History, chance, and adaptation: the evolution of *Silene vulgaris* in its native and introduced ranges

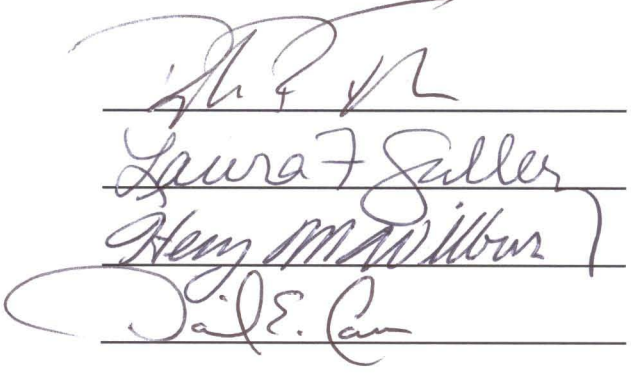
Stephen Robert Keller
Malvern, Pennsylvania

M.S., University of Alaska Fairbanks, 2002
B.S., Juniata College, 1997

A Dissertation presented to the Graduate Faculty
of the University of Virginia in Candidacy for the Degree of
Doctor of Philosophy

Department of Biology

University of Virginia
January, 2008



Abstract

Species expanding into new ranges will experience stochastic effects of colonization and must either be preadapted or evolve adaptations to survive and reproduce in a novel environment. Biological invasions provide natural experiments for investigating these issues. This dissertation used the 200 year old invasion of North America by *Silene vulgaris*, a weedy plant native to Europe, to address how invasions are impacted by (1) evolutionary history in the native range, (2) the genetic composition of the invasion inoculum, (3) the match in ecological niche between ranges, and (4) post-invasion evolution. A phylogenetic investigation of chloroplast DNA supported a history of range expansion in Europe from Mediterranean glacial refugia. However, there was no detectable signature of isolation among native range geographic regions in this locus. In a multi-locus analysis of the nuclear genome, amplified fragment length polymorphisms (AFLP) revealed five divergent demes that were regionally structured in Europe. Both cpDNA and AFLP supported a genetically diverse inoculum, probably the result of multiple introductions. Founder effects were evident in shifted deme frequencies and in the mismatching of genotypes from the ecological niche predicted by demes in the native range. As the invasion progressed, *S. vulgaris* expanded its range to fill and eventually exceed the predicted niche, consistent with a scenario of adaptive evolution. This hypothesis was tested by planting 1600 individuals from both ranges into common gardens in Ontario and Virginia. North American genotypes were genetically divergent for several traits, including having a faster reproductive maturity, and greater overall fecundity and survival compared to European genotypes. Phenotypic clines with latitude were also evident among families collected from both ranges. Incorporating null

expectations using AFLP confirmed that divergence between and within ranges was due in part to post-invasion evolution. However, demes themselves also differed in performance regardless of continent of origin. Coupled with the shift in deme frequencies due to founder effect, this indicated that total phenotypic divergence resulted from a combination of adaptive and stochastic processes. Collectively, these results demonstrate that evolution during invasion is a multi-dimensional process, affected by prior evolutionary history, chance sampling events, and adaptation to the introduced environment.

Acknowledgements

There is no substitute for working through a dissertation to cement the fact that science, and life for that matter, are collaborative processes. I owe a great debt of gratitude to the many people who have contributed to the work laid out on these pages.

This dissertation would not have been possible without the guidance and support of my advisor, Doug Taylor. Doug was immensely generous with his time and resouces, and was an unwaivering advocate for the work I was doing. He has become a true collaborator, one I look forward to continuing to do science with for years to come.

Thanks also to my committee, Janis Antonovics, Dave Carr, Laura Galloway, and Henry Wilbur. I appreciate their time, expertise, and encouragement. In particular, I'd like to thank Janis for pointing out early on the necessity of trying to ascertain which native range sources contributed genotypes during the introduction.

I'd like to thank 'Uncle' Dave McCauley for his generosity with seeds and ideas, and Eric Nagy for his willingness to assist with the borrowing of herbarium material. Many collaborators and friends contributed to the seed collection, especially Henry Frierson, Helena Storchova, and Lorne Wolfe. Nan Rothwell and Carter Smith have extended many kindesses to me, the greatest of which was the permission to set up an experiment in their field.

A special thanks to the Taylor labbies, whose "war room" comraderie has produced a wealth of intellectual breakthroughs, side-stitching laughs, and some great friendships. It's been a wonderful place to come to work each day.

My adoptive family, the Nevilles, not only encouraged what I was doing, but also understood how important it was for me. They contributed much to this dissertation, from

making the farm available, to putting up fence, and even taking data when I couldn't be in two places at once. I couldn't imagine a more loving, generous, and supportive family to marry into.

My Mom and Dad taught me early on that we all stand on the shoulders of those who have come before us. I now fully appreciate the sentiment in those words. I am so grateful for their wisdom, encouragement, and unfailing support. And besides all of that, they braved a cold rainy April day to help plant seedlings. Thanks also to the rest of the Keller family, who are stalwart, steadfast, and always there when you need them. Go Jimmies!

Finally, to my wife Julie and my sons Sam and Jonah, you are my greatest source of calm, perspective, and unconditional love. When we were first starting out, I was worried it might be difficult to be a graduate student and have a family. Now I can't imagine it any other way. My big book is finally done.

Table of Contents

History, chance, and adaptation: the evolution of *Silene vulgaris* in its native and introduced ranges

Introduction

The nature of limits to a species' geographic range is a central problem in ecology and evolutionary biology (Antonovics 1976, Kirkpatrick and Barton 1997, Holt 2003, Holt and Keitt 2005, Parmesan et al. 2005). Historical biogeography and comparative phylogenetics have illuminated our view that geographic ranges are heritable across closely related taxa, but can sometimes show tremendous spatial and temporal flux (Ricklefs and Latham 1992, Peterson et al. 1999, Losos et al. 2003, Weins and Graham 2005, Knouft et al. 2006). Why do some taxa experience rapid expansions of their geographic range limits? Does evolution during historical range expansion pre-adapt a species to spread elsewhere? To what extent do stochastic forces such as founder effect influence evolutionary change during range expansion compared to natural selection? These questions integrate fundamental issues in evolutionary biology such as the efficacy of selection (Antonovics 1976b), the importance of history and constraint (Gould and Lewontin 1979), the balance between selection and gene flow (Wright 1951), and the operation of higher levels of selection (Wade and Goodnight 1998). Indeed, understanding the evolutionary mechanisms that govern shifts in species' ranges has been called a "deeper mystery" (Kirkpatrick and Barton 1997) for which we have relatively few answers.

While evolution during range expansion has traditionally been an historical science, the unprecedented rate with which non-native species are being globally dispersed offers an opportunity to study range limits as they are actively changing (Sakai et al. 2001).

Given the importance of invasive species as threats to biodiversity and ecosystem functioning, major research efforts have recently focused around identifying the attributes that allow species to spread invasively. This work points to a central role for evolution during invasion (Ellstrand and Schierenbeck 2000, Garcia-Ramos and Rodriguez 2002, Lee 2002, Lambrinos 2004), with mounting evidence for both alterations in the genomic organization of populations and in selective regime as species expand into new ranges (de la Vega et al. 1991, Colautti et al. 2004, Kolbe et al. 2004, Bossdorf et al. 2005). These studies bring to the spotlight Antonovics' (1976) view that understanding the nature of range limits is a problem for genetics and evolution as well as ecology.

While the importance of evolution is increasingly recognized, many studies assume a somewhat caricatured view of the invasion process. A frequent hypothesis states that invasive species are initially benign members of native ecological communities that become released from their specialist enemies after introduction to a new range, thus spawning the rapid evolution of invasiveness (Blossey and Notzfold 1995). While the consequences of enemy release for the evolution of invasiveness are still unclear (Willis et al. 2000, Colautti et al. 2004), these instances are a subset of a more complex set of evolutionary forces. Namely, most invasions will involve a combination of (1) evolutionary history, whereby traits that predispose species toward rapid range expansion have evolved over prolonged evolutionary time in the native range, (2) stochastic sampling processes, where history and chance combine to affect which lineages are introduced, and (3) the non-random proliferation of invasive genotypes that may be pre-adapted to the colonized environment, and/or have responded to selection in the new environment. It is important to recognize that the evolutionary consequences of invasions

may be driven as much by historical processes that have occurred in the sources from which invasions are drawn, as they are by contemporary processes that influence the number and diversity of invasive propagules. Indeed, meta-analyses have shown that two of the most powerful predictors of whether a species becomes invasive are the size of the native range and whether the species is invasive elsewhere (Goodwin et al. 1999, Kolar and Lodge 2001)

The central biological question addressed in this dissertation is, how do evolutionary history, chance events, and adaptive evolution affect the invasion of new geographic ranges. In this context, 'history,' refers to the evolution of traits in the native range that pre-adapt species for invasive potential, as well as forces that create or destroy genetic structure in the native range. 'Chance,' refers to founder effects during colonization, as influenced by the history of population structure in the native range and the process of sampling that distribution during invasion. 'Adaptation,' refers to genetic change, post-introduction, that promotes the rapid expansion of the species' geographic range. While contingency and stochasticity, as well as natural selection, are inherent features of the evolutionary process, few studies concurrently address all three of these contributions (Travisano et al. 1995, Losos et al. 1998).

The Study System

Silene vulgaris L. (Moenche) Garke, or "bladder campion", is an herbaceous plant in the family Caryophyllaceae. *S. vulgaris* is native to Eurasia, where it has a widespread geographic range (Figure 1A) and is a common member of the temperate zone flora (Jalas and Suominen 1986). In its native range, S. *vulgaris* occurs as a highly

polymorphic complex with multiple subspecies recognized, distinguishable by their growth habit and ecological requirements (Marsden-Jones and Turrill 1957, Aeschimann and Bocquet 1980, Jalas and Suominen 1986). Many of these are local endemics, with restricted distributions in alpine (*S. v. prostrata*, *S. v. glareosa*), coastal (*S. v. maritima*, also recognized as *S. uniflora*), or Mediterranean habitats (*S. v. commutata*, *S. v. angustifolia*). In contrast, the weedy subspecies *S. v. vulgaris* is a widespread and abundant taxon that grows in a variety of natural and human-disturbed environments, including cliff faces, river banks, talus slopes, open woodlands, meadows, roadsides, hayfields, and occasionally cereal crops (Marsden-Jones and Turrill 1957). It is this weedy subspecies (hereafter referred to simply as *S. vulgaris*) that is the focus of this dissertation.

Compared to its closest relatives, the geographic range of weedy *S. vulgaris* in Europe shows a dramatic expansion in size (Figure 1B). In addition to being widespread in its native range, *S. vulgaris* has been transported globally by humans, and is naturalized in North and South America, South Africa, and Australia. In North America, *S. vulgaris* was introduced during the late eighteenth and early nineteenth centuries, occurring as a weed bordering agricultural fields near Boston and Quebec City (Cutler 1785, Pursh 1814). Later, the species was noted growing out of mounds of dirt, gravel, sand, and refuse from ship ballast piled around the docks at the port cities of Philadelphia and New York (Martindale 1876, 1877, Brown 1878). Thus, every indication is that the introduction was an unintentional byproduct of human commerce and immigration. Since its introduction, *S. vulgaris* has expanded its range across most of temperate North America, although the full extent of its distribution is not previously known.

References

Aeschimann, D., and G. Bocquet. 1980. Les types biologiques du *Silene vulgaris* s.l. Candollea **35**:451-495.

Antonovics, J. 1976. The nature of limits to natural selection. Annals of the Missouri Botanical Garden **63**:224-247.

Blossey, B., and R. Notzfold. 1995. Evolution of increased competitive ability in invasive nonindigenous plants: a hypothesis. Journal of Ecology **83**:887-889.

Bossdorf, O., H. Auge, L. Lafuma, W. E. Rogers, E. Siemann, and D. Prati. 2005. Phenotypic and genetic differentiation between native and introduced plant populations. Oecologia **144**:1-11.

Brown, A. 1878. Plants introduced with ballast and on made land. Bulletin of the Torrey Botanical Club **6**:255-258.

Colautti, R. I., A. Ricciardi, I. A. Grigorovich, and H. J. MacIsaac. 2004. Is invasion success explained by the enemy release hypothesis? Ecology Letters **7**:721-733.

Cutler, M. 1785. An account of some of the vegetable productions, naturally growing in this part of America, botanically arranged. Memoirs of the American Academy of Arts and Sciences **I**:396-493.

de la Vega, M., P. Garcia, and R. Allard. 1991. Multilocus genetic structure of ancestral Spanish and colonial California populations of *Avena barbata*. PNAS **88**:1202-1206.

Ellstrand, N. C., and K. A. Schierenbeck. 2000. Hybridization as a stimulus for the evolution of invasiveness. Proceedings of the National Academy of Sciences of the United States of America **97**:7043-7050.

Garcia-Ramos, G., and D. Rodriguez. 2002. Evolutionary speed of species invasions. Evolution **56**:661-668.

Goodwin, B. J., A. J. McAllister, and L. Fahrig. 1999. Predicting invasiveness of plant species based on biological information. Conservation Biology **13**:422-426.

Gould, S. J., and R. C. Lewontin. 1979. The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme. Proceedings of the Royal Society of London Series B **205**:581-598.

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

Holt, R. D., and T. H. Keitt. 2005. Species' borders: a unifying theme in ecology. Oikos **108**:3-6.

Jalas, J., and J. Suominen. 1986. Atlas Florae Europaeae. Cambridge University Press, Cambridge.

Kirkpatrick, M., and N. H. Barton. 1997. Evolution of a species' range. American Naturalist **150**:1-23.

Knouft, J. H., J. B. Losos, R. E. Glor, and J. J. Kolbe. 2006. Phylogenetic analysis of the evolution of the niche in lizards of the *Anolis sagrei* group. Ecology **87**:S29-S38.

Kolar, C. S., and D. M. Lodge. 2001. Progress in invasion biology: predicting invaders. Trends in Ecology and Evolution **16**:199-204.

Kolbe, J. J., R. E. Glor, L. R. Schettino, A. C. Lara, A. Larson, and J. B. Losos. 2004. Genetic variation increases during biological invasion by a Cuban lizard. Nature **431**:177-181.

Lambrinos, J. G. 2004. How interactions between ecology and evolution influence contemporary invasion dynamics. Ecology **85**:2061-2070.

Lee, C. E. 2002. Evolutionary genetics of invasive species. Trends in Ecology and Evolution **17**:386-391.

Losos, J. B., T. R. Jackman, A. Larson, K. de Queiroz, and L. Rodriguez-Schettino. 1998. Contingency and determinism in replicated adaptive radiations of island lizards. Science **279**:2115-2118.

Losos, J. B., M. Leal, R. E. Glor, K. de Queiroz, P. E. Hertz, L. R. Schettino, A. C. Lara, T. R. Jackman, and A. Larson. 2003. Niche lability in the evolution of a Caribbean lizard community. Nature **424**:542-545.

Marsden-Jones, E. M., and W. B. Turrill. 1957. The bladder campions. Ray Society, London.

Martindale, I. C. 1876. The introduction of foreign plants. Botanical Gazette **2**:55-58.

Martindale, I. C. 1877. More about ballast plants. Botanical Gazette **2**:127-128.

Parmesan, C., S. Gaines, L. Gonzalez, D. M. Kaufman, J. Kingsolver, A. T. Peterson, and R. Sagarin. 2005. Empirical perspectives on species borders: from traditional biogeography to global change. Oikos **108**:58-75.

Peterson, A. T., J. Soberon, and V. Sanchez-Cordero. 1999. Conservatism of ecological niches in evolutionary time. Science **285**:1265-1267.

Pursh, F. 1814. Flora Americae Septentrionalis. White, London.

Ricklefs, R. E., and R. E. Latham. 1992. Intercontinental correlation of geographical ranges suggests stasis in ecological traits of relict genera of temperate perennial herbs. The American Naturalist **139**:1305-1321.

Sakai, A. K., F. W. Allendorf, J. S. Holt, D. M. Lodge, J. Molofsky, K. A. With, S. Baughman, R. J. Cabin, J. E. Cohen, N. C. Ellstrand, D. E. McCauley, P. O'Neil, I. M. Parker, J. N. Thompson, and S. G. Weller. 2001. The population biology of invasive species. Annual Review of Ecology and Systematics 32:305-332.

Travisano, M., J. A. Mogols, A. F. Bennet, and R. E. Lenski. 1995. Experimental tests of the roles of adaptation, chance, and history in evolution. Science 267:87-90.

Wade, M. J., and C. J. Goodnight. 1998. Perspective: The theories of Fisher and Wright in the context of metapopulations: when nature does many small experiments. Evolution 52:1537-1553.

Weins, J. J., and C. H. Graham. 2005. Niche conservatism: integrating evolution, ecology, and conservation biology. Annual Review of Ecology, Evolution, and Systematics 36.

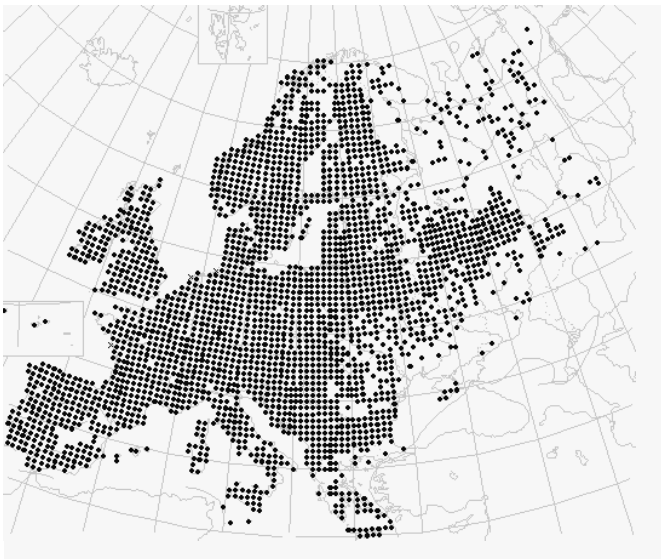
Willis, A. J., J. Memmott, and R. I. Forrester. 2000. Is there evidence for the post-invasion evolution of increased size among invasive plant species? Ecology Letters 3:275-283.

Wright, S. 1951. The genetical structure of populations. Annals of Eugenics 15.

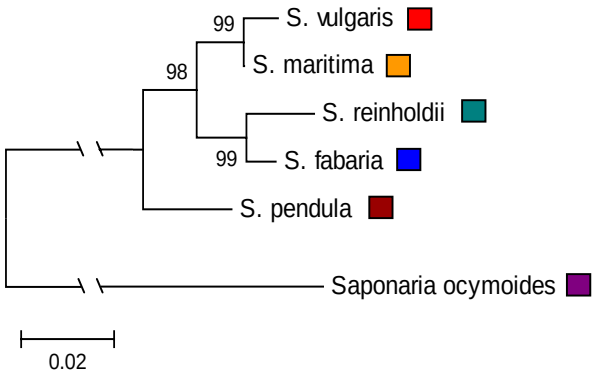
Figure Legends

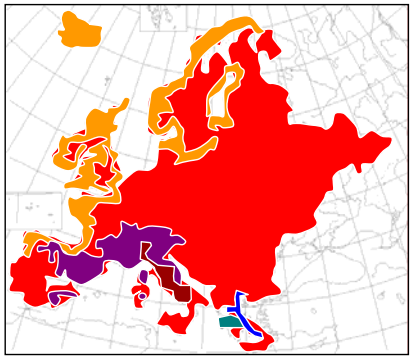
Figure 1. (A) Geographic range of *S. vulgaris* in Europe. Each dot indicates the historical occupancy of a 50 km x 50 km area, as represented in the Atlas Florae Europaeae (Jalas and Suominen 1986). (B) Phylogenetic position and range size of *Silene vulgaris* and its close congeners. Neighbor-joining phylogeny based on the internal transcribed spacers 1 and 2 between the 18S and 26S subunits of the nuclear ribosomal DNA (total aligned length of 575 bp; distances estimated by the Jukes-Cantor method). *Sapponaria ocymoides* was used as an outgroup; all others are genus *Silene.* Sequences were obtained from the Genbank database. Numbers next to nodes are bootstrap support from 1000 replicates. Geographic ranges are adapted from the Atlas Florae Europaeae (Jalas and Suominen 1986).

A



B


99
S. vulgaris
S. maritima
98
S. reinholdii
99
S. fabaria
S. pendula
Saponaria ocymoides
0.02



Chapter 1:

Separating stochastic phenotypic evolution from response to selection during biological invasion[1]

[1] Formatted for submission as a coauthored manuscript: Keller, S.R. and D.R. Taylor

Abstract

Introduced species are rapidly becoming recognized as promising systems for studying adaptive evolution over contemporary time scales. However, changes in adaptively important traits during invasion occur under non-equilibrium demographic conditions and are likely to reflect the influences of prior evolutionary history and chance events, as well as natural selection. We briefly review the evidence for phenotypic evolution and the role of selection during invasion. We then discuss the likelihood that stochastic events shift phenotypic distributions during invasion, and argue that hypotheses of adaptation must be tested against appropriate null models. Two experimental frameworks are suggested for separating stochastic evolution from adaptation: statistically accounting for phenotypic variation among putative invasion sources identified using phylogenetic or assignment methods, and comparing estimates of differentiation within and among ranges for both traits and neutral markers (Q_{ST} versus F_{ST}). Incorporating neutral expectations provides greater insights into contemporary evolution and the emergence of adaptations during invasion.

Keywords: colonization, invasive, founder effect, admixture, drift, assignment tests, F_{ST}, Q_{ST}

Introduction

The widespread introduction and proliferation of non-native species presents a unique opportunity to study evolutionary events accompanying rapidly changing range limits (Baker and Stebbins 1965). Introduced species can experience alterations in genetic diversity and organization, changes in selection regime, and genetically-based shifts in phenotypic traits as they expand into new ranges. In fact, introduced species are now recognized as some of the best model systems for understanding contemporary evolution as the distinction between ecological and evolutionary time becomes less apparent (Thompson 1998, Stockwell et al. 2003). These studies bring to the spotlight the view that explaining the abundance and distribution of organisms, and the nature of their range limits, is a problem for genetics as well as ecology (Antonovics 1976a, b).

While an experimental comparison of populations from a species' native and introduced ranges often reveals phenotypic change (Bossdorf et al. 2005), identifying causal mechanisms is a more challenging task. Changes in the distribution of phenotypic traits during dispersal, colonization, and range expansion occur under non-equilibrium demographic conditions and are affected by prior evolutionary history, chance events, and responses to selection (Figure 1). While it is tempting to ascribe divergence in phenotypic traits to adaptive evolution, this inference is too often made in the absence of an appropriate null model (e.g., Gould and Lewontin 1979). By analogy to community ecology, the neutral theory of biodiversity suggests communities may differ in species diversity because of the deterministic match between resident species and the available ecological niches, as well as the historical and stochastic nuances of dispersal and sampling processes. In this sense, the forces structuring genetic diversity and species

diversity are similar, and we must expect the joint influences of stochasticity and determinism to affect their outcomes (Antonovics 1976a, Alonso et al. 2006).

In this paper, we explore the likelihood that chance events, such as founder effect, interact with prior evolutionary history as major factors driving phenotypic evolution during the introduction and range expansion of introduced species. We first briefly review the potential role of selection and summarize studies that demonstrate phenotypic differentiation at some stage of the invasion process. Although empirical evidence for phenotypic evolution is strong, in most cases the experimental design does not permit an unambiguous, or even a probabilistic, assessment of whether chance sampling of evolutionary history or adaptive evolution has influenced the outcome. We then discuss why neutral phenotypic change is a probable outcome during invasion and offer suggestions for experimental frameworks that test hypotheses of adaptation against null models of neutral phenotypic evolution. Designs that statistically incorporate a null expectation can reveal the importance of history and chance in the evolutionary process, and allow a more precise understanding of the contemporary evolution of adaptations.

The importance of adaptation during invasion

In theory, natural selection during colonization and range expansion could be critical to invasion success (Carroll and Dingle 1996, Reznick and Ghalambor 2001, Sakai et al. 2001, Lee 2002). The ability of a species to respond to selection is thought to be an important determinant of geographic range (Antonovics 1976b, Kirkpatrick and Barton 1997, Holt et al. 2005a). Invasions thus represent an interesting challenge to the limits of natural selection as long-distance (often trans-oceanic) dispersal causes colonizing

genotypes to abruptly experience an environment that may differ dramatically from their place of origin. Thus, the transition to invasiveness could be constrained or delayed if colonizing genotypes are initially maladapted to their surroundings (Holt et al. 2005a). For invasions whose historical dynamics have been documented, the spread of introduced populations often starts out slowly before undergoing a rapid increase, resulting in a lag-phase between establishment and range expansion (Mack et al. 2000). While the lag phase may have multiple explanations, including the waiting time for pre-adapted genotypes to colonize, several studies suggest it may result from the time required to evolve adaptations to the new selective environment (Sakai et al. 2001, Lee 2002). Strong selection on the colonizing propagule pool is likely to truncate the phenotypic distribution of the establishing population (Simons 2003), yielding not only a shift in trait mean but also a sharp reduction in N_e, the effective population size. Indeed, a lag time following introduction may actually be *caused* by intense selection, as the few selected survivors begin the process of exponential growth from a meager population size. Thus selection during establishment may generate different genotypes as well as different numerical dynamics, relative to the native range (Antonovics 1976a).

Initial establishment may be facilitated by phenotypic plasticity (Baker 1974, Richards et al. 2006), with selection gaining importance later as populations reorganize their genetic variance through repeated introductions. Multiple introductions from the native range can create genetic admixture within introduced populations, that may influence the process of adaptation following establishment. Mating between previously isolated gene pools can produce recombinant genotypes that may be entirely absent from the native range (de la Vega et al. 1991, Ellstrand and Schierenbeck 2000). Such

admixture can produce transgressive segregation and enhance the response to local selection, setting the stage for an adaptive radiation across the introduced range (Kolbe et al. 2004, Novack and Mack 2005). However, the consequences of admixture also predict a null outcome – a mixture of phenotypes that are a weighted average of the source populations (Kolbe et al. 2007).

What traits are likely to be under selection during invasion, and what is the empirical evidence for phenotypic divergence? Adaptation for invasiveness should be operationally definable (Antonovics 1976a), yet there are currently few generalizations from empirical studies. Forces of selection may relate to changes in the abiotic components of the species' physiological niche, such as temperature, precipitation, or growing season length (Peterson and Vieglais 2001, Holt et al. 2005a). A response to physiological selection may be evident in clinal patterns of quantitative traits across the introduced range. Latitudinal clines in body size have been observed among introduced populations of house sparrows (Johnston and Selander 1964) and fruit flies (Huey et al. 2000). In invasive plants, latitudinal clines have been reported for biomass, flowering time, and fecundity (Weber and Schmid 1998, Kollmann and Banuelos 2004, Maron et al. 2004, Leger and Rice 2007). Because of the covariance between latitude and many aspects of the abiotic environment (e.g., temperature, growing season length), clines of traits with latitude may suggest physiological adaptation during the course of invasion; however, in most studies this cannot be unambiguously distinguished from neutral demographic scenarios (see below).

Selection may also act more broadly across invasive populations. Many invaders are ecological opportunists, and changes in disturbance regime may select for shifts towards

a "faster" or more *r*-selected life history (Lewontin 1965, Baker 1974, Sakai et al. 2001, Lambrinos 2004). Early reproduction and allocation for increased offspring number is predicted for colonizing species experiencing high levels of disturbance or environmental unpredictability (Lewontin 1965, Grime 1977). Demographic uncertainty during the founding of new populations also predicts invasion will select for self-compatibility, asexuality, or other means of reproductive assurance (i.e., Baker's Law: Baker 1955, Kolar and Lodge 2001). Finally, biotic interactions such as predation, pathogen attack, or mutualisms may also drive phenotypic evolution in the introduced range. In particular, the impact of natural enemies often differs systematically between the native and introduced ranges (reviewed in Colautti et al. 2004). The relaxation of selection from enemies may favor the evolution of traits conferring increased growth, competitive ability, or reproduction at the expense of defense (Blossey and Notzfold 1995). While evidence for the evolutionary consequences of enemy escape remains equivocal (Willis et al. 2000, Siemann and Rogers 2001, Blair and Wolfe 2004, Bossdorf et al. 2004, Wolfe et al. 2004, Agrawal et al. 2005, Genton et al. 2005a), few definitive tests have been conducted (Colautti et al. 2004, Bossdorf et al. 2005).

Taken together, there is good reason to believe that responding to selection may be a common or even prerequisite component of successfully established, self-perpetuating, and expanding populations. However, the role of selection has rarely been explicitly tested against a neutral model of phenotypic evolution.

The problem: $\Delta z \neq h^2 S$

To understand the extent to which invasion involves adaptive evolution, we must study changes in the distribution of genetically-based phenotypic traits associated with fitness. Traits affecting fitness often have a polygenic basis and can be analyzed using the methods of quantitative genetics (Lynch and Walsh 1998). Quantitative genetic designs parse out genetic and environmental influences on phenotypic traits by raising pedigreed individuals, usually full or half-sib families, under controlled conditions or in common garden experiments. Quantitative genetic theory gives us an inferential framework for analyzing the response of a trait (z) to selection, known originally as the breeder's equation: $R = h^2S$, where R is the response to selection (the cross-generational change in mean phenotype, Δz, before and after selection), h^2 is the narrow sense heritability (the proportion of total phenotypic variance attributable to additive genetic effects), and S is the selection differential (the within-generation difference in phenotype before and after selection). The implication of the breeder's equation is straightforward: if both h^2 and S are non-zero, the phenotypic distribution will shift in response to selection on the trait, assuming no countering effects from genetic correlations. By extension, populations that occupy different environments and are observed to differ genetically for a trait ($\Delta z > 0$) are often interpreted as having diverged in response to selection, although this may leave the agent(s) of selection unspecified .

The quantitative genetics of population divergence is naturally finding applications in invasion biology. Many studies use families collected from across the native and introduced ranges, rear them in a common environment, and demonstrate significant genetically-based divergence in trait means (reviewed in Bossdorf et al. 2005). In some cases, a history of adaptive evolution is a reasonable, though uncertain interpretation.

However, it is difficult to interpret genetically-based phenotypic divergence as a response to selection without additional information on the history of the samples.

The problem with adopting an adaptationist view is that invasions are inherently non-equilibrium demographic situations where the influences of historical events and stochastic processes are prominent. Founder effects (Eckert et al. 1996, Cabe 1998, Amsellem et al. 2000, Tsutsui et al. 2000, Taylor and Keller 2007), admixture among multiple introductions (de la Vega et al. 1991, Guinand and Esteal 1996, Neuffer and Hurka 1999, Gaskin and Schaal 2002, Kolbe et al. 2004, Durka et al. 2005, Wares et al. 2005, Taylor and Keller 2007), and metapopulation dynamics (McCauley et al. 1995, McCauley et al. 2003) all point to demographic perturbations causing changes in genetic diversity during invasion as the rule rather than the exception.

An evolutionary consequence of chance demographic events during invasion is the sampling, introduction, and redistribution of alleles from the native range that have potentially diverse evolutionary histories (Figure 1). An invasive species in its native range often displays population structure that has developed over its evolutionary history, both for neutral loci and those that code for quantitative traits. Structure in the native range then interacts with sampling processes to determine which alleles and traits are represented among the genotypes of each invasion inoculum. The number of inocula, their points of introduction and their genetic similarity to each other can have important consequences not only for the evolutionary trajectory of newly established populations, but also for our inference of what forces have shaped the current population structure. Differentiation among newly founded populations will often reflect the sources and patterns of colonization, rather than an equilibrium between genetic drift and gene flow

among established populations (Whitlock and McCauley 1999). For example, a clinal pattern of apparent isolation by distance may reflect the establishment of genetically distinct sources at different points of introduction, where genotypes in close proximity are more closely related than those farther away. This has clearly occurred during the invasion of North America by *Silene latifolia*, where haplotypes from divergent European clades established at geographically separated sites (Figure 2). Demographic processes such as these are capable of generating clines in both neutral markers and phenotypic traits. Therefore, even phenotypic clines should be interpreted cautiously, especially when they occur over geographic gradients (i.e., latitude) as well as environmental ones.

Among the growing number of studies testing hypotheses about phenotypic evolution during invasion, very few incorporate a null expectation that accounts for chance sampling of evolutionary history from the native range (but see Maron et al. 2004, Kliber and Eckert 2005, Kolbe et al. 2007). This approach is absolutely essential, since the loci contributing to variance in quantitative traits experience the same impact from chance events on average as neutral loci (Lande 1992, Whitlock 1999, Merila and Crnokrak 2001, McKay and Latta 2002). To see this clearly, consider a scenario of stochastic phenotypic evolution that is probably quite common during invasion (Figure 3). In the first scenario (Figure 3A), a species in its native range is structured into several discrete demes that have diverged for some quantitative trait. If the process of colonization samples only a subset of the native demes, then the mean phenotype of the colonists will be shifted, relative to the mean across the native range. In the second scenario (Figure 3B), we assume the same native range structure, but consider that sampling during invasion may change the relative frequencies of the demes, again resulting in a shift in

the phenotypic mean of the colonists relative to the native range. Thus, stochastic events may determine which individuals contribute to the invasion, with the phenotypes of those individuals reflecting a complex history of selection and drift in the native range. In these cases, a quantitative genetic comparison of the native and introduced ranges may reveal phenotypic divergence, but $\Delta z \neq h^2 S$. One way to account for stochastic divergence owing to sampling effects is to estimate phenotypic divergence directly between ancestral lineages in the native range and their descendants in the introduced range (Figure 3C). Fortunately, with the advent of high resolution molecular markers and analytical techniques, establishing ancestor-descendent relationships at the intraspecific level is increasingly feasible. This presents the challenge of integrating knowledge from neutral markers about the demographics of invasion with hypotheses about adaptive evolution of the phenotype.

Testing phenotypic evolution against neutral expectations

Separating the effects of history and chance from adaptation is a significant challenge (Barrett 2000), but is experimentally tractable. The key observation is that neutral loci are subject to the demographic effects of founder effect, genetic drift, and gene flow, while loci contributing to quantitative traits are subject to these same demographic effects, plus the action of natural selection. Thus, the genetic contributions of history and chance can be jointly controlled for by incorporating neutral molecular variation into experimental designs. To address this issue, we offer two experimental designs that combine neutral and quantitative genetic information to parse the relative roles of selection following introduction versus other evolutionary forces during species invasion.

Conceptual design 1: ancestor-descendent comparisons

One way to partition the effects of history and chance from selection is by directly assigning invasive genotypes to their putative source gene pools using molecular methods. If probable ancestor-descendent relationships can be identified, then ancestry information can be incorporated into experimental designs as a fixed main effect in a standard analysis of variance framework. A quantitative comparison of invasive and native range genotypes, identified as descendants from the same ancestral deme, can provide crucial insight into the causes of phenotypic divergence (Figure 3C).

*1.1. The Approac*h

Two molecular approaches show promise for generating ancestor-descendent comparisons: phylogenetics and multi-locus assignment methods. In the phylogenetic approach, DNA sequencing of individuals from native and introduced ranges identifies haplotypes and their relatedness (Schaal et al. 2003). If sufficient phylogeographic structure exists in the haplotype network, then invaders can be traced back to their ancestral sources by searching for shared haplotypes between ranges (Gaskin and Schaal 2002, Saltonstall 2002, Kolbe et al. 2004, Gaskin et al. 2005, Taylor and Keller 2007). Phylogenies from sequence data provide the most complete genealogical information for resolving ancestor-descendent relationships because they produce historically ordered relationships and in general are less prone to homoplasy than marker loci; however they may lack resolution depending on the rate of mutation, especially in plants. Since it is a priority to connect introduced genotypes as closely as possible to their native range

ancestors (Kolbe et al. 2007), anticipated phylogenetic resolution should be carefully considered.

As an alternative to phylogenetic methods, marker loci can be used to connect introduced genotypes to their ancestral gene pools. While most marker loci (e.g., AFLPs, allozymes, microsatellites, RAPDs) are not ideal for inferring phylogenetic relationships (due to problems of non-homology among similarly sized fragments), they are ideal for generating multi-locus genotypes for use in assigning sources under non-equilibrium conditions (Davies et al. 2000). Assignment methods are a growing class of analyses that share the general feature of using individual multi-locus genotypes to assign probability of membership to different putative sources. The approach is similar to forensic or parentage analyses (Manel et al. 2005), and is based on the idea that at any given locus, an individual has a probability of matching one or more sources. By combining information from many independent loci, these probabilities can be calculated with enough precision to exclude all but the most likely source(s).

Assignment methods can be divided into two types, distinguished by the assumptions they make about the source populations. First, assignment *tests* group individuals with their most likely population, chosen from an a priori group of putative source populations sampled by the investigator (Rannala and Mountain 1997). Simulations have shown assignment tests have considerable statistical power for producing correct assignments, provided the genetic structure among the set of source populations is not too low (Waples and Gaggiotti 2006) and of course that the source population is included in the sample. Second, a parallel approach uses model-based clustering to identify genetic structure and genotype membership while making few assumptions about the source populations

(Pritchard et al. 2000, Corander et al. 2003, Falush et al. 2003). Clustering methods work by utilizing information on the allelic associations (i.e., statistical linkage disequilibria) that develop within isolated demes. These methods first solve for the most likely number of genetically distinct demes, given a dataset of multi-locus genotypes, and then assign individuals to demes probabilistically based on the estimated allele frequencies (Pritchard et al. 2000). Clustering methods usually perform well when gene flow among populations is low to moderate and with mixed results when the level of gene flow is higher and hence structure is more cryptic (Pritchard et al. 2000, Manel et al. 2005, Waples and Gaggiotti 2006). Both phylogenetic surveys and multi-locus genotyping are routinely being used to examine the introduction of genetic diversity during invasion, though neither has been well integrated with studies of phenotypic evolution.

Once invasive genotypes are assigned to putative sources, this information can be incorporated into quantitative genetic experimental designs. To illustrate this, consider a straightforward test for phenotypic change during invasion by phenotyping individuals from multiple populations distributed across the native and introduced ranges (Blair and Wolfe 2004, Leger and Rice 2007). The hypothesis of interest is usually whether ranges have diverged in phenotype, and secondarily, if populations within ranges show divergence. Including the ancestral deme as a fixed effect in ANOVA (with *n* levels corresponding to the number of lineages or demes common to both ranges) permits the phenotypic divergence caused by shifts in deme frequencies during invasion to be statistically controlled. The principal test of the fixed effect of range then takes on a new interpretation. Phenotypic divergence between ranges that persists after controlling for

divergence among demes lends strong support to selection driving the change, or at least phenotypic change that has occurred within the introduced range following colonization. This design is flexible to a variety of biological and statistical outcomes (Figure 4). For example, not all demes introduced will experience selection following introduction, or the strength of selection and magnitude of response may vary among demes. This may occur because demes differ in phenotypic distance from the new adaptive optima in the introduced range ('pre-adaptation'), because demes are introduced to locations that differ in the strength or direction of selection, or because demes contain different amounts of genetic variance and hence differ in their evolutionary potential. Differential evolution of demes is captured by the range*deme interaction effect, indicating that adaptive evolution during invasion proceeded as a deme-specific process (Figure 4).

A slightly different approach is appropriate when testing the hypothesis of selection generating clinal patterns. In this case, multi-locus genotypes can be evaluated on a more continuous scale using distance-based ordination methods such as principal coordinates analysis (PCoA). Relatedness among invasive genotypes, including their affinity to native range genotypes, can then be described by their relative positions in multi-locus genetic space. Such ordination would be particularly suited to providing a null expectation when attempting to infer histories of clinal selection in the introduced range. For example, one or more axes from the PCoA could be used as independent variables in a multiple regression model along with the putative environmental gradients influencing trait evolution. Significant covariance between a trait and an environmental gradient, while

holding the effects of relatedness constant, would be compelling evidence for the action of clinal selection rather than a correlated effect of spatial genetic structure.

1.2. Issues and caveats

Several issues arise regarding these analyses, some minor and others more substantial. First, the majority of phylogenetic studies preferentially use organelle genomes (mitochondria in animals and chloroplasts in plants) because of their high mutation rate and lack of frequent recombination. While mitochondrial or chloroplast DNA sequences may be excellent for inferring demographic history, these histories may not reflect the history of the nuclear genome (which is presumably responsible for the majority of quantitative trait variation). To the extent that the nuclear and organelle genomes have experienced different histories either prior to or during the invasion, phylogenetic information from the organelles may be inappropriate neutral controls for phenotypic evolution. Admixture during invasion may exacerbate this issue, as nuclear-cytoplasmic disequilibria may dissipate. Advances in nuclear gene phylogeography offers a solution to this problem (Hare 2001, Gaskin and Schaal 2002). Finally, gene coalescence is a stochastic process, which makes inferences from a single locus subject to considerable variation around the "mean" demographic history of the species. Therefore, to obtain robust inferences of demographic history, it is preferable to use multiple loci.

A caveat is also necessary for the assignment method approach when admixture is present during invasion. Mating among individuals from divergent demes generates genotypes composed of loci from multiple sources. Although model-based clustering has been designed to take admixture into account, assignment methods perform best when

admixture is recent, but can be extended if data on physical linkage among the markers are available (Falush et al. 2003). Assignment methods lose effectiveness when admixture is very extensive, but this is entirely appropriate because there is no single ancestral deme to which genotypes should assigned. Even in this case, recombinant genotypes could be compared to the weighted mean of the sources from which they were derived.

A final issue is that while introducing a lineage effect into an ANOVA provides an appropriate null model for adaptation, rejecting the null does not reveal the phase of the invasion during which the response to selection occurred. Selection during establishment and selection to local conditions during expansion would both cause the phenotypic mean of invaders to deviate from their native range ancestors, leaving the precise timing of the selective events open to further experimentation.

Conceptual design 2: Q_{ST} vs. F_{ST}

Another way to partition history and chance from selection is to make the population the unit of comparison and describe the genetic variance at neutral loci relative to the variance in the phenotypic traits. This approach is appealing because it is directly tied to the methodologies of traditional population genetic surveys, which analyze hierarchical genetic structure within and between populations. Since it is frequency based and not dependent on individual genotypes, it is appropriate for any demographic scenario including ongoing admixture of differentiated demes. Finally, the hierarchical design allows an overall test for adaptation between ranges, as well as adaptive differentiation among populations following invasion.

2.1. The Approach

The neutral theory of phenotypic evolution poses that the additive genetic variance for a trait, σ^2_g can be partitioned into within $\sigma^2_{g(w)}$ and between $\sigma^2_{g(b)}$ population variance components in a manner analogous to single locus population genetics (Wright 1951, Lande 1992, Whitlock 1999). Ignoring for a moment the effects of mutation and selection, the balance between drift and gene flow will result in the hierarchical partitioning of σ^2_g in proportion to Wright's fixation coefficient for neutral loci, F_{ST} (Wright 1951). Based on results of the neutral phenotypic theory, the analogous fixation coefficient for quantitative traits is $Q_{ST} = \sigma^2_{g(b)} / (\sigma^2_{g(b)} + 2\ \sigma^2_{g(w)})$ (Spitze 1993). The important result is that, in the absence of selection, Q_{ST} and F_{ST} estimated from a set of populations are expected to have closely similar values (Whitlock 1999, Merila and Crnokrak 2001, McKay and Latta 2002). When F_{ST} is estimated from neutral loci distributed across the genome, it estimates the sum of the demographic processes that contribute to divergence, such as founder effect and genetic drift, but is less affected by the force of selection. In contrast, Q_{ST} summarizes divergence at loci affecting a phenotypic trait and will be affected by the same demographic forces as neutral loci, as well as potentially influenced by a history of selection on the phenotype. Therefore, F_{ST} provides the null expectation for divergence caused by chance, against which divergence at putatively selected traits (Q_{ST}) can be compared. If quantitative traits are evolving neutrally, their divergence will approximate that for neutral loci, and $Q_{ST} = F_{ST.}$ Therefore, $Q_{ST} > F_{ST}$ is evidence of a history of adaptive divergence, while $Q_{ST} < F_{ST}$

indicates a history of stabilizing selection (Merila and Crnokrak 2001, McKay and Latta 2002).

By assaying both types of genetic variation, it is possible to make this comparison at different hierarchical levels (e.g., among ranges, among populations within ranges) to test hypotheses regarding the history of selection and adaptive evolution. There are a number of historical scenarios for species invasions that can be disentangled by this method (Table 1). For example, invasion may have involved selection for "weedy" traits that promote productivity early in the life history, in which case $Q_{ST} > F_{ST}$ for invasive versus native range populations. A second possibility is that native populations are locally adapted to environments within the native range, but admixture during the invasion process has scrambled the distribution of adaptive traits within the introduced range. In this case, $Q_{ST} > F_{ST}$ among native populations, but for invasive populations, $Q_{ST} = F_{ST}$. As before, this conceptual design should be robust to a wide variety of possible outcomes (Table 1), making it a powerful means to decouple neutral phenotypic evolution from adaptation during invasion.

2.2 Issues and caveats

Comparisons of Q_{ST} and F_{ST} are currently an active area of research in population genetics, from both empirical and theoretical perspectives. While the behavior of the estimators and the sensitivity of their assumptions still receive attention, several recent insights are relevant to the conceptual design proposed here. First, Q_{ST} is formally a partitioning of additive genetic variance, which requires intensive breeding designs capable of isolating just the additive effects of genes (Lynch and Walsh 1998). Less

complicated breeding designs may produce variance components that include some amount of non-additive genetic effects such as dominance or epistasis. While simulation studies suggest the general effect of dominance variance may be to lower Q_{ST} and therefore avoid type I errors (Goudet and Buchi 2006), this area of research warrants additional attention. Another assumption that is infrequently discussed is that the rate of mutation is significantly lower than the migration rate and that the model of evolution for neutral and quantitative traits is similar (Hendry 2002). However, the mutation rate of hypervariable markers (e.g., microsatellites) may be quite different than that for quantitative traits and high enough to exert a downward bias on F_{ST}, potentially leading to type I errors of falsely rejecting the null hypothesis $F_{ST} = Q_{ST}$. A final issue relates to statistical power. The power of Q_{ST} estimates is affected by the number of populations sampled; some simulations suggest fewer than 20 populations may compromise the ability to detect the signature of selection (O'Hara and Merila 2005, Goudet and Buchi 2006). Similarly, studies of invasion that wish to make statistical statements about phenotypic divergence between the native versus introduced ranges also require a large and unbiased sample of populations for reliable inference (R.I. Colautti, J.L. Maron, and S.C.H. Barrett, unpublished manuscript). The need for many populations when estimating Q_{ST} makes for potentially large experimental designs, though Goudet and Buchi (2006) suggest replication within populations can be somewhat relaxed. For example, if we replicated the native and introduced ranges with 20 populations each, and performed a modest sized paternal half-sib design within each population (ex., 10 sires each mated to 2 dams and raising 5 offspring from each family), the experimental design would involve

phenotyping 4000 individuals. Less replication intensive designs exist (e.g., full-sib families), but will involve some amount of non-additive genetic variance.

Conclusions

Invasive species have great potential to reveal the process of adaptive evolution, but evidence for selection must be evaluated relative to null expectations based on neutral phenotypic evolution. The experimental designs presented here are meant to further this goal. The key advancement is that by incorporating demographic insights gained from neutral molecular markers, experiments can be designed that isolate demographic history and chance sampling events from the history of selection on the phenotype. We do not regard these as the only methods of accounting for demographic effects when studying adaptation, but rather view them as promising examples of the more general approach of incorporating neutral expectations for phenotypic change. Each of the conceptual designs presented has strengths and weaknesses, and which may be appropriate will depend on the exigencies of the study system, prior knowledge from others sources of evidence, and the resources available. It is also important to keep in mind that these approaches are necessarily statistical in nature (i.e., do not reveal the agents of selection), and are perhaps best used as a first step in the study of adaptation during invasion. Field experiments such as reciprocal transplants among the putative selective environments, coupled with direct measurements of contemporary selection on the traits, would complement the experiments described here. Nevertheless, some of the most interesting questions in biological invasions involve inferences of past selection shaping phenotypic

distributions. Tests of adaptation against null models of neutral evolution make this possible.

Acknowledgments

We appreciate the thoughtful comments of Janis Antonovics, Laura Galloway, and the Taylor lab group. This paper was supported by NSF DEB-0608358 Doctoral Dissertation Improvement and NSF DEB-444176.

References

Agrawal A.A., Kotanen P.M., Mitchell C.E., Power A.G., Godsoe W. & Klironomos J. (2005). Enemy release? An experiment with congeneric plant pairs and diverse above- and belowground enemies. *Ecology*, 86, 2979-2989.

Alonso D., Etienne R.S. & McKane A.J. (2006). The merits of neutral theory. *Trends Ecol. Evol.*, 21, 451-457.

Amsellem L., Noyer J.L., Le Bourgeois T. & Hossaert-McKey M. (2000). Comparison of genetic diversity of the invasive weed *Rubus alceifolius* Poir. (Rosaceae) in its native range and in areas of introduction, using amplified fragment length polymorphism (AFLP) markers. *Mol. Ecol.*, 9, 443-455.

Antonovics J. (1976a). The input from population genetics: "the new ecological genetics". *Syst. Bot.*, 1, 233-245.

Antonovics J. (1976b). The nature of limits to natural selection. *Ann. Miss. Bot. Gard.*, 63, 224-247.

Baker H.G. (1955). Self-compatibility and establishment after "long-distance" dispersal. *Evolution*, 9, 347-348.

Baker H.G. (1974). The evolution of weeds. *Ann. Rev. Ecol. Syst.*, 5, 1-24.

Baker H.G. & Stebbins G.L. (1965). *The genetics of colonizing species*. Academic Press, New York.

Barrett S.C.H. (2000). Microevolutionary influences of global changes on plant invasions. In: *Invasive species in a changing world* (eds. Mooney HA & Hobbs R). Island Press Covelo, CA, pp. 115-139.

Blair A.C. & Wolfe L.M. (2004). The evolution of an invasive plant: an experimental study with *Silene latifolia*. *Ecology*, 85, 3035-3042.

