollination environment (Briscoe Runquist and Moeller, 2013). Second, their study investigated differences between plants under natural pollination conditions and plants receiving supplemental pollen, while we compared plants under natural pollination and plants under strong pollen limitation. These differences in direction and magnitude of the treatments imposed may affect the degree to which a plant reallocation response can be detected. An alternative explanation for the apparent resource reallocation is that our pollinator exclusion tents protected plants from herbivores that might have removed fruits in the control plots. While herbivory of individual fruits (rather than entire plants) appears rare (M. Bontrager, personal observation), we can not rule out the possibility of a herbivore effect. Finally, it is also possible that the pollinator exclusion tents reduced heat or drought stress by increasing moisture retention or shading the plots, and this could have allowed plants to produce more fruits.

3.4.3 Implications for responses to climate change

If we assume that the correlations we found between traits and climate across sites can be generally extrapolated to future climates and future responses, our results would suggest that the projected temperature increases in coming decades (Figure 3.2) will have negative effects on reproduction via negative effects on fruit set (Figure 3.6B). However, it is important to be cautious about inferring future responses from current spatial patterns (Warren et al., 2014). Common garden experiments in the field and growth chamber (Chapter 5; Gamble et al., 2018) indicate that populations of *C. pulchella* are differentiated based on climate of origin, therefore population responses to changes in climate are likely to be individualized and will depend not only on a population's current climate optimum, but also its capacity for adaptive and plastic responses.

3.4.4 Conclusions and future directions

Populations of *Clarkia pulchella* from across the species' range are reliant on pollinator service to maintain high levels of seed production, which is likely an important demographic transition for this species. Our data support the hypothesis that populations in areas of the range that have undergone post-glacial expansion may have elevated levels of reproductive assurance, but alternative drivers of this pattern remain plausible. Future work should explore these drivers, and could begin by examining geographic variation in the phenology, abundance, and composition of pollinator communities, as well as the responses of these communities to changes in climatic conditions. In order to better understand how *C. pulchella* might respond to changes in pollinator service, future work should measure the capacity of populations to evolve higher rates of self-pollination in the absence of pollinators.

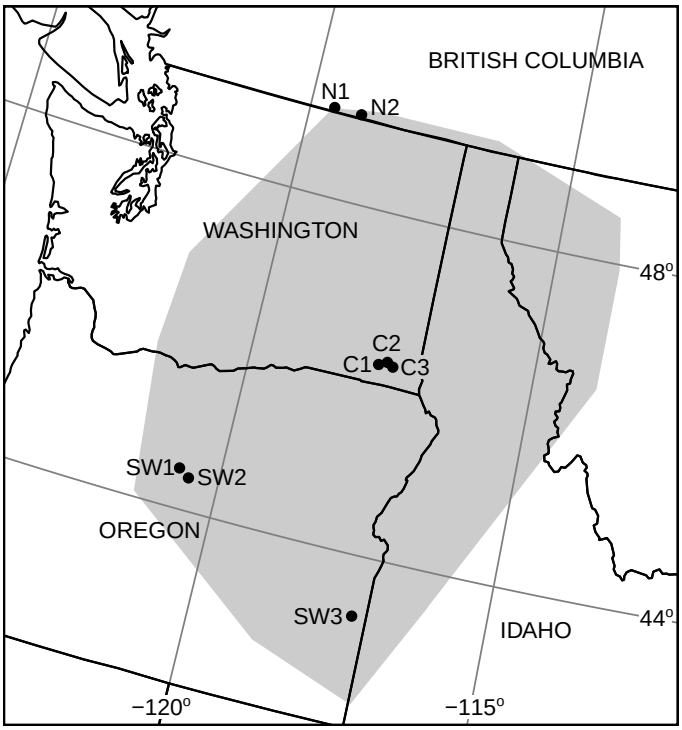

Figure 3.1: Experimental sites relative to the geographic range of *Clarkia pulchella* (shaded area). N1 and N2 are northern sites; S1, S2, and S3 are southwestern sites, and C1, C2, and C3 are central sites. For geographic coordinates and elevations, see Table 3.1.

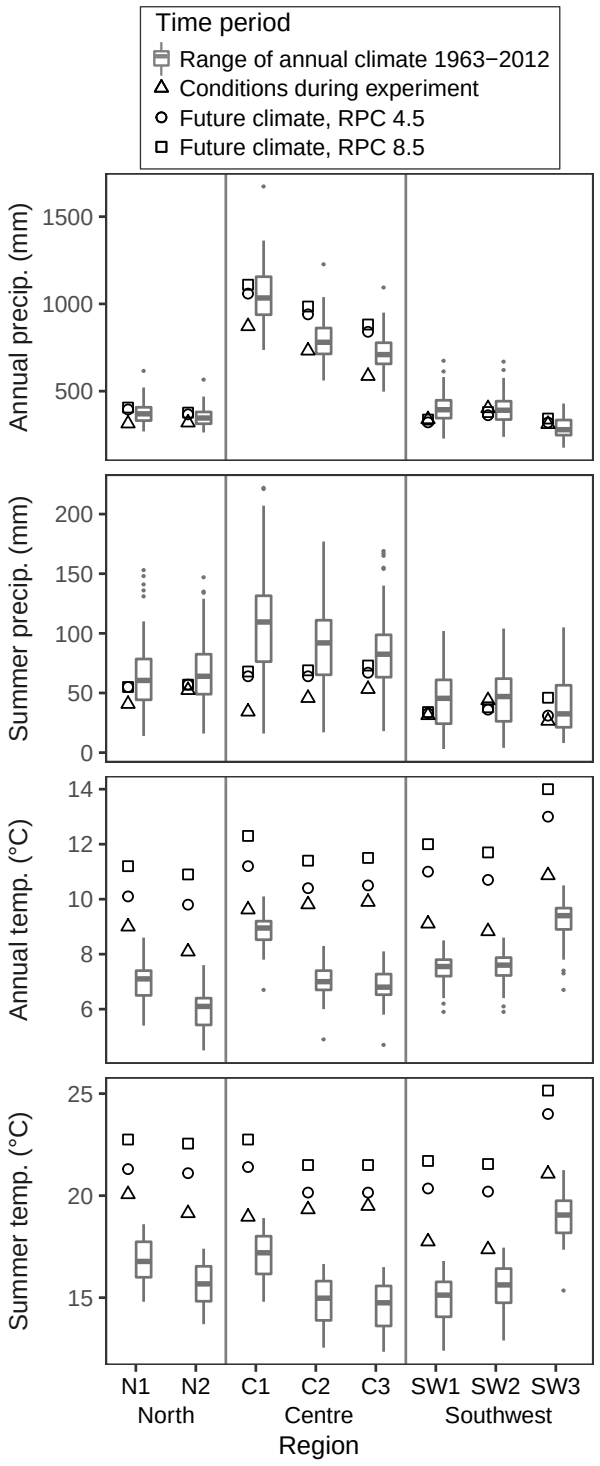

Figure 3.2: Climate conditions in experimental sites in each region. Boxplots summarize annual values over a 50-year time window (1963-2012). Triangles represent conditions during the experiment. Also shown are climate projections for 2055 under two different emissions scenarios (circles: CanESM RCP 4.5; squares: CanESM RCP 8.5). Historic and future values extracted from ClimateWNA (Wang et al., 2012), weather during the experiment was downloaded from PRISM (PRISM Climate Group, Oregon State University, prism.oregonstate.edu).

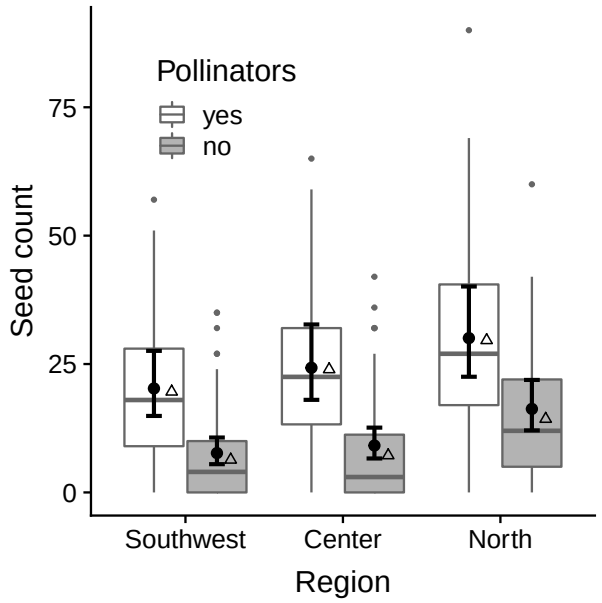

Figure 3.3: Seeds per fruit in plots with and without pollinators in each of three geographic regions within the range of *Clarkia pulchella*. Boxplots show the median, first and third quartiles, and range of the raw data; black points and error bars show the model-fitted means and 95% confidence intervals; open triangles are raw means of the data.

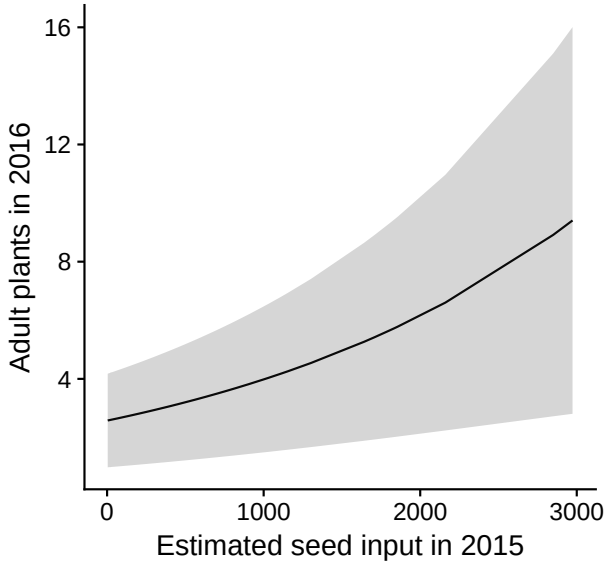

Figure 3.4: Model-fitted relationship and 95% confidence interval of the effect of plot-level seed input in 2015 on the number of adult plants present in 2016.

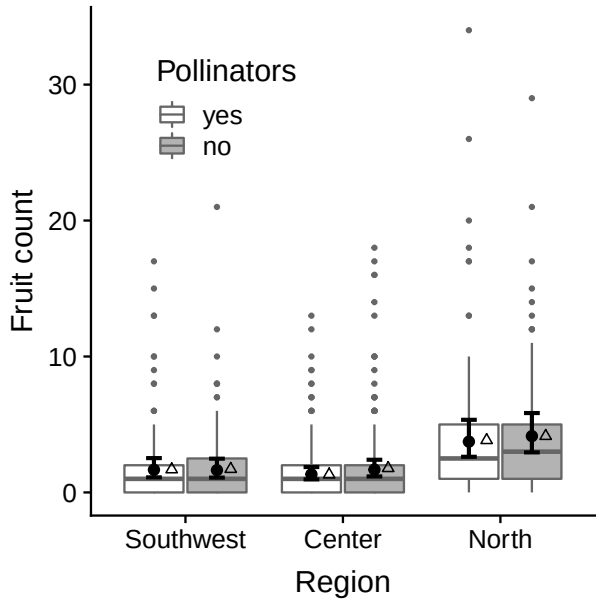

Figure 3.5: Fruits per plant in plots with and without pollinators in each of three geographic regions within the range of *Clarkia pulchella*. Boxplots show the median, first and third quartiles, and range of the raw data; black points and error bars show the model-fitted means and 95% confidence intervals; open triangles are raw means of the data.

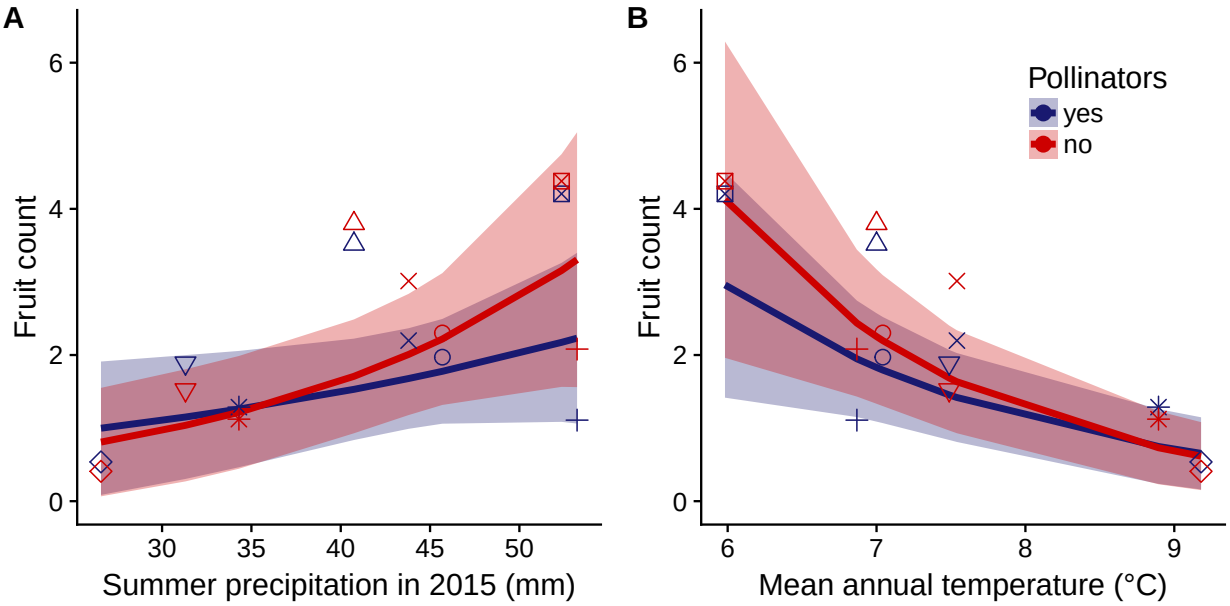

Figure 3.6: **(A)** Effects of summer precipitation during the experiment (2015) and pollinator exclusion on per-plant fruit production. **(B)** Effects of mean annual temperature (1963-2012) and pollinator exclusion on per-plant fruit production. Average per-plant fruit counts in plots with and without pollinators are also plotted. Each site is represented by a different shape.

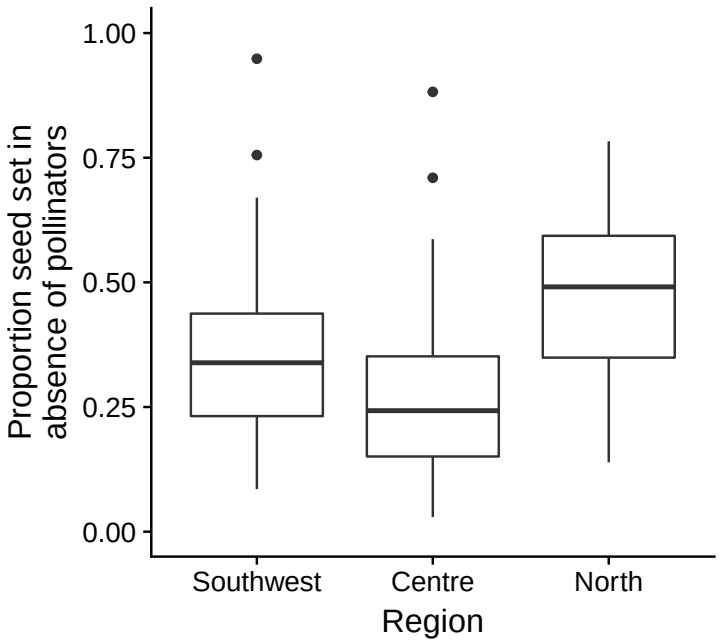

Figure 3.7: An alternative visualization of reproductive assurance in each of three regions within the range of *C. pulchella*. Rather than comparing total seeds per fruit in plots with and without pollinators, here we represent the average seed set in pollinator exclusion plots in each block as a proportion of the average number of seeds set in control plots in the same block.

Table 3.1: Geographic data for experimental sites. Coordinates are given in decimal degrees. Cross-reference ID refers to the identifying codes used in Chapter 4 and Chapter 5.

Name	Abbreviation	Cross-reference ID	Latitude	Longitude	Elevation (m)
Southwest 1	SW1	P15	44.47	-120.71	1128
Southwest 2	SW2	P16	44.38	-120.52	1134
Southwest 3	SW3	P17	43.30	-117.27	1043
Centre 1	C1	D12	46.24	-117.74	1022
Centre 2	C2	D11	46.28	-117.60	1457
Centre 3	C3	P14	46.24	-117.49	1445
North 1	N1	F1	49.05	-119.56	842
North 2	N2	F2	49.04	-119.05	866

Table 3.2: Effects of pollinator exclusion, region, and climate on seed set per fruit. Estimates, standard errors, and P-values are from zero-inflated negative binomial GLMMs. Effects of being in the northern or southwestern region are expressed relative to central populations. Significant main effects and interactions are indicated with bold font.

Climate/region predictor		Climate/region			Pollinator exclusion			Climate/region x pollinator exclusion		
		β	SE	P-value	β	SE	P-value	β	SE	P-value
Region	North	0.219	0.155	0.157	**-0.987**	**0.112**	**< 0.001**	**0.371**	**0.159**	**0.020**
	Southwest	-0.217	0.139	0.119				0.098	0.153	0.523
Mean annual precipitation		0.046	0.098	0.635	**-0.834**	**0.066**	**< 0.001**	-0.111	0.064	0.086
Mean annual temperature		-0.084	0.090	0.348	**-0.827**	**0.066**	**< 0.001**	-0.038	0.062	0.541
Summer precipitation (2015)		0.031	0.096	0.747	**-0.825**	**0.066**	**< 0.001**	-0.019	0.063	0.763
Summer temperature (2015)		0.037	0.093	0.688	**-0.840**	**0.067**	**< 0.001**	0.105	0.065	0.105

Table 3.3: Effects of pollinator exclusion, region, and climate on fruit number. Estimates, standard errors, and P-values are from negative binomial GLMMs. Effects of being in the northern or southwestern region are expressed relative to central populations. Significant main effects and interactions are indicated with bold font.

Climate/region predictor		Climate/region			Pollinator exclusion			Climate/region x pollinator exclusion		
		β	SE	P-value	β	SE	P-value	β	SE	P-value
Region	North	**1.156**	**0.442**	**0.009**	**0.302**	**0.096**	**0.002**	-0.272	0.146	0.063
	Southwest	0.001	0.399	0.997				-0.153	0.151	0.309
Mean annual precipitation		-0.126	0.236	0.594	**0.178**	**0.062**	**0.004**	0.020	0.064	0.752
Mean annual temperature		**-0.454**	**0.148**	**0.002**	**0.154**	**0.063**	**0.014**	-0.118	0.066	0.072
Summer precipitation (2015)		0.281	0.206	0.172	**0.148**	**0.062**	**0.017**	**0.211**	**0.068**	**0.002**
Summer temperature (2015)		-0.253	0.201	0.209	**0.174**	**0.062**	**0.005**	-0.033	0.066	0.617

Chapter 4

Genetic differentiation is determined by geographic distance in *Clarkia pulchella*

4.1 Introduction

Geographic distance is often a primary predictor of genetic differentiation among populations on the landscape. Populations that are near each other are often more genetically similar, while distant populations are often more divergent. This pattern arises when the dispersal distances of individuals and gametes are small relative to the distances separating populations; as a result, differences accumulate among populations due to drift faster than they are homogenized by gene flow (Slatkin, 1993; Wright, 1943). Isolation by distance is well-documented and prevalent (Sexton et al., 2014) to the extent that it is a reasonable null expectation for how genetic differentiation is structured at geographic scales.

However, geographic distance is not the only factor that structures dispersal and realized gene flow among populations (McRae, 2006; Epps et al., 2005). Not all geographic distances are equivalent in the extent to which they might facilitate or impede gene flow (Storfer et al., 2007). Landscape features between populations may impose barriers to gene flow beyond those predicted by geographic distance. Gaps in suitable habitat may be large enough that very few instances of gene flow occur across them, leading to differentiation of the populations on either side. For example, Reeves and Richards (2014) found genetic differentiation between populations of *Helianthus pumilus* that could be attributed to an unsuitable mountainous area interrupting the species' distribution. Other features of the landscape might act as corridors for the organisms themselves or for agents of gene flow (i.e., seed dispersers or pollinators). For example, wind and water flow along rivers may increase gene flow among populations situated along them (Lee et al., 2018). In these types of scenarios we expect to see deviations from a strict pattern of isolation by distance, and population genetic structure will be better described by membership in discrete groups on either side

of a barrier in the former case, or by patterns of admixture or increased similarity in populations connected by corridors in the latter.

Environmental differences between occupied sites may also contribute to the magnitude of genetic differentiation between populations (Slatkin, 1973; Wang and Bradburd, 2014). If populations are strongly locally adapted, then migrants that have moved between environments may be unable to survive to reproduction or may have low reproductive success (Nosil et al., 2005). In this case, realized gene flow may be low between different environments (Mosca et al., 2012). Similarly, vectors of gene flow such as pollinators and seed dispersers (or the organisms themselves, in the case of motile species) may have environmental preferences that lead to greater rates of gene flow among similar environments (Bolnick et al., 2009).

The current genetic structure of populations is also strongly influenced by past processes (Hewitt, 2004). In temperate regions including the Pacific Northwest, higher latitudes were glaciated until approximately 20,000 years ago (Booth et al., 2003) and this affected the distribution of many species, leaving lasting signatures on their genetic structure (Brunsfeld et al., 2001; Shafer et al., 2010). Species that previously had disjunct distributions—for example, those that occupied multiple refugia during glaciation—may exhibit multiple corresponding genetic clusters in the present day (Beatty and Provan, 2011; Carstens et al., 2013; Sproul et al., 2015). Populations that are the result of range expansions into previously glaciated areas may have lower levels of genetic diversity as a result of repeated founder events (Kuchta and Tan, 2005; Hewitt, 2004). These patterns may underlie (and sometimes confound) genetic structure that could also be attributed to isolation by distance or environment.

Despite the accumulation of numerous case studies, it is still challenging to draw generalizations about the extent to which the genetic structure of a given species is likely to be determined by geographic vs. environmental differences. A recent meta-analysis (Sexton et al., 2014) examined how the frequency of isolation by distance vs. by environment varied across broad taxonomic groups, and found that plants more frequently showed patterns of isolation by distance than vertebrates or invertebrates. However, in more than half of the plant species that displayed a pattern of isolation by distance, environmental similarity also contributed to genetic structure. In a small number of plant species, only environmental differences explained genetic structure. Although geography and environment may both have important effects on patterns of genetic differentiation, generalizations about when one will prevail over the other and what organismal traits determine their relative effect sizes remain elusive. The accumulation of more case studies and the development and use of more appropriate statistical methods will likely move this field forward (Wang and Bradburd, 2014; Bradburd et al., 2013).

The way that the landscape shapes genetic structure is of particular interest in the context of geographic range limits. Local adaptation may be constrained in range edge populations if these populations are inundated with gene flow from populations in dissimilar environments (Kirkpatrick and Barton, 1997). If populations are isolated by environmental differences, that might prevent swamping gene flow. Rather, gene flow between populations in similar environments could facilitate

local adaptation by increasing adaptive genetic diversity (Sexton et al., 2011). This might be of particular importance if species occupy spatially heterogeneous environments, where random dispersal would otherwise result in frequent gene flow between divergent environments.

In this study, we use RADseq data to investigate whether environmental differences between populations of the annual wildflower *Clarkia pulchella* contribute to their genetic differentiation, which we expected to also be strongly structured by geographic distances. Among the populations in our study, geographic distances are not highly correlated with environmental differences, allowing us to decouple these drivers. Further, we explored whether patterns of genetic differentiation are better described by admixture among distinct genetic groups or continuous genetic differentiation across the landscape. We expected that topographic features, such as the Rocky Mountains, might be an impediment to the movement of seed dispersers and pollinators, and that this might result in disjunct genetic groups. Finally, we explored whether genetic diversity varies geographically in this species. We predicted lower levels of genetic diversity at high latitudes if this species has undergone a range expansion northward after the last glacial maximum. These analyses will also inform our interpretation of the results of a field transplant experiment, in which we simulated gene flow using a subset of populations and evaluated performance in common gardens at the northern range edge (Chapter 5). We were interested in knowing the extent to which these populations are genetically differentiated and whether differentiation depended upon environmental differences between them.

4.2 Methods

4.2.1 Study species

Clarkia pulchella Pursh (Onagraceae) is a winter annual wildflower that grows east of the Cascade Mountains in the Pacific Northwest. It can be found in eastern Washington, eastern Oregon, Idaho, and western Montana (United States) and in southeastern British Columbia (Canada; Figure 4.1). It grows in large populations (i.e., thousands of flowering individuals) on open, south-facing slopes from 100 to 2200 meters elevation, though the majority of populations are found between 500 and 1600 m. While temperature generally decreases and precipitation generally increases from south to north and west to east across the range of *C. pulchella*, temperature and precipitation are also strongly influenced by elevation. Topographic complexity across the range creates large amounts of variation around geographic trends and appears to disrupt spatial autocorrelation in climate among populations of *C. pulchella* (Figure 4.2). This species has small seeds (c. 1 mm long) that lack an obvious dispersal mechanism. Flowers are visited by a diverse array of pollinators, including solitary bees, bee flies, bumblebees, and occasionally hummingbirds (M. Bontrager, personal observation).

4.2.2 Population selection, climate characterization, and seed collection

For this study, we selected populations that would allow us to decouple climatic and spatial axes of differentiation. For example, we wanted to include populations that were spatially near each other

but climatically different and populations that were geographically distant but climatically similar. Monthly temperature and precipitation data from 1951-1980 for all populations were obtained from PRISM (PRISM Climate Group, 2017). We calculated the average temperature across the months that encompass the *C. pulchella* life cycle (September-July) and average precipitation when *C. pulchella* is most likely to be water-limited (April-July) for each population. Based on field observations and common garden trials (Chapter 5), we considered these to be good candidates for variables that might have the potential to generate patterns of isolation by environment via selection against migrants. We first considered a set of 40 populations that we had located, then narrowed that set down to 32 populations that maximized variation in the relationship between spatial proximity and climatic similarity (Figure 4.1, Table 4.1). In July of 2014, we collected seeds from 12 plants separated by at least 0.5 m in each of those populations. Seeds from 17 populations were grown in the greenhouse beginning in December of 2014, and seeds from the remaining 15 populations were grown in growth chambers beginning in February of 2016.

4.2.3 DNA Extraction

Tissue was harvested from the first cohort of plants in May 2015. Leaf or bud tissue was collected into 2 mL tubes on ice, then frozen at -80°C until DNA extraction. Tissue from the second cohort was collected onto dry ice in April 2016 and stored at -80°C until DNA extraction. DNA was extracted using DNeasy Plant Mini kits and DNeasy Plant 96 kits (Qiagen), following the protocol for frozen tissues. DNA extractions that did not have satisfactory 260/230 or 260/280 ratios were cleaned with ethanol precipitation. DNA was eluted and stored in 10mM Tris-HCl pH 8.

