Conserv Genet (2013) 14:467–481
DOI 10.1007/s10592-012-0411-5

RESEARCH ARTICLE

Identifying genetic signatures of selection in a non-model species, alpine gentian (*Gentiana nivalis* L.), using a landscape genetic approach

Helen Bothwell · Sarah Bisbing · Nina Overgaard Therkildsen · Lindsay Crawford · Nadir Alvarez · Rolf Holderegger · Stéphanie Manel

Received: 7 February 2012 / Accepted: 8 September 2012 / Published online: 26 September 2012

Abstract It is generally accepted that most plant populations are locally adapted. Yet, understanding how environmental forces give rise to adaptive genetic variation is a challenge in conservation genetics and crucial to the preservation of species under rapidly changing climatic conditions. Environmental variation, phylogeographic history, and population demographic processes all contribute to spatially structured genetic variation, however few current models attempt to separate these confounding effects. To illustrate the benefits of using a spatially-explicit model for identifying potentially adaptive loci, we compared outlier locus detection methods with a recently-developed landscape genetic approach. We analyzed 157 loci from samples of the alpine herb *Gentiana nivalis* collected across the European Alps. Principle coordinates of neighbor matrices (PCNM), eigenvectors that quantify multi-scale spatial variation present in a data set, were incorporated into a landscape genetic approach relating AFLP frequencies with 23 environmental variables. Four major findings emerged. 1) Fifteen loci were significantly correlated with at least one predictor variable ($R^2_{adj} > 0.5$). 2) Models including PCNM variables identified eight more potentially adaptive loci than models run without spatial variables. 3) When compared to outlier detection methods, the landscape genetic approach detected four of the same loci plus 11 additional loci. 4) Temperature, precipitation, and solar radiation were the three major environmental factors driving potentially adaptive genetic variation in *G. nivalis*. Techniques presented in this paper offer an efficient method for identifying potentially adaptive genetic variation and associated environmental forces of selection, providing an important step forward for the conservation of non-model species under global change.

Electronic supplementary material The online version of this article (doi:10.1007/s10592-012-0411-5) contains supplementary material, which is available to authorized users.

H. Bothwell (✉)
Environmental Genetics and Genomics Laboratory, Department of Biological Sciences, Northern Arizona University, P.O. Box 5640, Flagstaff, AZ 86011-5640, USA
e-mail: Helen.Bothwell@nau.edu

S. Bisbing
Graduate Degree Program in Ecology, Colorado State University, 1472 Campus Delivery, Fort Collins, CO 80523, USA

N. O. Therkildsen
Section for Population Ecology and Genetics, National Institute of Aquatic Resources, Technical University of Denmark, Vejlsøvej 39, 8600 Silkeborg, Denmark

L. Crawford
Department of Biology, Western University, 1151 Richmond St, London, ON N6A 5B7, Canada

N. Alvarez
Department of Ecology and Evolution, Biophore Building University of Lausanne, 1015 Lausanne, Switzerland

R. Holderegger
WSL Swiss Federal Research Institute, Zürcherstrasse 111, 8903 Birmensdorf, Switzerland

S. Manel
Laboratoire Population Environnement Développement, Université Aix-Marseille, 3 Place Victor Hugo, 13331 Marseille Cedex 03, France

Keywords Adaptive genetic variation · Landscape genetics · Allele distribution models · Outlier locus detection · Principle coordinates of neighbor matrices · Climate change

Introduction

An emerging field of research within landscape genetics is the study of adaptive genetic variation in response to environmental change (Holderegger and Wagner 2008; Manel and Segelbacher 2009; Manel et al. 2010a; Schoville et al. 2012). Under rapidly changing climatic conditions, the capacity of organisms to adapt is crucial to their persistence. This is particularly important for sessile organisms, such as plants, living at range margins or under extreme environmental conditions (Crawford 2008). Arctic and alpine plants, for example, are expected to be particularly sensitive to environmental change (Walther et al. 2002; Thuiller 2007), and many are already showing upward elevational shifts of one to four meters per decade correlated with general warming trends (Walther et al. 2002; Körner 2003). The development of appropriate methods for identifying adaptive genetic variation in natural populations is essential to advancing our understanding of species' responses to global change and for conserving the evolutionary potential of natural populations (Allendorf et al. 2010; Manel et al. 2010b; Hansen et al. 2012).

Several genomic methods exist for identifying adaptively important genes and loci linked to genetic regions under selection (e.g. gene expression profiling, candidate gene association mapping, quantitative trait loci mapping). These methods, however, have limited applicability in natural populations where environmental heterogeneity and population genetic structure resulting from demographic history potentially confound the signature of selection (Joost et al. 2007; Excoffier et al. 2009; see below). New genetic approaches are currently being developed to address these limitations and provide alternative methodologies for studying adaptive genetic variation in the field. These approaches are particularly valuable for studying non-model species for which little prior genomic knowledge is available and for species whose timely conservation management of genetic resources is imperative. With these new techniques, landscape-level genetic data sampled across natural ecological gradients can be analyzed in conjunction with readily-available environmental data (e.g. downloadable GIS climate layers) to detect genetic regions of adaptive importance (Holderegger et al. 2008, 2010; Manel et al. 2010a, 2012; Coop et al. 2010). This approach is versatile and can be applied to any genetic data set for which a large number of loci and allele frequency data are available, including a wide variety of marker systems such as single nucleotide polymorphism (SNPs) and microsatellites (MSATs). Although SNPs are widely used to search for genome-wide locus by environment associations in model organisms (e.g. Fournier-Level et al. 2011), we focus on the utility of amplified fragment length polymorphism (AFLP) markers, because they can easily be applied to non-model organisms and used to generate hundreds of loci potentially widely-distributed across the genome (Meudt and Clarke 2007). AFLPs provide a quick and low-cost means of obtaining allele frequency data for large sample sizes and organisms for which little prior genetic knowledge is available (Bensch and Åkesson 2005; Poncet et al. 2010). Once allele frequency data has been compiled, loci linked to genetic regions under selection can be identified using one of the following two approaches: 1) outlier locus detection or 2) landscape genetic allele distribution modeling.

In outlier locus detection, loci that demonstrate significantly higher or lower among-population genetic differentiation than expected under neutrality are identified as outliers and are thus considered potentially under selection. Of the large number of markers usually surveyed in this method, less than 5 % are generally identified as outliers (Hoffmann and Willi 2008). Outlier locus detection is a population-level analysis that relies on estimates of population genetic differentiation (e.g., F_{ST}) and is thus potentially sensitive to departures from Hardy–Weinberg equilibrium (HWE) common in many natural populations that violate assumptions in the designation of neutral population genetic structure (Excoffier et al. 2009). This problem increases when we attempt to define clusters by minimizing HWE for populations that are not well differentiated. In fact in most of cases, natural populations are very difficult to define (Waples and Gaggiotti 2006). Additionally, selection caused by factors that vary on spatial scales smaller than that of population designation, may not be detectable with population-based methods. Finally, outlier locus detection provides no direct connection between genetic and environmental data; environmental factors are not used to identify outlier loci, although correlations with selected loci can be assessed post hoc (Storz 2005; Holderegger et al. 2008).

Alternatively, a landscape genetic approach can be used to detect potentially adaptive loci by directly correlating allele frequencies with environmental variation (Joost et al. 2007; Schmidt et al. 2008; Holderegger et al. 2010). This approach assumes that clinal variation in the environment produces measureable changes in allele frequencies in or near genomic regions under selection (Endler 1977; Hirao and Kudo 2004; Ingvarsson et al. 2006; Schmidt et al. 2008; Manel et al. 2010a). It has the advantage of being able to identify both markers potentially linked to genes under selection as well as the environmental factors driving

selection. As high-throughput technology increases the amount of genetic data to be analyzed, the calculation time it takes to process very large data sets using outlier locus detection methods may become an obstacle to users, necessitating alternative approaches. In contrast, the landscape genetic approach is able to process large amounts of data very rapidly. By explicitly including spatial and environmental variables when identifying loci under selection, allele distribution models are typically able to identify a larger percentage of potentially adaptive loci than outlier locus detection methods (e.g. Parisod and Joost 2010; Poncet et al. 2010, Manel et al. 2010b). Incorporation of additional environmental information allows the landscape genetic approach to retrieve loci that lie just outside the upper or lower limits of the neutral distribution and identified as false negatives by outlier locus detection methods. Use of both outlier locus detection and landscape genetic approaches together will complement and strengthen robustness of the final set of loci identified as potentially under selection. A direct comparison of these approaches, however, has as yet not been addressed.