Blossey B. & Notzfold R. (1995). Evolution of increased competitive ability in invasive nonindigenous plants: a hypothesis. *J. Ecol.*, 83, 887-889.

Bossdorf O., Auge H., Lafuma L., Rogers W.E., Siemann E. & Prati D. (2005). Phenotypic and genetic differentiation between native and introduced plant populations. *Oecologia*, 144, 1-11.

Bossdorf O., Prati D., Auge H. & Schmid B. (2004). Reduced competitive ability in an invasive plant. *Ecol. Lett.*, 7, 346-353.

Cabe P.R. (1998). The effects of founding bottlenecks on genetic variation in the European starling (*Sturnus vulgaris*) in North America. *Heredity*, 80, 519-525.

Carroll S.P. & Dingle H. (1996). The biology of post-invasion events. *Biol. Cons.*, 78, 207-214.

Colautti R.I., Ricciardi A., Grigorovich I.A. & MacIsaac H.J. (2004). Is invasion success explained by the enemy release hypothesis? *Ecol. Lett.*, 7, 721-733.

Corander J., Waldmann P. & Sillanpaa M.J. (2003). Bayesian analysis of genetic differentiation between populations. *Genetics*, 163, 367-374.

Davies N., Villablanca F.X. & Roderick G.K. (2000). Determining the source of individuals: multilocus genotyping in nonequilibrium population genetics. *Trends Ecol. Evol.*, 14, 17-21.

de la Vega M., Garcia P. & Allard R. (1991). Multilocus genetic structure of ancestral Spanish and colonial California populations of *Avena barbata*. *Proc. Natl. Acad. Sci. U.S.A.*, 88, 1202-1206.

Durka W., Bossdorf O., Prati D. & Auge H. (2005). Molecular evidence for multiple introductions of garlic mustard (*Alliaria petiolata*, Brassicaceae) to North America. *Mol. Ecol.*, 14, 1697-1706.

Eckert C.G., Manicacci D. & Barrett S.C.H. (1996). Genetic drift and founder effect in native versus introduced populations of an invading plant, *Lythrum salicaria* (Lythraceae). *Evolution*, 50, 1512-1519.

Ellstrand N.C. & Schierenbeck K.A. (2000). Hybridization as a stimulus for the evolution of invasiveness. *Proc. Natl. Acad. Sci. U.S.A.*, 97, 7043-7050.

Falush D., Stephens P. & Pritchard J.K. (2003). Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics*, 164, 1567-1587.

Gaskin J.F. & Schaal B.A. (2002). Hybrid *Tamarix* widespread in U.S. invasion and undetected in native Asian range. *Proc. Natl. Acad. Sci. U.S.A.*, 99, 11256-11259.

Gaskin J.F., Zhang D.-O. & Bon M.-C. (2005). Invasion of *Lepidum draba* (Brassicaceae) in the western United States: distributions and origins of chloroplast DNA haplotypes. *Mol. Ecol.*, 14, 2331-2341.

Genton B., Kotanen P., Cheptou P.-O., Adolphe C. & Shykoff J. (2005). Enemy release but no evolutionary loss of defence in a plant invasion: an inter-continental reciprocal transplant experiment. *Oecologia*, 146, 404-414.

Goudet J. & Buchi L. (2006). The effects of dominance, regular inbreeding and sampling design on Qst, an estimator of population differentiation for quantitative traits. *Genetics*, 172, 1337-1347.

Gould S.J. & Lewontin R.C. (1979). The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme. *Proc. Roy. Soc. Lond. Ser. B.*, 205, 581-598.

Grime J.P. (1977). Evidence for the existence of three primary strategies in plants and its relevance to ecological and evolutionary theory. *Am. Nat.*, 111, 1169-1194.

Guinand B. & Esteal S. (1996). Multivariate patterns of genetic differentiation support complex colonization schemes in *Bufo marinus* populations. *Evolution*, 50, 944-951.

Hare M.P. (2001). Prospects for nuclear gene phylogeography. *Trends Ecol. Evol.*, 16, 700-706.

Hendry A.P. (2002). $Q_{ST} > = <> < F_{ST}$? *Trends Ecol. Evol.*, 17, 502.

Holt R.D., Barfield M. & Gomulkiewicz R. (2005). Theories of niche conservatism and evolution: Could exotic species be potential tests? In: *Species invasions: insights into ecology, evolution, and biogeography* (eds. Sax DF, Stachowicz JJ & Gaines SD). Sinauer Associates, Inc. Sunderland, MA, pp. 259-290.

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

Johnston R.F. & Selander R.K. (1964). House sparrows: rapid evolution of races in North America. *Science*, 144, 548-550.

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

Kliber A. & Eckert C.G. (2005). Interaction between founder effect and selection during biological invasion in an aquatic plant. *Evolution*, 59, 1900-1913.

Kolar C.S. & Lodge D.M. (2001). Progress in invasion biology: predicting invaders. *Trends Ecol. Evol.*, 16, 199-204.

Kolbe J.J., Glor R.E., Schettino L.R., Lara A.C., Larson A. & Losos J.B. (2004). Genetic variation increases during biological invasion by a Cuban lizard. *Nature*, 431, 177-181.

Kolbe J.J., Larson A. & Losos J.B. (2007). Differential admixture shapes morphological variation among invasive populations of the lizard *Anolis sagrei*. *Mol. Ecol.*, 16, 1579-1591.

Kollmann J. & Banuelos M.J. (2004). Latitudinal trends in growth and phenology of the invasive alien plant *Impatiens glandulifera* (Balsaminaceae). *Div. Dist.*, 10, 377-385.

Lambrinos J.G. (2004). How interactions between ecology and evolution influence contemporary invasion dynamics. *Ecology*, 85, 2061-2070.

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

Lee C.E. (2002). Evolutionary genetics of invasive species. *Trends Ecol. Evol.*, 17, 386-391.

Leger E.A. & Rice K.J. (2007). Assessing the speed and predictability of local adaptation in invasive California poppies (*Eschscholzia californica*). *J. Evol. Biol.*, 20, 1090-1103.

Lewontin R.C. (1965). Selection for colonizing ability. In: *The genetics of colonizing species* (eds. Baker HG & Stebbins GL). Academic Press New York.

Lynch M. & Walsh B. (1998). *Genetics and the analysis of quantitative traits*. Sinauer Associates, Sunderland, MA.

Mack R.N., Simberloff D., Lonsdale W.M., Evans H., Clout M. & Bazzaz F.A. (2000). Biotic invasions: epidemiology, global consequences, and control. *Ecol. Appl.*, 10, 689-710.

Manel S., Gaggiotti O. & Waples R.S. (2005). Assignment methods: matching biological questions with appropriate techniques. *Trends Ecol. Evol.*, 20, 136-142.

Maron J.L., Vila M., Bommarco R., Elmendorf S. & Beardsley P. (2004). Rapid evolution of an invasive plant. *Ecol. Mon.*, 74, 261-280.

McCauley D.E., Raveill J. & Antonovics J. (1995). Local founding events as determinants of genetic structure in a plant metapopulation. *Heredity*, 75, 630-636.

McCauley D.E., Smith R.A., Lisenby J.D. & Hsieh C. (2003). The hierarchical spatial distribution of chloroplast DNA polymorphism across the introduced range of *Silene vulgaris*. *Mol. Ecol.*, 12, 3227-3235.

McKay J., K, & Latta R.G. (2002). Adaptive population divergence: markers, QTL and traits. *Trends Ecol. Evol.*, 17, 285-291.

Merila J. & Crnokrak P. (2001). Comparison of genetic differentiation at marker loci and quantitative traits. *J. Evol. Biol.*, 14, 892-903.

Neuffer B. & Hurka H. (1999). Colonization history and introduction dynamics of *Capsella bursa-pastoris* (Brassicaceae) in North America: isozymes and quantitative traits. *Mol. Ecol.*, 8, 1667-1681.

Novack S.J. & Mack R.M. (2005). Genetic bottlenecks in alien plant species: Influence of mating systems and introduction dynamics. In: *Species invasions: insights into ecology, evolution, and biogeography* (eds. Sax DF, Stachowicz JJ & Gaines SD). Sinauer Associates, Inc. Sunderland, MA, pp. 201-228.

O'Hara R.B. & Merila J. (2005). Bias and precision in Qst estimates: problems and some solutions. *Genetics*, 171, 1331-1339.

Peterson A.T. & Vieglais D.A. (2001). Predicting species invasions using ecologic niche modeling. *BioScience*, 51, 363-371.

Pritchard J.K., Stephens P. & Donnelly P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, 155, 945-959.

Rannala B. & Mountain J.L. (1997). Detecting immigration using multilocus genotypes. *Proc. Natl. Acad. Sci. U.S.A.*, 94, 9197-9201.

Reznick D.N. & Ghalambor C.K. (2001). The population ecology of contemporary adaptations: what empirical studies reveal about the conditions that promote adaptive evolution. *Genetica*, 112-113, 183-198.

Richards C.L., Bossdorf O., Muth N.Z., Gurevitch J. & Pigliucci M. (2006). Jack of all trades, master of some? On the role of phenotypic plasticity in plant invasions. *Ecol. Lett.*, 9, 981-993.

Sakai A.K., Allendorf F.W., Holt J.S., Lodge D.M., Molofsky J., With K.A., Baughman S., Cabin R.J., Cohen J.E., Ellstrand N.C., McCauley D.E., O'Neil P., Parker I.M., Thompson J.N. & Weller S.G. (2001). The population biology of invasive species. *Ann. Rev. Ecol. Syst.*, 32, 305-332.

Saltonstall K. (2002). Crptic invasion by a non-native genotype of the common reed, *Phragmites australis*, into North America. *Proc. Natl. Acad. Sci. U.S.A.*, 99, 2445-2449.

Schaal B.A., Gaskin J.F. & Caicedo A.L. (2003). Phylogeography, haplotype trees, and invasive plant species. *J. Hered.*, 94, 197-204.

Siemann E. & Rogers W.E. (2001). Genetic differences in growth of an invasive species. *Ecol. Lett.*, 4, 514-518.

Simons A.M. (2003). Invasive aliens and sampling bias. *Ecol. Lett.*, 6, 278-280.

Spitze K. (1993). Population-structure in *Daphnia obtusa* - quantitative genetic and allozyme variation. *Genetics*, 135, 367-374.

Stockwell C.A., Hendry A.P. & Kinnison M.T. (2003). Contemporary evolution meets conservation biology. *Trends Ecol. Evol.*, 18, 94-101.

Taylor D.R. & Keller S.R. (2007). Historical range expansion determines the phylogenetic diversity introduced during contemporary species invasion. *Evolution*, 61, 334-345.

Thompson J.N. (1998). Rapid evolution as an ecological process. *Trends Ecol. Evol.*, 13, 329-332.

Tsutsui N.D., Suarez A.V., Holway D.A. & Case T.J. (2000). Reduced genetic variation and the success of an invasive species. *Proc. Natl. Acad. Sci. U.S.A.*, 97, 5948-5953.

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. *Mol. Ecol.*, 15, 1419-1439.

Wares J.P., Hughes A.R. & Grosberg R.K. (2005). Mechanisms that drive evolutionary change: insights from species introductions and invasions. In: *Species invasions: insights into ecology, evolution, and biogeography* (eds. Sax DF, Stachowicz JJ & Gaines SD). Sinauer Associates, Inc. Sunderland, MA.

Weber E. & Schmid B. (1998). Latitudinal population differentiation in two species of *Solidago* (Asteraceae) introduced to Europe. *Am. J. Bot.*, 85, 1110-1121.

Whitlock M.C. (1999). Neutral additive genetic variance in a metapopulation. *Genet. Res.*, 74, 215-221.

Whitlock M.C. & McCauley D.E. (1999). Indirect measures of gene flow and migration: $F_{ST} \neq 1/(4Nm + 1)$. *Heredity*, 82, 117-125.

Willis A.J., Memmott J. & Forrester R.I. (2000). Is there evidence for the post-invasion evolution of increased size among invasive plant species? *Ecol. Lett.*, 3, 275-283.

Wolfe L.M., Elzinga J.A. & Biere A. (2004). Increased susceptibility to enemies following introduction in the invasive plant *Silene latifolia*. *Ecol. Lett.*, 7, 813-820.

Wright S. (1951). The genetical structure of populations. *Ann. Eug.*, 15.

Table 1. Hypothetical outcomes from a hierarchical comparison of genetic variance in quantitative traits (Q_{ST}) and neutral loci (F_{ST}). The outcomes and their interpretation are a non-exhaustive list assuming a nested experimental design involving comparisons of the native range (N.R.) and introduced range (I.R.) and comparisons among populations within each range.

‡Traits: N.R. vs. I.R.	Q_{ST} vs. F_{ST} between ranges	‡Traits: among N.R. populations	Q_{ST} vs. F_{ST} among N.R. populations	‡Traits: among I.R. populations	Q_{ST} vs. F_{ST} among I.R. populations	Biological interpretation
***	>					Selection during establishment or selection in the I.R. after establishment promotes adaptive divergence between ranges.
***	=	***				Unrepresentative sampling of traits during invasion. Evolution driven by stochastic events.
		***	>		=	Local adaptation in the N.R.; disrupted by stochastic processes during invasion.

		***	>	***	>	Diversifying selection drives local adaptation within both ranges. Invasions as “adaptive radiations”.
n.s.	<	n.s.	<	n.s.	<	Stabilizing selection maintains similar trait means within and between ranges.
		n.s.	<	***	>	Stabilizing selection within the N.R., but release from selective constraint promotes adaptive radiation in the I.R..

‡ Outcomes from an analysis of phenotypic divergence measured in a common environment.

***significant difference in trait means; n.s. no significant difference in trait means. Entries left blank reflect where the specific outcome does not qualitatively affect the overall interpretation.

Figure legends

Figure 1. Path diagram illustrating the contributions of evolutionary history, chance events, and natural selection to the genetics of introduced populations. The genetic diversity present among invasive populations has been shaped by a suite of historical, stochastic, and deterministic forces. Most of the path transitions leading up to and during invasion involve demographic events or intrinsic genetic effects (e.g., mutational input) that influence diversity at both neutral loci and quantitative traits (dotted arrows in diagram). These paths represent the contributions of historical and chance events that may influence quantitative trait evolution during invasion. Transitions involving selection (solid arrows) represent an additional influence on quantitative traits that may work in concert with or in opposition to chance and historical processes. Only a few stages of the invasion process are directly observable by empirical studies (boxes with outlines). Therefore, quantitative genetic studies must statistically control for the influence of unobserved stages in the invasion process when testing for adaptive evolution.

Figure 2 Isolation by distance among chloroplast sequence haplotypes in *Silene latifolia* generated by the spatial pattern of colonization by divergent native range haplotypes. Haplotypes show significant phylogeographic structure in Europe (1A; AMOVA: Φ_{ST} = 0.27; $P < 0.0001$). In North America, colonization of western European haplotypes (blue symbols) occurred primarily in the southeastern U.S.A., while haplotypes from eastern Europe (orange symbols) colonized elsewhere. This colonization pattern lead to a signature of isolation by distance in eastern North America (1B; Mantel's test: r = 0.63, $P < 0.0001$). Figure 1A adapted from Taylor and Keller (2007).

Figure 3 Stochastic sampling during invasion impacts phenotypic traits. Shown are phenotypic distributions for each of several divergent demes in the native range, and those that have invaded a new range. In each scenario, "$\bar{z}$" represents the mean of the phenotypic distribution in the native range, while "$\bar{z}'$ " that in the introduced range. Note that in A and B, continental means are shown, while C displays deme-specific means.

Figure 4 Hypothetical "norm of reaction" showing two possible outcomes of conceptual experimental design 1. In (A), plants from different demes posses divergent phenotypes in the native range prior to invasion (e.g., a significant effect of "Deme" but not "Range" in an ANOVA model). If stochastic sampling has changed the frequencies of demes during invasion, then phenotypic evolution occurs but is attributable to neutral processes. In contrast, (B) shows that invaders from some demes have evolved new phenotypic means, after controlling for differences due to common ancestry (e.g., a significant "Range" or Range*Deme" effect in ANOVA). This suggests invaders have evolved toward new phenotypic optima in response to selection during or since the invasion.

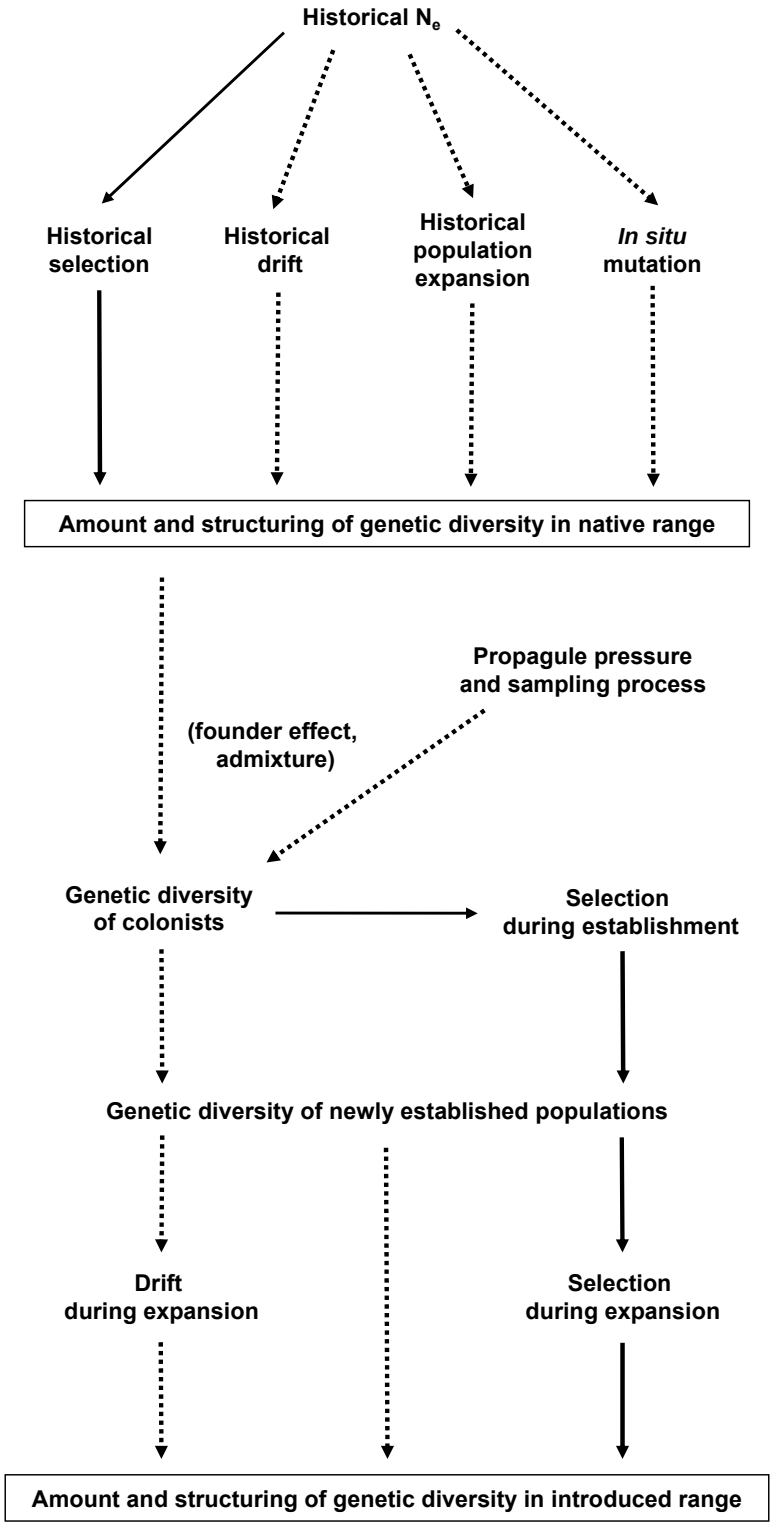
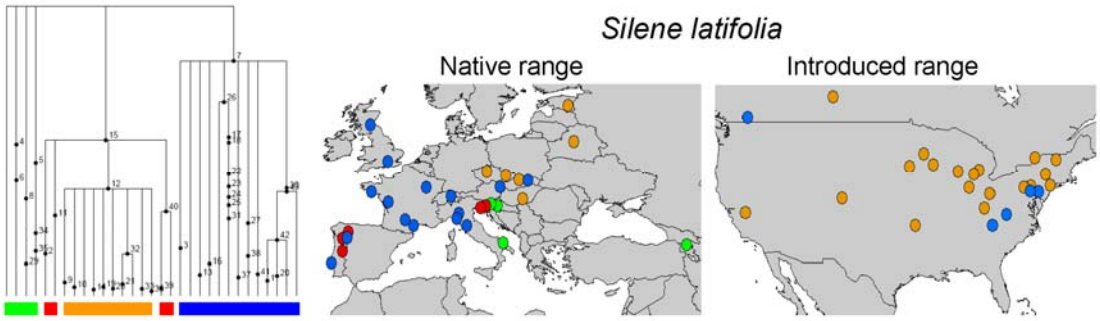

Historical N_e
Historical selection
Historical drift
Historical population expansion
In situ mutation
Amount and structuring of genetic diversity in native range
Propagule pressure and sampling process
(founder effect, admixture)
Genetic diversity of colonists
Selection during establishment
Genetic diversity of newly established populations
Drift during expansion
Selection during expansion
Amount and structuring of genetic diversity in introduced range

Figure 1.

(A)


Silene latifolia
Native range
Introduced range

(B)

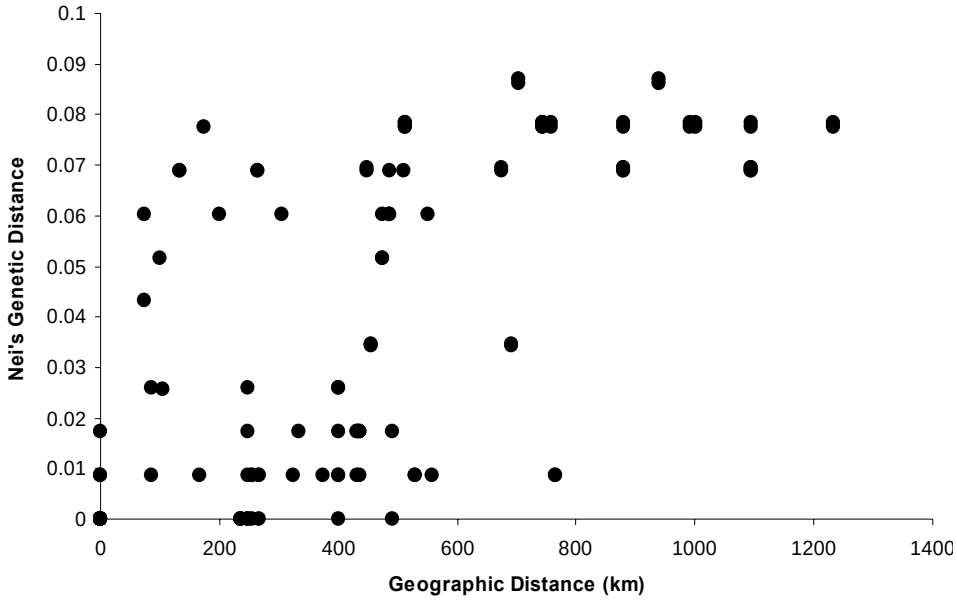

Nei's Genetic Distance
0.1
0.09
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
0
200
400
600
800
1000
1200
1400
Geographic Distance (km)

Figure 2.

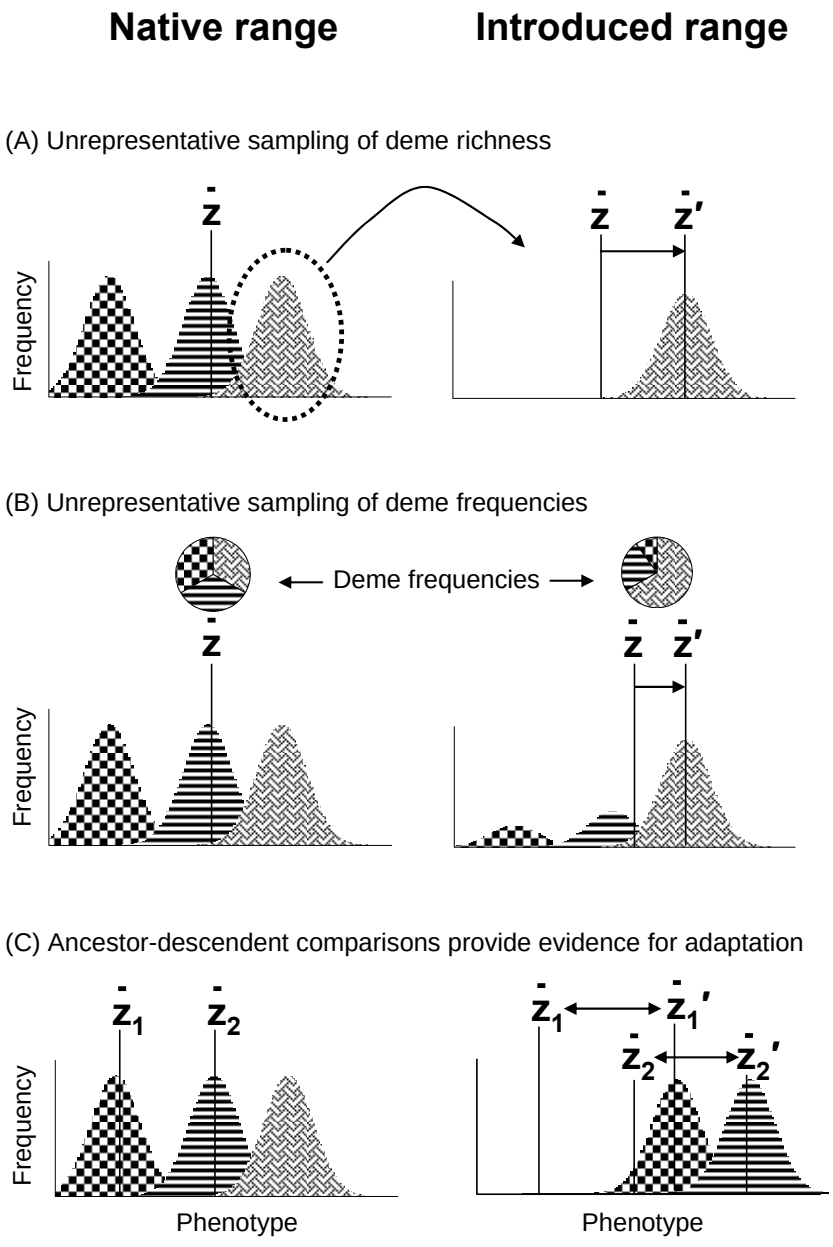

Native range
Introduced range
(A) Unrepresentative sampling of deme richness
Frequency
(B) Unrepresentative sampling of deme frequencies
Deme frequencies
Frequency
(C) Ancestor-descendent comparisons provide evidence for adaptation
Frequency
Phenotype
Phenotype

Figure 3.

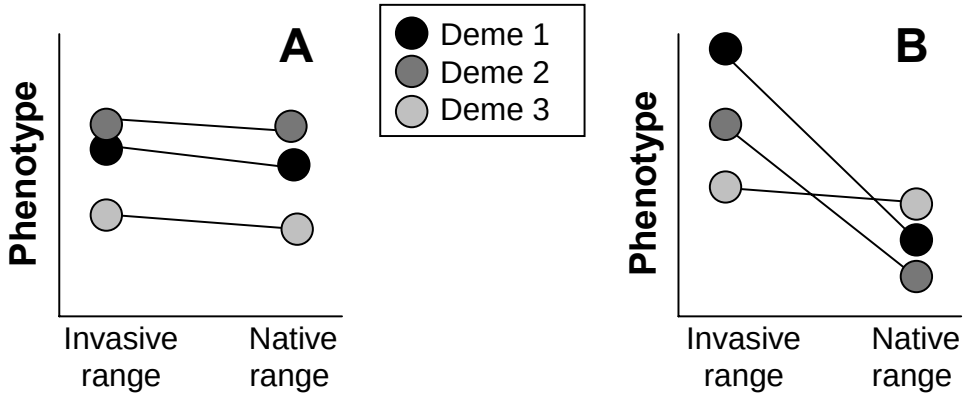

A
B
Deme 1
Deme 2
Deme 3
Phenotype
Phenotype
Invasive range
Native range
Invasive range
Native range

Figure 4.

Chapter 2:

Historical range expansion determines the phylogenetic diversity introduced during contemporary species invasion[2]

[2] Published manuscript: Taylor, D.R. and S.R. Keller. 2007. *Evolution* 61(2):334-345. Taylor was responsible for initiating the study and overseeing data collection. Keller was responsible for compiling the data, and designing and conducting the analyses. Keller and Taylor collaborated on writing the manuscript.

Abstract

For a species rapidly expanding its geographic range, such as during biological invasion, most alleles in the introduced range will have their evolutionary origins in the native range. Yet, how historical processes occurring over evolutionary time in the native range contribute to the diversity sampled during contemporary invasion is largely unknown. We used chloroplast DNA (cpDNA) gene genealogies and coalescent methods to study two congeneric plants, *Silene latifolia* and *S. vulgaris*. We examined how phylogenetic diversity was shaped by demographic growth and historical range expansions in the native European range, and how this history affected the diversity sampled during their recent invasion of North America. Genealogies from both species depart from neutrality, likely as a result of demographic expansion in the ancestral range, the timing of which corresponds to shortly after each species originated. However, the species differ in the spatial distribution of cpDNA lineages across the native range. *Silene latifolia* shows highly significant phylogeographic structure that most likely reflects different avenues of the post-glacial expansion into Northern Europe from Mediterranean refugia. By contrast, cpDNA lineages in *S. vulgaris* have been widely scattered across Europe during, or since, the most recent post-glacial expansion. These different evolutionary histories resulted in dramatic differences in how phylogenetic diversity was sampled during invasion of North America. In *S. latifolia*, relatively few, discrete invasion events from a structured native range resulted in a rather severe genetic bottleneck, but also opportunities for admixture among previously isolated lineages. In *S. vulgaris*, lack of genetic structure was accompanied by more representative sampling of phylogenetic diversity during invasion, and reduced potential for admixture. Our results provide clear insights into how historical

processes may feed forward to influence the phylogenetic diversity of species invading new geographic ranges.

Keywords: chloroplast DNA, coalescent, invasion, mismatch distribution, phylogeography, range expansion, *Silene*

Introduction

The dynamics of a species' geographic range is a central, yet understudied, problem in evolutionary biology. Phylogeography has provided many important insights into geographic range expansions because the pattern of colonization often leaves enduring signatures in the genome (Hewitt 2000; Petit et al. 2003). The evolutionary history of a species prior to expansion plays a central role in deciphering the phylogeography of subsequent range changes because it is precisely this history that determines the distribution of phylogenetic diversity from which colonists are drawn. Chance events during colonization then interact with prior evolutionary history to affect the amount and organization of phylogenetic diversity across a species' range.

Invasive species, defined as those that permanently establish and spread into ecosystems in which they were previously absent (Kolar and Lodge 2001, Lee 2002), provide interesting natural experiments for investigating how history and chance affect phylogenetic diversity during range expansions. Indeed molecular methods have proven indispensable for estimating the genetic diversity of inocula (Amsellem et al. 2000; Tsutsui et al. 2000; Gaskin et al. 2005), identifying the source of invading propagules (Vega et al. 1991; Novak and Mack 2001; Saltonstall 2002), and detecting when lineages that were geographically separate in the native range have come into contact within the area of introduction – so called "admixture" (Guinand and Easteal 1996; Gaskin and Schaal 2002; Kolbe et al. 2004).

However, while the genetics of contemporary invasions are often well characterized, the alleles present in invaders usually originate during, and are affected by, processes that occur over evolutionary time in the native range (Davies et al. 2000). For example, many

species have experienced periods of rapid demographic increases and range expansions during the warm interglacial periods of the Pleistocene (Hewitt 1996, 2000; Taberlet et al. 1998). We use the term "expansion" in this context instead of "invasion" to denote range changes that took place over evolutionary time without human intervention. These historical expansions can have lasting genetic effects (Tajima 1989; Rogers and Harpending 1992; Ray et al. 2003). It is important to recognize, therefore, that the genetic consequences of modern invasions may be affected as much by historical processes that have occurred in the sources from which invasions are drawn, as they are by contemporary processes that influence the number and diversity of invasive propagules (Travisano et al. 1995; Schaal et al. 2003).

The phylogeographic history of a species in its native range determines the distribution of diversity available for invasion to sample. The number of individuals and sources introduced (i.e., propagule pressure, Kolar and Lodge 2001) then determines whether invasion captures this diversity representatively. For example, low propagule pressure would likely sample within a phylogeographic region, producing low diversity inocula. Higher propagule pressure increases the likelihood of sampling across phylogeographic regions, creating opportunities for admixture in the introduced range. Alternatively, invasive propagules originating from an unstructured native range are more likely to be a representative sample of the diversity present, and will be less sensitive to the magnitude of propagule pressure. Thus, to understand factors affecting genetic diversity during invasion, we need to understand the evolutionary history of the distribution from which it is drawn. Otherwise, it will be unclear *why* genetic diversity becomes bottlenecked (or elevated), or if admixture is occurring, whether it is among

lineages with an ancient or a more recent history of isolation. Answers to these questions provide a more detailed view of the invasion process and reveal what implications the sample of diversity has for future evolutionary change.

Gene genealogies are a rich source of information into a species' evolutionary past that could be put to broader use for the study of invasions (Schaal et al. 2003). First, genealogies provide historically ordered alleles, which permit the evolutionary relationships among invasive and non-invasive lineages to be deduced. Second, applying coalescent theory to the distribution of mutations in gene genealogies reveals much about a species' demographic history (Tajima 1989; Hudson 1990; Griffiths and Tavare 1994; Schaal and Olsen 2000). In fact, some of the most significant insights into the genetics of invasions have been achieved by applying coalescent theory to studies of human migration history. Humans are an invasive species with a recent evolutionary history of colonization, range expansion, and rapid demographic growth (Cann et al. 1987; Rogers 1995; Watson et al. 1997; Thompson et al. 2000; Alonso and Armour 2001). These are many of the same phenomena population geneticists are studying during contemporary invasions by non-human species. Yet an evolutionary history of range expansion and population growth has never been explicitly connected to the diversity sampled during invasion.

In this paper, we use gene genealogies to study how evolutionary history affects the genetics of contemporary invasion in two related plant species, *Silene latifolia* and *S. vulgaris*. Both species are weeds of agriculture and disturbed land that have widespread native ranges in Eurasia and have recently invaded North America. Previous studies using morphology and flavonoid genes (Mastenbroek et al. 1983), RAPDs (Vellekoop et

al. 1996), and Y-chromosome sequence data (Ironside and Filatov 2005) support a major genetic subdivision between East-Central and Western Europe in *S. latifolia.* In *S. vulgaris*, population genetic analyses show high levels of cytoplasmic diversity among populations in East-Central Europe (Storchova and Olson 2004) and among North American populations (McCauley et al. 2003). We use genealogical data from throughout the native and introduced ranges of both species to show how evolutionary history in the native range has affected phylogenetic diversity introduced during invasion. We show that despite having similar ecologies and histories of demographic expansions in their native ranges, these two species show markedly different patterns of diversity sampled during invasion of North America. We also show that by taking a coalescent-based approach to the problem, we can begin to unravel some of the manifold complexities of species invading new geographic ranges.

Methods

Study Species

Silene latifolia Poiret (= *S. alba*) and *Silene vulgaris* (Moench) Garcke are short-lived herbaceous perennials with a history of human association. These species are commonly found in disturbed areas such as roadsides, railroad embankments, cultivated fields, and abandoned lots (Baker 1948; Marsden-Jones and Turrill 1957). While both species have similar ecological and life history characteristics, they differ in their breeding system: *S. latifolia* is dioecious (separate male and female individuals), while *S. vulgaris* is gynodioecious (populations are a mixture of female and hermaphrodite individuals).

Silene latifolia and *S. vulgaris* have widespread geographic ranges throughout their native Eurasia. Their ranges broadly overlap, extending from the Atlantic coast (Spain, France, U.K., Ireland) across continental Europe and becoming diffuse in Russia. From the south, the distributions extend from North Africa (Morocco to Egypt) and the Middle East northward to Scandinavia (Atlas Florae Europaeae: Jalas and Suominen 1986). Both are believed to have originated in the Middle East or Mediterranean region, and have since colonized most of Europe, possibly with the spread of agriculture (Baker 1948; Marsden-Jones and Turrill 1957; Mastenbroek et al. 1983; Vellekoop et al. 1996; Runyeon and Prentice 1997). Similarly, both species appear to have been introduced to North America during the late eighteenth or early nineteenth centuries (Cutler 1785; Pursh 1814; McNeill 1977), likely as a contaminant of clover seed or in ship ballast (Martindale 1876). Each invasion began along the eastern seaboard and spread rapidly south and west (Chapter 4; Antonovics et al. 2003). *Silene latifolia* and *S. vulgaris* are designated as invasive by the USDA (http://plants.usda.gov), and considered to be agricultural pests in Canada (Agriculture and Agri-Food Canada).

DNA Sequencing

Samples were collected as seeds or leaf tissue dried on silica gel from across the geographic ranges of the species in Europe and North America (Fig. 1). Genomic DNA was isolated from one haphazardly chosen plant per site using DNeasy plant miniprep kits (Qiagen, Valencia, CA).

To address the phylogenetics of expansion and invasion, we applied predictions from coalescent theory to gene genealogies based on chloroplast DNA (cpDNA) sequence

data. cpDNA sequence data are particularly suited to phylogeographic studies in plants because they 1) provide ordered alleles for inferring ancestor-descendent relationships (Schaal et al. 2003), 2) are essentially asexual, permitting the effects of population history to be inferred from the genealogy without the complicating effects of recombination, 3) are highly polymorphic in these species (McCauley 1994; Ingvärsson et al. 2003; McCauley et al. 2003), and 4) are uniparentally inherited and dispersed through seeds, the same propagules we trace when reconstructing the invasion process. Four cpDNA regions consisting of one intron and three intergenic spacers were PCR amplified and sequenced, following previously reported methods (Ingvärsson and Taylor 2002; Ingvärsson et al. 2003). All sequences for these four cpDNA regions from both species have been deposited in the Genbank database (accession numbers in Appendix A). Sequence alignments were made using the default options in CLUSTALW embedded within MEGA 3.1 (Kumar et al. 2004). Alignments were manually adjusted to minimize the number of distinct gaps caused by insertion/deletion events (indels) and nucleotide polymorphisms in the vicinity of the indels. Data from all four regions were concatenated for a total length of ~1800 bp (1876 bp in *S. latifolia* and 1799 bp in *S. vulgaris*), including indels. These regions were sequenced in a total of 62 *S. latifolia* and 50 *S. vulgaris* individuals.

Indels represented a large portion of the variation in these 4 cpDNA regions. Indels located in introns and intergenic spacers of cpDNA are known to possess useful phylogenetic information, and significantly increase the resolution of phylogenies at the intraspecific level for several taxa (Gielly and Taberlet 1994; Hamilton et al. 2003), including Silene (Ingvärsson et al. 2003). Length variation in the chloroplast genomes of *S. latifolia* and *S. vulgaris* includes single and di-nucleotide repeat motifs, as well as non-

repetitive indels. Repeat indels evolve at a faster rate than non-repetitive indels or nucleotides, and are more prone to homoplasy (Ingvärsson et al. 2003). In contrast, non-repetitive indels are thought to have similar substitution rates as nucleotides, and nearly identical levels of homoplasy. Therefore, repeat indels were excluded from analysis, while binary (0/1) coding was used for non-repetitive indels and nucleotide substitutions. Indels were coded as a single binary locus, regardless of length (Ingvärsson et al. 2003). When more than one length variant was nested completely within another, this was conservatively regarded as a single locus, with the full-length gap coded as either present or absent.

Analytical Methods

Haplotype networks were constructed using 95% statistical parsimony (Templeton et al. 1992) implemented in the software TCS 1.21 (Clement et al. 2000). Phylogeographic structure in the native range was investigated by dividing Europe into three geographical regions (Eastern, Western, and Southern; Fig. 1). The boundaries for these hypothesized genetic subdivisions were defined based on previous studies of *S. latifolia* in Europe, and comparative phylogeographic patterns in Europe among other species of plants and animals. The division between Eastern and Western Europe was based on evidence from *S. latifolia* that points to a genetic break beginning north of the Austrian Alps and extending northward through Germany to the North Sea (Mastenbroek et al.1983; Vellekoop et al. 1996; Ironside and Filatov 2005). This area also corresponds to a phylogeographic suture zone observed in studies of the post-glacial expansion of other European species (reviewed in Taberlet et al. 1998; Hewitt 2000). The boundary for

Southern Europe was defined by physiographic features that mark putative glacial refugia along the Mediterranean Sea (Iberia south of the Pyrenees, Italy south of the Alps, and the Balkan Peninsula). Phylogeographic structure was tested using analysis of molecular variance (AMOVA) with the software ARLEQUIN 2.001 (Schneider et al. 2000).

Several haplotypes in each species were recovered only in North America. Given the age of the invasion (ca. 200 year old), these haplotypes almost certainly originated in Europe and thus should be included when estimating the evolutionary history of each species in its native range. However, analyses based on coalescent theory make use of both the relatedness of haplotypes as well as their frequencies in the sample. Non-neutral processes such as rapid demographic growth or selection cause deviations from the neutral model by inflating the frequency of some haplotypes over others. In the current context, if we included all North American samples (with frequencies influenced by the demographics of the recent invasion), our analyses of long-term evolutionary history would likely be biased by the impact of invasion on the frequency distribution of mutations among lineages. To resolve this problem, we included a single sample of each unique North American haplotype in all coalescent simulations. We regard this approach as conservative in the current analyses, as the effect will be to include ancestral diversity known to be present in the genealogy while assigning a low frequency to haplotypes that are not regionally abundant in Europe (as inferred by our sampling).

We tested for the signature of historical demographic expansion in the shapes of the genealogies by calculating Tajima's D and Fu's F_S statistics (Tajima 1989; Fu 1997). Negative values of these statistics indicate an excess of young or rare alleles in the

genealogy – evidence of either population expansion or genetic hitchhiking. Calculations were performed in ARLEQUIN and used 1000 simulations to evaluate significance.

To explore the genealogical history of *Silene* in Europe, we performed coalescent simulations in GENETREE 9.0 (Griffiths and Tavare 1994; http://www.stats.ox.ac.uk/~griff/software.html). For each species, we used the other species as an outgroup to manually reconstruct the ancestral state at each site. Since GENETREE assumes the data fit the infinite sites model of evolution (each site may mutate only once), sites that violated this assumption were removed prior to analysis.

We used GENETREE to estimate the scaled mutation parameter, θ (= $2N_e\mu$ for haploids, where N_e is the effective size of the cpDNA population and μ is the mutation rate per sequence per generation). We searched the likelihood surface using 100,000 simulations for each of many values of θ and took the maximum likelihood value as our estimate of θ. To estimate μ, we obtained the silent substitution rate for chloroplasts ($1.1\text{-}2.9 \times 10^{-9}$ nucleotide substitutions per site per year; Wolfe et al. 1987). However, this mutation rate does not include indel evolution. We formulated a regression model based on our cpDNA dataset to estimate the rate of indel evolution from the nucleotide evolution observed. To accomplish this, statistical parsimony networks were built for each species using only the nucleotide polymorphisms. We then mapped the indels onto each network and obtained the number of nucleotide and indel changes for each connection between extant haplotypes. Linear regression was used to predict the number of indel changes, given the number of nucleotide changes, and yielded the equation: # indels = 0.711 + 0.748 * (# nucleotides). The 95% confidence interval for the slope (0.410-1.09) was significantly

greater than zero but not different than one, confirming previous observations that the two types of mutations occur at comparable rates (Ingvärsson et al. 2003). We applied this equation to the range of values for the cpDNA silent substitution rate (K_S) reported by Wolfe et al. (1987) to obtain the combined substitution rate for indels and nucleotides. From the resulting set of corrected estimates, we took the median value of $\mu = 2.8 \times 10^{-9}$ substitutions (nucleotides and indels) per site per year (range = 2.0-5.1). While uncertainty in μ exists due to a limited fossil record in plants (Wolfe et al. 1987), conclusions from our analyses were not qualitatively different when using the upper or lower range of estimates (data not shown). We adjusted μ to a per-sequence per-generation value by multiplying by the number of bases in the concatenated dataset and by the generation time (taken to be ~2 years, based on life history data from field experiments). Using the relationship $\theta = 2N_e\mu$, we then solved for species-wide chloroplast N_e.

Population growth can have a pronounced effect on θ; therefore, we used GENETREE to fit a population growth parameter (β) to the data. We explored the joint likelihood surface of θ and β by running 100,000 replicate simulations for each of many values of θ and β and taking the joint maximum likelihood values as our estimators (e.g., Thompson et al. 2000). We tested if population growth significantly improved the model using a likelihood ratio test (Neter et al. 1996). This test competes two nested models using the statistic $\Delta = -2\log_e(L_1/L_2)$, where L_1 is the likelihood of the model with β and L_2 is the likelihood of the model without β. Δ is approximately χ^2 distributed with 1 degree of freedom.

We then simulated the coalescent for 1,000,000 replicates using the ML estimates of θ and β. From these we obtained the time to the most recent common ancestor (T_{MRCA}) and the ages of individual mutations (in coalescent units, T). For haploids, time is measured in N_eTg years, where g is the generation time.

We explored the distribution of relatedness among haplotypes in North America and Europe by generating the mismatch distribution (the frequency of pairwise differences among individuals). Mismatch distributions are sensitive to demographic events in the evolutionary history of the sample, and thus are useful indicators of past population expansion (Rogers and Harpending 1992; Rogers 1995; Schneider and Excoffier 1999). Specifically, rapid demographic growth leads to a smooth, unimodal mismatch distribution, whose moments can be used to estimate the timing of the expansion, $\tau = 1/2\mu$ generations (Rogers 1995). We generated the mismatch distribution in *Silene* among European haplotypes and fit it to a model of sudden demographic expansion using ARLEQUIN. We also included unique North American haplotypes, as above. We converted our estimate of τ to years using our estimate of μ (obtained above) and multiplying by 2 (number of years per generation).

Mismatch distributions can also reveal recent admixture among divergent lineages. If admixture is occurring, we expect bimodal or multimodal frequency distributions of pairwise differences, attributable to few differences among haplotypes within clades, and greater differences among haplotypes between clades. We tested for admixture in the *Silene* invasions by calculating the mismatch distribution among all North American haplotypes.

To investigate the cpDNA diversity sampled during invasion, we calculated the total number of haplotypes, private haplotypes, segregating sites, average number of pairwise differences, and gene diversity for North America, Europe, and both continents pooled using ARLEQUIN. We determined whether the diversity sampled during invasion was a representative sample from the genealogy of each species by measuring phylogenetic diversity (PD), equivalent to the sum of the branch lengths represented in invasive haplotypes (Crozier 1997; Purvis et al. 2000). We compared the observed PD to a null expectation by constructing a neighbor-joining tree in PAUP* (Swofford 1998) and calculating the observed PD of invasive haplotypes using the software MESA (http://www.agapow.net/software/mesa/releases/1.9.22). Keeping that tree topology, we then randomly sampled individuals with replacement from the tree (keeping the proportion of "invasive" haplotypes constant) and recalculated PD for each pseudosample. This was repeated 1000 times to generate a null distribution against which the observed value of PD among invasive haplotypes was compared.

Results

Haplotype Networks and Population Structure

From the 62 *Silene latifolia* and the 50 *S. vulgaris* individuals sampled, we identified 38 and 33 cpDNA haplotypes, respectively. Each species had 54 segregating sites that defined the genealogical relationships among the haplotypes (Fig. 2). The networks of both species contained a large number of unique haplotypes, with a few common haplotypes, and possessed high species-wide gene diversity (Table 1).

The spatial distribution of *S. latifolia* haplotypes showed clear phylogeographic structure in Europe (Fig. 2). Haplotypes from Eastern and Western Europe formed divergent groups descended from haplotypes currently distributed in Southern Europe. AMOVA showed this structuring among European regions is highly significant (Φ_{ST} = 0.270, P<0.0001; Table 2). Invasive haplotypes originated from both Eastern and Western Europe, suggesting that North America experienced multiple introductions from each of these two phylogeographic regions.

By contrast, *S. vulgaris* showed no clustering of haplotypes with region (Φ_{ST} = 0.016, P = 0.27; Table 2). Haplotypes introduced to North America were also widely dispersed in the genealogy, suggesting either multiple introductions or a single introduction from a diverse source (Fig. 2).

Evolutionary History and Post-Glacial Expansion in Europe

Despite the clear difference in the degree of phylogeographic structure, *S. latifolia* and *S. vulgaris* had remarkably similar demographic histories in Europe. Coalescent simulations (assuming demographic equilibrium) returned values of the scaled mutation parameter, θ, in close agreement for the two species (Table 3). Incorporating population growth (β) into the simulations significantly improved the model fit for both *S. latifolia* (Δ = 12.70, P = 0.0007) and *S. vulgaris* (Δ = 13.64, P = 0.0004). From these models, the effective size of the cpDNA population for both species was estimated to be ~2.5 million individuals (Table 3). The maximum likelihood estimate of T_{MRCA} (mean ± SD) was T= 0.165 ± 0.013 for *S. latifolia* and T=0.171 ± 0.019 for *S. vulgaris*, placing coalescence to

the common ancestor of extant haplotypes over 800 kya (thousand years ago) for both species (Table 3).

Both species showed an excess of mutations that were either rare or recently derived, indicated by significantly negative values of Tajima's *D* and Fu's F_S, suggesting either a population expansion or a selective sweep (Table 4). Pairwise differences among European haplotypes showed a unimodal mismatch distribution that provided a good fit to a model of sudden demographic growth (*S. latifolia*: $P = 0.41$; *S. vulgaris*: $P = 0.86$), further supporting a history of population expansion (Fig. 3). The dates of onset for these demographic increases were similar for each species, estimated to be 521 kya for *S. latifolia* and 604 kya for *S. vulgaris* (Table 4).

