DISSERTATION

THE POTENTIAL FOR ADAPTATION IN THE MODEL PLANT, *ARABIDOPSIS*, AND ITS CLOSE RELATIVE, *BOECHERA*

Submitted by

John Thomson Lovell

Graduate Degree Program in Ecology

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Spring 2014

Doctoral Committee:

Advisor: John K. McKay

Cameron K. Ghalambor
William L. Bauerle
Amy L. Angert

ABSTRACT

THE POTENTIAL FOR ADAPTATION IN THE MODEL PLANT, *ARABIDOPSIS*, AND ITS CLOSE RELATIVE, *BOECHERA*

Populations of a species are found across diverse environments. Frequently, evolutionary responses to varying natural selection pressures across environments cause adaptive differentiation among populations. For plants with limited dispersal capability, adaptation is the primary way populations can persist through changing environments and climates. Therefore, factors that constrain adaptation can directly affect the conservation status and future distributions of species and populations.

Contemporary adaptations, via either selection on standing genetic variation or new beneficial mutations that sweep to fixation, have been observed across many taxonomic groups. However, these events of fast, streamlined adaptive evolution may be rare. Instead, adaptation is often constrained by both ecological and genetic forces. Determining the mechanisms and ecological manifestations of adaptive constraint remains a major challenge for evolutionary biologists, conservation biologists and crop breeders. The primary goal of my dissertation research is to address this challenge. The research projects described herein documented the extent and causes of evolutionary constraint by accomplishing three separate goals: (1) to document genes underlying drought adaptation in the model plant, *Arabidopsis thaliana*, to infer the adaptive effects of pleiotropy, (2) to determine the mechanisms constraining adaptation in rare species relative to widespread congeners, and (3) to assess the degree of adaptive differentiation across the genomic loci and populations.

By measuring quantitative and molecular diversity across several species, I determined the relative potential of adaptation at the gene, population, and species levels of organization. In chapter 2 I studied adaptation at the gene-level and demonstrated that the *Arabidopsis thaliana* gene *FRI* (*FRIGIDA*) exhibits adaptive pleiotropy. Through simultaneous genetic effects on many traits, variation at the single gene *FRI* produces trait correlations along an axis that is in line with the vector of selection. In this case, sequence polymorphism at *FRI* caused phenotypic co-variance of water-use-efficiency (WUE), relative growth rate and timing of flowering. This genetic correlation coincided with a well-described adaptive correlation found in natural and agricultural systems.

In chapter 3, I studied the processes that cause range size diversity across species, by comparing population ecology and genetics of species with broadly divergent range sizes. I assayed the heritability of potentially adaptive traits and other quantitative genetic statistics from multiple rare and widespread species and found that rare species lack heritable genetic variation and physiological plasticity. Combined, these factors place rare species at increased risk of extinction across changing environmental conditions.

In chapter 4, I studied adaptation at the population level. I examined how environmental variation impacted genomic structure, selection pressures and local adaptation in *Boechera spatifolia* (Brassicaceae), a species that contains both sexual and asexual (apomictic) individuals. I found that, despite occupying sympatric sites, apomictic lineages are both phenotypically and genetically distinct from sexuals. Additionally, while sexual populations formed strong clines (both genomic and physiological) along latitude and elevation gradients, apomicts showed no such signature of local adaptation.

ACKNOWLEDGMENTS

To accomplish this dissertation, I received support from many people. First, I would like to thank John McKay, my dissertation advisor, who intellectually challenged me and helped guide my research. This project would have not been possible without support from my other scientific mentors and collaborators, including Tim Sharbel, Jim Richards, Thomas Juenger, Saunak Sen, Patrick Alexander, James Beck, Donovan Bailey, David Siemens, Marcus Koch, Jack Mullen, Julius Mojica, and my committee members, Bill Bauerle, Cameron Ghalambor and Amy Angert. I owe a great deal to Kelsi Grogan, who as both an undergraduate and lab manager assisted in collection of data and was a great partner in research. Shea Roberts, Rico Moore and Chris Klingbeil assisted in data collection in the laboratory. I would also like to thank my fellow graduate students, especially Seema Sheth, Dave Hoover, Jamie Fuller, Derek Shook, Marco Pellino and Christa Fettig, who provided critical feedback and support throughout my graduate tenure. Andrew Norton's advice greatly improved my teaching skills and demonstrated to me the rewards of teaching. I am particularly grateful to my former mentors, especially Shane Heschel and Eric Menges, who inspired me to undertake scientific research as a profession.

I want to thank my friends and family for their support and in particular, my parents, Julie and Chris Lovell, whose encouragement gave me the opportunity to succeed. Lastly, I would like to express my deepest love and gratitude to my wife, Ashley. Her advice and critical feedback greatly improved all aspects of my dissertation and helped to make my graduate research a joy.

The research conducted herein was funded by NSF DEB-1022196 to John McKay, DFG SH337/7-1, a μMorph Training Fellowship, and a Myrna P. Steinkamp Grant.

TABLE OF CONTENTS

Chapter 1

INTRODUCTION

Background

Adaptation to environmental conditions in plants

> "In all species complexes that cover climatically different areas, the phenomenon of regional differentiation stands as a fundamental principle."
>
> -Clausen *et al.* (1940, p. 411)

In their classic study of adaptation in the species complex of *Achillea*, Clausen, Keck and Heisey determined that populations of plants found along an environmental gradient grow in dramatically different ways when cultivated in a common garden (Clausen *et al.* 1940). These genetically-based trait differences lead to greater fitness in local habitats, a phenomenon called "local adaptation". After >80 years of subsequent experimentation, it is commonly accepted that local adaptation is nearly ubiquitous among organisms, and especially higher plants (Kawecki & Ebert 2004).

Plants, as sessile organisms, cannot migrate to overcome stressful conditions, but instead, must cope with environmental stress through physiological adaptation (Agren & Schemske 2012) or environmental sensitivity (plasticity) (Pigliucci *et al.* 1999; Price *et al.* 2003). Since climactic conditions vary at small spatial and temporal scales, local adaptation among populations across environments is expected to be strong. Plants are an ideal experimental system because of their experimental utility and exposure to diverse and extreme environmental conditions. Classic experiments on the genetic basis of adaptation have advanced the study of

evolution by comparing plant populations across diverse environments, including, metal accumulation in the soil (Antonovics 1976), elevation (Clausen *et al.* 1940) and many other environmental gradients (Agren & Schemske 2012; Hall *et al.* 2010; Latta 2009; Leinonen *et al.* 2009; Linhart & Grant 1996; McKay *et al.* 2001; Nagy 1997).

Adaptation to diverse environmental conditions has permitted the range expansion of invasive species (Lavergne & Molofsky 2007; Leger *et al.* 2009; Colautti & Barrett 2013) and the maintenance of diverse geographic ranges in a variety of species (Agren & Schemske 2012; Angert & Schemske 2005; Heschel *et al.* 2002; Kawecki & Ebert 2004; Latta 2009; Leinonen *et al.* 2009; McKay *et al.* 2001). The historical events of successful adaptation in these species were driven by selection (Colautti & Barrett 2013) and permitted by the underlying raw material for evolution: genetic diversity (Davis *et al.* 2005). Current global environmental shifts, not only climactic but also local landscape use, are causing a major shift in the environmental conditions experienced by plant populations (Palumbi 2001). As their ancestors successfully adapted to the current local conditions, contemporary plant populations must now adapt to continuously changing environments (Ehrlich & Pringle 2008; Etterson & Shaw 2001). Failing to do so likely means the extinction of local populations. In rare species, with few extant populations, successful contemporary adaptation is critical to the future demographic persistence of the species.

Factors constraining adaptation in nature

Examples of adaptation to novel or changing environmental conditions abound in the literature. Contemporary adaptation, via either selection on standing genetic variation or new beneficial mutations that sweep to fixation, has been observed across many taxonomic groups.

This pattern has been documented in crops (Lukens & Doebley 2001; Wright *et al.* 2005), model systems (Bersaglieri *et al.* 2004; Voight *et al.* 2006) and many wild species (Bodbyl Roels & Kelly 2011; Kingsolver *et al.* 2001; Storz *et al.* 2009). Given these examples, it is possible that evolution proceeds unimpeded towards adaptive optima. However, a growing body of theory indicates that these events of fast, streamlined adaptive evolution are rare. Instead, it is generally accepted that adaptation is constrained by both ecological and genetic forces (Arnold 1992; Cheverud 1988; Cheverud 1984; Etterson & Shaw 2001; Mitchell-Olds 1996; van Kleunen & Fischer 2005). Determining the mechanisms and ecological manifestations of adaptive constraint remains a major challenge for both evolutionary biologists and crop breeders.

Natural selection acts proximately upon fitness (Lande 1976), where genotypes with elevated intrinsic population growth rates are favored over less prolific genotypes. However, fitness, the ultimate complex trait, is affected by many other traits (e.g. phenology, growth rate, disease resistance, etc.), thousands of genetic loci, and their interactions with the multitude of environmental conditions experienced by the individual (Falconer & Mackay 1996; Houle 1991; Lande 1984; Lande & Arnold 1983). As selection works towards optimizing fitness, the potential for a population to experience an evolutionary response to selection exists in the connection between fitness and its component phenotypes- the genetic variance-covariance matrix (Agrawal & Stinchcombe 2009; Cheverud 1988; Cheverud 1984; Houle 1991, 1992; Lande & Arnold 1983). In the simplest case, where all traits are uncorrelated and genetic loci segregate independently, the potential for adaptation (response to selection, R) will be affected by two factors: 1) the amount of additive genetic variation present within the population (narrow-sense heritability, h^2), and 2) the strength of selection (s) (Falconer & Mackay 1996).

"$Breeder's\ Equation$": $R = h^2 s$

In this case, the potential for adaptive evolution, given a constant strength of selection, is due to the heritability of phenotypes correlated with fitness- adaptive traits. Therefore, the adaptive potential of a population is affected by the genetic diversity and heritability of putatively adaptive traits.

Even if heritability is high and sufficient genetic diversity underlies adaptive phenotypes, genetic correlations can constrain adaptation (Houle 1991, 1992; Mitchell-Olds 1996; Rose 1982; Zhang & Hill 2003). Genetically based trait correlations can be caused by functional relationships among traits (e.g. individuals with larger leaves may also have increased growth rate) or linkage among loci that underlie physiologically distinct traits. In the simplest case, two structurally unrelated phenotypes directly affect fitness and are each affected by variation at a single locus. Recombination among loci disrupts genetic correlations, allowing selection to act efficiently and independently on each adaptive phenotype. However, if recombination is suppressed or the traits are structurally correlated, selection upon one trait will necessarily affect the other phenotype. These genetic correlations are often antagonistic due to the interfering effects of correlated loci.

While many other factors may inhibit adaptation, here I focus on genetic diversity, which is the source of all raw material for evolution. The broadest circumspection of genetic diversity includes 1) the extent of multivariate molecular and quantitative trait variance, 2) the degree to which these factors are correlated and 3) the potential for those correlations to break down. As such, I discuss three components of genetic diversity independently: chapter 1) antagonistic and

potentially adaptive correlations, chapter 2) heritability and total diversity, and chapter 3) the effects of recombination and asexual reproduction.

Study system

Due to its generally small stature, diverse agricultural utility and advanced genomic tools, the mustard (Brassicaceae) plant family is perhaps the most experimentally tractable of plant taxa. Most major groups within the Brassicaceae have one or more species with completely sequenced genomes (Haudry *et al.* 2013; Hu *et al.* 2011; Schranz *et al.* 2006; Schranz *et al.* 2007a). Furthermore, *Arabidopsis thaliana*, which has the first and most complete genome of all plants, shares broad scale sequence homology with other species of the Camilineae and *Boechera*e tribes (Schranz *et al.* 2007c; Song *et al.* 2009; Windsor *et al.* 2006). This common ancestry permits genomic inference of wild species by simply comparing sequences to the *A. thaliana* reference.

Boechera is quickly emerging as a model system to study speciation, adaptation and development (Lovell 2011; Rushworth *et al.* 2011). *Boechera*'s greatest advantage is its close phylogenetic relationship with *Arabidopsis* and the resulting high degree of sequence and physiological conservation between species (Schranz *et al.* 2007b; Windsor *et al.* 2006). This recent common ancestry with the premier system for functional genomics in plants provides a wealth of *a priori* hypotheses about genotype-to-phenotype processes and their potential role in adaptation.

In addition to being characterized by physiological and genomic resources, many *Boechera* species possess the ability to produce seeds asexually, which is known as apomixis

(Al-Shehbaz 2003; Dobeš *et al.* 2007; Lovell *et al.* 2013; Rushworth *et al.* 2011; Schranz *et al.* 2005; Sharbel *et al.* 2009). Apomictic species are derived from sexual species (Grossniklaus *et al.* 2001). Three independent developmental steps must be acquired for a sexual plant to produce seeds apomictically (Nogler 1984): 1) the formation of an unreduced megaspore (apomeiosis), 2) the subsequent development of an embryo in the absence of fertilization (parthenogenesis), and 3) fertilization of the binucleate central cell to form a functional endosperm (pseudogamy). The apomeiotically-derived embryo thus receives its entire genome through the female line. Polyploidy is ubiquitously associated with apomixis among most taxa; however, *Boechera* is characterized by naturally occurring diploid sexual and diploid apomictic forms (Aliyu *et al.* 2010). Furthermore, the relative rate of diploid apomixis (facultative, obligate or rare) can be analyzed through advanced methodologies developed specifically for this genus. Thus, the differences between and relative prevalence of apomictic and sexual reproduction can be compared without the confounding effects of polyploidy.

Summary of dissertation chapters

Using three experiments, I directly assess the adaptive causes and consequences of three potential forces that constrain adaptation in nature.

Chapter 1: Genetic trait correlations

Adaptation to drought and other environmental conditions in plants provides an excellent system to study the genetics of and constraints to multivariate selection. In natural and

agricultural systems, annual plants can be adapted to local drought conditions by either growing and reproducing before the onset of drought (drought escape) (Araus *et al.* 2002; Franks *et al.* 2007; Meyre *et al.* 2001; Sherrard & Maherali 2006) or by delaying reproduction, increasing water use efficiency (WUE) and conserving resources (dehydration avoidance) (Donovan *et al.* 2007; McKay *et al.* 2003; Rosenthal *et al.* 2010). Several studies have also suggested that single genes with multivariate phenotypic effects (pleiotropy) may also affect this correlation (Geber & Dawson 1990; McKay *et al.* 2003; Menendez & Hall 1995). Through laboratory and greenhouse experimentation I demonstrate that the *Arabidopsis thaliana* gene *FRI* (*FRIGIDA)* exhibits adaptive pleiotropy, conferring trait correlations along an axis that promotes adaptation and positive assortative mating simultaneously. This conclusion is in contrast to other studies that have both theoretically (Otto 2004) and empirically implicated pleiotropy as a process that promotes mal-adaptive evolution. I find that the ancestral dehydration-avoidant phenotype was recovered when genotypes with null *FRI* alleles were transformed with over-expressed alleles. This indicates that variation at *FRI* can induce a shift between stress adaptation strategies. I also found patterns of increased population structure between allele classes, which suggests that *FRI* may promote positive assortative mating, thus increasing the potential for adaptation by reducing mal-adaptive gene flow. Combined, *FRI*'s multivariate phenotypic affects and among-allele reproduction isolation through phenology simultaneously improve the potential for adaptation to drought and the likelihood of positive assortative mating.

Chapter 2: Rarity

Species can inhabit expansive or extremely localized ranges. The size and environmental diversity of a species' range reflects historical patterns, as well as current limitations to dispersal and/or physiological limits to environmental stresses (Brown *et al.* 1996). If adaptive evolution is occurring at the range margin, a species' range should expand like "tree rings", gradually filling all available ecological niches (Mayr 1963). All species occupy a subset of the available habitat, some a very narrow subset (rare) and others a broader range of environments. Therefore, evolution at the range margin must be constrained by genetic or ecological factors (Chevin & Lande 2011; Kawecki 2008). Furthermore, those constraining factors are likely to be stronger in rare species than widespread relatives because of a higher degree of habitat restriction. Genetic factors constraining biogeographic expansion may include genetic diversity, trait heritability and phenotypic plasticity (Angert 2006; Angert *et al.* 2008; Chevin & Lande 2011; Kawecki 2008). In this study I compared the extent of these three factors between two rare and two widespread species. I found that rare species had relatively similar levels of molecular genetic diversity, but highly depressed heritability and phenotypic plasticity. Combined these factors placed rare species in double jeopardy, limiting both the breadth of conditions tolerated by an individual and the potential for adaptation to novel environments.

Chapter 3: Asexuality and subsequent decreased recombination

I examined how environmental variation impacted genomic structure, selection pressures and local adaptation in *Boechera spatifolia*. There is substantial evidence that apomixis can

promote greater fitness and ecological niche breadth (van Dijk 2003). For example all North American *Hieracum aurantiacum* are apomictically derived (Loomis & Fishman 2009). However, most researchers find that apomixis reduces (or eliminates) recombination (Grossniklaus *et al.* 2001; Ozias-Akins & van Dijk 2007; Richards 2003), which causes interference among loci (Barton 2010; Hill & Robertson 1966), correlations among traits and reduced responses to selection (Barton & Charlesworth 1998; Burt 2000). As such, apomixis is generally thought to constrain adaptive evolution. In this experiment, I empirically tested this hypothesis by comparing patterns of genetic variation along environmental gradients in apomictic and sexual lineages. I found that, despite occupying sympatric sites, apomictic lineages were both phenotypically and phylogenetically distinct from sexuals. Additionally, while sexual populations formed strong clines (both genomic and physiological) along latitude and elevation gradients, apomicts showed no such signature of local adaptation. This result was consistent with one hypothesis for the evolutionary success of sexual reproduction despite the 2-fold cost of sex: asexual lineages may have poor responses to selection.

References

Agrawal AF, Stinchcombe JR (2009) How much do genetic covariances alter the rate of adaptation? Proceedings of the Royal Society of London. Series B: Biological Sciences 276, 1183-1191.

Agren J, Schemske DW (2012) Reciprocal transplants demonstrate strong adaptive differentiation of the model organism *Arabidopsis thaliana* in its native range. New Phytologist 194, 1112-1122.

Al-Shehbaz IA (2003) Transfer of most North American species of *Arabis* to *Boechera* (Brassicaceae). *Novon*, 381-391.

Aliyu OM, Schranz ME, Sharbel TF (2010) Quantitative variation for apomictic reproduction in the genus *Boechera* (Brassicaceae). American Journal of Botany 97, 1719-1731.

Angert AL (2006) Demography of central and marginal populations of monkeyflowers (*Mimulus cardinalis* and *M. lewisii*). Ecology 87, 2014-2025.

Angert AL, Bradshaw HD, Jr., Schemske DW (2008) Using experimental evolution to investigate geographic range limits in monkeyflowers. Evolution 62, 2660-2675.

Angert AL, Schemske DW (2005) The Evolution of Species' Distributions: Reciprocal Transplants across the Elevation Ranges *of Mimulus Cardinalis* and *M. Lew*isii. Evolution 59, 1671.

Antonovics J (1976) The nature of limits to natural selection. Annals of the Missouri Botanical Garden, 224-247.

Araus JL, Slafer GA, Reynolds MP, Royo C (2002) Plant breeding and drought in C-3 cereals: What should we breed for? Annals of Botany 89, 925-940.

Arnold SJ (1992) Constraints on phenotypic evolution. American Naturalist, S85-S107.

Barton NH (2010) Mutation and the evolution of recombination. Philosophical Transactions of the Royal Society of London- B Biological Sciences 365, 1281-1294.

Barton NH, Charlesworth B (1998) Why sex and recombination? Science 281, 1986-1990.

Bersaglieri T, Sabeti PC, Patterson N, *et al.* (2004) Genetic signatures of strong recent positive selection at the lactase gene. The American Journal of Human Genetics 74, 1111-1120.

Bodbyl Roels SA, Kelly JK (2011) Rapid evolution caused by pollinator loss in *Mimulus guttatus*. Evolution 65, 2541-2552.

Brown JH, Stevens GC, Kaufman DM (1996) The geographic range: Size, shape, boundaries, and internal structure. Annual Review of Ecology and Systematics 27, 597-623.

Burt A (2000) Perspective: sex, recombination, and the efficacy of selection—was Weismann right? Evolution 54, 337-351.

Cheverud J (1988) The evolution of genetic correlation and developmental constraints. In: Population genetics and evolution, pp. 94-101. Springer.

Cheverud JM (1984) Quantitative genetics and developmental constraints on evolution by selection. Journal of Theoretical Biology 110, 155-171.

Chevin LM, Lande R (2011) Adaptation to marginal habitats by evolution of increased phenotypic plasticity. Journal of Evolutionary Biology 24, 1462-1476.

Clausen JC, Keck DD, Hiesey WM (1940) *Effect of varied environments on western North American plants* Carnegie Institution.

Colautti RI, and Barrett SC (2013). Rapid adaptation to climate facilitates range expansion of an invasive plant. Science 342, 364-366.

Davis MB, Shaw RG, Etterson JR (2005) Evolutionary responses to changing climate. Ecology 86, 1704-1714.

Dobeš C, Sharbel TF, Koch M (2007) Towards understanding the dynamics of hybridization and apomixis in the evolution of the genus *Boechera* (Brassicaceae). Systematics and Biodiversity 5, 321-331.

Donovan LA, Dudley SA, Rosenthal DM, Ludwig F (2007) Phenotypic selection on leaf water use efficiency and related ecophysiological traits for natural populations of desert sunflowers. Oecologia 152, 13-25.

Ehrlich PR, Pringle RM (2008) Colloquium paper: where does biodiversity go from here? A grim business-as-usual forecast and a hopeful portfolio of partial solutions. Proceedings of the National Academy of Sciences of the United States of America 105 Suppl 1, 11579-11586.

Etterson JR, Shaw RG (2001) Constraint to adaptive evolution in response to global warming. Science 294, 151-154.

Falconer DS, Mackay TFC (1996) *Introduction to Quantitative Genetics,4th Edn.* Longman, Harlow.

Franks SJ, Sim S, Weis AE (2007) Rapid evolution of flowering time by an annual plant in response to a climate fluctuation. Proceedings of the National Academy of Sciences of the United States of America 104, 1278-1282.

Geber MA, Dawson TE (1990) Genetic-variation in and covariation between leaf gas-exchange, morphology, and development in *Polygonum arenastrum*, an annual plant. Oecologia 85, 153-158.

Grossniklaus U, Nogler GA, van Dijk PJ (2001) How to avoid sex: the genetic control of gametophytic apomixis. Plant Cell 13, 1491-1498.

Hall MC, Lowry DB, Willis JH (2010) Is local adaptation in *Mimulus guttatus* caused by trade-offs at individual loci? Molecular Ecology 19, 2739-2753.

Haudry A, Platts AE, Vello E, *et al.* (2013) An atlas of over 90,000 conserved noncoding sequences provides insight into crucifer regulatory regions. Nature Genetics 45, 891-898.

Heschel MS, Donohue K, Hausmann N, Schmitt J (2002) Population differentiation and natural selection for water-use efficiency *in Impatiens capensis* (Balsaminaceae). International Journal of Plant Sciences 163, 907-912.

Hill W, Robertson A (1966) The effect of linkage on limits to artificial selection. Genetical research 8, 269-294.

Houle D (1991) Genetic covariance of fitness correlates: what genetic correlations are made of and why it matters. Evolution, 630-648.

Houle D (1992) Comparing evolvability and variability of quantitative traits. Genetics 130, 195-204.

Hu TT, Pattyn P, Bakker EG, *et al.* (2011) The *Arabidopsis lyrata* genome sequence and the basis of rapid genome size change. Nature Genetics 43, 476-481.

Kawecki TJ (2008) Adaptation to Marginal Habitats. Annual Review of Ecology, Evolution, and Systematics 39, 321-342.

Kawecki TJ, Ebert D (2004) Conceptual issues in local adaptation. Ecology Letters 7, 1225-1241.

Kingsolver JG, Hoekstra HE, Hoekstra JM, *et al.* (2001) The strength of phenotypic selection in natural populations. The American Naturalist 157, 245-261.

Lande R (1976) Natural selection and random genetic drift in phenotypic evolution. Evolution, 314-334.

Lande R (1984) The genetic correlation between characters maintained by selection, linkage and inbreeding. Genetical Research 44, 309-320.

Lande R, Arnold SJ (1983) The measurement of selection on correlated characters. Evolution, 1210-1226.

Latta RG (2009) Testing for local adaptation in *Avena barbata*: a classic example of ecotypic divergence. Molecular Ecology 18, 3781-3791.

Lavergne S, Molofsky J (2007) Increased genetic variation and evolutionary potential drive the success of an invasive grass. Proceedings of the National Academy of Sciences of the United States of America 104, 3883-3888.

Leger EA, Espeland EK, Merrill KR, Meyer SE (2009) Genetic variation and local adaptation at a cheatgrass (*Bromus tectorum*) invasion edge in western Nevada. Molecular Ecology 18, 4366-4379.

Leinonen PH, Sandring S, Quilot B, *et al.* (2009) Local adaptation in European populations of *Arabidopsis lyrata* (Brassicaceae). Am J Bot 96, 1129-1137.

Linhart YB, Grant MC (1996) Evolutionary significance of local genetic differentiation in plants. Annual Review of Ecology and Systematics, 237-277.

Loomis ES, Fishman L (2009) A Continent-Wide Clone: Population Genetic Variation of the Invasive Plant *Hieracium aurantiacum* (Orange Hawkweed; Asteraceae) in North America. International journal of plant sciences 170, 759-765.

Lovell J, Aliyu O, Mau M, *et al.* (2013) On the origin and evolution of apomixis in *Boechera*. Plant reproduction 26, 309-315.

Lovell JT (2011) Meeting Review: *Boechera* Summit 2011. *Molecular Ecology* 20 4840-4842

Lukens L, Doebley J (2001) Molecular evolution of the teosinte branched gene among maize and related grasses. Molecular biology and evol*ution* 18, 627-638.

Mayr E (1963) *Animal species and evolution*. Harvard University Press.

McKay JK, Bishop JG, Lin J-Z, *et al.* (2001) Local adaptation across a climatic gradient despite small effective population size in the rare sapphire rockcress. Proceedings of the Royal Society of London. Series B: Biological Sciences 268, 1715-1721.

McKay JK, Richards JH, Mitchell-Olds T (2003) Genetics of drought adaptation in *Arabidopsis thaliana*: I. Pleiotropy contributes to genetic correlations among ecological traits. Molecular Ecology 12, 1137-1151.

Menendez CM, Hall AE (1995) Heritability of carbon-isotope discrimination and correlations with earliness in cowpea. Crop Science 35, 673-678.

Meyre D, Leonardi A, Brisson G, Vartanian N (2001) Drought-adaptive mechanisms involved in the escape/tolerance strategies of *Arabidopsis* Landsberg erecta and Columbia ecotypes and their F1 reciprocal progeny. Journal of Plant Physiology 158, 1145-1152.

Mitchell-Olds T (1996) Pleiotropy causes long-term genetic constraints on life-history evolution in *Brassica rapa*. Evolution, 1849-1858.

Nagy ES (1997) Selection for native characters in hybrids between two locally adapted plant subspecies. Evolution, 1469-1480.

Nogler GA (1984) *Gametophytic apomixis* Springer.

Otto SP (2004) Two steps forward, one step back: the pleiotropic effects of favoured alleles. Proceedings of the Royal Society of London Series B-Biological Sciences 271, 705-714.

Ozias-Akins P, van Dijk PJ (2007) Mendelian genetics of apomixis in plants. Annu Rev Genet 41, 509-537.

Palumbi, SR (2001). Humans as the world's greatest evolutionary force. Science 293(5536): 1786-1790.

Pigliucci M, Cammell K, Schmitt J (1999) Evolution of phenotypic plasticity a comparative approach in the phylogenetic neighbourhood of *Arabidopsis thaliana*. Journal of Evolutionary Biology 12, 779-791.

Price TD, Qvarnstrom A, Irwin DE (2003) The role of phenotypic plasticity in driving genetic evolution. Proceedings of the Royal Society B-Biological Sciences 270, 1433-1440.

Richards AJ (2003) Apomixis in flowering plants: an overview. Philosophical Transactions of the Royal Society of London- B Biological Sciences 358, 1085-1093.

Rose MR (1982) Antagonistic pleiotropy, dominance, and genetic variation. Heredity 48, 63-78.

