

1-1-2010

## Genetic patterns of dispersal and colonization during initial invasion and spread of an invasive grass, *Brachypodium sylvaticum*

Alisa Paulsen Ramakrishnan  
*Portland State University*

**Let us know how access to this document benefits you.**

Follow this and additional works at: [http://pdxscholar.library.pdx.edu/open\\_access\\_etds](http://pdxscholar.library.pdx.edu/open_access_etds)

---

### Recommended Citation

Ramakrishnan, Alisa Paulsen, "Genetic patterns of dispersal and colonization during initial invasion and spread of an invasive grass, *Brachypodium sylvaticum*" (2010). *Dissertations and Theses*. Paper 355.

10.15760/etd.355

This Dissertation is brought to you for free and open access. It has been accepted for inclusion in Dissertations and Theses by an authorized administrator of PDXScholar. For more information, please contact [pdxscholar@pdx.edu](mailto:pdxscholar@pdx.edu).

Genetic Patterns of Dispersal and Colonization  
During Initial Invasion and Spread of an Invasive Grass,

*Brachypodium sylvaticum*

by

Alisa Paulsen Ramakrishnan

A dissertation submitted in partial fulfillment of the  
requirements for the degree of

Doctor of Philosophy  
in  
Biology

Dissertation Committee:  
Mitchell B. Cruzan, chair  
Sarah Eppley  
Keith Karoly  
Susan E. Masta  
Alan Yeakley

Portland State University  
2010

## Abstract

Evolution of genotypes during range expansion is driven in part by colonization dynamics. I investigated genetic patterns of colonization and dispersal during initial expansion of an invasive bunchgrass, *Brachypodium sylvaticum*, into Oregon. Using microsatellite markers, I sampled plants at two different scales: at regular intervals along three parallel roads spanning about 30km, and in populations identified throughout Oregon. I also collected field-generated progeny from a subset of populations and used molecular identification of outcrossing events to estimate selfing rates in both central and peripheral populations.

Dispersal patterns were similar at both scales, with non-contiguous dispersal responsible for colonization of new populations. High levels of differentiation were observed at all scales, though newly-colonized populations were more differentiated than older populations. Corvallis populations were responsible for colonization of a majority of populations throughout Oregon, while individuals from Eugene were only occasionally found in new populations. Admixture occurs between Corvallis and Eugene populations, decreasing differentiation, and potentially creating novel phenotypes and increasing evolutionary potential of populations. Selfing rates were high, but two populations in the areas of original introduction had lower rates of selfing, suggesting that selfing rates may decrease as population density and diversity increases with age. The influences of founder effects and bottlenecks on phenotypic evolution during range expansion require further investigation, as inbreeding, lag times, and selection may influence evolutionary trajectories of populations.

## **Dedication**

Dedicated to my family

## **Acknowledgements**

Many people helped me complete my dissertation, but I would especially like to thank my husband Vivek, who supported me in every stage of my research, and my daughter Maya, who helped during the latter half of my degree program. I have been very fortunate in my advisor, who has been crucial in assisting me to develop and complete my research goals. Without the help of Trieste Musial and David Rosenthal, these data would have been woefully incomplete. I also thank numerous reviewers and editors including J. M. Rhode, D. Rosenthal, T. Musial, M. Paulsen, P. Sochaki, S. E. Jørgensen, and several anonymous reviewers. Thanks also to G. B. Hinckley and T. S. Monson. A special thanks to people who helped support me financially, including F. and V. Thompson, V. and T. Paulsen and L. Ramakrishnan. This work was funded in part by USDA CSREES grant 2005-35320-15317 to M. B. Cruzan and A. P. Ramakrishnan, and by the Center for Invasive Plant Management grant GC261-04-Z1138.

## Table of Contents

Abstract.....	i
Dedication.....	ii
Acknowledgements .....	iii
List of Tables .....	v
List of Figures.....	vi
Chapter 1: Review of Dispersal and Migration .....	1
Synopsis.....	1
Introduction .....	2
Dispersal Patterns .....	8
Modeling Dispersal .....	10
Measuring Dispersal.....	17
Chapter 2: Isolation and characterization of nine microsatellite markers for <i>Brachypodium sylvaticum</i> (Huds.) Beauv. ....	28
Abstract.....	28
Microsatellite development and characterization.....	28
Chapter 3: Shifting dispersal modes at an expanding species' range margin .....	34
Abstract.....	34
Introduction .....	35
Methods .....	40
Results .....	46
Discussion.....	50
Conclusions .....	56
Acknowledgements .....	57
Chapter 4: Founder effects and bottlenecks strongly influence colonization dynamics in a nascent invasive species .....	67
Abstract.....	67
Introduction .....	68
Materials and Methods .....	75
Results .....	83
Discussion.....	91
Chapter 5: Concluding remarks .....	116
Literature Cited.....	119

## List of Tables

Table 2.1: Locus information for microsatellites characterized in <i>Brachypodium sylvaticum</i> .....	32
Table 2.2: Locus characterization information in two populations .....	33
Table 3.1: Previously uncharacterized loci .....	59
Table 3.2: Analysis of molecular variance .....	59
Table 3.3: Assignment test results .....	60
Table 3.4: Relationships of site diversity and interallelic disequilibrium to site history and assignment tests .....	60
Table 4.1: Population information.....	102
Table 4.2: Analysis of molecular variance (AMOVA) .....	104
Table 4.3: Population pairwise differentiation of source populations.....	104
Table 4.4: Population pairwise differentiation of peripheral populations .....	105
Table 4.5: Assignment test probabilities .....	106
Table 4.6: Population history .....	107

## List of Figures

Fig. 1.1: Dispersal distributions.....	27
Fig. 3.1: Distribution of <i>Brachypodium sylvaticum</i> in Oregon, as reported by managers and by personal observation .....	61
Fig. 3.2: Hypothetical populations experiencing different types and amounts of dispersal.....	62
Fig. 3.3: Map of study area including Foster City and a log-processing pond.....	63
Fig. 3.4: Population similarities.....	64
Fig. 3.5: Diversity, heterozygosity and expansion direction.....	65
Fig. 3.6: Interallelic gametic disequilibrium .....	66
Fig. 4.1: Possible patterns of correlation between differentiation and genetic diversity .....	108
Fig. 4.2: Known distribution of <i>B. sylvaticum</i> in Oregon as of 2006 .....	109
Fig. 4.3: Comparison of diversity between parental and progeny samples .....	110
Fig. 4.4: Relationship between diversity (effective allele number) and average pairwise population differentiation ( $G'_{ST}$ ).....	111
Fig. 4.5: Population genetic distance.....	112
Fig. 4.6: Population clustering and admixture.....	113
Fig. 4.7: Selfing rates estimated using the program MLTR.....	114
Fig. 4.8: Effective number of alleles and selfing frequency.....	115



## **Chapter 1: Review of Dispersal and Migration**

Published as: Ramakrishnan AP (2008) Dispersal-Migration. In: Encyclopedia of Ecology eds. Jørgensen SE, Fath BD), pp. 930-938. Elsevier, Oxford. Used by permission.

### **Synopsis**

Population dynamics are directly affected by dispersal, through the immigration of individuals into populations and by the emigration of individuals out of populations.

Much of what we understand about dispersal patterns, their causes and effects comes from mathematical models. These models range in complexity from estimating the effect of simple diffusion processes on a population (i.e. simple reaction-diffusion models) to incorporating explicit information about multiple parameters into a detailed model (i.e. complex cellular automata models). Field measurements of dispersal can be difficult, depending on the level of detail desired. Ideally, demographic studies are combined with measurements of dispersal taken from individuals tracked in detail throughout their lifetimes. However, it is common practice to focus on only one or a few parameters of dispersal, depending on resources available to the researcher.

Methods including mark-recapture, seed traps, and genetic estimates of dispersal can be used to collect dispersal data. Each method has its strengths and weaknesses, which should be carefully evaluated by the researcher prior to utilization. As methods for modeling and detecting dispersal events improve, our ability to predict population dynamic responses to environmental perturbations will further benefit a broad spectrum of biological sciences.

## **Introduction**

Dispersal, or the movement and subsequent breeding of individuals from one area to another, strongly influences the population dynamics of a species. Dispersal can help regulate population size and density; many animals, such as aphids and female root voles, have increased dispersal rates under high density situations (Aars & Ims 2000; Mashanova *et al.* 2008). Sometimes low density instead of high density is associated with greater dispersal rates. For example, during range expansions, peripheral populations of some crickets may experience higher dispersal rates though they are of lower density than central populations, probably because of fitness costs associated with morphologies specialized for dispersal (Thomas *et al.* 2001).

Such dispersal events can have large effects on neighboring populations. Marginal populations that are subject to high rates of immigration may experience a rescue effect, where despite poor genetic or ecological conditions, populations are able to persist (Brown & Kodric-Brown 1977). On the other hand, high dispersal rates can inhibit adaptation to novel environments due to constant influx of nonadapted individuals (Slatkin 1987). Small populations that experience high rates of immigration may have a higher probability of extinction under such situations (Levin 1976).

Natural populations in highly fragmented areas, such as agricultural or urbanized settings, may not experience sufficient levels of dispersal (Cain *et al.* 2000). Lack of dispersal can lead to high rates of inbreeding (Wright 1946), which in turn can lead to decreased fitness in many species (Johnson & Gaines 1990; Wright 1931;

Wright 1980). Because dispersal can have such strong effects on populations, dispersal patterns and processes are important when considering the potential spread of a biocontrol agent, pathogen, or invasive species into a new range (Fagan *et al.* 2002; Sakai *et al.* 2001; Sax *et al.* 2007). Dispersal also has implications for species redistributions due to climate change, as the dispersal rates and distances of a species will affect its potential to shift its range in response to climate change (Cain *et al.* 2000; Clark 1998; Hewitt 1999; Higgins & Richardson 1999).

Two types of dispersal are commonly distinguished: natal dispersal, which is movement and subsequent breeding away from the birth territory or area, and breeding dispersal, which is movement from one area to another after the first breeding season (Johnson & Gaines 1990). Dispersal of spores, or haploid life stages (such as pollen), strongly affects patterns of gene flow in a species (Austerlitz *et al.* 2000; Ellstrand & Marshall 1985; Wright 1946), and can have consequences for population dynamics of species (Hanski 2001). Dispersal in plants is generally limited to natal dispersal, as little to no secondary movement is possible (Dlugosch & Parker 2007; but see Travis *et al.* 2002; Wolfenbarger 1975), while many animals disperse multiple times (Lidicker & Stenseth 1992).

All species disperse to some extent, in part because resources become limited locally as populations grow (Galloway 2005; Johnson & Gaines 1990). Seedlings of plants must grow at some distance from the parent plant in order to obtain enough water, nutrients, and light to survive (Nathan & Muller-Landau 2000). Similarly, animals must disperse to avoid competing for resources such as mates, food, and

territory (Lidicker & Stenseth 1992). Depending on intraspecific patterns of resource limitation, dispersal is often sex-biased. In mammals, females tend to disperse more often than males (Wolff 1997); the trend is reversed in birds (Clark & Low 2001; Wolff & Plissner 1998).

In areas with high temporal environmental variation, or in areas prone to frequent disturbances, species with greater dispersal abilities are expected to have a greater likelihood of survival (Ouborg *et al.* 1999). When one population's habitat is rendered untenable, if the species has a high dispersal rate, many individuals in that population will be able to move to a more suitable area (Matlack & Monde 2004). In the case of nonmotile organisms such as plants, high dispersal rates increase the likelihood that another population may be established even as the original population is rendered extinct (Clobert *et al.* 2004; Levin *et al.* 2003). When studying populations that specialize in habitats with high temporal environmental variation, it is sometimes appropriate to distinguish between spatial and temporal dispersal. For example, many animals and plants will produce desiccation-resistant embryos that delay maturity until favorable environmental conditions cue further development (Bohonak & Roderick 2001; Hairston *et al.* 1995; Rezende *et al.* 2008). Instead of traveling long distances to reach suitable habitat, the individuals produce offspring that are able to lie dormant until the habitat is once again suitable for survival and reproduction. Because dispersal can enable escape from low-quality environments and access to higher-quality resources, many species that specialize in colonizing disturbed areas tend to have

greater dispersal abilities than species that live in relatively stable habitats (Cain *et al.* 2000).

In some cases, dispersal can have a high cost associated with it, especially if individuals that disperse experience a higher mortality rate than those that do not disperse, or that disperse only a short distance (reviewed in Johnson & Gaines 1990). Because individuals are moving to an area that may not be as productive, and because they may have to travel through unsuitable habitats, mortality rates may be high during the dispersal process (Bohonak & Jenkins 2003; Johnson & Gaines 1990). The number of individuals that successfully establish in a new area may be far fewer than the number of individuals engaging in the dispersal process (Johnson & Gaines 1990). In plants and other organisms with no choice involved in the dispersal process (passive dispersal), many propagules may never establish simply because they land in an unsuitable habitat (Levin *et al.* 2003). In animals where some choice may be involved in the final dispersal location (active dispersal), survival of dispersing individuals may be higher than individuals of species with passive dispersal (Weisser 2001), but there are still risks associated with dispersal, such as locating an appropriate territory, finding a mate, and successfully breeding in the new area (Phillips *et al.* 2008; Travis & Dytham 2002). However, the benefits of dispersal can overcome the costs if mates and/or resources are limiting in the home range (Aars & Ims 2000; Johnson & Gaines 1990).

The process of dispersal is not necessarily as simple as suggested above, as it involves both emigration (leaving the original patch) and immigration (entering a new

patch). The entire process of dispersal can be divided into approximately four different stages: (1) emigration, (2) exploring or traveling through the surrounding habitat, (3) immigrating to a different patch, and (4) successfully breeding in the new patch (Lidicker & Stenseth 1992; Weisser 2001). Each of these stages has a cost involved. Leaving the original patch involves leaving an area where resources are known to exist, but may have become limiting. The exploratory phase of dispersal can involve a high risk of mortality (Weisser 2001), as the individual may have to travel through territories with inadequate resources (Andreassen *et al.* 2002). In many plants and other passive dispersers, the exploratory phase entails a high rate of mortality (Weisser 2001). Even when a propagule successfully disperses to a hospitable environment, it may not be able to establish there, due to mortality rates associated with establishment (Levin *et al.* 2003; Nathan & Muller-Landau 2000). The risks involved with emigration, exploratory movement, and settling in a new patch can be outweighed by the potential benefits of dispersal if successful dispersal significantly increases the fitness of the individual (Levin *et al.* 2003).

There are varying degrees of active and passive dispersal, with many species exhibiting intermediate levels of participation in the dispersal process. In many animals, dispersal is active, involving a high level of choice during the dispersal process (Stenseth & Lidicker 1992). In passive dispersal, there is little or no choice involved in selection of the final location. In many insects, many marine animals, and all plants, dispersal is largely passive, depending on air currents, water currents, or on the actions of vectors transporting the propagule (Lidicker & Stenseth 1992). Larvae

of many marine animals are often dispersed solely at the whims of the currents or in ship ballast. Insects are often at the mercy of the wind when entering a dispersal phase, especially if they cannot generate enough speed to overcome wind velocities (Wolfenbarger 1975). However, even dispersal of small insects need not be completely passive. Small insects, even if they are not large enough to overcome wind velocity, can have some level of choice as to where they land. They can begin exiting a wind stream when they decide to settle, then make short, self-powered trips to explore the surrounding area and find a suitable habitat (Wolfenbarger 1975).

Though considered passive dispersers, plants can regulate dispersal to some extent. Seed size, shape, and seed coat construction vary among species (Levin *et al.* 2003; Nathan & Muller-Landau 2000). Seed morphologies that aid dispersal include barbs (for attaching to animals), eliasomes (for attracting ants as dispersal vectors), and pappus scales (to assist in wind transport) (Greene & Calogeropoulos 2002; Martin & Cruzan 1999). However, because the seed itself is not actively involved in the decision process, it is still a passive process.

A species with little innate dispersal ability may be able to move greater distances and have higher survival than expected if it has the ability to be spread by a vector, such as ants, birds, or other animals (Greene & Calogeropoulos 2002). Plants commonly use vector-assisted dispersal, and there are many instances of adaptations by plants to use animals as dispersal agents (Levin *et al.* 2003). For example, mistletoe seeds are eaten by birds which then fly to another tree. The seeds are adapted to

survive the digestive tract, and are subsequently deposited on the tree where the bird lands, which is usually a suitable tree (Nathan & Muller-Landau 2000).

Most vector-associated dispersal regimes have evolved over hundreds of generations. Recently, however, many species of both plants and animals have serendipitously become associated with novel and extremely efficient dispersal vectors. Species associated with humans have always been dispersed in concert with human movements (Mack & Lonsdale 1985). However, the last few generations of humans have seen an exponential increase in the rates of movement around the globe (Mack *et al.* 2000). Many terrestrial and marine species have been spread at unprecedented rates through ship ballast and packing materials (Gelembiuk *et al.* 2006; Lee *et al.* 2004; Mack *et al.* 2000). In addition, ornamental plants and agriculturally associated species are deliberately transported from one location to another by humans, at distances and rates that would be impossible for each species to accomplish under its own power (Mack *et al.* 2000). Hundreds of species involved in these accidental experiments in dispersal and evolution have benefited tremendously, becoming the world's invasive species. Species such as cheatgrass in North America, *Caulerpa taxifolia* (an alga) in the Mediterranean, and the Nile Perch in Africa have successfully outcompeted hundreds of native species, sometimes driving them to extinction (Mack *et al.* 2000; Sax *et al.* 2002).

### **Dispersal Patterns**

There are several terms associated with dispersal patterns. Dispersal distributions, or dispersal curves, are frequency distributions of the proportion of



individuals moving different distances; dispersal kernels are probability density functions used in modeling dispersal (Nathan & Muller-Landau 2000). Because a dispersal kernel describes the probability of a seed landing within a particular region (Neubert *et al.* 1995; van den Bosch *et al.* 1990), dispersal curves can be calculated from dispersal kernels (Nathan & Muller-Landau 2000). Dispersal modeled with a simple diffusion equation (a Gaussian kernel) generates a normal distribution curve (Austerlitz *et al.* 2004; Levin *et al.* 2003), due to the random Brownian motion assumed in such simple diffusion models. Though accurate for some species, most species appear to have leptokurtic, or long-tailed dispersal curves (Austerlitz *et al.* 2004; Clark *et al.* 1998; Skalski & Gilliam 2003).

The shape of a dispersal curve drastically affects estimated rates of population expansion, with normal curves having lower expansion rates than dispersal curves with longer tails (Clark *et al.* 1998; Kot *et al.* 1996; Le Corre *et al.* 1997). The fatter the tail of the distribution, the greater is the speed of the range expansion (Clark *et al.* 1998; Kot *et al.* 1996). Accurately estimating the shape of the curve is important, for example, in predicting spread rates of spread of the emerald ash borer (Muirhead *et al.* 2006). Rates and pattern of spread are often consistent with simple diffusion; however, some infestations in Michigan do not spread via simple diffusion, but have a higher frequency of long-distance dispersal events (Muirhead *et al.* 2006). If all control efforts and spread rate predictions are based on simple diffusion, management programs will be unprepared for long-distance dispersal events and the efficacy of management efforts will be greatly diminished (Skalski & Gilliam 2003).

If the tail is exponentially bounded, as in curves with moderately fat tails (i.e., a Laplace distribution) (Figure 1), then models predict that the rate of expansion of a population will remain constant (Kot *et al.* 1996). However, in curves with unbounded tails (very fat tails, i.e., a Weibull distribution), the rate of expansion of a population can actually increase over time (Kot *et al.* 1996; Mollison 1977). Because different dispersal curves can drastically affect estimates of population expansion, it is important to choose appropriate dispersal curves when attempting to predict species' population dynamics.

### **Modeling Dispersal**

A simple model of a species consists of one population, infinite in size, randomly mating, with no immigration or emigration (Crow 1999; Fisher 1930; Hardy 1908; Wright 1931). This idealized situation is never found in nature, though populations can approach this equilibrium (Hey & Machado 2003). In real life, species are divided into populations that are subdivided to some extent, with dispersal occurring between the subdivisions at varying degrees. However, the idealized situation is often a good starting point for modeling species dispersal. Models investigating the effects of dispersal on population dynamics use several different approximations of natural situations. One of the simplest models of dispersal is a model proposed by Levins in 1969 (Levins 1969), one of the first metapopulation models. This model describes the colonization and extinction dynamics of sites under different rates of colonization (dispersal) and mortality. An individual produces offspring which then disperse to other sites. When a propagule reaches an empty site,

it occupies it completely, and any propagules reaching that same site are subsequently eliminated. This model is termed a metapopulation model because it deals with colonization and extinction of sites. Though originally formulated as entire sites that are either colonized or not colonized, it can be viewed as a collection of sites where each site is the size of a single individual (Tilman *et al.* 1997). In this way, a simple equation can be used to model the spread of a population through dispersal:

$$\frac{\partial s}{\partial t} = cs(1-s) - ms$$

In this equation, the basis of the Levins model (Levins 1969),  $s$  represents the proportion of sites colonized,  $c$  is the rate of propagule production, and  $m$  is the mortality rate (Tilman *et al.* 1997). Though not spatially explicit, spatial relationships are implied in this model by having a proportion of sites either available or not available. Some of the assumptions this model makes (Tschardtke & Brandl 2004) are that propagules are distributed randomly, and that each propagule can occupy any space in the habitat. A proportion of sites will be filled until the population reaches an equilibrium, determined by the relationship between the rates of mortality and propagule production (Tilman *et al.* 1997). The population will grow at a rate that increases with lower mortality rates and increased propagule production. As long as mortality rates are less than propagule production, the population can persist. Though extremely simplified, this model has led to some interesting predictions about population dynamics. Because the final density of individuals is dependent on mortality rate in addition to propagule production, it will be impossible for a species to

occupy every suitable patch in the population's habitat (Tilman & Kareiva 1997). When one individual dies, the point it occupied is empty until recolonized by another propagule. Thus, the population reaches an equilibrium where a certain percentage of sites are empty, rather than completely filling the entire habitat.

Though this model yields some interesting predictions about the way population dynamics depend on propagule production and mortality rates, one of the drawbacks of this simple approximation of dispersal is that propagules in real life cannot occupy any unoccupied site in a habitat (Hanski 1997; Tilman & Kareiva 1997). Real dispersal is restricted in distance, with a large spatial component that is only implied in the basic Levins model. The spatial component of the model is unrealistic, being essentially infinitely large and infinitely accessible (Hanski 1997).

To observe the effects of dispersal on population dynamics in more realistic situations, greater spatial detail is required. Incorporating spatial information into a model of dispersal is relatively simple by using a cellular automaton or lattice model (Hogweg 1988; Matlack & Monde 2004; Molofsky 1994; Tilman *et al.* 1997). This type of model does not have an explicit mathematical solution but can be evaluated through simulations that can lead to predictions about population dynamics based on certain assumptions. In this type of model, individuals are placed on a grid of polygons. Each individual has a certain chance of mortality and of producing propagules. A propagule can travel a certain distance away from its current position, and its direction can be determined by either having the propagule move through a side or corner of the parental polygon. The propagule can be restricted to land within a

certain area or not restricted at all. Again, any propagule that lands in an occupied site is lost. Parameters in these models are relatively simple to modify, and any combination of requirements can be included (Ferrière & Le Galliard 2001), including different dispersal distribution curves, rates of long-distance dispersal, habitat extinction, propagule production rates, and mortality rates. However, more complex models will require more computational power, and they may not be as applicable to multiple species and/or situations as are more general models (Wiens 2001). One of the interesting predictions made by relatively simple cellular automaton models is that the distributions of individuals will be aggregated, even in a homogenous environment (reviewed in Tilman *et al.* 1997). The amount of aggregation of individuals depends on the average dispersal distance, the mortality rate, and fecundity. Some of these models have demonstrated that patchy spread is likely during population expansion simply through the stochastic nature of mortality and propagule dispersal distances. This is an important prediction, as patchiness need not be the direct result of unsuitable habitat, but instead can be merely the product of different dispersal patterns. Cellular automata models are often used in epidemiology, and have also been applied to species invasion dynamics.

Cellular automaton models are accurate if space, time, and population dynamics are best represented as discrete variables. In a species where many parameters are known, it is possible to reach a high degree of specificity in the simulations, which may be useful when trying to predict what may occur for a particular species under different scenarios. Less-specific models, though less

applicable to a single species, can lead to more generalized predictions about the effects of dispersal on population dynamics.

Often, it is convenient to model a population or set of populations more generally, using mathematical equations. One such model involves a reaction-diffusion equation (Skellam 1951), describing the rate of change of the number of individuals in a population ( $N$ ) over time ( $t$ ):

$$\frac{\partial N}{\partial t} = rN \frac{(K - N)}{K} + D \frac{\partial^2 N}{\partial x^2}$$

This model consists of two parts: the reaction portion of the equation (which describes how the population acts in the absence of dispersal) and the diffusion portion (which is a partial differential equation that describes the movement of individuals in the population) (Tilman *et al.* 1997). The reaction portion can be as simple as the logistic equation for growth of a population, shown here immediately to the right of the equals sign, where  $r$  is the intrinsic rate of increase, and  $K$  is the carrying capacity. Other models of population growth can be used as well, such as adding an Allee effect, where propagule production rate declines at a lower than expected rate at low densities (e.g., when a mate is difficult to find) (Stephens & Sutherland 1999). The diffusion portion of the equation is usually the mathematical representation of simple passive diffusion

$$\frac{\partial}{\partial t} N(x, t) = D \frac{\partial^2 N}{\partial x^2}$$

where  $D$  is the diffusion coefficient,  $t$  is time,  $N$  is the number of individuals, and  $x$  represents space (Tilman *et al.* 1997). The diffusion portion of the equation can be more complex, incorporating directional movement, changes in velocity, or interactions among individuals (Tilman *et al.* 1997). Increasing complexity is not always necessary to make meaningful predictions, as the simplest version of this equation matches observations seen in mark–release–recapture studies in several animals (Kareiva 1983). In models of invasion dynamics, a reaction diffusion model produces waves of invaders that advance at a rate dependent in part on the rate of increase of the population (Tilman & Kareiva 1997). This type of equation is most appropriate when the environment is homogenous, all individuals have similar dispersal patterns, and reproduction/dispersal occurs constantly. Such models generally produce a smooth traveling wave front with a linear rate of spread (Liebhold & Tobin 2008).

To incorporate the discrete reproduction events (and hence dispersal events) frequently seen in plants and animals, it is appropriate to use an integrodifference equation (Kot *et al.* 1996; Lewis 1997). Integrals allow each moment in time to be dependent on the previous moment in time, thus incorporating discrete time intervals into the model, as opposed to differential equations, which assume that time is continuous. Integrodifference models consist of two main parts: a dispersal kernel (a function that describes the dispersal patterns of a population) and a density function.

$$N_{t+1}(x) = \int_{-\infty}^{+\infty} k(x-y)f[N_t(y)]dy$$

In this equation from Kot *et al.* (1996),  $k(x - y)$  represents the dispersal kernel, or an equation that describes the movement of a propagule from the natal territory,  $x$ , to the final breeding place,  $y$ . The density function describes the density of individuals at the location  $y$  and time  $t$ . The various dispersal kernels used with these types of models are often derived from empirical data. There are many different dispersal kernels that can be incorporated into these models, including long-tailed (leptokurtic) kernels and other non-normal dispersal distributions. When production of propagules is not continuous, as is true in many animals and plants, this type of integrodifference model will better approximate reality than a reaction-diffusion model (Clark *et al.* 2001; Kot *et al.* 1996). In contrast to reaction-diffusion models, integrodifference models can show an accelerating rate of spread over time if long-distance dispersal is relatively common (Kot *et al.* 2004). Integrodifference models are more consistent with observed patterns of range expansion than reaction-diffusion models because of the accelerating rate of spread often observed. The waves of invasions seen in these models have smooth fronts, similar to the reaction-diffusion models (Kot *et al.* 2004). Space is still represented as continuous in these models, and spatial relationships are implicit, not explicit as in cellular automata/lattice models (Molofsky 1994).

There are many other models that can be used to investigate the effects of dispersal on population dynamics. Stratified diffusion is a variant of the reaction-diffusion equation, where a proportion of individuals are assumed to travel long distances (Shigesada *et al.* 1995). Stratified diffusion models have accelerating rates of



spread, similar to the integrodifference models (Austerlitz & Garnier-Gere 2003; Liebhold & Tobin 2008). Metapopulation models are applicable to modeling dispersal patterns and rates of spread, and are also applied to other problems in ecology such as persistence of groups of populations in specific spatial arrangements under various scenarios of habitat destruction. They are commonly associated with biogeographical studies. A variation on integrodifference equations that includes stochastic population dynamics (nonlinear integrodifference models) shows waves of invasion that are patchy, with variable spread (Kot *et al.* 2004).

## **Measuring Dispersal**

### *Direct Measures*

Generating an accurate picture of the entire process of dispersal in a species involves detailed demographic analyses in addition to tracking emigrating and immigrating individuals (Cain *et al.* 2000). In order to know what demographic parameters drive effective dispersal, it is important to know how many individuals leave, survive the exploratory process, and breed successfully in the new area (Aars & Ims 2000). Ideally, all parameters of dispersal should be quantified. However, because the dynamics of a population are directly driven by effective dispersal, it may be unnecessary to conduct detailed studies of each stage of dispersal, depending on the particular goals of the researcher.

Mark–recapture methods and demographic analyses can assist in the estimation of many dispersal-related parameters, and though the route traveled by the individual captured in a new patch is often unknown, it is still possible to gain an estimate of

immigration and emigration rates. Many studies focus on relatively local effects of dispersal, studying population dynamics in a few interconnected populations that are spatially tractable (Koenig *et al.* 1996). These studies involve either mark–recapture methods, genetic methods, seed traps (for plants), or radio- or satellite tracking methods. Animals and seeds can both be marked using tags, paint, or dyes. Tracking methods, such as by radio telemetry or satellite, show great promise for obtaining detailed information on dispersal patterns, especially on the tail of the dispersal curve (Burland *et al.* 2001; Nathan 2001; Nathan *et al.* 2003; Phillips *et al.* 2008). Genetic methods are becoming popular as well, because they can detect effects of very low rates of dispersal over long distances (Cain *et al.* 2000; Koenig *et al.* 1996; Nathan *et al.* 2003). Each method has advantages and disadvantages, and all these methods have assumptions and uncertainties associated with them, which must be taken into account when analyzing data and estimating dispersal curves.