Several mutations in the *S. latifolia* genealogy defined the phylogeographic structure observed with AMOVA, and hence likely reflect important historical migration events (Fig. 4). The divergence of one group from an ancestral Southern European gene pool occurred 725 ± 127 kya. The haplotypes that colonized Western Europe represent a subset of this diversity. A second group in Southern Europe diverged from the ancestral pool 483 ± 61 kya. This group then gave rise to a subset of haplotypes that originated 330 ± 64 kya that went on to colonize Eastern Europe. Thus the origin of the phylogeographic subdivision between Eastern and Western Europe likely occurred >400,000 years ago.

Genetics of Invasion into North America

Silene latifolia in North America has undergone a substantial bottleneck of diversity, with fewer segregating sites, lower pairwise differences, and lower gene diversity compared to Europe (Table 1). Observed PD of North American haplotypes was only

51% of the diversity expected if the invasion were drawn randomly from the genealogy (P<0.001; Fig. 5). This bottleneck may result from either a founder effect during the colonization of North America by a set of genetically similar individuals, or from demographic events that occurred post-colonization that reduced the diversity initially introduced. Although an overall bottleneck in diversity is evident, there has also been an admixture caused by multiple introductions from Eastern and Western Europe. The signature of admixture is evident in the bimodal mismatch distribution for North America, which shows invasion by haplotypes originating from two highly differentiated sources (Fig. 3).

The situation for *S. vulgaris* was quite different. *Silene vulgaris* in North America had slightly more segregating sites and higher pairwise differences, and equivalent gene diversity compared to Europe (Table 1). North America also contained an amount of PD comparable to random sampling from the genealogy (P = 0.51; Fig. 5). The mismatch distribution of North American *S. vulgaris* was unimodal, although the overall shape of the distribution was more ragged than for Europe (Fig. 3). This is consistent with *S. vulgaris* being sampled representatively from an unstructured native range.

Both species had fewer private haplotypes in North America compared to Europe (Table 1). Consistent with other indices of diversity, *S. latifolia* showed a greater proportional reduction in private haplotypes (63%) compared to *S. vulgaris* (25%), although the magnitude of reduction was not statistically significant (Fisher's exact test: P = 0.28). Thus, some of the difference between species in PD bottlenecked during invasion may be attributable to the sampling of rare haplotypes, but this alone cannot explain the observed differences in diversity. Rather, the haplotypes of *S. latifolia*

introduced to North America originated more recently (on average) than haplotypes that were not introduced (300 kya vs. 589 kya; t = -3.53, df = 29, P = 0.001), because invaders mostly come from the younger Eastern European clade (Fig. 4). In contrast, the mean age of *S. vulgaris* haplotypes invading North America was more similar to haplotypes that contributed no colonists (403 kya vs. 524 kya; t = -1.35, df = 20, P = 0.19). This difference in haplotype age contributed to the bottlenecking of evolutionary history during invasion by *S. latifolia*, but not *S. vulgaris*.

Discussion

The results of this study provide clear insights into how the genetics of invasion by two closely related plant species, *Silene latifolia* and *S. vulgaris*, have been influenced by their different histories of post-glacial expansion in the native range. Although the two species have comparable genealogical histories, they differ in the spatial distribution of lineages in Europe. This has resulted in dramatic differences in how phylogenetic diversity was sampled during their invasion of North America.

Below, we discuss the phylogenetic history of the two species in their native range. We then interpret how this history interacted with the invasion process to determine the diversity present in the introduced range.

Post-Glacial Colonization of Europe

The genealogical results suggest that *Silene latifolia* and *S. vulgaris* have similar pre-glacial histories and post-glacial expansions into Europe, but rather different histories of

dispersal since that time. For both species, the T_{MRCA} is similar (~800-900 kya) with similar demographic histories of population expansion (significantly negative *D* and F_S). The T_{MRCA} for both can be traced to the mid-Pleistocene, while the estimated dates of expansion (~500-600 kya) suggest a proliferation of each species shortly after the onset of extreme climate oscillations which characterized the glacial cycles of the last 700 kyr (Webb and Bartlein 1992; Hewitt 1996). Initial demographic expansions such as these are likely to overwhelm the signature of subsequent expansions (Rogers 1995). Thus, although *Silene* almost certainly expanded and retreated several times during the glacial cycles of the Pleistocene (Hewitt 1996), the initial expansions would have obscured these subsequent events.

The species, however, are significantly different in how cpDNA lineages are currently distributed in space. The pattern of post-glacial range expansion in *S. latifolia* is particularly clear. Chloroplast DNA lineages in Eastern and Western Europe are different subsets of the lineages found in Southern Europe. This general pattern of post-glacial expansion is similar to that seen in a variety of plant and animal taxa (Taberlet et al. 1998; Hewitt 1996, 2000). The age of mutations that define these lineages suggest that Europe was colonized from genetically distinct refugia, probably located in the Iberian and Balkan Peninsulas. Our results for the chloroplast genome are consistent with phylogeographic structure present in the *S. latifolia* Y-chromosome (Ironside and Filatov 2005), suggesting similar patterns of seed and pollen flow occurred at the regional scale.

Our data confirm the existence of phylogeographic structure between Eastern and Western Europe for *S. latifolia* (Mastenbroek et al. 1983; Vellekoop et al. 1996), but are ambiguous as to the migration events that distributed the haplotypes. It has been

hypothesized that with the advent of human agricultural practices, evolution of a weedy lifeform occurred in Southern Europe (Baker 1948; Mastenbroek et al. 1983). Weedy genotypes from divergent refugia are then thought to have migrated northward with the spread of agriculture, resulting in the different "races" currently distributed in Eastern and Western Europe. (Mastenbroek et al. 1983; Vellekoop et al. 1996). Our results clearly point to divergence between Eastern and Western European lineages as a result of population subdivision that originated >400 kya. We cannot rule out that *S. latifolia* remained in these refugia until humans inadvertently dispersed them with the spread of agriculture into Northern Europe. However, the fact that *S. latifolia* in Southern Europe more often occupies natural habitats such as open woodlands and montane limestone screes (Mastenbroek et al. 1983), suggests that the evolution of weediness may have occurred outside of refugia. These data, along with the evidence that *S. latifolia* experienced rapid demographic expansions well before the spread of agriculture (ca. 6-8 kya), suggests that *S. latifolia* initially colonized Northern Europe in a post-glacial expansion unassisted by humans (*sensu* Hewitt 2000).

For *S. vulgaris*, the spatial pattern of expansion is less clear, with lineages widely scattered across the continent. The *S. latifolia* data suggest Europe was colonized from several distinct refugia. If *S. vulgaris* colonized Europe from a single refugium that was a melting pot of lineages, then the current differences in genetic structure could be a consequence of differences in the ancestral structure of those refugia. Although we cannot reject this idea, the similar distributions of mutations on the genealogies suggest similar histories for the two species during the course of the Pleistocene.

An alternative interpretation of the data is that differences in phylogeographic structure reflect relatively recent differences in dispersal across Europe. While both species are short-lived perennials with similar dispersal ecologies, *Silene vulgaris* is a self-compatible hermaphrodite whereas *S. latifolia* is dioecious and thus incapable of selfing. This difference in reproductive assurance is thought to be an important determinant of colonizing ability (i.e. Baker's Law: Baker 1955; Taylor et al. 1999; Kolar and Lodge 2001). Thus, the current phylogeographic structure of *S. vulgaris* in Europe could reflect a more successful colonizing ability that has subsequently diluted the signature of historical range expansion. *Silene vulgaris* also has cytoplasmic male sterility (or CMS), and differences in population structure of cpDNA lineages could reflect a role of selection on the distribution of the cytoplasmic genomes (Ingvärsson and Taylor 2002; Olson and McCauley 2002; Tsitrone et al. 2003).

Contemporary Invasion of North America

The haplotypes sampled during the invasion of North America by Silene represent a subset of the total phylogenetic diversity (PD) present. Both species exhibit similar levels of species-wide polymorphism (Table 1) and experienced a sudden demographic expansion into Europe (Table 4); thus the amount of PD available for invasion was roughly similar. However, the PD actually sampled during invasion differs markedly between the species.

North American haplotypes of *S. latifolia* come from a few local sections of the entire genealogy that correspond to the phylogeographic regions of Eastern and Western Europe

(Fig. 2). The mismatch distribution of North American *S. latifolia* clearly shows the presence of divergent lineages among the introduced haplotypes (Fig. 3). The result is an admixture of anciently separated (>400 kya) lineages in the introduced range, an increasingly common feature of biological invasions (e.g., Kolbe et al. 2004). However, even though propagule pressure was high in the sense that sampling occurred across phylogeographic subdivisions, the sampling within each region was sufficiently restricted to result in a strong overall reduction in PD. Furthermore, most (but not all) North American haplotypes were descended from a recently derived Eastern Europe clade. As a result, the invasion captured relatively little evolutionary history (Fig. 4), and a clear bottleneck occurred on a continent-wide basis (Fig. 5). This may represent either a founder effect from colonization of North America by a set of genetically similar individuals, or demographic effects that occurred post-colonization that reduced the diversity initially introduced. In either case, a strong bottleneck of PD occurred despite multiple introductions that admixed anciently separated lineages.

It may seem counter-intuitive that the invasion process can simultaneously result in bottlenecking and admixture, but the reasons become clear when we consider how diversity is structured in the native range along with what the data reveal about the amount of propagule pressure. The abundance of low pairwise differences in the North American mismatch distribution (the first mode in Fig. 3) demonstrates that each invasion episode sampled a specific subset of the available PD. In other words, a large proportion of the haplotypes introduced differed by few or no mutations, having come from the same phylogeographic region in Europe. The abundance of high pairwise differences (the second mode in Fig. 3) marks haplotypes separated by many more

mutations. Because phylogenetically distant haplotypes of *S. latifolia* are spatially separated in Europe, this points to the occurrence of a minimum of two invasion episodes. However, the amount of propagule pressure during invasion was not so great as to make the mismatch distribution continuous, as it is in Europe. Thus multiple introductions, even ones that each involved a genetic bottleneck, produced an admixture of divergent East and West European lineages within North America. The evolutionary consequences of admixture between historically separated, recently bottlenecked populations may be very different compared to admixture among recently separated, non-bottlenecked populations. For example, the degree of heterosis following hybridization between lineages will depend on how inbred the lineages are initially, an outcome that may bear directly on the evolution of invasiveness (Ellstrand and Schierenbeck 2000). This suggests an additional layer of complexity may exist for species experiencing admixture during invasion.

In *Silene vulgaris*, the lack of phylogeographic structure in the native range enhanced the probability that a genetically diverse inoculum was obtained, even if the introduced lineages originated from a single geographic region. This lack of structure precluded any novel admixture of divergent native range populations and avoided a genetic bottleneck in North America. These results are consistent with previous findings of high cpDNA diversity in both the native and introduced ranges of *S. vulgaris* (McCauley et al. 2003; Storchova and Olson 2004). Interestingly, North American haplotypes do show a slight overabundance of low pairwise differences in the mismatch distribution, suggesting that some local sampling may have occurred during invasion (Fig. 3). However, since cpDNA haplotypes in *S. vulgaris* are scattered randomly across Europe, there is little more we can

conclude definitively about the invasion process. The distribution of pairwise differences observed in North America is consistent with a large number of introductions, a single introduction from a diverse inoculum, and everything in between.

Taken together, our results suggest that the evolutionary history of these species had a profound influence on the phylogenetic diversity captured during their recent invasion. Two species that are otherwise similar in their ecology and genealogical history differ in the current spatial distribution of cpDNA lineages across their native ranges. This has resulted in markedly different population genetic patterns during invasion of North America. These data have implications for the evolution of species invading new geographic ranges if the cpDNA lineages are representative of the nuclear genomes, and hence the traits affecting invasiveness. If this is the case, then for species like *S. vulgaris* with an unstructured native range, invasion may result in less severe bottlenecks of species-wide diversity, fewer opportunities for admixture of previously isolated lineages, and less potential for the redistribution of genetic variance in the introduced range. Conversely, species with a structured native range, such as *S. latifolia*, are more likely to experience severe bottlenecks, have greater opportunities for admixture, and may therefore experience greater population genetic change during invasion. Because bottlenecks and admixture can have important phenotypic consequences (e.g., inbreeding depression, heterosis, outbreeding depression), these two scenarios present very different implications for the evolution of a species during invasion of a new geographic range.

Acknowledgements

We thank P. Ingvärsson, S. Ribstein, and W. Farnum for sequencing, and C. Barr, S. Freedberg, M. Neiman, D. Sloan, D. Sowell, D. Funk, and one anonymous reviewer for helpful comments on the manuscript. We are also grateful to the many colleagues who have donated to our seed collection. This research was supported by grants DEB-444176 from the National Science Foundation and Award 2002-35320-12174 from the United States Department of Agriculture (to D.R.T.), and a University of Virginia Presidential Fellowship (to S.R.K.).

References

Alonso, S. and J. A. L. Armour. 2001. A highly variable segment of human subterminal 16p reveals a history of population growth for modern humans outside of Africa. Proc. Natl. Acad. Sci. USA 98:864-869.

Amsellem, L., J. L. Noyer, T. Le Bourgeois, and M. Hossaert-McKey. 2000. Comparison of genetic diversity of the invasive weed *Rubus alceifolius* Poir. (Rosaceae) in its native range and in areas of introduction, using amplified fragment length polymorphism (AFLP) markers. Mol. Ecol. 9: 443-455.

Antonovics, J., M. E. Hood, P. H. Thrall, J. Y. Abrams, and G. M. Duthie. 2003. Herbarium studies on the distribution of anther-smut fungus (*Microbotryum violaceum*) and *Silene* species (Caryophyllaceae) in the eastern United States. Am. J. Bot. 90:1522-1531.

Baker, H. G. 1948. Stages in invasion and replacement demonstrated by species of *Melandrium*. J. Ecol. 36:96-119.

---. 1955. Self-compatibility and establishment after "long-distance" dispersal. Evolution 9:347-349.

Cann, R. L., M. Stoneking, and A. C. Wilson. 1987. Mitochondrial DNA and human evolution. Nature 325:31-36.

Clement, M., D. Posada, and K. A. Crandall. 2000. TCS: a computer program to estimate gene genealogies. Mol. Ecol. 9:1657-1659.

Crozier, R. H. 1997. Preserving the information content of species: genetic diversity, phylogeny, and conservation worth. Ann. Rev. Ecol. Syst. 28:243-268.

Cutler, M. 1785. An account of some of the vegetable productions, naturally growing in this part of America, botanically arranged. Memoirs of the American Academy of Sciences of the USA 1:396-493

Davies, N., F. X. Villablanca, and G. K. Roderick. 2000. Determining the source of individuals: multilocus genotyping in nonequilibrium population genetics. Trends Ecol. Evol. 14:17-21.

de la Vega, M. P., P. Garcia, and R. W Allard. 1991. Multilocus genetic structure of ancestral Spanish and colonial Californian populations of *Avena barbata*. Proc. Natl. Acad. Sci. USA 88:1202-1206.

Ellstrand, N. C., and K. A. Schierenbeck 2000. Hybridization as a stimulus for the evolution of invasiveness in plants? Proc. Natl. Acad. Sci. USA 97:7043-7050.

Gaskin, J. F., and B. A. Schaal. 2002. Hybrid *Tamarix* widespread in U.S. invasion and undetected in native Asian range. *Proc. Natl. Acad. Sci. USA* 99:11256-11259.

Gaskin, J. F., D.–O. Zhang and M.-C Bon. 2005. Invasion of *Lepidium draba* (Brassicaceae) in the western United States: distributions and origins of chloroplast DNA haplotypes. Mol. Ecol. 14:2331-2341.

Gielly, L., and P Taberlet. 1994. The use of chloroplast DNA to resolve plant phylogenies: noncoding versus *rbc*L sequences. Mol. Biol. Evol. 11:769-777.

Griffiths, R. C., and S. Tavaré. 1994. Simulating probability distributions in the coalescent. Theor. Pop. Biol. 46:131-159.

Guinand, B., and S. Easteal. 1996. Multivariate patterns of genetic differentiation support complex colonization schemes in *Bufo marinus* populations. Evolution 50:944-951.

Hamilton, M. B., J. M. Braverman, and D. F. Soria-Hernanz. 2003. Patterns and relative rates of nucleotide and insertion/deletion evolution at six chloroplast intergenic regions in New World species of the Lecythidaceae. Mol. Biol. Evol. 20:1710-1721.

Hewitt, G. M. 1996. Some genetic consequences of ice ages, and their role in divergence and speciation. Biol. J. Linn. Soc. 58:247-276.

---. 2000. The genetic legacy of the Quaternary ice ages. Nature 405:907-913.

Hudson, R. R. 1990. Gene genealogies and the coalescent process. Oxford Surv. Evol. Biol. 7:1-44.

Ingvärsson, P. K., and D. R. Taylor. 2002. Genealogical evidence for epidemics of selfish genes. Proc. Natl. Acad. Sci. USA 99:11265-11269.

Ingvärsson, P. K., S. Ribstein, and D. R. Taylor. 2003. Molecular evolution of insertions and deletion in the chloroplast genome of *Silene*. Mol. Biol. Evol. 20:1737-1740.

Ironside, J. E., and D. A. Filatov. 2005. Extreme population structure and high interspecific divergence of the *Silene* Y chromosome. Genetics 171:705-713.

Jalas, J., and J. Suominen. 1986. Atlas Florae Europaeae vol. 3 Cambridge Univ. Press, Cambridge, UK.

Kolar, C. S. and D. M. Lodge. 2001. Progress in invasion biology: predicting invaders. Trends Ecol. Evol. 16:199-204.

Kolbe, J. L., R. E. Glor, L. R. Schettino, A. C. Lara, A. Larson, and J. B. Losos. 2004. Genetic variation increases during biological invasion by a Cuban lizard. Nature 431:177-181.

Kumar, S., K. Tamura, and M. Nei. 2004. MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. Briefings in Bioinf. 5:150-163.

Lee, C. E. 2002. Evolutionary genetics of invasive species. Trends Ecol. Evol. 17:386-391.

Marsden-Jones, E. M., and W. B. Turrill. 1957. The Bladder Campions. Bartholomew Press, Dorking.

Martindale, I. C. 1876. The introduction of foreign plants. Bot. Gazette 2:55-58.

Mastenbroek, O., H. C. Prentice, R. Kamps-Heinsbroek, J. van Brederode, G. J. Niemann, and G. van Nigtevecht. 1983. Geographic trends in flavone-glycosylation genes and seed morphology in European *Silene pratensis* (Caryophyllaceae). Plant Syst. Evol. 141:257-271.

McCauley, D. E. 1994. Contrasting the distribution of chloroplast DNA and allozyme polymorphism among local populations of *Silene alba*: implications for studies of gene flow in plants. Proc. Natl. Acad. Sci. USA 91:8127-8131.

McCauley, D. E., R. A. Smith, J. D. Lisenby, and C. Hsieh. 2003. The hierarchical spatial distribution of chloroplast DNA polymorphism across the introduced range of *Silene vulgaris.* Mol. Ecol. 12:3227-3235.

McNeill, J. 1977. The biology of Canadian weeds. 25. *Silene alba* (Miller) E.H.L. Krause. Can. J. Plant Sci. 57:1103-1114.

Neter, J., M. H. Kutner, C. J. Nachtsheim, and W. Wasserman. 1996. Applied Linear Regression Models 4th ed. Irwin, Chicago.

Novak, S. J., and R. N. Mack. 2001. Tracing plant introduction and spread: genetic evidence from *Bromus tectorum* (Cheatgrass). Bioscience 51:114-122.

Olson, M. S., and D. E. McCauley. 2002. Mitochondrial DNA diversity, population structure, and gender association in the gynodioecious plant *Silene vulgaris*. Evolution 56:253-262.

Petit, R. J., I. Aguinagalde, J. –L. de Beaulieu, C. Bittkau, S. Brewer, R. Cheddadi, R. Ennos, S. Fineschi, D. Grivet, M. Lascoux, A. Mohanty, G. Muller-Starck, B. Demesure-Musch, A. Palme, J. P. Martin, S. Rendell, and G. G. Vendramin. 2003. Glacial refugia: hotspots but not melting pots of genetic diversity. Science 300:1563-1565.

Pursh, F. 1814. A systematic arrangement and description of the plants of North America. White, London.

Purvis, A., P.-M. Agapow, J. L. Gittleman, and G. M. Mace. 2000. Nonrandom extinction and the loss of evolutionary history. Science 288:328-330.

Ray, N., M. Currat, and L. Excoffier. 2003. Intra-deme molecular diversity in spatially expanding populations. Mol. Biol. Evol. 20:76-86.

Rogers, A. R. 1995. Genetic evidence for a Pleistocene population explosion. Evolution 49:608-615.

Rogers, A. R., and H. Harpending. 1992. Population growth makes waves in the distribution of pairwise genetic differences. Mol. Biol. Evol. 9:552-569.

Runyeon, H., and H. Prentice. 1997. Genetic differentiation in the Bladder campions, *Silene vulgaris* and *S. uniflora* (Caryophyllaceae), in Sweden. Biol. J. Linn. Soc. 61:559-584.

Sakai, A. K., F. W. Allendorf, J. S. Holt, D. M. Lodge, J. Molofsky, K. A. With, S. Baughman, R. J. Cabin, J. E. Cohen, N. C. Ellstrand, D. E. McCauley, P. O'Neil, I. M. Parker, J. N. Thompson, S. G. Weller. 2001. The population biology of invasive species. Ann. Rev. Ecol. Syst. 32:305-332.

Saltonstall, K. 2002. Cryptic invasion by a non-native genotype of the common reed, *Phragmites australis*, into North America. Proc. Natl. Acad. Sci. USA 99:2445-2449.

Schaal, B. A., and K. M. Olsen. 2000. Gene genealogies and population variation in plants. Proc. Natl. Acad. Sci. USA 97:7024-7029.

Schaal, B. A., J. F. Gaskin, and A. L. Caicedo. 2003. Phylogeography, haplotypes trees, and invasive plant species. J. Heredity 94:197-204.

Schneider, S., and L. Excoffier. 1999. Estimation of past demographic parameters from the distribution of pairwise differences when the mutation rates vary among sites: application to human mitochondrial DNA. Genetics 152:1079-1089.

Schneider, S., D. Roessli, and L. Excoffier. 2000. Arlequin: A software for population genetics data analysis. Genetics and Biometry Lab, Dept. of Anthropology, University of Geneva.

Storchova, H., and M. S. Olson. 2004. Comparison between mitochondrial and chloroplast DNA variation in the native range of *Silene vulgaris*. Mol. Ecol. 13:2909-2919.

Swofford, D. 1998. Phylogenetic analysis using parsimony (and other methods) PAUP 4.04 beta version. Sinauer Associates, Sunderland, MA.

Taberlet, P., L. Fumagalli, A.-G. Wust-Saucy, and J.-F. Cosson. Comparative phylogeography and postglacial colonization routes in Europe. Mol. Ecol. 7:453-464.

Tajima, F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123:585-595.

Taylor, D. R., S. Trimble, and D. E. McCauley. 1999. Ecological genetics of gynodioecy in *Silene vulgaris*: relative fitness of females and hermaphrodites during the colonization process. Evolution 53:745-751.

Templeton, A. R., K. A. Crandall, and C. F. Sing. 1992. A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. Genetics 132:619-633.

Thompson, R., J. K. Pritchard, P. Shen, P. J. Oefner, and M. W. Feldman. 2000. Recent common ancestry of human Y chromosomes: Evidence from DNA sequence data. Proc. Natl. Acad. Sci. USA 97:7360-7365.

Travisano, M., J. A. Mogold, A. F. Bennett, and R. E. Lenski. 1995. Experimental tests of the roles of adaptation, chence, and history in evolution. Science 267:87-90.

Tsitrone, A., M. Kirkpatrick, and D. A. Levin. 2003. A model for chloroplast capture. Evolution 57:1776-1782.

Tsutsui, N. D., A. V. Suarez, D. A. Holway and T. J. Case. 2000. Reduced genetic variation and the success of an invasive species. Proc. Natl. Acad. Sci. U.S.A. 97:5948-5953.

Vellekoop, P., J. B. Buntjer, J. W. Maas, and J. van Brederode. 1996. Can the spread of agriculture in Europe be followed by tracing the spread of the weed *Silene latifolia*. A RAPD study. Theor. Appl. Genet. 92:1085-1090.

Watson, E., P. Forster, M. Richards, and H.- J. Bandelt. 1997. Mitochondrial footprints of human expansions in Africa. Am. J. Hum. Genet. 61:691-704.

Webb III, T., and P. J. Bartlein. 1992. Global changes during the last 3 million years: climatic controls and biotic responses. Ann. Rev. Ecol. Syst. 23:141-173.

Wolfe, K. H., W. –H. Li, and P. M. Sharp. 1987. Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. Proc. Natl. Acad. Sci. USA 84:9054-9058.

Figure Legends

Figure 1. Map of collection sites in North America and Europe. Dotted lines show the phylogeographic groupings tested in AMOVA.

Figure 2. Statistical parsimony haplotype networks. Circles are sampled haplotypes (size proportional to frequency) and small filled squares are inferred haplotypes not recovered in the sample. Each link between haplotypes, regardless of length, represents one mutational step.

Figure 3. Mismatch distributions among cpDNA haplotypes in Europe and North America. Filled bars are the observed distributions. Solid lines with diamonds are the fitted distributions under the sudden demographic expansion model of Rogers 1995.

Figure 4. Evolutionary history of Silene in Europe and invasion of North America. Gene trees show the distribution of mutations among haplotypes, with mutation age measured along the Y-axis in coalescent units (see *Methods*). Major groups defined by mutations are color coded and mapped.

Figure 5. Phylogenetic diversity (PD) sampled during invasion of North America. Filled bars show the null distribution of PD under random sampling. Arrows show the observed PD encompassed by invasive haplotypes.

Table 1. Estimates of cpDNA genetic diversity in Europe, in North America, and overall for *Silene latifolia* and *S. vulgaris*.

	S. latifolia			*S. vulgaris*		
	Europe	North America	Total	Europe	North America	Total
# samples	36	26	62	28	22	50
# haplotypes (# private)	29 (24)	14 (9)	38	21 (16)	17 (12)	33
Segregating sites	48	17	54	34	36	54
Mean pairwise differences (SD)	5.14 (2.55)	3.35 (1.77)	4.90 (2.42)	4.87 (2.45)	6.56 (3.22)	5.64 (2.75)
Gene diversity (SD)	0.990 (0.009)	0.835 (0.070)	0.956 (0.018)	0.963 (0.024)	0.965 (0.028)	0.965 (0.015)

Table 2. Analysis of molecular variance (AMOVA) among Eastern, Western, and Southern European cpDNA haplotypes for *Silene latifolia* and *S. vulgaris*. Geographic regions are defined in Figure 1.

Species	Source of variation	d.f.	Sums of Squares	Variance component	Percentage of variation
S. latifolia	Among regions	2	21.397	0.76650	27.03***
	Within regions	32	66.232	2.06975	72.97
	Total	34	87.629		
S. vulgaris	Among regions	2	5.500	0.03809	1.56 ns
	Within regions	25	60.214	2.40857	98.44
	Total	27	65.714	2.44666	

*** $P<0.0001$; ns not significant

Table 3. Maximum likelihood estimates of coalescent parameters for *Silene latifolia* and *S. vulgaris* in Europe.

Species	Model†	θ	β	N_e	T_{MRCA} ± SD (kya)
S. latifolia	β = 0	19	--	$0.90*10^6$	298 ± 24
	β > 0	57	20	$2.71*10^6$	894 ± 72
S. vulgaris	β = 0	18	--	$0.89*10^6$	304 ± 34
	β > 0	48	18	$2.38*10^6$	811 ± 90

† Model refers to maximum likelihood coalescence simulations which included no population growth (β = 0) or exponential population growth (β > 0). Likelihood ratio tests indicate inclusion of β > 0 significantly improves the fit of the model to the data (see Results).

Table 4. Evidence for the occurrence and timing of population expansions of *Silene latifolia* and *S. vulgaris* in Europe.

			Sudden demographic expansion model		
Species	*D*	F_S	Model fit[†]	τ (95% C.I.)	Onset of expansion (95% C.I.)
S. latifolia	-2.01[*]	-25.46[***]	*P* = 0.41	5.469 (3.101-7.017)	521 kya (295-668)
S. vulgaris	-1.89[*]	-25.21[***]	*P* = 0.86	6.084 (3.484-8.318)	604 kya (346-826)

[†] Refers to the goodness of fit of a model of sudden demographic expansion to the data. Non-significant *P* values indicate a good fit of the presumed model to the data.

[*] $P<0.01$; [***] $P<0.0001$

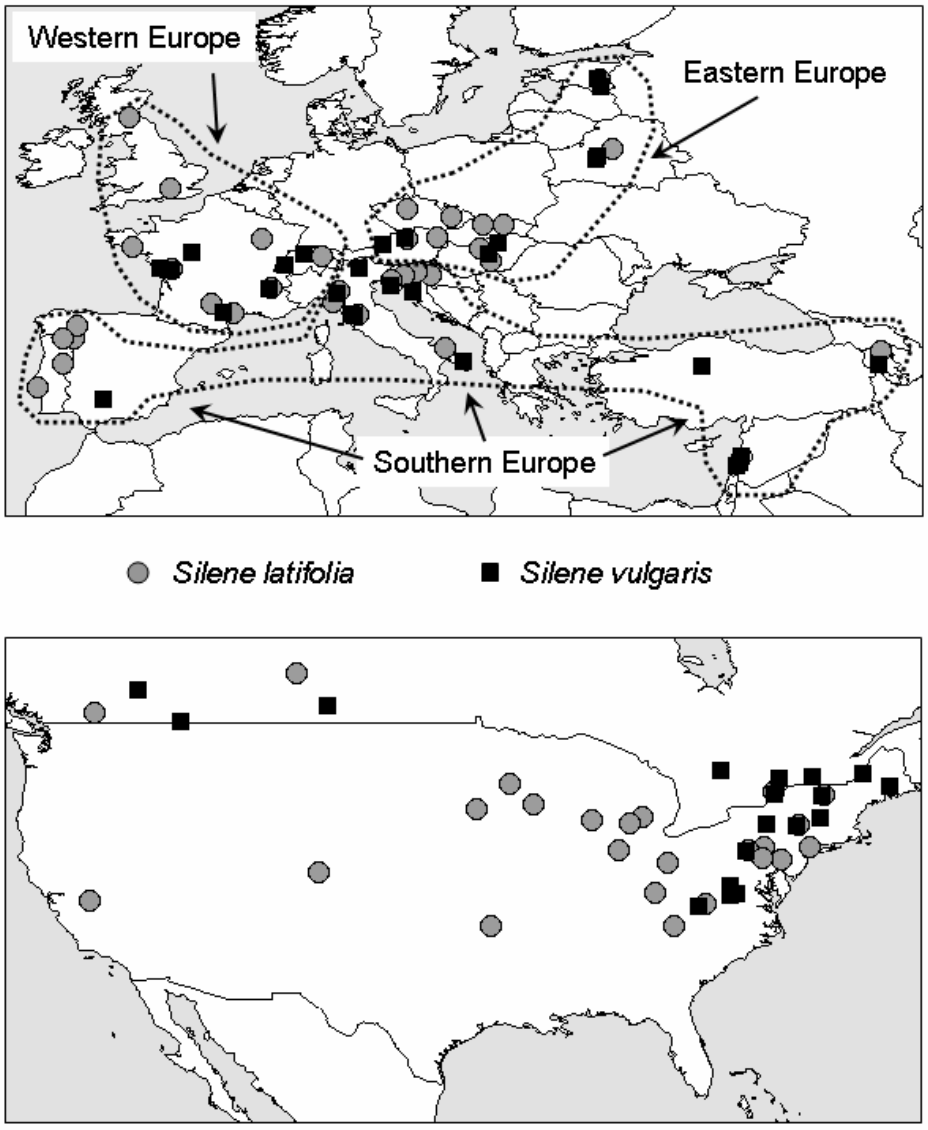

Western Europe
Eastern Europe
Southern Europe
Silene latifolia
Silene vulgaris

Figure 1.

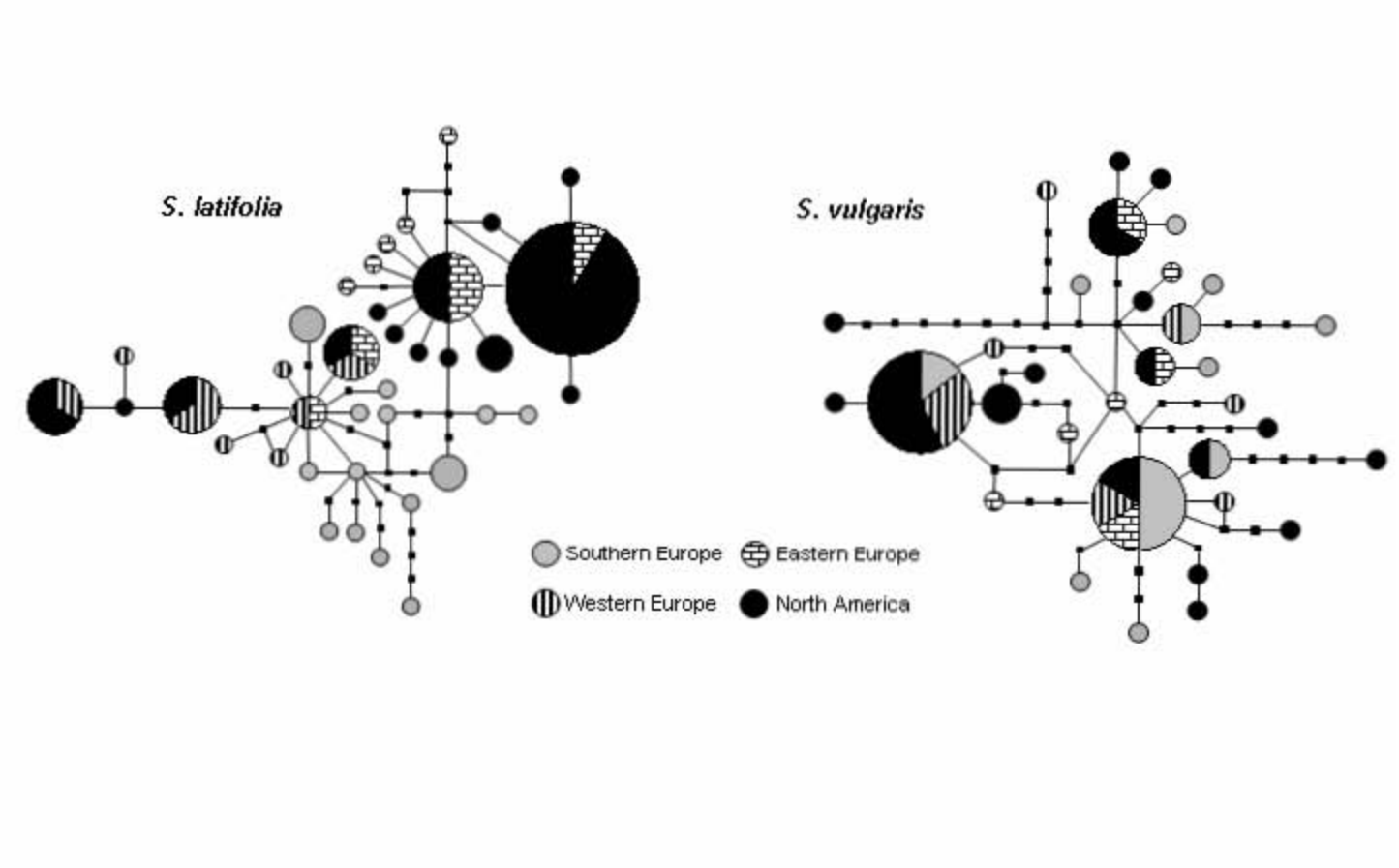

S. latifolia
S. vulgaris
Southern Europe
Eastern Europe
Western Europe
North America

Figure 2.

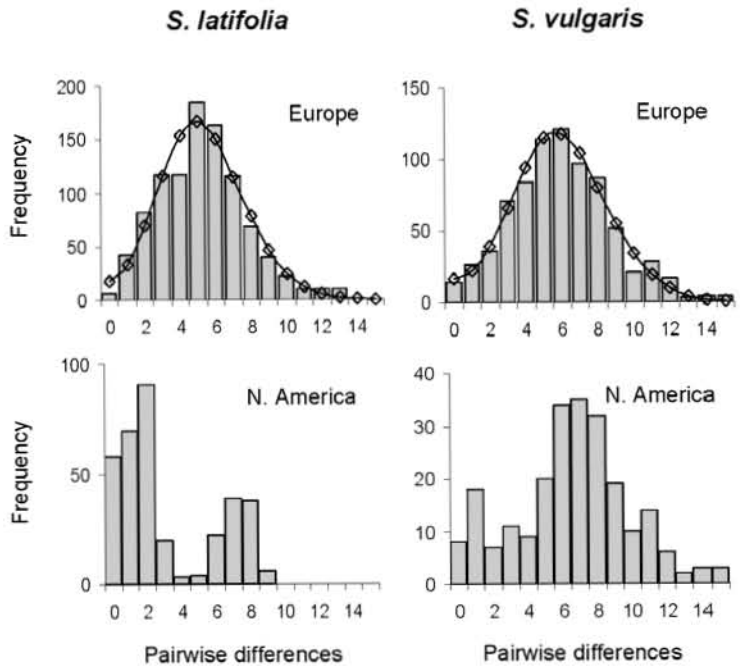

S. latifolia
S. vulgaris
Europe
Europe
N. America
N. America
Frequency
Frequency
Pairwise differences
Pairwise differences

Figure 3.

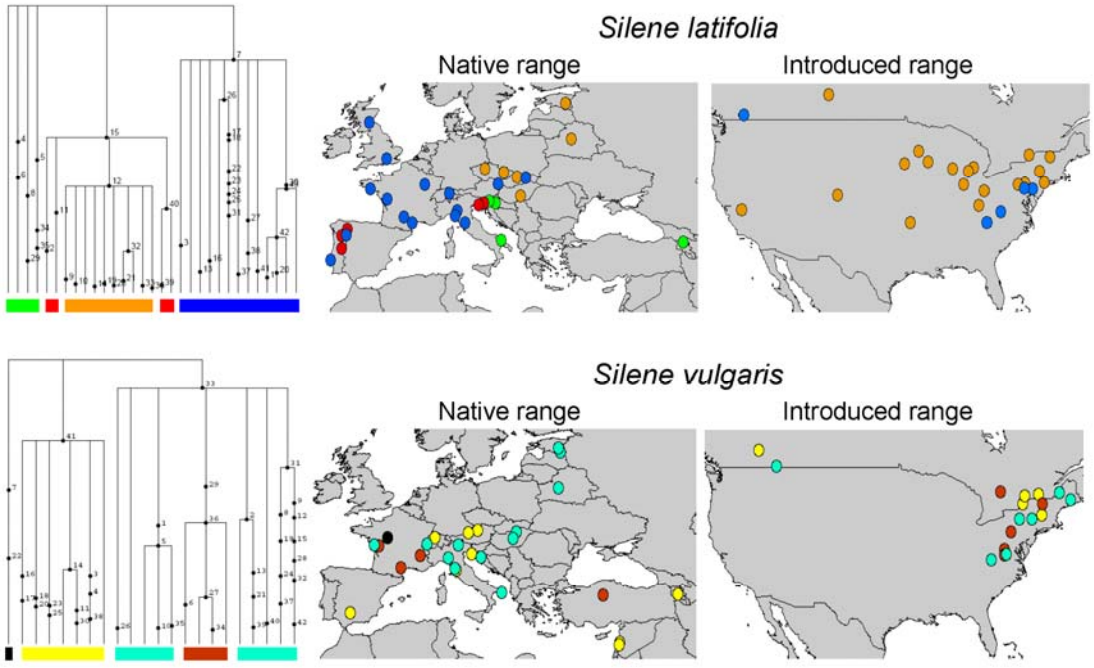

Silene latifolia
Native range
Introduced range
Silene vulgaris
Native range
Introduced range

Figure 4.

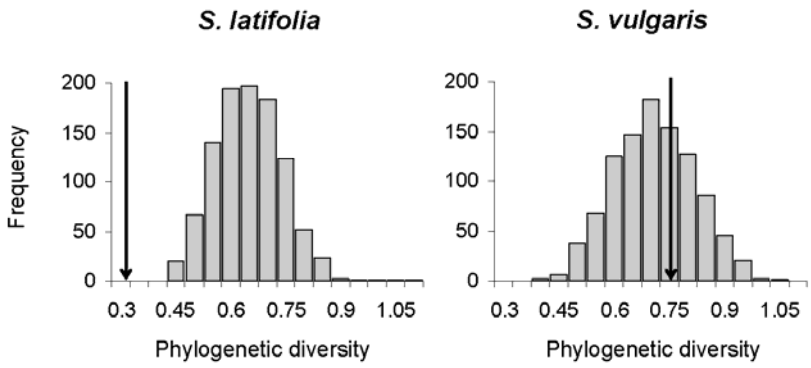

S. latifolia
S. vulgaris
Frequency
200
150
100
50
0
0.3
0.45
0.6
0.75
0.9
1.05
Phylogenetic diversity

Figure 5.

Appendix A. Accession numbers of samples archived on Genbank.

Genbank accession numbers for 4 cpDNA regions

Species	Continent	Sample ID	trnL-trnF	trnH-psbA	trnG-trnS	trnL intron
Silene latifolia	North America	LATNA1	EF088853	EF091495	EF088891	EF091533
Silene latifolia	North America	LATNA14	EF088863	EF091503	EF088893	EF091542
Silene latifolia	North America	LATNA15	EF088860	EF091504	EF088904	EF091543
Silene latifolia	North America	LATNA16	EF088857	EF091505	EF088902	EF091544
Silene latifolia	North America	LATNA19	EF088868	EF091506	EF088898	EF091545
Silene latifolia	North America	LATNA2	EF088872	EF091496	EF088896	EF091534
Silene latifolia	North America	LATNA20	EF088856	EF091507	EF088897	EF091546
Silene latifolia	North America	LATNA22	EF088865	EF091508	EF088911	EF091547
Silene latifolia	North America	LATNA27	EF088867	EF091509	EF088899	EF091548
Silene latifolia	North America	LATNA3	EF088862	EF091497	EF088894	EF091535
Silene latifolia	North America	LATNA31	EF088871	EF091510	EF088906	EF091549
Silene latifolia	North America	LATNA32	EF088869	EF091511	EF088905	EF091550
Silene latifolia	North America	LATNA37	EF088866	EF091512	EF088907	EF091551
Silene latifolia	North America	LATNA38	EF088861	EF091513	EF088903	EF091552
Silene latifolia	North America	LATNA41	EF088870	EF091514	EF088910	EF091553
Silene latifolia	North America	LATNA42	EF088878	EF091515	EF088909	EF091554
Silene latifolia	North America	LATNA5	EF088864	EF091498	EF088900	EF091536
Silene latifolia	North America	LATNA6	EF088854	EF091499	EF088895	EF091537
Silene latifolia	North America	LATNA62	EF088873	EF091517	EF088913	EF091555
Silene latifolia	North America	LATNA63	EF088875	EF091518	EF088914	EF091556
Silene latifolia	North America	LATNA64	EF088876	EF091519	EF088916	EF091557
Silene latifolia	North America	LATNA68	EF088877	EF091520	EF088915	EF091558
Silene latifolia	North America	LATNA7	EF088858	EF091500	EF088908	EF091538
Silene latifolia	North America	LATNA7A	EF088874	EF091516	EF088912	EF091539
Silene latifolia	North America	LATNA8	EF088859	EF091501	EF088892	EF091540
Silene latifolia	North America	LATNA9	EF088855	EF091502	EF088901	EF091541
Silene latifolia	Europe	LAT10	AF518880	AF518907	AF519048	AF519014
Silene latifolia	Europe	LAT10A	AF518879	AF518904	AF519047	AF519013
Silene latifolia	Europe	LAT11	AF518881	AF518908	AF519049	AF519015
Silene latifolia	Europe	LAT13	AF518882	AF518909	AF519050	AF519016
Silene latifolia	Europe	LAT15	AF518883	AF518910	AF519051	AF519017
Silene latifolia	Europe	LAT17	AF518884	AF518911	AF519052	AF519018
Silene latifolia	Europe	LAT18	AF518885	AF518912	AF519053	AF519019
Silene latifolia	Europe	LAT1A	AF518886	AF518905	AF519054	AF519020
Silene latifolia	Europe	LAT2	EF088884	EF091521	EF088923	EF091559
Silene latifolia	Europe	LAT21	AF518887	AF518913	AF519055	AF519021
Silene latifolia	Europe	LAT23	AF518888	AF518914	AF519056	AF519022
Silene latifolia	Europe	LAT28	AF518889	AF518915	AF519057	AF519023
Silene latifolia	Europe	LAT3	AF518894	AF518916	AF519062	AF519028
Silene latifolia	Europe	LAT30	AF518890	AF518917	AF519058	AF519024
Silene latifolia	Europe	LAT33	EF088882	EF091523	EF088917	EF091561
Silene latifolia	Europe	LAT35	AF518891	AF518918	AF519059	AF519025
Silene latifolia	Europe	LAT36	AF518892	AF518919	AF519060	AF519026
Silene latifolia	Europe	LAT37	EF088885	EF091524	EF088918	EF091562
Silene latifolia	Europe	LAT39	AF518893	AF518920	AF519061	AF519027
Silene latifolia	Europe	LAT3A	EF088887	EF091531	EF088919	EF091569
Silene latifolia	Europe	LAT40	AF518895	AF518921	AF519063	AF519029
Silene latifolia	Europe	LAT41	AF518896	AF518922	AF519064	AF519030
Silene latifolia	Europe	LAT42	AF518897	AF518923	AF519065	AF519031
Silene latifolia	Europe	LAT43	AF518898	AF518924	AF519066	AF519032
Silene latifolia	Europe	LAT46	EF088879	EF091525	EF088920	EF091563
Silene latifolia	Europe	LAT48	AF518900	AF518926	AF519068	AF519034
Silene latifolia	Europe	LAT49	AF518901	AF518927	AF519069	AF519035
Silene latifolia	Europe	LAT50	EF088881	EF091526	EF088922	EF091564
Silene latifolia	Europe	LAT51	EF088880	EF091527	EF088924	EF091565
Silene latifolia	Europe	LAT65	EF088888	EF091528	EF088928	EF091566
Silene latifolia	Europe	LAT66	EF088889	EF091529	EF088926	EF091567

Silene latifolia	Europe	LAT69	EF088890	EF091530	EF088927	EF091568
Silene latifolia	Europe	LAT8	AF518903	AF518928	AF519071	AF519037
Silene latifolia	Europe	LAT8A	AF518902	AF518906	AF519070	AF519036
Silene latifolia	Europe	LAT9	EF088883	EF091522	EF088921	EF091560
Silene latifolia	Europe	LAT9A	EF088886	EF091532	EF088925	EF091570
Silene vulgaris	North America	VLG_ITC-3	EF094028	EF094072	EF094050	EF094094
Silene vulgaris	North America	VLG_ITD-11	EF094029	EF094073	EF094051	EF094095
Silene vulgaris	North America	VLG_ITE-l	EF094030	EF094074	EF094052	EF094096
Silene vulgaris	North America	VLG_MEBDR	EF094031	EF094075	EF094053	EF094097
Silene vulgaris	North America	VLG_MLST	EF094032	EF094076	EF094054	EF094098
Silene vulgaris	North America	VLG_MONT	EF094033	EF094077	EF094055	EF094099
Silene vulgaris	North America	VLG_ORNO1	EF094034	EF094078	EF094056	EF094100
Silene vulgaris	North America	VLG_OTTA1	EF094035	EF094079	EF094057	EF094101
Silene vulgaris	North America	VLG_RAHS4	EF094038	EF094080	EF094058	EF094102
Silene vulgaris	North America	VLG_STFD2-5	EF094036	EF094081	EF094059	EF094103
Silene vulgaris	North America	VLG_STFD4-5	EF094037	EF094082	EF094060	EF094104
Silene vulgaris	North America	VLG_WGPP	EF094039	EF094092	EF094070	EF094114
Silene vulgaris	North America	VLG_WTGN2	EF094040	EF094093	EF094071	EF094115
Silene vulgaris	North America	VLGNA01	EF094049	EF094083	EF094061	EF094105
Silene vulgaris	North America	VLGNA02	EF094048	EF094084	EF094062	EF094106
Silene vulgaris	North America	VLGNA03	EF094047	EF094085	EF094063	EF094107
Silene vulgaris	North America	VLGNA04	EF094046	EF094086	EF094064	EF094108
Silene vulgaris	North America	VLGNA05	EF094045	EF094087	EF094065	EF094109
Silene vulgaris	North America	VLGNA06	EF094044	EF094088	EF094066	EF094110
Silene vulgaris	North America	VLGNA07	EF094043	EF094089	EF094067	EF094111
Silene vulgaris	North America	VLGNA08	EF094042	EF094090	EF094068	EF094112
Silene vulgaris	North America	VLGNA10	EF094041	EF094091	EF094069	EF094113
Silene vulgaris	Europe	VLG1	AF519072	AF518957	AF518929	AF518985
Silene vulgaris	Europe	VLG10	AF519073	AF518958	AF518930	AF518986
Silene vulgaris	Europe	VLG11	AF519074	AF518959	AF518931	AF518987
Silene vulgaris	Europe	VLG13	AF519075	AF518960	AF518932	AF518988
Silene vulgaris	Europe	VLG16	AF519076	AF518961	AF518933	AF518989
Silene vulgaris	Europe	VLG17	AF519077	AF518962	AF518934	AF518990
Silene vulgaris	Europe	VLG2	AF519078	AF518963	AF518935	AF518991
Silene vulgaris	Europe	VLG20	AF519079	AF518964	AF518936	AF518992
Silene vulgaris	Europe	VLG22	AF519080	AF518965	AF518937	AF518993
Silene vulgaris	Europe	VLG23	AF519081	AF518966	AF518938	AF518994
Silene vulgaris	Europe	VLG24	AF519082	AF518967	AF518939	AF518995
Silene vulgaris	Europe	VLG25	AF519083	AF518968	AF518940	AF518996
Silene vulgaris	Europe	VLG26	AF519084	AF518969	AF518941	AF518997
Silene vulgaris	Europe	VLG27	AF519085	AF518970	AF518942	AF518998
Silene vulgaris	Europe	VLG29	AF519086	AF518971	AF518943	AF518999
Silene vulgaris	Europe	VLG30	AF519087	AF518972	AF518944	AF519000
Silene vulgaris	Europe	VLG32	AF519088	AF518973	AF518945	AF519001
Silene vulgaris	Europe	VLG33	AF519089	AF518974	AF518946	AF519002
Silene vulgaris	Europe	VLG35	AF519090	AF518975	AF518947	AF519003
Silene vulgaris	Europe	VLG36	AF519091	AF518976	AF518948	AF519004
Silene vulgaris	Europe	VLG38	AF519092	AF518977	AF518949	AF519005
Silene vulgaris	Europe	VLG4	AF519093	AF518978	AF518950	AF519006
Silene vulgaris	Europe	VLG43	AF519094	AF518979	AF518951	AF519007
Silene vulgaris	Europe	VLG44	AF519095	AF518980	AF518952	AF519008
Silene vulgaris	Europe	VLG46	AF519096	AF518981	AF518953	AF519009
Silene vulgaris	Europe	VLG5	AF519097	AF518982	AF518954	AF519010
Silene vulgaris	Europe	VLG7	AF519098	AF518983	AF518955	AF519011
Silene vulgaris	Europe	VLG8	AF519099	AF518984	AF518956	AF519012
Silene latifolia	North America	LATNA1	EF088853	EF091495	EF088891	EF091533
Silene latifolia	North America	LATNA14	EF088863	EF091503	EF088893	EF091542
Silene latifolia	North America	LATNA15	EF088860	EF091504	EF088904	EF091543
Silene latifolia	North America	LATNA16	EF088857	EF091505	EF088902	EF091544
Silene latifolia	North America	LATNA19	EF088868	EF091506	EF088898	EF091545

Chapter 3:

Population structure and cytonuclear disequilibrium in the native range of *Silene vulgaris* reshaped by founder effects and admixture in the introduced range[3]

[3] Formatted for submission as a coauthored manuscript: Keller, S.R., C.F. Farnum, and D.R Taylor

Abstract

The establishment and spread of species into new ranges is an increasingly frequent consequence of human-mediated dispersal. The genetics of an invasion will reflect the evolutionary history of a species in its native range (population structure), the breadth of sampling and amount of propagule pressure (founder effects), and the degree of mixing between previously isolated lineages (admixture). These processes act during invasion to shape diversity at both the population and genome levels, leading not only to changes in the distribution of genotypes among sites, but also to new allelic combinations within and between genomes. In this study, we investigated the population genetics of *Silene vulgaris*, a weedy plant introduced to North America from Europe during the past two centuries. Amplified fragment length polymorphisms and chloroplast haplotypes were used to characterize the distribution of nuclear and cytoplasmic diversity in both ranges, assign introduced genotypes to their putative source regions, and examine changes in nuclear and cytonuclear linkage disequilibium. We found that S. *vulgaris* is genetically differentiated into five native range demes that showed different spatial patterns of occurrence in Europe, suggesting a recent isolation in refugia or drift during the last post-glacial expansion. The AFLP demes showed a significant association with chloroplast haplotypes in Europe. The introduction to North America has resulted in an undersampling sampling of deme richness and a shift in deme frequencies. Assignments of genotypes suggested multiple introductions originating from eastern and western Europe. Deme frequencies and historical records suggest the greatest amount of propagule pressure came from eastern Europe during the late 19th and early 20th centuries. Despite multiple introductions, admixture has not yet homogenized demes in North

America. In contrast, the associations between cpDNA and AFLP demes have been dissolved, and new pairwise associations among chloroplast and nuclear loci have developed. These results highlight how the invasion process sampled native range evolutionary history and reshaped the genetics in the introduced range at different scales of organization.

