GENETIC AND ECOLOGICAL FACTORS CONTRIBUTING TO RANGE LIMITS IN NARROW ENDEMIC PLANT SPECIES

by

LOUISA GRANT CARTER STATON

(Under the Direction of Shu-Mei Chang)

ABSTRACT

Species are finite in their abundances and distributions, and the processes that form distributional patterns are complex. In this dissertation, I investigated the factors that contribute to geographic range limits in a set of narrow endemic species. Specifically, I first constructed species distribution models to investigate the location of abiotic niches and their predicted stability under climate change for nine threatened/endangered species. This work demonstrated that suitable habitat is not predicted to remain stable under the predictions of climate change, and these species are at great risk of extinction under current climate projections. Second, I performed reciprocal transplant experiments to investigate local adaptation and niche constraints in two sister taxa with varying distributions, *Polygonella americana* (widespread) and *P. fimbriata* (narrow). Populations of the narrow species displayed no evidence of niche constraints, and very little evidence of local adaptation. In contrast, populations of *P. americana* appeared locally adapted to their home environments. Finally, I used microsatellite markers designed specifically for these same two species (*P. americana* and *P. fimbriata*) to investigate the role of genetic constraints in shaping their ranges. I

found that populations of the narrow species are genetically depauperate, while the widespread species had populations with relatively higher levels of genetic diversity. Additionally, the structure of diversity in the widespread species demonstrated a strong geographic signal. These results indicate that an edge group of populations is diverging in this species as a result of low gene flow between it and the rest of the populations we sampled. Overall, my research indicates that the factors contributing to range limits in *P. fimbriata* adhere mainly to an evolutionary genetic constraints model of range limits, and that the added effects of dispersal limitation are causing this species to remain geographically restricted. My work has implications for the field of range limits, and I put forward specific improvements in methodology for investigating questions related to geographic distributions.

INDEX WORDS: Polygonaceae, Polygonella, range limits, species distribution models, SDMs, microsatellites, population genetics, reciprocal transplants, local adaptation, niche constraints

GENETIC AND ECOLOGICAL FACTORS CONTRIBUTING TO RANGE LIMITS IN NARROW ENDEMIC PLANT SPECIES

by

LOUISA GRANT CARTER STATON

Bachelor of Science, Furman University, 2004

Master of Science, College of Charleston, 2008

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2014

GENETIC AND ECOLOGICAL FACTORS CONTRIBUTING TO RANGE LIMITS IN NARROW ENDEMIC PLANT SPECIES

by

LOUISA GRANT CARTER STATON

Major Professor:	Shu-Mei Chang
Committee:	Kelly Dyer Jim Hamrick David Moeller John Wares

Electronic Version Approved:

Julie Coffield
Interim Dean of the Graduate School
The University of Georgia
August 2014

DEDICATION

To my husband, the love and joy of my life whose support through this process and ability to make me laugh is what kept me going, and to my family, whose encouragement throughout my entire academic career is what made the difference in my making it this far.

ACKNOWLEDGEMENTS

I want to thank my advisor, Shu-Mei Chang, for allowing me the room to pursue my own research questions, and my lab mates for feedback and discussion on manuscripts along the way. I also want to thank Dr. Jeffrey Hepinstall-Cymerman for his time and guidance in the development of my first research chapter. Also, I want to thank the many people whose help at various field sites and natural areas made my field experiments possible: Nathan Klaus, Linda Lee, Dr. Rebecca Sharitz, Sean Poppy, Patti Donnellan, Dr. Ann Johnson, Dr. Tom Wendt, Dr. Bill Carr, Dr. Marco Antonio Gúzman, Dr. Eric Menges, and Dr. Carl Weekley. Finally, I want to recognize the tireless efforts of countless graduate students who selflessly volunteered their time and efforts in order to help me accomplish my fieldwork.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER ONE

INTRODUCTION AND LITERATURE REVIEW

Every species, be it plant or animal, is a collection of populations distributed in some pattern throughout geographical space. Some patterns of distribution are more frequent than others. For instance, Darwin noted that within many genera, most taxa are geographically limited in range, while one or two species are widespread (Darwin, 1959). Theoretical pursuits to understand such patterns and the processes contributing to them have historically been of interest to biologists (MacArthur & Wilson, 1967; Brown, 1984; Hubbell, 2001). However, the causative factors driving distributional patterns are still poorly understood (Gaston, 2009; Hargreaves *et al.*, 2014). This is because a species' interaction with its surrounding environment is complex, involving both abiotic and biotic factors (Brown, 1995). This relationship becomes even more convoluted when a species is widely distributed across a heterogeneous landscape. Given this complexity of species' ranges, multifacted approaches to assessing variation in distributions will be the most useful. These should incorporate an assessment of dispersal limitation (Nathan & Muller-Landau, 2000; Münzbergová & Herben, 2005), as well as investigations into what environmental (both biotic and abiotic), demographic, and evolutionary genetic factors contribute to shaping the range (Geber, 2008; Gaston, 2009; Geber, 2011).

The simplest and most obvious explanation of variation in geographic ranges is environmental variation; however, teasing apart which environmental factors explain a distribution is not simple—the environment itself is complex, especially as it applies to species' presences (Salisbury, 1926; Billings, 1952). Rather, it is straightforward to envision that a species' geographical range is distributed a certain way because it tracks

the spatial distribution of its ecological niche requirements. The ecological niche of a species comprises the set of biotic and abiotic conditions necessary for the persistence of a species (Hutchinson, 1957). This includes biotic interactions, which determine how much of the fundamental niche (the full set of requirements) is physically occupied (the realized niche; Hutchinson, 1957; Travis *et al.*, 2005; Araújo & Guisan, 2006). Specifically, the realized niche is governed by the amount of interspecific competition, which is negatively correlated with realized niche space, and the presence of beneficial interactions (pollinators, arbuscular mycorrhizae, etc.), which are positively correlated with niche space (Araújo & Guisan, 2006). These interactions are variable in time and space and will act in concert with the abiotic niche requirements to shape the geographic distribution of a species (Pulliam, 2000; Case *et al.*, 2005; Soberón & Peterson, 2005; Chamberlain *et al.*, 2014). The geographical distribution of these niche requirements is referred to as the suitable habitat for a species (MacArthur, 1972).

While experimentally measuring biotic interactions and their impact on suitable habitat can be difficult (though, see Connell, 1961), delineating the abiotic components of a species' suitable habitat has become more feasible with recent advances in modeling and technology. First, the application of generalized linear models (GLMs) to regress presence/absence location data with associated environmental data allows statistical inferences regarding which environmental factors are the best predictors of species' presence (Elith & Leathwick, 2009). Second, the advancement of powerful geographic information systems (GIS) platforms has enabled improved visualization, manipulation, and communication of these models (Franklin, 2009). Such models are referred to as ecological niche models or the recently more preferable term species distribution models

(SDMs), which avoids the confusion of niche definitions (see McInerny & Etienne, 2012a-c, 2013; Warren, 2012, 2013). The recent use of SDMs in the literature has risen dramatically. A search on Web of Science (Web of Knowledge) for articles on the topic of either SDMs or ecological niche models yielded more than 34,500 results, nearly one third of which have been published since the beginning of this decade.

Primarily, the use of SDMs has been geared towards determining which climatic variables correlate most strongly with the presence of a species (defining the components of the abiotic niche) and mapping the suitable habitat of a species (locating the boundaries of this niche) (Franklin, 2009). However, their utility has also been demonstrated for locating previously unknown populations of rare species (e.g., de Siqueira *et al.*, 2009) and as a predictive tool for conservation in the context of future climate change (for a recent example, see McCallum *et al.*, 2014). This latter use of SDMs has been criticized recently due to the general tendency of SDMs to overestimate habitat loss under climate change and their inability to incorporate biologically relevant data for conservation decisions (e.g., dispersal, genetic diversity, effective population size, interactions; Wiens *et al.*, 2009; Hellmann *et al*., 2012; Higgins *et al.*, 2012; Schwartz, 2012). Even so, SDMs can be used in conservation to identify putative regions for assisted migration—one of the proposed conservation strategies for dispersal-limited species that are threated by climate change (Hunter, 2007; McLachlan *et al*., 2007; Vitt *et al.*, 2010). Despite their imperfections, SDMs are useful tools for evaluating where suitable habitat exists and whether or not it is fully occupied by the species under consideration. In this regard, they can provide an indirect test of whether or not species geographic ranges are in equilibrium with (completely fill) their ecological niches. In this

way, they are an inexpensive approach for inferring such scenarios, and can help to guide more costly experimental efforts aimed at understanding why this disequilibrium exists.

If species ranges are in equilibrium with their ecological niches, then presumably niche constraints are what maintain the range boundary (Sexton *et al.*, 2009). Alternatively, if species geographic ranges are in disequilibrium with their ecological niches, this can be attributable to several factors (Sexton *et al*., 2009; Hargreaves *et al.*, 2014). For instance, the locations of the niche and species' range are variable in time and space (Brown *et al.*, 1996; Davis & Shaw, 2001). If a shift in suitable habitat occurs quickly enough, dispersal limitation can cause range/niche disequilibrium, as has been demonstrated with the role of dispersal lags in the post-glacial expansion of European trees (post-glacial migrational lag hypothesis; Svenning & Skov, 2007). This can be further complicated by fluctuations in environmental quality (Bahn *et al.*, 2006; Dytham, 2009; Hargreaves & Eckert, 2014). In contrast, if dispersal limitations are suspected not to contribute or only play a minor role in range/niche disequilibrium, then the focus falls upon determining what dynamics are at play at the edge of the species' range. The majority of range limits models are concerned with addressing these dynamics (Kirkpatrick & Barton, 1997; Holt & Keitt, 2000; Case *et al.*, 2005; Moeller *et al*., 2011), and typically invoke the failure of populations to adapt at the margin (Bridle & Vines, 2007; Kawecki, 2008; Bridle *et al.*, 2010).

The most powerful way to test whether populations are locally adapted at the margin is through the use of reciprocal transplant experiments (Kawecki & Ebert, 2004; Leimu & Fischer, 2008; Hargreaves *et al.*, 2014). Reciprocal transplant experiments are an extension of the common garden design, which was first used by Turesson (1922) to

identify different ecological races in a species ("ecotypes"). The utility of common garden experiments is that they eliminate environmental variation (V_E), such that observed differences in phenotype (V_P) are due solely to genetic variation (V_G; $V_P = V_G + V_E$). Classic experiments by Clausen *et al.* (1940, 1948) extended this experimental design to include multiple common gardens, which spanned altitudinal clines within the range. By transplanting populations of each species from each of these locations into every garden, including back into their "home" gardens, a direct test of local adaptation can be performed. The observation that "home" plants have higher fitness than "away" plants within a garden, and that "away" plants also have significantly decreased fitness compared to their own home sites indicates that populations are locally adapted (Kawecki & Ebert, 2004). In the context of range limits, studies that reciprocally transplant populations into gardens located at the range interior, edge, and beyond can effectively test for limits to adaptation (Geber & Eckhart, 2005; Angert & Schemske, 2005), adaptive trade-offs between environments (Angert *et al.*, 2008), and evolutionary constraints in ecologically important traits tied to range expansion (Griffith & Watson, 2006).

Additionally, reciprocal transplant experiments are powerful tools for investigating the presence of niche constraints (Gaston, 2003). First, observations of low to no fitness in edge transplant sites would infer that range limits exceed niche limits, and edge populations are possibly sustained by source-sink dynamics (Pulliam, 2000). Second, an observation that edge populations outperform interior populations at both edge and beyond transplant sites (situated outside of the current geographic range) would indicate an environmental gradient from the center to the edge of the range (Brown,

1984). It would suggest that range limits fall short of niche limits (Hargreaves *et al.*, 2014). Third, observations that edge population fitness does not differ at beyond versus edge transplant sites would also suggest that range limits fall short of niche limits, and perhaps are maintained by metapopulation dynamics (Holt & Keitt, 2000). Finally, in comparison to performance at the edge, decreased or no fitness of any populations transplanted beyond the margin would suggest that niche constraints play an important role in the maintenance of the range limit (Hargreaves *et al.*, 2014). In these scenarios, populations founded beyond the range margin should fail because their finite population growth will be insufficient ($\lambda<1$; Holt, 2003).

When niche constraints are not responsible for species' range limits, then this suggests that populations are failing to adapt at the edge of the range. There are several models that focus on why adaptation is not achieved in marginal populations, and these are grouped into two main categories: evolutionary genetic models and demographic models (Geber, 2008; Moeller *et al*., 2011). All of these models assume that the range edge is in equilibrium and not changing shape or size. An obvious alternative to this is that the range is actually expanding, but our temporal "snapshot" is too early in its evolution to capture this. Also, evolutionary genetics models assume the case of Brown's abundant center model (1984)—that a species' range exists along an environmental gradient where conditions are best at the center and deteriorate towards the edge. Under this model, populations will be larger and more densely concentrated in the center of the range, while populations on the edge will be smaller and more sparsely distributed.

Under the two assumptions of equilibrium and the abundant center model, evolutionary genetic models predict that peripheral populations are constrained from

adaptation due to low genetic diversity (Hoffman & Blows, 1994; Kawecki & Ebert, 2004) or maladaptive gene flow from the center of the range to the periphery (Kirkpatrick & Barton, 1997). Low genetic diversity in peripheral populations is expected to be a consequence of drift and inbreeding due to small effective population sizes, and isolation and subsequent differentiation among populations due to geographic isolation (Hoffman & Blows, 1994). A review by Eckert *et al.* (2008) demonstrated that experimental studies on peripheral populations do in fact tend to confirm these expectations (low diversity, increased differentiation), but that they fail to address the historical reasons for why this might be the case (e.g., postglacial expansion, human-mediated fragmentation). Additionally, genetic constraints (e.g., antagonistic pleiotropy in flowering phenology) have recently been shown to play a role in preventing local adaptation (Anderson *et al.*, 2013).

The other side of evolutionary genetic constraints models hinge on gene flow. In some cases, peripheral populations with small effective population sizes may be "rescued" from the consequences of low genetic diversity and can adapt if gene flow from the center to the edge of the range is moderate (Barton, 2001; Alleaume-Benharira *et al.*, 2006). Alternatively, density-dependent models predict the opposite (Bridle & Vines, 2007). These models predict that populations with high enough density will track the trait optimum for their environment, thereby expanding along the environmental gradient and maintaining high diversity. However, if the trait optimum changes too quickly and there is not sufficient genetic variation in the population to achieve this optimum, populations will decrease in fitness and in density. When this happens, central to marginal gene flow will introduce poorly adapted immigrants, and population density

will decrease as a consequence of selection (Kirkpatrick & Barton, 1997; Lenormand, 2002). Finally, interspecific gene flow at the range edge can result in hybridization, which can act to sharpen boundaries by preventing adaptation within marginal populations, regardless of environmental gradients (Goldberg & Lande, 2006). It does so by creating less-fit hybrid offspring as well as linkage disequilibrium, which can impede the fixation of beneficial alleles, all of which should in turn act to create sharp margins (Goldberg & Lande, 2006).

Alternative models of range limits focus on population demographics (e.g., population size and density, colonization/extinction) and their how fluctuations across homogeneous and heterogeneous landscapes act to prevent adaptation and form distinct boundaries (Holt & Keitt, 2000; Moeller *et al*., 2011). It is important to note that while range limit models are typically grouped into demographic versus genetic models, the two often act in concert to form range boundaries (Holt & Barfield, 2011). For instance, population sizes can fluctuate greatly through time and space, sometimes going extinct. When extinctions and recolonizations occur frequently, populations experience genetic bottlenecks, which can sometimes be "rescued" via gene flow from other local populations. Such demographic processes (frequent extinctions, recolonizations, and rescues by gene flow) are referred to as metapopulation dynamics, and have a distinct genetic signature within populations (Levins, 1969; Moeller *et al.*, 2011). The ecological gradient over which metapopulation dynamics occur is expected to have a strong influence on the shape of the range edge (Hanski & Gilpin, 1997). When suitable habitat is disjunct, high extinction and low colonization will prevent adaptation at the margin and form a sharp boundary (Holt & Keitt, 2000). Population extinctions are more likely when

environments are unstable, when populations have strong density-dependent growth rates (Allee effects; Keitt *et al*., 2001) and carrying capacity changes along an environmental gradient (Holt *et al.*, 2005), and when sex ratios are uneven such that finding a mate is difficult (Legendre *et al.*, 1999). These are all exacerbated when suitable habitat is patchy (Keitt *et al*., 2001), and when species interactions (competition, predation, parasitism, mutualisms) are at play (Case & Taper, 2001; Case *et al*., 2005). Thus, demographic effects can play a significant role in the ability of populations to succeed and adapt at the margin. However, they can be difficult to investigate via experimental observation due to the extended time frame over which they usually occur (Moeller *et al.*, 2011). Even though this is the case, certain demographic events, like frequent extinctions and recolonizations, are expected to affect the genetics of a population (e.g., bottlenecked diversity; Nei *et al.*, 1975). For this reason, assessments of population genetic characteristics (diversity, differentiation, effective population size) are the best way to infer both demographic and evolutionary genetic constraints.

As has been illustrated here, the factors that control geographic range limits are complex and not mutually exclusive. Given that this is the case, it is likely that in most systems some combination of dispersal, genetic, and demographic constraints contributes to the distinct margins we see occurring in species' ranges. An investigation into a single aspect of range limits, e.g. an assessment solely of local adaptation within the range, would fail to determine either why local adaptation fails at the range edge or why populations are locally adapted but not found beyond the range border. Attempts to decipher what factors are at play in shaping species' ranges should incorporate investigations into as many of the factors listed above as possible. By doing so,

evolutionary biologists can begin to understand more fully why species are distributed the way they are, and make better predictions about how species' ranges will change in the future, especially in the context of our rapidly changing climate.

Works Cited

Alleaume-Benharira, M., Pen, I.R. & Ronce, O. (2006) Geographical patterns of adaptation within a species' range: interactions between drift and gene flow. *Journal of Evolutionary Biology*, **19**, 203-15.

Anderson, J.T., Lee, C.-R., Rushworth, C.A., Colautti, R.I. & Mitchell-Olds, T. (2013) Genetic trade-offs and conditional neutrality contribute to local adaptation. *Molecular Ecology*, **22**, 699-708.

Angert, A.L., Bradshaw Jr., H.D. & Schemske, D.W. (2008) Using experimental evolution to investigate geographic range limits in monkeyflowers. *Evolution*, **62**, 2660-75.

Angert, A.L. & Schemske, D.W. (2005) The evolution of species' distributions: reciprocal transplants across the elevation ranges of *Mimulus cardinalis* and *M. lewisii*. *Evolution*, **59**, 1671-84.

Araújo, M.B. & Guisan, A. (2006) Five (or so) challenges for species distribution modeling. *Journal of Biogeography*, **33**, 1677-88.

Bahn, V., O'Connor, R.J. & Krohn, W.B. (2006) Effect of dispersal at range edges on the structure of species ranges. *Oikos*, **115**, 89-96.

Barton, N.H. (2001) Adaptation at the species' range. (Eds. J. Silvertown & J. Antonovics) *Integrating Ecology and Evolution in a Spatial Context*. Blackwell Science, Oxford, U.K.

Billings, W.D. (1952) The environmental complex in relation to plant growth and distribution. *The Quarterly Review of Biology*, **27**, 251-65.

Bridle, J.R., Polechová, J., Kawata, M. & Butlin, R.K. (2010) Why is adaptation prevented at ecological margins? New insights from individual-based simulations. *Ecology Letters*, **13**, 485-94.

Bridle, J.R. & Vines, T.H. (2007) Limits to evolution at range margins: when and why does adaptation fail? *Trends in Ecology and Evolution*, **22**, 140-7.

Brown, J.H. (1984) On the relationship between abundance and distribution of species. *The American Naturalist*, **124,** 255-79.

Brown, J.H. (1995) *Macroecology*. University of Chicago Press, Chicago, IL.

Brown, J.H., Stevens, G.C. & Kaufman, D.M. (1996) The geographic range: size, shape, boundaries, and internal structure. *Annual Review in Ecology and Systematics*, **27**, 597-623.

Case, T.J., Holt, R.D., McPeek, M.A. & Keitt, T.H. (2005) The community context of species' borders: ecological and evolutionary perspectives. *Oikos*, **108**, 28-46.

Case, T.J. & Taper, M.L. (2000) Interspecific competition, environmental gradients, gene flow, and the coevolution of species borders. *The American Naturalist*, **155**, 583 605.

Clausen, J.C., Keck, D.D. & Hiesey, W.M. (1940) Experimental studies on the nature of species. I. Effect of varied environments on western North American plants. *Carnegie Institution of Washington Publications*, **520**, 1-452.

Clausen, J.C., Keck, D.D. & Hiesey, W.M. (1948) Experimental studies on the nature of species. III. Environmental responses of climatic races of *Achillea*. *Carnegie Institution of Washington Publications*, **581**, 1-129.

Chamberlain, S.A., Bronstein, J.L. & Rudgers, J.A. (2014) How context dependent are species interactions? *Ecology Letters*, **17**, 881-90.

Connell, J.H. (1961) The influence of interspecific competition and other factors on the distribution of the barnacle *Chthamalus stellatus*. *Ecology*, **42**, 710-23.

Darwin, C. (1859) *On the origins of species by means of natural selection*. Jon Murray, London.Davis, M.B. & Shaw, R.G. (2001) Range shifts and adaptive responses to Quaternary climate change. *Science*, **292**, 673-9.

Davis, M.B. & Shaw, R.G. (2001) Range shifts and adaptive responses to Quaternary climate change. *Science*, **292**, 673-9.

de Siqueira, M.F., Durigan, G., de Marco Júnior, P. & Peterson, A.T. (2009) Something from nothing: Using landscape similarity and ecological niche modeling to find rare plant species. *Journal for Nature Conservation*. **17**, 25-32.

Dytham, C. (2009) Evolved dispersal strategies at range margins. *Proceedings of the Royal Society B: Biological Sciences*, **276**, 1407-13.

Eckert, C.G., Samis, K.E. & Lougheed, S.C. (2008) Genetic variation across species' geographical ranges: the central-marginal hypothesis and beyond. *Molecular Ecology*, **17**, 1170-88.

Elith, J. & Leathwick, J.R. (2009) Species distribution models: ecological explanation and prediction across space and time. *Annual Review in Ecology, Evolution, and Systematics,* **40**, 677-97.

Franklin, J. (2009) *Mapping species distributions: spatial inference and prediction*. Cambridge University Press, Cambridge, U.K.

Gaston, K.J. (2003) *The structure and dynamics of geographic ranges*. Oxford University Press, Oxford, U.K.

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

Geber, M.A. (2008) To the edge: studies of species' range limits. *New Phytologist*, **178**, 228-30.

Geber, M.A. (2011) Ecological and evolutionary limits to species geographic ranges. *The American Naturalist*, **178**, S1-S5.

Geber, M.A. & Eckhart, V.M. (2005) Experimental studies of adaptation in *Clarkia xantiana*. II. Fitness variation across a subspecies border. *Evolution*, **59**, 521-31.

Goldberg, E.E. & Lande, R. (2006) Ecological and reproductive character displacement on an environmental gradient. *Evolution*, **60**, 1344-57.

Griffith, T.M. & Watson, M.A. (2006) Is evolution necessary for range expansion? Manipulation reproductive timing of a weedy annual transplanted beyond its range. *The American Naturalist*, **167**, 153-64.

Hanski, I. & Gilpin, M. (1991) Metapopulation dynamics: brief history and conceptual domain. *Biological Journal of the Linnean Society*, **42**, 3-16.

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

Hargreaves, A.L, Samis, K.E. & Eckert, C.G. (2014) Are species' range limits simply niche limits writ large? A review of transplant experiments beyond the range. *The American Naturalist,* **183,** 157-73.

Hellmann, J.J., Prior, K.M. & Pelini, S.L. (2012) The influence of species interactions on geographic range change under climate change. *Annals of the New York Academy of Sciences*, **1249**, 18-28.

Higgins, S.I., O'Hara, R.B. & Römermann, C. (2012) A niche for biology in species distribution models. *Journal of Biogeography*, **39**, 2091-95.

Hoffman, A.A. & Blows, M.W. (1994) Species borders: ecological and evolutionary perspectives. *Trends in Ecology and Evolution*, **9**, 223-7.

Holt, R.D. (2003) On the evolutionary ecology of species' ranges. *Evolutionary Ecology Research*, **5**, 159-78.

Holt, R.D. & Barfield, M. (2011) Theoretical perspectives on the statics and dynamics of species' borders in patchy environments. *The American Naturalist*, **178**, S6-S25.

Holt, R.D. & Keitt, T.H. (2000) Alternative causes for range limits: a metapopulation perspective. *Ecology Letters*, **3**, 41-7.

Holt, R.D., Keitt, T.H., Lewis, M.A., Maurer, B.A. & Taper, M.L. (2005) Theoretical models of species' borders: single species approaches. *Oikos*, **108**, 18-27.

Hubbell, S.P. (2001) *The unified neutral theory of biodiversity and biogeography*. Princeton University Press, Princeton.

Hunter Jr., M.L. (2007) Climate change and moving species: furthering the debate on assisted colonization. *Conservation Biology*, **21**, 1356-58.

Hutchinson, G.E. (1957) Concluding remarks. *Cold Spring Harbor Symposia on Quantitative Biology*, **22**, 145-59.

Kawecki, T.J. (2008) Adaptation to marginal habitats. *Annual Reviews in Ecology, Evolution and Systematics*, **39**, 321-42.

Kawecki, T.J. & Ebert, D. (2004) Conceptual issues in adaptation. *Ecology Letters*, **7**, 1225-41.

Keitt, T.H., Lewis, M.A. & Holt, R.D. (2001) Allee effects, invasion pinning, and species' borders. *The American Naturalist*, **157**, 203-16.

Kirkpatrick, M. & Barton, N.H. (1997) Evolution of a species' range. *The American Naturalist*, **150**, 1-23.

Legendre, S., Clobert, J., Møller, A.P. & Sorci, G. (1999) Demographic stochasticity and social mating system in the process of extinction of small populations: the case of passerines introduced to New Zealand. *The American Naturalist*, **153**, 449-63.

Leimu, R. & Fischer, M. (2008) A meta-analysis of local adaptation in plants. *PLoS ONE,* **3**, e4010.

Lenormand, T. (2002) Gene flow and the limits to natural selection. *Trends in Ecology and Evolution*, **17**, 183-9.

Levins, R. (1969) Some demographic and genetic consequences of environmental heterogeneity for biological control. *Bulletin of the ESA*, **15**, 237-40.

MacArthur, R.H. (1972) *Geographical ecology*. Harper & Row, New York.

MacArthur, R.H. & Wilson, E.O. (1967) *The theory of island biogeography*. Princeton University Press, Princeton, NJ.

McCallum, K.P., Guerin, G.R., Breed, M.F. & Lowe, A.J. (2014) Combining population genetics, species distribution modeling, and field assessments to understand a species vulnerability to climate change. *Austral Ecology*, **29**, 17-28.

McInerny, G.J. & Etienne, R.S. (2012a) Ditch the niche—is the niche a useful concept in ecology or species distribution modeling? *Journal of Biogeography*, **39**, 2096-2102.

McInerny, G.J. & Etienne, R.S. (2012b) Stitch the niche—a practical philosophy and visual schematic for the niche concept. *Journal of Biogeography*, **39**, 2103-11.

McInerny, G.J. & Etienne, R.S. (2012c) Pitch the niche—taking responsibility for the concepts we use in ecology and species distribution modeling. *Journal of Biogeography*, **39**, 2112-18.

McInerny, G.J. & Etienne, R.S. (2013) 'Niche' or 'distribution' modeling? A response to Warren. *Trends in Ecology and Evolution*, **28**, 191-2.

McLachlan, J.S., Hellmann, J.J. & Schwartz, M.W. (2007) A framework for debate of assisted migration in an era of climate change. *Conservation Biology*, **21**, 297-302.

Moeller, D.A., Geber, M.A. & Tiffin, P. (2011) Population genetics and the evolution of geographic range limits in an annual plant. *The American Naturalist*, **178**, S44-S57.

Münzbergová, Z. & Herben, T. (2005) Seed, dispersal, microsite, habitat and recruitment limitation: identification of terms and concepts in studies of limitations. *Oecologia*, **145**, 1-8.

Nathan, R. & Muller-Landau, H.C. (2000) Spatial patterns of seed dispersal, their determinants and consequences for recruitment. *Trends in Ecology and Evolution,* **15**, 278-85.

Nei, M., Maruyama, T. & Chakraborty, R. (1975) The bottleneck effect and genetic variability in populations. *Evolution*, **29**, 1-10.

Pulliam, H.R. (2000) On the relationship between niche and distribution. *Ecology Letters*, **3**, 349-61.

Salisbury, E.J. (1926) The geographical distribution of plants in relation to climatic factors. *The Geographic Journal*, **67**, 312-35.

Schwartz, M.W. (2012) Using niche models with climate projections to inform conservation management decisions. *Biological Conservation*, **155**, 149-56.

Sexton, J.P., McIntyre, P.J., Angert, A.L. & Rice, K.J. (2009) Evolution and ecology of species range limits. *Annual Review of Ecology, Evolution and Systematics*, **40**, 415-36.

Soberón, J. & Peterson, A.T. (2005) Interpretations of models of fundamental ecological niches and species' distributional areas. *Biodiversity Information*, **2**, 1-10.

Svenning, J.-C. & Skov, F. (2007) Could the tree diversity pattern in Europe be generated by postglacial dispersal limitation? *Ecology Letters*, **10**, 453-60.

Travis, J.M.J., Brooker, R.W. & Dytham, C. (2005) The interplay of positive and negative species interactions across an environmental gradient: insights from an individual-based simulation model. *Biology Letters*, **1**, 5-8.

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

Vitt, P., Havens, K., Kramer, A.T., Sollenberger, D. & Yates, E. (2010) Assisted migration of plants: Changes in latitudes, changes in attitudes. *Biological Conservation*, **143**, 18-27.

Warren, D.L. (2012) In defense of 'niche modeling.' *Trends in Ecology and Evolution*, **27**, 497-500.

Warren, D.L. (2013) 'Niche modeling': that uncomfortable sensation means it's working. A reply to McInerny and Etienne. *Trends in Ecology and Evolution*, **28**, 193-4.

Wiens, J.A., Stralberg, D., Jongsomjit, D., Howell, C.A. & Snyder, M.A. (2009) Niches, models, and climate change: Assessing the assumptions and uncertainties. *Proceedings of the National Academy of Sciences*, **106**, 19729-36.

CHAPTER TWO

USING SPECIES DISTRIBUTION MODELS TO UNDERSTAND THE GEOGRAPHICAL RANGES OF THREATENED AND ENDANGERED PLANT SPECIES IN THE SOUTHEASTERN UNITED STATES

Louisa G.C. Staton[1], Jeffrey Hepinstall-Cymerman[2], and Shu-Mei Chang[1,3]

To be submitted to the Journal of Biogeography

ABSTRACT:

Aim The process of determining what factors limit geographic range is complicated due to the complex nature of understanding the specific niche of a given species, but this process may be simplified by utilizing modeling techniques, like species distribution models (SDMs). Additionally, SDMs may be used to predict future changes in distribution in the context of global climate change. We used SDMs and the abundant and thorough location data kept for threatened and endangered plant species to identify what environmental factors define the current distributions of nine plant species. Additionally, we used SDMs to model distributional changes and make comparisons between the distributions under current and future climate scenarios to identify which species should be considered top priority for conservation efforts.

Location Plant species in this study, listed either as threatened or endangered, occur naturally in the southeastern United States.

Methods We obtained presence data for nine plant species (*Croomia pauciflora,* (Stemonaceae), *Pachysandra procumbens* (Buxaceae), *Panax quinquefolius* (Araliaceae), *Polygonella basiramia* (Polygonaceae), *Polygonella macrophylla* (Polygonaceae), *Polygonella myriophylla* (Polygonaceae), *Silene polypetala* (Caryophyllaceae), *Trillium lancifolium* (Melanthiaceae), and *Waldsteinia lobata* (Rosaceae) from state agencies (e.g. Department of Natural Resources). Using these occurrence data, we generated SDMs under current, 2050 and 2080 climate predictions for the A1B1 ("middle-of-the-road") climate scenario using maximum entropy methods (Maxent). We compared model

performance in terms of area under the curve (AUC) scores from receiver operating characteristic curves. We calculated percent changes in suitable habitat from current to future climate scenarios using a strict and a liberal threshold, and used these percent changes to identify which species were at the greatest risk in the context of future climate change.

Results Models had strong predictive power, having AUC scores of 0.9 or above for all but one of the species. Temperature correlates and soil type were the most common explanatory factors underlying current geographic distributions of species. Six of the nine species were predicted to experience dramatic amounts of loss in suitable habitat in future climate scenarios, most notably with one species (*W. lobata*) predicted to lose 100% of its suitable habitat by 2080. Additionally, the suitable habitat of four of the nine species was predicted to geographically shift under future climate scenarios. Of these four species, three were predicted to also experience an increase in the amount of suitable habitat by 2080.

Main Conclusions Based on our SDM predictions, *C. pauciflora, P. basiramia,* and *W. lobata* are the most at-risk species in the context of future climate changes in this study. Additionally, while some species are predicted to experience a gain in the amount of suitable habitat under future climate scenarios, the geographic center of these suitable habitat areas is predicted to shift considerably outside of the current range of these species. Likely, dispersal in these species will not track quickly enough with these changes in suitable habitat locations to allow populations of these species to persist. We

recommend population genetic investigations for these species in order to identify genetically diverse populations of these species for use in conservation efforts. Additionally, we recommend the listing of *W. lobata* be elevated to "Critically Endangered" following IUCN guidelines.

Keywords Species distribution models, SDM, maximum entropy, Maxent, *Croomia pauciflora, Pachysandra procumbens*, *Panax quinquefolius*, *Polygonella basiramia, Polygonella macrophylla*, *Polygonella myriophylla*, *Silene polypetala*, *Trillium lancifolium*, *Waldsteinia lobata,* conservation, climate change, narrow endemics

Introduction

All species are finite in their ranges, with few that have truly cosmopolitan distributions. The range in which a species can be found geographically is often attributed to the combination of abiotic and biotic factors in the environment, usually referred to as that species' niche. However, defining the specific niche for a species remains a major challenge for ecologists trying to understand the principles underlying the geographic distributions of species. This is because many factors may contribute to the niche (Pulliam, 2000), and the relative importance of these factors is generally not immediately clear (also see the discussion on the use of "niche": Vandermeer, 1972; Soberón 2007). However, the ability to delineate the ecological niche of a species (*sensu* Grinnell, 1917) has recently gained speed as more computational tools have become available. This ability has great potential to allow us to evaluate how the ecological niche requirements for a given species may change as a result of global climate change, and thus address conservation concerns.

Narrow endemic species, defined as those unique to their geographic locations and not widely geographically distributed, are a priority for determining what factors (both abiotic and biotic) may limit geographic range expansion. This is because not only are these species expansion-limited by some external factor, they are also at a high extinction risk due to having small population sizes. As a result of effective population sizes are often decreased, and drift and inbreeding become stronger forces that can prevent adaptation (Kruckeberg & Rabinowitz, 1985; Newman & Pilson, 1997). In addition, global climate change is expected to exacerbate conditions for endemic species in several different ways. For example, anthropogenic climate change is predicted to

increase plant and animal extinction rates roughly 20-35% across all species by the year 2050 under several emissions scenarios (Thomas *et al.*, 2004). The environmental change predicted by these estimates may drastically influence the future survival of endemic species. Specifically, endemic plant species extinctions due to global warming are projected to be on average 11.6% worldwide by the year 2100 (Malcolm *et al.*, 2006). These projections are of particular concern for the southeastern United States, which is a hotspot for endemic species due to Pleistocene glaciation events (Estill & Cruzan, 2001; Petit *et al.*, 2003; Bennett & Provan, 2008; Rull, 2009). Regional temperatures in the southeastern United States are projected to increase 4.5ºF by the 2080s under the lower emissions scenario B1, and 9ºF by the higher emissions scenario A2 (Canadell *et al.*, 2007; IPCC, 2007; USGCRP, 2009). Additionally, overall rainfall is expected to decrease, drought length and severity is predicted to increase, and sea level rise in the form of coastal inundation and shoreline retreat is expected to change the landscape in the southeast, all of which will have severe effects on plant and animal habitats along the coast (USGCRP, 2009). These anticipated changes make endemic species in the southeastern U.S. a cogent target for conservation studies. Conservation efforts with the maximal impact will require a strategy for identifying which endemic species are at highest risk for being lost due to these predicted environmental and climatic changes.

Traditional approaches for determining range-limiting factors like reciprocal transplant experiments tend to be time-consuming. These approaches seek to investigate one or both of the two broad reasons proposed for range limitations, which are (1) a dispersal limitation, which is controlled by ecological and geographical factors, and (2) a limitation to adaptation across heterogeneous environments, which is controlled by

genetic variation (Eckhart *et al.*, 2011). Both limitations likely apply to endemic species with small populations that are geographically disjunct. On its face, determining potential limitations to the act of dispersal itself may seem straightforward. One would expect that an assessment of seed morphology would indicate whether or not plants have the potential for long-distance dispersal. However, the act of dispersal is challenged by dense vegetation, and the establishment and success of all dispersal propagules depends on the microsite habitat upon their arrival (Nathan & Muller-Landau, 2000; Münzbergová & Herben, 2005). The interaction between seed and habitat at the microsite level is what is most challenging to tease apart, and certainly warrants more attention (Nathan & Muller-Landau, 2000; Donohue *et al.*, 2010).

Similarly, the assessment of genetic variation in small populations of endemic species is a process that is both time consuming and financially challenging, even with recent technological advances. Species that have not previously been studied require the development of genetic markers (e.g., microsatellites or single nucleotide polymorphisms), which then must be analyzed in many individuals and populations. Additionally, monitoring populations for adaptive genetic responses requires repeated measurements over long periods of time (Hansen *et al.*, 2012). While this approach yields highly valuable biological insights, it may not be practical for conservation management in some cases given the short amount of time in which dramatic changes in climate are predicted to occur. This is particularly germane for species that already are narrow in range and are predicted to be significantly impacted by changes in climate, like narrow endemic species. A more feasible and likely fruitful first step to identifying putative factors responsible for shaping geographical distributions in a limited time frame is to

focus on identifying the ecological and geographical constraints using modeling approaches.

A powerful way to identify the most at-risk species is by determining what environmental factors shape the geographic ranges of a selection of species, and then evaluating the predictive change of those environmental factors. Species distribution models (SDMs) are an ideal tool for such a task. SDMs encompass a broad range of heuristic and statistical approaches, and have become an increasingly useful tool because they can reveal what environmental factors correlate strongly with species presence in a given location. SDMs incorporate ecological niche (*sensu* Grinnell, 1917) concepts into predictive modeling techniques to provide a comprehensive understanding of a species' geographical distribution (Kearny, 2006). SDMs, like ecological niche modeling approaches, have been criticized for their failure to incorporate biotic interactions and gene flow into their models (Araújo & Guisan, 2006; McInerny & Etienne, 2012). However, it is necessary to simplify the modeling process on some level due to paucity of such data at the landscape level as well as constraints in computer processing time and power. In addition to identifying what environmental factors explain a species given range, SDMs have the potential to identify new areas for further exploration and discovery of novel populations (see de Siqueira *et al.*, 2009 for an example with the rare plant, *Byrsonima subterranea*). This utility is particularly significant in the context of conservation and restoration of rare and threatened species.

In this study, we utilized bioclimatic and soil data in conjunction with presence data for threatened and endangered species to generate SDMs for nine plant species. Our objectives were to 1) identify ecogeographical factors integral to explaining the current

distributions of threatened plant species, 2) model distributional changes based on SDMs generated from objective (1), and 3) compare current and predicted future ranges (size, location, etc.) to determine which species should be a top conservation priority.

Materials and Methods

Study Species

A broad phylogenetic sampling of plant species, comprising a total of nine species spanning seven families, were used for this study. These species are the following: *Croomia pauciflora* (Stemonaceae), *Pachysandra procumbens* (Buxaceae), *Panax quinquefolius* (Araliaceae), *Polygonella basiramia* (Polygonaceae), *Polygonella macrophylla* (Polygonaceae), *Polygonella myriophylla* (Polygonaceae), *Silene polypetala* (also known as *Silene catesbaei*, Caryophyllaceae), *Trillium lancifolium* (Melanthiaceae), and *Waldsteinia lobata* (also known as *Geum lobata*, Rosaceae). These species were selected based on their range and conservation status—all species occur in the southeastern United States and are either federally or state listed (or both) as threatened, endangered, or rare. Every species, with the exception of *P. quinquefolius*, is considered threatened/rare across its entire range. Within the area of study (southeastern United States, Figure 1), *P. quinquefolius* is classified as rare in Mississippi, Georgia, South Carolina, North Carolina, and Tennessee. Outside of the southeast, *P. quinquefolius* is listed as threatened/rare in fifteen states. Of these species, *Panax quinquefolius* has the broadest geographic range, spanning thirty-three states, and conversely, *Polygonella basiramia* has the most severely limited range, found in only two counties in Florida. Because each of these species is federally or state listed, populations of each are carefully

tracked on a county-level basis by each state's Natural Heritage Programs (or equivalent programs). For this reason, we were able to obtain exhaustive population location records for each of our species from the states in which they occur as threatened species, thereby overcoming the problem of incomplete population sampling with respect to SDMs (Reddy & Dávalos, 2003; Phillips *et al.*, 2009).