Collecting data from the tail of the dispersal curve can be difficult, either hampered by the difficulty of maintaining sampling densities, or due simply to the rare and stochastic nature of long-distance dispersal events. The importance of long-distance dispersal in estimating the spread of populations was highlighted in scientific literature when the rate of post-Pleistocene expansion of trees in Europe estimated with models neglecting long-distance dispersal could not account for the rapid expansion rates observed as the glaciers receded (Reid 1899; Skellam 1951). Incorporating long-distance dispersal by modeling spread with leptokurtic dispersal curves matched the estimated rates of spread more closely (Clark *et al.* 1998).

Unfortunately long-distance dispersal events are extremely difficult to measure empirically, and hence estimating them has since received much attention.

For animals, one way of estimating dispersal patterns involves marking and releasing animals, then observing the animals when they are collected, usually during an annual harvest. In the case of mark–harvest methods, animals are only viewed twice, once during the marking process, and once when harvested. This type of data may be useful for estimating mortality rates associated with movement from one site to another if it is possible to assume that the animals in question always return to either the original marking site or to the final capture site. Prior knowledge of movement patterns is important; mortality cannot be estimated if a significant number of animals disperse outside of the sample area. If multiple mark–capture episodes are accomplished in one season, it is possible to estimate probabilities of survival and movement for a specific area. However, no models currently can estimate dispersal from one area to another using this type of data.

In a similar method, animals are marked, released, then re-sighted or recaptured and released again. Animals may be sighted multiple times with this method, and with a robust sampling design, immigration and temporary emigration rates can be estimated. If this type of method is employed on multiple sites, with site-specific markers, immigration and emigration probabilities in addition to transition rates can be estimated. If possible to employ, this type of design is quite useful, as it provides data necessary for estimating population dynamic parameters associated with

dispersal. Long-distance dispersal events are difficult to detect with this method, due to the stochasticity associated with the occurrence and detection of such events.

Seed dispersal is often measured using seed traps to capture seeds at varying distances from the source. Seed traps usually involve pit traps or sticky traps placed in or near the ground. To identify seeds' origins, individual fruits can be marked, a chemical tagging method can be used, or a rare genetic variant can be used as a marker. The most common method is to measure the densities of seed deposited at various distances from a source. Because individual plants are not identified when only density of seeds can be recorded, likelihood methods are used to model dispersal curves. Seed traps work well for estimating dispersal curves near the source, but as distance from the source increases it becomes more and more difficult to detect dispersal events. If enough traps are used, long distance dispersal events can be detected; however, such events will be rare, and their detection will be dependent on the resources available to the researcher.

Radio telemetry and satellite tracking provide excellent data, when practical. Such studies have documented that long-distance dispersal events are more common than estimates from mark–recapture methods suggest (Koenig *et al.* 1996). Most studies involve large- to medium-sized animals, including marine mammals. Invaluable information about the long-distance travels of these animals has been collected, including information about movements of some seabirds. Ideally, a large proportion of a population could be followed individually, and detailed analyses made of their movements. In order to accomplish this, the radio or satellite transmission

units should not inhibit movement or survival, and the batteries should be strong enough to allow signal detection at a distance for a long period of time. As technology advances, smaller tags can be used. For example, very small radar tags have lately been adapted for use on bumblebees, showing promise for generating detailed dispersal data for larger insects.

### *Indirect Methods*

Genetic methods hold promise for estimating dispersal patterns, though it is important to remember that genetic methods only measure effective dispersal, and not dispersal of individuals that did not successfully breed in the new population (Koenig *et al.* 1996). In addition, for organisms with motile gametes, genetic patterns will likely reflect the movement of gametes among populations as well as the movement of diploid individuals. Most genetic methods involve collecting DNA from immature or mature individuals, then analyzing the DNA to try and identify the origin of a particular individual. If a dispersed individual has the same genetic signature as the individuals in the new population, the dispersal event will be undetected (Berry *et al.* 2004); this becomes less likely when highly variable markers are used (Berry *et al.* 2004). Another possible drawback to genetic data is the potential for unsampled source populations to contribute to apparent gene flow estimates between two sampled populations (Beerli 2004). Also, it is difficult to generate a detailed dispersal curve using solely genetic data, especially at local distances, due to the large amounts of data that would have to be collected and the heavy expense involved. Nevertheless, genetic data enable estimation of many parameters of interest, such as historical amounts of

gene flow (Barton & Charlesworth 1984; Dobzhansky & Wright 1941; Wright 1943), effective dispersal rates among differentiated populations (Waples & Gaggiotti 2006; Wilson & Rannala 2003; Wright 1931), and dispersal rates over long distances (Nathan *et al.* 2003).

Genetic data are also often (but not always) easier and faster to collect than detailed demographic data (Nathan *et al.* 2003). These advantages can outweigh the potential difficulties with genetic data, depending on the parameters of interest (Koenig *et al.* 1996). The first genetic estimates of dispersal were derived from Wright's equation

$$N_e m = \frac{1}{4} \left( \frac{1}{F_{ST}} - 1 \right)$$

where  $N_e$  is effective population size (an estimate of the number of individuals in an idealized population that would show the same patterns of genetic diversity or levels of inbreeding),  $m$  is migration (dispersal) rate, and  $F_{ST}$  is a measure of population structure (Wright 1951). This equation can be used to estimate dispersal among populations using DNA sequence data, DNA markers of variable length, or allozyme (protein) markers (Ouborg *et al.* 1999). However, more recently the use of this equation to estimate dispersal rates has come into question based both on unlikely assumptions made when calculating  $F_{ST}$ , and on the applicability of the above equation to natural populations as opposed to idealized populations (Pearse & Crandall 2004; Whitlock & McCauley 1999). Some of the assumptions made include equal and constant population sizes, and equilibrium between gene flow and genetic drift

(stochastic variation in genetic frequencies over time) (Ouborg *et al.* 1999). In addition,  $F_{ST}$  -based methods do not distinguish between historical and contemporary gene flow (Pearse & Crandall 2004). In some cases, however, as in well-established, large populations,  $F_{ST}$  -based methods may be sufficient for estimates of dispersal rates (Gaggiotti 1999; Waples & Gaggiotti 2006).

Parentage analyses and assignment methods are both techniques that use genetic marker data to estimate dispersal (Berry *et al.* 2004; Burland *et al.* 2001; Cain *et al.* 2000; Ellstrand & Marshall 1985; Telfer *et al.* 2003). They both assume that source populations are discrete and that there is no genetic linkage disequilibrium (nonrandom associations of genetic marker variants due to inbreeding or chromosomal proximity) among the different markers used. Parentage analyses are based on multilocus genotypes, and can generate data about local animal movement and both seed and pollen dispersal. However, parentage analyses require extensive sampling to ensure that all possible parents in the source area have been included in the study. If a parent present in source populations is not sampled, dispersal from an unsampled (possibly quite far away) population may be inferred, potentially altering dispersal estimates. Depending on the situation, parentage analysis can be expensive enough to outweigh the potential advantages of genetic analyses over demographic studies.

Assignment methods, on the other hand, use allele frequencies, or frequencies of different variants of genetic markers, to predict from which source a particular individual came. This means that exhaustive sampling of source populations is unnecessary. It is still desirable to have representatives from all possible sources. In

order to distinguish sex-biased dispersal, sex-specific markers must be used. Assignment methods assume discrete source populations and no linkage disequilibrium, and often assume equilibrium of populations. However, methods that enable dispersal estimates when populations are not in equilibrium are being developed.

A method that has come into use recently for many questions in genetics involves Bayesian analyses. Bayesian methods can be used in nonequilibrium situations, such as during range expansions, with high levels of inbreeding, or with unequal population sizes. Information known about the populations in question is entered in the analysis in the form of prior probability distributions. This information, commonly known as a prior, is basically a guess about how the populations might act based on data already available, such as experimental data. If no information is already known about the populations, an uninformative prior can be used. Then the genetic data is used, in conjunction with the prior, to calculate posterior probabilities of the data using a maximum likelihood algorithm, given the parameters currently in place in the maximum likelihood model. Markov chain Monte Carlo resampling is then used to explore parameter space and find values that optimize the fit of the parameters in the model to the data. The accuracy of detecting dispersal events with these methods is still being explored. Factors affecting accuracy of dispersal detection include the level of genetic diversity in the populations, amount of dispersal occurring among populations, how many genetic markers are used, and the level of variability in the markers themselves. Depending on the population structure in the system in question,



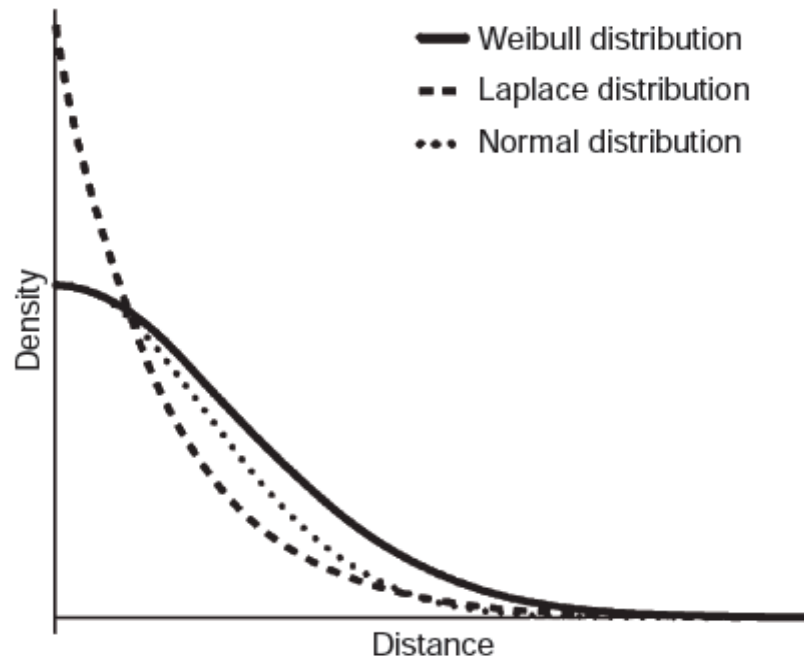
these methods may be equally viable both for detection of local dispersal and long-distance dispersal events.

As mentioned in the beginning of this article, dispersal in the context of population dynamics usually refers to the dispersal of diploid individuals. However, especially when genetic methods are used, movement of haploid gametes can affect dispersal estimates. Male and female gametic dispersal patterns often differ, and if not accounted for, can skew dispersal estimates from genetic data. Nuclear genetic markers come from both paternal and maternal sources, due to the combination of nuclear genetic material during fertilization. If gametic gene flow is significantly different than diploid dispersal patterns, such as with pollen in many plants, care must be taken not to confuse gametic patterns of gene flow with movements of diploid individuals. In plants, for example, the male pollen often travels farther than seeds, especially if the species in question is wind-pollinated. In order to distinguish the dispersal of diploid seeds from haploid pollen, sex-specific genetic markers must be used. In plants, the chloroplast genome is generally (but not always) maternally inherited, and comparing patterns of genetic differentiation between the chloroplast and the nucleus can be used to estimate pollen versus seed dispersal patterns. If only seed dispersal is of interest, it may be sufficient to focus on chloroplast genetic markers, if the chloroplast is indeed maternally inherited in the species in question. Similarly, markers based on mitochondrial DNA will only show dispersal patterns of the female in animals. If dispersal patterns of both males and females are desired, both nuclear and mitochondrial markers must be used.

Because different analysis techniques have different strengths, weaknesses, and assumptions, it is important to consider what the goals of a given study are, and what methods are best suited to the questions at hand. Genetic methods hold promise for dispersal estimations, as data can be gathered relatively quickly and with less cost than demographic data. However, depending on the species and/or populations under consideration, estimates may require more genetic data than are currently available to optimize the parameters of a genetic analysis technique. If this is the case, then the cost of genetic analyses could equal or exceed the costs of demographic data collection. In addition, some data must be gathered through observation, such as detailed movement patterns among sites or populations and breeding success rates. Other data can best be estimated using genetic data, such as historical patterns of gene flow or long-distance dispersal events, depending on the organism in question. As techniques for measuring dispersal and its consequences improve, we will be better able to predict the survival, extinction, range expansion, and range contraction of populations and species. These predictions will improve estimates of the effects of habitat destruction and fragmentation on populations, a growing concern at multiple scales worldwide. In addition, accurate predictions of range expansion and contraction rates based on models of global warming will enable me to prepare for the possible effects of a rapidly changing climate on both marine and terrestrial species.

Fig. 1.1: Dispersal distributions

Normal curves are drawn in each graph for comparison. The Laplace distribution is more peaked than the normal curve, and leptokurtic. The Weibull distribution shown here has a fatter tail than both the normal and the Laplace distribution. Graph by Hyrum Paulsen.



## **Chapter 2: Isolation and characterization of nine microsatellite markers for *Brachypodium sylvaticum* (Huds.) Beauv.**

Published as: Ramakrishnan AP, Rosenthal DM, Dobberstein T, Cruzan MB (2008) Isolation and characterization of nine microsatellite markers for *Brachypodium sylvaticum* (Huds.) Beauv., a recently invasive grass species in Oregon. *Molecular Ecology Resources* 8, 1297-1299. Used by permission.

### **Abstract**

The patterns of genetic diversity caused by rapid range expansions following recent colonizations are best observed using highly polymorphic genetic markers. I characterized nine microsatellite markers for *Brachypodium sylvaticum*, a bunchgrass invasive in the Northwestern United States and native to Eurasia. Loci exhibited from two to ten alleles, and generally had high  $F_{IS}$  values. These loci will help identify sources of new populations in the region, and they will be useful for studying patterns of genetic diversity during rapid range expansions.

### **Microsatellite development and characterization**

*Brachypodium sylvaticum* is a non-rhizomatous bunchgrass that is newly invasive in the Western United States (Clayton *et al.* 2002 onwards). It is usually diploid (Kahn & Stace 1999), self-compatible (Judd 1983), and is undergoing rapid range expansion in Oregon (Kaye 2001; Rosenthal *et al.* 2008).

To isolate microsatellite loci, I extracted DNA from three leaves from a single plant in Oregon using a DNeasy Plant Mini Kit (QIAGEN). I constructed a genomic library using Dynabeads (Invitrogen), following the protocol outlined in Hauswaldt & Glenn (2003) and available from T. C. Glenn (glenn@sre1.edu).

I digested genomic DNA with *Rsa*I, ligated it to Super SNX linkers, and hybridized the fragments to two different mixes of biotinylated oligonucleotides

(available from T. C. Glenn: Mix 2: (TG)12, (AG)12, (AAG)8, (ATC)8, (AAC)8, (AAT)12, (ACT)12; Mix 3: (AAAC)6, (AAAG)6, (AATC)6, (AATG)6, (ACCT)6, (ACAG)6, (ACTC)6, (ACTG)6). The library was double-enriched using Dynabeads (Invitrogen). Each oligonucleotide mix was hybridized to linker-ligated DNA fragments, then washed twice with 1 x SSC, 0.1% SDS at 50°C, and twice with 1 xSSC and 0.1% SDS at 55°C. After amplification, the enriched library was cloned into *E. coli* using a TOPO-TA cloning kit for sequencing, version J (Invitrogen). PCR products of the inserts were sequenced using M13 forward and reverse primers with BigDye Terminator v.3.0 or 3.1 (Applied Biosystems). T3/T7 primers yielded poor quality sequences. Sequenced PCR products were visualized on either a 310 or a 3100 capillary genetic analyzer (Applied Biosystems). I edited sequences in GeneScan (Applied Biosystems) and BioEdit v7.0.5 (Hall 1999). Sequences with sufficient flanking DNA and containing tandem repeats of five or more for dinucleotide repeats and three or more for tri- and tetranucleotide repeats were input into Primer3 (Rozen & Skaletsky 2000) to develop PCR primers with melting temperatures between 55-60°C that would yield fragments between 200-500bp.

The genetic library had about 2700 clones, stored in 12 96-well plates at -70°C. In a subset of clones from oligonucleotide Mix 2 (see above), 72 clones contained 40 unique fragments, and 15 fragments contained repeating elements. In Mix 3, 42 clones contained 35 unique fragments, with 20 fragments containing repeating elements. I randomly selected six plates to sequence; when any step in the process failed, I discarded the clone. In total, I successfully sequenced approximately

300 clones, and was able to develop primers for 46 loci. Twenty loci amplified cleanly and were polymorphic, with primer pairs designed to amplify between 160-300bp (Table 1).

Primers were initially tested on approximately 12 individuals from multiple populations in Oregon and Europe. Genomic DNA was extracted on a MixerMill (QIAGEN) using the Dneasy 96 Plant Kit (QIAGEN). For PCR, I used either an MJ Research P-100 thermal cycler or an Eppendorf Master Gradient cycler. Reactions were carried out in 10 $\mu$ l or 7.5 $\mu$ l reactions with HotStarTaq Master Mix (QIAGEN) containing 0.5 or 0.38 units HotStarTaq, 1.5 mM MgCl<sub>2</sub>, 200  $\mu$ M each dNTP, and 0.25 $\mu$ M each primer. After an initial denaturing step of 95°C for 15 minutes, the PCR profile was: 95°C for 30 seconds, 59°C for 30 seconds, 72°C for 45 seconds, cycled 30-35 times, with a final extension of 72°C for 2 minutes. For initial screening, fragments were visualized on 5% MetaPhor agarose. Fragments that amplified cleanly were initially tested for polymorphism on about 40 individuals from at least two Oregon populations and at least three European collections, and screened on MetaPhor agarose prior to developing labeled primers. Forward or reverse primers were labeled with Hex, Tamra, or 6-Fam, and used at a 1:10 ratio of labeled to unlabeled primer. Tamra-labeled primers were generally quite dim, and were re-labeled with either Hex or 6-Fam. Fluorescently-labeled fragments were visualized on an Applied Biosystems 310 genetic analyzer with ROX500 as the internal lane standard (Applied Biosystems).

For locus characterization, I assayed nine loci on 40 individuals from two populations: 24 individuals from the MacDonald Research Forest in Corvallis, OR, and 16 individuals from a USDA accession from Spain (PI 237792). I used Genepop on the Web (Raymond & Rousset 1995) to calculate heterozygosity,  $F_{IS}$  and  $F_{ST}$  values. Numbers of amplified alleles ranged from two to ten; two loci were monomorphic in Oregon. Four loci in Oregon and one locus in Spain deviated from Hardy-Weinberg equilibrium, and tested significant for heterozygote deficiency (Table 1); the locus deficient in Spain was not deficient in Oregon. Heterozygote deficiency could be due to null alleles (e.g. Chapuis & Estoup 2007), but in Oregon it is probably due to inbreeding and recent population colonization. Six loci had high  $F_{ST}$  values (Table 1). Several locus comparisons exhibited linkage disequilibrium in Oregon. This is not surprising considering the high  $F_{IS}$  values, the self-compatible nature of the plant, and the recent colonization of this population. One locus comparison in Spain was significant, but the same two loci were not linked in Oregon. These markers will be invaluable for both identifying source populations and for studying population genetic consequences of rapid range expansions.

I developed several multiplex reactions using FastPCR (Kalendar 2005) with the same reaction conditions mentioned above to minimize costs and time involved. A few loci amplified in a multiplex of four (3-2B2, 2-6H1, 2-6C3, and 3-4F9), but most loci worked best in combinations of two or three, such as: a) 3-2E3, 3-2G2 and 2-6C3, b) 2-3A1 and 3-4F9, and c) 2-6E8 and 2-6E6.

Table 2.1: Locus information for microsatellites characterized in *Brachypodium sylvaticum*

Reverse primers were labeled with fluorophores unless otherwise indicated (\*), in which case forward primers were labeled.

Locus (Genbank ID)	Primer sequences	Length	Motif and Fluorophore
2-3A1 (EF450746)	F: AGAGGGATTGCA TTGTCATCA G R: TTCGGA GGATA GCTTGGTCACTC	300	(GAA)21 6-FAM
3-4F9 (EF450748)	F: GCTCAGCTTGTCTTTTACCCATATC R: TTGCCA CCGCCTCTTAACATAC	251	(GATT)3 (GATT)3 HEX
3-2E3 (EF450751)	F: TTGCGA GGCA CGTATGTCTA R: ATCTTGTGCTTCATGGCAGA	160	(GGTT)3 6-FAM
3-2G2 (EF450752)	F: TACA GACGAA CCTGGCAGAC R: GCCTACCTCAACTTGCTTGG	174	(AAAC)5 HEX*, 6-FAM
3-2B2 (EF450754)	F: GACAACTCTACTGTGCATGAATTTG R: AGGCTTGGAGCTCATACCA G	122	(GTTT)5 FAM
2-6C3 (EF450756)	F: AGCAA CCACCAAACCCTTC R: CTCGTCGTCTCCAACCTCTC	218	(CT)16 6-FAM
2-6E6 (EF450757)	F: TATGAACCA CAA GCCCAGA G R: TCCATGTGCCTGAATCTTGA	225	(CAA)14 6-FAM
2-6H1 (EF450759)	F: ATGATCCCTGCATTCTCGTC R: CGTCGTTTCTGCTTGGATTT	160	(CTT)23 6-FAM, HEX
2-6E8 (EF450765)	F: CTGCTTCCTTGCCCACTAAC R: ATTTATGCCGTGTGGGA GAA	214	(GA)18 6-FAM, HEX



Table 2.2: Locus characterization information in two populations  
 Sample size (n), expected heterozygosity ( $H_e$ ), Observed heterozygosity ( $H_o$ ), Inbreeding coefficients ( $F_{IS}$ ), and differentiation ( $F_{ST}$ ) for one population each in Spain and Oregon.

Locus	Size Range	Number of Alleles	n	$H_e$ Spain	$H_o$ Spain	$F_{IS}$ Spain	$H_e$ OR	$F_{IS}$ OR	$H_o$ OR	$F_{ST}$
2-3A1	259-357	8	38	0.59	0.38	0.37†	0.54	0.5	0.27	0.41
3-4F9	247-251	2	40	0.06	0.06	0		Only one allele		-0.01
3-2E3	153-169	3	38	0.4	0.13	0.7		Only one allele		0.19
3-2G2	164-172	3	38	0.28	0.19	0.35	0.09	0.02	0.09	0.02
3-2B2	105-129	5	40	0.46	0.56	-0.22	0.49	0.92†	0.04	0.31
2-6C3	198-220	10	40	0.52	0.31	0.41	0.61	1.0†	0	0.41
2-6E6	195-247	9	40	0.81	0.63	0.23	0.59	0.86†	0.08	0.03
2-6H1	117-153	7	40	0.58	0.5	0.15	0.34	1.0†	0	0.53
2-6E8	192-228	5	40	0.49	0.5	-0.02	0.04	0	0.04	0.2

### **Chapter 3: Shifting dispersal modes at an expanding species' range margin**

Published as: Ramakrishnan AP, Musial T, Cruzan MB (2010) Shifting dispersal modes at an expanding species' range margin. *Molecular Ecology* 19, 1134-1146.

Used by permission.

#### **Abstract**

While it is generally recognized that non-contiguous (long-distance) dispersal of small numbers of individuals is important for range expansion over large geographic areas, it is often assumed that colonization on more local scales proceeds by population expansion and diffusion dispersal (larger numbers of individuals colonizing adjacent sites). There are few empirical studies of dispersal modes at the front of expanding ranges, and very little information is available on dispersal dynamics at smaller geographic scales where I expect contiguous (diffusion) dispersal to be prevalent. I used highly polymorphic genetic markers to characterize dispersal modes at a local geographic scale for populations at the edge of the range of a newly-invasive grass species (*Brachypodium sylvaticum*) that is undergoing rapid range expansion in the Pacific Northwest of North America. Comparisons of Bayesian clustering of populations, patterns of genetic diversity, and gametic disequilibrium indicate that new populations are colonized ahead of the invasion front by non-contiguous dispersal from source populations, with admixture occurring as populations age. This pattern of non-contiguous colonization was maintained even at a local scale. Absence of evidence for dispersal among adjacent pioneer sites at the edge of the expanding range of this species suggests that pioneer populations undergo an establishment phase during which they do not contribute emigrants for colonization of neighboring sites. My data indicate that dispersal modes change as the invasion

matures: initial colonization processes appear to be dominated by non-contiguous dispersal from only a few sources, while contiguous dispersal may play a greater role once populations become established.

## **Introduction**

Dispersal dynamics during range expansions and the resultant population genetic patterns are of interest to many scientists, including those studying human biology, epidemiology, historical biogeography, and invasive species biology. Much of our current understanding of dispersal dynamics during range expansions has been generated from observations of post-Pleistocene range expansions; the consequent genetic structure of such populations has been studied for a number of species (Comes & Abbott 1998; Hewitt 1996; Maroja *et al.* 2007; Masta *et al.* 2003; Petit *et al.* 1997; Soltis *et al.* 1997). Contiguous (diffusion) dispersal (i.e., a larger number of migrants moving short distances) is common in these studies, and low-frequency non-contiguous (long-distance) dispersal is thought to be an important component of range expansions over larger geographic areas (Clark *et al.* 1998; Levin 2003a).

The common model of dispersal dynamics, with frequent contiguous dispersal and infrequent non-contiguous dispersal, may not hold true at the extreme edge of an expanding range. Pioneer colonies at the forefront of a range expansion will be more isolated than populations in the main body of the species' range, and may not experience the same dispersal dynamics. If recently colonized populations also undergo an establishment phase during which they are unable to contribute migrants to range expansion (lag phase), source/sink metapopulation dynamics may prevail

(Thomson 2007), with a few source populations responsible for colonization and maintenance of the majority of pioneer populations.

Dispersal dynamics may change as a range expansion progresses from establishment phase to expansion phase. Changing dynamics in different phases of expansion could affect models of the consequences of range expansion for new mutations and overall patterns of population genetic variation. In this study, I assess patterns of colonization and dispersal processes contributing to a rapid expansion at the range edge of an invasive grass species. In my choice of organism and sampling, I attempted to form a comprehensive, widely-applicable model system in which to test hypotheses about expanding edge dispersal dynamics. *Brachypodium sylvaticum* (Hudson) Beauv. (Gramineae), or false brome, is a perennial bunchgrass, and is self-compatible and diploid (Judd 1983; Kahn & Stace 1999). This plant lends itself to evolutionary research with traits such as non-rhizomatous growth habit (Kaye 2001; A. Ramakrishnan, personal observation) and greenhouse suitability; these traits will simplify future investigations into fitness-related traits.

*Brachypodium sylvaticum* has a broad native range, spanning North Africa and Eurasia (Hitchcock *et al.* 1969). This grass appears to be in the early stages of expansion in Western North America, and has become noxious in the Willamette Valley region of Oregon over the past 20 years (Kaye 2001). It may have been first introduced 70 to 80 years ago by scientists in the United States Department of Agriculture for testing as a more productive range grass (Hull 1974; Rosenthal *et al.* 2008). Recombinant hybrids from the original plants have since become widespread in

mixed coniferous forest edges and understories, and it is listed as a noxious weed in western Oregon (Roy 2010).

*Brachypodium sylvaticum* has a large native range globally; however, its distribution in the United States is limited. It has naturalized and is common in Western Oregon and Southern Washington, with a few populations in Central California (Roy 2010). It has also naturalized in Virginia (Roy 2010). Initial introductions may have been for range grass experiments; my analyses focus on the Oregon populations. Recent genetic analyses of native and invasive populations of *B. sylvaticum* indicate that the original sources of the invasion in Oregon were derived from Western Europe (Rosenthal *et al.* 2008). Invasive populations of the grass are less genetically diverse than native populations both in terms of allelic richness and expected heterozygosity. Populations in Oregon were established from multiple introductions and appear to be recombinant genotypes, with several European regions contributing to their genomic composition (Rosenthal *et al.* 2008). It has been reported that *B. sylvaticum* (the seeds of which are covered in microscopic barbs) is dispersed by wild ungulates in the native range (Heinken & Raudnitschka 2002), and by humans via recreational activities such as fishing and logging equipment in Oregon (Kaye & Blakeley-Smith 2006).

Different dispersal patterns will have varying effects on the genetic structure of populations of *B. sylvaticum* during range expansion. Contiguous dispersal leads to a gradual increase in the geographic extent of a population, while non-contiguous dispersal will establish pioneer foci relatively far from the main population. Pioneer

foci established from non-contiguous colonization events may expand via contiguous dispersal after establishment (e.g. Petit *et al.* 1997). The term stratified dispersal is used in the latter situation (Shigesada *et al.* 1995). Dispersal modes as well as the number of newly introduced migrants from source populations are known to affect genetic structure (Bialozyt *et al.* 2006; Nei *et al.* 1975). Dispersal mode and dispersal rate produce various patterns of genetic differentiation that develop as a species expands its range (Fig. 1.2) (Clark 1998; Hewitt 1996; Hutchison & Templeton 1999; Le Corre *et al.* 1997; Shigesada *et al.* 1995; Skellam 1951).

Despite many theoretical investigations of dispersal during range expansions (Austerlitz *et al.* 1997; Bialozyt *et al.* 2006; Fix 1997; Ibrahim *et al.* 1996; Le Corre & Kremer 1998; Nei *et al.* 1975; Neubert & Caswell 2000; Petit *et al.* 2004; Shigesada *et al.* 1995; Skellam 1951; Whitlock & McCauley 1990), there are few empirical studies describing local scale dispersal dynamics at the edge of an expanding range using genetic markers (Dlugosch & Parker 2008b; Estoup *et al.* 2004; Husseneder *et al.* 2005; Ingvarsson & Giles 1999). A few studies have utilized phylogenetic techniques with molecular markers to identify colonization events and the genetic patterns generated by colonizing or early-successional species on islands (Erickson *et al.* 2004; Litrico *et al.* 2005), volcanic areas (Litrico *et al.* 2005; Yang *et al.* 2008), and forests (Franceschinelli & Kesseli 1999; Jones *et al.* 2006), but genetic characterizations of expanding range edges are rare.

Studies of population establishment generally assess levels of genetic variation within populations and allelic distributions among genotypes; however, gametic

(linkage) disequilibrium can also result when genetically distinct individuals play large roles in colonization and establishment of new populations (Jain *et al.* 1981; Slatkin 1994). When two genetically distinct populations act as sources for a newly founded population, random allele pairs in the new population will appear to be linked (Nei & Li 1973); allele pairs from certain loci will often (or rarely) appear together in any one individual. Such nonrandom associations are often observed in hybrid zones (Barton & Hewitt 1985; Cruzan 2005; Szymura & Barton 1986). This linkage is not necessarily due to physical chromosomal distance, but because random mating among individuals from the two sources has not yet broken apart the original genotypic signatures of the founder populations. Patterns of gametic disequilibrium could yield valuable insights into population dynamic processes.