Keywords: population structure, AFLP, chloroplast, admixture, colonization, founder effects, invasion, range expansion

Introduction

The largely stochastic nature of human-mediated dispersal and colonization can profoundly alter the amount and spatial distribution of genetic diversity in invading species. Demographic events during colonization can result in founder effects (Nei et al. 1975, Amsellem et al. 2000, Tsutsui et al. 2000) and/or cause admixture of previously isolated lineages (Guinand and Esteal 1996, Ellstrand and Schierenbeck 2000, Kolbe et al. 2004, Williams et al. 2005). Founder effects may reduce or otherwise shift the representation of genotypes compared to the native range. Admixture can yield two related outcomes: demographically, it results in genotypes from separate introductions co-occuring within local populations (Kolbe et al. 2004), and genetically, it results in new genotypic combinations of alleles through mating and recombination (de la Vega et al. 1991, Lee 2002, Lavergne and Molofsky 2007). The intensity of sampling from the native range may influence both founder effect and admixture, reflecting an interaction between the historical events that shape genetic diversity in the native range and the contemporary events that sample from and redistribute that diversity (Taylor and Keller 2007).

In addition to understanding how invasions redistribute diversity at the population and genome levels, it is often desirable to reveal which native range regions contributed invasive genotypes. For example, by knowing introduction sources, studies can account for neutral phenotypic evolution during invasion due to founder effect (Chapter 1, Chapter 5). Knowing the sources of invasion also allows an assessment of the degree of environmental matching between the native and introduced ranges, and hence the importance of preadaptation to invasion success (Peterson and Vieglais 2001, Sakai et al.

2001, Chapter 4). Finally, knowledge of invasion sources permits a more informed assessment of biocontrol agents for the management of particularly noxious invaders (Gaskin et al. 2005).

Tracing genotypes back to their source regions is not always a straightforward task. Many invasive species are widespread and excellent dispersers across their native geographic ranges (Kolar and Lodge 2001), which dissipates genetic structure in the native range structure. Additionally, invasions by definition are non-equilibrium demographic scenarios, and individuals can occur in populations that do not resemble the gene pools to which they historically belonged (Davies et al. 1999). In general, tracing invasion sources with 'population' as the sampling unit will often obscure ancestor-descendent relationships in the presence of admixture (Kolbe et al. 2004, Durka et al. 2005, Williams et al. 2005, Lavergne and Molofsky 2007).

A solution to this problem is to use individual-based approaches for identifying genetic structure (Davies et al. 1999, Manel et al. 2005, Waples and Gaggiotti 2006), such as Bayesian model-based clustering (Pritchard et al. 2000, Falush et al. 2003). The general philosophy behind model-based clustering is to use patterns of allele covariance to identify divergent populations (or demes), and assign individuals to their putative sources based on their multi-locus genotypes. This greatly increases statistical power over frequency-based methods, as individuals and not populations are the sampling unit. Second, even under conditions of moderate gene flow (N_m~5) or cryptic divergence (F_{ST}~0.05), the ability to detect population structure is fairly powerful, given enough polymorphic loci (Pritchard et al. 2000, Rosenberg et al. 2002, Hauser et al. 2006, Ostrowski et al. 2006, Waples and Gaggiotti 2006). The ability to assign introduced

genotypes without having to define populations *a priori* is especially useful when there is little knowledge of which native populations should be considered as possible sources. A related approach, the assignment test, relies on having a user-defined set of candidate source populations (Rannala and Mountain 1997, Waser and Strobeck 1998, Manel et al. 2005). While a few recent studies on species introductions have employed the assignment test approach (Gaskin et al. 2005, Genton et al. 2005b), their effectiveness at tracing invasion origins has been limited. To our knowledge, no study has applied model-based clustering to identify structured demes in the native range and to trace the origins of introduced genotypes back to those demes.

Silene vulgaris is a weedy plant native to Eurasia and introduced to North America about 200 years ago. Previous work using chloroplast DNA (cpDNA) showed that haplotype diversity was relatively unstructured at regional scales in both ranges and that North America contained a high level of gene diversity (McCauley et al. 2003, Taylor and Keller 2007). While cpDNA variation is useful for tracing patterns of migration because it is generally maternally inherited through seeds, all loci on the chloroplast genome are thought to effectively represent a single linkage group with a common evolutionary history. This causes analyses of cpDNA to be subject to the stochastic variance of gene coalescence that comes from sampling a single locus (Knowles 2004). So while it seems clear that the introduction of *S. vulgaris* to North America did not involve a severe bottleneck, the population genetics of the invasion process are still relatively unknown.

In this study, we surveyed amplified fragment length polymorphisms (AFLPs) from a geographically widespread sample of *S. vulgaris* in its native and introduced ranges. Our

goal was to combine the method of Bayesian model-based clustering with the high resolution of AFLP multi-locus genotyping to determine if nuclear genetic diversity exhibits structure in the native range and how the invasion sampled from and redistributed that diversity. Specifically, we ask (1) did introduced *S. vulgaris* originate from a single diverse source or multiple sources, (2) what were the geographic locations of the source region(s) in the native range, (3) is there evidence for a founder effect, or did the process of invasion sample nuclear diversity representatively, as appears true for cpDNA, and (4) has genetic admixture altered the pattern of allelic associations, both within the nuclear genome or intergenomically through associations with the chloroplast?

Methods

Study Species

Silene vulgaris L. (Moench) Garcke (= *S. cucubalus, S. inflata*) is a widespread, abundant perennial herb in the Caryophyllaceae native to temperate climates in Eurasia. Its range extends as far south as Morocco and the Middle East at higher elevations, northwards into Scandinavia, and from the Atlantic coast eastward into Russia and western Asia where it becomes diffuse (Marsden-Jones and Turrill 1957, Jalas and Suominen 1986). Throughout its native range, *S. vulgaris* inhabits a variety of natural environments including rocky outcrops, wood margins, river banks, and cliff faces, as well as human-disturbed areas such as roadsides, railroad embankments, abandoned lots, and hayfields. It is believed that *S. vulgaris* was restricted to Mediterranean refugia during times of Pleistocene glaciations and recolonized Europe in post-glacial expansions during the warm inter-glacial periods (Marsden-Jones and Turrill 1957, Runyeon and

Prentice 1997, Taylor and Keller 2007). In North America, *S. vulgaris* was introduced during the late eighteenth or early nineteenth centuries, with the earliest collections originating in Boston, Quebec City, and Philadelphia (Cutler 1785, Pursh 1814, Chapter 4). The introduction appears to have been accidental and early American botanists on the East Coast observed *S. vulgaris* growing out of mounds of soil and agricultural waste used to ballast ships during the trans-oceanic trip to North America (Martindale 1876, 1877, Brown 1878). Since this time, the invasion has spread out from its centers of origin in the northeast to occupy much of temperate North America (Chapter 4).

Collections and DNA Isolation

We adopted a sampling design that involved collecting *S. vulgaris* from many locations across a broad geographic area spanning the breadth of the known ranges in Europe and North America (Appendix A). Between 1 and 9 individuals were sampled per site (mean = 3), for a total of sample size of 404 plants (233 individuals across 95 European sites, 171 individuals across 42 North American sites). Individuals were either collected as seeds or as leaf tissue dried on silica gel. Many samples came from human disturbed sites such as fields and roadsides where *S. vulgaris* is abundant, especially in North America. However, samples were also collected from more natural sites such as wildflower meadows, montane scree slopes, riverbanks, and rocky cliff faces. Included in the sample were several specimens of *S. uniflora* ($N = 5$), an interfertile sister taxon which is sometimes considered a subspecies of *S. vulgaris*, as well as *S. vulgaris* ssp. *angustifolium* ($N = 2$) (Marsden-Jones and Turrill 1957). The breadth of sampling across the native range, in terms of both geography and habitat, should permit a confident

assessment of the distribution of genetic diversity and the manner in which invasion sampled from this diversity. Seeds were germinated at the University of Virginia greenhouse and genomic DNA was isolated from either fresh or dried leaf tissue using DNeasy plant miniprep kits (Qiagen, Valencia, CA).

Amplified Fragment Length Polymorphisms (AFLPs)

AFLP reactions were performed according to a general protocol (Vos et al. 1995) with slight modifications using reagents from the Applied Biosystems Plant Mapping Kit for large genomes (www.appliedbiosystems.com). Specifically, ca. 50-100 ng of genomic DNA was digested at room temperature overnight with a pair of restriction enzymes (*Eco*RI and *Mse*I; New England Biolabs) and ligated to double-stranded adapters in a single 10 µl reaction. Restriction-ligation products were diluted 1:10 with $TE_{0.1}$ buffer. Preselective amplification was then performed in 10 µl reaction volumes. PCR conditions during preselective amplification were an initial hold for 2 min at 72ºC, 20 cycles of denaturing at 94ºC for 20 sec, annealing at 56°C for 30 sec, extension at 72°C for 2 min, and a final extension at 60°C for 30 min. Preselective amplifications were then diluted 1:10 with $TE_{0.1}$ buffer. Selective amplifications were performed using four different primer pair combinations on each sample, including two FAM labeled primer pairs (*Eco*RI: ACA / *Mse*I: CTC, *Eco*RI: ACA / *Mse*I: CAC) and two TAMRA labeled primer pairs (*Eco*RI: ACC / *Mse*I: CAG, *Eco*RI: AGC / *Mse*I: CTG). Amplifications were done in separate 10 µl reactions. PCR conditions were an initial hold for 2 min at 94ºC, 10 cycles of denaturing at 94ºC for 20 sec, annealing at 66°C for 30 sec (decreasing by 1°C per cycle to 57°C on the final cycle), and extension at 72°C for 2 min, followed by 20

cycles of denaturing at 94ºC for 20 sec, annealing at 56°C for 30 sec, and extension at 72°C for 2 min, with a final extension at 60°C for 30 min. FAM and TAMRA-labeled selective amplification products were then multiplexed and separated by electrophoresis on an ABI 377 or ABI 3130xl automated sequencer. Negative controls were run at each step in the AFLP protocol and several duplicate samples were run to evaluate reproducibility.

Fluorescent peaks were sized against the ROX-500 standard using GENEMAPPER v4.0 software (Applied Biosystems). All samples were normalized to the sum of the sample signal within GENEMAPPER, calculated across the entire project. We initially observed that products run on different electrophoresis platforms (377 vs. 3130xl sequencers) appeared to have different patterns of mobility among homologously sized fragments. This was subsequently verified using sample duplicates and necessitated a method for "mapping" homologous bins across platforms and defining a platform-specific size range for each bin. To accomplish this, we ran 30-40 sample duplicates per primer pair on each platform and used GENEMAPPER to size the fragments (scoring threshold 50 RFU). We then manually compared duplicate samples across platforms, matched homologous bins, and obtained a bin set definition for each platform that we used to score the remaining samples. Automated scoring was manually verified for all samples by the same observer. Peak presence/absence was then converted to binary (0/1) coding for further analyses. Several samples exhibited amplification products for some primer pairs characterized by few strongly amplified peaks. These outcomes were consistent even after re-amplification and analysis. Null alleles in these samples were assigned missing data values.

Chloroplast DNA (*cpDNA*) *Polymorphism*

From previously published sequence data (Taylor and Keller 2007), we identified three clades in the cpDNA genealogy for *S. vulgaris* (Figure 1A). We searched the sequence alignment for polymorphisms in genomic insertions/deletions (indels) or restriction fragment length polymorphisms (RFLPs) that defined these regions of the genealogy. Of the three clades present, two could be positively assigned using indel or RFLP polymorphism, while the third was assigned by process of elimination. Clade A was defined by a 7 bp insertion in an intergenic spacer between *trn*L (UAA) and *trn*F (GAA). Primers were designed to amplify a 123 bp region that flanks this indel polymorphism (forward: 5' tttaggtcttcaaaaagaggaactc 3'; reverse: 5' gcaggcagtactccgttgag 3'). PCR was performed in 25 µl reactions with an initial hold at 96 ºC for 30 sec, 40 cycles of denaturing at 96 ºC for 30 sec, annealing at 50 ºC for 15 sec, and extension at 60 ºC for 4 min. From the resulting PCR product, 8.5 µl was run out on a 3% agarose gel for 80 min. at 105 V. Individuals belonging to Clade A had a product 7 bp longer than those not belonging to Clade A (Figure 1B). Clade B was defined by a single nucleotide polymorphism (g/t) that created a *Psi*I RFLP within the same *trn*L-*trn*F intergenic spacer. This region was PCR amplified in a 25 µl reaction using the universal primers e and f as previously described (Ingvarsson and Taylor 2002). Following amplification, 8.9 µl of PCR product was combined with 0.1 µl of *Psi*I enzyme and 1 µl of enzyme buffer (New England Biolabs) and digested at 37 ºC. The resulting product was run on a 3% agarose gel for 45 min. at 105 V. Individuals belonging to Clade B had fragments of size 287 bp, 68 bp, 12 bp, while those not belonging to Clade B had fragments 287 bp and 80 bp. During electrophoresis, the 12 bp fragment typically migrated off the end of the gel, leaving a clearly defined size difference between the 68

and 80 bp bands (Figure 1C). Samples lacking the 7 bp indel and the *Psi*I restriction site belonged to Clade C.

Gel banding patterns were scored blind by two independent observers and verified for agreement. A number of samples that had been previously sequenced for the *trn*L-*trn*F region ($N = 32$) were also typed to serve as positive controls on the method and the scoring. A total of 224 samples were typed for their cpDNA clade, including 153 European and 71 North American samples.

Statistical Analyses

To test for native range genetic structure and assign introduced genotypes to their ancestral demes, we used Bayesian model-based clustering (Pritchard et al. 2000, Falush et al. 2003, 2007) implemented in the software STRUCTURE v2.2. This version of STRUCTURE accounts for the presence of null alleles and thus accommodates the use of dominant markers such as AFLP (Falush et al. 2007). Briefly, the algorithm behind STRUCTURE attempts to partition multi-locus genotypes among a number of demes (K), where the value of K is chosen by the user. Clustering is accomplished by finding groups in the sample that maximize the likelihood function under the assumption of linkage equilibrium within demes. Multiple models are run which vary in K and a "best" model is selected from among the candidate models. Allele frequencies for each of the K demes are then used to probabilistically assign membership to individual genotypes.

For most STRUCTURE parameters, default settings were used. We chose the correlated allele frequencies model, which assumes demes drifted apart recently from a common ancestral population, and gives good performance when divergence is low (Falush et al.

2003). Genetic admixture among demes was accounted for by estimating the admixture parameter, α. Values of $\alpha << 1$ indicates most individuals come from a single deme while $\alpha >> 1$ indicates the presence of extensive admixture in the sample. We conducted an initial search for the most likely value of K by performing five replicate simulations at each value of K ranging from 1 to 15 demes. Runs had a burn-in of 10,000 iterations, followed by parameter estimation over an additional 10,000 iterations. Both α and model likelihood were monitored during these runs and were found to converge during burn-in. We followed this procedure for each of several different parameter sets that held different assumptions about the data. First, we analyzed the entire dataset (pooled across continents) using an admixture model that allows for past migration among demes to have generated genotypes with partial membership in more than one source. Second, we used a no-admixture mode that constrains genotypes to belong to a single deme. Although the no-admixture model is less biologically plausible for most species (and hence less preferable), it is generally more powerful at detecting genetic structure when divergence among demes is cryptic (Pritchard et al. 2007). Finally, we used an admixture model on just the native range samples from Europe. This was done in case extensive admixture in the invasive range had resulted in samples that were largely composed of recombinant genotypes, which we reasoned might obscure a signature of native range structure. Model selection proceeded by choosing the model that maximized the second order rate of change in the Ln(probability of the data|K) as a function of increasing K (Evanno et al. 2005). To obtain more precise parameter estimates of allele frequencies and assignment probabilities, we performed an additional run with K fixed at the value from the best model for 10,000 burn-in and 100,000 sampling iterations. For all

runs, we visually inspected model parameters to ensure the MCMC chain had reached convergence. Longer runs (10^6) were also explored, but were not found to affect convergence or improve the precision of parameter estimates (see also Larson et al. 2004).

Simulation studies have shown that when genetic structure is nested or hierarchical, STRUCTURE tends to identify the highest level in the hierarchy (Evanno et al. 2005). Thus, it is possible to identify additional nested structure by re-running the algorithm within each of the initial demes STRUCTURE identifies. We used this approach by taking the demes identified under the admixture model and re-running STRUCTURE on each deme separately. Initial searches of the parameter space for K were done as above. When additional structure was found, we chose the best model and performed one additional run at this fixed value of K for the purpose of parameter estimation, as described above. Individual genotypes were assigned to demes based on a majority rule criterion of their assignment scores from STRUCTURE. To examine how deme frequencies have shifted between the native and introduced ranges, and across different cpDNA backgrounds, we analyzed the distribution of individuals assigned to each deme with *G*-tests of independence (SAS Institute 1999).

We determined the magnitude and significance of genetic divergence among the demes using the Bayesian estimator of population structure for dominant data, θ^B, as implemented in the software HICKORY (Holsinger et al. 2002). Divergence among demes was estimated for the multi-locus AFLP data and the cpDNA haplotypes (the latter was a measure of cytonuclear disequilibrium) in separate analyses for Europe and North America. Default settings were used in HICKORY (including uninformative priors), with

5,000 burn-in iterations followed by 25,000 sampling iterations. Longer burn-in times (up to 50,000) and sampling iterations (500,000; thinned at every 100 iterations) on a subset of models produced the same results. We ran full models, which estimate the inbreeding coefficient *f*, as well as *f*-free models (which estimate θ^B over the full range of possible values of *f*), and compared the posterior distributions of θ^B to examine if it was necessary to take uncertainty about the rate inbreeding into account. In all cases, the difference in the posterior distributions of θ^B between models overlapped zero, suggesting that our ability to estimate *f* does not significantly affect estimates of genetic differentiation. Given that dominant markers may produce unreliable estimates of *f*, we report θ^B values for just the *f*-free models (Holsinger and Lewis 2003).

In addition to examining divergence among demes, we tested for a geographic component of genetic structure in Europe by dividing the sample into three physiographic regions: eastern Europe, southern Europe, western Europe, according to the boundaries defined in Taylor and Keller (2007). These groupings were chosen because they reflect regional features defined by physical barriers and known dispersal corridors for plants and animals since the last ice age (Hewitt 1996, Taberlet et al. 1998, Hewitt 2000, Petit et al. 2003). First, we partitioned the genetic variance in AFLP and cpDNA among the geographic regions in Europe using HICKORY and ARLEQUIN v3.0, respectively (Excoffier et al. 2005). Second, we analyzed the frequency distribution of demes and clades across geographic regions using *G*-tests (SAS Institute 2000).

We examined how invasion may have resulted in genetic admixture among demes. If mating among demes generated recombinant genotypes, North American samples would show (1) differences in the composition of multi-locus genotypes, (2) changes in the level

of linkage disequilibrium among loci, and (3) more heterogeneous assignment scores. We tested for changes in multi-locus genotype space by performing principal coordinates analysis on the binary AFLP data matrix using the R PACKAGE v4.0 (http://www.bio.umontreal.ca/Casgrain/en/labo/R/v4/index.html). Genetic distances among individuals were defined as 1- Jaccard's similarity index, J_{ij}, which calculates similarity based on shared bands among individuals and ignores shared absences ($J_{ij} = a/(a+b+c)$, where a is the number of bands shared by individuals *i and j*, b is the number of bands in i but not j, and c is the number of bands in j but not i). We used the first two eigenvectors from the principal coordinates analysis in two-factor ANOVAs testing for differences among demes and continents. Changes in linkage disequilibrium were tested by performing exact tests in ARELQUIN, with significance assessed via 10,000 iterations of the MCMC chain after a burn-in of 10,000 iterations. Finally, the maximum assignment value for each individual was obtained from the STRUCTURE output and used to test for differences between continents in the degree to which genotypes were mixtures of different demes using a Wilcoxon two-sample test (SAS Institute 1999).

Results

A total of 267 loci were scored, averaging 67 loci per primer pair. Band frequencies ranged from 0.01-0.98, though 67% of loci had frequencies <0.10. All 404 individuals examined possessed unique multi-locus genotypes. In Europe, all 267 loci were polymorphic, while North American samples were monomorphic for five of these loci. Pairwise divergence and gene diversity over loci were also slightly higher among European samples (Table 1).

All three initial STRUCTURE models showed the presence of genetic structure within the AFLP data, as indicated by the rise in model likelihood with increasing K (Figure 2A). All models showed a decelerating likelihood function after K = 3, suggesting the highest level of structure consisted of three demes (Figure 2B). Since all three models returned very similar inference on K, and since there was evidence of weak admixture across runs ($\alpha = 0.1707 \pm 0.0068$, SE), we chose the admixture model with both continents for further analysis. Searches for additional structure nested within each deme showed an increase in model likelihood with K for demes 1 and 2 (Figure S1). In deme 1, the likelihood plateaued at K=2; however, genotype assignment probabilities showed each individual had roughly equal probability (1/K) of belonging to each subdeme, an indication that the structure is artefactual (Pritchard et al. 2007). Deme 3 also contained no indication of nested structure. In contrast, model likelihoods clearly indicated additional structure nested within deme 2 (Figure S1). Analysis of the rate of change in likelihood as a function of K showed a maximum at K=3, supporting the splitting of deme 2 into three additional demes (hereafter referred to as demes 2.1 - 2.3). Therefore, we consider the data to support the existence of a total of five demes.

Genetic Structure in the Native Range

The five AFLP demes showed marked frequency differences in their distribution across Europe ($G = 151.75$, d.f. = 6, $P < 0.0001$; Figure 3). This geographic clustering was evident despite the fact that STRUCTURE was provided no information on where samples originated from, which strongly supports divergence in the nuclear genome as a result of historical isolation in Europe. STRUCTURE supported deme 2.3 as having the

lowest divergence from the estimated ancestral allele frequencies (not shown), suggesting this widely distributed southern European deme may represent the closest extant deme to the common ancestor of all *S. vulgaris* demes. The two other lineages nested within deme 2 are deme 2.1, which is narrowly distributed in Croatia and southern Italy, and deme 2.2 which has a wider distribution extending from northern Italy into most of central and eastern Europe (Figure 3). Deme 2.1 contains one representative of *S. vulgaris* ssp. *angustifolium*, while the other *angustifolium* specimen was assigned to deme 2.2. Another widespread but less frequent lineage is deme 1, which occurs in the Caucuses of southern Russia, in the central Alps, and among samples of *S. uniflora* along the Atlantic coast. Deme 3 shows a distribution primarily centered in western Europe, north of the Pyrenees, and also contains representatives of *S. uniflora* from the coast of Ireland and Spain. Interestingly, an apparent suture zone is present in central Europe, suggesting the region just north of the Alps may represent an area of secondary contact among the demes (Figure 3).

The magnitude of AFLP allele frequency variance among demes in Europe was moderate but significant ($\theta^B = 0.0791$; Table 2). Genotypes with membership in different demes also separated along the first two axes of the principal coordinates analysis (Figure 4), and were highly differentiated according to ANOVA (Table 3). Somewhat surprisingly, when divergence was considered from the standpoint of variance in AFLP allele frequencies among geographic regions, structure was much lower ($\theta^B = 0.0056$; Table 2), despite the pronounced differences in deme frequencies among regions. Thus, spatial structure in the native range appears to result largely from the geographic

distribution of demes, which are themselves defined by covariances among sets of loci, and less from the locus-by-locus variances in allele frequency among regions.

Similar results for the variance among demes and geographic regions were obtained for the chloroplast genome. Divergence in cpDNA clade frequency among demes in Europe was very close to the multi-locus average for AFLP loci, but was non-significant among geographic regions (Table 3).

Genetic Diversity of the Invasion Inoculum

Four of the five demes present in Europe have invaded North America, with the exception of the restricted deme 2.1. While this has resulted in a slight reduction in deme richness, deme frequencies have experienced a more pronounced shift during the invasion process. The eastern European deme 2.2 has increased in frequency from 23% to 52%, while the southern European deme 2.3 has decreased from 28% to 8%. Overall, deme frequencies show a significant shift between continents ($G = 52.18$, d.f. = 4, $P < 0.0001$). This shift in frequencies is still highly significant when the rare deme 2.1 is not included in the analysis ($G = 46.52$, d.f. = 3, $P < 0.0001$).

Despite strong sampling effects on among-deme diversity during invasion, the sampling of diversity within-demes was representative of the source. The PCoA analysis indicated no evidence of a shift in the allelic composition of multi-locus genotypes (Figure 4), and ANOVA supported that genotypes from both continents occupied the same multi-locus genetic space (Table 2). Consistent with this, the maximum assignment probabilities from STRUCTURE were high for both North American genotypes (mean ± SE: 0.80 ± 0.01) and European genotypes (0.78 ± 0.01), and continents showed no

significant difference in uniformity of assignments (Wilcoxon test: $Z = 1.0853$, $P = 0.2778$). Differentiation among demes in North America was significant and of the same magnitude observed for Europe ($\theta^B = 0.087$; Table 3). The difference in the posterior distribution between continents was not significantly different from zero (mean change between continents in $\theta^B = 0.0076$; 95% C.I. -0.0141-0.0292), indicating that the invasion has not altered the divergence among demes.

Chloroplast clades showed a slight but significant shift in frequencies during invasion ($G = 7.41$, d.f. = 2, $P = 0.0247$). This is largely in agreement with Taylor and Keller (2007) who showed no bottleneck in phylogenetic diversity, but a slight shift in the distribution of pairwise haplotype differences compared to Europe, suggesting of some sampling effects during the invasion.

Nuclear and Cytonuclear Linkage Disequilibria

Levels of linkage disequilibrium (LD) within the nuclear genome were about 2.5 times higher among pairwise comparisons than expected by chance (Table 4). The proportion of total LD tests that were significant was slightly lower (<1%) for genotypes from North America relative to Europe, suggesting allelic associations have persisted over the course of the invasion. Cytonuclear LD was also significant on both continents, although the proportion of significant associations was reduced from ca. 14% (38 of 267 tests) in Europe to ca. 10% (26 of 250 tests) among North American samples (Table 4). Interestingly, associations between specific AFLP and cpDNA markers were not well conserved. Among the 250 cytonuclear pairwise comparisons common to both continents, 49 out of the 53 significant associations were significant only on one

continent or the other. Thus, the overall similarity in cytonuclear disequilibria does not seem due to conservation of particular pairwise associations, but to the formation of novel associations among nuclear and cytoplasmic loci.

Dissolution of cytonuclear LD during invasion is also apparent in the associations among demes and clades. In Europe, there was a significant association between AFLP demes and cpDNA clades (G = 19.74, d.f. = 6, P = 0.0031), while the pattern of association was consistent with random expectations among North American samples (G = 10.62; P = 0.1009). However, small expected sample sizes for deme 1 in some cells may bias goodness of fit tests. When deme 1 is removed from the analysis, the overall pattern is even more pronounced (Europe: G = 18.76, d.f. = 4, P = 0.0009; North America: G = 2.98, d.f. = 4, P = 0.5595). Thus, on average different nuclear-based demes of *S. vulgaris* have been experiencing different cytoplasmic backgrounds since invading North America (Figure 5).

Discussion

In this study, we found that the native range of *Silene vulgaris* is subdivided into spatially structured demes as a consequence of the evolutionary history of the species, probably in relation to past climate changes. These demes are now intermixed and in significantly different frequencies within the introduced range. However, despite this overall shift in deme occurrence since invasion, there is little evidence that extensive genomic integration has dissolved the associations present in the native range. One exception might be found in the pattern of disequilibria between the nuclear and chloroplast genomes. Below, we interpret these findings in light of historical knowledge

of *S. vulgaris* in its native and introduced ranges, and discuss mechanisms that shape and re-shape genetic structure leading up to and during the process of invasion.

Native Range Structure and Evolutionary History

The overall geographic distribution of demes in Europe suggests a phylogeographic history related to the Pleistocene glacial period (Figure 3). Specifically, eastern and western Europe are dominated by different demes (2.2 and 3, respectively). The distribution of the demes is consistent with many species of plants and animals that expanded out of glacial refugia located near the Iberian and Balkan peninsulas (Hewitt 1996, Taberlet et al. 1998, Hewitt 2000). In addition, an area of apparent admixture among AFLP demes in central Europe matches a well documented zone of secondary contact between populations of many species expanding out of different southern refugia (Petit et al. 2003).

The overall pattern is also reminiscent of the congener, *S. latifolia*, which shows an east-west phylogeographic break in Europe (Ironside and Filatov 2005, Taylor and Keller 2007). In *S. latifolia*, Taylor and Keller (2007) estimated that eastern and western clades split 483-725 kya (thousand years ago), probably a reflection of divergence that occurred in ancient refugia well before the most recent glacial maximum (ca. 18-21 kya, Webb and Bartlein 1992). This raises the question in *S. vulgaris* of whether divergence among demes reflects a similarly ancient history of isolation or a more recent divergence. While we cannot apply a molecular clock to the AFLP data, the magnitude of θ^B suggests that isolation among demes happened more recently, perhaps during or since the most recent glacial cycle.

How does recent nuclear genome divergence fit with our understanding of the chloroplast genome? While *S. vulgaris* cpDNA lacks structure between eastern and western Europe that might indicate past isolation in refugia (this study, Taylor and Keller 2007), there is a significant association between AFLP demes and cpDNA clades (Figure 5). The main source of non-independence between genomes is caused by the higher frequency of cpDNA clade A and reduced frequency of clade C among western European deme 3 individuals, whereas the eastern and southern European demes 2.2 and 2.3 show the opposite trend. The timing of the split between clade A and the other clades in the cpDNA genealogy is bracketed by a pair of mutations which date to 448–730 kya (Taylor and Keller 2007), a time very consistent with the east-west split in *S. latifolia*. Thus a plausible scenario is that *S. vulgaris* did experience divergence in ancient refugia, but that expansion during the ensuing interglacial periods partially admixed cpDNA haplotypes (and probably homogenized nuclear divergence through recombination). Cytonuclear associations could then have gradually developed as both nuclear and cpDNA loci drifted in frequency during the most recent refugial isolation or during post-glacial expansion into northern Europe.

It is noteworthy that the estimate of chloroplast F_{ST} among the demes is very close to the multi-locus average from AFLP (Table 3). Basing our inference solely on the chloroplast acting as a single locus would have provided little information on genetic structure in Europe (Table 3; see also Taylor and Keller 2007). This serves as a reminder that analyses based on a single locus can be misleading, and that incorporating a large number of loci (in this case, using AFLP) along with methods that take advantage of covariances among loci is a much more powerful approach.

Together, these results suggest an evolutionary history of *S. vulgaris* in its native range that is characterized by repeated range contractions, expansions, periods of admixture followed by spatially restricted gene flow, the development of intra-, inter-genomic disequilibria. It is difficult to ignore how the recurrent theme of range expansion and admixture in the face of repeated bouts of global change might have pre-adapted *Silene vulgaris*, along with many other temperate zone taxa, for invasion success (Dynesius and Jansson 2000). For certain, the evolutionary history in the native range shaped the distribution of diversity from which the contemporary invasion of North America sampled and re-distributed.

Founder effects during invasion

The genetics of the *S. vulgaris* invasion can best be described by multiple introductions that sampled across native range demes but also involved the action of founder effects. The founder effects included both a subsampling of deme richness and a shift in the relative frequencies of the demes. These founder effects are probably a consequence of structure in Europe, since sampling processes should more often obtain a biased sample of diversity when a range is structured. The fact that all but the lowest frequency demes in Europe were introduced to North America, however, points to a fairly broad geographic sampling process that contributed introduced genotypes. The shift in deme frequencies is then most likely attributable to variation in the propagule pressure coming from different demes into North America.

During the time of initial establishment of *S. vulgaris* in North America (pre-1850), immigration was dominated by colonists from western Europe (Gibson and Lennon 1999,

Chapter 4). These colonists probably contributed genotypes from deme 3, which occur in high abundance in western Europe, and possibly genotypes of deme 1 as well. During the second half of the 1800's and into the early 1900's, a large influx of Eastern Europeans settled in the northeastern and north-central United States (Gibson and Lennon 1999). These colonists probably brought deme 2.2 and possibly deme 2.3 genotypes with them. American botanists in Philadelphia and New York City first noticed *S. vulgaris* (often recorded as *S. inflata* or *S. cucubalus*) growing out of mounds of ballast waste from European ships piled near the docks (Martindale 1876, 1877, Brown 1878). Though earlier collections exist (Cutler 1785, Pursh 1814), this period corresponds to a large increase in the observed density and range extent of *S. vulgaris* in North America (Chapter 4), perhaps as a result of a surge in propagule pressure associated with the mass immigration from eastern Europe. This increase in propagule pressure may have caused deme 2.2 to rise in frequency relative to the other demes. Since the introduction of *S. vulgaris* appears accidental, the change in deme frequencies during invasion was essentially a stochastic byproduct of the pattern of human migration.

While a stochastic colonization process is a plausible explanation for the shift in deme frequencies, we cannot rule out the potential for interdemic selection (Lewontin 1965). Consider the possibility that deme 2.2 possesses attributes that permit it to become more widespread and abundant in North America compared to the other demes. For example, in a companion study, we measured plant performance in the field at common gardens within North America and found that members of deme 2.2 do show a faster life history schedule and increased fecundity compared to other demes (Chapter 5). Whether these life history differences translate into differential deme productivity is a difficult

hypothesis to test. Some insight might be possible through the study of historical collections, to see if current deme frequencies are proportional to frequencies earlier in the invasion, or if instead deme 2.2 has been increasing in relative abundance since its time of introduction.

Interestingly, the diversity of founders from a deme was representative of the overall diversity of the source (Figure 4). The lack of additional structure within demes is evidence that most genetic variation is apportioned within as opposed to between populations within demes. Therefore, even fairly restricted sampling would likely obtain a diverse inoculum of the genetic variation from within a deme.

Genomic Disequilibria

The nuclear genome of this widespread collection of *S. vulgaris* contains significant levels of linkage disequilibrium (LD), roughly 2.5 times more than expected based on the number of pairwise tests conducted. A significant fraction of this LD probably stems from population structure, as allelic associations form the principal source of information for clustering-based analyses such as STRUCTURE. When multiple introductions during an invasion result in admixture, recombinant genotypes and changes in LD can be expected. However, we found levels of LD among North American genotypes were roughly similar to those in Europe. While genotypes with mixed assignments were found from both continents, assignment probabilities were on average as high or higher among North American genotypes, indicating persistence of the allelic associations since invasion (Figure 3). It would seem that even though demes are not spatially separated in North America, the frequency of interdemic mating is still low. While the invasion is about 200

years old, (roughly 100 generations), much of the demographic expansion has happened within the last 50-100 years (Chapter 4). Thus, we might expect the frequency of interdemic mating to increase in the future, which has possible consequences for plant fitness (Bailey and McCauley 2006).

The dissolution of cytonuclear associations in North America indicates at least partial mating and re-assortment of AFLP demes across different cpDNA backgrounds since the invasion. How does this reconcile with the nuclear results which show no major change in LD? One possibility is chloroplast capture from one deme onto the nuclear genetic background of another deme (Tsitrone et al. 2003). For example, when a migrant carrying a novel cytoplasm arrives in a population and mates with its residents, its seed offspring will repeatedly backcross with resident genotypes, leading to asymmetric introgression. This process is accelerated when non-neutrality of the chloroplast causes the migrant cytoplasm to have higher seed fitness. In *S. vulgaris*, the cytoplasmic genomes may be under direct or hitchhiking selection as a result of cytoplasmic male sterility (CMS). Previous work has shown variance among cpDNA haplotypes in male sterility, fecundity, and seed viability, presumably due to LD with CMS genes in the mitochondria (McCauley and Olson 2003). Thus selection may play a role in the pattern of cytonuclear associations in *S. vulgaris*.

In conclusion, we demonstrate that *Silene vulgaris* possesses an evolutionary history in Europe that has structured nuclear genetic variation into spatially distinct demes. The invasion of North America sampled from these demes unrepresentatively, resulting in a shift in deme frequencies between the continents. Clustering of multi-locus genotypes strongly suggested multiple introductions predominantly from eastern and western

Europe, and coincided with human propagule pressure from these areas. The persistence of allelic associations suggests admixture has not yet run its course. The break-up of cytonuclear associations points to introgression of nuclear and chloroplast genomes as a consequence of invasion. Our results will be important for future studies directed at measuring adaptive evolution during invasion. By comparing genotypes from the native and introduced ranges descended from the same deme, it will be possible to assess phenotypic evolution among introduced populations while controlling for the founder effects that accompany the invasion process (Chapter 1, Chapter 5).

Acknowledgements

We thank the many collaborators who have generously contributed tissue for DNA analysis over the years, especially Camille Barr, Dave McCauley, Dexter Sowell, Helena Storchova, and Lorne Wolfe. Allen Gunn helped with the AFLP genotyping. This work was supported by NSF DEB-0608358, NSF DEB-444176, and a Graduate School of Arts and Sciences Dissertation Year Fellowship.

References

Amsellem, L., J. L. Noyer, T. Le Bourgeois, and M. Hossaert-McKey. 2000. Comparison of genetic diversity of the invasive weed *Rubus alceifolius* Poir. (Rosaceae) in its native range and in areas of introduction, using amplified fragment length polymorphism (AFLP) markers. Molecular Ecology 9:443-455.

Bailey, M. F., and D. E. McCauley. 2006. The effects of inbreeding, outbreeding, and long-distance gene flow on survivorship in North American populations of *Silene vulgaris*. Journal of Ecology 94:98-109.

Brown, A. 1878. Plants introduced with ballast and on made land. Bulletin of the Torrey Botanical Club 6:255-258.

Cutler, M. 1785. An account of some of the vegetable productions, naturally growing in this part of America, botanically arranged. Memoirs of the American Academy of Arts and Sciences I:396-493.

Davies, N., F. X. Villablanca, and G. K. Roderick. 1999. Determining the source of individuals: multilocus genotyping in nonequilibrium population genetics. Trends in Ecology and Evolution 14:17-21.

de la Vega, M., P. Garcia, and R. Allard. 1991. Multilocus genetic structure of ancestral Spanish and colonial California populations of *Avena barbata*. PNAS 88:1202-1206.

Durka, W., O. Bossdorf, D. Prati, and H. Auge. 2005. Molecular evidence for multiple introductions of garlic mustard (*Alliaria petiolata*, Brassicaceae) to North America. Molecular Ecology 14:1697-1706.

Dynesius, M., and R. Jansson. 2000. Evolutionary consequences of changes in species' geographical distributions driven by Milankovitch climate oscillations. Proceedings of the National Academy of Sciences of the United States of America 97:9115-9120.

Ellstrand, N. C., and K. A. Schierenbeck. 2000. Hybridization as a stimulus for the evolution of invasiveness. Proceedings of the National Academy of Sciences of the United States of America 97:7043-7050.

Evanno, G., S. Regnaut, and J. Goudet. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular Ecology 14:2611-2620.

Excoffier, L., G. Laval, and S. Schneider. 2005. Arlequin ver 3.0: An integrated software package for population genetics data analysis. http://cmpg.unibe.ch/software/arlequin3.

Falush, D., M. Stephens, and J. K. Pritchard. 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics 164:1567-1587.

Falush, D., M. Stephens, and J. K. Pritchard. 2007. Inference of population structure using multilocus genotype data: dominant markers and null alleles. Molecular Ecology Notes.

Gaskin, J. F., D.-O. Zhang, and M.-C. Bon. 2005. Invasion of *Lepidum draba* (Brassicaceae) in the western United States: distributions and origins of chloroplast DNA haplotypes. Molecular Ecology 14:2331-2341.

Genton, B. J., J. A. Shykoff, and T. Giraud. 2005. High genetic diversity in French invasive populations of common ragweed, *Ambrosia artemisiifolia*, as a result of multiple sources of introduction. Molecular Ecology 14:4275-4285.

Gibson, C. J., and E. Lennon. 1999. Historical census statistics on the foreign-born population of the United States: 1850-1990. U.S. Bureau of the Census. Population Division Working Paper No. 29.

Guinand, B., and S. Esteal. 1996. Multivariate patterns of genetic differentiation support complex colonization schemes in *Bufo marinus* populations. Evolution 50:944-951.

Hauser, L., T. R. Seamons, M. Dauer, K. A. Naish, and T. P. Quinn. 2006. An empirical verification of population assignment methods by marking and parentage data: hatchery and wild steelhead (*Oncorhynchus mykiss*) in Forks Creek, Washington, USA. Molecular Ecology 15:3157-3173.

Hewitt, G. M. 1996. Some genetic consequences of ice ages, and their role in divergence and speciation. Biological Journal of the Linnaean Society 58:247-276.

Hewitt, G. M. 2000. The genetic legacy of the Quaternary ice ages. Nature 405:907-913.

Holsinger, K. E., and P. O. Lewis. 2003. Hickory: a package for analysis of population genetic data v1.0.

Holsinger, K. E., P. O. Lewis, and D. K. Dey. 2002. A Bayesian approach to inferring population structure from dominant markers. Molecular Ecology 11:1157-1164.

Ingvarsson, P. K., and D. R. Taylor. 2002. Genealogical evidence for epidemics of selfish genes. Proceedings of the National Academy of Sciences of the United States of America 99:11265-11269.

Ironside, J. E., and D. A. Filatov. 2005. Extreme population structure and high interspecific divergence of the Silene Y chromosome. Genetics 171:705-713.

Jalas, J., and J. Suominen. 1986. Atlas Florae Europaeae. Cambridge University Press, Cambridge.

Knowles, L. L. 2004. The burgeoning field of statistical phylogeography. Journal of Evolutionary Biology 17:1-10.

Kolar, C. S., and D. M. Lodge. 2001. Progress in invasion biology: predicting invaders. Trends in Ecology and Evolution 16:199-204.

Kolbe, J. J., R. E. Glor, L. R. Schettino, A. C. Lara, A. Larson, and J. B. Losos. 2004. Genetic variation increases during biological invasion by a Cuban lizard. Nature 431:177-181.

Larson, S. R., T. A. Jones, and K. B. Jensen. 2004. Population structure in *Pseudoroegneria spicata* (Poaceae: Triticeae) modeled by Bayesian clustering of AFLP genotypes. American Journal of Botany 91:1789-1801.

Lavergne, S., and J. Molofsky. 2007. Increased genetic variation and evolutionary potential drive the success of an invasive grass. PNAS 104:3883-3888.

Lee, C. E. 2002. Evolutionary genetics of invasive species. Trends in Ecology and Evolution 17:386-391.

Lewontin, R. C. 1965. Selection for colonizing ability. *in* H. G. Baker and G. L. Stebbins, editors. The genetics of colonizing species. Academic Press, New York.

Manel, S., O. Gaggiotti, and R. S. Waples. 2005. Assignment methods: matching biological questions with appropriate techniques. Trends in Ecology and Evolution 20:136-142.

Marsden-Jones, E. M., and W. B. Turrill. 1957. The bladder campions. Ray Society, London.

Martindale, I. C. 1876. The introduction of foreign plants. Botanical Gazette 2:55-58.

Martindale, I. C. 1877. More about ballast plants. Botanical Gazette 2:127-128.

McCauley, D. E., and M. S. Olson. 2003. Associations among cytoplasmic molecular markers, gender, and components of fitness in *Silene vulgaris*, a gynodioecious plant. Molecular Ecology 12:777-787.

McCauley, D. E., R. A. Smith, J. D. Lisenby, and C. Hsieh. 2003. The hierarchical spatial distribution of chloroplast DNA polymorphism across the introduced range of *Silene vulgaris*. Molecular Ecology 12:3227-3235.

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

Ostrowski, M.-F., J. Davis, S. Santoni, H. Mckhann, X. Reboud, V. Le Corre, C. Camilleri, D. Brunel, D. Bouchez, B. Faure, and T. Bataillon. 2006. Evidence for a large-scale population structure among accessions of *Arabidopsis thaliana*: possible causes and consequences for the distribution of linkage disequilibrium. Molecular Ecology 15:1507-1517.

Peterson, A. T., and D. A. Vieglais. 2001. Predicting species invasions using ecologic niche modeling. BioScience 51:363-371.

Petit, R., I. Aguinagalde, J.-L. de Beaulieu, C. Bittkau, S. Brewer, R. Cheddadi, R. Ennos, S. Fineschi, D. Grivet, M. Lascoux, A. Mohanty, G. Muller-Starck, B. Demesure-Musch, A. Palme, J. P. Martin, S. Rendell, and G. G. Vendramin.

2003. Glacial refugia: hotspots but not melting pots of genetic diversity. Science 300:1563-1565.

Pritchard, J. K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. Genetics 155:945-959.

Pritchard, J. K., X. Wen, and D. Falush. 2007. Documentation for *structure* software: version 2.2.

Pursh, F. 1814. Flora Americae Septentrionalis. White, London.

Rannala, B., and J. L. Mountain. 1997. Detecting immigration using multilocus genotypes. Proceedings of the National Academy of Sciences of the United States of America 94:9197-9201.

Rosenberg, N. A., J. K. Pritchard, J. L. Weber, H. M. Cann, K. K. Kidd, L. A. Zhivotovsky, and M. W. Feldman. 2002. Genetic Structure of Human Populations. Science 298:2381-2385.

Runyeon, H., and H. C. Prentice. 1997. Genetic differentiation in the bladder campions, *Silene vulgaris* and *Silene uniflora* (Caryophyllaceae) in Sweden. Biological Journal of the Linnaean Society 61:559-584.

Sakai, A. K., F. W. Allendorf, J. S. Holt, D. M. Lodge, J. Molofsky, K. A. With, S. Baughman, R. J. Cabin, J. E. Cohen, N. C. Ellstrand, D. E. McCauley, P. O'Neil, I. M. Parker, J. N. Thompson, and S. G. Weller. 2001. The population biology of invasive species. Annual Review of Ecology and Systematics 32:305-332.

SAS Institute. 1999. SAS/STAT version 8. SAS Institute, Cary, North Carolina, U.S.A.

Taberlet, P., L. Fumagalli, A.-G. Wust-Saucy, and J.-F. Cosson. 1998. Comparative phylogeography and postglacial colonization routes in Europe. Molecular Ecology 7:453-464.

Taylor, D. R., and S. R. Keller. 2007. Historical range expansion determines the phylogenetic diversity introduced during contemporary species invasion. Evolution 61:334-345.

Tsitrone, A., M. Kirkpatrick, and D. A. Levin. 2003. A model for chloroplast capture. Evolution 57:1776-1782.