Regardless of the approach used to identify loci under selection, it is often difficult to disentangle correlations between identified loci and environmental factors from correlations with other spatially structured processes (Manel et al. 2010a). Spatially structured genetic variation can arise as a result of three major processes. 1) Induced spatial dependence of genetic variation arises when species respond to spatially structured external variables, such as environmental clines in temperature or moisture. 2) Spatial autocorrelation is generated by species themselves, through demographic processes such as restricted mating, limited dispersal, and interspecific interactions (Borcard et al. 2011). 3) Lastly, the historical signature of broad-scale, phylogeographic structuring patterns associated with glacial refugia and migration routes can give rise to spatially structured patterns of genetic variation (Dray et al. 2006; Jombart et al. 2009). Understanding species' genetic adaptations to changing environmental conditions requires separating the signature of environmental forces of selection from both spatial autocorrelation and broad-scale phylogeographic structuring that results from neutral genetic processes. Spatial autocorrelation, whereby observations exhibit a greater or lesser degree of correlation with geographic location than expected under a neutral distribution, violates assumptions of error structure in a given dataset, resulting in high rates of Type I error (Wagner and Fortin 2005; Dray et al. 2006). Left unaddressed, spatially structured environmental heterogeneity introduces statistical concerns when comparing replicate samples across the landscape (Wagner and Fortin 2005). Consequently, when analyzing locus by environment associations, models that explicitly account for the spatial component of variation inherent in environmental datasets are needed (e.g., partialling out the spatial component from residual variation, adding a spatial error term in regression analysis) (Wagner and Fortin 2005).

Principle coordinates of neighbor matrices (PCNM) is a method that allows partitioning out of the spatial component intrinsic to ecological datasets (Borcard and Legendre 2002). PCNM is a special form of Moran's eigenvector maps (MEM); both utilize Moran's index of spatial autocorrelation (Moran's *I*) to characterize pairwise spatial relationships among sampling locations (Dray et al. 2006). While MEM quantifies both negative and positive autocorrelation, PCNM only measures the latter. Negative autocorrelation tends to capture very small-scale local demographic processes; excluding negative eigenvectors, PCNM and MEM converge (Dray et al. 2006; Dray pers comm). As we were primarily interested in the influence of broad-scale environmental forces of selection on patterns of genetic variation in this study, we focused on PCNM. Collectively, PCNM eigenvectors describe a continuum of variation across all scales that can be ascertained by a studies' sampling design. They allow one to account for the spatial structure present in a given dataset that is due to spatially autocorrelated environmental and demographic patterns, without those variables having to be directly measured or identified. PCNM eigenvectors can be used as spatial predictors, e.g., in multiple regression (Dray et al. 2006). Quantifying this unmeasured spatial variation allows researchers to salvage information from residual variation, thereby reducing model error and increasing the likelihood of discovering ecologically-meaningful relationships (Borcard et al.1992; Borcard and Legendre 2002). Broad-scale PCNM eigenvectors are thought to represent the spatial structure inherent in climatic gradients, such as temperature or precipitation, while smaller-scale PCNM eigenvectors correspond to intermediate- to small-scale biotic processes, such as dispersal and competition (Wagner and Fortin 2005; Jombart et al. 2009).

Over the last two decades, PCNM has been utilized as a surrogate for spatial variation in analysis of a wide variety of ecological patterns, including spruce budworm (*Choristoneura* spp.) defoliation (Bellier et al. 2007), hotspots of plant community diversity (Gibson et al. 2010), and intertidal metacommunity structure (Okuda et al. 2010). Incorporation of PCNM as explanatory variables in regression analysis has, however, only very recently been applied in a genetic context for the identification of loci under selection (Manel et al. 2010b). Allele distribution models that include broad-scale PCNM variables offer a promising new tool that requires further exploration (Holderegger et al. 2010).

We apply these methods to a data set of the arctic-alpine species *Gentiana nivalis*, using AFLP allele frequencies

analyzed from samples collected throughout the European Alps (Gugerli et al. 2008). In this study, we aim to validate and gain a deeper understanding of the landscape genetic method recently proposed by Manel et al. (2010b) by addressing the following three main goals. 1) Identify loci exhibiting potentially adaptive patterns of genetic variation by applying multiple linear regression analysis and testing for correlations between allele frequencies and environmental variables, PCNM variables, and population genetic history. 2) Identify major environmental factors potentially driving local adaptive genetic variation in *G. nivalis* at each locus identified in part 1. 3) Compare results from this landscape genetic approach with two outlier locus detection methods to corroborate results and demonstrate the effectiveness of accounting for unmeasured environmental variation when identifying adaptive genetic variation.

Materials and methods

Study species

Gentiana nivalis L. (alpine gentian; Gentianaceae) is a widespread, herbaceous annual plant distributed across arctic and alpine regions of Europe's major mountain ranges, including the Alps, as well as more northern latitudes in Greenland, Scandinavia, and Iceland (Hultén and Fries 1986). *Gentiana nivalis* typically grows under cool, moist conditions, such as those typical of mountain meadows, pastures, heaths, fens, and bogs (Hegi 1957; Kozuharova and Anchev 2006). Soils are commonly characterized as meager, rocky, and usually calcareous (Hess et al. 1972). Flowering occurs from June through September (Aeschimann et al. 2004). The small, blue flowers of *G. nivalis* are self-compatible (Kozuharova and Anchev 2002) and produce numerous, lightweight seeds that can be transported by wind over long distances (Hegi 1957).

Study area, genetic data, and population structure

We utilized a data set of *G. nivalis* generated from the IntraBioDiv project (Gugerli et al. 2008). In an effort to support large-scale conservation planning, this international, interdisciplinary research collaborative compiled genetic data sets for 45 herbaceous species (Gugerli et al. 2008; Taberlet et al. 2012). As part of this project, leaf samples of *G. nivalis* were collected within a regular grid system (resolution of 22.3 × 25 km) across the European Alps (covering an area >200,000 km^2) in 2004 (Fig. 1). Within every second cell of the sampling grid, one location was selected; along a transect established at this location, leaves of three *G. nivalis* individuals were collected, with individuals sampled at 10 m intervals (for further details on sampling see Gugerli et al. 2008). The current study includes 218 individuals sampled from 74 locations. Alvarez et al. (2009) successfully genotyped these samples at 157 polymorphic AFLP loci and found reproducibility of genotypes was greater than 95 %.

Demographic history, as reflected in population genetic structure, can be a dominant factor shaping allele frequencies, and numerous methods for identifying genetic signatures of selection have been shown to be sensitive to this influence (Nielsen 2005, Excoffier et al. 2009). To control for the confounding effects of population genetic history when identifying potential signatures of selection (Nielsen 2005; Joost et al. 2007), we used a Bayesian clustering method implemented in STRUCTURE 2.3.3 to analyze population structure of the *G. nivalis* data set and ascertain if the sampled individuals grouped into distinct genetic lineages (Pritchard et al. 2000, Falush et al. 2007). Assuming admixture and an independent alleles model, we determined the most plausible number of populations ($K = 1$–10, five runs for each K) and with what probability each individual could be assigned to each cluster (200,000 generation burn-in period followed by 1,000,000 Monte Carlo Markov chain (MCMC) generations). Following methods outlined in the STRUCTURE manual, we found the strongest support for a model of $K =$ four ancestral populations (mean posterior -log likelihood $= -8641.5$) (Falush et al. 2007), and subsequently removed all individuals with probability of population membership <0.75 from further analyses (Table 1). To ensure that we identified true correlations between loci and environmental factors and not spurious relationships due to spatially structured patterns of demographic history, we included cluster membership as a covariate in later regression analyses. While admixed individuals likely harbor important information, they also introduce an added source of uncertainty, increasing model variance and reducing our ability to test the robustness of the landscape genetic model to population history. We chose a moderate threshold of >0.75 so as to exclude highly admixed samples, while also retaining valuable genetic information that would be lost with a very high threshold (e.g., 0.99 as is commonly used for assignment tests) (Manel et al. 2002).This resulted in a final data set of 199 individuals; 88 % of these individuals (176 out of 199) exhibited >0.9 probability of belonging to a given population. Figure 1 illustrates the geographic distribution of the four populations.