Rosenthal DM, Stiller V, Sperry JS, Donovan LA (2010) Contrasting drought tolerance strategies in two desert annuals of hybrid origin. Journal of Experimental Botany 61, 2769-2778.

Rushworth C, Song BH, Lee CR, Mitchell-Olds T (2011) *Boechera*, a model system for ecological genomics. Molecular Ecology 20(23): 4843-4857

Schranz ME, Dobeš C, Koch MA, Mitchell-Olds T (2005) Sexual reproduction, hybridization, apomixis, and polyploidization in the genus *Boechera* (Brassicaceae). American Journal of Botany 92, 1797-1810.

Schranz ME, Lysak MA, Mitchell-Olds T (2006) The ABC's of comparative genomics in the Brassicaceae: building blocks of crucifer genomes. Trends in Plant Science 11, 535-542.

Schranz ME, Song B-H, Windsor AJ, Mitchell-Olds T (2007a) Comparative genomics in the Brassicaceae: a family-wide perspective. Current opinion in plant biology 10, 168-175.

Schranz ME, Windsor AJ, Song BH, Lawton-Rauh A, Mitchell-Olds T (2007b) Comparative genetic mapping in *Boechera stricta*, a close relative of *Arabidopsis*. Plant Physiology 144, 286-298.

Sharbel TF, Voigt ML, Corral JM, *et al.* (2009) Molecular signatures of apomictic and sexual ovules in the *Boechera holboellii* complex. Plant Journal 58, 870-882.

Sherrard ME, Maherali H (2006) The adaptive significance of drought escape in *Avena barbata*, an annual grass. Evolution 60, 2478-2489.

Song BH, Windsor AJ, Schmid KJ, *et al.* (2009) Multilocus patterns of nucleotide diversity, population structure and linkage disequilibrium in *Boechera stricta*, a wild relative of *Arabidopsis*. Genetics 181, 1021-1033.

Storz JF, Runck AM, Sabatino SJ, *et al.* (2009) Evolutionary and functional insights into the mechanism underlying high-altitude adaptation of deer mouse hemoglobin. Proceedings of the National Academy of Sciences of the United States of America 106, 14450-14455.

van Dijk PJ (2003) Ecological and evolutionary opportunities of apomixis: insights from *Taraxacum* and *Chondrilla*. Philosophical Transactions of the Royal Society of London-Biological Sciences 358, 1113-1121.

van Kleunen M, Fischer M (2005) Constraints on the evolution of adaptive phenotypic plasticity in plants. New Phytologist 166, 49-60.

Voight BF, Kudaravalli S, Wen X, Pritchard JK (2006) A map of recent positive selection in the human genome. PLoS biology 4, e72.

Windsor AJ, Schranz ME, Formanova N, *et al.* (2006) Partial shotgun sequencing of the *Boechera stricta* genome reveals extensive microsynteny and promoter conservation with *Arabidopsis*. Plant Physiology 140, 1169-1182.

Wright SI, Bi IV, Schroeder SG, *et al.* (2005) The effects of artificial selection on the maize genome. Science 308, 1310-1314.

Zhang X-S, Hill WG (2003) Multivariate stabilizing selection and pleiotropy in the maintenance of quantitative genetic variation. Evolution 57, 1761.

Chapter 2

PLEIOTROPY OF *FRIGIDA* ENHANCES THE POTENTIAL FOR MULTIVARIATE ADAPTATION

Overview

An evolutionary response to selection requires genetic variation; however, even if it exists, the genetic details of the variation can constrain adaptation. In the simplest case, unlinked loci and uncorrelated phenotypes respond directly to multivariate selection and permit unrestricted paths to adaptive peaks. In contrast, "antagonistic" pleiotropic loci may constrain adaptation by affecting variation of many traits and limiting the direction of trait correlations to vectors that are not favored by selection. However, certain pleiotropic configurations may improve the conditions for adaptive evolution. Here we present evidence that the *Arabidopsis thaliana* gene *FRI* (*FRIGIDA)* exhibits "adaptive" pleiotropy, producing trait correlations along an axis that results in two adaptive strategies. Derived, low expression *FRI* alleles confer a "drought escape" strategy due to fast growth, low water use efficiency and early flowering. In contrast, a dehydration avoidance strategy is conferred by the ancestral phenotype of late flowering, slow growth and efficient water use during photosynthesis. The dehydration avoidant phenotype was recovered when genotypes with null *FRI* alleles were transformed with functional alleles. Our findings indicate that the well-documented affects of *FRI* on phenology result from differences in physiology, not only a simple developmental switch.

Introduction

Populations of a species are frequently distributed across climatic gradients, where natural selection can lead to adaptation to local conditions. The environmental conditions that cause local adaptation have been well documented through reciprocal transplants and studies of clines (Huey et al. 2000; Reznick and Ghalambor 2001; Stinchcombe et al. 2004; Moore and Hendry 2005; Hall and Willis 2006; Agren and Schemske 2012). These experiments show that divergent patterns of selection cause shifts in the mean values of many traits leading to a multivariate response. Such a response to selection improves fitness and promotes successful adaptation to local conditions. Despite a large body of research, it remains a challenge to determine the specific genetic loci which respond to selection and confer local adaptation (Lande and Arnold 1983; Lande 1984; Orr 1998; Rockman 2012).

Long-term breeding programs and quantitative genetic studies have demonstrated variation in nearly all traits, and thus a simple lack of additive genetic variation is not expected to constrain adaptation (Lynch and Walsh 1998). Instead, a limited amount of genetic variation along vectors of selection has been shown to limit adaptive evolution (Schluter 1996; Etterson and Shaw 2001; Kirkpatrick 2009). Theoretically, independence of all loci and phenotypes will improve the potential for adaptation by optimizing evolvability (Wagner and Altenberg 1996) and the response to selection (R) (Lande 1982; Schluter 1996; Blows and Hoffmann 2005). However, certain genetic correlations can disrupt the optimal genetic architecture by reducing the amount of genetic variation which is available to selection and causing correlated responses of non-adaptive traits (Mitchell-Olds 1996; Kirkpatrick 2009). Although this mal-adaptive role for genetic correlations is not universal (Agrawal and Stinchcombe 2009), genetic correlations

may affect R by limiting the dimensionality of the genetic (co)variance matrix or restricting genetic variation to vectors which are not aligned with selection(Mitchell-Olds 1996; Etterson and Shaw 2001; Blows et al. 2004; Kirkpatrick 2009).

The combined effects of the pleiotropic loci, which cause genetic correlations, may have a profound impact on patterns of local adaptation. As pleiotropy can constrain multivariate adaptation and cause correlated evolution of adaptive and deleterious phenotypic values, these loci are typically considered "antagonistic" (Rose 1982; Houle 1991; Curtsinger et al. 1994; Hedrick 1999; Scarcelli et al. 2007). Adaptation is especially constrained when pleiotropic gene action limits phenotypic correlations along a vector orthogonal to that of selection and reduces R (Etterson and Shaw 2001). Antagonistic pleiotropy is well documented and has led to the belief that all pleiotropy is maladaptive (Otto 2004). However, recent theoretical work has countered this viewpoint by demonstrating that intermediate levels of pleiotropy may actually improve the conditions for adaptation and evolution of complexity (Wang et al. 2010; Hill and Zhang 2012; Ostman et al. 2012).

To study the adaptive value of pleiotropic loci, it is necessary to assess the effects of genetic variation on the structure of many phenotypes which are subject to correlational selection in nature. Adaptation to drought in plants provides an ideal system to achieve this goal (Sultan and Bazzaz 1993; Bray 1997; Meyre et al. 2001; Touchette et al. 2007). In natural and agricultural systems, annual plants can be adapted to local drought conditions by either growing and reproducing before the onset of drought (drought escape) (Meyre et al. 2001; Araus et al. 2002; Sherrard and Maherali 2006; Franks et al. 2007) or by delaying reproduction, increasing water use efficiency (WUE) and conserving resources (dehydration avoidance) (McKay et al. 2003; Donovan et al. 2007; Rosenthal et al. 2010). For example, accessions which exhibit early

flowering time (FT) and low WUE were selected for in consistently wet soil and late season drought conditions (Heschel and Riginos 2005; Sherrard and Maherali 2006), while direct selection on increased WUE favored a dehydration avoidance strategy in environments with early season drought (Heschel et al. 2002). Therefore, adaptation to different local soil moisture conditions and seasonal rainfall patterns contributes to the observed strong correlations between FT, growth rate and WUE within and among species (Sultan and Bazzaz 1993; Heschel et al. 2002; Chaves et al. 2003; McKay et al. 2003; Stinchcombe et al. 2004; Angert et al. 2007). Several studies have suggested that pleiotropy may also affect this correlation (Geber and Dawson 1990; Menendez and Hall 1995; McKay et al. 2003).

Here, we provide empirical evidence for an adaptive role of pleiotropy. Using genome-wide approaches, allelic variants and transgenic manipulation, we demonstrate that the "flowering time" gene, *FRIGIDA* (*FRI*) pleiotropically affects phenotypic variation in growth rate, WUE and FT. Derived, null *FRI* alleles produce a drought escape phenotype (decreased WUE, increased growth rate, decreased FT) relative to the ancestral adaptive strategy. This phenomenon, which we term "adaptive pleiotropy", enhances the likelihood of adaptation by increasing adaptive responses to selection.

Methods

Arabidopsis thaliana genetic resources

We utilized four sets of genetic variants: recombinant inbred lines (TK RILs), a panel of 317 physiologically diverse *A. thaliana* accessions, a nearly isogenic line (FRI-NIL) and *FRI*

transgenic overexpression lines (tr-FRI). The TK RILs are the product of a bi-directional cross between two physiologically divergent accessions: TSU-1 (low WUE, short FT) and KAS-1 (high WUE, long FT) (McKay et al. 2008). The TK RILs mapping population consists of 343 F9 lines each genotyped at 166 genomic loci. In addition to the published loci, all RILs were genotyped at *FRI* via fragment analysis of PCR product generated across the promoter (primer F: 5°-AGTACTCACAAGTCACAAC-3°; primer R: 5°-GAAGATCATCGAATTGGC-3°) (Johanson et al. 2000). The 317 accession panel was genotyped at this marker (FRIdel1) and two additional markers: FRIdel2 (primer F: 5°- AGATTTGCTGGATTTGATAAGG -3°; primer R: 5°- ATATTTGATGTGCTCTCC -3°) and FRIcap (primer F: 5°- CCATAGACGAATTAGCTGC -3°; primer R: 5°- AGACTCCAGTATAAGAAG -3°). The 317 accessions and TK RILs are listed in Table S1-S2 and are available from the *Arabidopsis* stock center (*Arabidopsis*.org).

The FRI-NIL was generated by introgressing a functional *FRI* allele from the Sf-2 line into wild-type (WT) Col-0, the reference *A. thaliana* accession with a null *FRI* allele (Lee et al. 1994; Lee and Amasino 1995; Johanson et al. 2000). The tr-FRI transgenic over expressed line was generated by ligating FRI-GFP into the Xma I and Xho I sites of 35SpBARN vector and then transformed a into Col-0 background using the floral dip method. We utilized only *FRI* transgenic lines which exhibited a late flowering phenotype. The FRI-NIL and the transgenic line "FRI-GFP Col T2 #20" are available from S. Michaels and X. Yu. We also present WUE data from FRI-NIL and Columbia genotypes with knocked-out *FLC* alleles. See Michaels and Amasino (1999) (Michaels and Amasino 1999) for details on these lines.

Plant growth and phenotypic analysis

Phenotypic analyses of the TK RILs, the FRI-NIL, tr-FRI and Col-0 were conducted in a Conviron ATC60 growth chamber (Controlled Environments, Winnipeg, MB, Canada) at Colorado State University (CSU). All plants were grown at 12h, 40% humidity, 23°C days and 12h, 50% humidity, 18°C nights. Photosynthetic photon flux density during daylight was approximately 330 μmol m^{-2} s^{-1}. All plants except those analyzed for gas exchange were grown in 2" plastic pots containing Fafard 4p mix (Conrad Fafard Inc. Agawam, MA, USA). Gas exchange measurements were taken on plants grown in the same conditions in modified Conetainer™ pots (Stuewe and Sons, Tangent, OR, USA). A 195-line subset of the 317 line panel was grown at University of Texas, Austin in promix BT potting soil and 164-mL Conetainer™ pots under long-day photoperiod conditions (16L/8D) at ca. 18-21°C. Consistent with previous studies, these long-day environmental conditions induced flowering much more quickly than the 12h/12h conditions at CSU.

Gas exchange physiology was measured with an LI-6400 Photosynthesis System (LiCor Inc, Lincoln, NB, USA) equipped with a custom whole-plant gas-exchange cuvette. A total of 20 measurements were taken over a two minute period for each of 20 plants (ten replicates/genotype) at two time points (14 and 21 days post-germination). The photosynthetic parameters (A, c_i and g_s) were estimated following von Caemmerer and Farquhar (1981) (Caemmerer and Farquhar 1981). Gas exchange data were analyzed in a mixed-model framework where genotype was fixed and measurement and date were nested within individual as a random effect in JMP Genomics 5.0 (SAS Institute, Cary, NC, USA). We also generated A/c_i curves by measuring photosynthetic rate across nine levels of external CO_2 concentrations

using a different set of plants grown hydroponically. We compared A between the FRI-NIL and Col-0 controlling for variation in c_i with a mixed effect ANOVA. The genotype was a fixed effect and c_i was a continuous, random covariate.

We measured WUE, growth rate, and FT for each plant (n/genotype=10). Flowering initiation was recorded when a visible bolting structure first appeared at the apical meristems; FT is calculated as the number of days between germination and initiation of flowering. We analyzed carbon isotope composition ($\delta^{13}C$), a surrogate measure of WUE (McKay et al. 2003; Juenger et al. 2005), on lyophilized, finely ground rosette leaves at the Stable Isotope Facility at University of California, Davis (UCD). Leaves were harvested before the onset of flowering of the earliest accession at a single time point for all lines. Using images taken directly over the rosette, we assessed leaf area in image/J. These data were used to calculate relative growth rate of leaf area ($GR_{la}= [\ln(LA_{t2}) - \ln(LA_{t1})]/(t2-t1)$) where LA is leaf area at time 1 (t1) and time 2 (t2). We analyzed the affect of genotype on WUE, FT and GR_{la} via one-way ANOVA in JMP Genomics 5.0.

QTL analysis

We analyzed QTLs for WUE and FT in R/qtl (Broman et al. 2003) using the following settings: 1) imputations (256 draws) to generate a complete and even genomewide pseudomarker grid of 2cM for mapping, 2) 10,000 permutations to calculate QTL incorporation thresholds at an experiment wise α=0.05, 3) stepwise model selection scanning for epistatic and additive QTL at each step (Manichaikul et al. 2009), 4) iterative position refinement analysis by holding all but one QTL constant and varying the position of the focal QTL and re-calculating the penalized

LOD score for the model, and 5) fitting the refined model via ANOVA to calculate the effect size, percent variance explained and LOD score for each QTL. The allelic affect at *FRI* was compared via one-way ANOVA in JMP Genomics 5.0. To further refine the bi-phenotype QTL position we standardized the LOD scores by the largest value for each phenotype (st. LOD) and summed the bivariate scores for each point on the genotype grid, then calculated the bivariate QTL interval as the point where the summed LOD scores decreased to the average single phenotype odds ratio at a given map position.

Quantification of the FRI-NIL Sf-2 introgression

Whole genome sequence was obtained by paired-end Illumina sequencing at the UCD Genome Center. A reference based assembly of the TAIR 9 Columbia genome was conducted in SHORE to call SNPs (Ossowski et al. 2008) and identify the size of the introgression.

Gene expression analysis

Genome-wide gene expression was determined via Affymetrix (Affymetrix Inc. Santa Clara, CA, USA) AthSNPtile arrays for all TK RILs. We screened for all genes within 50kb of the QTL point estimate (CH4 237060-337060bp) and compared expression levels between TSU-1 and KAS-1 alleles at each gene, then corrected for multiple comparisons via q-value calculations (R package qvalue).

Analysis of population structure at FRIGIDA

We conducted three separate population genetic analyses using the publicly available genome-wide SNP data (Atwell et al. 2010; Horton et al. 2012). We imputed the functionality of *FRI* for all lines by extracting all SNPs within 100kb of *FRI* and training a classification model, support vector machines (SVM) with a radial basis function (Lasky et al. 2012),using data on SNPs and *FRI* functionality for the 317 genotypes in our panel that had both SNP and *FRI* data. After a grid search of tuning parameter values, our final SVM model predicted *FRI* functionality with 95% accuracy in four-fold cross-validation. We tested the accuracy of the SVM model using n-fold cross validation: after selecting *n* accessions at random, we tested the accuracy of SVM models in *n*-fold cross validation (i.e. leave-one-out cross validation) for *n*= 10, 15, 20, 25, 30, 35, and 40. For each value of n, we cross-validated the SVM predictions for 20 random subsets of accessions. For *n* = 10, SVM models were on average 87% accurate in cross-validation. By *n* = 20, models were 93% accurate in cross-validation. This signifies that the SNP associations with *FRI* functionality are easily observable in even small samples of accessions. Using this model, we then imputed the allelic state (binned into functional or null categories) for all accessions in the SNP database.

We calculated genome wide F_{ST} in PLINK (Purcell et al. 2007) by classifying the accessions as "functional" or "non-functional" *FRI* and calculating the molecular variance between and within these allele classes. We generated 5,000 random divisions at the same frequency as the *FRI* allele classes. These permutations allow us to assess the significance of the F_{ST} measure compared to random evolution. We conducted two additional analyses with subsets of the available accessions. Ten of the 574 sites sampled by Horton et al (2012) showed within

population variation at *FRI*. Using these populations and geographic clusters at the country level (Horton et al. 2012), we calculated an average heterozygosity over SNPs sampled at 50kb intervals (H_t). Then we split the population based on *FRI* phenotype, calculated an average heterozygosity within each subpopulation (H_s) in the same way and took the mean of those. We used these H_t and H_s values to calculate genome-wide F_{ST} based on the *FRI* phenotype. We bootstrapped to calculate significance by dividing the data at a random subset of 5000 SNPs with similar frequency to *FRI* and recalculating F_{ST}.

Comparison of climatic variables associated with FRI variation

FRI functionality calls, latitude, and longitude for each line were input into DIVA-GIS (www.diva-gis.org). The 19 BIOCLIM (www.bioclim.org) climatic variables were extracted for each point. *FRI* allelic association with these variables was made via t-tests with significance corrected for multiple comparisons by Bonferroni adjustments. The distribution of the climate under each allele was compared by ranking the climate variables and plotting the relative position of each allele relative to its rank.

Results and Discussion

Mapping the WUE-FT correlation

We measured FT and WUE of 195 *A. thaliana* accessions in a common garden. The genetic correlation between WUE and FT is positive and significant: WUE explains nearly 40%

of FT variation (n=195, r^2=0.395, p<0.0001; Fig. 2.1). If this correlation results from many loci independently affecting each phenotype, recombination between differently adapted lines will break down this favorable correlation. To test the cause of the WUE-FT correlation, we used recombinant inbred lines (TK RILs) from two phenotypically divergent accessions, TSU-1 (low WUE, short FT) and KAS-1 (high WUE, long FT) (McKay et al. 2008) (Fig. 2.2). Experimental crosses induce recombination and break up linkage disequilibrium across these genomes. Despite a large reduction in linkage disequilibrium, FT and WUE remained significantly correlated (n=304, r^2=0.138, *P*<0.0001; Fig. 2.1a) in the TK RILs, demonstrating that either tight genetic linkage or pleiotropy caused WUE and FT to co-vary.

To determine the genetic basis of the remaining WUE-FT correlation in the RILs we conducted a Quantitative Trait Locus (QTL) analysis by simultaneously scanning for genomic loci significantly associated with both phenotypes. Stepwise model selection (α=0.05) revealed a total of 11 different QTLs across both traits (Table S3). Only one QTL was found that affected both phenotypes and this QTL was the largest for each individual trait (Fig. 2.1b, Table S3). Lines with the KAS-1 QTL allele had later FT (df=1, F=338.8, *P*<0.0001), and higher WUE (df=1, F=19.14, *P*<0.0001) than TSU-1 alleles (Fig. 2.1c). The genetic correlation between WUE and FT is well documented in agricultural breeding populations and studies of local adaptation in nature (Geber and Dawson 1990; Menendez and Hall 1995; McKay et al. 2003). High WUE decreases photosynthetic assimilation rates and the amount of fixed carbon available for flowering. As the initiation of flowering is affected, in part, by resource availability (Reekie and Bazzaz 1987; Nord et al. 2011), a physiological connection between WUE and FT is plausible.

Cloning the WUE-FT QTL

To identify all possible causal variants underlying the main QTL we re-sequenced both parents and analyzed gene expression in the TK RILs for loci within a 100kb region surrounding the QTL. Within this region, only *FRI* (*FRIGIDA*) is differentially expressed between TSU-1 and KAS-1 (Fig. 2.1d). Re-sequencing of both parents revealed a 376 bp deletion within the promoter of the TSU-1 *FRI* allele, but a functional allele in KAS-1. We genotyped the *FRI* deletion in all TK RILs (Table S2). After adding the *FRI* polymorphism to the linkage map, we conducted a multi-trait position refinement analysis. The QTL maps to a single pleiotropic locus at the nearest pseudo-marker to *FRI*: chromosome 4, position 4.0cM (Fig. 2.1e).

FRI is a particularly good candidate gene underlying the FT QTL. Derived mutations that reduce expression have been involved in the evolution of spring annual types from the ancestral state of a fully functional *FRI* and a winter annual life history (Johanson et al. 2000; Shindo et al. 2005); allelic variation at *FRI* contributes to variation in FT across diverse accessions (Lee et al. 1994; Gazzani et al. 2003; Korves et al. 2007; Geraldo et al. 2009). *FRI* is also a candidate for WUE (McKay et al. 2003; Christman et al. 2008). Biogeographical analyses have associated lines with functional *FRI* alleles, such as KAS-1, to regions with lower precipitation; these environments would favor drought adaptation via dehydration avoidance (Stinchcombe et al. 2004; Brock et al. 2009; Mendez-Vigo et al. 2011).

To further assess the pleiotropic effects of the *FRI* locus, we also genotyped 195 *A. thaliana* accessions at *FRI* to determine functionality (Table S1). Consistent with pleiotropy and our observation in the TK RILs (Fig. 2.3a), phenotypic variation in both WUE (ANOVA df=1, F=51.705, $P<0.0001$) and FT (ANOVA df=1, F=34.643, $P<0.0001$) is predicted by functional

variation at *FRI* in natural populations (Fig. 2.3b). In the 195 accessions, *FRI* explains 30.2% and 24.7% of the total phenotypic variation of WUE and FT respectively. Interestingly, this "flowering time" gene explains less phenotypic variation in FT than WUE in wild accessions. Null *FRI* alleles represent a derived state of early FT and lower WUE relative to functional alleles; a drought escape strategy.

Physiological pleiotropy of FRI

To test for the phenotypic effects of *FRI*, we compared the phenotypes of a near isogenic line with a functional *FRI* allele (FRI-NIL) to the wild type progenitor Col-0 (WT) which contains a null *fri* allele. We phenotyped the three major physiological determinants of WUE: photosynthetic rate (A), leaf internal CO_2 concentration (c_i) and stomatal conductance (g_s). Stomatal conductance, which directly alters leaf water-loss dynamics, also affects A by regulating the supply of CO_2 and thus, c_i. For example, low g_s reduces A by limiting c_i, resulting in increased WUE, decreased growth rate and delayed FT (Angert et al. 2007) (Fig. 2.4a).

The FRI-NIL (functional *FRI*) had decreased g_s (contrast df=1, F=208.48, P<0.0001), marginally lower c_i (contrast df=1, F=3.28, P=0.072) and significantly lower A (contrast df=1, F=255.23, P<0.0001) relative to WT (null *fri*), indicating stomatal limitation to growth through A (Fig. 2.4b). Additionally, WUE (df =1, F=43.14, P<0.0001), GR_{la} (df =1, F=22.47, P<0.0001), and FT (df =1, F=125.22, P<0.0001) all significantly differ between FRI-NIL and WT (Fig. 2.4c). To determine the physiological mechanism for increased WUE, we modulated c_i and repeatedly measured A. Supporting a decrease in g_s as the basis for increased WUE, no

significant difference in photosynthetic capacity was found between the WT and FRI-NIL while controlling for c_i (df=288, F=1.701, *P*=0.1932; Fig. 2.5).

Many functional analyses have found that *FRI* produces a transcription factor which induces expression of *FLC*, inhibiting floral development (Michaels and Amasino 1999; Michaels and Amasino 2001; Gazzani et al. 2003; Caicedo et al. 2004; Shindo et al. 2005; Korves et al. 2007). To place our analyses in the context of these results, we analyzed WUE for WT and FRI-NIL lines which have knocked out *FLC* alleles. Consistent with the epistasis observed to affect flowering time, *FRI* confers increased WUE only in the presence of a functional *FLC* (contrast df= 1, F=44.77, *P*<0.0001), but not the when associated with a null *flc* allele (contrast df =1, F=0.79, *P*=0.38) (Fig. 2.6).

The FRI-NIL (also referred to as "Sf-2 *FRI* in Col" or "Col-*FRI*") is one of the most utilized genetic resources in the flowering time literature (Lee et al. 1994; McKay et al. 2003; Wilczek et al. 2009). These studies assume that the FRI-NIL carries a single, narrow, introgression of the Sf-2 genome which contains a functional *FRI* allele; however, this assumption has never been tested. To assess the size of the Sf-2 introgression, we resequenced the FRI-NIL, aligned the reads to the Tair-9 Columbia genome, called single nucleotide polymorphisms (SNPs) and mapped SNP density to the reference genome. Many SNPs exist between Sf-2 and Col-0 (Gan et al. 2011). High SNP density between the FRI NIL and Col-0 exists solely on proximate Chr. 4 (Fig. 2.7a-b). The region of elevated SNP density represents a single 1.070Mb (+/- 10kb) Sf-2 introgression which contains *FRI* as well as the other 325 gene models between AT4G00005 and AT4G02710. Although most studies that utilize the FRI-NIL assume the only genotypic divergence exists at *FRI*, this is obviously not the case.

To unambiguously determine if the effects observed in the FRI-NIL were due to *FRI*, we compared WUE, FT and GR_{la} between WT Col-0 and transgenic lines (Col-0 overexpressing *FRI*: tr-FRI). Under well-watered conditions, tr-FRI had greater WUE (df =1, F=57.25, $P<0.0001$), decreased GR_{la} (df=1, F=22.32, $P<0.0001$), and later FT (df =1, F=179.1, $P<0.0001$) than WT (Fig. 2.4d). As *FRI* functionality is the only DNA sequence difference between these lines, *FRI is pleiotropic and controls covariation of three traits along a vector shown to be adaptive*. Our conclusion is supported by QTL, natural accession, NIL and transgenic comparisons.

The population genetics of adaptive pleiotropy

Population genetic models are at odds about the role of pleiotropy in maintaining variation within and among populations. Pleiotropic gene action may cause non-adaptive and adaptive phenotypes to covary, thus reducing the efficacy of correlational selection and permitting the persistence of multiple allelic states within populations (Zhang and Hill 2003). However, where the effects of pleiotropy are more aligned with the direction of selection, within population variation can be purged by strong directional selection (Waxman and Peck 1998). Therefore, we predicted low levels of within population variation at *FRI*, since multivariate selection would favor either a functional (drier habitats) or non-functional allele (wetter habitats). In addition, if variation at *FRI* can lead to local adaptation, we predicted increased population structure (across the entire genome) between functional and non-functional *FRI* classes.

A population genetic test for adaptive pleiotropy is complicated in our study as *FRI* may cause population structure through both adaptive pleiotropy and allochrony: FRI-NILs and tr-FRI lines flowered at least 28 and 32 days later than Col-0 respectively. All main-raceme Col-0 flowers had been pollinated and produced fruits before any FRI-NIL or tr-FRI lines produced open flowers. In the greenhouse environment single mutations at *FRI* can produce a reproductive isolation index near 1.0. However, assortative mating due to variation at *FRI* may be tempered in nature as the environment has a profound effect on phenology (Wilczek et al. 2009).