Tsutsui, N. D., A. V. Suarez, D. A. Holway, and T. J. Case. 2000. Reduced genetic variation and the success of an invasive species. Proceedings of the National Academy of Sciences of the United States of America 97:5948-5953.

Vos, P., R. Hogers, M. Bleeker, M. Reijans, T. Van de Lee, M. Hornes, A. Frijters, J. Pot, J. Peleman, M. Kuiper, and M. Zabeau. 1995. AFLP: a new technique for DNA fingerprinting. Nucleic Acids Research 23:4407-4414.

Waples, R. S., and O. Gaggiotti. 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.

Waser, P. M., and C. Strobeck. 1998. Genetic signatures of interpopulation dispersal. Trends in Ecology & Evolution 13:43-44.

Webb, T. I., and P. J. Bartlein. 1992. Global changes during the last 3 million years: climate controls and biotic responses. Annual Review of Ecology and Systematics 23:141-173.

Williams, D. A., W. A. Overholt, J. P. Cuda, and C. R. Hughes. 2005. Chloroplast and microsatellite DNA diversities reveal the introduction history of Brazilian peppertree (*Schinus terebinthifolius*) in Florida. Molecular Ecology 14:3643-3656.

Table 1. AFLP molecular diversity in the native European and introduced North American ranges.

	Europe	**North America**
Number of individuals	233	171
Number polymorphic loci	267	262
Mean (SD) pairwise restriction site differences	67.28 (29.11)	58.52 (25.39)
Mean (SD) gene diversity over loci	0.2500 (0.1206)	0.2192 (0.1053)

Table 2. Analysis of variance of the first two axes of a principal coordinates analysis on ALFP genotypes using Jaccard's distance.

	Source	d.f.	Type III SS	*F*	*P*
Principal Coordinate 1 (PC1)	Continent	1	0.0062	1.40	0.2366
	Deme	3	3.9514	298.04	<0.0001
	Continent*deme	3	0.0241	1.82	0.1428
	Error	390	1.7235		
Principal Coordinate 2 (PC2)	Continent	1	0.0048	1.28	0.2579
	Deme	3	3.8552	342.07	<0.0001
	Continent*deme	3	0.0065	0.57	0.6324
	Error	390			

Table 3. Summary of genetic structure analyses on AFLP and cpDNA datasets. Two hypotheses were tested: molecular variance among the demes identified by STRUCTURE, and among three physiographic regions in Europe (eastern, western, and southern Europe; see Methods). No regional structure was hypothesized for North American samples. Negative variance components are interpreted as $F_{ST} = 0$ (Excoffier et al. 2005).

	Europe	**North America**
AFLP: $\theta^B \pm$ SD (95% C.I.)		
Among demes	0.0791 ± 0.0073	0.0868 ± 0.0082
	(0.0652 - 0.0928)	(0.0707 - 0.1027)
Among geographic regions	0.0056 ± 0.0017	--
	(0.0031 - 0.0094)	
cpDNA: F_{ST}		
Among demes	0.0779^{**}	-0.0140^{ns}
Among geographic regions	0.0142^{ns}	--

^{ns} not significant; ** $P < 0.01$

Table 4. Nuclear-nuclear and cytonuclear linkage disequilibrium

	Europe	**North America**
Among AFLP loci:		
Number of individuals	233	171
Number of pairwise tests	35511	34191
Number observed significant	4542	4035
Number expected significant by chance	1775.55	1709.55
Ratio of significant : expected	2.56	2.36
$\chi^2_{1\ d.f.}$	4537.21	3329.73
P	<0.0001	<0.0001
Among cpDNA and AFLP loci:		
Number of individuals	153	71
Number of pairwise tests	267	250
Number observed significant	38	26
Number expected significant by chance	13.35	12.5
Ratio of significant : expected	2.85	2.08
$\chi^2_{1\ d.f.}$	47.91	15.35
P	<0.0001	<0.0001

Figure Legends

Figure 1. Genealogical relationships among chloroplast haplotypes showing the two markers used in this study to define clades. (A) Neighbor-joining tree (Jukes-Cantor distances) based on a concatenated sequence of 1799 bp among four non-coding cpDNA regions (see Taylor and Keller (2007) for more details). Three main clades are recognized within the tree. (B) 7 bp indel in the intergenic space between *trn*L-*trn*F used to define clade A. (C) *Psi*I RFLP used to define clade B. Clade C was defined by mutual exclusion.

Figure 2. Results of STRUCTURE simulations. Each plot shows the Ln(probability of the data|K) as an increasing function of the number of demes, K. (A) Model likelihood across the full range of initial K values. (B) Model likelihood in the neighborhood of highest increase. Three different parameter sets were evaluated: (*i*) an admixture model using genotypes from both continents, (*ii*) a no admixture model using genotypes from both continents, and (*iii*) an admixture model using just European genotypes. In (B), the likelihood scores for (*iii*) are plotted on the secondary *y*-axis to facilitate comparison with the other parameter sets, although this model had higher overall likelihood scores.

Figure 3. Deme assignments based on an admixture model of genotypes from both continents. The five demes are plotted by color (deme 1 = blue, deme 2.1 = magenta, deme 2.2 = yellow, deme 2.3 = red, deme 3 = green). Bars plot the assignment scores for each individual genotype, showing the proportion of total ancestry assigned to each deme. Pie diagrams show the proportion ancestry contained within each sampling site, with pie size proportional to the number of genotypes sampled.

Figure 4. Principal coordinates analysis of AFLP genotypes based on Jaccard's index. Shown are the first two PC axes, with deme membership from STRUCTURE assignments indicated by color. European genotypes are open symbols while North American genotypes are solid symbols.

Figure 5. Association between AFLP demes and cpDNA clades. Frequencies of cpDNA clades within demes are plotted for Europe and North America. Deme 1 is omitted because of low sample size in North America (N = 4).

Figure S1. Results of STRUCTURE simulations searching for nested structure within the three original demes. Each plot shows the Ln(probability of the data|K) as an increasing function of the number of nested demes, K. All simulations used an admixture model that included genotypes from both continents.

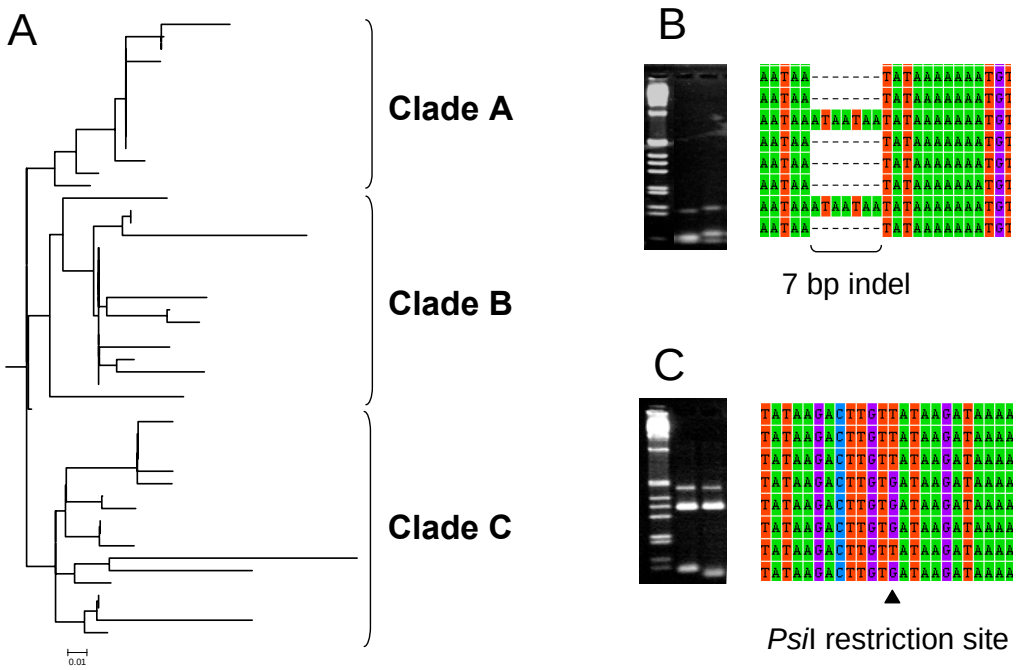

A
Clade A
Clade B
Clade C
0.01
B
7 bp indel
C
PsiI restriction site

Figure 1

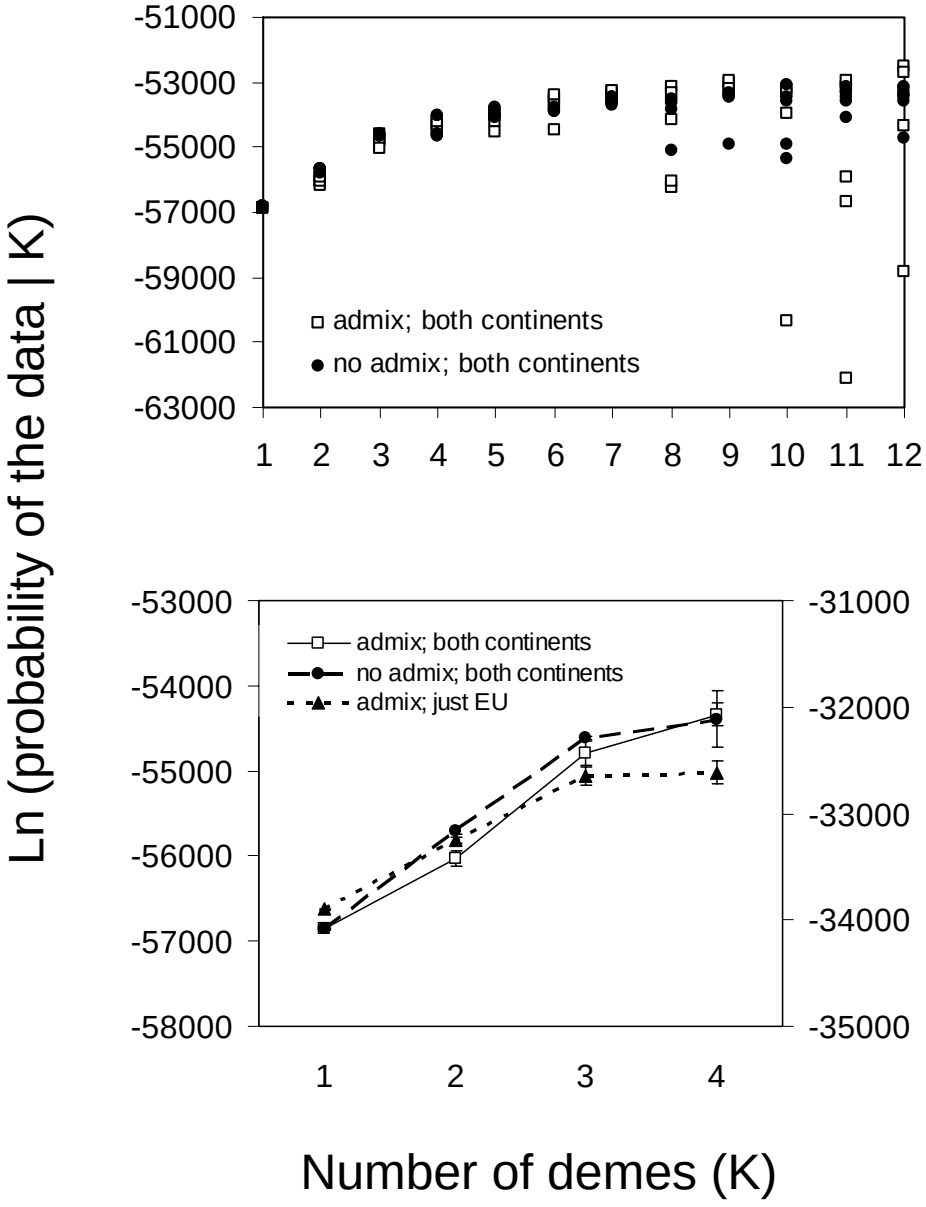

-51000
-53000
-55000
-57000
-59000
-61000
-63000
1 2 3 4 5 6 7 8 9 10 11 12
admix; both continents
no admix; both continents
-53000
-54000
-55000
-56000
-57000
-58000
-31000
-32000
-33000
-34000
-35000
1 2 3 4
admix; both continents
no admix; both continents
admix; just EU
Ln (probability of the data | K)
Number of demes (K)

Figure 2

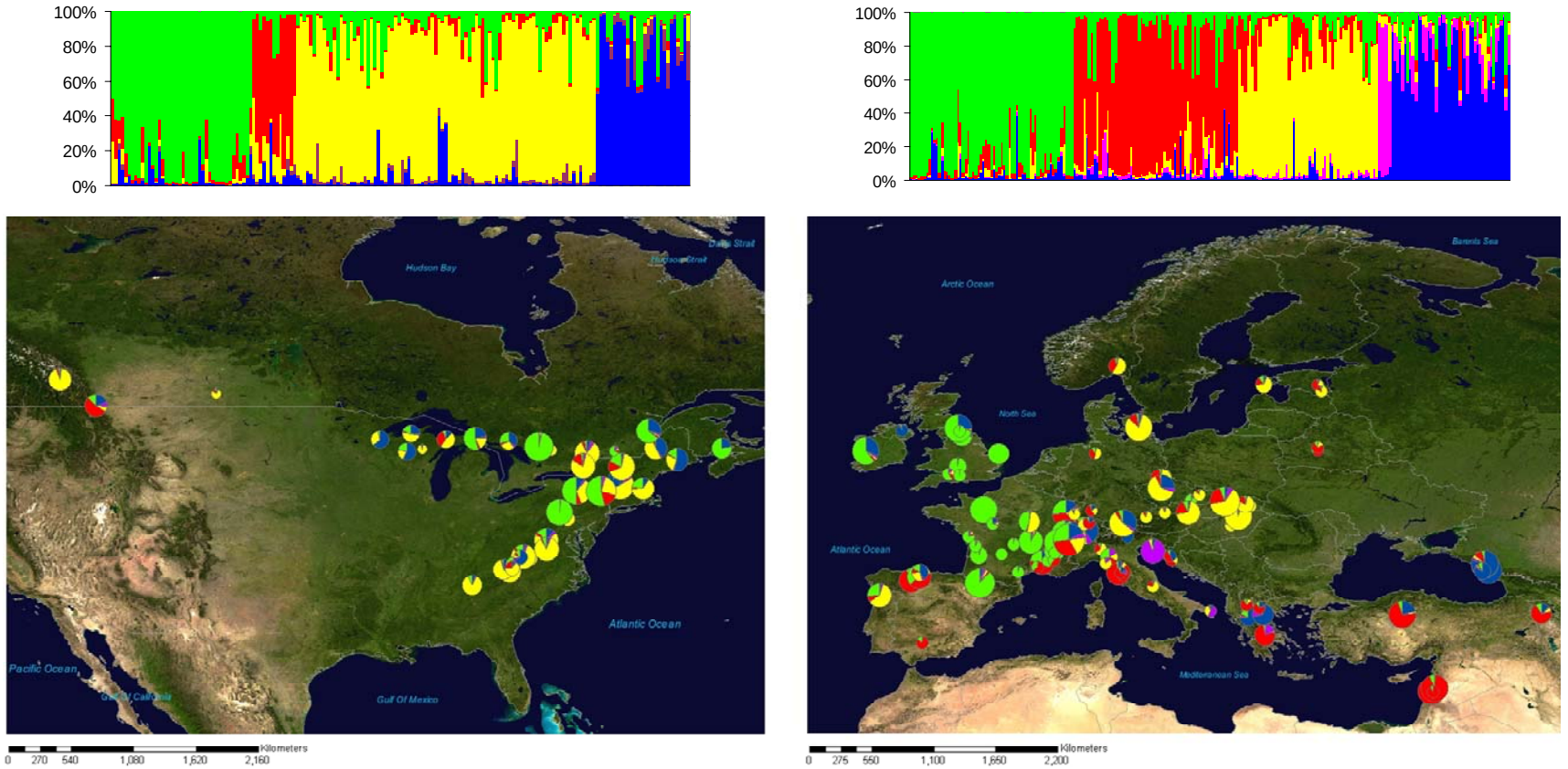

100%
80%
60%
40%
20%
0%
Hudson Bay
Atlantic Ocean
Pacific Ocean
Gulf Of Mexico
Kilometers
0 270 540 1,080 1,620 2,160
100%
80%
60%
40%
20%
0%
Arctic Ocean
Barents Sea
North Sea
Atlantic Ocean
Mediterranean Sea
Kilometers
0 275 550 1,100 1,650 2,200

Figure 3

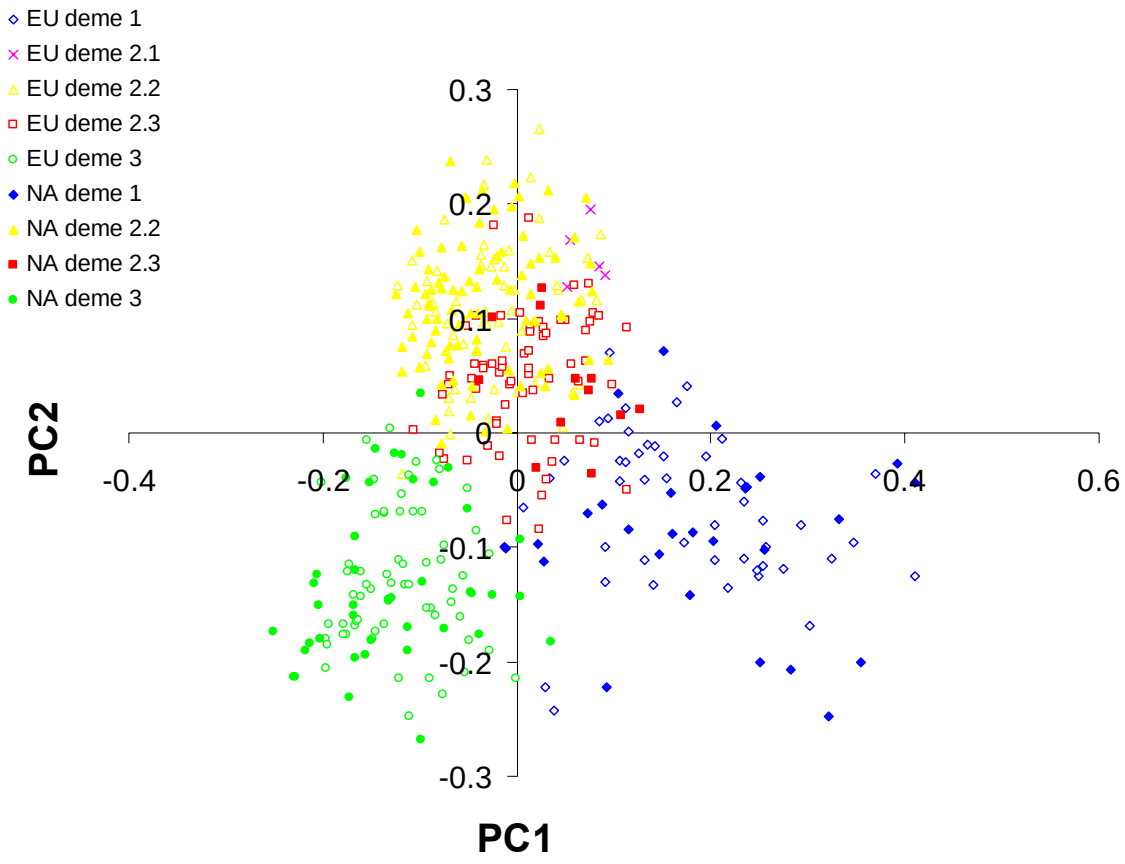

EU deme 1
EU deme 2.1
EU deme 2.2
EU deme 2.3
EU deme 3
NA deme 1
NA deme 2.2
NA deme 2.3
NA deme 3
0.3
0.2
0.1
0
-0.1
-0.2
-0.3
-0.4
-0.2
0
0.2
0.4
0.6
PC2
PC1

Figure 4

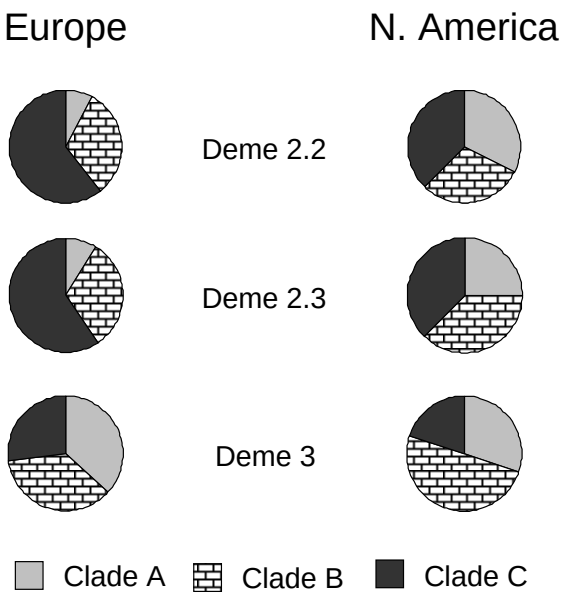


Figure 5

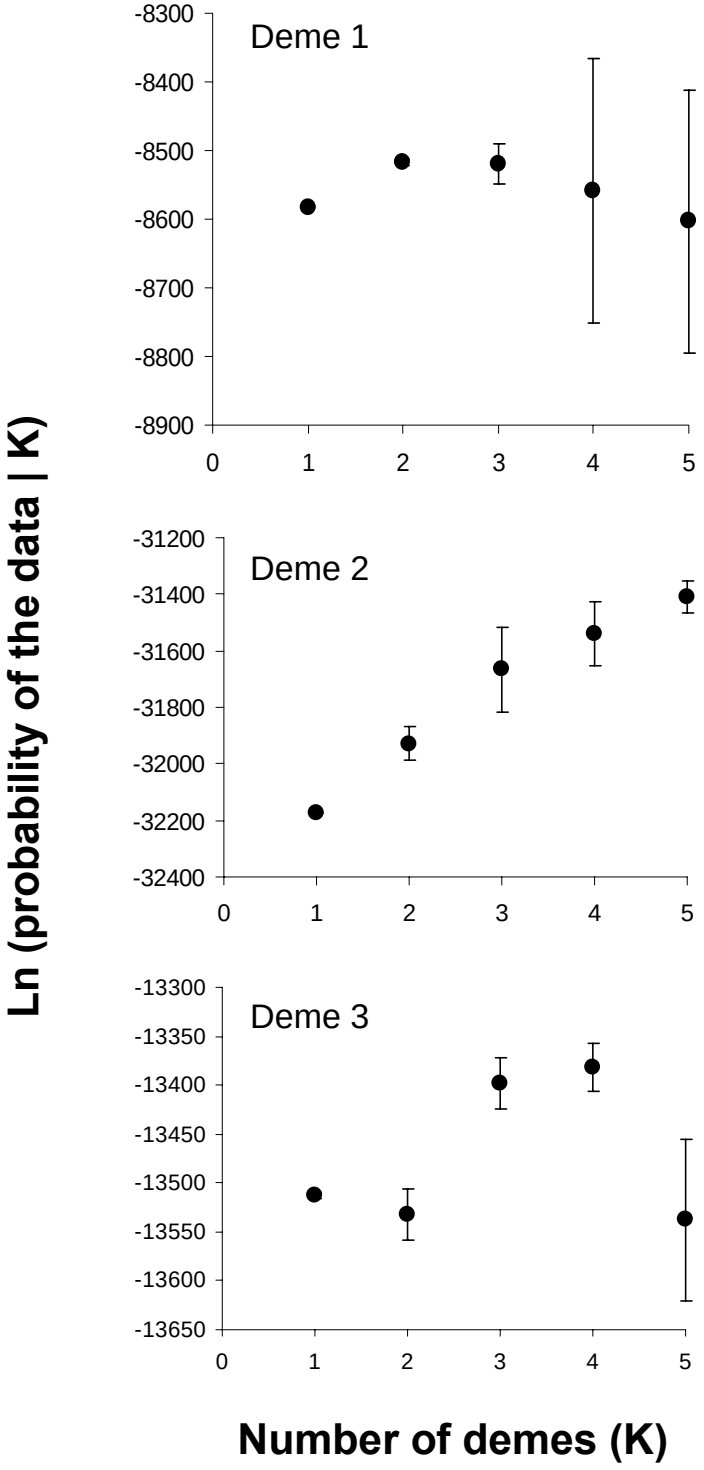

Deme 1
-8300
-8400
-8500
-8600
-8700
-8800
-8900
0
1
2
3
4
5
Deme 2
-31200
-31400
-31600
-31800
-32000
-32200
-32400
Deme 3
-13300
-13350
-13400
-13450
-13500
-13550
-13600
-13650
Ln (probability of the data | K)
Number of demes (K)

Figure S1

Appendix A. Sample locality information. Some latitude (degrees North) and longitude (degrees West) coordinates are approximate, based on locality description.

Continent	Country	Site	Location	Date	Collectors[†]	Latitude	Longitude
N. America	Canada	ALGP	Algonquin Park, Ontario. Past Lake of Two Rivers. Along hwy 60, just west of km 38 marker. Alt. 396 m.	10/9/2005	SK	45.5702	-78.4333
Europe	France	ALN	Alencon. Just S of Alencon, in suburban area next to a park field alongside of rd (N138). Alt. 149 m.	7/18/2004	SK, JK	48.4140	0.0893
Europe	Turkey	ANK	Ankara. From the Anakara area. A4 Ankara, Kalecik, Kecideresi. Serpantin. Alt. 800 m.	6/7/1999	MC	40.0796	33.4622
N. America	U.S.A.	ARCK	Armstrong Creek, WI. Approx. 3 miles W of town, along Hwy 8, at intersection with Janczewski Rd. Alt. 455 m.	10/6/2005	SK	45.6376	-88.4873
Europe	Germany	BAD	Bad Schandau. Rte. 62 heading N towards Dresdin, just N of the German/Czech border on E side of rd. Alt. 250 m.	7/29/2004	CB, SK, DS	50.9044	14.2167
N. America	U.S.A.	BAKRV	Bakersville, NC. On NC 261, 7.3 miles N of Bakersville. Alt. 1099 m.	6/9/2005	SK, VP	36.0743	-82.1021
N. America	U.S.A.	BDWY	Broadway, VA. W of I-81 on Woodlands Church Rd., just after crossing over Tide Springs Branch.	8/8/1999	SK	38.5752	-78.8472

Europe	U.K.	BKNP	Belas Knap, Gloucestershire. Plants growing on ancient burial mound, and in the immediate vicinity. Alt 302 m.	8/6/2005	DT, BP, MN, DS	51.9276	-1.9711
N. America	Canada	BLCK	Blackiston Creek, Alberta. About 1 mi. from Waterton. Off road.	7/1/2004	LW	49.0558	-113.9114
N. America	U.S.A.	BNMNT	Just S of Roanoke, VA. Along rte. 612 (off rte. 221), near underpass with Blue Ridge Parkway.	6/12/2005	SK, VP	37.1614	-80.1014
N. America	U.S.A.	BNTN	Bennington, VT. Population on both sides of rte. 9 heading W from the town of Bennington.	8/20/1999	SK	42.8774	-73.1031
Europe	France	BOL	Bollene.	7/22/2004	CB, DS	44.2753	4.7744
Europe	Germany	BRE	Bremen. 10 km SE from Bremen, along border of RR.	6/18/2004	HK	52.9816	9.0079
Europe	Czech Republic	BRN	Brno. Empty corner lot outside the train station, across the street, 200 m to the left.	7/8/1999	DT	49.2030	16.6162
N. America	U.S.A.	BRPW	Blue Ridge Parkway, NC. Near the town of Blowing Rock. MP 294 off Parkway; Cove family cemetery.	9/11/2004	WF	36.1553	-81.6860
N. America	U.S.A.	BURK	Burkes Garden, VA. Along rte. 623 on the way up the mountain. Alt. 959 m.	6/8/2005	SK, VP	37.0775	-81.3007
N. America	U.S.A.	CARL	Carlisle, PA. Roadside shaley bank at the Carlisle interchange of the PA turnpike. Alt. 148 m.	10/15/2005	SK	40.2296	-77.1530

Europe	U.K.	CAT	Catterick, England. N of town, on the E side of A6136, parallel to the A1. Adjacent to abandoned field. Alt. 66 m.	7/15/2004	SK, JK	54.3810	-1.6393
Europe	Hungary	CEG	Cegled.	7/3/1999	DT	47.1703	19.7921
Europe	Slovakia	CER	Cerveny'brek, Kosice, eastern Slovakia	9/26/2004	HS	48.7476	21.3350
Europe	France	CHA	Chardonne. Exit off the N9 W toward Lausanne. Turn right at first traffic circle just after leftward bend.	6/19/1999	DT	46.4767	6.8271
N. America	U.S.A.	CHEL	Chelmsford, MA. Plants in weed patch in middle of industrial area, located off Plain, Payton Streets. Alt. 62m.	8/15/2004	SK	42.6234	-71.3169
Europe	Ireland	COM	Cliffs of Moher. W coast, outside of Doolin. Plants at cliffs' edge, near pedestrian trail. *Silene uniflora.*	7/8/2004	SK, JK	52.9747	-9.4262
Europe	France	CRE	Creon -1	7/20/2004	CB, DS	44.8336	-0.2817
N. America	U.S.A.	CRMN	Carter's Mountain, VA. Apple orchard, just SE of Charlottesville. Off rte 53, just W of Michie Tavern.	11/15/1999	SK	37.9948	-78.4695
Europe	U.K.	CRO	Cromer, Norfolk. Near parking area on A149 on east end of town. Plants along eroding cliff face. Alt. 41 m.	7/16/2004	SK, JK	52.9337	1.2901
Europe	Austria	DAN	Danube, Vienna. 100 m N of Donanuinsel subway station (U1 line); 50 m N of the overpass.	6/30/1999	DT	48.2092	16.3727
Europe	Denmark	DK3	Isle of Mons. 0.5-1 km E of Borre on the way to Mons	8/18/2001	DT	54.9956	12.4473

			Kilt. 50m before marker #12/7.				
Europe	Ireland	DOB	Doolin Beach. On the W coast, near the town of Doolin. Plants growing on beach. *Silene uniflora.*	7/9/2004	SK, JK	53.0774	-9.3437
Europe	Greece	DPI	Delphi, Greece.	?/?/2002	SR	38.4697	22.5026
Europe	Hungary	EGR	Eger. Plants around railway station, at the termination of the tracks by old rail cars W of the station.	7/4/1999	DT	47.8951	20.3827
Europe	Lebanon	ELB	el Barouk. 2 km SW of El Barouk in Southern Lebanon. Alt. 1400 m.	6/6/1999	DW	33.7000	35.7000
Europe	Switzerland	EMP	Baden-Wuttemberg, Empfingen	8/22/2004	Unknown	48.3920	8.7083
Europe	France	EPI	Epinal SE of Docelles on D11c, past Xamontaru Pt on Chemin de Fossard. Near the recycling area.	7/21/2004	SK, JK	48.1187	6.6680
Europe	Switzerland	ERL	Central Switzerland, between Chateau D'oex, Interlaken.	7/23/2004	CB, SK, JK, DS	46.6619	7.5680
N. America	Canada	ESPA	Espanola, Ontario. On hwy 17, near intersection with hwy 6. Plants around Esso station. Alt. 205 m.	10/4/2005	SK	46.2865	-81.7751
Europe	Lebanon	FAK	Fakra (Faqura). 2 km SW of Ferraya (Faraya) in Central Lebanon. Alt. 1450 m.	6/17/1999	DW	34.0000	35.7833
Europe	Ireland	FAN	Fanore. Hwy R477, just S of Fanore along roadside.	7/9/2004	SK, JK	53.1160	-9.2845

N. America	U.S.A.	FS77	Tennessee. US Forest Service Rd. 77; off hwy 64. E side of rd, ca. 0.9 miles from intersection of FS77 and hwy 64.	6/10/2005	SK, VP	35.1015	-84.6092
Europe	Switzerland	GAS	Gasthaus Kulm Between Chateau D'Oex and Interlaken. Plants growing in a lush forb community. Alt. 1050 m.	7/23/2004	SK, JK	46.7582	8.1351
Europe	Slovakia	GEL	Gelnica, Cecly, eastern Slovakia	8/8/2004	HS	48.8550	20.9466
Europe	France	GOG	On D94 west. Before intersection of D94/D64 approaching la Bonte.	6/21/1999	DT	44.3888	5.2328
Europe	France	GRE	Grenoble. 2km N of Grenoble. Foothills, open area near a *Quercus pubescens* forest. Alt. 300 m.	9/9/1999	PC	45.1943	5.7316
N. America	U.S.A.	HAMPV	Hampton, TN. Abandoned lot off First Ave., just after bridge over Laurel Fork of the Roan River. Alt. 550 m.	6/9/2005	SK, VP	36.2883	-82.1717
N. America	U.S.A.	HDFL	Hudson Falls, NY. W of rte. 4 in town of Hudson Falls. Parking lot, approx. 0.5 km from Hudson River.	9/1/2004	DO	43.2845	-73.5878
N. America	U.S.A.	HNKS	Hanksville, VT. Unnamed backroad between VT 17, VT 116. Near Warren/Waitsfield area.	8/20/2003	SK	44.3075	-72.9845
Europe	Croatia	IGA	Island Galopun, UL90, rocks. Alt. 2 m.	6/18/1999	MV	45.1500	13.6000
N. America	U.S.A.	ITHA	Ithaca, NY. Populations reported as MDSW (Meadowsweet Farm) in McCauley et al. 2003		DM	42.4405	-76.4957

Europe	Croatia	JZV	Jablanac, Zavratnica, VK 95, path edge, Mainland across from Island Rab near Pag, Alt. 10 m.	7/21/1999	MV	44.7098	14.8953
Europe	Lebanon	KDS	Kadischa. 0.5km SW of Bcharre in Kadischa valley, north Lebanon. Alt. approx. 1100 m.	6/15/1999	DW	34.2333	35.9834
Europe	Russia	KRA	Krasnaya Polyana, Russia. Alt. 2201 m.	7/9/2004	HF	43.6319	40.2884
Europe	Spain	LAR	La Robla	7/13/2004	CB, DS	42.7986	-5.6903
Europe	Switzerland	LAU	Lausanne. In a grassy field at the park along the Lac Leman shoreline, near the University of Lausanne.	7/22/2004	SK, JK	46.5166	6.5952
Europe	Russia	LOG6	Logonaky Mountain, Russia. Alt. 1761 m.	7/13/2004	HF	44.0762	40.0125
Europe	Spain	LTR	La Trassiera. 15 km from Cordoba.	1/1/1992		37.8863	-4.7769
N. America	U.S.A.	ME1	Maine. Rte. 9, ca. 5 km W of Amherst, ca. 24 km ESE of Orono. Alt 134 m.	7/13/2005	DS	44.8288	-68.4340
N. America	U.S.A.	MEBDR	Maine border, ME. Rte. 201, S of Quebec border but before Jackman, ME. Roadside cutbank. Alt. 570 m	8/6/2004	SK	45.7500	-70.3076
Europe	Greece	MET	Near Metsovo. Alt. 1639 m.	8/2/2005	HF	39.8475	21.1746
Europe	Russia	MEZ5	Near Mezmay, Russia. Alt. 1199 m.	7/13/2004	HF	44.1667	40.0143
Europe	Italy	MH	Lamole, Italy.	8/1/1998	MH	43.4000	10.8500
Europe	Belarus	MIN	Minsk province, Nesvizh region, near RR station	9/12/1999	MD	53.2189	26.6825

			Pogoreltsy; 200 m to the Baranovitchi station. Alt. 200 m.				
N. America	Canada	MLST	Milestone, Saskatchewan. Along hwy 70.	7/1/2004	LW	49.9909	-104.5130
N. America	Canada	MONT	Montreal	7/1/2000	DM	45.5088	-73.5541
N. America	U.S.A.	MTLK	Near Mountain Lake, VA. Rte. 730, W of Eggleston, before intersection with rte. 621. Alt. 628 m.	8/28/2004	SK, CB	37.2815	-80.6395
Europe	Greece	MTOP	Mt Olympus, Greece. One site near refuge at alt. 1908 m, another on W side of mountain at alt. 982 m.	7/21/2005	HF	40.0726	22.4619
N. America	U.S.A.	MTROG	Near Mt. Rogers, VA. On rte 16, near mailbox #9238. Alt. 920 m.	6/8/2005	SK, VP	36.6853	-81.4263
N. America	U.S.A.	MUNI	Munising, MI. Along abandoned RR parallel to rte. 28. Near veteran's memorial. Alt. 195 m.	10/5/2005	SK	46.4116	-86.6514
Europe	Greece	MVRM	Mt. Vermio, near Edessa. Alt. 817 m.	8/8/2005	HF	40.6354	22.0307
N. America	Canada	NB1	New Brunswick. Rte. 114, NW of Fundy National Park, ca. 25 km NW of Alma. Alt 319 m.	7/13/2005	DS	45.7224	-65.1911
Europe	Lebanon	NBS	Naba Sannine. 4 km E of Beskinta. Alt. 1700 m.	6/24/1999	DW	33.9334	35.8667
Europe	Czech Republic	OLO	Olomouc. Train station, across from the platforms, eastward ~100m. Weed patches among the shacks.	7/8/1999	DT	49.5972	17.2621
N. America	U.S.A.	ONTO	Ontonagon, MI. 4 mi. W of Ontonagon, opposite the	10/5/2005	SK	46.8543	-89.3782

			Sunshine Motel. West of residence #23940. Alt. 165 m.				
Europe	Italy	OP	Overpass on SS357 at intersection with SS9 toward Milan. Signs to Fontana Letto, Piacenza (to the S).	6/18/1999	DT	45.3268	9.4069
Europe	Russia	ORL	Orlynye Skaly, Russia. Eagle's rocks, near Sochi. Alt. 424 m.	7/10/2004	HF	43.5597	39.8228
N. America	U.S.A.	ORNO	Orono, ME. Main St., just N of bridge over Penebscot River, across from the University Inn. Alt. 27 m.	8/8/2004	SK	44.8870	-68.6692
Europe	Spain	OSE	Oseja de Sajambre.	7/14/2004	CB, DS	43.1290	-5.0267
Europe	Norway	OSL	Oslo, Norway		Unknown	59.9138	10.7388
N. America	Canada	OTTA	Ottawa. Uplands Drive, near intersection with Hunt Club Rd., near airport. Alt. 99 m.	8/4/2004	SK	45.3430	-75.6659
Europe	Croatia	PAG	Island Pag, Drazice, VK84, by fence. Alt. 5 m.	6/22/1999	MV	44.4412	15.0534
Europe	Italy	PAN	Panchia, Trentino region, Italy. 1000 m.	Oct-99	FP	46.2867	11.5436
Europe	France	PCH	Pont Charron. 85 La tabariere, Vendee, Couteaux sec.	unknown	NBG	46.2936	-1.1057
Europe	Spain	PCO	Puerto de Cotefablo. Just E of tunnel on rte. 260.	7/16/2004	CB, DS	42.6223	-0.1938
Europe	France	PIG	le Pigeon.Parking area/pull off W of le Pigeon on D703. On N side of D 703, 0.7 km east of km post #8.	6/23/1999	DT	44.9042	1.5187
Europe	France	PIS	le Pisseux, 11.9 km S of Annecy. On D31 W of N201.	6/20/1999	DT	45.8313	6.0224

Europe	France	PLB	Ponte la Barque. Intersection of D994, N75 between towns of Veynes, Serres, Parking area at W side of intersection.	6/21/1999	DT	44.4495	5.7235
Europe	Spain	PLGUA	Playa de Guadamia. On the N coast, near Ribadesella.	7/15/2004	CB, DS	43.4607	-4.9757
Europe	France	PON	Pons. Small road W of Pons, off D732 @ sign to dump (Dechetterie). 0.9 km to site, just before dump entrance.	6/24/1999	DT	45.5728	-0.5748
Europe	Czech Republic	PRG	Prague. At the Prirodin Pamatka natural monument, in the Praha 6 region. Alt. 274 m.	7/28/2004	CB, SK, DS, HS	50.1857	14.2520
Europe	Greece	PRL	Prespa Lakes, Psarades. Alt. 907 m.	8/5/2005	HF	40.8306	21.0232
Europe	France	PSN	Pont St. Nicholas. Ca. 14 km S of Uzes on the road to Nimes. D979 S of bridge, near intersection with D135.	6/21/1999	DT	43.9169	4.3877
N. America	Canada	QUEB	Quebec City, Quebec. Near Mt. Ste. Anne. On rte. 360E, near intersection with Rue de la Ferreolaise. Alt. 155 m.	8/6/2004	SK	47.0808	-70.8947
N. America	Canada	RAHS	Radium Hot Springs, British Columbia. Ca. 70 km N of town on hwy 95. On shoulder of highway. 2395-2405.	7/29/2004	LW	51.0939	-116.6868
Europe	France	ROD	Rodomouls. 9 km SE of St. Pons on N112. SE of intersection with D176 near Rodomouls.	6/22/1999	DT	43.4725	2.8492
Europe	France	RON	Route de Napoleon. N85 S toward La Mure. Top of mountain, pulloff 3.3 km from tourist information sign.	6/20/1999	DT	45.0100	5.7725

Europe	Italy	ROS	Rosia, Tuscany. .0.8 km on SS 223 off of SS73.	6/16/1999	DT	43.2499	11.2234
Europe	Italy	RS	Rest stop on Autostrade from La Spezia toward Milan. Ca. 17 km W of Fornovo exit. On eastbound side.	6/18/1999	DT	44.6930	10.1020
Europe	Slovakia	RUZ	Ruzomberok. Across tracks from train station. Road runs parallel to the tracks. Plants E and W of the station.	7/10/1999	DT	49.0788	19.3030
Europe	Estonia	SAA	Saaremaa Island. Sorve Peninsula, Estonia. Fallow Land.	9/2/1999	VK	58.4287	22.4056
Europe	Switzerland	SCN	Les Sciernes. N of Bulle on rte. 8 towards Chateau D'Oex. Along side road in residential area. Alt. 761 m.	7/23/2004	CB, SK, JK, DS	46.5161	7.0541
Europe	Italy	SCT	Santa Cesarea. Term-Lecce. Plants found on rocky cliffs. *S. vulgaris* subsp. *angustifolia*	5/6/1999	SM	40.3533	18.1740
Europe	Austria	SEE	Seefeld. NW of Seefeld towards Leutasch on Olympia Strasse. In field by abandoned hotel. Alt. 1205 m.	7/25/2004	CB, SK, DS	47.3301	11.1821
Europe	U.K.	SFD	Straffordstown, N. Ireland. On Ballynamillan Rd. close to N shore of Logh Neagh.	7/11/2004	SK, JK	54.7204	-6.3871
Europe	Italy	SG	San Gimignano, Tuscany. Via G. Matteotti. In parking lot across from Hotel de Graziano.	6/16/1999	DT	43.4049	11.0198
Europe	Austria	SLZ	Salzburg. Old lot just SW of train station.	7/1/1999	DT	47.8005	13.0444
Europe	France	SOS	SOS phone 68031 on N83. Along W side of road. 2.7 km	6/28/1999	DT	48.0584	7.3086

			S of intersection between N83 and D417.				
N. America	U.S.A.	SRT	Sugar Run Trail. SE VA, top of Sugar Run rd. up to look out tower.	7/28/1999	DT	37.2476	-80.8540
N. America	Canada	SRVR	South River, Ontario. Along road to the town dump.	8/14/2003	SK	45.8605	-79.4212
N. America	U.S.A.	SSPG	Solon Springs, WI. Near town, at intersection of county roads P and A. Alt. 328 m.	10/6/2006	SK	46.3848	-91.7788
N. America	U.S.A.	SSTM	Sault Ste Marie, MI. In town; along RR tracks under road bridge to Canada; just W of Soo Locks. Alt. 169 m.	10/5/2005	SK	46.4998	-84.3645
N. America	U.S.A.	STCL	State College, PA. Along hedgerows bordering private property off Bailey Lane; outside of Boalsburg, PA.	8/10/2003	SK, DO	40.7697	-77.8026
N. America	U.S.A.	STFD	Stamford, NY. Along old RR tracks converted to a multiuse path (Rails to Trails) at E side of Stamford.	8/21/2003	SK	42.4113	-74.6042
Europe	U.K.	STH	Stonehenge, Wiltshire. Ca. 250 m from Stonehenge, along fenceline adjacent to crop fields, S of A360. Alt 98 m.	8/6/2005	DT, BP, MN, DS	51.1793	-1.8306
Europe	France	STJ	St. Just	7/18/2004	CB, DS	45.9082	3.8514
Europe	U.K.	STK	Stocksfield. Ca. 1 mi. S of Stocksfield, England on B6309. Alt. 64 m.	7/14/2004	SK, JK	54.9373	-1.9181
Europe	Belarus	STO	Minsk province, Stolbtzy region, Okyntchitzy forest parcel	7/24/1999	MD	53.4107	26.7249

			#66. Pine forest, along forest road. Alt. ca. 160 m.				
Europe	Italy	SVAN	*S. vulgaris* var. *angustifolium*. Alt. 1123 m.	11/2/2004	JA, MH	42.3508	13.5726
Europe	Austria	SVL	St. Valentin. On main rd, just S of A1, next to RR tracks and on roadside opposite. Alt. 283 m.	7/26/2004	CB, SK, DS	48.1869	14.5181
Europe	Switzerland	TGN	Triengen. Ca. 9km S of Reiden on rte 2. Intersection with road to Triengen. Plants along road just N of intersection.	6/28/1999	DT	47.2339	8.0769
N. America	Canada	THIS	Thousand Islands, Ontario. Roadside after Canadian customs; accessed from U.S. I-81, near Watertown, NY.	8/11/2003	SK	44.3673	-75.9815
Europe	Italy	TI	Tivegna, Tuscany. 1 km after turn off to Tivegna, on the way to Follo.	6/17/1999	DT	44.1818	9.8392
N. America	U.S.A.	TOMA	Tomahawk, WI. On CR A, just E of intersection with hwy 51, ca. 0.1 km W Edgewater Country Club. Alt. 452 m.	10/6/2005	SK	45.4965	-89.6835
Europe	France	TRU	Truyes. .N of Truyes, S of les Granlineries. D45 at PR1 mile post.	6/25/1999 ; 7/19/04	DT (1999); SK (2004)	47.2798	0.8390
Europe	Estonia	TTU	Tartu.	9/27/1999	VK	58.3612	26.7159
Europe	France	VAL	Vallorcine. In Alps near Swiss border. Alt. 4413 m.	7/23/2004	CB, DS	46.0235	6.9178
Europe	Portugal	VDM	Vieira Do Minho. Ca. 2.3 km E of rte. 103, towards town of Vieira Do Minho. Alt. 1721 m	7/12/2004	CB, DS	41.6414	-8.1670

Europe	France	VEY	Veyrieres. Alt. 2360 m	7/19/2004	CB, DS	45.6552	2.5107
Europe	France	VEZ	Vezelay, Burgandy. Grassy banks of visitor parking lot off D951 on the W side of town. Alt. 245 m.	7/20/2004	SK, JK	47.4641	3.7403
Europe	Estonia	VVI	Voru, Viitina. South Estaonia, south from Voru, Viitina	9/17/1999	VK	57.8442	27.0137
Europe	Croatia	VZA	Velebit, Zavizan, VK95, path edge, Alt. 1400 m.	7/31/1999	MV	44.8105	14.9823
N. America	Canada	WGPP	Waterton-Glacier, Alberta. Hwy 5 leaving Waterton Park, E of bridge and just S of Red Rock Rd. In dry river wash.	7/4/2004	LW	49.0774	-113.8812
N. America	U.S.A.	WJEFF	West Jefferson, NC. Off rte. 221, in waste place at edge of parking lot for the Visitor's Center. Alt. 910 m.	6/8/2005	SK, VP	36.4044	-81.4912
Europe	U.K.	WLW	Witton le Wear. On E side of A68, just S of town; across from a few houses and a gas station. Alt. 159 m.	7/14/2004	SK, JK	54.6792	-1.7813
Europe	Russia	WSIB	Western Siberia. Vicinity of Krasnojarsk. Surroundings of the city. Natural reserve Sbelt.	7/1/2002	HS	56.2000	92.5000
N. America	U.S.A.	WTGN	Wintergreen, VA. Adial Rd off VA6 in Nelson Co. Between rtes. 29 and 151. In field by transfer station.	6/23/2004	SK	37.8703	-78.8226
Europe	Armenia	YER	Yerevan. Territory of the botanical garden of the Institute of Botany of the Armenian National Academy of Sciences.	6/2/1999	AN	40.1596	44.5090
Europe	Switzerland	ZUR	Zurich. W side of the mouth of the river in Zurich, Ca. 200	6/29/1999	DT	47.3691	8.5380

m from the mouth of the river.

† Collectors: A. Nersesyan (AN); B. Penna (BP); C. Barr (CB); D. McCauley (DM); D. Ombalski (DO); D. Sowell (DS); D. Taylor (DT); D. West (DW); F. Piccoli (FP); H. Frierson (HF); H. Kuhbier (HK); H. Storchova (HS); J. Antonovics (JA); J. Keller (JK); L. Wolfe (LW); M. Dzhus (MD); M. Hood (MH); M. Neiman (MN); M. Vrbek (MVr); M. Vural (MVu); Nantes Botanical Garden Staff (NBG); P. Choler (PC); S. Keller (SK); S. Marchiori (SM); S. Ribstein (SR); V. Kuusk (VK); V. Panjeti (VP); W. Farnum (WF).