We analyzed pairwise population differentiation among the four genetic clusters with phi-statistics, analogues of Wright's F-statistics for binary data. Analysis of molecular variance (AMOVA) was conducted in GenAlEx 6.4 using 9,999 permutations (Peakall and Smouse 2006).

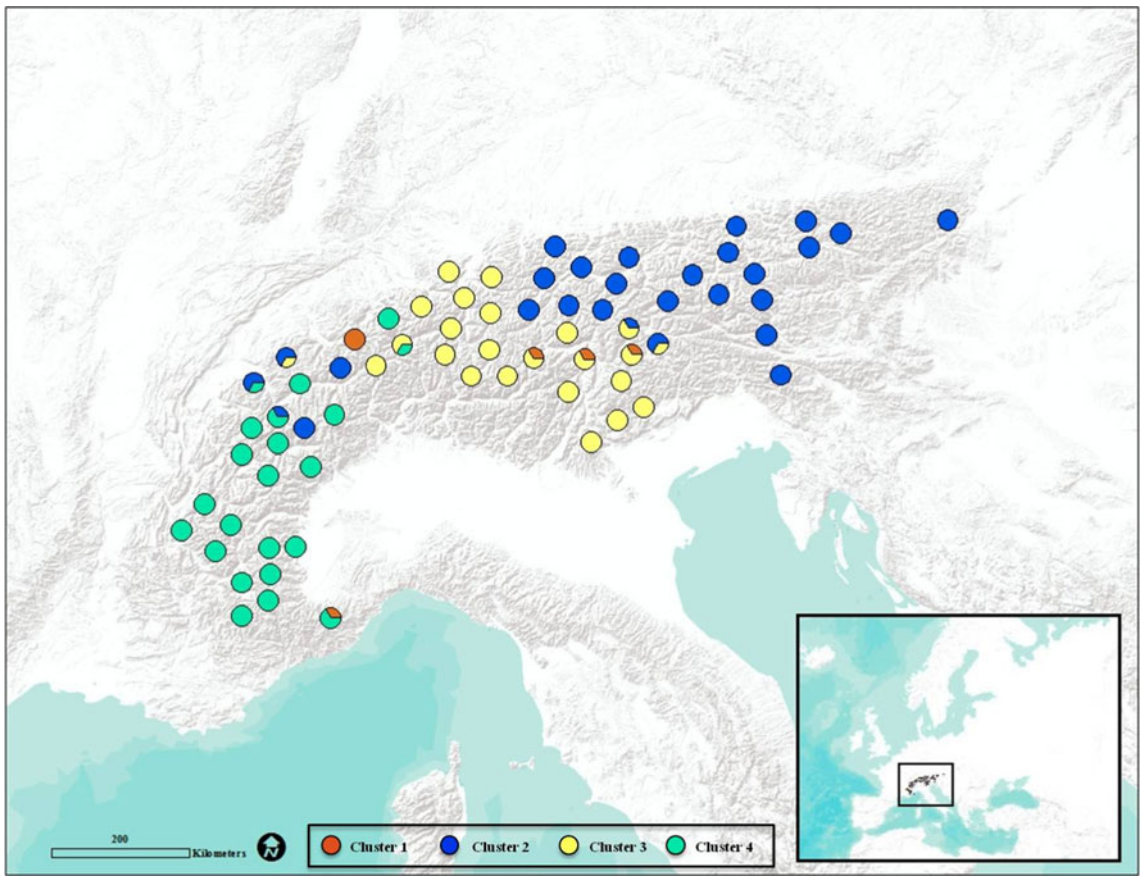

Fig. 1 Sampling locations of *Gentiana nivalis* across the European Alps. Genetic clusters identified by STRUCTURE show strong geographic differentiation, suggesting distinct phylogeographic history

Environmental and spatial variables

We used 200 m resolution climate data interpolated from records collected daily at 977 climate stations located throughout the Alps (see Gugerli et al. 2008 for details). Table 1 provides an overview of the environmental variables used in this study. Following the methods of Manel et al. (2010b), 14 environmental variables related to temperature, precipitation, and topography were extracted from published GIS eco-climatic layers (Gugerli et al. 2008).

Next, we applied a principal component analysis (PCA) to identify highly correlated variables, defined in this study as those variables with $|r| > 0.8$ and ecologically related. Including all 14 original environmental variables, the first two axes of the PCA explained 84.1 % of the variation in environmental predictors across our study area (PC1 = 69.1 %; PC2 = 15 %). Of these, we retained eight final variables that were both uncorrelated and were the most biologically relevant for explaining local adaptation of *G. nivalis*. These environmental variables were then utilized to identify potentially adaptive loci and to determine the selective forces likely operating on each locus. Variables retained included: *radiation, spring precipitation, summer precipitation, mean annual minimum temperature, slope, aspect, topography*, and *potential soil humidity*.

To account for non-linear relationships between AFLP allele frequencies and these environmental factors, we transformed uncorrelated environmental variables into cubic polynomials (Legendre and Legendre 1998; Manel et al. 2010b), while aspect was converted to its sin and cos for use in linear models. These transformations resulted in a total of 23 untransformed and transformed environmental explanatory variables for use in identifying potentially adaptive loci.

To quantify unmeasured spatial variation and create spatial predictors for inclusion in multiple linear regression, we computed PCNM from the geographic coordinates of sample locations using R packages "AEM" (Stéphane Dray) and "PCNM" (Pierre Legendre) (available at https://r-forge.r-project.org/R/?group_id=195). First, a spatial weighting matrix **W** was calculated from pairwise Euclidean distances measured among geographic coordinates of sampling locations. PCNM eigenvectors were then produced based on this weighting matrix. All PCNM

Table 1 We used the topo-climatic variables included in Table 1 to identify AFLP loci under selection in *Gentiana nivalis* (following the methods of Manel et al. 2010b)

	Description	Abbreviation
Topography	Altitude	alt
	Slope (%)[a]	**slope**
	Aspect[b]	**aspect**
	Topography (integrated topographic exposure map)[a]	**topo**
	Soil (potential soil humidity)[a]	**soil**
Seasonal climate layers[c]	**Spring seasonal precipitation** (number of rain days from March–May)[a]	**spr_prcp**
	Summer seasonal precipitation (number of rain days from June–August)[a]	**sum_prcp**
Yearly climate layers[c]	Annual degree days above 0 °C (°C × days)[d]	ddeg
	Mean annual precipitation sum (cm)	prcpavgy
	Radiation (annual mean of daily global radiation; kJ/m^2/day)[a, e]	**srad**
	Number of days with maximum temperature below freezing	tmaxavgiy
	Maximum average temperature (mean annual maximum temp; °C)	tmax
	Number of days with minimum temperature below freezing	tminavgty
	Minimum average temperature (mean annual minimum temp; °C)[a]	**tmin**

Eight variables were retained for use in final analyses (along with their transformations); these are indicated in bold

[a] First, second, and third order polynomials included for these variables

[b] sin and cos transformed

[c] Period 1980–1989

[d] Calculated from daily climate maps

[e] Horizon-terrain-corrected

Table 2 Phi-statistics revealed significant pairwise population differentiation among all four of the genetic clusters identified from STRUCTURE

	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Cluster 1		0.003	0.001	0.000
Cluster 2	0.098		0.000	0.000
Cluster 3	0.156	0.041		0.000
Cluster 4	0.232	0.138	0.113	

Φ_{PT} values are presented below the diagonal, and accompanying probabilities based on 9,999 permutations are shown above the diagonal

eigenvectors are orthogonal and uncorrelated with each other, indicating that each eigenvector represents a different spatial scale of ecological variation. We used the first half of this matrix (nine PCNM eigenvectors) as a surrogate for broad-scale environmental variation (e.g., temperature and precipitation gradients).