To test for evidence of reproductive isolation between accessions and populations that differ at *FRI*, we first imputed *FRI* functionality of 1188 accessions (Horton et al. 2012) then compared the group of individuals with derived, weak alleles (i.e. null Col-0 missense and Ler deletion alleles) to the group of individuals with functional, ancestral-type *FRI*. We then calculated F_{ST} between *FRI* allele functional classes in PLINK (Purcell et al. 2007). F_{ST} values averaged across 216,130 SNPs are significantly greater between the *FRI* functionality classes than is expected from genome-wide sub-sampling ($p<0.0001$) (Fig. 2.8). To control for geographic population structure, we divided the global sample into 11 geographic regions according to Horton *et al.* (2012). Ten of 11 geographic regions showed elevated F_{ST} at *FRI* compared to a genome-wide sample of sites with the same allele frequencies as the *FRI* functional variants (Table S5). These results show that elevated global F_{ST} when sorting by *FRI* is due to a lack of within population variation in *FRI*. Less than 2% of 574 local populations harbored functional variants at *FRI*.

While extremely low within-population variation is present at *FRI*, functionally divergent alleles have gone to fixation in geographically proximate populations. Several authors have shown that an abundance of derived null *FRI* alleles are present in nature, far more than would

be expected by chance (Le Corre et al. 2002; Toomajian et al. 2006). Here we demonstrate that these mutations cause a phenotypic leap between drought adaptation strategies which may promote adaptation to novel ecological conditions. Combined, the strong signature of selection, high levels of population structure and lack of within population variation observed at *FRI* suggests an adaptive role of this pleiotropy.

FRI and drought adaptation

Previous studies have found that the early flowering, low WUE phenotypes associated with drought escape are adaptive in sites without consistent low soil moisture (Sherrard and Maherali 2006). Although we did not directly measure selection in this study, we utilized the large body of work on drought adaptation to infer the adaptive value of specific trait combinations. We predicted that due to the drought escape strategy conferred by derived loss of function mutations at *FRI*, accessions with these alleles would inhabit environments with consistently wetter growing seasons, relative to accessions with functional *FRI* alleles. To confirm the allelic association with drought, we generated a climate envelope for both *FRI* allele classes (Table S5). Functional alleles tend to be present in areas with lower growing season precipitation than non-functional alleles (t=-3.68, P=0.0003) (Fig. 2.9).

We have demonstrated that lines that diverged only at *FRI* exhibit altered positions along an adaptive phenotypic correlation. Scarcelli et al. (2007) found antagonism between the floral morphology traits affected by *FRI,* and we cannot rule out that a portion of *FRI*'s pleiotropic gene action is mal-adaptive. However, analyses presented here demonstrate a strong adaptive role of the physiological and phenological phenotypic correlations conferred by *FRI*. Given our

results, it is not surprising that *FRI* is associated with strong population genetic signatures of diversifying selection (Le Corre et al. 2002; Toomajian et al. 2006; Korves et al. 2007). Studies demonstrating historical selection on *FRI* invoke the timing of flowering as the phenotype under selection (Korves et al. 2007). Our results indicate that the observed signature of selection is not only an effect of FT variation, but may also be due to upstream physiological effects.

Conclusions

We have presented a mechanistic understanding of how *FRI* alters physiology, phenology and confers local adaptation. Phenology, growth rate and water use physiology have been mapped to similar genomic loci or correlated in natural or experimental populations (Meyre et al. 2001; McKay et al. 2003; Juenger et al. 2005; Christman et al. 2008; McKay et al. 2008). Here we have demonstrated that *FRI* causes these adaptive correlations to be heritable. Although we present a situation where pleiotropy controls phenotypic variation along a vector known to be adaptive, we have not measured the efficacy of or response to selection in the field. Fitness measures in diverse common gardens with watering treatments would allow for direct inference of the adaptive value of *FRI*.

To date, most gene annotation and characterization is conducted by forward or reverse genetics whereby a single gene or trait is under consideration. Our results indicate that a more holistic approach to phenotyping and whole plant, integrative approaches for annotating gene function may reveal complex patterns of pleiotropy among ecologically correlated phenotypes. It is possible that many trait associations are not purely a product of correlational selection, but also affected by adaptive pleiotropy.

Tables

Table 2.1 Effects of the drought and heat wave treatments and sampling date on mean daily volumetric water content (VWC) and canopy temperature (CT) during the two-week period in which the heat wave treatments were applied. F-statistics and p-values from mixed-model repeated measures ANOVAs are reported. Bold text indicates significance at $p \leq 0.05$.

WUE

ANOVA	df	SS	MS	LOD	%var	Pvalue(F)
Model	5	76.99	15.4	19.67	23.76	>.0001
Error	328	247.08	0.75			
Total	333	324.07				

QTL	df	IIISS	LOD	%var	F	Pvalue(F)
3@140.0	1	9.84	2.83	3.04	13.07	>.0001
4@4.0	1	27.91	7.76	8.61	37.05	>.0001
4@56.0	1	15.41	4.39	4.76	20.46	>.0001
5@19.2	1	13.25	3.79	4.09	17.59	>.0001
5@128.0	1	17.91	5.08	5.53	23.78	>.0001

FT

ANOVA	df	SS	MS	LOD	%var	Pvalue(F)
Model	8	249.94	31.2	106.5	76.47	>.0001
Error	330	76.93	0.23			
Total	338	326.87				

QTL	df	IIISS	LOD	%var	F	Pvalue(F)
1@88.0	1	3.95	3.69	1.21	16.96	>.0001
1@143.0	1	19.32	16.5	5.91	82.89	>.0001
4@4.0	1	212.45	97.5	65.0	911.4	>.0001
4@92.0	1	7.09	6.49	2.17	30.41	>.0001
5@34.2	1	23.82	19.9	7.29	102.2	>.0001
5@110.0	1	3.04	2.85	0.93	13.03	>.0001
5@118.0	1	3.68	3.44	1.13	15.78	>.0001

Figures

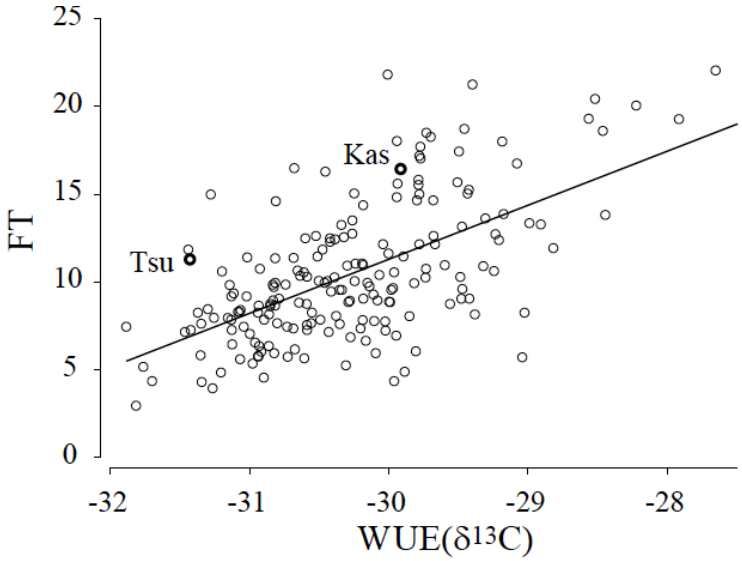

Figure 2.1 The phenotypes of 195 A. *thaliana* accessions grown in a common garden. The positive correlation, denoted by a linear regression line, is significant (r^2=0.395, p<0.0001). The two parents of the TK RILs are labeled with darker points and their respective identifications.

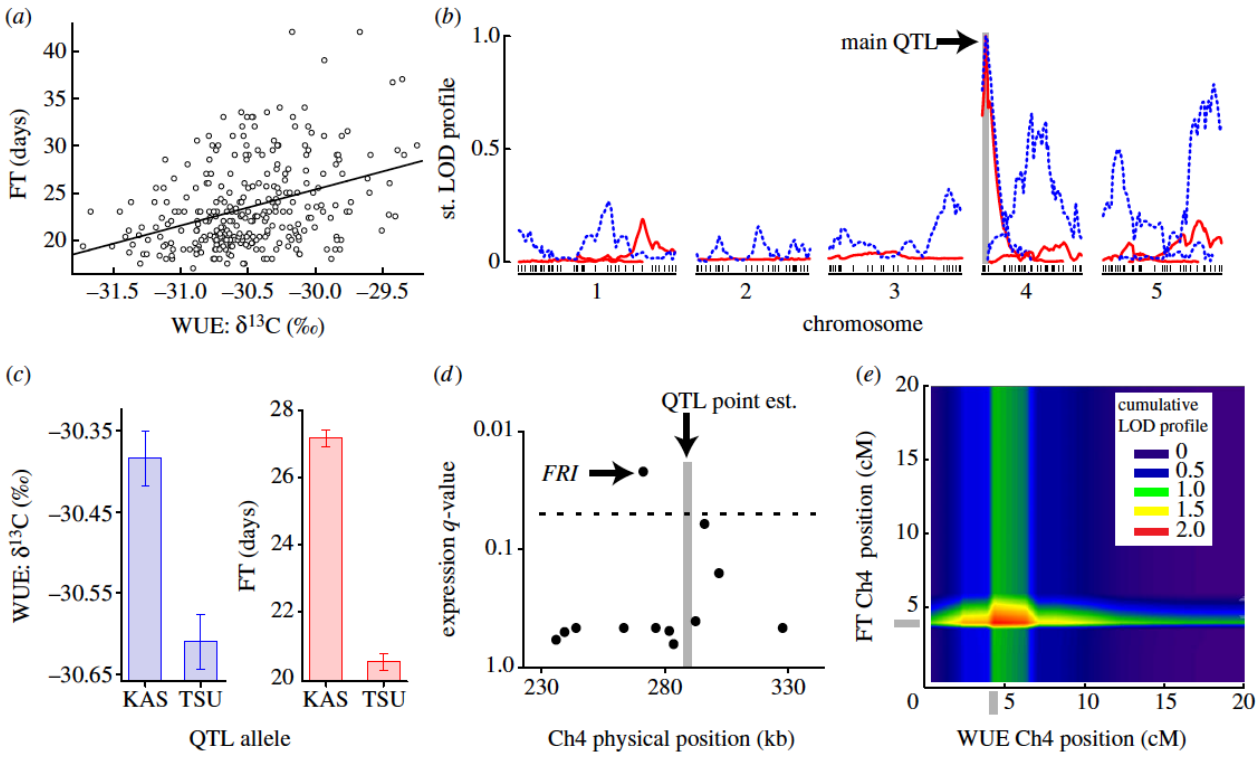

Figure 2.2 One QTL affects both WUE and FT and is associated with functionally divergent *FRI* alleles. The WUE-FT correlation observed in nature is present within our recombinant mapping population; bi-variate breeding values for each TK RIL (hollow points) and the linear model (solid line) are plotted (**a**). Standardized LOD profile scores for both WUE (dashed blue) and FT (solid red) colocalize at 4cM on chromosome 4 (**b**). WUE and FT of each RIL are split by genotype and plotted with means +/- standard errors of the mean (**c**). The only significant expression difference within 100kb of the QTL point estimate (Chromosome 4,287.06kb) is at *FRI* (labeled). Horizontal line: FDR=0.05. (**d**). Precise co-localization of the main QTL for WUE and FT is shown: standardized, summed LOD profiles are plotted for each pair wise locus combination across the first 20cM of Ch4. Gray bars on the axes indicate the point where the maximum score is achieved (**e**).

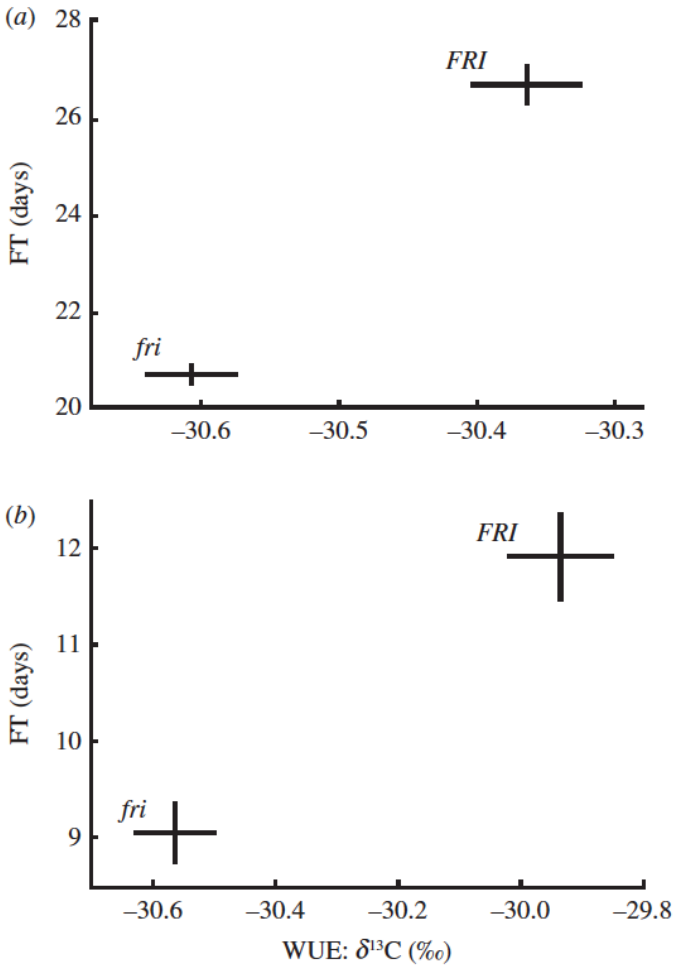

Figure 2.3 Phenotypic variance of both WUE and FT is due to natural variation at FRI. Means +/- standard errors are plotted for each allele class (FRI=functional, fri=null alleles). FRI functional variation significantly explains WUE and FT variation in the TSU-1xKAS-1 mapping population (a) and 195 natural accessions (b) when grown in a common garden.

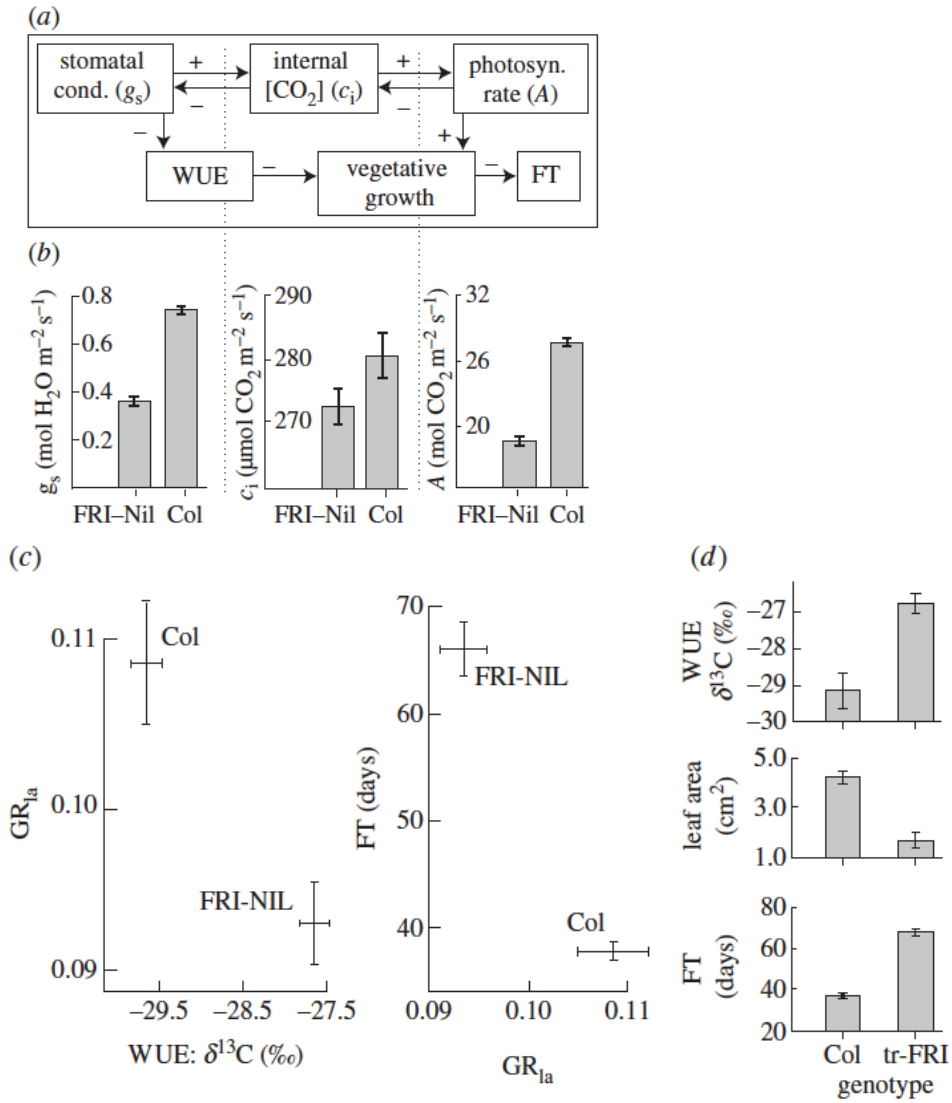

Figure 2.4 *FRI* pleiotropically affects WUE, GR_{la} and FT. The conceptual model demonstrates the mechanism of pleiotropy, upstream variation in gas exchange causes subsequent changes in GR_{la} and FT (**a**). FRI-NILs show reduced gas exchange compared to WT Col-0 (**b**). FRI-NILs have increased WUE and reduced GR_{la} compared to WT Col-0 (**c**). Transgenic overexpression lines show the same pattern as the FRI-NIL (**d**). Least square means +/- standard errors are presented in panels b-d.

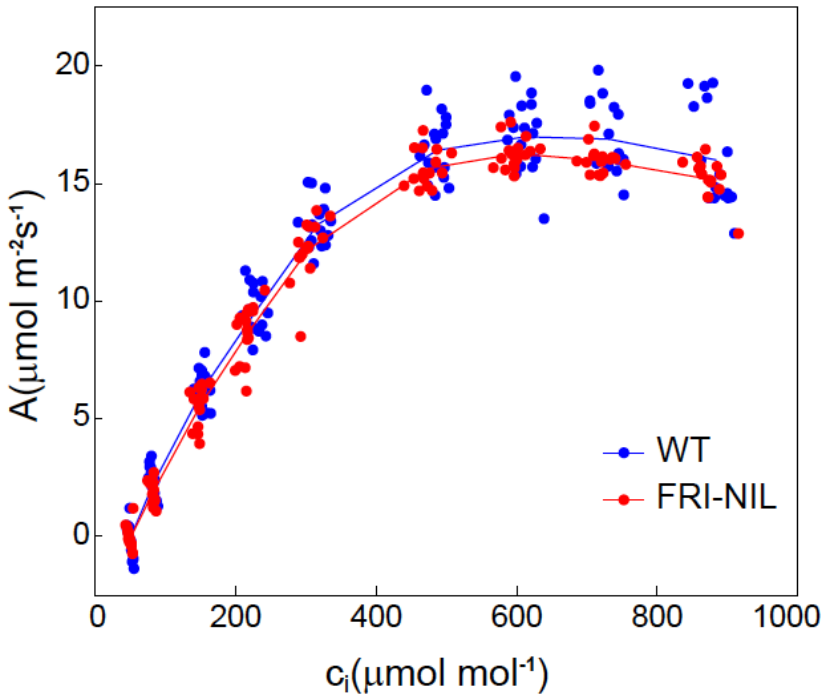

Figure 2.5 Whole-rosette photosynthetic rate (A) as a function of internal CO_2 concentration (c_i). There was no significant difference in A between WT and Col-FRI as a function of leaf internal CO_2 concentration (C_i).

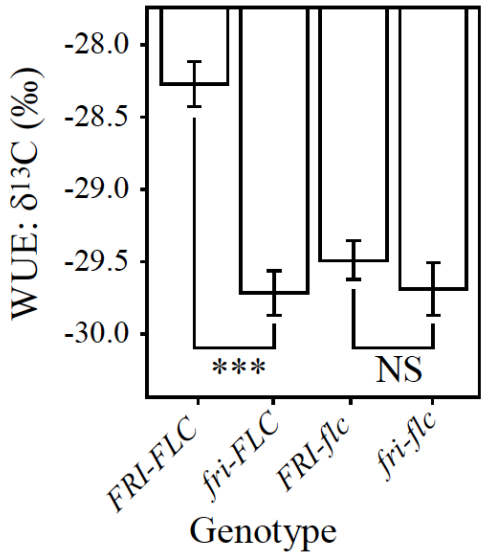

Figure 2.6 The phenotypic effects of *FLC* and *FRI*. Mean +/- standard error WUE of the four factorial combinations of *FRI* (functional, Sf-2 allele), *fri* (null, Col-0 allele), *FLC* (functional, Col-0 allele) and *flc* (flc-3 knocked out t-DNA insertion) are plotted. Bars labeled with the same letter are not statistically significantly different (α=0.05).

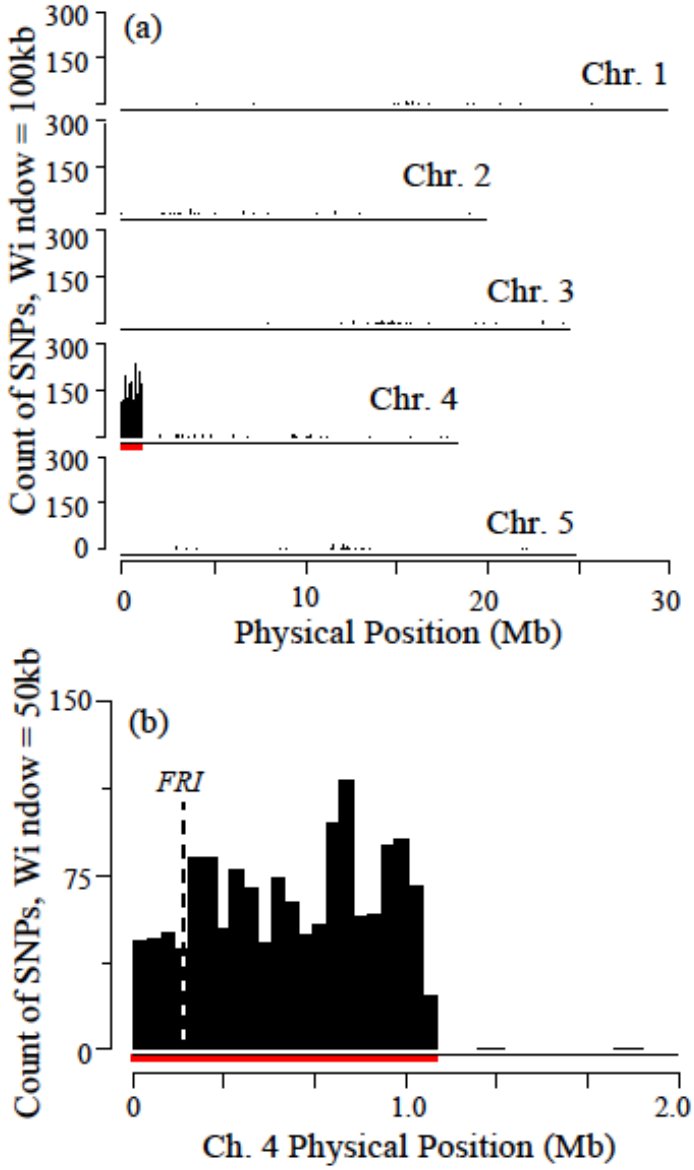

Figure 2.7 SNPs between the FRI-NIL and Col-0 are concentrated in the first 1.070MB of chromosome 4. SNPs are calculated by comparisons of paired end Illumina reads which map to the same physical position in Columbia. The position of *FRI* is highlighted by the dashed line in panel b.

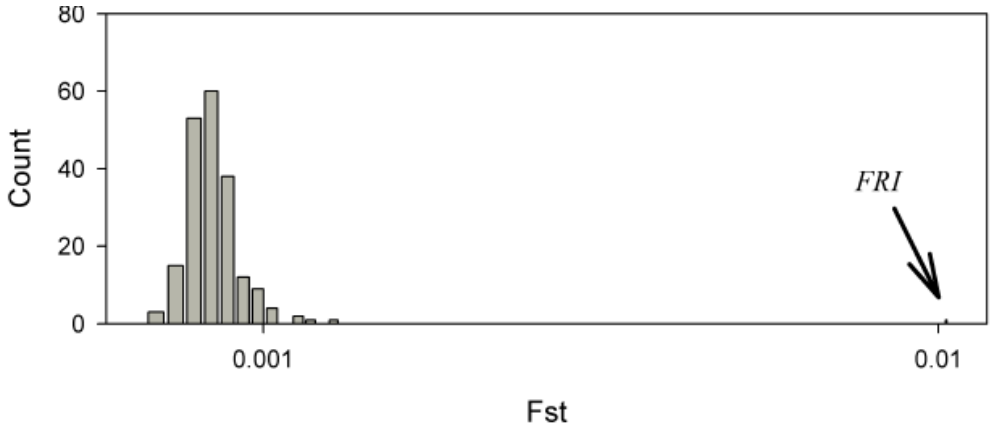

Figure 2.8 *FRIGIDA* variation produces significantly greater population structure than is expected. Permutations were conducted and F_{ST} was analyzed in PLINK; permutation results are concentrated between 0.0005 and 0.001.

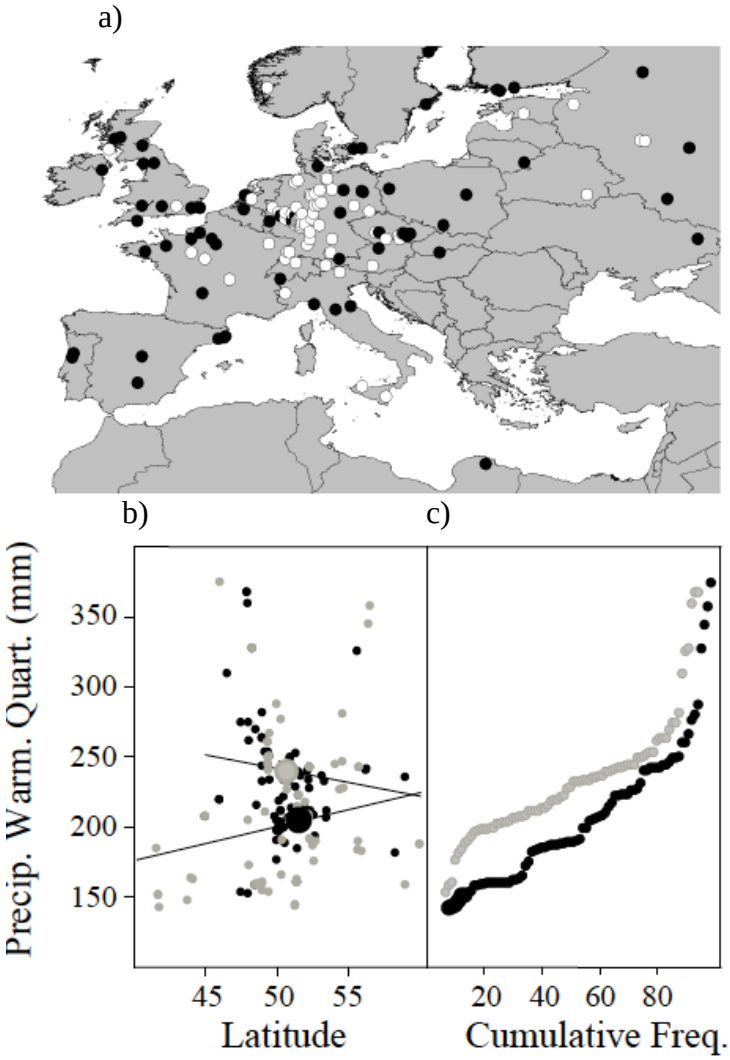

Figure 2.9 The geographic distribution of null (open circles) and functional (closed circles) *FRI* alleles in Europe and Western Asia (a). Growing season precipitation is significantly associated with *FRI* functional variation. Multiple regression of precipitation and latitude are plotted for both FRI allele classes (b). The rank of growing season precipitation is plotted against the value for each FRI class. Although very wet sites are inhabited by both allele classes, only functional alleles are found at the driest sites (c).

References

Agrawal, A. F. and J. R. Stinchcombe. 2009. How much do genetic covariances alter the rate of adaptation? Proceedings of the Royal Society B-Biological Sciences 276:1183-1191.