Chapter 4:

Departures from niche conservatism during a biological invasion[4]

[4] Formatted for submission as a coauthored manuscript: Keller, S.R. and D.R. Taylor

Abstract

The presence of range limits within species, and niche similarity between closely related taxa suggest multiple constraints may operate to conserve ecological niches. Yet recent theory suggests niches can evolve under a broader range of demographic and genetic circumstances than previously thought. When a species colonizes and spreads throughout a new geographic range, one possibility is that the ecological niche is evolutionarily conserved and populations are simply invading previously unexploited habitats. Alternatively, the ecological niche may be more evolutionarily labile, and range expansion may reflect a true expansion of a species' ecological breadth. Invasive species represent a unique opportunity to test the hypothesis of niche conservatism during range expansion by generating niche predictions based on a species' native range and examining how occurrences in the introduced range match these predictions. In this study, we tested the hypothesis of niche conservatism during the invasion of North America by the weedy European plant, *Silene vulgaris*. Niche predictions based on the European distribution were compared with observed occurrences in North America generated from a broad survey of herbarium collections. We also adopted a genetical approach by modeling intraspecific niche variation among native range demes identified from a previous analysis of neutral genetic variation, and generated unique predictions for deme occurrence within the introduced range. The North American distribution of *S. vulgaris* showed significant evidence for niche conservatism (73% of occurrence points within the ancestral niche prediction), but also a large region of niche expansion. Expansion outside the ancestral niche primarily occurred in western North America, and this portion of the species distribution has increased in relative frequency over time (from

13% before 1950 to 30% after 1950). Four geographically distinct demes within Europe showed differences in their niche predictions, as evidenced by the size and spatial location of niche predictions in North America. Introduced genotypes were often mismatched from the niche predicted by their deme, probably as a result of a stochastic colonization process. The relative frequency of niche mismatching varied significantly among the demes. These results broaden our understanding of ecological niches in two ways. First, the apparent niche expansion in western North America suggests the paradigm of niche conservatism over short time periods may not always hold. This may be especially true during biological invasion, where constraints on the realized niche are often removed (e.g., loss of natural enemies) and unique genotypes may result from the admixture of separate introductions. Second, the presence of intraspecific niche variation coupled with a stochastic colonization process means that genotypes will often become mismatched from their niche, even if they occur within the species-wide prediction. This may reduce genotype fitness, resulting in newly established populations that have low intrinsic rates of increase and experience selection for new niche dimensions during invasion.

Keywords: ecological niche, invasive species, range expansion, *Silene vulgaris*, GARP, AFLP

Introduction

As to the mechanism of geographic limitation, the adjustments to the various critical factors are inevitably forever in process, though reduced to a minimum at times of slow environmental change. The refined method of individual "preference" or "choice" is superior to the wasteful process of wholesale destruction which would be experienced by individuals finding themselves out of place as the result of a haphazard selection of locality.

J. Grinnell (1917)

What determines a species' ability to colonize and spread throughout a new geographic range? Since the recognition that species distributions occur within definable environmental contexts (Grinnell 1917, 1924) and the conceptualization of the multidimensional ecological niche (Hutchinson 1957), biologists have sought correlations between suites of environmental variables and species occurrence. But the translation between environmental context and species occurrence may become complicated when species' borders are in flux; perturbed from equilibrium by the demographic effects of colonization, immigration, and extinction, and the genetic effects of adaptation and maladaptation within newly founded populations. Viewing the ecological niche as a property of a population, subject to variation, selection, and evolutionary change, is key to understanding the forces governing geographic range expansions and the establishment of their limits (Antonovics 1976, Holt and Gaines 1992, Holt and Keitt 2005).

The environmental conditions where a species has the capacity to establish and maintain populations (the fundamental niche), the conditions actually occupied given interactions with other species (the realized niche), and the spatial extent of occurrence within habitable space (the "realized" geographic range) will often yield inconsistencies, suggesting the operation of one or more constraints on distribution in ecological and

geographic space (Soberon and Peterson 2005). Interspecific interactions such as competition, predation, or disease may act deterministically to constrain distributions within the fundamental niche (Connell 1961, Hochberg and Ives 1999). Constraints on geographic ranges may also reflect historical or stochastic processes, such as dispersal limitation (Svenning and Skov 2007) or metapopulation processes (Holt and Keitt 2000). An evolutionary constraint is also suggested by the fact that species have borders at all, and in theory natural selection should promote adaptation in marginal populations that enable range limits to advance (Antonovics 1976b). While natural selection in marginal populations may be rendered ineffective by lack of heritable variation or the immigration of maladapted genes from demographically more abundant parts of the species' range (Holt and Gaines 1992, Kirkpatrick and Barton 1997), there are also many realistic conditions under which the ecological niche can readily evolve (Holt and Gomulkiewicz 1997, Holt 2003, Holt et al. 2003), and hence range limits should be expected to advance.

Niche conservatism is the tendency for evolutionary stasis in the ecological requirements of related taxa. Evidence for niche conservatism stems from biogeographical studies that show stasis among taxa for geographic range size (Ricklefs and Latham 1992) and position in ecological space (Huntley et al. 1989, Prinzing et al. 2001). Recently, the use of geographic information systems (G.I.S.), natural history collections, and computational advances in the modeling of ecological niches have catalyzed this field of study (Graham et al. 2004a, Parmesan et al. 2005), opening a host of new opportunities to test the hypothesis of niche conservatism and its implications for past, present, and future species distributions (Peterson et al. 2004, Peterson 2006, Thuiller et al. 2006). Interspecific comparisons of ecological niches have uncovered

further evidence of niche conservatism among closely related species (Peterson et al. 1999, reviewed in Weins and Graham 2005), but also examples of niche lability (Losos et al. 2003, Knouft et al. 2006). Intraspecific analyses are much less frequent (Peterson and Holt 2003), perhaps reflecting the view from niche conservatism that species are relatively immutable. At some point this becomes an issue of scale, as the occurrence of locally adapted populations in many species is well established, suggesting that in principle, intraspecific niche variation does exist. Therefore, a population genetic approach to variation within a species may prove fruitful to the study of niche evolution.

The introduction of species to new ranges provide natural experiments for observing the influences of historical contingency, chance events, biotic interactions, and local adaptation on the ecological niche and the resulting implications for geographic range limits (Hochberg and Ives 1999, Peterson and Vieglais 2001, Sakai et al. 2001, Peterson 2003, Holt et al. 2005a). One approach is to first develop a niche model based on occurrence points in the native range, and then predict where a species ought to occur within the introduced range. Such niche models conditioned on native range occurrences have been used to predict the distribution of species in new ranges and identifying areas susceptible to future invasion (Peterson and Vieglais 2001, Peterson 2003, Kriticos et al. 2005, Mau-Crimmins et al. 2006, Broennimann et al. 2007). The implication of these models is that ecological niches are conserved, and by extension, that evolutionary processes are not of paramount importance for understanding the distributions of invasive species (but see Broennimann et al. 2007). There are several reasons why this may not always hold. First, range-wide niche models assume species are genetically uniform in both ranges with respect to the niche (Mau-Crimmins et al. 2006). Yet many species in

their native ranges are subdivided into geographical regions composed of lineages with unique evolutionary histories (Gaskin and Schaal 2002, Kolbe et al. 2004, Taylor and Keller 2007, Chapter 3). Historical isolation among lineages may also involve divergence along one or more axes of the ancestral niche space. Thus, different lineages within a species may not only produce distinct predicted areas of occurrence, but may also be sampled from differently during the invasion process. Second, there is growing evidence that phenotypic evolution in response to environmental gradients within the introduced range may be common (Johnston and Selander 1964, Huey et al. 2000, Maron et al. 2004, Leger and Rice 2007). Thus a contrarian view to niche conservatism would instead suggest that the evolutionary forces operating during invasion will drive the ecological distribution of the species in the introduced range.

In this paper we use ecological niche modeling to generate predictions for the geographic range of the weedy plant, *Silene vulgaris*, introduced to North America from Europe about 200 years ago. These predictions are compared to its known area of occurrence in North America, as inferred from a survey of museum records. We then use North American occurrence points to model the niche in the introduced range and ask whether it is representative of the native range niche, or if instead the species' niche in North America has become bottlenecked or expanded relative to Europe. Finally, we develop a novel population genetic perspective by modeling the predicted and occupied range for each of several genetic demes, and ask if demes differ in the degree to which they are matched or mismatched from their range prediction.

The results show a substantial expansion of *S. vulgaris* in North America outside the area predicted by the ancestral European niche, possibly due to evolutionary change.

Additionally, the distribution of genotypes belonging to different intraspecific demes is often mismatched from the corresponding prediction, although this varies significantly among demes. This pattern of expansion and mismatching highlights the importance of historical and stochastic processes to invasions and the potential for novel selection regimes to promote niche evolution during invasion of a new range.

Methods

Study Species

Silene vulgaris (bladder campion) is a short-lived perennial native throughout much of Europe and parts of Asia. It occurs across a large geographic range that extends from Northern Africa (Morocco) and the Mediterranean, north to Scandinavia, and from the Atlantic coast of Europe east into Russia. It typically occupies disturbed habitats, both natural (riverbanks, rocky talus slopes, cliff faces, meadows) and anthropogenic (roadsides, fields, mine tailings, railroad ballast) (Marsden-Jones and Turrill 1957).

S. vulgaris was introduced to North America with European settlement. The earliest recorded occurrences in Floras are from the northeastern United States (Boston) and Canada (fields around Quebec City) (Cutler 1785, Pursh 1814). Since that time, *S. vulgaris* has become a common weed in North America, being reported across much of the United States and Canada, although the true extent of its range is largely unknown.

Sampling of Occurrence Points

The geographic range of *S. vulgaris* in Europe is well known (Atlas Florae Europeae: Jalas and Suominen 1986); therefore it was not necessary to undertake a comprehensive sampling strategy to establish the limits of its distribution, as in the introduced range (see below). Rather, we obtained occurrence points for analysis during extensive collecting trips done by ourselves and our colleagues between 1999 and 2006. These observations were supplemented with literature accounts of the species when details of the study location were provided. The resulting dataset included 116 observations and covered the full breadth of the known distribution in Europe and the Near East (Appendix A).

To establish the geographic range of *Silene vulgaris* in North America, we queried 59 North American herbaria for information on historically collected specimens of *S. vulgaris* (Appendix A). Herbaria were selected to provide coverage across the anticipated range extent of the species. In order to minimize bias, we also queried herbaria located in areas where *S. vulgaris* was not anticipated to occur (Appendix A). Specimen label data on the locality and date of collection were obtained either during personal visits to the herbaria, from curated electronic databases, or from specimens loaned to the University of Virginia for this study. All localities were georeferenced to the nearest 10^{-4} degree, using either GPS coordinates supplied by the collector or by matching the written locality description to on-line geodatabases (within the United States: www.topozone.com; outside the United States: www.multimap.com). In total, we obtained data on 2314 recorded occurrences of *S. vulgaris* in North America, of which 1906 were unique collections (not sample duplicates or localities sampled by > 1 collector in a given year)

and 1642 were spatially unique sites (excluding samples from the same locality collected across multiple years).

Niche Modeling: Range-Wide Models

Ecological niche hypotheses were modeled using the Genetic Algorithm for Ruleset Prediction (GARP: Stockwell and Peters 1999), implemented in DesktopGARP (http://nhm.ku.edu/desktopgarp/). GARP is a machine-learning algorithm that employs multiple rule types (e.g., atomic rules, bioclimatic envelopes, logistic regression) in an iterative search for a combination of rules that maximizes the predictive accuracy of known occurrence points, given a set of environmental layers. It proceeds by dividing the input observations into two sets, with one set used as training points and the remaining being reserved for independent model validation. At each iteration, GARP produces a dataset of presence and pseudo-absence points by sampling 1250 presence points, with replacement, from the training dataset and 1250 pseudo-absence points picked randomly from the background. Since true absences are typically not available from occurrence data, such as museum specimens, pseudo-absences drawn from random points within the defined study area are instead used to discriminate between environments at localitites known to be occupied and those where occupation is unknown. GARP then attempts to predict species occurrence by fitting a set of rules to the resampled dataset. Rules that contribute positively to model accuracy are retained while others are discarded, and the procedure continues iteratively until either convergence or a maximum number of

iterations are reached. Model accuracy of the final set of rules following convergence is then tested using the held back data.

Environmental variables consisted of global climate data based on monthly means averaged over a 30 year period (1961-1990) from the International Panel on Climate Change (New et al. 1999) and the United States Geological Survey mapped to a resolution of 0.1 degrees. To avoid problems of over prediction that may accompany models with many autocorrelated environmental variables (Parmesan et al. 2005, Weins and Graham 2005), we chose to analyze a non-exhaustive set of climatic variables that we believe are important in defining the distribution of *S. vulgaris*. Initially, we included data on ground-frost frequency (days per year), wet day frequency (days per year), precipitation (mm/day), solar radiation ($W \cdot m^{-2}$), average minimum temperature (°C), average mean temperature (°C), average maximum temperature (°C), and elevation derived from a digital elevation model (DEM). Next we used a jackknife procedure to reduce the model to those layers that contributed positively to model accuracy, following the methods of Peterson and Cohoon (1999) (Appendix B). This procedure is essentially a least-squares approach to model selection, whereby measures of model performance are correlated with the inclusion/eclusion of particular environmental variables, and variables which detract from model performance are omitted. The final set of variables consisted of six environmental layers, including ground frost frequency, amount and frequency of precipitation, minimum temperature, mean temperature, and maximum temperature; DEM was removed due to its negative association with model accuracy (Table S1).

Simulation studies suggest predictive accuracy during niche modeling approaches 90% of maximum with 10 observations and is near 100% of maximum with 50 observations (Stockwell and Peterson 2002). To ensure our native range sample size was adequate to saturate model accuracy, we randomly sampled from our 116 European occurrence points to produce replicate datasets each consisting of 3, 5, 10, 20, 30, 50, 75, and 100 observations. These models clearly show an asymptote in model accuracy after 50 observations, suggesting our sample size of 116 points is more than adequate for developing accurate niche models in the native range (Figure S1).

Each implementation of the GARP algorithm is a stochastic simulation that produces a prediction of the species' niche. To account for variability among runs, we employed a model selection procedure. Specifically, we ran 200 replicate models for each project. Model selection proceeded by taking the 20 models with the lowest extrinsic omission errors (known presence points falsely predicted as absent). We then calculated the average commission error (predicted presence when no point is observed) and retained the 10 models with values closest to the mean. These 10 models were then displayed within ArcGIS v9 (Environmental Systems Research Institute 2004) and used to conduct a weighted overlay to determine areas of model agreement. A strict consensus niche prediction was developed based on areas of agreement among all 10 of the best subset models for each project.

Because models with a large prediction area would be expected to contain more occurrences just by chance, a null expectation was generated by multiplying the number of occurrences by the proportion of the total study area (i.e., the United States and Canada) that the consensus niche predicted species presence (e.g., Peterson et al. 1999).

The number of occurrence points observed within the consensus niche was then compared to the number expected by chance, and significance evaluated using a χ^2 test with 1 d.f..

In addition to predicting the range in North America based solely on European occurrences, a measure of niche conservatism during invasion, we were also interested in detecting niche divergence between ranges. If *S. vulgaris* occupies different portions of the environmental niche space within each range, then we would expect a reciprocal loss of model accuracy at predicting occurrences across ranges relative to predicting within each range. In other words, under the hypothesis of divergence, the frequency of correctly predicted occurrences should follow the pattern: (NA predicts NA) > (EU predicts NA) and (EU predicts EU) > (NA predicts EU). Niche conservatism then forms the null model: (NA predicts NA) = (EU predicts NA) and (EU predicts EU) = (NA predicts EU). We developed all four sets of models and tested the null hypothesis that accuracy in each cross-range prediction did not deviate from expected values based on the corresponding within range prediction using χ^2 tests with 1 d.f.

Niche Modeling: Population Genetic Models

In order to assess intraspecific variation in niche requirements, we developed niche predictions for each of four genetically divergent demes (deme 1, deme 2.2, deme 2.3, deme 3) based on a previous study of AFLP multilocus genotypes (Chapter 3). This previous work identified five demes in Europe, four of which have been introduced to North America. The fifth deme (deme 2.1) is rare and spatially restricted in Europe to an area bordering the Adriatic Sea. Since this deme does not occur in North America, and its

sample size is too small to permit the development of robust niche predictions, it was not included in the current analyses. Sample sizes in Europe for the remaining demes were: deme 1 $N = 46$; deme 2.2 $N = 54$; deme 2.3 = 63; deme 3 $N = 62$. We ran 200 GARP models per deme, using the same protocol and environmental layers as described above. The 10 best models for each deme were imported into ArcGIS and used to define each deme's consensus niche prediction within North America. North American genotypes were then categorized as occurring either (1) within the niche predicted by its own deme, (2) within the niche predicted by another deme, or (3) outside the niche predicted by any demes. Heterogeneity among demes in niche matching during invasion was evaluated with a *G*-test of independence.

Results

Range-Wide Models

The ecological niche of *S. vulgaris* in its native range showed evidence of conservatism during invasion (Figure 1). Of the 1642 spatially unique localities within North America, the European consensus successfully predicted 1203, a highly significant difference relative to the null expectation ($\chi^2 = 5399$, d.f. = 1, $P < 0.0001$). The areas of initial introduction appear to have been within the ancestral European niche. The earliest collections were from Philadelphia (1825), Nova Scotia (1840), and Boston (1842) (Figure 2A). Spread from these and other early sites of introduction (e.g., New York City, New Haven, Montreal) began slowly then accelerated, with particularly dense collections from the Philadelphia area (Figure 2, B-D). By 1950, *S. vulgaris* had spread throughout

most of its eventual range. An area of apparently suitable but unoccupied niche space was observed in the southeastern and south-central United States.

In addition to evidence for conservatism, there was also a notable expansion outside of the ancestral niche. Expansion outside the niche was most pronounced in portions of the Rocky Mountains and west-central regions of the United States and Canada, and became more frequent as the invasion progressed (G = 71.73, d.f. = 1, $P < 0.0001$; Figure 3). This difference was not due to dispersal limitation, as the majority of observations outside the niche occured in the western states and provinces, which were already well colonized by 1950 (Figure 3). Rather, changes in the relative abundance of observations inside and outside the consensus niche appear to have occurred over the course of invasion.

In contrast to the European prediction, niche models based on North American occurrences covered the full breadth of the species' distribution (Figure 4). The consensus of the 10 best models successfully predicted 1518 of 1642 North American observations (χ^2 = 2741, d.f. = 1, $P < 0.0001$). Only two northern localities in Canada fell outside the prediction of any of these best subset models (near Dawson City, Yukon Territory, and near Schefferville, Quebec). The consensus niche predicted by European samples fell almost entirely within the North American consensus, with the latter successfully predicting the area of expansion in central and western North America (Figure 5). Of the 1642 spatially unique localities in North America, 1203 (73 %) were predicted by both continents, 315 (19 %) were predicted by North America but not Europe, 0 (0 %) were predicted by Europe but not North America, and 124 (8 %) were not predicted by either continent. Thus, while there was substantial evidence for niche

expansion since the invasion, there was no evidence that the ancestral niche had been bottlenecked.

Evidence for niche divergence between continents was evident in the reciprocal range predictions. When cross-continent niche models (NA predicts EU; EU predicts NA) were compared to within continent models (EU predicts EU; NA predicts NA), the highest accuracy was observed in the range where the models were developed, and accuracy was reciprocally lower in each case for the predicted range, indicating a difference in the ecological niches occupied across ranges (Table 1).

Population Genetic Models

European demes produced spatially divergent niche predictions within the introduced range, indicating the presence of intraspecific niche variation (Figure 6). Of the 800 models run (four demes x 200 models per deme), 99.5% were significant predictors of deme occurrence in Europe (χ^2 test: $P < 0.05$ for 796 of 800 models), and when combined across demes very closely approximated the range-wide niche predicted by the European consensus. However, the partial geographic separation of deme predictions combined with the haphazard distribution of genotypes resulted in many individuals located outside their niche. The frequency of niche mismatches varied significantly among the demes ($G = 26.78$, d.f. = 6, $P = 0.0002$; Table 2). Genotypes from deme 2.2, whose native range ancestors occurred primarily in eastern Europe, exhibited the highest degree of niche matching in North America (Table 2; Figure 6). Their predicted niche was distributed primarily throughout the northeastern U.S. and southern Canada, where the density of herbarium collections indicated propagule pressure had been most intense (Figure 2).

Deme 2.3 had the broadest area of niche prediction in North America, coincident with its broad distribution in Europe (Chapter 3). However, despite the favorable niche conditions available for establishment, it occurs in the lowest frequency of any deme in North America (N = 13 genotypes in North America). Genotypes from deme 3, whose ancestors occurred primarily in France and the U.K., often occurred outside of their niche which was centered in the southeastern U.S. (Figure 6). Thus, while there was evidence for niche divergence among demes, stochastic processes during invasion redistributed genotypes among the ancestral niches.

Discussion

This study used an introduced species to uncover two important results pertaining to the ecological niche during range expansion. First, while the native range niche predicted the majority of occurrences in the introduced range, there was a substantial frequency of observations that occurred outside the predicted niche area. Expansion outside of the ancestral niche space increased as the invasion progressed, suggesting either an evolutionary adjustment of the niche or an ecological release from a former constraint as a result of the invasion process. Second, genetic structure in the native range caused intraspecific variation in niche predictions during invasion. Introduced genotypes varied in their match between predicted and occupied area, presumably because of the stochastic nature of the colonization process. These two findings have important consequences for our understanding of the forces controlling niche evolution, and how introduced genotypes may be matched (or mismatched) with their ancestral niches during range expansion.

Niche Conservatism versus Expansion during Invasion

Historically, the evolution of ecological niches has been proposed to be under considerable constraint, perhaps resulting in relatively stable species borders and geographic range sizes (Antonovics 1976b, Kirkpatrick and Barton 1997). Empirical studies have provided examples of both long-term niche conservatism(Ricklefs and Latham 1992, Peterson et al. 1999, Prinzing et al. 2001) and more recent niche lability (Losos et al. 2003, Graham et al. 2004b, Knouft et al. 2006). Our results suggest both conservatism and expansion of the ecological niche may accompany species introductions in a new range (Figure 5). There may be several mechanisms at work here. First, the realized niche in Europe may reflect historical constraints not present in North America. For example, range expansion in Europe following the last glaciation may still be incomplete for some species owing to dispersal barriers, such as recent work has convincingly shown for forest trees (Svenning and Skov 2007). However, the near ubiquitous distribution of *S. vulgaris* in Europe, and its global occurrence across 6 continents (S. R. Keller, unpublished data), would seem to refute major barriers to dispersal as limitations on its regional distribution.

Second, since niche models are based on known occurrences, they effectively model the realized niche of the species, which may change for ecological reasons during invasion. In particular, invasions are often accompanied by the loss of specialist and/or generalist predators and pathogens (Wolfe 2002, Callaway et al. 2004, Colautti et al. 2004, DeWalt et al. 2004). If interspecific interactions act to restrict the realized niche in the native range, then enemy release could permit invaders to expand into previously

uninhabited regions of their fundamental niche. In Europe, *S. vulgaris* is attacked by several species of the Lepidopteran seed predator, *Hadena*, all of which are absent from North America (Kephart et al. 2006). While *Hadena* spp. have been observed to reduce fruit set locally within European populations (S.R. Keller, pers. obs), none show specificity on *S. vulgaris*, such as the relationship between *S. latifolia* and *H. bicuris* (Wolfe 2002, Wolfe et al. 2004), and it is not clear what impacts, if any, *Hadena* spp. have on regulating the regional distribution of *S. vulgaris* in Europe. Additionally, field experiments in North America suggest introduced genotypes of *S. vulgaris* do not benefit from enemy release, relative to a North American congener (Agrawal et al. 2005). In fact, introduced genotypes may actually suffer higher attack rates than native range genotypes by several generalist enemies (S.R. Keller unpublished data), a counterintuitive result if enemy release were responsible for niche expansion. Thus, while a role for interspecific interactions in the niche shift of introduced *S. vulgaris* populations cannot be ruled out, there is no supporting evidence.

Third, apparent niche divergence may represent combinations of environmental variables in North America that that are outside the range of or simply have no analog to those occurring in Europe (Peterson and Holt 2003, Parmesan et al. 2005). In other words, even though niche models are developed based on *ecological space*, they are tested in *geographic space* (Peterson 2006). When the niche model prediction is projected onto a new geographic area, it may encounter variable combinations that were unavailable and hence untested during model development (Soberon and Peterson 2005). In fact, investigators concerned with predicting where future invasions may occur often argue for the inclusion of occurrences from other invaded ranges when developing niche

models, since they may contain unique environments not found in the native range (e.g., Kriticos et al. 2005). Because niche models describe a complex multidimensional parameter space, direct tests of this possibility are not straightforward, and appropriate computational methods remain to be developed (Soberon and Peterson 2005). However, to provide a rough assessment of whether the zone of expansion in western North America contains environmental conditions also present in Europe, we developed niche models for occurrence points just in this region and projected these back onto Europe. We found that in general, these models were highly congruent with the native range niche predicted by European occurrence points (results not shown). This suggests that niche divergence between ranges was at least partly attributable to the presence of environmental conditions (either individually or in various combinations) in North America that only partially overlap with the environments found in Europe (see also Figure 5).

Fourth, adaptation to local conditions may permit introduced species to expand the boundaries of fundamental niche itself. Through a series of theoretical studies, R.D. Holt and colleagues have shown how the boundaries of a species' niche can evolve in response to selection under realistic demographic assumptions and over contemporary time scales (Holt and Gomulkiewicz 1997, Holt 2003, Holt et al. 2003, Holt et al. 2005a). Several studies have now demonstrated that adaptive evolution can occur quickly during invasion. For example, clines in size and physiological traits have been observed in the introduced ranges of birds, flies, and several species of plants (Johnston and Selander 1964, Weber and Schmid 1998, Huey et al. 2000, Kollmann and Banuelos 2004, Maron et al. 2004). In other cases, genotypes from the introduced and native ranges show clear

genetically-based differences in fitness related traits (Blair and Wolfe 2004, Lavergne and Molofsky 2007). In a companion study, we show that genetically-based divergence in several key phenotypic traits has occurred between North American and European *S. vulgaris*, and that after controlling for founder effects, this divergence is most likely a response to selection during the invasion process (Chapter 5). Thus, it is possible that the niche expansion we observed in this study may reflect adaptive evolution of *S. vulgaris* to its North American environments, although demonstrating this requires a more explicit experimental approach (Chapter 1). It is interesting to note that the region of putative niche expansion also corresponds to areas where infestations of *S. vulgaris* have been reported to reach very high densities in western Canada (Wall and Morrison 1990).

Until very recently, examples of niche expansion during invasion had been lacking (Broennimann et al. 2007). However, using much of the same climatic information as we did, Broennimann et al. (2007) demonstrated that the niche of Spotted Knapweed (*Cantaurea maculosa*) has undergone a climatic shift after colonizing western North America from Europe. The pattern they reported is similar to our results with *Silene vulgaris*. Both species have widespread (indeed largely overlapping) native geographic ranges in Europe. Both were introduced to regions in North America within the predicted niche and expanded into regions outside the predicted niche as the invasion progressed, and both show expansion outside the ancestral niche having occurred in western North America. They interpret their results by suggesting two of the four possible mechanisms we considered above, namely release from an ecological constraint (such as natural enemies), or evolutionary modification of the niche within the introduced range. We suggest the commonality between our two studies derives from either (1) a parallel

response to similar selection pressures exerted on these species within this geographic region, or (2) the underlying presence of environmental parameter space in the North American climate variables that does not occur in Europe.

Niche Population Structure and Stochastic Colonization

Our analysis is the first attempt to explicitly account for genetic structure in ecological niche predictions during invasion. Genetically divergent AFLP demes in Europe contain both widespread and spatially structured demes (Chapter 3). As a result, the area in the introduced range predicted to provide appropriate matches to the native environments also varied (Figure 6). Stochastic elements of the colonization history of *S. vulgaris* in North America involve both founder effects, where demes were sampled differentially from the native range, and the location of initial introduction. Two of the most widespread demes in Europe (demes 1 and 2.3) are present at relatively low frequencies in North America, making robust interpretation of niche matching problematic due to small sample size. Among the remaining two demes, the relative frequency of the eastern European deme 2.2 (24%) has greatly increased during or since its introduction to North America (51%), while that of the western European deme 3 (27%) has remained approximately the same (25%). This may be the result of a founder effect shifting initial deme frequencies (Chapter 3). However, the present results raise the possibility that the points of introduction may have been more favorable for one deme relative to the other. Historical reconstruction of the invasion suggests initial introductions happened in the northeast, and subsequently spread south and west (Figure 2, Cutler 1785, Pursh 1814). Deme 2.2 may have increased in frequency because the

initial invasion occurred in an area within its ancestral niche. By contrast, the invasion of deme 3 appears to have occurred largely in areas where it is mismatched from its ancestral niche (Table 2, Figure 6).

Our results present a situation that may be a common but unexplored feature of successful invasions, and perhaps responsible for failed invasions as well. An invasive species in its native range often has an evolutionary history of isolation among lineages that has lead to regional genetic structure (Neuffer and Hurka 1999, Meekins et al. 2001, Novak and Mack 2001, Schaal et al. 2003, Kolbe et al. 2004, Gaskin et al. 2005, Williams et al. 2005, Ostrowski et al. 2006, Taylor and Keller 2007). The spatial isolation among lineages in the native range may also reflect the occupation of distinct regions within the species-wide niche (this study). As a result, the zone of expected preadaptation may vary in location, size, degree of connectivity, etc., and colonists may match or mismatch their niche. Thus, the geographic distribution of an invasive may be affected by genetic divergence and evolutionary history in the native range, especially when the sampling process is stochastic. If instead colonization results in admixture and hybridization between previously isolated lineages (de la Vega et al. 1991, Gaskin and Schaal 2002, Frankham 2005, Genton et al. 2005b, Lavergne and Molofsky 2007), then gene flow may produce genotypes that can expand into ecological niches that were previously inaccessible in their native range.

The consequences of niche matching for establishment success during invasion will depend critically on the rate of population growth outside the ancestral niche (Holt et al. 2005a). The key question is whether variance among genotypes in the matching of ancestral niches translates into functional differences in fitness, or if it instead reflects

historical contingency and neutral structure. Testing this question can be accomplished through manipulative experiments in the introduced range (Chapter 5).

Acknowledgements

We appreciate the many herbaria and their curators for specimen loans. We gratefully acknowledge Eric Nagy for assistance in obtaining herbarium loans, and Christi Howell for help with data collection.

References

Agrawal, A. A., P. M. Kotanen, C. E. Mitchell, A. G. Power, W. Godsoe, and J. Klironomos. 2005. Enemy release? An experiment with congeneric plant pairs and diverse above- and belowground enemies. Ecology 86:2979-2989.

Antonovics, J. 1976. The nature of limits to natural selection. Annals of the Missouri Botanical Garden 63:224-247.

Blair, A. C., and L. M. Wolfe. 2004. The evolution of an invasive plant: an experimental study with *Silene latifolia*. Ecology 85:3035-3042.

Broennimann, O., U. A. Treier, H. Muller-Scharer, W. Thuiller, A. T. Peterson, and A. Guisan. 2007. Evidence of climatic niche shift during biological invasion. Ecology Letters 10:701-709.

Callaway, R. M., G. C. Thelen, A. Rodriguez, and W. E. Holben. 2004. Soil biota and exotic plant invasion. Nature 427:731-733.

Colautti, R. I., A. Ricciardi, I. A. Grigorovich, and H. J. MacIsaac. 2004. Is invasion success explained by the enemy release hypothesis? Ecology Letters 7:721-733.

Connell, J. H. 1961. The influence of interspecific competition and other factors on the distribution of the barnacle *Chthamalus stellatus*. Ecology 42:710-723.

Cutler, M. 1785. An account of some of the vegetable productions, naturally growing in this part of America, botanically arranged. Memoirs of the American Academy of Arts and Sciences I:396-493.

de la Vega, M., P. Garcia, and R. Allard. 1991. Multilocus genetic structure of ancestral Spanish and colonial California populations of *Avena barbata*. PNAS 88:1202-1206.

DeWalt, S. J., J. S. Denslow, and K. Ickes. 2004. Natural-enemy release facilitates habitat expansion of the invasive tropical shrub *Clidemia hirta*. Ecology 85:471-483.

Environmental Systems Research Institute. 2004. ArcGIS version 9. *in*. Environmental Systems Research Institute, Redlands, CA, U.S.A.

Frankham, R. 2005. Resolving the genetic paradox in invasive species. 94:385.

Gaskin, J. F., and B. A. Schaal. 2002. Hybrid *Tamarix* widespread in U.S. invasion and undetected in native Asian range. PNAS 99:11256-11259.

Gaskin, J. F., D.-O. Zhang, and M.-C. Bon. 2005. Invasion of *Lepidum draba* (Brassicaceae) in the western United States: distributions and origins of chloroplast DNA haplotypes. Molecular Ecology 14:2331-2341.

Genton, B. J., J. A. Shykoff, and T. Giraud. 2005. High genetic diversity in French invasive populations of common ragweed, *Ambrosia artemisiifolia*, as a result of multiple sources of introduction. Molecular Ecology 14:4275-4285.

Graham, C. H., S. Ferrier, F. Huettman, C. Moritz, and A. T. Peterson. 2004a. New developments in museum-based informatics and applications in biodiversity analysis. Trends in Ecology & Evolution 19:497-503.

Graham, C. H., S. R. Ron, J. C. Santos, C. J. Schneider, and C. Moritz. 2004b. Integrating phylogenetics and environmental niche models to explore speciation mechanisms in Dendrobatid frogs. Evolution 58:1781-1793.

Grinnell, J. 1917. Field tests of theories concerning distributional control. The American Naturalist 51:115-128.

Grinnell, J. 1924. Geography and evolution. Ecology 5:225-229.

Hochberg, M. E., and A. R. Ives. 1999. Can natural enemies enforce geographical range limits? Ecography 22:268-276.

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

Holt, R. D., M. Barfield, and R. Gomulkiewicz. 2005. Theories of niche conservatism and evolution: Could exotic species be potential tests? Pages 259-290 *in* D. F. Sax, J. J. Stachowicz, and S. D. Gaines, editors. Species invasions: insights into ecology, evolution, and biogeography. Sinauer Associates, Inc., Sunderland, MA.

Holt, R. D., and M. S. Gaines. 1992. Analysis of adaptation in heterogenous landscapes -- implictions for the evolution of fundamental niches. Evolutionary Ecology 6:433-447.

Holt, R. D., and R. Gomulkiewicz. 1997. How does immigration influence local adaptation? A reexamination of a familiar paradigm. The American Naturalist 149:563-572.

Holt, R. D., R. Gomulkiewicz, and M. Barfield. 2003. The phenomenology of niche evolution via quantitative traits in a 'black-hole' sink. Proceedings of the Royal Society of London Series B 270:215-224.

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

Holt, R. D., and T. H. Keitt. 2005. Species' borders: a unifying theme in ecology. Oikos 108:3-6.

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

Huntley, B., P. J. Bartlein, and I. C. Prentice. 1989. Climatic control of the distribution and abundance of Beech (Fagus L.) in Europe and North America. Journal of Biogeography 16:551-560.

Hutchinson, G. E. 1957. Concluding remarks. Cold Spring Harbor Symposium in Quantitative Biology 22:415-427.

Jalas, J., and J. Suominen. 1986. Atlas Florae Europaeae. Cambridge University Press, Cambridge.

Johnston, R. F., and R. K. Selander. 1964. House sparrows: rapid evolution of races in North America. Science 144:548-550.

Kephart, S., R.J. Reynolds, M.T. Rutter, C.B. Fenster, and M.R. Dudash. 2006. Pollination and seed predation by moths on *Silene* and allied Caryophyllaceae: evaluating a model system to study the evolution of mutualisms. New Phytologist 169:667-680.

Kirkpatrick, M., and N. H. Barton. 1997. Evolution of a species' range. American Naturalist 150:1-23.

Knouft, J. H., J. B. Losos, R. E. Glor, and J. J. Kolbe. 2006. Phylogenetic analysis of the evolution of the niche in lizards of the *Anolis sagrei* group. Ecology 87:S29-S38.

Kolbe, J. J., R. E. Glor, L. R. Schettino, A. C. Lara, A. Larson, and J. B. Losos. 2004. Genetic variation increases during biological invasion by a Cuban lizard. Nature 431:177-181.

Kollmann, J., and M. J. Banuelos. 2004. Latitudinal trends in growth and phenology of the invasive alien plant *Impatiens glandulifera* (Balsaminaceae). Diversity and Distributions 10:377-385.

Kriticos, D. J., T. Yonow, and R. E. McFadyen. 2005. The potential distribution of *Chromolaena odorata* (Siam weed) in relation to climate. Weed Research 45:246-254.

Lavergne, S., and J. Molofsky. 2007. Increased genetic variation and evolutionary potential drive the success of an invasive grass. PNAS 104:3883-3888.

Leger, E. A., and K. J. Rice. 2007. Assessing the speed and predictability of local adaptation in invasive California poppies (*Eschscholzia californica*). Journal of Evolutionary Biology 20:1090-1103.

Losos, J. B., M. Leal, R. E. Glor, K. de Queiroz, P. E. Hertz, L. R. Schettino, A. C. Lara, T. R. Jackman, and A. Larson. 2003. Niche lability in the evolution of a Caribbean lizard community. Nature 424:542-545.

Maron, J. L., M. Vila, R. Bommarco, S. Elmendorf, and P. Beardsley. 2004. Rapid evolution of an invasive plant. Ecological Monographs 74:261-280.

Marsden-Jones, E. M., and W. B. Turrill. 1957. The bladder campions. Ray Society, London.

Mau-Crimmins, T. M., H. R. Schussman, and E. L. Geiger. 2006. Can the invaded range of a species be predicted sufficiently using only native-range data?: Lehmann lovegrass (*Eragrostis lehmanniana*) in the southwestern United States. Ecological Modelling 193:736-746.

Meekins, J. F., H. E. Ballard, and B. C. McCarthy. 2001. Genetic variation and molecular biogeography of a North American invasive plant species (*Alliaria petiolata*) Brassicaceae. International Journal of Plant Sciences 162:161-169.

Neuffer, B., and H. Hurka. 1999. Colonization history and introduction dynamics of *Capsella bursa-pastoris* (Brassicaceae) in North America: isozymes and quantitative traits. Molecular Ecology 8:1667-1681.

New, M., M. Hulme, and P. Jones. 1999. Representing twentieth-century space-time climate variability. Part I: Development of a 1961-1990 mean monthly terrestrial climatology. Journal of Climate 12:829-856.

Novak, S. J., and R. N. Mack. 2001. Tracing plant introduction and spread: genetic evidence from *Bromus tectorum* (Cheatgrass). BioScience 51:114-122.

Ostrowski, M.-F., J. Davis, S. Santoni, H. Mckhann, X. Reboud, V. Le Corre, C. Camilleri, D. Brunel, D. Bouchez, B. Faure, and T. Bataillon. 2006. Evidence for a large-scale population structure among accessions of Arabidopsis thaliana: possible causes and consequences for the distribution of linkage disequilibrium. Molecular Ecology 15:1507-1517.

Parmesan, C., S. Gaines, L. Gonzalez, D. M. Kaufman, J. Kingsolver, A. T. Peterson, and R. Sagarin. 2005. Empirical perspectives on species borders: from traditional biogeography to global change. Oikos 108:58-75.

Peterson, A. T. 2003. Predicting the geography of species' invasion via ecological niche modeling. Quarterly Review of Biology 78:419-433.

Peterson, A. T. 2006. Uses and requirements of ecological niche models and related distributional models. Biodiversity Informatics 3:59-72.

Peterson, A. T., and K.C. Cohoon. 1999. Sensitivity of distribution prediction algorithms to geographic data completeness. Ecological Modelling 117:159-164.

Peterson, A. T., and R. D. Holt. 2003. Niche differentiation in Mexican birds: using point occurrences to detect ecological innovation. Ecology Letters 6:774-782.

Peterson, A. T., E. Martinez-Meyer, C. Gonzalez-Salazar, and P. W. Hall. 2004. Modeled climate change effects on distributions of Canadian butterfly species. Canadian Journal of Zoology 82:851-858.

Peterson, A. T., J. Soberon, and V. Sanchez-Cordero. 1999. Conservatism of ecological niches in evolutionary time. Science 285:1265-1267.

Peterson, A. T., and D. A. Vieglais. 2001. Predicting species invasions using ecologic niche modeling. BioScience 51:363-371.

Prinzing, A., W. Durka, S. Klotz, and R. Brandl. 2001. The niche of higher plants: evidence for phylogenetic conservatism. Proceedings of the Royal Society of London Series B 268:2383-2389.

Pursh, F. 1814. Flora Americae Septentrionalis. White, London.

Ricklefs, R. E., and R. E. Latham. 1992. Intercontinental correlation of geographical ranges suggests stasis in ecological traits of relict genera of temperate perennial herbs. The American Naturalist 139:1305-1321.

Runyeon, H., and H. C. Prentice. 1997. Genetic differentiation in the bladder campions, *Silene vulgaris* and *Silene uniflora* (Caryophyllaceae) in Sweden. Biological Journal of the Linnaean Society 61:559-584.

Sakai, A. K., F. W. Allendorf, J. S. Holt, D. M. Lodge, J. Molofsky, K. A. With, S. Baughman, R. J. Cabin, J. E. Cohen, N. C. Ellstrand, D. E. McCauley, P. O'Neil, I. M. Parker, J. N. Thompson, and S. G. Weller. 2001. The population biology of invasive species. Annual Review of Ecology and Systematics 32:305-332.

Schaal, B. A., J. F. Gaskin, and A. L. Caicedo. 2003. Phylogeography, haplotype trees, and invasive plant species. Journal of Heredity 94:197-204.

Soberon, J., and A. T. Peterson. 2005. Interpretation of models of fundamental ecological niches and species' distributional areas. Biodiversity Informatics 2:1-10.

Stockwell, D., and D. Peters. 1999. The GARP modelling system: problems and solutions to automated spatial prediction. International Journal of Geographic Information Science 13:143-158.

Stockwell, D. R. B., and A. T. Peterson. 2002. Effects of sample size on accuracy of species distributional models. Ecological Modelling 148:1-13.

Storchova, H., and M. S. Olson. 2004. Comparison between mitochondrial and chloroplast DNA variation in the native range of *Silene vulgaris*. Molecular Ecology 13:2909-2919.

Svenning, J.-C., and F. Skov. 2007. Could the tree diversity pattern in Europe be generated by postglacial dispersal limitation? Ecology Letters 10:453-460.

Taylor, D. R., and S. R. Keller. 2007. Historical range expansion determines the phylogenetic diversity introduced during contemporary species invasion. Evolution 61:334-345.

Thuiller, W., S. Lavorel, M. T. Sykes, and M. B. Araujo. 2006. Using niche-based modelling to assess the impact of climate change on tree functional diversity in Europe. Diversity and Distributions 12:49-60.

van de Staaij, J. W. M., E. Bolink, J. Rozema, and W. H. O. Ernst. 1997. The impact of elevated UV-B (280-320 nm) radiation levels on the reproduction biology of a highland and a lowland population of *Silene vulgaris*. Plant Ecology 128:173-179.

Wall, D. A., and I. N. Morrison. 1990. Phenological development and biomass allocation in *Silene vulgaris* (Moenche) Garke. Weed Research 30:279-288.

Weber, E., and B. Schmid. 1998. Latitudinal population differentiation in two species of *Solidago* (Asteraceae) introduced to Europe. American Journal of Botany 85:1110-1121.

Weins, J. J., and C. H. Graham. 2005. Niche conservatism: integrating evolution, ecology, and conservation biology. Annual Review of Ecology, Evolution, and Systematics 36.

Williams, D. A., W. A. Overholt, J. P. Cuda, and C. R. Hughes. 2005. Chloroplast and microsatellite DNA diversities reveal the introduction history of Brazilian peppertree (*Schinus terebinthifolius*) in Florida. Molecular Ecology 14:3643-3656.

Wolfe, L. M. 2002. Why alien invaders succeed: support for the escape from enemies hypothesis. American Naturalist 160:705-711.

Wolfe, L. M., J. A. Elzinga, and A. Biere. 2004. Increased susceptibility to enemies following introduction in the invasive plant *Silene latifolia*. Ecology Letters 7:813-820.

Table 1. Comparison of the ability of niche models developed in one range to predict occurrences of *S. vulgaris* in the other range. Occurrence points were categorized as either falling inside or outside the consensus niche predicted by the 10 best models in each comparison (see Methods). χ^2 tests the null hypothesis that the frequencies of occurrences in each cross-range prediction do not deviate from expected values based on the corresponding within range prediction. Significant tests indicate divergence in the niche between range predictions.

	N. America predicts N. America	Europe predicts N. America	Europe predicts Europe	N. America predicts Europe
Frequency inside the consensus niche	1518 / 1642 (92%)	1203 / 1642 (73%)	88 / 116 (76%)	69 / 116 (59%)
	$\chi^2_{1\ d.f.} = 865.57$; $P < 0.0001$		$\chi^2_{1\ d.f.} = 17.00$; $P < 0.0001$	

Table 2. Niche matching among demes during invasion of North America. Genotypes belonging to different demes were categorized as falling either within the niche prediction of their source, within a niche predicted by another deme, or outside any predicted niche.

	Native range region of origin	Total *N*	Source niche	Other niche (any deme)	Outside any niche
Deme 1	Widespread	28	13 (46%)	5 (18%)	10 (36%)
Deme 2.2	Eastern Europe	88	50 (57%)	26 (29%)	12 (14%)
Deme 2.3	Widespread; mostly southern Europe	13	6 (46%)	3 (23%)	4 (31%)
Deme 3	Western Europe	42	10 (24%)	27 (64%)	5 (12%)

G-test between demes and category of occurrence: $G = 26.78$, d.f. = 6, $P = 0.0002$.

Table S1. ANOVA of model accuracy using jackknifed environmental layers. "Effect" indicates the direction of effect that inclusion of a variable had on model accuracy. See also Appendix B.

Source	**df**	**Mean Square**	***F*-ratio**	***P***	**Effect**
Frost-free days	1	0.00012	8.80	0.0031	+
DEM	1	0.00018	12.84	0.0003	-
Precipitation	1	0.00103	72.97	< 0.0001	+
Solar radiation	1	0.00002	1.11	0.2921	+
Minimum temperature	1	0.00014	10.10	0.0015	+
Mean temperature	1	0.00019	13.61	0.0002	+
Maximum temperature	1	0.00018	12.72	0.0004	+
Wet days	1	0.00047	32.92	< 0.0001	+
Error	1514	0.00001			

Figure Legends

Figure 1. Niche predictions for Europe and North America based on European occurrences. Darker shading indicates greater model agreement among the 10 best models for each range.

Figure 2. Time series of invasion overlaid on the consensus niche prediction from Europe (in green). Arrows indicate the earliest occurrences of *S. vulgaris* in North America among the herbaria sampled.

Figure 3. Comparison of occurrences inside and outside of the European consensus niche prediction (shown in green) before and after 1950.

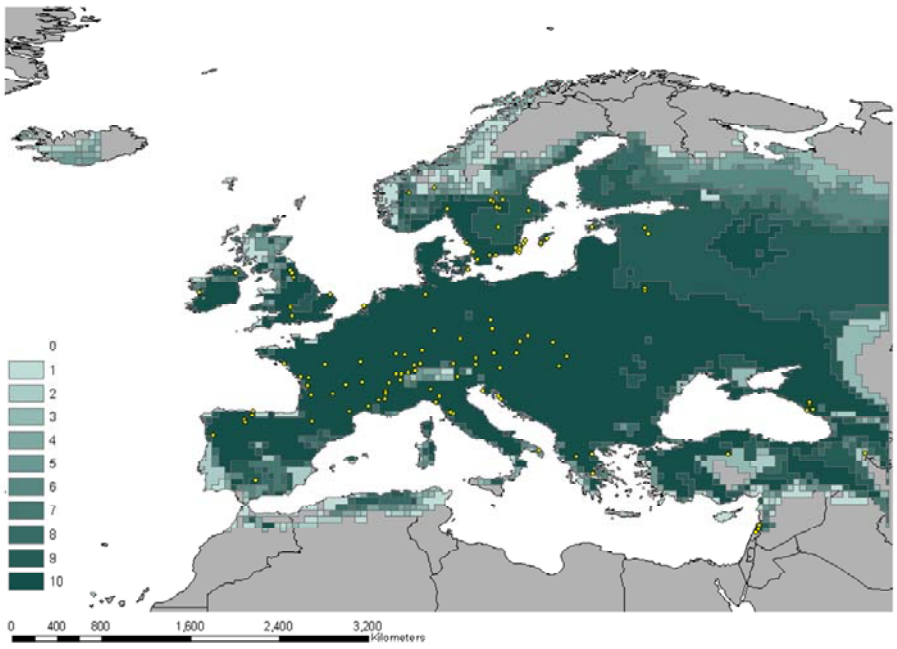
Figure 4. Niche prediction developed from North American occurrence points. Darker shading indicates greater model agreement.

Figure 5. The distribution of North American occurrences in relation to the consensus niche predictions based on European (light green) versus North America (dark green) models. The pie diagram shows the frequency of points falling within the predicted areas of each model.

Figure 6. Matching between genotype occurrences and consensus niche predictions based on the occurrence of AFLP demes in Europe. A-D represent deme 1, deme 2.2, deme 2.3, and deme 3, respectively. Black circles represent sampled genotypes belonging

to each deme, with diameter proportional to sample size. Areas in dark green represent the combined niche prediction when all demes are considered together.

Figure S1. Plot of model accuracy versus sample size for European occurrence points. Three datasets were generated for each value of *N* (3, 5, 10, 20, 30, 50, 75, and 100) by randomly sampling without replacement from the full dataset (*N* = 116). Error bars are ± 1 SE.