Habitats of these species vary, but many occur as deciduous forest sub-story species. The exceptions to this are *Polygonella basiramia, P. macrophylla,* and *P. myriophylla*, which are herbaceous to subshrub species that occur, respectively, in sand hill scrub, sand pine-oak scrub, and Florida rosemary scrub habitats. Each species in this study clearly has a unique combination of habitat preferences, range specifics, conservation status, and lifespan (Table 1 and Figure 1). All but one of the species in this study are perennials (the exception being *Polygonella basiramia*). Though not all of these species have been investigated for levels of population genetic variation, several species have been assessed for within-population levels of genetic diversity (H_e in Table 1; Lewis & Crawford, 1995; Grubbs & Case, 2004; Cruse-Sanders *et al.*, 2005; Hamrick & Pattavina, unpublished data). Of those that have been assessed, all have relatively low levels of genetic diversity, with the exception of *Silene polypetala* (H_e = 0.262, 20 populations; Hamrick & Pattavina, unpublished data) when compared to mean levels of genetic variation for plant species with narrow/endemic ranges and selfing or mixed-mating systems (Hamrick & Godt, 1996). This high level of genetic diversity in *S. polypetala* is explained most likely by its polyploidy state.

Species Distribution Modeling

Species Population Data

It is logical to begin the process of identifying the most at-risk plant species candidates by targeting species that have already been identified as threatened or endangered by state and federal agencies. We took advantage of the exhaustive and careful surveys done by state agencies like the Department of Natural Resources as a starting point for conceptualizing the distribution of a species. These organizations endeavor to track carefully and thoroughly the distribution of populations of rare and threatened species. We utilized species population data collected by such agencies and converted these data to a uniform format. Data sources for each state and species are provided in the appendix (Table S1; Supplemental Materials). These data were obtained either in the format of ArcGIS shapefiles (polygon or point), or as Excel files of GPS coordinates, which were converted to point shapefiles in ArcMap 10.1 (ESRI 2012). Any data in polygon shapefiles were converted to point data using the "Feature to Points" tool in ArcGIS v10.1, which locates a point at the centroid of the polygon.

Ecogeographical Data and Modeling Conditions

Elevation and climate data at a resolution of ~1 km were obtained from the Bioclim data set from WorldClim (www.worldclim.org). The BioClim data set is composed of 19 bioclimatic layers representing climatic trends, which are derived from monthly values of precipitation and temperature data (e.g., maximum temperature of the warmest month). Because these layers are not independent climatic layers, a subset of 1000 values from each layer (sampled using a random points layer generated via the "Sample" tool in ArcMap v10.1) were extracted from all 19 BioClim layers and a

Principal Components Analysis (PCA) was performed in R (R Core Team, 2009). We retained layers for the two axes with the largest eigenvalues (that explained the largest proportion of the variance). We narrowed our BioClim layers down to a subset of six—Temperature Seasonality (Bio4), Mean Temperature of the Wettest Quarter (Bio8), Annual Precipitation (Bio12), Precipitation of the Wettest Quarter (Bio16), Precipitation of the Driest Quarter (Bio17), Precipitation of the Warmest Quarter (Bio18). We also utilized a soil texture layer derived from the STATSGO data from the Natural Resources Conservation Service (NRCS, USDA). Our steps for configuring the soils data may be found in the supplementary materials (Figure S1). For various soil types, the proportion of components is most important for the drainage capacity. It is worth noting that these soil layers are currently modeled to remain constant for all SDMs generated for future climate scenarios (2050 and 2080), and we have not analyzed dynamic models for soil conditions.

To evaluate how projected global climate change might affect range distributions of our study species, SDMs for each of the nine species were generated for current climate conditions and for 2050 and 2080 using the A1B emissions scenario and the UKMO:HADGEM1 climate model from the 4th Intergovernmental Panel on Climate Change (IPCC, 2007) in Maxent (Phillips *et al.*, 2004). Maxent (maximum entropy) is a machine-learning method that uses presence-only data and environmental data across a given landscape to model an estimate of where suitable habitat outside the current distribution may be located, and mathematically is equivalent to a maximum likelihood approach (Phillips *et al.*, 2006). Our Maxent runs used default values for all settings with the exception of including 15 subsampled replicates to generate a confidence envelope, a

25% random test percentage using the random seed option to generate the AUC scores and 5000 iterations. We converted the output from Maxent into binary suitability/unsuitability maps using two Maxent-defined values, one liberal and one conservative, as thresholds. Our liberal threshold, Minimum Training Presence, sets the cutoff for suitability equal to the logistic suitability value associated with the least suitable training presence record, so that all locations with these least suitable attributes or better are classified as suitable (Table S2). Our strict threshold on the other hand, the 10 Percentile Training Gain value, uses the suitability of the training presence record found in the tenth percentile as a cutoff, such that 10% of the presence records have suitability values that will fall below that cutoff, making it a more conservative outcome of the model (Table S2). Henceforth, we will refer to these two categories of output as liberal and conservative, respectively.

The amount (km^2) of suitable habitat for each of the climate scenarios (current, 2050, and 2080) was calculated for both the conservative and liberal threshold estimates, and percent suitable habitat loss was estimated for current versus future climate scenarios. We evaluated the performance of our model using Area Under the Curve (AUC) scores, which are generated by setting aside a randomly selecting 25% of the presence data during each model iteration and using these data to test the generated model from each iterative run. AUC scores evaluate the predictive performance of SDMs, and a value of 0.5 is associated with random predictions, whereas values greater than 0.5 are associated with better-than-random performance of the model (Franklin, 2009). Generally speaking, models with >0.9 AUC scores are considered to have highly accurate predictive power, and models with AUC scores ranging from 0.7 – 0.9 have moderately

accurate predictive power (Swets, 1988; Manel *et al.*, 2001). Additionally, Maxent provides estimates of how important each environmental layer was to the development of the model. These estimates can, in turn, be interpreted as which environmental layers most strongly correlate with the presence of a given species (and thus are predictors of its presence in other locations). Because Maxent gives a rank of importance for environmental layers, we are able to establish which environmental layers are the most and the second most important predictors for species presence. We will refer to these as the primary and secondary predictors, respectively.

Results

Environmental Predictors

Temperature correlates and soil type were the most common primary environmental predictors for habitat suitability (Table 2). Temperature seasonality (the greatest annual range of temperatures in a geographical area) was the strongest predictive variable for each of the *Polygonella* species, with values of 42ºC of fluctuation for *P. basiramia* and *P. myriophylla* (which have nearly identical ranges) and 65ºC of fluctuation for *P. macrophylla*. For *Panax quinquefolius* and *Waldsteinia lobata*, mean temperature in the wettest quarter (MTWQ) followed by elevation were the two most important environmental predictors of presence, with values of 3ºC and 7ºC, respectively for MTWQ. Soil type (specifically, some derivative of loam) was the most important environmental predictor for *Croomia pauciflora*, *Silene polypetala,* and *Trillium lancifolium*. *Pachysandra procumbens* was the only species for which precipitation was the primary environmental predictor, with an average of 430 mm of rainfall during the

wettest months of the year strongly correlating with presence of *P. procumbens* populations. Soil type was a secondary predictor for *P. procumbens*, with loam derivatives again being the most highly correlated soil types. Finally, elevation was a secondary environmental predictor for two species—*Panax quinquefolius* and *Waldsteinia lobata*, with higher elevations being strongly correlated with species presences. Across these nine species, those that were most limited in suitable habitat predictions were most strongly correlated with temperature correlates (all three *Polygonella* species, *Silene polypetala* (secondary predictor), and *Waldsteinia lobata*).

Predicted Suitable Habitat in Current and Future Climate Scenarios

All models had strong predictive power, with AUC scores falling above 0.9 with one exception (*Pachysandra procumbens*, AUC=0.887; Table 2). The species with the greatest amount of suitable habitat estimated under current climate conditions were *Pachysandra procumbens* (ranging from ~348,000-683,000 km^2 depending on conservative vs. liberal estimates; Figures 2b1 and 3, Table 3), *Panax quinquefolius* (~282,000-879,000 km^2; Figures 2c1 and 3, Table 3), and *Trillium lancifolium* (~148,000-454,000 km^2; Figures 2h1 and 3, Table 3). Conversely, *Polygonella macrophylla* had the least amount of suitable habitat predicted under current climate conditions, with roughly 2,500 km^2 under conservative estimates and 12,400 km^2 under liberal estimates (Figures 2e1 and 4, Table 3).

The majority of species (six of the nine) demonstrated drastic amounts of predicted habitat loss for the 2050 and 2080 climate scenarios, the most drastic of which was the predicted loss for *Waldsteinia lobata*. Its total amount of suitable habitat for the current climate is estimated at ~32,000 km^2 for the conservative threshold and ~116,000

km^2 for the liberal threshold. For both 2050 and 2080 projections under conservative estimates, *W. lobata* is predicted to experience a 100% loss of suitable habitat (Figure 5, Table 2). Liberal estimates for *W. lobata* are not much better, with predicted loss by 2050 at 95.6%, and 100% loss predicted by 2080. The species with the least predicted habitat loss is *Panax quinquefolius*, with 75-90% loss from current amounts predicted for 2050-2080 respectively under conservative estimates, and 45-73% (2050-2080) loss under liberal estimates (Figure 5, Table 2).

Interestingly, three species (*Silene polypetala, Trillium lancifolium,* and *Polygonella myriophylla*) are expected to experience an actual increase in the amount of suitable habitat for 2050 and 2080 predictions in at least one of the thresholds used (Figures 2(f-h), 5 and 6; Tables 3 and 4). *S. polypetala*, which is estimated to have roughly 34,000-43,000 km^2 (conservative-liberal) of suitable habitat under current conditions, is predicted to experience up to a 24.1% gain by 2080 under conservative estimates and up to a 15.2% gain by 2080 under liberal estimates. *T. lancifolium* is predicted to fluctuate under conservative estimates, gaining 58.7% in area that is suitable by 2050, and gaining 27.7% by 2080 compared to current suitable habitat. Alternatively, under liberal estimates, *T. lancifolium* is predicted to experience a 32.2% decrease in suitable habitat by 2080. Finally, *P. myriophylla* is expected to undergo substantial increases in suitable habitat by 2080, with a predicted 48.1% increase under conservative estimates and 116.3% increase under liberal estimates. It is worth noting that *P. macrophylla* is predicted to experience minor increases in suitable habitat by 2050 under both threshold estimates, but both liberal and conservative estimates predict a substantial decrease in suitable habitat by 2080 (37.5% and 79.1%, respectively). This species is

driven by temperature seasonality and mean precipitation in the driest quarter (Table 2). Likely, this asymptotic response is due to the fact that temperature seasonality should take longer to be affected by predicted changes in climate, while precipitation decreases should occur much more quickly.

In addition to changes in total amounts of suitable habitat, the location of habitat that is suitable for some of the species in this study is predicted to shift in geographical location. This is the case for *P. basiramia*, *P. myriophylla*, *S. polypetala*, and *T. lancifolium* ("New" area denoted in Figures 2 (d, f-h), 5, and 6). For these species, the percentage amount of new area predicted as suitable, whether by liberal or conservative estimates, is greater than 50% of total suitable area for 2050 and 2080 estimates. These amounts of new area for 2050 and 2080 respectively are as follows: 73% and 77% for *P. basiramia,* 51% and 57% for *P. myriophylla*, 83% and 78% for *S. polypetala*, and 63% and 62% for *T. lancifolium*.

Discussion

The Environmental Predictors of Geographical Distributions

Our models indicate that correlates of temperature, such as temperature seasonality (each of the *Polygonella* species) and mean temperature in the wettest quarter (*Panax quinquefolius* and *Waldsteinia lobata*), were most often the primary environmental predictors correlated with species presence data. Given climate change predictions for temperature in the southeastern U.S. (increases in temperature by as much as 9ºF by the 2080s), one would expect that those species for which temperature correlates are a primary predictor in their respective SDMs would be at a high risk for

projected habitat loss, and indeed this was the case with our models. In addition to decreases in suitable habitat, one study predicts that increases in temperature will greatly alter flowering phenology by shifting flowering date an average of 2.4 days earlier per 1ºC increase in temperature, based on trends in north-central U.S. (Calinger *et al.*, 2013). This change in phenology could potentially create mismatches among plants and their respective pollinators, particularly in plant species that are specialist pollinated (Memmott *et al.*, 2007; Hegland *et al.*, 2009; González-Varo *et al.*, 2013).

Soil type was the second most common primary predictor that emerged for the SDMs, specifically of the models for *Croomia pauciflora*, *Silene polypetala*, and *Trillium lancifolium.* Each of these species was strongly correlated with some derivative of loam, which is a high nutrient- and humus-content soil composed of sand, silt, and clay in roughly equivalent proportions (Kaufman & Cleveland, 2007). However, variations of loam may yield sandy loam or silty loam, which would have higher draining capacity than clay loam. Additionally, loam could contain other components like gravel, which would also serve to increase drainage capacity. The soil types of the above-mentioned species are all loam derivatives that incorporate these types of drainage-increasing components. As a result of the increase in temperatures and decrease in overall precipitation levels predicted for climate change, soil aridity is expected to increase (USGCRP, 2009). As previously mentioned, we expect that soils with higher drainage capacities are likely to be more affected by increases in aridity. If this is the case, plant populations that are likely adapted to current soil moisture and nutrient conditions in loamy soils should be more greatly affected by predicted changes in climate. As a result, we would expect that species correlating strongly with soil type should experience lower

amounts of predicted habitat loss due to climate change. This is the case for two of the three species (*S. polypetala* and *T. lancifolium*), but not for *C. pauciflora*. This result may be influenced by soil type's being a static variable in our models. This is not because soil type is not expected to change with future global climate changes, but rather because these data predicting soil changes on a landscape level do not exist. One way to greatly improve future modeling efforts will be to develop dynamic soil layers that reflect the resultant soil characteristics expected from climatic changes. The inclusion of such layers may allow more realistic predictions of habitat change for threatened plant species.

Because the soil layers in this study are static, any changes we see in habitat amounts in future predictions for species that have soil as their primary predictor could be driven by secondary environmental predictors. The secondary predictor for *C. pauciflora* is precipitation in the driest quarter. Predicted amounts of suitable habitat loss for this species are 84-95% by 2080, depending on conservative vs. liberal estimates. Given the predicted decrease in overall precipitation amounts for climate change scenarios, this change is not surprising. In contrast, *Silene polypetala* is predicted to experience an increase in suitable habitat for conservative and liberal future climate estimates. The secondary predictor for this species is temperature seasonality, which appears to be driving the distinctive shifts in the location of suitable habitat for this species.

Habitat Loss Predictions and At-Risk Species

The species in this study that are predicted to experience the greatest amounts of suitable habitat loss are *Waldsteinia lobata* (100%, 100%) and *Croomia pauciflora* (94%, 84%) under both conservative and liberal estimates, respectively, and *Polygonella*

basiramia under conservative estimates (99%). To our knowledge, there have not been any studies that have analyzed levels of population genetic variation in *W. lobata* or *C. pauciflora*. We argue that following the prioritization of species at a certain risk level, this is an important next step before making a decision regarding conservation strategy for these species (see discussion below on assisted migration). Within-population genetic diversity (He) for *P. basiramia* was 0.0845 (3 populations, allozymes; Lewis & Crawford, 1995); this estimate is low relative to what is expected for an annual plant that is narrowly distributed and/or endemic (Hamrick and Godt, 1996). That said, while increasing the number of populations evaluated would yield a more reliable result, we do not expect that this value would likely vary much as populations of this plant tend to be small and isolated. Accurately assessing levels of genetic variation is imperative for rare species, particularly those that are not capable of long-distance dispersal, in that a species' ability to evolve in response to environmental change is inherently tied to its level of genetic variation (Lande & Shannon, 1996). Given that populations with higher diversity levels will likely have a better chance at survival in the long term, the consideration of genetic diversity in populations of at-risk species should help guide conservation efforts (Holsinger & Gottlieb, 1991). We recommend that based on the predicted amounts of habitat loss for these species, investigations into their population genetic diversity be initiated to help target which populations are the best candidates for future conservation efforts.

While *W. lobata, C. pauciflora,* and *P. basiramia* are predicted to experience the most habitat loss, *S. polypetala, T. lancifolium* and *P. basiramia* face a different challenge. Both conservative and liberal estimates for 2050 and 2080 predict the

geographic center of the species range will shift, indicating that many new areas will need to be colonized to realize this change in the geographic location of suitable habitat. The ability of a plant population to colonize new area is dependent on the frequency and distance of dispersal events, and very few plant species disperse outside the average 10-1500 m range from the maternal plant (Corlett & Westcott, 2013). For gravity-dispersed plant species, the expected dispersal distance is likely ≤10 m. Even for myrmecochorous species like many species of the *Trillium* genus, maximum dispersal distances are generally <10 m (Cain *et al.*, 2000). This means that even when we have predictions of increased suitable habitat like those for *S. polypetala*, the likelihood that populations of these species would be able to shift to these new areas of suitable habitat is very low. Even in the event that these species were able to track suitable climate and shift their ranges to these new suitable areas, low levels of standing genetic variation could prevent adaptation to these new locations, and colonizations could result in a confounding effect of founder events due to strong drift (Klopfstein *et al.*, 2006; McInerny *et al.*, 2009). Finally, these species are already found only in few extant populations (40 known populations of *S. polypetala*, 37 of *T. lancifolium*, and 137 of *P. basiramia*). So, even though our initial predictions are positive with respect to the amount of suitable habitat available by 2050 and 2080 for *S. polypetala* and *T. lancifolium*, these species still face considerable challenges. For *P. basiramia*, which is predicted to experience both a range shift and considerable decreases in suitable habitat (50-100% for liberal and conservative estimates, respectively), it is unlikely that these populations will be capable of surmounting dispersal barriers (mainly dense human development), and our models suggest that this species is facing extinction in the absence of human intervention. The

one species standing in exception to all of the above listed challenges is *P. myriophylla.* This species is predicted to experience an increase in suitable habitat (under both conservative and liberal estimates) by 2080 via expansion of the area that it currently inhabits. This is the only species in this study that is not considered at immediate risk.

Conservation and Climate Change

It is possible that evaluating levels of population genetic diversity in the high-risk species identified in this study would take too long to be relevant, given that climatic changes are predicted to occur within the next 80-100 years. Climate projections are likely to outpace rates of niche evolution in vertebrate species and require an evolutionary rate >10,000 times typical rates based on historical data for roughly 500 species (Quintero & Wiens, 2013). One would expect this number to be even more severe for perennial endemic plant species that are geographically isolated and likely have low levels of genetic diversity. For this reason, a more immediate and practical conservation strategy that can be tested in the short term is crucial to the persistence of many of these species. One such strategy is assisted migration (AM).

AM has received much contention in the literature, but there are very few alternative approaches with respect to the conservation of highly threatened plant species. While many argue the utility of habitat improvement and landscape connectivity for animal species (Gillson *et al.*, 2013), these options are not practical for plants. Despite dissent in the literature, the overwhelming majority of recent studies support the idea of AM. The primary objection to AM is that it involves the introduction of non-native species, though a more reasonable view of AM is the relocation of plants to the broadening edges of the geographic range (Corlett & Westcott, 2013). This process

would allow at-risk species to potentially expand, and continuing this process may allow the establishment of permanent populations. Sax *et al.* (2013) similarly argue that there are different extents to AM, distinguishing among distances of relocation efforts—short, medium, and long-distance AM. Deciding on a relocation distance is dependent upon how much of the fundamental niche is actually occupied (the realized niche), and how this relates to the tolerance niche (the set of ecological conditions where a population can reside but not establish in the long term; Sax *et al.*, 2013). Recent studies suggest that in the face of rapidly changing climates and low dispersal capacities of plant populations, the only truly feasible conservation option is AM (Thomas, 2011). To help with the decision-making process, Shoo *et al.* (2013) have developed a decision framework upon which to base research and conservation decisions given the amount of known information for a species of interest. Ultimately, however, they agree that in the absence of frequent long-distance dispersal and population genetic variation, some form of AM is the optimal conservation approach, and populations with the highest amounts of genetic diversity should be targeted for this strategy.

Our findings of projected severe losses are in agreement with previous studies that have examined potential habitat loss for narrow endemic species in the context of future climate change (e.g., see Dullinger *et al.*, 2012; Guisan *et al.*, 2013; McCallum *et al.*, 2014). In particular, McCallum *et al.* (2014) generated SDMs in conjunction with performing population genetic analyses with AFLP and chloroplast microsatellites for the narrow endemic Needle Bottlebrush (*Callistemon teretifolius*). Populations in three of the seven geographic regions studied were predicted to have extreme habitat loss of roughly 90% or greater. The authors recommended immediate seed collection from populations

with greatest diversity for restoration efforts (McCallum *et al.*, 2014). It is important to note, however, that these studies have been criticized due to the uncertainty in the predictions made by SDMs (Dormann, 2007; Schwartz, 2012; Conlisk *et al.*, 2013). For example, Schwartz (2012) noted that these predictive studies that focus on narrow endemics are predisposed to generate geographically small areas of habitat suitability, because these species initially have a narrow climatic envelope (see also Ohlemüller *et al.*, 2008). As a result, these types of species are more likely to be projected as vulnerable in the context of climate change. Schwartz (2012) cited the inability of such studies to incorporate factors like phenotypic plasticity, dispersal capacity, and biotic interactions. In addition, Schwartz (2012) suggested that another weakness of these predictive models is their reliance on the assumption that limiting factors will remain constant and that species will not adapt to the climatic distributional predictors. As a result, SDMs will inherently overestimate the severity of habitat loss.

While these criticisms are valid, we argue that SDMs are appropriate in the context of this study for several reasons. First, perennial plant species can have relatively long generational times, and the species in this study are not capable of long-distance dispersal to our knowledge. These factors should constrain dispersal rates in these plants. Under current climate predictions, Davis & Shaw (2001) estimate that range shifts will have to occur at a rate of 300-500 km per century. This is twice the rate of the fastest documented range shift involving trees capable of long-distance dispersal (white spruce; Ritchie & MacDonald, 1986). Additionally, adaptation is generally restricted at the trailing edge of a species where gene flow from more adapted populations at the leading edge is often restricted (Davis & Shaw, 2001; Anderson *et al*. 2009). Thus, we predict

that evolutionary adaptation to climatic drivers is unlikely. Second, while it is reasonable to expect that limiting factors for animal species may not remain constant, plants are sessile. Therefore, it is logical to assume that their climatic drivers are more likely to remain constant. In this study, this means that the list of primary and secondary predictors should be the same at any give time point (current vs. future climates). Finally, because SDMs do not incorporate information pertaining to genetic diversity, phenotypic plasticity, dispersal capacity, and biotic interactions, it is likely that predictions pertaining to future suitable habitat are actually broader than may be the case for the future realized areas of occurrence.

Despite the criticisms listed above, we suggest that SDMs have great utility for conservation and restoration efforts (see Guisan *et al.*, 2013). SDMs can help to delineate between closely related congeners with respect to niche requirements and taxonomic status (Lipsen *et al.*, 2013). Additionally, SDMs more efficiently delineate the boundaries of extent of occurrence (EOO) polygons than do α–hulls (Marcer *et al.*, 2013; de Castro Pena *et al.*, 2014), which are the current standard suggested by the IUCN for the estimation of EOOs and for determining the threatened/endangered status of species. With respect to the species in this study, we suggest that several of those we analyzed are eligible for inclusion on the IUCN red list for conservation under Criterion B1 (IUCN, 2014). These are *Polygonella basiramia, P. macrophylla, P. basiramia*, and *Waldsteinia lobata* (Table 4). In particular, *W. lobata* meets the criteria for being critically endangered by the year 2100. It is important to note that Criteria B require that in addition to species meeting habitat range minimums (<20,000 km^2 for "Vulnerable," <5,000 km^2 for "Endangered," and <100 km^2 for "Critically Endangered"), species distributions must

also be highly fragmented and/or have a projected reductions in habitat range (IUCN, 2014).

Conclusions

The majority of species in this study (six of the nine) are predicted to experience dramatic reductions in the amount of available suitable habitat under both conservative and liberal estimates by the year 2080. Those few species (with one exception: *P. myriophylla*) that are predicted to experience an increase in suitable habitat are not expected to track along with suitable climate and environment, be it from dispersal limitations and/or genetic limitations. The species at greatest risk in this study, according to our models, are *Waldsteinia lobata, Croomia pauciflora,* and *Polygonella basiramia*. We argue that these species are in need of immediate conservation efforts, and we propose that assisted migration is the best approach for the conservation of these species. Additionally, given the immediacy with which climatic changes are occurring, we suggest that utilizing SDMs will be the most rapid and efficient way to identify at-risk species with respect to global climate change. This study has identified areas that will be suitable habitat for AM efforts for several at-risk species, and this approach can be applied to any plant or animal system where habitat reduction may be a product of climate change.

Literature Cited

Anderson, B.J., Akçakaya, H.R., Araújo, M.B., Fordham, D.A., Martinez-Meyer, E., Thuiller, W. & Brook, B.W. (2009) Dynamics of range margins for metapopulations under climate change. *Proceedings of the Royal Society B: Biological Sciences*, **276**, 1415-20.

Araújo, M.B. & Guisan, A. (2006) Five (or so) challenges for species distribution modeling. *Journal of Biogeography,* **33,**1677-88.

Bennett, K.D. & Provan, J. (2008) What do we mean by 'refugia'? *Quaternary Science Reviews*, **27**, 2449-55.

Cain, M.L., Milligan, B.G., & Strand, A.E. (2000) Long-distance seed dispersal in plant populations. *American Journal of Botany*, **87,** 1217-27.

Calinger, K.M., Queensborough, S. & Curtis, P.S. (2013) Herbarium specimens reveal the footprint of climate change on flowering trends across north-central North America. *Ecology Letters,* **16,** 1037-44.

Canadell, J.G., Le Quéré, C., Raupach, M.R., Field, C.B., Buitenhuis, E.T., Ciais, P., Conway, T.J., Gillett, N.P., Houghton, R.A. & Marland, G. (2007) Contributions to accelerating atmospheric CO2 growth from economic activity, carbon intensity, and efficiency of natural sinks. *PNAS,* **104,** 18866-70.

Conlisk, E., Syphard, A.D, Franklin, J., Flint, L., Flint, A. & Regan, H. (2013) Uncertainty in assessing the impacts of global change with coupled dynamic species distribution and population models. *Global Change Biology,* **19,** 858-69.

Corlett, R.T. & Westcott, D.A. (2013) Will plant movements keep up with climate change? *Trends in Ecology and Evolution***, 28,** 482-8.

Cruse-Sanders, J.M., Hamrick, J.L., & Ahumada, J.A. (2005) Consequences of harvesting for genetic diversity in American ginseng (*Panax quinquefolius* L.): a simulation study. *Biodiversity and Conservation,* **14,** 493-504.

Davis, M.B. & Shaw, R.G. (2001) Range shifts and adaptive responses to Quaternary climate change. *Science*, **292**, 673-9.

de Castro Pena, J.C., Kamino, L.H.Y., Rodrigues, M., Mariano-Neto, E. & de Siqueira, M.F. (2014) Assessing the conservation status of species with limited available data and disjunct distribution. *Biological Conservation,* **170**, 130-136.

de Siqueira, M.F., Durigan, G., Junior, P.dM. & Peterson, A.T. (2009) Something from nothing: Using landscape similarity and ecological niche modeling to find rare plant species. *Journal for Nature Conservation,* **17,** 25-32.

Donohue, K., de Casas, R.R., Burghardt, L., Kovach, K. & Willis, C.G. (2010) Germination, postgermination, adaptation, and species ecological ranges. *Annual Review of Ecology, Evolution, and Systematics,* **41**, 293-319.

Dormann, C.F. (2007) Promising the future? Global change projections of species distributions. *Basic and Applied Ecology*, **8**, 387-397.

Dullinger, S., Gattringer, A., Thuiller, W., Moser, D., Zimmerman, N.E., Guisan, A., Willner, W., Plutzar, C., Leitner, M., Mang, T., Caccianiga, M., Dirnböck, T., Ertl, S., Fischer, A., Lenoir, J., Svenning, J.C., Psomas, A., Schmatz, D.R., Silc, U., Vittoz, P. & Hülber, K. (2012) Extinction debt of high-mountain plants under twenty-first-century climate change. *Nature Climate Change,* **2,** 619-22.

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

ESRI (Environmental Systems Resource Institute). (2012) *ArcMap, version 10.1*. ESRI, Redlands, CA.

Estill, J.C. & Cruzan, M.B. (2001) Phytogeography of rare plant species endemic to the southeastern United States. *Castanea*, **66**, 3-23.

Franklin, J. (2009) *Mapping species distributions*. Cambridge University Press, Cambridge.

Gillson, L., Dawson, T.P., Jack, S. & McGeoch, M.A. (2013) Accomodating climate change contingencies in conservation strategy. *Trends in Ecology and Evolution,* **28**, 135-42.

González-Varo, J.P., Biesmeijer, J.C., Bommarco, R., Potts, S.G., Schweiger, O., Smith, H.G., Steffan-Dewenter, I., Szentgyörgyi, H., Woyciechowski, M. & Vilà, M. (2013) Combined effects of global climate change pressures on animal-mediated pollination. *Trends in Ecology and Evolution,* **28**, 524-30.

Grinnell, J. (1917) Field tests of theories concerning distributional control. *The American Naturalist,* **51**, 115-28.

Grubbs, H.J. & Case, M.A. (2004) Allozyme variation in American ginseng (*Panax quinquefolius* L.): Variation, breeding system, and implications for current conservation practice. *Conservation Genetics,* **5,** 13-23.

Guisan, A., Tingley, R., Baumgartner, J.B., Naujokaitis-Lewis, I., Sutcliffe, P.R., Tulloch, A.I.T., Regan, T.J., Brotons, L., McDonald-Madden, E., Mantyka Pringle, C., Martin, T.G., Rhodes, J.R., Maggini, R., Setterfield, S.A., Elith, J., Schwartz, M.W., Wintle, B.A., Broennimann, O., Austin, M., Ferrier, S., Kearney, M.R., Possingham, H.P. & Buckley, Y.M. (2013) Predicting species distributions for conservation decisions. *Ecology Letters,* **16,** 1424-1435.

Hamrick, J.L. & Godt, M.J.W. (1996) Effects of life history traits on genetic diversity in plant species. Philosophical Transactions: *Biological Sciences,* **351,** 1291-8.

Hansen, M.M., Olivieri, I., Waller, D.M., Nielsen, E.E. & The GeM Working Group. (2012) Monitoring adaptive genetic responses to environmental change. *Molecular Ecology,* **21,** 1311-29.

Hegland, S.J., Nielsen, A., Lázaro, A., Bjerknes, A.-L. & Totland, Ø. (2009) How does climate warming affect plant-pollinator interactions? *Ecology Letters,* **12,** 184-95.

Holsinger, K.E. & Gottlieb, L.D. (1991) Conservation of rare and endangered plants: principles and prospects. *Genetics and Conservation of Rare Plants.* (Eds. D.A. Falk and K.E. Holsinger), pp. 195-208. Oxford University Press, New York.

IPCC (Intergovernmental Panel on Climate Change). (2007) Climate change 2007: climate change impacts, adaptation, and vulnerability. Summary for policymakers. www.ipcc.ch

IUCN Standards and Petitions Subcommittee. (2014) Guidelines for using the IUCN red list categories and criteria. Version 11. Prepared by the Standards and Petitions Subcommittee. URL: http://www.iucnredlist.org/documents/RedListGuidelines.pdf

Kaufman, R. & Cleveland, C. (2007) *Environmental Science*. McGraw-Hill, New York, NY.

Klopfstein, S., Currat, M. & Excoffier, L. (2006) The fate of mutations surfing the wave of a range expansion. *Molecular Biology and Evolution,* **23,** 482-90.

Kruckeberg, A.R. & Rabinowitz, D. (1985) Biological aspects of endemism in higher plants. *Annual Review of Ecology and Systematics*, **16,** 447-79.

Lande, R. & Shannon, S. (1996) The role of genetic variation in adaptation and population persistence in a changing environment. *Evolution,* **50,** 434-7.

Lewis, P.O. & Crawford, D.J. (1995) Pleistocene refugium endemics exhibit greater allozymic diversity than widespread congeners in the genus *Polygonella* (Polygonaceae). *American Journal of Botany,* **82,**141-9.

Lipsen, L.P.J., Lee, C. & Whitton, J. (2013) Genetic and ecological differences between two Utah endemics: US federally threatened *Townsendia aprica* and its close congener *T. jonesii* var. *lutea* (Asteraceae). *Botany,* **91,** 242-50.

Malcolm, J.R., Liu, C., Nielson, R.P., Hansen, L. & Hannah, L. (2006) Global warming and extinctions of endemic species from biodiversity hotspots. *Conservation Biology,* **20,** 538-48.

Manel, S., Williams, H.C. & Ormerod, S.J. (2001) Evaluation presence-absence models in ecology: The need to account for prevalence. *Journal of Applied Ecology,* **38,** 921-31.

Marcer, A., Sáez, L., Molowny-Horas, R., Pons, X. & Pino, J. (2013) Using species distribution modeling to disentangle realized versus potential distributions for rare species conservation. *Biological Conservation,* **166,** 221-30.

McCallum, K.P., Guerin, G.R., Breed, M.F. & Lowe, A.J. (2014) Combining population genetics, species distribution modeling and field assessments to understand a species vulnerability to climate change. *Austral Ecology,* **39,**17-28.

McInerny, G.J., Turner, J.R.G., Wong, H.Y., Travis, J.M.J. & Benton, T.G. (2009) How range shifts induced by climate change affect neutral evolution. *Proceedings of the Royal Society B: Biological Sciences,* **276,** 1527-37.

McInerny, G.J. & Etienne, R.S. (2012) Pitch the niche—taking responsibility for the concepts we use in ecology and species distribution modeling. *Journal of Biogeography,* **39,** 2112-18.

Memmott, J., Craze, P.G., Waser, N.M. & Price, M.V. (2007) Global warming and the disruption of plant-pollinator interactions. *Ecology Letters,* **10,** 710-17.

Münzbergová, Z. & Herben, T. (2005) Seed, dispersal, microsite, habitat and recruitment limitation: identification of terms and concepts in studies of limitations. *Oecologia*, **145,** 1-8.

Nathan, R. & Muller-Landau, H. (2000) Spatial patterns of seed dispersal, their determinants and consequences for recruitment. *Trends in Ecology and Evolution*, **15**, 278-85.

Newman, D. & Pilson, D. (1997) Increased probability of extinction due to decreased genetic effective population size: experimental populations of *Clarkia puchella*. *Evolution*, **51**, 354-62.

Ohlemüller, R., Anderson, B.J., Araújo, M.B., Butchart, S.H.M., Kudrna, O., Ridgely, R.S. & Thomas, C.D. (2008) The coincidence of climatic and species rarity: high risk to small range species from climate change. *Biology Letters* 4: 568-72.

Petit, R.J., Aguinagalde, I., de Beaulieu, J.-L., Bittkau, C., Brewer, S., Cheddadi, R., Ennos, R., Fineschi, S., Grivet, D., Lascoux, M., Mohanty, A., Müller-Starck, G., Demesure-Musch, B., Palmé, A., Martín, J.P., Rendell, S. & Vendramin, G.G. (2003) Glacial refugia: hotspots but not melting pots of biodiversity. *Science*, **300**, 1563-5.

Phillips, S.J., Dudík, M. & Schapire, R.E. (2004) A maximum entropy approach to species distribution modeling. *Proceedings of the 21st International Conference on Machine Learning*. pp83.

Phillips, S.J., Anderson, R.P. & Schapire, R.E. (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling,* **190,** 231-59.

Phillips, S.J., Dudík, M., Elith, J., Graham, C.H., Lehmann, A., Leathwick, J. & Ferrier, S. (2009) Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. *Ecological Applications,* **19,** 181-97.

Pulliam, H.R. (2000) On the relationship between niche and distribution. Ecology Letters, **3,** 349-61.

Quintero, I. & Wiens, J.J. (2013) Rates of projected climate change dramatically exceed past rates of climatic niche evolution among vertebrate species. *Ecology Letters*, **16**, 1095-1103.

R Development Core Team. (2009) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org

Reddy, S. & Dávalos, L.M. (2003) Geographical sampling bias and its implications for conservation priorities in Africa. *Journal of Biogeography,* **30,**1719-27.

Ritchie, J.C. & MacDonald, G.M. (1986) The patterns of post-glacial spread of white spruce. *Journal of Biogeography*, **13**, 527-40.

Rull, V. (2009) Microrefugia. *Journal of Biogeography*, **36**, 481-4.

Sax, D.F., Early, R. & Bellemare, J. (2013) Niche syndromes, species extinction risks, and management under climate change. *Trends in Ecology and Evolution,* **28,** 517-23.

Schwartz, M.W. (2012) Using niche models with climate projections to inform conservation management decisions. *Biological Conservation*, **155**, 149-156.

Shoo, L.P., Hoffman, A.A., Garnett, S., Pressey, R.L., Williams, Y.M., Taylor, M., Falconi, L., Yates, C.J., Scott, J.K., Algador, D. & Williams, S.E. (2013) Making decisions to conserve species under climate change. *Climatic Change,* **119,** 239-46.

Soberón, J. (2007) Grinnellian and Eltonian niches and geographic distributions of species. *Ecology Letters,* **10**, 1115-23.

Swets, J.A. (1988) Measuring the accuracy of diagnostic systems. *Science,* **240,** 1285-93.

Thomas, C.D., Cameron, A., Green, R.E., Bakkenes, M., Beaumont, L.J., Collingham, Y.C., Erasmus, B.F.N., de Siqueira, M.F., Grainger, A., Hannah, L., Hughes, L., Huntley, B., van Jaarsveld, A.S., Midgley, G.F., Miles, L., Ortega Huerta, M.A., Peterson, A.T., Phillips, O.L. & Williams, S.E. (2004) Extinction risk from climate change. *Nature,* **427,** 145-8.

Thomas, C.D. (2011) Translocation of species, climate change, and the end of trying to recreate past ecological communities. *Trends in Ecology and Evolution,* **26,** 216-21.

U.S. Global Change Research Program (USGCRP). (2009) Global climate change impacts in the United States. Thomas R. Karl, Jerry M. Melillo, and Thomas C. Peterson (eds.) Cambridge University Press. (http://www.globalchange.gov/usimpacts)

Vandermeer, J.H. (1972) Niche theory. *Annual Review of Ecology and Systematics*, **3**, 107-32.

Table 2.1. Study species and their conservation status (state and/or federal), lifespan (annual or perennial), habitat description, levels of within-population genetic diversity (He) if known, and the number of population records used in this study.

Species (Family)	Conservation Status*	Lifespan**	Habitat Specifics	Within-pop He (#pops)	No. of Records
Croomia pauciflora (Stemonaceae)	End: FL Thr: GA	Per.	Loess-derived soils, low-altitude temperate deciduous forests	n.d.	55
Pachysandra procumbens (Buxaceae)	End: FL, IN	Per.	Woodlands; in moist, acidic, organic, well-drained soil	n.d.	94
Panax quinquefolius (Araliaceae)	End: ME, RI Thr: MI, NH S.C.: CN, MA, NC, TN	Per.	Understory of deciduous forests; calcium-rich soils	0.110 (32)[1] 0.071 (16)[2]	780
Polygonella basiramia (Polygonaceae)	End: FL, **U.S.**	Ann.	Sandhill scrub, Lake Wales Ridge only, loose sandy soils	0.0845 (3)[3]	137
Polygonella macrophylla (Polygonaceae)	Thr: FL	Per.	Sand pine-oak scrub ridges (coastal habitat)	0.1141 (5)[3]	179
Polygonella myriophylla (Polygonaceae)	End: FL, **U.S.**	Per.	Florida Rosemary Scrub, loose sandy soils	0.1666 (4)[3]	120
Silene polypetala (Caryophyllaceae)	End: FL, GA, **U.S.**	Per.	Deciduous hardwood forests and rich wooded ravines	0.262 (20)[4]	40
Trillium lancifolium (Melanthiaceae)	End: FL, TN	Per.	Heavy shade or thin, open woods or brushy thickets; alluvial soils	n.d.	57
Waldsteinia lobata (Rosaceae)	Thr: GA	Per.	Stream banks; damp, acidic soil	n.d.	80

*End = endangered, Thr = Threatened, S.C. = Special Concern
"**U.S.**" indicates federal status of the listing
**Per. = perennial, Ann. = annual
[1]Grubbs HJ and MA Case. 2004. Conservation Genetics 5:13-23.
[2]Cruse-Sanders JM, JL Hamrick, and JA Ahumada. 2005. Biodiversity and Conservation 14: 493-504.
[3]Lewis PO and DJ Crawford. 1995. American Journal of Botany 82: 141-9.
[4]Hamrick JL and P Pattavina. Unpublished data.

Table 2.2. Results of Maxent runs. Envir. predictor = environmental predictor.

Species	AUC	1º Envir. Predictor	Value	2º Envir. Predictor	Value
Croomia pauciflora	0.933	STATSGO*	Channery Silt Loam Loamy Sand Sandy Clay Loam Gravelly Sandy Loam	Precipitation, Driest Quarter	265 mm
Pachysandra procumbens	0.887	Precipitation, Wettest Quarter	430 mm	STATSGO	Clay Loam Sandy Loam Silt Loam
Panax quinquefolius	0.904	Mean Temp, Wettest Quarter	3ºC	Altitude	1000 m
Polygonella basiramia	0.993	Temperature Seasonality**	Range of 42ºC	Precipitation, Driest Quarter	150 mm
Polygonella macrophylla	0.994	Temperature Seasonality	Range of 65ºC	Precipitation, Wettest Quarter	550 mm
Polygonella myriophylla	0.994	Temperature Seasonality	Range of 42ºC	Precipitation, Driest Quarter	150 mm
Silene polypetala	0.979	STATSGO	Sandy Loam Loamy Sand	Temperature Seasonality	Range of 67ºC
Trillium lancifolium	0.928	STATSGO	Clay Loam Cobbly Loam Gravelly Sandy Loam Loamy Sand	Mean Temperature, Wettest Quarter	7ºC, 27ºC
Waldsteinia lobata	0.986	Mean Temperature, Wettest Quarter	7ºC	Altitude	400 m

*For categorical data (STATSGO only), it is possible for more than one variable value (soil type in this case) to have a high correlation with presence. In these cases, we reported values with a greater than 50% probability of presence value.