Agren, J. and D. W. Schemske. 2012. Reciprocal transplants demonstrate strong adaptive differentiation of the model organism *Arabidopsis thaliana* in its native range. New Phytologist 194:1112-1122.

Angert, A. L., T. E. Huxman, G. A. Barron-Gafford, K. L. Gerst, and D. L. Venable. 2007. Linking growth strategies to long-term population dynamics in a guild of desert annuals. Journal of Ecology 95:321-331.

Araus, J. L., G. A. Slafer, M. P. Reynolds, and C. Royo. 2002. Plant breeding and drought in c-3 cereals: What should we breed for? Annals of Botany 89:925-940.

Atwell, S., Y. S. Huang, B. J. Vilhjalmsson, G. Willems, M. Horton, Y. Li, D. Z. Meng, A. Platt, A. M. Tarone, T. T. Hu, R. Jiang, N. W. Muliyati, X. Zhang, M. A. Amer, I. Baxter, B. Brachi, J. Chory, C. Dean, M. Debieu, J. de Meaux, J. R. Ecker, N. Faure, J. M. Kniskern, J. D. G. Jones, T. Michael, A. Nemri, F. Roux, D. E. Salt, C. L. Tang, M. Todesco, M. B. Traw, D. Weigel, P. Marjoram, J. O. Borevitz, J. Bergelson, and M. Nordborg. 2010. Genome-wide association study of 107 phenotypes in *Arabidopsis thaliana* inbred lines. Nature 465:627-631.

Blows, M. W., S. F. Chenoweth, and E. Hine. 2004. Orientation of the genetic variance-covariance matrix and the fitness surface for multiple male sexually selected traits. American Naturalist 163:329-340.

Blows, M. W. and A. A. Hoffmann. 2005. A reassessment of genetic limits to evolutionary change. Ecology 86:1371-1384.

Bray, E. A. 1997. Plant responses to water deficit. Trends in Plant Science 2:48-54.

Brock, M. T., J. R. Stinchcombe, and C. Weinig. 2009. Indirect effects of *FRIGIDA*: Floral trait (co)variances are altered by seasonally variable abiotic factors associated with flowering time. Journal of Evolutionary Biology 22:1826-1838.

Broman, K. W., H. Wu, S. Sen, and G. A. Churchill. 2003. R/qtl: Qtl mapping in experimental crosses. Bioinformatics 19:889-890.

Caemmerer, S. and G. D. Farquhar. 1981. Some relationships between the biochemistry of photosynthesis and the gas exchange of leaves. Planta 153:376-387.

Caicedo, A. L., J. R. Stinchcombe, K. M. Olsen, J. Schmitt, and M. D. Purugganan. 2004. Epistatic interaction between *Arabidopsis* fri and flc flowering time genes generates a

latitudinal cline in a life history trait. Proceedings of the National Academy of Sciences of the United States of America 101:15670-15675.

Chaves, M. M., J. P. Maroco, and J. S. Pereira. 2003. Understanding plant responses to drought - from genes to the whole plant. Functional Plant Biology 30:239-264.

Christman, M. A., J. H. Richards, J. K. McKay, E. A. Stahl, T. E. Juenger, and L. A. Donovan. 2008. Genetic variation in *Arabidopsis thaliana* for night-time leaf conductance. Plant Cell and Environment 31:1170-1178.

Curtsinger, J. W., P. M. Service, and T. Prout. 1994. Antagonistic pleiotropy, reversal of dominance, and genetic-polymorphism. American Naturalist 144:210-228.

Donovan, L. A., S. A. Dudley, D. M. Rosenthal, and F. Ludwig. 2007. Phenotypic selection on leaf water use efficiency and related ecophysiological traits for natural populations of desert sunflowers. Oecologia 152:13-25.

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

Franks, S. J., S. Sim, and A. E. Weis. 2007. Rapid evolution of flowering time by an annual plant in response to a climate fluctuation. Proceedings of the National Academy of Sciences of the United States of America 104:1278-1282.

Gan, X., O. Stegle, J. Behr, J. G. Steffen, P. Drewe, K. L. Hildebrand, R. Lyngsoe, S. J. Schultheiss, E. J. Osborne, V. T. Sreedharan, A. Kahles, R. Bohnert, G. Jean, P. Derwent, P. Kersey, E. J. Belfield, N. P. Harberd, E. Kemen, C. Toomajian, P. X. Kover, R. M. Clark, G. Ratsch, and R. Mott. 2011. Multiple reference genomes and transcriptomes for *Arabidopsis thaliana*. Nature 477:419-423.

Gazzani, S., A. R. Gendall, C. Lister, and C. Dean. 2003. Analysis of the molecular basis of flowering time variation in *Arabidopsis* accessions. Plant Physiology 132:1107-1114.

Geber, M. A. and T. E. Dawson. 1990. Genetic variation in and covariation between leaf gas-exchange, morphologhy, and development in *Polygonum arenastrum*, an annual plant. Oecologia 85:153-158.

Geraldo, N., I. Baurle, S. Kidou, X. Y. Hu, and C. Dean. 2009. *FRIGIDA* delays flowering in *Arabidopsis* via a cotranscriptional mechanism involving direct interaction with the nuclear cap-binding complex. Plant Physiology 150:1611-1618.

Hall, M. C. and J. H. Willis. 2006. Divergent selection on flowering time contributes to local adaptation in *Mimulus guttatus* populations. Evolution 60:2466-2477.

Hedrick, P. W. 1999. Antagonistic pleiotropy and genetic polymorphism: A perspective. Heredity 82:126-133.

Heschel, M. S., K. Donohue, N. Hausmann, and J. Schmitt. 2002. Population differentiation and natural selection for water-use efficiency in *Impatiens capensis* (Balsaminaceae). International Journal of Plant Sciences 163:907-912.

Heschel, M. S. and C. Riginos. 2005. Mechanisms of selection for drought stress tolerance and avoidance in impatiens capensis (balsaminacea). American Journal of Botany 92:37-44.

Hill, W. G. and X. S. Zhang. 2012. Assessing pleiotropy and its evolutionary consequences: Pleiotropy is not necessarily limited, nor need it hinder the evolution of complexity. Nature Reviews Genetics 13:295-295.

Horton, M. W., A. M. Hancock, Y. S. Huang, C. Toomajian, S. Atwell, A. Auton, N. W. Muliyati, A. Platt, F. G. Sperone, B. J. Vilhjalmsson, M. Nordborg, J. O. Borevitz, and J. Bergelson. 2012. Genome-wide patterns of genetic variation in worldwide *Arabidopsis thaliana* accessions from the regmap panel. Nature Genetics 44:212-216.

Houle, D. 1991. Genetic covariance of fitness correlates- what genetic correlations are made of and why it matters. Evolution 45:630-648.

Huey, R. B., G. W. Gilchrist, M. L. Carlson, D. Berrigan, and L. Serra. 2000. Rapid evolution of a geographic cline in size in an introduced fly. Science 287:308-309.

Johanson, U., J. West, C. Lister, S. Michaels, R. Amasino, and C. Dean. 2000. Molecular analysis of *FRIGIDA*, a major determinant of natural variation in *Arabidopsis* flowering time. Science 290: 344-347.

Juenger, T. E., J. K. McKay, N. Hausmann, J. J. B. Keurentjes, S. Sen, K. A. Stowe, T. E. Dawson, E. L. Simms, and J. H. Richards. 2005. Identification and characterization of qtl underlying whole-plant physiology in *Arabidopsis thaliana*: Delta c-13, stomatal conductance and transpiration efficiency. Plant Cell and Environment 28:697-708.

Kirkpatrick, M. 2009. Patterns of quantitative genetic variation in multiple dimensions. Genetica 136:271-284.

Korves, T. M., K. J. Schmid, A. L. Caicedo, C. Mays, J. R. Stinchcombe, M. D. Purugganan, and J. Schmitt. 2007. Fitness effects associated with the major flowering time gene *FRIGIDA* in *Arabidopsis thaliana* in the field. American Naturalist 169:E141-E157.

Lande, R. 1982. A quantitative genetic theory of life-history evolution. Ecology 63:607-615.

Lande, R. 1984. The genetic correlation between characters maintained by selection, linkage and inbreeding. Genetical Research 44:309-320.

Lande, R. and S. J. Arnold. 1983. The measurement of selection on correlated characters. Evolution 37:1210-1226.

Lasky, J. R., D. L. Des Marais, J. K. McKay, J. H. Richards, T. E. Juenger, and T. H. Keitt. 2012. Characterizing genomic variation of *Arabidopsis thaliana*: The roles of geography and climate. Molecular Ecology 21:5512-5529.

Le Corre, V., F. Roux, and X. Reboud. 2002. DNA polymorphism at the *FRIGIDA* gene in *Arabidopsis thaliana*: Extensive nonsynonymous variation is consistent with local selection for flowering time. Molecular Biology and Evolution 19:1261-1271.

Lee, I. and R. M. Amasino. 1995. Effect of vernalization, photoperiod, and light quality on the flowering phenotype of *Arabidopsis* plants containing the *FRIGIDA* gene. Plant Physiology 108:157-162.

Lee, I., S. D. Michaels, A. S. Masshardt, and R. M. Amasino. 1994. The late-flowering phenotype of *FRIGIDA* and mutations in luminidependens is suppressed in the landsberg erecta strain of *Arabidopsis*. The Plant Journal 6:903-909.

Lynch, M. and B. Walsh. 1998. Genetics and analysis of quantitative traits. Sinauer, Sunderland, MA.

Manichaikul, A., J. Y. Moon, S. Sen, B. S. Yandell, and K. W. Broman. 2009. A model selection approach for the identification of quantitative trait loci in experimental crosses, allowing epistasis. Genetics 181:1077-1086.

McKay, J. K., J. H. Richards, and T. Mitchell-Olds. 2003. Genetics of drought adaptation in *Arabidopsis thaliana*: I. Pleiotropy contributes to genetic correlations among ecological traits. Molecular Ecology 12:1137-1151.

McKay, J. K., J. H. Richards, K. S. Nemali, S. Sen, T. Mitchell-Olds, S. Boles, E. A. Stahl, T. Wayne, and T. E. Juenger. 2008. Genetics of drought adaptation in *Arabidopsis thaliana* ii. Qtl analysis of a new mapping populations, kas-1 x tsu-1. Evolution 62:3014-3026.

Mendez-Vigo, B., F. X. Pico, M. Ramiro, J. M. Martinez-Zapater, and C. Alonso-Blanco. 2011. Altitudinal and climatic adaptation is mediated by flowering traits and *fri*, *flc*, and *phyc* genes in *Arabidopsis*. Plant Physiology 157:1942-1955.

Menendez, C. M. and A. E. Hall. 1995. Heritability of carbon-isotope discrimination and correlations with earliness in cowpea. Crop Science 35:673-678.

Meyre, D., A. Leonardi, G. Brisson, and N. Vartanian. 2001. Drought-adaptive mechanisms involved in the escape/tolerance strategies of *Arabidopsis* landsberg erecta and columbia ecotypes and their f1 reciprocal progeny. Journal of Plant Physiology 158:1145-1152.

Michaels, S. D. and R. M. Amasino. 1999. *Flowering locus c* encodes a novel mads domain protein that acts as a repressor of flowering. The Plant Cell Online 11:949-956.

Michaels, S. D. and R. M. Amasino. 2001. Loss of *flowering locus c* activity eliminates the late-flowering phenotype of *FRIGIDA* and autonomous pathway mutations but not responsiveness to vernalization. Plant Cell 13:935-941.

Mitchell-Olds, T. 1996. Pleiotropy causes long-term genetic constraints on life-history evolution in brassica rapa. Evolution 50:1849-1858.

Moore, J. S. and A. P. Hendry. 2005. Both selection and gene flow are necessary to explain adaptive divergence: Evidence from clinal variation in stream stickleback. Evolutionary Ecology Research 7:871-886.

Nord, E. A., K. Shea, and J. P. Lynch. 2011. Optimizing reproductive phenology in a two-resource world: A dynamic allocation model of plant growth predicts later reproduction in phosphorus-limited plants. Annals of Botany 108:391-404.

Orr, H. A. 1998. The population genetics of adaptation: The distribution of factors fixed during adaptive evolution. Evolution 52:935-949.

Ossowski, S., K. Schneeberger, R. M. Clark, C. Lanz, N. Warthmann, and D. Weigel. 2008. Sequencing of natural strains of *Arabidopsis thaliana* with short reads. Genome Research 18:2024-2033.

Ostman, B., A. Hintze, and C. Adami. 2012. Impact of epistasis and pleiotropy on evolutionary adaptation. Proceedings of the Royal Society of London Series B-Biological Sciences 279:247-256.

Otto, S. P. 2004. Two steps forward, one step back: The pleiotropic effects of favoured alleles. Proceedings of the Royal Society of London Series B-Biological Sciences 271:705-714.

Purcell, S., B. Neale, T.-B. K., L. Thomas, M. Ferreira, D. Bender, J. Maller, P. Sklar, P. de Bakker, M. Daly, and P. Sham. 2007. Plink: A toolset for whole-genome association and population-based linkage analysis. American Journal of Human Genetics 81:559-575.

Reekie, E. G. and F. A. Bazzaz. 1987. Reproductive effort in plants .3. Effect of reproduction on vegetative activity. American Naturalist 129:907-919.

Reznick, D. N. and C. K. Ghalambor. 2001. The population ecology of contemporary adaptations: What empirical studies reveal about the conditions that promote adaptive evolution. Genetica 112:183-198.

Rockman, M. V. 2012. The qtn program and the alleles that matter for evolution: All that's gold does not glitter. Evolution 66:1-17.

Rose, M. R. 1982. Antagonistic pleiotropy, dominance, and genetic variation. Heredity 48:63-78.

Rosenthal, D. M., V. Stiller, J. S. Sperry, and L. A. Donovan. 2010. Contrasting drought tolerance strategies in two desert annuals of hybrid origin. Journal of Experimental Botany 61:2769-2778.

Scarcelli, N., J. M. Cheverud, B. A. Schaal, and P. X. Kover. 2007. Antagonistic pleiotropic effects reduce the potential adaptive value of the *FRIGIDA* locus. Proceedings of the National Academy of Sciences of the United States of America 104:16986-16991.

Schluter, D. 1996. Adaptive radiation along genetic lines of least resistance. Evolution 50:1766-1774.

Sherrard, M. E. and H. Maherali. 2006. The adaptive significance of drought escape in *Avena barbata*, an annual grass. Evolution 60:2478-2489.

Shindo, C., M. J. Aranzana, C. Lister, C. Baxter, C. Nicholls, M. Nordborg, and C. Dean. 2005. Role of *FRIGIDA* and *flowering locus c* in determining variation in flowering time of *Arabidopsis*. Plant Physiology 138:1163-1173.

Stinchcombe, J. R., C. Weinig, M. Ungerer, K. M. Olsen, C. Mays, S. S. Halldorsdottir, M. D. Purugganan, and J. Schmitt. 2004. A latitudinal cline in flowering time in *Arabidopsis thaliana* modulated by the flowering time gene *FRIGIDA*. Proceedings of the National Academy of Sciences of the United States of America 101:4712-4717.

Sultan, S. E. and F. A. Bazzaz. 1993. Phenotypic plasticity in *Polygonum persicaria* 2. Norms of reaction to soil moisture and the maintenance of genetic diversity. Evolution 47:1032-1049.

Toomajian, C., T. T. Hu, M. J. Aranzana, C. Lister, C. L. Tang, H. G. Zheng, K. Y. Zhao, P. Calabrese, C. Dean, and M. Nordborg. 2006. A nonparametric test reveals selection for rapid flowering in the *Arabidopsis* genome. Plos Biology 4:732-738.

Touchette, B. W., L. R. Iannacone, G. E. Turner, and A. R. Frank. 2007. Drought tolerance versus drought avoidance: A comparison of plant-water relations in herbaceous wetland plants subjected to water withdrawal and repletion. Wetlands 27:656-667.

Wagner, G. P. and L. Altenberg. 1996. Perspective: Complex adaptations and the evolution of evolvability. Evolution 50:967-976.

Wang, Z., B.-Y. Liao, and J. Zhang. 2010. Genomic patterns of pleiotropy and the evolution of complexity. Proceedings of the National Academy of Sciences of the United States of America 107: 18034-18039

Waxman, D. and J. R. Peck. 1998. Pleiotropy and the preservation of perfection. Science 279:1210-1213.

Wilczek, A. M., J. L. Roe, M. C. Knapp, M. D. Cooper, C. Lopez-Gallego, L. J. Martin, C. D. Muir, S. Sim, A. Walker, J. Anderson, J. F. Egan, B. T. Moyers, R. Petipas, A. Giakountis, E. Charbit, G. Coupland, S. M. Welch, and J. Schmitt. 2009. Effects of genetic perturbation on seasonal life history plasticity. Science 323:930-934.

Zhang, X. S. and W. G. Hill. 2003. Multivariate stabilizing selection and pleiotropy in the maintenance of quantitative genetic variation. Evolution 57:1761-1775.

Chapter 3

QUANTITATIVE TRAIT VARIATION AND HERITABILITY, BUT NOT NEUTRAL GENETIC VARIATION, PREDICT RANGE SIZE IN *BOECHERA SPP.*

Overview

The question of what causes range size variation among species is central to the study of evolution and ecology as well as efforts to conserve biodiversity. Many taxonomic groups contain both rare and widespread species, revealing that range size can evolve as quickly as other quantitative traits. Comparisons of groups of rare and widespread species that share a recent common ancestor can elucidate the ecological and evolutionary processes that underlie range size diversity. Here, we report quantitative and molecular genetic attributes of four related *Boechera* (Brassicaceae) species with divergent range sizes and discuss patterns that may explain the causes and evolutionary consequences of rarity. Microsatellite polymorphism does not show a strong relationship with range size. In contrast, quantitative genetic diversity and partitioning, and the degree of phenotypic plasticity, are positively correlated with range size. In particular, heritability is significantly lower in rare species for all eight traits measured. We also found evidence of elevated levels of both neutral molecular (F_{ST}) and quantitative genetic (Q_{ST}) population structure in widespread species. These low levels of population divergence, phenotypic plasticity and trait heritability may reduce evolutionary potential and seriously impede future adaptation of rare species.

Introduction

"Who can explain why one species ranges widely and is very numerous, and why another allied species has a narrow range and is rare?"

-Charles Darwin, *The Origin of Species* (6th edition, p.12)

The geographic range is one of the most fundamental ecological characteristics of a species (Darwin 1872; MacArthur 1972; Brown et al. 1996; Sexton et al. 2009), defining its conservation status, evolutionary dynamics, and interactions with the physical and biological environment (Gaston 1996; Geber 2011). Underlying the geographic range is a set of environmental conditions in which a species can maintain demographic stability (Hutchinson 1957; Brown 1984; Warren et al. 2008). Under equilibrium conditions, the range is a spatial manifestation of the ecological niche (Hutchinson 1957; Pulliam 2000; Holt 2003). Therefore, range size is taken as a general indicator of the diversity of environmental conditions that can be tolerated by a species (Brown 1984; Futuyma and Moreno 1988). Within most taxonomic groups there exists a continuum between rare species with narrow ranges and those that are widespread across a large range of ecological conditions (Darwin 1872; Brown et al. 1996; Gaston 1996). As rare species are prone to the effects of environmental stochasticity and face the potential of extinction (Lande 1993; Payne and Finnegan 2007), understanding the processes underlying range size evolution is of great interest to conservation and evolutionary biologists (Kruckeberg and Rabinowitz 1985). Groups of closely related taxa that differ greatly in range size provide an opportunity to investigate the evolutionary and ecological processes underlying these differences.

While widespread species have adapted to a diverse array of habitats, rare species remain constrained to narrow geographic distributions. Therefore, rare species' small geographic ranges may be a result of processes that are known to constrain evolution at the range margin, namely decreased evolutionary potential, low genetic diversity or maladaptive gene flow (Angert and Schemske 2005; Blows and Hoffmann 2005; Angert et al. 2008; Gaston 2009; Eckhart et al. 2011). The likelihood of an adaptive response to selection (evolutionary potential) is affected by the amount of inherited trait variation within a population (Etterson and Shaw 2001; Etterson 2004; Pujol and Pannell 2008). A lack of evolutionary potential in rare species inhibits adaptation, which may eventually limit niche breadth and increase the likelihood of extinction (Blows and Hoffmann 2005).

Range size may alternatively be a consequence of species' physiological tolerances to environmental heterogeneity and stress. Many ecological studies have shown that niche breadth is directly related to the putative diversity of phenotypes within a species (Brown 1984; Gaston and Spicer 2001; Slatyer et al. 2013), and the range of ecological conditions that permit a non-negative population growth rate (λ) across the landscape (Hutchinson 1957; Levins 1968; Spicer and Gaston 2009). Phenotypic diversity can be partitioned into the relative degree of population (or individual) differentiation and the extent of phenotypic plasticity. For example, widespread species, with broad ecological niches may be characterized by highly differentiated populations (Bolnick et al. 2007; Nakazato et al. 2008; Agren et al. 2013). Conversely, species with large ranges may represent a generalist strategy (Baker and Stebbins 1965). In this case, range size should be positively correlated with the capability of individuals to respond to diverse ecological conditions (plasticity) (Whitlock 1996; Pohlman et al. 2005). Since labile traits may permit

homeostasis of fitness in diverse habitats, broad niches of widespread species may be due to increased phenotypic plasticity (Baker and Stebbins 1965; Kellermann et al. 2009).

To test these hypotheses about the causes of rarity, ecological and evolutionary studies commonly compare genetic or physiological attributes between related pairs of rare and widespread species. For example, in an analysis of 34 species pairs, Gitzendanner and Soltis (2000) showed evidence of weakly elevated average genetic diversity in widespread species; however, 26% of the rare species are more genetically diverse than their widespread congeners. These data, and other reviews (Karron 1987; Cole 2003), support the conclusion that elevated genetic diversity is generally, but inconsistently, associated with increased range size. There are also many examples of ecological studies that compare niche breadth and phenotypic diversity among closely related species with divergent range sizes (Baskauf and Eickmeier 1994; Baskauf 2001; Richards et al. 2003). In two broader comparisons, one of 20 (Lavergne et al. 2004) and another of five species pairs (Pohlman et al. 2005), rare species' phenotypic values were not consistently divergent from widespread species. These findings lead the authors to conclude that phenotypic differentiation between rare and widespread species is taxon and context dependent. Therefore, despite the wealth of both phenotypic and genetic data, the biological factors that contribute to the diversity of range sizes among species remain, for the most part, unknown.

Here, we conducted a combined quantitative and neutral molecular genetic comparison of two rare and two widespread species in the genus *Boechera* (Brassicaceae). Among our study species were the two Colorado state-listed (S2), imperiled members of the genus, *B. crandallii* and *B. vivariensis* (formerly *B. fernaldiana ssp. vivariensis*). This study marked the first genetic analysis and ex-situ conservation efforts for these two species. We also included a moderately widespread relative, *B. spatifolia* and the well-studied cosmopolitan species, *B. stricta*, as a

baseline to compare results. Using these species, we directly assessed three hypotheses about the ecological and evolutionary causes and consequence of rarity. First, we predicted that individuals from widespread species should display elevated phenotypic plasticity relative to rare species. Second, we expected a greater level of population structure and total diversity in widespread species. Third, we expected rare species to have lower levels of genetically-based trait variation (heritability). Combined, these factors may limit the range of environments that permit demographic stability of rare species and impede the potential for future adaptation to novel environmental conditions.

Methods

Study Species

The genus *Boechera* is ideal for between-species comparisons due to recent common ancestry, ecological divergence of many species (Beck et al. 2012; Alexander et al. 2013), and the availability of diverse molecular and ecological tools (Lovell 2011; Rushworth et al. 2011). We quantified the geographic range of four *Boechera* species by summing the number of 0.1°latitude x0.1°longitude (123 km^2) grid cells that contain at least a single georeferenced herbarium specimen (e.g. Lloyd et al. (2002)) in DIVA-GIS (www.diva-gis.org). Here, we classify *B. vivariensis* and *B. crandallii* as rare due to their restricted range sizes and listing in the Colorado rare plant index (http://www.cnhp.colostate.edu).

The four study species are sexual diploids that inhabit montane and semi-arid environments in the rocky mountain region of North America (Roy 1995, Al-Shehbaz 2003;

Windham and Al-Shehbaz 2006; Alexander et al. 2013). Populations of the widespread species, *B. spatifolia* (n_{pop} = 26) and the rare species, *B. crandallii* (n_{pop} = 7), and *B. vivariensis* (n_{pop} = 6) were sampled across the entire geographic range. For the widespread species, *B. stricta* (n_{pop} = 17), sampling was conducted less densely, across a subset of the geographic range (Fig. 1). To confirm that all lines were sexual diploids, we screened all seed families using the Flow Cytometric Seed Screen (FCSS: Matzk et al. (2000)) on a Partec PAII flow cytometer (Partec GmbH, Münster, Germany) following methods of Lovell et al. (2013a). Maternal plants that produced a signature of apomixis or triploidy were excluded from this experiment.

Phenotypic Analysis

We planted four reps (sibs) of each of 391 maternal families across the four species. The germination rate was 90%, resulting in 1401 plants that we phenotyped: 186 *B. crandallii* (52 families from 7 populations), 103 *B. vivariensis* (35 families from 6 populations), 717 *B. spatifolia* (193 families from 26 populations), and 394 *B. stricta* (111 families from 13 populations).

Plants were grown in 1” diameter RLC-4 conetainers (Steuwe and Sons, Tangent, OR, USA) filled with Fafard 4P soil mix. Three seeds were placed directly on the soil and germinated following 14 days of cold stratification. After ten days of growth, seedlings were thinned to one plant/conetainer. Growth conditions were designed to mimic those experienced by winter-annual *Boechera* species: germination in early fall (23/18°C, 12/12h day/night), growth during the fall (18/8°C, 12/12h day/night), vernalization over the winter (8/4°C, 8/16h day/night), then growth

in the spring (23/18°C, 12/12h day/night). All plants were grown in a single Conviron ATC60 growth chamber at Colorado State University, Ft. Collins, CO, USA.

We measured eight traits: height, leaf area and leaf number at 21 days post-germination (pre-vernalization); height leaf area and leaf number at 55 days post-germination (post-vernalization); inter-node distance (leaf number/height at 55 days), and leaf size (leaf area/leaf number at 55 days). Height and leaf number were measured directly, while leaf area was calculated by extracting the canopy area from photographs taken directly above the plant. Image processing was completed in Photoshop CS5.1 (Adobe Corporation, San Jose, CA, USA) and analysis of leaf area was conducted in ImageJ (http://rsbweb.nih.gov/ij/). We chose these traits because leaf morphology and vegetative rosette architecture vary considerably across differently adapted populations in other plant species (McKay et al. 2001; Leinonen et al. 2009). Furthermore, mean values of these phenotypes are relatively similar for each species, increasing our power to detect differences in within-species variance components.

To calculate quantitative genetic statistics, we fit a random effects model to the data with two factors: population of origin (V_{POP}) and seed family [nested within population] (V_{FAM}). From this model, we extracted two species-level statistics: Total phenotypic variance ($V_{TOT}=V_{POP}+V_{FAM}+V_{RESIDUAL}$), and the proportion of genetic variance partitioned among populations ($Q_{ST}=V_{POP}/(V_{FAM}+V_{POP})$). Heritability, the degree to which variation of a trait is due to heritable genetic factors, is the major determinant of the response to selection (breeders' equation: $R=h^2S$). We calculated broad-sense heritability (H^2) for each of the eight phenotypes from a random-effect model where family [nested within population] is the only term ($H^2=V_{FAM}/r(V_{TOT})$ and total variance is scaled by the relatedness (r) of the sibs within a family. In designs such as ours, where within population sampling is small, individual population estimates

can be inaccurate. This method of nesting families within populations provides an appropriate mechanism to improve the accuracy of estimates of genetic variance (B. Walsh, Personal Communication). We estimated self-pollination rates with population genetic data (see below), and used this information to inform our estimates of quantitative genetic parameters. The selfing rate places a lower limit of the likelihood of a full sib mating. Furthermore, selfing rate can be used to directly infer expected relatedness of a family. For example, a selfing rate of 30% leads to expected relatedness of approximately 0.4 (Jordan et al. 1999). If selfing rates were higher than 50%, we assumed families to be full sibs (r=0.5).