In this study, I used nuclear microsatellite markers to examine patterns of dispersal and colonization over a relatively small geographic region at the leading edge of the invasive range of *B. sylvaticum*. Microsatellite markers have high polymorphism, which assists detection of dispersal events. Populations in my study area do not extend beyond roadsides; I sampled populations along three roughly parallel roads to produce a distribution of sample sites that allowed relatively independent assessment of the contributions of geographic distance and “road distance” to genetic relatedness among sites. The primary goals of this study were: a) to test whether roads are serving as migration corridors that facilitate range expansion; b) to reveal which sites are serving as sources for the local spread of this species, and c) to determine whether or not *B. sylvaticum* sites at the edge of the range are

spreading primarily through contiguous or non-contiguous dispersal events. I used Bayesian analyses of population clusters, assignment tests and analyses of isolation by distance to assess the contributions of contiguous and non-contiguous dispersal to expansion processes at the invasion front. Characterization of gametic disequilibrium allowed me to make inferences about colonization histories of sample sites.

## **Methods**

### *Sample Sites*

Large numbers of *B. sylvaticum* populations have been found in recently logged areas and along roadsides of adjoining regions. My study area spanned about 30km at the edge of this species' range in Oregon (Fig. 1.1), in mixed coniferous forest edges of low mountains. In the region I sampled, *B. sylvaticum* abundance and time since establishment decreases from west to east, due to recent invasion of the area by *B. sylvaticum* as indicated by land managers in the area (the False Brome Working Group, Corvallis, Oregon). Anecdotal evidence suggests that sites near the Foster City log processing pond (see Fig. 1.1) may have been established from the McDonald Research Forest near Corvallis, Oregon (D. Johnson, personal communication). This experimental forest is near Corvallis, one of the first North American sites where *B. sylvaticum* became established (Rosenthal *et al.* 2008).

### *Sampling regime*

Information from managers in the False Brome Working Group and personal observations indicate that plants in this area occur along roads. On the north side of Foster Reservoir, *B. sylvaticum* grows in forest understories, but in my study area, it



does not extend beyond roadsides. I collected samples in 2004 from three roads that run parallel to each other for about 5.5 km (Fig. 1.3). Occurrence of *B. sylvaticum* in the area was more or less continuous, though more sporadic towards the east, near the leading edge of the invasion. To facilitate analyses of allele frequency changes by distance, and to avoid biasing sample collection by density, I collected leaf samples from 15-25 plants every 1.5 km. Where plants were absent, such as along Highway 20 (Fig. 1.3), I collected wherever I observed plants. Plants were separated by at least 1.5 m where possible to avoid resampling the same individual. I sampled a total of 423 individuals from 18 sample sites.

Sample sites M1 through M5 are the only locations along Highway 20 where *B. sylvaticum* was found; this highway is a popular state highway with few turnouts (Fig. 1.3). Other roads I sampled included Wiley Creek Drive, Cedar Creek Road, and Whiskey Butte Drive. Wiley Creek Drive and Cedar Creek Road run roughly parallel to Whiskey Butte Drive for a portion of their length, and sample sites between roads were roughly equidistant to sample sites along roads (0.9-1.5 km, see Fig. 1.3). The three residential roads traverse rural residential neighborhoods where houses are generally on large lots of several hectares; these roads have more turnouts and appear to be subject to greater roadside disturbance than Highway 20, and they are adjacent to actively logged areas.

#### *DNA extraction and amplification*

A QIAGEN MixerMill was used to grind dried leaf samples; genomic DNA was extracted using the DNeasy 96 Plant Kit (QIAGEN). I used seven microsatellite

loci (2-3A1, 3-4F9, 2-6C3, 2-6E6, 2-6E8, 3-2E3 and 3-2G2) previously characterized for *B. sylvaticum* (Ramakrishnan *et al.* 2008), multiplexed in groups of two or three. I also used three loci not previously reported (Table 1). Reverse primers were labeled with fluorescent dyes and used in PCR reactions at a ratio of 1:10 with unlabeled primers.

Amplification reactions were performed in 6.0-7.5  $\mu$ L volume with HotStarTaq Master Mix (QIAGEN) containing 0.38 or 0.5 units HotStarTaq, 1.5 mM  $MgCl_2$ , 200  $\mu$ M each dNTP, 0.25  $\mu$ M each primer, and 2  $\mu$ g genomic DNA, and were conducted either on an MJ Research P-100 thermal cycler or an Eppendorf Master Gradient cycler. The initial denaturing step was at 95° C for 15 minutes; amplification cycles were 30-40 cycles of 95° C for 30 seconds, 57-60° C of annealing for 30 seconds and two minutes of extension at 72° C, with a final extension at 72° C for three minutes. Fragments were visualized on an Applied Biosystems 310 Automated Sequencer. Genotyper and Genescan software (Applied Biosystems) were used to score alleles. Raw peak lengths were graphed and binned into discrete allele sizes to ease determination of allele sizes. When a locus was amplified with different fluorescent markers in different samples, peak lengths were shifted by two to three basepairs. Allele sizes were corrected for such differences by comparing allele profiles for different dyes.

#### *Hardy-Weinberg equilibrium and site differentiation*

Samples that did not amplify PCR products successfully for more than two loci (34 samples out of 423, or 8%) were excluded from genetic analyses. I tested for

Hardy-Weinberg equilibrium with 3,000 steps burn-in and 100,000 steps in the Markov chain in Arlequin 3.11 (Excoffier *et al.* 2005). I calculated observed heterozygosities and Nei's unbiased gene diversity (Nei 1987) using Microsatellite Toolkit (Park 2001).

To describe patterns of differentiation in my study site, I compared variance components among clusters of sites (as defined by BAPS, see below), sites and individuals using an analysis of molecular variance (AMOVA),  $F_{ST}$  and  $F_{IS}$  (Excoffier *et al.* 1992) in Arlequin with 1000 permutations.  $F_{ST}$ -associated statistics assume equilibrium conditions and are often inaccurate for estimation of gene flow (Hutchison & Templeton 1999); however,  $F_{ST}$  remains useful as a method of estimating population differentiation (Neigel 2002).

#### *Population structure*

To define populations, sample sites were clustered using Bayesian genetic mixture analysis in BAPS (Corander *et al.* 2008). The BAPS program stochastically clusters sites, then splits clusters based on divergence between sites in a cluster (Corander & Marttinen 2006). This method is preferred to jointly inferring clusters and individual admixture when there are relatively few loci and when groups of samples are collected from distinct geographic locations (Corander & Marttinen 2006). I ran 10 iterations at each level of k (number of clusters) between one and sixteen. The optimal k was chosen by maximization of the marginal likelihood among multiple runs at different levels of k. To describe similarities among genetic clusters, I made a UPGMA diagram in BAPS using Nei's standard distance (Nei 1972).

### *Dispersal patterns*

Assignment tests are appropriate for non-equilibrium situations when recent dispersal events are of interest (Cornuet *et al.* 1999; Davies *et al.* 1999; Pearse & Crandall 2004). In order to describe patterns of migration among sites, I used a Bayesian assignment test as implemented in GeneClass2 (Bandouin *et al.* 2004; Piry *et al.* 2004) to assign individuals to sample sites using the method of Rannala and Mountain (1997); this method is preferred when some populations have not been sampled (Berry *et al.* 2004). Prior to calculating assignment probabilities, I calculated pairwise  $D_{LR}$  between sites (Paetkau *et al.* 2004) using the online calculator Doh available at <http://www2.biology.ualberta.ca/jbrzusto/Doh.php> (Brzustowski 2002).  $D_{LR}$  is based on the average likelihood of individuals from one site belonging to another site. Values under three are associated with low power to distinguish among the sites in question; I grouped all sites with  $D_{LR}$  values under three to improve assignment test accuracy. Probability computations were calculated as recommended by Paetkau *et al.* (2004), with 10,000 simulated individuals used in Monte Carlo resampling. Individuals with probabilities less than 0.07 were assumed to originate from unsampled source populations. Individuals with assignment to multiple sources (less than 0.05 difference in probabilities between putative sources) were assigned to both sites; if they assigned to their own site in addition to another site, I conservatively labeled them as belonging to the site where they were sampled.

### *Dispersal along roads*

To test whether or not dispersal followed a pattern of isolation by distance, I plotted pairwise  $F_{ST}$  against two distance measures: Euclidean distance and distance calculated using roads. The use of roads to calculate distance reflected the tendency of sites to be clustered along roadsides. Isolation by distance tests were performed using a Mantel test with 1000 permutations for significance testing in NTS YSpc v2.1 (Applied Biostatistics, Inc).

Statistical analyses may be able to detect a pattern of dispersal along roads if adjacent populations along roads are more similar than populations located on different roads. To specifically test for road-biased dispersal, I conducted a general linear model test in SPSS 12.0. I compared  $F_{ST}$  values between adjacent sites along two parallel roads to  $F_{ST}$  values between adjacent sites situated between the roads, using geographic distance as a covariate.

### *Genetic patterns of site history*

Sites recently colonized by a low number of founders can be expected to have low genetic diversity and high differentiation. Correlations between genetic diversity and longitude, and between average pairwise  $F_{ST}$  and longitude, were tested with Spearman's rank correlation in SPSS 12.0. Colonization history also affects patterns of interallelic gametic disequilibrium. I tested for gametic disequilibrium using MIDAS with Yates's correction for multiple comparisons (Gaunt *et al.* 2006). Detection of gametic disequilibrium was originally developed for low-polymorphism systems; however, multiallelic high-polymorphism loci are very useful for observing

disequilibria, as each allelic pair can potentially be associated at a higher frequency than expected (Zapata *et al.* 2001). Information is lost when gametic disequilibrium results are reported simply by locus, instead of the more detailed allele-based information available in multiallelic systems (Zapata *et al.* 2001).

Gametic disequilibrium is closely associated with levels of admixture in a population. I estimated individual admixture proportions using BAPS, based on the clustering analysis;  $p$ -values associated with each individual were calculated by using suggested values for admixture analysis (100 iterations per individual, 200 reference individuals simulated for each population, and 20 iterations per simulated reference individual). Individuals were considered admixed if they contained more than 10% contribution from a different cluster, with  $p < 0.05$ .

## **Results**

### *Genetic diversity and differentiation among sites*

The ten loci amplified from one to 16 alleles each (Appendix 1). Loci in many sites were out of Hardy-Weinberg equilibrium after adjustment for multiple comparisons in each site. The average Nei's unbiased gene diversity ( $H_z$ , equivalent to expected heterozygosity) ranged from zero to 0.37 per site, and was higher than average observed heterozygosity, which ranged from zero to 0.18 per site (Appendix 1).

In the AMOVA, over half the variation I observed was partitioned among sites, while 25% was partitioned among individuals and 20% within individuals (Table 2). All pairwise  $F_{ST}$  tests were significant; low values were 0.04 and 0.08 between M6

and M16, and M6 and M7 respectively. The highest value was 0.93, observed between three pairs of sites: M1 and M3, M1 and M11, and M3 and M4 (Appendix 2). Highly differentiated sites were found at the eastern side of my study area, where colonization was more recent than in the western sites.

### *Population Structure*

The BAPS Bayesian clustering method detected 13 genetic clusters of sites (Fig. 1.4A). Four clusters were composed of multiple sample sites, but only one such cluster grouped adjacent sample sites. The clustered sites M9, M10 and M11 are adjacent sites along Whiskey Butte Drive, roughly in the center of my study area (Fig. 1.3). The other three multiple-site clusters were composed of nonadjacent sites. All multiple-site clusters were among older sites (Fig. 1.4). I detected significant differentiation among genetic clusters ( $F_{ST}=0.430$ ).

Robustness of the inferred clusters can be described by the change in marginal likelihood when a site is moved from one cluster to another. Some sites could be moved from one cluster to another without drastically decreasing marginal likelihood: For example, M2 could be clustered with M9, M10 and M11, or it could be clustered with M5 or with M14 without lowering the likelihood more than -9. Clusters at the eastern side of my study area, on the other hand, were strongly supported; for example, the average change in marginal likelihood was -177 or lower for site M13 (Appendix 3).

The UPGMA diagram based on Nei's standard genetic distance illustrates similarities among these clusters. Clusters composed of multiple sites and with

relatively small changes in marginal likelihood are grouped in one clade, and are situated in the older part of my study area (Fig. 1.4C). The other clades, as highlighted in Figs. 1.4B and 1.4(C), are composed of more dissimilar sites, and tend to be situated towards the more recently colonized areas of my study. All clades of clusters are geographically noncontiguous, echoing the pattern observed in the initial BAPS clustering, where all but one genetic cluster of sites were composed of nonadjacent sample sites.

### *Dispersal Patterns*

The assignment test implemented in GeneClass2 is designed to detect recent migration. Twelve out of 153 pairwise site  $D_{LR}$  values were under three, indicating that the assignment test would have difficulty distinguishing among these sites. Grouping these sites led to construction of a cluster of ten sites to be considered a single population for purposes of the assignment test; these sites are the same as the sites in the center clade of the BAPS UPGMA diagram (Fig. 1.4B).

Most individuals in my study (95.4%) assigned to their home site. The quality index, a reflection of the likelihood scores for individuals in their home site, was 77.3%. Ten individuals had maximum likelihood scores under 0.07. These individuals likely originated from unsampled source populations, and were scattered throughout my study area (Table 3). All migrants from known sources originated from the grouped sites in the older part of my study area. Newly colonized sites such as M1, M3 and M4 did not appear to contribute any recent migrants (Table 3).



### *Isolation by distance and dispersal along roads*

I did not detect a significant overall pattern of isolation by distance at any scale using either geographic distance or distance along roads. When I tested among various contiguous groups of sample sites using Mantel tests, only sample sites along a portion of Wiley Creek Drive (M15, M16, M17 and M18) exhibited a significant pattern of isolation by distance ( $r=0.65$ ,  $p=0.00$ ).

When I compared pairwise  $F_{ST}$  values for adjacent sites either along roads or between roads using GLM in SPSS, there was an insignificant trend ( $p>0.5$ ) for smaller values between adjacent sites situated along roads (mean  $F_{ST}=0.347$ ,  $n=6$ ) than sites between roads (mean  $F_{ST}=0.419$ ,  $n=8$ ).

### *Genetic patterns of site history*

Newly colonized sites are expected to have low genetic diversity when dispersal is limited; they will also be genetically distinct due to founder effects. In my study, there was a significant correlation between genetic diversity and longitude (Fig. 1.5A) and between average pairwise  $F_{ST}$  and longitude (Fig. 1.5B) (Spearman's rank correlation tests,  $p<0.001$ ); genetic diversity decreased dramatically in the newer, eastern sites, and newer sites were more genetically distinct (Fig. 1.4). These observations are reflected in the number of allelic pairs available for testing of gametic disequilibrium: new sites had fewer allele pairs than old sites (Fig. 1.6A).

Further insights into population history can be gained by examining patterns of interallelic gametic disequilibrium. The proportion of allele pairs non-randomly associated (gametic disequilibrium) varied widely among sites (Fig. 1.6B). No two

loci pairs were linked in all sample sites, so the linkage I observe is unlikely to be physical linkage. For sites with gametic disequilibrium, the average absolute value of  $D'$  was 0.92 (range=0.55-1.0); a value of one indicates complete linkage. Sites in the extreme front of the expansion had no interallelic gametic disequilibrium, and the oldest sites had only a few allele pairs nonrandomly associated (Fig. 1.6B).

Intermediate sites varied greatly in observed interallelic disequilibrium; almost half the allele pairs in site M8 were associated more or less often than expected, while site M11 had no gametic disequilibrium at all.

When progeny are produced from parents from separate, genetically distinct populations, the progeny will show signs of admixture. The admixture test in BAPS identified ten individuals with significant genetic contributions from non-home clusters (outside the area in which the individual was sampled, Fig. 1.4A). Sites M8, M10 and M18 had the highest fractions of admixed individuals (Fig. 1.6C). No admixed individuals were detected in sites where there was no gametic disequilibrium. Sites where gametic disequilibrium was detected had varying fractions of admixed individuals (0-0.10). Patterns of genetic diversity and gametic disequilibrium are summarized in Table 4.

## **Discussion**

The *Brachypodium sylvaticum* invasion in Oregon is quite recent; this grass was introduced about 80 years ago, became invasive only in the last 20 years, and is still confined primarily to the Willamette Valley of Oregon (Kaye & Blakeley-Smith 2006). I was able to document the contributions of contiguous and non-contiguous

dispersal to establishment and invasion at the edge of the range because of the dispersal dynamics of *B. sylvaticum* and the small scale of my sampling (every 1-2 km for about 20 km along three roads). I found contiguous dispersal to be restricted to older areas of the invasion, while non-contiguous dispersal dominated newly colonized areas. Along local roads, most colonization could be explained by dispersal from sites within the area sampled, while most sites along the highway were colonized via dispersal from more distant sources. Close investigation of patterns of gametic disequilibrium enabled me to draw more detailed conclusions about colonization history of individual sites. Lack of dispersal from newly colonized sites suggests that new populations need to become established before they contribute to the pool of migrants, which would promote non-contiguous expansion at the leading edge of the geographic range, and may result in a secondary lag phase that slows the rate of invasion.

#### *Power of genetic analyses for inferring population history*

Populations that have recently been established from a few founders often have very low polymorphism. This will decrease power to detect departures from equilibrium for both linkage equilibrium and Hardy-Weinberg equilibrium (Guo & Thompson 1992; Morrell *et al.* 2005; Slatkin & Excoffier 1996). A few sites in my study with extremely low polymorphism and few genotypes failed to reject the null hypothesis for tests of Hardy-Weinberg equilibrium, but this is likely due low information content. Monomorphic sites such as M3 have no information available with which to test allelic associations and assumptions of equilibria; however,

combining cluster analyses with assignment tests and related analyses allows a rough reconstruction of the dispersal history of the area.

Bayesian clustering of genetic data is notoriously difficult, especially if admixture and genetic structure are inferred concurrently; the programs Structure (Pritchard *et al.* 2000) and InStruct (Gao *et al.* 2007) were hampered by multimodal genetic structure in my data (results not shown). In BAPS, the use of geographic information simplified cluster analysis, and enabled detection of structure despite the complex nature of my data. The number of clusters,  $k$ , is likely overestimated due to presence of family structure (Waples & Gaggiotti 2006); however, comparing changes in marginal likelihoods among clusters allowed me to make broader inferences regarding population structure.

Assignment test results may be affected by inbreeding and family structure; some computations in the Rannala and Mountain (1997) algorithm assume Hardy-Weinberg and linkage equilibrium, both of which are violated to various degrees in my dataset. I also analyzed my data using a method that does not assume equilibrium, the Cavalli-Sforza & Edwards' distance (1967). The results had a much lower quality index, and many individuals assigned to two or more sites equally (not shown). This indicates that the distance-based method was not able to distinguish among sites. Preliminary simulations by Cornuet *et al.* (1999) suggested that likelihood-based approaches still outperform distance-based methods even when some assumptions are violated.

### *Invasion dynamics*

Levels of genetic diversity I observed on a small scale in the non-native range of *B. sylvaticum* were comparable to those observed throughout Oregon (Rosenthal *et al.* 2008). Gene diversity in my study (0.163) was lower than that observed throughout the non-native range (0.203), and much lower than that observed in the native range (0.404). Similarly, the most recently colonized sites in my study area were much less diverse than the older sites I sampled. The pattern of loss of diversity I observed in my small scale study reflects structure observed at larger scales (in preparation).

There are very few studies that focus on genetic patterns at the expanding edge of an invasive species (Dlugosch & Parker 2008b; Estoup *et al.* 2004; Husseneder *et al.* 2005). The sites in my study exhibited high differentiation ( $F_{ST}=0.430$ ) with evidence of both contiguous and non-contiguous dispersal. High differentiation was also observed among populations in the invasive range of *Hypericum canariense*, a recently invasive plant with few non-native populations; however, the relative contributions of long-distance and contiguous dispersal to expansion of populations are unclear (Dlugosch 2006). High among-population differentiation was also found in the giant hogweed in Switzerland (Henry *et al.* 2009). Levels of differentiation among newly colonized cane toad populations in Australia were very low; this is consistent with the dispersal mechanisms observed in cane toads, which involve little if any long-distance dispersal events (Estoup *et al.* 2004). Among these studies, the study on cane toads is the only one to specifically address dispersal patterns; my results differ

significantly, as *B. sylvaticum* has distinctly different dispersal mechanisms than those used by cane toads.

Frequent non-contiguous dispersal can lead to low levels of differentiation among populations; this has been observed in several colonizing species (Erickson *et al.* 2004; Ingvarsson & Giles 1999; Jones *et al.* 2006; Litrico *et al.* 2005; Yang *et al.* 2008). In *Brachypodium sylvaticum*, non-contiguous dispersal is not so common as to homogenize sites. The high spatial structure I observed among recently-colonized sites could be affected by kin-structured colonization, as in *Quercus rubra* during primary forest succession (Jones *et al.* 2006). Low gene flow among pioneer sites would also contribute to high levels of differentiation; this was observed in a study of tropical forest succession with *Helicteres brevispira* (Franceschinelli & Kesseli 1999). My observation of a transition from high differentiation in new sites to relatively lower differentiation in old sites was not observed in the above studies, but is consistent with theory regarding development of age structure among demes in a metapopulation (Cruzan 2005; Pannell & Charlesworth 1999; Wade & McCauley 1988; Wakeley & Aliacar 2001).

Inferences of colonization history need not be based solely on observations of changes in differentiation among sites; gametic disequilibrium is also affected by population history. Genetic drift, admixture among distinct sources and selection will affect levels of gametic disequilibrium (Hill & Robertson 1968; Kruglyak 1999; Nei & Li 1973). Selection probably does not significantly influence the genetic patterns I observed, because I intensely sampled a geographically limited area, and

environmental differences among sites were minimal. Factors that would affect levels of gametic disequilibrium in my study are genetic drift and admixture; these are closely related, as dispersal from distinct sources will have a greater effect in recently established, low-diversity sites.

The sites of original introduction at Foster Dam have high genetic diversity and low levels of disequilibrium, likely due to ongoing migration from diverse sources. Sites in the middle of my study area display a range of characteristics; several sites have been recently established from a single source. Other sites have experienced significant amounts of dispersal from distinct sources, or have yet to recover from colonization by genetically distinct individuals, raising levels of gametic disequilibrium in these sites. Interallelic gametic disequilibrium is absent in the easternmost sites in my study area, consistent with recent colonization from a single source, though not necessarily the same source for each site. Site M2 is an exception; though adjacent to the most recent and least diverse sample sites in this area, it was either established around the same time as older sites such as M5 and M17, or it was established by a comparatively large number of founders. Site M8 is also an exception to the patterns I observe; it is adjacent to the oldest sites, but it was likely established relatively recently, and from several sources.

As new populations become established, they may reach a demographic threshold, or a threshold of genetic diversity, after which they become actively involved in colonization. If diversity-dependent source/sink dynamics are in force, a lag phase may result, where newly established populations do not contribute to the

establishment of other new populations, and range expansion is effectively slowed. No recent migrants were detected via assignment tests from newly established sites, and the admixture tests did not identify any allelic contributions between marginal sites. The large numbers of migrants detected from older source sites and the lack of dispersal among marginal sites support the hypothesis that pioneer populations in my study system are established via migration from older, established sources rather than from adjacent, newly colonized sites.

An apparent increase in the rate of non-contiguous dispersal as a site ages could simply be the result of a low overall dispersal rate. Low rates of dispersal would inhibit nascent sites from contributing migrants, due to small census sizes and the stochastic nature of dispersal. I observed ample dispersal among older sites in my study area, but there was no dispersal among adjacent, nascent sites. Either contiguous dispersal rates are low enough that initial sizes of sites are below a threshold size required for contribution of migrants, or nascent sites may be inhibited from contributing propagules to neighboring sites. The relationships between genetic drift, inbreeding depression, and colonization dynamics will become clearer as I analyze genetic and life history data from throughout the invaded range.

## **Conclusions**

Dispersal and colonization processes represent critical aspects of the ecology and population genetics of all species, but they are intrinsically difficult to study. Minimizing the confounding factors of multiple sources and unsampled populations is facilitated in species that have recently become invasive but are still restricted in



range. The *B. sylvaticum* invasion in Oregon has provided important insights into processes of dispersal, colonization, and range expansion. By examining sites at the fringe of the expanding range using genetic data, I found evidence that initial colonization events occur through non-contiguous dispersal; admixture from multiple sources then creates gametic disequilibrium that is subsequently reduced as the population ages. Later in the invasion process, populations with relatively high levels of genetic diversity contribute more emigrants. Populations begin to coalesce as contiguous dispersal expands the geographic area of previously established sites. These dynamics are suggestive of a lag period after initial colonization, during which levels of polymorphism are low and new populations do not contribute migrants to the invasion. Demographic studies are needed to clarify the relationship between colonization, establishment, migration, and a lag period. Further empirical research into population dynamics at the front of an invasion wave is necessary to improve our understanding of the complex processes behind the often explosive population growth and range expansion seen in invasive species.

### **Acknowledgements**

The first two authors contributed equally to this manuscript. I thank H. Machado, J. Picotte, B. Benz, P. Sochaki, and K. Anton for support in the laboratory, T. Cheeke-İçöz, D. Jolles, J. M. Rhode, M. Bahmer, S. Eppley, P. Sochaki, and six anonymous reviewers for editorial comments, and D. Rosenthal and J. M. Rhode for assistance

with data collection and helpful discussions on data analysis. This work was funded by USDA CSREES grant 2005-35320-15317 to M. B. Cruzan and A. P. Ramakrishnan.

Table 3.1: Previously uncharacterized loci

Locus	Genbank accession number	Repeat sequence and fluorofore	Amplified alleles size range in bp. (allele no.)	Primer sequences
3-4E8	EF450747	(CAAA) <sub>9</sub> 6-FAM, TAMRA	241-329 (7)	For: ACATGGTAA GAACCA GAATCGG Rev: TGAATTCGGCACGTCTGGATCC
3-2A7	EF450750	(ATCT) <sub>6</sub> HEX	166-182 (5)	For: CTTATGCCTTTCCA GGACGA Rev: CCTGCACTGCTAATCAACCA
2-3D12	EF450749	(GAGT) <sub>3</sub> (GAAAA) <sub>2</sub> 6-FAM	220-300 (3)	For: TGTGACAGCCATAGATATCGGC Rev: ATCACTCGTTAATATTCCTACTAGTG

59

Table 3.2: Analysis of molecular variance

ANOVA within and among genetic clusters and sites,  $P < 0.0001$ .

Sample	Sum of squares	Variance components	Percentage
Among clusters	590.16	0.6876	43.0
Among sites within clusters	45.90	0.1814	11.3
Within sites	422.51	0.4068	25.4
Within individuals	126.50	0.3252	20.3

Table 3.3: Assignment test results

Rows assign to columns. Low  $D_{LR}$  sites include M2, M5, M6, M7, M9, M10, M11, M15, M16, M17; these were the only sites detected as sources other than unsampled sites.

Sample Site	Unsampled		Total sample size
	Sites	Low $D_{LR}$ sites	
Low $D_{LR}$ sites	7	217	224
M8	1	2	24
M18	0	2	23
M12	0	6	24
M13	1	2	23
M14	0	1	19
M4	1	0	14
M3	0	0	23
M1	0	0	15

Table 3.4: Relationships of site diversity and interallelic disequilibrium to site history and assignment tests

Nei's unbiased gene diversity is  $H_z$ , and disequilibrium is represented as the fraction of allele pairs exhibiting significant disequilibrium in a site ( $\%D$ ).

	Low Disequilibrium ( $\%D < 0.05$ )	High Disequilibrium ( $\%D > 0.05$ )
Low Diversity ( $H_z < 0.1$ )	Recently established site colonized from a single source. (M1 <sup>‡</sup> , M3 <sup>‡</sup> , M4 <sup>‡</sup> , M11, M12 <sup>‡</sup> , M14 <sup>*‡</sup> )	Recently established from a small number of founders; multiple sources (M13 <sup>‡</sup> )
Intermediate Diversity ( $H_z = 0.1-0.25$ )	Established site, source for other sites in my study (M17, M18 <sup>‡</sup> , M2, M5)	Recent admixture; arrival of unrelated genotypes in a somewhat recently established site; act as sources (M8 <sup>‡</sup> , M9, M10)
High Diversity ( $H_z > 0.25$ )	Established site, source for other sites in my study (M6, M7, M15, M16)	Not observed

\*M14  $H_z = 0.11$ ; <sup>‡</sup>No emigrants detected

Fig. 3.1: Distribution of *Brachypodium sylvaticum* in Oregon, as reported by managers and by personal observation  
Darker shades indicate areas where *B. sylvaticum* is more abundant.

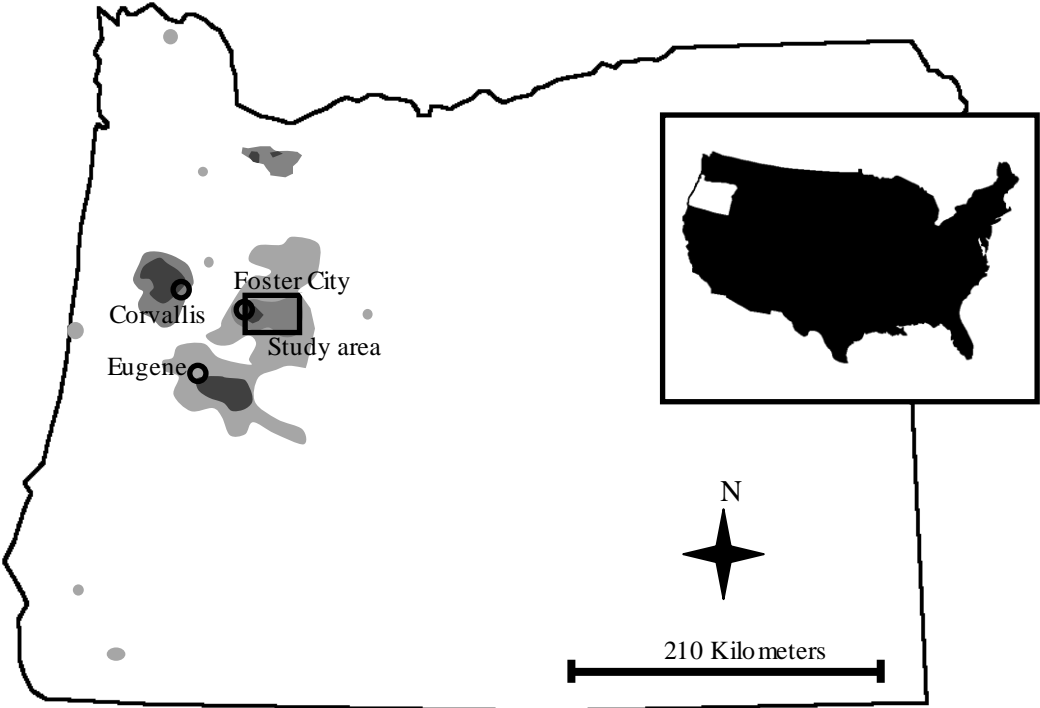


Fig. 3.2: Hypothetical populations experiencing different types and amounts of dispersal

Populations homogenized by high amounts of dispersal (A); populations with moderate amounts of dispersal, exhibiting a classic isolation by distance pattern (B); dispersal dominated by non-contiguous dispersal from a single source population (C); stratified dispersal where older populations are connected by dispersal and exhibit isolation by distance, while younger populations have recently been colonized by non-contiguous dispersal from a source (D). Shades of grey indicate population differentiation.  $F_{ST}$  x geographic distance patterns adapted from Hutchison and Templeton (1999).