Detecting potentially adaptive loci using regression analyses

To address our first goal of detecting potentially adaptive loci in *G. nivalis*, we used multiple linear regression analysis to identify loci whose allele frequencies were correlated with the 23 explanatory environmental variables and nine broad-scale PCNM variables defined above. (i) First, using a multiple linear model, we regressed allele frequencies at the 70 sampling locations for each AFLP locus on the 23 environmental variables alone. (ii) Next, we ran a model including all environmental variables plus the nine broad-scale PCNM variables to test whether more loci were identified as potentially under selection by accounting for spatial variation in the model. (iii) Finally, we ran both of the above regression models again, adding a predictor variable for population membership as identified in STRUCTURE analysis (see above) to test if phylogeographic history had a significant influence on identification of potentially adaptive variation. To consider loci for further exploration, we required that a locus be significantly correlated to at least one explanatory variable and also required that environmental predictors explained at least 50 % of the genetic variation at that locus (as indicated by Ohtani's (2000) unbiased estimator R^2_{adj}; Manel et al. 2010b). We also employed a null model to explore how many significant positive correlations the regression model identified simply due to chance. By randomly shuffling genotypes with respect to sample identity and associated explanatory variables, and then running the regression model on the randomly shuffled dataset, we identified an expected percentage of positive spurious correlations.

After identifying those loci for which environmental predictors explained at least 50 % of the genetic variation,

we examined which particular environmental variables were correlated with allele frequencies at each of the potentially adaptive loci. For each locus exhibiting a potential signature of selection, we applied an inferential approach to determine the best model and relative contribution of each variable. Akaike information criterion (AIC) was utilized to weight each explanatory variable and for model selection and averaging (Burnham and Anderson 2002). Analyses were conducted using the R package "MuMIn" (Kamil Bartoń, available at http://mumin.r-forge.r-project.org/). We ran candidate models, including all possible combinations of the 23 environmental variables and nine broad-scale PCNM variables, with a penalization parameter (k = 2) for addition of model parameters. We used AICc to correct for small sample sizes (Burnham and Anderson 2004) and considered models with ΔAIC values ≤2 as equally well supported. Analyses were conducted with R version 2.10.1 (R Foundation for Statistical Computing, Vienna, Austria; R Development Core Team (2008/2009)).

Redundancy analysis

We then used redundancy analysis (RDA) to investigate the multivariate relationship between explanatory variables and allele frequencies at potentially adaptive loci in ordination space. RDA allowed us to visualize the ecological patterns represented by broad-scale PCNM variables in this study (Borcard et al. 2011). To test for collinearity among the 23 environmental variables and nine broad-scale PCNM variables, we calculated variance inflation factors (VIF) and removed any variables with VIFs >20 (denoting highly correlated canonical coefficients) (Borcard et al. 2011). We included the remaining PCNM eigenvectors in a co-variable matrix to investigate the effect of broad-scale spatial and environmental patterns on genetic variation after partitioning out the effect of small-scale demographic patterns. We applied 'type 2' scaling to preserve correlations among loci (Borcard et al. 2011). VIF and RDA were conducted with the R package 'rda.Test', version 1.2 (Legendre; available at http://www.bio.umontreal.ca/legendre/).

Outlier locus detection

To compare results from the landscape genetic approach with the broadly applied method of outlier locus detection, we conducted outlier tests among the genetic clusters identified from the STRUCTURE analysis. The basic premise of this approach is that loci under selection should exhibit levels of population differentiation (F_{ST}) that fall outside of the range of F_{ST} values for neutral loci. Due to the small sample size ($n = 7$) of one cluster, outlier locus detection tests were only carried out between three of the four genetic clusters (each containing between 58 and 70 individuals).

We applied two commonly used outlier locus detection methods, namely BAYESCAN and DFDIST. First, BAYESCAN utilizes the Bayesian model of Beaumont and Balding (2004), decomposing F_{ST} into locus- and population-specific components (Foll and Gaggiotti 2008). Using a reversible jump MCMC algorithm, BAYESCAN computes the posterior probability of a selection-based model versus a neutral model of differentiation for each locus. We conducted multiple runs, varying the prior odds (from 10:1 odds to even 1:1 odds) in favour of a neutral model. For each set of priors, we ran 20 short pilot runs followed by 100,000 iterations with a 50,000 burn-in.

DFDIST, as implemented in the program MCHEZA (Antao and Beaumont 2011), is an extension of Beaumont and Nichols' (1996) FDIST model, which allows for the use of dominant markers. The neutral distribution was modeled based on 500,000 data points generated through coalescent simulations under a symmetric island model. This model was parameterized with an estimate of neutral differentiation obtained from the observed data (after removing non-neutral loci) and corrected for the small number of populations. Based on these parameters, the simulated neutral distribution was then used to identify loci exhibiting unusually high or low levels of differentiation.

To avoid spurious findings, we excluded loci where global frequency of the minor allele was <0.05 across all individuals, as recommended in the BAYESCAN manual. For each parameter setting, we conducted three independent runs with each program to ensure convergence of results and corrected for multiple testing by setting a false discovery rate FDR < 0.1 (Benjamini and Hochberg 1995).

Results

Population genetic differentiation

Phi-statistics (Φ_{PT}) revealed significant pairwise differentiation among all four of the genetic clusters identified by STRUCTURE (see Table 2). We found that levels of Φ_{PT} were low to moderate (ranging from 0.098 to 0.232, $p < 0.003$) as would be expected from *G. nivalis*'s life history characteristics (e.g., wide-spread distribution, long-distance wind dispersal).

Detecting potentially adaptive loci using regression analyses

The first regression model (i) including only environmental variables found seven (4.5 %) of the 157 loci tested were

identified as potentially under selection (those with $R^2_{adj} \geq 50$ %; Table 3). The second regression model (ii) including all environmental variables and the nine broad-scale PCNM variables detected an additional eight loci (15 of the 157 loci (9.6 %); Fig. 2; Table 3). Finally (iii) when both of the above regression models were run with the addition of a predictor variable representing membership to one of the four genetic clusters identified by STRUCTURE, we found no effect of population genetic structure on the identification of potentially adaptive loci. However, we note that using the same genetic data to estimate both population genetic structure and detection of loci introduces a multiple testing bias. Lastly, results from the null model indicate that, on average, 2 % of loci were found to be significantly correlated with environmental variables simply due to chance.

To identify the environmental and PCNM variables correlated with each selected locus, we applied AIC_C inferential model selection. The relative importance of predictor variables for explaining genetic variation at each locus as well as across all selected loci are presented in Table 4. Among the environmental variables, average minimum temperature, solar radiation, and spring and summer precipitation were most strongly associated with potentially adaptive loci.

Redundancy analysis

VIF results indicated very low collinearity among all 23 environmental variables and nine broad-scale PCNM variables (mean = 2.8, SD = 2.0). From RDA results, environmental and broad-scale PCNM variables together explained 77 % of the variation in allele frequencies at potentially adaptive loci. Intermediate to small-scale PCNM eigenvectors, however, accounted for 0 % of the variance, indicating that population demographic processes did not have a significant influence on the variation within our matrix of selected AFLP loci (canonical $R^2 = 0.77$). Online Resource 1 shows the relationships between allele frequencies at selected loci and spatial and environmental variables. Note that the first principal coordinate (PCNM1) appears to be highly correlated with three selected loci, suggesting potential adaptation to a common, unidentified environmental factor.

Outlier locus detection

Using the conservative prior odds of 10:1 in favor of the neutral model, BAYESCAN identified no significant outlier loci among the three STRUCTURE groups. For all loci, the posterior probability of being under selection was

Table 3 We summarized AFLP loci identified as potentially adaptive for *Gentiana nivalis* using both a landscape genetic approach and outlier locus detection approaches

Selected loci	Allele distribution method		F_{ST} outlier detection method	
	MLR	MLR incl. PCNMs	DFDIST	BAYESCAN
ACT_CAC_103.1	x	x	x[b]	x[b]
ACT_CAC_218.4			x[a]	
ACT_CAC_320.8			x[a]	
ATC_CAC_100.0	x	x		
ATC_CAC_144.5		x		
ATC_CAC_151.7		x	x[b]	
ATC_CAC_177.8			x[a]	
ATC_CAC_221.4	x	x		
ATC_CAC_225.9	x	x	x[b]	
ATC_CAC_300.0	x	x		
ATC_CAC_302.0	x	x		
ATC_CAC_362.7		x		
ATG_CTG_71.3	x	x	x[b]	
ATG_CTG_113.4		x		
ATG_CTG_140.8			x[b]	
ATG_CTG_199.5		x		
ATG_CTG_200.8		x		
ATG_CTG_236.7		x		
ATG_CTG_458.8			x[a]	
ATG_CTG_489.1		x		
Total	7	15	9	1