We measured the response to changing environmental conditions within each individual as a measure of phenotypic plasticity. This measure of plasticity, also referred to as flexibility in animals (Bradley 1978), or acclimatization, provides a direct estimate of the magnitude of response by genotypes to changing environmental conditions (Pelletier et al. 2007). Individuals were grown in both fall and winter temperature and photoperiod conditions; therefore, we calculated plasticity as the within-individual phenotypic variance (Falconer and Mackay 1996; van de Pol and Wright 2009) between fall and winter growth conditions for two traits: growth rate of leaf area (GR_{LA}) (Lovell et al. 2013b) and stem elongation growth rate (growth rate = $(\ln(\text{trait}_{t2})-\ln(\text{trait}_{t1}))/(\text{days}_{t2}-\text{days}_{t1})$). In outbred or long-lived systems such as *Boechera*, where cloning genotypes is not feasible, this repeated-measures approach is the most robust method available to calculate plasticity.

Genetic Analysis

DNA was extracted from lyophilized leaf tissue using the ChargeSwitch gDNA plant kit (Invitrogen Corp. Carlsbad, CA, USA). We followed PCR and genotyping protocols optimized by Beck et al. (2012) for 15 SSR markers known to amplify well across most *Boechera* species. Three-primer set multiplexed PCR and genotyping were conducted on all samples. Primers were constructed with "FAM" and "HEX" labeled dyes and genotyped on an ABI 3130xL Genetic Analyzer at the Colorado State University proteomics and metabolomics facility. Alleles were called using the ABI "GeneMapper v 4.0" software. PCR conditions and primer information can be found in Table 3.1. Three loci gave null alleles in *B. crandallii*. One of these was also null in *B. vivariensis*. These alleles were coded as null and incorporated into the analysis. Individuals with missing data for >4 SSRs or signatures of duplication were excluded. Following inclusion of the null alleles and exclusion of individuals with poor amplification, our genotyping success was >98%. Genotype data are archived online in DRYAD (www.datadryad.org, doi:___).

We calculated summary statistics at three scales: 1) within-population diversity, A_{POP} (effective number of alleles/population) and H_S (within population gene diversity, a.k.a. expected heterozygosity), 2) species-level diversity, A_{SP} (total effective number of alleles), and H_T (total gene diversity), and 3) among-population genetic structure, F_{ST} (the degree of population structure) using the R package "heirfstat" (Goudet 2005). We also calculated the rate of self pollination (selfing-rate, "s") through a maximum likelihood estimation procedure in RMES (David et al. 2007).

Statistical Analyses

All statistical comparisons were conducted in the R environment for statistical computing 3.0.2 (R Core Team 2013). To infer the statistical significance of observed differences between rare and widespread species we conducted two non-parametric approaches. First, we calculated the mean value of each statistic (e.g. genetic diversity, heritability, plasticity) for each species and trait/locus combination. Using these means, we tested whether rare species were less diverse, were less structured or had lower heritability than widespread species through a one-tailed wilcoxon ranked sum test paired by trait or locus. Second, we calculated the linear correlation coefficient of each trait/locus statistic and the log-transformed estimated range size of each species. The correlation coefficients were categorized as positive or negative, and the statistical significance of an excess of positively signed values was assessed using an exact binomial sign test. *P*-values exceeding the $\alpha = 0.05$ threshold indicated that range size is positively correlated with the statistic in question. The significance of each estimate of plasticity was assessed by 1000 bootstrapped wilcoxon ranked sum (non-paired) tests on the genotype means. Significance of total phenotypic plasticity was assessed by a multiple regression across species.

To remove any confounding role that spatial distribution or sample size may introduce into our analyses, we conducted a spatially explicit population sub-sampling approach by a custom script in R version 3.0.2. First, we extracted sets of seven populations from the widespread species; these sets were spatially constrained to geographic distributions similar to that of *B. crandallii*. Second, we recalculated the molecular and quantitative genetic statistics for the sub-sampled populations. To determine the extensibility of this approach, we also conducted an automated routine that carried out non-independent iterative subsampling of 100 population

groups. We compared the statistical output from both manual and automated subsampling routines to the entire species values through non-parametric Wilcoxon tests, implemented in R version 3.0.2.

Results

The four species had very different geographic ranges; area occupied by each species in our study region (Fig. 1) was 1113km^2 (*B. vivariensis*), 3339km^2 (*B. crandallii*), 27080km^2 (*B. spatifolia*) and 98924km^2 (*B. stricta*). Additionally, geographic range size was highly correlated with the range of environmental variation underlying the geographic distribution (Table 3.2).

Microsatellite variation indicated relatively equivalent genetic distance between species, without evidence of phylogenetic species pairs (Fig. S1). This was consistent with a recently published phylogenetic analysis of *Boechera* (Alexander et al. 2013). Summary data from the 15 SSR loci can be found in Table 3.3.

Observed heterozygosity (H_o) and the proportion of variation in a sub-population found within individuals (F_{IS}) were both significantly associated with range size (H_o: $binom_{14} = 2$, P=0.013, F_{IS}: $binom_{15} = 14$, P=0.001), where rare species had higher H_o (V_{15}=97, P=0.006), but lower F_{IS} (V_{15}=2, P=0.002. These effects are primarily driven by the highly divergent statistical values of *B. vivariensis*, which exhibited lower self-fertilization rates (34%), relative to the other three species (Table 3.3).

Molecular genetic diversity is not strongly correlated with range size

Neither within population (A_{pop} $binom_{15}$ = 6, P<0.1; H_S: $binom_{15}$ = 6, P<0.1), nor total genetic diversity statistics (H_t: $binom_{15}$ = 11, P<0.1, A_{sp}: $binom_{15}$ = 10, P<0.1) were correlated with range size. To determine whether rare species had lower neutral genetic variation than widespread species, we averaged across species types and compared rare to widespread species by a Wilcoxon test (Fig. 2). No statistically significant effects of distribution classification were found for any of the four statistics (A_{pop}: V_{15}=87, P <0.1; H_t: V_{15}=73, P <0.1; A_{SP}: V_{15} = 26.5, P <0.1; H_T: V_{15} = 31, P <0.1).

It is possible that diversity estimates widespread species may be distorted by increased sample size and the spatial distribution of our sampled population. To test this, we controlled for the geographic scale and number of sampled populations by spatially sub-sampling populations from the widespread species to mimic the geographic distribution and sample size of the rare species. While no statistically significant differences were found following sub-sampling, measures of A_{SP} in widespread species regressed towards the mean of the rare species (Fig. S2a). Inferences by within population molecular diversity through A_{pop} and H_s, and species-level diversity by H_t were unaffected by sub-sampling (Fig. S2b).

Quantitative genetic partitioning and diversity are associated with range size

Total phenotypic variance (V_{TOT}) was calculated for all species and traits. The correlation between ($\log_{10}$) range size and V_{TOT} was positive for all eight traits, resulting in a significantly positive multi-trait correlation ($binom_8$ = 0, P=0.008). Furthermore, on average, rare species had

significantly lower V_{TOT} than widespread species (V_8=4, P=0.027, Fig. 3a). V_{TOT} was marginally negatively correlated with H_s (n=4, r=-0.93, P=0.069, Fig. S3a) and negatively, but not significantly, associated with the other indices of molecular diversity (A_{sp}: n=4, r= -0.14, P=0.85; A_{pop}: n=4, r= -0.56, P=0.44; H_t: n=4, r= -0.15, P=0.29; Fig. S3b-d).

Genotypic estimates of the two measures of phenotypic plasticity, GR_{LA} and stem elongation, were not correlated (n=368, r=-0.0655, P=0.21); however, there existed a strong positive correlation between range size and plasticity across traits (multiple regression$_{1,6}$, r^2=0.5542, P=0.034). Widespread species had the highest level of plasticity in both traits. *B. stricta* displayed the greatest plasticity of GR_{LA}, (t=8.55, bootstrap P<0.0001; Fig. 3b) and *B. spatifolia* had the greatest plasticity of stem elongation growth rate (t=8.96, bootstrap P<0.0001; Fig. 3c).

Rare species have decreased evolutionary potential

Estimation of quantitative genetic parameters is affected by the relationship among siblings. To infer family structure, we calculated a single rate of outcrossing for each species (Table 3.3). The 37% selfing rate of the rare species, *B. vivariensis*, results in a predicted relatedness of progeny of 0.4. We scaled our heritability estimates accordingly. Rare species had lower heritability of each trait; *B. spatifolia* or *B. stricta* demonstrated the greatest heritability for all eight of the traits (Fig. 3d). Across the eight traits, heritability was strongly positively correlated with range size ($binom_8$ = 0, P=0.008) and was significantly greater in widespread species (V_8=2, P=0.018).

The proportion of phenotypic variance partitioned among populations (Q_{ST}) is highly positively correlated with range size ($binom_8 = 0$, P=0.008). Across species types, rare species displayed lower levels of Q_{ST} than widespread species (V_8=0, P=0.008, Fig. 3e). The neutral molecular equivalent of Q_{ST}, F_{ST}, was also strongly correlated with range size ($binom_{15}$= 0, P=0.001) and was significantly greater in widespread than rare species (V_{15}=9, P=0.002, Fig. 3f). The H^2-Q_{ST}, H^2-F_{ST} and F_{ST}-Q_{ST} correlations were all positive and strong (Fig. 4a-c). This correlation generally held among species, where widespread species had much greater Q_{ST}, F_{ST} and H^2 than rare species. Additionally, the correlations between neutral molecular genetic diversity and heritability were all negatively correlated or non-significantly different from zero (Fig. S3e-h).

As in our measures of molecular genetic diversity, we also controlled for sampling spatial scale and number of populations. However, this procedure had no statistical effect. Even when sampled across a similar distribution, rare species continued to exhibit decreased V_{TOT} ($F_{1,125.5}$= 29.16, P<0.0001), Q_{ST} ($F_{1,125.7}$= 8.73, P=0.004) and H^2 ($F_{1,123.4}$= 6.85, P=0.010) relative to widespread species.

As in our measures of molecular genetic diversity, we also controlled for sampling spatial scale and number of populations. However, this procedure had no statistical effect. Even when sampled across a similar distribution, rare species continued to exhibit decreased V_{TOT} ($F_{1,125.5}$= 29.16, P<0.0001), Q_{ST} ($F_{1,125.7}$= 8.73, P=0.004) and H^2 ($F_{1,123.4}$= 6.85, P=0.010) relative to widespread species.

Discussion

Our results were consistent with the hypothesis that widespread species have broader ecological niches (Slatyer et al. 2013) and should therefore display elevated phenotypic diversity relative to rare species. The increased phenotypic variation found in widespread species was due to both strong spatial structure (>3x increase in among population variance relative to rare species) and increased phenotypic plasticity. However, our neutral molecular analysis ran contrary to the hypothesis that rare species should be genetically homogenous. Molecular genetic diversity was not significantly different between rare and widespread species across all species and diversity indices. While total diversity was similar among species types, structure of this variation among populations (F_{ST}) was much lower in rare than widespread species.

From an evolutionary perspective, rare species may have narrower ranges because of an inability to adapt to novel conditions at the range margin (Sexton et al. 2009). Therefore, we hypothesized that rare species have limited heritability of potentially adaptive traits. Our data are strongly aligned with this hypothesis: widespread species exhibited much greater heritability across eight phenotypes. Additionally, both molecular and quantitative population differentiation were highly positively correlated with heritability and were elevated in widespread species. Since rare species were characterized by decreased evolutionary potential and exhibited lower levels of phenotypic plasticity, these species may be more exposed than their widespread relatives to the short and long term effects of climate change and anthropogenic landscape modification.

Molecular diversity within and across populations

Across all four indices, there were no consistent associations between genetic diversity and range size (Fig. 2a-d). While total genetic diversity was higher, but not significantly so, in widespread species, we observed the opposite pattern for within-population diversity. As expected, within population measures of genetic diversity were robust to the geographic extent of population sampling; however, there was a significant effect of sub-sampling on the species-level diversity index, A_{sp}, where sub sampled distributions displayed decreased diversity relative to the full population sample. It is important to note that *B. spatifolia* and *B. stricta* exhibit little isolation by distance (IBD). In highly structured species with high IBD, a larger reduction in diversity could be expected following sub-sampling. Many comparisons of molecular diversity utilize a wider distribution and higher sample size for widespread species relative to the rare congener (e.g. Song and Mitchell-Olds (2007); Takahashi et al. (2011)). In these cases, it is possible that species-level diversity may be confounded by sampling scheme.

In her seminal review of rarity, Rabinowitz (1981) argued that rare species could not be classified by a single demographic pattern. Distinguished by opposing values of population connectivity, census size and spatial extent, *B. crandallii* and *B. vivariensis* represented opposite ends of the spectrum of narrowly endemic rare species. Despite having the narrowest range, *B. vivariensis* maintained the largest and densest populations of all study species (J.T. Lovell personal observation). Also, its highly connected habitat and small spatial distribution may have increased the amount of genetic exchange among populations. Conversely, *B. crandallii* maintained small, dispersed populations (J.T. Lovell personal observation). Measures of molecular diversity are often strongly affected by effective population size (N_e); therefore, it is

possible that the demography of these rare species drove the observed within-population diversity patterns. These processes and a lower level of self-pollination may have generated the relatively high number of alleles and expected heterozygosity observed in *B. vivariensis*. The presence of broadly different molecular genetic signatures between the two rare species raises the possibility that rare species can exhibit widely varying population genetic attributes.

Phenotypic diversity and correlations with range size

The widespread species studied here exhibited geographic ranges 1-3 orders of magnitude larger and occupied considerably more heterogeneous ecological and climactic habitats than the rare species. Given this broader geographic and environmental distribution, we hypothesized that widespread species would exhibit elevated total phenotypic diversity. Our results confirm this hypothesis (Fig. 3a).

The total diversity of phenotypes is one potential driver of ecological niche breadth and range size; however, patterns of phenotypic partitioning, particularly within individuals (plasticity) and among populations, may be better predictors of ecological tolerance. While the total amount of variation was only slightly different among species, a much greater proportion of variance was found among populations (Q_{ST}) of widespread species than of rare congeners. Due to lower relative responses to selection, decreased intensity of diversifying selection and/or increased gene flow, rare species display decreased population differentiation relative to widespread congeners.

Plasticity can have contradictory effects on adaptation (Whitlock 1996; Price et al. 2003; Ghalambor et al. 2007) and population persistence at the range margin. It is logical that a high

degree of plasticity will expand the diversity of suitable habitats (Pohlman et al. 2005; Baker and Stebbins 1965), and there are several instances where increased plasticity is beneficial when invading new sites (Loomis and Fishman 2009), surviving stress (Heschel et al. 2004) or persisting through changing environmental conditions (Chevin and Lande 2011). However, plasticity may be non-adaptive (Relyea 2002), reducing the local fitness as well as the strength of selection.

By calculating differences in growth rate across environmental conditions, we found that genotypes of widespread species have significantly greater plasticity of both leaf area growth rate and stem elongation rate than those of rare species (Fig. 3b-c). It is possible that the faster and more extreme phenotypic adjustments of widespread species would permit homeostasis of fitness across more diverse environments, leading to increased niche breadth. However, future studies that measure fitness and physiology across diverse conditions are needed to directly address this hypothesis.

Heritability and the correlates of evolutionary potential

An adaptive response to selection, especially at the range margin, permits range expansion (Angert and Schemske 2005; Angert 2006; Sexton et al. 2009; Chevin and Lande 2011). Therefore, rare species are hypothesized to be geographically restricted by poor responses to changing environments (adaptive potential) and an inability to adapt to conditions outside of their narrow range. For example, if species differ in the amount of additive genetic variation, the relative response to selection will also vary. In the case of range margins, species with a greater potential for adaptive evolution will be more likely to experience range expansion, while species

with a low amount of heritable trait variation will experience static or contracting ranges. Therefore, it is possible that rare species are less capable of range expansion because of decreased heritability.

Consistent with the hypothesis that rare species should have decreased evolutionary potential, growth rate, leaf size and stem elongation, were significantly less heritable in rare species than in their widespread congeners (Fig. 3d). Furthermore, estimates of heritability for all eight traits were significantly positive in the widespread species; however, only two and seven traits were significantly heritable in *B. vivariensis* and *B. crandallii* respectively. This lack of heritable quantitative trait variation, especially in the rarest species, may have constrained adaptation and the range size of these lineages.

Additional considerations for the use of molecular and quantitative genetic tools in the study of rarity

Summary statistics describing molecular polymorphism within populations and species are routinely used to make decisions about conservation concerns of rare species. However, the majority of studies comparing such population genetic parameters between rare and widespread species do not find strong associations between genetic diversity and range size. Our analyses also find genetic diversity to be a poor correlate with rarity. Consistent with Reed and Frankham (2001) we found a slight negative correlation between heritability and molecular diversity. Therefore, molecular genetic indices are poorly correlated with those traits that could actually *cause* rarity. While well-situated for characterization of population structure and relatively inexpensive to obtain, neutral genetic diversity indices do not provide an adequate statistic with

which to test the hypothesis that rare species lack the potential to adapt to novel environmental conditions (Frankham et al. 1999; Vitt and Havens 2004), nor do they provide direct inference about local adaptation or the potential for future adaptive evolution (McKay et al. 2001). Population genetic analysis can provide valuable information about the demographic history of a species, while quantitative genetic variation directly affects evolutionary potential and may provide insight into the causes of rarity.

Tables

Table 3.1 List of primer names, labels and sequences used in the SSR analysis. PCR was conducted with 5-PRIME HotStart Master Mix (Gaithersburg, MD, USA) in 12ml reactions using the following PCR conditions: initial denaturation (95C, 120sec), [denaturation (94C, 30sec) annealing (53C, 90sec), extension (65C, 60sec)], number of cycles (25), final extension (65C, 30mins).

Multiplex Set	Primer Name	Label	Forward Primer	Reverse Primer
1	I3	6-FAM	gactaatcatcaccgactcagccac	attcttcttcacttttcttgatcccg
	B20	HEX	ttctcgggaaagtaatgaggag	gcaaatctgaccaatgcaag
	A1	6-FAM	gtctattcgaggacgcc	aggttgggtaggtgaag
2	B11	6-FAM	tcctccattgtagagcagagc	ccattgcttaaaccctaaacc
	I14	HEX	tcgaggtgctttctgaggtt	tacctcacccttttgaccca
	C8	6-FAM	ttccgggtatcattcctag	gttgtaagttctttctcag
3	B9	6-FAM	aaacacattcccgtcagctc	ttgattgaatcctgcgtttg
	B18	HEX	aacctcccaagattcgcttc	ttcgccattgttgtgatttg
	E9	6-FAM	aggaaaggacaaaagacatg	gcttccatggaaggagaccc
4	BF3	6-FAM	tttttagacagtagtggctgtgag	acttcgttccaggctcgtc
	BF19	HEX	accgcattggtgttgtgtc	ataacggacgcgaccaaag
	B6	6-FAM	gcaaaagatcttcatgggac	tgccatttctttccctagtg
5	BF15	6-FAM	cagcatctcctttgggtttg	acttgctcctttgcatgacc
	B266	HEX	tttaatttgtgcgtttgatcc	caaaatcgcagaatgagagg
	A3	6-FAM	agctttgtttgcaatggag	gtgagaataatattgacc

Table 3.2 Environmental diversity underlying the geographic range of each species. The range (maximum value – minimum value) of each of the 19 BIOCLIM variables (bio1-19) and elevation above sea level (alt) were presented for each species. Wilcoxon test *p*-values were reported.

Species	*vivariensis*	*crandallii*	*spatifolia*	*stricta*
Range (km2)	1113	3339	27080	98924
bio1	21	90	116	147
bio2	22	47	63	82
bio3	2	3	14	13
bio4	753	1592	2509	2875
bio5	37	125	131	182
bio6	11	84	127	148
bio7	47	97	131	155
bio8	97	81	242	316
bio9	10	174	213	277
bio10	28	100	119	167
bio11	10	73	115	119
bio12	86	379	588	843
bio13	5	34	64	91
bio14	6	26	38	45
bio15	6	36	57	50
bio16	22	84	161	229
bio17	18	97	127	157
bio18	25	86	198	237
bio19	18	140	189	280
alt	471	1488	1911	2256

Table 3.3 Summary of population genetic diversity. For each species and locus, we calculated the total number of alleles (*A*), % of individuals with missing data (%NA), and observed heterozygosity (H_o). Species level statistics, expected heterozygosity (H_e), inbreeding coefficient (F_{IS}), and rate of self pollination are also reported. See Beck et al. (2012) for genus wide statistics.

	B. vivariensis			*B. crandallii*			*B. spatifolia*			*B. stricta*		
Locus	%NA	*A*	H_O	%NA	*A*	H_O	%NA	*A*	H_O	%NA	*A*	H_O
a1	0.00	2	0.11	0.00	2	0.00	5.70	2	0.02	0.79	2	0.00
bf11	0.00	1	0.00	0.00	1	0.00	3.63	3	0.15	0.00	4	0.00
bf18	0.00	1	0.00	0.00	1	0.00	1.04	2	0.00	3.15	4	0.00
b6	0.00	2	0.03	0.00	1	0.00	0.52	11	0.09	3.94	1	0.00
a3	0.00	3	0.14	0.00	1	0.00	0.00	1	0.00	9.45	3	0.20
bf20	0.00	6	0.49	1.96	4	0.13	5.18	8	0.18	0.79	5	0.02
c8	6.38	5	0.08	0.00	2	0.00	6.74	3	0.18	14.96	6	0.01
bf9	17.02	4	0.21	0.00	3	0.12	1.04	6	0.13	3.15	12	0.02
bf19	6.38	17	0.77	0.00	1	0.00	12.44	5	0.14	3.94	10	0.06
bdru266	61.70	10	0.33	0.00	1	0.00	0.52	17	0.18	2.36	12	0.04
ice3	0.00	12	0.50	1.96	6	0.10	6.22	7	0.18	0.79	7	0.01
ice14	2.13	3	0.05	0.00	1	0.00	3.63	2	0.00	0.00	3	0.00
e9	0.00	5	0.42	0.00	3	0.06	1.04	5	0.13	3.15	12	0.03
bf3	4.26	20	0.77	1.96	1	0.00	0.00	8	0.13	3.94	11	0.00
bf15	8.51	3	0.35	3.92	5	0.23	0.00	5	0.17	2.36	4	0.01
Across Loci												
H_e	0.355			0.082			0.142			0.2311		
F_{IS} (+/- SE)	0.270 (0.071)			0.660 (0.085)			0.710 (0.098)			0.903 (0.054)		
Self. rate (+/- SE)	0.337 (0.082)			0.831 (0.0698)			0.950 (0.010)			0.833 (0.0594)		

Figures

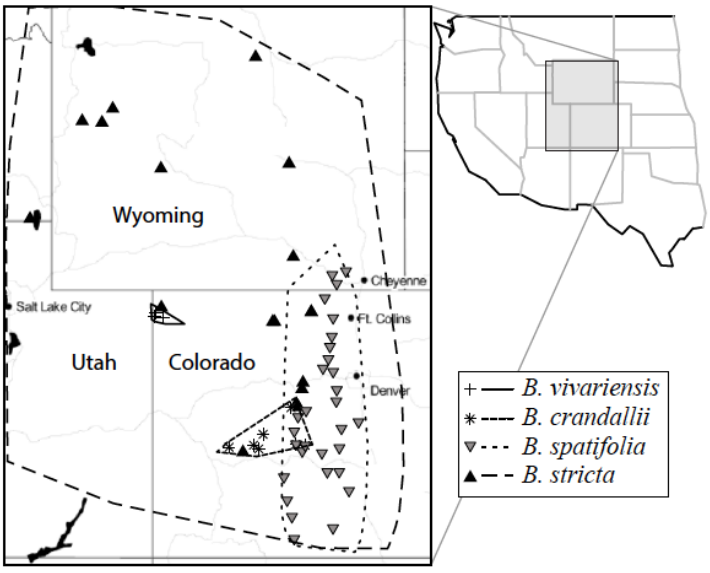

Figure 3.1 The sample populations (markers) and geographic distribution (polygons) of the four species in Utah, Wyoming and Colorado, USA. Minimum convex polygons (MCP) were generated after excluding all herbarium specimens outside the extent of this map and plotted in the associated line characteristics for each species. The total distribution of *B. stricta* extends beyond this map. Locations of populations collected for this experiment are plotted as the respective marker for each species.

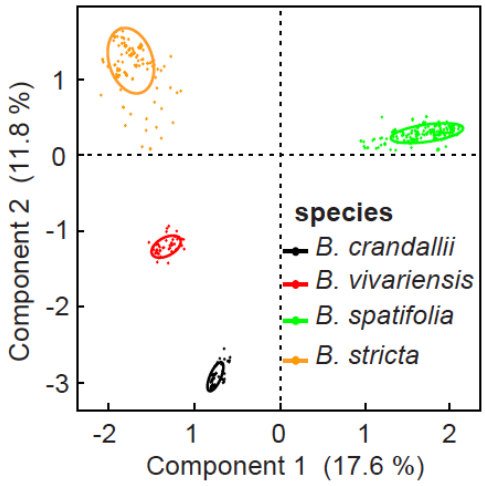

Figure 3.2 Principal component analysis derived from 15 SSR loci demonstrated the amount of divergence between the species. PCA axis 1 and 2 scores for each individual were plotted by species with the associated 95% confidence ellipse.

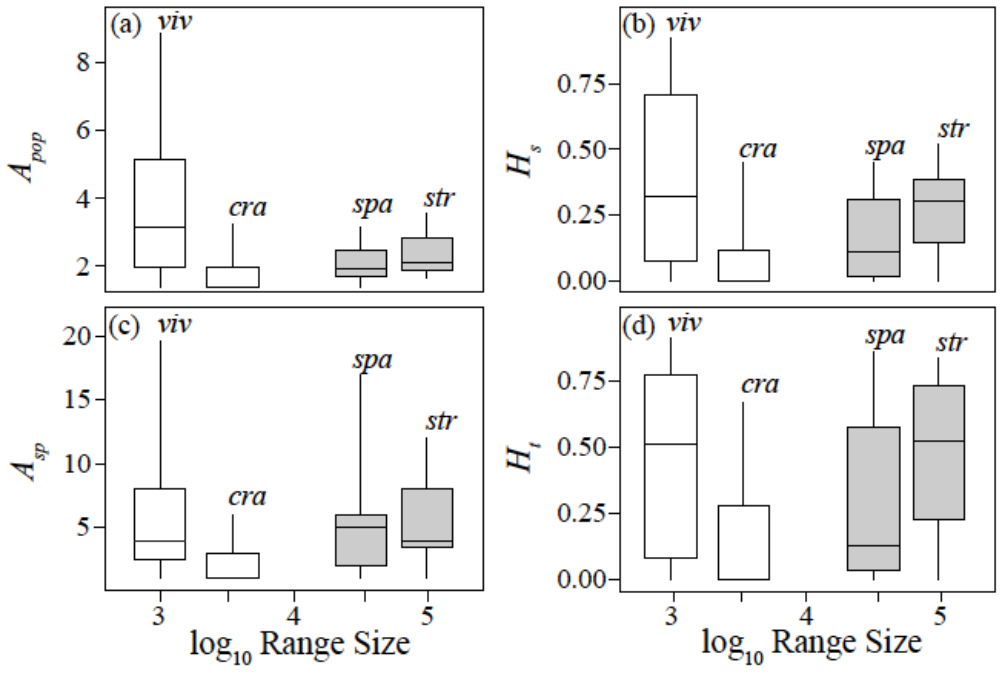

Figure 3.3 Molecular diversity of rare and widespread *Boechera* species. The distribution of the two within population diversity indices, effective number of alleles (A_{pop}) and gene diversity (H_s) (a-b), and the species-level genetic diversity, alleles/locus (A_{sp}) and total gene diversity (H_t) (c-d) are presented. The boxplots display the inter-quartile range (box) and total range of observations, including outliers (lines). Species were labeled as *viv* (*B. vivariensis*), *cra* (*B. crandallii*), *spa* (*B. spatiflia*) and *str* (*B. stricta*); rare species were indicated by hollow bars, widespread were solid gray. The position on the x-axis represents the $\log_{10}$ geographic range size (km^2). Other figures followed this notation

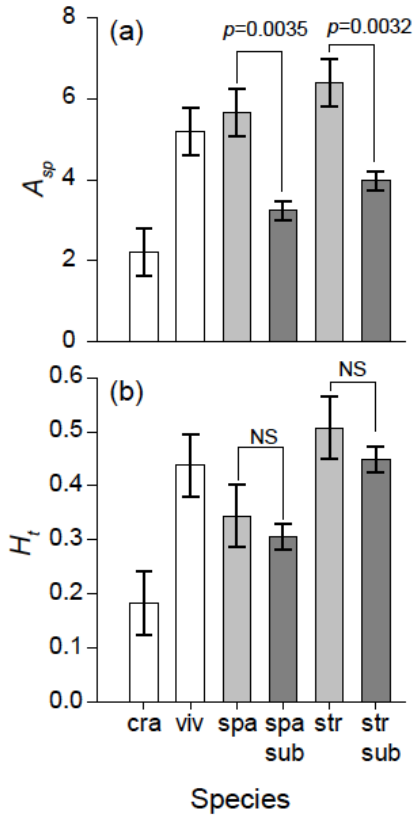

Figure 3.4 The effect of spatial sub-sampling on molecular and phenotypic diversity estimates. Total number of alleles/locus was highly affected by sampling distribution (a); however, total gene diversity was relatively unaffected (b). The darker shaded bars, labeled with "sub" indicate the statistical value of sub-sampled distributions.