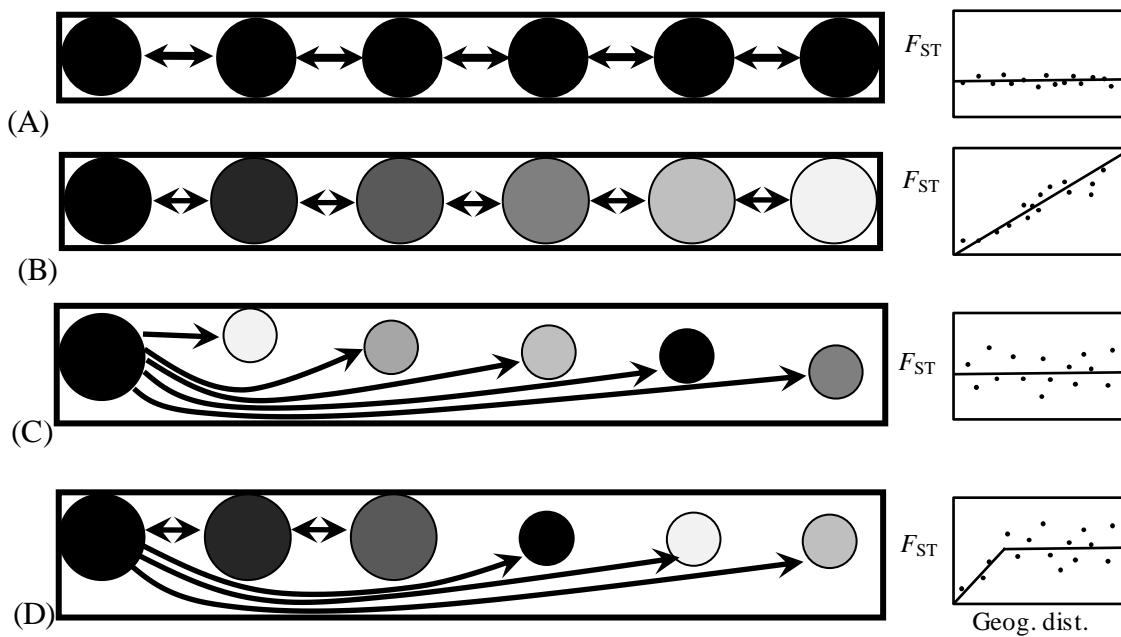


Fig. 3.3: Map of study area including Foster City and a log-processing pond. Open triangles are sample sites, the open circle is Foster City, and sites where *B. sylvaticum* has been observed but was not sampled are marked 'x'.

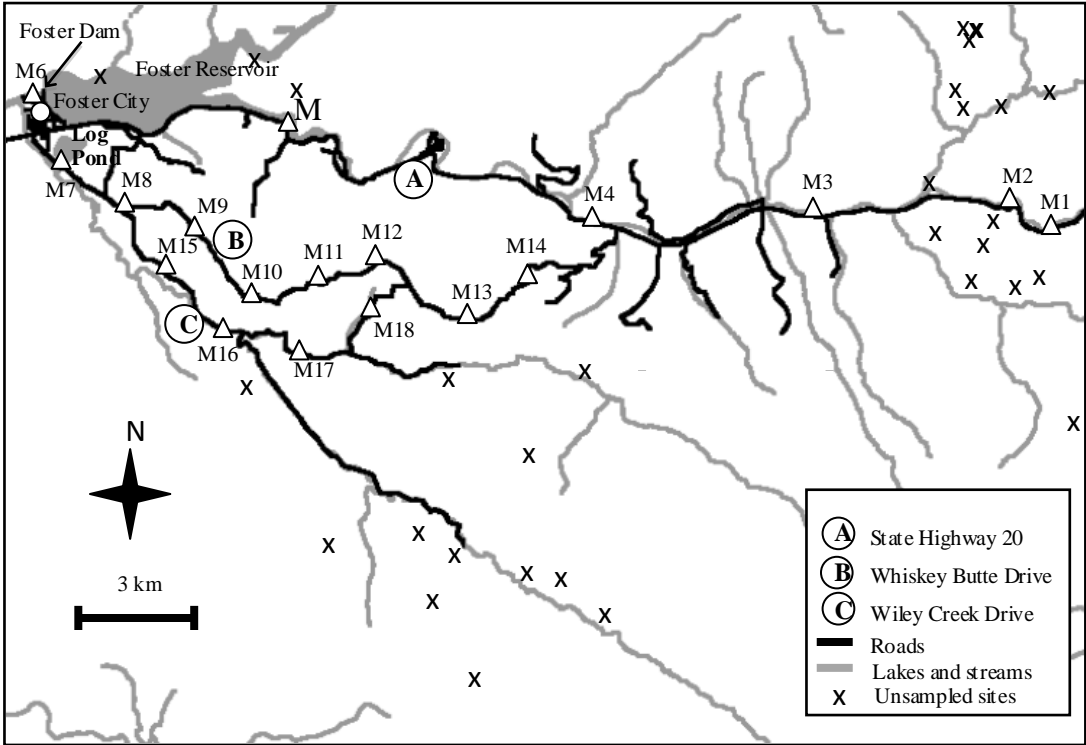


Fig. 3.4: Population similarities

Cluster assignments from BAPS, and admixed individuals identified from BAPS admixture analysis ( $p < 0.05$ ) (A). UPGMA on Nei's  $D_a$  pairwise distance between BAPS-identified genetic clusters (B): All sites that were grouped during cluster analysis are in one clade (light grey box). Colonization dynamics at the edge of the range (not to scale) (C): Marker size is proportional to genetic diversity ( $H_Z$ ). Dashed lines are roads.

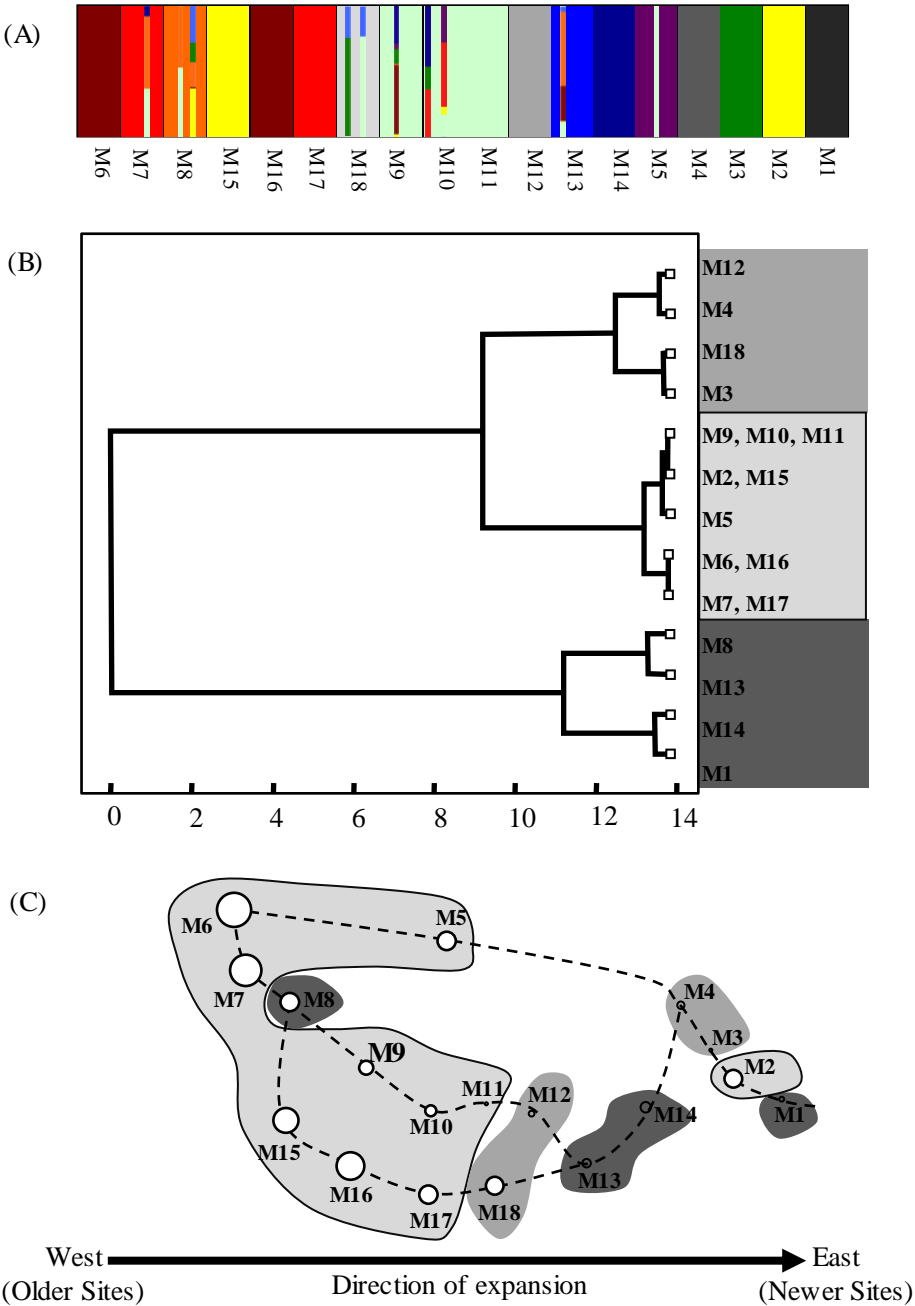




Fig. 3.5: Diversity, heterozygosity and expansion direction

Relationships between site age (approximated using longitude) and observed and expected heterozygosities ( $H_o$  and  $H_z$ , respectively) (A), and average pairwise  $F_{ST}$  (B). Site names on the x-axis are ordered strictly by longitude. Open site names are on Wiley Creek Drive, boxed site names are on Whiskey Butte Drive, and shaded names are on Highway 20.

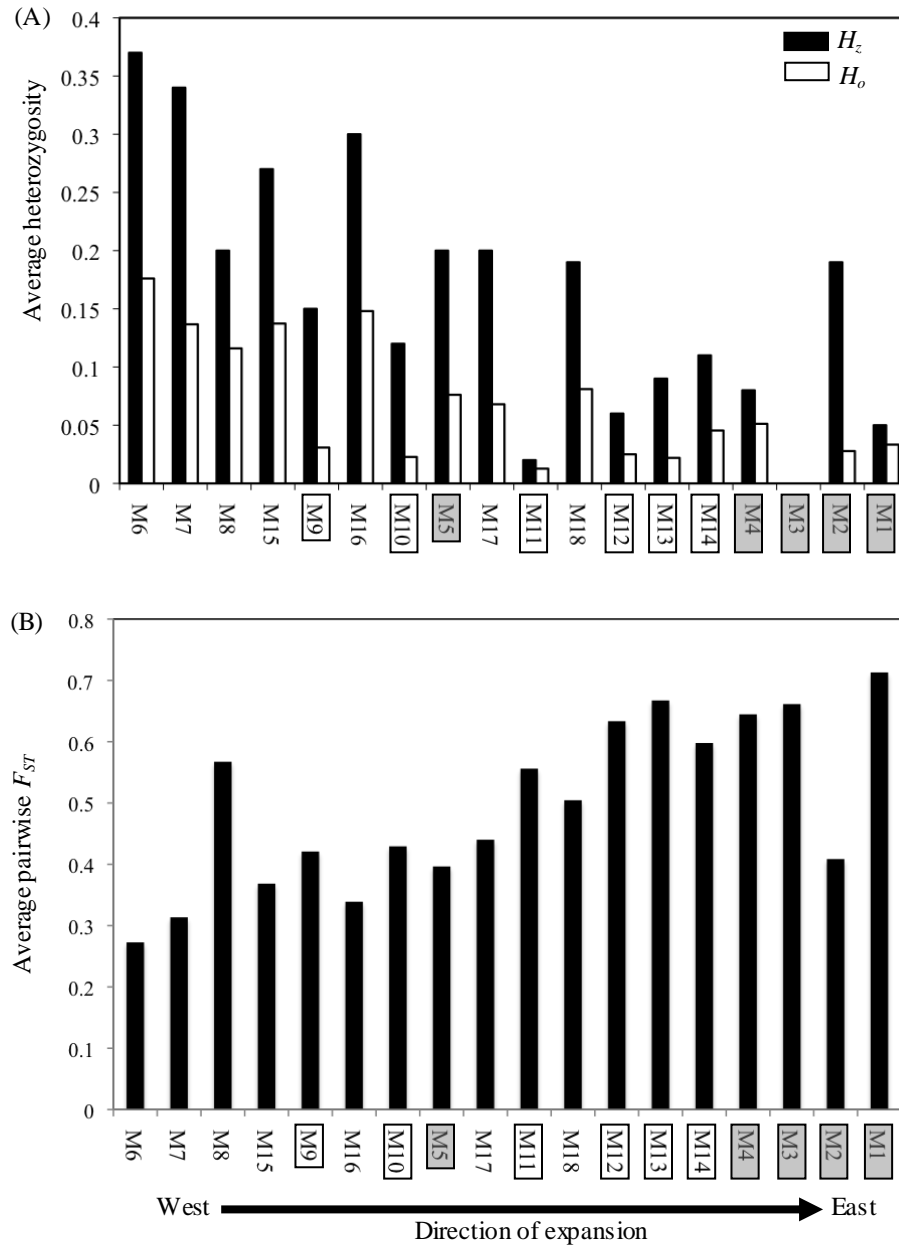
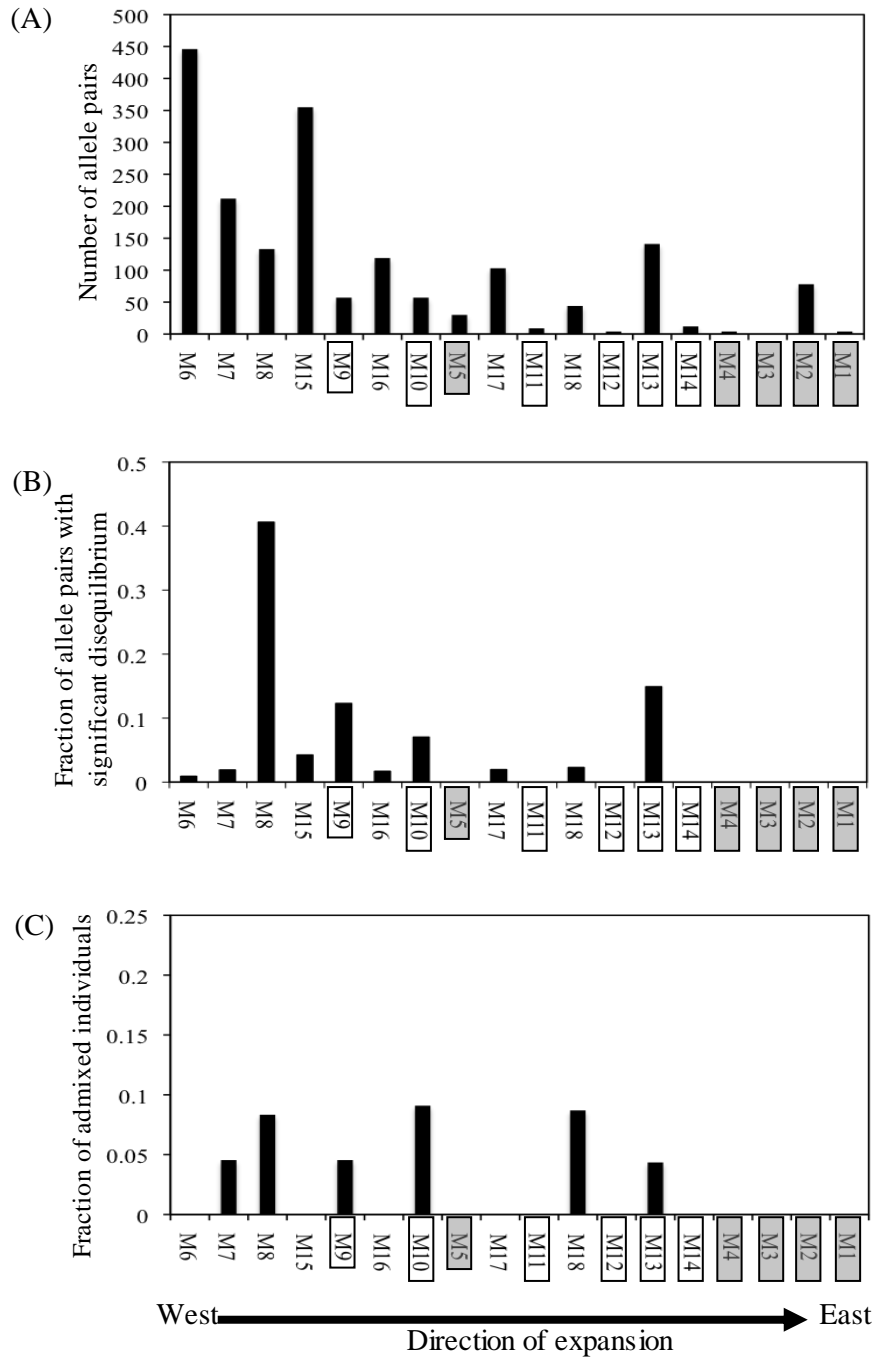


Fig. 3.6: Interallelic gametic disequilibrium

Number of allele pairs (A). Percentage of allele pairs testing positive for gametic (linkage) disequilibrium (B). Fraction of admixed individuals, identified using BAPS admixture analysis (C). Open site names are on Wiley Creek Drive, boxed site names are on Whiskey Butte Drive, and shaded names are on Highway 20.



## **Chapter 4: Founder effects and bottlenecks strongly influence colonization dynamics in a nascent invasive species**

### **Abstract**

I studied patterns of colonization, dispersal and trait evolution in the newly invasive grass *Brachypodium sylvaticum* using eight microsatellite markers. Successful colonization events in this species have led to rapid range expansion in Oregon. In this study, I hypothesize that dispersal patterns throughout the invasive range will be similar to patterns previously observed on a small scale, that the correlation between genetic diversity and differentiation will be consistent with bottlenecks during colonization, and that selfing rates will be high in peripheral populations. Results indicate that genetic patterns of colonization dynamics were similar at small and large scales; patchy genetic patterns formed by colonization events in *B. sylvaticum* decrease as populations age, likely due to increased migration among more established populations. Contribution of sources to colonization events was highly imbalanced: one of the two sites of introduction in Oregon (Corvallis, OR) was responsible for colonizing most of the peripheral populations I studied. Populations in the Eugene area as yet do not contribute high numbers of colonists, though this may change as the invasion progresses. I found evidence that higher selfing rates may be favored during colonization events; however, potential selfing rates of new populations are likely dependent on selfing rates in source populations. This study presents evidence that lower levels of population differentiation could be indicative of range expansion and evolution of invasiveness in introduced species.

## **Introduction**

The history of all species includes some periods of range expansion, but scientists rarely have the opportunity to observe rapid expansion in progress. Knowledge of evolutionary processes that can precipitate range expansion may help combat invasive species as well as promote population growth in endangered species (Wilson *et al.* 2009). The process of range expansion begins when individuals of a species are introduced to a novel environment or when suitable habitat becomes available. New colonists must be able to survive despite presence of competitors, herbivores, and a limited reproductive pool (Baker 1955). Invasive species, characterized by aggressive range expansion in non-native environments, provide an opportunity to examine the ecological and genetic factors affecting successful colonization events.

Many studies have investigated evolutionary genetic processes in invasive species, often relying on post-hoc phenotypic and genetic comparisons between putative source populations and widespread, invasive populations (Aketarawong *et al.* 2007; Astanei *et al.* 2005; DeWalt & Hamrick 2004; Saltonstall 2003) reviewed in (Bossdorf *et al.* 2005; Dlugosch & Parker 2008a; Ellstrand & Schierenbeck 2000; Sax *et al.* 2007) but see (Dlugosch & Parker 2008b; Estoup *et al.* 2004). Evolutionary processes and rates may differ between well-established and recently-colonized populations (Davis *et al.* 2004; Phillips *et al.* 2008; Travis & Dytham 2002), potentially confounding inferences about range expansion dynamics. Examining evolutionary processes in recently introduced, actively expanding species can provide

useful insights concerning successful geographic expansion (Dlugosch & Parker 2008b; Ramakrishnan *et al.* 2010).

I explore the evolutionary consequences of colonization processes in *Brachypodium sylvaticum*, a bunchgrass that was introduced to Western North America in the early 1900's (Fletcher & Swaller 1939; Roy 2010). Native to Europe and Asia, it was naturalized in parts of the ME by 1966 (Roy 2010), and invaded areas are as yet restricted to a few locations, simplifying intensive collection throughout the non-native range. *Brachypodium sylvaticum* is a long-lived perennial in gardens (C. A. Stace, personal communication), seeds can germinate in summer or fall conditions (De Frenne *et al.* 2009; Long 1989), and plants can reproduce after the first winter (personal observation). *Brachypodium sylvaticum* is diploid and self-compatible (Hasterok *et al.* 2004). Initial genetic research detected at least two distinct introductions of *Brachypodium sylvaticum* into Oregon (Rosenthal *et al.* 2008). Because rapid geographic expansion has only recently begun in this species, genetic patterns I observe may elucidate evolutionary mechanisms by which any species could become invasive.

Patterns of genetic differentiation during initial stages of range expansion can be affected by several factors. Important components of range expansion dynamics include the frequencies of long-distance and diffusion dispersal, the extent of bottlenecks and founder effects during colonization, and the potential for admixture among genetically distinct sources. The relationship between population diversity and average pairwise population differentiation may elucidate important colonization

dynamics. In addition, increased rates of self-pollination in newly colonized populations may help new populations become established when few congeners are present. Characterization of the contributions of these traits to a nascent range expansion may help predict potential for development of rapid range expansion in other species.

Long-distance and diffusion dispersal can have distinct effects on genetic diversity. In the absence of selection, frequent rates of either long-distance or diffusion dispersal will tend to homogenize genetic diversity (Epperson & Allard 1989; Hutchison & Templeton 1999; Le Corre & Kremer 1998). If dispersal is limited, diffusion dispersal will lead to a pattern of isolation by distance, where nearby populations are less differentiated than distant populations (Balloux & Lugon-Moulin 2002; Levin *et al.* 2003; Nichols & Hewitt 1994; Paetkau *et al.* 2004; Slatkin 1993; Walker *et al.* 2003; Wright 1943). On the other hand, if long-distance dispersal is common, there may be little or no relationship between genetic similarity and geographic distance. If long-distance dispersal is responsible for colonization events ahead of an expanding wave, a mosaic of distinct populations will result when dispersal is limited (Dlugosch & Parker 2008b; Ficetola *et al.* 2008). This patchy genetic pattern can then persist for many generations, unless it is swamped by high rates of gene flow among populations (Le Corre & Kremer 1998).

Bottlenecks and founder events during colonization of new populations may strongly influence successful establishment rates (Newman & Pilson 1997; Saccheri *et al.* 1998). Low initial levels of genetic diversity can increase the probability of

extinction for a new population (Ahlroth *et al.* 2003; Crawford & Whitney 2010), and can also increase differentiation among newly-colonized populations (Pannell & Dorken 2006; Wade & McCauley 1988). Many separate introduction events may be necessary before a species successfully expands its range, especially if few individuals are involved in each introduction (Drake & Lodge 2006). Although stochastic processes in evolution are widely acknowledged, there have been few opportunities to empirically explore the consequences of such processes for range expansion dynamics and invasion success (Huey *et al.* 2005).

Multiple introductions during colonization of invasive species can drastically alter patterns of genetic diversity in invasive vs. native populations, contributing to increased intrapopulation diversity in many populations of invasive species (Dlugosch & Parker 2008a; Kolbe *et al.* 2007; Petit *et al.* 2004). Admixture among genetically distinct populations is common in invasive species (Dlugosch & Parker 2008a), and may affect the ability of a species to establish new populations in a novel environment due to released additive genetic variation (Ellstrand & Schierenbeck 2000; Kolbe *et al.* 2004). The range expansion of *Brachypodium sylvaticum* in Oregon may be strongly influenced by genotypes formed by admixture between distinct European sources (Rosenthal *et al.* 2008); further admixture may increase the ability of *B. sylvaticum* to evolve novel traits during subsequent range expansion.

The pattern of correlation between genetic diversity and average pairwise population differentiation has rarely been investigated in field populations during initial colonization of invasive species (Ramakrishnan *et al.* 2010). This relationship

could help elucidate patterns of colonization dynamics (Pannell & Dorken 2006). For example, if colonization events involve few numbers of founders, and numbers of migrants increase proportionally as populations grow in density and/or size, then genetic diversity will increase as populations age. If dispersal is limited, then population differentiation would be high among new populations due to random sampling of a limited number of individuals from source populations; however, as populations merge and migrant numbers increase, differentiation would decrease (Fig. 4.1a). Species traits could affect this expectation: a species with high levels of selfing and limited dispersal could exhibit high levels of differentiation throughout an invasive range (Fig. 4.1b). On the other hand, if newly colonized populations do not experience significant bottlenecks or founder effects due to high dispersal rates, especially if long-distance dispersal rates are high, then pairwise differentiation levels would be low whether populations were newly-colonized or well-established (Fig. 4.1c). It is difficult to imagine a scenario where a single species' range could have both low-diversity and high-diversity populations while maintaining constant levels of differentiation (dotted lines Fig. 4.1b and 4.1c). It may be possible to observe a negative association between average pairwise population differentiation and genetic diversity if a majority of colonists are composed of a cohort of genetically similar individuals adapted for colonization. In this scenario, differentiation would be low among new populations, while differentiation could increase between populations as populations become established especially in the presence of selection in contrasting environments, if there is some sort of selective sweep occurring (Teschke *et al.*



2008)(Fig. 4.1d). Examining the relationship between diversity and differentiation during early stages of range expansion can give useful clues about the dispersal dynamics of a region.

Founder effects, bottlenecks, and admixture can affect the evolution of selfing as well as patterns of molecular genetic diversity. Populations at the edge of an expanding range in several self-compatible species have been shown to have increased selfing rates (Arnaud-Haond *et al.* 2006; Barton & Charlesworth 1984; Levin 2003b; Pannell & Dorken 2006), but see Herlihy and Eckert (2005). More studies are needed to investigate the evolution of traits in areas where an invasive species is in the nascent stages of range expansion (e.g., Dlugosch & Parker 2008b; Phillips *et al.* 2008), and at early stages of expansion into novel habitats (Leger *et al.* 2009; Parker *et al.* 2003; Thomson 2007).

In a previous study performed on a small geographic scale (~30km) in *Brachypodium sylvaticum*, I found that long-distance dispersal was common during colonization events, and that colonizing individuals generally originated from distant source populations rather than from adjacent populations (Ramakrishnan *et al.* 2010). Diffusion dispersal was more common in older, established populations than in new populations. Genetic bottlenecks often occurred during colonization, as newly colonized populations had lower genetic diversity and higher levels of genetic differentiation than older populations (Ramakrishnan *et al.* 2010). The small-scale study helped set some expectations for patterns I might see throughout Oregon.

In the current study I investigate four hypotheses. First, dispersal dynamics observed on a small scale will be representative of range-wide dispersal patterns. Second, new populations will likely exhibit high differentiation due to founder effects, bottlenecks and isolation, and some populations will experience admixture between distinct sources. Third, well-established populations that contribute genotypes to the migrant pool will be less differentiated and more genetically diverse than new populations if dispersal rates increase as populations grow. Fourth, I hypothesize that selfing rates will be lower in older, high-diversity populations than in newly-colonized populations. I use levels of linkage disequilibrium, clustering analyses, assignment tests, and multilocus progeny analysis to investigate differences in patterns of genetic diversity, dispersal, and selfing rates.

I compared levels of genetic structure and selfing rates among older source populations for two distinct areas of introduction (Corvallis, OR and Eugene, OR) and among newer, peripheral populations (populations that were geographically isolated from the primary introductions) throughout Oregon. The Eugene and Corvallis areas are sites where *B. sylvaticum* was first observed in Oregon; genetic data also indicate presence of two genetically distinct introductions. The genotypes spreading in Oregon are recombinants from several sites in Europe (Rosenthal *et al.* 2008), but the Rosenthal (2008) study did not attempt to characterize dispersal dynamics in Oregon. To determine the source(s) of each peripheral population, I conducted assignment tests and Bayesian clustering analyses. Populations with hybrid individuals were identified through assignment tests, calculation of a hybrid index, and Bayesian admixture

analyses. Evaluation of selfing rates was conducted by analyzing field-collected seeds. My results suggest that invasion dynamics can be chaotic, involving founder events, admixture, and lag times. Any one of these factors has the potential to drastically affect the evolutionary trajectory of an invasive species.

## **Materials and Methods**

### *Populations*

I sampled populations throughout Oregon to represent both central and peripheral regions of its current range. The False Brome Working Group is a large network of land managers interested in tracking the spread of *Brachypodium sylvaticum*. Working with these managers, several major areas of infestation were identified. In 2004, the only known *Brachypodium sylvaticum* populations in Oregon were in three major areas: Corvallis, Eugene, and Sweet Home (Foster City). The McDonald Research Forest, near Corvallis, is one of the first sites where the grass was introduced and became invasive. Herbarium records from Eugene indicate that the grass was established there around the same time it was found in Corvallis, though more detailed history of the invasion in the Eugene area is not available. Sites near Foster City may have first been established via logs sent to be processed from the McDonald forest (Ramakrishnan *et al.* 2010). There was also an infestation in the San Francisco Bay area of California that represents an independent introduction from Europe (Rosenthal *et al.* 2008).

I sampled populations within several main areas (Fig. 4.2): the McDonald forest (five populations: C1, C2, C3, C4, C5), the forests near Eugene (seven

populations: E1, E3, E4, E6, E7, E8, E9), and the area near Foster City (3 populations: M6, M11, M16). Foster City sites were included in a previous study where I compared genetic patterns among groups of samples taken every 2km along transects (roads) (Ramakrishnan *et al.* 2010). Samples from one population near San Francisco, California (S3) that is genetically distinct from Oregon populations (Rosenthal *et al.* 2008) was also included in the current study. Peripheral populations of *Brachypodium sylvaticum* were reported from various places throughout Oregon. I sampled eight of these sites (C10, C6, C7, C8, CY, S1, Metolius, E15) including sites at the extremes of the known distribution (Metolius, E15), in order to compare recently established sites to source populations. I collected 2-3 leaves from each plant for 25 plants from each site when possible. Plants were at least 1.5m apart when possible to avoid resampling the same plant. To test for within-year genetic variation, I re-sampled three sites (E1, E4, Metolius) in July 2004 and in September 2004. Leaves were dried on silica gel prior to DNA extraction.