Loci identified by both multiple linear regression and DFDIST are indicated in bold

MLR multiple linear regression, *PCNM* principle coordinates of neighbor matrices

[a, b] Loci demonstrating >95 % probability for balancing selection and diversifying selection, respectively; after controlling for FDR < 0.1, these outlier loci were no longer significant

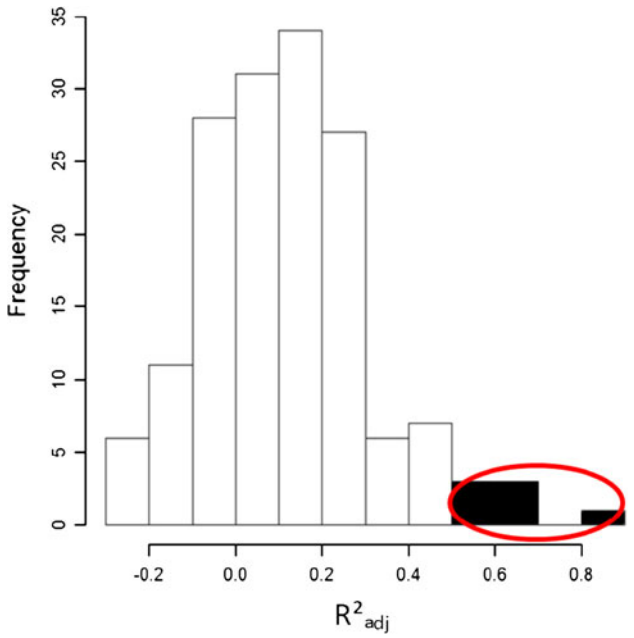

Fig. 2 Frequency of R^2_{adj} values from multiple linear regression of *Gentiana nivalis* AFLP allele frequencies at 157 loci associated with 23 environmental and nine broad-scale PCNM variables

less than 0.16. When the analysis was run with even prior odds for both models (1:1), meaning that every locus was as likely to be under selection as neutral, a single locus (ACT_CAC_103.1; Table 3) showed a posterior probability = 0.79 for divergent selection, indicating substantial evidence according to Jeffreys' scale (Jeffreys 1961). Although this result was no longer significant after controlling for false discovery rate, locus ACT_CAC_103.1 was also identified as potentially adaptive in the above landscape genetic approach.

Using DFDIST, the same locus (ACT_CAC_103.1), along with four others showed high probability (>95 %) for divergent selection among the three STRUCTURE groups (Table 3). Four of these five loci under divergent selection were also identified in the landscape genetic approach above (Table 3). This analysis also indicated four loci with a high probability (>95 %) for balancing selection (data not presented here). Again, after controlling for FDR, none of the outlier loci detected by DFDIST remained significant. Results for both BAYESCAN and DFDIST are presented in Fig. 3.

Discussion

This study applied the landscape genetic approach proposed by Manel et al. (2010b) to an AFLP dataset developed for the alpine plant *Gentiana nivalis* (Gugerli et al. 2008; Alvarez et al. 2009), identifying potential loci under selection by correlating allele frequencies with environmental data and broad-scale PCNM variables. Incorporation of PCNM variables allowed us to model the influence of spatially-structured environmental variation on this data set. We also applied powerful AICc model selection to determine the relationship between environmental variables and selected loci. From this analysis, we identified 15 loci of potential adaptive relevance, which were primarily correlated with temperature, solar radiation, and precipitation as well as with broad-scale PCNM variables. Four of these same loci were also identified as being under divergent selection using outlier locus detection analysis.

Incorporation of PCNM and genetic structure into a landscape genetic approach to detect potentially adaptive loci

PCNM analysis identified nine broad-scale PCNM variables, representing large-scale, unmeasured, spatio-environmental variation across our study area of the European Alps. In fact, the spatial variation quantified by the PCNM eigenvectors was the most important factor influencing allele frequency distribution in *G. nivalis* (Table 4). The significant correlation of PCNM variables with all loci identified as potentially adaptive in the landscape genetic approach suggests that additional unmeasured, broad-scale environmental factors are influencing potentially adaptive genetic variation in *G. nivalis*. To gain a better understanding of the factors driving adaptation, future research is needed to identify what environmental patterns the unmeasured variation represents. Yet, even with unlimited resources, the complexity of natural ecosystems makes it unlikely that researchers will be able to identify and measure all dimensions of environmental variation influencing adaptation of species. This emphasizes the value of the PCNM approach, as it enables the identification of loci significantly correlated to spatio-environmental gradients even if we have not measured the causative environmental factor. Including PCNM in a landscape genetic approach allows for a more realistic assessment of the proportion of the genome affected by selection and a better understanding of the specific dynamics of potentially adaptive variation across spatial scales.

To test for the potentially confounding effect of population genetic history on detecting loci under selection, we included a predictor variable for population membership to one of the four genetic clusters defined from STRUCTURE analysis. Alvarez et al. (2009, 2012) also found that *G. nivalis* clustered into four populations across the European Alps and Carpathians, and they suggest that these four well-defined clades likely correspond to four distinct glacial refugia. Although we found very strong and geographically distinct neutral genetic structure in this study

Table 4 Identification of the best AIC_C models and the relative importance of the environmental and PCNM variables acting as drivers of genetic variation in selected AFLP loci

Selected loci	srad	sum_prcp	sp_prcp	tmin	Slope	topo	Soil	Aspect	PCNM	Second best model	Δ_i
ACT_CAC_103.1	0.23	1.00, **0.65**	1.00						**1.00**	pcnms	1.25, **1.77**
										srad + sum_prcp + sp_prcp	
ATC_CAC_100.0			1.00, **1.00**		0.32,**0.32**					sum_prcp + sp_prcp + slope	1.47, **1.47**
										sum_prcp + sp_prcp + slope	
ATC_CAC_144.5				**0.70**					**1.00**	**pcnms**	**1.71**
ATC_CAC_151.7	**1.00**				**0.11**	**0.93**			**1.00**	**srad + topo + pcnms + slope**	**3.95**
ATC_CAC_221.4	0.28	1.00, **1.00**	0.71	1.00	**0.72**				**1.00**	sum_prcp + tmin + topo + soil	1.08, **1.88**
										srad + pcnms	
ATC_CAC_225.9		1.00, **1.00**	1.00, **1.00**			1.00	0.87, **0.23**		**1.00**	sum_prcp + sp_prcp + topo	3.89, **1.80**
										pcnms + soil	
ATC_CAC_300.0	1.00, **1.00**			0.66, **0.66**	1.00,**1.00**	**1.00**				srad + slope + topo	1.29, **1.29**
										srad + slope + topo	
ATC_CAC_302.0	1.00, **1.00**			1.00, **1.00**		**0.13**		0.14, **0.13**		srad + tmin + aspect	3.03, **3.03**
										srad + tmin + aspect	
ATC_CAC_362.7			**0.29**			**0.38**			**1.00**	**sp_prcp + pcnms**	**0.31**
ATG_CTG_71.3		1.00, **1.00**	1.00, **0.66**				0.75, **0.62**			srad + sum_prcp + sp_prcp	1.56, **1.56**
										srad + sum_prcp + sp_prcp	
ATG_CTG_113.4		**1.00**						**0.28**	**1.00**	**sum_prcp + pcnms + aspect**	**1.92**
ATG_CTG_199.5				**1.00**			**0.65**		**1.00**	**tmin + pcnms**	**1.26**
ATG_CTG_200.8				**1.00**			**0.48**	**0.33**	**1.00**	**tmin + pcnms**	**0.02**
ATG_CTG_236.7			**1.00**					**0.36**	**1.00**	**sp_prcp + pcnms + aspect**	**1.15**
ATG_CTG_489.1								**0.44**	**1.00**	**pcnms + aspect**	**0.51**
Sum of AIC_C weights	2.23, **3.28**	5.00, **3.65**	4.71, **3.95**	2.66, **4.36**	1.32,**2.15**	1.50, **2.44**	1.62, **1.98**	0.14, **1.54**	0.00, **11.00**		
Avg. of AIC_C weights	0.74, **0.82**	1.00, **0.91**	0.94, **0.79**	0.89, **0.87**	0.66,**0.54**	0.75, **0.61**	0.81, **0.50**	0.14, **0.31**	0.00, **1.00**		