Temperature Seasonality = 100*(standard deviation of monthly mean temperature); annual range in temperature

Table 2.3. Total amounts of current suitable habitat under conservative and liberal estimates, and changes under 2050 and 2080 climate predictions.

	Current Suitable Area (km^2)		% Change 2050*		% Change 2080*	
Species	**Conservative**	**Liberal**	**Conservative**	**Liberal**	**Conservative**	**Liberal**
Croomia pauciflora	135,019.07	329,021.50	-85.6	-70.2	-94.1	-83.6
Pachysandra procumbens	347,968.27	682,920.95	-82.2	-47.0	-97.6	-69.7
Panax quinquefolius	282,246.87	878,622.77	-74.8	-45.2	-89.7	-73.0
Polygonella basiramia	3,019.13	7,976.84	-95.3	-31.6	-99.8	-49.8
Polygonella macrophylla	2,498.83	12,429.95	+8.4	+11.1	-79.1	-37.5
Polygonella myriophylla	4,181.25	10,733.52	-15.5	+82.6	+48.1	+116.3
Silene polypetala	34,364.89	43,004.12	+11.2	+4.8	+24.1	+15.2
Trillium lancifolium	147,809.55	454,280.59	+58.7	-9.7	+27.7	-32.2
Waldsteinia lobata	32,346.70	116,263.09	-100.0	-95.9	-100.0	-100.0

*% Change = change in amount of suitable habitat (50-100% suitability) from current to 2050 or 2080
Sign indicates whether change is expected to be a gain (+) or a loss (-).

Table 2.4. Total amounts of lost, retained, and new habitat for current vs. 2050 predictions and current vs. 2080 predictions. IUCN refers to suggested IUCN status upgrade based on the amount of area combined from "Retained" and "New" estimates under each climate scenario. The IUCN categories listed here are Not Evaluated ("NE"), Vulnerable ("VU"), Endangered ("EN"), and Critically Endangered ("CE"). Categorizations are based on IUCN red list guidelines for criterion B1 found in Table 2.1 (IUCN 2014).

	Current – 2050 (km^2)				Current – 2080 (km^2)			
Species	**Lost**	**Retained**	**New**	**IUCN**	**Lost**	**Retained**	**New**	**IUCN**
Croomia pauciflora	253,892	74,538	23,225	NE	301,609	26,821	27,057	NE
Pachysandra procumbens	323,503	359,082	2,632	NE	476,811	205,774	1,002	NE
Panax quinquefolius	417,323	461,300	19,761	NE	661,057	217,566	19,652	NE
Polygonella basiramia	6,517	1,456	3,996	EN	7,041	932	3,063	EN
Polygonella macrophylla	3,136	8,301	5,340	VU	6,343	5,094	2,592	EN
Polygonella myriophylla	1,143	9,586	10,010	VU	693	10,035	13,177	VU
Silene polypetala	35,162	7,8412	37,128	NE	32,395	10,609	38,701	NE
Trillium lancifolium	300,545	152,769	256,580	NE	335,548	117,766	189,505	NE
Waldsteinia lobata	112,199	4,065	754	EN	116,263	0	0	CE

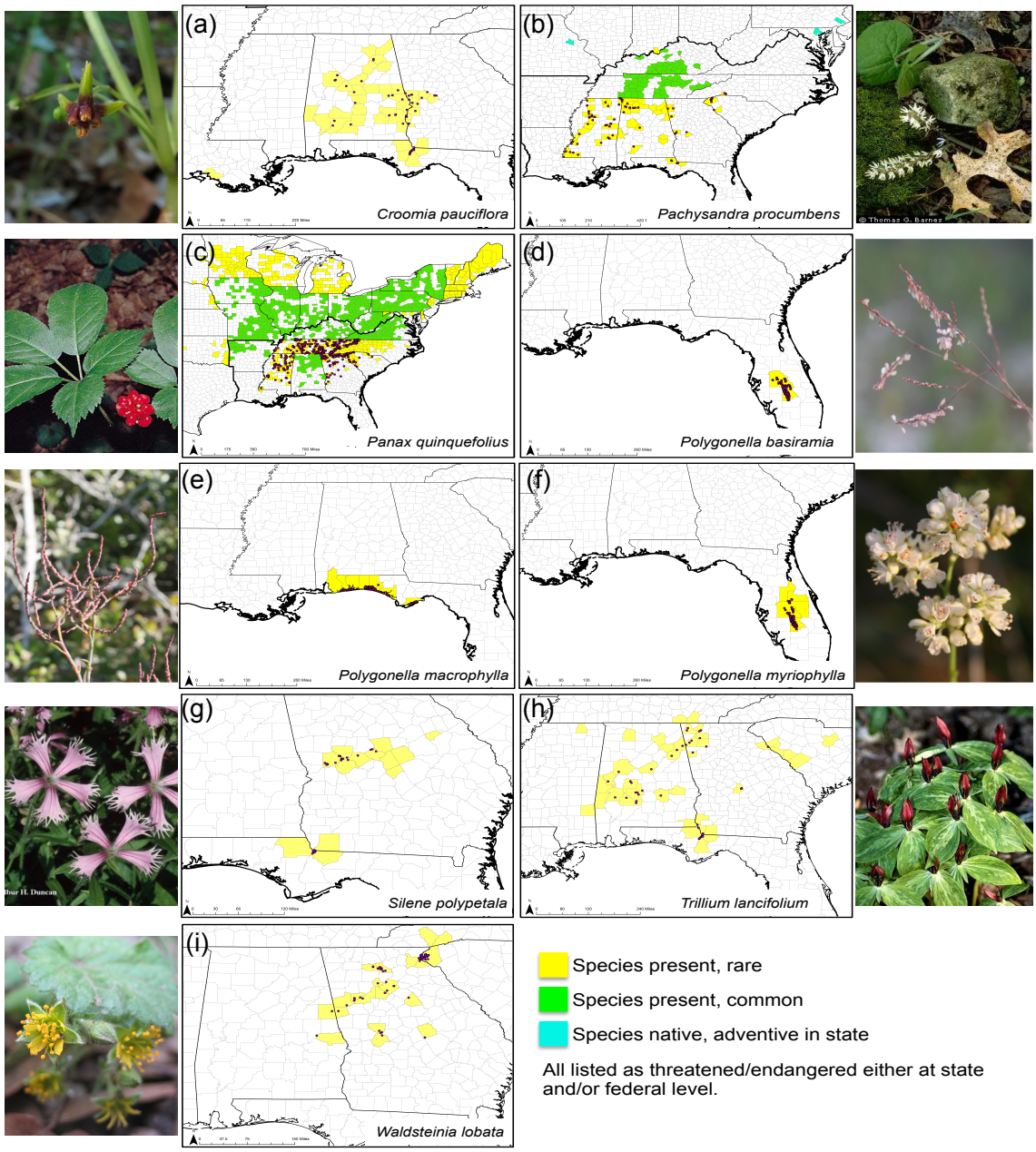


Figure 2.1. Geographic range maps and photos of the study species. Geographic range data were obtained from the Biota of North America Program (www.bonap.org).

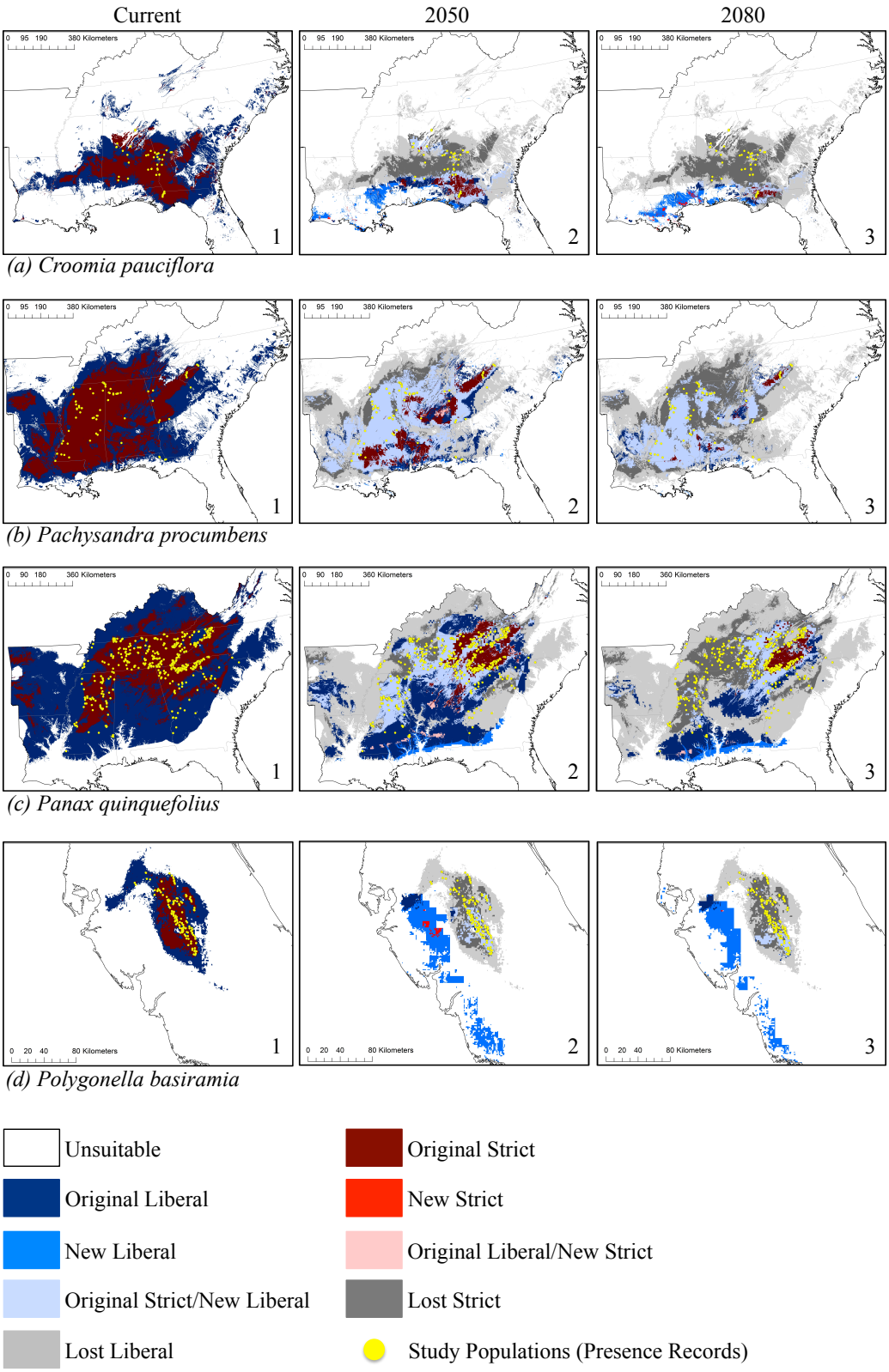
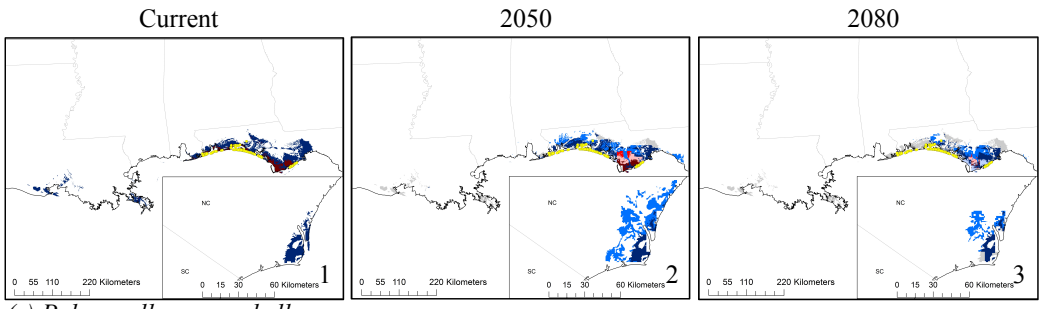
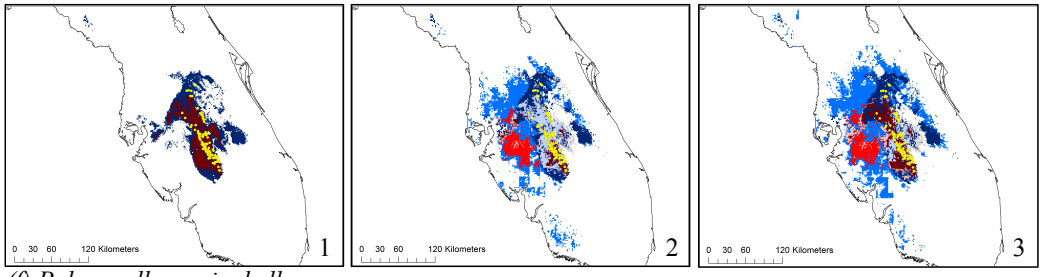


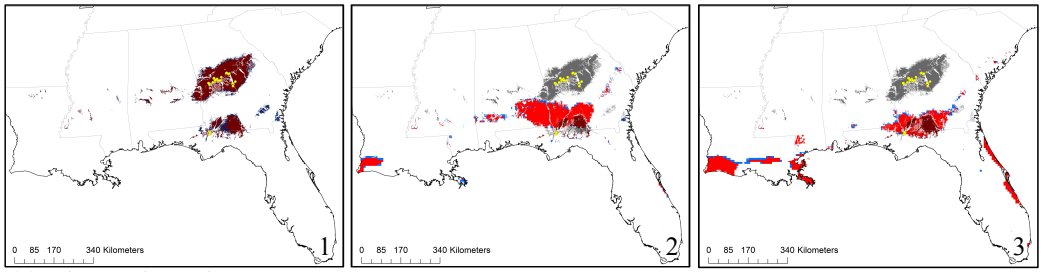
Figure 2.2. Predicted suitable habitat under current, 2050, and 2080 climate estimates. Colors correspond to area that is suitable under liberal (blues) and conservative (reds) thresholds, area that is retained as suitable when comparing between current and future climate scenarios (pale blues and reds), new suitable area in 2050 or 2080 that is outside of suitable area under current climate (bright blues and reds), or area that is lost when contrasting between current and future climates.



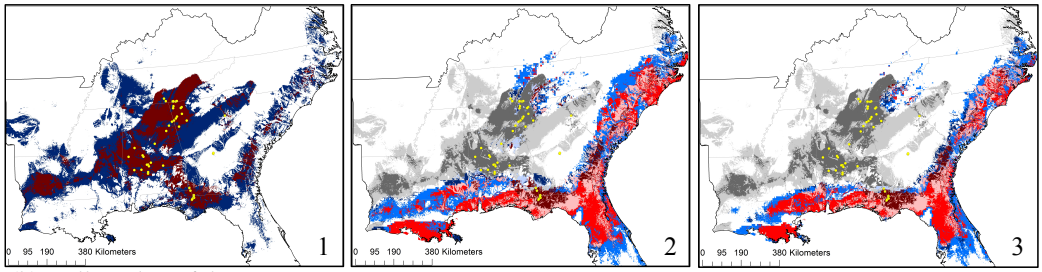
(e) Polygonella macrophylla



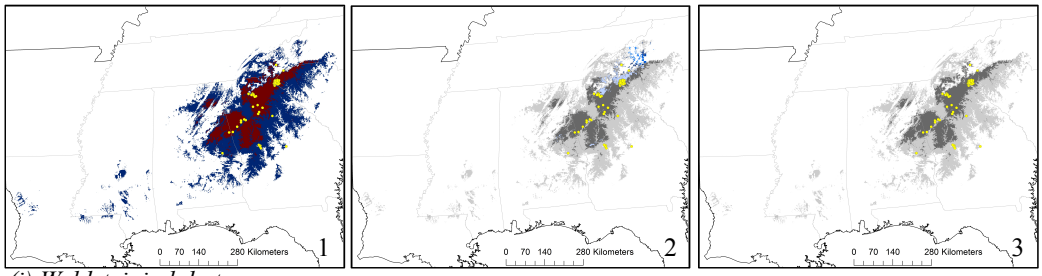
(f) Polygonella myriophylla



(g) Silene polypetala



(h) Trillium lancifolium



(i) Waldsteinia lobata

Figure 2.2 continued.

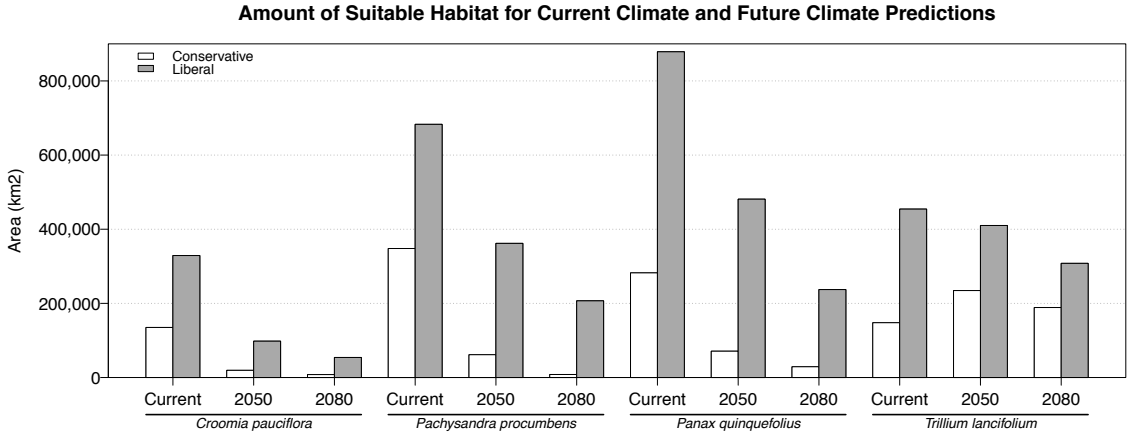


Figure 2.3. Amount of suitable habitat (km^2) under current and future climate scenarios for conservative (white bars) and liberal (gray bars) thresholds. Species shown here are those for which estimates of suitable habitat were highest in this study.

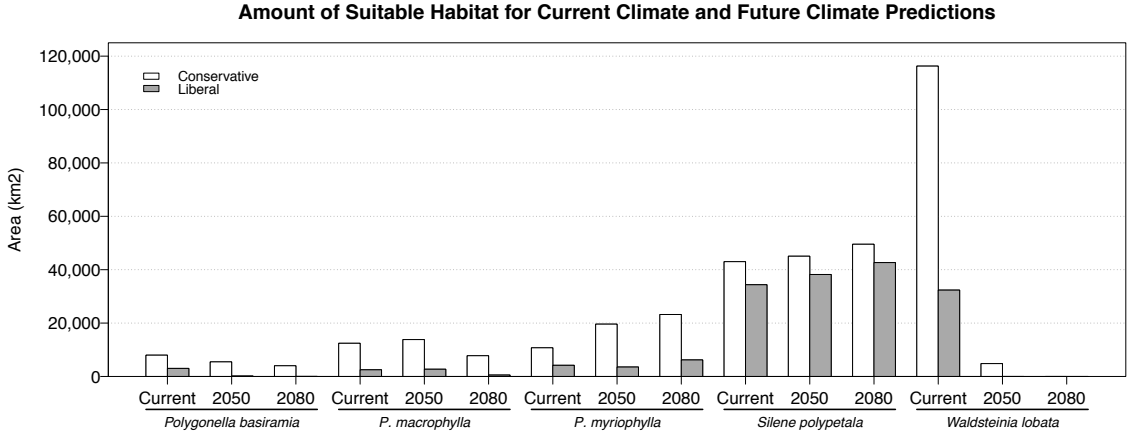


Figure 2.4. Amount of suitable habitat (km^2) under current and future climate scenarios for conservative (white bars) and liberal (gray bars) thresholds. Species shown here are those for which estimates of suitable habitat were lowest in this study.

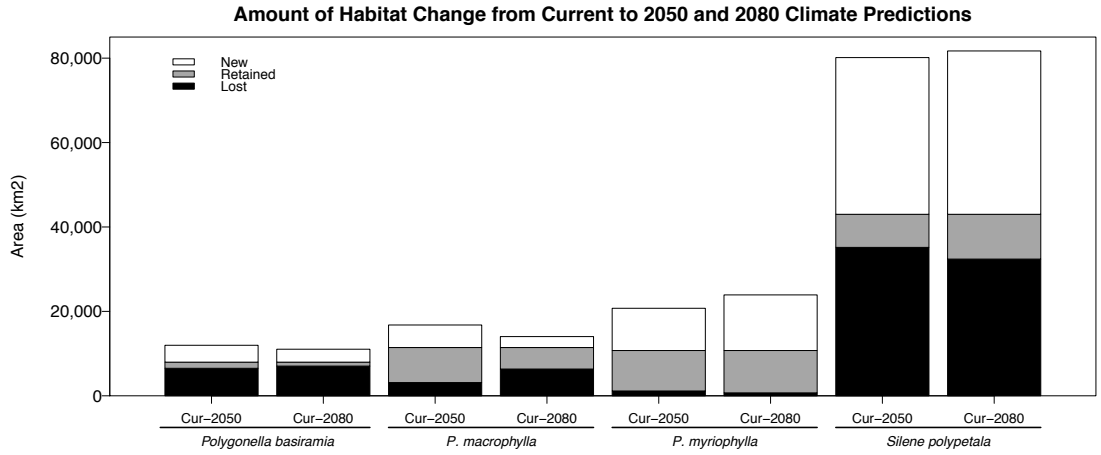


Figure 2.5. Amount of habitat change (km^2) for current-versus-2050 ("Cur-2050") and current-versus-2080 ("Cur-2080") climate scenarios for the species in this study that had the least amount of suitable habitat predicted overall. Colors correspond to area that is new (white), retained (gray) or lost (black) when suitability maps were overlaid for each comparison.

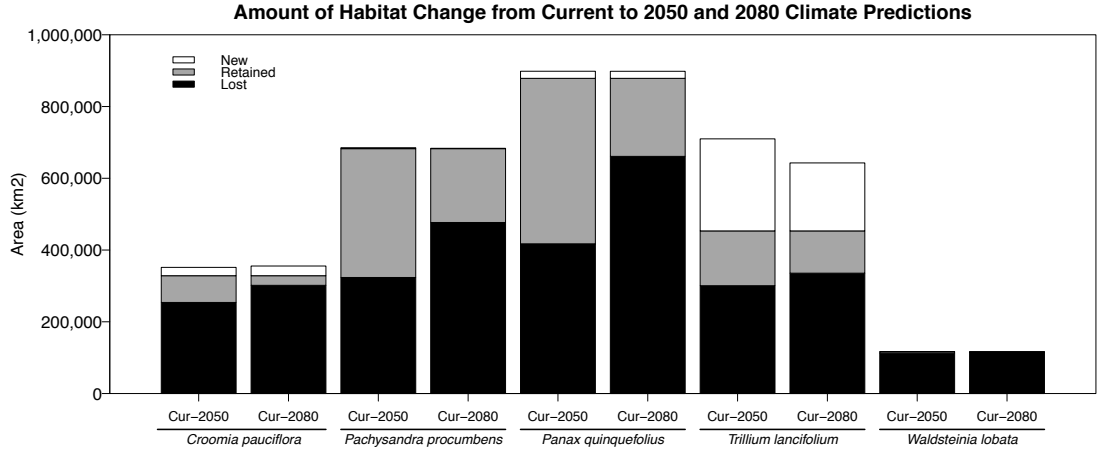


Figure 2.6. Amount of habitat change (km^2) for current-versus-2050 ("Cur-2050") and current-versus-2080 ("Cur-2080") climate scenarios for the species in this study that had the greatest amount of suitable habitat predicted overall. Colors correspond to area that is new (white), retained (gray) or lost (black) when suitability maps were overlaid for each comparison.

CHAPTER THREE

THE ROLE OF LOCAL ADAPTATION AND NICHE CONSTRAINTS IN THE RANGE LIMITS OF A NARROW AND WIDESPREAD PAIR OF SISTER SPECIES IN *POLYGONELLA* (POLYGONACEAE)

Louisa G.C. Staton and Shu-Mei Chang

To be submitted to Ecology Letters

Abstract

Determining what factors shape the geographic range limits of plant species is important for understanding how certain taxa will respond to external pressures such as urban development and anthropogenic climate change. This project aims to investigate levels of local adaptation and identify potential factors associated with geographic range limits in two species of the plant genus *Polygonella.* Species of this genus are mainly confined to the southeastern United States. This genus is a suitable study system for investigating questions related to range limits due to its geographic distribution (nine species are narrowly distributed, while the other two are widely distributed). We performed reciprocal transplant experiments using seeds from seven populations of a widespread *Polygonella* species (*P. americana*) and seven populations of a narrowly distributed congener, *P. fimbriata*. Our results indicate that signals of local adaptation are better detectable in populations of the widespread species, *P. americana*, as compared to populations of *P. fimbriata.* Additionally, comparisons of germination rates between these two species reveal that individuals of *P. americana* demonstrate higher germination rates that are consistent with other common and widely distributed species. Finally, populations of *P. fimbriata* appear to be limited in their ranges due more to dispersal and evolutionary genetic constraints as opposed to niche constraints. This work contributes to understanding the most effective ways to determine what factors are limiting range expansion in narrowly distributed plant species. Additionally, it is informative with respect to conservation work in this and other geographically limited species, and points to the need for completely reciprocal transplant experiments that incorporate population genetic information.

Introduction

An important question in ecology is why species are distributed in certain patterns and what shapes these distributions. The topic of geographic range limits has historically been of interest to biologists (Darwin, 1859; MacArthur & Wilson, 1967; Brown, 1984; Hubbell, 2001), and yet the causative factors driving distributional patterns are still poorly understood (Gaston, 2009; Hargreaves *et al.*, 2014). Why are some taxa severely limited in their range while closely related sister species are geographically widespread? Are these narrow species trapped as the result of a lack of escape opportunities, or have populations become so adapted to their current habitats that relocation would be detrimental to their survival? Answering these questions will provide crucial information for the conservation of narrowly distributed, at-risk species.

Conceptually, a species is expected to occupy all of the geographic areas suitable for its persistence, or its habitat—the geographical locations coinciding with its ecological niche, defined as the set of biotic and abiotic conditions necessary for that species to persist (Hutchinson, 1957; MacArthur, 1972). In most species, a portion of the suitable habitat for a species remains unoccupied, suggesting disequilibrium between the geographic range limits and the fundamental niche (Holt, 2003; Sax *et al.*, 2013). This unoccupied suitable habitat, known as the tolerance distribution (Sax *et al.*, 2013), generally occurs geographically just beyond the edges of the realized distribution. Why does the tolerance distribution remain unoccupied even though it comprises suitable habitat? One possibility is that populations on the edge of the range are actually expanding, and our temporal snapshot of the range is simply too early in the evolution of these range boundaries to capture this expansion (Geber, 2008). Another possibility

points towards evolutionary genetic constraints. Populations on the edge of the range, tend to be smaller and more isolated than populations at the center of the distribution (Brown, 1984). Small edge populations would more likely be at risk for inbreeding and stronger drift (Kruckeberg & Rabinowitz, 1985), as well as center-to-edge gene flow of maladapted alleles (Gomulkiewicz *et al*., 1999; Bridle & Vines, 2007). If any or all of these conditions are true, then dispersal propagules from these edge populations into the tolerance distribution are more likely to have severely decreased genetic variation and/or maladapted alleles. As a result, they will have a low chance of establishing self-sustaining populations (their intrinsic population growth rate, λ, will be less than one; Holt & Gomulkiewicz, 1997). Finally, dispersal is no small task (Nathan & Muller-Landau, 2000). Limits to dispersal could be the primary cause for the persistence of species' boundaries (Holt *et al*., 2005; Geber, 2008), and dispersal could be the sole explanation for the persistence of populations outside of the tolerance distribution (Hargreaves *et al.*, 2014). It is important to note that the above mechanisms are not mutually exclusive, and often occur in tandem (Geber, 2008).

Though less common, the geographic range of the species may in fact be at equilibrium with the suitable habitat. In other words, the species' distribution boundary and the ecological niche boundary are the same. The question then becomes what generates the end of a suitable habitat? Even within the habitat considered suitable, the ecological niche requirements for a species may not be uniformly distributed across space. As a result, habitat quality within a species' range may exist as a gradient such that the quality of the habitat is highest in the center of the geographic distribution where populations of the species are most abundant and gradually degrades towards the edge of

the species distribution (Brown, 1984). In this scenario, dispersal beyond the geographic range would not serve any utility in expansion, because dispersal propagules would be located beyond the ecological limits necessary for their survival. Thus, when investigating explanations for what may be limiting the geographical expansion of a species' range, it is important to be able to distinguish between whether or not a species realized distribution and the geographic distribution of its ecological niche are in equilibrium.

The reciprocal transplant experiment is the most powerful approach for understanding whether or not the realized distribution and distribution of suitable habitat for a species overlap. It can also help identify which factors may be responsible for disequilibrium between the two if it exists (Gaston, 2003; Hargreaves *et al.*, 2014). These experiments involve transplanting individuals out of their home environments into "away" transplant sites, as well as back into their own home environments. This is done across several transplant locations such that within a single transplant site, a collection of both home and away individuals may be evaluated for their performance. The observation that away individuals have low overall fitness and home individuals perform better than away individuals within their home sites is evidence of local adaptation (Kawecki & Ebert, 2004). In addition to individual responses within sites, population-level responses in transplant experiments can allow us to gain a better understanding of the likelihood of range expansion, as has been demonstrated in responses to climate change (Davis & Shaw, 2001). By using populations that are distributed across the entire range of a species in transplant experiments, we can delineate what factors are contributing to range/niche disequilibrium (Hargreaves *et al*., 2014). Because edge

populations are those integral to whether or not expansion can occur, specifically transplanting individuals from these populations to transplant sites that are beyond the current range of the distribution and evaluating their performance can elucidate whether or not niche constraints are driving range limitations (Hargreaves *et al.*, 2014). Additionally, the performance of individuals from edge populations transplanted into sites in the interior of their range can provide us with an estimation of adaptation and genetic constraints.

Based on the various models of range limits (see Holt, 2003; Moeller *et al.*, 2011), we can make predictions for expected outcomes of reciprocal transplant experiments. For example, if populations on the range margins are serving as sink populations such that gene flow from interior populations is the only sustaining force (Holt & Gomulkiewicz, 1997), then individuals of these populations should have equivalent performance (similar reaction norms) to individuals from interior source populations when transplanted at the same interior site. In the absence of frequent gene flow, interior and marginal populations have the greater potential to adaptively diverge via natural selection, and the slope and direction of their reaction norms should be significantly different (Alleaume-Benharira *et al*., 2006; Murren *et al.*, 2014). Extreme levels of divergence, equivalent to habitat specialization, would result in significant decreases in performance at away transplant sites, and reaction norms should have very steep slopes. Thus, the degree of local adaptation, which can be imagined as a sliding scale from not locally adapted to habitat specialization, can be inferred from the slope of the reaction norms of populations. Subsequently, this can be an indicator of the potential success of dispersal propagules. We expect that the higher the degree of local adaptation (the steeper the reaction norm),

the lower the likelihood of dispersal propagules to successfully establish and colonize in areas beyond their current geographical boundaries.

Comparative studies using congeneric species in reciprocal transplant experiments, particularly ones with drastically different geographic ranges, can potentially allow for even more powerful inferences regarding geographic range limits due to their shared evolutionary history. For example, Lavergne *et al.* (2004) measured ecological and physiological traits for 20 congeneric pairs (across 17 families) of narrow endemic and widespread species to assess whether there were significant differences to explain their disparity in geographic range size. They found that narrow endemic plant species had unique ecological specificities, and as a result some physiological differences, but no major competitive disadvantages when compared to their widespread congeners. Using such comparative approaches can allow us to more quickly zero in on what factors are important for the geographic range distribution of these species (Debussche & Thompson, 2003; Lavergne *et al.*, 2004; Maliakal-Witt *et al.*, 2005).

In order to investigate what factors underlie the geographic range limits of the narrow endemic plant species, *Polygonella fimbriata*, we performed reciprocal transplant studies utilizing populations of this plant in conjunction with populations from its widespread congener, *P. americana*. Transplants were performed within the contemporary distribution of *P. americana*, and within and beyond the contemporary distribution of *P. fimbriata*. The majority of *P. fimbriata*'s range is sympatric with that of *P. americana,* but a small portion in the southeast of Georgia is allopatric (Figure 1a). In this study, we address the following questions regarding the range limits of these species. First, are populations of *Polygonella* species locally adapted to their home sites? If so, we

would expect to see these populations having higher fitness at their home sites as opposed to at away transplant sites. Second, are range limits and niche limits in *P. fimbriata* in disequilibrium? If so, this would be shown by edge populations, which are presumably better adapted to conditions further from the interior, out-performing individuals from interior populations when both are moved beyond the range edge. Finally, are niche constraints driving range limits in *P. fimbriata*? If this is the case, individuals from edge populations should experience a decrease in fitness when moved beyond the range (Hargreaves *et al.,* 2014). Alternatively, if individuals from interior or edge populations experience no decrease in fitness when moved beyond the range, we would infer that range limits are better explained by dispersal limitation, metapopulation dynamics, population genetic constraints, or some combination thereof (Hargreaves *et al.*, 2014). When narrow and patchy distributions of species complicate the classification of populations as edge or interior, indirect tests of niche constraints are still possible. If populations are transplanted to sites beyond the current distribution, their failure or decreased fitness in these sites would not be informative for niche constraints, but their success in these sites would be evidence against it. **We hypothesize that populations of the widespread species, *P. americana*, will fall on the lower end of the adaptation spectrum and show trends more consistent with common species, while populations of *P. fimbriata* will have greater degrees of local adaptation and potentially be habitat specialists**. Additionally, **we hypothesize that narrow endemic species are limited in their ranges by niche constraints,** and individuals from edge populations will experience marked decreases in fitness when moved beyond the range edge**.**

Study System

Polygonella (Polygonaceae) is a small genus comprising eleven species, all of which are endemic to North America. Seven of these are geographically narrowly restricted to the southeastern United States (found in only 1-3 states), two are considered geographically intermediate in their ranges (spanning 6-7 states), and the remaining two are geographically widespread (found in 10 or more states). Species in this genus are hermaphroditic, dioecious, gynodioecious, or gynomonoecious (Horton, 1963; Hong & Smets, 2004), and there are both perennial and annual taxa in the genus. *Polygonella* is distinct from other members of the Polygonaceae in that branches are adnate to the stem, causing them to appear internodal, and ocrea are present. Previous work in this genus has demonstrated that, based on allozyme analyses, the two geographically widespread species (*P. americana* and *P. articulata*), have decreased within-population genetic variation when compared to the rest of the species in this genus (Lewis & Crawford, 1995). Additionally, a morphological phylogeny exists for this genus, but it has yet to be fully tested using molecular data (Lewis, 1991). A molecular phylogeny for the Polygoneae including eight species of *Polygonella* was constructed (Schuster *et al*., 2011), and differs significantly from the morphological phylogeny Lewis & Crawford (1995).

This study utilizes seeds from populations of two species of *Polygonella*: *P. fimbriata* and *P. americana* (Figure 1). *P. fimbriata* is an annual plant that germinates in March, flowers from July to October and dies around December, with gynomonoecious individuals (plants have both hermaphroditic and female flowers on the same plant). It is narrowly distributed in the southern half of the state of Georgia with some putative

populations in the panhandle of Florida. *P. fimbriata* is easily distinguished from *P. americana* based on its ciliated ocrea and linear leaves (*P. americana* has spatulate leaves and its ocrea are not ciliated). Alternatively, *P. americana* is a hermaphroditic perennial species that flowers from June to October (Figure 2). It is geographically widespread throughout the southern and southwestern states of the United States.

These two species were chosen for this study because they represent the varying range distributions seen in the genus and because of their availability. Four of the narrowly endemic species in *Polygonella* are listed as threatened or rare on either the state or federal level (or both). Among the other three narrowly distributed species, populations of *P. fimbriata* are the most easily accessible, and overlap in part of its range with populations of *P. americana*, making it possible to establish reciprocal transplant gardens within the range of both species.

Materials and Methods

Seed Collections

Seeds from seven populations each of *P. americana* and *P. fimbriata* were collected during the fall of 2011 and 2012. Populations for seed collections were located using herbarium records from the University of Georgia and other universities in the southeast and through personal communications with local botanists. The locations of these populations are shown in Figure 1. Seeds were collected from a minimum of twenty maternal individuals per population, spaced at least one meter apart to avoid collecting from highly related individuals. The number of seeds collected from each maternal individual ranged from 20-1000 seeds, depending on the amount of material available.

Reciprocal Transplant Gardens

Three gardens were planted in December of 2011, one in Guntersville State Park in Guntersville, AL (henceforth, "AL"), one in Fall Line Sandhills Natural Area in Butler, GA (henceforth, "GA"), and one at the Savannah River Ecology Laboratory Conference Center in Aiken, SC (henceforth, "SC"; Figure 1). Individuals (seeds) from seven populations of each of our two species were reciprocally planted into one of these three 'home' and 'away' gardens, the classification of which depended upon the origin of the source population (see below). We will refer to these gardens as "sites," and the populations from which seeds were obtained for planting purposes will be referred to as "source populations." All transplant sites are within the current range of *P. americana* with natural population of *P. americana* nearby. Additionally, the SC transplant site is on the edge of the range margin for this species while the GA and AL sites are interior range for this species. Alternatively, the GA site is within the current range of *P. fimbriata*, but the AL and SC site are beyond its geographic range limits. Transplanted individuals were tracked and measured for two years in the case of *P. americana* (perennial), or two 1-year growing seasons in the case of *P. fimbriata* (annual). Within each site there were three subplots the first year, followed by the addition of a fourth subplot the second year for the additional *P. fimbriata* seeds that were planted in the second round. To keep track of seeds planted in the field and to distinguish between our experimental plants and any native germinants, we sank 6.5cm-diameter paper cups, with the bottoms removed and filled with local soil into the ground, and planted two to five seeds (depending on the availability) from a single maternal individual into each cup. We randomized the planting locations of populations and maternal individuals in each field subplot. For the first year

of this study, we planted 1500 seeds of *P. americana* in each of the three transplant sites along with 215 seeds of *P. fimbriata* in each site in Dec 2011. Fewer seeds of *P. fimbriata* were planted due to limited seed availability in 2011. However, we found a very low germination rate of *P. fimbriata* in the field, and as a result, 1900 more seeds of *P. fimbriata* populations (15 seeds per maternal individual per each of six populations) were planted at the garden sites in SC and GA in January 2013. Also, one of our 2011 populations (PF6) was unavailable for the recollection of seeds in 2012 so we replaced it with another population (PF7) that is no further in geographic distance from the GA garden site. Finally, our AL site was ultimately excluded from this study as mortality was 100% within the first year. All data collections presented here are based on values from our GA and SC sites only.

Population Clustering

To detect local adaptation, we classified each source population in this study as either home or away at a given site based on its proximity to that site. If a population is considered home in one site, then by default it is considered away for any other site. For *P. americana*, we classified three populations as home in the SC site (PAM1, PAM2, PAM3), two populations as home in the GA site (PAM26, PAM27; Table 1). The remaining two populations were considered away at all sites (PAM12, PAM24; Table 1). For *P. fimbriata*, four populations were classified as home at the GA site (PF1, PF2, PF6, PF7). The remaining three populations were considered away at all transplant sites (PF3, PF4, PF5; Table 2).

Because the identification of populations can be somewhat arbitrary at times (Berryman, 2002), we wanted a way to define our populations that would be biologically

relevant and go beyond simple field observations from collection locations. Based on microsatellite data from another study in this system, we have evidence that the populations utilized in this study formed two genetic clusters (Staton & Chang, unpublished; Chapter 4). Grouping populations by genetic similarity and analyzing the performance of these genetic clusters makes biological sense (Waples & Gaggiotti, 2006). Because we are interested in the transplant response of a biological unit, we can investigate statistical site by genetic cluster interactions. These interactions should be similar to genotype by environment interaction effects, and yield more statistical power in contrast to population by site interaction effects. For this reason, in addition to our field observation-defined populations and home/away clustering, we also grouped populations by their genetic similarity. We refer to these as genetic clusters, and they are defined as follows for each species. In *P. americana*, we grouped populations PAM1, PAM2, and PAM3 into cluster A, populations PAM26 and PAM27 into cluster B, and populations PAM12 and PAM24 into cluster C (Table 1). For *P. fimbriata*, populations PF1, PF2, PF6, PF7, and PF4 form cluster X, and populations PF3 and PF5 form cluster Y (Table 2).

Data Collection

We measured percent germination for each species as well as three vegetative traits for *P. americana* (PAM) and five for *P. fimbriata* (PF), eight reproductive traits for PAM and six for PF, and three fitness traits for both species (Table 3). Measurements were repeated for two growing seasons for the perennial PAM plants and for the two cohorts of annual PF planted in consecutive years. Vegetative characteristics were measured once every two weeks. Percent germination was calculated as the proportion of

seeds planted in a given cup that germinated. Plant area was calculated as the product of the vegetative width and height of the plant. Seedset refers to the total number of seeds divided by the number of flowers (will always be ≤ 1 because each flower produces one seed at most), and fecundity was measured by the total number of non-aborted seeds per plant. Additionally, growth rate was calculated from a bi-weekly measure of height taken on each individual using the formula GR = $(h_2-h_1)/(\Delta t)$ where h is the height at a given time point and Δt is the difference in time (days) between the two time points.