4.2.4 Library preparation and sequencing

Libraries were prepared using 100 ng starting material. We prepared for two lanes of sequencing, with six individually barcoded samples from each population in each lane (191 or 192 individuals per lane, because we only had DNA of a high enough quality from a total 11 individuals from one population). Our library preparation protocol was based on Poland et al. (2012) with modification by M. Todesco, K. Ostevik, and B. Moyers (Rieseberg Lab, University of British Columbia). DNA was digested in a 20 µL reaction using 8 units each of the enzymes MspI and Pst I-HF (New England Biolabs) in the supplied buffer. Digestion was carried out for 5 hours at 37°C, followed by 20 minutes at 65°C. Reactions were then stored overnight at 4°C. Ligation was performed in a 40 µL reaction in the same buffer as the digestion with 200 units of T4 DNA ligase (New England Biolabs) using 192 barcoded adapters and 12 common adapters on the opposite end. Ligation was performed for 3 hours at 22°C followed by a 20 minute hold at 65°C. Reactions were then cleaned with 1.6 volumes of SPRI beads and two 80% ethanol washes and resuspended in 12 µL of Tris-HCl pH 8.

Amplification was carried out in 10 µL reactions using 4 µL of cleaned ligation product, Kapa HIFI HotStart master mix (Kapa Biosystems), and primers from Poland et al. (2012). Amplification began at 98°C (30 s), followed by 14 cycles of 98°C (30 s), 62°C (20 s), 72°C (30 s), and a 72°C hold

for 5 minutes. After amplification, samples were quantified using fluorometry, then each plate was pooled according to individual concentrations to yield a final product with equal amounts of library from each individual. This pooled library was run out on a 1.5% agarose gel and bands containing fragments 400 to 600 bp long were excised and cleaned using a gel extraction kit (Qiagen). The eluted product was cleaned and concentrated using SPRI beads.

Finally, we reduced the number of high copy fragments from our library using a protocol modified by M. Todesco from Shagina et al. (2010) and Matvienko et al. (2013). We began with 480 ng of each library in a 3 µL volume. To this we added 1 µL of hybridization buffer (200 mM HEPES pH 7.5, 2M NaCl, 0.8 mM EDTA), covered the reaction with mineral oil, heated it to 98°C for 2 minutes, then held it at 78°C for 3 hours. We then added 5 µL of duplex specific nuclease buffer (0.1 M Tris pH 8, 10mM $MgCl_2$, 2mM DTT) and incubated at 70°C for 5 minutes. We then added 0.2 µL of duplex specific nuclease and incubated at 70°C for another 15 minutes, then stopped the reaction with 10 µL of 10 mM EDTA. We then reamplified the library using the same reagents as above in a 25 µL reaction with 2-4 µL of template and cleaned again with SPRI beads. Libraries were stored at -20°C until sequencing. Libraries were sequenced with paired-end 100 bp reads on the Illumina HiSeq 2000 platform at the Biodiversity Research Centre at UBC.

4.2.5 Alignment and SNP calling

Sequences were processed and aligned using components of the Stacks pipeline (version 1.40, Catchen et al., 2011, 2013). Reads with uncalled bases or low quality scores (average quality in a 14-base sliding window <10) were discarded. After cleaning and demultiplexing, ten samples had far fewer reads than the rest (<300k reads) and these were excluded. All other samples had between 507k and 3.2 million reads (mean read number = 1.5 million). Paired end reads were pooled with first end reads, i.e., during alignment and SNP (single nucleotide polymorphism) detection the two ends of each read were treated as if they were independent loci (we later checked for linkage disequilibrium among SNPs). During initial "stacking" and catalog building we allowed sequences to diverge at 3 bases, and set the minimum depth of coverage required to create a stack at 3 (Rochette and Catchen, 2017). Modifications to these parameters did not result in substantial differences in values of pairwise F_{ST} (data not shown). The maximum number of stacks per locus was set to 3, and gapped alignments were not allowed. We enabled the removal algorithm, which drops highly repetitive stacks (removes initial stacks that have >2 SD coverage relative to individual sample mean), and the deleveraging algorithm, which breaks up or removes over-merged sequences. Our catalog was built using all samples. We employed the rxstacks corrections module to correct or omit loci with putative sequencing errors, loci with low log-likelihoods (<-10), confounded loci, and loci with excess haplotypes.

SNP tables were generated using the populations module of Stacks. Initial inspection of PCA plots using SNPRelate (Zheng et al., 2012) revealed three individuals that were not clustering with the other individuals from their populations. We consider it more plausible that these represent mis-labeled samples in the field, greenhouse, or lab than long-distance migration events. Down-

stream analyses were performed without these individuals. Therefore, in our final dataset, seven populations had only 11 individuals, one population had only 10, one population had only 8, and the remaining 23 populations were each represented by 12 individuals. In our analyses we included only loci that had coverage of at least 12x in 75% of individuals in 75% of populations, with a minimum minor allele frequency of 0.05 and a maximum heterozygosity of 70% across all populations. We checked that pairwise F_{ST} was not sensitive to these parameter choices. In case of multiple SNPs occurring in a single locus, we kept just the first one. After applying these filters, 2982 SNPs were retained. Linkage disequilibrium was generally low among our loci (r^2 <0.2 for 26639 pairs, $0.2 < r^2$ <0.55 for the remaining 22 pairs of SNPs). F_{ST} was calculated using the implementation of Weir and Cockerham (1984) and expected heterozygosity (within-population gene diversity) was calculated using methods from Nei (1987) in the R package hierfstat (Goudet and Jombart, 2015). Because populations varied in the average proportion of loci that were successfully genotyped (three populations had <60% success; among all populations the median success rate was 78% and the range was 23-92%), we checked that expected heterozygosity did not correlate with genotyping success rate ($r = 0.27$, $P = 0.13$).

4.2.6 Quantifying isolation by environment vs. isolation by distance

We used BEDASSLE (Bradburd et al., 2013) to estimate the relative contributions of geographic distance and climatic differences to genetic differentiation. BEDASSLE is implemented in R (R Core Team, 2017), and it employs a Markov chain Monte Carlo (MCMC) algorithm to estimate the relative effect sizes of geographic distance and environmental differences on covariance in allele frequencies among populations. As environmental covariates, we used pairwise differences in average September-July temperature and average spring/summer precipitation (April-July). We initially generated resistance-weighted distances between populations using projected habitat suitability (Chapter 2) as a conductance matrix, but these distances were highly correlated with actual geographic distances and did not produce better model fits in preliminary analyses, so we did not use them in these models. We estimated effect sizes of geography, temperature, and precipitation differences using all 32 populations, but also ran BEDASSLE for subsets consisting of populations clustered in the central and northern parts of the range (indicated in Table 4.1) to see if we could detect effects of the environment that may be obscured or weakened at large geographic scales. Prior to analysis, we divided pairwise geographic distance and precipitation differences by their standard deviations so that these predictors were on a scale more similar to pairwise temperature differences. We ran these models for 10 million generations, and thinned the chains by sampling every 1000 generations. We visually inspected MCMC traces and marginal distributions to ensure that models reached stationary distributions. All results are reported after a burn-in of 20%, with effect sizes back-transformed to the scale of the original data. We checked these results against partial Mantel tests of pairwise geographic, temperature, and precipitation differences on pairwise F_{ST} using the R package phytools (Revell, 2012). We did not rely upon partial Mantel tests as our main analytical method because of their potential to have inflated Type I error rates (Guillot and

Rousset, 2013).

4.2.7 Assessment of spatially continuous vs. discrete genetic differentiation

We were interested in evaluating whether population structure was well-described by modelling populations as admixtures between multiple discrete genetic groups, as might be caused by geographic barriers (e.g., the Rocky Mountains) or historic phylogeographic processes. We evaluated how well models prescribing various numbers of discrete genetic groups described differentiation and similarity among our populations using conStruct (Bradburd et al., 2017). conStruct is implemented in R (R Core Team, 2017), and is similar to the frequently-used program Structure (Pritchard et al., 2000) but allows genetic differentiation to increase with geographic distance between populations even when these populations draw from the same genetic groups. In the spatial implementation of this program, populations are composed of admixture from a user-specified number of discrete layers (K), and genetic similarity decays with geographic distance within each of these layers. We ran conStruct for 1000 iterations setting the number of layers to 1, 2, 3, 4, and 5. We compared the fits of each of these different parameterizations using cross-validation and by evaluating the contribution of each additional layer to the total covariance of these loci. For cross-validation, we fit models with subsets containing 90% of loci and evaluated the resulting model fit by calculating the log likelihood of the remaining loci. We performed 100 replicate cross-validation runs.

4.2.8 Exploring spatial patterns in genetic diversity

We examined whether population genetic diversity (as estimated by expected heterozygosity) exhibited geographic trends. We used linear models in R (R Core Team, 2017) to test whether expected heterozygosity was predicted by latitude or by proximity to the range edge (as measured by the distance of a population to the nearest edge of a polygon drawn around all localities of the species; Figure 4.1).

4.3 Results

4.3.1 Isolation by environment vs. geographic distance

Overall F_{ST} among these populations is 0.135; the distribution of per-locus F_{ST} is presented in Figure A.1. Genetic differentiation between populations of *Clarkia pulchella* is primarily structured by geographic distance, with no apparent contribution of the environmental variables that we have considered here (Figure 4.3). The effect size of a temperature difference of one degree (C) relative to the effect of 100 km of geographic distance is 1.18×10^{-7} (95% credible interval = 8.52×10^{-8} - 1.58×10^{-7}; Figure 4.4A), and the effect of 10 mm of spring/summer precipitation difference relative to the effect of 100 km of geographic distance is 5.84×10^{-7} (95% credible interval = 1.50×10^{-8} - 2.98×10^{-6}; Figure 4.4B). The scales at which these ratios are presented are arbitrary, but they were chosen so that the range of values among populations is on the same order of magnitude:

100 km represents about one sixth of the maximum pairwise geographic distance, 1°C represents approximately one fourth of the maximum pairwise temperature difference, and 10 mm precipitation represents about one fourth of the maximum pairwise precipitation difference. The climatic effect sizes we found are so small that the effects of these variables can be considered nonexistent in terms of their biological importance; they are non-zero due to the priors for these effects being unsupported below zero. Effects of environmental differences did not emerge at smaller geographic scales in subsets of populations in the north (Figure A.2AB; Figure A.3) or centre (Figure A.2CD; Figure A.4). These conclusions are consistent with the results of partial Mantel tests, in which only pairwise geographic distance is a significant predictor of pairwise F_{ST} (Table 4.2).

4.3.2 Genetic structure of populations

The genetic structure of these populations is explained slightly better by a model of admixture between two genetic groups than by a model of continuous genetic differentiation across space, as indicated by the increase in predictive accuracy in models where two layers were allowed rather than one (Figure 4.5). Northern populations primarily belong to one genetic group, while southern populations belong to another, and populations from mid-latitudes are a mix of the two (Figure 4.6). Allowing more than two layers did not improve predictive accuracy (Figure 4.5). Note that populations east of the Rocky Mountains (populations D9, D10, and P12) never formed a separate group, regardless of the number of layers allowed (results not shown). Although models with two layers did have greater predictive accuracy than those with one, when $K = 2$ the amount of covariance contributed by the second layer was small relative to the first (Table 4.3).

4.3.3 Geographic trends in genetic diversity

Genetic diversity increases with latitude among these populations (estimate = 0.0104, SE = 0.0019, df = 30, $P < 0.0001$, Figure 4.7A), but is not related to distance from the range edge (df = 30, $P = 0.811$). Genetic diversity appears to be lower in populations in the southern half of the range, and also in populations near the eastern range edge, but is higher in central and northern populations (Figure 4.7B).

4.4 Discussion

We contrasted the relative effects of geographic vs. climatic distances on genetic differentiation in *Clarkia pulchella*, examined whether geographic structure in this species could be described by assigning populations to distinct genetic groups, and tested for geographic gradients in genetic diversity. Our analyses revealed a genetic structure that is predominantly shaped by geographic distances between populations. In addition to this pattern of isolation by distance, populations partition into northern and southern groups, with admixed populations in the centre of the range. Genetic diversity was highest in northern and central populations, resulting in a trend of increasing genetic diversity with latitude.

4.4.1 Populations of *Clarkia pulchella* are isolated by distance

At the scale of the geographic range in *Clarkia pulchella*, isolation by distance is the dominant pattern. This likely reflects gene flow that is strongly restricted by geographic distances between populations. This is perhaps not surprising, given that this species has no obvious mechanism for seed dispersal and our best guess is that gene flow between populations is facilitated by occasional pollen movement by bumblebees, hummingbirds, and other floral visitors. In the case of an absence of climatically structured seed and pollen movement, selection against migrants and their offspring is the remaining mechanism that could drive isolation by environment. While *C. pulchella* does appear to be locally adapted to historic climate (Chapter 5), selection against foreign genotypes may not be strong enough to preempt the spread of neutral loci, even as recently-arrived loci that confer poor performance in a given environment are purged. This could lead to a signal of isolation by distance at neutral loci, while populations are still adaptively differentiated based on their local climate.

It is possible that the absence of an effect of temperature and precipitation differences on genetic structure is the result of our experimental design, and that environmental differences might matter in other contexts. There may be environmental variables other than those we have considered here that are more important in determining the movement of genes or the realized rate of gene flow among populations. These could be climatic, but also could include soil characteristics, or local adaptation to competitors, pollinators, or soil biota. It is also possible that the effects of environmental differences are more detectable at smaller spatial scales. For example, in some plant species, differences in phenological timing along local snowmelt gradients structure gene flow to a greater extent than geographic distances (Hirao and Kudo, 2004; Shimono et al., 2009). Similar processes may play out in *C. pulchella* as well, possibly along local elevation gradients.

4.4.2 Populations are admixtures of northern and southern genetic groups

Rather than mountain ranges separating populations into genetic groups, we detected underlying population structure that divides the species into northern and southern groups, with admixed populations in the middle. This suggests that perhaps the Columbia Basin, a low-elevation, relatively flat area in south-central Washington (Figure 4.1), is a barrier to gene flow in this species. Species distribution models indicate that it is an area of low suitability (Chapter 2) and few occurrences of *Clarkia pulchella* have been recorded in this region. Most studies of population genetic structure in the Pacific Northwest focus on mesic forest species that occupy the wet western slopes of both the coastal and Rocky Mountains (Shafer et al., 2010), and these studies often find differentiation between western and eastern populations. Phylogeographic research on species occupying the arid inter-mountain region is less common. In the Great Basin pocketmouse, a species with a range that overlaps with that of *C. pulchella*, a north-south split in genetic structure was detected in approximately the same location as in our results (Riddle et al., 2014). It is possible that the Columbia Basin (or some geographic feature within it) represents a barrier to gene flow, either past or ongoing, for a variety of taxa that occupy the dry intermountain region. The habitat affinity

of species can influence the effect of glaciation events on genetic structure (Massatti and Knowles, 2014), therefore further work on *C. pulchella*, including paleoclimate modelling or modelling demographic history, might allow for an interesting contrast with the relatively well-studied mesic flora of the Pacific Northwest.

4.4.3 Genetic diversity increases with latitude

We expected we would see lower genetic diversity at higher latitudes, but we detected the opposite: genetic diversity was highest in north-central and northern populations (though the total magnitude of variation in expected heterozygosity was not large). This latitudinal pattern is somewhat surprising, because northern populations are in areas that were under glaciers during the last glacial maximum, and we expected that range expansion into this area after their retreat would result in a signature of lower genetic diversity. When high levels of genetic diversity are present in areas of past range expansion, this can sometimes be attributed to the mixing of populations that had previously been persisting in multiple refugia (Petit et al., 2003; Brunsfeld and Sullivan, 2005). Species in the northern Rocky Mountains that are presumed to have occupied multiple refugia often exhibit some degree of contemporary differentiation between northern and southern populations (Brunsfeld et al., 2001; Brunsfeld and Sullivan, 2005), a pattern consistent with what we have found in *Clarkia pulchella*. Regardless of the location or number of refugia that *C. pulchella* previously occupied, it is also possible that range expansion was not accompanied by reductions in genetic diversity in this species, as is sometimes the case in other systems (Vandepitte et al., 2017). A further possible explanation for the observed patterns in genetic diversity is that variation in genetic diversity could be driven by demographic expansions upslope, rather than northward. Our southern populations tended to be from higher elevations than our northern populations (Table 4.1), so this could result in apparent regional variation.

The more common expectation for geographic patterns in genetic diversity is that range edge populations will have lower genetic diversity (Vucetich and Waite, 2003). This prediction is based on the assumption of an abundant centre distribution pattern, in which edge populations are small, and may experience frequent turnover or constant directional selection (if they are far from the phenotypic optima of an extreme environment). Our results are not consistent with this being the case for *C. pulchella*, at least not at all range edges. We note however that populations at southern and eastern edges do appear to have lower genetic diversity relative to the northern and north-central populations, and further work could be done to investigate the processes that might generate this pattern.

4.4.4 Conclusions

Our investigation of the genetic structure of *Clarkia pulchella* has revealed some intuitive patterns, as well as surprising ones. Despite substantial heterogeneity in climate across the species' range, genetic similarity is primarily determined by geographic proximity. Though a signal of isolation by distance is not surprising in a sessile organism studied at a large spatial scale, the absence of any

effect of environment indicates that to the extent that populations experience gene flow, it may be from both similar and divergent environments. This species does not exhibit geographic patterns of genetic diversity consistent with our expectations for a recently expanded northern range edge nor a range limited by adaptation. These results would be complemented by future work examining mechanisms of contemporary gene flow and historic demographic processes in *Clarkia pulchella*.

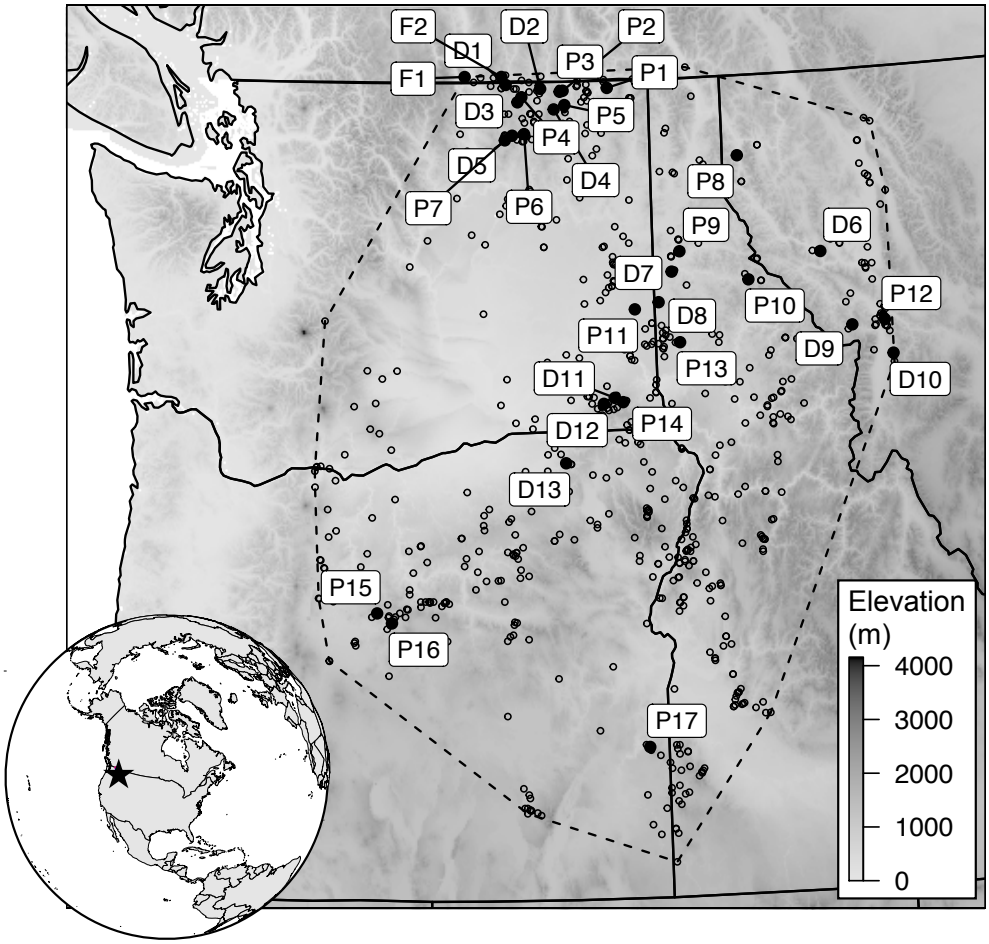

Figure 4.1: The geographic range of *Clarkia pulchella* across the interior of the Pacific Northwest. Small open points mark the locations of all herbarium records of *C. pulchella* from the Consortium of Pacific Northwest Herbaria that could be accurately assigned coordinates. The dashed line marks the maximum convex polygon drawn around these points. Larger filled points are populations that were sampled for this project. Labels correspond to population IDs in Table 4.1 and are consistent with Chapter 5. Background shading shows elevation. The Columbia Basin is the unsampled area west of population D11.

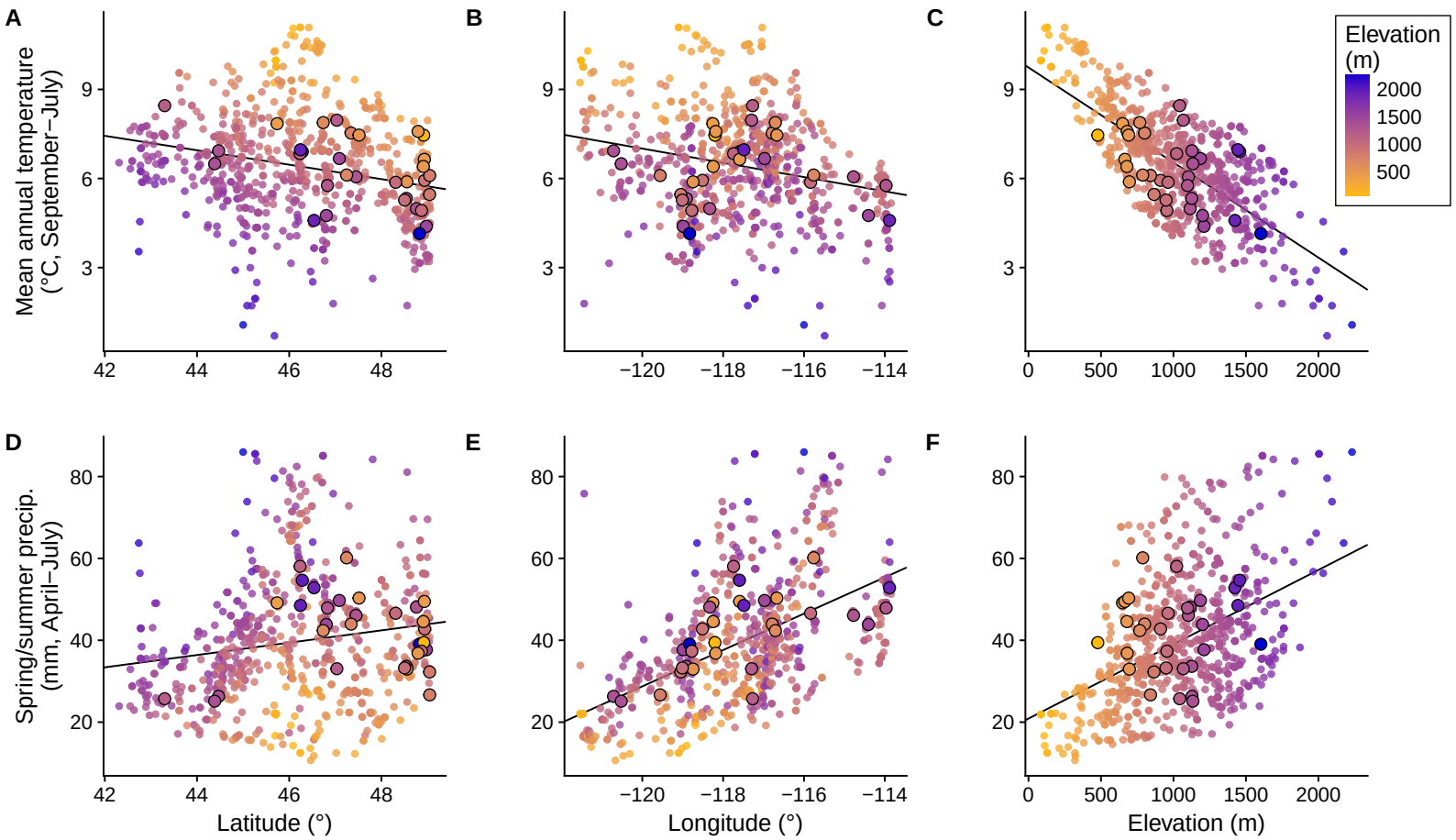

Figure 4.2: Relationships of climate and geography across the range of *Clarkia pulchella*. Small points represent all herbarium localities of *C. pulchella*, larger outlined points represent populations included in this study. Points are coloured according to elevation. Temperature is influenced by **(A)** latitude, **(B)** longitude, and **(C)** elevation. Precipitation is also influenced by **(D)** latitude, **(E)** longitude, and **(F)** elevation. However, the interaction of these drivers results in climate that is heterogeneous across space. Climate data are 1951-1980 averages from PRISM (PRISM Climate Group, 2017). Trend lines are slopes from linear regression.

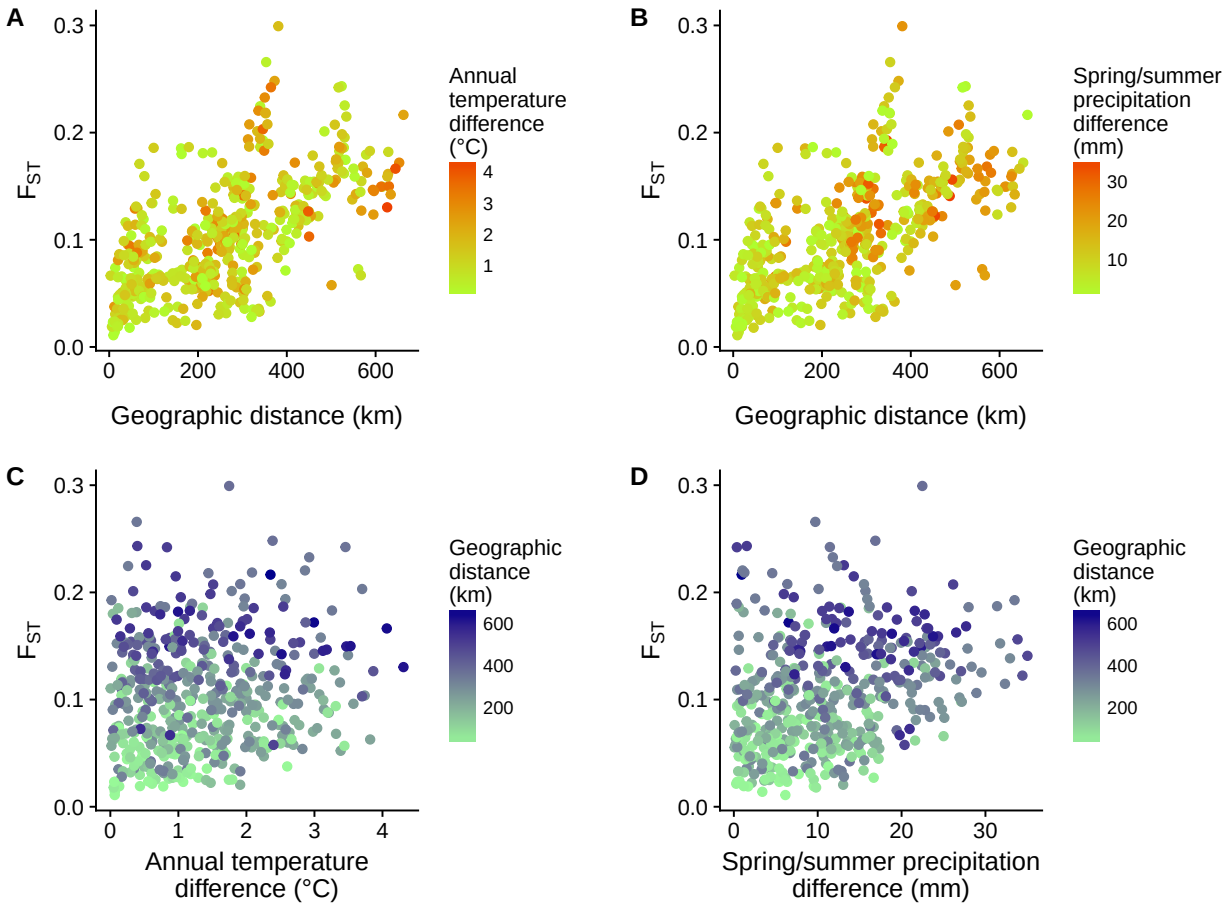

Figure 4.3: Pairwise genetic differentiation (F_{ST}) of populations of *Clarkia pulchella* increases with geographic distance (x-axis in **A** and **B**), but shows no discernible relationship to temperature differences (colour in **A**) or precipitation differences (colour in **B**). An alternative visualization is presented in **(C)** and **(D)**, in which climate differences are plotted on the x-axis and geographic distance is indicated with colour. Climate data are 1951-1980 averages from PRISM (PRISM Climate Group, 2017). Statistical tests are presented in Section 4.3.1 and Table 4.2.