Results of the non-spatial and spatial (includes nine broad-scale PCNM variables) analyses are denoted by plain and bold text, respectively. For each locus, the second best model and simple difference values (Δ_i) among top models are also presented. Relative importance of each variable across all loci is indicated by the average of AIC_C weights

PCNM principle coordinates of neighbor matrices, *sp_prcp* spring precipitation, *srad* radiation, *sum_prcp* summer precipitation, *tmin* minimum average temperature, *topo* topography

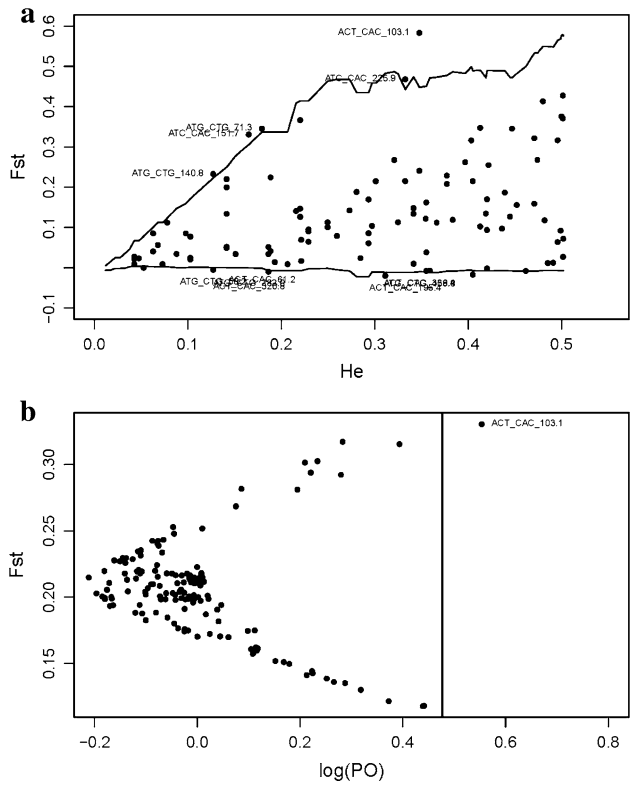

Fig. 3 Results from outlier locus detection methods. **a** Output from DFDIST shows the F_{ST} of each locus plotted against its heterozygosity. Lines illustrate the 0.025 and 0.975 quantiles of the neutral distribution simulated under the island model. **b** Output from BAYESCAN shows the F_{ST} of each locus plotted against the posterior odds for a model including selection (based on even prior odds 1:1). The vertical bar illustrates the threshold for substantial evidence of selection according to Jeffreys' scale

(Fig. 1), with 88 % of individuals showing >90 % inferred ancestry to one of the four populations, our regression analysis indicated that detection of potentially adaptive loci was robust to the influence of demographic history.

Redundancy analysis results similarly suggest that small-scale processes have little, if any, influence on potentially adaptive genetic variation at selected loci. Finding no effect of small-scale PCNM eigenvectors on genetic variation supports our decision to include only the first half (nine) of broad-scale PCNM eigenvectors in identifying loci correlated with environmental gradients.

Identifying environmental variables affecting potentially adaptive loci

Temperature, precipitation, and solar radiation appeared to be the environmental variables most strongly influencing (i.e. affecting the largest number of loci) allele frequencies at selected loci, after incorporating PCNM variables and thus modelling unmeasured spatial variation. Additionally, of the potentially adaptive loci, several demonstrated associations with second and third order polynomial transformations of environmental predictors, suggesting that it is important to consider non-linear relationships in studying the complexity of adaptive responses to environmental clines. In these respects, our results are in agreement with those recently reported for numerous other plant species also collected under the IntraBioDiv Consortium and analyzed using the landscape genetic approach presented in this paper (Manel et al. 2010b, 2012). Applying another approach, generalized estimating equations, Poncet et al. (2010) also found non-linear relationships to be important. These studies found temperature and precipitation to be the major environmental variables driving the distribution of allele frequencies at potentially adaptive loci in alpine species. Consistent with these findings, results from *G. nivalis* support the conclusion that specific environmental factors act as general selective forces on groups of species distributed across the same landscape. AIC_C analysis identified additional strong associations between several loci and other environmental variables in *G. nivalis*, including solar radiation, topography, and soil humidity. Thus, while some variables are consistently identified as strongly influencing alpine species across the European Alps, unique selective forces appear to be affecting potentially adaptive genetic variation in *G. nivalis* as compared to other alpine species, suggesting species-specific adaptive response to a common habitat.

Comparing a landscape genetic approach with outlier locus detection

In agreement with the conclusions of Manel et al. (2010b), our research indicates that using PCNM variables in allele distribution models has the capacity to identify more potentially adaptive loci than traditional outlier locus detection methods in natural populations. By directly comparing these two methods, we aimed to investigate whether the traditional outlier locus detection approach could be used to provide supporting evidence for loci identified as potentially adaptive via the recently developed landscape genetic approach. Using outlier detection methods, we found only weak evidence for a small number of selected loci, and results were not significant after correcting for multiple testing. Despite weak support, we did, however, find some consistency across methodologies, with both analyses identifying four of the same loci as candidates for divergent selection. Although these methods were based on very different assumptions, corroboration across methods provides compelling support for the hypothesis that these loci are of adaptive relevance and warrant further exploration. In addition to the four loci identified by both methods, the landscape genetic approach

was able to identify 11 more loci associated with environmental variation.

It is important to note that direct comparison of these methods can be misleading, and methodological issues do exist (Manel et al. 2009). First, outlier locus detection methods rely on a specific evolutionary and demographic hypothesis, while the landscape genetic approach makes a simple assumption that clinal gradients in environmental variables can result in gradual changes in allele frequencies (Schmidt et al. 2008). Second, outlier locus detection methods utilize allele frequency estimates from randomly sampled individuals within a priori genetically-defined populations and are thus not ideally suited for the analysis of samples continuously distributed across the landscape, such as would be the case for many widespread and common organisms (Manel et al. 2010b). Third, outlier detection methods are unable to directly incorporate spatial and environmental information. As a result, the inherent variation of natural systems adds to residual noise, making it less likely that significant genetic variation is detected. Indeed, these limitations were motivating factors in the development of landscape genetic approaches.

Conclusions and conservation perspectives

Identifying environmental heterogeneity associated with adaptive genetic variation is of great value in assisting the development of conservation plans for managing species' evolutionary potential. By including PCNM spatial variables, we showed that the landscape genetic approach presented in this paper was able to detect more than twice the number of potentially adaptive loci when compared with commonly used non-spatial methods for the same *G. nivalis* data set. Inclusion of PCNM variables enabled retrieval of residual variation that remained as model error in outlier locus detection methods, thereby allowing us to take advantage of information that was not utilized in the latter case. Additional benefits of the spatially-explicit landscape genetic approach include the ability to directly incorporate spatial and environmental information when identifying potentially adaptive loci as opposed to associating environmental patterns post hoc, as well as the ability to identify likely causal agents of selection. The ability to identify such relationships is increasingly valuable as organisms respond to accelerating changes in local and regional climatic conditions. More importantly than increasing the number of potentially adaptive loci identified, the landscape genetic approach also allowed us to determine potential causal agents of selection in associated environmental variation.

Identifying the ecological and evolutionary limits to species' ranges remains one of the largest gaps in our understanding of evolutionary biology (Eckert et al. 2008), restricting our capacity to provide for the long-term preservation of species (but see Jay et al. 2012). By identifying environmental factors associated with potentially adaptive genetic variation, this method allowed us to better understand what forces constrain the distribution of *G. nivalis* and determine conditions essential for its growth and survival. This approach can be used to generate hypotheses regarding functional mechanisms operating at potentially adaptive loci. Hypotheses can then be tested through a combination of (i) sequencing and (ii) common garden studies to determine if genetic variation at these loci is indeed adaptive. Genetic signatures of selection arise not only as a result of functional regions that are direct targets of selection, but also from genetic hitchhiking of neutral portions of the genome linked to genes under selection (Barton 2000). (i) Sequencing of identified loci and comparison with known candidate gene databases can help in determining if identified loci fall in known candidate gene regions and yield insight into potential homologous function across species. By identifying allele frequency changes along environmental gradients, researchers can locate plant genotypes that may have different physiological tolerances and are locally adapted to different bio-climatic conditions. (ii) To better understand how environmental conditions may limit an organism's range of tolerance, these genotypes can then be reciprocally transplanted in common gardens across environmental gradients. Concurrently implementing these methods may ultimately allow for greater power in determining the amount of genetic variation existing for adaptively important traits within populations and across a species' range.