Some traits were only measured for one species and not the other for the following reasons. First, number of leaves was only measured for PF, because the number of leaves on PAM individuals rapidly exceeded a value that was easily measured. Second, longest branch length is equivalent to the plant height in PF and plant area was an irrelevant measure for PF—these plants grow mainly vertically and do not produce much, if any, horizontal growth. Finally, measures relating to fructescences were not measured in PF, because each of the branches in a PF individual becomes a flowering branch, and there is not a flowering/fruiting body separate from the vegetative structures.

Inflorescences in *Polygonella* are determinate with basipetal maturation (terminal flowers mature first and the floral maturation progresses towards the base of the inflorescence branch). Because visits to sites more frequent than biweekly were not possible and flowers do not all mature simultaneously, we were unable to collect every single flower following maturation. Thus, inflorescence branches that were open on the day of a visit were bagged with small organza swatches before departing from the site at the end of the day (late afternoon). This allows natural pollination prior to the application of a bag, and will also hold mature seeds until they are collected. We collected bagged

flowers/achenes and flowering branches and brought them back to the lab for measurement. After achenes have fallen from the inflorescence branch, the pedicel is left behind within the remaining ocrea. This makes it possible to count ocrea/pedicels on a given inflorescence branch to determine the number of flowers originally formed on the branch, even if some achenes were lost. We took ocrea count data in conjunction with inflorescence branch length measurements on a subset of inflorescences branches on all plants in order to create population-by-family estimates for the total number of flowers formed. These were based on the length measurements of all inflorescence branches (regression methods below). At the end of the study, we harvested all plant materials from the field.

Data Analysis

Each of the two transplant sites (GA and SC) were chosen based on their locations within the ranges of each species, and on their similarities in soil type and. To evaluate potential climatological differences that were not immediately obvious at the beginning of this study, we performed multivariate analysis of variance (MANOVA) followed by subsequent univariate ANOVAs on climatological data for each site as well as source population. We calculated mean values from monthly data to generate the following for each site and source: mean precipitation, mean precipitation in the wettest quarter, mean precipitation in the driest quarter, mean temperature, mean temperature in the hottest quarter, mean temperature in the coldest quarter, maximum temperature in the hottest quarter, and minimum temperature in the coldest quarter. These are meant to capture the temperature extremes and norms for each site. These values were calculated for the year 2012 (all twelve months) and for the year 2013 from the months of January to November

(when the plants were removed from the field). All raw data were obtained from the PRISM Climate Group (PRISM, 2013). We performed MANOVAs and ANOVAs on these climate data in R v2.10.1 (R Core Team, 2009) with year and site as predictors.

Because we were unable to count all individual flowers, we estimated the total number of flowers each plant produced by using the inflorescence length measurements and corresponding flower count data to parameterize a predictive model (a total of 326 observations for PAM and 225 observations for PF, using R v2.10.1 (R Core Team, 2009)). We used transplant site, population, family, length, block (for PAM only), and four interaction terms (site*length, population*length, site*population, family*length) as explanatory variables, manually dropping nonsignificant factors from the model in a step-wise manner. For PAM, the best fit model included length, site, population, family, block, and two interaction effects (adjusted R^2=0.9331, $F_{8,317}$=567.9, p<<0.0001). For PF, the best fit model included length, site, population, family and all four interaction terms (adjusted R^2=0.9262, $F_{8,216}$=352.4, p<<0.0001). We used the parameterized model to estimate the flower production for all individuals that produce any flowers. Additionally, we calculated seedset for each plant as the ratio of seeds to flowers and the ratio of normal or aborted seeds to the total number of seeds. From these calculations, we were then able to estimate the total number of seeds per plant and the number of normal or aborted seeds per plant.

We analyzed data measurements for germination, vegetative growth, reproductive measures and fitness using the generalized linear mixed model procedure (PROC GLIMMIX) in SAS v9.3 (SAS, 2011). All data were checked for normality, and non-normal data were log-transformed prior to analysis to assure normality and

homoscedasticity. Because we measured multiple traits from each life history stage and these might be correlated, we first performed MANOVAs for traits grouped by vegetative and reproductive stages using PROC GLM in SAS. If any effects from the MANOVAs were significant, we ran univariate analyses for the full models that are described next. For our "main model," we held site, region (geographical region (TX, AL, GA, SC) for each source population; PAM only), and population (nested within region for PAM) as fixed effects, and treated family as a random effect. Home vs. away effects were tested using orthogonal contrasts. Additionally, we included a population*site interaction effect, and treated year (for germination data) and plot (within transplant sites) as block effects. We excluded block terms whenever they came out as statistically insignificant in our models. To more thoroughly investigate genotype by environment (GxE) effects, we ran separate models using genetic clusters (GC). For these, we treated site, population (nested with GC), and GC as fixed effects. Finally, GC by site interactions were included, which if significant would support a significant GxE interaction. We will refer to this model as our "GC model."

Results

Site Similarity Analysis

Our GA and SC transplant sites did not significantly differ from each other for any of the climatological variables we tested. We did, however, detect a significant difference in precipitation between the two study years, such that 2013 was a much wetter and slightly colder year for all locations measured. Mean precipitation in 2013 was 220.9 mm as compared to 122.8 mm in 2012 ($F_{1,30}$=104.29, p<0.001), and mean temperature in the coldest quarter in 2013 was 0.8ºC cooler than that of 2012 (10.1ºC and

10.9ºC, respectively; $F_{1,30}$=4.96, p=0.034). Additionally, the hottest temperatures seen during the hottest quarter were significantly higher in 2012 (32.1ºC) than in 2013 (31.1ºC; $F_{1,30}$=9.109, p=0.005). Despite these differences, study year did not have a significant effect on germination for *P. fimbriata*, the only species that was planted in the two successive years.

Germination

Out of 3006 seeds total for *P. americana* (1502 seeds in GA, 1504 seeds in SC), average percent germination rate over the two-year period was 13.79 ± 0.94%. We found a significant effect on germination from transplant site (GA=12.07 ± 0.02%, SC=16.65 ± 0.02%; $F_{1,659}$=4.70, p=0.031), source population ($F_{3,99}$=2.78, p=0.045), and source region ($F_{3,99}$=14.75, p<0.001), as well as a significant interaction effect among population and site ($F_{6,659}$=2.18, p=0.044). Post-hoc tests for population revealed varying differences among populations (Figure 3), but generally population PAM12 (TX) had significantly higher percent germination when compared to the three populations from SC (PAM1, PAM2, and PAM3) and the population from AL (PAM24; Figure 3). Additionally, post-hoc tests for region effects revealed that the population from TX had significantly higher germination (25.3 ± 3.3%) when compared to populations from AL (2.5 ± 4.6%; t_{99}=-4.02, p<0.001) and SC (9.1 ± 1.7%; t_{99}=-4.36, p<0.001), but not in contrast with populations from GA (20.5 ± 1.4%). Also, populations from GA had significantly higher germination than populations from AL (t_{99}=-3.74, p<0.001) and SC (t_{99}=5.27, p<0.001). Orthogonal contrasts revealed a significant home-site advantage for germination (t_{659}=2.41, p=0.016), specifically within the SC site (t_{659}=3.93, p<0.001; Table 4).

We observed significantly lower germination rates across both years for *P. fimbriata* when compared to *P. americana* (PAM=13.8 ± 3.7%, PF=6.1 ± 0.9%; $F_{1,1063}$=33.99, p<0.001). The average germination rate across all populations and sites was 6.2 ± 1.5% (n=640) for the first year and 5.9 ±1.1% (n=4818) for the second year. Though we had two rounds of planting for *P. fimbriata*, only plants from the second round went on to flower. Therefore, we analyzed both years of germination data combined as well as only the second year. For both years, source population had a significant effect on germination rates ($F_{6,119}$=10.29, p<0.001; Figure 3), and post-hoc tests revealed that population PF6 was significantly higher than all other populations (p<0.001). Additionally, population PF6 had a trend towards higher germination in the GA site in comparison to the SC site, but after correcting for multiple comparisons, this difference was not significantly different (t_{405}=1.44, p=0.149). In year two, source population remained a significant factor impacting germination rates ($F_{6,119}$=11.39, p<0.001, Table 5)—population PF7 (7.0 ± 1.3%) was significantly higher than populations PF3 (1.2 ± 1.5%; t_{109}=-2.85, p=0.005) and PF4 (2.3 ± 1.6%; t_{109}=-2.77, p=0.007). Finally, there was a significant home-site advantage for germination where home individuals outperformed away individuals. This was significant within the GA site (t_{405}=3.87, p<0.001) and marginally significant in the SC site (t_{405}=1.77, p=0.078).

Vegetative Traits

For *P. americana*, a significant site effect on branch number ($F_{1,99}$=4.71. p=0.032), longest branch length ($F_{1,99}$=10.40, p=0.001), and plant area ($F_{1,99}$=8.34, p<0.001) revealed that plants tended to be larger when grown in the SC transplant site (Table 4). We observed significant population effects for these three traits as well (branch

number: $F_{3,61}$=3.44, p=0.022; longest branch length: $F_{3,61}$=3.85, p=0.014; plant area $F_{3,61}$=5.69, p=0.002). Also, transplant site by population interaction effects were significant for branch number ($F_{5,99}$=2.72, p=0.024) and plant area ($F_{5,99}$=2.92, p=0.017), but not longest branch length.

While *P. fimbriata* individuals were not significantly different in their widths, we observed site differences such that SC had taller plants (marginally significant; SC=20.6 ± 1.2 cm, GA=10.0 ± 1.2 cm; $F_{1,16}$=3.64, p=0.074; Table 5), with a greater number of leaves per plant (SC=15.7 ± 1.2, GA=8.6 ± 1.2; $F_{1,16}$=6.34, p=0.023) and a marginally faster growth rate (SC=0.170 ± 1.17 cm/day, GA=0.095 ± 1.18 cm/day; $F_{1,13}$=2.96, p=0.100; Table 5). Additionally, population by site interaction effects were significant for height and growth rate ($F_{5,16}$=2.88, p=0.048 and $F_{5,13}$=5.04, p=0.009, respectively; Table 5). Even though *P. fimbriata* plants in the SC site (an "away" site for all *P. fimbriata* individuals) tended to be taller, have more leaves and grow faster, we observed a significant home-site advantage within the GA site for height (Table 5 and Figure 4; t_{16}=2.62, p=0.019), growth rate (t_{16}=2.34, p=0.036), and leaf number (marginally significant; t_{16}= 1.95, p=0.069). These home-site advantages, where 'home' individuals outperform 'away' individuals, is evidence of plasticity in these traits. While source population did not have a significant effect on the number of leaves or branches for any given plant, source population was significant for height and growth rate. Individuals from population PF3 tended to be significantly shorter than those from other populations (5.8 ± 0.1 cm vs. 8.9-28.4 cm; F5,51=3.19, p=0.014), and individuals from population PF5 grew significantly faster (0.57 ± 1.69 cm/day) when compared to most other populations that grew at an average rate of 0.08-0.12 cm/day ($F_{5,45}$=4.72, p=0.002).

Reproductive Traits

A total of 33 individuals from across five source populations of *P. americana* flowered during this study. While there were no significant differences among populations or sites for any of the reproductive traits measured, region had a significant impact on the number of fructescences produced per plant ($F_{2,61}$=5.07, p=0.009) and a marginally significant effect on the percent of seeds per plant that were aborted ($F_{1,15}$=4.48, p=0.052). Individuals from source populations in GA produced significantly more fructescences (1.57 ± 0.14 per plant), but had higher frequencies of aborted seeds (0.17 ± 0.02) than SC individuals. Alternatively, SC individuals produced 0.83 ± 0.22 fructescences per plant, but had lower frequencies of aborted seeds (0.10 ± 0.01), regardless of transplant site. Finally, we observed a significant home vs. away effect for number of fructescences (t_{99}=-2.15, p=0.034), specifically with home plants producing more fructescences than away plants in the SC site (H: 0.47 ± 0.18 per plant, A: 0.22 ± 0.13 per plant; t_{99}=-1.75, p=0.083).

Of the *P. fimbriata* individuals that germinated in the second year, 52 survived to flower. The only significant main effects we observed were on seedset and mean per-seed weight. For seedset, source population had a significant impact ($F_{5,33}$=7.84, p=0.001). Specificaly, where population PF5 consistently had significantly lower seedset values (0.38 ± 0.05 seeds/flowers), than all other populations but PF4, regardless of transplant site. For mean per-seed weight, site had a significant effect in that plants growing in the GA site had higher per-seed weights (0.32 ± 0.13 mg/seed) than those in SC (0.19 ± 0.12 mg/seed; $F_{1,6}$=6.16, p=0.048). We also observed significant home-site advantages for

these two traits (seedset: t_6=2.57, p=0.042; per-seed weight: t_7=3.23, p=0.014), but this was only maintained within the site level for seedset within GA (t_7=2.49, p=0.042).

Fitness

Because *P. americana* is a perennial species, survival (to the end of the two-year study), biomass, and a composite value of early lifetime fitness (germination*survival*biomass) were considered for fitness estimates. Survival of *P. americana* individuals was most greatly affected by source region. Significant effects of region on survival differences ($F_{3,72}$=3.33, p=0.024), appeared to be driven by higher survival of individuals from the TX population (90.3 ± 1.09) in comparison to individuals from GA populations (69.0 ± 1.04; t_{72}=-2.88, p=0.014). In contrast, survival of individuals from SC populations was not significantly different from either of these other two regions (79.2 ± 1.08). We also observed a marginally significant site by source population interaction effect on survival (Figure 5a; $F_{5,268}$=2.16, p=0.059). Additionally orthogonal contrasts revealed a marginally significant home-site advantage within the SC site alone (t_{268}=1.83, p=0.069). This may be evidence for local adaptation, but absence of a similar effect within the GA site complicates this inference. There were no other significant home-site advantages for the remaining fitness traits.

In addition to survival, we observed a significant effect on biomass of region ($F_{2,61}$=3.55, p=0.0347), as well as a significant transplant site by population interaction effect (Figure 5b; $F_{5,87}$=2.35, p=0.048), and a significant impact by population ($F_{3,61}$=3.58, p=0.019). Individuals from population PAM12 (TX) tended to have significantly lower biomass compared to most other populations, and individuals from population PAM26 (GA) tended to have significantly higher biomass than most other

populations (Figure 6). Similarly, when considering region as a whole, individuals from TX (0.44 g, SE=1.47) had significantly lower biomass when compared to individuals from SC (2.02 g, SE=1.53) and GA (0.99 g, SE=1.22; $F_{2,61}$=3.55, p=0.0347). Finally for *P. americana*, all main effects were marginally significant for our composite early lifetime fitness measure (Table 4). For the 33 *P. americana* individuals that flowered, we did not observe any significant effects on fecundity.

For *P. fimbriata*, our fitness measures for individuals that flowered in the second cohort were post-germination survival, total dried biomass, fecundity, and lifetime fitness (percent germination*fecundity). Survival in *P. fimbriata* was significantly twice that of *P. americana* (PAM=29.8 ± 0.05%, PF=62.9 ± 0.06%; $F_{1,299}$=18.00, p<0.001; Table 5). Also, within *P. fimbriata*, there was a significant home-site advantage for survival at the GA transplant site (Figure 4d; t_{16}=2.18, p=0.045). No other fitness traits were significant for any of the main effects in our main model.

Performance of Genetic Clusters and GxE

Our analyses with genetic clusters allowed us to identify more significant effects consistent with GxE for both species. For *P. americana* germination, there was no significant GC by site interaction (Figure 7a), but for vegetative traits we found significant GC by site effects for branch number ($F_{2,101}$=3.63, p=0.030; Figure 7b) and plant area ($F_{2,101}$=4.52, p=0.013; Figure 7c), suggesting significant GxE effects on these two traits (Table 4). With respect to reproductive traits in *P. americana*, we detected a significant effect of GC on number of frucescences ($F_{2,101}$=4.96, p=0.009; Figure 7d) and a marginally significant effect of GC on the number of flowers per plant ($F_{1,5}$=5.29,

p=0.070; Figure 7e). Plants from cluster B (PAM26 and PAM27) had marginally more flowers per plant (1217.4 ± 328.6) than cluster A (PAM1, PAM2, and PAM3; 486.5 ± 146.4), but significantly fewer fructescences per plant (B: 0.50 ± 0.07, A: 0.96 ± 0.15; t_{101}=3.12, p=0.007). Additionally, cluster B had fewer fructescences than cluster C (PAM12 and PAM24; 0.69 ± 0.05 fructescences/plant; t_{101}=-2.14, p=0.034), which did not produce any flowers at all. Finally with respect to the fitness of *P. americana* clusters, we observed significant GxE on total dried biomass ($F_{2,89}$=2.93, p=0.049; Figure 7f) and marginally significant GxE on early lifetime fitness ($F_{2,89}$=2.68, p=0.074; Figure 7h).

In *P. fimbriata*, we observed significant GC by site interactions for all vegetative traits measured, suggesting significant GxE effects on these traits (Table 5, Figure 8). While cluster X performance did not to vary significantly across sites, cluster Y performed differentially better in SC than in GA in that it had more branches and leaves, and grew taller and faster (Figure 8). Additionally, there was significant GxE for survival ($F_{1,19}$=6.25, p=0.022, Table 5), wherein cluster Y survived better in SC than in GA, where survival was nearly zero (t_{19}=-2.71, p=0.014). In contrast, cluster X survival did not differ between sites. Finally, GC had a significant effect on germination and seedset across sites, in that cluster X was higher than cluster Y for both ($F_{1,410}$=10.08, p=0.002 and $F_{1,9}$=6.22, p=0.034, respectively).

Discussion

An important first step to understanding the dynamics controlling range limits is to establish whether species' ranges are in disequilbrium with their suitable habitat. Our goals for this study were to establish whether or not disequilibrium exists, and if not, to

identify the putative reasons for this. One piece of this puzzle is whether or not populations are locally adapted to their home sites. If not, they may not have the necessary genetic "toolbox" to succeed beyond their present locations, which would warrant further investigation. In addition, comparisons between widespread and narrow congeners allow us to discount differences attributable to having different evolutionary histories, while possibly illuminating fundamental differences between species of different range sizes. Specifically for this study, our objectives were to compare levels of local adaptation between the widespread *Polygonella americana* and its narrow congener, *P. fimbriata,* and investigate niche constraints as a plausible explanation for the narrow range of *P. fimbriata*.

In addition to traditional population by site and home versus away effects, we utilized the performance of genetic cluster to investigate local adaptation. The use of genetic clusters in this study is a better way to estimate GxE due to the potential complicating effects of gene flow. The principles of local adaptation dictate that populations will perform best in their home sites compared to all other transplants, and have decreased fitness when moved away from their home environment (Kawecki & Ebert, 2004; Leimu & Fischer, 2008). If gene flow occurs among geographically dissimilar populations, this could create a scenario where away populations perform better or no worse than home populations in a given site, and confound the signal of local adaptation. By grouping populations based on their genetic similarity, not only will any effects we see should be closer to true GxE, we will have better statistical power to detect these GxE effects.

Evidence for Local Adaptation in Widespread Perennial

Overall in *P. americana,* we found evidence for local adaptation based on significant home-site advantage in germination, number of fructescences, and survival. Also, we found evidence for significant GxE for plant area, biomass, and early lifetime fitness based on significant GC by site interactions. Cluster A (plants of SC genetic identity) performed better in the SC site in comparison to other genetic clusters. Furthermore, we observed patterns of differential success between the two species in this study, wherein *P. americana* out-performed *P. fimbriata* across transplant sites. These patterns support our hypothesis that *P. americana* will be lower on the local adaptation spectrum than *P. fimbriata*, which would be demonstrated by better overall performance of *P. americana* in transplant sites.

Specifically with respect to germination, *P. americana* had nearly five-fold germination rates than those of *P. fimbriata*. This is consistent with patterns observed in other common and widespread species, which are expected to have a wider range of ecological conditions under which germination is possible (germination niche breadth; Grubb, 1977). This pattern is not confined in the literature to comparisons across genera; it has also been demonstrated in comparisons between invasive and non-invasive congeners (Grime *et al*., 1981; Radford & Cousens, 2000; Brändle *et al*., 2003). While we are unable to fully characterize the germination niche breadth of either of these two species with only two transplant sites, our data suggest that the germination niche breadth of *P. fimbriata* is narrower than that of *P. americana*. If this were true, it would be consistent with the idea that dispersal propagules from populations with high levels of local adaptation or possible habitat specialization will be unable to persist outside of the

narrow ecological range of tolerances and cues necessary for germination (Donohue et al. 2010).

Generally speaking, we would expect that the widespread distribution of *P. americana* is indicative of many populations being locally adapted to their habitats, as has been observed in other widely distributed plant species (Joshi *et al*., 2001). This is because there are implied costs (energetic, genetic, etc.) inherent in maintaining a genotype that would allow a species to thrive across a wide distribution of habitats (DeWitt *et al.,* 1998). In our study, population-level signals of local adaptation (which would be most visible in population by site interaction effects) are difficult to interpret. This may be due to low sample sizes or confounding effects of differential site performance (SC > GA). However, when considering our genetic clustering of populations, clearer signals begin to emerge. In particular, cluster A (consisting of source populations from SC, that are on the edge of *P. americana's* range) exhibited significantly better performance than other clusters within the SC site. This is evidence of adaptive divergence in these populations.

There are at least a couple of reasons for why we were unable to detect population level adaptation. First, the inclusion of more transplant sites nearer to each of the source population locations would have been a more thorough and reciprocal test. There is increasing evidence across various taxa that local adaptation may often occur at a finer geographic scale than would otherwise be expected (Skelly, 2004; Jump & Peñuelas, 2005; Heggenes & Røed, 2006; Willi & Hoffmann, 2012). This phenomenon, deemed microgeographic adaptation (Selander & Kaufman, 1974), has recently been defined as fine scale adaptation that occurs within the dispersal neighborhood of the organism

(Richardson *et al.*, 2014). Microgeographic adaptation may be promoted by a number of external factors, particularly those that serve to strengthen selection regimes or decrease gene flow of maladapted alleles into a population (Richardson *et al*., 2014). In the case of *P. americana*, which is widely distributed but patchy within its distribution, gene flow among populations is expected to be significantly decreased such that the arrival of a maladapted migrant is unlikely. In this experiment, the smallest transplant distance from any *P. americana* source population to transplant site was 11.4 km (Euclidean distance from PAM27 to the GA site). This distance is greater than what we would expect the dispersal neighborhood to be for a gravity-dispersed plant (Vittoz & Engler, 2007; Thomson *et al.*, 2011). If the populations of *P. americana* in this study are in fact undergoing microgeographic divergence, then our transplant distances may be too coarse to detect it. Second, our definition of population is an arbitrary unit. If there is high gene flow among populations, or habitat fragmentation served to split a population into two genetically similar units, we would be diluting the effect of the "true" population. Instead, by using genetic clusters, it would seem that we are more closely approximating the "true" population effects in our transplant sites. This is supported by the fact that we detected significant GC by site interaction effects in the SC site, indicating that populations in the A cluster are locally adapted to their home sites.

Evidence for Local Adaptation in Narrow Endemic

Significant home-site advantage in conjunction with significantly depressed germination rates overall in *P. fimbriata* support the hypothesis of habitat specialization for this specific fitness correlate. In *P. fimbriata*, this early life history stage trait appears

to act as a sieve, after which the fitness of individuals who have successfully germinated performed similarly across sites. However, we did observe significant home-site advantage effects on survival within the GA site. Germination and post-germination characteristics are expected to be co-adapted in many cases (Donohue *et al.*, 2010). Because selection on post-germination traits occurs subsequent to successful germination, this selection occurs against the background that is already established as favorable for germinants, and thus these traits can become correlated (Kalisz, 1986; Kalisz & Wardle, 1994; Blows, 2007). However, this is not always the case, and plants are expected to be divergent from this expectation particularly when the germination niche breadth is narrow but the post-germination niche breadth is relatively wide (Donohue *et al*., 2010). This mismatch of different life history stage niche breadths can decouple germination and post-germination correlations, and result in the type of outcomes we observed in our own study. Our findings are consistent with those of another study, in which germination in *Gilia tricolor* was particularly sensitive to changes in soil conditions, but post-germination survival and success were unaffected by transplant conditions (Baack *et al*., 2006). This investigation of the range limited *G. tricolor* demonstrated that failure to germinate beyond the distinct population range margins of the species were the primary cause for failure to expand its geographic range (Baack *et al*., 2006). Substantially low germination rates, even when transplant distances were <1 km, suggest that specific germination cues and tolerances for *P. fimbriata* are a limiting factor in its success.

Similar to *P. americana*, weak evidence of differential population by site performance for *P. fimbriata* was strengthened when considering the performance of genetic clusters across sites. However, when considering vegetative traits, GC

performance did not show home-site advantage. Under the auspices of local adaptation, we would have expected cluster X, which comprised populations with genetic identities most commonly found near the GA site, to outperform cluster Y within GA. Instead, Cluster B performance was not significantly different from cluster X within the GA site. Also, cluster Y significantly outperformed cluster X in the SC site. In contrast, fitness-correlated traits (germination, seedset, survival) under GC clustering was impacted mainly by GC, and generally did not differ across sites. Typically, vegetative traits are expected to correlate with overall plant fitness (). However, a meta-analysis of local adaptation studies in plants demonstrated that the regression of local adaptation on vegetative phenotypic differences in plants did not explain any of the variation observed in local adaptation across sites (Hereford, 2009). Overall, there is no clear evidence to support a hypothesis of habitat specialization when considering all fitness traits for this species as a whole.

Factors Constraining the Expansion of P. fimbriata

Because marginal populations in range-limited species are expected to play an integral role in range expansion, a consideration of edge versus interior population performance in *P. fimbriata* is important. However, because the distribution of *P. fimbriata* is so narrow and geographically disjunct, defining populations as "edge" or "interior" is not straightforward. Not only are there are no clear guidelines for this classification process (Hargreaves *et al.*, 2014), there is dissension in the literature regarding the appropriate construction of a simple range map to aid in making such distinctions (Fortin *et al.*, 2005; Getz *et al*., 2007; Kie *et al.*, 2010). To avoid this

confusion, we approach the classification of edge/interior populations from a biological perspective. Edge and interior populations are characterized not only by their location in the range, but also by their density across the range (Brown, 1984; Bahn *et al.*, 2006), and by expected levels and direction of gene flow among populations, which is expected to be biased towards the edge (Slatkin, 1985; García-Ramos & Kirkpatrick, 1997). If we place *P. fimbriata* populations into this framework, populations centered around the GA site (PF1, PF2, PF6, and PF7) can be classified as interior, and those further away and more sparsely distributed (PF3, PF4, and PF5) as edge populations. If niche constraints were preventing the expansion of *P. fimbriata*'s range, then edge populations planted beyond the range margins should have exhibited significantly decreased performance. Ultimately, both within this framework and from our GC clustering perspective, our data do not support the hypothesis. While we do not have an edge transplant site for more thorough comparisons, we do have the performance of interior populations at our two transplant sites as a point of comparison to edge populations. Similarly to our GC analysis, edge populations outperformed interior populations at our 'beyond' site with respect to vegetative traits, while no differences across sites were observed in the fitness data (germination and seedset). Additionally, interior performance for these traits was not significantly different between site. Taken together, these data suggest that niche constraints are not playing a significant role in the range limitation of *P. fimbriata*.

Overall, when considering GC clusters in conjunction with edge/interior classifications of *P. fimbriata* populations, it seems more likely that dispersal, population genetic constraints, and/or metapopulation dynamics play a more important role in the range edge. If genetic constraints are a factor in *P. fimbriata* range limits, then most

likely it is because they are small, isolated populations, and are subject to increased genetic drift and inbreeding (Kruckeberg & Rabinowitz, 1985). Also, it is likely that *P. fimbriata* populations are dispersal limited, as they have gravity-dispersed seeds, the dispersal distance of which is generally considerably less than 1 m (Vittoz & Engler, 2007; Thomson *et al.*, 2011). Ultimately, more work is needed to determine what role dispersal and genetic constraints may be playing in maintaining the range edge of *P. fimbriata*.

Putative Factors for Differential Performance

While site effects were significant for some traits of both species in this study, our site similarity analysis revealed no significant climatological differences. Additionally, while we were unable to perform soil core sample analyses, a coarse evaluation of the soil type and drainage amounts from the NRCS soils database did not reveal any noticeable differences among sites (NRCS, USDA). There are two things we could not control in this study, which may have contributed to this. First, while we did not control competing vegetation within sites during the two-year study, we did have the SC site disked prior to planting. In contrast, the GA site was not disked due to the presence of a state-threatened plant species (*Stylisma pickeringii*). This may have boosted initial germination rates in SC, but is not expected to have had long-term effects on the success of plants in the field as both the settling of the soil and the recolonization of the native vegetation occurred quickly. Second, the SC site was surrounded by tall pine stands and parts of it were significantly shaded for portions of the day. In contrast, the GA site received full-sun all day. While our knowledge of these plants would dictate their

preference for greater sun exposure, lighting difference between sites was something we were unable to control. When studies have found significant differences in performance between transplant sites, these transplants generally incorporated only a single species or the incorporation of ecological or latitudinal clines was an objective of the experiment (e.g., Antonovics & Primack, 1982; Angert & Schemske, 2007). Though, one study demonstrated that performance differences between reciprocally transplanted ramets from central and marginal populations of *Ranunculus* reflected site differences in herbivory and fungal infections (Johansson, 1994). Ultimately, deciphering the cause for these differences will require further investigation.

Finally, it is interesting to note that even though we observed a significant home-site advantage for growth rate within the GA site for *P. fimbriata*, one population that was considered 'away' at both transplant sites had a significantly higher growth rate than the rest. Plants from population PF5 grew six times faster than plants from any other population, regardless of transplant site. This population is unique in that it grows along the top and steep side of an exposed sand dune, and is sympatric with a population of another species of *Polygonella*, *P. gracilis.* It is possible that the faster growth rate in PF5 is a reflection either of competition with *P. gracilis* for pollinators and other resources in this particular site or interspecific gene flow from *P. gracilis* to *P. fimbriata*. In the event that either is true, it would be reasonable to expect that individuals from PF5 would have different phenology compared compared to other *P. fimbriata* populations. *P. gracilis* is a dioecious species (compared to gynomonoecious *P. fimbriata*). In the event intraspecific gene flow is occurring, it would be expected to be more heavily biased into *P. fimbriata* due to the ratios of female:male function in *P. fimbriata* (3:1) versus *P. gracilis* (1:1).

Conclusions

Overall, we were able to detect stronger evidence of local adaptation in populations of *P. americana* than in populations of *P. fimbriata*. Failure to detect stronger signals of adaptation in *P. fimbriata* may have been a result of low numbers of flowering individuals, and thus low sample sizes. That said, germination patterns in this narrow endemic are consistent with expectations for habitat specialist species. Finally, these data fail to support a hypothesis of niche constraints in *P. fimbriata*. Instead, it is possible that populations of this species are dispersal limited, under metapopulation dynamics, and/or face population genetic constraints to expansion, which would prevent them from establishing stable populations in the tolerance distribution. Ultimately, further information is needed to test these ideas. Smaller-scale transplant experiments incorporating transplant sites at the range edge would allow us to better detect dispersal limitations and potential maladaptation from drift or interior-to-edge gene flow. Additionally, an investigation into the levels of gene flow, genetic diversity, and inbreeding in these populations would elucidate potential population genetic constraints.

Literature Cited

Alleaume-Benharira, M., Pen, I.R. & Ronce, O. (2006) Geographical patterns of adaptation within a species' range: interactions between drift and gene flow. *Journal of Evolutionary Biology*, **19**, 203-15.

Angert, A.L. & Schemske, D.W. (2007) The evolution of species' distributions: reciprocal transplants across the elevation ranges of *Mimulus cardinalis* and *M. lewisii*. *Evolution*, **59,** 1671-84.

Antonovics, J. & Primack, R.B. (1982) Experimental ecological genetics in *Plantago*: VI. The demography of seedling transplants in *P. lanceolata*. *Journal of Ecology*, **70,** 55-75.

Bahn, V., O'Connor, R.J. & Krohn, W.B. (2006) Effect of dispersal at range edges on the structure of species ranges. *Oikos*, **115**, 89-96.

Berryman, A.A. (2002) Population: a central concept for ecology? *Oikos*, **97**, 439-42.

Blows, M.W. (2007) A tale of two matrices: multivariate approaches in evolutionary biology. *Journal of Evolutionary Biology*, **20**, 1-8.

Brändle, M., Stadler, J., Klotz, S. & Brandl, R. (2003) Distributional range size of weedy plant species is correlated to germination patterns. *Ecology*, **84**, 136-4.

Bridle, J.R. & Vines, T.H. (2007) Limits to evolution at range margins: when and why does adaptation fail? *Trends in Ecology and Evolution*, **22**, 140-7.

Brown, J.H. (1984) On the relationship between abundance and distribution of species. *The American Naturalist*, **124,** 255-79.

Darwin, C. (1859) *On the origins of species by means of natural selection*. Jon Murray, London.

Davis, M.B. & Shaw, R.G. (2001) Range shifts and adaptive responses to quarternary climate change. *Science*, **292**, 673-79.

Debussche, M. & Thompson, J.D. (2003) Habitat differentiation between two closely related Mediterranean plant species, the endemic *Cyclamen balearicum* and the widespread *C. repandum*. *Acta Oecologica,* **24**, 35-45.

DeWitt, T.J., Sih, A. & Wilson, D.S. (1998) Costs and limits of phenotypic plasticity. *Trends in Ecology and Evolution*, **13,** 77-81.

Donohue, K., de Casas, R.R., Burghardt, L., Kovach, K. & Willis, C.G. (2010) Germination, postgermination adaptation, and species ecological ranges. *Annual Review of Ecology, Evolution, and Systematics,* **41**, 293-319.

Fortin, M.-J., Keitt, T.H., Maurer, B.A., Taper, M.L., Kaufman, D.M. & Blackburn, T.M. (2005) Species' geographical ranges and distributional limits: pattern analysis and statistical issues. *Oikos*, **108**, 7-17.

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

Gaston, K.J. (2003) *The structure and dynamics of geographic ranges*. Oxford University Press, Oxford, U.K.

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

Geber, M. (2008) To the edge: studies of species' range limits. *New Phytologist*, **178**, 228-30.

Getz, W.M., Fortmann-Roe, S., Cross, P.C., Lyons, A.J., Ryan, S.J. & Wilmers, C.C. (2007) LoCoH: Nonparameteric kernel methods for constructing home ranges and utilization distributions. *PlosOne*, **2,** e207, 1-11.

Gomulkiewicz, R., Holt, R.D. & Barfield, M. (1999) The effects of density dependence and immigration on local adaptation and niche evolution in a black-hole sink environment. *Theoretical Population Biology*, **55**, 283-96.

Grime, J.P., Mason, G., Curtis, A.V., Rodman, J., Band, S.R., Mowforth, M.A.G., Neal, A.M. & Shaw, S. (1981) A comparative study of germination characteristics in a local flora. *Journal of Ecology*, **69**, 1017-59.

Grubb, P.J. (1977) The maintenance of species-richness in plant communitites: the importance of the regeneration niche. *Biological Review*, **52**, 107-45.

Hargreaves, A.L, Samis, K.E. & Eckert, C.G. (2014) Are species' range limits simply niche limits writ large? A review of transplant experiments beyond the range. *The American Naturalist,* **183,** 157-73.

Heggenes, J. & Røed, K.H. (2006) Do dams increase genetic diversity in brown trout (*Salmo trutta*)? Microgeographic differentiation in a fragmented river. *Ecology of Freshwater Fish*, **15**, 366-75.

Hereford, J. (2009) A quantitative survey of local adaptation and fitness trade-offs. *The American Naturalist*, **173**, 579-88.

Holt, R.D. (2003) On the evolutionary ecology of species' ranges. *Evolutionary Ecology Research*, **5,** 159-78.

Holt, R.D. & Gomulkiewicz, R. (1997) How does immigration influence local adaptation? A reexamination of a familiar paradigm. *The American Naturalist*, **149,** 563-72.

Holt, R.D., Keitt, T.H., Lewis, M.A., Maurer, B.A. & Taper, M.L. (2005) Theoretical models of species borders: single species approaches. *Oikos*, **108**, 18-27.

Hong, S.P. & Smets, E.F. (2004) What is the taxonomic status of *Polygonella*? Evidence from floral morphology. *Annals of the Missouri Botanical Gardens*, **91,** 320-45.

Horton, J.H. (1963) A taxonomic revision of *Polygonella* (Polygonaceae). *Brittonia,* **15,** 177-203.

Hubbell, S.P. (2001) *The unified neutral theory of biodiversity and biogeography*. Princeton University Press, Princeton.

Hutchinson, G.E. (1957) Concluding remarks. *Cold Spring Harbor Symposium on Quantitative Biology*, **22**, 415-27.

Johansson, M.E. (1994) Life history differences between central and marginal populations of the clonal aquatic plant *Ranunculus lingua*: a reciprocal transplant experiment. *Oikos*, **70**, 65-72.

Joshi, J., Schmid, B., Caldeira, M.C., Dimitrakopoulos, P.G., Good, J., Harris, R., Hector, A., Huss-Danell, K., Jumpponen, A., Minns, A., Mulder, C.P.H., Pereira, J.S., Prinz, A., Scherer-Lorenzen, M., Siamantziouras, A.-S.D., Terry, A.C., Troumbis, A.Y. & Lawton, J.H. (2001) Local adaptation enhances performance of common plant species. *Ecology Letters*, **4,** 536-44.

Jump, A.S. & Peñuelas, J. (2005) Running to stand still: adaptation and the response of plants to rapid climate change. *Ecology Letters*, **8**, 1010-20.

Kalisz, S. (1986) Variable selection on the timing of germination in *Collinsia verna* (Scrophulariaceae). *Evolution*, **40**, 479-91.

Kalisz, S. & Wardle, G.M. (1994) Life history variation in *Campanula americana* (Campanulaceae): population differentiation. *American Journal of Botany*, **81**, 521-27.

Kingsolver, J.G., Hoekstra, H.E., Hoekstra, J.M., Berrigan, D., Vignieri, S.N., Hill, C.E., Hoang, A., Gibert, P. & Beerli, P. (2001) The strength of phenotypic selection in natural populations. *The American Naturalist*, **157**, 245-61.

Kie, J.G., Matthiopoulos, J., Fieberg, J., Powell, R.A., Cagnacci, F., Mitchell, M.S., Gaillard, J.-M. & Moorcroft, P.R. (2010) The home-range concept: are traditional estimators still relevant with modern telemetry technology? *Philosophical Transactions of the Royal Society B*, **365**, 2221-2231.

Kruckeberg, A.R. & Rabinowitz, D. (1985) Biological aspects of endemism in higher plants. *Annual Review of Ecology and Systematics*, **16**, 447-79.

Lavergne, S., Thompson, J.D., Garnier, E. & Debussche, M. (2004) The biology and ecology of narrow endemic and widespread plants: a comparative study of trait variation in 20 congeneric pairs. *Oikos,* **107**, 505-18.

Leimu, R. & Fischer, M. (2008) A meta-analysis of local adaptation in plants. *PLoS ONE*, **3**: pp4010.

Lewis, P.O. (1991) *Allozyme variation and evolution in Polygonella (Polygonaceae)*. Dissertation, Ohio State University.

Lewis, P.O. & Crawford, D.J. (1995) Pleistocene refugium endemics exhibit greater allozymic diversity than widespread congeners in the genus *Polygonella* (Polygonaceae). *American Journal of Botany,* **82**, 141-9.

MacArthur, R.H. (1972) *Geographical ecology*. Harper & Row, New York.

MacArthur, R.H. & Wilson, E.O. (1967) *The theory of island biogeography*. Princeton University Press, Princeton, NJ.

Maliakal-Witt, S., Menges, E.S. & Denslow, J.S. (2005) Microhabitat distribution of two Florida scrub endemic plants in comparison to their habitat-generalist congeners. *American Journal of Botany*, **92**, 411-21.

Moeller, D.A., Geber, M.A. & Tiffin, P. (2011) Population genetics and the evolution of geographic range limits in an annual plant. *The American Naturalist*, **178**, S44-S57.

Murren, C.J., Maclean, H.J., Diamond, S.E., Steiner, U.K., Heskel, M.A., Handelsman, C.A., Ghalambor, C.K., Auld, J.R., Callahan, H.S., Pfennig, D.W., Relyea, R.A., Schichting, C.D. & Kingsolver, J. (2014) Evolutionary change in continuous reaction norms. *The American Naturalist*, **183**, 453-67.

Nathan, R. & Muller-Landau, H.C. (2000) Spatial patterns of seed dispersal, their determinants and consequences for recruitment. *Trends in Ecology and Evolution*, **15**, 278-85.

Natural Resources Conservation Service (NRCS), USDA. Web Soil Survey. http://websoilsurvey.sc.egov.usda.gov, data accessed December 2013.

PRISM Climate Group, Oregon State University. (2013) http://prism.oregonstate.edu, data created 9 Sep - 19 Dec 2013.

R Development Core Team. (2009) R: A Language and Environment for Statistical Computing. The R Foundation for Statistical Computing. Vienna, Austria. http://www.R-project.org

Radford, I.J. & Cousens, R.D. 2000. Invasiveness and comparative life-history traits of exotic and indigenous *Senecio* species in Australia. Oecologia 125:531-42.

SAS Core Team. (2011) Base SAS® 9.3 Procedures Guide. Cary, NC: SAS Institute Inc.

Sax, D.F., Early, R. & Bellemare, J. (2013) Niche syndromes, species extinction risks, and management under climate change. *Trends in Ecology and Evolution*, **28**, 517-23.

Schuster, T.M., Reveal, J.L. & Kron, K.A. (2011) Phylogeny of Polygoneae (Polgonaceae: Polygonoideae). *Taxon*, **60,** 1653-66.

Selander, R.K. & Kaufman, D.W. (1975) Genetic structure of populations of the brown snail (*Helix aspersa*). I. Microgeographic variation. *Evolution*, **29,** 385-401.