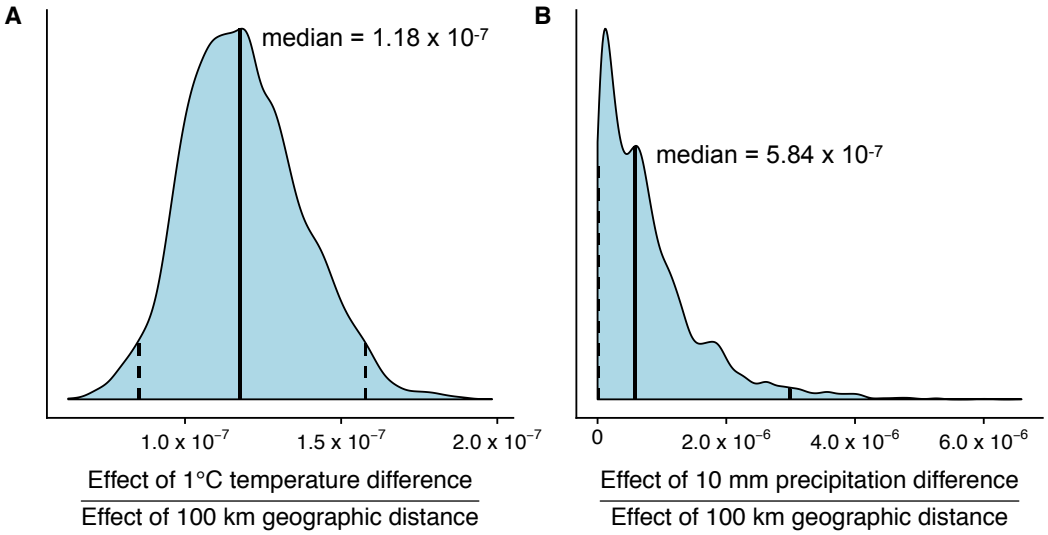

Figure 4.4: Marginal posterior distributions, median values (solid lines) and 95% credible intervals (dashed lines) of the ratio of the effect sizes of **(A)** temperature vs. geographic distance and **(B)** spring/summer precipitation vs. geographic distance on genetic differentiation of populations of *Clarkia pulchella* after a burn-in of 20%.

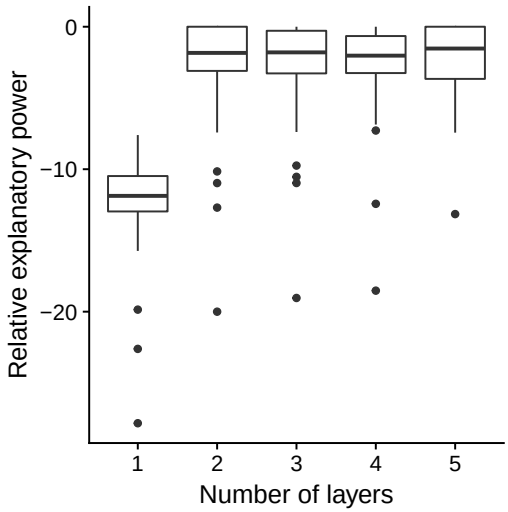

Figure 4.5: Results of 100 replicate cross-validation runs of conStruct with the number of layers set to 1, 2, 3, 4, or 5. In each replicate, the model is built using 90% of loci, and the log-likelihood of the remaining loci is calculated. Predictive accuracy is then calculated as the difference in log-likelihood between each model and the best model (i.e., the best number of layers) in each replicate. These results indicate that models constructed with two layers are best, because they provide as much explanatory power as other models without further complexity.

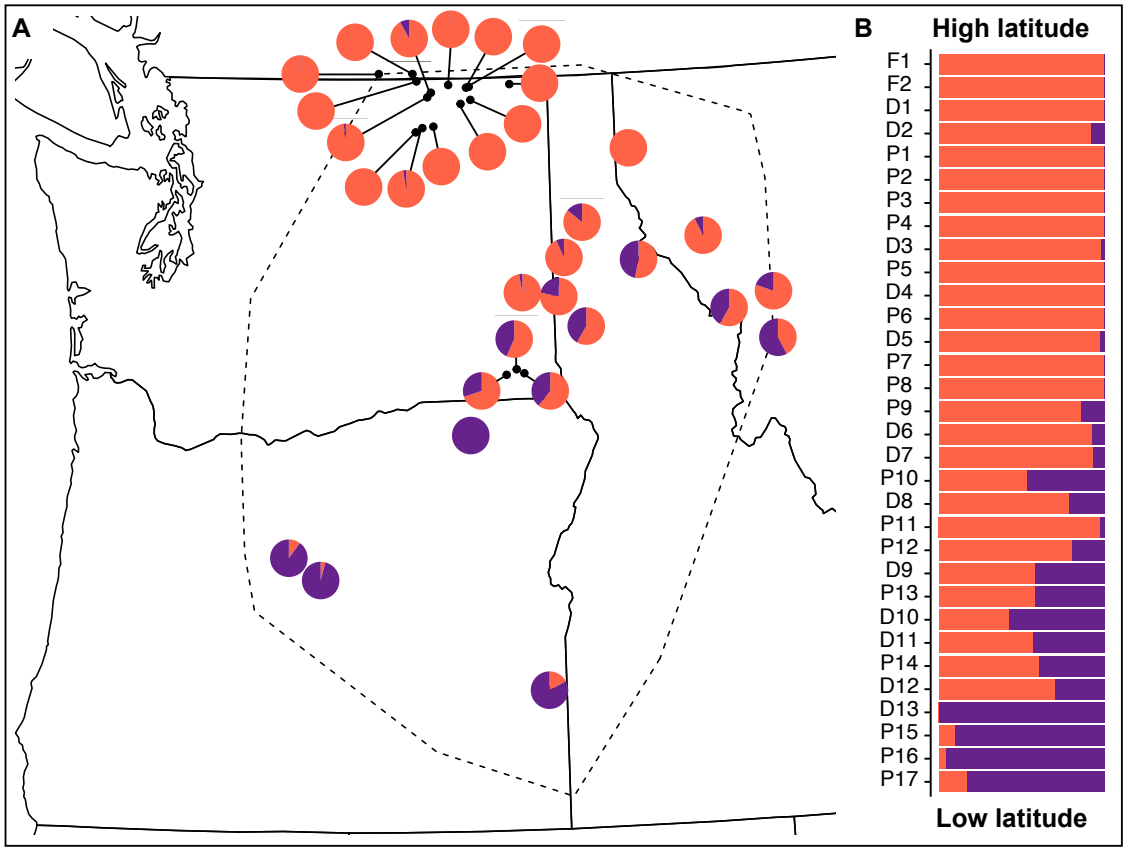

Figure 4.6: Admixture proportions of each of 32 populations of *Clarkia pulchella* estimated from by conStruct with K = 2. **A** Admixture proportions are shown in geographic space and **(B)** arranged by latitude . Population ID codes are consistent with Table 4.1 and Figure 4.1.

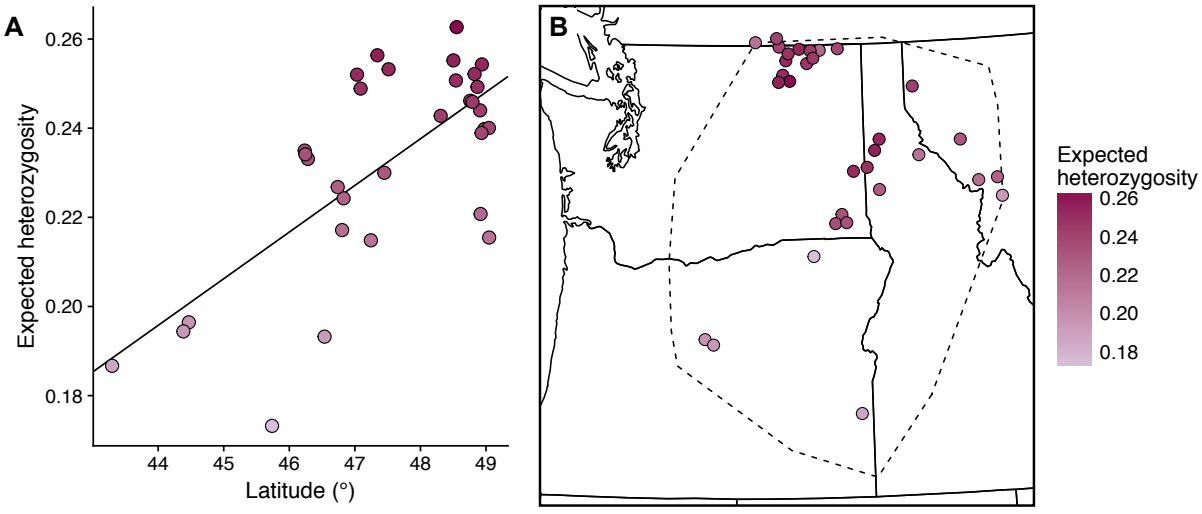

Figure 4.7: **(A)** Expected heterozygosity increases with latitude across the range of *Clarkia pulchella*. **(B)** Expected heterozygosity appears to be higher in central and northern parts of the range, but lower in the south and east.

Table 4.1: Geographic locations and elevations of populations of *Clarkia pulchella* included in these analyses. Population IDs are consistent with Figure 4.1, Chapter 3, and Chapter 5. The populations included in analyses of geographic subsets are indicated.

Population ID	Geographic subset	Latitude	Longitude	Elevation (m)
F1	North	49.05	-119.56	842
F2	North	49.04	-119.05	866
D1	North	48.98	-118.99	1211
D2	North	48.94	-118.51	911
P1	North	48.93	-117.59	665
P2	North	48.92	-118.20	478
P3	North	48.91	-118.25	679
P4	North	48.87	-118.77	955
D3	North	48.83	-118.83	1603
P5	North	48.79	-118.18	681
D4	North	48.76	-118.33	1115
P6	North	48.55	-118.74	696
D5	North	48.54	-118.91	1126
P7	North	48.50	-119.01	949
P8	-	48.31	-115.84	963
P9	Centre	47.51	-116.67	691
D6	-	47.45	-114.77	1103
D7	Centre	47.34	-116.79	801
P10	-	47.24	-115.76	788
D8	Centre	47.09	-116.98	1186
P11	Centre	47.03	-117.30	1068
P12	-	46.83	-113.97	1097
D9	-	46.80	-114.41	1201
P13	Centre	46.74	-116.71	768
D10	-	46.54	-113.89	1424
D11	Centre	46.28	-117.60	1457
P14	Centre	46.24	-117.49	1445
D12	Centre	46.24	-117.74	1022
D13	Centre	45.74	-118.25	649
P15	-	44.47	-120.71	1128
P16	-	44.38	-120.52	1134
P17	-	43.30	-117.27	1043

Table 4.2: Results of partial Mantel tests of pairwise geographic distance (km), pairwise temperature differences (°C, September-July, 1951-1980 averages), and pairwise precipitation differences (mm, April-July, 1951-1980 averages) on pairwise genetic differentiation (F_{ST}) among populations of *Clarkia pulchella*. Climate data are 1951-1980 averages from PRISM (PRISM Climate Group, 2017).

Region	R^2	*P*-value	Predictor	Coefficient	t-statistic	*P*-value
Entire range	0.42	**0.001**	Geographic distance	0.0002	15.73	**0.001**
			Temperature differences	0.0028	1.44	0.486
			Precipitation differences	0.0006	2.46	0.209
North	0.36	**0.008**	Geographic distance	0.0006	4.65	**0.006**
			Temperature differences	0.0061	1.63	0.377
			Precipitation differences	0.0004	0.57	0.692
Centre	0.44	0.06	Geographic distance	0.0006	4.87	**0.001**
			Temperature differences	0.0111	0.56	0.463
			Precipitation differences	-0.0005	-0.45	0.737

Table 4.3: Covariance contributions of each layer in conStruct models with the number of layers (K) set to 1, 2, 3, 4, or 5.

Number of layers	1	2	3	4	5
Layer contributions	1.000	0.9004	0.8062	0.8014	0.8925
	-	0.0996	0.1043	0.1541	0.0795
	-	-	0.0895	0.0438	0.0204
	-	-	-	0.0007	0.0055
	-	-	-	-	0.0021

Chapter 5

Gene flow disrupts local adaptation but improves performance at the northern range edge of *Clarkia pulchella*

5.1 Introduction

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. There are a number of processes that could generate a pattern of reduced genetic variation at range edges. If range edge populations are small (either because of maladaptation, or low carrying capacity at range edges) or if they experience frequent or severe fluctuations in population size, genetic variation may be lost to drift (Vucetich and Waite, 2003). Similarly, populations at equilibrial range margins may have lower genetic variation if they experience frequent founder events due to higher rates of extinction and colonization (Lande, 1992; Holt and Keitt, 2000). Populations that are on the leading edge of range expansions may exhibit similar 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 (though 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 core 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 centre of that gradient. Because of this asymmetry in abundance, net gene flow is asymmetric from the centre towards the range edge 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 centre to edge as well as the steepness of the environmental gradient (i.e., the magnitude of environmental difference 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 swamping gene flow may limit adaptation along geographic gradients in some systems (Paul et al., 2011) and 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). Fitness consequences may arise as a result of gene flow between highly diverged populations with genetic incompatibilities, however, these effects may not be as important as they were once thought to be (Frankham et al., 2011). Outbreeding depression may appear similar to the effects that are predicted when local adaptation is disrupted, 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 centres 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 en-

vironmentally divergent neighbouring 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-centre 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 maladaptive homozygosity caused by drift or inbreeding. Populations at range margins are thought to have smaller population sizes and to be more isolated than central populations (Vucetich and Waite, 2003). Because of this, individuals in these populations may mate with relatives more frequently than individuals in central populations, increasing homozygosity by inbreeding. Small populations are also more likely to fix deleterious alleles through drift. In either of these scenarios, gene flow from other populations can 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, though 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 heterozygosity is positively related to fitness we would expect a greater benefit from gene flow between more genetically divergent populations.

5.2 Methods

5.2.1 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, when temperatures are cool and rains begin, 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 (including solitary bees, bee flies, and bumblebees), though it has some capacity to self-pollinate in the absence of pollinators or mates (MacSwain et al., 1973; Palladini and Maron, 2013). Individual plants typically produce fewer than 10 flowers, though some larger individuals may produce up to *c.* 100 on occasion.

Seeds of *C. pulchella* were collected from 15 populations in July of 2014 (Figure 5.1A; Table 5.1). Collection sites were located based on herbarium records from the Consortium of Pacific Northwest Herbaria (www.pnwherbaria.org) and targeted surveys. We used the two northwestern-most localities 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; Figure 5.1B) across the northern half of the species' range. In each of the populations used in the experiment, seeds were collected haphazardly from at least 22 plants spaced >0.5 m apart. Seeds were stored in paper envelopes in the lab until a greenhouse generation was planted.

5.2.2 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. Seeds were planted in the greenhouse 9-11 December, 2014 in conetainers (Stuewe and Sons, Tangent, Oregon, USA) filled with Sunshine Mix No. 4 (Sun Gro Horticulture, Agawam, MA, USA). For each of 22 maternal families per population, 3-5 seeds were

planted on the soil surface in each pot. For families from each of the two focal populations, three replicate pots were planted per family because larger quantities of flowers would be needed from these families; other populations were represented by one replicate per family. Pots were arranged into randomized blocks, with each block containing one family from each population (one pot from each donor population and three replicate pots from each of the two focal populations). The soil was kept moist until germination, then plants were hand watered every 1-3 days as needed to prevent wilting. After germination, plants were thinned randomly to one per cone and pumice was added to the soil surface to prevent fungal growth. Plants began to flower in March 2015. Plants were bagged to prevent unintentional pollination, and flowers were emasculated upon opening to prevent self-pollination. For the crosses, 20 of the 22 blocks were used, the other two were maintained in the same growing conditions to provide alternate plants in case of mortality or sterility.

Two types of crosses were performed: "within-population" crosses and "between-population" crosses (Figure A.5). For within-population crosses, dams were pollinated using pollen from the plant of the same population in the subsequent block in a "daisy-chain" design. Each plant from each population 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. These crosses simulate one 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 performed as many crosses as possible using a single focal plant, but if flower production was too low on that plant we also used one of the replicate focal plants from the same family. Most crosses had to be performed 2-3 times to obtain adequate numbers of seeds for the experiment. Some crosses could not be performed due to mortality, sterility, or limited flower production. As ripening progressed, the ends of fruits were taped shut to prevent seed loss. Upon ripening, fruits were collected and stored in coin envelopes in the lab. Crosses were performed March-May 2015 and we collected fruits March-June 2015.

5.2.3 Common garden design and installation

For the transplant, we used 15 families of each cross type from each population. In other words, 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 Rock Creek plots only contained between-population crosses performed on Rock 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 Rock Creek and 836 toothpicks at Blue Lake. In total, our design included 16,680 toothpicks and 32,755 seeds.

Seeds attached to toothpicks were planted in the ground 18-21 September 2015. Plots were prepared by removing litter, large rocks, and dried remains of herbaceous perennial plants. The ground surface was minimally levelled to allow for placement of planting grids that aided in consistently spacing the plants. Each toothpick was inserted into the ground gently so that seeds were not dislodged or damaged until seeds were ∼3 mm below the soil surface. Toothpicks were inserted at 5 cm spacing into ∼1 m by 2 m blocks. Block shape was varied to accommodate rocks and shrubs surrounding the planting area. After planting, each block was protected with 20 cm high hardware cloth cages supported by rebar. These cages were intended to prevent trampling by larger animals but did not prevent entry of rodents and other small animals. The area surrounding the plots at each site was sprayed with deer repellent several times during the course of the experiment. To ensure germination, plots were watered at a rate of ∼10 L per plot 27-29 October 2015, though at that time most seeds that were checked already had radicles emerging. In May 2016 cattle fencing was put around the plots at the Blue Lake site before cattle were released into the area for grazing; this fencing succeeded in keeping the cows off the plots. No cattle were present at the Rock Creek site.

5.2.4 Monitoring and measuring

Germination was censused 16-20 November 2016. We documented the emergence of either 0, 1, or 2 germinants at each toothpick. If two germinants were present, these were randomly thinned so that just one remained. The size of the remaining seedling was measured to the nearest millimetre as the distance from one cotyledon tip to the other. At 23 out of 16,680 grid points (0.14%), we censused one more germinant than the number of seeds that we planted. This gives an estimate of the minimum rate at which naturally occurring seedlings were indistinguishable from our planted seedlings. So, while it is probable that some naturally occurring plants were mistaken for experimental plants, we consider the frequency of possible misidentification to be acceptably low.

Overwinter survival was assessed 17-21 March 2016. At this time, seedlings typically had just one pair of leaves, so size was measured as the length from the tip of one leaf to the tip of the other, to the nearest millimetre. Some plots were affected by frost-heave and seedlings were uprooted from their planting locations when their toothpick was forced out of the ground (1901/16680 grid points, 11.4%). In lightly affected areas, toothpicks and seedlings were gently settled back into the soil. In more heavily affected areas, individual identity could no longer be determined confidently and individuals were excluded from further measurements and analyses (95/16680 grid points, 0.57%).

On 12-13 June 2016 we censused survival of all plants. Censuses of reproduction began on 2 June 2016. Once flowering began, we placed bridal-veil nets over the hardware cloth on each plot to prevent pollen escape into local populations. In June we censused each plot every 2-3 days. We

recorded the date of first flowering of each plant at this temporal resolution. During each census, the immature ovary length of each new flower was recorded to be used as a proxy for maximum seed set. Flowers were marked as they were measured with a permanent marker and a running flower count was kept for each plant to avoid double-counting as flowers senesced. We continued these assessments as flowering slowed in July, but reduced the census interval to once a week. Damage to plants, such as rodent activity or herbivory, was noted during monitoring. Any plants with uncertain identities (due to frost damage as mentioned above, being far from their toothpick, or the toothpick disappearing; n = 201/16680 toothpicks, 1.2%) were excluded from all analyses. Plants that were killed by gophers, browsers, or galling insects were excluded from analyses that involved lifestages downstream from these events (n = 525/16479 plants, 3.2%) because we do not think that this mortality is related to population origin but rather to block-specific factors.

Pollinations were performed on a subset of plants to calibrate a conversion from immature ovary lengths to seed production. On 596 flowers (mean = 29.8 per plot, range = 0–126) stigmas were dusted with an ample pollen load using all four anthers from another plant in the plot. These flowers were marked with strings around the pedicles and fruits were collected when ripe. Seeds in each of these pollinated fruits were later counted in the lab. Total seed production per individual was estimated by multiplying the total ovary length of each plant by the average number of seeds per millimetre of immature ovary, as determined from the pollinated fruits. This resulted in an estimate of 4.75 seeds per mm based on a linear regression of number of seeds predicted by ovary length with the intercept set to 0 ($R_0^2 = 0.87$). We pollinated only a maximum of one flower per plant, so these may be overestimates because they do not account for potential resource limitation of seed set. However, we checked whether variation in seeds per mm of fruit was associated with individual fitness (the overall fruit production per individual) or block quality (estimated based on the average fruit production of a block), and we could not attribute variation in seeds per mm of fruit to either of these factors. Therefore, while our conversion from fruit length to seeds may not be exact, we do not expect it to be systematically biased.

5.2.5 Climate data

We compiled monthly temperature and precipitation data from 1951-1980 for all seed sources, as well as the gardens during the months of the experiment (September 2015-July 2016) from PRISM (PRISM Climate Group, 2017). We calculated historic (pre-warming, 1951-1980) climate averages for each site, which we compared to conditions experienced by plants during the transplant in our analyses. 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.

5.2.6 Population genetic data

Pairwise population differentiation (F_{ST}) was calculated from 2982 SNPs that were genotyped in up to 12 parental individuals from each population (tissue was collected during the controlled crossing phase in the greenhouse). F_{ST} was calculated using the implementation of Weir and

Cockerham (1984) in the R package hierfstat (Goudet and Jombart, 2015). See Chapter 4 for methods describing the construction of libraries and generation of SNP tables.

5.2.7 Statistical analyses

Did local populations outperform foreign populations?

First, to investigate whether the focal populations were locally adapted to conditions during the experiment, we used only fitness components from the within-population crosses. A comprehensive assessment of local adaptation requires a fully reciprocal transplant design so that fitness trade-offs can be identified between populations and environments (Hereford, 2009; Kawecki and Ebert, 2004). The presence of local adaptation is indicated by both local populations outperforming foreign populations and populations performing best at home when compared to other environments. Our design only allows us to infer local adaptation based on the former: local population performance (i.e., the performance of each of the two focal populations in their respective home sites) relative to the performance of foreign populations.

We tested whether local populations were, on average, superior to foreign populations by comparing lifetime fitness of local vs. 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 once zero-inflation has been accounted for). 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. Seed production differs from fruit count because seed production takes into account the size of fruits as well as the number. Plant size was modelled with a Gaussian response distribution. Germination and survival were modelled using binomial response distributions and logit link functions. Fruit counts and seed production were modelled using zero-inflated negative binomial response distributions and log link functions. We used negative binomial distributions because they are appropriate for overdispersed count data. 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 sire population and block within site.

Does climate of origin explain performance in common gardens?

We built GLMMs using the methods described above to evaluate the effects of climatic parameters on lifetime fitness and all fitness components, using only within-population crosses. If populations have adapted to their historic climatic conditions, we expected provenances with historic climates that most closely matched the experimental conditions to perform best in our common gardens. To test this, 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 (hotter or cooler, wetter or drier) will negatively impact fitness. However, 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 only included spring and summer precipitation differences because these are the seasons when we expect precipitation to be limiting, and variation among sources in these drier seasons might be obscured by larger amounts of precipitation in winter. We isolated the lifestages affected by each of these climatic predictors using the lifestage-specific analytical methods described for the tests of local adaptation. In these analyses, we used size during the previous census as a covariate to account for earlier lifestages (as in our tests for local advantage), and calculated climate differences using only the months between each census and the previous census (or the months between planting and the first census, for germination and size after germination). In all analyses, all continuous predictors were scaled and centred (by subtracting the mean and dividing by the standard deviation) and checked for co-linearity. Correlations between predictors in each model were low ($r < |0.5|$), except in one case (discussed in Section 5.3.3).

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. We wanted to evaluate whether the climatic drivers that were important for determining performance of within-population crosses also held true in between-population crosses and to evaluate whether there were benefits of gene flow that were independent of the effects of climate of origin. To this end, we calculated the midparent historic temperature average for all individuals (that is, the average temperature of dam and sire sites) and then calculated the absolute difference between this temperature and the experimental temperature. For within-population plants, this midparent temperature difference is equivalent to the temperature difference described in the previous section, that is, the absolute difference between the garden conditions and the historic temperature of each source population. For between-population plants this calculation resulted in a narrower range of temperature differences, because all plants had one parent from one

of the focal populations (Figure A.6). We calculated a metric of absolute midparent precipitation difference (averaged over April-July) in the same manner. We used GLMMs as described in the previous sections 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 discussion in Section 5.3.2), 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 we are using a different subset of plants than in previous analyses (between-population crosses only). We built zero-inflated GLMMs as described above using lifetime fitness from all between-population individuals 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).

5.3 Results

5.3.1 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, indicating that local populations were not well-adapted to conditions during the experiment (Figure 5.2, Table 5.2).

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 (Figure 5.3A). This occurred via effects on both the probability of producing any seeds (the zero-inflation part of the model; $\beta = -0.337$, SE $= 0.035$, $P < 0.001$; Figure 5.3B), and the number of seeds produced (the conditional part of the model; $\beta = -0.114$, SE $= 0.050$, $P = 0.022$; Figure 5.3C). Note that all parameter estimates are reported and plotted untransformed, that is, on the link scale. Local populations, which are historically intermediate in temperature (Figure 5.1B), were mismatched from the experiment conditions and performed worse

than populations from warmer sites that were more climatically similar to the garden conditions.

Analyses of component lifestages support these inferences (Figure 5.3D, Table 5.3): 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 = 0.0495, $P = 0.178$, Figure 5.3C), they did have a negative effect on seed production among plants surviving the winter (Figure 5.3D, Table 5.3). The significant effect of precipitation on seed production, but not fruit production, indicates that adaptation to precipitation conditions affects the size of fruits, rather than just their number.