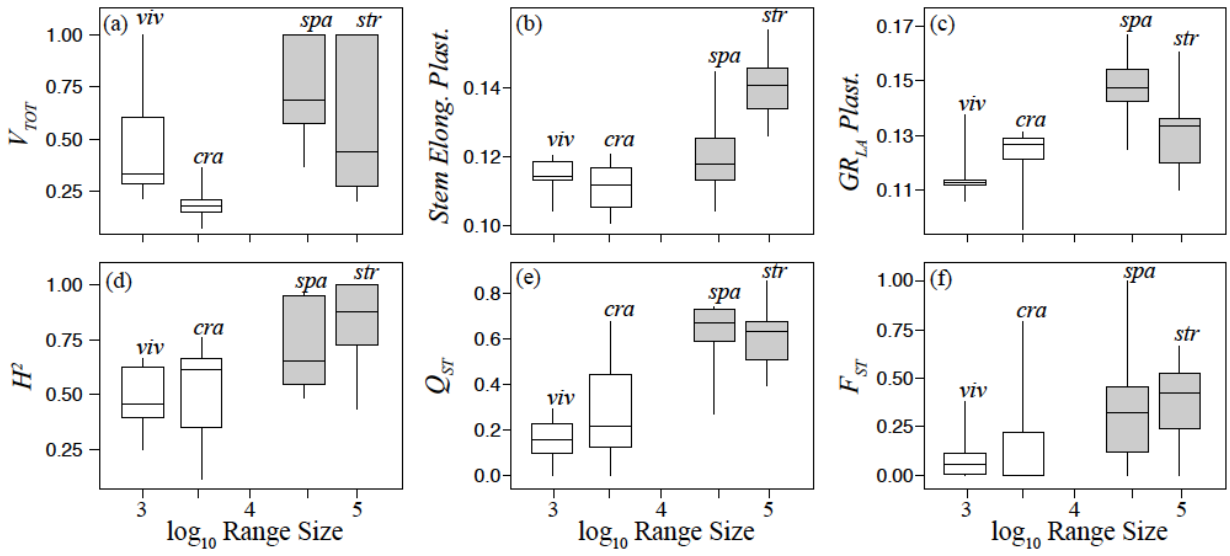

Figure 3.5 Structure and quantitative genetic attributes differ between rare and widespread species. Boxplots, following notation in Fig. 2, present the amount and structure of genetic variance for each species. The total amount (V_{TOT}) of phenotypic variance, standardized within each trait (a) and the two plasticity measures (b-c) were significantly elevated in widespread species. Plasticity was presented as the within-individual variance for each growth rate phenotype. Heritability, (H^2) and quantitative (Q_{ST}) and molecular genetic (F_{ST}) structure were also significantly associated with range size (d-f).

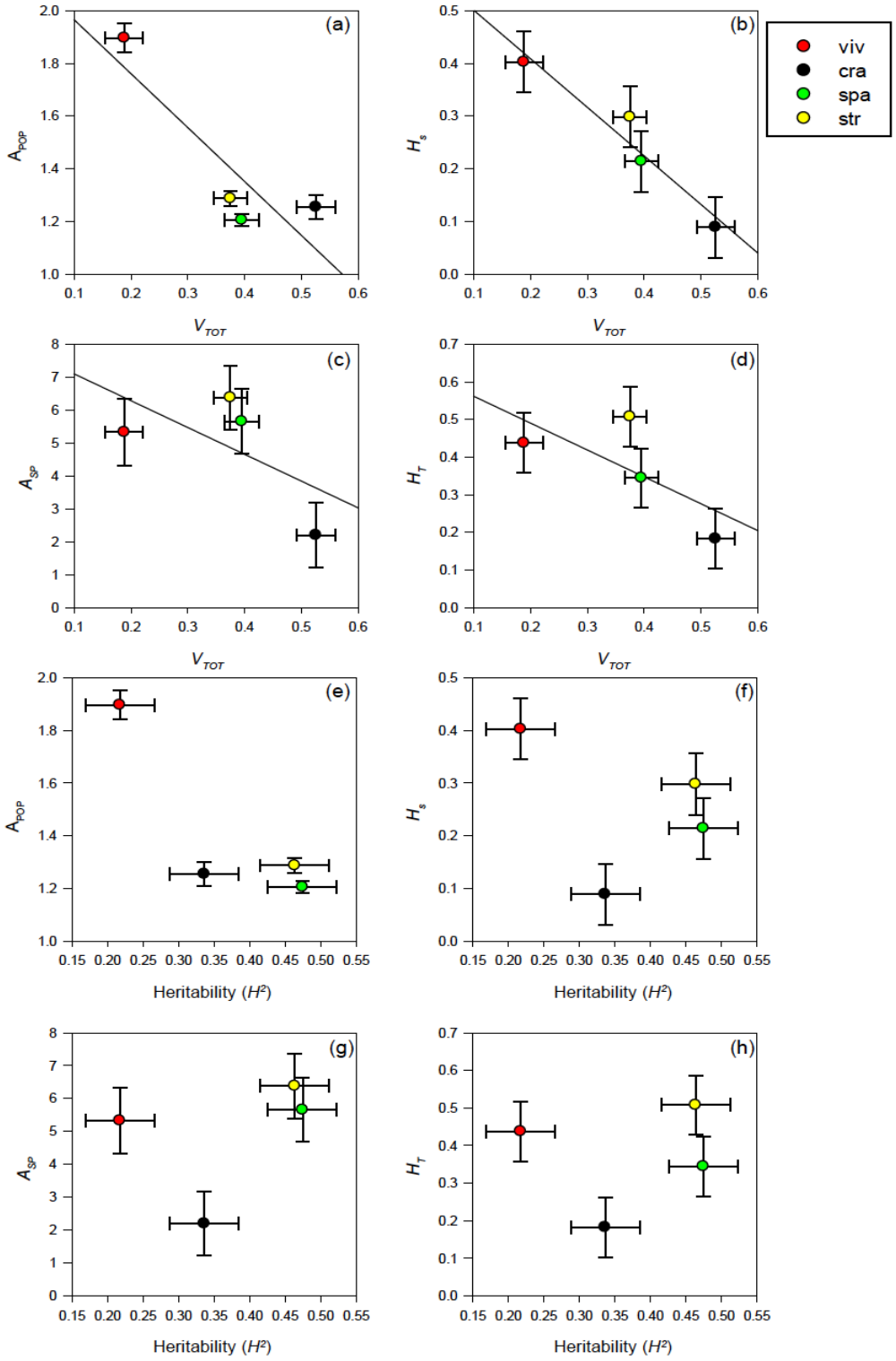

Figure 3.6 Correlations between the four molecular genetic diversity indices, total phenotypic variance (a-d) and heritability (e-h). Least square means +/- SE constructed across the 15 loci and 8 phenotypes were reported

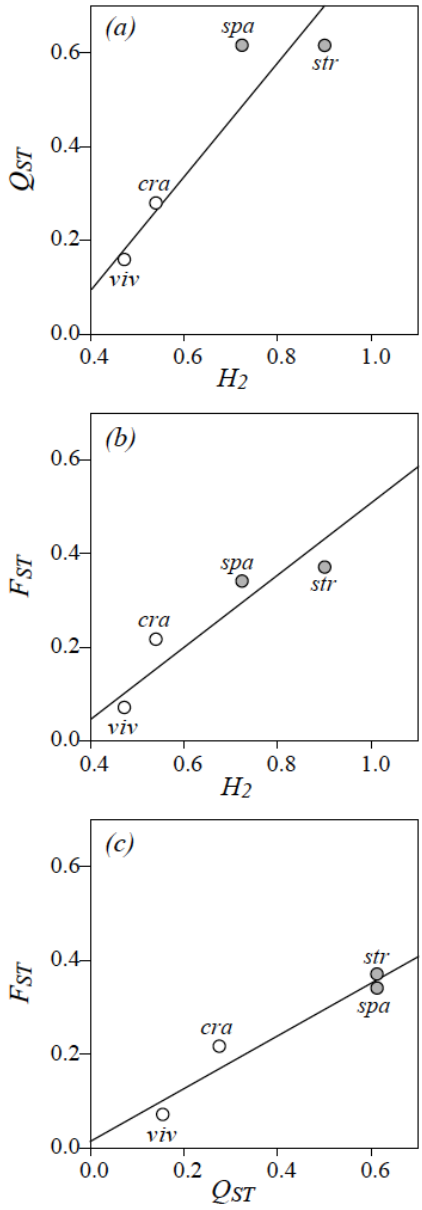

Figure 3.7 Correlations between molecular and quantitative genetic partitioning among the four species. Mean heritability (a-b) and Q_{ST} (a,c) across phenotypes and mean FST (b-c) across loci were plotted. Rare species had significantly lower values of all three statistics. Correlations among these statistics (solid line) were significantly positive.

References

Agren, J., C. G. Oakley, J. K. McKay, J. T. Lovell, and D. W. Schemske. 2013. Genetic mapping of adaptation reveals fitness trade-offs in *Arabidopsis thaliana*. Proceedings of the National Academy of Sciences of the United States of America 110: 21077-21082.

Al-Shehbaz, I. A. 2003. Transfer of most North American species of Arabis to *Boechera* (Brassicaceae). Novon:381-391.

Alexander, P. J., M. D. Windham, J. B. Beck, I. A. Al-Shehbaz, L. Allphin, and C. D. Bailey. 2013. Molecular phylogenetics and taxonomy of the genus *Boechera* and related genera (Brassicaceae: Boechereae). Systematic Botany 38:192-209.

Angert, A. L. 2006. Demography of central and marginal populations of monkeyflowers (*Mimulus cardinalis* and *M. lewisii*). Ecology 87:2014-2025.

Angert, A. L., H. D. Bradshaw, Jr., and D. W. Schemske. 2008. Using experimental evolution to investigate geographic range limits in monkeyflowers. Evolution 62:2660-2675.

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 59:1671-1684.

Baker, H. G., and G. Stebbins. 1965. Characteristics and modes of origin of weeds. Pp 147-168 in H. G. Baker and G. Stebbins, ed. The genetics of colonizing species. Academic Press.

Baskauf, C. J. 2001. Examining rarity through comparisons with widespread congeners: a genetic and ecophysiological example from limestone glade endemics. Castanea:126-133.

Baskauf, C. J. and W. G. Eickmeier. 1994. Comparative ecophysiology of a rare and a widespread species of *Echinacea* (Asteraceae). American Journal of Botany:958-964.

Beck, J. B., P. J. Alexander, L. Allphin, I. A. Al-Shehbaz, C. Rushworth, C. D. Bailey, and M. D. Windham. 2012. Does hybridization drive the transition to asexuality in diploid *Boechera*? Evolution 66:985-995.

Blows, M. W. and A. A. Hoffmann. 2005. A reassessment of genetic limits to evolutionary change. Ecology 86:1371-1384.

Bolnick, D. I., R. Svanback, M. S. Araujo, and L. Persson. 2007. Comparative support for the niche variation hypothesis that more generalized populations also are more heterogeneous. Proceedings of the National Academy of Sciences of the United States of America 104:10075-10079.

Bradley, B. P. 1978. Genetic and physiological adaptation of the copepod *Eurytemora affinis* to seasonal temperatures. Genetics 90:193-205.

Brown, J. H. 1984. On the relationship between abundance and distribution of species. American Naturalist 124: 255-279.

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 27:597-623.

Chevin, L. M. and R. Lande. 2011. Adaptation to marginal habitats by evolution of increased phenotypic plasticity. Journal of Evolutionary Biology 24:1462-1476.

Cole, C. T. 2003. Genetic variation in rare and common plants. Annual Review of Ecology, Evolution, and Systematics 34:213-237.

Darwin, C. 1872. The origin of species. London: Murray.

David, P., B. Pujol, F. Viard, V. Castella, and J. Goudet. 2007. Reliable selfing rate estimates from imperfect population genetic data. Molecular Ecology 16:2474-2487.

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. American Naturalist 178 Suppl 1:S26-43.

Etterson, J. R. 2004. Evolutionary potential of *Chamaecrista fasciculata* in relation to climate change. I. Clinal patterns of selection along an environmental gradient in the Great Plains. Evolution 58:1446-1456.

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

Falconer, D. S. and T. F. C. Mackay. 1996. Introduction to Quantitative Genetics,4th Edn. Longman, Harlow.

Frankham, R., K. Lees, M. E. Montgomery, P. R. England, E. H. Lowe, and D. A. Briscoe. 1999. Do population size bottlenecks reduce evolutionary potential? Animal Conservation 2:255-260.

Futuyma, D. J. and G. Moreno. 1988. The evolution of ecological specialization. Annual Review of Ecology and Systematics 19:207-233.

Gaston, K. J. 1996. Species-range-size distributions: patterns, mechanisms and implications. Trends in Ecology & Evolution 11:197-201.

Gaston, K. J. 2009. Geographic range limits: achieving synthesis. Proceedings of the Royal Society B-Biological Sciences 276:1395-1406.

Gaston, K. J. and J. I. Spicer. 2001. The relationship between range size and niche breadth: a test using five species of *Gammarus* (Amphipoda). Global Ecology and Biogeography 10:179-188.

Geber, M. A. 2011. Ecological and evolutionary limits to species geographic ranges. American Naturalist 178 Suppl 1:S1-5.

Ghalambor, C. K., J. K. McKay, S. P. Carroll, and D. N. Reznick. 2007. Adaptive versus non-adaptive phenotypic plasticity and the potential for contemporary adaptation in new environments. Functional Ecology 21:394-407.

Gitzendanner, M. A. and P. S. Soltis. 2000. Patterns of genetic variation in rare and widespread plant congeners. American Journal of Botany 87:783-792.

Goudet, J. 2005. HIERFSTAT, a package for R to compute and test hierarchical F-statistics. Molecular Ecology Notes 5:184-186.

Heschel, M. S., S. E. Sultan, S. Glover, and D. Sloan. 2004. Population differentiation and plastic responses to drought stress in the generalist annual *Polygonum persicaria*. International Journal of Plant Sciences 165:817-824.

Holt, R. D. 2003. On the evolutionary ecology of species' ranges. Evolutionary Ecology Research 5:159-178.

Hutchinson, G. E. 1957. Concluding Remarks. Cold Spring Harbor Symposia on Quantitative Biology 22:415-427.

Jordan, G. J., B. M. Potts, and R. J. Wiltshire. 1999. Strong, independent, quantitative genetic control of the timing of vegetative phase change and first flowering in *Eucalyptus globulus ssp. Globulus* (tasmanian blue gum). Heredity 83:179-187.

Karron, J. D. 1987. A comparison of levels of genetic polymorphism and self-compatibility in geographically restricted and widespread plant congeners. Evolutionary Ecology 1:47-58.

Kellermann, V., B. van Heerwaarden, C. M. Sgrò, and A. A. Hoffmann. 2009. Fundamental evolutionary limits in ecological traits drive *Drosophila* species distributions. Science 325:1244-1246.

Kruckeberg, A. R. and D. Rabinowitz. 1985. Biological aspects of endemism in higher plants. Annual Review of Ecology and Systematics 16:447-479.

Lande, R. 1993. Risks of population extinction from demographic and environmental stochasticity and random catastrophes. American Naturalist 142: 911-927.

Lavergne, S., J. D. Thompson, E. Garnier, and M. Debussche. 2004. The biology and ecology of narrow endemic and widespread plants: a comparative study of trait variation in 20 congeneric pairs. Oikos 107:505-518.

Leinonen, P. H., S. Sandring, B. Quilot, M. J. Clauss, T. Mitchell-Olds, J. Agren, and O. Savolainen. 2009. Local adaptation in European populations of *Arabidopsis lyrata* (Brassicaceae). Am J Bot 96:1129-1137.

Levins, R. 1968. Evolution in Changing Environments: Some Theoretical Explorations. Princeton University Press.

Lloyd, K. M., W. G. Lee, and J. B. Wilson. 2002. Growth and reproduction of New Zealand *Acaena* (Rosaceae) species in relation to rarity and commonness. New Zealand Journal of Ecology 26:149-160.

Loomis, Eli S. and L. Fishman. 2009. A Continent-Wide Clone: Population Genetic Variation of the Invasive Plant, *Hieracium aurantiacum* (Orange Hawkweed; Asteraceae), in North America. International Journal of Plant Sciences 170:759-765.

Lovell, J. T. 2011. *Boechera* Summit 2011. Molecular Ecology 20:4840-4842.

Lovell, J., O. Aliyu, M. Mau, M. Schranz, M. Koch, C. Kiefer, B. Song, T. Mitchell-Olds, and T. Sharbel. 2013a. On the origin and evolution of apomixis in *Boechera*. Plant reproduction 26: 309-315.

Lovell, J. T., T. E. Juenger, S. D. Michaels, J. R. Lasky, A. Platt, J. H. Richards, X. Yu, H. M. Easlon, S. Sen, and J. K. McKay. 2013b. Pleiotropy of *FRIGIDA* enhances the potential for multivariate adaptation. Proceedings of the Royal Society of London. Series B: Biological Sciences 280:20131043.

MacArthur, R. H. 1972. Geographical ecology: patterns in the distribution of species. Princeton University Press.

Matzk, F., A. Meister, and I. Schubert. 2000. An efficient screen for reproductive pathways using mature seeds of monocots and dicots. The Plant Journal 21:97-108.

McKay, J. K., J. G. Bishop, J.-Z. Lin, J. H. Richards, A. Sala, and T. Mitchell-Olds. 2001. Local adaptation across a climatic gradient despite small effective population size in the rare sapphire rockcress. Proceedings of the Royal Society of London. Series B: Biological Sciences 268:1715-1721.

Nakazato, T., M. Bogonovich, and L. C. Moyle. 2008. Environmental factors predict adaptive phenotypic differentiation within and between two wild Andean tomatoes. Evolution 62:774-792.

Payne, J. L. and S. Finnegan. 2007. The effect of geographic range on extinction risk during background and mass extinction. Proceedings of the National Academy of Sciences of the United States of America 104:10506-10511.

Pelletier, F., D. Reale, D. Garant, D. W. Coltman, and M. Festa-Bianchet. 2007. Selection on heritable seasonal phenotypic plasticity of body mass. Evolution 61:1969-1979.

Pohlman, C. L., A. B. Nicotra, and B. R. Murray. 2005. Geographic range size, seedling ecophysiology and phenotypic plasticity in Australian *Acacia* species. Journal of Biogeography 32:341-351.

Price, T. D., A. Qvarnstrom, and D. E. Irwin. 2003. The role of phenotypic plasticity in driving genetic evolution. Proceedings of the Royal Society of London. Series B: Biological Sciences 270:1433-1440.

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

Pulliam, H. R. 2000. On the relationship between niche and distribution. Ecology Letters 3:349-361.

R Core Team. 2013. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rabinowitz, D. 1981. Seven forms of rarity. Pp. 205-217 in H. Synge, ed. The Biological Aspects of Rare Plant Conservation. Riley.

Reed, D. H. and R. Frankham. 2001. How closely correlated are molecular and quantitative measures of genetic variation? a meta-analysis. Evolution 55:1095-1103.

Relyea, R. A. 2002. Costs of phenotypic plasticity. The American Naturalist 159:272-282.

Richards, A. E., A. Shapcott, J. Playford, B. Morrison, C. Critchley, and S. Schmidt. 2003. Physiological profiles of restricted endemic plants and their widespread congenors in the North Queensland wet tropics, Australia. Biological Conservation 111:41-52.

Roy, B. 1995. The breeding systems of six species of *Arabis* (Brassicaceae). American Journal of Botany 82: 869-877.

Rushworth, C. A., B. H. Song, C. R. Lee, and T. Mitchell-Olds. 2011. *Boechera*, a model system for ecological genomics. Molecular Ecology 20:4843-4857.

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 40:415-436.

Slatyer, R. A., M. Hirst, and J. P. Sexton. 2013. Niche breadth predicts geographical range size: A general ecological pattern. Ecology Letters 16: 1104-1114.

Song, B. H. and T. Mitchell-Olds. 2007. High genetic diversity and population differentiation in *Boechera fecunda*, a rare relative of *Arabidopsis*. Molecular Ecology 16:4079-4088.

Spicer, J. and K. Gaston. 2009. Physiological Diversity: Ecological Implications. Wiley.

Takahashi, Y., H. Takahashi, and M. Maki. 2011. Comparison of genetic variation and differentiation using microsatellite markers among three rare threatened and one widespread toad lily species of *Tricyrtis sect. Flavae* (Convallariaceae) in Japan. Plant Species Biology 26:13-23.

Van de Pol, M. and J. Wright. 2009. A simple method for distinguishing within- versus between-subject effects using mixed models. Animal Behaviour 77: 753-758.

Vitt, P. and K. Havens. 2004. Integrating quantitative genetics into ex situ conservation and restoration practices. Ex situ plant conservation: supporting species survival in the wild. Island Press, Washington:286-304.

Warren, D. L., R. E. Glor, and M. Turelli. 2008. Environmental niche equivalency versus conservatism: quantitative approaches to niche evolution. Evolution 62:2868-2883.

Whitlock, M. C. 1996. The red queen beats the jack-of-all-trades: the limitations on the evolution of phenotypic plasticity and niche breadth. American Naturalist 148: S65-S77.

Whitlock, M. C. 2008. Evolutionary inference from QST. Molecular Ecology 17:1885-1896.

Windham, M. D. and I. A. Al-Shehbaz. 2006. New and noteworthy species of *Boechera* (Brassicaceae) I: sexual diploids. Harvard Papers in Botany 11:61-88.

Chapter 4

MATING SYSTEM AND ENVIRONMENTAL VARIATION DRIVE PATTERNS OF ADAPTATION IN *BOECHERA SPATIFOLIA* (BRASSICACEAE)

Overview

Determining the relative contribution of population genetic processes to the distribution of natural variation is a major goal of evolutionary biology. Here, we take advantage of variation in mating system to test the hypothesis that local adaptation is constrained by asexual reproduction. We explored patterns of variation in ecological traits and genome-wide molecular markers in *Boechera spatifolia* (Brassicaceae), a species that contains both apomictic (asexual) and sexual individuals. Using a combination of quantitative genetics, neutral genetic (SSR) and genome-wide single nucleotide polymorphism, we assessed the hypothesis that asexual lineages should have reduced signatures of adaptation relative to sexual conspecifics. All three measures (traits, SSRs, SNPs) demonstrated that apomicts are divergent from sexuals, regardless of population location. Additionally, phylogenetic clustering revealed that the apomictic group shared a single common ancestor, separate from sexuals. Across the landscape, sexual populations showed very strong congruence between genome-wide SNP variation and latitude (r^2>0.9), indicating that sexual populations have differentiated across an environmental gradient. Furthermore, flowering time and growth rate, as assessed in a common garden, strongly covary with the elevation and latitude of the source population. Despite a wide geographic distribution that largely overlaps with sexual populations, there was little evidence for differentiation in molecular markers or quantitative characters among apomictic populations. Combined, these

data indicated that, in contrast to asexual populations, sexual populations show evidence of local adaptation.

Introduction:

Adaptation to local environmental conditions is the best and longest studied manifestation of an evolutionary response to selection in nature (Clausen *et al.* 1940; Kawecki & Ebert 2004; Turesson 1922). Documented by reciprocal transplant experiments, common garden studies, and ex-situ experimentation (Agren & Schemske 2012; Fournier-Level *et al.* 2011), local adaptation is often viewed as ubiquitous (Kawecki & Ebert 2004). Physiological adaptation to local conditions is especially important, and well documented, in sessile organisms such as plants (Berry & Bjorkman 1980; Turesson 1930). Despite these examples, adaptation is commonly constrained by genetic and biological factors (Arnold 1992; Etterson & Shaw 2001; van Kleunen & Fischer 2005), and drift and other neutral processes may drive evolution, even under strong selection regimes. Determining the mechanisms and ecological manifestations of adaptive constraint remains a major goal of evolutionary biology.

The relative contribution of genetic drift and selection can also be inferred by comparing patterns of genomic and phenotypic differentiation with landscape and geographic variables (Manel *et al.* 2003). If genomic and/or phenotypic differentiation is largely correlated with geographic distance (isolation-by-distance: IBD), a balance of gene flow and genetic drift is the primary mode of evolution (Slatkin 1993). However, if differentiation is associated with environmental variables, evolution may be driven by responses to natural selection (Lasky *et al.* 2012; Lee & Mitchell-Olds 2013).

The mating system of a population can affect the relative likelihood that evolution is due to either drift or responses to natural selection (Charlesworth & Wright 2001). For example, lower effective recombination rates in self-pollinating species increases the extent of linkage disequilibria across the genome and reduces the efficacy of selection (Conway *et al.* 1999; Nordborg 2000; Qiu *et al.* 2011). Asexual lineages offer an extreme example where the entire genome is in linkage disequilibrium due to a lack of recombination (Henry *et al.* 2012). Genome-wide interference among loci in asexual lineages constrains adaptation by reducing population genetic variation and the efficacy of selection (Barton & Charlesworth 1998; Hill & Robertson 1966; Lynch & Blanchard 1998; Otto & Lenormand 2002).

Genetic constraints and reduced evolutionary potential of asexual lineages forms the basis for many theories that explain the maintenance and evolution of sex, despite the fitness costs of sexual reproduction (Barton 2010; Burt 2000; Nautiyal *et al.* 2002). For example, sexual reproducing natural enemies may gain an advantage over asexual hosts ("red queen" dynamics, reviewed by Lively (2010)) or slightly deleterious mutations may not be purged by selection in asexual lineages (e.g. "muller's ratchet") (Barton 2010).

Despite a large body of theoretical work, the manifestation of reduced responses to selection in asexual lineages has only been tested in a handful of multicellular organisms and no vascular plants. This lack of experimentation is due, in part, to the fixation of hybridity and/or polyploidy in all obligately asexual plant species (Mogie 1992). These factors, which covary with mating system, confound any experimentation. *Boechera spatifolia*, a close relative of the model plant, *Arabidopsis thaliana*, is unique among plants because it contains both diploid sexual and diploid apomictic (asexual reproduction through seed) lineages. As such, comparisons

between mating systems can be conducted without the confounding effects of hybridization or polyploidy.

Here, we analyze genomic (~10,000 SNPs), neutral molecular genetic (14 SSR markers) and quantitative genetic differentiation of *B. spatifolia* along environmental gradients in both apomictic and sexual lineages. We test the hypothesis that divergence of apomictic populations is caused primarily by drift, while that of sexual populations is dominated by natural selection. This hypothesis predicts that: 1) apomictic lineages will display low quantitative genetic structure among populations and little phenotypic differentiation along environmental gradients; and 2) sexual lineages will exhibit signatures of adaptive evolution, including strong genomic and quantitative genetic structure, and covariance of phenotypic values with environmental variables.

Methods:

Plant Material

Boechera spatifolia (Rydb.) is a diploid (2n=2x=14), winter annual (or short lived perennial) species of the central Rocky Mountains, USA (Windham & Al-Shehbaz 2006). Small stature and short life span make *B. spatifolia* a promising species for combining population genetics and field trials. We collected seeds from maternal plants from 30 populations across Colorado and southern Wyoming, USA over the summers of 2011 and 2012 (Fig. 4.1, Table 2.1). These populations cover >90% of the geographic range of herbaria collections. To reduce the potential of collecting cryptic hybrids, where possible, populations were chosen from locations used in the systematic documentation of *B. spatifolia* (Alexander *et al.* (2013); P.