Using the example of *G. nivalis*, we outline how results from the spatially-explicit landscape genetic approach can be directly applied in a conservation context to increase our understanding of how specific environmental factors limit species' ranges. Several of the loci identified as potentially adaptive in this study exhibited strong, directional clines in allele frequency with respect to environmental gradients. For example, locus ATC_CAC_144.5 showed an allele frequency of 0.7 under low summer precipitation (mean = 33 rain days) and decreased to a frequency of 0.05 under conditions of high summer precipitation (mean = 58 rain days) (Fig. 4a). Genetic variation at this locus could be linked to phenotypic variation in water use efficiency, drought tolerance, or other such mechanisms regulating growth and maintenance. As future precipitation patterns are expected to change, quantifying the adaptive genetic variation and associated physiological tolerances of different *G. nivalis* populations to various moisture regimes will aid managers in assessing climate change risk. Another locus, ACT_CAC_103.1, exhibited a strong cline in allele frequency with respect to latitude, with one allele present at

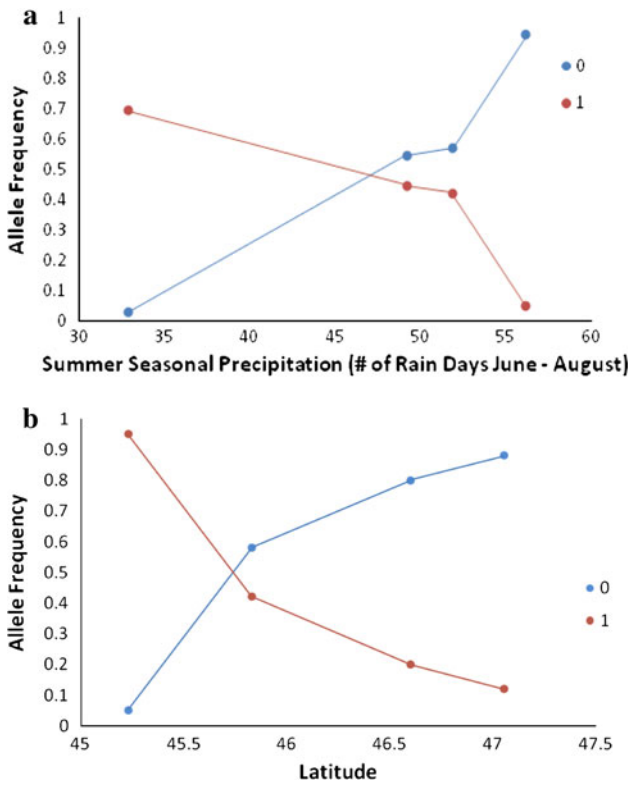

Fig. 4 **a** Locus ATC_CAC_144.5 shows an allele frequency for band presence of 0.7 under low summer precipitation (mean = 33 rain days) and decreases to a frequency of 0.05 under conditions of high summer precipitation (mean = 58 rain days). **b** Locus ACT_CAC_103.1 exhibits a strong clinal pattern in allele frequency with respect to a gradient in latitude. Band presence = 1; band absence = 0

a frequency of 0.05 in low latitudes and nearly fixed (frequency = 0.95) at high latitudes (Fig. 4b). Previous studies have found clinal genetic variation across latitudinal gradients to be associated with key phenological traits such as day-length-induced growth cessation and bud set (Ingvarsson et al. 2006). Future common garden studies could assist researchers in assessing the range of phenotypic variation for important phenological traits potentially associated with this locus in *G. nivalis*. By linking functional mechanisms to genetic variation at selected loci, we can gain a better understanding of different population's ranges of tolerance to environmental variation. The landscape genetic approach presented in this paper can therefore assist managers in determining key niche requirements necessary for species growth and survival as well as help managers identify genetically distinct forms of local adaptation harboured within different populations.

Inclusion of PCNM spatial predictor variables allowed us to identify more potentially adaptive loci, but how do we then determine what these unmeasured selective forces represent? The first PCNM eigenvector provides a good example. PCNM1 corresponds to the maximum spatial scale operating in this study, with each subsequent PCNM eigenvector representing increasingly smaller scales of spatial variation. In the upper right quadrant of the RDA ordination (Online Resource 1), PCNM1 is strongly correlated with three loci as well as with the quadratic polynomial of summer precipitation. PCNM1 is operating at a similarly broad spatial scale as summer precipitation, and the strong correlation between the three selected loci and PCNM1 suggests potential adaptation to a common, unidentified environmental factor. Sequencing and blasting these three loci in GenBank's genetic sequence database may provide insights into gene function and the corresponding selective force behind PCNM1 (Benson et al. 2005). From hypotheses of potential functional mechanisms, one can then apply PCNM eigenvector associations with clinal patterns in allele frequency in much the same way as for known environmental variables. For example, PCNM2 shows a strong, directional cline that looks very similar to the allele frequency clines mentioned above for summer precipitation and latitude. Common garden studies would then be employed to test physiological tolerances to hypothesized relationships.

Our exploration of PCNM variables as a means of controlling for unmeasured spatial variation when identifying potentially adaptive loci is the most comprehensive to date. With this study, we present a statistical framework that can be readily applied to available genetic data sets for any broadly distributed species and used to investigate the relationship between allele frequencies and site-specific environmental factors. Application of these methodologies will be especially useful for species experiencing dramatic environmental change. Identifying the environmental factors potentially driving patterns of adaptive genetic variation and interpreting these patterns across the landscape offers a significant contribution towards the conservation of evolutionary potential in non-model organisms and species of conservation concern.

Acknowledgments This research was made possible through the National Center for Ecological Analysis and Synthesis (NCEAS) at the University of California in Santa Barabara (Distributed Graduate Seminar on Landscape Genetics). We thank the IntraBioDiv Consortium for the use of their *G. nivalis* genetic data set. We are grateful to J. Bregy, D. Bühler, S. Dray, P. Legendre, the Cottonwood Ecology Group, and two anonymous reviewers for thoughtful discussions and comments on earlier versions of the manuscript. NA is funded by the Swiss National Science Foundation (Ambizione fellowship PZ00P3-126624). SM was supported by the Institut Universitaire de France.

References

Aeschimann D, Lauber K, Moser DM, Theurillat J (2004) Flora alpina, vol 2. Haupt, Berne

Allendorf FW, Hohenlohe PA, Luikart G (2010) Genomics and the future of conservation genetics. Nat Rev Genet 11:697–709

Alvarez N, Manel S, Schmitt T, IntraBioDiv Consortium (2012) Contrasting diffusion of quaternary gene pools across Europe: the case of the arctic-alpine *Gentiana nivalis* L. (Gentianaceae). Flora 207:408–413

Alvarez N, Thiel-Egenter C, Tribsch A et al (2009) History or ecology? Substrate type as a major driver of spatial genetic structure in Alpine plants. Ecol Lett 12:632–640

Antao T, Beaumont MA (2011) MCHEZA: a workbench to detect selection using dominant markers. Bioinformatics 27:1717–1718

Barton NH (2000) Genetic Hitchhiking. Philos Trans R Soc Lond B 355:1553–1562

Beaumont MA, Balding DJ (2004) Identifying adaptive genetic divergence among populations from genome scans. Mol Ecol 13:969–980

Beaumont MA, Nichols RA (1996) Evaluating loci for use in the genetic analysis of population structure. Proc Royal Soc London B 263:1619–1626

Bellier E, Monestiez P, Durbec J-P, Candau J-N (2007) Identifying spatial relationships at multiple scales: principal coordinates of neighbor matrices (PCNM) and geostatistical approaches. Ecography 30:385–399

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. J Royal Stat Soc B 57:289–300

Bensch S, Åkesson M (2005) Ten years of AFLP in ecology and evolution: why so few animals? Mol Ecol 14:2899–2914

Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Wheeler DL (2005) GenBank. Nucleic Acids Res. doi:10.1093/nar/gki063

Borcard D, Legendre P (2002) All-scale spatial analysis of ecological data by means of principal coordinates of neighbor matrices. Ecol Model 153:51–68