5.3.2 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 5.4A; Figure 5.4AB). Gene flow (i.e., being a between-population vs. a within-population cross) did not have a significant effect in the lifetime fitness model that also included both temperature and precipitation 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 (Figure A.6B). This was not an issue with temperature differences, because our focal populations are intermediate to other provenances in terms of temperature (Figure A.6A). 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 5.4B).

The potential for a small positive effect of gene flow, independent of climatic differences, is supported by analyses of some lifestage components (Figure 5.4C; Table 5.4C). 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.

5.3.3 Genetic differentiation between parental populations is positively correlated with fitness

Both midparent temperature difference from the garden conditions and genetic differentiation between parental populations had significant relationships with fitness when analyzed in separate models (Table 5.5AB). Genetic differentiation between parental populations had a positive relationship with lifetime fitness via both the probability of producing seeds (the zero-inflation part of the model) and the number of seeds made (the conditional part of the model). The effects of

genetic differentiation between parental populations on lifetime fitness are mirrored in the analyses of these effects on single lifestages: F_{ST} had a positive relationship with germination, size after winter, fruit count, and seed production (Table 5.5A).

The negative relationship between performance and 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 are well-matched to temperatures during the experiment are 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 5.5B). Midparent temperature differences did not significantly affect any single lifestage, but had marginally significant negative relationships with size after germination, size after winter, and fruit number (Table 5.5B). Midparent precipitation differences did not significantly affect lifetime fitness or component lifestages (Table 5.5C).

When both temperature differences and genetic differentiation were put into the same model (along with precipitation differences), only F_{ST} had a significant relationship with lifetime fitness (Table 5.6A). Offspring of crosses with more genetically differentiated parents were more likely to produce seeds (Figure 5.5). In full models of component lifestages, genetic differentiation between parental populations had a positive relationship with germination and seed production, and a marginally significant positive relationship with fruit number (Table 5.6B).

It is important to note that genetic differentiation is not highly correlated with temperature differences ($r = -0.25$), though genetic differentiation and precipitation differences are correlated ($r = 0.64$), with plants whose parents are more genetically differentiated also having larger differences between their historic midparent precipitation and conditions during the experiment. We think it is unlikely that the significant positive effects of genetic differentiation are actually driven by precipitation differences, because we would expect high precipitation differences to negatively affect fitness. The 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 (Figure A.6).

5.4 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 were most closely matched to conditions during the experiment performed best, even better than local populations. 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 seemed to confer some benefits independent of

climate, as evidenced by the potential positive effect of gene flow when controlling for temperature differences and the positive effect of increasing genetic differentiation between the parental populations.

5.4.1 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 (Figure 5.3). 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 the potential to disrupt local adaptation, as indicated by the somewhat negative effects of midparent temperature differences on between-population plants (Figure 5.5, Table 5.5). If we view our within-population plants as simulated dispersal from other sites, the negative effects of temperature and precipitation deviations were more pronounced. Our results highlight that gene flow and dispersal need not be from populations that are geographically distant (or from the centre of the range) to be climatically divergent from historic or current conditions. Rather, two of the coolest populations used in the experiment are from sites nearest to our common gardens (populations D1 and D3; Figure 5.1).

However, just because gene flow has the potential to limit adaptation in peripheral populations, that does not mean that gene flow occurs between populations in a manner that serves to limit ranges. A full test of the swamping gene flow hypothesis should also examine whether this type of gene flow actually occurs across populations. Landscape genetic analyses (see Chapter 4) indicate that genetic differentiation of populations of *C. pulchella* is generally moderate (overall $F_{ST} = 0.14$), and is primarily structured by geographic distances between populations, at least on the scale of our sampling. This means that in climatically heterogeneous parts of the range, populations experiencing quite different selection pressures are likely to be connected via gene flow (though adaptive differentiation will differ from patterns of differentiation at neutral markers). Future theoretical and empirical investigations of swamping gene flow at range edges may benefit from considering variation in the magnitude of gene flow and environmental heterogeneity in regional population crosses (i.e., moving windows) from the range centre to the range edge. Perhaps strong, spatially heterogeneous selection with gene flow among populations can cause maladaptation and suppress demographic performance in a manner similar to asymmetric gene flow across an environmental gradient (Kirkpatrick and Barton, 1997).

Under climate change, local adaptation to historic climate regimes may generate local maladaptation in field trials. We see this in our results, where populations from warmer locations performed best in our gardens (Figure 5.3, Table 5.3), and gene flow from warmer locations had positive effects on some lifestages (Figure 5.5, Table 5.5). This type of 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 emphasize that dispersal and gene flow may be 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 (Figure 5.1), though 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 favoured another set of populations. Both of these climate dimensions seem to exert their effects on fitness via phenology: populations from warm places began flowering earlier in our gardens, and populations from dry places kept flowering longer (results not shown). 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.

5.4.2 Gene flow confers benefits independent of climate

We saw some additional positive effects of gene flow once the effects of climate are controlled for, though statistical support for these effects is limited (Figure 5.4, Table 5.4). These positive effects may be the result of increased heterozygosity when parental plants come from two different populations; this inference is supported by the positive effect of genetic differentiation between parental populations on performance (Figure 5.5, Table 5.5). This result is 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 individ-

uals 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 (for example, 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 (results not shown) though our experiment was not especially well-suited to test this because we have only two local populations. However, if there is an additional benefit of being locally adapted along other environmental axes that we have not detected here, it could be an alternative explanation for the apparent benefit of gene flow at some lifestages. Perhaps that benefit is not due to benefits of outbreeding, but rather to having one parent from the local populations, while most of the within-population plants have two foreign parents.

5.4.3 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. While it seems clear that gene flow from warm sites is likely to accelerate adaptation to warming conditions, 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 lab (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 lab (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-axis on Figure 5.3A, Figure 5.5A), so we have no evidence that gene flow has the potential to drive populations extinct, or to turn them from demographic sources into sinks. The high lifetime fitness we observed during our experiment could be due to several factors. First, perhaps warm conditions over the entire season are favourable for all sources, but are more favourable for warm-adapted populations. Alternatively, we could have increased fitness by limiting antagonistic biotic interactions, in particular with large herbivores, which may be consequential. These interactions were recently found to be important in another *Clarkia* species (Benning et al., 2018), but note that *C. pulchella* is smaller than the focal species of that study and frequent damage from grazers has not been observed in natural populations of *C. pulchella* (M. Bontrager, personal observation). Finally, and perhaps most plausibly, our plot placement may have upwardly biased our germination and reproductive estimates. We placed plots in patches that appeared favourable to *C. pulchella*, but naturally dispersing seeds are likely to land in a mix of favourable and unfavourable patches. We do not know whether any of these

factors might interact with provenance, in which case they might change the relative performances of populations in our experiment.

5.4.4 Conclusions

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 may no longer be at equilibrium with climate. In a climate year that was more characteristic of historic conditions at our common garden sites, we expect we would have seen a signal of fitness declines caused by gene flow from both warmer and cooler sites. Even interannual variation in climate that is not explicitly attributed to warming may affect the results and inferences from common garden experiments. 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.

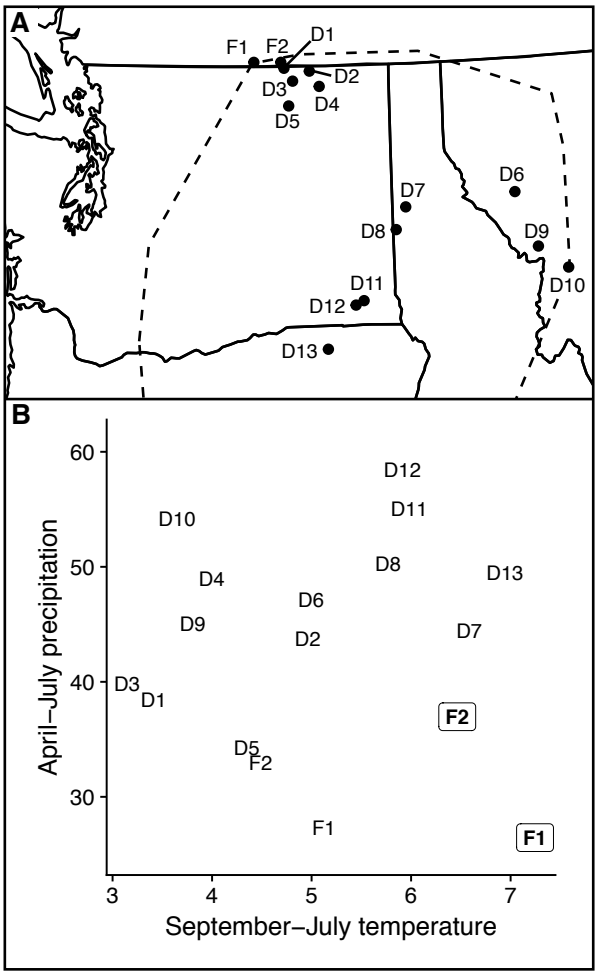

Figure 5.1: Geographic locations and climate averages of populations used in this experiment. **(A)** Populations span the northern half of the geographic range of *Clarkia pulchella* (indicated by the dashed line). Focal population sites (where common gardens were installed) are indicated by "F" and donor population sites are indicated by "D". Identifying codes for each population correspond to the map ID column in Table 5.1. **(B)** Temperature (°C) and precipitation (mm) conditions in common gardens during the experiment and averages in each population's home site. Bold labels in boxes represent weather conditions during the experiment and unboxed labels represent the 1951-1980 average in the home site of each population. Focal populations are historically intermediate relative to donor populations in average historic temperature (x-axis), but are the 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.

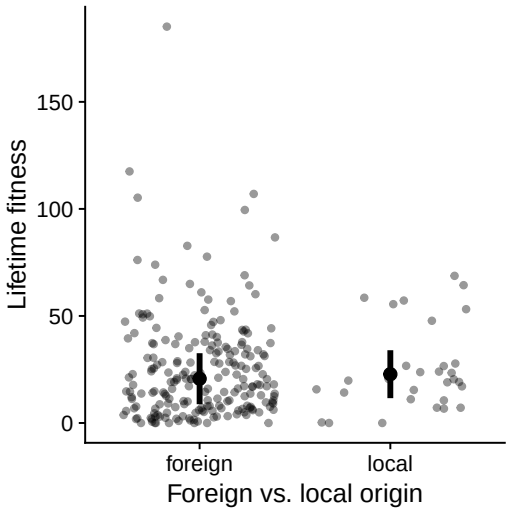

Figure 5.2: Lifetime fitness (seeds produced per seed planted) from populations of *Clarkia pulchella* with foreign vs. local parents. This analysis includes within-population plants only from the two focal populations and the 13 donor populations (no gene flow). Local populations are the focal populations in their home sites. Each point represents the average of a single family. Error bars are 95% confidence intervals of model estimated means, omitting variation from random effects.

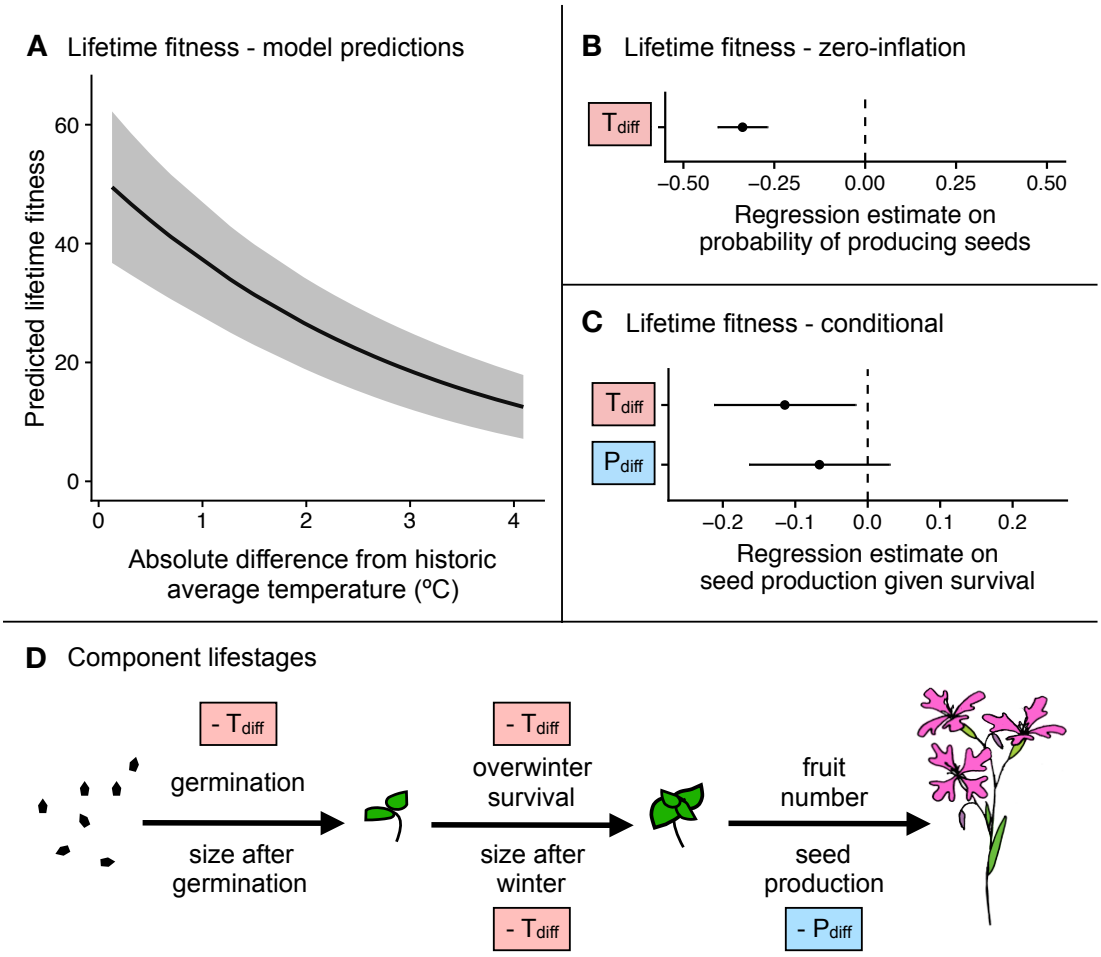

Figure 5.3: Effects of absolute temperature difference (September-July; T_{diff}) and absolute precipitation difference (April-July; P_{diff}) on performance of *Clarkia pulchella* in common gardens. These analyses include within-population plants only (no gene flow). **(A)** Lifetime fitness declines with increasing differences in temperature between the historic average of the source population and the experimental conditions in the transplant gardens. The shaded area represents the 95% confidence interval of the model estimate 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. **(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 5.3.

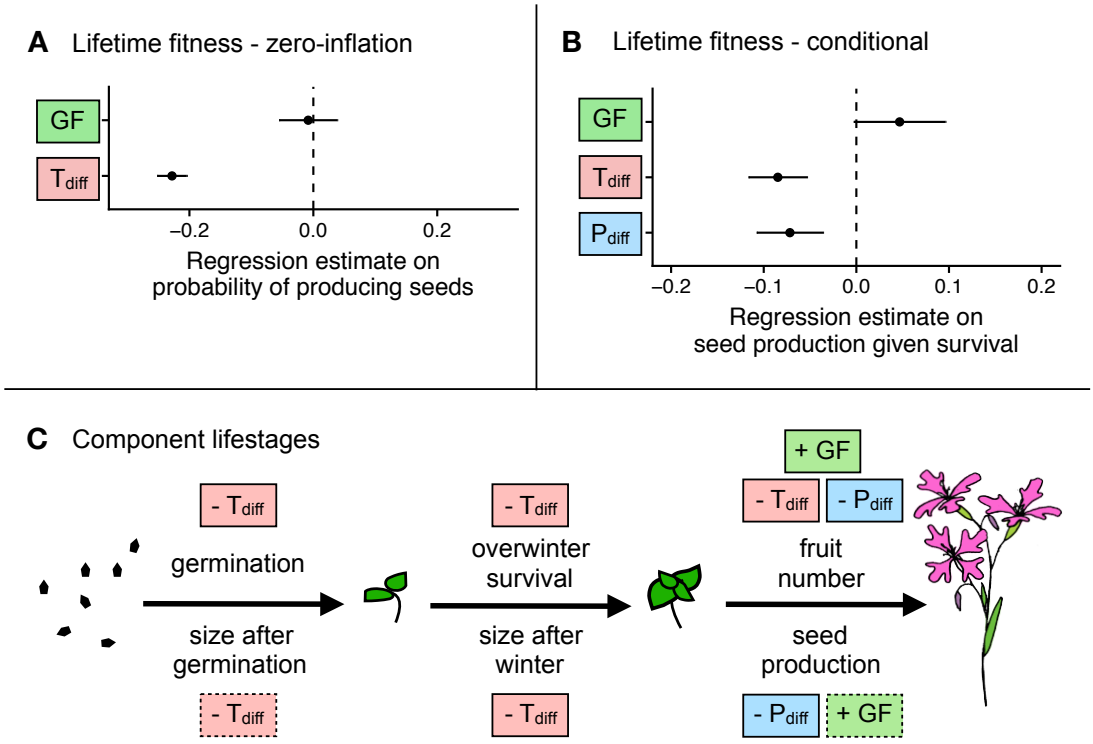

Figure 5.4: Effects of gene flow (differences between between-population and within-population crosses) on performance of *Clarkia pulchella*, accounting for midparent temperature and precipitation. **(A)** Regression estimates and standard errors from the zero-inflation part of a model of lifetime fitness. **(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 "-". Marginally 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 5.4.

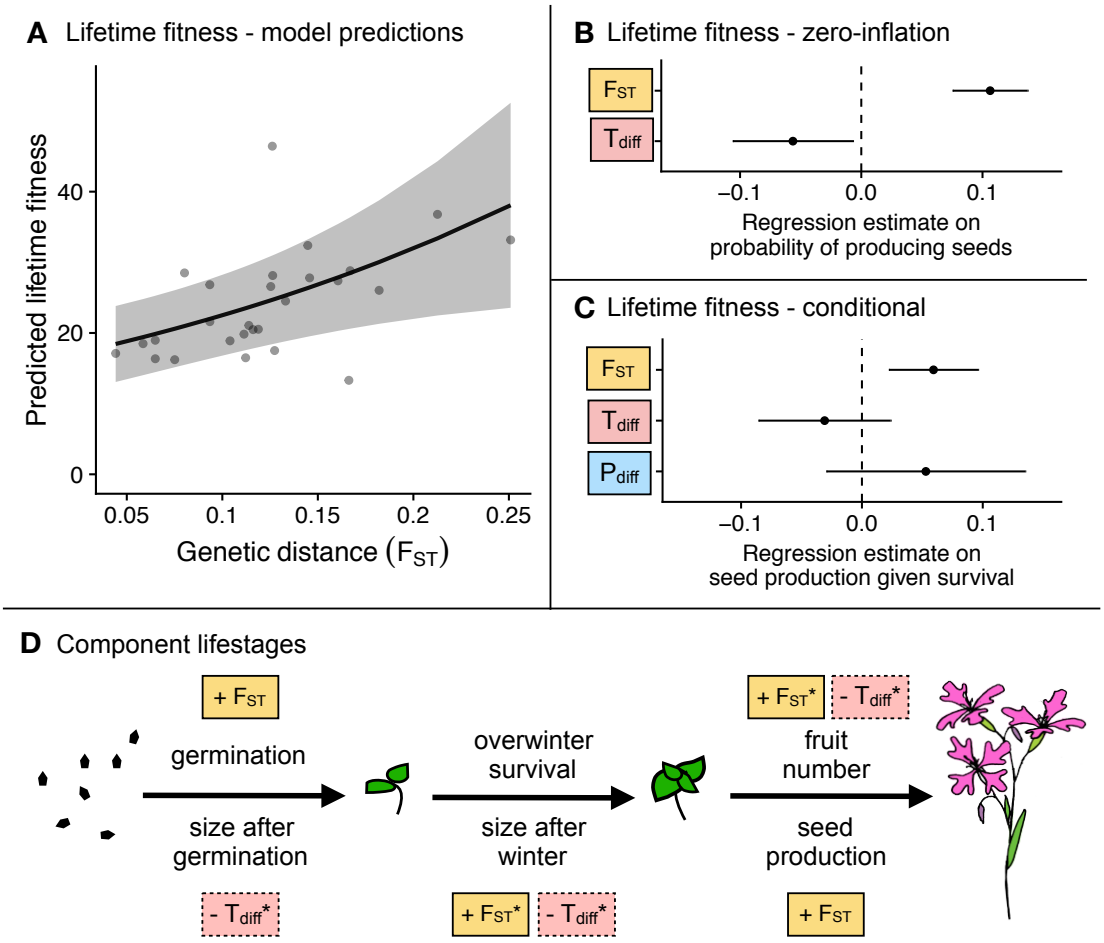

Figure 5.5: Effects of genetic differentiation between parental populations, as well as midparent temperature and precipitation on performance of *Clarkia pulchella*. **(A)** Among between-population crosses, increased genetic divergence between parental populations had a positive effect on lifetime fitness. Each point represents the average for a combination of parental populations. These effects manifested through both conditional seed production and the probability of producing seeds. **(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 "-". Marginally 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 5.5 and Table 5.6.

Table 5.1: Geographic information for the populations of *Clarkia pulchella* used in this experiment.

Population name	Map ID	Type	Latitude	Longitude	Elevation (m)
Blue Lake	F1	Focal	49.05	-119.56	842
Johnstone Creek	F2	Focal	49.04	-119.05	866
Border	D1	Donor	48.98	-118.99	1211
Day Creek Road	D2	Donor	48.94	-118.51	911
Bodie Mountain	D3	Donor	48.83	-118.83	1603
Boulder Creek Road	D4	Donor	48.76	-118.33	1115
Aeneas Valley	D5	Donor	48.54	-118.91	1126
Henry Creek Road	D6	Donor	47.45	-114.77	1103
Heyburn State Park	D7	Donor	47.34	-116.79	801
McCrosky State Park	D8	Donor	47.09	-116.98	1186
Graves Creek Road	D9	Donor	46.80	-114.41	1201
Bitterroot	D10	Donor	46.54	-113.89	1424
Abel's Ridge	D11	Donor	46.28	-117.60	1457
Tucannon	D12	Donor	46.24	-117.74	1022
Pendleton	D13	Donor	45.74	-118.25	649

Table 5.2: Results of generalized linear mixed effects models for the effect of local vs. foreign origin on performance of *Clarkia pulchella* in common gardens. There are no significant differences between populations of local vs. foreign origin in fitness components or lifetime fitness. Size during the previous census (November for overwinter survival and size, March for fruit counts and estimated seed production) is always a significant predictor of performance in subsequent lifestages.

Response	Foreign vs. local origin			Size during previous census		
	Estimate	SE	*P*-value	Estimate	SE	*P*-value
Germination	-0.034	0.088	0.702	-	-	-
Size after germination	0.027	0.073	0.708	-	-	-
Overwinter survival	0.176	0.126	0.161	**0.377**	**0.033**	**< 0.001**
Size after winter	-0.068	0.182	0.708	**0.750**	**0.042**	**< 0.001**
Fruit count	0.089	0.146	0.544	**0.544**	**0.041**	**< 0.001**
Seed production	-0.005	0.128	0.966	**0.388**	**0.032**	**< 0.001**
Lifetime fitness - zero inflation	0.042	0.129	0.742	-	-	-
Lifetime fitness - conditional	0.061	0.127	0.629	-	-	-

Table 5.3: Results of generalized linear mixed effects models of the effects of absolute precipitation and temperature differences on component lifestages of *Clarkia pulchella*. Temperature and precipitation differences refer to absolute differences between the historic conditions that a population experienced and the conditions in the common gardens during the experiment. These differences were calculated using climate data from only the months of that census period (i.e., September-November for germination and size after germination, December-March for overwinter survival and size after winter, and April-July for fruit counts and seed production). Analyses were conducted using only plants surviving the previous census window. Whenever applicable, size in the previous census was included as a covariate to account for differences accumulated during earlier lifestages. Significant parameters are indicated with bold text.

	Absolute temperature difference			Absolute precipitation difference			Size in previous census		
Response	Estimate	SE	P-value	Estimate	SE	P-value	Estimate	SE	P-value
Germination	**-0.094**	**0.035**	**0.007**	-	-	-	-	-	-
Size after germination	-0.027	0.026	0.299	-	-	-	-	-	-
Overwinter survival	**-0.078**	**0.032**	**0.015**	-	-	-	**0.366**	**0.033**	**< 0.001**
Size after winter	**-0.518**	**0.095**	**< 0.001**	-	-	-	**0.748**	**0.042**	**< 0.001**
Fruit count	-0.114	0.071	0.108	-0.091	0.058	0.119	**0.543**	**0.041**	**< 0.001**
Seed production	-0.055	0.047	0.249	**-0.110**	**0.044**	**0.011**	**0.389**	**0.033**	**< 0.001**

Table 5.4: Results of generalized linear mixed effects models of the effects of being a within-population cross vs. a between-population cross, while accounting for effects of absolute precipitation and temperature differences. Temperature and precipitation differences refer to absolute differences between the average historic conditions of an individual's parental populations and the conditions in the common gardens during the experiment. Positive estimates of the effects of between-population vs. within-populations indicate that having parents from two different populations ("gene flow") is beneficial. **(A)** Effects on lifetime fitness of *Clarkia pulchella*. **(B)** Effects on lifetime fitness when midparent precipitation differences are not included in the model. **(C)** Effects on component lifestages. Climate differences were calculated using climate data from only the months of that census period (i.e., September-November for germination and size after germination, December-March for overwinter survival and size after winter, and April-July for fruit counts and seed production). Analyses were conducted using only plants surviving the previous census window. Whenever applicable, size in the previous census was included as a covariate to account for differences accumulated during earlier lifestages. Significant parameters are indicated with bold text.

	Between-populations vs. within-populations			Absolute midparent temperature difference			Absolute midparent precipitation difference			Size during previous census		
Response	Estimate	SE	*P*-value	Estimate	SE	*P*-value	Estimate	SE	*P*-value	Estimate	SE	*P*-value
A. Lifetime fitness												
Lifetime fitness - zero inflation	0.008	0.047	0.860	**-0.228**	**0.024**	**< 0.001**	-	-	-	-	-	-
Lifetime fitness - conditional	0.047	0.049	0.345	**-0.085**	**0.032**	**0.007**	**-0.072**	**0.036**	**0.045**	-	-	-
B. Lifetime fitness without precipitation in model												
Lifetime fitness - zero inflation	0.008	0.047	0.860	**-0.228**	**0.024**	**< 0.001**	-	-	-	-	-	-
Lifetime fitness - conditional	**0.114**	**0.036**	**0.002**	**-0.059**	**0.029**	**0.041**	-	-	-	-	-	-
C. Component lifestages												
Germination	-0.035	0.078	0.653	**-0.088**	**0.029**	**0.003**	-	-	-	-	-	-
Size after germination	-0.033	0.041	0.421	-0.028	0.015	0.058	-	-	-	-	-	-
Overwinter survival	0.012	0.047	0.794	**-0.058**	**0.022**	**0.009**	-	-	-	**0.336**	**0.023**	**< 0.001**
Size after winter	-0.052	0.127	0.678	**-0.274**	**0.068**	**< 0.001**	-	-	-	**0.721**	**0.031**	**< 0.001**
Fruit count	**0.125**	**0.055**	**0.021**	**-0.081**	**0.032**	**0.011**	**-0.092**	**0.034**	**0.007**	**0.493**	**0.030**	**< 0.001**
Seed production	0.092	0.047	0.052	-0.030	0.028	0.285	**-0.092**	**0.030**	**0.002**	**0.357**	**0.024**	**< 0.001**

Table 5.5: Results of generalized linear mixed effects models separately testing the effects of **(A)** genetic differentiation, **(B)** absolute midparent temperature differences, and **(C)** absolute midparent precipitation differences on performance of *Clarkia pulchella* in common gardens. Absolute midparent temperature and precipitation differences refer to absolute differences between the conditions in the common gardens during the experiment and the average historic conditions of an individual's parental populations. These analyses were performed using between-population crosses only, that is, every plant has one parent from a focal population and one parent from a donor population. For analyses of lifetime fitness, temperature differences were calculated using the duration of the experiment and precipitation differences were calculated using April-July values. Precipitation differences are only included as an effect in the conditional part of the model of lifetime fitness because precipitation effects are expected to manifest at later lifestages. For analyses of component lifestages, climate differences were calculated using climate data from only the months of that census period (i.e., September-November for germination and size after germination, December-March for overwinter survival and size after winter, and April-July for fruit counts and seed production). Component lifestage analyses were conducted using only plants surviving the previous census window. Whenever applicable, size in the previous census was included as a covariate to account for differences accumulated during earlier lifestages. Significant parameters are indicated with bold text.