0
1
2
3
4
5
6
7
8
9
10
0 400 800 1,600 2,400 3,200
Kilometers

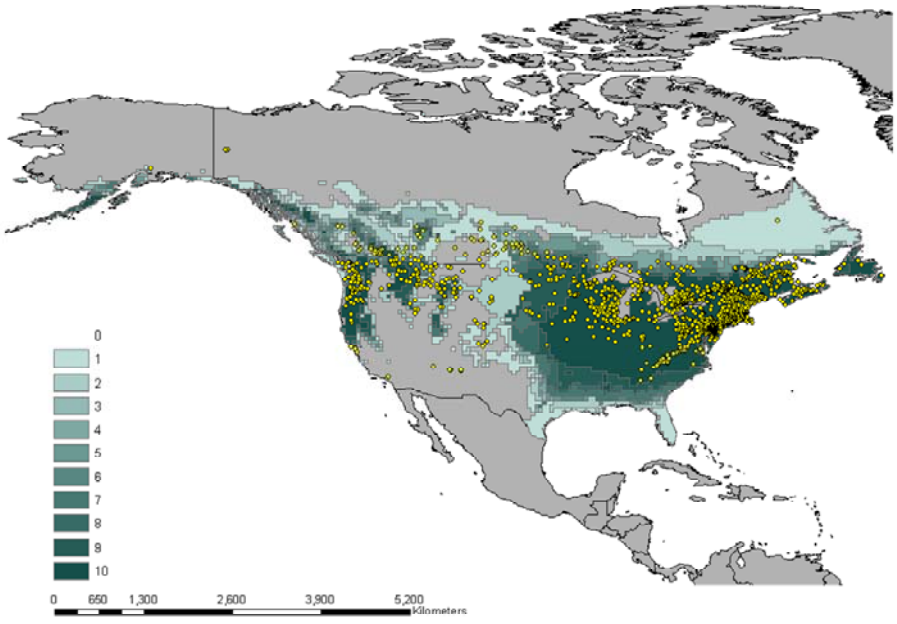

0
1
2
3
4
5
6
7
8
9
10
0 650 1,300 2,600 3,900 5,200
Kilometers

Figure 1.

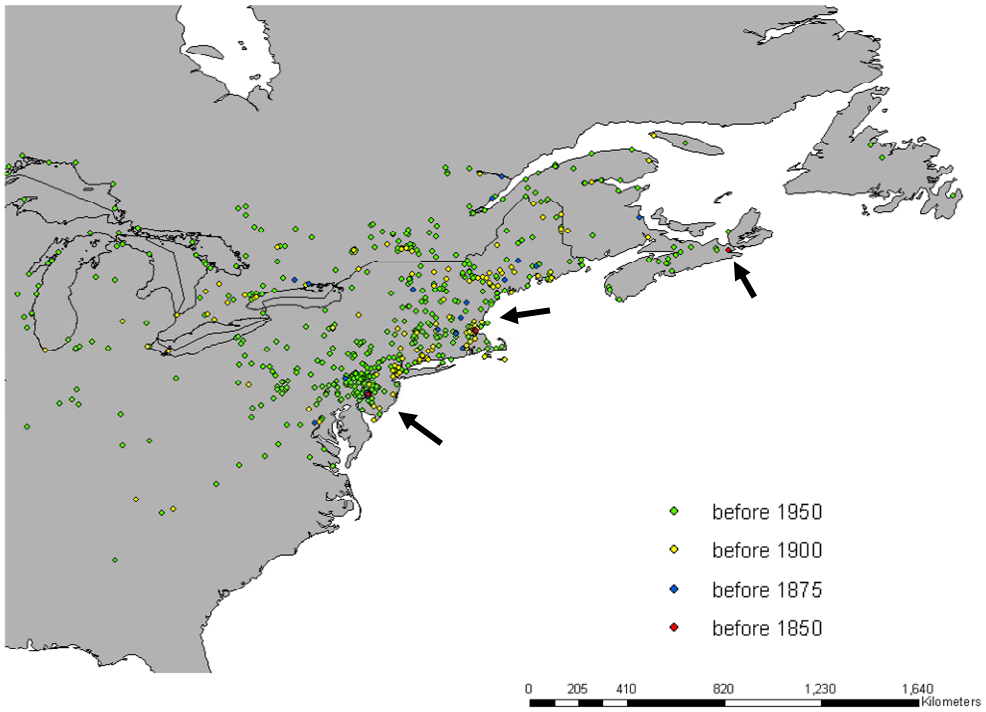

before 1950
before 1900
before 1875
before 1850
0
205
410
820
1,230
1,640
Kilometers

Figure 2.

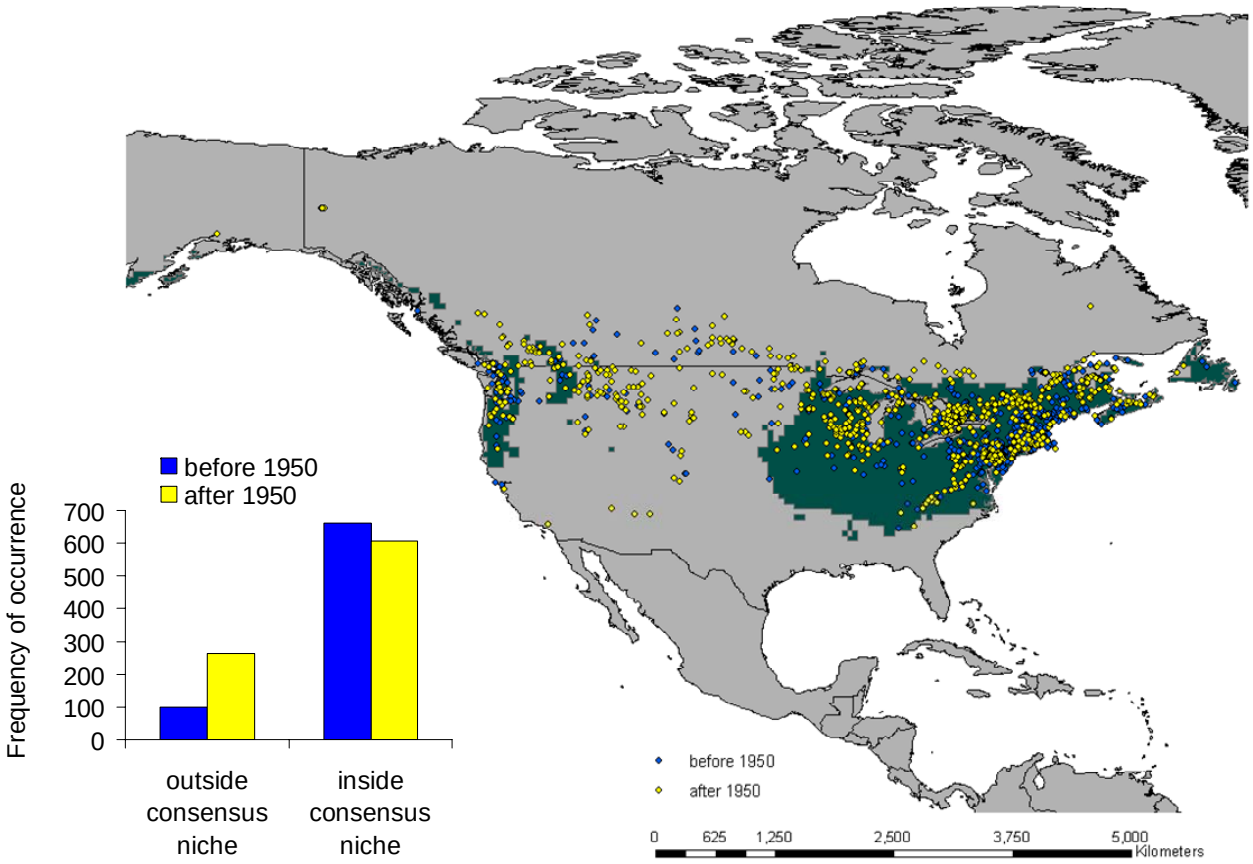

before 1950
after 1950
Frequency of occurrence
700
600
500
400
300
200
100
0
outside consensus niche
inside consensus niche
before 1950
after 1950
0
625
1,250
2,500
3,750
5,000
Kilometers

Figure 3.

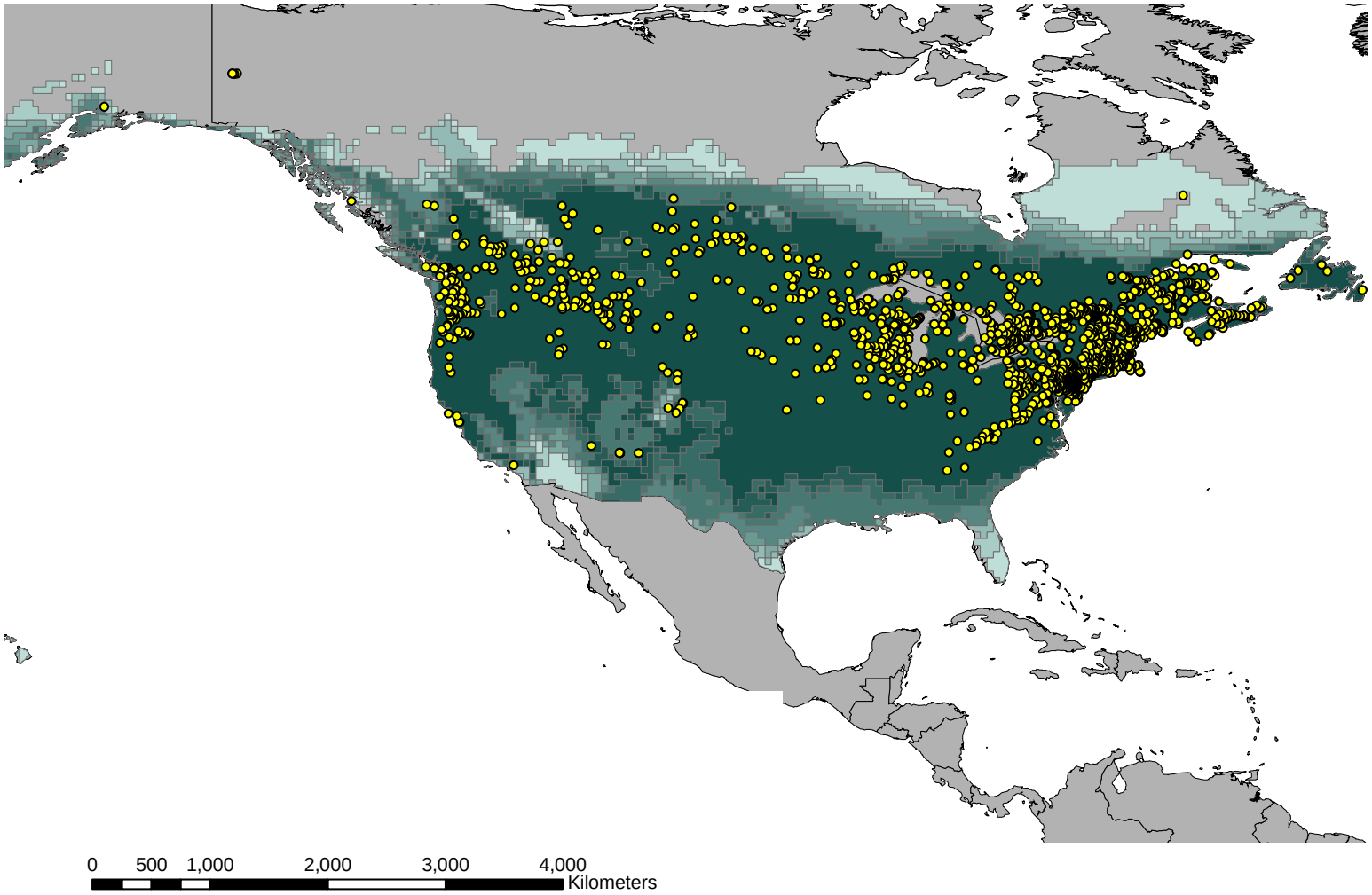


Figure 4.

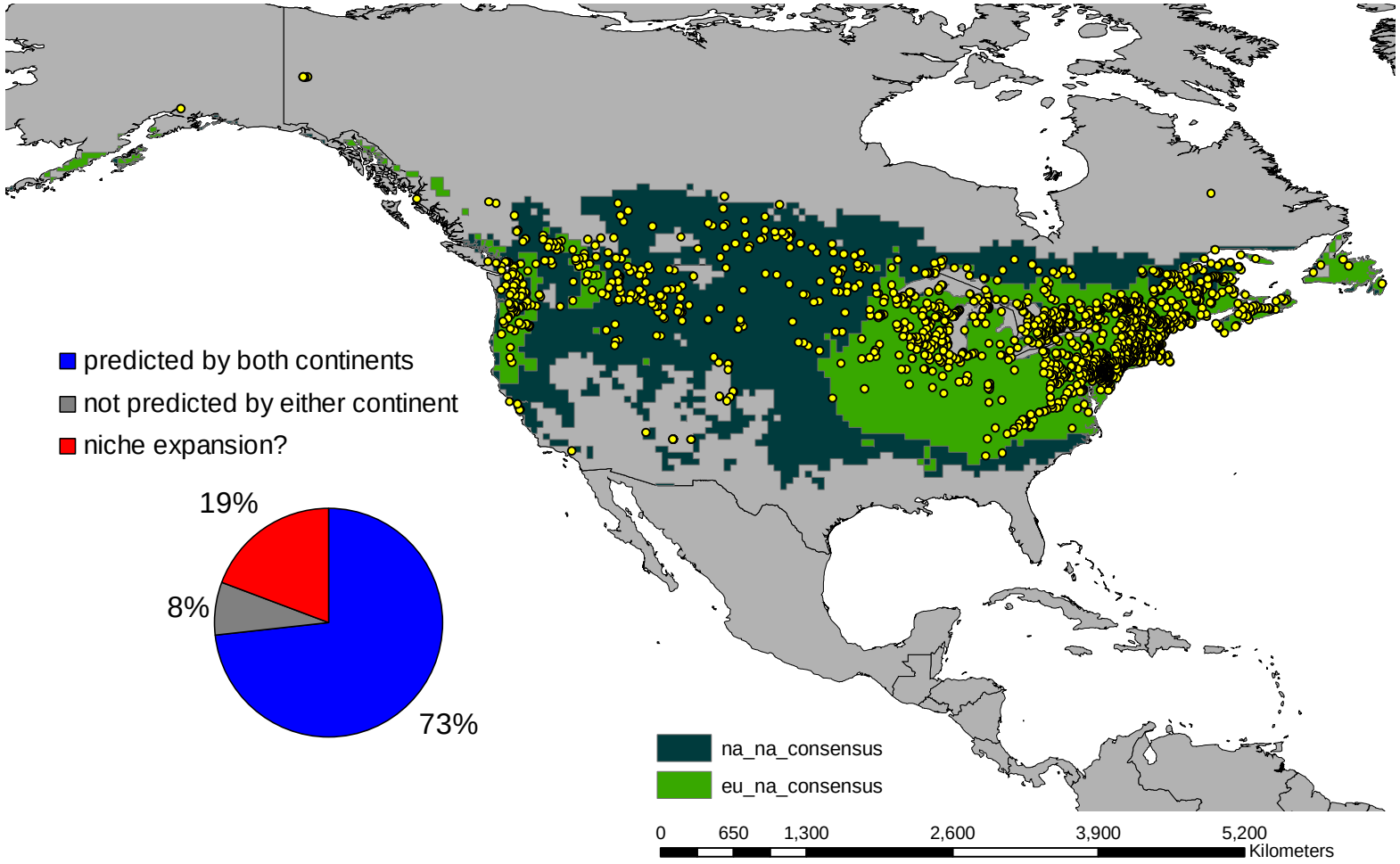

predicted by both continents
not predicted by either continent
niche expansion?
19%
8%
73%
na_na_consensus
eu_na_consensus
0
650
1,300
2,600
3,900
5,200
Kilometers

Figure 5.

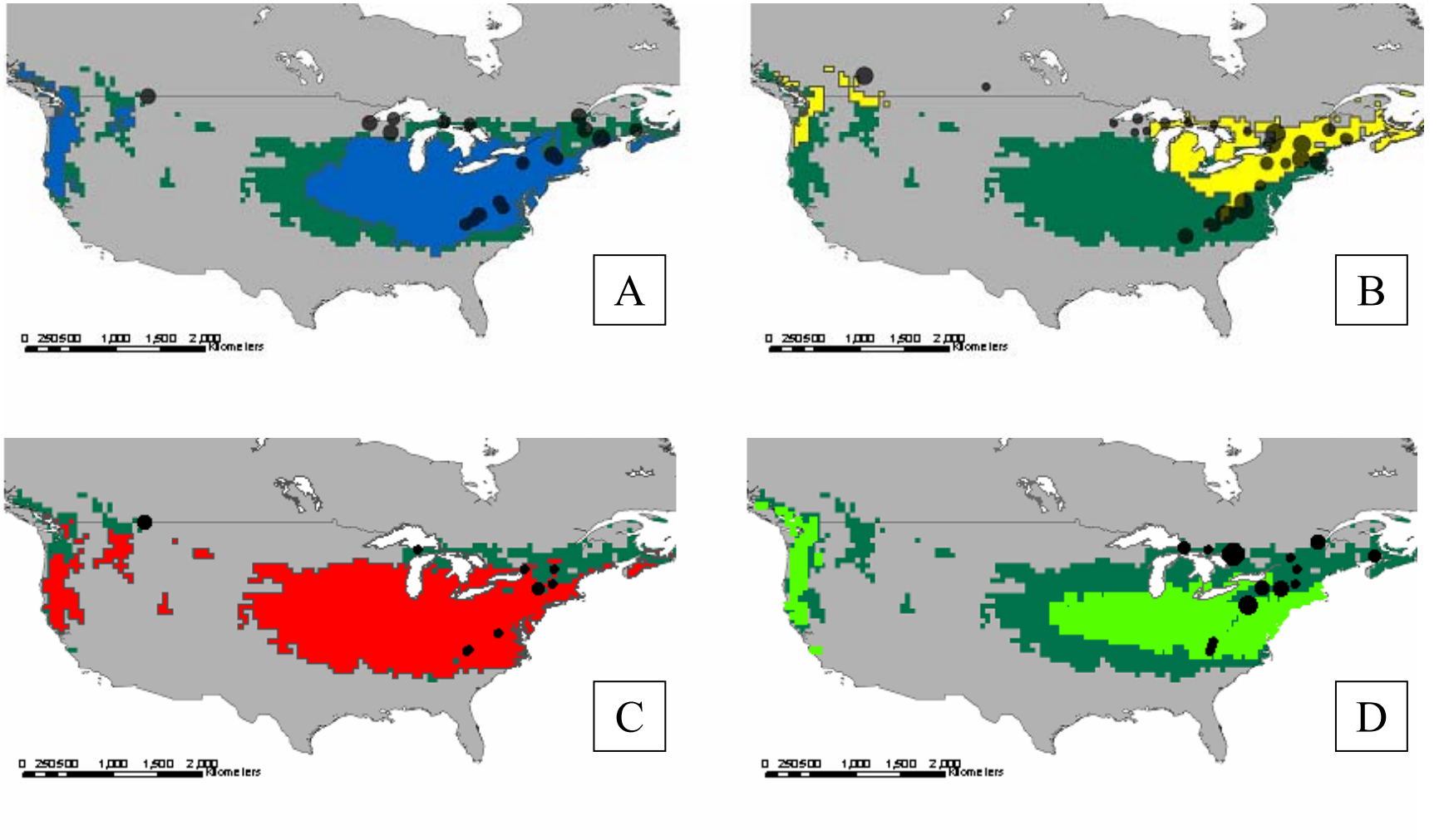


Figure 6.

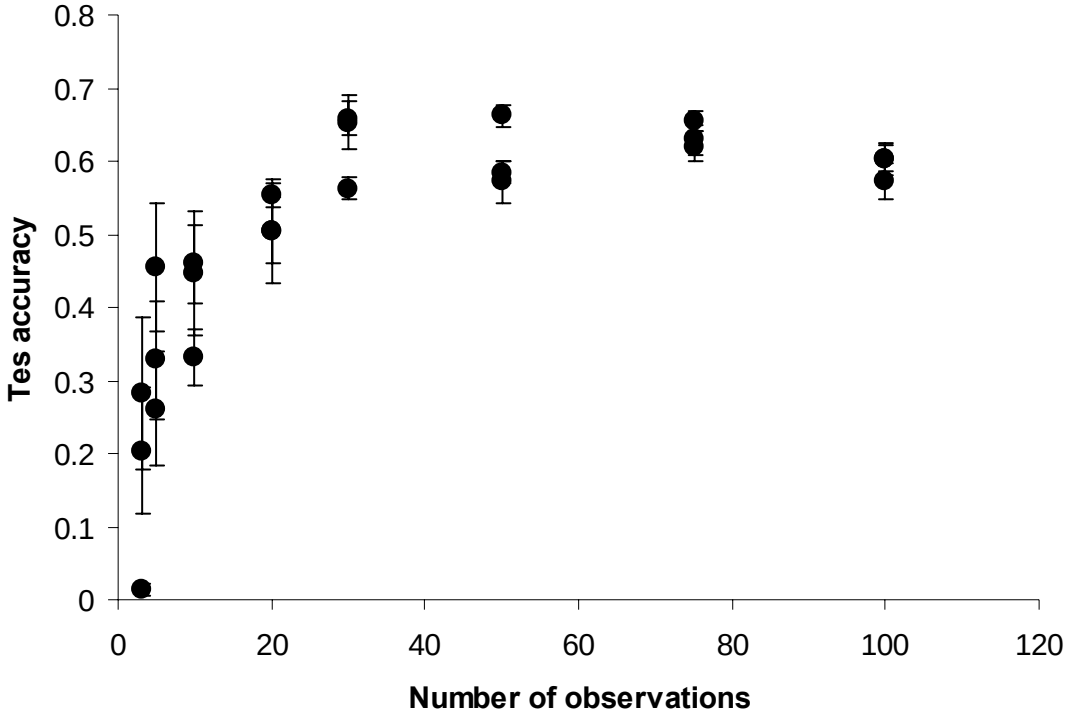

0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Tes accuracy
0
20
40
60
80
100
120
Number of observations

Figure S1.

Appendix A: Sources of European and North American occurrence points used in niche modeling.

Native European range

Source	Countries	Comments	N
Collecting trips	Armenia, Austria, Belarus, Croatia, Czech Republic, Denmark, Estonia, France, Germany, Greece, Hungary, Ireland, Italy, Lebanon, Norway, Portugal, Russia, Slovakia, Spain, Switzerland, Turkey, United Kingdom		86
Literature accounts	Austria[1,3,4], Germany[4], Netherlands[3], Norway[2], Sweden[1,2]	1,2,3,4	30
Total			116

Introduced North American range

Source	Herbarium Code	Comments	N
Academy of Natural Sciences	PH	25	395
Acadia University	ACAD	19	28
Algoma University College	AUC	5	4
Arizona State University	ASU	10	0
Canadian Museum of Nature	CAN	26	135

Colorado State University	CS	25	5
Delaware State University	DOV	26	2
Eastern Kentucky University	EKY	26	2
Forcella and Harvey 1988	n/a	6	31
Great Lakes Forestry Research Centre, Canadian Forest Service	SSMF	5	18
Harvard University Gray Herbarium	GH	26	275
Illinois Natural History Survey	ILLS	26	26
Iowa State University	ISC	26	34
Louisiana State University	LSU	11	4
Montana Weed Seed Free Forage Program	n/a	6	1
Morton Arboretum	MOR	26	33
National Biological Information Infrastructure (NBII) Crisis maps	n/a	12	8
New York Botanical Garden	NY	20	6
Northern Arizona University	ASC	10	0
Northern Kentucky University	KNK	26	11
Oklahoma Vascular Plants Database	OVPD	13	0
Ontario Forest Research Initiative	n/a	5	2
Ontario Ministry of Natural Resources Natural Heritage Information Center	NHIC	5	3

Oregon State University	OSC	6	32
R. Old (Private collector)	n/a	6	20
Saint Mary's College of California	n/a	9	1
San Jose State University	SJSU	14	1
Texas A&M University	TAMU	15	0
United States National Herbarium	USNH	25	90
University of Alabama	UNA	16	0
University of Alaska Fairbanks	ALA	17	1
University of Arizona	ARIZ	10	0
University of British Columbia	UBC	7	79
University of California Berkeley	JEPS	26	42
University of Colorado Boulder	COLO	21	6
University of Connecticut	CONN	26	14
University of Florida	FLAS	18	0
University of Georgia	GA	26	17
University of Idaho	ID	26	23
University of Illinois	ILL	26	58
University of Kansas	KANU	26	30

University of Kentucky	KY	26	4
University of Maine	MAINE	22	19
University of Manitoba	WIN	26	26
University of Massachusetts	MASS	26	75
University of Minnesota	MIN	26	28
University of Montana	MONTU	6	12
University of Montreal	MT	26	237
University of New Brunswick	UNB	23	31
University of Notre Dame	ND	26	6
University of Oregon	ORE	6	2
University of Saskatchewan	SASK	26	41
University of Toronto	TRT	26	146
University of Washington	WTU	6	37
University of Wisconsin	WIS	24	129
University of Wyoming	RM	6	12
Virginia Polytechnic Institute	VPI	25	26
Washington State University	WS	6	25
Yale University Herbarium	YU	26	21

Total	59	2316

[1] (Prentice 1992)
[2] (Runyeon and Prentice 1997)
[3] (van de Staaij et al. 1997)
[4] (Storchova and Olson 2004)
[5] Accessed via: http://www.northernontarioflora.ca/index.cfm
[6] Accessed via: http://invader.dbs.umt.edu/linksearch/linksearch.asp
[7] Accessed via: http://herbarium.botany.ubc.ca/index.html
[8] Accessed via: http://www.flmnh.ufl.edu/natsci/herbarium/cat/catsearch.htm
[9] Accessed via: http://www.calclim.dri.edu/cgi-bin/rawMAIN.pl?caCSCT
[10] Accessed via: http://seinet.asu.edu/seinet/collections/selection2.jsp
[11] Accessed via: http://www.herbarium.lsu.edu/search.php
[12] Accessed via: http://cain.ice.ucdavis.edu
[13] Accessed via: http://www.coordinatesolutions.com/ovpd/ovpd.aspx
[14] Accessed via: http://www.sjsu.edu/depts/herbarium/dbform.html
[15] Accessed via: http://www.csdl.tamu.edu/FLORA/tracy/main1.html
[16] Accessed via: http://serfis.by.ua.edu/herbdb/AccessingUNAdata.html
[17] Accessed via: http://arctos.database.museum/SpecimenSearch.cfm
[18] Accessed via: http://www.flmnh.ufl.edu/natsci/herbarium/cat/catsearch.htm
[19] Accessed via: http://herbarium.acadiau.ca/search.html
[20] Accessed via http://sciweb.nybg.org/science2/vii2.asp
[21] Accessed via http://cumuseum.colorado.edu/Research/Botany/Databases/search.php
[22] Accessed via http://herbaria.umaine.edu/index.php?action=plants
[23] Accessed via: http://www.unb.ca/herbarium/databases.html
[24] Accessed via: http://www.botany.wisc.edu/wisflora/specimen
[25] Data collected during personal visit to herbarium
[26] Specimens and/or specimen data sent directly to University of Virginia

Appendix B: Jackknifing of environmental variables

The jackknife analysis on the original 8 environmental variables was performed as previously described (Peterson and Cohoon 1999, Mau-Crimmins et al. 2006). Three replicate GARP models were run for all possible combinations of environmental variables (including variables by themselves). We then assessed the importance of each variable with analysis of variance (ANOVA), using model accuracy as our response variable. Model accuracy was assessed from the resampled dataset of the test points (N = 1250 presences and 1250 psuedo-absences), and was calculated as (A + B) / (A+B+C+D), where A = the number of points predicted present and observed present, B = the number of points predicted absent and observed absent, C = the number of points predicted present and observed absent, and D = the number of points predicted absent and observed present. We performed the jackknife procedure for both Europe and North America so as not to exclude environmental variables which may be influential on one continent and not the other, and ran ANOVA on the pooled dataset with continent as a blocking factor.

All environmental layers were significantly positive contributors to model accuracy, with two exceptions (Table S1). DEM significantly reduced model accuracy, while solar radiation increased accuracy but not significantly so. Therefore, we eliminated DEM and retained all other environmental variables when performing subsequent models runs in GARP.

Chapter 5:

Phenotypic evolution during biological invasion: adaptation overwhelms the founder effect [5]

[5] Formatted for submission as a coauthored manuscript: Keller, S.R. and D. R. Taylor

Abstract

Natural selection can be a powerful diversifying mechanism, producing phenotypes that functionally match local environmental variation. Yet historical processes and chance events may also exert strong influences on diversification of the phenotype. Many empirical studies provide evidence for the adaptive nature of phenotypic variation within a species, but few have directly considered how neutral evolution may contribute to the structuring of phenotypic diversity. Introduced species provide ideal testing grounds to assess the roles of selection and chance during phenotypic evolution. In their native range, many species are subdivided into lineages which may experience variance in sampling intensity during invasion. When these lineages are phenotypically divergent, it becomes critical to control for sampling frequency during invasion to separate neutral from adaptive mechanisms of change. In this study, we examined phenotypic evolution in the weedy plant *Silene vulgaris* since its introduction to North America from Europe approximately 200 years ago. We grew genotypes from both continents at two common garden sites within the introduced range, and combined phenotypic measurements with genetic ancestry inferred from a previous AFLP survey. We also conducted a novel test for preadaptation by comparing the relative fitness of European genotypes grown inside or outside the ecological niche predicted by their region of origin. Over a two year period, North American genotypes had faster germination rates, produced fewer more dominant stems, had higher survivorship, were more likely to become reproductive, and exhibited significantly greater flowering and fruit production compared to European genotypes. Genotypes from within each range also showed clinal variation between phenotypic traits and their latitude of origin. These clines were often opposite in sign

between North America and Europe, suggesting different patterns of selection across environmental gradients. The phenotypic divergence between continents and clinal diversification within continents was significant, even after using AFLP to control for founder effects. Thus, even though the evolutionary history of *S. vulgaris* in Europe resulted in genetic lineages that are phenotypically diverse, and these lineages occur in a biased sample in North America, natural selection has resulted in a pattern of phenotypic evolution that greatly exceeds expectations due to chance. Historical niche divergence among lineages did not influence the relative fitness of European genotypes grown at different sites. Based on these results, natural selection emerges as the primary driving force behind the contemporary evolution of *S. vulgaris* in North America. By controlling for ancestry and sampling effects, this study provides a much needed test of the assumed adaptive nature of phenotypic evolution during invasion.

Keywords: Selection, adaptation, genetic drift, founder effect, invasion, *Silene*, preadaptation.

Introduction

The importance of selective versus neutral processes as drivers of phenotypic diversity has been a recurring theme in evolutionary biology, at times even a tension, since the modern synthesis (Provine 1971, Coyne et al. 1997, Wade and Goodnight 1998, Coyne et al. 2000, Goodnight and Wade 2000). Natural selection provides a mechanism for producing phenotypes that confer high fitness within local environments. Yet historical processes and chance events may also exert strong influences on diversification of the phenotype. Past evolutionary events can produce preadaptations and exaptations (traits later co-opted for a novel purpose), but can also present constraints on future modification and enables chance events to create new populations that are maladapted to their environments (Gould and Lewontin 1979). In addition, many natural populations are effectively small, subdivided, and ephemeral; conditions that may favor chance sampling of the phenotypic landscape (Wade and Goodnight 1998), but which may also limit the effectiveness of selection (Willi et al. 2006). While the influence of history, chance, and adaptation are conceptually well appreciated (Gould 2002), there remain comparatively few empirical works that concurrently address these different aspects of the evolutionary process (Travisano et al. 1995, Losos et al. 1998, Taylor and McPhail 2000).

Over the last decade, introduced species have increasingly come to represent for biologists both a headache and an opportunity. Some aggressive invaders are negatively impacting native biota, and are forecasted to exacerbate the loss of biodiversity accompanying global change (Mooney and Cleland 2001). But other less virulent species provide intriguing examples of long-distance (often trans-oceanic) dispersal, population establishment in novel environments with different sets of interacting species, and

eventually the transition to demographic proliferation and range expansion. These events must surely impact introduced populations at a both ecological and genetic levels. For these reasons, species invasions have become recognized as promising phenomena for observing the evolutionary process (Baker and Stebbins 1965, Barrett 2000, Ellstrand and Schierenbeck 2000, Sakai et al. 2001, Lee 2002, Lambrinos 2004, Wares et al. 2005).

In particular, invasions are ideal testing grounds for observing the impacts of history, chance, and adaptation on the evolutionary process. The native range of an introduced species will often contain structured populations that harbor unique evolutionary histories (Meekins et al. 2001, Gaskin and Schaal 2002, Kolbe et al. 2004, Ostrowski et al. 2006, Chapter 3). Chance will tend to interact with history via the sampling process, for example by introducing a biased sample of diversity to the new range (Taylor and Keller 2007). The strength and effectiveness of selection following introduction will depend on the diversity of the founders and the extent to which their ecological niche is a preadaptation to the introduced range (Holt et al. 2005a).

Introduced species have already been recognized as model systems for studying the contemporary evolution of phenotypic traits (Thompson 1998, Reznick and Ghalambor 2001, Garcia-Ramos and Rodriguez 2002, Stockwell et al. 2003). Genetically-based differences in phenotypes between the native and introduced ranges (Blair and Wolfe 2004, Bossdorf et al. 2005, Lavergne and Molofsky 2007), and clinal variation along environmental gradients within the introduced range (Johnston and Selander 1964, Weber and Schmid 1998, Huey et al. 2000, Kollmann and Banuelos 2004, Maron et al. 2004, Leger and Rice 2007), all argue for evolutionary changes to the phenotype as a result of invasion. But it is often unclear what changes are attributable to adaptation *per se*, apart

from history and chance. Because the stochastic nature of colonization often creates genetic divergence among populations (Wade and McCauley 1988, McCauley et al. 1995, Whitlock 1999), invasions should be particularly prone to sampling effects that create neutral divergence in phenotypic traits (Chapter 1). In fact, the few published studies that have tested for neutral phenotypic evolution during invasion have concluded that it was of primary importance (Eckert et al. 1996, Kliber and Eckert 2005, Kolbe et al. 2007).

In this study, we make use of detailed knowledge about the evolutionary history of *Silene vulgaris*, a weedy plant native to Europe, to examine the causes of phenotypic evolution in its introduced range of North America. We consider history (the structuring of lineages in Europe and their preadaptation to North American environments), chance (the sampling of European lineages and their associated phenotypes), and adaptation (post-invasion divergence away from the ancestral phenotype) as unique contributors to the evolutionary changes that have accrued since its introduction.

Methods

Study Species and Seed Collections

Silene vulgaris L. (Moenche) Garke is a short-lived perennial plant in the family Caryophyllaceae. It is native to Eurasia, where it occupies a widespread geographic range (Chapter 4). *S. vulgaris* is believed to have been restricted to Mediterranean refugia during the glacial advances of the Pleistocene, and has since recolonized most of Europe in a post-glacial expansion. It is especially frequent along roadsides, railroads, orchards, and hayfields, but populations can also be found along river banks, forest edges, cliff

faces, and talus slopes in the mountains (Marsden-Jones and Turrill 1957). The seeds possess no specialized mechanism for dispersal and germinate exclusively in spring (Marsden-Jones and Turrill 1930). Reproduction is principally by sexual means, though fragmentation of belowground meristems can occasionally result in vegetative reproduction (Lubke and Cavers 1970).

The introduction phase of *S. vulgaris* to North America first began approximately 200 years ago. Plants were initially observed around the docks and nearby agricultural areas of the port cities of Boston and Philadelphia, and in hayfields outside of Quebec City (Cutler 1785, Pursh 1814, Martindale 1876). Following an extended lag period up to the late 1800's, *S. vulgaris* experienced a rapid range expansion throughout much of temperate North America (Chapter 4) and has since become a common member of the North American flora.

Seed collections were obtained from 36 European localities and 23 North American localities that are a subset of the locations analyzed in a previous genetic study of AFLP markers (Chapter 3). Collection localities were always > 5km distant from each other, and were selected haphazardly across the breadth of the native and introduced ranges (Figure 1). At each locality, we collected open-pollinated seed capsules from up to 7 maternal plants, and dispersed our collecting within sites so as to avoid over-sampling genetic neighborhoods. Seed families from each plant all shared a maternal parent, and may or may not have shared a paternal parent; thus they represent families of full and potentially half sib relatedness. In all, 100 families from North America and 100 families from Europe were initially included in the experiment (averaging 3-4 families/population)..

Common Garden Field Experiment

Plants were grown in common gardens within North America to assess plant performance under field conditions. While ideally, performance would be evaluated at many common gardens sites, logistical considerations prevented this. Instead, we chose to grow plants at two sites that encompass the range in environmental conditions inhabited by the species. This should reveal any genotype*environment interactions present for growth traits, and help control for plastic responses when inferring adaptive divergence. One site "ON" was located in Ontario, Canada (N 45.8642, W -79.4362) and another "VA" in Virginia, U.S.A. (N 37.8577, W -78.8208; Figure 2). The ON site was located in a farm field no longer under cultivation. A small population of *S. vulgaris* (< 5 plants) grew in the field and was destroyed prior to planting. Natural populations were located ca. 1 km away on the roadside. The VA site was located in a hayfield near the Rockfish River, Nelson County. There were no *S. vulgaris* previously growing in the field, but natural populations existed along the roadsides and adjacent hayfields ca. 1 km away. These two sites are in strategic locations for several reasons: (1) they are near the observed latitudinal limits of *S. vulgaris* in eastern North America, (2) they provide an overall contrast in climatic growing conditions, and (3) they are located in different zones of the predicted distribution in North America, as defined by ecological niche models conditioned on different native range lineages (Figure 2; Chapter 4). Thus, comparing traits between the two common gardens should provide an effective means for evaluating how sensitive our inference of past evolutionary change in *S. vulgaris* is to different environmental contexts.

Six weeks prior to planting at each site, seeds were surface sown in a completely randomized design into plug trays filled with a standard potting mix (Promix HP). Trays were placed into one of two Percival growth chambers at the University of Virginia set to a diurnal photoperiod of 12:12 (light:dark) and a temperature cycle of 21 °C and 12 °C during light and dark phases, respectively. Trays were checked daily, misted as necessary, and rotated within chambers to reduce position effects. Germination was scored when the radicle had penetrated the seed coat. Four seedlings per family were randomly chosen for planting into the field (total N = 2 sites x 2 continents of origin x 100 families/cont x 4 plants/family = 1600 plants). In addition, a full replicate of seeds were germinated to replace transplant mortality after initial planting.

Seedlings were transplanted into the field during spring 2005. We determined the timing of transplanting relative to the growing season at each site based on the number of growing degree days > 5 °C (gdd). Planting times were determined such that on average ~85 % of the gdd were left in the season at each site (based on climate normals from the Southeast Regional Climate Center (VA) and Environment Canada (ON)). Planting dates were April 29, 2005 at VA and May 24, 2005 at ON. The common gardens consisted of four rectangular subplots (blocks), each containing 200 plants. Plants were spaced 0.5 meters apart within rows, and 0.75 m apart between rows. Prior to planting, the existing vegetation was cut back to ground level, and individual holes ca. 15 cm in diameter were created with a post-hole digger. Seedlings were planted directly into the ground in these holes and watered in. Seedlings lost due to transplant mortality that occurred up to 2 weeks post-planting were replaced with a replicate seedling from the same family at the same stage of growth. Plants were watered as needed during early summer 2005 at ON to

offset drought conditions when plants were young. The existing vegetation was cut back periodically at each site to avoid competition for light.

Weekly censuses were conducted throughout the 2005 growing season and again in 2006. During each census, plants were checked for survival and reproductive status (vegetative or flowering). The week a plant first became reproductive during each season, we recorded the julian day of flowering, height of the tallest stem (cm), the number of primary stems, the number of leaves > 1 cm long, and the number of open flowers. During each subsequent census, we counted the number of new open flowers and the number of mature fruits. Flowers were marked on their calyces with a small dot from a liquid paint pen to avoid recounting in subsequent censuses. Mature fruits were harvested prior to dehiscence to avoid seed contamination at the sites. The sum of fruit production across years, which integrates both survivorship and fecundity, was used as an index of fitness.

Genetic Assignments

To obtain a robust test for adaptive evolution during invasion, we needed to account for phenotypic differences due to the chance sampling of native range lineages that differ in phenotype. We combined our phenotypic observations from the common gardens with estimates of membership to different historical lineages present in the native range. We used amplified fragment length polymorphisms (AFLP) from a previous study to examine genetic variation in the native and introduced ranges and to trace North American genotypes back to their ancestral origins in Europe (Chapter 3). These analyses included a representative from all 200 families in the current study.

Two methods were used to estimate genetic similarity among genotypes: Bayesian model-based clustering (STRUCTURE) and principal coordinates analysis (PCoA). Using STRUCTURE v2.2 (Pritchard et al. 2000, Pritchard et al. 2007), the native range was found to be subdivided into five geographically and genetically divergent demes. Four of these demes contained genotypes introduced to North America (deme 1, deme 2.2, deme 2.3, and deme 3; Figure 2). However, the demes in North America occur in significantly different frequencies compared to Europe ($G = 46.52$, d.f. = 3, $P < 0.0001$; Chapter 3), suggesting a founder effect during invasion. We used the assignment probabilities from the STRUCTURE analysis to assign families used in the field experiment to demes and to test for phenotypic differences among the demes that would contribute to founder effects on the phenotype. In addition to STRUCTURE, we used the first two axes from a principal coordinates analysis (PCoA) to ordinate each family in continuous multi-locus genotype space. The deme assignments and PC scores were combined with the phenotypic measurements obtained from the common gardens to generate null expectations for phenotypic divergence due to founder effects (Chapter 1; see below).

Statistical Analyses

Analyses of phenotypic traits measured during 2005 and 2006 were performed using repeated measures mixed-model analysis of variance (PROC MIXED: SAS Institute 1999). Continent of origin (= cont), deme, and common garden (= site) were fixed effects; family (nested within cont, deme) and subplot (nested within site) were random effects. Random effects were tested against a χ^2 distribution by specifying the "covtest" option in the PROC statement. Individual plants were the subjects measured over years.

Population of origin was not analyzed in the models because our design explicitly sought to replicate each range with many families from across many different collection sites, at the expense of replication within sites. Two of the families were found to be *S. latifolia* (both European), and so were excluded from all analyses. Traits were either left untransformed (juvenile number of leaves, stem number at flowering), natural log transformed (time to germination, date of flowering, number of leaves at flowering, number of flowers, flowering period, number of fruits), or square root transformed (height) prior to analysis to improve normality and homoscedasticity.

Analyses on the discrete traits survivorship and probability of flowering (assessed cumulatively across years) were performed using a general linear model (GLM) assuming a binomial distribution of errors and a logit link (PROC GENMOD: SAS Institute 1999). GLM models were of the same structure as ANOVA, except for the exclusion of family(cont, deme) after preliminary analyses showed models with this effect did not converge.

By including continent and ancestral deme effects in all statistical models, we tested for post-invasion evolutionary change (continent effect) while controlling for phenotypic shifts due to founder effects (deme effect)(see Chapter 1). Specifically, we interpreted a significant effect of deme to indicate that historical isolation in the native range affected divergence in phenotypic traits as well as neutral genetic diversity. A significant effect of continent, after controlling for phenotypic differences among demes, indicates divergence between ranges has occurred when deme effects are removed, that is, divergence in excess of the null expectation based on the relative frequencies of demes and their associated phenotypes. Finally, the continent*deme interaction indicates that post-

invasion phenotypic evolution is a deme-specific process. This can potentially arise from either founder effects or genetic drift occurring within some demes and not others, or from deme-specific patterns of response to selection. Previous analyses showed that demes occur in North America essentially intact, with no evidence for founder effects within demes during or since the introduction (Chapter 3). Thus, a significant effect of continent or continent*deme provides evidence in support of adaptive evolution at some point during the invasion process. Interactions with site (e.g., cont*site, deme*site, cont*deme*site) tested whether observing genetic divergence was wholly or partially contingent on the influence of the environment on trait expression.

In addition to detecting the presence of evolutionary change since invasion, we also determined the rate of phenotypic divergence between the native and introduced ranges (Hendry and Kinnison 1999, Bone and Farres 2001, Stockwell et al. 2003). Estimating rates between the native and introduced ranges is a synchronic design, because we are comparing two extant populations that shared a common ancestor some time in the past. Therefore, our rates are expressed correctly as divergences, since they incorporate evolution occurring within both ranges, and not a time series of evolutionary change within a single population. We calculated rates in units of darwins and haldanes (Hendry and Kinnison 1999). The darwin measures the proportional rate of divergence in phenotypic means between populations over absolute time: $[\ln(x_1) - \ln(x_2)] / t$, where x_1 and x_2 are the phenotypic means in two populations, and t is the time since separation in units of 10^6 years. The haldane measures divergence in units of standard deviations per generation: $[(x_1/s_p) - (x_2/s_p)] / g$, where x_1 and x_2 are the population means as before, s_p is the pooled standard deviation, and g is the number of generations since separation (years

/ generation time). We used two estimates of *t*, one based on the earliest published record of *S. vulgaris* in North America (Cutler 1785), and another based on when vouchered specimens of *S. vulgaris* began appearing in herbaria (ca. 1850; Chapter 4). We estimated generation time by using observations of age at first reproduction from our common gardens. Since *S. vulgaris* is a short-lived perennial, some plants were still alive at the end of our two year study but had yet to reproduce. To incorporate these plants into our estimate of the generation time, we projected the life history to obtain approximate estimates of first reproduction for subsequent age classes. We calculated the expected number of plants becoming reproductive for each future age class by multiplying the number of vegetative plants alive at the end of age x by the product of the probabilities of surviving to age x+1 (estimated from our data to be 0.81) and becoming reproductive (estimated from our data to be 0.48). Note that this assumes that the transition probabilities do not vary with age. We projected our experimental population until all individuals had either died or reproduced. This gave us an approximate distribution for the age at first reproduction across the life history, from which we took the mean value for our estimate of generation time (= 1.5 years).

We tested for phenotypic diversification within ranges by assessing clinal variation between phenotypes and latitude using linear regression. Trait means were calculated initially over years for each plant, and then over plants for each family. If lineages assorted with latitude, then clinal variation could be due to founder effects, but significance of latitude, after controlling for lineage effects, is evidence for adaptive evolution along one or more environmental gradients. We tested for adaptive clines against a null expectation of founder effects by including as predictor variables the first

two axes from the PCoA of the AFLP data along with the latitude of collection in a multiple regression (PROC REG: SAS Institute 1999). Regression models were formulated separately for each common garden site and continent of origin. Significance levels were assessed following a table-wide Bonferroni correction ($\alpha = 0.05 / 24$ tests = 0.002).

Finally, we tested whether an evolutionary history of ecological niche divergence among demes in Europe predisposed those lineages to different environments within North America. In a previous study, we produced ecological niche models separately for each deme, conditioned on deme occurrence in Europe, and used these models to predict potential area of occurrence in North America (Chapter 4). The ON site was found to be within the niche predicted for deme 2.2, but outside the niche of the other three demes (Figure 2). Likewise, the VA site was within the niche of demes 1, 2.3, and 3, but not deme 2.2. We used this pattern to generate predictions for relative fitness at our two common garden sites. Specifically, we predicted that for preadaptation to be important, European genotypes should have higher relative fitness when grown at a site within versus outside the ancestral niche. Guided by this expectation, each European family was categorized at the two common garden sites as growing inside or outside of its niche. We then analyzed differences in relative fitness among niche status, demes, and the niche*deme interaction using repeated measures ANOVA on family means, with family as the repeated subject. Relative fitness was defined as the mean cumulative fruit set per family, divided by the mean across families for each site. Only European families were used for this analysis because we wished to assess the potential importance of

preadaptation, without the confounding effects of evolutionary changes that have occurred since the invasion.

Results

Evolutionary changes have accrued since the introduction of *S. vulgaris* to North America. Continent of origin was a significant predictor of phenotypic variation for 8 out of 11 traits (Table 1). Genetically-based phenotypic divergence between continents was evident for both growth related traits (Figure 3) and reproductive traits (Figure 4). In comparison with European genotypes, North American genotypes germinated faster, had a higher probability of flowering, had fewer but taller stems at reproductive maturity, produced more flowers and fruits over a longer reproductive period, and had greater survivorship (Table 2). These differences in growth, phenology, and reproductive effort were not due to chance events such as founder effects, as phenotypic differences among demes were controlled for in the models. Rather, the results strongly support a history of adaptive evolution at some stage in the establishment and spread of *S. vulgaris* in North America. Rates of phenotypic divergence between continents ranged from 0.02×10^{-3} darwins for leaf number to 4.43×10^{-3} darwins for flower number (Table 3). When scaled by the generation time, rates ranged from 0.0001 haldanes for stem number to 0.009 haldanes for germination time. The largest rates we observed were within the range reported for other introduced plants (Table 3).

In addition to differences between continents, mean differences among demes were also highly significant for 6 of the 11 traits, suggesting that the history of neutral genetic divergence in Europe prior to invasion had also been accompanied by divergence in

phenotypic traits (Table 1). There was little indication that demes had experienced different responses to selection since invasion (i.e., no cont*deme interaction; Table 1). Differences among demes were generally of similar or reduced magnitude compared to differences between continents (Table 1; Figures 3 and 4).

The two common garden sites had a large impact on plant performance (Table 1). However, evidence for context dependency on the phenotypic differences between continents or demes was scant (Table 1). One exception was julian day of flowering, which showed that while European demes exhibited variability in flowering time across sites, North American demes converged to a common flowering times within each site (Figure 4).

Examining divergence within ranges, we observed significant phenotypic clines with latitude among both European and North American families (Table 4). Phenotypic clines with latitude were significant after controlling for the combined effects of evolutionary history and chance sampling (using the first two PCoA axes), strongly suggesting adaptive diversification within each range in response to one or more environmental gradients. Interestingly, the direction of the clines was typically reversed between continents (Table 4, Figure 5). Among European families, plants from northern latitudes typically grew larger and flowered longer compared to those from southern latitudes, while North American families showed the opposite trend. Only the time to flowering was consistently later among northern latitude families from both continents.