Skelly, D.K. (2004) Microgeographic countergradient variation in the wood frog, *Rana sylvatica*. *Evolution*, **58,** 160-5.

Slatkin, M. (1985) Gene flow in natural populations. *Annual Review of Ecology and Systematics*, **16**, 393-430.

Thomson, F.J., Moles, A.T., Auld, T.D. & Kingsford, R.T. (2011) Seed dispersal distance is more strongly correlated with plant height than with seed mass. *Journal of Ecology*, **99**, 1299-1307.

Vittoz, P. & Engler, R. (2007) Seed dispersal distances: a typology based on dispersal modes and plant traits. *Botanica Helvetica*, **117,** 109-24.

Waples, R.S. & Gaggiotti, O. (2006) What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. *Molecular Ecology*, **15**, 1419-1439.

Willi, Y. & Hoffmann, A.A. (2012) Microgeographic adaptation linked to forest fragmentation and habitat quality in the tropical fruit fly *Drosophila birchii*. *Oikos*, **121**, 1627-37.

Table 3.1. Pairwise Euclidean distances (km) between all populations of P. americana and the two transplant sites, "GA" and "SC." Populations considered 'home' within a given transplant site are shown in bold. Letters in parentheses indicate genetic cluster membership.

	PAM1(A)	PAM2(A)	PAM3(A)	PAM12(C)	PAM24(C)	PAM26(B)	PAM27(B)	GA	SC
PAM1	—								
PAM2	23.73	—							
PAM3	39.53	28.27	—						
PAM12	1466.58	1443.14	1434.86	—					
PAM24	620.27	598.60	583.64	880.99	—				
PAM26	276.94	253.66	256.73	1208.97	421.29	—			
PAM27	286.66	262.93	260.72	1185.04	378.69	48.75	—		
GA	275.80	252.09	250.31	1196.36	388.91	**43.61**	**11.35**	—	
SC	**13.41**	**14.66**	**26.49**	1454.45	607.08	267.55	275.76	265.01	—

Table 3.2. Pairwise Euclidean distances (km) between all populations of *P. fimbriata* and the two transplant sites, "GA" and "SC." Populations considered 'home' within the "GA" site are in bold (no populations of this species were considered 'home' in the SC site), and populations considered 'edge' sites are boxed with a dashed line. Letters in parentheses indicate genetic cluster membership.

	PF1(X)	PF2(X)	PF3(Y)	PF4(X)	PF5(Y)	PF6(X)	PF7(X)	GA	SC
PF1	—								
PF2	8.73	—							
PF3	150.79	156.57	—						
PF4	221.39	228.06	75.98	—					
PF5	124.63	127.89	54.66	127.55	—				
PF6	6.35	15.07	146.40	216.33	122.14	—			
PF7	7.99	3.11	157.83	228.91	130.01	14.20	—		
GA	**8.85**	**3.15**	158.58	229.71	130.57	**15.05**	**0.85**	—	
SC	256.41	265.07	178.24	143.71	228.59	250.06	264.17	265.01	—

Table 3.3. Traits measured for each plant in the GA and SC transplant sites.

Trait	***P. americana*** # Observations	***P. fimbriata*** # Observations
Percent Germination	772	538
Vegetative		
Number of branches	172	80
Number of leaves	-	80
Longest branch length	172	-
Area (cm^2)	172	-
Height (cm)	-	80
Growth Rate (mm/day)	-	70
Reproductive		
Total number of flowers per plant	31	52
Total number of seeds	31	52
Total seed weight	31	52
Total per-seed weight	31	52
Percent of aborted seeds	31	52
Seedset	31	52
Number of fructescences	31	-
Fructescence height (cm)	31	-
Fitness		
Percent Survival	353	81
Total dried biomass (mg)	160	56
Fecundity	31	52

Table 3.4. Environmental (Site), population, region, genetic cluster (GC) and home (H) vs. away (A) effects on plant life-history traits measured for *P. americana*. Table entries are *p* values and F-statistics (with degrees of freedom in subscript) from generalized linear mixed models (for Full Model and Genotypic Clusters), and *p* values and *t*-statistics (with degrees of freedom in subscript) from orthogonal contrasts for home vs. away effects tested within the full model. Bold numbers indicate significance ($p<0.05$ level); underlined values indicate marginal significance ($p<0.10$). Arrows on orthogonal contrast *p*-values indicate the direction of the orthogonal contrast if significant: H >A (↑), H<A (↓), GA>SC (↑), SC>GA (↓).

P. americana	Site	Population	Site*Population	Region	H vs. A	(in GA) H vs. A	(in SC) H vs. A	H (GA) vs. H (SC)	A (GA) vs. A (SC)
Full Model					**Orthogonal Contrasts**				
Germination	**0.031** $F_{1,659}$=4.70	**0.045** $F_{3,99}$=2.78	**0.044** $F_{6,659}$=2.18	**<0.001** $F_{3,99}$=14.75	**0.016**↑ t_{659}=2.41	0.841 t_{659}=-0.20	**<0.001**↑ t_{659}=3.93	**0.024**↓ t_{659}=-2.26	0.120 t_{659}=-2.26
Number of branches	**0.032** $F_{1,99}$=4.71	**0.022** $F_{3,61}$=3.44	**0.024** $F_{5,99}$=2.72	0.368 $F_{2,61}$=1.02	0.432 t_{99}=-0.79	0.315 t_{99}=-1.01	0.914 t_{99}=-0.11	0.951 t_{99}=-0.06	0.438 t_{99}=-0.78
Longest branch length	**0.002** $F_{1,99}$=10.40	**0.014** $F_{3,61}$=3.85	0.291 $F_{5,99}$=1.25	0.166 $F_{2,61}$=1.85	0.285 t_{99}=-1.07	0.695 t_{99}=-0.39	0.365 t_{99}=-0.91	0.213 t_{99}=-1.25	0.558 t_{99}=0.59
Area	**0.005** $F_{1,99}$=8.34	**0.002** $F_{3,61}$=5.69	**0.017** $F_{5,99}$=2.92	0.062 $F_{2,61}$=2.91	0.285 t_{99}=-1.08	0.419 t_{99}=-0.81	0.474 t_{99}=-0.72	0.861 t_{99}=0.18	0.756 t_{99}=-0.31
Number of Flowers	0.629 $F_{1,5}$=0.26	0.657 $F_{3,15}$=0.57	0.100 $F_{4,5}$=0.87	0.539 $F_{1,15}$=4.05	0.178 t_{5}=1.57	0.285 t_{5}=1.20	0.478 t_{5}=0.77	0.328 t_{5}=-1.08	0.713 t_{5}=-0.39
Total seed weight	0.575 $F_{1,6}$=0.35	0.515 $F_{3,16}$=0.81	0.585 $F_{4,6}$=0.76	0.835 $F_{1,16}$=0.04	0.585 t_{6}=0.58	0.934 t_{6}=0.09	0.409 t_{6}=0.89	0.432 t_{6}=0.84	0.981 t_{6}=-0.02
Mean per-seed weight	0.502 $F_{1,6}$=0.51	0.291 $F_{3,16}$=1.36	0.460 $F_{4,6}$=1.04	0.115 $F_{1,16}$=2.77	0.164 t_{6}=-1.58	0.159 t_{6}=-1.61	0.739 t_{6}=-0.35	0.209 t_{6}=1.41	0.923 t_{6}=-0.10
% of aborted seeds	0.335 $F_{1,4}$=1.20	0.486 $F_{3,15}$=0.85	0.661 $F_{3,4}$=0.58	0.052 $F_{1,15}$=4.48	0.397 t_{4}=0.95	0.051↑ t_{4}=2.75	0.093↑ t_{4}=-2.19	**0.005**↑ t_{4}=-5.65	0.814 t_{4}=0.25
Seedset	0.639 $F_{1,5}$=0.25	0.370 $F_{3,15}$=1.13	0.905 $F_{4,5}$=0.24	0.224 $F_{1,15}$=1.61	0.357 t_{5}=-1.01	0.182 t_{5}=-1.55	0.827 t_{5}=0.23	0.225 t_{5}=1.39	0.627 t_{5}=-0.52
# of fructescences	0.548 $F_{1,99}$=0.36	0.205 $F_{3,61}$=1.57	0.684 $F_{5,99}$=0.62	**0.009** $F_{2,61}$=5.07	**0.034**↑ t_{99}=-2.15	0.156 t_{99}=-1.43	0.083↑ t_{99}=-1.75	0.212 t_{99}=1.26	0.590 t_{99}=-0.54
Fructescence height	0.695 $F_{1,5}$=0.17	0.853 $F_{3,15}$=0.26	0.720 $F_{4,5}$=0.53	0.734 $F_{1,15}$=0.12	0.264 t_{5}=1.26	0.259 t_{5}=1.27	0.590 t_{5}=0.58	0.819 t_{5}=-0.24	0.672 t_{5}=0.45
Percent Survival	0.569 $F_{1,268}$=0.32	0.957 $F_{3,72}$=0.10	0.059 $F_{5,268}$=2.16	**0.024** $F_{3,72}$=0.10	0.835 t_{268}=-0.21	0.207 t_{268}=-1.27	0.069↑ t_{268}=1.83	**0.002**↓ t_{268}=-3.15	0.556 t_{268}=-0.59
Total dried biomass	0.060 $F_{1,87}$=3.64	**0.019** $F_{3,61}$=3.58	**0.048** $F_{5,87}$=2.35	**0.035** $F_{2,61}$=3.55	0.158 t_{87}=-1.42	0.239 t_{87}=-1.19	0.369 t_{87}=-0.90	0.549 t_{87}=0.60	0.504 t_{87}=-0.67
Fecundity	0.655 $F_{1,5}$=0.23	0.586 $F_{3,15}$=0.67	0.563 $F_{4,5}$=0.82	0.098 $F_{1,15}$=3.12	0.232 t_{5}=1.36	0.407 t_{5}=0.90	0.466 t_{5}=0.79	0.455 t_{5}=-0.81	0.649 t_{5}=-0.48
Early lifetime fitness	0.061 $F_{1,87}$=3.59	0.064 $F_{3,61}$=2.56	0.052 $F_{5,87}$=2.29	0.061 $F_{2,61}$=2.93	0.222 t_{87}=-1.23	0.211 t_{87}=-1.26	0.569 t_{87}=-0.57	0.547 t_{87}=0.61	0.350 t_{87}=-0.94

Table 3.4 (continued).

Genetic Clusters	Site	Population	GC	Site*GC
Germination	**0.002** $F_{1,633}$=9.77	**<0.001** $F_{4,633}$=6.10	**<0.001** $F_{2,633}$=14.09	0.478 $F_{2,633}$=0.74
Number of branches	**<0.001** $F_{1,101}$=17.49	0.052 $F_{3,101}$=2.67	0.442 $F_{2,101}$=0.82	**0.030** $F_{2,101}$=3.63
Longest branch length	**<0.001** $F_{1,101}$=23.31	**0.016** $F_{3,101}$=3.63	0.165 $F_{2,101}$=1.84	0.201 $F_{2,101}$=1.63
Area	**<0.001** $F_{1,101}$=26.41	**0.002** $F_{3,101}$=5.13	0.075 $F_{2,101}$=2.66	**0.013** $F_{2,101}$=4.52
Number of Flowers	0.659 $F_{1,5}$=0.22	0.721 $F_{3,5}$=0.46	0.070 $F_{1,5}$=5.29	0.344 $F_{1,5}$=1.09
Total seed weight	0.815 $F_{1,6}$=0.06	0.321 $F_{3,6}$=1.44	0.730 $F_{1,6}$=0.13	0.150 $F_{1,6}$=2.72
Mean per-seed weight	0.131 $F_{1,6}$=3.05	0.094 $F_{3,6}$=3.40	0.237 $F_{1,6}$=1.73	0.196 $F_{1,6}$=2.12
% of aborted seeds	0.566 $F_{1,4}$=0.39	0.240 $F_{3,4}$=2.13	0.130 $F_{1,4}$=3.61	0.910 $F_{1,4}$=0.01
Seedset	0.358 $F_{1,5}$=1.03	0.372 $F_{3,5}$=1.30	0.154 $F_{1,5}$=2.83	0.519 $F_{1,5}$=0.48
# of fructescences	0.310 $F_{1,101}$=1.04	0.243 $F_{3,101}$=1.42	**0.009** $F_{2,101}$=4.96	0.441 $F_{2,101}$=0.83
Fructescence height	0.797 $F_{1,5}$=0.07	0.664 $F_{3,5}$=0.56	0.905 $F_{1,5}$=0.02	0.326 $F_{1,5}$=1.19
Percent Survival	0.250 $F_{1,271}$=1.33	0.400 $F_{4,271}$=1.01	0.292 $F_{2,271}$=1.24	0.640 $F_{2,271}$=0.45
Total dried biomass	**0.001** $F_{1,89}$=3.59	**0.041** $F_{3,89}$=2.56	0.051 $F_{2,89}$=2.29	**0.049** $F_{2,89}$=2.93
Fecundity	0.704 $F_{1,5}$=0.16	0.623 $F_{3,5}$=0.64	0.098 $F_{1,5}$=4.14	0.357 $F_{1,5}$=1.03
Early lifetime fitness	**<0.001** $F_{1,89}$=15.41	0.116 $F_{3,89}$=2.02	0.092 $F_{2,89}$=2.45	0.074 $F_{2,89}$=2.68

Table 3.5. Environmental (Site), population, region, genetic cluster (GC) and home (H) vs. away (A) effects on plant life-history traits measured for *P. fimbriata*. Table entries are *p* values and F-statistics (with degrees of freedom in subscript) from generalized linear mixed models (for Full Model and Genetic Clusters), and *p* values and *t*-statistics (with degrees of freedom in subscript) from orthogonal contrasts for home vs. away effects tested within the full model. Bold numbers indicate significance ($p<0.05$ level); underlined values indicate marginal significance ($p<0.10$). Arrows on orthogonal contrast *p*-values indicate the direction of the orthogonal contrast if significant: H >A (↑), H<A (↓), GA>SC (↑), SC>GA (↓).

P. fimbriata	Site	Population	Site*Population	H v. A	(in GA) H v. A	A (GA) v. A (SC)
Population * Site Interactions						
Germination	0.373 $F_{1,405}$=0.80	**<0.001** $F_{6,119}$=11.39	0.904 $F_{6,405}$=0.36	**0.001**↑ t_{405}=3.51	**<0.001**↑ t_{405}=3.87	0.078 t_{405}=1.77
Number of branches	0.509 $F_{1,16}$=0.45	0.664 $F_{5,52}$=0.65	0.088 $F_{5,16}$=2.35	0.288 t_{16}=1.10	0.140 t_{16}=1.55	0.239 t_{16}=1.22
Number of leaves	**0.023** $F_{1,16}$=6.34	0.256 $F_{5,52}$=1.36	0.131 $F_{5,16}$=2.02	0.897 t_{16}=0.13	0.069↑ t_{16}=1.95	**0.015**↓ t_{16}=2.71
Height	0.074 $F_{1,16}$=3.64	**0.014** $F_{5,51}$=3.19	**0.048** $F_{5,16}$=2.88	0.222 t_{16}=1.27	**0.019**↑ t_{16}=2.62	**0.019**↓ t_{16}=2.60
Growth Rate	0.100 $F_{1,13}$=2.96	**0.002** $F_{5,45}$=4.72	**0.009** $F_{5,13}$=5.04	0.230 t_{13}=1.26	**0.036**↑ t_{13}=2.34	**0.038**↓ t_{13}=2.30
Number of Flowers	0.226 $F_{1,7}$=1.76	0.532 $F_{5,33}$=0.84	0.430 $F_{5,7}$=1.12	0.892 t_{7}=0.14	0.322 t_{7}=1.06	0.205 t_{7}=1.40
Total seed weight	0.235 $F_{1,6}$=1.74	0.677 $F_{5,29}$=0.63	0.473 $F_{4,6}$=1.01	0.229 t_{6}=1.34	0.999 t_{6}=0.00	0.397 t_{6}=-0.91
Total per-seed weight	**0.048** $F_{1,6}$=6.16	0.153 $F_{5,29}$=1.76	0.366 $F_{4,6}$=1.31	**0.042**↓ t_{6}=2.57	0.345 t_{6}=1.03	0.695 t_{6}=-0.41
Percent of aborted seeds	0.413 $F_{1,7}$=0.76	0.692 $F_{5,33}$=0.61	0.491 $F_{4,7}$=0.95	0.061↓ t_{7}=-3.14	0.071↓ t_{7}=-2.13	0.593 t_{7}=-0.56
Seedset	0.608 $F_{1,7}$=0.29	**<0.001** $F_{5,33}$=7.84	0.244 $F_{4,7}$=1.74	**0.014**↑ t_{7}=3.23	**0.042**↑ t_{7}=2.49	0.439 t_{7}=0.82
Percent Survival	0.267 $F_{1,16}$=1.32	0.155 $F_{5,52}$=1.68	0.169 $F_{5,16}$=1.81	0.185 t_{16}=1.38	**0.045**↑ t_{16}=2.18	0.082↓ t_{16}=1.86
Total dried biomass	0.152 $F_{1,9}$=2.45	0.377 $F_{5,35}$=1.10	0.886 $F_{5,9}$=0.32	0.189 t_{9}=-1.42	0.883 t_{9}=-0.15	0.355 t_{9}=0.98
Fecundity	0.295 $F_{1,7}$=1.28	0.352 $F_{5,33}$=1.15	0.480 $F_{5,7}$=1.00	0.503 t_{7}=0.71	0.197 t_{7}=1.43	0.194 t_{7}=1.44
Lifetime fitness	0.341 $F_{1,7}$=1.04	0.139 $F_{5,33}$=1.80	0.802 $F_{5,7}$=0.45	0.940 t_{7}=-0.08	0.534 t_{7}=0.65	0.351 t_{7}=1.00

Table 3.5 (continued)

Genetic Clusters	Site	Population	GC	GC * Site
Germination	0.763 $F_{1,410}$=0.09	**<0.001** $F_{5,410}$=12.42	**0.002** $F_{1,410}$=10.08	0.970 $F_{1,410}$=0.00
Number of branches	0.313 $F_{1,19}$=1.07	0.602 $F_{4,19}$=0.70	0.495 $F_{1,19}$=0.48	**0.010** $F_{1,19}$=8.25
Number of leaves	**0.012** $F_{1,19}$=7.68	0.353 $F_{4,19}$=1.18	0.753 $F_{1,19}$=0.10	**0.023** $F_{1,19}$=6.17
Height	**0.011** $F_{1,19}$=7.84	0.144 $F_{4,19}$=1.95	0.649 $F_{1,19}$=0.21	**0.013** $F_{1,19}$=7.60
Growth Rate	**0.002** $F_{1,16}$=13.50	**0.020** $F_{4,16}$=3.97	0.867 $F_{1,16}$=0.03	**<0.001** $F_{1,16}$=21.70
Number of Flowers	0.091 $F_{1,9}$=3.59	0.570 $F_{4,9}$=0.77	0.829 $F_{1,9}$=0.05	0.278 $F_{1,9}$=1.33
Total seed weight	0.603 $F_{1,7}$=0.30	0.594 $F_{4,7}$=0.74	0.709 $F_{1,7}$=0.15	0.788 $F_{1,7}$=0.08
Mean per-seed weight	0.558 $F_{1,7}$=0.38	0.343 $F_{4,7}$=1.34	0.767 $F_{1,7}$=0.10	0.822 $F_{1,7}$=0.05
Percent of aborted seeds	0.799 $F_{1,9}$=0.07	0.500 $F_{4,9}$=0.91	0.388 $F_{1,9}$=0.82	0.374 $F_{1,9}$=0.88
Seedset	0.470 $F_{1,9}$=0.57	**0.013** $F_{4,9}$=5.97	**0.034** $F_{1,9}$=6.22	0.257 $F_{1,9}$=1.46
Percent Survival	**0.026** $F_{1,19}$=5.84	0.500 $F_{4,19}$=0.87	0.102 $F_{1,19}$=2.96	**0.022** $F_{1,19}$=6.25
Total dried biomass	0.344 $F_{1,11}$=0.98	0.540 $F_{4,11}$=0.82	0.182 $F_{1,11}$=2.03	0.962 $F_{1,11}$=0.00
Fecundity	0.162 $F_{1,9}$=2.32	0.339 $F_{4,9}$=1.30	0.754 $F_{1,9}$=0.10	0.466 $F_{1,9}$=0.58
Lifetime fitness	0.308 $F_{1,9}$=1.17	0.133 $F_{4,9}$=2.34	0.895 $F_{1,9}$=0.02	0.566 $F_{1,9}$=0.35

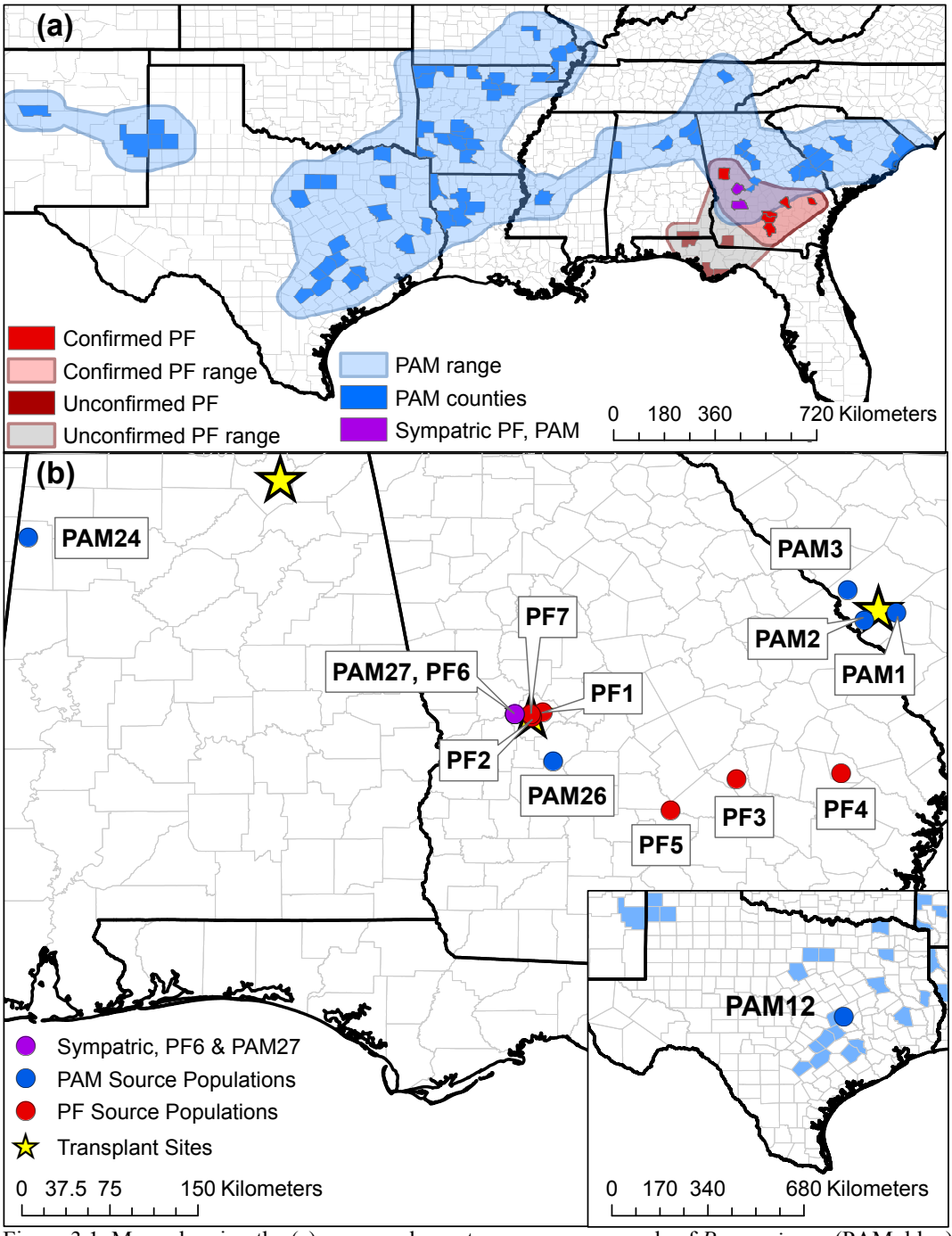


Figure 3.1. Maps showing the (a) range and county occurrence records of *P. americana* (PAM, blue) and confirmed and putative range and county occurrence records of *P. fimbriata* (PF, red), and (b) the locations of transplant sites (stars) and source populations of *P. americana* (PAM, blue) and *P. fimbriata* (PF, red). One location ("Sympatric, PAM27 & PF6"; purple), is the source location for both a *P. americana* population (PAM27) and a *P. fimbriata* population (PF6). Inset in (b) shows the location of the one *P. americana* source population from Texas.

Figure 3.2. Flowers of (a) *P. americana* and (b) *P. fimbriata.*

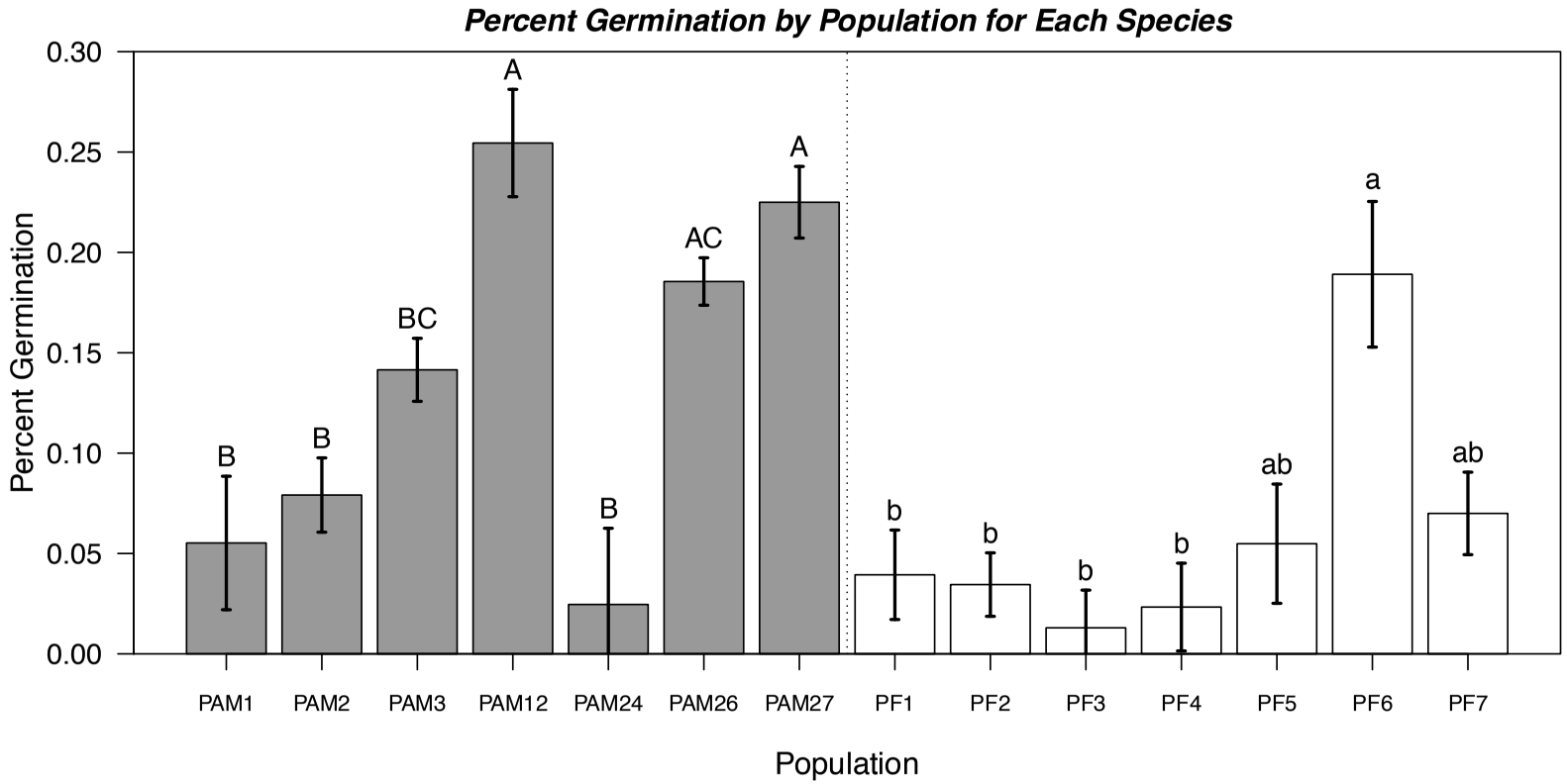


Figure 3.3. Percent germination by population for each species. *P. americana* populations are in gray, *P. fimbriata* populations are in white. Error bars represent standard error of the mean, and bars sharing the same letter (upper case for *P. americana* or lower case for *P. fimbriata*) are not significantly different from one another at the $p<0.05$ level.

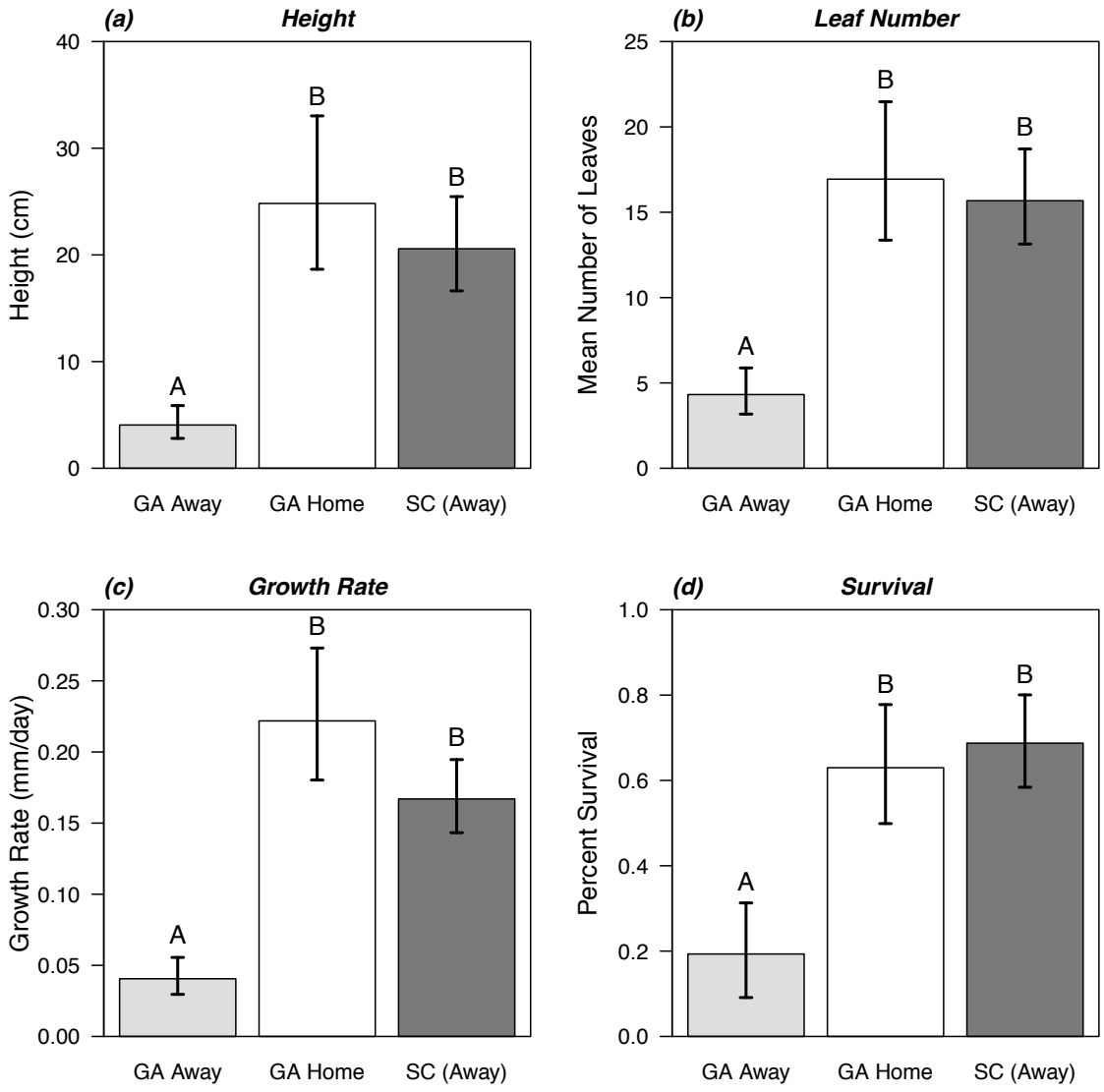


Figure 3.4. Home vs. away effects for traits in *P. fimbriata*. (a) Height (cm), (b) leaf number, (c) growth rate (mm/day), and (d) percent survival. Within a single plot/frame, bars sharing the same letter are not significantly different from one another at the $p<0.05$ level. X-axis designates plants that are grouped as home/away in the GA transplant site and away in the SC site. Error bars represent standard error of the mean.

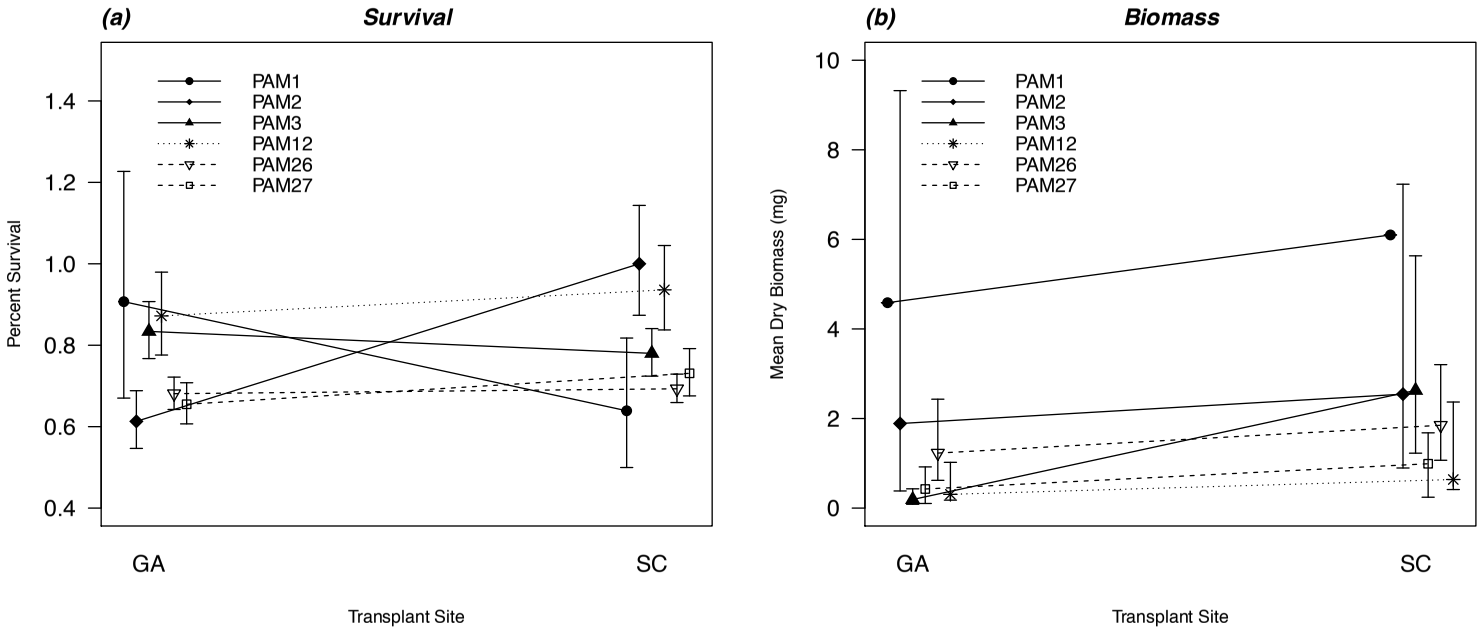


Figure 3.5. Reaction norms (both sets represent significant population by site interaction effect at the $p<0.05$ level) for (a) survival and (b) biomass in *P. americana*. Dashed lines with open symbols represent source populations from GA considered 'home' in the GA site, and solid lines with filled symbols represent source populations from SC that are considered 'home' in the SC site. The one dotted line with star symbols represents the population from TX (PAM12) that is considered 'away' at all sites. Error bars represent standard error of the mean.

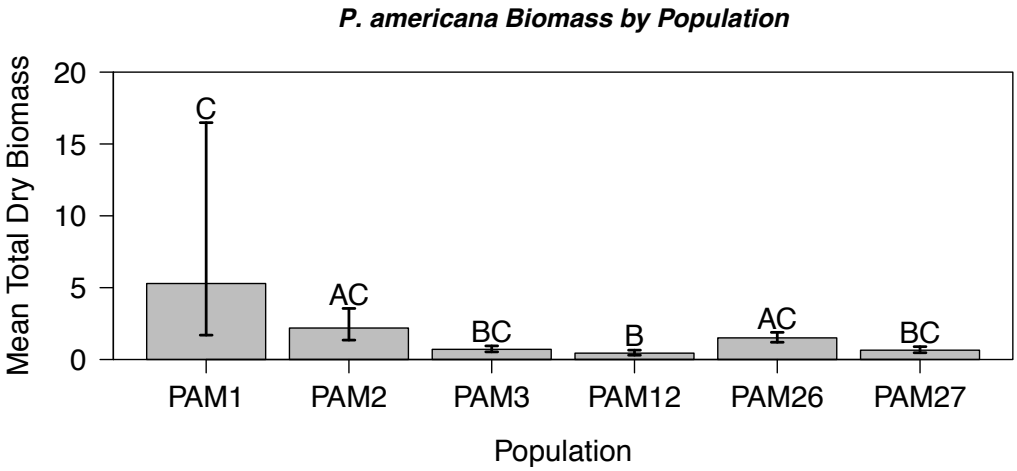


Figure 3.6. Total dry biomass (mg) of *P. americana* by population. Error bars represent standard error of the mean, and bars sharing the same letter are not significantly different from one another at the $p<0.05$ level.

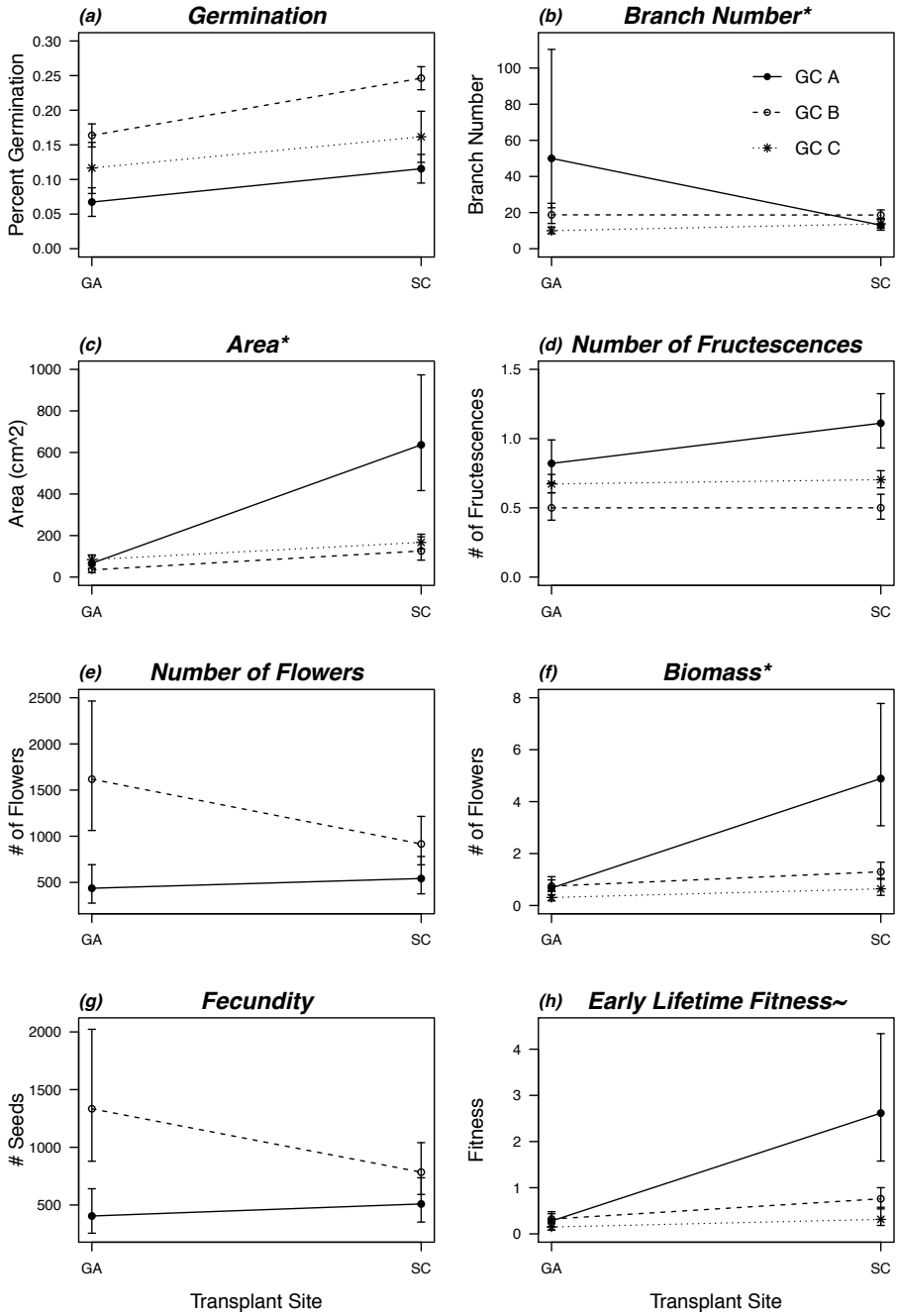


Figure 3.7. Reaction norms for genetic clusters (GCs) of *P. americana* for (a) germination and (b) branch number, (c) area, (d) number of fructescences, (e) number of flowers, (f) biomass, (g) fecundity, and (h) early lifetime fitness. Legend in (b) applies to all plots: GC A (populations from SC), GC B (populations from GA), and GC C (the AL and TX populations). Asterisks in title indicate significant GC by site interaction effects at the significance levels of $p<0.05$ (*); tilde represents marginal significance ($p<0.10$). Error bars represent standard error of the mean. Plots (e) and (g) only have two GCs, because GC C did not flower or set seeds.