To collect progeny for selfing rate analysis, I collected a single mature spike from 15 individuals in each of 10 Oregon populations distributed throughout the invasive range (C10, C1, C2, C4, C6, C7, E1, E4, E9, Met). Seeds were germinated in trays in the greenhouse and placed outside in early spring 2005 following germination. Eight to 16 seedlings per family from 12 Oregon populations (mean 12.8 families per population) were randomly selected and repotted into individual pots. Leaves were harvested, frozen in liquid nitrogen and stored in -80C prior to DNA extraction.

### *DNA extraction and amplification*

A QIAGEN MixerMill was used to grind leaf samples; genomic DNA was extracted using the DNeasy 96 Plant Kit (QIAGEN). I used eight unlinked microsatellite loci (2-6C3, 2-6E6, 2-6E8, 2-6H1, 3-2B2, 3-2E3, 3-2G2, 3-4F9) previously characterized for *B. sylvaticum* (Ramakrishnan *et al.* 2008), multiplexed in groups of two or three. Reverse primers were labeled with fluorescent dyes and used in PCR reactions at a ratio of 1:10 with unlabeled primers.

Amplification reactions were performed in 6.0-7.5  $\mu$ L volume with HotStarTaq Master Mix (QIAGEN) containing 0.38 or 0.5 units HotStarTaq, 1.5 mM  $MgCl_2$ , 200  $\mu$ M each dNTP, 0.25  $\mu$ M each primer, and 2  $\mu$ g genomic DNA, and were conducted either on an MJ Research P-100 thermal cycler or an Eppendorf Master Gradient cycler. The initial denaturing step was at 95° C for 15 minutes; amplification cycles were 30-40 cycles of 95° C for 30 seconds, 57-60° C of annealing for 30 seconds and two minutes of extension at 72° C, with a final extension at 72° C for three minutes. Fragments were visualized on an Applied Biosystems 310 Automated Sequencer. Genotyper and Genescan software (Applied Biosystems) were used to score alleles. Raw peak lengths were plotted on a graph and binned into discrete allele sizes to simplify determination of allele sizes. When a locus was amplified with different fluorescent markers in different samples, peak lengths were shifted by two to three basepairs. Allele sizes were corrected for such differences by comparing allele profiles for different dyes.

### *Descriptive genetic analyses and population structure*

Individuals that amplified fewer than six loci were excluded from analysis (49/663). Genotypes were compared among progeny and parental plants using seven loci (see “Evolution of outcrossing rate” below) to estimate scoring error rates. I compared average number of alleles per locus between parental data and progeny in 11 populations to observe any differences made by increased sample sizes. I tested for Hardy-Weinberg equilibrium with 1,000 dememorization steps and 100,000 steps in the Markov chain in Arlequin 2.0 (Excoffier *et al.* 2005). I calculated observed heterozygosities and Nei’s unbiased gene diversity (Nei 1987) using Microsatellite Toolkit (Park 2001). Effective numbers of alleles were calculated in PopGene 1.31 (Yeh *et al.* 1999). I tested for isolation by distance using a Mantel test in NTS YSpC 2.21c (Applied Biostatistics, Inc.) with 500 permutations.

To test for differences in population structure between the Corvallis and Eugene invasions, I used AMOVAs (Excoffier *et al.* 1992) in Arlequin 2.0 (Schneider *et al.* 2000) to assess partitioning of within- and among-population variance for each region of introduction, and within and among all other populations. Significance was tested with 1000 permutations. To estimate levels of differentiation between and among populations, I calculated Hedrick’s standardized  $G_{ST}$ , indicated as  $G'_{ST}$  (Hedrick 2005), using SMOGD (available at <http://www.ngcrawford.com/django/jost/>). I used a standardized version of  $G_{ST}$  because nonstandardized measures of  $G_{ST}$  were unsuitable for my purposes; populations with unique alleles and similar levels of heterozygosity can have very low

levels of differentiation as measured by  $G_{ST}$  even if they have no alleles in common (Hedrick 2005; Jost 2008).

Some populations were missing data for an entire locus due to experimental error (locus 3-2B2 in pop C10, locus 2-6H1 in pop M1, loci 2-6E6 and 2-6E8 in pop E3, and locus 3-2E3 in pop E15). For pairwise  $G'_{ST}$  comparisons involving these populations, I excluded the missing loci. Three other studied populations had over 50% of individuals missing one or two loci (17/25 individuals in population M6 missing locus 3-2B2; 19/31 individuals in C2 and 15/30 individuals in E4 missing loci 2-6E6 and 2-6E8). For pairwise  $G'_{ST}$  comparisons involving these populations, I excluded the individuals missing the loci in question.

Populations with increased genetic diversity may be older than populations with little or no diversity (Haag *et al.* 2005; Wade & McCauley 1988). To test for correlation between relative population age (approximated by the effective number of alleles in a population) and differentiation measured as  $G'_{ST}$ , I used linear regression between pairwise  $G'_{ST}$  and effective number of alleles in SigmaPlot 8.02a.

I tested for population clustering using nonmetric multidimensional scaling analysis (NMDS) and Bayesian cluster analysis. For NMDS, I first calculated Nei's standard genetic distance in Populations 1.2.30 ([http://bioinformatics.org/project/?group\\_id=84](http://bioinformatics.org/project/?group_id=84)). The resulting distance matrix was reduced to three dimensions using NMDS in NTSYSpc 2.21c (Applied Biostatistics, Inc.). NMDS iterates points in three dimensions while attempting to maximize fit of the calculated distances among points to the original matrix. Fit of the distance matrix

to the resultant graph was measured using default stress options. I tested for presence of genetic clusters using Bayesian methods in BAPS (Corander *et al.* 2008). The BAPS program stochastically clusters populations, then splits clusters based on divergence between populations in a cluster (Corander & Marttinen 2006). This method is preferred when there are relatively few loci and when groups of samples are collected from distinct geographic locations (Corander & Marttinen 2006). I ran 10 iterations at each level of k (number of clusters) between one and 24. The optimal k was chosen by maximization of the marginal likelihood among multiple runs at different levels of k (Corander & Marttinen 2006).

#### *Gametic disequilibrium*

Establishment of a new population from multiple, distinct sources will temporarily increase levels of interallelic gametic disequilibrium (GD). I tested for GD using MIDAS with Yates's correction for multiple comparisons (Gaunt *et al.* 2006). Detection of GD was originally developed for low-polymorphism systems, however, multiallelic high-polymorphism loci are very useful for observing disequilibria, as each allelic pair can potentially be associated at a higher or lower frequency than expected (Zapata *et al.* 2001).

#### *Identification of sources*

I used Bayesian assignment tests as implemented in GeneClass2 (Bandouin *et al.* 2004; Piry *et al.* 2004) to assign individuals to populations using the method of Rannala and Mountain (1997); this method is preferred when some source populations have not been sampled (Berry *et al.* 2004). To test for potential difficulties using



assignment tests, I first calculated  $D_{LR}$  (Paetkau *et al.* 1997), a distance method based on the average likelihood of an individual assigning to a population. Values under 5 are associated with low power (Paetkau *et al.* 1997). Because the assignment test did not distinguish well among several peripheral and source populations (see Results), I ran two different tests. The first test included only individuals in source populations (Corvallis, Eugene, and CA), testing for recent migrants from one source to another. The second test assigned individuals in peripheral populations (C6, C7, C8, C9, C10, CY, Metolius, E15, M11, M6, M16 and S1) to the primary invasion foci (Corvallis, Eugene, and California). Probability computations were calculated as recommended by Paetkau *et al.* (2004), with 10,000 simulated individuals used in Monte Carlo resampling. Individuals with probabilities less than 0.15 were assumed to originate from unsampled source populations. Individuals with less than 0.1 difference in probabilities between putative sources were assigned to both sites.

#### *Detection of admixture between sources in Oregon*

The two areas of original introduction in Oregon are genetically distinct (Rosenthal *et al.* 2008) and preliminary analyses indicated that admixture is occurring among genotypes derived from Corvallis and Eugene. Based on allele frequencies in the populations sampled from Corvallis and Eugene, I estimated the percent contribution of Corvallis and Eugene regions to individuals throughout Oregon. Percent contribution of sources was estimated using the software Hindex (Buerkle 2005), which uses maximum likelihood to estimate the percent contribution of

parental populations to individuals. Six alleles not present in source populations were excluded.

Because Hindex does not allow for more than two source populations, and alleles not present in parental populations were excluded from these analyses, I also tested for admixture using BAPS (Corander *et al.* 2008). The Bayesian clustering software BAPS includes an option for estimating admixture among individuals, and allows inclusion of more than two potential source populations. I estimated admixture proportions for individuals based on contributions from Corvallis, Eugene, and/or California. Default parameters for iteration of individuals and populations (minimum size of a population = 5, number of iterations = 50, number of reference individuals from each population = 50, number of iterations for reference individuals = 10).

#### *Evolution in outcrossing rate*

I estimated outcrossing rates via analyses of field-generated progeny in a subset of ten populations, chosen to represent both source and peripheral populations. DNA was extracted and selectively amplified with PCR as described above. Progeny were assayed for seven of the eight loci (all but locus 3-4F9), and the population-level multilocus outcrossing rate,  $t$ , was estimated using an expectation-maximization algorithm in MLTR (Ritland 2002). Individuals that were missing more than 2 loci were excluded; after this exclusion, each family consisted of an average of 6.4-7.3 progeny per maternal plant. Single locus outcrossing rates ( $t_m-t_s$ ) were also estimated; the difference between multilocus and single locus outcrossing rates is an estimate of biparental inbreeding, or mating between relatives. I also estimated the correlation of

selfing among loci, which approximates the percentage of inbreeding due to selfing, and is less dependent on the number of loci than either multilocus selfing rates or the difference between multilocus and single locus selfing rates (Ritland 2002). Standard errors were calculated using results from 100 bootstraps, resampling families in populations. All progeny were included in the analyses, as missing data had no significant effect on the parameter estimates (not shown). Maternal genotypes were inferred from the most likely parent based on genotypes of the progeny.

## **Results**

Progeny samples amplified an average of 0.81 alleles/locus more than the parental samples, except for population C10, which only differed by 0.05 alleles/locus (Fig. 4.3). By comparing genotypes among progeny and parents, I estimated a 2.7% error rate due largely to scoring errors. In other words, 156 out of 5713 loci-individual amplifications were scored incorrectly, due to either poor amplification and/or user error. All populations except S1 (Table 4.1) were deficient for heterozygotes for at least one locus; many populations were out of HW equilibrium (Table 4.1). While heterozygote deficiency can be a result of null alleles, it is also commonly observed in selfing species and in populations with extremely low polymorphism. There was no evidence for isolation by distance ( $r=0.016$ ,  $p=0.825$ ).

Percent variation among populations was greatest among peripheral populations, while Corvallis populations had the least percent variation among populations (Table 4.2). Genetic structure differed between the two Oregon invasion foci, Corvallis and Eugene; populations in the McDonald Forest (Corvallis) were less

differentiated ( $G_{ST}=0.211$ ,  $G'_{ST}=0.44$ ) than were populations near Eugene ( $G_{ST}=0.334$ ,  $G'_{ST}=0.513$ ).

Pairwise  $G'_{ST}$  values reflect the complex genetic structure observed during the early invasion process: the smallest value was 0.01, observed both between two populations in the McDonald Forest (C2 and C5) and between two populations near Foster (M6 and M16). The most highly differentiated populations had values of 1.0 between three pairs of populations (C6 and E4, C6 and S1, and E4 and S1) (Tables 4.3 and 4.4). Populations identified as being colonized by Corvallis (see assignment test results below) also had low  $G'_{ST}$  values: three of six peripheral populations colonized by Corvallis (C10, E15 and C7) had comparatively low pairwise  $G'_{ST}$  values with one or more of the Corvallis sources. Some populations identified as containing high levels of admixture (M6, M16; described below), or as being sourced from Eugene (M1), had low pairwise  $G'_{ST}$  values with population E9 in Eugene. Average pairwise  $G'_{ST}$  values were significantly correlated with effective allele number ( $r^2=0.539$ ,  $p<0.001$ , Fig. 4.4), consistent with expectations for an expanding range when dispersal is limited.

#### *Population clustering*

Samples taken from the same population at different times in the same year clustered together in the BAPS clustering analysis; they were combined in all other analyses. Three populations in the McDonald Forest in Corvallis, and a fourth population (newer in origin) nearby also clustered together (C2, C4, C5 and C10). All other populations clustered separately, making 21 clusters in all. This is probably an

overestimate of the true number of genetic clusters, as presence of family structure tends to increase the apparent number of clusters (Waples & Gaggiotti 2006), and high selfing rates (see Outcrossing Rates, below) probably reflect high levels of inbreeding in *B. sylvaticum*. The strong clustering in Corvallis reflects the comparatively low population structure in this invasion; no populations clustered together in the Eugene invasion, where differentiation is higher.

Differences among populations are represented by the three-dimensional NMDS graph made from Nei's standard genetic distance. Stress levels indicate how well the data fit the graph: a stress of 0.0 indicates a perfect fit, while stress of 0.4 indicates poor fit (Kruskal 1964). Minimum stress was achieved in this NMDS at 0.295, due to the complex relationships among populations. Groups of populations from Corvallis and Eugene were in different parts of the graph (Fig. 4.5). Populations that contained individuals with admixture between the Eugene and/or Corvallis area were often located in the center of the NMDS graph. Most non-hybrid peripheral populations clustered with the Corvallis populations.

#### *Gametic disequilibrium and gene diversity*

Many populations exhibited measurable amounts of gametic disequilibrium (GD, Table 4.1), though no two loci amplified alleles that were linked in all populations. Populations with the highest numbers of allele pairs (number of pairwise comparisons of alleles in different loci) displayed relatively little disequilibrium (e.g. C5, M6, C10, CY). Some populations with very low numbers of allele pairs had high amounts of GD (e.g. E4, E7), while other low diversity populations exhibited none

(e.g. E8, M11, C8). This may be the result of small sample size; if there are few allele pairs in a population, the ability of these tests to detect disequilibrium is also very low.

In general, older populations, such as those in Corvallis or Eugene, had higher GD than newer populations. Populations in the Foster area also exhibited low levels of disequilibrium, as did the Rogue River (CY) population and the Head of Metolius (Met) population. The presence of disequilibrium in both source and peripheral populations is consistent with ongoing dispersal from distinct sources.

#### *Power of the assignment test*

Low  $D_{LR}$  values ( $D_{LR} < 3$ ) are indicative of low power of the assignment test to distinguish between potential source populations, while values over five have high distinguishing power (Paetkau *et al.* 2004). Several population pairs within the Corvallis area and within the Eugene area had low  $D_{LR}$  values ( $D_{LR} < 4$ ). There was also low differentiation when comparing some peripheral populations to Corvallis populations (S1b, C6, C7, Metolius, and E15). To clarify patterns of dispersal in the presence of low differentiation, two separate assignment tests were completed (see below).

#### *Assignment of populations among source populations*

When individuals in the three main invasion foci were used in an assignment test, individuals from Corvallis, Eugene and California largely self-assigned (Fig. 4.6). Three individuals in Corvallis (population C4) were migrants from Eugene, but no individuals in Eugene assigned unambiguously to Corvallis. Some individuals assigned both to Corvallis and Eugene; four individuals in Corvallis and one

individual in Eugene assigned roughly equally to both sources. California individuals assigned mainly to California, with a few assigning to unsampled sources.

Unsampled sources contributed migrants to sampled populations; 8% of all individuals likely originated from one or more unsampled sources. A high percentage of individuals in the Eugene population E3 assigned to an unsampled source, however, when I reran the assignment test including Foster as a potential source, four of the five previously unassigned samples assigned to Foster. Two other individuals, one each in Eugene populations E7 and E9, assigned to Foster when Foster populations were included in the assignment test. It is unclear whether Foster populations or CY were the source for these individuals, because the assignment test cannot distinguish well between CY and M6 or M16 (see Table 4.4). Because populations in Foster and CY were likely originally established from Corvallis and Eugene, I present details only for assignments to original introduction events.

#### *Assignment of peripheral populations to source populations*

Peripheral populations originated mainly from Corvallis (Fig. 4.6). Six populations had migrants originating from Eugene (C10, C8, CY, M1, M6, M16). One Foster City population (M11) had a large number of individuals that assigned ambiguously to both Corvallis and Eugene, and there were a large number of individuals from an unsampled source. All areas that received migrants from Eugene also had individuals that assigned ambiguously to both Corvallis and Eugene (except C10), possibly due to hybridization among distinct sources.

### *Hybrid detection*

Six alleles were excluded from the hybrid index analysis because they were not detected in the parental populations. These included two alleles from C8 (frequencies of 0.08 and 0.26), one allele from Metolius (frequency of 0.02), two alleles from M6 (frequencies of 0.12 and 0.02), and one allele from S1 (frequency of 0.02). Large numbers of hybrid individuals were found in the Foster populations, to the west of Eugene and Corvallis (M6, M16, M11) (Fig. 4.6). Rogue River (CY), in Southern Oregon, also had a significant number of hybrid individuals.

### *BAPS admixture test*

Populations near Corvallis (C1, C2, C3, C4, C5) and in Eugene (E1, E3, E4, E6, E7, E8, E9) were defined as source populations in the admixture test (see Methods). The admixture test identified some individuals Corvallis as admixed between Corvallis and Eugene. Significant admixture was detected in individuals from populations throughout the range, both in source (C4) and peripheral (C10, CY, C8, M6, M11, M16) locations. Three populations near Foster City (M6, M11, M16) had high amounts of admixture. This test identified California as contributing to admixture in some populations, but this is likely due to the contributions of unsampled source populations; BAPS does not allow for unsampled sources. Individuals in Oregon that assigned to California in this test assigned to unsampled populations in the GeneClass assignment test. The GeneClass assignment test is the only test I used that allows for presence of unsampled populations.



### *Consistency among different methods of hybrid detection*

Individuals identified as hybrids in HINDEX were not always the same individuals as those identified as admixed by the BAPS admixture test, nor the same as those that assigned to two or more populations. This is likely due to the widely different methods employed by the different tests, and by the complex genetic relationships among populations in this dataset. At the population level, however, populations with high fractions of hybrid individuals identified by HINDEX or by BAPS also had high fractions of individuals assigned ambiguously to multiple populations.

Levels of disequilibrium were not always consistent with presence of admixed individuals. No admixed individuals were identified in one population in the Eugene source area (E4), despite presence of low levels of GD. In addition, two populations identified as having large numbers of admixed individuals (CY and C8) had little to no GD. In the assignment test for C8, individuals in population C8 had relatively low maximum probabilities (average maximum probability = 0.252) compared to other populations (average maximum probability for populations without hybrids = 0.840, see Table 4.5). Average maximum probabilities were above criteria for an unsampled source (0.15, see methods), but were still low, and may indicate presence of an unsampled source. A similar explanation is likely for population M11, which also had large numbers of hybrids.

### *Outcrossing rates*

These data suggest that there is variation among populations in outcrossing rates. Multilocus *t*-values (estimates of fraction of outcrossed progeny) ranged between 0.2-0.9 (Fig. 4.7). An average of 1.2 loci per individual did not amplify (0.73-1.46 loci per individual in each population), but this did not affect estimated selfing estimates (data not shown). Estimates of biparental inbreeding ranged between -0.048 and 0.098 (Fig. 4.7). These estimates are not comparable among populations with different numbers of loci and alleles (Ritland 2002).

To compare selfing rates among populations that have different numbers of loci and alleles, I estimated the correlation of selfing among loci (Ritland 2002). This estimate is reflective of the percentage of biparental inbreeding that is due to self-pollination. The correlation of selfing among loci was over 0.85 for all populations (Figs. 4.7 and 4.8); Eugene and Corvallis appear to have similar levels of selfing rates. The highest proportion of selfing was found in C7, in northern Oregon, and the lowest proportions of selfing were in C4 and C10 near Corvallis, C6, north of Foster City, and E1, in Eugene. The standard errors of these estimates often overlapped, but it appears that populations with lower selfing rates were often in central, more diverse populations, while the highest selfing rate was observed in a single peripheral population.

### *Results summary*

Summarizing results from diversity, assignment tests, and outcrossing studies (Table 4.6) allows several patterns to emerge. In general, peripheral populations were

less diverse than central populations, and varied widely in genetic composition. Populations exhibited a significant negative association between genetic diversity and differentiation. Central populations in Corvallis all had high levels of genetic diversity; however, some populations differed in selfing rate. Central populations in Eugene varied more widely than Corvallis populations in levels of diversity, but all populations I sampled in Eugene had similar outcrossing levels. All measured variables differed among populations, with genetic measures distinguishing among many populations, while outcrossing rate was different for two central and one peripheral population. The peripheral population with a lower selfing rate originated from Corvallis.

## **Discussion**

I investigated evolution of molecular genetic patterns and selfing rates among populations in the expanding range of *Brachypodium sylvaticum*. Results from this study echo the results from my previous study conducted on a small scale (Ramakrishnan *et al.* 2010). Long-distance dispersal was common throughout Oregon; it was also common in my small-scale study. Source populations have relatively low levels of differentiation and high levels of diversity, while newly-established populations exhibit the opposite pattern, with high levels of differentiation and low genetic diversity; this was also observed in the small scale study. The Corvallis region may be the initial source of range expansion, as it established the highest number of peripheral populations in Oregon. Analyses with genetic admixture models suggest that hybridization has occurred between the two genetically distinct invasions centered

in Corvallis and Eugene. Most populations had high selfing rates, but lower selfing rates were observed in two central populations and a single Corvallis-sourced peripheral population. Further insights into evolution during expansion will be explored in a later study comparing phenotypic and genetic variance distributions. Though results were similar at both scales I investigated, building an accurate picture of colonization dynamics especially in nonequilibrium situations is profoundly benefited by studying both small-scale and range-wide patterns of dispersal and population structure.

In this study, new populations had experienced bottlenecks and founder effects, similar to what I observed in my study conducted on a small scale in the Foster area (Ramakrishnan *et al.* 2010). Other aspects of dispersal dynamics, discussed below, were also quite similar at the different scales I used in these studies. The observed effect of geographic scale on observed dispersal patterns varies in the literature (Arnaud-Haond *et al.* 2006; Beck *et al.* 2008; Berry *et al.* 2004; Berthier *et al.* 2006; Bohonak 1999; Peakall *et al.* 2003; Young *et al.* 1993), due to various species-specific traits. Sampling at multiple scales can give valuable insights to the colonization dynamics of an invasive species, as small-scale analyses enable observations of colonization events comprising only one or a few colonists, which would be difficult to detect range-wide.

Long-distance or non-contiguous dispersal was common in the invaded range of *B. sylvaticum*, on both small and large scales. Long-distance dispersal is an important component of the population dynamics involved in range expansions (Clark

1998; Levin *et al.* 2003; Ramakrishnan *et al.* 2010). Some invasive species do not exhibit long-distance dispersal (e.g. Estoup *et al.* 2004), but long-distance dispersal enables a species to establish colonies ahead of the main wave of expanding populations (Ibrahim *et al.* 1996), increasing the rate of range expansion exponentially (Clark 1998; Skellam 1951). *Brachypodium sylvaticum* exhibits a commonly expected pattern of dispersal during rapid range expansion, with long-distance dispersal playing a significant role in colonization of new populations.

Source regions in the invaded range of *Brachypodium sylvaticum* have high diversity and low differentiation, consistent with high levels of dispersal, while newly-colonized regions exhibit low diversity and high differentiation, consistent with limited dispersal and metapopulation dynamics (Pannell & Dorken 2006). Extremely high levels of differentiation and low diversity observed in the invasive range of a closely-related species, *Brachypodium distachyon*, were likely caused by founder effects or bottlenecks and exacerbated by lack of outcrossing in this selfing species (Bakker *et al.* 2009). Other species that are obligately outcrossing can also exhibit population differentiation during range expansion: *Silene latifolia* is dioecious, but exhibits genetic differences among regions, while the closely-related *S. vulgaris* is self-compatible, yet exhibits no significant genetic variance at the regional level (Taylor & Keller 2007). Increased structure at the edge of a species' range is common (Arnaud-Haond *et al.* 2006; Austerlitz *et al.* 1997; Barton & Charlesworth 1984; Eckert *et al.* 2008), but has rarely been tested at the fringe of an active range expansion (but see Pannell & Dorken 2006).

All but one of the peripheral populations I sampled likely originated from Corvallis. The small-scale study had similar results, where two sites appeared to be the sources for the majority of other sites studied. The high contributions to colonization events of Corvallis compared to Eugene may be due to a higher density metapopulation structure in Corvallis, where populations exchange more migrants than in Eugene. Eugene populations are more geographically isolated than in Corvallis perhaps due to lower density logging activities. Some peripheral populations in Oregon were quite similar to Corvallis, due to either a large number of founders or high migration frequencies through several seasons. Selection could also play a role in development of population genetic structure in this region (Bakker *et al.* 2009; Charlesworth 2003). Comparing patterns of variance distribution between morphological traits and genetic markers in future studies may help clarify the complimentary roles of selection and dispersal in *B. sylvaticum* (D. Rosenthal, in prep).

Hybridization between distinct European sources may have had a strong impact on initial evolution of invasiveness in this species (Rosenthal *et al.* 2008). In the present study, I observed ongoing admixture between Corvallis and Eugene populations. Populations in the Foster region in particular had high levels of admixture; all peripheral populations that had genetic contributions from Eugene also exhibited admixture, except for one population that has since gone extinct (M1). This population was almost completely monomorphic, was established solely from Eugene, and progeny had low biomass, implying low growth rates (unpublished data).

Inadequate additive genetic variation for traits under selection has been shown to inhibit establishment (Ellstrand & Schierenbeck 2000; Kolbe *et al.* 2004; Lee 2002); on the other hand, the presence of admixture in many populations indicates that there is potential for further evolution of novel traits, as has been documented in several other invasive species (Brown & Eckert 2005; Kolbe *et al.* 2007; Lavergne & Molofsky 2007; Lindholm *et al.* 2005; Phillips *et al.* 2008).

Admixture between genetically distinct populations causes gametic disequilibrium (GD), which disappears quickly with random mating, especially in rapidly expanding populations (Nei & Li 1973; Slatkin 1994; Zapata *et al.* 2001). If there is inadequate genetic variation, GD will not be evident (Zhang *et al.* 2004). Several peripheral populations of *B. sylvaticum* in Oregon had minimal genetic diversity, making it difficult to detect GD in those populations. Other populations that displayed no GD may be expanding rapidly, erasing initial patterns of disequilibrium. As a result, absence of GD is somewhat uninformative in this study, while presence of GD indicates admixture or drift-related population dynamic processes. Several source regions with high levels of diversity displayed GD in the absence of admixture as detected by assignment tests, Bayesian admixture tests, and the hybrid index. These formal tests of admixture likely identify older admixture events, while GD identifies extremely recent admixture and/or founder events. GD could also be high in some populations if some plants have high selfing rates, creating cryptic maternal lineages in a population (Hassel *et al.* 2005); however, several high diversity source

populations with high levels of self-fertilization exhibited no GD, indicating that dispersal history likely plays a significant role in generating GD.

Most populations had high selfing rates, consistent with selection for increased selfing during colonization, though comparative selfing estimates of plants from Europe are unknown. The population in this study with the highest selfing rate was a peripheral population, which agrees with similar observations by previous researchers that selfing may be under selection during range expansion (Daehler 1998; Holsinger 1986; Ingvarsson 2002; Lande & Schemske 1985; Schemske & Lande 1985). This selection pressure may be reduced in older, high diversity populations, where outcrossing could provide an advantage. Though most populations had similar selfing rates, the lowest selfing rates were found in central populations near Corvallis, while a single Corvallis-sourced peripheral population also had a relatively low selfing rate. Interactions between selection and genetic drift will likely play a major role in further evolution of this species during range expansion.

There are many similarities between Corvallis and Eugene populations; populations in both regions have similar levels of genetic diversity, differentiation, and selfing rates. Despite the similarities, Eugene populations are more differentiated than Corvallis populations, and the Eugene area includes populations with widely varying levels of genetic diversity. Corvallis populations are genetically similar to each other, but some Corvallis populations differed in selfing rates. The Eugene populations were highly variable genetically, but more similar in selfing rates than were Corvallis populations. Divergent phenotypes among populations can be influenced by founder



effects and bottlenecks during colonization, especially if dispersal is limited, as the assignment tests in this paper and previous results (Ramakrishnan *et al.* 2010) have demonstrated.

### *Limitations of analyses*

Analyses of populations in non-equilibrium situations can be difficult. Though there are currently computer programs available that take into account factors like nonrandom mating and cryptic genetic structure, there are still areas of uncertainty when analyzing data from extreme situations such as those I observed in *B. sylvaticum*. First, assignment tests have low power when differentiation is low (Paetkau *et al.* 2004). My research suggests that assignment tests may also be confounded in areas characterized by high levels of immigration from multiple distinct sources. Second, estimation of outcrossing rates using field-generated seeds is not straightforward when using populations with highly different numbers of polymorphic loci (Ritland 2002). In addition, identification of small, recently-founded populations is difficult when geographic surveys are limited in scope. Finally, though I did not observe major differences in environment among most populations, selection may affect the genetic structure I observed in this system. Though other researchers have developed many ways of addressing non-equilibrium situations, I highlight here some areas that need further development.

For assignment tests, high similarities among some source populations and peripheral populations made it difficult to determine whether peripheral populations are contributing to the migrant pool. When all populations were included in a single

assignment test, the test was not able to distinguish between several Corvallis and peripheral populations as potential sources (population pairs with low  $D_{LR}$  values, Tables 4.3 and 4.4). It may be that some peripheral populations were established not from Corvallis itself, but from a well-established peripheral population that was itself colonized from Corvallis. Combining microsatellite data with AFLP markers, or with SNP markers developed for a closely-related species, could help clarify dispersal patterns among genetically similar populations.

Another confounding effect I observed in assignment tests is that populations receiving many migrants from multiple sources may have increased genetic diversity compared to other populations; this could bias the assignment test to misidentify high diversity sink populations as sources. Two populations that were very small, geographically isolated, and distant from historical accounts of source regions, were identified as sources for a large majority of individuals when all populations were included equally in a single assignment test (results not shown). This may be either because these populations receive migrants from many other populations, increasing diversity and making them appear to be sources (e.g., Petit *et al.* 2003), or because they were populations on the edges of previously unknown, large, unsampled source populations. Relaxing the assumption of random mating during analyses, or being able to account for differences in migration rates and levels of polymorphism among populations, may help clarify this issue in other, similar studies.