Alexander, pers. comm.). We characterized the mating system of all seed families using the flow cytometric seed screen (FCSS: (Matzk *et al.* 2000)), following methods of both Lovell et al. (2013a) and Aliyu *et al.* (2010).

Phenotypic analysis- glasshouse and field experimentation

We grew four replicates (sibs) from each of 190 seed families from 29 source populations (n = 1564 plants). Germination, planting and general growth conditions follow Lovell and McKay (in review). We simulated the natural temperature and photoperiod experienced by *B. spatifolia* plants in the wild, but on an accelerated time-scale: four weeks of fall and winter conditions, then steadily lengthening days to mimic spring-summer conditions.

We assayed ten phenotypes in the growth chamber. The physiological traits, specific leaf area (*SLA*) and leaf water content (H_2O content), are related to stress tolerance, especially drought (e.g. Nautiyal *et al.* (2002)). To measure growth, we calculated the relative growth rate of both stem elongation in "spring" conditions (*GRht*) and of total rosette leaf area (*GRla*, following Lovell *et al.* (2013b)). Rosette morphology, which is associated with adaptive differentiation in *Boechera* (Lee & Mitchell-Olds 2013), was assessed by the number and size of leaves and the height of the whole rosette. Finally, phenology was measured as the days to anthesis ("anthesis"), the height of the bolting structure and the days from the end of vernalization to the initiation of bolting (*FT*).

We analyzed the growth chamber phenotypes to determine the extent of phenotypic divergence between mating systems. After standardizing and normalizing the growth chamber phenotypes, we calculated breeding values for each family and utilized a discriminant function

analysis to determine multivariate differences between mating systems; significance was assessed via "Pillai's trace" multivariate tests. All analyses were conducted in JMP 10.1 Pro (SAS Corporation, Cary, NC, USA).

For two traits, *FT* and *GRla*, we correlated the mean phenotypic values for each population with the latitude and elevation of the site where the population was collected. We chose to focus on *FT* and *GRla* because these traits vary considerably between differentially adapted populations of other species (Angert *et al.* 2009; Banta *et al.* 2012; Lee & Mitchell-Olds 2013; Wilczek *et al.* 2009). To determine the extent of ecological differentiation we correlated population breeding values of *FT* with *GRla* and the environmental variables, latitude and elevation. These linear regressions were conducted separately for apomictic and sexual lineages in R Environment for Statistical Computing, version 2.15.1 (R Core Team 2013). We also analyzed the extent of ecological differentiation between apomixis and sex by comparing regression coefficients and r^2 values. To correct for differences in power due to sample size (9 apomictic, 27 sexual populations), we randomly subsampled the sexuals into sets of nine populations (after excluding groups of geographically proximate populations) and re-conducted the analysis 5000 times. The relative likelihood of a true difference between the two mating systems was calculated as the proportion of subsamples that were more extreme than the apomictic r^2 of each trait and correlation coefficients (r) between *GRla* and *FT*.

To further explore ecological trait variation, we planted a subset of 12 populations in two experimental gardens, one at the upper and another at the lower elevation range margin of *B. spatifolia*. The "sub-alpine" garden was located at 3022m (40.0364°N, -105.5442°W), within 100m of the Niwot Ridge C1 weather/climate station at the University of Colorado Mountain Research Station, Nederland, CO, USA. The "foothills" site was located in Fort Collins, CO

(40.5702°N, 105.0630°W) at 1519m (Fig. 4.1a). In each site, we planted 6 seedlings from 8 families of 12 populations (experiment-wise n = 1152) in September 2011. These plants overwintered as rosettes and flowered after snowmelt.

Fruit (silique) length and number of seeds/silique were calculated for >10 individuals from each population and garden. For each individual, we counted the number of siliques produced and calculated total seed number (absolute fitness) by multiplying silique number by the number of seeds/fruit. These calculations were conducted separately for each garden and source population. To calculate relative fitness, the absolute fitness measure from each individual was divided by the mean absolute fitness within each garden. Genotype means were calculated as the mean trait value of each family. We calculated linear selection coefficients of *GRla* and *FT* in both gardens following Lande & Arnold (1983) using quantitile normalized genotypic means

Population genetic analysis- neutral molecular structure and diversity

DNA was extracted from lyophilized leaf tissue using the ChargeSwitch gDNA plant kit (Invitrogen Corp. Carlsbad, CA, USA). We followed PCR and genotyping protocols optimized by Beck *et al.* (2012) for 14 SSR markers known to amplify well across most *Boechera* species. Basic statistics for each marker were calculated in the R package "adegenet" (Jombart 2008) (Table 2.2).

We analyzed the patterns of genetic differentiation between apomicts and sexuals by comparing the multi-locus (SSR) principal component (PC) positions and phylogenetic clustering between mating systems. A "Bruvo" distance matrix (Bruvo *et al.* 2004), which is

robust to differences in ploidy and mating systems, was used to calculate PC positions of all genotypes in the R package "polysat" (Clark & Jasieniuk 2011). We conducted phylogenetic clustering in the "ape" package (Paradis *et al.* 2004) by generating a neighbor-joining tree from this distance matrix rooted to an accession of *B. stricta* (Natrona County, WY, USA, 42.74451°N, 106.32512°W).

Analysis of genome-wide single nucleotide polymorphism

We extracted DNA of one sexual individual per population (as assessed by flow cytometry) using the Qiagen DNEasy Plant Miniprep kit (Qiagen Corp. Germantown, MD, USA) following manufacturer protocols (qiagen.com). Samples were analyzed at the Institute for Genomic Diversity, Cornell University, Ithaca, NY, USA, through the "genotyping-by-sequencing" (GBS) analytical protocol (Elshire *et al.* 2011). Single nucleotide polymorphisms (SNPs) were called using the published *A. lyrata* genome (Hu *et al.* 2011) as a reference. We analyzed GBS data by processing raw SNP calls in TASSEL (Bradbury *et al.* 2007). A UPGMA tree was constructed from the resulting distance matrix.

Results

Of the 30 populations surveyed, 19 were obligately sexual, 2 were apomictic and 9 were mixed (Table 2.1). Apomixis was found exclusively in the southern 3/4 of the *B. spatifolia* range, but was highly dispersed within this region (Fig. 4.1a).

Genomic variation was structured by mating system and latitude

Genotyping-by-sequencing of one individual from each of the 30 populations (3 apomicts, 27 sexuals) resulted in ca.82 million barcoded high quality reads with 9126 SNPs mapping to all eight *A. lyrata* chromosomes. While efforts were made to screen only sexual genotypes, three apomictic individuals ("Poso", "Alvarado", "Hondo") were also genotyped by GBS (Fig. 4.1a). Principal component (PC) analysis demonstrated that these three apomictic lines were highly divergent from each other and from the sexual accessions, which formed a tight cluster (Fig. 4.1b). "Poso", the only triploid individual analyzed was widely differentiated from its 2x apomictic counterparts. Tree-based analyses indicate that the three highly divergent apomicts form a single cluster distinct from and potentially ancestral to the sexual lineages (Fig. 4.2). Clustering analysis of SSR variation further supports this conclusion (Fig. 4.3).

The GBS PC scores of the sexual individuals showed extremely strong congruence with the environmental distribution of the sampled populations (Fig. 4.1c). In particular, latitude (the spatial axis with the greatest variance) was highly collinear with PC axis #1 (n=27, r^2=0.903, P<0.0001; Fig. 4.1d).

Across all SSR genotyped individuals, apomictic lineages were clearly divergent from sexuals. Apomicts were assigned to a single phylogenetic cluster within a neighbor-joining tree, rooted to the widespread relative, *B. stricta* (Fig. 4.4a). This strong genetic differentiation was also documented by principal component analyses (Fig. 4.3). There was a strong absence of population structure in apomictic lineages. All apomictic or mixed populations except "Poso" were composed of individuals found across the genetic diversity of apomicts (Fig. 4.4b). The

absence of structure among apomictic populations was evidenced by much lower F_{ST} (0.1661) than that among sexual populations (0.3958).

Phenotypic divergence between mating systems

Discriminant-function analysis of genotype means of the ten growth chamber phenotypes revealed highly significant differences between mating systems (Fig. 4.5a; Pillai test: $F_{11,180}$=45.08, P<0.0001). Phenotypic differentiation occurred along the first canonical axis and was primarily influenced by rosette architecture. *FT* and other phenological traits, which affected the second canonical axis, did not differ significantly between mating systems.

In addition to the growth chamber analysis, we also surveyed phenotypes in two experimental gardens. Two populations of the 12-population subset analyzed in the field experiment contained both sexual and apomictic lineages. Of the 87 genotypes, seven were apomictic (8%). Comparisons of genotypic means between mating systems after controlling for variation among populations and gardens were consistent with growth chamber analyses: mating system variation was associated with significant differences between rosette growth traits (*GRla*, $F_{1,29}$=23.67, *P*<0.0001), but not phenology ($F_{1,29}$=0.244, *P*=0.63). Interestingly, apomicts have greater *GRla*, but similar *FT*, which leads to slightly elevated fitness ($F_{1,29}$=23.67, *P*=0.058).

Selection on FT and GRla was conserved across populations and mating systems

To measure genetic correlations and the strength of selection, we compared correlation structures of the 80 sexual genotypic means of *FT*, *GRla* and relative fitness measured in the

experimental gardens. We detected strong directional selection on *GRla* and *FT* in both field sites, as evidenced by highly significant linear selection gradients (Table 2.3, Fig. 4.6). We estimated the selection gradient of *GRla* in the foothills site to be more than twice as strong as that found in the subalpine site (Table 2.3). Alternately, selection on *FT* was slightly weaker in the sub-alpine site than the foothills garden (Table 2.3). Despite differences in the strength of selection between sites, the direction and linear shape of the selection gradients are strongly conserved: early flowering and high growth rate lines were favored in both environments. Non-linear selection gradients were not significant for any site-trait combination. Apomictic phenotypes were also subjected to these selection gradients as evidenced by similar residual variance to that of sympatric sexuals in both gardens (Sub-alpine: t_{14}=1.25, *P*>0.1, Foothills: t_{12}=0.47, *P* >0.1).

Sexual, but not apomictic, phenotypic variance was highly correlated with environmental variables

We associated population-level breeding values (growth chamber) of *FT* and *GRla* with two geographical variables, latitude and elevation, separately for each mating system. Among sexual populations, flowering time was strongly correlated with the latitude of the source population (r=0.721, *P* <0.0001), but not elevation (*P* =0.202; Fig. 4.5b). Growth rate was driven by latitude of the source population (r=-0.434, *P* =0.0235) and was also significantly associated with elevation (r=0.395, *P* =0.024; Fig. 4.5c). Growth rate and flowering time were negatively correlated (n=26, r=-0.526, *P*=0.0058).

In contrast, among apomictic populations, no significant associations between phenotypes and elevation or latitude were found (*FT*-elevation: P =0.719; FT-latitude: P =0.574; *GRla*-elevation: p=0.393; *GRla*-latitude: p=0.261, Fig. 4.5d-e). Furthermore, the two phenotypes were positively, but not significantly correlated (P =0.404).

To overcome the difference in power between sexual and apomictic regressions, we subsampled 5000 nine-population sets from the sexual distribution and recalculated the model statistics. For >99.9% of the *FT* and >82.4% of *GRla* subsamples, sexual r^2 values exceeded apomicts. Correlation coefficients between *FT* and *GRla* were negative in >98.6% of all sexual subsamples (Fig. 4.7). These data indicated that the observed difference between mating systems were not a statistical artifact.

These correlations were also found among sexual populations planted in the field gardens. In the sub-alpine site, elevation of the source population was negatively associated with *FT* (higher populations flower earlier) but positively with *GRla* (higher populations grow faster; Table 2.4). Latitude only marginally affected *FT* and was non-significantly associated with *GRla*. In the foothills site, latitude strongly affected *FT* (more northern populations flower later) and marginally affects *GRla* (northern populations grow faster). Elevation was strongly associated with *GRla* (higher elevation populations grow faster) but not *FT* (Table 2.4).

Discussion

The evolution and overwhelming prevalence of sexual reproduction in multicellular organisms has formed one of the fundamental questions in evolutionary biology: what factors have permitted the maintenance of sexual reproduction despite its apparent costs? The major

hypotheses, such as “red queen” dynamics, or the accumulation of deleterious mutations (“muller’s ratchet”), have invoked an assumption that, with all else being equal, asexual lineages will exhibit a relatively weaker response to selection than sexual lineages (Barton 2010; Burt 2000; Butlin 2002). Experimental evolution of rapidly cycling organisms and *in silico* inquiry (Otto & Lenormand 2002) has confirmed that evolution in asexual lineages is in fact less affected by selection than sexual lineages (Engelstadter 2008). However, these comparisons are not possible in higher plants or most multicellular taxa because polyploidy, hybridity or other confounding factors are fixed in asexual relatives of sexual diploids (Carman 1997; Paun *et al.* 2006; Robertson *et al.* 2010).

Here we presented ecological genomic and quantitative genetic comparisons between asexual and sexual lineages of *Boechera spatifolia*, a species that contains both diploid sexual and diploid apomictic lineages that were neither genetically nor phenotypically identified as hybrids. This situation is unique among angiosperms and may provide an opportunity to investigate the causes and consequences of the evolution of asexual reproduction.

By collecting populations from across the genetic and geographic distribution of *B. spatifolia*, we were able to (1) compare the degree of genetic and physiological divergence among mating systems, and (2) assess what factors drive population differentiation. Our data showed strong evidence of molecular and quantitative differentiation among sexual populations across environmental gradients, implicating adaptation as the primary cause of population structure. In contrast, apomictic populations were not structured in a consistent manner across environmental gradients, had high within population diversity, and demonstrated non-adaptive phenotypic correlations. Combined, these data indicate that apomictic lineages were relatively less affected by selection.

Biogeography and evolutionary history of apomixis

Geographic parthenogenesis, where asexual lineages exhibit divergent ecological characteristics and geographic ranges compared to sexual relatives, is nearly omnipresent across plants (Horandl 2008; Kearney 2005; Mráz *et al.* 2009). However, our data showed limited evidence of geographic parthenogenesis. The geographic range of apomictic *B. spatifolia* was broad, and overlapped with approximately 70% of the sexual distribution (Fig. 4.1). While apomixis was on average found in higher elevation and lower latitude sites than sexual (Fig. 4.1, Table 2.1), apomixis was still widely distributed across elevation (>1100 m) and geographic area (>20,000 km^2). Taken together with evidence for a single origin of the apomictic lineages (Fig. 4.1b, Fig. 4.4), this implies a geographic spread of apomixis over much of the range of sexuals.

Across the genus *Boechera*, it has been hypothesized that apomictic lineages were recurrently derived by independent hybridization events (Beck *et al.* 2012; Dobeš *et al.* 2007); however, it appeared that *B. spatifolia* has not followed this trend. The phylogenetic and phenotypic analyses here showed strong structure among mating systems. Genome-wide genotyping indicated that apomictic lineages were highly diverged from their sexual relatives. This observation was confirmed by multi-locus SSR data, which indicated that apomictic lineages were a monophyletic group within the diversity of sexual *B. spatifolia* (Fig. 4.1b, Fig. 4.4a). Despite geographic proximity to sexual conspecifics, and wide distances among populations, all apomictic populations clustered together.

Many studies have documented phenotypic and ecological differentiation between asexual and sexual lineages (Horandl 2008; Mráz *et al.* 2009). Here, apomictic *B. spatifolia* lineages also displayed ecological trait divergence from sexuals, especially in rosette

morphological and growth rate phenotypes (Fig. 4.5a). While on a broad geographic scale these phenotypic shifts did not coincide with altered ecological niches, micro-climactic differentiation may have accompanied the divergent phenotypes of apomicts.

Relative impacts of selection and drift in apomictic and sexual lineages

Selection gradients, inferred by correlated genotype means on relative fitness, demonstrated that selection is acting in a similar fashion on populations across a wide range of environments. At both the high and low elevation extremes of the *B. spatifolia* range, those lines that flowered early and/or grew vegetative tissue most quickly displayed the greatest relative fitness. *GRla* was under stronger selection in the more benign "foothills" site, while *FT* was under the strongest selection in the highly stressful sub-alpine site. Despite a small sample size, genotype means of apomicts followed these correlations as strongly as sexuals. Given this result and the highly overlapping distributions of the two mating systems, it appeared that apomicts were subject to similar selection regimes as sexual population.

Given comparable selection regimes, the distribution of genetic variance across the landscape can permit inference about the relative effect of selection and neutral process on the evolution of populations. Here, we conducted two analyses to assess the hypothesis that the selection is the evolutionary driving force in sexual, but not apomictic populations. First, apomictic genotypes deviated from a strong ecological correlation among traits. Selection in nature causes a negative ecological correlation between flowering and growth rate (Angert *et al.* 2009; Lovell *et al.* 2013b). This adaptive correlation was strongly present in sexuals; however, the *FT*-*GRla* correlation was positive, but not significantly so, in apomicts. The non-adaptive

sign of the correlation was indicative of a deviation from a selectively advantageous suite of trait values. Second, while sexual population breeding values were highly correlated with the environmental characteristics of the local habitat, the same is not true of apomicts. These comparisons were partially affected by differential power of regressions within apomicts (n=9) and sexuals (n=27). However, our results held true, even after subsampling the sexuals into groups of nine (with similar ranges of latitude and elevation). These results indicated that, even when controlling for sample size, phenotypes of apomicts showed a weaker association with environmental factors than those of sexuals.

Population and landscape genomics of B. spatifolia

Despite strong divergence between mating systems, there was very little genetic structure among apomictic populations (Fig. 4.5b). The cause of decreased genetic structure, even across wide environmental gradients, was not clear. Recent gene flow or high rates of dispersal may have geographically distributed diverse genotypes. Alternatively, it was possible that the diverse genotypes in apomictic population were maintained by a relative lack of evolutionary responses to directional or purifying selection. The latter scenario seems more likely because the rate of inconspicuous sexual reproduction in apomictic *Boechera* is thought to be very low (Aliyu *et al.* 2010), and there was no obvious difference in dispersal capability between mating systems.

Across our study system, latitude was a major driver of climactic variation. Given our 470km (~4 degrees of latitude) sampling gradient, it was not surprising that we observed a very strong genomic cline across latitudes. In fact, latitude explained >46% of the total genomic variance from GBS. However, latitude was also a predictor of geographic distance. Isolation-by-

distance, which is driven primarily by neutral processes of drift and gene flow (Slatkin 1987; Slatkin 1993), may have also promoted this cline.

Conclusions

Apomixis has played a major role in the evolution of *Boechera* (Beck *et al.* 2012; Lovell *et al.* 2013a). Our data indicated a single phylogenetic origin of apomixis in *B. spatifolia*. Since divergence, apomictic lineages experienced relatively less quantitative and molecular genetic differentiation among populations than sexuals. More importantly, apomictic population divergence was not correlated with environmental variation, and co-variation among traits was in the opposite direction than has been shown to be adaptive in other species. Conversely, genomic structure and quantitative traits of sexual lineages were highly correlated with latitude and to a lesser extent elevation. The negative growth rate-flowering time correlation is adaptive in *A. thaliana* and is conserved in sexual *B. spatifolia*, but is not present among apomictic lineages. Combined, these data pointed to a lack of adaptive evolution in apomictic relative to sexual *B. spatifolia* lineages.

We presented data on a previously un-studied species that displays remarkable physiological and ecological diversity. Data transfer from related model systems (*Arabidopsis thaliana, A. lyrata*), field manipulation and quantitative genetic analyses are simple in this species, making it an ideal system with which to answer questions related to adaptive evolution in nature. Furthermore, the presence of sympatric diploid apomictic and sexual lineages in *B. spatifolia* provides a unique opportunity to understand the formative processes and consequences of apomixis.

Tables

Table 4.1 Population descriptions. The 30 sampled population names, geographic position, and total sample size (apomictic/sexual individuals) are presented.

ID	Lat.	Lon.	elev (m)	Mating Sys.	N Fam	% apo	ploidy
103	40.0145	-105.5106	2792	Sex	8	0	2
105	40.1692	-105.4739	2596	Sex	8	0	2
377	38.9373	-106.1466	2886	Sex	8	0	2
Alvarado	38.0775	-105.5637	3031	Mixed	6	0.875	2
Antero	39.0317	-105.9862	2804	Sex	8	0	2
Barr	38.8551	-104.941	2370	Sex	6	0	2
Bellaire	40.7507	-105.6126	2606	Sex	8	0	2
Brookvale	39.6306	-105.4464	2636	Sex	6	0	2
Central	39.8266	-105.5418	2859	Sex	8	0	2
Chicago	39.6834	-105.6487	3087	Mixed	7	0.625	2
Chiquito	37.3723	-106.2677	2663	Mixed	8	0.5	2
Cotopaxi	38.3735	-105.6705	2133	Sex	7	0	2
Cripple	38.7543	-105.281	2474	Mixed	8	0.25	2
Crosier	40.4499	-105.4468	2463	Sex	7	0	2
Crystal	41.1537	-105.1917	2159	Sex	5	0	2
Delnorte	37.6322	-106.3631	2493	Sex	7	0	2
Gardner	37.783	-105.1305	2128	Sex	6	0	2
Goose	39.1762	-105.3843	2803	Sex	8	0	2
Green	38.5108	-106.1842	2722	Mixed	5	0.75	2
Hondo	37.0234	-106.2138	2805	Apomictic	8	1	2
Pinegrove	39.3526	-105.3745	2414	Sex	6	0	2
Poso	37.9081	-106.4276	2904	Apomictic	8	1	3
Prince	38.7135	-106.2216	3294	Mixed	7	0.125	2
Rosita	38.0795	-105.3274	2714	Mixed	5	0.375	2
Round	38.4015	-106.0626	2706	Mixed	7	0.125	2
Royal	38.4531	-105.3226	1942	Mixed	3	0.67	2
Salida	38.4901	-106.0015	2770	Sex	5	0	2
Sanluis	37.1922	-105.4493	2475	Sex	8	0	2
Tiesiding	41.1001	-105.465	2419	Sex	7	0	2
Virginiadale	40.9658	-105.3761	2238	Sex	5	0	2

Table 4.2 Descriptive statistics and primer sequences of 14SSR markers. Number of alleles (*A*), the percent of missing data (*%NA*), observed heterozygosity (H_o), and expected heterozygosity (H_e) are presented for each locus across all sexual accession.

Name	*A*	%NA	H_o	H_e
a1	2	5.699	0.011	0.011
b6	11	0.518	0.094	0.729
bdru266	17	0.518	0.156	0.881
bf11	3	3.627	0.124	0.209
bf15	5	0.000	0.150	0.142
bf18	2	1.036	0.000	0.010
bf19	5	12.435	0.148	0.444
bf20	8	5.181	0.131	0.369
bf3	8	0.000	0.130	0.676
bf9	6	1.036	0.131	0.163
c8	3	6.736	0.161	0.548
e9	5	1.036	0.131	0.163
ice14	2	3.627	0.000	0.072
ice3	7	6.218	0.133	0.684

Table 4.3 Genotypic selection gradients for flowering time (*FT*) and relative growth rate of leaf area (*GRla*) are presented for each garden. Statistical significance of the whole model is reported.

Phenotype (site)	Estimate	*t*	*P*
FT (Foothills)	-0.2994	-3.427	>0.001
FT (Sub-Alpine)	-0.3841	-3.829	>0.001
GRla (Foothills)	0.5707	6.548	>0.001
GRla (Sub-Alpine)	0.2794	2.792	>0.001

Table 4.4 Statistical analyses of relationships among phenotypes and the source latitude/elevation of populations. F-statistics are generated from a mixed effect model controlling for variation within families. The direction of effect is reported, (+) or (-), from associated t-test of all significant relationships.

Site (Phenotype)	Latitude (effect direction)	Elevation (effect direction)
Foothills (FT)	$f_{1,82.81}$= 8.92, p=0.0037 (+)	$f_{1,81.9}$= 0.123, p=0.7265
Foothills (*GRla*)	$f_{1,85.77}$= 3.045, p=0.0845 (+)	$f_{1,84.3}$= 6.39, p=0.0133 (+)
Sub-Alpine (FT)	$f_{1,69.6}$= 2.617, p=0.11 (-)	$f_{1,75.8}$= 6.196, p=0.0150 (-)
Sub-Alpine (*GRla*)	$f_{1,81.6}$= 0.304, p=0.583	$f_{1,84.01}$= 4.36, =0.039 (+)

Figures

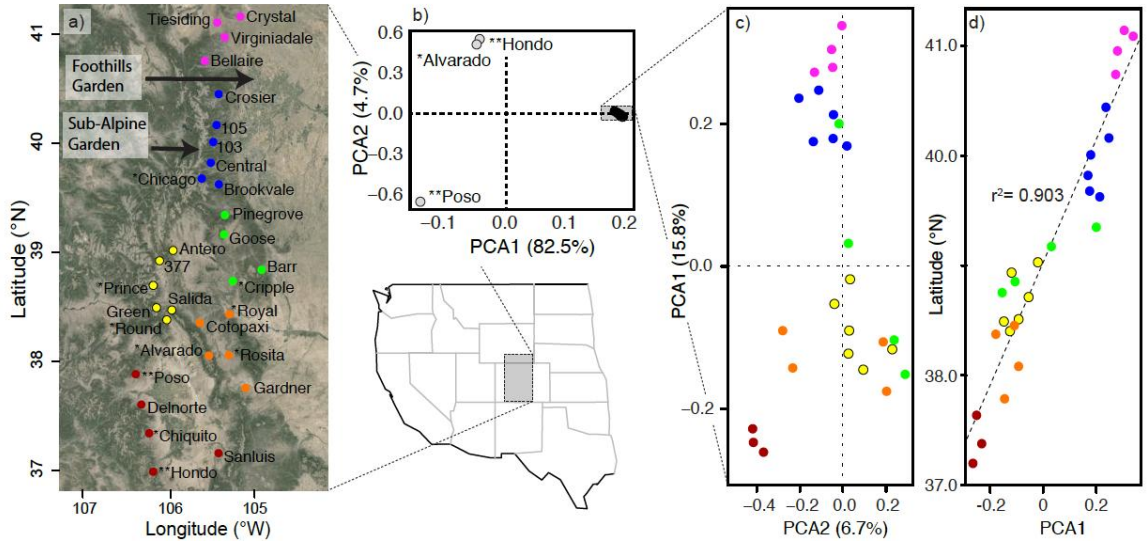

Figure 4.1 Spatial and genetic distribution of sampled populations. a) The geographic positions of seed-source populations are color-coded by mountain range: Northern Foothills (Purple), Mt. Evans/Indian Peaks (Blue), Rampart Range (Green), Collegiate Peaks (Yellow), Sange de Cristo/Wet Mts. (Orange), San Juans Mts./San Luis Valley (Red). Other panels follow this coloring scheme. The mating system is indicated in the population label: (**) obligate apomictic, (*) mixed sexuality, (no asterisk) obligate sexual. b) Position in genomic PC space of all 30 sampled populations. Sexual accessions are filled circles, apomictic accessions are highly diverged, labeled and represented by hollow circles. c) Genetic PC positions of the sexual *B. spatifolia* populations mirror their geographic distribution. d) The correlation between genetic PC axis #1 and latitude is strong and positive; the linear model and reported r^2 is overlayed as a dashed line.

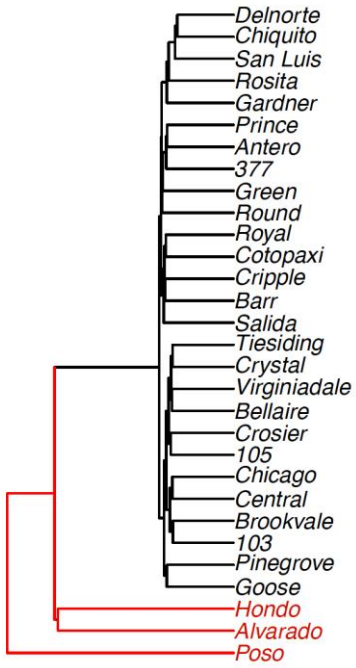

Figure 4.2 UPGMA tree of GBS data from all accessions colored by mating system. Apomictic lineages (red) are highly diverged from sexual lineages (black) and display much longer branches.