A. F_{ST}

	F_{ST}			Size during previous census		
Response	Estimate	SE	P-value	Estimate	SE	P-value
Lifetime fitness - zero inflation	**0.117**	**0.029**	**< 0.001**	-	-	-
Lifetime fitness - conditional	**0.080**	**0.031**	**0.010**	-	-	-
Germination	**0.145**	**0.065**	**0.025**	-	-	-
Size after germination	0.014	0.014	0.306	-	-	-
Overwinter survival	0.054	0.035	0.124	**0.336**	**0.037**	**< 0.001**
Size after winter	**0.104**	**0.050**	**0.035**	**0.743**	**0.044**	**< 0.001**
Fruit count	**0.067**	**0.031**	**0.028**	**0.451**	**0.044**	**< 0.001**
Seed production	**0.081**	**0.029**	**0.005**	**0.334**	**0.035**	**< 0.001**

B. Absolute midparent temperature differences

	Temperature difference			Size during previous census		
Response	Estimate	SE	P-value	Estimate	SE	P-value
Lifetime fitness - zero inflation	**-0.108**	**0.048**	**0.024**	-	-	-
Lifetime fitness - conditional	-0.090	0.056	0.104	-	-	-
Germination	-0.078	0.075	0.293	-	-	-
Size after germination	-0.029	0.017	0.096	-	-	-
Overwinter survival	-0.054	0.049	0.275	**0.335**	**0.037**	**< 0.001**
Size after winter	-0.132	0.070	0.060	**0.743**	**0.044**	**< 0.001**
Fruit count	-0.121	0.063	0.057	**0.457**	**0.044**	**< 0.001**
Seed production	-0.071	0.066	0.283	**0.336**	**0.035**	**< 0.001**

C. Absolute midparent precipitation differences

	Precipitation difference			Size during previous census		
Response	Estimate	SE	P-value	Estimate	SE	P-value
Lifetime fitness - conditional	0.123	0.078	0.113	-	-	-
Fruit count	0.041	0.083	0.621	**0.445**	**0.044**	**< 0.001**
Seed production	0.099	0.076	0.193	**0.334**	**0.035**	**< 0.001**

Table 5.6: Results of generalized linear mixed effects models of the effects of genetic differentiation between parental populations on performance of *Clarkia pulchella* in common gardens. Effects of absolute precipitation and temperature differences are also included in these models. Temperature and precipitation differences refer to the absolute midparent differences, i.e., the absolute differences between the conditions in the common gardens during the experiment and the average historic conditions of an individual's parental populations. These analyses were performed using between-population crosses only, that is, every plant has one parent from a focal population and one parent from a donor population. **(A)** Effects on lifetime fitness. **(B)** Effects on component lifestages. Climate differences were calculated using climate data from only the months of that census period (i.e., September-November for germination and size after germination, December-March for overwinter survival and size after winter, and April-July for fruit counts and seed production). Analyses were conducted using only plants surviving the previous census window. Whenever applicable, size in the previous census was included as a covariate to account for differences accumulated during earlier lifestages. Significant parameters are indicated with bold text.

	F_{ST}			Temperature difference			Precipitation difference			Size during previous census		
Response	Estimate	SE	*P*-value	Estimate	SE	*P*-value	Estimate	SE	*P*-value	Estimate	SE	*P*-value
A. Lifetime fitness												
Lifetime fitness - zero inflation	**0.106**	**0.031**	**0.001**	-0.056	0.050	0.255	-	-	-	-	-	-
Lifetime fitness - conditional	0.059	0.037	0.109	-0.031	0.055	0.572	0.053	0.082	0.520	-	-	-
B. Component lifestages												
Germination	**0.138**	**0.070**	**0.047**	-0.021	0.080	0.798	-	-	-	-	-	-
Size after germination	0.005	0.016	0.770	-0.026	0.020	0.179	-	-	-	-	-	-
Overwinter survival	0.047	0.044	0.286	-0.014	0.060	0.809	-	-	-	**0.335**	**0.037**	**< 0.001**
Size after winter	0.072	0.061	0.232	-0.073	0.082	0.376	-	-	-	**0.742**	**0.044**	**< 0.001**
Fruit count	0.058	0.034	0.089	-0.093	0.058	0.110	-0.010	0.084	0.905	**0.455**	**0.044**	**< 0.001**
Seed production	**0.071**	**0.033**	**0.031**	-0.035	0.055	0.519	0.028	0.077	0.719	**0.334**	**0.035**	**< 0.001**

Chapter 6

Conclusions

In this dissertation, I have investigated how floral traits, reproduction, and reproductive assurance vary with climate and geography across the range of the winter annual plant *Clarkia pulchella* (Chapter 2, Chapter 3). I also described the genetic structure of populations of *C. pulchella* across the range and tested whether environmental differences contribute to genetic differentiation between populations (Chapter 4). Finally, I examined the effects of gene flow on populations at the northern range edge using a controlled crossing design and common gardens, and I tested how the effects of gene flow are modulated by climatic and genetic differentiation (Chapter 5). In this concluding chapter, I first summarize the findings and emphasize the major contributions of each of these projects (Section 6.1). I later discuss conclusions that can be drawn about range limiting processes in *C. pulchella* and highlight directions for future work in this system (Section 6.2). Finally, I reflect on the present state of research on geographic range limits, discuss current disconnects between empirical research and the theory that inspires it, and consider how this field will be impacted by disequilibrium imposed by climate change (Section 6.3).

6.1 Major findings

6.1.1 Chapter 2: Associations of climate and geography with herbarium specimen characteristics

In Chapter 2, I tested for effects of climate and estimated isolation (derived from a species distribution model) on plant characteristics measured on herbarium specimens of *Clarkia pulchella*. The conditions that are thought to promote self-pollination are quite similar to predicted characteristics of range edge populations (small, low density, or more limited by their environments), and this prompted me to investigate whether range edge populations also have traits consistent with higher rates of self-pollination. I measured reproductive output and floral traits that are often associated with mating system (petal length and herkogamy) on specimens collected across the species' range. I extracted climate data associated with specimens and derived a population isolation metric from a species distribution model. This isolation metric was based on the average predicted suitability

of the area around the location where a specimen was collected, and it assumes that the density of populations or individuals is predicted by this distribution model. This work illustrates the potential for leveraging existing herbarium collections to investigate trait variation over a larger geographic area and over a wider variety of climatic conditions than would typically be possible in field surveys.

Across the range of *C. pulchella*, some climatic axes correlate with geographic range position (Table 2.5, Figure 2.5AD), but there is still a great deal of variation around these trends. Therefore, in this project I first considered the effects of environment on reproduction and floral traits, then considered whether the environment varied systematically with range position, and finally examined whether this resulted in geographic trends in traits or reproduction. A strength of this approach is that it identifies putative drivers (climate variables) of geographic variation in plant characteristics first, then attempts to map variation in both the trait and the drivers back on to the geographic range. Some spatial patterns emerged in my results: reproductive output was positively correlated with summer precipitation (Table 2.1, Figure 2.5E) and reproductive output declined from the centre of the range towards the western range edge (Table 2.6, Figure 2.5F). In contrast, although climate was related to both herkogamy and geographic range position, the residual variance around each of these relationships was large enough that I did not identify spatial trends in that trait (Figure 2.5ABC). So, while herkogamy was lower in sites with warm spring and summer temperatures, herkogamy was not predicted by geography (Figure 2.5C).

These results indicate that low precipitation is possibly a factor limiting reproduction of *C. pulchella* at the southern and western edges of its geographic range. The role of summer precipitation as a selective force in this species is consistent with the results of Chapter 5, in which provenances that were best matched to precipitation conditions during the experiment performed best. The association that I found between reduced herkogamy and warm temperatures led me to expect that I might detect greater rates of reproductive assurance in populations in warm sites during my field exclusion of pollinators (Chapter 3), but this was not the case (possible reasons discussed below).

6.1.2 Chapter 3: Exclusion of pollinators in natural populations of *Clarkia pulchella*

In this project, I followed up on results of Chapter 2 by manipulating pollinator access to plants in eight sites spanning the geographic range of *Clarkia pulchella*. My goal was to investigate geographic and climatic drivers of autonomous reproductive assurance (seed set in the absence of pollinators) and fruit production. I examined how reproductive assurance and fruit production varied with the positions of sites within the range of the species, as well as with temperature and precipitation. I found that reproductive assurance in *C. pulchella* was greatest in the northern part of the species' range (Figure 3.3) and was not well-explained by any of the climate variables that I considered. Despite some degree of reproductive assurance in all populations, pollinators are important for seed production in this species, and recruitment appears to be sensitive to the magnitude of seed input.

The results of this study contrast with my expectation (based on Chapter 2) that reproduc-

tive assurance might be greater in warmer sites, where evaporative stress is high and flowering times might be compressed. There are several factors that might explain this discrepancy. First, herkogamy may not actually be a consistent predictor of capacity for autonomous self-pollination in this species. If this is the case, the variation in this trait measured in Chapter 2 would not result in variation in seed production when pollinators are excluded. Second, there was a great deal of variation around the relationship between herkogamy and temperature in Chapter 2, so the limited number of sites at which I conducted manipulations may not have allowed for enough statistical power to detect a relationship. Finally, biogeographic processes, such as range expansion, may also contribute to geographic differentiation in traits, and this could result in patterns that are predicted broadly by geography and only driven by environmental variation at a finer spatial scale (if at all). I did not have enough replication within regions to evaluate whether climate affects traits within regions of the species' range.

Consistent with Chapter 2, fruit production of *C. pulchella* was positively correlated with summer precipitation (Figure 3.6). In the absence of pollinators, some populations of *C. pulchella*, particularly those in wetter sites, appear to have the capacity to increase fruit production, perhaps through resource reallocation. While populations appear to be adapted to average precipitation conditions (Chapter 5), individuals are also able to respond to precipitation availability during the flowering season. This is also consistent with the positive correlation between fruit production and deviations from average precipitation in Chapter 2 (Table 2.1).

6.1.3 Chapter 4: Genetic structure across the geographic range of *Clarkia pulchella*

Both of the previous chapters and my transplant experiment can be interpreted more fully with some knowledge of the genetic structure of populations *Clarkia pulchella* across the landscape. In Chapter 4, I sampled 32 populations from across the range and tested whether climatic differences between populations correlated with their genetic differentiation. I found no notable contribution of climatic differences, indicating that any processes that might operate to differentiate populations based on temperature or precipitation are not affecting the putatively neutral loci in these analyses (Figure 4.4). Rather, these results support seed and pollen movement at limited distances relative to the species' range and that this movement and the subsequent incorporation of immigrants into the local gene pool are not influenced by temperature or precipitation similarities among populations. I also investigated patterns of population structure and geographic gradients in genetic diversity. I found that populations in the northern and southern parts of the range mostly belonged to distinct genetic groups and that central and eastern populations were admixed between these two groups (Figure 4.6). This could be the result of a past or current geographic barrier associated with the Columbia Plateau, or it could be the result of spread from separate sets of refugia after the last glacial maximum.

One possible explanation for the increased capacity for self-pollination in the absence of pollinators at the northern range edge (Chapter 3) is that small population sizes during post-glacial

range expansion favoured individuals with greater capacity for self-pollination. However, in light of the results from my population genetics analyses, this explanation seems less plausible. I found an increase in genetic diversity towards the northern range edge (Figure 4.7), rather than the decline that might be expected if populations spread north at very low densities. Both of these patterns are surprising, and further investigation is necessary to understand them. It is possible that differences in traits can be attributed to different phylogeographic histories in different parts of the range and may not be the result of selection. Alternatively, regional differences in pollinator community composition or pollinator phenology could be explored as drivers of trait divergence.

6.1.4 Chapter 5: Effects of gene flow on performance at the northern range edge

I designed this transplant experiment to test two competing predictions about the effects of gene flow on range edge populations. Gene flow might inhibit edge populations by disrupting adaptation to local conditions. Alternatively, if range edge populations are small or isolated, gene flow may provide beneficial genetic variation. I simulated gene flow in the greenhouse, using 13 populations from across the northern half of the range of *Clarkia pulchella* to pollinate plants local to two sites at the northern range edge. I then planted the progeny of these crosses into common gardens in these sites and monitored them over their lifespan, from germination to reproduction. During the experiment, conditions were very warm, and this raised an additional question: what are the effects of gene flow when local populations experience climates that strongly diverge from those that they have historically experienced? My results indicate that populations are locally adapted to temperature and precipitation in their sites of origin. However, the anomalously warm conditions during the experiment resulted in the disruption of local adaptation: plants that had one or both parents from warmer provenances outperformed individuals with two local parents (Figure 5.3, Figure 5.4). 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.

The extent to which fall and winter temperatures predicted fitness differences among populations in this experiment surprised me. In previous projects I had dismissed the importance of these seasons, but they are evidently quite important for growth and establishment. Future investigations into whether populations are locally adapted in their germination cues and which environmental variables trigger germination could be interesting.

With regard to the questions that initially motivated the experiment, I found a benefit of gene flow that was independent of effects of climate matching (Figure 5.4). Relief from homozygosity (or benefit of heterozygosity) is consistent with predictions of positive effects of gene flow on range edge populations beyond just providing alleles that are adaptive in the edge environment and is supported by the result that there were benefits of having parents from more genetically differentiated populations (Figure 5.5). However, it is unclear over how many generations these benefits might persist. It is also possible that these benefits are not unique to range edge populations. It would have been interesting to perform a parallel experiment in the interior of the range to see

if central populations show similar or different responses to gene flow compared to populations at the range edge. While gene flow from cooler populations had negative effects, which could be considered support for the potential of swamping gene flow, it is unlikely that gene flow among natural populations is ever as great in magnitude as what I have simulated, and it seems likely that selection against alleles that are maladaptive in the local climate might prevent them from negatively affecting the overall population growth rate.

6.2 What limits the range in *Clarkia pulchella*? Synthesis and future directions

In the chapters presented here, I have taken a variety of approaches to study processes playing out at large spatial scales among populations of the species *Clarkia pulchella*. While identifying the definitive causes of range limits for *C. pulchella* will require more work, my results identify some important factors influencing population dynamics and local adaptation across the species' geographic range. Much of my work can only be interpreted in the context of range limits with some caution and assumptions. I only studied features of populations within the range—a more direct test of range limits (though one with its own set of caveats) would be to move individuals beyond the range edge and try to identify what (if anything) limits their performance (Gaston, 2003; Lee-Yaw et al., 2016). Additionally, while I tried in my transplant experiment (and to a very small extent in my pollinator exclusion study) to consider cumulative effects across multiple lifestages, my results only explain differences in relative fitness components among populations and cannot be directly extended to inferences about population persistence. Despite these limitations, in this section I draw some conclusions about what might limit the geographic range in this species and what work could be done to further test these ideas.

Pollinators are important for seed production in *C. pulchella* (Chapter 3), and populations may be somewhat differentiated in traits that are often related to mating system (Chapter 2), but pollinator availability did not seem to strongly limit seed production in any of the populations where I conducted pollinator exclusions. To really know this, I would need to do hand pollinations to assess maximum seed set (Knight et al., 2005; Eckert et al., 2010), so I base that statement solely on the fact that flowers that were exposed to pollinators set approximately three times as many seeds as those with pollinators excluded. The average number of seeds in fruits in control plots varied (non-significantly) across the range and was lowest in the Southwest (Figure 3.3). The southern and western range edges are also the parts of the range where summer precipitation may be most limiting of fruit production (Table 2.1), as summer precipitation near these edges is both lower (Figure 4.2) and more variable (Table 2.5) than in the range centre. Therefore, the southern and western edges may be the places where populations of *C. pulchella* have shorter flowering seasons and where the total number of flowers on display in a population is low. Were I to continue studying whether pollination limits the geographic range of this species, I would work at these edges, and use experimental arrays of different flower numbers and densities both within and beyond the range to investigate whether Allee effects might limit pollinator attraction and

subsequent colonization success near these range edges, similar to work done by Groom (1998) in the congener *Clarkia concinna*. These array experiments could be accompanied by small scale over-the-edge transplants to determine the flowering phenology in sites beyond the range so that arrays could be placed during the appropriate time window. These projects would address the potential for declines in colonization rates to limit the range at these edges, as discussed in Section 1.3.3. Should these range edges appear to be limited by pollinator visitation, it would be interesting to see if herkogamy and dichogamy are linked to rates of self-pollination in populations near the margin, and if so, to measure the genetic variance and heritability of these traits (Opedal et al., 2017).

Populations at the northern edge of the range of *C. pulchella* do not seem strongly limited by low genetic variance, as might be expected if they have a history of small population size. This conclusion is supported by the fact that they did not show declines in neutral genetic diversity with increasing latitude, as might be expected with small historic population sizes or frequent bottlenecks (Hoffmann and Blows, 1994). Further evidence comes from the fact that they did no worse in the common gardens than populations from other parts of the range, once climate of origin was controlled for (Section 5.4.2); there did not seem to be anything inherently bad about being from a northern edge population. It also seems unlikely that swamping gene flow is leading to maladaptation in these populations—they performed as expected based on their climate of origin. Rather, the northern range edge of *C. pulchella* may be dispersal limited. This may have historically been the case, or it may be a scenario induced by recent climate change. Either way, this hypothesis could be tested with transplants into potentially suitable habitat beyond the northern range edge. However, the results of Chapter 5 highlight the importance of understanding not only what limits the species beyond its current distribution but also how extant populations will adapt to the rapidly changing sites that they already occupy. Fitness is affected by both temperature and precipitation in this species (likely in addition to environmental variables that I haven't considered in this thesis), and populations may experience novel combinations of these facets of climate in the future (Williams and Jackson, 2007; Mahony et al., 2017). A future direction that could inform both climate change responses and range limits would be to examine whether populations have adequate genetic variance to adapt to these diverse selection pressures, and if they do, whether the phenotypes favoured by each climatic axis are antagonistic or correlated.

Based on the results of Chapter 5, and my growing understanding of the theory of swamping gene flow, I think that swamping is unlikely to be an important process at the scale of the geographic range in *C. pulchella*. Were I to continue to investigate its potential role in this system, I would work along steeper environmental gradients at smaller spatial scales where swamping is more likely to be relevant, such as along elevation gradients. At these smaller spatial scales, gene flow between populations is likely to occur more frequently, perhaps frequently enough to swamp local adaptation. However, temperature and precipitation changes along these gradients would likely also generate phenology variation, and it would be interesting to simultaneously investigate to what extent differences in phenology might prevent gene flow via pollen.

In addition to the contributions that these case studies have made to our body of knowledge

on geographic range limits and spatial variation in plant mating systems, I also hope that this work has shown that *C. pulchella* is a tractable system for studying geographic range limits and adaptation to climate. Populations are large and generally easy to find, it can be grown in the greenhouse and transplanted as seeds in the field, and its annual life history makes it possible to study each component lifestage within a reasonable time span. The natural history of the species is very interesting, in particular the fact that it has a winter annual life history despite growing in sites that have both long, cold winters and extremely dry summers. I would like to do more work in this study system and hope that others might be inspired to do so as well.

6.3 Next steps in range limit research

Range limits are inherently difficult to study. A common thread I have noticed in the range limit literature (which is exemplified by my own dissertation work) is that research that seeks to understand geographic range limits often results in findings that further our understanding of the ecology and evolution of a given species, but these findings frequently fall short of explaining the range limit in question. We still lack the ability to generalize broadly about in which taxa, at which edges, we expect a given factor to be limiting. It also remains challenging to find case studies that unequivocally support some of the classic theoretical predictions. This is certainly not to say that researchers are doing a poor job. Rather, this reflects the facts that research on geographic range limits requires collection of data of many types at large spatial scales, that the theory about geographic range limits sets up predictions that are quite difficult to test, and that the predictions associated with one causal mechanism are often not mutually exclusive from those of another.

Inferences about low genetic variance (discussed in Section 1.3.1) that are drawn from studies of neutral markers (reviewed in Eckert et al., 2008; Pironon et al., 2017) can inform us to some extent about historical or contemporary demography of range edge populations, but they do not tell us whether heritable variation for adaptive traits declines towards margins. To understand whether range margins are limited by adaptation requires knowing which traits are ecologically relevant in habitats at and beyond the range edge and measuring heritable variation in these traits (Hoffmann et al., 2003; Blows and Hoffmann, 2005); this process is generally labor-intensive and is intractable for some species. The swamping gene flow hypothesis (discussed in Section 1.3.2) also presents empirical challenges. To comprehensively test it requires measuring rates of gene flow across an environmental gradient, measuring whether populations at the peripheries of the gradient are demographic sinks, and evaluating whether peripheral populations are not at the phenotypic optimum for their environment, but are instead displaced from this optimum towards that of central populations (Kirkpatrick and Barton, 1997). Knowing any one of these things in the absence of the others does not allow for differentiation between a range edge limited by swamping gene flow and one limited by other processes. For example, low demographic rates and suboptimal phenotypes are also expected if adaptive variance is limited due to drift or strong selection (Hoffmann and Blows, 1994). Finally, the data required to test whether a range limit conforms to expectations of metapopulation models (discussed in Section 1.3.3) are quite difficult to obtain. Measuring

extinction, colonization, and dispersal success at the scale of the geographic range is generally a prohibitively challenging task. This is not to say that we should not continue to try to gather the data needed to test these hypotheses, but the field might benefit from better communication between theoreticians and empiricists so that theory is tested at relevant spatial scales and in scenarios where its critical assumptions are met.

Empiricists (including myself) are frequently inspired to test range limit theories in natural systems, but often we do not consider (or do not have the data necessary to know) whether or not our systems conform to the assumptions of a particular theory. My work (Chapter 5) would have benefited from a more thoughtful (if somewhat qualitative) evaluation of whether or not the environmental differences between populations relative to the likely frequency of gene flow between populations fell within a range that is expected to produce a range limit (Kirkpatrick and Barton, 1997). Generalizing forward from theory about what measurable preconditions make a given hypothesis ripe for testing in a natural system may be more efficient than attempts to generalize backwards about which theories have the most support in empirical tests. Our understanding of geographic range limits will also be advanced if theoretical explorations are extended to be more applicable to natural landscapes, for example, by incorporating temporal variability in selection pressures, or by transitioning from assumptions of smooth environmental gradients to models that allow for environmental heterogeneity to be more broadly defined (Polechova, 2018). Using simulations to assess how robust a theory is to violations of assumptions will also help us understand whether we can expect it to apply in a given study system (Bridle et al., 2010).

A potentially fruitful complement to work in natural systems is the development of experimental evolution systems that could be used to test theoretical predictions for equilibrial range limits in the lab. Lab systems are currently being used to study range expansions (Ochocki and Miller, 2017; Williams et al., 2016) and adaptation under demographic decline (Bell and Gonzalez, 2011), and it seems possible to extend these types of experiments to generate ranges at equilibrium. Experiments that explore adaptation along artificial environmental gradients with the possibility of controlling rates and distances of dispersal, population sizes, and starting genetic variance might allow us to identify the ranges of conditions under which predictions of theory are met and might reveal parameters of importance that are not currently considered. Individual based simulations offer similar advantages, and have been used to explore the effect of carrying capacity on adaptation along a gradient (Bridle et al., 2010) and to investigate how genetic architecture influences rates of range expansion on an environmentally patchy landscape (Gilbert and Whitlock, 2017).

Finally, in a time of rapid climate change and extensive habitat modification, it is important to revise our expectations for whether and when we expect to find range limits at equilibrium. Many species are shifting their ranges in response to climate change (Chen et al., 2011; Parmesan et al., 1999); however, many others will be prevented from tracking their climatic niche due to limited dispersal rates (Midgley et al., 2006; Schloss et al., 2012). Rather than focusing primarily on the ecological and evolutionary factors that limit species in space, it is increasingly important to investigate what limits adaptation to conditions changing in time, as many species lag behind

their optimal phenotype in a changing environment (McGraw et al., 2015; Wilczek et al., 2014). Themes that have been on the forefront of research on geographic range limits—such as the limits to adaptation in novel environments, metapopulation dynamics in heterogeneous environments, and the importance of spatially varying biotic interactions—are all the more interesting and important to understand in our changing world.