Borcard D, Legendre P, Drapeau P (1992) Partialling out the spatial component of ecological variation. Ecology 73:1045–1055

Borcard D, Gillet F, Legendre P (2011) Numerical ecology with R. Springer, New York

Burnham KP, Anderson DR (2002) Model selection and multimodel inference: a practical information-theoretic approach. Springer, New York

Burnham KP, Anderson DR (2004) Multimodel inference: understanding AIC and BIC in model selection. Sociol Method Res 33:261–304

Coop G, Witonsky D, Di Rienzo A, Pritchard JK (2010) Using environmental correlations to identify loci underlying local adaptation. Genetics 185:1411–1423

Crawford RMM (2008) Plants at the margin. Ecological limits and climate change. Cambridge University Press, Cambridge

Dray S, Legendre P, Peres-Neto PR (2006) Spatial modelling: a comprehensive framework for principal coordinate analysis of neighbour matrices (PCNM). Ecol Model 196:483–493

Eckert CG, Samis KE, Lougheed SC (2008) Genetic variation across species' geographical ranges: the central-marginal hypothesis and beyond. Mol Ecol 17:1170–1188

Endler JA (1977) Geographic variation, speciation and clines. Princeton University Press, Princeton

Excoffier L, Hofer T, Foll M (2009) Detecting loci under selection in a hierarchically structured population. Heredity 103:285–298

Falush D, Stephens M, Pritchard JK (2007) Inference of population structure using multilocus genotype data: dominant markers and null alleles. Mol Ecol Notes 7:574–578

Foll M, Gaggiotti O (2008) A genome scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. Genetics 180:977–993

Fournier-Level A, Korte A, Cooper MD, Nordborg M, Schmitt J, Wilczek AM (2011) A map of local adaptation in *Arabidopsis thaliana*. Science 334:86–89

Gibson N, Yates CJ, Dillon R (2010) Plant communities of the ironstone ranges of South Western Australia: hotspots for plant diversity and mineral deposits. Biodiv Conserv 19:3951–3962

Gugerli Z, Englisch T, Niklfeld H et al (2008) Relationships among levels of biodiversity and the relevance of intraspecific diversity in conservation – a project synopsis. Perspect Plant Ecol Evol Syst 10:259–281

Hansen MM, Olivieri I, Waller DM, Nielsen EE (2012) The GeM Working Group. Monitoring adaptive genetic responses to environmental change. Mol Ecol. doi: 10.1111/j.1365-294X.2011.05463.x

Hegi G (1957) Illustrierte flora von Mittel-Europa, vol 3. Lehmanns, München

Hess HE, Landolt E, Hirzel R (1972) Flora der Schweiz und angrenzender Gebiete, vol 3. Birkhäuser, Basel

Hirao AS, Kudo G (2004) Landscape genetics of alpine-snowbed plants: comparisons along geographic and snowmelt gradients. Heredity 93:290–298

Hoffmann AA, Willi Y (2008) Detecting genetic responses to environmental change. Nat Rev Genet 9:421–432

Holderegger R, Wagner HH (2008) Landscape genetics. Bioscience 58:199–207

Holderegger R, Hermann D, Poncet B et al (2008) Land ahead: using genome scans to identify molecular markers of adaptive relevance. Plant Ecol Divers 1:273–283

Holderegger R, Bühler D, Gugerli F, Manel S (2010) Landscape genetics of plants. Trends Plant Sci 15:675–683

Hultén E, Fries M (1986) Atlas of North European vascular plants: north of the Tropic of Cancer. Koeltz, Königstein

Ingvarsson PK, García MV, Hall D, Luquez V, Jansson S (2006) Clinal variation in *phyB2*, a candidate gene for day-length-induced growth cessation and bud set, across a latitudinal gradient in European aspen (*Populus tremula*). Genetics 172: 1845–1853

Jay F, Manel S, Alvarez N et al (2012) Forecasting changes in population genetic structure of alpine plants in response to global warming. Mol Ecol 21:2354–2368

Jeffreys H (1961) The theory of probability, 3rd edn. Oxford University Press, New York, p 432

Jombart T, Dray S, Dufour A-B (2009) Finding essential scales of spatial variation in ecological data: a multivariate approach. Ecography 32:161–168

Joost S, Bonin A, Bruford MW et al (2007) A spatial analysis method (SAM) to detect candidate loci for selection: towards a landscape genomics approach to adaptation. Mol Ecol 16:3955–3969

Körner C (2003) Alpine plant life: functional plant ecology of high mountain ecosystems. Springer, New York

Kozuharova E, Anchev M (2002) Floral biology, pollination ecology and breeding systems in *Gentiana verna*, *G. utriculosa* and *G. nivalis* (sect. Calatianae, Gentianaceae). God Sofiisk Univ St. Kliment Ohridski Biol Fak 2 Bot 92:57–71

Kozuharova E, Anchev ME (2006) Nastic corolla movements of nine *Gentiana* species (Gentianaceae), presented in the Bulgarian flora. Phytol Balcanica 12:255–265

Legendre P, Legendre L (1998) Numerical ecology. Elsevier, Amsterdam

Manel S, Conord C. Despres L (2009) Genome scan to assess the respective role of host - plant and environmental constraints on the adaptation of a widespread insect. BMC. Evol Biol 9:288. http://www.biomedcentral.com/1471-2148/9/288

Manel S, Gugerli F, Thuiller W, Alvarez N, Legendre P, Holderegger R, Gielly L, Taberlet P, IntraBioDiv Consortium (2012) Broad-scale adaptive genetic variation in alpine plants is driven by temperature and precipitation. Mol Ecol 21:3729–3738

Manel S, Segelbacher G (2009) Perspectives and challenges in landscape genetics. Mol Ecol 18:1821–1822

Manel S, Berthier P, Luikart G (2002) Detecting wildlife poaching: identifying the origin of individuals with Bayesian assignment tests and multilocus genotypes. Conserv Biol 3:650–659

Manel S, Joost S, Epperson B et al (2010a) Perspectives on the use of landscape genetics to detect genetic adaptive variation in the field. Mol Ecol 19:3760–3772

Manel S, Poncet NB, Legendre P, Gugerli F, Holderegger R (2010b) Common factors drive genetic variation of adaptive relevance at different spatial scales in *Arabis alpina*. Mol Ecol 19:3824–3835

Meudt HM, Clarke AC (2007) Almost forgotten or latest practice? AFLP applications, analyses and advances. Trends Plant Sci 12:106–117

Nielsen R (2005) Molecular signatures of natural selection. Annu Rev Genet 39:197–218

Ohtani K (2000) Bootstrapping R^2 and adjusted R^2 in regression analysis. Econ Model 17:473–483

Okuda T, Noda T, Yamamoto T, Hori M, Nakaoka M (2010) Contribution of environmental and spatial processes to rocky intertidal metacommunity structure. Acta Oecol 36:413–422

Parisod C, Joost S (2010) Divergent selection in trailing- versus leading-edge populations of *Biscutella laevigata*. Ann Bot 105:655–660

Peakall R, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol Ecol Notes 6:288–295

Poncet B, Herrmann D, Gugerli F et al (2010) Tracking genes of ecological relevance using a genome scan: application to *Arabis alpina*. Mol Ecol 19:2896–2907

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–959

R Development Core Team (2008/2009) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. (http://www.R-project.org.)

Schmidt PS, Serrao EA, Pearson GA et al (2008) Ecological genetics in the north Atlantic: environmental gradients and adaptation at specific loci. Ecology 89:S91–S107

Schoville S, Bonin A, Francois O, Lobreaux S, Melodelima C, Manel S (2012) Adaptive genetic variation on the landscape: methods and cases. Annu Rev Ecol Evol Syst 43:23–43

Storz JF (2005) Using genome scans of DNA polymorphism to infer adaptive population divergence. Mol Ecol 14:671–688

Taberlet P, Zimmermann NE, Englisch T et al (2012) Genetic diversity in widespread species is not congruent with species richness in alpine plant communities. Ecol Lett (in press)

Thuiller W (2007) Biodiversity: climate change and the ecologist. Nature 448:550–552

Wagner HH, Fortin M-J (2005) Spatial analysis of landscapes: concepts and statistics. Ecology 86:1975–1987

Walther GR, Post E, Convey P et al (2002) Ecological responses to recent climate change. Nature 416:389–395

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. Mol Ecol 15:1419–1439