European families showed no evidence for preadaptation when grown at sites within the ecological niche versus outside the niche ($F_{1,87} = 0.99$, $P = 0.32$; Figure 6). Relative fitness within sites was also not affected by differences among demes ($F_{3,92} = 1.11$, $P =$

0.35) or by differences among demes in response to niche matching ($F_{3,87}$ = 0.73, P = 0.54)

Discussion

In the 200 years since being introduced by humans to North America, *Silene vulgaris* has undergone a suite of genetically-based phenotypic changes suggesting the action of natural selection. On average, North American genotypes possess a faster phenology, an allocation pattern of fewer more dominant stems, and increased reproductive output within the first two years of growth compared to European genotypes. These differences are consistent with theoretical studies of selection on colonizing ability, which predict a life history shift towards increased early life reproduction (Lewontin 1965). The enhanced survivorship of North American versus European genotypes when grown in North America may point to the presence of local adaptation to conditions within the introduced range. However, it is often difficult to know whether phenotypic divergence during invasion is due to adaptation *per se*, or chance events sampling among divergent native range populations (Chapter 1). Indeed, we found that the different genetic demes we previously identified often possessed different mean phenotypes. In conjunction with the observed shift in deme frequencies (Chapter 3), this would suggest that some component of the uncorrected phenotypic differences between European and North American genotypes must be attributable to founder effects. However, in this study we controlled for neutral phenotypic evolution by isolating the contribution that variance among demes makes to the distribution of phenotypes within each range. After

accounting for neutral expectations, we found widespread evidence for phenotypic change attributable to natural selection.

Given out best estimate for the age of the invasion (ca. 150 - 215 years), the rate of divergence for most traits was within the range observed for other introduced plant species, which were themselves towards the slow end of the range observed across plant taxa (Bone and Farres 2001). This raises the question of whether the "rapid" evolutionary changes being reported for many introduced species are really very rapid when placed in the context of a distribution of evolutionary rates (Hendry and Kinnison 1999). Rather than being perceived as exceptional events, species such as *S. vulgaris* may represent a very general phenomenon of adaptive evolution during invasion, reinforcing the notion that evolutionary processes are occurring continuously over ecological time scales (Antonovics 1976a).

In addition to the overall divergence observed between ranges, we found evidence for adaptive diversification within the introduced range in the form of latitudinal clines. These clines were significant after controlling for lineage effects, and strongly point to adaptive diversification within ranges along environmental gradients. Previous studies have also reported phenotypic clines within the introduced ranges of invasive species (Weber and Schmid 1998, Huey et al. 2000, Kollmann and Banuelos 2004, Maron et al. 2004, Leger and Rice 2007). However, ours is the first to statistically incorporate a null expectation for clinal variation in phenotypes.

When repeatability across ranges has been assessed, most studies report the development of parallel clines (Huey et al. 2000, Maron et al. 2004, Leger and Rice 2007). In our study, we found clines in North America that were in the opposite direction

of the European clines for multiple traits. This may be a function of the growing season *S. vulgaris* experiences in each range. In Europe, the southern distribution of *S. vulgaris* is in the Mediterranean region where hot dry summers likely begin the growing season early, but restrict flowering to late spring and early summer. At more northerly European latitudes, the season starts later but favorable growing conditions extend for a longer period. In contrast, the southern distribution of *S. vulgaris* in North America experiences a growing season that starts earlier and lasts longer than conditions found at more northerly latitudes. These qualitative differences in climate between the continents may be responsible for the observed reversal in clines. For example, the probability that a plant flowers versus remaining vegetative is a trait that may reflect the width of opportunity for successful reproduction at a given site. Consistent with this idea, European plants were more likely to flower within their first two years if they came from higher latitudes, whereas plants from southerly latitudes in North America were more likely to become reproductive (Table 3).

Given the presence of clinal variation within the native range, one expectation for new invasions is that individuals should have relatively higher fitness when experiencing environments more like their region of origin, i.e., there should be a detectable signal of preadaptation during the establishment phase of invasion. We tested the importance of matching the ecological niche to preadaptation by growing European genotypes inside and outside of the niche predicted by their deme. Contrary to expectation, we found no evidence for preadaptation affecting relative fitness in *S. vulgaris*. On average, families grown in a common garden outside their predicted niche performed as well as families grown within their niche. To our knowledge, only one other study has explicitly tested

the role of preadaptation in colonization success (Maron 2006). It is noteworthy that Maron's study of *Hypericum perforatum* (St. John's wort) led to conclusions broadly similar to our own: despite strong evidence for local adaptation across the native range, the environmental match between source populations and the common gardens mattered little to establishment success or plant fitness. However, both studies were conducted over a 2 year period. Selection imposed by temporal variance (such as unpredictable frosts, or variance among years in drought conditions) may be important in determining the long-term persistence of populations outside their niche (Holt et al. 2005b), but may also go unobserved over the period of most field experiments. Another possibility relates to the basis for defining an expectation for preadaptation. We used deme occurrence in Europe to generate ecological niche models based on multiple climatic variables at the locations where genotypes were sampled. However, it is possible that niche predictions developed from geographical occurrences largely reflect historical contingency (for ex., dispersal limitation) and not functional constraints imposed by the fundamental niche (Svenning and Skov 2007). Thus, while ecological niche models can be useful for predicting areas susceptible to invasion (Peterson and Vieglais 2001, Mau-Crimmins et al. 2006), they should be regarded as conservative predictions until experimental tests can be conducted.

History and chance are sources of variance in the predictability of the evolutionary process. Despite the utility of natural selection at producing traits that are functionally well matched to their environment, chance events such as founder effects may distribute phenotypes into environments to which they have no history of adaptation. Biological invasions should be especially subject to these non-selective forces, by nature of the fact

that species are often subdivided in their native ranges and sampling processes during invasion are largely stochastic. Yet our data suggest that selection has been the dominant evolutionary force driving phenotypic diversification of *S. vulgaris* since its introduction to North America. While a response to selection in North America seems evident, questions regarding the agents of selection, the magnitude of the selection episodes, and the potential for correlated responses among traits remain unanswered. Future work will begin to address these issues by examining paths to fitness that emerge from selection on individual traits mediated through the life history.

Acknowledgements

We thank the many people who helped with the setup and data collection on the common gardens, including C. Barr, R. Carr, W. Farnum, S. Freedberg, C. Jenkins, C. Keller, J. Keller, K. Keller, M. Neiman, A. Neville, R. Neville, T. Poole, V. Panjeti, N. Priest, D. Sloan, and D. Sowell. We are especially grateful to the landowners N. Rothwell and C. Smith in VA, and M. and R. Neville in ON for allowing us to conduct these experiments. Lastly, we acknowledge the many colleagues who have contributed to our seed collection over the years. This work was supported by NSF DEB-0608358 Doctoral Dissertation Improvement Grant and NSF DEB-444176.

References

Antonovics, J. 1976. The input from population genetics: "the new ecological genetics". Systematic Botany **1**:233-245.

Baker, H. G., and G. L. Stebbins. 1965. The genetics of colonizing species. Academic Press, New York.

Barrett, S. C. H. 2000. Microevolutionary influences of global changes on plant invasions. Pages 115-139 *in* H. A. Mooney and R. Hobbs, editors. Invasive species in a changing world. Island Press, Covelo, CA.

Blair, A. C., and L. M. Wolfe. 2004. The evolution of an invasive plant: an experimental study with *Silene latifolia*. Ecology **85**:3035-3042.

Bone, E., and A. Farres. 2001. Trends and rates of microevolution in plants. Genetica **112-113**:165-182.

Bossdorf, O., H. Auge, L. Lafuma, W. E. Rogers, E. Siemann, and D. Prati. 2005. Phenotypic and genetic differentiation between native and introduced plant populations. Oecologia **144**:1-11.

Coyne, J. A., N. H. Barton, and M. Turelli. 1997. Perspective: A critique of Sewall Wright's shifting balance theory of evolution. Evolution **51**:643-671.

Coyne, J. A., N. H. Barton, and M. Turelli. 2000. Is Wright's shifting balance process important in evolution? Evolution **54**:306-317.

Cutler, M. 1785. An account of some of the vegetable productions, naturally growing in this part of America, botanically arranged. Memoirs of the American Academy of Arts and Sciences **I**:396-493.

Eckert, C. G., D. Manicacci, and S. C. H. Barrett. 1996. Genetic drift and founder effect in native versus introduced populations of an invading plant, *Lythrum salicaria* (Lythraceae). Evolution **50**:1512-1519.

Ellstrand, N. C., and K. A. Schierenbeck. 2000. Hybridization as a stimulus for the evolution of invasiveness. Proceedings of the National Academy of Sciences of the United States of America **97**:7043-7050.

Garcia-Ramos, G., and D. Rodriguez. 2002. Evolutionary speed of species invasions. Evolution **56**:661-668.

Gaskin, J. F., and B. A. Schaal. 2002. Hybrid *Tamarix* widespread in U.S. invasion and undetected in native Asian range. PNAS **99**:11256-11259.

Goodnight, C. J., and M. J. Wade. 2000. The ongoing synthesis: a reply to Coyne, Barton, and Turelli. Evolution **54**:317-324.

Gould, S. J. 2002. The structure of evolutionary theory. Belknap Press.

Gould, S. J., and R. C. Lewontin. 1979. The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme. Proceedings of the Royal Society of London Series B **205**:581-598.

Hendry, A. P., and M. T. Kinnison. 1999. Perspective: The pace of modern life: measuring rates of contemporary microevolution. 53 **6**:1637-1653.

Holt, R. D., M. Barfield, and R. Gomulkiewicz. 2005a. Theories of niche conservatism and evolution: Could exotic species be potential tests? Pages 259-290 *in* D. F. Sax, J. J. Stachowicz, and S. D. Gaines, editors. Species invasions: insights into ecology, evolution, and biogeography. Sinauer Associates, Inc., Sunderland, MA.

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

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

Johnston, R. F., and R. K. Selander. 1964. House sparrows: rapid evolution of races in North America. Science **144**:548-550.

Kliber, A., and C. G. Eckert. 2005. Interaction between founder effect and selection during biological invasion in an aquatic plant. Evolution **59**:1900-1913.

Kolbe, J. J., R. E. Glor, L. R. Schettino, A. C. Lara, A. Larson, and J. B. Losos. 2004. Genetic variation increases during biological invasion by a Cuban lizard. Nature **431**:177-181.

Kolbe, J. J., A. Larson, and J. B. Losos. 2007. Differential admixture shapes morphological variation among invasive populations of the lizard *Anolis sagrei*. Molecular Ecology **16**:1579-1591.

Kollmann, J., and M. J. Banuelos. 2004. Latitudinal trends in growth and phenology of the invasive alien plant *Impatiens glandulifera* (Balsaminaceae). Diversity and Distributions **10**:377-385.

Lambrinos, J. G. 2004. How interactions between ecology and evolution influence contemporary invasion dynamics. Ecology **85**:2061-2070.

Lavergne, S., and J. Molofsky. 2007. Increased genetic variation and evolutionary potential drive the success of an invasive grass. PNAS **104**:3883-3888.

Lee, C. E. 2002. Evolutionary genetics of invasive species. Trends in Ecology and Evolution **17**:386-391.

Leger, E. A., and K. J. Rice. 2007. Assessing the speed and predictability of local adaptation in invasive California poppies (*Eschscholzia californica*). Journal of Evolutionary Biology **20**:1090-1103.

Lewontin, R. C. 1965. Selection for colonizing ability. *in* H. G. Baker and G. L. Stebbins, editors. The genetics of colonizing species. Academic Press, New York.

Losos, J. B., T. R. Jackman, A. Larson, K. de Queiroz, and L. Rodriguez-Schettino. 1998. Contingency and determinism in replicated adaptive radiations of island lizards. Science **279**:2115-2118.

Lubke, M., and P. B. Cavers. 1970. Studies of vegetative regeneration in *Saponaria officinalis* L. (Soapwort) and *Silene cucubalus* (Bladder Campion). The Canadian Field-Naturalist **84**:43-47.

Maron, J. L. 2006. The relative importance of latitude matching and propagule pressure in the colonization success of an invasive forb. Ecography **29**:819-826.

Maron, J. L., M. Vila, R. Bommarco, S. Elmendorf, and P. Beardsley. 2004. Rapid evolution of an invasive plant. Ecological Monographs **74**:261-280.

Marsden-Jones, E. M., and W. B. Turrill. 1957. The bladder campions. Ray Society, London.

Martindale, I. C. 1876. The introduction of foreign plants. Botanical Gazette **2**:55-58.

Mau-Crimmins, T. M., H. R. Schussman, and E. L. Geiger. 2006. Can the invaded range of a species be predicted sufficiently using only native-range data?: Lehmann

lovegrass (*Eragrostis lehmanniana*) in the southwestern United States. Ecological Modelling **193**:736-746.

McCauley, D. E., J. Raveill, and J. Antonovics. 1995. Local founding events as determinants of genetic structure in a plant metapopulation. Heredity **75**:630-636.

Meekins, J. F., H. E. Ballard, and B. C. McCarthy. 2001. Genetic variation and molecular biogeography of a North American invasive plant species (*Alliaria petiolata*) Brassicaceae. International Journal of Plant Sciences **162**:161-169.

Mooney, H. A., and E. E. Cleland. 2001. The evolutionary impact of invasive species. Proceedings of the National Academy of Sciences of the United States of America **98**:5446-5451.

Ostrowski, M.-F., J. Davis, S. Santoni, H. Mckhann, X. Reboud, V. Le Corre, C. Camilleri, D. Brunel, D. Bouchez, B. Faure, and T. Bataillon. 2006. Evidence for a large-scale population structure among accessions of Arabidopsis thaliana: possible causes and consequences for the distribution of linkage disequilibrium. Molecular Ecology **15**:1507-1517.

Peterson, A. T., and D. A. Vieglais. 2001. Predicting species invasions using ecologic niche modeling. BioScience **51**:363-371.

Pritchard, J. K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. Genetics **155**:945-959.

Pritchard, J. K., X. Wen, and D. Falush. 2007. Documentation for *structure* software: version 2.2.

Provine, W. B. 1971. The origins of theoretical population genetics. University of Chicago Press, Chicago.

Pursh, F. 1814. Flora Americae Septentrionalis. White, London.

Reznick, D. N., and C. K. Ghalambor. 2001. The population ecology of contemporary adaptations: what empirical studies reveal about the conditions that promote adaptive evolution. Genetica **112-113**:183-198.

Sakai, A. K., F. W. Allendorf, J. S. Holt, D. M. Lodge, J. Molofsky, K. A. With, S. Baughman, R. J. Cabin, J. E. Cohen, N. C. Ellstrand, D. E. McCauley, P. O'Neil, I. M. Parker, J. N. Thompson, and S. G. Weller. 2001. The population biology of invasive species. Annual Review of Ecology and Systematics **32**:305-332.

SAS Institute. 1999. SAS/STAT version 8. SAS Institute, Cary, North Carolina, U.S.A.

Stockwell, C. A., A. P. Hendry, and M. T. Kinnison. 2003. Contemporary evolution meets conservation biology. Trends in Ecology and Evolution **18**:94-101.

Svenning, J.-C., and F. Skov. 2007. Could the tree diversity pattern in Europe be generated by postglacial dispersal limitation? Ecology Letters **10**:453-460.

Taylor, D. R., and S. R. Keller. 2007. Historical range expansion determines the phylogenetic diversity introduced during contemporary species invasion. Evolution **61**:334-345.

Taylor, E. B., and J. D. McPhail. 2000. Historical contingency and ecological determinism interact to prime speciation in sticklebacks, *Gasterosteus*. Proceedings of the Royal Society of London Series B **267**:2375-2384.

Thompson, J. N. 1998. Rapid evolution as an ecological process. Trends in Ecology and Evolution **13**:329-332.

Travisano, M., J. A. Mogols, A. F. Bennet, and R. E. Lenski. 1995. Experimental tests of the roles of adaptation, chance, and history in evolution. Science **267**:87-90.

Wade, M. J., and C. J. Goodnight. 1998. Perspective: The theories of Fisher and Wright in the context of metapopulations: when nature does many small experiments. Evolution **52**:1537-1553.

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

Wares, J. P., A. R. Hughes, and R. K. Grosberg. 2005. Mechanisms that drive evolutionary change: insights from species introductions and invasions. *in* D. F. Sax, J. J. Stachowicz, and S. D. Gaines, editors. Species invasions: insights into ecology, evolution, and biogeography. Sinauer Associates, Inc., Sunderland, MA.

Weber, E., and B. Schmid. 1998. Latitudinal population differentiation in two species of *Solidago* (Asteraceae) introduced to Europe. American Journal of Botany **85**:1110-1121.

Whitlock, M. C. 1999. Neutral additive genetic variance in a metapopulation. Genetical Resources **74**:215-221.

Willi, Y., J. Van Buskirk, and A. A. Hoffmann. 2006. Limits to the Adaptive Potential of Small Populations. Annual Review of Ecology, Evolution, and Systematics **37**:433-458.

Table 1. Summary of mixed-model anlaysis of variance (ANOVA) and general linear models (GLM) on plant performance. Germination time and juvenile size were measured once per plant, and hence have no effect of year. The block effect for germination time and juvenile size was planting cohort, and for all other traits was subplot(site). Family and block were random effects; all others were fixed effects. Significance levels are given from *F*-ratios (ANOVA) or likelihood ratio χ^2 tests (GLM). Degrees of freedom for ANOVA fixed effects (numerator, denominator) and GLM tests (χ^2 d.f.) are shown. † $P< 0.1$, * $P<0.05$, ** $P<0.01$, *** $P<0.001$, **** $P<0.0001$.

Effect	Germ time	Juv. leaf number	Prob. of flower	Julian flower	Stem height	Stem number	Leaf number	Flower number	Flower period	Fruit number	Survival
Continent	****	ns	****	†	**	*	ns	****	****	**	****
	(1,187)	(1,188)	(1)	(1,183)	(1,183)	(1,183)	(1,183)	(1,179)	(1,183)	(1,177)	(1)
Deme	***	***	****	ns	**	ns	ns	**	****	ns	†
	(3,187)	(3,188)	(3)	(3,183)	(3,183)	(3,183)	(3,183)	(3,179)	(3,183)	(3,177)	(3)
Site			*	***	*	**	**	*	ns	*	****
			(1)	(1,6)	(1,6)	(1,6)	(1,6)	(1,6)	(1,6)	(1,6)	(1)

Cont*deme	ns	†	ns	ns	ns	ns	ns	ns	ns	ns	ns
	(3,187)	(3,188)	(3)	(3,183)	(3,183)	(3,183)	(3,183)	(3,179)	(3,183)	(3,177)	(3)
Cont*site			ns	ns	ns	ns	ns	*	**	ns	ns
			(1)	(1,1509)	(1,1475)	(1,1475)	(1,1473)	(1,756)	(1,1509)	(1,1131)	(1)
Deme*site			ns	**	ns	ns	ns	ns	*	ns	ns
			(3)	(3,1509)	(3,1475)	(3,1475)	(3,1473)	(3,756)	(3,1509)	(3,1131)	(3)
Cont*deme*site			†	*	*	ns	ns	ns	ns	ns	*
			(3)	(3,1509)	(3,1475)	(3,1475)	(3,1473)	(3,756)	(3,1509)	(3,1131)	(3)
Year			1	****	****	****	****	2	**	****	1
				(1,1509)	(1,1475)	(1,1475)	(1,1473)		(1,1509)	(1,1131)	
Fam(cont,deme)	****	****		****	****	**	****	****	****	****	
Block	ns	ns	**	†	†	†	*	ns	†	†	**
			(6)								(6)

[1] Probability of flowering and survivorship were defined over the entire two year period.

[2] Flower number was only measured during 2005

Table 2. Means ± SE of phenotypic traits by continent and common garden site (ON = Ontario, Canada; VA = Virginia, U.S.A.). Means were first calculated over years for each plant before averaging by continent and site. Probability of flowering and survivorship report estimates (sample size).

	Europe ON	N. America ON	Europe VA	N. America VA
Germination time	7.0 ± 0.2	4.6 ± 0.1	9.6 ± 0.2	7.6 ± 0.2
Juvenile leaf number	7.8 ± 0.1	7.8 ± 0.1	8.2 ± 0.1	8.3 ± 0.1
Probability of flowering	0.70 (358)	0.60 (384)	0.81 (353)	0.82 (392)
Julian day of 1st flower	187.2 ± 0.9	190.1 ± 0.6	181.3 ± 1.1	182.7 ± 0.9
Stem height (cm)	35.72 ± 0.67	38.96 ± 0.57	36.90 ± 0.65	42.19 ± 0.55
Stem number	12.8 ± 0.5	11.6 ± 0.4	3.9 ± 0.1	3.9 ± 0.1
Adult leaf number	676.1 ± 59.0	681.0 ± 47.9	77.0 ± 4.3	98.2 ± 3.6
Flower number	12.8 ± 1.3	21.9 ± 1.9	13.3 ± 1.1	29.3 ± 1.6
Flower period (wks)	2.9 ± 0.1	3.7 ± 0.1	2.7 ± 0.1	4.1 ± 0.1
Fruit number	53.3 ± 7.3	101.4 ± 10.1	3.3 ± 0.3	7.4 ± 0.6
Survivorship	0.78 (358)	0.86 (384)	0.40 (353)	0.59 (392)

Table 2. Evolutionary rates for *Silene vulgaris* introduced to North America. Divergence from the ancestral population was estimated from the earliest recorded occurrence in publication (Cutler 1785) and the time of appearance in herbarium collections (1850). Generation time was taken to be 1.5 yrs (see Methods). Traits marked with (*) were significantly different between continents in ANOVA (Table 1).

Trait	Darwins (* 10^{-3})		Haldanes	
	1785-2000	1850-2000	1785-2000	1850-2000
Germination time *	-1.60	-2.29	-0.0063	-0.0090
Juvenile size	0.11	0.16	0.0012	0.0017
Time to flowering	0.42	0.60	0.0032	0.0046
Stem height *	-0.27	-0.39	-0.0008	-0.0012
Stem number *	0.05	0.08	0.0001	0.0002
Leaf number	0.02	0.04	0.0007	0.0009
Flower number *	3.09	4.43	0.0050	0.0072
Flower period *	1.43	2.04	0.0050	0.0072
Fruit number *	0.98	1.41	0.0016	0.0023
Mean absolute value for *S. vulgaris*	0.89	1.27	0.0027	0.0038
Mean absolute value(range) for other introduced plants †	5.07 (0.64 – 9.44)		0.157 (0.0021 – 0.8082)	

† Mean values taken from Bone and Farres (2001) who review estimated rates for 8 traits among 6 introduced plants.

Table 3. Clinal variation among European and North American plants grown in two common garden sites. Values are partial regression coefficients from multiple regressions of family trait means on latitude and the first two axes of a principal coordinates analysis of multilocus AFLP genotypes. Coefficients significant after Bonferroni correction are in bold.

	Europe			North America		
Ontario	Latitude	PC1	PC2	Latitude	PC1	PC2
Probability of flowering	**0.1288**	-0.8518	1.4323	**-0.1496**	-0.2116	0.5503
Time to flower	0.0007	-0.0781	0.0344	0.0017	0.0515	0.0049
Height at flowering	**0.0463**	0.1397	1.7295	**-0.0431**	-0.1505	0.2237
Number of stems at flowering	**0.5212**	-0.1448	4.1550	-0.0220	-5.2998	-7.6392
Number of leaves at flowering	**0.0728**	0.1209	0.0337	-0.0196	0.2521	-1.0119
Flowering period (wks)	**0.0275**	0.3125	0.3741	-0.0210	-0.3838	-0.0185
Virginia						
Probability of flowering	**0.0585**	0.1310	**3.1388**	**-0.1630**	1.1312	**2.6265**
Time to flower	**0.0073**	0.0465	0.1381	**0.0052**	-0.0066	-0.0378
Height at flowering	0.0065	-0.8246	1.5587	**-0.0792**	0.2988	1.0051
Number of stems at flowering	0.0745	7.2509	**-6.0334**	-0.1070	1.1411	1.2557

Number of leaves at flowering	**0.0555**	1.1854	-1.0721	**-0.0428**	0.3251	0.6920
Flowering period	**0.0421**	0.7927	0.7740	**-0.0517**	0.2164	**1.1163**

Figure Legends

Figure 1. Map showing the sampling locations for seed families used in this study. Each family was previously analyzed for its AFLP multi-locus genotype, and assigned membership to one of five demes using model-based clustering method (Chapter 3). Circle size is proportional to the number of families sampled at a given locality. For clarity, some symbols are offset from their true locations (shown with arrows).

Figure 2. Map showing the location of common garden sites (ON and VA) within the ecological niches predicted for each deme by their native range distribution (Chapter 4).

Figure 3. Norm of reaction for three morphological traits measured on the week a plant first flowered. Values are least-square means ± 1 SE for each combination of deme and continent of origin.

Figure 4. Norm of reaction for three reproductive traits. See Figure 3.

Figure 5. Clinal variation in flowering period in Europe and North America. Response values represent residuals form ANOVA after removing main and interaction effects due to site and deme.

Figure 6. Difference in relative fitness (w) for European families when grown inside and outside conditions matching their native range ecological niche. The x-axis is an index variable indicating family membership.

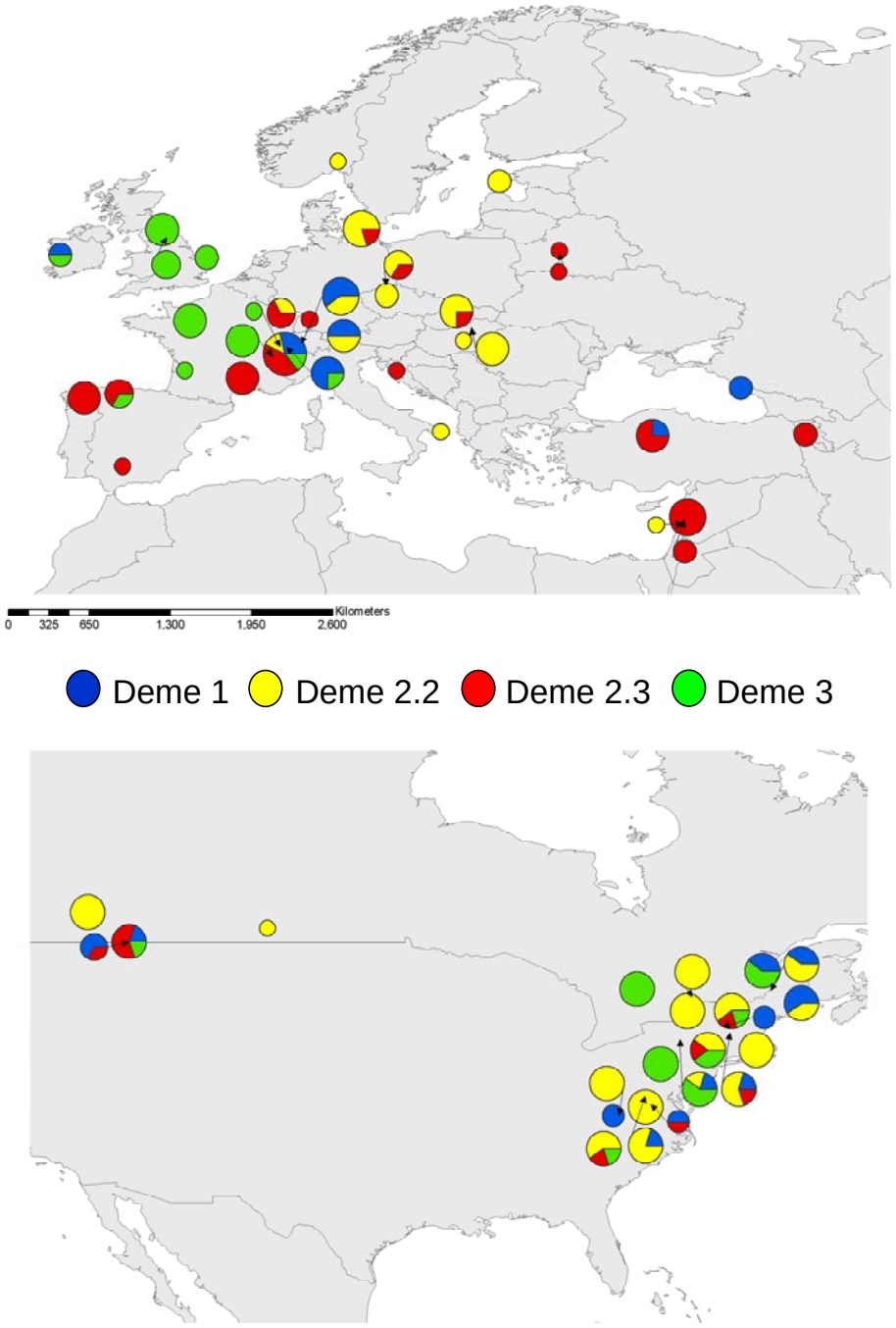

Kilometers
0 325 650 1.300 1.950 2.600
Deme 1
Deme 2.2
Deme 2.3
Deme 3

Figure 1.

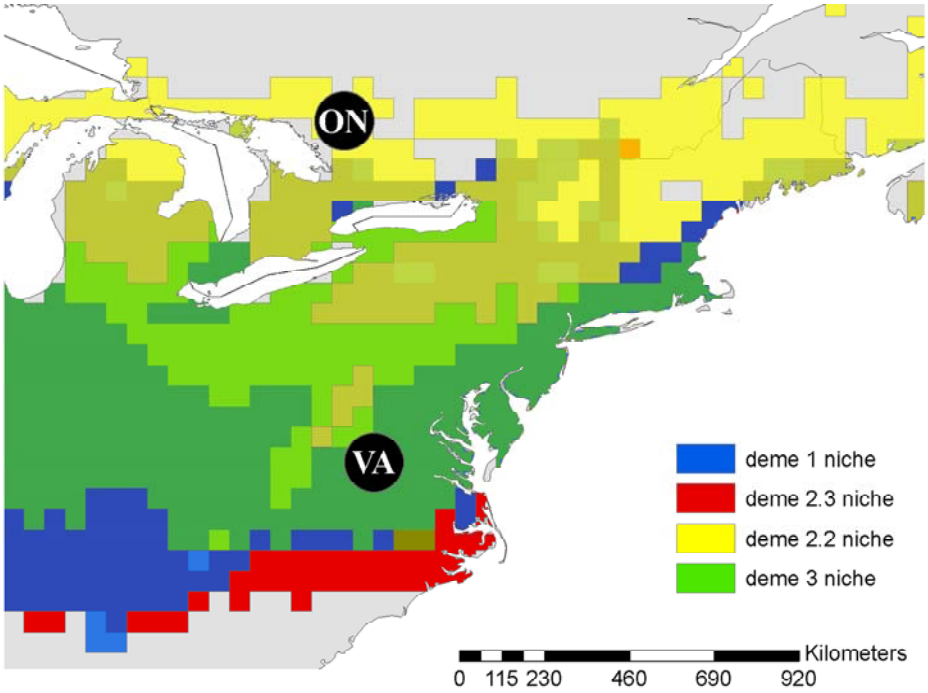

ON
VA
deme 1 niche
deme 2.3 niche
deme 2.2 niche
deme 3 niche
Kilometers
0 115 230 460 690 920

Figure 2.

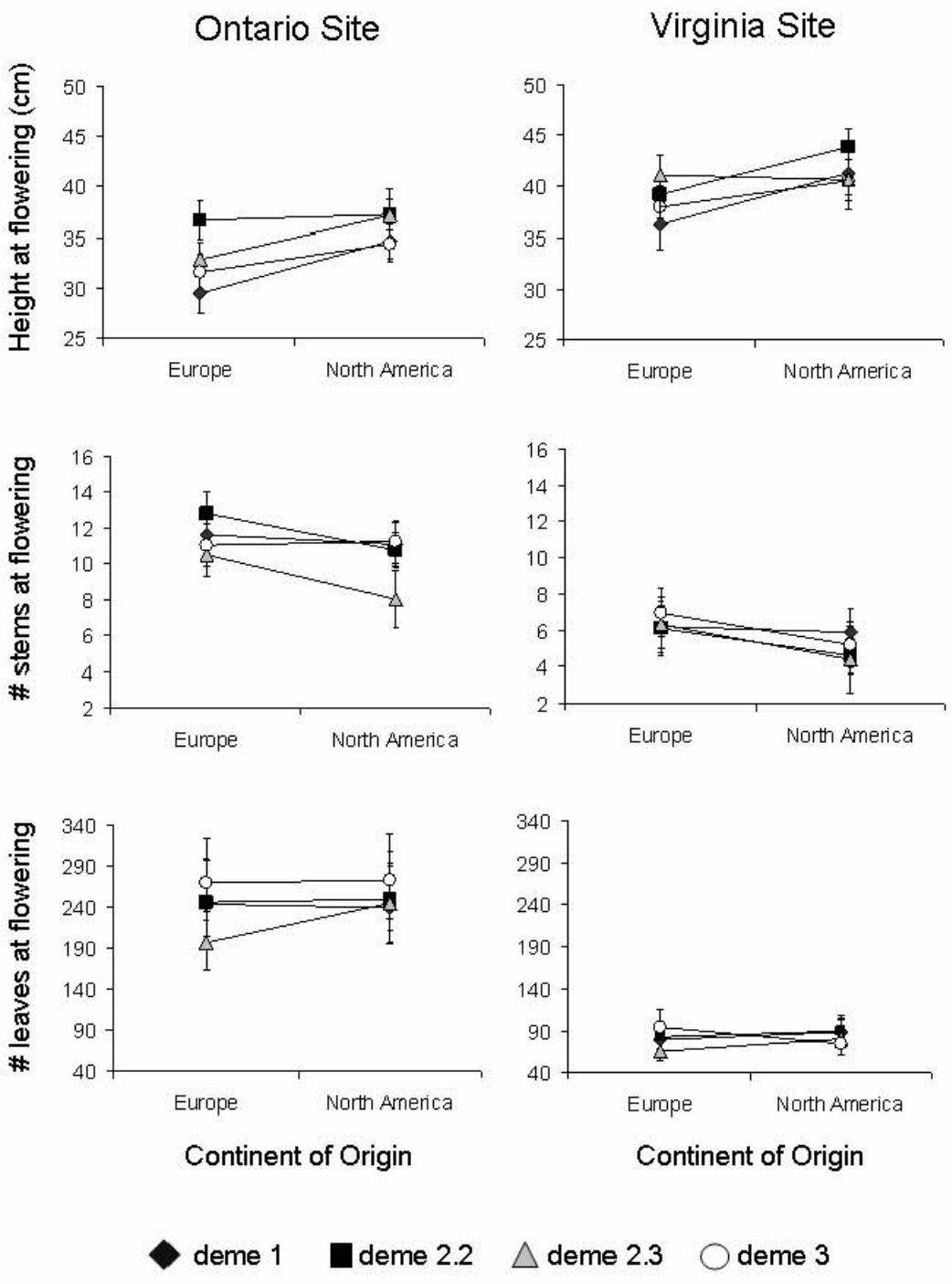

Ontario Site
Virginia Site
Height at flowering (cm)
stems at flowering
leaves at flowering
Europe
North America
Continent of Origin
deme 1
deme 2.2
deme 2.3
deme 3

Figure 3.

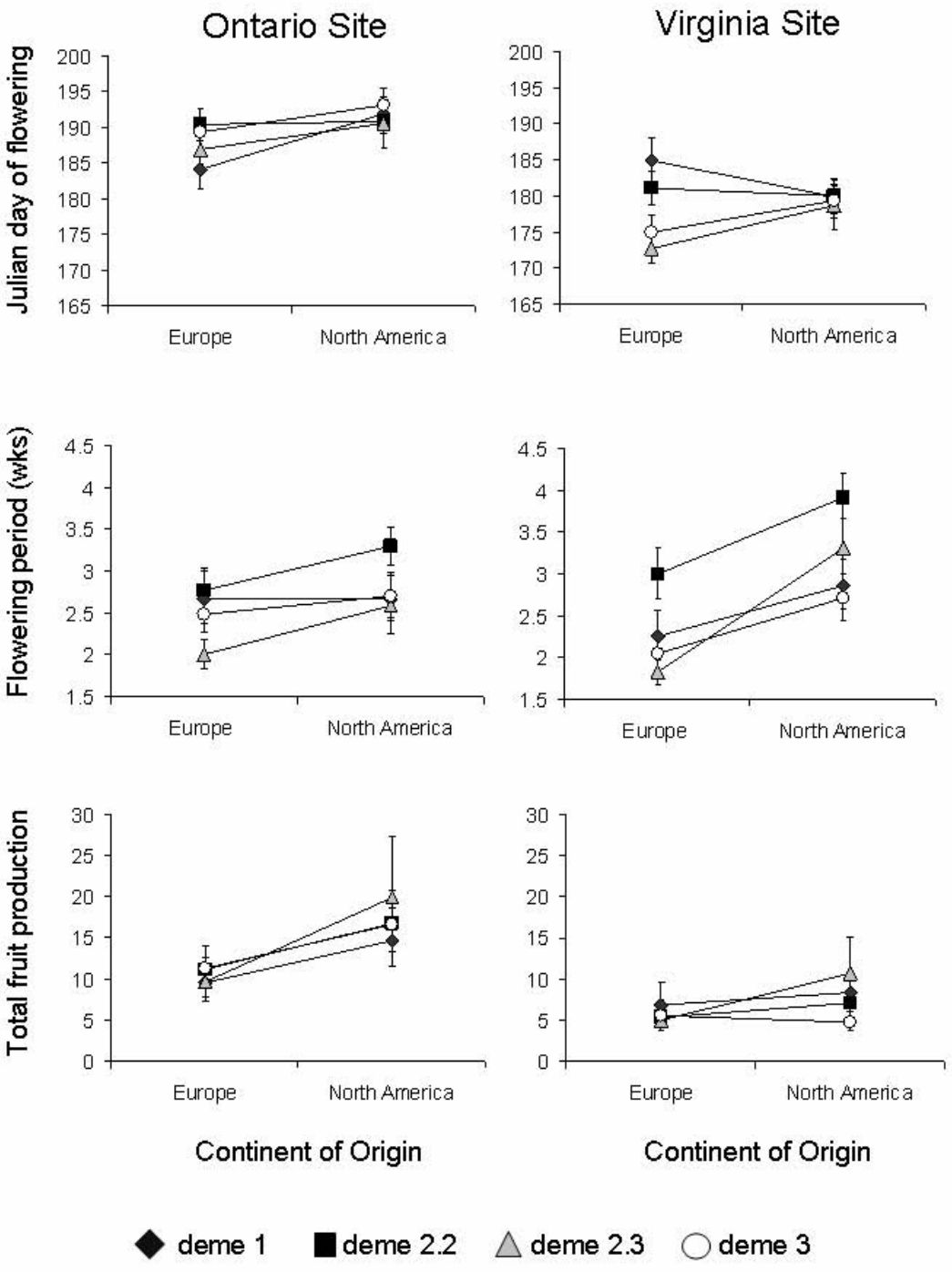

Ontario Site
Virginia Site
Julian day of flowering
Flowering period (wks)
Total fruit production
Europe
North America
Continent of Origin
deme 1
deme 2.2
deme 2.3
deme 3

Figure 4.

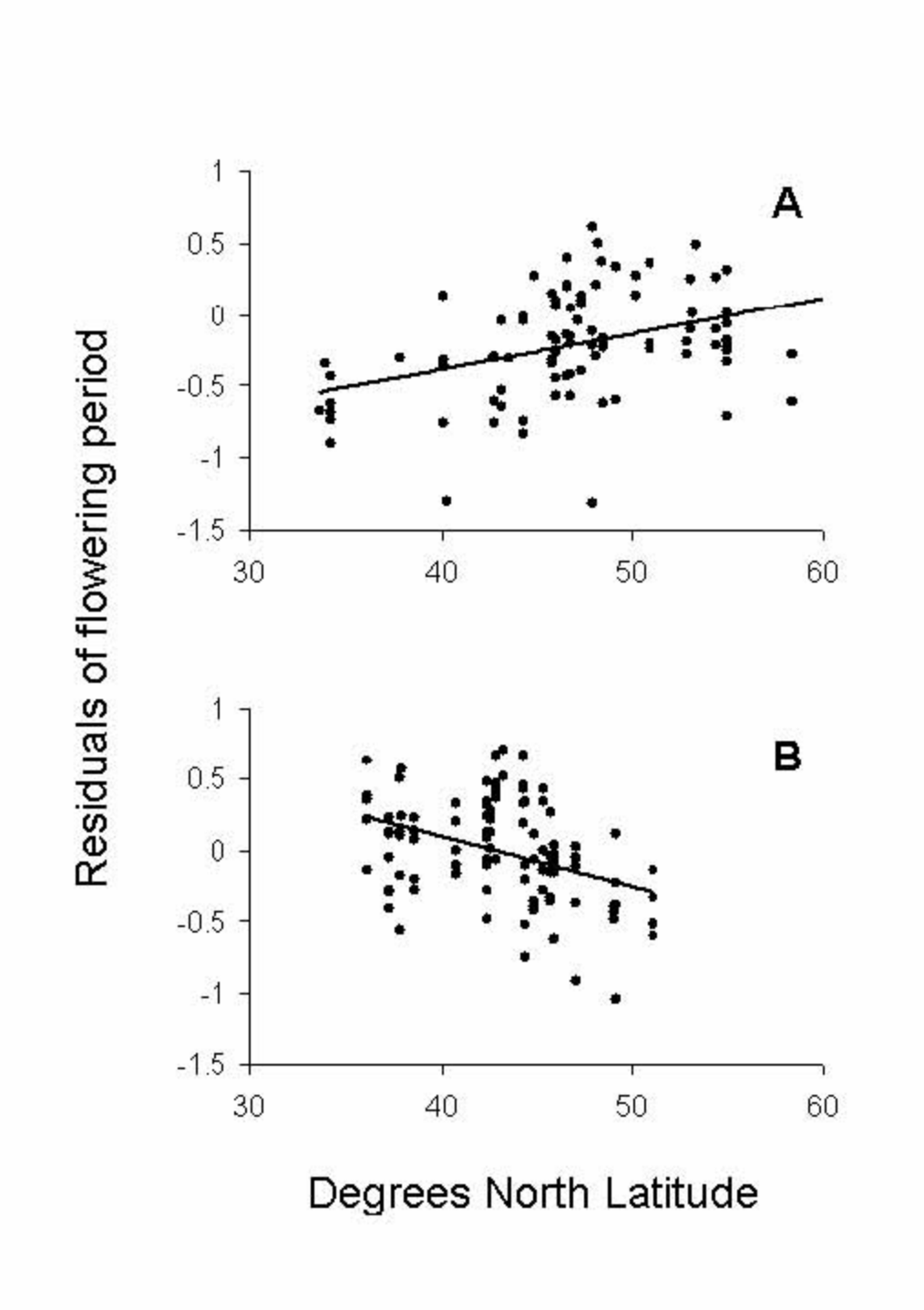

A
B
1
0.5
0
-0.5
-1
-1.5
30
40
50
60
Residuals of flowering period
Degrees North Latitude

Figure 5.

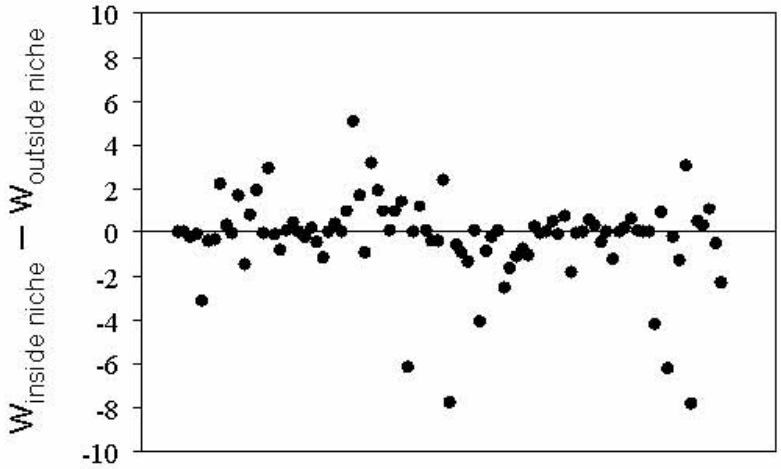


Figure 6.

Synthesis

This dissertation focused on the invasion of North America by *Silene vulgaris*, a weedy plant native to Europe, to address how this contemporary expansion of the species' range has been impacted by its evolutionary history in Europe, the genetic composition of the colonizing migrant pool, the novel selective environments encountered, and adaptive responses since the introduction.

In Chapter 1, I reviewed the literature on evolution during invasion and identified an area previously neglected in empirical studies: testing hypotheses of adaptive evolution during invasion against neutral expectations due to founder effects and drift. I developed two conceptual experimental designs that define null models for evaluating the hypothesis of adaptation against neutral expectations.

In Chapter 2, I characterized the evolutionary history of *S. vulgaris* and a closely related congener, *S. latifolia*, in Europe, and asked what consequences this history had for the sample of diversity introduced to North America. Genealogies of the maternally inherited chloroplast genome supported a demographic proliferation and range expansion of *S. vulgaris* over the last 600,000 years, conincident with the glacial cycles of the Pleistocene. Expansions during inter-glacial periods resulted in a departure of chloroplast genome diversity from equilibrium expectations, manifested as an excess of closely related lineages that have became widely dispersed across Europe. Chloroplast lineages invading North America were found to be a representative sample of European diversity, probably reflecting the lack of structure in Europe.

In Chapter 3, I used biparentally inherited genetic markers (Amplified Fragment Length Polymorphisms, or AFLPs) to show that, in contrast to the chloroplast genome,

the nuclear genome of *S. vulgaris* contained a signature of geographic structure among five genetically divergent lineages (i.e., "demes") differentially distributed in eastern, western, and southern Europe. The disagreement between genomes may reflect differences in sampling error around a single locus (the chloroplast) versus multiple loci (AFLPs), or possibly selection affecting the spatial distribution of chloroplast lineages. However, in agreement with the chloroplast data, AFLP genotypes showed a genetically diverse invasion, a result of multiple introductions from four of the five European demes. The structure observed with AFLP enabled invasive genotypes to be assigned back to their ancestral demes in Europe. These data suggested a stochastic colonization process had altered the frequency of demes in North America. Genotypes from eastern Europe were overrepresented in North America, while genotypes from western and southern Europe were underrepresented, relative to their frequencies in the native range. Despite the absence of spatial structure among demes in North America, an apparent lack of recombination has kept the genetic differences among demes intact.

In Chapter 4, I used ecological niche modeling to analyze the spatial pattern of colonization and expansion in the context of predictions from the ancestral niche in Europe. Museum specimens indicated the earliest introductions occurred in northeastern North America during the mid-1800s. A novel population genetic niche analysis showed that colonization often mismatched genotypes from their predicted distribution. After a lag phase, *S. vulgaris* rapidly spread west and south across much of continental U.S. and Canada. While the ancestral niche was a significant predictor of occurrence in the introduced range, a large area of niche expansion was evident in western North America. This is consistent with either *S. vulgaris* experiencing an evolutionary adjustment of the

ancestral niche since invasion, or North America occupying dimensions of the fundamental niche not present in Europe.

In Chapter 5, I experimentally tested whether there had been phenotypic evolution since invasion by planting individuals from each deme and continent combination into two experimental gardens in North America. These data conclusively demonstrated the evolution of several key fitness traits, including faster emergence times, allocation for apical dominance, and increased survival and reproductive effort compared to European genotypes. Incorporating a null model using AFLPs confirmed these differences were in excess of neutral expectations, strongly supporting a response to selection. Thus, genetically-based changes in fitness traits since invasion have been in part an adaptive process. However, demes were also divergent for many traits, regardless of continent of origin, and experienced shifts in relative frequency. This indicates that phenotypic divergence also occured due to processes acting at the deme level, such as stochastic sampling during colonization.

Collectively, these results demonstrate that a species rapidly expanding its geographic range is critically affected by prior evolutionary history, stochastic sampling events, and selection pressures in the introduced environment. In *S. vulgaris*, the contribution of **history** was largely driven by historical isolation among genetic lineages in Europe. This isolation had two important consequences: it spatially subdivided genetic diversity, creating opportunities for sampling error and founder effects during invasion (Chapter 3), and it created intraspecific divergence in the ecological niche (Chapter 4) and in phenotypic traits important to fitness (Chapter 5).

These historical factors interacted with the sampling process during invasion to determine the effects of **chance**. A clear result that emerges is that when population structure exists in the native range, it tends to magnify opportunities for unrepresentative sampling. This was evident in the comparison between the structured and unstructured chloroplast genealogies of *S. latifolia* and *S. vulgaris* (Chapter 2), and between the chloroplast and nuclear genomes of *S. vulgaris* (Chapter 3). A second result is that historical isolation creates niche divergence, which increases the likelihood that genotypes are randomly introduced outside their zone of preadaptation (Chapter 4). Thirdly, phenotypic divergence in the native range means sampling may alter the distribution of traits affecting fitness, with or without the action of selection.

The effects of **adaptation** emerged as selection acted within the context provided by history and chance (Chapter 1). The invasion of *S. vulgaris* was genetically diverse (Chapters 2 and 3), suggesting selection had adequate variation to act upon. Divergence between ranges indicated that on a broad scale, selection within the introduced range acted on phenology, growth, and life history. Clinal variation within both ranges suggested a finer-scale of adjustment to local climate conditions. Finally, it should be noted that the deme that is in highest frequency in North America (deme 2.2) consistently possessed phenotypic means in Europe that were closest to the direction of change that demes exhibited since colonizing North America. This hints at the possibility that the change in frequency may reflect a component of interdemic selection (Chapters 3 and 5).

These results encapsulate a major advance in our understanding of evolutionary processes acting during dynamic changes to a species' range. Because many introduced species present a serious threat to native biota and human economy, the findings

presented here also have additional conservation implications. In particular, while it is not clear if phenotypic evolution will *always* accompany invasion, it does appear to be a frequent process occurring at rates of change within the norm of many taxa. This would seem to argue for viewing evolution during invasion as expectation, not as exception. Furthermore, different stages of invasion are likely to experience selection episodes with distinctive outcomes. For example, selection during establishment is much more likely to canalize the phenotypes that become invasive, acting essentially like a range-wide selective sweep. These are the types of evolved changes that produce differences between continents (shifts in life history, reproductive effort, allocation to defense, etc.). Since this is a stage where many invasions fail, species passing through this constriction are likely to either already possess or to have evolved the traits that will make them problematic. In contrast, selection during or after the range expansion will tend to act more locally, fine tuning phenotypes to their local conditions. The latter process is more likely to be benign, as it heralds an evolutionary integration with the recipient environment (biotic and abiotic) and a shift away from the colonizing syndrome. The results of this dissertation suggest that these basic considerations from evolutionary biology are broadly applicable to the conservation problem of invasive species.