Bibliography

Afkhami, M. E., P. J. McIntyre, and S. Y. Strauss (2014). Mutualist-mediated effects on species' range limits across large geographic scales. *Ecology Letters 17*(10), 1265–1273. → page 4

Aitken, S. N. and M. C. Whitlock (2013). Assisted gene flow to facilitate local adaptation to climate change. *Annual Review of Ecology, Evolution, and Systematics 44*. → page 68

Alleaume-Benharira, M., I. Pen, and O. Ronce (2006). Geographical patterns of adaptation within a species range: interactions between drift and gene flow. *Journal of Evolutionary Biology 19*(1), 203–215. → page 3

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). *The American Naturalist 186*(S1), S60–S73. → page 78

Angert, A. L. (2006). Demography of central and marginal populations of monkeyflowers (*Mimulus cardinalis* and *M. lewisii*). *Ecology 87*(8), 2014–2025. → pages 5, 6

Angert, A. L. (2009). The niche, limits to species' distributions, and spatiotemporal variation in demography across the elevation ranges of two monkeyflowers. *Proceedings of the National Academy of Sciences 106*(Supplement 2), 19693–19698. → page 5

Angert, A. L., M. Bayly, S. N. Sheth, and J. R. Paul (2018). Testing range-limit hypotheses using range-wide habitat suitability and occupancy for the scarlet monkeyflower (*Erythranthe cardinalis*). *The American Naturalist 191*(3), E000–E000. → page 6

Angert, A. L., H. Bradshaw Jr, and D. W. Schemske (2008). Using experimental evolution to investigate geographic range limits in monkeyflowers. *Evolution 62*(10), 2660–2675. → page 5

Angert, A. L. and D. W. Schemske (2005). The evolution of species' distributions: reciprocal transplants across the elevation ranges of *Mimulus cardinalis* and *M. lewisii*. *Evolution 59*(8), 1671–1684. → page 5

Antonovics, J. (1976). The nature of limits to natural selection. *Annals of the Missouri Botanical Garden*, 224–247. → pages 1, 2, 66

Ashman, T.-L., T. M. Knight, J. A. Steets, P. Amarasekare, M. Burd, D. R. Campbell, M. R. Dudash, M. O. Johnston, S. J. Mazer, R. J. Mitchell, et al. (2004). Pollen limitation of plant reproduction: ecological and evolutionary causes and consequences. *Ecology 85*(9), 2408–2421. → page 32

Baker, H. G. (1955). Self-compatibility and establishment after 'long-distance' dispersal. *Evolution 9*(3), 347–349. → pages 20, 21, 33

Barton, N. (2001). Adaptation at the edge of a species' range. *Special Publication–British Ecological Society 14*, 365–392. → page 3

Bayly, M. (2015). *Translocations of Mimulus cardinalis beyond the northern range limit show that dispersal limitation can invalidate ecological niche models.* Ph. D. thesis, University of British Columbia. → page 6

Beatty, G. E. and J. Provan (2011). Phylogeographic analysis of North American populations of the parasitic herbaceous plant *Monotropa hypopitys* L. reveals a complex history of range expansion from multiple late glacial refugia. *Journal of Biogeography 38*(8), 1585–1599. → page 49

Bell, G. and A. Gonzalez (2011). Adaptation and evolutionary rescue in metapopulations experiencing environmental deterioration. *Science 332*(6035), 1327–1330. → page 99

Benning, J., V. M. Eckhart, M. A. Geber, and D. A. Moeller (2018). Biotic interactions limit the geographic range of an annual plant: herbivory and phenology mediate fitness beyond a range margin. *bioRxiv*, 300590. → pages 7, 80

Bivand, R., T. Keitt, and B. Rowlingson (2014). rgdal: Bindings for the geospatial data abstraction library. R package version 0.9-1. → page 14

Bivand, R. and C. Rundel (2013). rgeos: Interface to geometry engine-open source (GEOS). R package version 0.3-8. → page 14

Bjerknes, A.-L., Ø. Totland, S. J. Hegland, and A. Nielsen (2007). Do alien plant invasions really affect pollination success in native plant species? *Biological Conservation 138*(1), 1–12. → page 32

Blows, M. W. and A. A. Hoffmann (2005). A reassessment of genetic limits to evolutionary change. *Ecology 86*(6), 1371–1384. → page 98

Bodbyl Roels, S. A. and J. K. Kelly (2011). Rapid evolution caused by pollinator loss in *Mimulus guttatus*. *Evolution 65*(9), 2541–2552. → page 32

Bolnick, D. I., L. K. Snowberg, C. Patenia, W. E. Stutz, T. Ingram, and O. L. Lau (2009). Phenotype-dependent native habitat preference facilitates divergence between parapatric lake and stream stickleback. *Evolution 63*(8), 2004–2016. → page 49

Bontrager, M. and A. L. Angert (2016). Effects of range-wide variation in climate and isolation on floral traits and reproductive output of *Clarkia pulchella*. *American Journal of Botany 103*(1), 10–21. → page 34

Booth, D. B., K. G. Troost, J. J. Clague, and R. B. Waitt (2003). The Cordilleran ice sheet. *Developments in Quaternary Sciences 1*, 17–43. → page 49

Bradburd, G., G. Coop, and P. Ralph (2017). Inferring continuous and discrete population genetic structure across space. *bioRxiv*, 189688. → page 54

Bradburd, G. S., P. L. Ralph, and G. M. Coop (2013). Disentangling the effects of geographic and ecological isolation on genetic differentiation. *Evolution 67*(11), 3258–3273. → pages 49, 53

Bridle, J. R., S. Gavaz, and W. J. Kennington (2009). Testing limits to adaptation along altitudinal gradients in rainforest *Drosophila*. *Proceedings of the Royal Society of London B: Biological Sciences 276*(1661), 1507–1515. → page 7

Bridle, J. R., J. Polechová, M. Kawata, and R. K. Butlin (2010). Why is adaptation prevented at ecological margins? New insights from individual-based simulations. *Ecology Letters 13*(4), 485–494. → page 99

Bridle, J. R. and T. H. Vines (2007). Limits to evolution at range margins: when and why does adaptation fail? *Trends in Ecology & Evolution 22*(3), 140–147. → pages 2, 66

Briscoe Runquist, R. D. and D. A. Moeller (2013). Resource reallocation does not influence estimates of pollen limitation or reproductive assurance in *Clarkia xantiana* subsp. *parviflora* (Onagraceae). *American Journal of Botany 100*(9), 1916–1921. → page 40

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. *The R Journal 9*(2), 378–400. → pages 37, 73

Brown, J. H. (1984). On the relationship between abundance and distribution of species. *The American Naturalist 124*(2), 255–279. → pages 2, 4

Brown, J. H., G. C. Stevens, and D. M. Kaufman (1996). The geographic range: size, shape, boundaries, and internal structure. *Annual Review of Ecology and Systematics 27*(1), 597–623. → page 11

Broxton, P. D., X. Zeng, W. Scheftic, and P. A. Troch (2014). A MODIS-based global 1-km maximum green vegetation fraction dataset. *Journal of Applied Meteorology and Climatology 53*(8), 1996–2004. → page 15

Bruckman, D. and D. R. Campbell (2016). Pollination of a native plant changes with distance and density of invasive plants in a simulated biological invasion. *American Journal of Botany 103*(8), 1458–1465. → page 32

Brunsfeld, S., J. Sullivan, D. Soltis, and P. Soltis (2001). Comparative phylogeography of northwestern North America: a synthesis. *Special Publication–British Ecological Society 14*, 319–340. → pages 49, 57

Brunsfeld, S. J. and J. Sullivan (2005). A multi-compartmented glacial refugium in the northern Rocky Mountains: evidence from the phylogeography of *Cardamine constancei* (Brassicaceae). *Conservation Genetics 6*(6), 895–904. → page 57

Brys, R., B. Geens, T. Beeckman, and H. Jacquemyn (2013). Differences in dichogamy and herkogamy contribute to higher selfing in contrasting environments in the annual *Blackstonia perfoliata* (Gentianaceae). *Annals of Botany 111*(4), 651–661. → page 20

Burd, M. (1994). Principle and plant reproduction: the role of pollen limitation in fruit and seed set. *The Botanical Review 60*(1), 83–139. → page 32

Busch, J. W. (2005). The evolution of self-compatibility in geographically peripheral populations of *Leavenworthia alabamica* (Brassicaceae). *American Journal of Botany 92*(9), 1503–1512. → pages 33, 39

Button, L., A. L. Villalobos, S. R. Dart, and C. G. Eckert (2012). Reduced petal size and color associated with transitions from outcrossing to selfing in *Camissoniopsis cheiranthifolia* (Onagraceae). *International Journal of Plant Sciences 173*(3), 251–260. → page 13

Carr, D. E. and C. B. Fenster (1994). Levels of genetic variation and covariation for *Mimulus* (Scrophulariaceae) floral traits. *Heredity 72*(6), 606. → page 13

Carstens, B. C., R. S. Brennan, V. Chua, C. V. Duffie, M. G. Harvey, R. A. Koch, C. D. McMahan, B. J. Nelson, C. E. Newman, J. D. Satler, et al. (2013). Model selection as a tool for phylogeographic inference: an example from the willow *Salix melanopsis*. *Molecular Ecology 22*(15), 4014–4028. → page 49

Case, T. J. and M. L. Taper (2000). Interspecific competition, environmental gradients, gene flow, and the coevolution of species' borders. *The American Naturalist 155*(5), 583–605. → page 4

Catchen, J., P. A. Hohenlohe, S. Bassham, A. Amores, and W. A. Cresko (2013). Stacks: an analysis tool set for population genomics. *Molecular Ecology 22*(11), 3124–3140. → page 52

Catchen, J. M., A. Amores, P. Hohenlohe, W. Cresko, and J. H. Postlethwait (2011). Stacks: building and genotyping loci de novo from short-read sequences. *G3: Genes, Genomes, Genetics 1*(3), 171–182. → page 52

Chen, I.-C., J. K. Hill, R. Ohlemüller, D. B. Roy, and C. D. Thomas (2011). Rapid range shifts of species associated with high levels of climate warming. *Science 333*(6045), 1024–1026. → page 99

Coulson, T., E. A. Catchpole, S. D. Albon, B. J. Morgan, J. Pemberton, T. H. Clutton-Brock, M. Crawley, and B. Grenfell (2001). Age, sex, density, winter weather, and population crashes in Soay sheep. *Science 292*(5521), 1528–1531. → page 33

Dart, S. R., K. E. Samis, E. Austen, and C. G. Eckert (2011). Broad geographic covariation between floral traits and the mating system in *Camissoniopsis cheiranthifolia* (Onagraceae): multiple stable mixed mating systems across the species' range? *Annals of Botany 109*(3), 599–611. → pages 13, 33

Darwin, C. (1859). The origin of species: By means of natural selection or the preservation of favoured races in the struggle for life. Technical report, London: John Murray. → page 1

Dixon, A. L., C. R. Herlihy, and J. W. Busch (2013). Demographic and population-genetic tests provide mixed support for the abundant centre hypothesis in the endemic plant *Leavenworthia stylosa*. *Molecular Ecology 22*(7), 1777–1791. → page 11

Doak, D. F. and W. F. Morris (2010). Demographic compensation and tipping points in climate-induced range shifts. *Nature 467*(7318), 959. → page 33

Eckert, C., K. Samis, and S. Lougheed (2008). Genetic variation across species' geographical ranges: the central–marginal hypothesis and beyond. *Molecular Ecology 17*(5), 1170–1188. → pages 66, 98

Eckert, C. G., S. Kalisz, M. A. Geber, R. Sargent, E. Elle, P.-O. Cheptou, C. Goodwillie, M. O. Johnston, J. K. Kelly, D. A. Moeller, et al. (2010). Plant mating systems in a changing world. *Trends in Ecology & Evolution 25*(1), 35–43. → pages 33, 96

Eckhart, V. M., M. A. Geber, W. F. Morris, E. S. Fabio, P. Tiffin, and D. A. Moeller (2011). The geography of demography: long-term demographic studies and species distribution models reveal a species border limited by adaptation. *The American Naturalist 178*(S1), S26–S43. → page 6

Emery, N. C. (2009). Ecological limits and fitness consequences of cross-gradient pollen movement in *Lasthenia fremontii*. *The American Naturalist 174*(2), 221–235. → page 67

Epps, C. W., P. J. Palsbøll, J. D. Wehausen, G. K. Roderick, R. R. Ramey, and D. R. McCullough (2005). Highways block gene flow and cause a rapid decline in genetic diversity of desert bighorn sheep. *Ecology Letters 8*(10), 1029–1038. → page 48

Etterson, J. R. and R. G. Shaw (2001). Constraint to adaptive evolution in response to global warming. *Science 294*(5540), 151–154. → page 80

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. *Journal of Evolutionary Biology 25*(8), 1676–1685. → page 67

Felsenstein, J. (1977). Multivariate normal genetic models with a finite number of loci. Technical report, Washington Univ., Seattle (USA). Dept. of Genetics. → page 3

Fenster, C. B. and L. F. Galloway (2000). Population differentiation in an annual legume: genetic architecture. *Evolution 54*(4), 1157–1172. → page 79

Fishman, L. and J. H. Willis (2008). Pollen limitation and natural selection on floral characters in the yellow monkeyflower, *Mimulus guttatus*. *New Phytologist 177*(3), 802–810. → page 10

Fishman, L. and R. Wyatt (1999). Pollinator-mediated competition, reproductive character displacement, and the evolution of selfing in *Arenaria uniflora* (Caryophyllaceae). *Evolution 53*(6), 1723–1733. → page 20

Frankham, R., J. D. Ballou, M. D. Eldridge, R. C. Lacy, K. Ralls, M. R. Dudash, and C. B. Fenster (2011). Predicting the probability of outbreeding depression. *Conservation Biology 25*(3), 465–475. → page 67

Galen, C., R. A. Sherry, and A. B. Carroll (1999). Are flowers physiological sinks or faucets? costs and correlates of water use by flowers of *Polemonium viscosum*. *Oecologia 118*(4), 461–470. → page 32

Gamble, D. E., M. Bontrager, and A. L. Angert (2018). Floral trait variation and links to climate in the mixed-mating annual *Clarkia pulchella*. *Botany 96*(7), 425–435. → page 41

García-Ramos, G. and M. Kirkpatrick (1997). Genetic models of adaptation and gene flow in peripheral populations. *Evolution 51*(1), 21–28. → page 3

Gaston, K. J. (2003). *The structure and dynamics of geographic ranges*. Oxford University Press. → pages 11, 96

Geber, M. A. and V. M. Eckhart (2005). Experimental studies of adaptation in *Clarkia xantiana*. II. Fitness variation across a subspecies border. *Evolution 59*(3), 521–531. → page 6

Geospatial Information Authority of Japan, Chiba University, and collaborating organizations (2008). Global map-percent tree cover. https://www.iscgm.org/gmd/. → page 15

Gilbert, K. and M. Whitlock (2017). The genetics of adaptation to discrete heterogeneous environments: frequent mutation or large-effect alleles can allow range expansion. *Journal of Evolutionary Biology 30*(3), 591–602. → page 99

Glémin, S., E. Bazin, and D. Charlesworth (2006). Impact of mating systems on patterns of sequence polymorphism in flowering plants. *Proceedings of the Royal Society of London B: Biological Sciences 273*(1604), 3011–3019. → page 10

Goldberg, E. E. and R. Lande (2006). Ecological and reproductive character displacement of an environmental gradient. *Evolution 60*(7), 1344–1357. → page 4

Goodwillie, C., R. D. Sargent, C. G. Eckert, E. Elle, M. A. Geber, M. O. Johnston, S. Kalisz, D. A. Moeller, R. H. Ree, M. Vallejo-Marin, et al. (2010). Correlated evolution of mating system and floral display traits in flowering plants and its implications for the distribution of mating system variation. *New Phytologist 185*(1), 311–321. → page 13

Gottlieb, L. and V. Ford (1988). Genetic studies of the pattern of floral pigmentation in *Clarkia gracilis*. *Heredity 60*(2), 237. → page 19

Goudet, J. and T. Jombart (2015). hierfstat: Estimation and tests of hierarchical F-statistics. R package version 0.04-22. → pages 53, 73

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*. *Journal of Ecology 102*(1), 95–107. → pages 6, 7, 67

Groom, M. J. (1998). Allee effects limit population viability of an annual plant. *The American Naturalist 151*(6), 487–496. → pages 11, 97

Guillot, G. and F. Rousset (2013). Dismantling the Mantel tests. *Methods in Ecology and Evolution 4*(4), 336–344. → page 53

Hargreaves, A. L. and C. G. Eckert (2014). Evolution of dispersal and mating systems along geographic gradients: implications for shifting ranges. *Functional Ecology 28*(1), 5–21. → page 10

Hereford, J. (2009). A quantitative survey of local adaptation and fitness trade-offs. *The American Naturalist 173*(5), 579–588. → page 73

Herlihy, C. R. and C. G. Eckert (2005). Evolution of self-fertilization at geographical range margins? a comparison of demographic, floral, and mating system variables in central vs. peripheral populations of *Aquilegia canadensis* (Ranunculaceae). *American Journal of Botany 92*(4), 744–751. → page 33

Herlihy, C. R. and C. G. Eckert (2007). Evolutionary analysis of a key floral trait in *Aquilegia canadensis* (Ranunculaceae): genetic variation in herkogamy and its effect on the mating system. *Evolution 61*(7), 1661–1674. → page 13

Hewitt, G. (2004). Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society B: Biological Sciences 359*(1442), 183–195. → page 49

Hijmans, R., S. Phillips, J. Leathwick, and J. Elith (2016). dismo: Species distribution modeling. → page 14

Hijmans, R. J., S. E. Cameron, J. L. Parra, P. G. Jones, and A. Jarvis (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology 25*(15), 1965–1978. → page 15

Hijmans, R. J. and J. van Etten (2014). raster: Geographic data analysis and modeling. *R package version 2.3-12*. → page 14

Hirao, A. and G. Kudo (2004). Landscape genetics of alpine-snowbed plants: comparisons along geographic and snowmelt gradients. *Heredity 93*(3), 290. → page 56

Hoffmann, A., R. Hallas, J. Dean, and M. Schiffer (2003). Low potential for climatic stress adaptation in a rainforest *Drosophila* species. *Science 301*(5629), 100–102. → pages 7, 8, 80, 98

Hoffmann, A. A. and M. W. Blows (1994). Species borders: ecological and evolutionary perspectives. *Trends in Ecology & Evolution 9*(6), 223–227. → pages 2, 10, 97, 98

Holliday, J. A., H. Suren, and S. N. Aitken (2012). Divergent selection and heterogeneous migration rates across the range of Sitka spruce (*Picea sitchensis*). *Proceedings of the Royal Society of London B: Biological Sciences 279*(1734), 1675–1683. → page 67

Holt, R. D. (2003). On the evolutionary ecology of species ranges. *Evolutionary Ecology Research 5*(2), 159–178. → page 2

Holt, R. D. and T. H. Keitt (2000). Alternative causes for range limits: a metapopulation perspective. *Ecology Letters 3*(1), 41–47. → pages 2, 3, 4, 11, 20, 66

Holtsford, T. P. and N. C. Ellstrand (1992). Genetic and environmental variation in floral traits affecting outcrossing rate in *Clarkia tembloriensis* (Onagraceae). *Evolution 46*(1), 216–225. → page 13

Hutchinson, M. (1957). Concluding remarks. In *Cold Spring Harbour Symposia on quantitative biology: population studies: animal ecology and demography. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, USA*, pp. 415–427. → page 1

Ingvarsson, P. K. and M. C. Whitlock (2000). Heterosis increases the effective migration rate. *Proceedings of the Royal Society of London B: Biological Sciences 267*(1450), 1321–1326. → pages 68, 79

Jenkins, N. L. and A. A. Hoffmann (1999). Limits to the southern border of *Drosophila serrata*: cold resistance, heritable variation, and trade-offs. *Evolution 53*(6), 1823–1834. → page 8

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*(8), 1585–1595. → page 79

Johnson, S. D., C. I. Peter, L. A. Nilsson, and J. Ågren (2003). Pollination success in a deceptive orchid is enhanced by co-occurring rewarding magnet plants. *Ecology 84*(11), 2919–2927. → page 20

Jorgensen, R. and H. Arathi (2013). Floral longevity and autonomous selfing are altered by pollination and water availability in *Collinsia heterophylla*. *Annals of Botany 112*(5), 821–828. → page 11

Karron, J. D., R. T. Jackson, N. N. Thumser, and S. L. Schlicht (1997). Outcrossing rates of individual *Mimulus ringens* genets are correlated with anther–stigma separation. *Heredity 79*(4), 365. → page 13

Kawecki, T. J. and D. Ebert (2004). Conceptual issues in local adaptation. *Ecology Letters 7*(12), 1225–1241. → page 73

Kay, K. M. and D. A. Picklum (2013). Drought alters the expression of mating system traits in two species of *Clarkia*. *Evolutionary Ecology 27*(5), 899–910. → page 33

Keitt, T. H., M. A. Lewis, and R. D. Holt (2001). Allee effects, invasion pinning, and species' borders. *The American Naturalist 157*(2), 203–216. → pages 3, 11

Kellermann, V., V. Loeschcke, A. A. Hoffmann, T. N. Kristensen, C. Fløjgaard, J. R. David, J.-C. Svenning, and J. Overgaard (2012). Phylogenetic contstraints in key functional traits behind species climate niches: patterns of dessication and cold resistance across 95 *Drosophila* species. *Evolution 66*(11), 3377–3389. → page 8

Kellermann, V., J. Overgaard, A. A. Hoffmann, C. Fløjgaard, J.-C. Svenning, and V. Loeschcke (2012). Upper thermal limits of *Drosophila* are linked to species distributions and strongly constrained phylogenetically. *Proceedings of the National Academy of Sciences 109*(40), 16228–16233. → page 8

Kellermann, V., B. van Heerwaarden, C. M. Sgrò, and A. A. Hoffmann (2009). Fundamental evolutionary limits in ecological traits drive *Drosophila* species distributions. *Science 325*(5945), 1244–1246. → pages 4, 8

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*(5), 1104–1108. → page 80

Kirkpatrick, M. and N. H. Barton (1997). Evolution of a species' range. *The American Naturalist 150*(1), 1–23. → pages 2, 3, 4, 5, 49, 67, 78, 98, 99

Knight, T. M. (2003). Floral density, pollen limitation, and reproductive success in *Trillium grandiflorum*. *Oecologia 137*(4), 557–563. → page 11

Knight, T. M., J. A. Steets, J. C. Vamosi, S. J. Mazer, M. Burd, D. R. Campbell, M. R. Dudash, M. O. Johnston, R. J. Mitchell, and T.-L. Ashman (2005). Pollen limitation of plant reproduction: pattern and process. *Annual Review of Ecology, Evolution, and Systematics 36*, 467–497. → pages 33, 96

Kramer, A. T., J. B. Fant, and M. V. Ashley (2011). Influences of landscape and pollinators on population genetic structure: examples from three *Penstemon* (Plantaginaceae) species in the Great Basin. *American Journal of Botany 98*(1), 109–121. → pages 4, 33

Kuchta, S. R. and A.-M. Tan (2005). Isolation by distance and post-glacial range expansion in the rough-skinned newt, *Taricha granulosa*. *Molecular Ecology 14*(1), 225–244. → page 49

Kudo, G. and T. Y. Ida (2013). Early onset of spring increases the phenological mismatch between plants and pollinators. *Ecology 94*(10), 2311–2320. → page 32

Kudo, G., Y. Nishikawa, T. Kasagi, and S. Kosuge (2004). Does seed production of spring ephemerals decrease when spring comes early? *Ecological Research 19*(2), 255–259. → page 32

Lande, R. (1992). Neutral theory of quantitative genetic variance in an island model with local extinction and colonization. *Evolution 46*(2), 381–389. → page 66

Lee, S.-R., Y.-S. Jo, C.-H. Park, J. M. Friedman, and M. S. Olson (2018). Population genomic analysis suggests strong influence of river network on spatial distribution of genetic variation in invasive saltcedar across the southwestern United States. *Molecular Ecology 27*(3), 636–646. → page 48

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. *Ecology Letters 19*(6), 710–722. → pages 1, 66, 96

Lennartsson, T. (2002). Extinction thresholds and disrupted plant–pollinator interactions in fragmented plant populations. *Ecology 83*(11), 3060–3072. → pages 4, 33

Lennon, J. J., J. R. Turner, and D. Connell (1997). A metapopulation model of species boundaries. *Oikos*, 486–502. → pages 2, 3, 20

Lenormand, T. (2002). Gene flow and the limits to natural selection. *Trends in Ecology & Evolution 17*(4), 183–189. → page 67

Lewis, H. (1953). The mechanism of evolution in the genus *Clarkia*. *Evolution 7*(1), 1–20. → pages 12, 13

Lewis, H. (1955). The genus *Clarkia*. *University of California Publications in Botany 28*, 241–392. → pages 15, 40

Lira-Noriega, A. and J. D. Manthey (2014). Relationship of genetic diversity and niche centrality: a survey and analysis. *Evolution 68*(4), 1082–1093. → page 11

Lüdecke, D. (2018). *ggeffects: Create tidy data frames of marginal effects for 'ggplot' from model outputs*. R package version 0.3.2. → page 37

Luo, Y. and A. Widmer (2013). Herkogamy and its effects on mating patterns in *Arabidopsis thaliana*. *PLoS One 8*(2), e57902. → page 13

MacArthur, R. H. (1972). *Geographical ecology: patterns in the distribution of species*. Princeton University Press. → page 1

MacSwain, J. W., P. H. Raven, and R. W. Thorp (1973). Comparative behavior of bees and Onagraceae: 4. *Clarkia* bees of the Western United States. *University of California Publications in Entomology 70*. → pages 34, 39, 69

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 Biology 23*(9), 3934–3955. → pages 79, 97

Massatti, R. and L. L. Knowles (2014). Microhabitat differences impact phylogeographic concordance of codistributed species: Genomic evidence in montane sedges (*Carex* L.) from the Rocky Mountains. *Evolution 68*(10), 2833–2846. → page 57

Matvienko, M., A. Kozik, L. Froenicke, D. Lavelle, B. Martineau, B. Perroud, and R. Michelmore (2013). Consequences of normalizing transcriptomic and genomic libraries of plant genomes using a duplex-specific nuclease and tetramethylammonium chloride. *PLoS One 8*(2), e55913. → page 52

Mazer, S. J., L. S. Dudley, A. A. Hove, S. K. Emms, and A. S. Verhoeven (2010). Physiological performance in *Clarkia* sister taxa with contrasting mating systems: do early-flowering autogamous taxa avoid water stress relative to their pollinator-dependent counterparts? *International Journal of Plant Sciences 171*(9), 1029–1047. → page 19

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 Biology 21*(10), 3827–3835. → pages 32, 79, 100

McRae, B. H. (2006). Isolation by resistance. *Evolution 60*(8), 1551–1561. → page 48

Merow, C., M. Smith, T. Edwards, A. Guisan, S. McMahon, S. Normand, W. Thuiller, R. Wüest, N. Zimmermann, and J. Elith (2014). Back to the basics of species distribution modeling: What do we gain from complex versus simple models? *Ecography 37*, 1267–1281. → page 15

Meyer, L., S. Brinkman, L. van Kesteren, N. Leprince-Ringuet, and F. van Boxmeer (2014). IPCC Climate Change 2014: Synthesis report. contribution of working groups I, II and III to the fifth assessment report of the Intergovernmental Panel on Climate Change. pp. 3–87. → page 35

Midgley, G., G. Hughes, W. Thuiller, and A. Rebelo (2006). Migration rate limitations on climate change-induced range shifts in Cape Proteaceae. *Diversity and Distributions 12*(5), 555–562. → page 99

Miller-Struttmann, N. E., J. C. Geib, J. D. Franklin, P. G. Kevan, R. M. Holdo, D. Ebert-May, A. M. Lynn, J. A. Kettenbach, E. Hedrick, and C. Galen (2015). Functional mismatch in a bumble bee pollination mutualism under climate change. *Science 349*(6255), 1541–1544. → page 32

Mimura, M. and S. Aitken (2007). Adaptive gradients and isolation-by-distance with postglacial migration in picea sitchensis. *Heredity 99*(2), 224. → page 33

Mitchell, R. J., R. J. Flanagan, B. J. Brown, N. M. Waser, and J. D. Karron (2009). New frontiers in competition for pollination. *Annals of Botany 103*(9), 1403–1413. → page 20

Moeller, D. A. (2006). Geographic structure of pollinator communities, reproductive assurance, and the evolution of self-pollination. *Ecology 87*(6), 1510–1522. → pages 13, 19

Moeller, D. A., R. D. Briscoe Runquist, A. M. Moe, M. A. Geber, C. Goodwillie, P.-O. Cheptou, C. G. Eckert, E. Elle, M. O. Johnston, S. Kalisz, et al. (2017). Global biogeography of mating system variation in seed plants. *Ecology Letters 20*(3), 375–384. → page 39

Moeller, D. A. and M. A. Geber (2005a). Ecological context of the evolution of self-pollination in *Clarkia xantiana*: population size, plant communities, and reproductive assurance. *Evolution 59*(4), 786–799. → pages 10, 11, 20

Moeller, D. A. and M. A. Geber (2005b). Ecological context of the evolution of self-pollination in *Clarkia xantiana*: population size, plant communities, and reproductive assurance. *Evolution 59*(4), 786–799. → page 33

Moeller, D. A., M. A. Geber, V. M. Eckhart, and P. Tiffin (2012). Reduced pollinator service and elevated pollen limitation at the geographic range limit of an annual plant. *Ecology 93*(5), 1036–1048. → pages 7, 19

Moeller, D. A., M. A. Geber, and P. Tiffin (2011). Population genetics and the evolution of geographic range limits in an annual plant. *The American Naturalist 178*(S1), S44–S61. → pages 6, 7