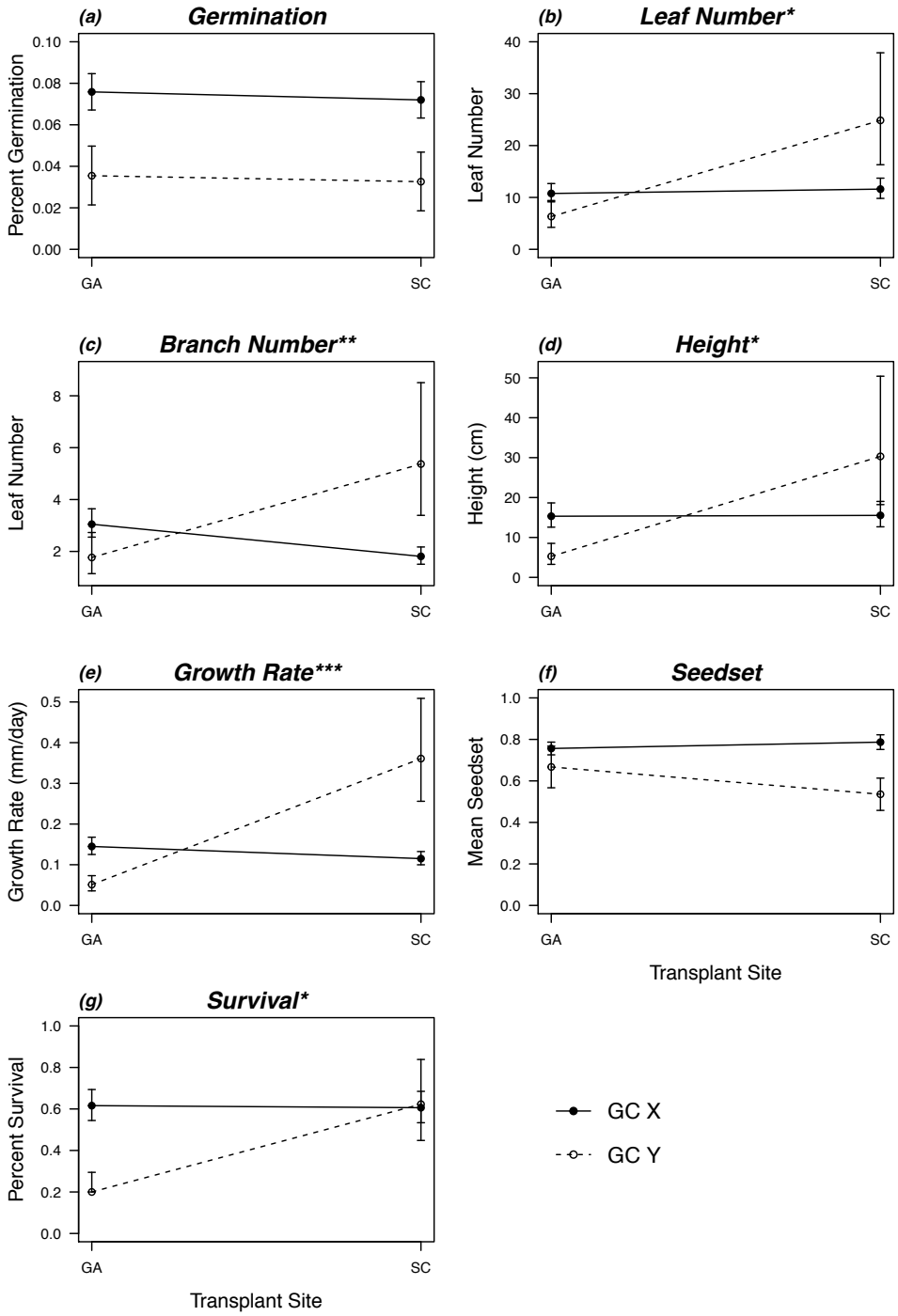


Figure 3.8. Reaction norms for genetic clusters (GCs) of *P. americana* for (a) germination and (b) leaf number, (c) branch number, (d) height, (e) growth rate, (f) seedset, and (g) percent survival. Legend at bottom applies to all plots: GC X (PF1, PF2, PF4, PF6, PF7), GC Y (PF3 and PF5). Asterisks in title indicate significant GC by site interaction effects at the levels of $p<0.05$ (*), $p<0.01$ (**), $p<0.001$ (***). Error bars represent standard error of the mean.

CHAPTER FOUR

THE ROLE OF GENETIC CONSTRAINTS IN THE RANGE LIMITS OF A NARROW ENDEMIC SPECIES AND ITS WIDESPREAD CONGENER IN *POLYGONELLA* (POLYGONACEAE)

Louisa G.C. Staton and Shu-Mei Chang

To be submitted to the journal of Molecular Ecology

Abstract

Determining what factors shape species' geographical distributions is important for making predictions about their evolutionary fates, particularly in the context of the changing climate and landscape due to human-mediated development. A failure to expand the geographic range into all suitable habitat is often attributed to failures in local adaptation. Therefore, a full understanding of what factors shape the range requires knowledge of population-level genetics. This project aims to investigate levels of genetic diversity and differentiation within populations of a narrow endemic annual plant (*Polygonella fimbriata*) and its widespread perennial congener (*P. americana*) in the southeastern United States. Previous work in this system has shown that allozyme diversity in widespread *Polygonella* species tends to be lower than that in narrowly distributed species. Using microsatellites designed for each of these two species, our results indicate that populations of the narrow *P. fimbriata* tend to have low amounts of within-population genetic diversity and polymorphism (), and population differentiation is highest among structure clusters. A putative long-distance dispersal event in this species eliminates any geographic signal to the structure of differentiation. Our evidence suggest that a model of evolutionary genetic constraints in conjunction with dispersal limitation best explains range limits in this species. In contrast, populations of the widespread *P. americana* exhibit higher levels of within-population genetic diversity and polymorphism (). There is a strong geographic signal to the patterns of differentiation among populations in this species, and limited gene flow between the eastern-most group of populations and the rest of the range suggests that these populations are on a divergent evolutionary trajectory.

Introduction

Understanding the abundance and distribution of species, and specifically principles controlling their geographic range limits, has historically been of interest to biologists and ecologists (Darwin, 1859; MacArthur & Wilson, 1967; Brown, 1984; Hubbell, 2001). Range limit theory is primarily described by two main sets of models: demographic and evolutionary genetic models (Moeller *et al.,* 2011). Demographic models incorporate the fluctuation of population dynamics (size, density, structure, growth rate) and how their impacts vary across heterogeneous versus homogeneous landscapes. In contrast, evolutionary genetic models are concerned with the genetic explanations of why populations fail or succeed at expanding beyond the range edge, including unidirectional gene flow and genetic paucity. Due to the timescale over which most demographic effects occur, range-limiting factors in this category can be hard to detect via field studies and experiments. Alternatively, the use of molecular genetic techniques can allow us to not only investigate factors of main influence in evolutionary genetic models, but also detect the signature of demographic history in many cases (Moeller *et al.*, 2011).

The specific outcomes of demographic events are dependent upon the frequency and rate of population dynamics (colonization/extinction, expansion/contraction, dispersal) and the ecological gradients over which they occur. However, regardless of the specific situations, some general genetic consequences in populations are expected if demographic effects are primarily maintaining a species' range. First, frequent extinction and recolonization should serve to reduce genetic variance within populations due to founder effects (Mayr, 1963). It should also decrease their effective population sizes,

leading to even further loss of genetic diversity (Glémin *et al*., 2003). Second, source-sink dynamics and metapopulation effects at the edge of the range should act to strongly reduce the amount of rare genetic variants within marginal populations, as these two forces are characterized by high rates of gene flow (Holt & Keitt, 2000; Moeller *et al.*, 2011). Additionally, as rare variants are lost to drift within these edge populations, the frequency of common allelic and sequence variants should increase (Kawecki, 2008). Also, under source-sink and metapopulation dynamics, genetic structure will be dictated by the mode of dispersal (McCauley, 1991). If colonists all come from a single, randomly selected parental population (as in the propagule-pool model), population differentiation (F_{ST}) is expected to be elevated (Wade & McCauley, 1988). In contrast, if founders are large in number and come from multiple parental populations (as in the migrant-pool model), then F_{ST} is expected to decrease (Wade & McCauley, 1988). Thus, the demography of a population can shape its genetic profile. By estimating pertinent population genetic parameters (e.g., effective population size, genetic diversity, F_{ST}), historical demographic signatures can be revealed.

The genetic characteristics of a population will also dictate its evolutionary and adaptive potential. This potential will subsequently determine a population's capacity to play a role in the expansion or limitation of a species' range. In particular, genetic variation and gene flow are critical in shaping the geographic range, as they will influence the degree of local adaptation achievable within a population (Kirkpatrick & Barton, 1997). Genetically diverse populations are more likely to have alleles that are favorable to a certain environment (Barton, 2001; Bridle & Vines, 2007). Selection may subsequently act upon this genetic variation to drive populations towards higher levels of

adaptation and ultimately specialization, assuming that effective population sizes are large enough to make the effects of genetic drift negligible. In contrast, populations with low genetic variation may lack the favorable alleles necessary for adaptation to their local environments (Kawecki, 2008). This is often the case in marginal populations with low effective population sizes, which lose important genetic variance due to drift (Kawecki, 2008).

Gene flow is critical in importing potentially favorable genetic variation into populations on the range edge. If marginal populations are small and harbor low amounts of genetic variation, the mutational input necessary to generate favorable alleles will often be lacking. When this is true, the genetic variants necessary for adaptation are more likely to arise in other, larger populations (Kawecki, 2008). These may then be delivered to genetically depauperate populations via gene flow (Al-Hiyaly *et al*., 1993; Kawecki, 2008). In these cases, the amount of gene flow will be crucial to determining the success of the marginal populations. Intermediate levels of gene flow, if they act to increase genetic variation and effective population size, can subsequently lead to increase the speed at which local adaptation occurs (Gomulkiewicz *et al*., 1999). Alternatively, high levels of gene flow may cause an influx of maladaptive alleles (Mayr, 1963; Lenormand, 2002; Bridle & Vines, 2007). Such factors bear great importance in the marginal populations that are at the distributional edge of a species (Bahn *et al*., 2006), whose adaptability determines the potential for range expansion. An assessment of gene flow alone will not elucidate what model best describes the dynamics driving species' range limits. Instead, population estimates of gene flow and genetic variation in tandem allow us to identify which range limits scenario is most fitting for a given species.

Comparing population genetic characteristics, thus, provides a strong test for the predictions from demographic and evolutionary genetic models of range limits. When assessed within populations of distributionally limited species, these comparisons can provide valuable insight into what is constraining or promoting range expansion. Additionally, comparisons among closely related geographically widespread and narrow species could help to explain further why narrowly restricted species have remained limited in their distributions. When combined with experimental investigations into local adaptation, we can gain valuable insight into the complex dynamics of range limits (Gaston, 2009).

For this study, we utilized microsatellite markers to investigate genetic variation, gene flow, and structure in populations of two closely related species: one geographically widespread (*Polygonella americana*) and one geographically narrow (*P. fimbriata*) species. In a previous study (Staton & Chang, unpublished; Ch. 3), we showed that levels of local adaptation within populations of these two species were generally low. In addition, we have observed little evidence of niche constraints maintaining the border of geographic range limits in these species (Staton & Chang, unpublished; Ch. 3). Based on these results, **we hypothesize that evolutionary genetic constraints (generated by small effective population sizes and high resultant levels of drift and/or inbreeding) are limiting expansion of the narrow species, *P. fimbriata***. If this is the case, we would expect to see significant reductions in heterozygosity within populations (Ho vs. He), high amounts of among-population differentiation (population pairwise-F_{ST}), and low frequencies of private alleles (unique variation). Alternatively, if we observe that populations of this narrow species are not genetically depauperate and detect moderate to

high levels of gene flow, **we would conclude that source-sink or metapopulation dynamics play a more significant role in maintaining this species' range**. In conjunction with these observations, the amount and direction of gene flow would dictate the frequency and location of private and common alleles among populations as well as their differentiation. Whether or not this gene flow is maladaptive would depend on the fitness of the populations involved. Also, the effects of gene flow on effective population size are dependent upon the fitness of the offspring (Holt & Gomulkiewicz, 1997), and neither are within the scope of this study. Finally, if we observe that populations are not genetically depauperate and are also gene flow among them is low or undetectable, then **we would conclude that dispersal limitation plays a more significant role in the range limitation of *P. fimbriata***. Because the widespread species in this study is not geographically limited, **we hypothesize that populations of *P. americana* are not limited by evolutionary genetic constraints**. If this is the case, we would expect higher amounts of genetic diversity (heterozygosity) and larger effective population sizes in comparison to *P. fimbriata*. Additionally, we would expect that any observed population differentiation should be driven largely by geographical factors (isolation by distance).

Materials and Methods

Study Species

Polygonella (Polygonaceae) is a small genus comprising eleven species, all of which are endemic to North America. Seven of these are geographically narrowly restricted to the southeastern United States (found in only 1-3 states, depending on the species), two are considered geographically intermediate in their ranges (spanning 6-7

states), and the remaining two are geographically widespread (found in 10 or more states). Species in this genus are hermaphroditic, dioecious, gynodioecious, or gynomonoecious (Horton, 1963; Hong & Smets, 2004), and there are both perennial and annual taxa in the genus. *Polygonella* is distinct from other members of the Polygonaceae in that branches are adnate to the stem, causing them to appear internodal, and ocrea are present. Previous work in this genus has demonstrated that, based on allozyme analyses, the two geographically widespread species (*P. americana* and *P. articulata*), have decreased within-population genetic variation when compared to the rest of the species in this genus (Lewis & Crawford, 1995). Additionally, a morphological phylogeny exists for this genus, but it has yet to be fully tested using molecular data (Lewis, 1991). A molecular phylogeny for the Polygoneae including eight species of *Polygonella* was constructed (Schuster *et al*., 2011), and differs significantly from the morphological phylogeny of Lewis & Crawford (1995).

This study utilizes populations of two species of *Polygonella*: *P. fimbriata* and *P. americana*, representing the varying range distributions found in the *Polygonella* genus (Figure 1). *P. fimbriata* is an annual, gynomonoecious plant (individuals have both hermaphroditic and female flowers on the same plant) that flowers from July to October. It is narrowly distributed in the southeast, and mainly restricted to the southern half of the state of Georgia with some putative populations in the panhandle of Florida. Alternatively, *P. americana* is a hermaphroditic, perennial plant, which flowers from June to October, and is widely geographically distributed throughout the southern and southwestern regions of the United States.

Tissue Collection, DNA Extraction

Populations for tissue collections were located using herbarium records from the University of Georgia and other universities in the southeast and through personal communications with local botanists. We collected leaf tissue from 20-30 maternal individuals of each of nine populations of *P. americana* and seven populations of *P. fimbriata* beginning in the fall of 2010 (Figure 1). Maternal individuals were spaced at least one meter apart to avoid collecting from highly related individuals. Additionally, we collected plant vouchers from each population and submitted them to the University of Georgia Herbarium. Upon tissue collection, we immediately stored leaf tissue in silica desiccant and kept leaf materials dry until extracting DNA. We extracted total genomic DNA from 20-25 individuals per population using the Qiagen DNeasy Plant Mini Kit (Qiagen, Valencia, CA, USA).

Microsatellite Genotyping

Microsatellites markers are appropriate for this type of study in that they provide estimates of genetic variation, gene flow, migration, and structure, can detect recent bottlenecks, and can be used to infer the relatedness of individuals (Spencer *et al*, 2000; Balloux & Lugon-Moulin, 2002; Selkoe & Toonen, 2006; Anderson *et al.*, 2010). Potential drawbacks of these markers have been thoroughly addressed in the literature, allowing evolutionary biologists to account for these complications in their analyses (Valdes *et al.*, 1993; Spencer *et al.*, 2000; Estoup *et al.,* 2002; Dakin & Avise, 2004; Dewoody *et al.*, 2006; Chapuis & Estoup, 2007; Väli *et al.*, 2008).

We sampled populations of these two species for genotyping at two species-specific sets of ten microsatellite loci for *P. americana* and eleven microsatellite loci for

P. fimbriata (Staton *et al.,* in prep). Loci were amplified using a 3-primer touchdown PCR. *P. americana* primers utilized in this study were PAMms8, PAMms11, PAMms20, PAMms30, PAMms32, PAMms38, PAMms42, PAMms54, PAMms56, and PAMms60. The *P. fimbriata* primers were PFms11, PFms13, PFms15, PFms23, PFms24, PFms26, PFms33, PFms35, PFms36, PFms39, and PFms41. For PCR reaction conditions, see Staton *et al*. (in prep). PCR products were analyzed on a 3730xl capillary sequencer (Applied Biosystems, Carlsbad, CA, USA) using a ROX-labeled internal size standard (GGF500R, Georgia Genomics Facility, Athens, GA, USA). We performed allele calls from chromatograms using Geneious v7.1.3 (Biomatters, San Francisco, CA, USA) and confirmed them by visual inspection. In total, 200 individuals from nine populations of *P. americana* were genotyped at ten microsatellite loci, and 169 individuals of seven populations of *P. fimbriata* were genotyped at ten microsatellite loci, six of which were polymorphic and are reported here.

Data Analyses

To assess genetic diversity, we calculated standard population genetic statistics for populations of each species separately using GenAlEx v6.501 (Peakall & Smouse, 2006) and FSTAT (Goudet, 1995). For each population, we generated the following statistics: mean number of individuals sampled per locus per population (N), estimated proportion of polymorphic loci (%P), mean estimated number of alleles at polymorphic loci (Na), mean number of effective alleles (NE), mean number of private alleles per population (PA), mean number of locally common alleles (frequency ≥5%) found in at least 50% of populations (CA), observed heterozygosity (Ho), expected heterozygosity (He), fixation index (F), and pairwise F_{ST} and gene flow matrices. For each locus, we also

estimated N, Na, Ne, Ho, He, and F as well as the mean estimated number of migrants per microsatellite locus. We estimated effective population size (Ne) using the molecular coancestry method of Nomura (2008), as implemented in NeEstimator v2.01 (Do *et al.*, 2014). The molecular coancestry method has been demonstrated to perform best in comparison to other Ne methods implemented for microsatellite data (Nomura, 2008) for inbred populations, and thus is suitable for our study system. Additionally, tests for significant deviations from Hardy-Weinberg equilibrium (HWE) per locus and significant excesses or deficiencies of heterozygotes per population were performed in GENEPOP (Raymond & Rousset, 1995).

To examine the distribution of genetic diversity across populations of each of the two species, we performed hierarchical analyses of molecular variance in GenAlEx by partitioning genetic variation within and among populations of each species separately (R_{ST}; Slatkin, 1995). In order to assess isolation by distance, we performed correlation tests in R (R Core Team, 2009), which compared the geographic (Euclidean) distance among populations and both the population pairwise conditional genetic distance (Dyer & Nason, 2004) and the pairwise population Cavalli-Sforza chord distance (Cavalli-Sforza & Edwards, 1967) corrected for the presence of null alleles (INA correction; Chapuis & Estoup, 2007).

To further assess the structuring of genetic data, we used the Bayesian clustering program InStruct (Gao *et al.*, 2007). InStruct is similar to the commonly used program STRUCTURE (Pritchard *et al.*, 2000), but is more robust to violations of Hardy-Weinberg assumptions, reduces false positive rates, and corrects for spurious admixture. Additionally, InStruct eliminates the assumption of Hardy-Weinberg equilibrium within

clusters and allows admixture and inbreeding. For these reasons, InStruct is suitable for *P. americana* and *P. fimbriata*, which comprise isolated, mixed-mating populations.

InStruct performs joint inference of population structure using a Markov Chain Monte Carlo (MCMC) algorithm, which probabilistically assigns individuals/populations to one of K subpopulations (clusters). For each species, we ran InStruct with K-values between one and nine for *P. americana* and one and seven for *P. fimbriata* (the number of sampled populations), using five independent MCMC chains, a 500000 burn-in and 500000 MCMC iterations with trimming every 100 generations, and estimating both admixture and the cluster's selfing rate. We re-ran each chain three times, and ran the full InStruct model a total of five times for each species in order to generate mean posterior log-likelihood, standard deviation, and delta K values for each level of K. To infer the optimal number of clusters, we used the delta K method of Evanno *et al*. (2005). Individual assignments to clusters were visualized using the program *distruct* (Rosenberg, 2004). Population assignments to clusters were visualized using ArcGIS v10.1 (ESRI, Redlands, CA, USA).

Because genetic structure and the relationships among populations vary over space and time, we wanted to incorporate a way to evaluate these characteristics that goes beyond approaches based on summary statistics (principle coordinates analysis, Mantel tests based on F_{ST}, etc.) and is more time and resource efficient than coalescence-based methods. To do this, we used PopGraph (Dyer & Nason, 2004), which is a multivariate graph-theoretic approach that performs analysis on populations simultaneously (as opposed to traditional pair-wise approaches). Generally speaking, PopGraph generates a matrix of among-population conditional genetic distances (cGD), which are the

conditionally independent genetic distances among populations. This matrix is then used to create a network of the populations whose topology is genetically informative. Each node represents a single population, the size of which is a representation of the relative within-population genetic variance. The length of the edges connecting each node corresponds to the cGD among each connected population. Additionally, connectivity among nodes is indicative of either contemporary or historical gene flow, isolated populations represent independently evolving entities, and closed loops within a network are the likely result of reticulate gene flow (Dyer & Nason, 2004). This program is useful for a priori hypothesis testing of principles like genetic divergence due to vicariance, and can also allow us to identify patterns (like a putative long-distance dispersal event and its origin) that otherwise might not be evident from traditional population genetic methods. We used this software to create population graphs for each of the two study species, and to perform separate tests of isolation by distance using both cGD and Cavalli-Sforza chord distance (Cavalli-Sforza & Edwards, 1967). Also, to visualize the evolutionary relationships of populations in *P. americana*, we used these chord distances to construct a neighbor-joining tree using the ape package in R.

Additionally, we estimated selfing rates in populations of *P. americana* and *P. fimbriata* using two separate approaches. Traditional approaches use mean population fixation indices (F_{IS}) to calculate selfing rates (Wright, 1965). This is a within-locus estimate based upon heterozygote deficiencies that is averaged across all loci within populations, and it is sensitive to misscoring, null alleles, inbreeding, and population substructure. Because this is the case, we opted to instead use an among-locus method of selfing rate estimation, robust multilocus estimation of selfing (RMES; David *et al*.

2007). This approach calculates selfing from a two-locus heterozygosity disequilibrium estimate ($\hat{g}_2$) based on the expectation that partial selfing will result in correlations in heterozygosity among different loci (identity disequilibrium). As a result, this estimate is F_{IS}-independent, and not sensitive to misscoring. However, RMES requires more than two loci that are heterozygous within a population, and based on the nature of our smaller data set in *P. fimbriata*, this was not the case for two of our populations. Therefore, we included a second measure, multilocus t and r (MLTR; Ritland, 2002), which is a within- and among-locus estimation of selfing. MLTR provides both a multi-locus selfing rate (s_m) and a single-locus selfing rate (s_s), the difference of which is an accurate measure of the amount of biparental inbreeding (sBPI) when the same loci are used for all populations within a species (Lu, 2000; Ritland, 2002).

Finally, we checked each locus for null alleles using FreeNA (Chapuis & Estoup, 2007). If loci had significant HWE deviations or frequencies of null alleles, we took several steps to assure the validity of our results. First, we ran analyses both with and without loci containing null alleles at frequencies >0.20. Second, we performed ENA corrections in FreeNA according to Chapuis & Estoup (2007) to calculate corrected Cavalli-Sforza chord distances for IBD analyses, to evaluate per-locus differences in F_{ST}, and to calculate subsequent population pairwise F_{ST} and gene flow matrices provided pre- and post-corrected F_{ST} values were significantly different from one another. Third, to validate our structure data, we performed Bayesian clustering analyses in Geneland (Guillot *et al*., 2008), which performs clustering analyses similarly to how STRUCTURE and InStruct perform, but is robust to the presence of null alleles. In the event that we

observed no differences in our analyses with and without null alleles, we proceeded with the full set of loci for each species.

Results

Genetic Diversity

Of the microsatellite loci utilized in this study, all ten loci sampled for *P. americana* were polymorphic (Figure 2). In contrast, five of the eleven loci for *P. fimbriata* were monomorphic across populations (PFms11, PFms15, PFms24, PFms39, and PFms41), and the remaining six loci had generally lower levels of polymorphism (Figure 3). Monomorphic loci were included in calculating descriptive statistics, but removed for InStruct and PopGraph analyses. For the 90 possible locus-by-population combinations in *P. americana*, more than half had null alleles at very low frequency (<0.05) or missing entirely. Additionally, 21 locus-by-population combinations had null alleles at intermediate frequencies (0.05 to 0.2), and the remaining 20 had null alleles at high frequencies (between 0.2 and 0.41). For the 42 possible locus-by-population combinations for *P. fimbriata*, 26 did not have null alleles or had them at low frequencies (<0.05), two had then at intermediate frequencies (between 0.05 and 0.2), and 14 had null alleles at high frequencies (between 0.2 and 0.75). Per-locus null allele frequencies can be found in Table 1. Overall, null alleles were observed at higher frequencies in *P. fimbriata* than in *P. americana.* It is likely that the signal of null alleles is biased upwards due to self-fertilization. We utilized null allele corrections suggested by Chapuis & Estoup (2007) for per-locus F_{ST} estimations, among-population F_{ST} and gene flow estimations, and for Cavalli-Sforza chord distances to assess isolation-by-distance in

populations for each of these species. However, because we had low sample sizes for polymorphic loci in *P. fimbriata*, all loci were retained for all other analyses. Additionally, for *P. americana* we performed analyses with and without null allele loci and observed no differences in our results other than decreased statistical power. For this reason, analyses for *P. americana* were performed upon the full data set as well, and InStruct and PopGraph results for both species were validated using Geneland.

Populations of *P. americana* had more polymorphic loci (78.9 ± 4.2%), more alleles per locus, higher observed and expected heterozygosity, and more private alleles than populations of *P. fimbriata* (Table 2). The average inbreeding coefficient for *P. americana* was negative (-0.130 ± 0.051), while for *P. fimbriata* it was positive (0.160 ± 0.118). Effective population sizes tended to be larger in *P. americana* than in *P. fimbriata*, but the confidence envelopes on these estimates were quite large (Table 2). Additionally, for several populations of both species (four in *P. americana* and two in *P. fimbriata*), the molecular coancestry method generated values of "infinity" for Ne. This method is based upon a parameter that assesses allele sharing among individuals. Estimates of infinity most likely are due to sampling error rather than genetic drift, and are a result of insufficient data and low sample sizes (Waples & Do, 2010).

Within *P. americana*, several populations exhibited a significant excess of heterozygotes (Ho vs. He, $p<0.05$; PAM2, PAM3, PAM26, and PAM27). Of these populations, all but PAM27 exhibited the lowest observed frequencies of private alleles (Table 2). In contrast, we observed highly significant heterozygote deficiencies within all populations of *P. fimbriata* sampled (Ho vs. He, $p<0.01$). Among the polymorphic loci for *P. fimbriata* (Table 1), we observe significant heterozygote deficiencies for four of

the six. These four loci also have the highest frequencies of null alleles within *P. fimbriata*. The remaining two either exhibit no difference in observed heterozygosity compared to expectations (PFms26) or have a significant excess of heterozygosity (PFms23). Additionally, half of the PF loci have He values of 0.5 or greater, while the remaining three have very low He values of 0.029-0.057, suggesting that while there were inconsistent patterns of Ho in sampled individuals, overall genetic diversity in these populations is low.

In addition, the populations of *P. fimbriata* sampled in this study showed moderate to high rates of selfing (s=0.373 to 1 depending on the estimation method; Table 3), which can increase the null allele signal. The population patterns of among-loci selfing estimates obtained from RMES were generally lower than those obtained from MLTR, and two populations (PF3 and PF4) did not have sufficient variation among loci to allow estimation of selfing using RMES. Additionally, population PF6 was estimated by RMES to be fully outcrossing ($\hat{s}=0.0 \pm 0.0$, $p=0.901$), while MLTR estimated it to have a moderately high selfing rate ($s_m=0.648$; Table 3). However, the RMES value was not significantly different from the null hypothesis ($\hat{s}=\hat{g}_2=0$), and may be due to the use of few loci. In contrast to *P. fimbriata*, populations of *P. americana* exhibited high outcrossing rates (s=0 to 0.402 depending on estimation method; Table 4). Estimates from RMES and MLTR were in better agreement for these populations with one exception (PAM11), likely as a result of greater loci sampling in *P. americana*. Additionally, biparental inbreeding appears to be the primary form of inbreeding observed in this species, whereas *P. fimbriata* was primarily selfing.

Genetic Structure and Gene Flow

For both *P. americana* and *P. fimbriata*, there is significant genetic differentiation among populations, and the majority of variation resides within populations (*P. americana* R_{ST}= 0.097, p=0.001; *P. fimbriata* R_{ST}=0.136, p=0.001). Population-pairwise F_{ST} analyses corrected for null alleles (ENA correction; Chapuis & Estoup, 2007) and subsequent gene flow estimates in *P. americana* reveal that there are two strongly differentiated groups of populations (PAM1-PAM2-PAM3 and PAM11-PAM12-PAM14-PAM24; Table 5). Within these two groups, differentiation is low and levels of gene flow are relatively high. These two groups are geographically distinct as well, with PAM1-3 confined to South Carolina, and PAM11-24 distributed from eastern Texas to western Alabama, and the differentiation among populations of these two groups reflects this. The remaining two populations, PAM26 and PAM27, are differentiated from the two aforementioned population groups and from each other, despite their relative geographic propinquity. Among *P. americana* populations, there was a moderate and marginally significant, effect of isolation by distance (Spearman's ρ=0.315, p=0.062), illustrated by the relationship between cGD and geographical distance (Figure 4a). This relationship became a strong and highly significant effect (Spearman's ρ=0.649, p<0.001) when considering genetic distance corrected for the presence of null alleles (Cavalli-Sforza chord distance with INA correction; Chapuis & Estoup, 2007; Figure 4b). In *P. fimbriata*, we also found two highly differentiated groups of populations (PF1-PF2-PF6-PF7-PF4 and PF3-PF5; Table 6), but this was not consistent with geographic location. For *P. fimbriata*, there was no significant isolation by distance effect when considering either cGD or chord distance corrected for null alleles (Figure 4c,d). Pairwise

matrices for cGD, corrected Cavalli-Sforza chord distance, and geographic distance for each species can be found in the appendix (Tables S1, S2).

The structure analyses from InStruct and PopGraph are consistent with the population F_{ST} data. For both *P. americana* and *P. fimbriata*, InStruct identified K=3 as the optimal number of clusters ($\Delta K=71.76$ and $\Delta K=38.04$, respectively; Table 7 and Figures 5 and 6). There is a very strong geographic signal among clusters of *P. americana* (Figure 5a), and little admixture within these clusters (Figure 5b). In contrast, there is not a clear geographic pattern to genetic structuring in *P. fimbriata*, and there are high amounts of admixture (Figure 6a,b). Because null alleles were prevalent in our data set, we checked our InStruct results against those from Geneland, which accounts for null alleles in the data set. Our results for *P. americana* are consistent for both analyses, so Geneland results for this species are not shown. Alternatively for *P. fimbriata*, when null alleles are taken into account, the optimal number of clusters is K=2 (Figure 7). While there still is, surprisingly, not a clear geographic signal in the distribution of these clusters, the posterior probability of assignment of each population to its cluster is 1.00 (PF1, PF2, PF6, PF7, and PF4 to cluster 1, and PF3 and PF5 to cluster 2). These latter structuring assignments are congruent with our results from PopGraph.

For *P. americana*, PopGraph analyses grouped populations into a single network comprising two connected subgraphs (Figure 8a). The main subgraph consists of the Texas, Alabama, and Georgia populations (PAM11, PAM12, PAM14, PAM24, PAM26, and PAM27), while the second subgraph is a cycle saturated for the three South Carolina populations (PAM1, PAM2, and PAM3). The two subgraphs are connected by gene flow between PAM3 and PAM26. Additionally, PAM14 seems to have the most within-

population genetic variation, while PAM27 has the lowest level of within-population genetic variation, relatively speaking (all of which is consistent with InStruct results). When placed within its geographic context, the population graph for *P. americana* demonstrates evidence of (likely historical) long-distance gene flow across the landscape (Figure 8b). The PopGraph results for *P. fimbriata* are most congruent with the structuring results from Geneland in that this species comprises two distinct networks that are not connected by gene flow (Figure 9a). The larger of the two population graphs is composed of the same populations that comprise cluster 1 of the Geneland clusters (PF1, PF2, PF6, PF7, and PF4), and is not saturated. The smaller population graph is composed of the same populations that make up cluster 2 of the Geneland clusters (PF3 and PF5), and due to its genetic isolation from the first population graph. When placed in its geographic context, this species' population graph reveals a pattern that is strongly suggestive of a long-distance dispersal and founder event (Figure 9b). Finally, a neighbor-joining tree constructed from the null-allele corrected Cavalli-Sforza chord distances among populations of *P. americana* suggests that there are two reciprocally monophyletic groups, one comprising populations from Texas and Alabama, and the other comprising populations from South Carolina and Georgia (Figure 10).

Discussion

Evidence for Population Genetic Constraints in P. fimbriata

Overall, our data support the hypothesis that populations of *P. fimbriata* face evolutionary genetic constraints. Under this hypothesis, we expected to observe low effective population sizes, significant reductions in heterozygosity, high population-

pairwise F_{ST}, and low frequencies of private alleles. Consistent with these expectations, we found first that populations of this species had very low estimated effective population sizes (between 1.0 - 4.1). These are on par with Ne estimates from a rare, biennial endemic plant based on census size (Königer *et al.*, 2012). Estimates of Ne for natural populations of endemic plant species based on molecular data are rare in the literature. Considering the importance of Ne for population fitness and survival (Newman & Pilson, 1997), there is a real need for more work in this area. Second, we observed significant low observed and expected genetic diversity (Ho, 0.095 – 0.167; He, 0.151 – 0.192) as well as low proportions of polymorphic loci (27.3 – 36.4%) across populations of *P. fimbriata*. These patterns are consistent with expectations for allozyme diversity in geographically restricted species (Karron, 1987; Hamrick *et al.*, 1991; Gitzendanner & Soltis, 2000) that are insect-pollinated and maintain low population sizes (Loveless & Hamrick, 1984), and are mixed-mating and have gravity dispersed seeds (Hamrick & Godt, 1996). Additionally, our findings are consistent with genetic diversity estimates for this species based on allozymes (Lewis & Crawford, 1995), as well as diversity estimates based on microsatellites for another narrow endemic member of the Polygonaceae family (*Eriogonum ovalifolium* var. *vineum*; Neel & Ellstrand, 2003). Third, private allele frequencies in populations of *P. fimbriata* were low. Most populations had only one private allele, though population PF3 had none and population PF4 had three. Moeller *et al.* (2011) found similar private allele frequencies in edge (0 private alleles) and intermediate/interior populations (≥3 private alleles) of an annual endemic plant that appears to adhere to an evolutionary genetic model of range limitation. Finally, while

population differentiation was not high among every individual population pairwise comparison, it was high among the predicted structure groups.

One caveat from our data is the presence of null alleles. Null alleles will cause populations to appear more homozygous (Chapuis & Estoup, 2007), and can complicate the interpretation of some of our results. Specifically, allelic diversity, assignment testing, heterozygosity estimates and measures typically based on heterozygosity (F_{ST}, gene flow, selfing rates) can be affected by the presence of null alleles. However, null alleles calculated from heterozygosity estimates (Chapuis & Estoup, 2007) might be inflated by selfing within populations since selfing generally leads to reduced heterozygosity (Chakraborty *et al.*, 1992). The high frequencies of null alleles that we observed in populations of *P. fimbriata* are, therefore, probably overestimated. Several reasons suggest that this is likely the case. First, our estimates of selfing rates are based on multilocus methods (citation here for the methods), which are robust to scoring errors that include null alleles. One method indicates that these populations have low to intermediate levels of selfing (RMES method, 0 - 0.446), only one value of which was significantly different from the null hypothesis ($\hat{s}$ is no different from zero). In scenarios where few loci are used to generate $\hat{g}_2$ estimates of selfing, it is possible to obtain nonsignificant underestimates compared to the true selfing rate (Szulkin *et al*., 2010). Because RMES estimates used four loci at most, they are likely less accurate than MLTR estimates. Selfing rates estimated in MLTR were moderate to high (0.648 - 1) for populations of *P. fimbriata*. Additionally, null alleles are expected to be more frequent in populations with large effective population sizes (Chapuis & Estoup, 2007), but we observed significantly

low Ne in populations of *P. fimbriata*. For these reasons, estimates of null alleles are probably overinflated within this species.

Among the *P. fimbriata* populations we studied, there are two main groups of populations between which there was strong differentiation and within which gene flow estimates were high. The smaller group (populations PF3 and PF5) is genetically isolated from the rest of the populations in this study, and these are likely on a path of evolutionary divergence (Dyer & Nason, 2004). Second, non-significant isolation by distance tests indicate that there is little geographic signal to the structuring of variation in *P. fimbriata* populations. This is attributable to the genetic connection between the eastern-most population (PF4) and the western cluster of populations (PF1, PF2, PF6, PF7) revealed in the analysis. Based on comparisons of diversity of PF4 to PF6 and PF7 (the two populations most closely connected to PF4 based on structure and gene flow estimates), the most parsimonious explanation for this pattern is the occurrence of a rare, long-distance dispersal (LDD) event. In comparison to PF6 and PF7, PF4 contains a smaller subset of genetic diversity, is less genetically differentiated from these two populations, and is estimated to be entirely selfing. This pattern is often observed from founder effects (Mayr, 1963), and fits the ecological explanations for the persistence of selfing in plant populations (Baker's Law and reproductive assurance; Stebbins, 1957; Jarne & Charlesworth, 1993). Additionally, selfing in populations with reduced effective population sizes may act to decrease the costs of inbreeding (Lande & Schemske, 1985). While LDD events in this species are unexpected given the seed dispersal mechanism (gravity dispersal), rare LDD events have been observed in plants before, particularly in the colonization of islands (Ridley, 1930; Carlquist, 1967). Additionally, the importance

of LDD events, effected via nonstandard dispersal mechanisms including extreme weather events such as tropical storms (Higgins *et al.*, 2003; Nathan, 2006), has been stressed with regards to achieving range expansion (Cain *et al.*, 2000; Nathan & Muller-Landau, 2000; Nathan, 2006). While there is not enough information to make conjectures regarding the nature of this suspected LDD event in *P. fimbriata*, the landscape in the southeastern U.S. is characterized by frequent tropical storms that enter through the Gulf of Mexico and travel east to northeast up the coast, and occur during the flowering and fruiting time in these plants.

Evidence for Greater Genetic Diversity in P. americana

Overall, our data showed no evidence that populations of the widely distributed *P. americana* are under population genetic constraints. In comparison to *P. fimbriata*, populations of this species had significantly higher levels of polymorphism (%P) and genetic diversity (Ho and He), and larger effective population sizes. Additionally, these populations exhibited high levels of outcrossing, which can explain the occurrence of excess heterozygosity observed in some of these populations. The levels of genetic diversity detected here are consistent with those found in widespread, hermaphroditic, perennial plant species based on allozyme data (Loveless & Hamrick, 1984; Karron, 1987; Hamrick *et al.*, 1991; Hamrick & Godt, 1996; Gitzendanner & Soltis, 2000), as well as diversity estimates based on microsatellites from a geographically widespread perennial member of the Polygonaceae family (*Rheum tanguticum*; Chen *et al*., 2009).

Our findings are, however, in contrast to previous findings in this system (Lewis & Crawford, 1995). In their allozyme survey of genetic diversity in species of

Polygonella, Lewis & Crawford (1995) observed lower levels of genetic diversity in widespread species in contrast to narrow endemics. Specifically, *P. americana* had lower average genetic diversity and proportions of polymorphic loci across populations (D=0.0166, %P=14.5, respectively; five populations, mean number of loci =8.2) in comparison to *P. fimbriata* (D=0.0819, %P=34.8; three populations, mean number of loci =11.3). They posited that decreased diversity in widespread species was a byproduct either of large-scale migrations during glaciation or inbreeding depression due to increased rates of selfing. The recovery of genetic diversity in populations of *P. americana* in the past twenty years would require high mutation and migration rates, large effective population sizes, and strong selection (Barrett & Kohn, 1991). This is unlikely, given that dispersal distances for this species are low and strong selection in nature is rare (Kingsolver *et al.*, 2001). While rapid recovery from bottleneck events has been documented in some animal species (e.g. Keller *et al.*, 2001; Colson & Hughes, 2004), we are unaware of any examples of such rapid recovery in perennial plant species. Alternatively, the contradictory results observed in our study can potentially be better explained by two things: population sampling and/or choice of genetic marker, the more plausible of which is the latter. First, the five populations of *P. americana* sampled by Lewis & Crawford (1995) are different from those in our study. There is only one ("Aiken, SC") that is potentially the same as one of our populations (PAM2 and PAM3 are from Aiken County, SC), but we cannot confirm this without exact coordinates from the allozyme study. Also, the Aiken population in the Lewis & Crawford study was not polymorphic for any of the allozymes sampled, whereas PAM2 and PAM3 were highly polymorphic and diverse. Instead, the use of different markers is the best explanation for

differences between these two studies. Due to their high mutation rates, microsatellites are more variable and polymorphic than allozymes (Spencer *et al.*, 2000; Li *et al.*, 2002). There are several examples where microsatellites have demonstrated high levels of heterozygosity in contrast to allozymes used in the same species (e.g. Hughes & Queller, 1993; Spencer *et al*., 1999).