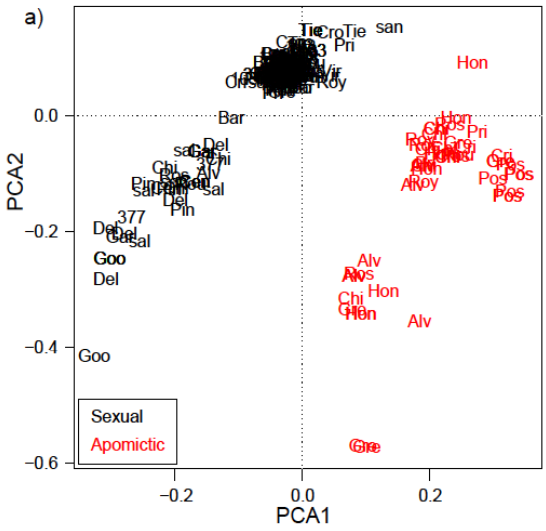

Figure 4.3 SSR clustering of apomictic and sexual individuals. Apomictic lineages (red) are highly diverged from sexual lineages (black).

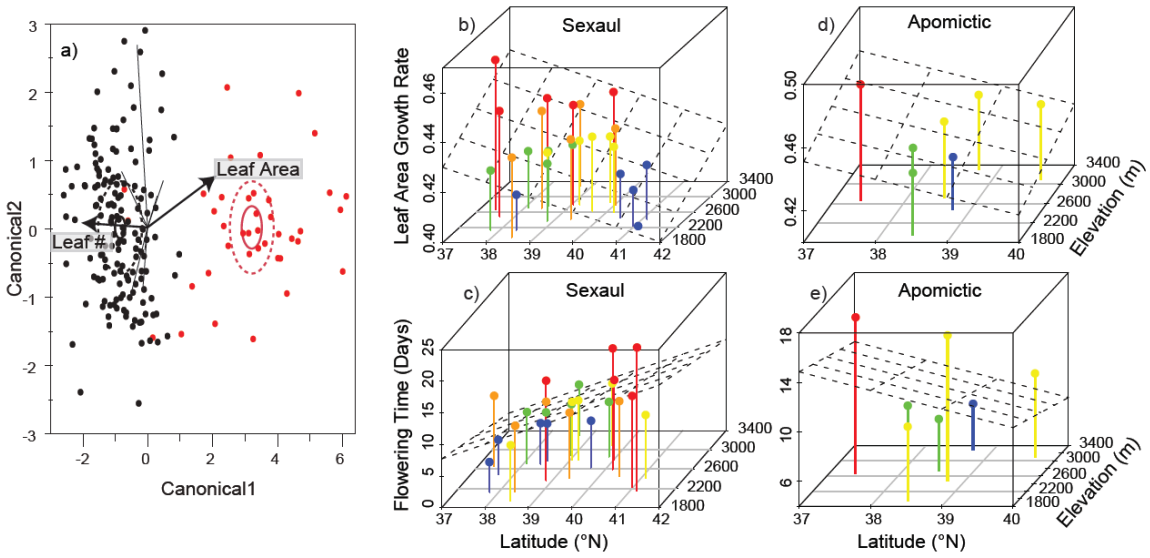

Figure 4.4 Molecular genetic clustering of apomixis in *B. spatifolia*. a) The neutral molecular genetic population structure of *B. spatifolia* demonstrates sharp distinctions between apomictic and sexual individuals. These clusters are also manifested in a neighbor-joining tree rooted to a sexual *B. stricta* accession. Apomictic lineages are depicted by red edges and tip indicators. Sexuals individuals are black. b) Most apomictic populations are not monophyletic. All but the population "Poso" contain individuals found across the genomic diversity of apomictic lineages.

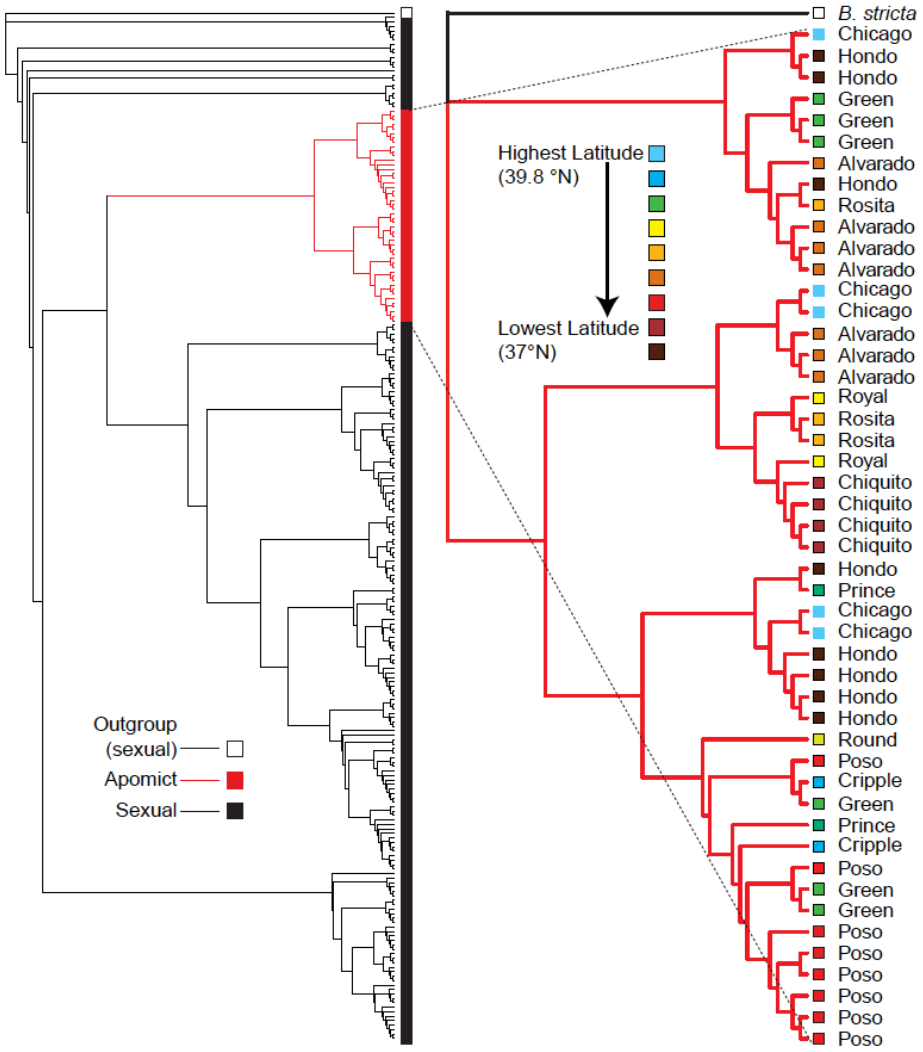

Figure 4.5 Phenotypic variation of sexual and apomictic lineages across the landscape. a) Discriminant function analysis reveals strong multivariate phenotypic differentiation between apomictic (red points) and sexual (black) individuals. Vectors from the origin depict the loading of each of the 10 phenotypes. Of these, only the arrowed and labeled vectors are significantly associated with mating system divergence. b-e) Phenotypic correlations latitude and elevation for sexual (b-c) and apomictic (d-e) populations. The geographic variables latitude and elevation are plotted on the x and z axes. Population-level least square means for flowering time (c,e) and growth rate (b,d) are plotted on the respective y-axes. Populations are ranked by breeding values for each trait and color coded: top 20%- Red, 80-61%-Orange, 60-41%-Yellow, 40-21%-Green, bottom 20%-Blue. A linear model (phenotype=latitude + elevation) was fit to the data and plotted as the dashed line plane.

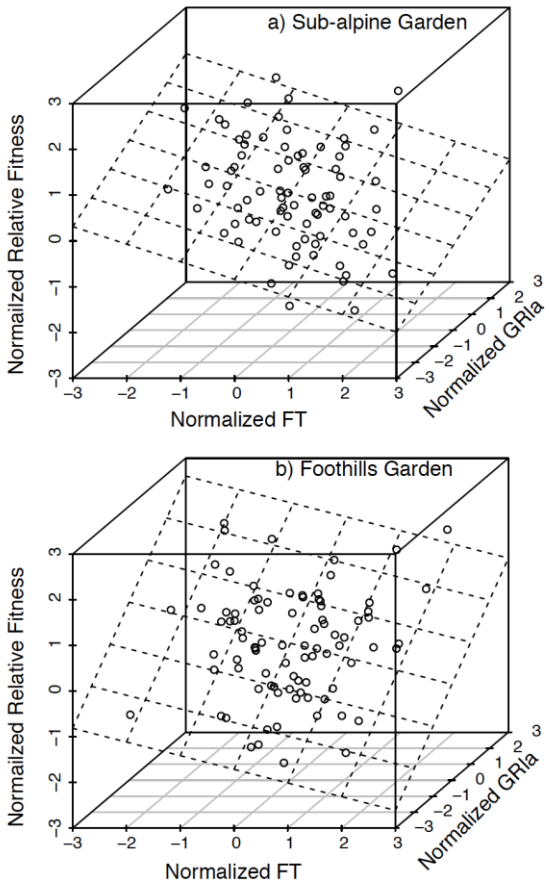

Figure 4.6 Linear selection surfaces for FT and *GRla*. The quantitile-normalized trait values of *GRla*, relative fitness and *FT* are plotted on the x and z axes respectively. The linear, least square mean regression plane is plotted as an additive function of *FT* and *GRla*. This analysis was conducted independently for the high elevation (a) and low elevation (b) gardens.

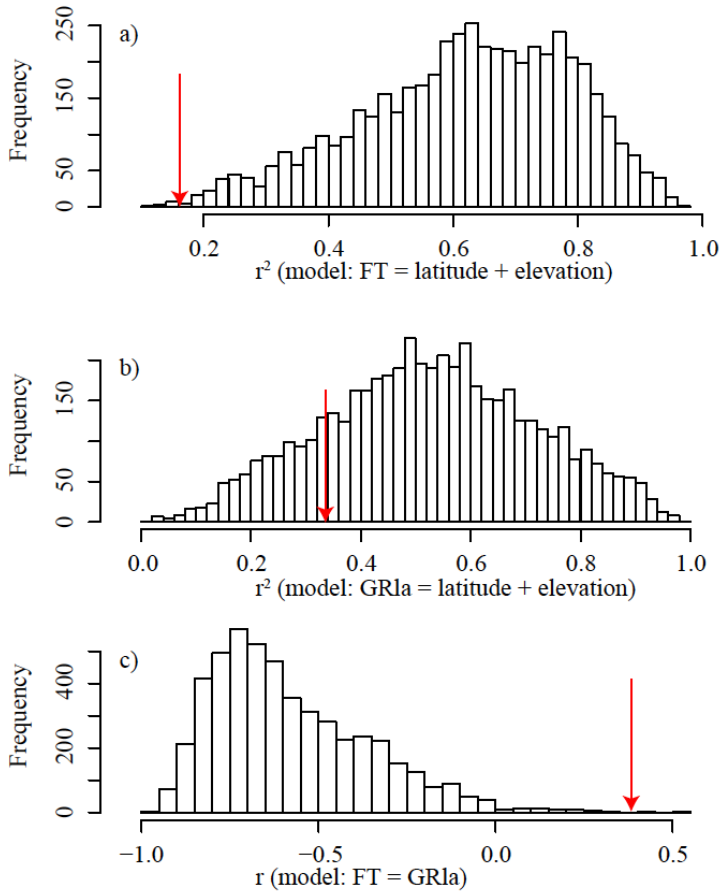

Figure 4.7 Distribution of correlation statistics following sub-sampling of sets of nine sexual populations. The correlation coefficient (r) or r^2 values are reported from 5000 runs of the models FT= Latitude + Elevation (a), GRla= Latitude + Elevation (b) *FT=GRla* (c). b) Proportion of variance explained by the regression (r^2) is reported for the multiple regression. For each panel, a red arrow represents the statistic from the apomictic analysis.

References:

Agren, J. and D. W. Schemske. 2012. Reciprocal transplants demonstrate strong adaptive differentiation of the model organism *Arabidopsis thaliana* in its native range. New Phytologist 194:1112-1122.

Alexander, P. J., M. D. Windham, J. B. Beck, I. A. Al-Shehbaz, L. Allphin, and C. D. Bailey. 2013. Molecular phylogenetics and taxonomy of the genus *Boechera* and related genera (Brassicaceae: Boechereae). Systematic Botany 38:192-209.

Aliyu, O. M., M. E. Schranz, and T. F. Sharbel. 2010. Quantitative variation for apomictic reproduction in the genus *Boechera* (brassicaceae). Am J Bot 97:1719-1731.

Angert, A. L., T. E. Huxman, P. Chesson, and D. L. Venable. 2009. Functional tradeoffs determine species coexistence via the storage effect. Proceedings of the National Academy of Sciences of the United States of America 106:11641-11645.

Arnold, S. J. 1992. Constraints on phenotypic evolution. American Naturalist:S85-S107.

Banta, J. A., I. M. Ehrenreich, S. Gerard, L. Chou, A. Wilczek, J. Schmitt, P. X. Kover, and M. D. Purugganan. 2012. Climate envelope modelling reveals intraspecific relationships among flowering phenology, niche breadth and potential range size in *Arabidopsis thaliana*. Ecol Lett 15:769-777.

Barton, N. H. 2010. Mutation and the evolution of recombination. Philosophical Transactions of the Royal Society of London- B Biological Sciences 365:1281-1294.

Barton, N. H. and B. Charlesworth. 1998. Why sex and recombination? Science 281:1986-1990.

Beck, J. B., P. J. Alexander, L. Allphin, I. A. Al-Shehbaz, C. Rushworth, C. D. Bailey, and M. D. Windham. 2012. Does hybridization drive the transition to asexuality in diploid *Boechera*? Evolution 66:985-995.

Berry, J. and O. Bjorkman. 1980. Photosynthetic response and adaptation to temperature in higher-plants. Annual Review of Plant Physiology and Plant Molecular Biology 31:491-543.

Bradbury, P. J., Z. Zhang, D. E. Kroon, T. M. Casstevens, Y. Ramdoss, and E. S. Buckler. 2007. Tassel: Software for association mapping of complex traits in diverse samples. Bioinformatics 23:2633-2635.

Bruvo, R., N. K. Michiels, T. G. D'SOUZA, and H. Schulenburg. 2004. A simple method for the calculation of microsatellite genotype distances irrespective of ploidy level. Molecular Ecology 13:2101-2106.

Burt, A. 2000. Perspective: Sex, recombination, and the efficacy of selection—was Weismann right? Evolution 54:337-351.

Butlin, R. 2002. The costs and benefits of sex: New insights from old asexual lineages. Nature Reviews Genetics 3:311-317.

Carman, J. G. 1997. Asynchronous expression of duplicate genes in angiosperms may cause apomixis, bispory, tetraspory, and polyembryony. Biological Journal of the Linnean Society 61:51-94.

Charlesworth, D. and S. I. Wright. 2001. Breeding systems and genome evolution. Current Opinion in Genetics & Development 11:685-690.

Clark, L. V. and M. Jasieniuk. 2011. Polysat: An R package for polyploid microsatellite analysis. Molecular Ecology Resources 11:562-566.

Clausen, J. C., D. D. Keck, and W. M. Hiesey. 1940. Effect of varied environments on western north american plants. Carnegie Institution.

Conway, D. J., C. Roper, A. M. Oduola, D. E. Arnot, P. G. Kremsner, M. P. Grobusch, C. F. Curtis, and B. M. Greenwood. 1999. High recombination rate in natural populations of *Plasmodium falciparum*. Proceedings of the National Academy of Sciences of the United States of America 96:4506-4511.

Dobeš, C., T. F. Sharbel, and M. Koch. 2007. Towards understanding the dynamics of hybridization and apomixis in the evolution of the genus *Boechera* (Brassicaceae). Systematics and Biodiversity 5:321-331.

Elshire, R. J., J. C. Glaubitz, Q. Sun, J. A. Poland, K. Kawamoto, E. S. Buckler, and S. E. Mitchell. 2011. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PloS One 6:e19379.

Engelstadter, J. 2008. Constraints on the evolution of asexual reproduction. Bioessays 30:1138-1150.

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

Fournier-Level, A., A. Korte, M. D. Cooper, M. Nordborg, J. Schmitt, and A. M. Wilczek. 2011. A map of local adaptation in *Arabidopsis thaliana*. Science 334:86-89.

Henry, L., T. Schwander, and B. J. Crespi. 2012. Deleterious mutation accumulation in asexual timema stick insects. Molecular biology and evolution 29:401-408.

Hill, W. and A. Robertson. 1966. The effect of linkage on limits to artificial selection. Genetical Research 8:269-294.

Horandl, E. 2008. Evolutionary implications of self-compatibility and reproductive fitness in the apomictic *Ranunculus auricomus* polyploid complex (ranunculaceae). International Journal of Plant Sciences169:1219-1228.

Hu, T. T., P. Pattyn, E. G. Bakker, J. Cao, J.-F. Cheng, R. M. Clark, N. Fahlgren, J. A. Fawcett, J. Grimwood, and H. Gundlach. 2011. The *Arabidopsis lyrata* genome sequence and the basis of rapid genome size change. Nature Genetics 43:476-481.

Jombart, T. 2008. Adegenet: A r package for the multivariate analysis of genetic markers. Bioinformatics 24:1403-1405.

Kawecki, T. J. and D. Ebert. 2004. Conceptual issues in local adaptation. Ecology Letters 7:1225-1241.

Kearney, M. 2005. Hybridization, glaciation and geographical parthenogenesis. Trends in Ecology and Evolution 20:495-502.

Lande, R. and S. J. Arnold. 1983. The measurement of selection on correlated characters. Evolution:1210-1226.

Lasky, J. R., D. L. Des Marais, J. McKAY, J. H. Richards, T. E. Juenger, and T. H. Keitt. 2012. Characterizing genomic variation of *Arabidopsis thaliana*: The roles of geography and climate. Molecular Ecology 21:5512-5529.

Lee, C. R. and T. Mitchell-Olds. 2013. Complex trait divergence contributes to environmental niche differentiation in ecological speciation of *Boechera stricta*. Molecular Ecology 22:2204-2217.

Lively, C. M. 2010. A review of red queen models for the persistence of obligate sexual reproduction. Journal of Heredity 101:S13-S20.

Lovell J, Aliyu O, Mau M, *et al.* (2013) On the origin and evolution of apomixis in *Boechera*. Plant Reproduction **26**, 309-315.

Lovell, J. T., T. E. Juenger, S. D. Michaels, J. R. Lasky, A. Platt, J. H. Richards, X. Yu, H. M. Easlon, S. Sen, and J. K. McKay. 2013b. Pleiotropy of *FRIGIDA* enhances the potential for multivariate adaptation. Proceedings of the Royal Society of London. Series B: Biological Sciences 280:20131043.

Lynch, M. and J. L. Blanchard. 1998. Deleterious mutation accumulation in organelle genomes. Genetica 102:29-39.

Manel, S., M. K. Schwartz, G. Luikart, and P. Taberlet. 2003. Landscape genetics: Combining landscape ecology and population genetics. Trends in Ecology & Evolution 18:189-197.

Matzk, F., A. Meister, and I. Schubert. 2000. An efficient screen for reproductive pathways using mature seeds of monocots and dicots. The Plant Journal 21:97-108.

Mogie, M. 1992. The evolution of asexual reproduction in plants. London: Chapman and Hall 276p.

Mráz, P., J. Chrtek, and B. Šingliarová. 2009. Geographical parthenogenesis, genome size variation and pollen production in the arctic-alpine species *Hieracium alpinum*. Botanica Helvetica 119:41-51.

Nautiyal, P., N. R. Rachaputi, and Y. Joshi. 2002. Moisture-deficit-induced changes in leaf-water content, leaf carbon exchange rate and biomass production in groundnut cultivars differing in specific leaf area. Field Crops Rresearch 74:67-79.

Nordborg, M. 2000. Linkage disequilibrium, gene trees and selfing: An ancestral recombination graph with partial self-fertilization. Genetics 154:923-929.

Otto, S. P. and T. Lenormand. 2002. Resolving the paradox of sex and recombination. Nature Reviews Genetics 3:252-261.

Paradis, E., J. Claude, and K. Strimmer. 2004. Ape: Analyses of phylogenetics and evolution in R language. Bioinformatics 20:289-290.

Paun, O., T. F. Stuessy, and E. Horandl. 2006. The role of hybridization, polyploidization and glaciation in the origin and evolution of the apomictic *Ranunculus cassubicus* complex. New Phytologist 171:223-236.

Qiu, S., K. Zeng, T. Slotte, S. Wright, and D. Charlesworth. 2011. Reduced efficacy of natural selection on codon usage bias in selfing *Arabidopsis* and *Capsella* species. Genome Biology and Evolution 3:868.

R Core Team. 2013. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Robertson, A., T. C. Rich, A. M. Allen, L. Houston, C. Roberts, J. R. Bridle, S. A. Harris, and S. J. Hiscock. 2010. Hybridization and polyploidy as drivers of continuing evolution and speciation in *Sorbus*. Molecular Ecology 19:1675-1690.

Slatkin, M. 1987. Gene flow and the geographic structure of natural. Science 3576198:236.

Slatkin, M. 1993. Isolation by distance in equilibrium and non-equilibrium populations. Evolution:264-279.

Turesson, G. 1922. The species and the variety as ecological units. Hereditas 3:100-113.

Turesson, G. 1930. The selective effect of climate upon the plant species. Hereditas 14:99-152.

van Kleunen, M. and M. Fischer. 2005. Constraints on the evolution of adaptive phenotypic plasticity in plants. New Phytologist 166:49-60.

Wilczek, A. M., J. L. Roe, M. C. Knapp, M. D. Cooper, C. Lopez-Gallego, L. J. Martin, C. D. Muir, S. Sim, A. Walker, J. Anderson, J. F. Egan, B. T. Moyers, R. Petipas, A. Giakountis, E. Charbit, G. Coupland, S. M. Welch, and J. Schmitt. 2009. Effects of genetic perturbation on seasonal life history plasticity. Science 323:930-934.

Windham, M. D. and I. A. Al-Shehbaz. 2006. New and noteworthy species of *Boechera* (Brassicaceae) i: Sexual diploids. Harvard Papers in Botany 11:61-88.

Chapter 5

CONCLUSION

This dissertation included three interconnected projects that assessed the constraints to adaptive evolution by comparing biological elements that were *a priori* predicted to constrain adaptation. The objectives of my dissertation were: (1) to document genes underlying drought adaptation in the model plant, *Arabidopsis thaliana*, to infer the adaptive effects of pleiotropy, (2) to determine the mechanisms constraining adaptation in rare species relative to widespread congeners, and (3) to assess the degree of adaptive differentiation in *Boechera spatifolia* and compare patterns of adaptation between sexually and asexually reproducing lineages. By achieving these objectives, I gained a novel perspective on those factors that constrain or promote adaptive evolution.

In chapter 2, I analyzed the multivariate phenotypic effects of a single locus. Pleiotropic loci may constrain adaptation by affecting variation of many traits and fixing the direction of trait correlations along vectors that are orthogonal to that favored by selection (non-adaptive, antagonistic pleiotropy) (Otto 2004; Rose 1982). However, the right kind of pleiotropy may improve the conditions for adaptive evolution (Eizaguirre *et al.* 2009; Ostman *et al.* 2012). *FRI* exhibits adaptive pleiotropy, affecting multiple traits and directing trait correlations along vectors that are predicted to be favored by selection in nature and experimental crosses. Because *FRI* confers phenotypic values along a selectively advantageous correlation, mutations at *FRI* may enhance successful adaptation to novel environmental conditions. Although they have been hypothesized to play a major role in speciation and local adaptation, few cases of causal loci that

pleiotropically affect phenological reproductive isolation and adaptive traits have been studied (Gavrilets 2003; Sobel *et al.* 2010). This study provided the first evidence of such a "magic" gene in *A. thaliana*. Furthermore, I presented a mechanistic understanding of how *FRI* altered physiology, phenology and conferred local adaptation. Phenology, growth rate and water use physiology have been mapped to similar genomic loci or correlated in natural populations. Here I demonstrated that *FRI* caused these adaptive correlations to be heritable.

In chapter 3 I assessed the hypothesis that rare species should have reduced evolutionary potential. The presence of closely related rare and widespread species in many taxa provides an ideal experimental system with which to understand the causes and consequences of range size evolution. By comparing the biological properties of species with divergent range sizes, I assessed the processes and factors that limit the geographic range and cause rarity. Species distributions may be limited by a lack of variation for ecologically important traits or low genetic diversity. Through quantitative (trait-based) and molecular genetic analyses, I found that, while molecular diversity was not different between species types, quantitative genetics revealed several significant differences between rare species and their widespread relatives. First, widespread species had much greater population structure than rare species, indicating that high levels of gene flow characterize rare species. It is possible that genetic exchange among locally adapted populations reduced the potential for range expansion. Second, rare species maintained significantly lower levels of phenotypic plasticity than widespread species. Third, rare species had significantly depressed levels of heritability, which permits strong responses to selection. Decreased heritability may directly inhibit adaptation of rare species and confine their geographic distribution. Combined, rare species exhibit decreased tolerances and adaptive potential to changing environments, placing these species at increased risk of extinction.

My 4th chapter addressed the hypothesis that asexual lineages have decreased potential for adaptive evolution, relative to sexual conspecifics. This hypothesis was derived from a series of theoretical and laboratory inquiries (Barton 2010; Burt 2000; Engelstadter 2008; Hill & Robertson 1966; Neiman *et al.* 2010). Combined, these studies demonstrated that reduced recombination in asexuals fixed linkage among loci, caused interference and decreased the efficacy of selection. I examined how environmental variation impacted genomic structure, selection pressures and local adaptation in sexual and asexual (apomictic) *Boechera spatifolia*. I found that, despite occupying sympatric sites, apomictic lineages are both phenotypically and phylogenetically distinct from sexuals. Additionally, while sexual populations formed strong clines (both genomic and physiological) along latitude and elevation gradients, apomicts did not show a similar signature of local adaptation. This result was consistent with one hypothesis for the evolution of sexual reproduction despite the "2-fold" cost of sex: asexual lineages may have poor responses to selection.

In conclusion, the potential to experience adaptive evolution differed greatly among species, populations and genotypes. Genetic correlations, physiological diversity and recombination each affected the total genetic diversity available to selection. These factors may directly affect a population's potential to adapt to environmental stresses (e.g. *FRI* and drought) or heterogeneity (e.g. apomixis and latitude). Additionally, reduced efficacy of selection generally constrains adaptation, which may have inhibited range expansion and effected patterns of biodiversity and geography. As genetic diversity is the driver for adaptation via evolutionary responses to natural selection, those factors constraining genetic diversity may have profound impacts on the evolutionary trajectory of species. By directly studying the factors that constrain

adaptive genetic diversity and responses to selection, it is possible to more effectively manage threatened populations and understand the evolutionary process in nature.

References

Barton NH (2010) Mutation and the evolution of recombination. Philosophical Transactions of the Royal Society of London- B Biological Sciences 365, 1281-1294.

Burt A (2000) Perspective: sex, recombination, and the efficacy of selection—was Weismann right? Evolution 54, 337-351.

Eizaguirre C, Lenz TL, Traulsen A, Milinski M (2009) Speciation accelerated and stabilized by pleiotropic major histocompatibility complex immunogenes. Ecology Letters 12, 5-12.

Engelstadter J (2008) Constraints on the evolution of asexual reproduction. Bioessays 30, 1138-1150.

Gavrilets S (2003) Perspective: models of speciation: what have we learned in 40 years? Evolution 57, 2197-2215.

Hill W, Robertson A (1966) The effect of linkage on limits to artificial selection. Genetical research 8, 269-294.

Neiman M, Hehman G, Miller JT, Logsdon JM, Jr., Taylor DR (2010) Accelerated mutation accumulation in asexual lineages of a freshwater snail. Molecular Biology and Evolution 27, 954-963.

Ostman B, Hintze A, Adami C (2012) Impact of epistasis and pleiotropy on evolutionary adaptation. Proceedings of the Royal Society of London. Series B: Biological Sciences 279, 247-256.

Otto SP (2004) Two steps forward, one step back: the pleiotropic effects of favoured alleles. Proceedings of the Royal Society of London. Series B: Biological Sciences 271, 705-714.

Rose MR (1982) Antagonistic pleiotropy, dominance, and genetic variation. Heredity 48, 63-78.

Sobel JM, Chen GF, Watt LR, Schemske DW (2010) The biology of speciation. Evolution 64, 295-315.