Estimating outcrossing rates using genetic data from field-collected seeds can be more problematic than when using plants organized in common garden outcrossing

arrays. Differences in numbers of polymorphic loci among populations will strongly affect outcrossing rate estimates (Ritland 2002). Estimates of outcrossing made with few loci and alleles will be less accurate than estimates using many loci and alleles, as the chance of detecting a single outcrossing event is low (Ritland 2002). High levels of disequilibrium also lower estimates of outcrossing rates (Shaw *et al.* 1981). To increase confidence in estimates of outcrossing, future studies should include a preliminary estimate of population diversity to use for an estimate of the numbers of progeny necessary for good outcrossing rate estimates. Populations with extremely low levels of diversity, such as M1, may completely inhibit estimation of outcrossing frequencies if field-generated progeny are used. Comparing results presented in this paper to results from controlled crossing experiments will help ascertain whether testing outcrossing rates of field-generated progeny is viable when low-diversity populations are included.

To identify traits affecting success during initial stages of colonization, populations that are still small and genetically depauperate should be identified. Identification of populations that are geographically isolated from the main source populations but that have not yet fully established would require detailed surveys of many unoccupied areas. This is an important consideration for researchers wanting to do similar studies; I recommend conducting initial colonization research on a small scale, where recently colonized populations that are prone to extinction can be easily identified.

Genetic drift through random sampling of source populations during colonization events causes a high degree of differentiation among colonizing populations, while selection can counter effects of drift, making populations more similar to each other (Barton & Charlesworth 1984; Whitlock 1992). This study includes newly-colonized populations in which selection has had little time to act. Differences in environments among populations could cause selection to occur during colonization events; however, the populations sampled here were mostly in quite similar environments. Genetic similarities among some populations in this study are probably due to unequal contribution of sources to the migrant pool instead of to selection. More detailed demographic studies will help clarify the relative effects of drift and selection in *B. sylvaticum* populations.

### *Conclusions*

Using a representative invasive species, I have found support for several hypotheses regarding evolution during rapid range expansion. First, dispersal dynamics on a large scale mimic the dynamics I observed on a small scale, and long-distance dispersal is a key component of colonization dynamics in this system. High frequencies of long-distance dispersal are common in other invasive species as well; results from studying the *B. sylvaticum* invasion may help predict evolution of invasiveness in similar species. Founder effects, bottlenecks and admixture occur during establishment of new populations, all of which factors provide opportunities for populations to rearrange population genetic structure in favor of evolution of novel traits. It appears that not all populations are equally invasive: the Corvallis region has

established many more peripheral populations than has the Eugene region. Further studies should be undertaken at different points in time to track development of invasive traits throughout the region, especially to determine whether Eugene begins to colonize large numbers of novel populations as the invasion progresses. Throughout both small and large scales, well-established populations that contribute to colonization of new populations have less differentiation than new populations; this may help managers decide which populations on which to focus control efforts. Detailed demographic research is required to confirm my hypothesis that low levels of differentiation are associated with high dispersal ability in different populations of invasive species. Finally, selfing rates are generally high in new populations, though low selfing rates at the source can affect selfing rates during colonization. My results have implications for management of both invasive and rare and endangered species; endangered or threatened species may simply be experiencing an invasive-type evolutionary pathway in reverse. Knowledge of the evolutionary effects of long-distance dispersal, bottlenecks and founder effects may enable managers to both inhibit growth of undesired populations while increasing fitness in populations that are at risk of becoming extinct.

Table 4.1: Population information

Information about populations sampled in my study, including elevation (Elev.), sample size (n), number of loci (No. Loci), number of polymorphic loci (poly), Nei's unbiased gene diversity (H<sub>Z</sub>), observed heterozygosity (H<sub>O</sub>), number of allele pairs (No. allele pairs), the number of allele pairs in gametic disequilibrium (No. in GD), the number of unique alleles (No. unique alleles) and the fraction of unique alleles (fraction).

Code	Name	Latitude	Longitude	Ecoregion	Elev.	n	No. Loci (poly)	% missing data	H <sub>Z</sub>	H <sub>O</sub>	No. allele pairs (No. in GD)	No. unique alleles (fraction)
C1*	McDonald Forest	44.65983	-123.239	Valley Foothills	160m	24	8 (6)	0.03	0.265	0.033†	106 (25)	0
C2*	McDonald Forest	44.63124	-123.305	Valley Foothills	365m	31	8 (6)	0.16	0.361	0.125†	124 (0)	1 (0.083)
C3*	McDonald Forest	44.71016	-123.316	Valley Foothills	~150m	32	8 (6)	0.04	0.317	0.034†	138 (27)	1 (0.016)
C4*	McDonald Forest	44.6858	-123.295	Valley Foothills	374m	25	8 (5)	0.04	0.261	0.081†	111 (2)	0
C5*	McDonald Forest	44.64313	-123.336	Valley Foothills	272m	30	8 (5)	0.04	0.331	0.109†	168 (2)	0
E1*	Fall Creek Reservoir	43.95978	-122.736	Valley Foothills	211m	44	8 (6)	0.04	0.307	0.089†	129 (18)	1 (0.029)
E3*	Dolly Varden Campground	43.9635	-122.618	Western Cascade Lowlands and Valleys	275m	11	8 (6)	0.16	0.348	0.182†	84 (0)	0
E4*	Panorama Rd	44.01325	-122.875	Valley Foothills	386m	30	8 (2)	0.13	0.079	0.022†	6 (4)	0
E6*	Hill Creek Rd	43.99635	-122.799	Valley Foothills	326m	24	8 (6)	0.00	0.354	0.078†	131 (12)	1 (0.438)
E7*	Big Fall Creek	43.9748	-122.646	Western Cascade Lowlands and Valleys	375m	22	8 (4)	0.03	0.113	0.053†	37 (13)	0
E8*	Pengra Rd	43.95701	-122.843	Valley Foothills	189m	11	8 (5)	0.02	0.173	0.200	24 (0)	0
E9*	Rock Quarry	43.9738	-122.873	Valley Foothills	175m	22	8 (7)	0.07	0.316	0.169†	74 (4)	1 (0.068)
S3*	San Francisco	37.38706	-122.262			24	8 (1)	0.01	0.053	0.026†	0 (0)	1 (1.00)
M1	Fish Ck Campground	44.39827	-122.345	Western Cascade Lowlands and Valleys	375m	15	7 (1)	0.13	0.066	0.038	0 (0)	0
M11	Whiskey Butte #3	44.38125	-122.599	Western Cascade Lowlands and Valleys	448m	25	8 (3)	0.03	0.069	0.030†	12 (0)	0

Code	Name	Latitude	Longitude	Ecoregion	Elev.	n	No. Loci (poly)	% missing data	H <sub>Z</sub>	H <sub>O</sub>	No. allele pairs (No. in GD)	No. unique alleles (fraction)
M16	Cedar and Wiley Junction	44.37187	-122.621	Western Cascade Lowlands and Valleys	228m	25	8 (5)	0.04	0.296	0.193†	53 (4)	0
M6	Wiley Park	44.41396	-122.675	Valley Foothills	165m	25	8 (6)	0.12	0.360	0.174†	211 (4)	1 (0.020)
C10	Bellfountain Rd	44.39131	-123.365	Valley Foothills	176m	12	8 (6)	0.14	0.349	0.225	145 (0)	0
S1	Fisherman's Bend	44.75384	-122.518	Valley Foothills	230m	31	8 (2)	0.00	0.012	0.012†	4 (0)	1 (0.016)
C6	Hwy 22	44.75532	-122.388	Western Cascade Lowlands and Valleys	295m	24	8 (3)	0.00	0.122	0.078†	26 (0)	0
C7	Blodgett Logging Tract	46.05544	-123.292	Volcanics	521m	24	8 (5)	0.03	0.232	0.131†	77 (0)	0
C8	Trappist Monastery	45.28296	-123.097	Valley Foothills	~90m	25	8 (4)	0.00	0.174	0.120†	36 (0)	2 (0.170)
Met	Head of Metolius River	44.43516	-121.640	Cascade Crest Montane Forest	908m	47	8 (5)	0.04	0.087	0.065†	34 (4)	1 (0.022)
CY	Rogue River	42.67544	-123.952	Coastal Siskiyou	71m	24	8 (6)	0.01	0.238	0.141†	146 (2)	0
E15	Cape Perpetua	44.28711	-124.108	Volcanics	248m	22	7 (6)	0.16	0.354	0.148†	129 (0)	0

\*Implicated as initial sites of introduction from native source(s), †Heterozygote deficient at least one locus  $p < 0.05$

Table 4.2: Analysis of molecular variance (AMOVA)  
Results for AMOVAs for different groups of populations

Source of Variance	Corvallis	Eugene	Secondary Populations
Among populations	27.7%	40.8%	49.2%
Within populations	72.4%	59.2%	50.8%
Number of populations	5	7	8

Table 4.3: Population pairwise differentiation of source populations  
Differentiation as measured by  $G'_{ST}$  (lower diagonal) and power of assignment tests as indicated by  $D_{LR}$  (upper diagonal). † is under 5

	C1	C2	C3	C4	C5	E1	E3	E4	E6	E7	E8	E9	S3
C1	0	8.05	7.19	7.42	5.86	14.5	9.85	16.7	9.08	7.91	11.8	6.35	13.3
C2	0.46	0	7.2	3.5†	1.83†	13.4	7.72	13.4	16.1	10	11.6	9.18	12.5
C3	0.45	0.48	0	8.35	5.29	14	11.8	13.8	12.1	13.5	6.15	9.44	12.3
C4	0.54	0.12	0.53	0	0.59†	9.43	8.87	12.3	13.3	9.79	8.94	7.39	7.26
C5	0.45	0.01	0.5	0.13	0	9.99	7.01	11.9	11.7	8.64	7.24	6.24	8.85
E1	0.56	0.45	0.53	0.43	0.42	0	4.87†	5.42	9.44	8.54	7.23	4.21†	14.5
E3	0.38	0.41	0.54	0.5	0.39	0.4	0	2.25†	8.46	5.7	7.55	4.11†	13.9
E4	0.66	0.61	0.61	0.55	0.58	0.31	0.4	0	10.9	10.3	7.32	5.86	16.3
E6	0.41	0.53	0.45	0.54	0.51	0.49	0.3	0.66	0	12.4	7.56	6.36	13.7
E7	0.39	0.48	0.53	0.5	0.49	0.57	0.5	0.96	0.57	0	10.7	5.81	13.7
E8	0.35	0.44	0.37	0.43	0.41	0.42	0.58	0.55	0.45	0.65	0	5.43	11.1
E9	0.3	0.3	0.38	0.3	0.3	0.25	0.28	0.35	0.32	0.29	0.27	0	11
S3	0.6	0.49	0.57	0.34	0.47	0.55	0.7	0.78	0.64	0.74	0.51	0.34	0
M1	0.39	0.4	0.5	0.35	0.38	0.43	0.5	0.75	0.52	0.34	0.33	0.12	0.54
M11	0.61	0.42	0.47	0.4	0.43	0.48	0.56	0.75	0.55	0.6	0.46	0.36	0.53
M16	0.45	0.33	0.28	0.36	0.34	0.33	0.34	0.53	0.31	0.57	0.3	0.18	0.49
M6	0.41	0.33	0.31	0.39	0.35	0.34	0.29	0.53	0.34	0.53	0.34	0.11	0.5
C10	0.32	0.05	0.22	0.23	0.08	0.44	0.35	0.59	0.43	0.47	0.36	0.29	0.43
S1	0.61	0.46	0.46	0.48	0.48	0.56	0.7	1	0.65	0.75	0.56	0.35	0.71
C6	0.38	0.31	0.59	0.37	0.34	0.69	0.54	1	0.65	0.75	0.67	0.46	0.99
C7	0.38	0.08	0.53	0.18	0.12	0.5	0.4	0.63	0.52	0.49	0.49	0.37	0.52
C8	0.5	0.32	0.46	0.29	0.29	0.55	0.67	0.75	0.54	0.74	0.51	0.34	0.54
Met	0.54	0.36	0.65	0.29	0.33	0.57	0.57	0.61	0.65	0.61	0.57	0.41	0.55
CY	0.46	0.25	0.22	0.32	0.29	0.49	0.58	0.56	0.47	0.49	0.26	0.32	0.4
E15	0.36	0.13	0.48	0.24	0.11	0.49	0.35	0.55	0.47	0.44	0.38	0.37	0.43

† Value is under 5, indicating low assignment test power



Table 4.4: Population pairwise differentiation of peripheral populations  
 Differentiation as measured by  $G'_{ST}$  (lower diagonal) and power of assignment tests as indicated by  $D_{LR}$  (upper diagonal). ‡ is under 5

	M1	M11	M16	M6	C10	S1	C6	C7	C8	Met	CY	E15
C1	10.5	14.3	7.41	8.22	7.16	14	5.59	4.16 <sup>‡</sup>	12.5	8.94	10.1	6.1
C2	14.6	10.4	7.89	11.7	6.5	10.6	6.73	2.96	10.9	5.44	8.74	7.12
C3	13.7	9.88	5.23	9.02	4.68 <sup>‡</sup>	7.75	8.95	7.5	9.98	9.57	4.94 <sup>‡</sup>	9.1
C4	10.1	10	8.41	8.31	4.7 <sup>‡</sup>	9.33	4.39 <sup>‡</sup>	2.6 <sup>‡</sup>	7.01	3.41 <sup>‡</sup>	8.65	4.86 <sup>‡</sup>
C5	8.73	10.5	6.62	6.98	3.03 <sup>‡</sup>	8.93	3.83 <sup>‡</sup>	1.29 <sup>‡</sup>	6.61	3.36 <sup>‡</sup>	5.78	3.65
E1	9.75	13.7	7.4	8.38	11.7	13.9	15.9	12.6	13.1	14.6	9.68	16.1
E3	10.5	10.6	5.92	6.49	9.9	14.3	12	8.84	12.4	9.81	11	11
E4	13.2	12.7	6.3	10.2	13.7	16.4	17.1	15	15	15.1	13.7	18.5
E6	12.1	14.3	5.74	8.25	12	14.3	14	11.7	10.9	15.2	9.53	14.9
E7	7.15	12.8	9.8	10.6	11.6	14.2	13.2	10.3	12.3	13.6	10.2	13.4
E8	9.92	10.1	3.65 <sup>‡</sup>	6.4	7.32	10.8	11.5	11.1	8.75	10.5	5.2	13.4
E9	6.53	9.06	4.62 <sup>‡</sup>	7.05	8.36	9.76	12.2	8.69	8.91	10.9	9.67	10.4
S3	11.6	11.1	11.4	12.5	12.3	11.5	13.9	11.7	10.6	12.8	11.1	12.6
M1	0	11.1	9.23	9.3	12.8	11.5	14	12.5	10.6	11.2	11.9	14.4
M11	0.43	0	6.6	8.58	11.1	6.65	14.2	11.7	10	11.8	6.57	15.7
M16	0.34	0.41	0	4.15 <sup>‡</sup>	5.94	8.5	9.44	7.94	9.91	10.3	3.01 <sup>‡</sup>	10.9
M6	0.28	0.34	0.01	0	2.83 <sup>‡</sup>	11.3	9.12	8.33	11.8	9.32	4.75 <sup>‡</sup>	11
C10	0.32	0.42	0.23	0.18	0	8.47	14.8	12.8	8.14	11.5	7.3	14.3
S1	0.65	0.44	0.43	0.48	0.36	0	5.78	4.55 <sup>‡</sup>	9.3	7.07	4.44 <sup>‡</sup>	7.42
C6	0.99	0.75	0.61	0.58	0.28	1	0	1.89 <sup>‡</sup>	10.3	6.11	9.24	7.14
C7	0.43	0.56	0.42	0.37	0.13	0.56	0.16	0	7.18	4.38 <sup>‡</sup>	7.1	4.9 <sup>‡</sup>
C8	0.44	0.41	0.47	0.5	0.34	0.48	0.59	0.35	0	7.85	8.79	13.1
Met	0.48	0.44	0.47	0.52	0.38	0.53	0.35	0.28	0.35	0	9.79	8.11
CY	0.29	0.25	0.11	0.16	0.16	0.29	0.48	0.3	0.38	0.45	0	9.21
E15	0.33	0.47	0.32	0.32	0.1	0.44	0.2	0.16	0.34	0.38	0.31	0

<sup>‡</sup>Value is under 5, indicating low assignment test power

Table 4.5: Assignment test probabilities  
Average maximum probability for individuals in peripheral populations assigned to  
Corvallis, Eugene or San Francisco.

Population	Maximum probability
C10	0.733
C6	0.885
C7	0.845
C8	0.253
C9	0.904
CY	0.572
E15	0.752
M1	0.931
M11	0.138
M16	0.584
M6	0.482
S1b	0.754
S2	0.916

Table 4.6: Population history

Diversity was measured as effective number of alleles ( $N_E$ ). Bold populations were central populations. Underlined populations were included in outcrossing analysis. Populations that had gametic disequilibrium are noted as having admixture; populations that had high fractions of hybrids are also noted.

	Low Differentiation ( $G'_{ST} < 0.4$ )	Medium Differentiation ( $0.4 < G'_{ST} < 0.5$ )	High Differentiation ( $G'_{ST} > 0.5$ )
Low Diversity ( $N_E < 1.2$ )		Recently established from a single source ( <u>M1</u> ) or hybrids (M11)	Recently established sites colonized from a single (S1, S3, <u>C6</u> <sup>†</sup> ) or multiple ( <u>Met</u> , <u>E4</u> <sup>‡</sup> , <u>E7</u> ) sources.
Intermediate Diversity ( $N_E = 1.2-1.8$ )	Established sites ( <u>C7</u> <sup>†</sup> , <u>C4</u> <sup>†</sup> , E15, <u>E9</u> <sup>**</sup> ) with hybrids (CY, M16 <sup>*</sup> )	Established sites ( <u>E8</u> , <u>E3</u> <sup>‡</sup> , Met) with hybrids (C8) and admixture ( <u>E1</u> , <u>C1</u> , <u>C3</u> )	
High Diversity ( $N_E > 1.8$ )	Established sites ( <u>C2</u> , <u>C5</u> , <u>C10</u> <sup>†</sup> ) with hybrids (M6 <sup>*</sup> )	Established site ( <u>E6</u> ) with admixture	

\*\* local source for Eugene populations, \* identified as local sources in Ramakrishnan 2010, † divergent selfing rate, ‡ central location but small and geographically isolated

Fig. 4.1: Possible patterns of correlation between differentiation and genetic diversity  
In a single region or set of populations, as genetic diversity increases after an initial  
colonization event, average pairwise population differentiation could have a negative  
(a), constant (b and c) or positive correlation with average pairwise population genetic  
differentiation (See text). Dotted lines represent a pattern not investigated in this  
paper.

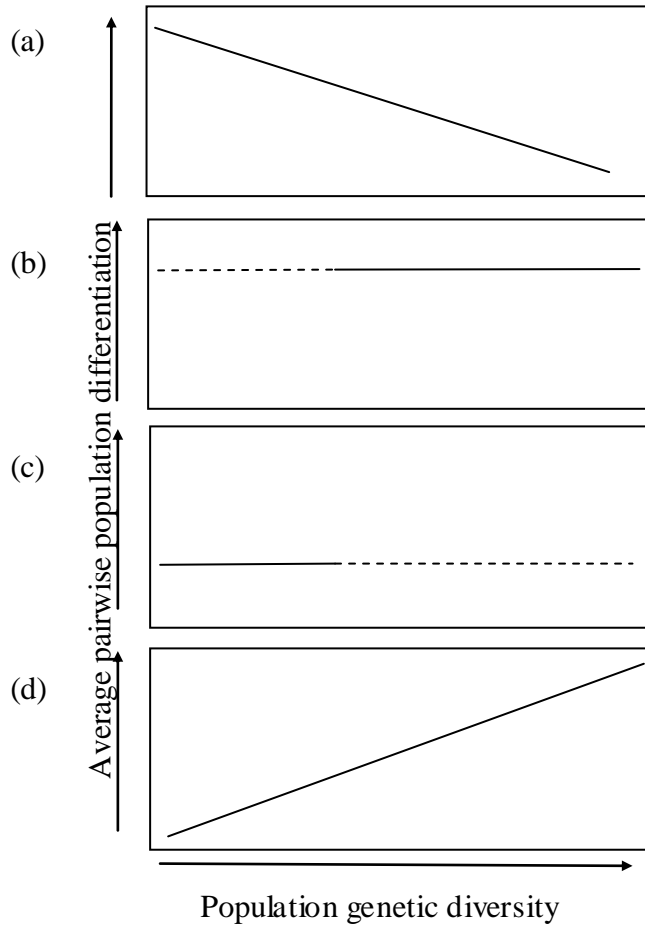


Fig. 4.2: Known distribution of *B. sylvaticum* in Oregon as of 2006  
 Inset maps include sampled populations (black triangles) and locations where *B. sylvaticum* was observed (x). Shading indicates areas of greater population density; grey lines are lakes and streams.

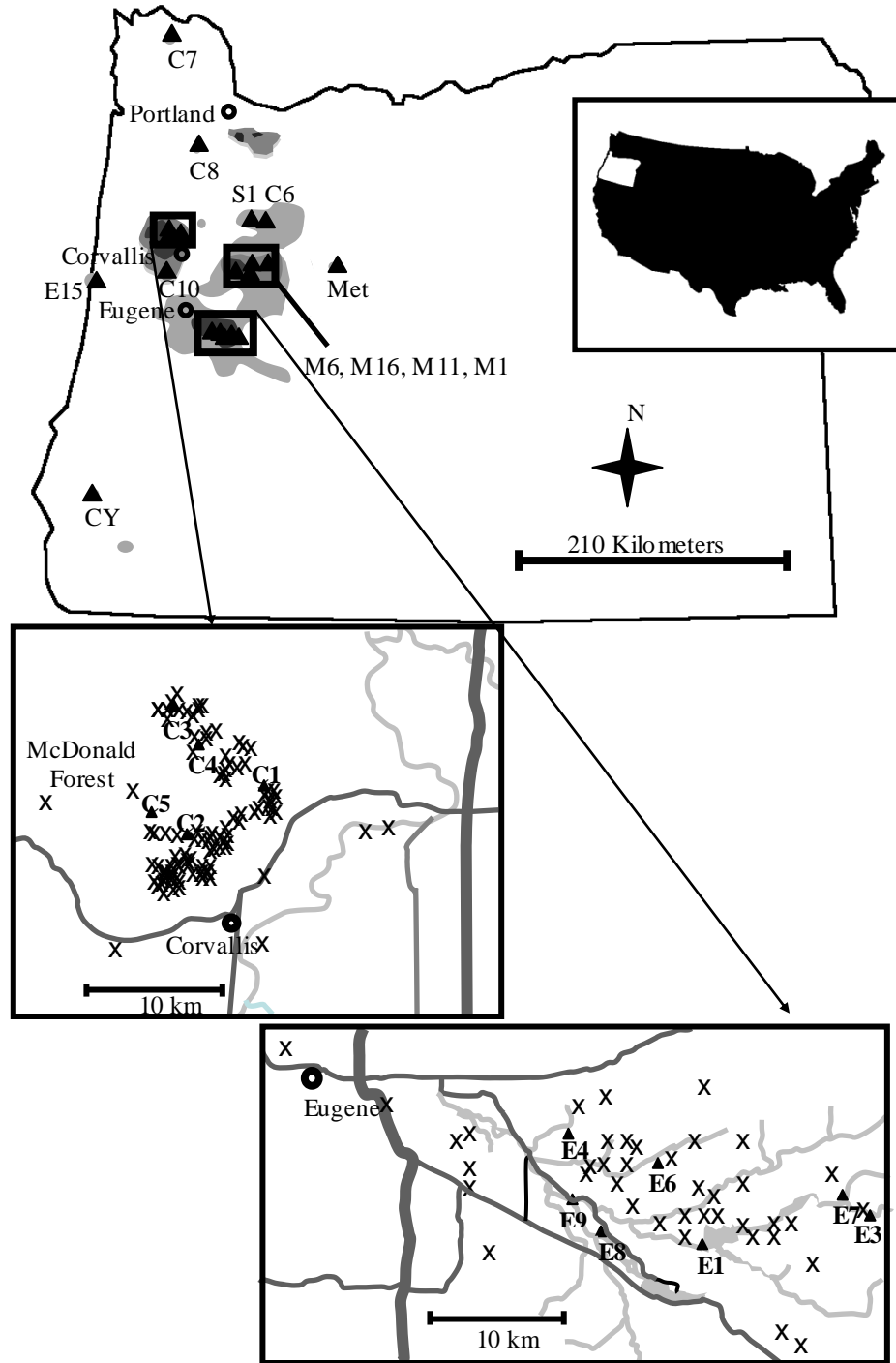


Fig. 4.3: Comparison of diversity between parental and progeny samples  
Average number of alleles per locus is compared between original population samples and progeny genetic analyses. The black line is the 1:1 ratio.

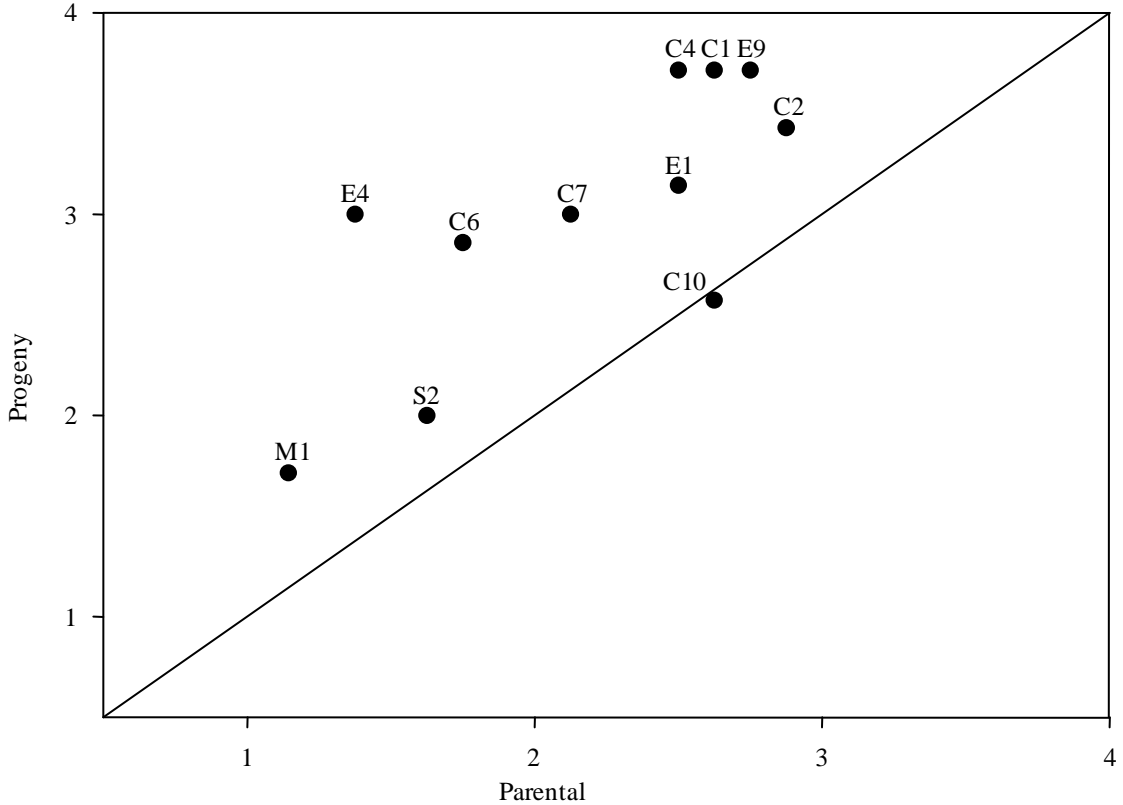


Fig. 4.4: Relationship between diversity (effective allele number) and average pairwise population differentiation ( $G'_{ST}$ )  
 Populations with low diversity have higher differentiation, while populations with high diversity have lower differentiation.

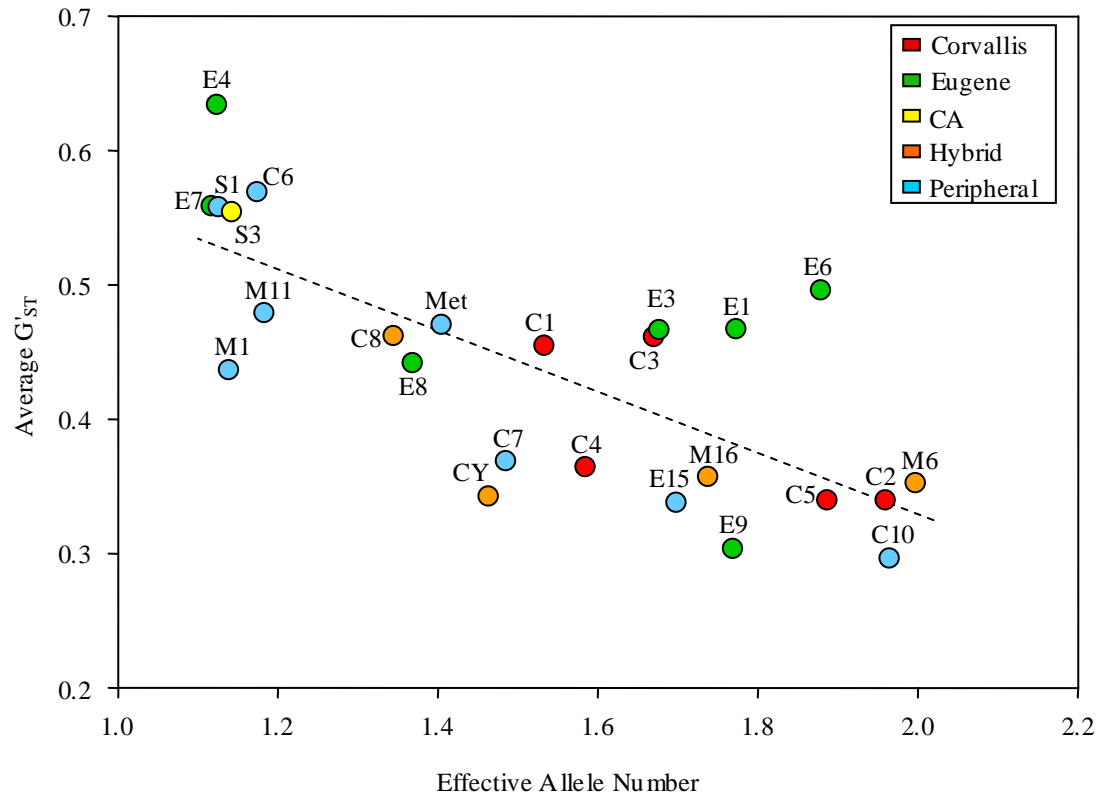


Fig. 4.5: Population genetic distance  
NMDS of Nei's standard distance plotted onto three axes.