The strong structuring of diversity observed across populations of *P. americana* is best explained by a significant isolation-by-distance effect. Among structure clusters, there is evidence for very low levels of gene flow, as demonstrated by the presence of few putative migrants in the individual structure assignments. The disjunct distribution and dispersal limited nature of these populations suggests that this gene flow is likely historical, although we do not have definitive evidence for that. Additionally, the strong geographic signal to clustering, the connectivity patterns in our population graph, and the identity of migrants (or individuals with migrant ancestry) in populations suggest that gene flow has occurred in an east to west fashion. More specifically, this pattern emerges when considering the structure bar plot of individual assignments (Figure 5b)—individuals with "eastern heritage" (yellow and blue clusters) show up within populations that are to the west, but individuals with "western heritage" (green clusters) do not appear in populations to the east. Finally, the strong differentiation among structure clusters and the single edge in the population graph (Figure 8) connecting the eastern populations and the rest of the group suggest that these eastern populations are diverging. Generally, plant population divergence is most common across latitudinal and elevational gradients (e.g. Eo *et al.,* 2008; Liu *et al.*, 2014). Due to the alignment of biogeographical features in North America, elevational gradients in the southern U.S. do not occur longitudinally.

However, a comparative phylogeography study of the southern U.S. demonstrated that phylogeographic breaks in both plant and animal taxa tend to be congruent with major river systems in this region (Soltis *et al.*, 2006). If one considers the evolutionary relationships of these populations (Figure 10), there is a break between the eastern group (SC and GA populations) and the western group (AL and TX populations). One hypothesis that could explain this break is that the Mississippi river has acted as a vicariant barrier between these two groups of populations. A subsequent LDD event could explain the presence of west-of-the-Mississippi genotypes occurring to the east of the river. This seems even more plausible when considering the range map of *P. americana* and the proximity of presence records to the Mississippi river (Figure 1). To test this hypothesis, more thorough population sampling from the full range of this species is necessary.

Implications for Range Limits in P. fimbriata

It has been widely demonstrated that understanding the dynamics at play in marginal populations is among the most important tasks in understanding range limits, because this is where expansion of the range usually occurs (Guo *et al.*, 2005; Bridle & Vines, 2007; Anderson *et al.*, 2010; Moeller *et al*., 2011). In dispersal-limited species, like those of the *Polygonella* genus, assessments of population-level genetic constraints are important for understanding why populations at the edge sometimes fail to colonize further beyond the border (Geber, 2008; Kawecki *et al*., 2008; Hargreaves *et al.*, 2014). A complication of this approach, however, is the definition of populations as marginal versus core. In species that are narrow and disjunct in their distribution, like *P. fimbriata*,

there may not be any apparent core region. Human-mediated fragmentation of the landscape and the subsequent extirpation of populations from the historical range can lead to the misclassification of populations as being edge or interior. For this reason, we made no *a priori* assumptions regarding the margin/core status of *P. fimbriata* populations in this study. Instead we treat each population as having the potential to expand the range in any direction. By doing so, we can assess the capacity of each population for its putative role in range expansion/limitation of the species.

All populations of *P. fimbriata* in this study exhibited low expected levels of genetic diversity compared to its widespread congener, but consistent with other narrow endemics (Hamrick *et al.*, 1991; Gitzendanner & Soltis, 2000; Neel & Ellstrand, 2003). Also, we observed significant heterozygote deficiencies and small effective population sizes. Rates of gene flow were relatively high among some populations (specifically within genetic cluster assignments), but low among others. Populations PF3 and PF5 exhibited some of the lowest genetic diversity, had few (or no, for PF3) unique genetic variants, and there was no evidence of gene flow into these populations from outside of the cluster. To our knowledge, there are no other populations of *P. fimbriata* located geographically between those that we collected for this study. This suggests that these two populations are too genetically depauperate to participate in expansion of the range. In contrast, while genetic diversity is still considerably low in the remaining *P. fimbriata* populations sampled in this study, increased levels of gene flow among these populations may increase the rate of spread of favorable novel genetic variants if they were to arise.

In conclusion, our data for *P. fimbriata* are consistent with an evolutionary genetics model of range limits. This, in combination with dispersal-limitations, best

explains the maintenance of range-limits in *P. fimbriata*. In contrast, populations of *P. americana* show overall higher levels of genetic diversity, but strong structuring of this diversity suggests that regional divergence is occurring. Additionally, evolutionary relationships between western and eastern suggest the presence of a phylogeographic break, but much thorough sampling is needed to test this. As available habitat for these species becomes rarer due to landscape development and climate change, considerations of these population genetic traits will be important for making predictions regarding range shifts. Additionally, these characteristics should be carefully considered in the preservation of *P. fimbriata*, which may be at substantial risk due to low population numbers, effective population sizes, and strong evidence for decreased heterozygosity due to inbreeding.

Literature Cited

Al-Hiyaly, S.A.K., McNeilly, T., Bradshaw, A.D. & Mortimer, A.M. (1993) The effect of zinc contamination from electricity pylons. Genetic constraints on selection from zinc tolerance. *Heredity*, **70**, 22-32.

Anderson, C.D., Epperson, B.K., Fortin, M.-J., Holderegger, R., James, P.M.A., Rosenberg, M.S., Scribner, K.T & Spear, S. (2010) Considering spatial and temporal scale in landscape genetic studies of gene flow. *Molecular Ecology*, **19**, 2565-75.

Bahn, V., O'Connor, R.J. & Krohn, W.B. (2006) Effect of dispersal at range edges on the structure of species ranges. *Oikos*, **115**, 89-96.

Balloux, F. & Lugon-Moulin, N. (2002) The estimation of population differentiation with microsatellite markers. *Molecular Ecology*, **11**, 155-165.

Barrett, S.C.H. & Kohn, J.R. (1991) Genetic and evolutionary consequences of small population size in plants: Implications for conservation. *Genetics and Conservation of Rare Plants* (Eds. D.A. Falk and K.E. Holsinger) Oxford University Press, New York, NY.

Barton, N.H. (2001) Adaptation at the edge of a species' range. *Integrating ecology and evolution in a spatial context*. (Eds. J. Silvertown & J. Antonovics) Blackwell Science, Oxford, U.K.

Bridle, J.R. & Vines, T.H. (2007) Limits to evolution at range margins: when and does adaptation fail? *Trends in Ecology and Evolution*, **22**, 140-7.

Brown, J.H. (1984) On the relationship between abundance and distribution of species. *The American Naturalist*, **124,** 255-79.

Cain, M.L., Milligan, B.G. & Strand, A.E. (2000) Long-distance seed dispersal in plant populations. *American Journal of Botany*, **87**, 1217-27.

Carlquist, S. (1967) The biota of long-distance dispersal. V. Plant dispersal to Pacific islands. *Bulletin of the Torrey Botanical Club*, **94,** 129-62.

Cavalli-Sforza, L.L. & Edwards, A.W. (1967) Phylogenetic analysis. Models and estimation procedures. *American Journal of Human Genetics*, **19,** 233-57.

Chakraborty, R., DeAndrade, M., Daiger, S.P. & Budowle, B. (1992) Apparent heterozygote deficiencies observed in DNA typing data and their implications in forensic applications. *Annals of Human Genetics*, **56**, 45-57.

Chapuis, M.-P. & Estoup, A. (2007) Microsatellite null alleles and estimation of population differentiation. *Molecular Biology and Evolution*, **24**, 621-31.

Chen, F., Wang, A., Chen, K., Wan, D. & Liu, J. (2009) Genetic diversity and population structure of the endangered and medically important *Rheum tanguticum* (Polygonaceae) revealed by SSR markers. *Biochemical Systematics and Ecology*, **37**, 613-21.

Colson, I. & Hughes, R.N. (2004) Rapid recovery of genetic diversity of dogwhelk *Nucella lapillus* L.) populations after local extinction and recolonization contradicts predictions from life-history characteristics. *Molecular Ecology*, **13**, 2223-33.

Dakin, E.E. & Avise, J.C. (2004) Microsatellite null alleles in parentage analysis. *Heredity*, **93**, 504-9.

Darwin, C. (1859) *On the origins of species by means of natural selection*. Jon Murray, London.

David, P., Benoît, P., Viard, F., Castella, V. & Goudet, J. (2007) Reliable selfing rate estimates from imperfect population genetic data. *Molecular Ecology*, **16**, 2474-2487.

Dewoody, J., Nason, J.D. & Hipkins, V.D. (2006) Mitigating scoring errors in microsatellite data from wild populations. *Molecular Ecology Notes*, **6**, 951-7.

Do, C., Waples, R.S., Peel, D., Macbeth, G.M., Tillett, B.J. & Ovenden, J.R. (2014) NeEstimator V2: re-implementation of software for the estimation of contemporary effective population size (Ne) from genetic data. *Molecular Ecology Resources*, **14**, 209-14.

Dyer, R.J. & Nason, J.D. (2004) Population Graphs: the graph theoretic shape of genetic structure. *Molecular Ecology*, **13,** 1713-1727.

Eo, S.H., Wares, J.P. & Carroll, J.P. (2008) Population divergence in plant species reflects latitudinal biodiversity gradients. *Biology Letters*, **4**, 382-84.

Estoup, A., Jarne, P. & Cornuet, J.-M. (2002) Homoplasy and mutation model at microsatellite loci and their consequences for population genetic analysis. *Molecular Ecology*, **11**, 1591-1604.

Evanno, G., Regnaut, S. & Goudet, J. (2005) Detecting the number of cluster of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, **14,** 2611-20.

Gao, H., Williamson, S. & Bustamante, C.D. (2007) A Markov Chain Monte Carlo approach for joint inference of population structure and inbreeding rates from multilocus genotype data. *Genetics*, **176,** 1635-51.

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

Geber, M.A. (2008) To the edge: studies of species' range limits. *New Phytologist*, **178**, 228-30.

Gitzendanner, M.A. & Soltis, P.A. (2000) Patterns of genetic variation in rare and widespread plant congeners. *American Journal of Botany*, **87**, 783-92.

Glémin, S. (2003) How are deleterious mutations purged? Drift versus nonrandom mating. *Evolution*, **57,** 2678-87.

Gomulkiewicz, R., Holt, R.D. & Barfield, M. (1999) The effects of density dependence and immigration on local adaptation and niche evolution in a black-hole sink environment. *Theoretical Population Biology*, **55**, 283-96.

Goudet, J. (1995) FSTAT version 1.2: a computer program to calculate Fstatistics. *Journal of Heredity*, **86,** 485-6.

Guillot, G., Santos, F. & Estoup, A. (2008) Analysing georeferenced population genetics data with Geneland: a new algorithm to deal with null alleles and a friendly graphical user interface. *Bioinformatics*, **24**, 1406-7.

Guo, Q., Taper, M., Schoenberger, M. & Brandle, J. (2005) Spatial-temporal population dynamics across species ranges: from centre to margin. *Oikos*, **108**, 47-57.

Hamrick, J.L. & Godt, M.J.W. (1996) Effects of life history traits on genetic diversity in plant species. *Philosophical Transactions: Biological Sciences*, **351**, 1291-98.

Hamrick, J.L., Godt, M.J.W., Murawski, D.A. & Loveless, M.D. (1991) Correlations between species traits and allozyme diversity: Implications for conservation biology. *Genetics and Conservation of Rare Plants* (Eds. D.A. Falk and K.E. Holsinger) Oxford University Press, New York, NY.

Hargreaves, A.L, Samis, K.E. & Eckert, C.G. (2014) Are species' range limits simply niche limits writ large? A review of transplant experiments beyond the range. *The American Naturalist,* **183,** 157-73.

Higgins, S.I., Nathan, R. & Cain, M.L. (2003) Are long-distance dispersal events in plants usually caused by nonstandard means of dispersal? *Ecology*, **84,** 1945-56.

Holt, R.D. & Gomulkiewicz, R. (1997) How does immigration influence local adaptation? A reexamination of a familiar paradigm. *The American Naturalist*, **149,** 563-72.

Holt, R.D. & Keitt, T.H. (2000) Alternative causes for range limits: a metapopulation perspective. *Ecology Letters*, **3**, 41-47.

Hong, S.P. & Smets, E.F. (2004) What is the taxonomic status of *Polygonella*? Evidence from floral morphology. *Annals of the Missouri Botanical Gardens*, **91,** 320-45.

Horton, J.H. (1963) A taxonomic revision of *Polygonella* (Polygonaceae). *Brittonia,***15,** 177-203.

Hubbell, S.P. (2001) *The unified neutral theory of biodiversity and biogeography*. Princeton University Press, Princeton.

Hughes, C.R. & Queller, D.C. (1993) Detection of highly polymorphic microsatellite loci in a species with little allozyme polymorphism. *Molecular Ecology*, **2**, 131-7.

Jarne, P. & Charlesworth, D. (1993) The evolution of the selfing rate in functionally hermaphrodite plants and animals. *Annual Review in Ecology and Systematics*, **24**, 441-66.

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.

Kawecki, T.J. (2008) Adaptation to marginal habitats. *Annual Review in Ecology, Evolution, and Systematics*, **39**, 321-42.

Keller, L.F., Jeffrey, K.J., Arcese, P., Beaumont, M.A., Hochchka, W.M., Smith, J.N.M. & Bruford, M.W. (2001) Immigration and the ephemerality of a natural population bottleneck: evidence from molecular markers. *Proceedings of the Royal Society B: Biological Sciences*, **268**, 1387-94.

Kingsolver, J.G., Hoekstra, H.E., Hoekstra, J.M., Berrigan, D., Vignieri, S.N., Hill, C.E., Hoang, A., Gibert, P. & Beerli, P. (2001) The strength of phenotypic selection in natural populations. *The American Naturalist*, **157**, 245-61.

Kirkpatrick, M. & Barton, N.H. (1997) Evolution of a species' range. *The American Naturalist*, **150**, 1-23.

Königer, J., Rebernig, C.A., Brabec, J., Kiehl, K. & Greimler, J. (2012) Spatial and temporal determinants of genetic structure in *Gentianella bohemica*. *Ecology and Evolution*, **2**, 636-48.

Lande, R. & Schemske, D.W. (1985) The evolution of self-fertilization and inbreeding depression in plants. 1. Genetic models. *Evolution*, **29**, 24-40.

Lenormand, T. (2002) Gene flow and the limits to natural selection. *Trends in Ecology and Evolution*, **17**, 183-9.

Lewis, P.O. (1991) *Allozyme variation and evolution in Polygonella (Polygonaceae)*. Dissertation, Ohio State University.

Lewis, P.O. & Crawford, D.J. (1995) Pleistocene refugium endemics exhibit greater allozymic diversity than widespread congeners in the genus *Polygonella* (Polygonaceae). *American Journal of Botany,* **82**, 141-9.

Li, Y.C., Korol, A.B., Fahima, T., Beiles, A. & Nevo, E. (2002) Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review. *Molecular Ecology*, **11**, 2453-2465.

Liu, C., Tsuda, Y., Shen, H., Hu, L., Saito, Y. & Ide, Y. (2014) Genetic structure and hierarchical population divergence history of *Acer mono* var. *mono* in south and northeast China. *PLoS ONE*, **9**, e87187.

Loveless, M.D. & Hamrick, J.L. (1984) Ecological determinants of genetic structure in plant populations. *Annual Review of Ecology and Systematics*, **15**, 65-95.

Lu, Y.Q. (2000) Effects of density on mixed mating systems and reproduction in natural populations of *Impatiens capensis*. *International Journal of Plant Sciences*, **161**, 671-81.

MacArthur, R.H. & Wilson, E.O. (1967) *The theory of island biogeography*. Princeton University Press, Princeton.

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

McCauley, D.E. (1991) Genetic consequences of local population extinction and recolonization. *Trends in Ecology and Evolution*, **6**, 5-8.

Moeller, D.A., Geber, M.A. & Tiffin, P. (2011) Genetics and the evolution of geographic range limits in an annual plant. *The American Naturalist*, **178**, S44-S57.

Nathan, R. (2006) Long-distance dispersal of plants. *Science*, **313**, 786-8.

Nathan, R. & Muller-Landau, H.C. (2000) Spatial patterns of seed dispersal, their determinants and consequences for recruitment. *Trends in Ecology and Evolution*, **15**, 278-85.

Neel, M.C. & Ellstrand, N.C. (2003) Conservation of genetic diversity in the endangered plant *Eriogonum ovalifolium* var. *vineum* (Polygonaceae). *Conservation Genetics*, **4**, 337-52.

Newman, D. & Pilson, D. (1997) Increased probability of extinction due to decreased genetic effective population size: experimental populations of *Clarkia pulchella*. *Evolution*, **51**, 354-62.

Nomura, T. (2008) Estimation of effective number of breeders from molecular coancestry of single cohort sample. *Evolutionary Applications*, **1**, 462-74.

Peakall, R. & Smouse, P.E. (2006) GENALEX 6: Genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, **6**, 288-95.

Pritchard, J.K., Stephens, M. & Donnelly, P. (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945-59.

R Development Core Team. (2009) R: A Language and Environment for Statistical Computing. The R Foundation for Statistical Computing. Vienna, Austria. http://www.R-project.org

Raymond, M. & Rousset, F. (1995) GENEPOP (version 1.2): population genetic software for exact tests and ecumenicism. *Journal of Heredity*, **86,** 248-9.

Ridley, H.N. (1930) *The dispersal of plants throughout the world*. Reeve & Co. Publishing, Ashford, U.K.

Ritland, K. (2002) Extensions of models for the estimation of mating systems using *n* independent loci. *Heredity*, **88**, 221-28.

Rosenberg, N.A. (2004) *Distruct*: a program for the graphical display of population structure. *Molecular Ecology Notes*, **4**, 137-8.

Schuster, T.M., Reveal, J.L. & Kron, K.A. (2011) Phylogeny of Polygoneae (Polgonaceae: Polygonoideae). *Taxon*, **60,** 1653-66.

Selkoe, K.A. & Toonen, R.J. (2006) Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecology Letters*, **9**, 615-29.

Silvertown, J.W. & Doust, J.L. (1993) *Introduction to plant population biology*. Blackwell Scientific Publications, Oxford, U.K.

Slatkin, M. (1995) A measure of population subdivision based on microsatellite allele frequencies. *Genetics*, **139,** 457-62.

Spencer, C.C., Chlan, C.A., Neigel, J.E., Scribner, K.T., Wooten, M.C. & Leberg, P.L. (1999) Polymorphic microsatellite markers in the western mosquitofish, *Gambusia affinis*. *Molecular Ecology*, **8**, 157-8.

Spencer, C.C., Neigel, J.E. & Leberg, P.L. (2000) Experimental evaluation of the usefulness of microsatellite DNA for detecting demographic bottlenecks. *Molecular Ecology*, **9**, 1517-28.

Stebbins, G.L. (1957) Self-fertilization and population variability in the higher plants. *American Naturalist*, **91**, 337-54.

Szulkin, M., Bierne, N. & David, P. (2010) Heterozygosity-fitness correlations: a time for reappraisal. *Evolution*, **64-5**, 1202-17.

Valdes, A.M., Slatkin, M. & Freimer, N.B. (1993) Allele frequencies at microsatellite loci: the stepwise mutation model revisited. *Genetics*, **133**, 737-49.

Väli, Ü., Einarsson, A., Waits, L. & Ellegren, H. (2008) To what extent do microsatellite markers reflect genome-wide diversity in natural populations? *Molecular Ecology*, **17**, 3808-17.

Wade, M.J. & McCauley, D.E. (1988) Extinction and recolonization: their effects on the genetic differentiation of local populations. *Evolution*, **42**, 995-1005.

Waples, R.S. & Do, C. (2010) Linkage disequilibrium estimates of contemporary Ne using highly variable genetic markers: a largely untapped resource for applied conservation and evolution. *Evolutionary Applications*, **3**, 244-62.

Wright, S. (1965) The interpretation of population structure by F-statistics with special regard to systems of mating. *Evolution*, **19**, 395-420.

Table 4.1. Summary statistics for polymorphic microsatellite loci across all populations of *P. americana* (nine) and *P. fimbriata* (seven). Abbreviations: N, mean number of individuals sampled per locus per population; Na, mean estimated number of alleles at polymorphic loci; NE, number of effective alleles; Ho, observed heterozygosity; He, expected heterozygosity; F, fixation index; F_{ST}, population genetic differentiation corrected for null alleles (ENA correction; Chapuis & Estoup 2007); Nm, estimated number of migrants per generation based on corrected F_{ST} values; and NF, the frequency of null alleles per locus. Stars on Ho indicate a significant excess of heterozygotes in the case of *P. americana* loci and significant heterozygote deficiency for *P. fimbriata* loci (* $p<0.05$, ** $p<0.01$, *** $p<0.001$).

Species/Loci	N	Na	NE	Ho	He	F	F_{ST}	Nm	NF
P. americana	21.466 (0.307)	3.044 (0.237)	1.958 (0.146)	0.383 (0.036)	0.334 (0.028)	-0.126 (0.051)	0.283 (0.074)	2.146 (1.003)	0.125 (0.02)
PAMms8	21.888	1.111	1.004	0.004	0.004	-0.021	0.140	1.535	0.179
PAMms11	22.000	2.111	1.095	0.061	0.083	0.220	0.095	2.373	0.121
PAMms20	22.111	2.222	1.302	0.235*	0.167	-0.060	0.674	0.121	0.087
PAMms30	21.777	4.666	2.097	0.466	0.487	0.097	0.202	0.986	0.088
PAMms32	21.555	3.777	1.583	0.267	0.289	0.043	0.109	2.036	0.116
PAMms38	22.000	4.333	3.24	0.415	0.469	0.015	0.500	0.250	0.089
PAMms42	18.444	2.444	1.926	0.786	0.479	-0.630	0.022	10.943	0.197
PAMms54	21.888	3.111	2.603	0.837***	0.590	-0.491	0.152	1.397	0.065
PAMms56	21.555	5.222	3.435	0.549	0.600	-0.030	0.233	0.834	0.094
PAMms60	21.444	1.444	1.293	0.213***	0.172	-0.227	0.200	1.000	0.213
P. fimbriata	22.642 (0.606)	2.809 (0.336)	1.999 (0.197)	0.252 (0.055)	0.324 (0.047)	0.373 (0.313)	0.250 (0.150)	137.34 (104.03)	0.195 (0.03)
PFms13	24.143	1.143	1.037	0.000***	0.029	1.000	0.905	0.026	0.186
PFms23	24.143	2.000	1.998	0.983***	0.499	-0.967	0.000	808.81	0.105
PFms26	24.143	1.714	1.065	0.060	0.057	-0.064	0.023	10.537	0.106
PFms33	21.286	5.714	3.472	0.167***	0.654	0.765	0.120	1.840	0.334
PFms35	24.143	1.143	1.067	0.000***	0.046	1.000	0.300	0.582	0.182
PFms36	21.857	5.286	3.044	0.306*	0.647	0.538	0.101	2.223	0.257

Table 4.2. Summary statistics for populations of *P. americana* and *P. fimbriata* based on ten and six polymorphic microsatellite loci, respectively. Abbreviations: N, mean number of individuals sampled per locus per population; %P, estimated proportion of polymorphic loci; Na, mean estimated number of alleles at polymorphic loci; NE, number of effective alleles; PA, number of private alleles; CA, number of locally common alleles (≥5% frequency) found in at least 50% of populations; Ho, observed heterozygosity; He, expected heterozygosity; F, fixation index; Ne, effective population size. Values at the species level represent the mean value for all populations. Values in parentheses are standard error. Stars on Ho indicate a significant excess of heterozygotes in the case of *P. americana* and significant heterozygote deficiency for *P. fimbriata* (* $p<0.05$, ** $p<0.01$, *** $p<0.001$). Infinity symbols (∞) for Ne generally are interpreted as a result of sampling error causing the variation observed in the genetic characteristic used to infer Ne.

Species/Pop	N	%P	Na	NE	PA	CA	Ho	He	F	Ne
P. americana	21.5 (0.3)	78.9% (4.2)	3.044 (0.238)	1.958 (0.147)			0.384 (0.037)	0.334 (0.029)	-0.130 (0.051)	
PAM1	24.1	100.0%	3.100	1.709	0.600	0.700	0.487	0.361	-0.180	318.8 (408.1)
PAM2	24.2	90.0%	2.300	1.546	0.100	0.700	0.414***	0.292	-0.290	∞
PAM3	22.8	70.0%	2.200	1.650	0.000	0.700	0.458***	0.311	-0.453	∞
PAM11	19.0	80.0%	3.400	1.935	0.600	1.400	0.281	0.344	0.125	∞
PAM12	19.4	80.0%	4.100	2.652	0.500	1.700	0.302	0.364	0.129	∞
PAM14	18.2	70.0%	3.800	2.390	0.600	1.600	0.237	0.325	0.289	163.6 (209.5)
PAM24	18.1	70.0%	3.500	2.328	0.500	1.100	0.337	0.342	-0.050	4.8 (3.75)
PAM26	23.8	90.0%	2.600	1.731	0.300	0.800	0.468*	0.351	-0.221	17.1 (15.9)
PAM27	23.6	60.0%	2.400	1.680	0.500	0.700	0.467***	0.314	-0.521	13.2 (12.4)
P. fimbriata	23.7 (0.3)	33.8% (1.7)	2.000 (0.215)	1.517 (0.112)			0.138 (0.033)	0.176 (0.031)	0.160 (0.118)	
PF1	24.0	36.4%	2.091	1.529	0.091	0.455	0.159**	0.176	-0.002	2.7 (1.3)
PF2	25.0	36.4%	1.909	1.531	0.091	0.091	0.124***	0.190	0.366	∞
PF3	25.2	36.4%	1.636	1.319	0.000	0.182	0.095***	0.151	0.468	1.0 (0.31)
PF4	21.2	27.3%	2.000	1.409	0.273	0.182	0.111***	0.157	0.228	∞
PF5	21.6	36.4%	2.182	1.645	0.091	0.636	0.151***	0.192	0.051	4.1 (3.85)
PF6	23.8	36.4%	2.091	1.486	0.091	0.273	0.167**	0.178	-0.020	3.6 (1.86)
PF7	24.8	27.3%	2.091	1.706	0.091	0.273	0.159***	0.186	0.006	4.1 (1.56)

Table 4.3. Population estimates of selfing rates in *P. americana* based upon two different estimation approaches: (a) among-loci estimates obtained using RMES and generated either from the value of $\hat{g}_2$, an estimator of the two-locus heterozygosity disequilibrium ($\hat{s}$ ($\hat{g}_2$)), and (b) within- and among-loci estimates obtained from MLTR for the multi-locus selfing rate (s_m) and the single-locus selfing rate (s_s), as well as an estimate of the amount of selfing due to biparental inbreeding (s BPI). Also provided are the number of individuals per population (N), and the number of loci used (L) for each estimation procedure. Stars indicate significance at the $p<0.05$ level.

		(a) Among-loci estimates			*(b) Within- and among-loci estimates*					
Pop	N	L	$\hat{s}$ ($\hat{g}_2$)	se	L	s_m (mltr)	se	s_s (mltr)	se	s BPI
PAM1	25	8	0.219*	0.024	10	0	0.001	0.131	0.097	0.131
PAM2	25	9	0.213*	0.022	10	0	0.001	0.088	0.022	0.088
PAM3	23	5	0	0.000	10	0	0.001	0.192	0.044	0.192
PAM11	20	8	0.402*	0.032	10	0.008	0.019	0.074	0.07	0.066
PAM12	20	8	0.102	0.037	10	0.044	0.051	0.099	0.046	0.055
PAM14	19	6	0	0.000	10	0	0.002	0.195	0.191	0.195
PAM24	19	7	0.292*	0.030	10	0	0.004	0.168	0.129	0.168
PAM26	25	9	0.057	0.017	10	0	0.001	0.201	0.1	0.201
PAM27	24	5	0.110*	0.021	10	0	0.001	0.260	0.024	0.260

Table 4.4. Population estimates of selfing rates in *P. americana* based upon two different estimation approaches: (a) among-loci estimates obtained using RMES and generated either from the value of $\hat{g}_2$, an estimator of the two-locus heterozygosity disequilibrium ($\hat{s}$ ($\hat{g}_2$)), and (b) within- and among-loci estimates obtained from MLTR for the multi-locus selfing rate (s_m) and the single-locus selfing rate (s_s), as well as an estimate for the percentage of selfing due to biparental inbreeding (s BPI). Also provided are the number of individuals per population (N), and the number of loci used (L) for each estimation procedure. Stars indicate significance at the p<0.01 level.

		(a) Among-loci estimates			*(b) Within- and among-loci estimates*					
Pop	N	L	$\hat{s}$ ($\hat{g}_2$)	se	L	s_m (mltr)	se	s_s (mltr)	se	s BPI
PF1	24	4	0.446**	0.024	6	0.675	0.149	0.763	0.131	0.088
PF2	25	3	0.373	0.027	6	0.711	0.302	0.766	0.269	0.055
PF3	26	1	--	--	6	1	0.652	1	0.383	--
PF4	22	1	--	--	6	1	0.001	1	0.001	--
PF5	23	3	0.410	0.121	6	0.815	0.100	0.940	0.032	0.125
PF6	24	3	0	0.000	6	0.648	0.190	0.753	0.134	0.105
PF7	25	3	0.372	0.098	6	1	0.001	1	0.001	--

Table 4.5. Population pairwise F_{ST} (below diagonal) and gene flow (Nm; above diagonal) estimates corrected for null alleles using the ENA correction Chapuis & Estoup (2007) for *P. americana*. Dotted-lined boxes highlight groups of populations among which there is low differentiation and relatively high levels of gene flow.

	PAM1	PAM2	PAM3	PAM11	PAM12	PAM14	PAM24	PAM26	PAM27
PAM1	--	18.29	3.32	0.47	0.50	0.52	0.46	0.94	0.46
PAM2	0.013	--	3.18	0.41	0.45	0.47	0.41	0.81	0.40
PAM3	0.070	0.073	--	0.37	0.41	0.38	0.35	0.60	0.36
PAM11	0.348	0.380	0.405	--	3.16	3.97	1.28	0.68	0.67
PAM12	0.335	0.359	0.381	0.073	--	7.29	3.28	0.69	0.71
PAM14	0.325	0.349	0.394	0.059	0.033	--	2.92	0.76	0.66
PAM24	0.353	0.379	0.418	0.164	0.071	0.079	--	0.59	0.59
PAM26	0.210	0.236	0.293	0.269	0.265	0.247	0.296	--	0.56
PAM27	0.351	0.383	0.407	0.271	0.262	0.275	0.299	0.307	--

Table 4.6. Population pairwise F_{ST} (below diagonal) and gene flow (Nm; above diagonal) estimates corrected for null alleles using the ENA correction Chapuis & Estoup (2007) for *P. fimbriata*. Dotted-lined boxes highlight groups of populations among which there is low differentiation and relatively high levels of gene flow.

	PF1	PF2	PF6	PF7	PF4	PF3	PF5
PF1	--	4.85	9.01	7.33	3.48	0.84	0.76
PF2	0.049	--	4.65	4.30	2.69	0.81	0.82
PF6	0.027	0.051	--	8.68	4.06	0.80	0.86
PF7	0.033	0.055	0.028	--	4.06	0.68	0.85
PF4	0.067	0.085	0.058	0.058	--	0.67	0.73
PF3	0.23	0.235	0.237	0.27	0.271	--	2.30
PF5	0.248	0.234	0.225	0.227	0.254	0.098	--

Table 4.7. Mean posterior log-likelihood values, standard deviations, and delta K values for each possible level of K in *P. americana* (9) and *P. fimbriata* (7) (the number of populations sampled for each), based on five independent runs in InStruct.

	P. americana			*P. fimbriata*		
K	Mean LnL	s.d.	ΔK	Mean LnL	s.d.	ΔK
1	-3191.64	0.09		-1210.93	0.35	
2	-2627.70	8.42	35.16	-1033.42	7.19	8.43
3	-2359.88	2.28	**71.76**	-916.53	1.58	**38.04**
4	-2255.74	3.00	0.41	-859.60	1.52	9.91
5	-2152.82	1.39	71.21	-817.74	1.24	12.04
6	-2148.53	1.99	1.40	-790.80	0.81	32.89
7	-2147.03	2.22	1.86	-790.65	0.59	
8	-2149.66	1.00	0.99			
9	-2151.30	1.09				

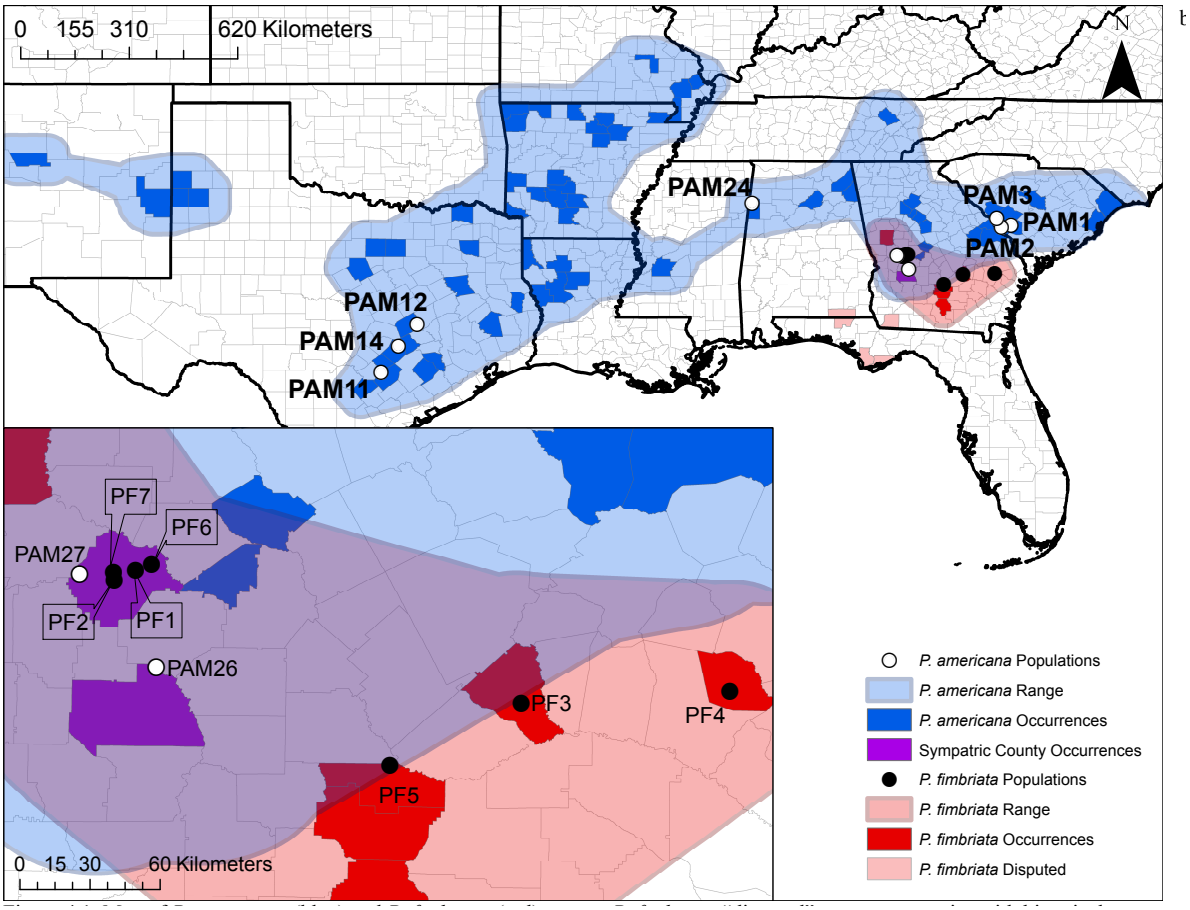


Figure 4.1 Map of *P. americana* (blue) and *P. fimbriata* (red) ranges. *P. fimbriata* "disputed" areas are counties with historical presence records that could not be located. Sympatric county occurrences (purple counties) are counties where the ranges of both species overlap. Tissue collections were taken from *P. americana* (white circles) and *P. fimbriata* (black circles) populations.

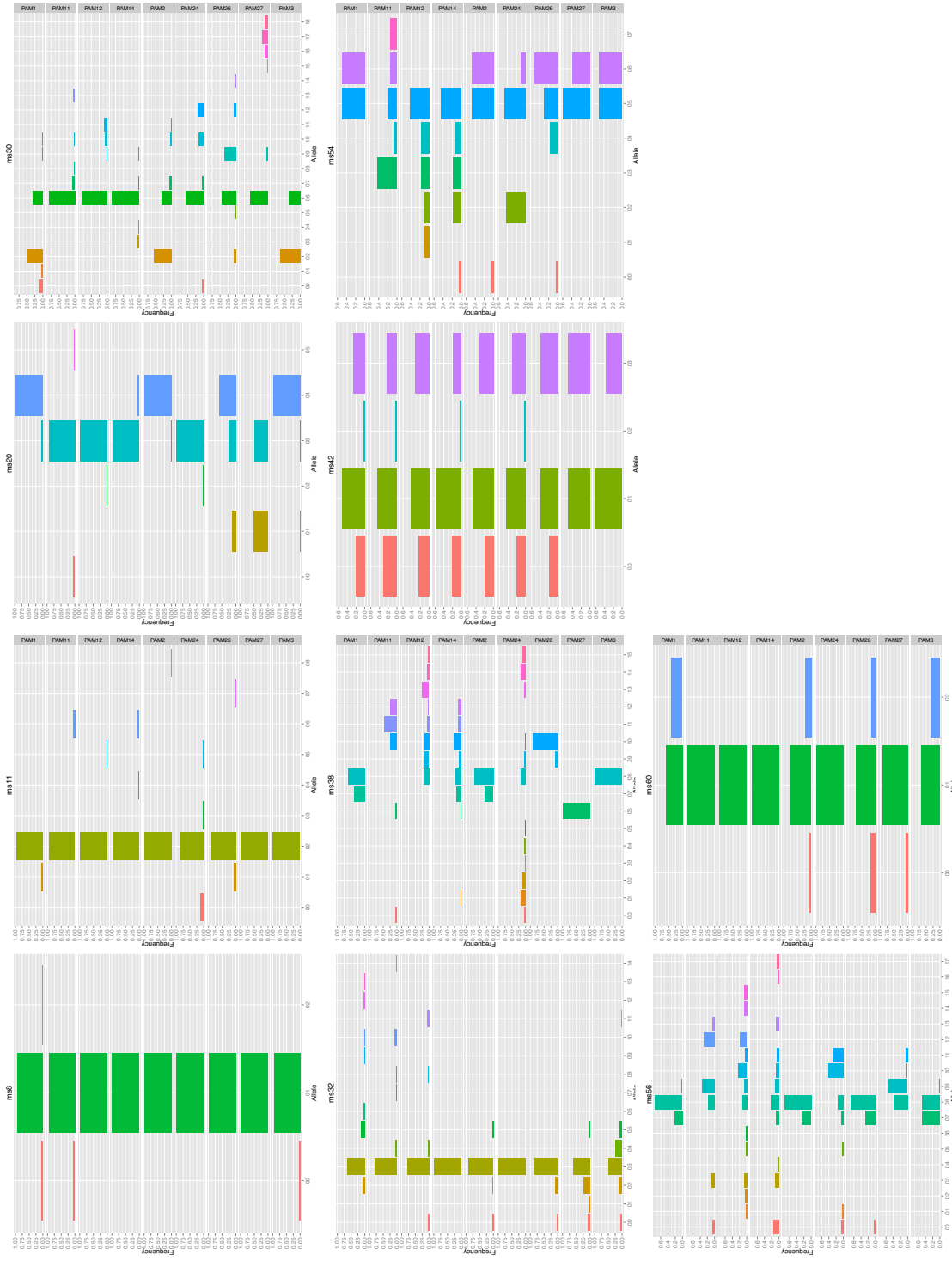


Figure 4.2. Allele frequencies for each locus in *P. americana* by population.