Moore, J.-S. and A. P. Hendry (2009). Can gene flow have negative demographic consequences? Mixed evidence from stream threespine stickleback. *Philosophical Transactions of the Royal Society B: Biological Sciences 364*(1523), 1533–1542. → page 67

Morjan, C. L. and L. H. Rieseberg (2004). How species evolve collectively: implications of gene flow and selection for the spread of advantageous alleles. *Molecular Ecology 13*(6), 1341–1356. → page 6

Mosca, E., A. Eckert, E. Di Pierro, D. Rocchini, N. La Porta, P. Belletti, and D. Neale (2012). The geographical and environmental determinants of genetic diversity for four alpine conifers of the European Alps. *Molecular Ecology 21*(22), 5530–5545. → page 49

Mosquin, T. (1964). Chromosomal repatterning in *Clarkia rhomboidea* as evidence for post-pleistocene changes in distribution. *Evolution 18*(1), 12–25. → page 19

Muir, C. D. and A. L. Angert (2017). Grow with the flow: a latitudinal cline in physiology is associated with more variable precipitation in *Erythranthe cardinalis*. *Journal of Evolutionary Biology 30*(12), 2189–2203. → page 5

Nei, M. (1987). *Molecular Evolutionary Genetics*. Columbia University Press. → page 53

Newman, D. and D. Pilson (1997). Increased probability of extinction due to decreased genetic effective population size: experimental populations of *Clarkia pulchella*. *Evolution 51*(2), 354–362. → pages 12, 19, 79

Nosil, P., T. H. Vines, and D. J. Funk (2005). Perspective: reproductive isolation caused by natural selection against immigrants from divergent habitats. *Evolution 59*(4), 705–719. → page 49

Ochocki, B. M. and T. E. Miller (2017). Rapid evolution of dispersal ability makes biological invasions faster and more variable. *Nature Communications 8*, 14315. → pages 80, 99

Opedal, Ø. H., G. H. Bolstad, T. F. Hansen, W. S. Armbruster, and C. Pélabon (2017). The evolvability of herkogamy: quantifying the evolutionary potential of a composite trait. *Evolution 71*(6), 1572–1586. → page 97

Palladini, J. D. and J. L. Maron (2013). Indirect competition for pollinators is weak compared to direct resource competition: pollination and performance in the face of an invader. *Oecologia 172*(4), 1061–1069. → pages 12, 20, 34, 40, 69

Pannell, J. R., J. R. Auld, Y. Brandvain, M. Burd, J. W. Busch, P.-O. Cheptou, J. K. Conner, E. E. Goldberg, A.-G. Grant, D. L. Grossenbacher, et al. (2015). The scope of Baker's law. *New Phytologist 208*(3), 656–667. → page 33

Pannell, J. R. and S. C. Barrett (1998). Baker's law revisited: reproductive assurance in a metapopulation. *Evolution 52*(3), 657–668. → page 20

Parmesan, C., N. Ryrholm, C. Stefanescu, J. K. Hill, C. D. Thomas, H. Descimon, B. Huntley, L. Kaila, J. Kullberg, T. Tammaru, et al. (1999). Poleward shifts in geographical ranges of butterfly species associated with regional warming. *Nature 399*(6736), 579. → page 99

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*. *The American Naturalist 178*(S1), S62–S79. → pages 5, 67

Pebesma, E. and R. Bivand (2005). sp: Classes and methods for spatial data. → page 14

Petit, R. J., I. Aguinagalde, J.-L. de Beaulieu, C. Bittkau, S. Brewer, R. Cheddadi, R. Ennos, S. Fineschi, D. Grivet, M. Lascoux, et al. (2003). Glacial refugia: hotspots but not melting pots of genetic diversity. *Science 300*(5625), 1563–1565. → page 57

Phillips, S. J., R. P. Anderson, and R. E. Schapire (2006). Maximum entropy modeling of species geographic distributions. *Ecological Modelling 190*(3-4), 231–259. → page 14

Phillips, S. J., M. Dudík, and R. E. Schapire (2004). A maximum entropy approach to species distribution modeling. In *Proceedings of the Twenty-first International Conference on Machine Learning*, pp. 83. Association for Computing Machinery. → page 14

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. *Biological Reviews 92*(4), 1877–1909. → pages 4, 66, 98

Poland, J. A., P. J. Brown, M. E. Sorrells, and J.-L. Jannink (2012). Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. *PloS One 7*(2), e32253. → page 51

Polechova, J. (2018). Is the sky the limit? on the expansion threshold of a species' range. *bioRxiv*, 234377. → page 99

Polechová, J. and N. H. Barton (2015). Limits to adaptation along environmental gradients. *Proceedings of the National Academy of Sciences 112*(20), 6401–6406. → page 2

Price, T. D. and M. Kirkpatrick (2009). Evolutionarily stable range limits set by interspecific competition. *Proceedings of the Royal Society of London B: Biological Sciences*, rspb–2008. → page 4

PRISM Climate Group (2017). Oregon state university. Accessed: 2017-09-30. → pages 51, 60, 61, 65, 72, 127, 128

Pritchard, J. K., M. Stephens, and P. Donnelly (2000). Inference of population structure using multilocus genotype data. *Genetics 155*(2), 945–959. → page 54

Pujol, B. and J. R. Pannell (2008). Reduced responses to selection after species range expansion. *Science 321*(5885), 96–96. → page 66

Pulliam, H. R. (2000). On the relationship between niche and distribution. *Ecology Letters 3*(4), 349–361. → page 1

R Core Team (2013). *R: A language and environment for statistical computing, version 3.0.2*. Vienna, Austria: R Foundation for Statistical Computing. → page 14

R Core Team (2017). *R: a language and environment for statistical computing, version 3.4.4*. Vienna, Austria: R Foundation for Statistical Computing. → pages 37, 53, 54, 75

Randle, A. M., J. B. Slyder, and S. Kalisz (2009). Can differences in autonomous selfing ability explain differences in range size among sister-taxa pairs of *Collinsia* (Plantaginaceae)? an extension of Baker's Law. *New Phytologist 183*(3), 618–629. → page 39

Rasband, W. (2012). ImageJ. version 1.46 r. US National Institutes of Health, Bethesda, MD. → page 14

Reeves, P. A. and C. M. Richards (2014). Effect of a geographic barrier on adaptation in the dwarf sunflower (*Helianthus pumilus* Nutt.). *International Journal of Plant Sciences 175*(6), 688–701. → page 48

Revell, L. J. (2012). phytools: An R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution 3*, 217–223. → page 53

Riddle, B. R., T. Jezkova, M. E. Eckstut, V. Oláh-Hemmings, and L. N. Carraway (2014). Cryptic divergence and revised species taxonomy within the Great Basin pocket mouse, *Perognathus parvus* (Peale, 1848), species group. *Journal of Mammalogy 95*(1), 9–25. → page 56

Rochette, N. C. and J. M. Catchen (2017). Deriving genotypes from RAD-seq short-read data using Stacks. *Nature Protocols 12*(12), 2640. → page 52

Runions, C. J. and M. A. Geber (2000). Evolution of the self-pollinating flower in *Clarkia xantiana* (onagraceae). I. Size and development of floral organs. *American Journal of Botany 87*(10), 1439–1451. → page 19

Sagarin, R. D. and S. D. Gaines (2002). The 'abundant centre' distribution: to what extent is it a biogeographical rule? *Ecology Letters 5*(1), 137–147. → pages 2, 4, 11, 68

Sagarin, R. D., S. D. Gaines, and B. Gaylord (2006). Moving beyond assumptions to understand abundance distributions across the ranges of species. *Trends in Ecology & Evolution 21*(9), 524–530. → page 11

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*(11), 2520–2536. → page 67

Schloss, C. A., T. A. Nuñez, and J. J. Lawler (2012). Dispersal will limit ability of mammals to track climate change in the western hemisphere. *Proceedings of the National Academy of Sciences 109*(22), 8606–8611. → page 99

Schoen, D. J. and A. Brown (1991). Intraspecific variation in population gene diversity and effective population size correlates with the mating system in plants. *Proceedings of the National Academy of Sciences 88*(10), 4494–4497. → page 10

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), 1–15. → pages 48, 49, 68

Sexton, J. P., P. J. McIntyre, A. L. Angert, and K. J. Rice (2009). Evolution and ecology of species range limits. *Annual Review of Ecology, Evolution, and Systematics 40*. → pages 2, 11, 66, 67

Sexton, J. P., S. Y. Strauss, and K. J. Rice (2011). Gene flow increases fitness at the warm edge of a species range. *Proceedings of the National Academy of Sciences 108*(28), 11704–11709. → pages 50, 68

Shafer, A., C. I. Cullingham, S. D. Cote, and D. W. Coltman (2010). Of glaciers and refugia: a decade of study sheds new light on the phylogeography of northwestern North America. *Molecular Ecology 19*(21), 4589–4621. → pages 49, 56

Shagina, I., E. Bogdanova, I. Z. Mamedov, Y. Lebedev, S. Lukyanov, D. Shagin, et al. (2010). Normalization of genomic DNA using duplex-specific nuclease. *Biotechniques 48*(6), 455. → page 52

Sheth, S. N. and A. L. Angert (2018). Demographic compensation does not rescue populations at a trailing range edge. *Proceedings of the National Academy of Sciences 115*(10), 2413–2418. → page 6

Shimono, Y., M. Watanabe, A. S. Hirao, N. Wada, and G. Kudo (2009). Morphological and genetic variations of *Potentilla matsumurae* (Rosaceae) between fellfield and snowbed populations. *American Journal of Botany 96*(4), 728–737. → page 56

Slatkin, M. (1973). Gene flow and selection in a cline. *Genetics 75*(4), 733–756. → page 49

Slatkin, M. (1993). Isolation by distance in equilibrium and non-equilibrium populations. *Evolution 47*(1), 264–279. → page 48

Sproul, J. S., D. D. Houston, C. R. Nelson, R. P. Evans, K. A. Crandall, and D. K. Shiozawa (2015). Climate oscillations, glacial refugia, and dispersal ability: factors influencing the genetic structure of the least salmonfly, *Pteronarcella badia* (Plecoptera), in western North America. *BMC Evolutionary Biology 15*(1), 279. → page 49

Storfer, A., M. Murphy, J. Evans, C. Goldberg, S. Robinson, S. Spear, R. Dezzani, E. Delmelle, L. Vierling, and L. Waits (2007). Putting the 'landscape' in landscape genetics. *Heredity 98*(3), 128. → page 48

Svenning, J.-C., S. Normand, and F. Skov (2008). Postglacial dispersal limitation of widespread forest plant species in nemoral europe. *Ecography 31*(3), 316–326. → page 1

Takebayashi, N. and L. F. Delph (2000). An association between a floral trait and inbreeding depression. *Evolution 54*(3), 840–846. → page 13

Takebayashi, N. and P. L. Morrell (2001). Is self-fertilization an evolutionary dead end? Revisiting an old hypothesis with genetic theories and a macroevolutionary approach. *American Journal of Botany 88*(7), 1143–1150. → page 10

van Heerwaarden, B., V. Kellermann, M. Schiffer, M. Blacket, C. M. Sgro, and A. A. Hoffmann (2009). Testing evolutionary hypotheses about species borders: patterns of genetic variation towards the southern borders of two rainforest *Drosophila* and a related habitat generalist. *Proceedings of the Royal Society of London B: Biological Sciences*, rspb–2008. → page 8

Van Kleunen, M., J. C. Manning, V. Pasqualetto, and S. D. Johnson (2007). Phylogenetically independent associations between autonomous self-fertilization and plant invasiveness. *The American Naturalist 171*(2), 195–201. → page 21

Vandepitte, K., K. Helsen, K. Van Acker, J. Mergeay, and O. Honnay (2017). Retention of gene diversity during the spread of a non-native plant species. *Molecular Ecology 26*(12), 3141–3150. → page 57

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. *Conservation Genetics 4*(5), 639–645. → pages 57, 66, 68

Wang, I. J. and G. S. Bradburd (2014). Isolation by environment. *Molecular Ecology 23*(23), 5649–5662. → pages 11, 49

Wang, T., A. Hamann, D. L. Spittlehouse, and T. Q. Murdock (2012). ClimateWNA: high-resolution spatial climate data for western North America. *Journal of Applied Meteorology and Climatology 51*(1), 16–29. → pages 15, 35, 37, 43, 123

Warren, D. L., M. Cardillo, D. F. Rosauer, and D. I. Bolnick (2014). Mistaking geography for biology: inferring processes from species distributions. *Trends in Ecology & Evolution 29*(10), 572–580. → page 41

Weir, B. S. and C. C. Cockerham (1984). Estimating F-statistics for the analysis of population structure. *Evolution 38*(6), 1358–1370. → pages 53, 72

Wilczek, A. M., M. D. Cooper, T. M. Korves, and J. Schmitt (2014). Lagging adaptation to warming climate in *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences 111*(22), 7906–7913. → pages 78, 79, 100

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. *Proceedings of the Royal Society of London B: Biological Sciences 274*(1623), 2357–2364. → page 79

Williams, J. L., B. E. Kendall, and J. M. Levine (2016). Rapid evolution accelerates plant population spread in fragmented experimental landscapes. *Science 353*(6298), 482–485. → pages 80, 99

Williams, J. W. and S. T. Jackson (2007). Novel climates, no-analog communities, and ecological surprises. *Frontiers in Ecology and the Environment 5*(9), 475–482. → pages 79, 97

Williams, P. H., M. B. Araújo, and P. Rasmont (2007). Can vulnerability among British bumblebee (*Bombus*) species be explained by niche position and breadth? *Biological Conservation 138*(3), 493–505. → page 32

Wright, S. (1943). Isolation by distance. *Genetics 28*(2), 114–138. → page 48

Wright, S. I., S. Kalisz, and T. Slotte (2013). Evolutionary consequences of self-fertilization in plants. *Proceedings of the Royal Society of London B: Biological Sciences 280*(1760), 20130133. → page 33

Yeaman, S. and A. Jarvis (2006). Regional heterogeneity and gene flow maintain variance in a quantitative trait within populations of lodgepole pine. *Proceedings of the Royal Society of London B: Biological Sciences 273*(1594), 1587–1593. → page 67

Zheng, X., D. Levine, J. Shen, S. M. Gogarten, C. Laurie, and B. S. Weir (2012). A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics 28*(24), 3326–3328. → page 52

Appendix A

Supporting Materials

Table A.1: Sensitivity analyses of tests involving isolation to the methods used to build the species distribution model in Chapter 2. We varied the number of background points and their extent, as well as the features that MaxEnt could use to fit predictors to presence/absence data. RP = range position (the distance of a specimen from the centre of the range). ISO = isolation. Petal length was always square-root transformed before analysis.

Model	Description	AUC	Cross-validated AUC	Isolation buffer size	Statistical test	n	Slope	Slope SE	F	df	*P*	R^2
Basic	310 localities, 3100 background points over 100 km buffered points, default MaxEnt features	0.836	0.805	1 km	ISO on herkogamy	120	0.11	0.37	0.087	1,118	0.77	0.00
					ISO on petal length	120	-0.037	0.16	0.055	1,118	0.81	0.00
					RP on ISO - all	260	0.00034	0.00016	4.1	1,258	0.043	0.02
					RP on ISO - north	81	-0.00033	0.00026	1.5	1,79	0.22	0.02
					RP on ISO - west	37	0.0011	0.00039	7.4	1,35	0.01	0.18
					RP on ISO - east	58	-0.00035	0.00033	1.1	1,56	0.29	0.02
					RP on ISO - south	84	0.0013	0.00031	18.5	1,82	<0.001	0.18
				5 km	ISO on herkogamy	120	0.24	0.47	0.26	1,118	0.61	0.00
					ISO on petal length	120	-0.09	0.2	0.21	1,118	0.65	0.00
					RP on ISO - all	260	0.0003	0.00014	4.6	1,258	0.033	0.02
					RP on ISO - north	81	-0.00016	0.00024	0.43	1,79	0.51	0.01
					RP on ISO - west	37	0.001	0.00033	9.6	1,35	0.004	0.22
					RP on ISO - east	58	-0.000051	0.00026	0.039	1,56	0.84	0.00
					RP on ISO - south	84	0.001	0.00026	14.6	1,82	<0.001	0.15
				10 km	ISO on herkogamy	120	0.032	0.53	0.0037	1,118	0.95	0.00
					ISO on petal length	120	-0.12	0.22	0.3	1,118	0.59	0.00
					RP on ISO - all	260	0.0003	0.00013	5.6	1,258	0.018	0.02
					RP on ISO - north	81	0.000016	0.00021	0.0064	1,79	0.94	0.00
					RP on ISO - west	37	0.00094	0.00032	8.6	1,35	0.006	0.20
					RP on ISO - east	58	0.000085	0.00028	0.09	1,56	0.77	0.00
					RP on ISO - south	84	0.00082	0.00023	13	1,82	<0.001	0.14
Smaller background extent	310 localities, 3100 background points over 50 km buffered points, default MaxEnt features	0.796	0.75	1 km	ISO on herkogamy	120	0.16	0.43	0.14	1,118	0.71	0.00
					ISO on petal length	120	-0.12	0.18	0.45	1,118	0.5	0.00
					RP on ISO - all	260	-0.000045	0.00015	0.091	1,258	0.76	0.00
					RP on ISO - north	81	-0.00076	0.00025	9.2	1,79	0.003	0.10
					RP on ISO - west	37	0.00069	0.00034	4.2	1,35	0.049	0.11
					RP on ISO - east	58	-0.00047	0.00031	2.3	1,56	0.14	0.04
					RP on ISO - south	84	0.00089	0.00025	12.4	1,82	<0.001	0.13
				5 km	ISO on herkogamy	120	0.38	0.52	0.52	1,118	0.47	0.00
					ISO on petal length	120	-0.2	0.22	0.85	1,118	0.36	0.01

Model	Description	AUC	Cross-validated AUC	Isolation buffer size	Statistical test	n	Slope	Slope SE	F	df	*P*	R^2
					RP on ISO - all	260	-0.000056	0.00013	0.19	1,258	0.66	0.00
					RP on ISO - north	81	-0.00061	0.00023	6.9	1,79	0.01	0.08
					RP on ISO - west	37	0.00068	0.0003	5.2	1,35	0.029	0.13
					RP on ISO - east	58	-0.00014	0.00024	0.32	1,56	0.57	0.01
					RP on ISO - south	84	0.00062	0.00023	7.2	1,82	0.009	0.08
				10 km	ISO on herkogamy	120	0.22	0.57	0.14	1,118	0.71	0.00
					ISO on petal length	120	-0.24	0.24	1	1,118	0.31	0.01
					RP on ISO - all	260	-0.00005	0.00012	0.17	1,258	0.68	0.00
					RP on ISO - north	81	-0.00047	0.00021	5	1,79	0.028	0.06
					RP on ISO - west	37	0.00057	0.00029	3.7	1,35	0.062	0.10
					RP on ISO - east	58	0.000021	0.00026	0.0068	1,56	0.93	0.00
					RP on ISO - south	84	0.00048	0.0002	5.9	1,82	0.018	0.07
1x background points	310 localities, 310 background points over 100 km buffered points, default MaxEnt features	0.683	0.678	1 km	ISO on herkogamy	120	-0.09	0.59	0.023	1,118	0.88	0.00
					ISO on petal length	120	-0.006	0.25	0.00059	1,118	0.98	0.00
					RP on ISO - all	260	0.00032	0.000089	12.9	1,258	<0.001	0.05
					RP on ISO - north	81	-0.0001	0.00011	0.72	1,79	0.4	0.01
					RP on ISO - west	37	0.00063	0.00022	8.2	1,35	0.007	0.19
					RP on ISO - east	58	-0.000093	0.00016	0.35	1,56	0.56	0.01
					RP on ISO - south	84	0.001	0.00018	32.8	1,82	<0.001	0.29
				5 km	ISO on herkogamy	120	0.28	0.68	0.17	1,118	0.68	0.00
					ISO on petal length	120	-0.032	0.29	0.012	1,118	0.91	0.00
					RP on ISO - all	260	0.00026	0.00009	8.3	1,258	0.004	0.03
					RP on ISO - north	81	-0.000089	0.00011	0.61	1,79	0.44	0.01
					RP on ISO - west	37	0.00067	0.00021	9.8	1,35	0.003	0.22
					RP on ISO - east	58	0.00013	0.0002	0.43	1,56	0.51	0.01
					RP on ISO - south	84	0.00077	0.00019	16.5	1,82	<0.001	0.17
				10 km	ISO on herkogamy	120	0.17	0.71	0.057	1,118	0.81	0.00
					ISO on petal length	120	-0.083	0.3	0.077	1,118	0.78	0.00
					RP on ISO - all	260	0.00023	0.000091	6.6	1,258	0.011	0.03
					RP on ISO - north	81	-0.000028	0.00011	0.069	1,79	0.79	0.00
					RP on ISO - west	37	0.00063	0.0002	9.8	1,35	0.004	0.22
					RP on ISO - east	58	0.00021	0.00025	0.72	1,56	0.4	0.01
					RP on ISO - south	84	0.00066	0.00017	14.2	1,82	<0.001	0.15
2x background points	310 localities, 620 background points over 100 km	0.745	0.732	1 km	ISO on herkogamy	120	-0.056	0.49	0.013	1,118	0.91	0.00
					ISO on petal length	120	-0.016	0.21	0.0064	1,118	0.94	0.00
					RP on ISO - all	260	0.00016	0.00011	2.1	1,258	0.15	0.01

Model	Description	AUC	Cross-validated AUC	Isolation buffer size	Statistical test	n	Slope	Slope SE	F	df	*P*	R^2
	buffered points, default MaxEnt features				RP on ISO - north	81	-0.0004	0.00015	6.8	1,79	0.011	0.08
					RP on ISO - west	37	0.00045	0.0003	2.2	1,35	0.14	0.06
					RP on ISO - east	58	-0.00034	0.00021	2.6	1,56	0.11	0.05
					RP on ISO - south	84	0.0012	0.00022	27.9	1,82	<0.001	0.25
				5 km	ISO on herkogamy	120	0.11	0.58	0.033	1,118	0.86	0.00
					ISO on petal length	120	-0.079	0.24	0.1	1,118	0.75	0.00
					RP on ISO - all	260	0.00013	0.00011	1.3	1,258	0.25	0.01
					RP on ISO - north	81	-0.00033	0.00016	4.6	1,79	0.035	0.06
					RP on ISO - west	37	0.00042	0.0003	1.9	1,35	0.17	0.05
					RP on ISO - east	58	-0.000043	0.00022	0.039	1,56	0.85	0.00
					RP on ISO - south	84	0.00091	0.00021	18.2	1,82	<0.001	0.18
				10 km	ISO on herkogamy	120	-0.027	0.63	0.0018	1,118	0.97	0.00
					ISO on petal length	120	-0.11	0.26	0.19	1,118	0.66	0.00
					RP on ISO - all	260	0.00011	0.00011	1.1	1,258	0.29	0.00
					RP on ISO - north	81	-0.00024	0.00014	2.8	1,79	0.1	0.03
					RP on ISO - west	37	0.00036	0.00029	1.6	1,35	0.21	0.04
					RP on ISO - east	58	0.000057	0.00026	0.048	1,56	0.83	0.00
					RP on ISO - south	84	0.00078	0.00019	16.8	1,82	<0.001	0.17
4x background points	310 localities, 1240 background points over 100 km buffered points, default MaxEnt features	0.796	0.777	1 km	ISO on herkogamy	120	0.032	0.42	0.0056	1,118	0.94	0.00
					ISO on petal length	120	-0.039	0.18	0.049	1,118	0.83	0.00
					RP on ISO - all	260	0.00032	0.00014	5	1,258	0.027	0.02
					RP on ISO - north	81	-0.00035	0.00021	2.6	1,79	0.11	0.03
					RP on ISO - west	37	0.001	0.00036	7.2	1,35	0.011	0.17
					RP on ISO - east	58	-0.00022	0.00028	0.64	1,56	0.43	0.01
					RP on ISO - south	84	0.0013	0.00027	20.8	1,82	<0.001	0.20
				5 km	ISO on herkogamy	120	0.16	0.52	0.099	1,118	0.75	0.00
					ISO on petal length	120	-0.12	0.22	0.28	1,118	0.6	0.00
					RP on ISO - all	260	0.00027	0.00012	4.6	1,258	0.033	0.02
					RP on ISO - north	81	-0.00023	0.0002	1.4	1,79	0.24	0.02
					RP on ISO - west	37	0.001	0.00032	9.3	1,35	0.004	0.21
					RP on ISO - east	58	0.000024	0.00023	0.011	1,56	0.92	0.00
					RP on ISO - south	84	0.001	0.00025	15.2	1,82	<0.001	0.16
				10 km	ISO on herkogamy	120	0.037	0.58	0.004	1,118	0.95	0.00
					ISO on petal length	120	-0.14	0.24	0.36	1,118	0.55	0.00
					RP on ISO - all	260	0.00024	0.00012	4.2	1,258	0.042	0.02
					RP on ISO - north	81	-0.00011	0.00018	0.39	1,79	0.54	0.00
					RP on ISO - west	37	0.00085	0.00031	7.5	1,35	0.01	0.18

Model	Description	AUC	Cross-validated AUC	Isolation buffer size	Statistical test	n	Slope	Slope SE	F	df	*P*	R^2
					RP on ISO - east	58	0.00011	0.00026	0.17	1,56	0.68	0.00
					RP on ISO - south	84	0.00079	0.00021	13.6	1,82	<0.001	0.14
No hinge features	310 localities, 3100 background points over 100 km buffered points, no hinge features	0.839	0.807	1 km	ISO on herkogamy	120	0.14	0.37	0.14	1,118	0.71	0.00
					ISO on petal length	120	-0.035	0.16	0.05	1,118	0.82	0.00
					RP on ISO - all	260	0.00035	0.00017	4.3	1,258	0.039	0.02
					RP on ISO - north	81	-0.00032	0.00027	1.5	1,79	0.23	0.02
					RP on ISO - west	37	0.0011	0.00041	6.9	1,35	0.013	0.16
					RP on ISO - east	58	-0.00038	0.00033	1.3	1,56	0.26	0.02
					RP on ISO - south	84	0.0013	0.00031	17.6	1,82	<0.001	0.18
				5 km	ISO on herkogamy	120	0.28	0.47	0.35	1,118	0.56	0.00
					ISO on petal length	120	-0.07	0.2	0.13	1,118	0.72	0.00
					RP on ISO - all	260	0.00032	0.00014	5.4	1,258	0.021	0.02
					RP on ISO - north	81	-0.00012	0.00024	0.25	1,79	0.62	0.00
					RP on ISO - west	37	0.0011	0.00035	9.5	1,35	0.004	0.21
					RP on ISO - east	58	-0.00009	0.00026	0.12	1,56	0.73	0.00
					RP on ISO - south	84	0.001	0.00027	14.5	1,82	<0.001	0.15
				10 km	ISO on herkogamy	120	0.061	0.53	0.013	1,118	0.91	0.00
					ISO on petal length	120	-0.094	0.22	0.18	1,118	0.67	0.00
					RP on ISO - all	260	0.00033	0.00013	6.7	1,258	0.01	0.03
					RP on ISO - north	81	0.000059	0.0002	0.084	1,79	0.77	0.00
					RP on ISO - west	37	0.00098	0.00033	8.8	1,35	0.005	0.20
					RP on ISO - east	58	0.000053	0.00028	0.035	1,56	0.85	0.00
					RP on ISO - south	84	0.00082	0.00023	13.2	1,82	<0.001	0.14
No hinge or threshold features	310 localities, 3100 background points over 100 km buffered points, no hinge or threshold features	0.798	0.792	1 km	ISO on herkogamy	120	0.2	0.4	0.25	1,118	0.62	0.00
					ISO on petal length	120	0.019	0.17	0.013	1,118	0.91	0.00
					RP on ISO - all	260	0.0002	0.00017	1.4	1,258	0.24	0.01
					RP on ISO - north	81	-0.00063	0.00026	6	1,79	0.016	0.07
					RP on ISO - west	37	0.00045	0.00038	1.4	1,35	0.25	0.04
					RP on ISO - east	58	0.0000052	0.00046	0.00013	1,56	0.99	0.00
					RP on ISO - south	84	0.0012	0.00025	22.5	1,82	<0.001	0.22
				5 km	ISO on herkogamy	120	0.67	0.5	1.8	1,118	0.18	0.02
					ISO on petal length	120	0.042	0.21	0.039	1,118	0.84	0.00
					RP on ISO - all	260	0.00027	0.00012	4.7	1,258	0.03	0.02
					RP on ISO - north	81	-0.0003	0.00021	2	1,79	0.16	0.03
					RP on ISO - west	37	0.00047	0.00035	1.8	1,35	0.19	0.05
					RP on ISO - east	58	0.00041	0.00027	2.4	1,56	0.13	0.04

Model	Description	AUC	Cross-validated AUC	Isolation buffer size	Statistical test	n	Slope	Slope SE	F	df	*P*	R^2
					RP on ISO - south	84	0.0009	0.00022	17.2	1,82	<0.001	0.17
				10 km	ISO on herkogamy	120	0.72	0.54	1.8	1,118	0.19	0.02
					ISO on petal length	120	0.034	0.23	0.022	1,118	0.88	0.00
					RP on ISO - all	260	0.00024	0.00011	4.7	1,258	0.031	0.02
					RP on ISO - north	81	-0.00011	0.00019	0.32	1,79	0.57	0.00
					RP on ISO - west	37	0.00032	0.00031	1	1,35	0.32	0.03
					RP on ISO - east	58	0.00049	0.00027	3.3	1,56	0.076	0.06
					RP on ISO - south	84	0.00067	0.00019	12.4	1,82	<0.001	0.13

Table A.2: Pearson correlation coefficients among precipitation and temperature variables associated with experimental sites used in Chapter 3. For **(A)** precipitation, **(B)** temperature, and **(C)** precipitation and temperature, correlations are shown between annual, fall (September-November), winter (December-February), spring (March-May), and summer (June-July, because all plants senesce before August). Normal values were calculated over 50 years (1963-2012), while 2014-2015 values are from the growing season of plants in the experiment. Normal climate data is from ClimateWNA (Wang et al., 2012) and 2014-2015 variables are from PRISM (PRISM Climate Group, Oregon State University, prism.oregonstate.edu). Variables used in models and their correlations are indicated in bold text.