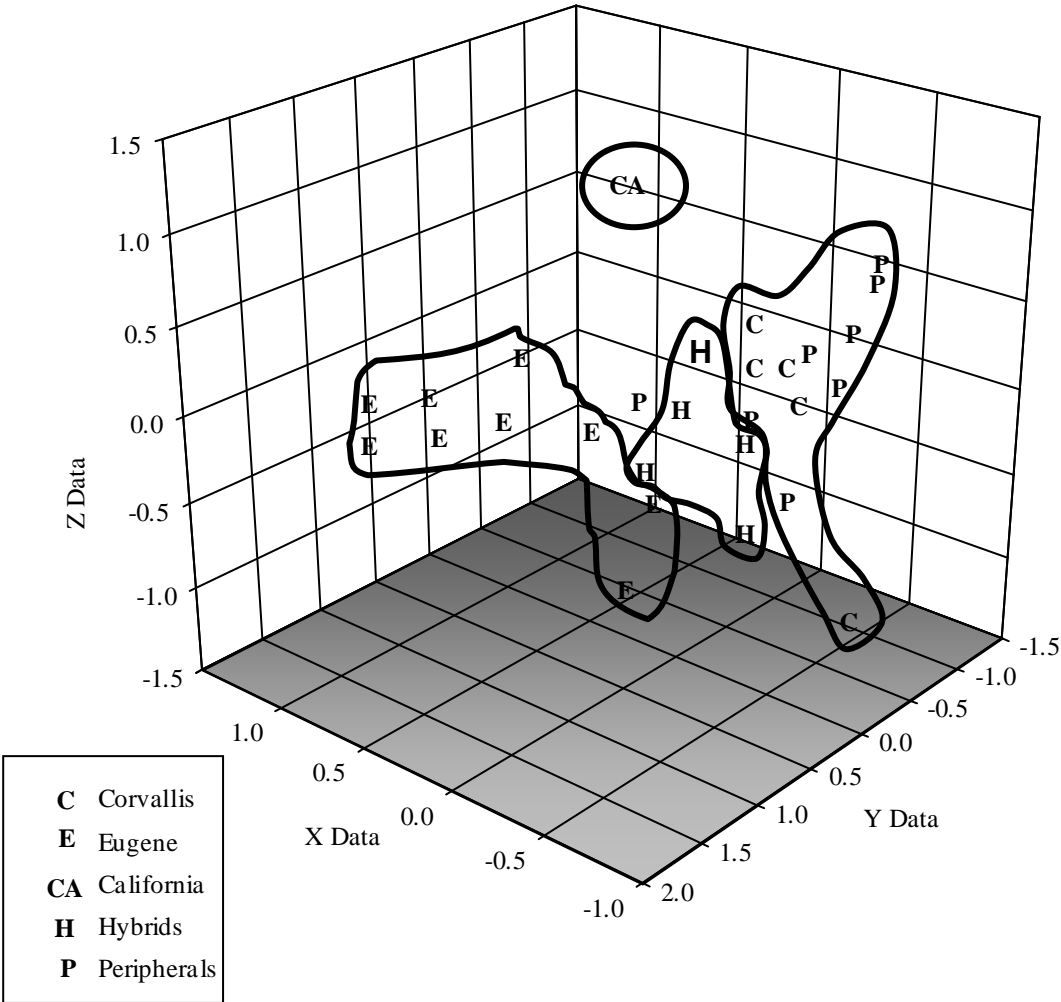




Fig. 4.6: Population clustering and admixture

Results from assignment tests (A), Hybrid index (B), BAPS admixture test (C), and gametic disequilibrium (D). C&E indicates individuals with contributions from both Corvallis and Eugene, while C&E&CA indicates individuals with Corvallis, Eugene and San Francisco contributions. GD is gametic disequilibrium.

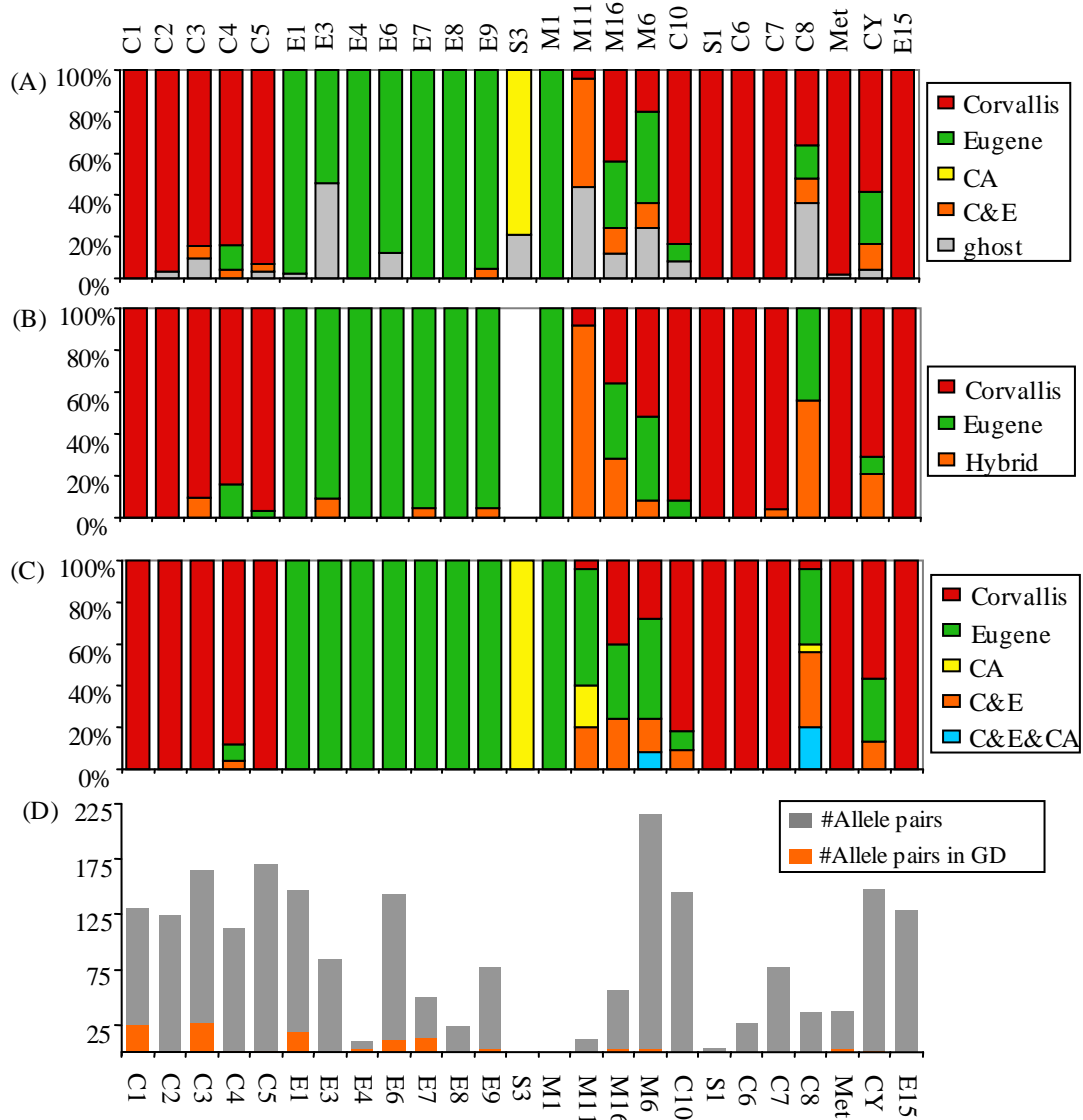


Fig. 4.7: Selfing rates estimated using the program MLTR  
 Self-pollination rates as estimated by correlation of selfing among loci (A) in MLTR. The multilocus outcrossing estimate (B) and difference between multilocus and single locus selfing estimates (C) are included, also estimated in MLTR. Non-overlapping standard errors as estimated by bootstrapping are indicated.

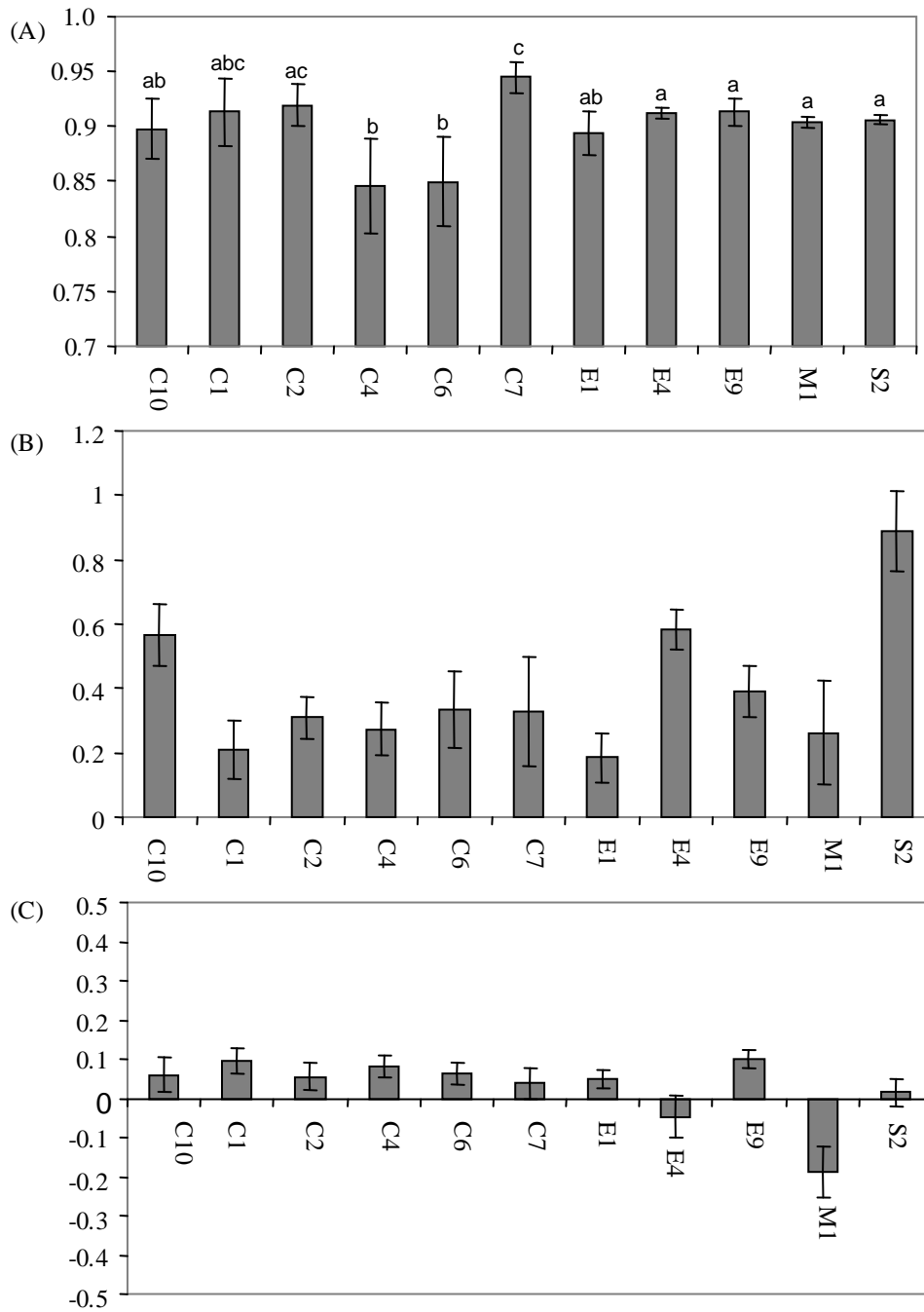
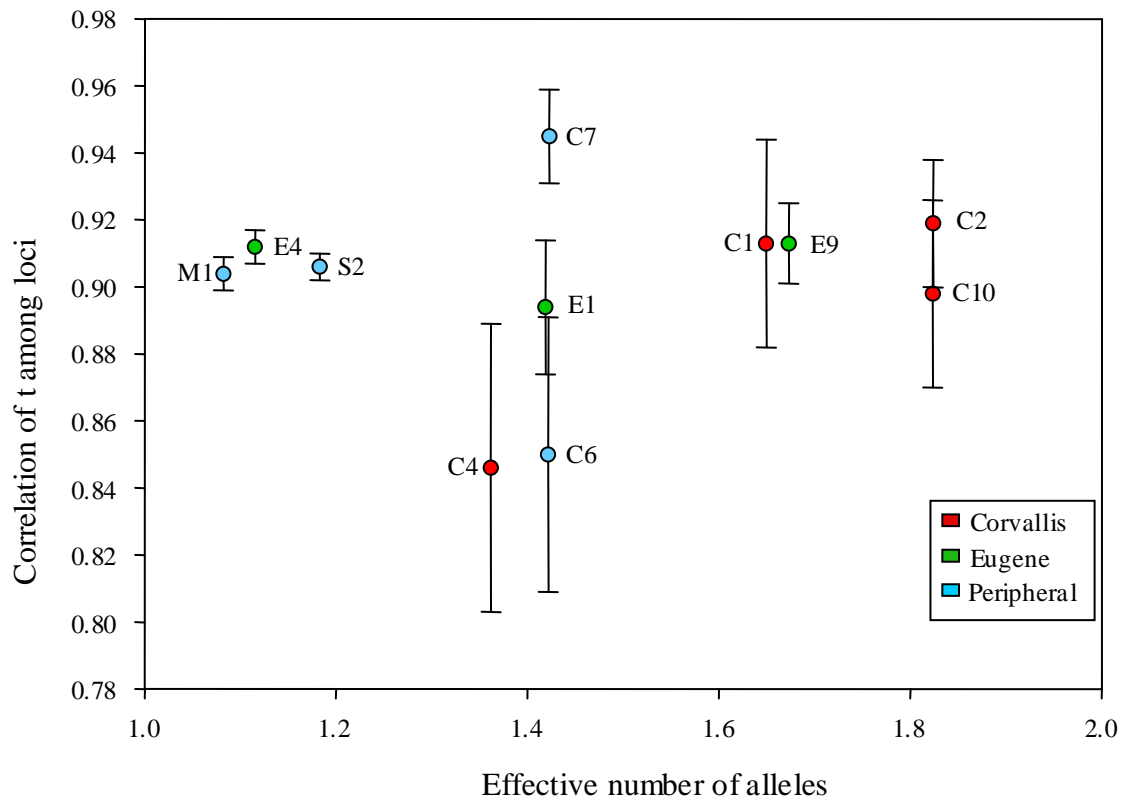


Fig. 4.8: Effective number of alleles and selfing frequency  
Correlation of selfing among loci compared to effective number of alleles (an estimator of population age). Effective number of alleles calculated using progeny samples. Standard errors estimated from bootstraps are indicated.



## **Chapter 5: Concluding remarks**

My study was one of the first to detail genetic patterns of colonization during a nascent range expansion. By sampling intensively on a small scale at the fringe of the invasion and comparing those results to patterns observed throughout the invaded range, I was able to address all my hypotheses, but the results I gathered generated several more questions that should be addressed in later research.

1. The lack of dispersal observed from newly-colonized populations in the small-scale study points toward the presence of a lag phase during colonization, where each newly-colonized population undergoes a period of lower fitness. The lag phase could be caused by low levels of diversity (e.g. Ahlroth *et al.* 2003; Crawford & Whitney 2010) or density (Davis *et al.* 2004; Keitt *et al.* 2001; Stephens & Sutherland 1999), with increasing migration and concomitant increasing diversity and finally sending the population to a new fitness peak. An experimental common garden design modifying both diversity and density should be prepared to investigate relative effects of these two factors in colonization success. Experimental arrays would also help clarify relationships between population age, genetic diversity and fitness.

2. Current computational genetic methods used to identify source populations of individuals assume some level of population equilibrium, be it random mating or thorough sampling of all source populations. To investigate the effects of perturbations of these assumptions on accuracy of source identification, simulated

populations with varying degrees of self-fertilization, immigration and emigration, and missing source populations should be generated. Many programs are currently available, and detailed investigation into the power of these tests in different situations should be undertaken.

3. My small-scale study was informative, but because the Foster region contained high numbers of hybrids and contributions from both Corvallis and Eugene, similar studies should be conducted in similar habitats with different source populations. Ideally, several study areas with populations originating from either Corvallis or Eugene, spanning different densities at the fringe of the range, and several study areas in central regions should be identified. Because of the low population differentiation I observed in the source regions Corvallis and Eugene, it would probably be necessary to increase the numbers of molecular markers in order to increase power of the assignment tests. Either AFLP markers, or SNP markers recently developed for *Brachypodium distachyon* (Garvin *et al.* 2008), could potentially help tease apart relationships among sample areas in regions of low differentiation.

4. I sampled groups of individuals in populations and used methods of moments to make inferences about population structure; however, several analysis methods make use of individual-level data to infer details about within- and among-population dispersal patterns (Peakall *et al.* 2003). I strongly recommend mapping sampled

individuals to enable more detailed investigations into within-population genetic structure.

5. Outcrossing estimates made from field-generated seeds can be unreliable, especially when populations have low numbers of polymorphic loci (Ritland & Jain 1981).

Plants sampled from populations throughout the range should be subjected to common garden outcrossing tests to determine the level of differences in outcrossing among populations. In addition, within-population variation may occur in selfing rates (Cruzan & Arnold 1994): a larger number of progeny would enable estimation of individual-level selfing rates.

In conclusion, I was able to document genetic patterns generated by initial colonization events during a nascent range expansion. Because my research was a pioneering study using a plant for which very little genetic information was available, much more research remains to be done. As the evolutionary factors leading to the success of this recent invader are further investigated, patterns will emerge that can help researchers predict what processes and/or traits increase the probability of evolution of aggressively invasive traits in non-native species. In addition, further investigation into factors necessary for colonization success could help researchers to more accurately model shifts in species' ranges during climate change or habitat disturbance.

## Literature Cited

- Aars J, Ims RA (2000) Population dynamic and genetic consequences of spatial density-dependent dispersal in patchy populations. *The American Naturalist* **155**, 252-265.
- Ahloth P, Alatalo RV, Holopainen A, Kumpalainen T, Suhonen J (2003) Founder population size and number of source populations enhance colonization success in waterstriders. *Oecologia* **137**, 617-620.
- Aketarawong N, Bonizzoni M, Thanaphum S, *et al.* (2007) Inferences on the population structure and colonization process of the invasive oriental fruit fly, *Bactrocera dorsalis* (Mendel). *Molecular Ecology* **16**, 3522-3532.
- Andreassen HP, Stenseth NC, Ims RA (2002) Dispersal behavior and population dynamics of vertebrates. In: *Dispersal Ecology* (eds. Bullock JM, Kenward RE, Hails RS), pp. 237-256. Blackwell Science Ltd, Reading, U.K.
- Arnaud-Haond S, Teixeira S, Massa SI, *et al.* (2006) Genetic structure at range edge: low diversity and high inbreeding in Southeast Asian mangrove (*Avicennia marina*) populations. *Molecular Ecology* **15**, 3515-3525.
- Astaneh I, Gosling E, Wilson J, Powell E (2005) Genetic variability and phylogeography of the invasive zebra mussel, *Dreissena polymorpha* (Pallas). *Molecular Ecology* **14**, 1655-1666.
- Austerlitz F, Dick CW, Dutech C, *et al.* (2004) Using genetic markers to estimate the pollen dispersal curve. *Molecular Ecology* **13**, 937-954.
- Austerlitz F, Garnier-Gere PH (2003) Modelling the impact of colonisation on genetic diversity and differentiation of forest trees: interaction of life cycle, pollen flow and seed long-distance dispersal. *Heredity* **90**, 282-290.
- Austerlitz F, JungMuller B, Godelle B, Gouyon PH (1997) Evolution of coalescence times, genetic diversity and structure during colonization. *Theoretical Population Biology* **51**, 148-164.
- Austerlitz F, Mariette S, Machon N, Gouyon PH, Godelle B (2000) Effects of colonization processes on genetic diversity: Differences between annual plants and tree species. *Genetics* **154**, 1309-1321.
- Baker HG (1955) Self-compatibility and establishment after "long-distance" dispersal. *Evolution* **9**, 347-349.
- Bakker EG, Montgomery B, Nguyen T, *et al.* (2009) Strong population structure characterizes weediness gene evolution in the invasive grass species *Brachypodium distachyon*. *Molecular Ecology* **18**, 2588-2601.
- Balloux F, Lugon-Moulin N (2002) The estimation of population differentiation with microsatellite markers. *Molecular Ecology* **11**, 155-165.
- Bandouin L, Piry S, Cornuet JM (2004) Analytical Bayesian approach for assigning individuals to populations. *Journal of Heredity* **95**, 217-224.
- Barton NH, Charlesworth B (1984) Genetic revolutions, founder effects, and speciation. *Annual Review of Ecology and Systematics* **15**, 133-164.

- Barton NH, Hewitt GM (1985) Analysis of hybrid zones. *Annual Review of Ecology and Systematics* **16**, 113-148.
- Beck JB, Schmuths H, Schaal BA (2008) Native range genetic variation in *Arabidopsis thaliana* is strongly geographically structured and reflects Pleistocene glacial dynamics. *Molecular Ecology* **17**, 902-915.
- Berli P (2004) Effect of unsampled populations on the estimation of population sizes and migration rates between sampled populations. *Molecular Ecology* **13**, 827-836.
- Berry O, Tocher MD, Sarre SD (2004) Can assignment tests measure dispersal? *Molecular Ecology* **13**, 551-561.
- Berthier K, Charbonnel N, Galan M, Chaval Y, Cosson JF (2006) Migration and recovery of the genetic diversity during the increasing density phase in cyclic vole populations. *Molecular Ecology* **15**, 2665-2676.
- Bialozyt R, Ziegenhagen B, Petit RJ (2006) Contrasting effects of long distance seed dispersal on genetic diversity during range expansion. *Journal of Evolutionary Biology* **19**, 12-20.
- Bohonak AJ (1999) Dispersal, gene flow, and population structure. *The Quarterly Review of Biology* **74**, 21-45.
- Bohonak AJ, Jenkins DG (2003) Ecological and evolutionary significance of dispersal by freshwater invertebrates. *Ecology Letters* **6**, 783-796.
- Bohonak AJ, Roderick GK (2001) Dispersal of invertebrates among temporary ponds: are genetic estimates accurate? *Israel Journal of Zoology* **47**, 367-386.
- Bossdorf O, Auge H, Lafuma L, *et al.* (2005) Phenotypic and genetic differentiation between native and introduced plant populations. *Oecologia* **144**, 1-11.
- Brown JH, Kodric-Brown A (1977) Turnover rates in insular biogeography: effect of immigration on extinction. *Ecology* **58**, 445-449.
- Brown JS, Eckert CG (2005) Evolutionary increase in sexual and clonal reproductive capacity during biological invasion in an aquatic plant *Butomus umbellatus* (Butomaceae). *American Journal of Botany* **92**, 495-502.
- Brzustowski J (2002) Doh assignment test calculator.
- Buerkle CA (2005) Maximum-likelihood estimation of a hybrid index based on molecular markers. *Molecular Ecology Notes* **5**, 684-687.
- Burland TM, Barratt EM, Nichols RA, Racey PA (2001) Mating patterns, relatedness and the basis of natal philopatry in the brown long-eared bat, *Plecotus auritus*. *Molecular Ecology* **10**, 1309-1321.
- Cain ML, Milligan BG, Strand AE (2000) Long-distance seed dispersal in plant populations. *American Journal of Botany* **87**, 1217-1227.
- Cavalli-Sforza LL, Edwards AWF (1967) Phylogenetic analysis: models and estimation procedures. *American Journal of Human Genetics* **19**, 233-257.
- Chapuis M-P, Estoup A (2007) Microsatellite Null Alleles and Estimation of Population Differentiation. *Molecular Biology and Evolution* **24**, 621-631.
- Charlesworth D (2003) Effects of inbreeding on the genetic diversity of populations. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* **358**, 1051-1070.



- Clark AL, Low BS (2001) Testing evolutionary hypotheses with demographic data. *Population and Development Review* **27**, 633-660.
- Clark JS (1998) Why trees migrate so fast: Confronting theory with dispersal biology and the paleorecord. *American Naturalist* **152**, 204-224.
- Clark JS, Fastie C, Hurtt G, *et al.* (1998) Reid's paradox of rapid plant migration: Dispersal theory and interpretation of paleoecological records. *Bioscience* **48**, 13-24.
- Clark JS, Lewis M, Horvath L (2001) Invasion by extremes: Population spread with variation in dispersal and reproduction. *American Naturalist* **157**, 537-554.
- Clayton WD, Harman KT, Williamson H (2002 onwards) World Grass Species: Descriptions, Identification, and Information Retrieval.
- Clobert J, Ims RA, Rousset F (2004) Causes, mechanisms and consequences of dispersal. In: *Metapopulation Ecology, Genetics and Evolution* (eds. Hanski I, Gaggiotti OE), pp. 305-333. Academic Press, San Diego.
- Comes H, Abbott R (1998) The relative importance of historical events and gene flow on the population structure of a Mediterranean ragwort, *Senecio gallicus* (Asteraceae). *Evolution* **52**, 355-367.
- Corander J, Marttinen P (2006) Bayesian identification of admixture events using multilocus molecular markers. *Molecular Ecology* **15**, 2833-2843.
- Corander J, Marttinen P, Sirén J, Tang J (2008) Enhanced Bayesian modelling in BAPS software for learning genetic structures of populations. *BMC Bioinformatics* **9**, 539-553.
- Cornuet JM, Piry S, Luikart G, Estoup A, Solignac M (1999) New methods employing multilocus genotypes to select or exclude populations as origins of individuals. *Genetics* **153**, 1989-2000.
- Crawford KM, Whitney KD (2010) Population genetic diversity influences colonization success. *Molecular Ecology* **19**, 1253-1263.
- Crow JF (1999) Hardy, Weinberg and language impediments. *Genetics* **152**, 821-825.
- Cruzan MB (2005) Patterns of introgression across an expanding hybrid zone: analysing historical patterns of gene flow using nonequilibrium approaches. *New Phytologist* **167**, 267-278.
- Cruzan MB, Arnold ML (1994) Assortative mating and natural selection in hybridizing irises: effects of phenology and floral densities on family outcrossing rates. *Heredity* **72**, 95-105.
- Daehler CC (1998) Variation in self-fertility and the reproductive advantage of self-fertility for an invading plant (*Spartina alterniflora*). *Evolutionary Ecology* **12**, 553-568.
- Davies N, Villablanca FX, Roderick GK (1999) Determining the source of individuals: multilocus genotyping in nonequilibrium population genetics. *Trends in Ecology & Evolution* **14**, 17-21.
- Davis HG, Taylor CM, Civille JC, Strong DR (2004) An Allee effect at the front of a plant invasion: *Spartina* in a Pacific estuary. *Journal of Ecology* **92**, 321-327.

- De Frenne P, Kolb A, Verheyen K, *et al.* (2009) Unravelling the effects of temperature, latitude and local environment on the reproduction of forest herbs. *Global Ecology and Biogeography* **18**, 641-651.
- DeWalt SJ, Hamrick JL (2004) Genetic variation of introduced Hawaiian and native Costa Rican populations of an invasive tropical shrub, *Clidemia hirta* (Melastomataceae). *American Journal of Botany* **91**, 1155-1162.
- Dlugosch KM (2006) *Adaptation and colonization in Hypericum canariense: Past and present invasions of an island endemic* Ph.D., University of California, Santa Cruz.
- Dlugosch KM, Parker IM (2007) Molecular and quantitative trait variation across the native range of the invasive species *Hypericum canariense*: evidence for ancient patterns of colonization via pre-adaptation? *Molecular Ecology* **16**, 4269-4283.
- Dlugosch KM, Parker IM (2008a) Founding events in species invasions: genetic variation, adaptive evolution, and the role of multiple introductions. *Molecular Ecology* **17**, 431-449.
- Dlugosch KM, Parker IM (2008b) Invading populations of an ornamental shrub show rapid life history evolution despite genetic bottlenecks. *Ecology Letters* **11**, 701-709.
- Dobzhansky T, Wright S (1941) Genetics of Natural Populations. V. Relations between mutation rate and accumulation of lethals in populations of *Drosophila pseudoobscura*. *Genetics* **26**, 23-51.
- Drake JM, Lodge DM (2006) Allee effects, propagule pressure and the probability of establishment: risk analysis for biological invasions. *Biological Invasions* **8**, 365-375.
- Eckert CG, Samis KE, Loughheed SC (2008) Genetic variation across species' geographical ranges: the central-marginal hypothesis and beyond. *Molecular Ecology* **17**, 1170-1188.
- Ellstrand NC, Marshall DL (1985) Interpopulation gene flow by pollen in wild radish, *Raphanus sativus*. *The American Naturalist* **126**, 606-616.
- Ellstrand NC, Schierenbeck KA (2000) Hybridization as a stimulus for the evolution of invasiveness in plants? *Proceedings of the National Academy of Sciences of the United States of America* **97**, 7043-7050.
- Epperson BK, Allard RW (1989) Spatial autocorrelation analysis of the distribution of genotypes within populations of lodgepole pine. *Genetics* **121**, 369-377.
- Erickson DL, Hamrick JL, Kochert GD (2004) Ecological determinants of genetic diversity in an expanding population of the shrub *Myrica cerifera*. *Molecular Ecology* **13**, 1655-1664.
- Estoup A, Beaumont M, Sennedot F, Moritz C, Cornuet JM (2004) Genetic analysis of complex demographic scenarios: Spatially expanding populations of the cane toad, *Bufo marinus*. *Evolution* **58**, 2021-2036.
- Excoffier L, Laval G, Schneider S (2005) Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* **1**, 47-50.