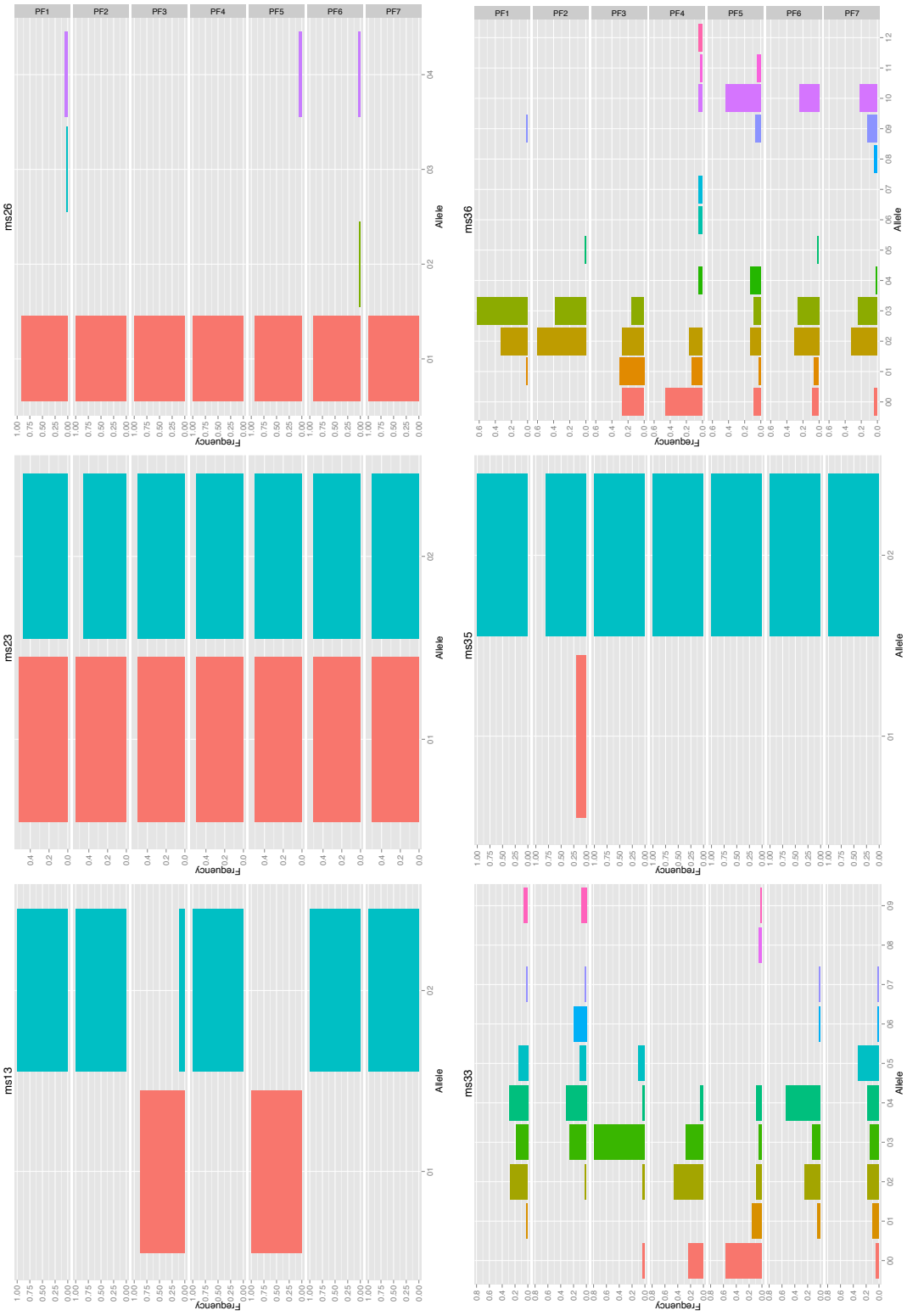


Figure 4.3. Allele frequencies for each locus in *P. fimbriata* by population.

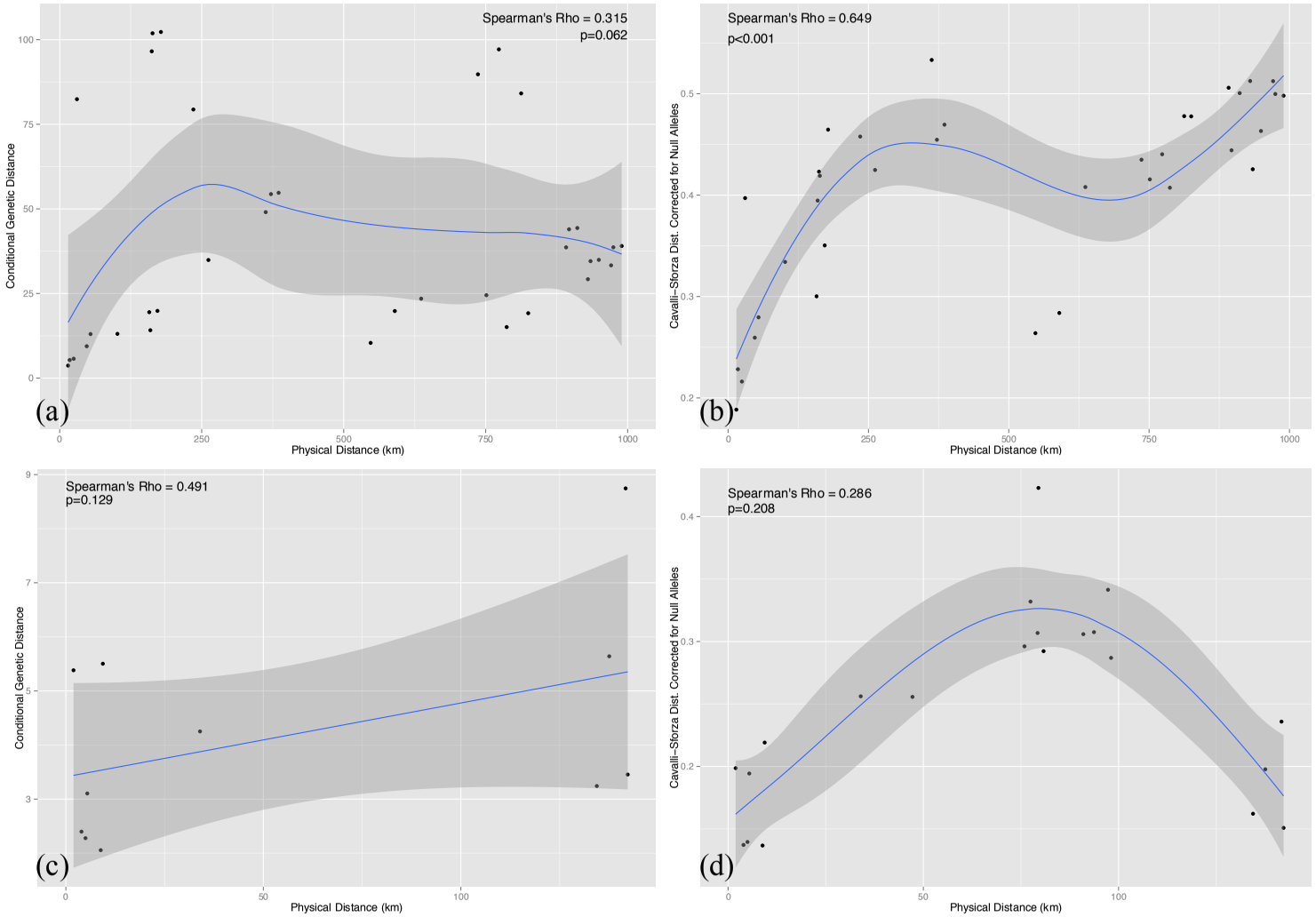


Figure 4.4. Plots for isolation by distance tests for *P. americana* (a,b) and *P. fimbriata* (c,d) using either conditional genetic distance (a,c) or Cavalli-Sforza chord distance corrected for the presence of null alleles (b,d) using the INA correction of Chapuis & Estoup (2007). Smoothing and curves applied using the loess method (a, b, and d) and the linear method (c) in R.

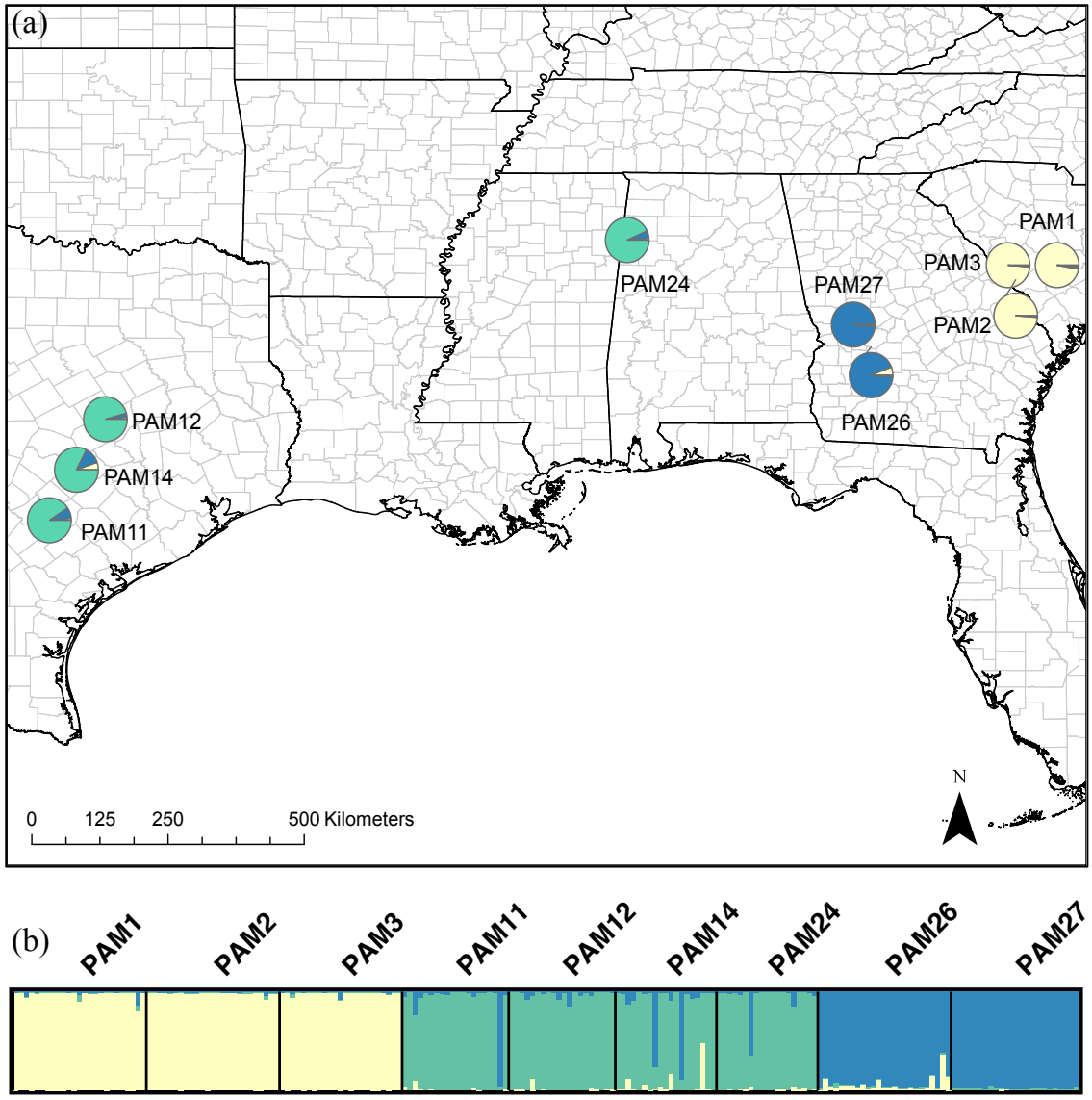


Figure 4.5. InStruct results for *P. americana* for K=3. (a) The geographic distribution of genetic structure in populations of *P. americana* and (b) individual assignments to structure clusters organized by population.

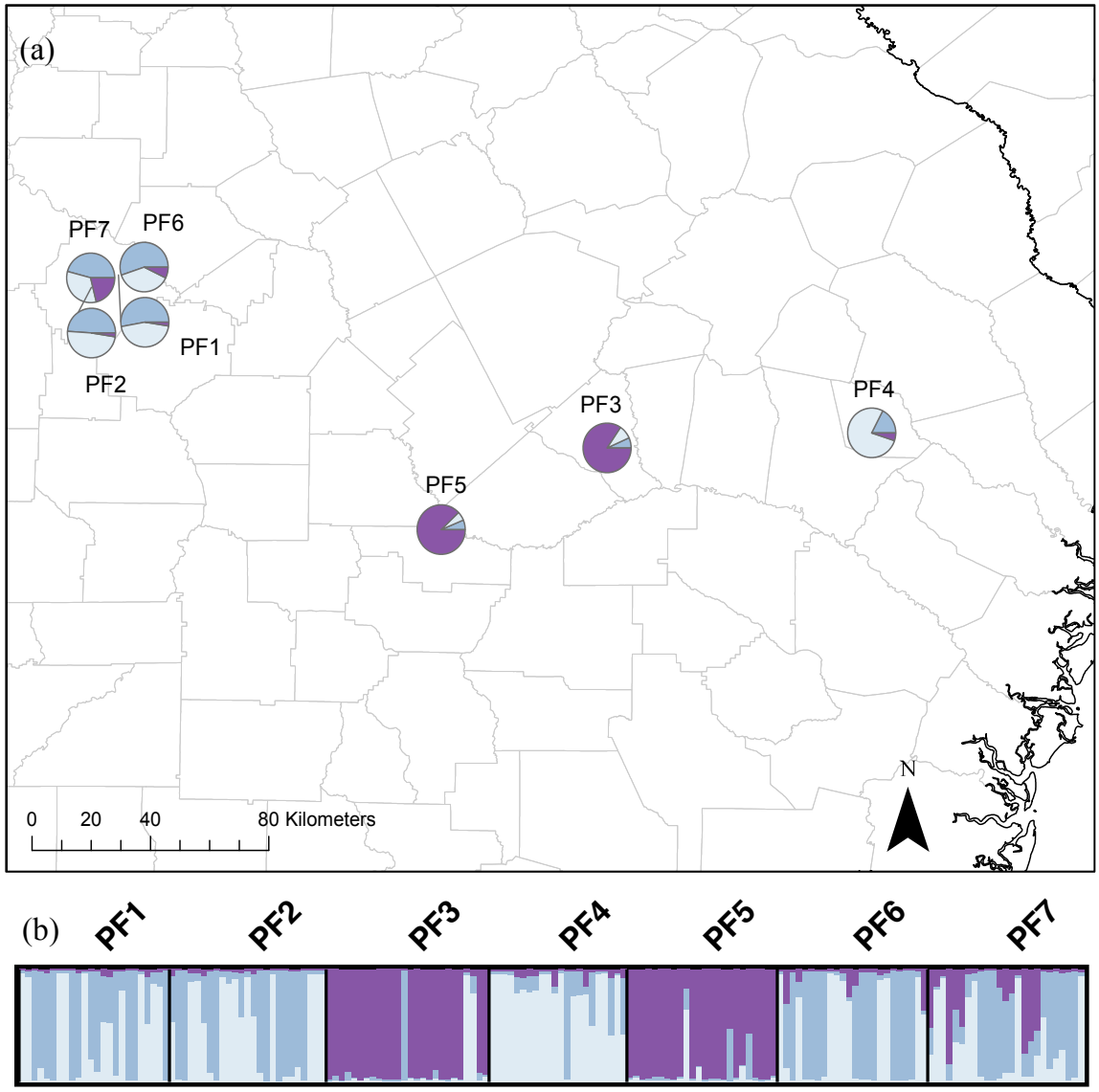


Figure 4.6. InStruct results for *P. fimbriata* for K=3. (a) The geographic distribution of genetic structure in populations of *P. fimbriata* and (b) individual assignments to structure clusters organized by population.

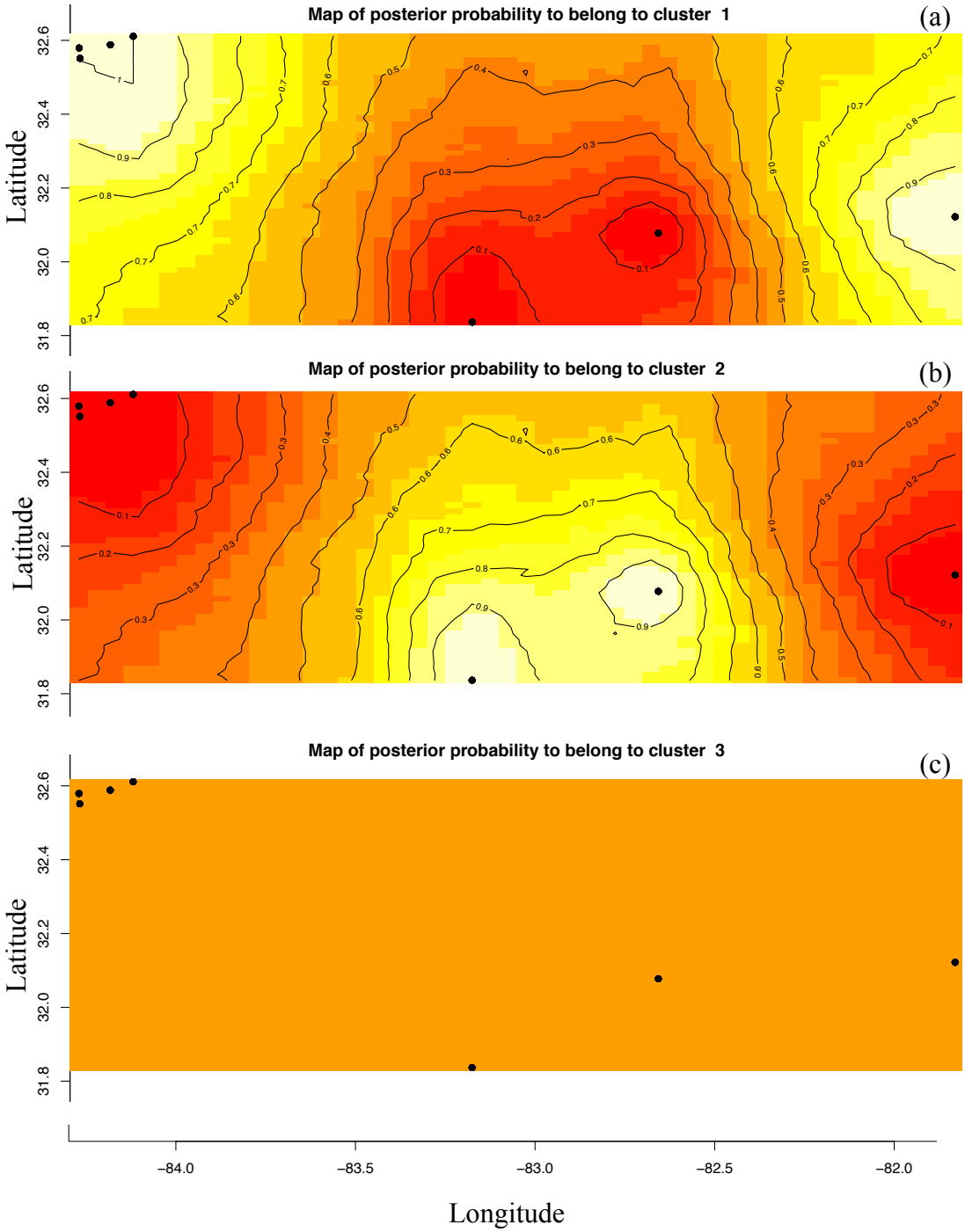


Figure 4.7. Heat maps of the posterior probability of belonging to (a) cluster 1, (b) cluster 2, or (c) cluster 3 from Geneland Bayesian analyses of structure for *P. fimbriata*. Black dots are each of the *P. fimbriata* populations, whiter colors indicate highest probability while red colors indicate lowest probability. Contour lines demarcate different zones of posterior probability.

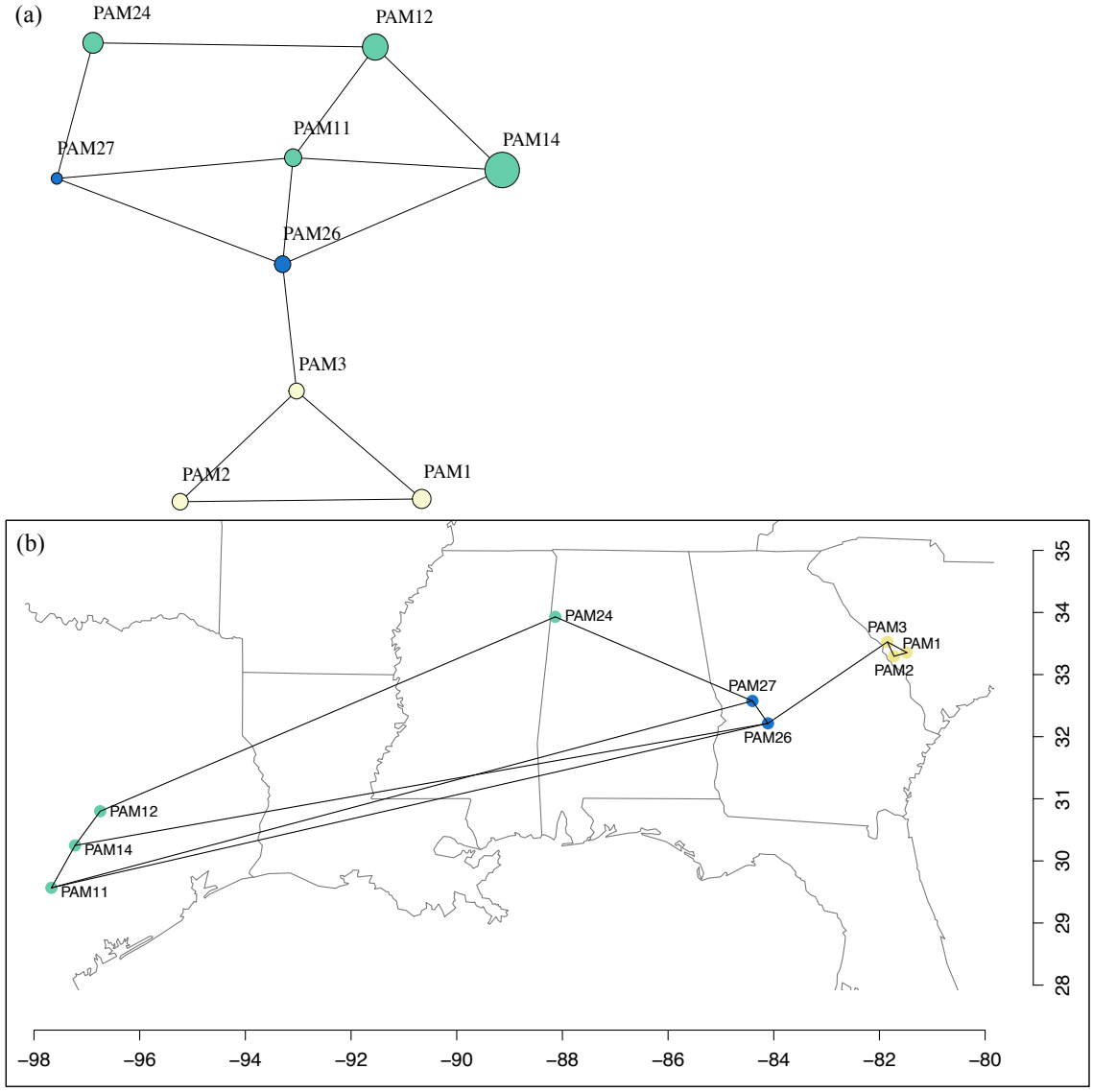


Figure 4.8. PopGraph results for *P. americana*. (a) The population graph created from conditional genetic distances (cGD) where edges between nodes infer gene flow, their lengths represent cGD among nodes, and the relative size of each node is indicative of the amount of genetic variation within that population. (b) The population graph for *P. americana* in its geographical context. In (b), edge lengths and node sizes are not genetically meaningful.

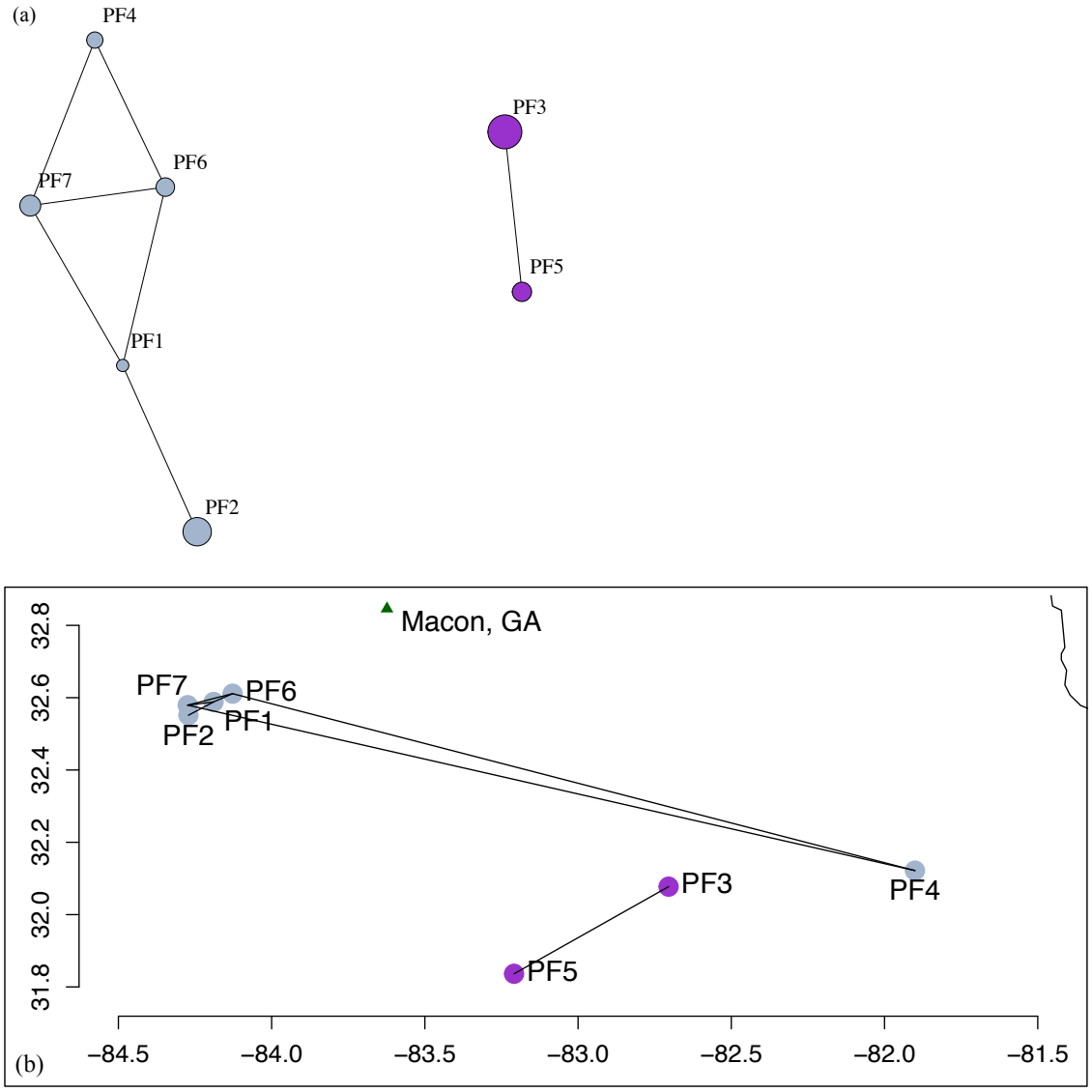


Figure 4.9. PopGraph results for *P. fimbriata*. (a) The population graph created from conditional genetic distances (cGD) where edges between nodes infer gene flow, their lengths represent cGD among nodes, and the relative size of each node is indicative of the amount of genetic variation within that population. (b) The population graph for *P. fimbriata* in its geographical context. In (b), edge lengths and node sizes are not genetically meaningful.

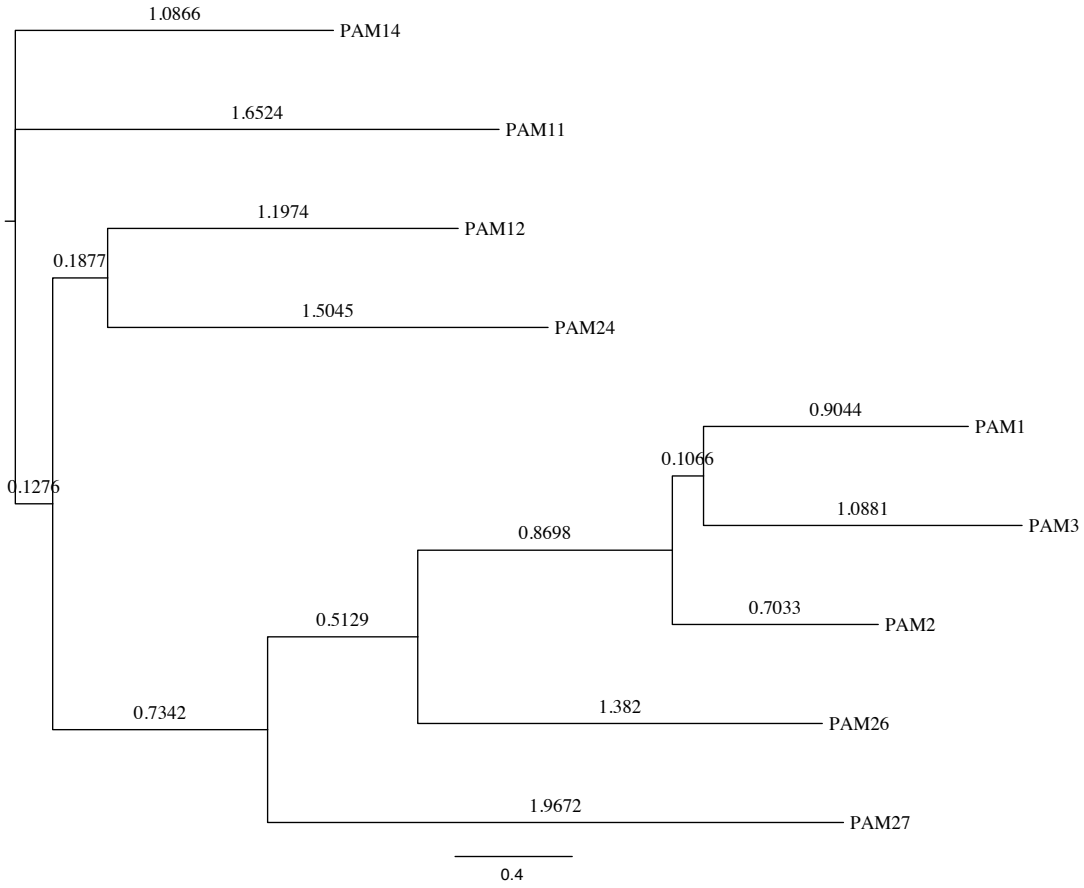


Figure 4.10. Unrooted neighbor joining tree for *P. americana* based on null allele corrected Cavalli-Sforza chord distances. Branch labels indicate genetic distances.

CHAPTER FIVE

CONCLUSIONS AND FUTURE DIRECTIONS

Species are finite in their abundances and distributions. While this observation is a simple one, the processes that determine species' distributions are a complex web of interacting forces. Species' ranges tend to follow the geographic distribution of their ecological niches (Araújo & Guisan, 2006), and when species fill their suitable habitat entirely, niche constraints are assumed to maintain the edge of the range (Sexton *et al.*, 2009). However, many species fail to track their ecological niches in space (Holt, 2003; Sax *et al.*, 2013). When this is the case, attention falls upon the dynamics affecting populations at the range margin (Barton, 2001). Failure to expand the range into the full distribution of suitable habitat may be the result of dispersal limitations, failures in local adaptation due to evolutionary genetic constraints or demographic processes, or any combination thereof (Kirkpatrick & Barton, 1997; Holt & Keitt, 2000; Kawecki, 2008; Case *et al.*, 2005; Bridle & Vines, 2007; Moeller *et al.*, 2011).

Plants are useful study systems for investigations into range limits because, due to their sessile nature, their ranges are less labile than those of most animals. In this dissertation, I set out to determine some of the factors contributing to geographic range limits in a set of narrow, endemic plant species. In chapter two, I took advantage of the abundant presence data available for rare and endangered species in order to construct spatial distribution models (SDMs) for nine plant species in the southeastern U.S. My objective was to identify what abiotic factors define the niches of the rare and threatened species in this study, evaluate whether these species' ranges were in equilibrium with their abiotic niches, and evaluate the predicted stability of these niches in the context of

future climate change. I demonstrated that the abiotic niches of rare species in the southeastern U.S. tend to be defined primarily by correlates of temperature and soil characteristics. In addition, these models predicted dramatic amounts of suitable habitat loss for most of these species in response to climate change, indicating that these abiotic niches will not remain stable as the climate shifts. For the small number of species for which suitable habitat was predicted to increase, this future habitat is predicted to shift away from the current geographic location of suitable habitat. Given dispersal and genetic constraints in these species, they are unlikely to track this habitat at the rate at which it is predicted to shift, similar to what has been demonstrated in trees that have failed to expand fully into suitable habitat following glaciation as a result of dispersal lags (Davis & Shaw, 2001; Svenning & Skov, 2007). Finally, I identified several plant species at severe risk of extinction due to climate change (*Croomia pauciflora, Polygonella basiramia,* and *Waldsteinia lobata*). For these species, the SDMs constructed in this study may be used to identify areas appropriate for assisted migration, following the suggestions of McLachlan *et al.* (2007).

In the following chapters, I shifted focus to range limits in a narrow endemic annual, *Polygonella fimbriata*, and its widespread perennial congener, *Polygonella americana*. In chapter three, I used reciprocal transplant experiments to demonstrate that niche constraints most likely are not contributing to range limits in the narrow endemic species, *P. fimbriata*. Additionally, I showed that there is very little evidence for local adaptation in this species. In contrast, I demonstrated that populations of *P. americana* show evidence of being locally adapted to their home sites, particularly populations from the edge of the *P. americana* range. Finally, I showed that the use of genetic clusters

(groups of populations that are more similar based on genetic analyses) in place of populations is a more appropriate and powerful way of inferring genotype by environment interactions in these types of studies when there is evidence of gene flow among populations. This last point is cause for reconsidering the way performance in reciprocal transplant experiments is evaluated.

In chapter four, I used microsatellites to investigate evidence for evolutionary genetic constraints in the same populations of *P. fimbriata* and *P. americana* utilized in chapter three. Overall, I showed that estimates of genetic diversity in populations of these species were similar to those in other widely and narrowly distributed species based on microsatellite data. Specifically, populations of *P. fimbriata* demonstrated low amounts of genetic diversity, high differentiation among structure clusters, and low effective population sizes. These findings are consistent with the expectations of Hoffman & Blows (1994) for species that are range limited due to evolutionary genetic constraints. Based on these findings, I argue that the most likely explanation for range limits in *P. fimbriata* is the combination of dispersal limitation and evolutionary genetic constraints. In contrast, widespread *P. americana* demonstrated higher levels of genetic diversity and a very strong geographic signal to population differentiation (specifically, isolation by distance). Based on my findings, I argue that regional divergence may be occurring within this species.

This work has implications for investigations into range limits in general, and specifically suggests some improvements for the use of reciprocal transplant experiments in these pursuits. First, the use of SDMs can improve the planning and design of reciprocal transplant studies, which require much in the way of time and resources.

Specifically, SDMs should be useful when exploring the role niche constraints play in shaping the range edge. In these cases, the location of transplant gardens beyond the range edge can be guided by SDM predictions of where abiotic niche limits occur in the landscape. This practice would also serve as a "ground truth" test of the predictions made by SDMs regarding the location of suitable habitat. As SDM techniques continue improve, specifically to incorporate relevant biological information for the species being modeled (e.g., biotic interactions and genetic diversity), this practice could make reciprocal transplants an even more powerful tool.

Second, this work argues strongly for careful consideration in how data from reciprocal transplant experiments are analyzed. The identification of true populations in nature is at times arbitrary (Berryman, 2002), and can be further confounded by human-mediated fragmentation and development. Because of this, the use of genetic information regarding connectivity via gene flow and differentiation has been argued for in the defining of populations (Waples & Gaggiotti, 2006). For this reason, in my investigation of local adaptation in populations of *P. americana* and *P. fimbriata*, I categorized response data in three different ways to detect adaptation: 1) according to geographic population assignments based on where seed and tissue collections were made, 2) based on classifications of populations as home or away within each transplant site, and 3) according to clustering of populations based on genetic similarity (genetic clusters, "GC"). This GC classification scheme demonstrated better success at distinguishing significant signals of local adaptation in comparison to the first two classification schemes. Most likely, this is because of gene flow among populations, and increased statistical power. I argue that future transplant experiments should incorporate cursory

investigations into gene flow among the populations involved in the experiment, so that subsequent transplant data analyses are not misinterpreted.

Finally with respect to experimental design, my transplant experiment demonstrates the importance of incorporating germination when considering local adaptation. Some of the more influential reciprocal transplant studies in the range limits literature have utilized seedlings in place of seeds in their experimental designs (e.g., Angert & Schemske, 2005; Griffith & Watson, 2006; Angert *et al*., 2008). While it is understandable that practical considerations will play a role in making this decision, the act of germination requires a very specific set of cues and has downstream effects that affect the entire life cycle of the plant (Donohue *et al*., 2010). Transplanting seedlings that have germinated under favorable, controlled conditions (fertilized soil, plenty of water, etc.) will artificially decouple the germination niche and the post-germination niche, for which there is some evidence of coadaptation (Donohue, 2002; Donohue *et al.*, 2005). Experiments geared towards elucidating factors important for range limits should include an assessment of germination within and beyond the limits of the range.

Not only does this work suggest areas for improvement in current work, it highlights some opportunities for future work. First, the structuring of genetic variation in *P. americana* (chapter four) is possible evidence for a phylogeographic break within this species. However, population sampling was not thorough enough to test this. Phylogeographic comparisons between this species and the other widespread *Polygonella* species that is distributed to the north (*P. articulata*) would be informative regarding *P. americana*'s postglacial expansion (or rather, its apparent lack thereof). Second, a fully-resolved and complete molecular phylogeny for the *Polygonella* genus that includes all

species still does not exist. This genus comprises eleven species that are variable in their ranges, breeding systems, lifespans, and their growth habits. Ancestral state reconstructions investigating the evolution of this incredible phenotypic diversity would be informative on a number of levels. Additionally, a phylogenetic context for future experimental work will improve conservation decision-making in this system, as well as help elucidate some putative instances of hybridization among some of these species, many of which are known to be sympatric. Finally, four of the *Polygonella* species are listed either on the state or federal level as threatened or endangered. Future climate change predictions demand that more work be done in this system to mitigate future biodiversity loss. The work done here is the first step towards achieving this.

Works Cited

Angert, A.L. & Schemske, D.W. (2005) The evolution of species' distributions: reciprocal tranplants across the elevation ranges of *Mimulus cardinalis* and *M. lewisii*. *Evolution*, **59**, 1671-84.

Angert, A.L., Bradshaw Jr., A.D. & Schemske, D.W. (2008) Using experimental evolution to investigate geographic range limits in monkeyflowers. *Evolution*, **62**, 2660-2675.

Araújo, M.B. & Guisan, A. (2006) Five (or so) challenges for species distribution modeling. *Journal of Biogeography*, **33**, 1677-88.

Bahn, V., O'Connor, R.J. & Krohn, W.B. (2006) Effect of dispersal at range edges on the structure of species ranges. *Oikos*, **115**, 89-96.

Barton, N.H. (2001) Adaptation at the edge of a species' range. (Eds. J. Silvertown & J. Antonovics) *Integrating Ecology and Evolution In a Spatial Context*. Blackwell Science, New York, NY.

Berryman, A.A. (2002) Population: a central concept for ecology? *Oikos*, **97**, 439-42.

Bridle, J.R. & Vines, T.H. (2007) Limits to evolution at range margins: when and why does adaptation fail? *Trends in Ecology and Evolution*, **22**, 140-7.

Case, T.J., Holt, R.D., McPeek, M.A. & Keitt, T.H. (2005) The community context of species' borders: ecological and evolutionary perspectives. *Oikos*, **108**, 28-46.

Davis, M.B. & Shaw, R.G. (2001) Range shifts and adaptive responses to Quaternary climate change. *Science*, **292**, 673-9.

Donohue, K. (2002) Germination timing influences natural selection on life-history characters in *Arabidopsis thaliana*. *Ecology*, **83**, 1006-1016.

Donohue, K,. de Casas, R.R., Burghardt, L., Kovach, K. & Willis, C.G. (2010) Germination, postgermination adaptation, and species ecological ranges. *Annual Review of Ecology, Evolution, and Systematics,* **41**, 293-319.

Donohue, K., Dorn, L., Griffith, C., Kim, E., Aguilera, A., Polisetty, C.R. & Schmitt, J. (2005) Environmental and genetic influence on the germination of *Arabidopsis thaliana* in the field. *Evolution*, **59**, 740-57.

Griffith, T.M. & Watson, M.A. (2006) Is evolution necessary for range expansion? Manipulating reproductive timing of a weedy annual transplanted beyond its range. *The American Naturalist*, **167**, 153-64.

Hoffman, A.A. & Blows, M.W. (1994) Species borders: ecological and evolutionary perspectives. *Trends in Ecology and Evolution*, **9**, 223-7.

Holt, R.D. (2003) On the evolutionary ecology of species' ranges. *Evolutionary Ecology Research*, **5**, 159-78.

Holt, R.D. & Keitt, T.H. (2000) Alternative causes for range limits: a metapopulation perspective. *Ecology Letters*, **3**, 41-7.

Kawecki, T.J. (2008) Adaptation to marginal habitats. *Annual Reviews in Ecology, Evolution and Systematics*, **39**, 321-42.

Kirkpatrick, M. & Barton, N.H. (1997) Evolution of a species' range. *The American Naturalist*, **150**, 1-23.

McLachlan, J.S., Hellmann, J.J. & Schwartz, M.W. (2007) A framework for debate of assisted migration in an era of climate change. *Conservation Biology*, **21**, 297-302.

Moeller, D.A., Geber, M.A. & Tiffin, P. (2011) Population genetics and the evolution of

geographic range limits in an annual plant. *The American Naturalist*, **178**, S44-S57.

Sax, D.F., Early, R. & Bellemare, J. (2013) Niche syndromes, species extinction risks, and management under climate change. *Trends in Ecology and Evolution*, **28**, 517-23.

Sexton, J.P., McIntyre, P.J., Angert, A.L. & Rice, K.J. (2009) Evolution and ecology of species range limits. *Annual Review of Ecology, Evolution and Systematics*, **40**, 415-36.

Svenning, J.-C. & Skov, F. (2007) Could the tree diversity pattern in Europe be generated by postglacial dispersal limitation? *Ecology Letters*, **10**, 453-60.

Waples, R.S. & Gaggiotti, O. (2006) What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. *Molecular Ecology*, **15**, 1419-1439.