A. Temperature

	MAT (normal)	Fall temp. (normal)	Winter temp. (normal)	Spring temp. (normal)	Summer temp. (normal)	MAT (2014-15)	Fall temp. (2014)	Winter temp. (2014-15)	Spring temp. (2015)
Fall temp. (normal)	0.97								
Winter temp. (normal)	0.69	0.83							
Spring temp. (normal)	0.86	0.73	0.23						
Summer temp. (normal)	0.72	0.54	0.01	0.94					
MAT (2014-2015)	0.71	0.68	0.62	0.44	0.48				
Fall temp. (2014)	0.70	0.75	0.82	0.30	0.26	0.95			
Winter temp. (2014-2015)	0.60	0.74	0.95	0.11	-0.04	0.71	0.90		
Spring temp. (2015)	0.28	0.05	-0.38	0.57	0.78	0.40	0.08	-0.35	
Summer temp. (2015)	**0.25**	0.05	-0.24	0.40	0.65	0.59	0.30	-0.14	0.93

B. Precipitation

	MAP (normal)	Fall precip. (normal)	Winter precip. (normal)	Spring precip. (normal)	Summer precip. (normal)	MAP (2014-15)	Fall precip. (2014)	Winter precip. (2014-15)	Spring precip. (2015)
Fall precip. (normal)	1.00								
Winter precip. (normal)	1.00	1.00							
Spring precip. (normal)	0.99	0.99	0.98						
Summer precip. (normal)	0.92	0.88	0.88	0.89					
MAP (2014-2015)	0.99	0.99	0.99	0.98	0.88				
Fall precip. (2014)	0.98	0.98	0.98	0.97	0.87	0.98			
Winter precip. (2014-2015)	0.98	0.98	0.98	0.98	0.86	1.00	0.98		
Spring precip. (2015)	0.96	0.97	0.96	0.98	0.82	0.98	0.95	0.99	
Summer precip. (2015)	**0.13**	0.10	0.11	0.05	0.38	0.11	0.04	0.09	0.01

C. Precipitation and temperature

	MAT (normal)	Fall temp. (normal)	Winter temp. (normal)	Spring temp. (normal)	Summer temp. (normal)	MAT (2014-15)	Fall temp. (2014)	Winter temp. (2014-15)	Spring temp. (2015)	**Summer temp. (2015)**
MAP (normal)	**0.20**	0.30	0.47	-0.05	-0.19	0.23	0.31	0.37	-0.18	**-0.08**
Fall precip. (normal)	0.23	0.34	0.53	-0.05	-0.21	0.22	0.33	0.41	-0.24	-0.14
Winter precip. (normal)	0.22	0.33	0.53	-0.06	-0.22	0.22	0.33	0.42	-0.25	-0.15
Spring precip. (normal)	0.30	0.37	0.52	0.03	-0.10	0.34	0.41	0.42	-0.09	0.03
Summer precip. (normal)	-0.08	-0.04	0.09	-0.17	-0.21	0.04	0.03	0.01	0.01	0.09
MAP (2014-2015)	0.25	0.34	0.51	-0.02	-0.16	0.28	0.37	0.43	-0.19	-0.07
Fall precip. (2014)	0.28	0.37	0.50	0.04	-0.13	0.21	0.30	0.38	-0.18	-0.12
Winter precip. (2014-2015)	0.28	0.37	0.53	0.00	-0.14	0.31	0.40	0.46	-0.18	-0.06
Spring precip. (2015)	0.34	0.43	0.61	0.02	-0.12	0.40	0.49	0.53	-0.17	-0.02
Summer precip. (2015)	**-0.84**	-0.80	-0.51	-0.79	-0.67	-0.50	-0.48	-0.39	-0.28	**-0.17**

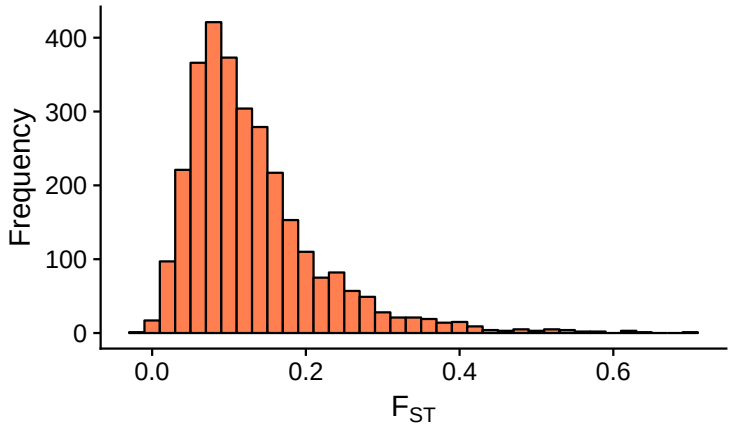

Figure A.1: Distribution of per-locus F_{ST} across 2982 SNPs from 32 populations of *Clarkia pulchella*.

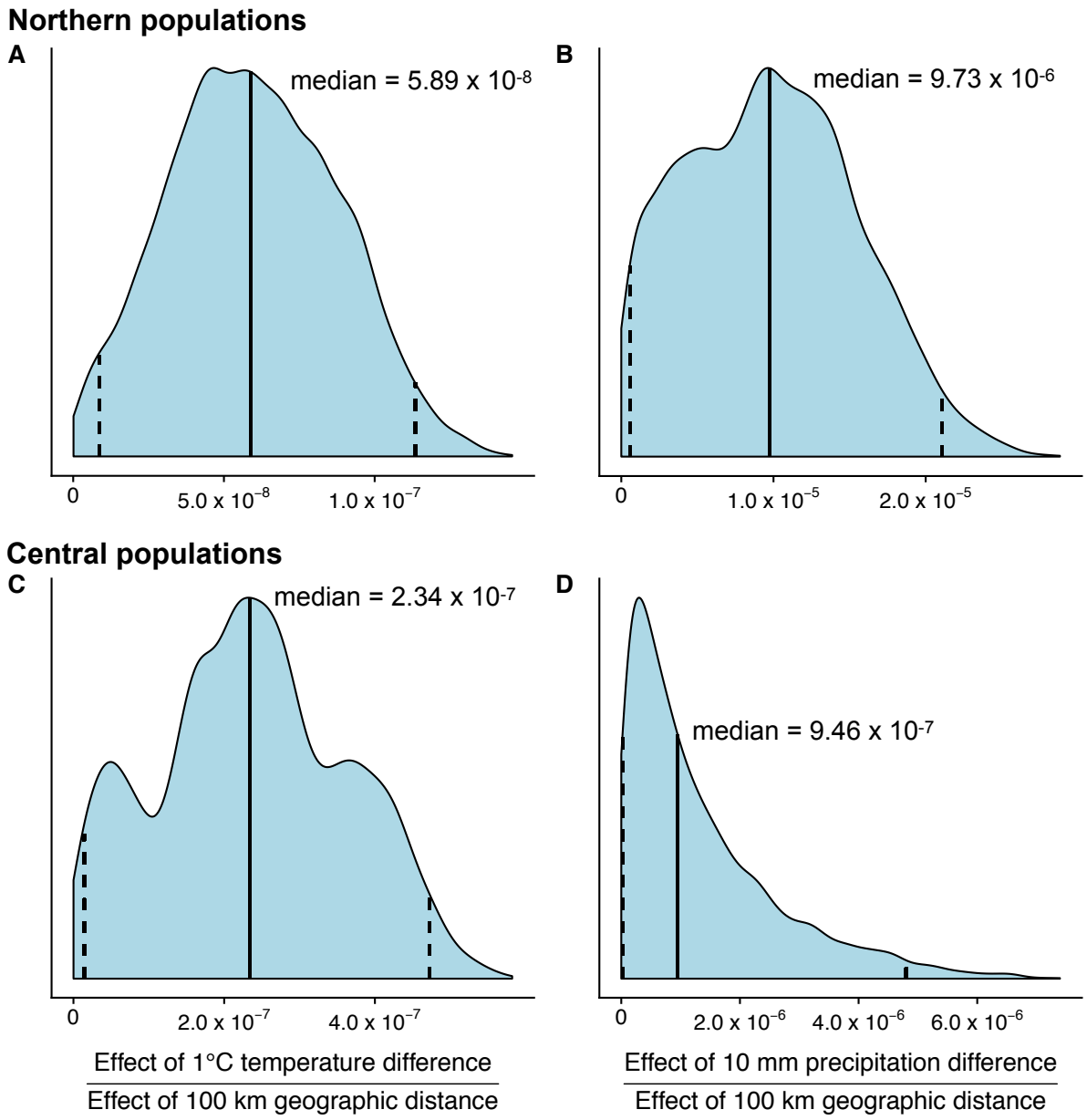

Figure A.2: Marginal posterior distributions, median values (solid lines) and 95% credible intervals (dashed lines) of effects of climate and geography on genetic differentiation of populations of *Clarkia pulchella* after a burn-in of 20%. **(A)** Temperature vs. geographic distance in northern populations only, **(B)** spring/summer precipitation vs. geographic distance in northern populations only, **(C)** Temperature vs. geographic distance in central populations only, **(D)** spring/summer precipitation vs. geographic distance in central populations only.

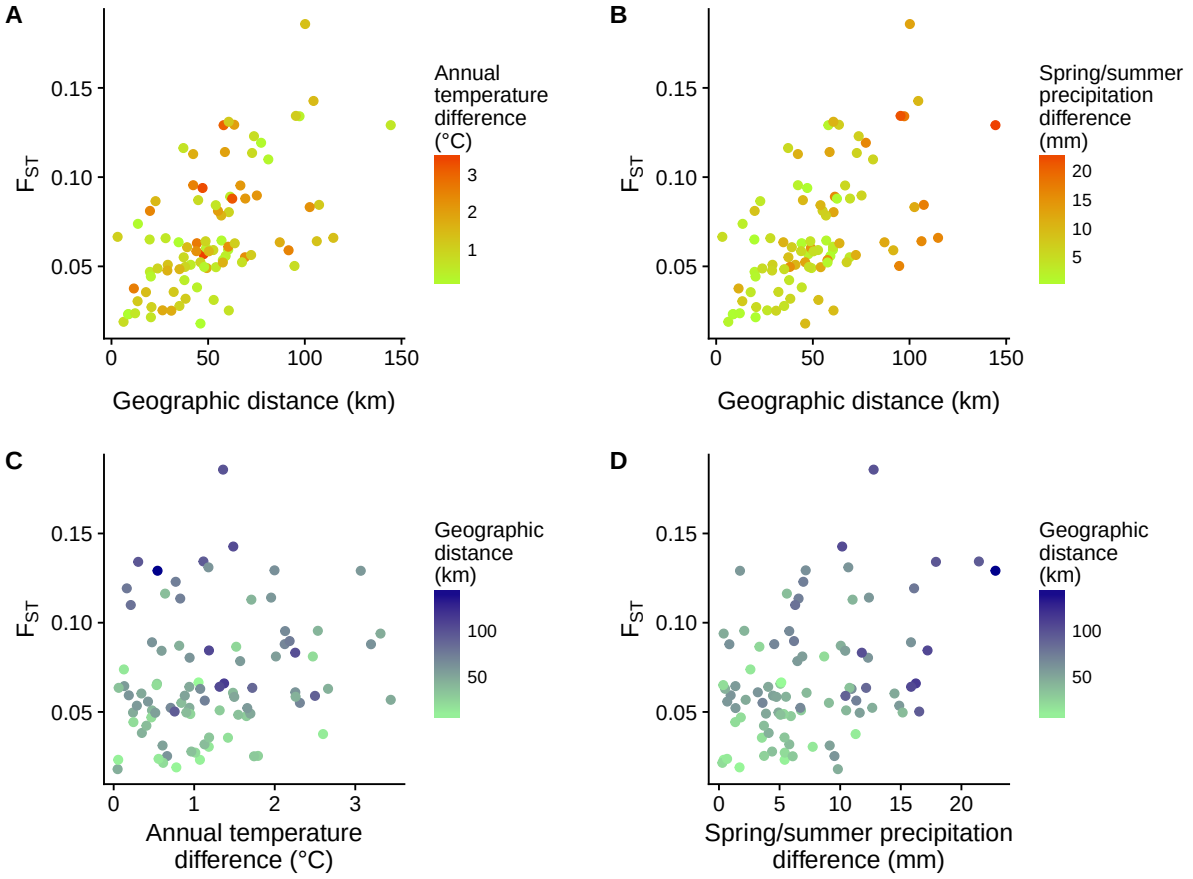

Figure A.3: Relationship between pairwise geographic distance (x-axis in **A** and **B**), temperature differences (colour in **A**) or precipitation differences (colour in **B**), and genetic differentiation (F_{ST}) among populations in the northern part of the geographic range of *Clarkia pulchella*. An alternative visualization is presented in **(C)** and **(D)**, in which climate differences are plotted on the x-axis and geographic distance is indicated with colour. Climate data are 1951-1980 averages from PRISM (PRISM Climate Group, 2017).

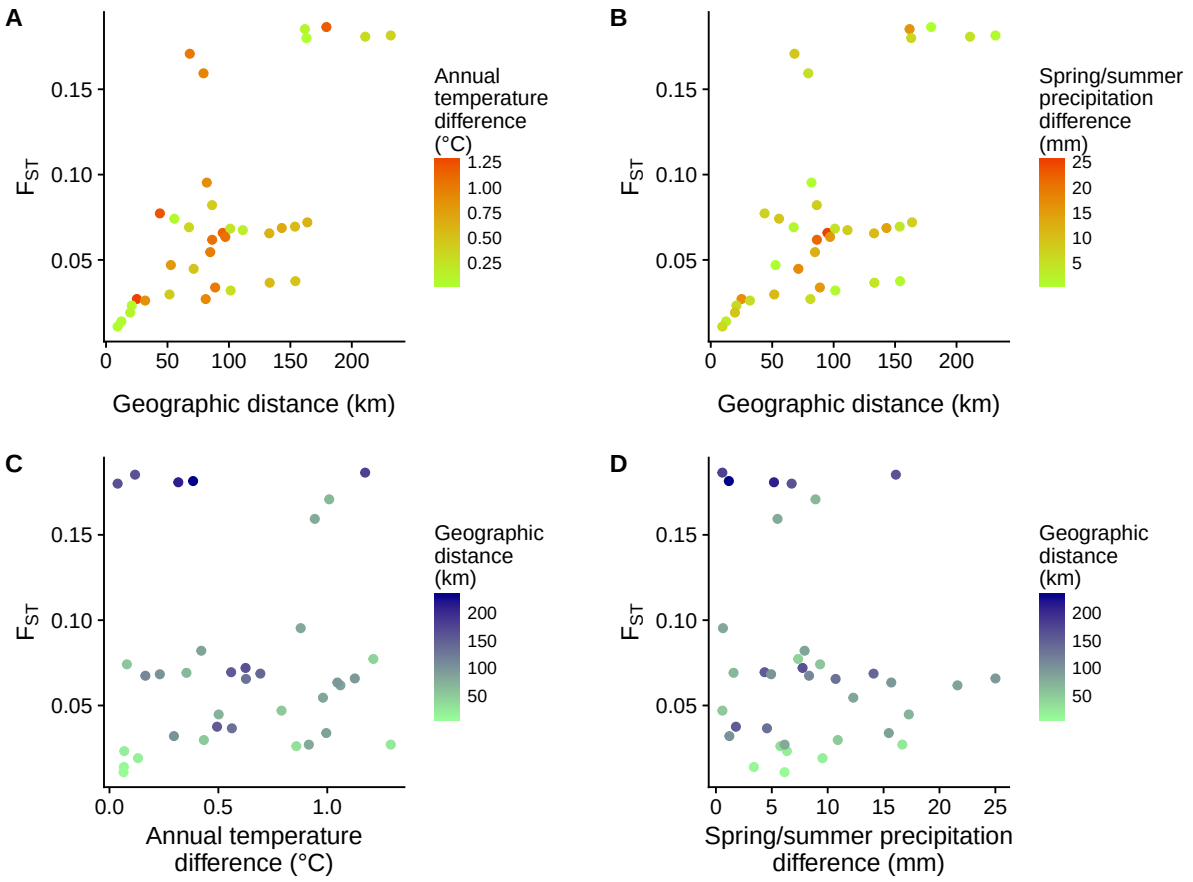

Figure A.4: Relationship between pairwise geographic distance (x-axis in **A** and **B**), temperature differences (colour in **A**) or precipitation differences (colour in **B**), and genetic differentiation (F_{ST}) among populations in the central part of the geographic range of *Clarkia pulchella*. An alternative visualization is presented in **(C)** and **(D)**, in which climate differences are plotted on the x-axis and geographic distance is indicated with colour. Climate data are 1951-1980 averages from PRISM (PRISM Climate Group, 2017).

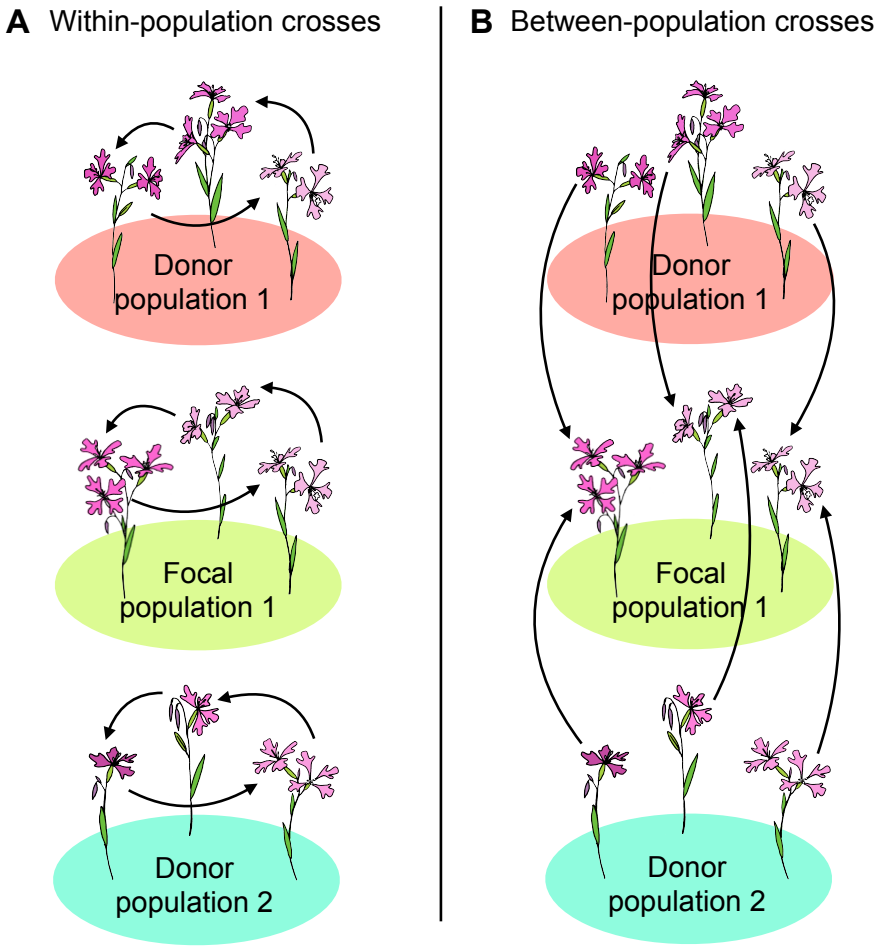

Figure A.5: Schematic of the greenhouse crossing design to generate seeds for our common gardens. **(A)** Within-population crosses: for each of 15 seed families (each represented by one plant) from each of 15 populations, plants were crossed in a "daisy chain" design, in which each plant was hand-pollinated using pollen from another individual of the same population. **(B)** Between-population crosses: we used pollen from 15 seed families (each represented by one plant) from each donor population to pollinate flowers on each of 15 seed families in each of the two focal populations. 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. We had greater replication of families per population during the greenhouse generation, but in this caption we refer to the numbers of families that were transplanted into the field.

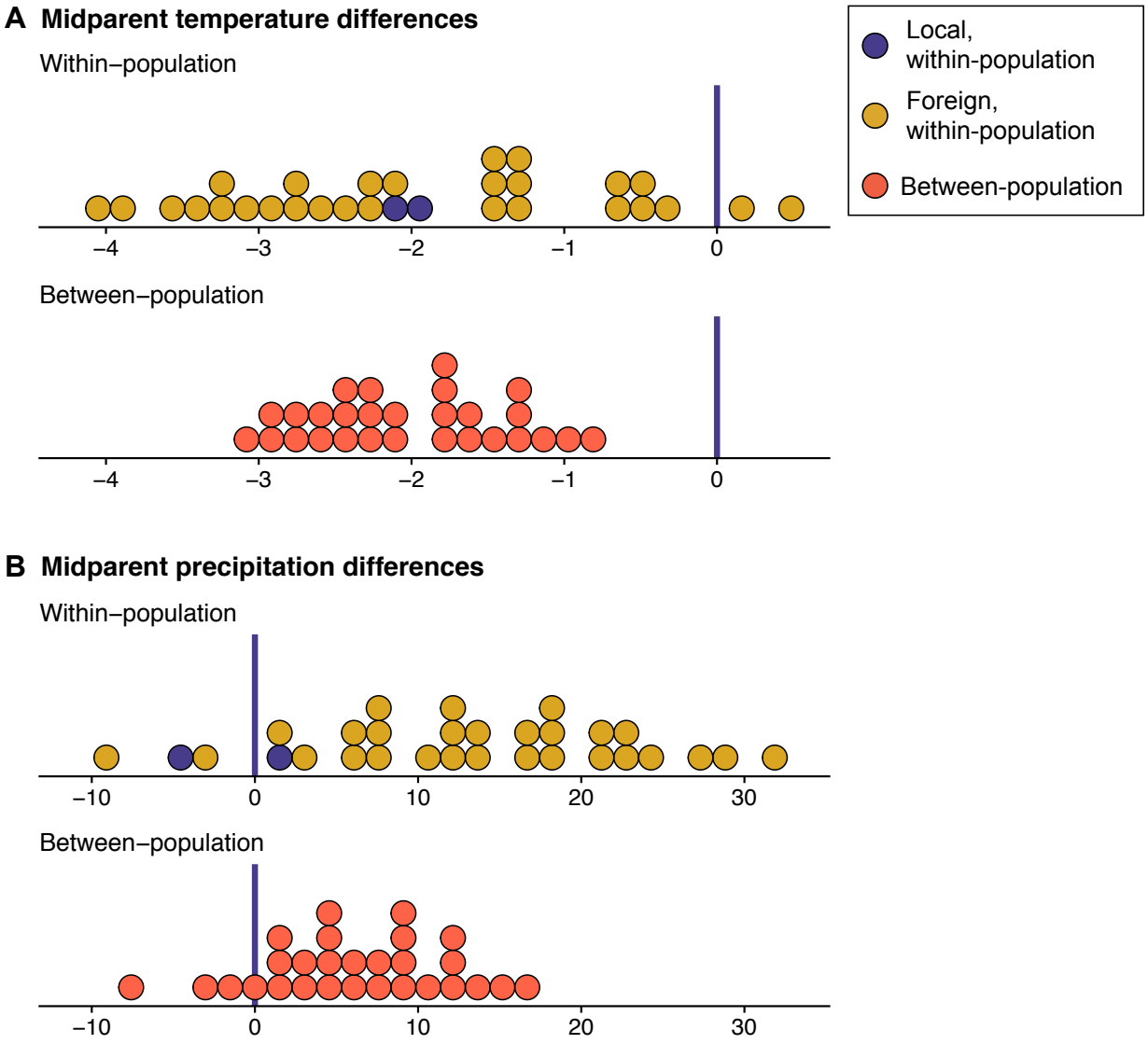

Figure A.6: Distribution of climate differences of within- vs. between-population crosses of *Clarkia pulchella* relative to conditions during the experiment. Each dot represents a combination of maternal population, paternal population, and transplant site. Dark blue dots are within-population crosses of focal populations transplanted into their home sites. Gold dots are within-population crosses from donor populations planted into each of the two gardens, as well as the focal populations planted into each other's sites. Red dots are between-population crosses. Vertical blue bars are placed at zero, indicating where populations would be perfectly matched to the temperature or precipitation conditions during the experiment. **(A)** Distribution of the differences between the average temperature in the home sites of parental populations and conditions during the experiment. Focal populations are intermediate in temperature relative to other populations in the experiment; this results in similar average differences in temperature in between-population crosses and within-population crosses. **(B)** Distribution of differences between the average precipitation in the home sites of parental populations and conditions during the experiment. Focal populations are among the driest in the experiment; this results in smaller average differences in precipitation in between-population crosses compared to within-population crosses. Note that figures in the main text use absolute temperature differences: the absolute value of the midparent differences as they are plotted in this figure.