- Excoffier L, Smouse P, Quattro J (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics* **131**, 479-491.
- Fagan WF, Lewis MA, Neubert MG, van den Driessche P (2002) Invasion theory and biological control. *Ecology Letters* **5**, 148-157.
- Ferrière R, Le Galliard J (2001) Invasion fitness and adaptive dynamics in spatial population models. In: *Dispersal* (eds. Clobert J, Danchin E, Dhondt AA, Nichols JD), pp. 57-79. Oxford University Press Inc., New York.
- Ficetola GF, Bonin A, Miaud C (2008) Population genetics reveals origin and number of founders in a biological invasion. *Molecular Ecology* **17**, 773-782.
- Fisher R (1930) *The genetical theory of natural selection* Clarendon, Oxford.
- Fix AG (1997) Gene frequency clines produced by kin-structured founder effects. *Human Biology* **69**, 663-673.
- Fletcher OS, Swaller JR (1939) *Brachypodium sylvaticum* Huds. Herbarium of Oregon State College, Eugene, OR.
- Franceschinelli EV, Kesseli R (1999) Population structure and gene flow of the Brazilian shrub *Helicteres brevispira*. *Heredity* **82**, 355-363.
- Gaggiotti OE (1999) A comparison of two indirect methods for estimating average levels of gene flow using microsatellite data. *Molecular Ecology* **8**, 1513-1520.
- Galloway LF (2005) Maternal effects provide phenotypic adaptation to local environmental conditions. *New Phytologist* **166**, 93-100.
- Gao H, Williamson S, Bustamante CD (2007) A Markov Chain Monte Carlo Approach for Joint Inference of Population Structure and Inbreeding Rates From Multilocus Genotype Data. *Genetics* **176**, 1635-1651.
- Garvin DF, Gu Y, Hasterok R, *et al.* (2008) Development of genetic and genomic research resources for *Brachypodium distachyon*, a new model system for grass crop research. *Crop Science* **48**, S69-S84.
- Gaunt TR, Rodriguez S, Zapata C, Day INM (2006) MIDAS: software for analysis and visualization of interallelic disequilibrium between multiallelic markers. *BMC Bioinformatics* **7**, 227.
- Gelembiuk G, May GE, Eunmilee C (2006) Phylogeography and systematics of zebra mussels and related species. *Molecular Ecology* **15**, 1033-1050.
- Greene DF, Calogeropoulos C (2002) Measuring and modelling seed dispersal of terrestrial plants. In: *Dispersal Ecology* (eds. Bullock JM, Kenward RE, Hails RS), pp. 3-23. Blackwell Science Ltd, Reading, U.K.
- Guo S, Thompson E (1992) Performing the exact test of Hardy-Weinberg proportion for multiple alleles. *Biometrics* **48**, 361-372.
- Haag CR, Riek M, Hottinger JW, Pajunen VI, Ebert D (2005) Genetic diversity and genetic differentiation in *Daphnia* metapopulations with subpopulations of known age. *Genetics* **170**, 1809-1820.
- Hairston NG, Jr., Van Brunt RA, Kearns CM, Engstrom DR (1995) Age and survivorship of diapausing eggs in a sediment egg bank. *Ecology* **76**, 1706-1711.

- Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* **41**, 95-98.
- Hanski I (1997) Predictive and practical metapopulation models: The incidence function approach. In: *Spatial Ecology: The role of space in population dynamics and interspecific interactions* (eds. Tilman D, Kareiva P), pp. 21-45. Princeton University Press, Princeton, New Jersey.
- Hanski I (2001) Population dynamic consequences of dispersal in local populations and in metapopulations. In: *Dispersal* (eds. Clobert J, Danchin E, Dhondt AA, Nichols JD), pp. 283-298. Oxford University Press, Inc., New York.
- Hardy GH (1908) Mendelian proportions in a mixed population. *Science* **28**, 49-50.
- Hassel K, Såstad SM, Gunnarsson U, Söderström L (2005) Genetic variation and structure in the expanding moss *Pogonatum dentatum* (Polytrichaceae) in its area of origin and in a recently colonized area. *American Journal of Botany* **92**, 1684-1690.
- Hasterok R, Draper J, Jenkins G (2004) Laying the cytotaxonomic foundations of a new model grass, *Brachypodium distachyon* (L.) Beauv. *Chromosome Research* **12**, 397-403.
- Hauswaldt JS, Glenn TC (2003) Microsatellite DNA loci from the Diamondback terrapin (*Malaclemys terrapin*). *Molecular Ecology Notes* **3**, 174-176.
- Hedrick PW (2005) A standardized genetic differentiation measure. *Evolution* **59**, 1633-1638.
- Heinken T, Raudnitschka D (2002) Do Wild Ungulates Contribute to the Dispersal of Vascular Plants in Central European Forests by Epizoochory? A Case Study in NE Germany. . *Forstwissenschaftliches Centralblatt* **121**, 179-194.
- Henry P, Le Lay G, Goudet J, *et al.* (2009) Reduced genetic diversity, increased isolation and multiple introductions of invasive giant hogweed in the western Swiss Alps. *Molecular Ecology* **18**, 2819-2831.
- Herlihy CR, Eckert CG (2005) Evolution of self-fertilization at geographical range margins? A comparison of demographic, floral, and mating system variables in central vs. peripheral populations of *Aquilegia canadensis* (Ranunculaceae). *American Journal of Botany* **92**, 744-751.
- Hewitt GM (1996) Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society* **58**, 247-276.
- Hewitt GM (1999) Post-glacial re-colonization of European biota. *Biological Journal of the Linnean Society* **68**, 87-112.
- Hey J, Machado CA (2003) The study of structured populations - New hope for a difficult and divided science. *Nature Reviews Genetics* **4**, 535-543.
- Higgins SI, Richardson DM (1999) Predicting plant migration rates in a changing world: The role of long-distance dispersal. *American Naturalist* **153**, 464-475.
- Hill WG, Robertson A (1968) Linkage disequilibrium in finite populations. *Theoretical and Applied Genetics* **38**, 226-231.

- Hitchcock CL, Cronquist A, Ownbey M, Thompson JW (1969) *Vascular Plants of the Pacific Northwest, Part I: Vascular Cryptogams, Gymnosperms, and Monocotyledons*. University of Washington Press, Seattle.
- Hogweg P (1988) Cellular automata as a paradigm for ecological modeling. *Applied Mathematics and Computation* **27**, 81-100.
- Holsinger KE (1986) Dispersal and plant mating systems: The evolution of self-fertilization in subdivided populations. *Evolution* **40**, 405-413.
- Huey RB, Gilchrist GW, Hendry AP (2005) Using invasive species to study evolution: case studies with *Drosophila* and salmon.
- Hull ACJ (1974) Species for seeding mountain rangelands in Southeastern Idaho, Northeastern Utah, and Western Wyoming. *Journal of Range Management* **27**, 150-153.
- Husseneder C, Messenger MT, Su N, Grace JK, Vargo EL (2005) Colony social organization and population genetic structure of an introduced population of Formosan subterranean termite from New Orleans, Louisiana. *Journal of Economic Entomology* **98**, 1421-1434.
- Hutchison DW, Templeton AR (1999) Correlation of pairwise genetic and geographic distance measures: Inferring the relative influences of gene flow and drift on the distribution of genetic variability. *Evolution* **53**, 1898-1914.
- Ibrahim KM, Nichols RA, Hewitt GM (1996) Spatial patterns of genetic variation generated by different forms of dispersal during range expansion. *Heredity* **77**, 282-291.
- Ingvarsson PK, Giles BE (1999) Kin-structured colonization and smallscale genetic differentiation in *Silene dioica*. *Evolution* **53**, 605-611.
- Ingvarsson PK (2002) A metapopulation perspective on genetic diversity and differentiation in partially self-fertilizing plants. *Evolution* **56**, 2368-2373.
- Jain SK, Rai KN, Singh RS (1981) Population biology of *Avena* XI. Variation in peripheral isolates of *A. barbata*. *Genetica* **56**, 213-215.
- Johnson ML, Gaines MS (1990) Evolution of dispersal: theoretical models and empirical tests using birds and mammals. *Annual Review of Ecology and Systematics* **21**, 449-480.
- Jones FA, Hamrick JL, Peterson CJ, Squiers ER (2006) Inferring colonization history from analyses of spatial genetic structure within populations of *Pinus strobus* and *Quercus rubra*. *Molecular Ecology* **15**, 851-861.
- Jost L (2008)  $G_{ST}$  and its relatives do not measure differentiation. *Molecular Ecology* **17**, 4015-4026.
- Judd DR (1983) *Intraspecific variation in response to shade*. Ph.D., University College.
- Kahn MA, Stace CA (1999) Breeding relationships in the genus *Brachypodium* (Poaceae: Pooideae). *Nordic Journal of Botany* **19**, 257-269.
- Kalendar R (2005) FastPCR. Institute of Biotechnology, University of Helsinki, Finland, Helsinki, Finland.
- Kareiva PM (1983) Local movement in herbivorous insects: applying a passive diffusion model to mark-recapture field experiments. *Oecologia* **57**, 322-327.

- Kaye T (2001) *Brachypodium sylvaticum* (Poaceae) in the Pacific Northwest. In: *Botanical Electronic News*.
- Kaye TN, Blakeley-Smith M (2006) False-brome (*Brachypodium sylvaticum*). In: *Invasive Species in the Pacific Northwest* (eds. Boersma PD, Reichard SE, van Buren AN), pp. 80-81. University of Washington Press, Seattle.
- Keitt TH, Lewis MA, Holt RD (2001) Allee effects, invasion pinning, and species' borders. *American Naturalist* **157**, 203-216.
- Koenig WD, Vuren DV, Hooge PN (1996) Detectability, philopatry, and the distribution of dispersal distances in vertebrates. *Trends in Ecology & Evolution* **11**, 514-517.
- Kolbe JJ, Glor RE, Schettino LRG, *et al.* (2004) Genetic variation increases during biological invasion by a Cuban lizard. *Nature* **431**, 177-181.
- Kolbe JJ, Larson A, Losos JB (2007) Differential admixture shapes morphological variation among invasive populations of the lizard *Anolis sagrei*. *Molecular Ecology* **16**, 1579-1591.
- Kot M, Lewis MA, vandenDriessche P (1996) Dispersal data and the spread of invading organisms. *Ecology* **77**, 2027-2042.
- Kot M, Medlock J, Reluga T, Walton DB (2004) Stochasticity, invasions, and branching random walks. *Theoretical Population Biology* **66**, 175-184.
- Kruglyak L (1999) Prospects for whole-genome linkage disequilibrium mapping of common disease genes. *Nature Genetics* **22**, 139-144.
- Kruskal JB (1964) Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika* **29**, 1-27.
- Lande R, Schemske DW (1985) The evolution of self-fertilization and inbreeding depression in plants. I. Genetic models. *Evolution* **39**, 24-40.
- Lavergne S, Molofsky J (2007) Increased genetic variation and evolutionary potential drive the success of an invasive grass. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 3883-3888.
- Le Corre V, Kremer A (1998) Cumulative effects of founding events during colonisation on genetic diversity and differentiation in an island and stepping-stone model. *Journal of Evolutionary Biology* **11**, 495-512.
- Le Corre V, Machon N, Petit RJ, Kremer A (1997) Colonization with long-distance seed dispersal and genetic structure of maternally inherited genes in forest trees: a simulation study. *Genetical Research* **69**, 117-125.
- Lee CE (2002) Evolutionary genetics of invasive species. *Trends in Ecology & Evolution* **17**, 386-391.
- Lee PLM, Patel RM, Conlan RS, Wainwright SJ, Hipkin CR (2004) Comparison of genetic diversities in native and alien populations of hoary mustard (*Hirschfeldia incana* L. Lagreze-Fossat). *International Journal of Plant Sciences* **165**, 833-843.
- Leger EA, Espeland EK, Merrill KR, Meyer SE (2009) Genetic variation and local adaptation at a cheatgrass (*Bromus tectorum*) invasion edge in western Nevada. *Molecular Ecology* **18**, 4366-4379.

- Levin DA (2003a) Ecological speciation: Lessons from invasive species. *Systematic Botany* **28**, 643-650.
- Levin DA (2003b) The ecological transition in speciation. *New Phytologist* **161**, 91-96.
- Levin SA (1976) Population dynamic models in heterogeneous environments. *Annual Review of Ecology and Systematics* **7**, 287-310.
- Levin SA, Muller-Landau HC, Nathan R, Chave J (2003) The ecology and evolution of seed dispersal: A theoretical perspective. *Annual Review of Ecology Evolution and Systematics* **34**, 575-604.
- Levins RL (1969) Some demographic and genetic consequences of environmental heterogeneity for biological control. *Bulletin of the Entomological Society of America* **15**, 237-240.
- Lewis MA (1997) Variability, patchiness, and jump dispersal in the spread of an invading population. In: *Spatial Ecology: The role of space in population dynamics and interspecific interactions* (eds. Tilman D, Kareiva P), pp. 46-69. Princeton University Press, Princeton, New Jersey.
- Lidicker WZ, Jr., Stenseth NC (1992) To disperse or not to disperse: who does it and why? In: *Animal dispersal: small mammals as a model* (eds. Stenseth NC, Lidicker WZ, Jr.), pp. 21-36. Chapman & Hall, London.
- Liebholt AM, Tobin PC (2008) Population ecology of insect invasions and their management. *Annual Review of Entomology* **53**, 387-408.
- Lindholm AK, Breden F, Alexander HJ, *et al.* (2005) Invasion success and genetic diversity of introduced populations of guppies *Poecilia reticulata* in Australia. *Molecular Ecology* **14**, 3671-3682.
- Litrico I, Ronfort J, Verlaque R, Thompson JD (2005) Spatial structure of genetic variation and primary succession in the pioneer tree species *Antirhea borbonica* on La Réunion. *Molecular Ecology* **14**, 1575-1584.
- Long GM (1989) *Morphological and physiological variation in Brachypodium sylvaticum*. Ph.D. thesis, University of Wales.
- Mack RN, Lonsdale WM (1985) Humans as global plant dispersers: getting more than we bargained for. *Bioscience* **51**, 95-102.
- Mack RN, Simberloff D, Lonsdale WM, *et al.* (2000) Biotic invasions: Causes, epidemiology, global consequences, and control. *Ecological Applications* **10**, 689-710.
- Maroja LS, Bogdanowicz SM, Wallin KF, Raffa KF, Harrison RG (2007) Phylogeography of spruce beetles (*Dendroctonus rufipennis* kirby) (Curculionidae: Scolytinae) in North America. *Molecular Ecology* **16**, 2560-2573.
- Martin LJ, Cruzan MB (1999) Patterns of hybridization in the *Piriqueta caroliniana* complex in central Florida: evidence for an expanding hybrid zone. *Evolution* **53**, 1037-1049.
- Mashanova A, Gange AC, Jansen VAA (2008) Density-dependent dispersal may explain the mid-season crash in some aphid populations. *Population Ecology* **50**, 285-292.

- Masta SE, Laurent NM, Routman EJ (2003) Population genetic structure of the toad *Bufo woodhousii*: an empirical assessment of the effects of haplotype extinction on nested cladistic analysis. *Molecular Ecology* **12**, 1541-1554.
- Matlack GR, Monde J (2004) Consequences of low mobility in spatially and temporally heterogeneous ecosystems. *Journal of Ecology* **92**, 1025-1035.
- Mollison D (1977) Spatial contact models for ecological and epidemic spread. *Journal of the Royal Statistical Society B* **39**, 283-326.
- Molofsky J (1994) Population dynamics and pattern formation in theoretical populations. *Ecology* **75**, 30-39.
- Morrell PL, Toleno DM, Lundy KE, Clegg MT (2005) Low levels of linkage disequilibrium in wild barley (*Hordeum vulgare* ssp. *spontaneum*) despite high rates of self-fertilization. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 2442-2447.
- Muirhead JR, Leung B, van Overdijk C, et al. (2006) Modelling local and long-distance dispersal of invasive emerald ash borer *Agrilus planipennis* (Coleoptera) in North America. *Diversity and Distributions* **12**, 71-79.
- Nathan R (2001) The challenges of studying dispersal. *Trends in Ecology & Evolution* **16**, 481-482.
- Nathan R, Muller-Landau HC (2000) Spatial patterns of seed dispersal, their determinants and consequences for recruitment. *Trends in Ecology & Evolution* **15**, 278-285.
- Nathan R, Perry G, Cronin JT, Strand AE, Cain ML (2003) Methods for estimating long-distance dispersal. *Oikos* **103**, 261-273.
- Nei M (1972) Genetic distance between populations. *American Naturalist* **106**, 283-292.
- Nei M (1987) *Molecular Evolutionary Genetics* Columbia University Press, New York.
- Nei M, Li W (1973) Linkage disequilibrium in subdivided populations. *Genetics* **75**, 213-219.
- Nei M, Maruyama T, Chakraborty R (1975) The bottleneck effect and genetic variability in populations. *Evolution* **29**, 1-10.
- Neigel JE (2002) Is  $F_{st}$  obsolete? *Conservation Genetics* **3**, 167-173.
- Neubert MG, Caswell H (2000) Demography and dispersal: Calculation and sensitivity analysis of invasion speed for structured populations. *Ecology* **81**, 1613-1628.
- Neubert MG, Kot M, Lewis MA (1995) Dispersal and pattern formation in a discrete-time predator-prey model. *Theoretical Population Biology* **48**, 7-43.
- Newman D, Pilson D (1997) Increased probability of extinction due to decreased genetic effective population size: Experimental populations of *Clarkia pulchella*. *Evolution* **51**, 354-362.
- Nichols RA, Hewitt GM (1994) The genetic consequences of long distance dispersal during colonization. *Heredity* **72**, 312-317.
- Ouborg NJ, Piquot Y, Van Groenendael JM (1999) Population genetics, molecular markers and the study of dispersal in plants. *Journal of Ecology* **87**, 551-568.



- Paetkau D, Slade R, Burden M, Estoup A (2004) Genetic assignment methods for the direct, real-time estimation of migration rate: a simulation-based exploration of accuracy and power. *Molecular Ecology* **13**, 55-65.
- Paetkau D, Waits LP, Clarkson PL, Craighead L, Strobeck C (1997) An empirical evaluation of genetic distance statistics using microsatellite data from bear (Ursidae) populations. *Genetics* **147**, 1943-1957.
- Pannell JR, Charlesworth B (1999) Neutral genetic diversity in a metapopulation with recurrent local extinction and recolonization. *Evolution* **53**, 664-676.
- Pannell JR, Dorken ME (2006) Colonisation as a common denominator in plant metapopulations and range expansions: effects on genetic diversity and sexual systems. *Landscape Ecology* **21**, 837-848.
- Park SDE (2001) *Trypanotolerance in West African cattle and the population genetic effects of selection* Ph.D., University of Dublin.
- Parker IM, Rodriguez J, Loik ME (2003) An evolutionary approach to understanding the biology of invasions: Local adaptation and general-purpose genotypes in the weed *Verbascum thapsus*. *Conservation Biology* **17**, 59-72.
- Peakall R, Ruibal M, Lindenmayer DB (2003) Spatial autocorrelation analysis offers new insights into gene flow in the Australian bush rat, *Rattus fuscipes*. *Evolution* **57**, 1182-1195.
- Pearse DE, Crandall KA (2004) Beyond  $F_{ST}$ : Analysis of population genetic data for conservation. *Conservation Genetics* **5**, 585-602.
- Petit RJ, Aguinagalde I, de Beaulieu JL, *et al.* (2003) Glacial refugia: hotspots but not melting pots of genetic diversity. *Science* **300**, 1563-1565.
- Petit RJ, Bialozyt R, Garnier-Gere P, Hampe A (2004) Ecology and genetics of tree invasions: from recent introductions to Quaternary migrations. *Forest Ecology and Management* **197**, 117-137.
- Petit RJ, Pineau E, Demesure B, *et al.* (1997) Chloroplast DNA footprints of postglacial recolonization by oaks. *Proceedings of the National Academy of Sciences of the United States of America* **94**, 9996-10001.
- Phillips BL, Brown GP, Travis JMJ, Shine RM (2008) Reid's paradox revisited: The evolution of dispersal kernels during range expansion. *The American Naturalist* **172**, S34-S48.
- Piry S, Alapetite A, Cornuet JM, *et al.* (2004) GENECLASS2: A software for genetic assignment and first-generation migrant detection. *Journal of Heredity* **95**, 536-539.
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* **155**, 945-959.
- Ramakrishnan AP, Musial T, Cruzan MB (2010) Shifting dispersal modes at an expanding species' range margin. *Molecular Ecology* **19**, 1134-1146.
- Ramakrishnan AP, Rosenthal DM, Dobberstein T, Cruzan MB (2008) Isolation and characterization of nine microsatellite markers for *Brachypodium sylvaticum* (Huds.) Beauv., a recently invasive grass species in Oregon. *Molecular Ecology Resources* **8**, 1297-1299.

- Rannala B, Mountain JL (1997) Detecting immigration by using multilocus genotypes. *Proceedings of the National Academy of Sciences of the United States of America* **4**, 9197-9201.
- Raymond M, Rousset F (1995) GENEPOP, version 1.2: population genetics software for exact tests and ecumenicism. *Journal of Heredity* **86**, 248-249.
- Reid C (1899) *The Origin of the British Flora*, Dulau, London.
- Rezende GL, Martins AJ, Gentile C, *et al.* (2008) Embryonic desiccation resistance in *Aedes aegypti*: presumptive role of the chitinized serosal cuticle. *BMC Developmental Biology* **8**, 82.
- Ritland K (2002) Extensions of models for the estimation of mating systems using  $n$  independent loci. *Heredity* **88**, 221-228.
- Ritland K, Jain S (1981) A model for the estimation of outcrossing rate and gene frequencies using  $n$  independent loci. *Heredity* **47**, 35-52.
- Rosenthal DM, Ramakrishnan AP, Cruzan MB (2008) Evidence for multiple sources of invasion and intraspecific hybridization in *Brachypodium sylvaticum* (Hudson) Beauv. in North America. *Molecular Ecology* **17**, 4657-4669.
- Roy B (2010) *Brachypodium sylvaticum* datasheet In: *CABI Invasive Species Compendium*. CABI.
- Rozen S, Skaletsky H (2000) Primer3 on the WWW for general users and for biologist programmers. In: *Bioinformatics Methods and Protocols: Methods in Molecular Biology* (eds. Krawetz S, Misener S), pp. 365-386. Humana Press, Totowa, NJ.
- Saccheri I, Kuussaari M, Kankare M, *et al.* (1998) Inbreeding and extinction in a butterfly metapopulation. *Nature* **392**, 491-494.
- Sakai AK, Allendorf FW, Holt JS, *et al.* (2001) The population biology of invasive species. *Annual Review of Ecology and Systematics* **32**, 305-332.
- Saltonstall K (2003) Microsatellite variation within and among North American lineages of *Phragmites australis*. *Molecular Ecology* **12**, 1689-1702.
- Sax DF, Gaines SD, Brown JH (2002) Species invasions exceed extinctions on islands worldwide: A comparative study of plants and birds. *American Naturalist* **160**, 766-783.
- Sax DF, Stachowicz JJ, Brown JH, *et al.* (2007) Ecological and evolutionary insights from species invasions. *Trends in Ecology & Evolution* **22**, 465-471.
- Schemske DW, Lande R (1985) The evolution of self-fertilization and inbreeding depression in plants. II. Empirical observations. *Evolution* **39**, 41-52.
- Schneider S, Roessli D, Excoffier L (2000) Arlequin, Genetics and Biometry Laboratory, University of Geneva, Switzerland.
- Shaw DV, Kahler AL, Allard RW (1981) A multilocus estimator of mating system parameters in plant populations. *Proceedings of the National Academy of Science, USA* **78**, 1298-1302.
- Shigesada N, Kawasaki K, Takeda Y (1995) Modeling stratified diffusion in biological invasions. *American Naturalist* **146**, 229-251.
- Skalski GT, Gilliam JF (2003) A diffusion-based theory of organism dispersal in heterogeneous populations. *American Naturalist* **161**, 441-458.

- Skellam JG (1951) Random dispersal in theoretical populations. *Biometrika* **38**, 196-218.
- Slatkin M (1987) Gene flow and the geographic structure of natural populations. *Science* **236**.
- Slatkin M (1993) Isolation by distance in equilibrium and non-equilibrium populations. *Evolution* **47**, 264-279.
- Slatkin M (1994) Linkage disequilibrium in growing and stable populations. *Genetics* **137**, 331-336.
- Slatkin M, Excoffier L (1996) Testing for linkage disequilibrium in genotypic data using the Expectation-Maximization algorithm. *Heredity* **76**, 377-383.
- Soltis DF, Gitzendanner MA, Strenge DD, Soltis PS (1997) Chloroplast DNA intraspecific phylogeography of plants from the Pacific Northwest of North America. *Plant Systematics and Evolution* **206**, 353-373.
- Stenseth NC, Lidicker WZ, Jr. (1992) The study of dispersal: a conceptual guide. In: *Animal dispersal: small mammals as a model* (eds. Stenseth NC, Lidicker WZ, Jr.), pp. 5-20. Chapman & Hall, London.
- Stephens PA, Sutherland WJ (1999) Consequences of the Allee effect for behaviour, ecology and conservation. *Trends in Ecology & Evolution* **14**, 401-405.
- Szymura JM, Barton NH (1986) Genetic analysis of a hybrid zone between the fire-bellied toads, *Bombina orientalis* and *B. orientalis*, near Cracow in southern Poland. *Evolution* **40**, 1141-1159.
- Taylor DR, Keller SR (2007) Historical range expansion determines the phylogenetic diversity introduced during contemporary species invasion. *Evolution* **61**, 334-345.
- Telfer S, Pieltney SB, Dallas JF, *et al.* (2003) Parentage assignment detects frequent and large-scale dispersal in water voles. *Molecular Ecology* **12**, 1939-1949.
- Teschke M, Mukabayire O, Wiehe T, Tautz D (2008) Identification of selective sweeps in closely related populations of the house mouse based on microsatellite scans. *Genetics* **180**, 1537-1545.
- Thomas CD, Bodsworth EJ, Wilson RJ, *et al.* (2001) Ecological and evolutionary processes at expanding range margins. *Nature* **411**, 577-581.
- Thomson DM (2007) Do source-sink dynamics promote the spread of an invasive grass into a novel habitat? *Ecology* **88**, 3126-3134.
- Tilman D, Kareiva P (1997) Spatial ecology, the role of space in population dynamics and interspecific interactions. Princeton University Press, Princeton, New Jersey.
- Tilman D, Lehman CL, Kareiva P (1997) Population dynamics in spatial habitats. In: *Spatial Ecology: The role of space in population dynamics and interspecific interactions* (eds. Tilman D, Kareiva P), pp. 3-20. Princeton University Press, Princeton, New Jersey.
- Travis JMJ, Dytham C (2002) Dispersal evolution during invasions. *Evolutionary Ecology Research* **4**, 1119-1129.
- Travis SE, Proffitt CE, Lowenfeld RC, Mitchell TW (2002) A comparative assessment of genetic diversity among differently-aged populations of *Spartina*

- alterniflora* on restored versus natural wetlands. *Restoration Ecology* **10**, 37-42.
- Tscharntke T, Brandl R (2004) Plant-insect interactions in fragmented landscapes. *Annual Review of Entomology* **49**, 405-430.
- van den Bosch F, Metz JAJ, Diekmann O (1990) The velocity of population expansion. *Journal of Mathematical Biology* **28**, 529-565.
- Wade MJ, McCauley DE (1988) Extinction and recolonization: Their effects on the genetic differentiation of local populations. *Evolution* **42**, 995-1005.
- Wakeley J, Aliacar N (2001) Gene genealogies in a metapopulation. *Genetics* **159**, 893-905.
- Walker NF, Hulme PE, Hoelzel AR (2003) Population genetics of an invasive species, *Heracleum mantegazzianum*: implications for the role of life history, demographics and independent introductions. *Molecular Ecology* **12**, 1747-1756.
- Waples RS, Gaggiotti OE (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.
- Weisser WW (2001) The effects of predation on dispersal. In: *Dispersal* (eds. Clobert J, Danchin E, Dhondt AA, Nichols JD), pp. 180-188. Oxford University Press, Inc., New York.
- Whitlock MC (1992) Nonequilibrium population structure in forked fungus beetles: Extinction, colonization, and the genetic variance among populations. *The American Naturalist* **139**, 952-970.
- Whitlock MC, McCauley DE (1990) Some Population Genetic Consequences of Colony Formation and Extinction: Genetic Correlations within Founding Groups. *Evolution* **44**, 1717-1724.
- Whitlock MC, McCauley DE (1999) Indirect measures of gene flow and migration:  $F_{ST}$  not equal  $1/(4Nm+1)$ . *Heredity* **82**, 117-125 Part 112.
- Wiens JA (2001) The landscape context of dispersal. In: *Dispersal* (eds. Clobert J, Danchin E, Dhondt AA, Nichols JD), pp. 96-109. Oxford University Press Inc., Oxford.
- Wilson GA, Rannala B (2003) Bayesian inference of recent migration rates using multilocus genotypes. *Genetics* **163**, 1177-1191.
- Wilson JRU, Dormontt EE, Prentis PJ, Lowe AJ, Richardson DM (2009) Something in the way you move: dispersal pathways affect invasion success. *Trends in Ecology & Evolution* **24**, 136-144.
- Wolfenbarger DO (1975) *Factors affecting dispersal distances of small organisms* Exposition Press, Hicksville, NY.
- Wolff JO (1997) Population regulation in mammals: an evolutionary perspective. *Journal of Animal Ecology* **66**, 1-13.
- Wolff JO, Plissner JH (1998) Sex biases in avian natal dispersal: an extension of the mammalian model. *Oikos* **83**, 327-330.
- Wright S (1931) Evolution in Mendelian populations. *Genetics* **16**, 97-159.
- Wright S (1943) Isolation by Distance. *Genetics* **28**, 114-138.

- Wright S (1946) Isolation by distance under diverse systems of mating. *Genetics* **31**, 39-59.
- Wright S (1951) The genetical structure of populations. *Annals of Eugenics* **15**, 323-354.
- Wright S (1980) Genic and organismal selection. *Evolution* **34**, 825-843.
- Yang S, Bishop JG, Webster MS (2008) Colonization genetics of an animal-dispersed plant (*Vaccinium membranaceum*) at Mount St Helens, Washington. *Molecular Ecology* **17**, 731-740.
- Yeh F, Yang RC, Boyle T (1999) POPGENE. Molecular Biology and Biotechnology Centre, University of Alberta, Edmonton, Canada.
- Young AG, Warwick SI, Merriam HG (1993) Genetic variation and structure at three spatial scales for *Acer saccharum* (sugar maple) in Canada and the implications for conservation. *Canadian Journal of Forest Research* **23**, 2568-2578.
- Zapata C, Rodríguez S, Visedo G, Sacristán F (2001) Spectrum of nonrandom associations between microsatellite loci on human chromosome 11p15. *Genetics* **158**, 1235-1251.
- Zhang W, Collins A, Gibson J, *et al.* (2004) Impact of population structure, effective bottleneck time, and allele frequency on linkage disequilibrium maps. *Proceedings of the National Academy of Science, USA* **101**, 18075-18080.