Genome-wide association to fine-scale ecological heterogeneity within a continuous population of *Biscutella laevigata* (Brassicaceae)

Christian Parisod and Pascal-Antoine Christin

Department of Ecology and Evolution, Biophore, quartier Sorges, University of Lausanne, CH–1015 Lausanne, Switzerland

Author for correspondence:
Christian Parisod
Tel: +33 1 30 83 30 24
Fax: +33 1 30 83 30 99
Email: cparisod@versailles.inra.fr

Received: *19 September 2007*
Accepted: *29 November 2007*

Summary

• Gene flow, drift and selection can be detected through different signatures across the genome and the landscape. Genetic discontinuities along with their correlation to environmental features can be used to tease out isolation-by-distance and isolation-by-time from processes related to selection.

• Using spatial statistics (spatial autocorrelation methods, canonical correspondence analysis and partial Mantel tests) dealing with genome-wide amplified fragment length polymorphism (AFLP) under unlikely Hardy–Weinberg assumptions, this study investigates 124 individuals within a continuous population of the autopolyploid *Biscutella laevigata* (Brassicaceae).

• Fine-scale spatial genetic structure was strong and the mosaic-like distribution of AFLP genotypes was consistently associated with habitat factors, even when controlled for geographical distances. The use of multivariate analyses enabled separation of the factors responsible for the repartition of the genetic variance and revealed a composite effect of isolation by distance, phenological divergence and local adaptation to habitats characterised by different solar radiation regimes.

• These results suggest that the immigrant inviability barrier facilitated the maintenance of adapted subpopulations to distinct environmental conditions at the local scale.

Key words: amplified fragment length polymorphism (AFLP), autopolyploidy, between-group eigenanalysis, canonical correspondence analysis, landscape genomics, local adaptation, partial Mantel test, phenology.

New Phytologist (2008) **doi**: 10.1111/j.1469-8137.2007.02361.x

Introduction

The increased availability of highly variable genetic markers led to the development of population genomics, a discipline investigating the evolution of different regions across the genome in natural population (Luikart *et al.*, 2003). Meanwhile, improved accessibility of spatially accurate ecological data resulted in the emergence of landscape genetics, which maps genetic discontinuities among population with environmental features with aim of a better understanding of the factors influencing the distribution and evolution of natural populations (Manel *et al.*, 2003). These closely related fields are currently merging into landscape genomics, which simultaneously investigates numerous genetic markers across the genome together with numerous sites across the landscape (Joost *et al.*, 2007). However, despite guidelines and sophisticated tools, the links between genetic and environmental variables remain difficult to unambiguously quantify, because of the numerous disadvantages of conventional analytical tools (Sork *et al.*, 1999; Escudero *et al.*, 2003; Luikart *et al.*, 2003; Manel *et al.*, 2003; Storfer *et al.*, 2007).

Although differentiation of gene pools is easily envisioned in spatial isolation when a barrier to gene flow leads to the independent evolution of lineages and thereby genetic

discontinuities, divergence is not as straightforward within populations (Levin, 2000). In continuous populations, recurrent processes, such as gene flow, genetic drift and selection, act in concert to shape the genetic structure (Lenormand, 2002; Latta, 2003). Theoretically, gene flow is expected to homogenize the distribution of genetic variation, unless selection or drift is strong. Fine-scale genetic differentiation has often been reported in plant populations (Vekemans & Hardy, 2004), even under substantial gene flow, suggesting that strong selective pressure promotes local adaptation at small scale in heterogeneous landscapes (Linhart & Grant, 1996). Nevertheless, restricted gene dispersal also increases genetic differentiation among populations as a function of geographical distance (Slatkin, 1993) and several evolutionary processes can create genetic heterogeneity, leading to a within-population structure similar to the expectations under selection (Latta, 2003). Purely demographic processes can be distinguished from selection, because they are expected to have different consequences across the genome and across the landscape. Demographic processes such as migration or genetic drift are expected to similarly affect all loci across the genome, while selection is supposed to act upon simple loci or small portions of the genome. Drift and selection are also expected to leave different genetic signatures across the landscape. On one hand, drift generates a strong, but stochastic spatial structure, which is therefore not expected to match environmental conditions. On the other hand, a consistent association of particular genotypes with environmental conditions suggests that populations are locally adapted (Latta, 2003). It has also been recently stressed that differential gene exchange influenced by plastic variation in the flowering phenology among local habitats may contribute to a fine-scale genetic variation associated with environmental heterogeneity (Hendry & Day, 2005). Unlike local adaptation, it is anticipated that differences in phenology lead to genome-wide structure associated with environmental factors and it is crucial to explore procedures that may distinguish local adaptation from phenological divergence.

Population genomics tools without Hardy–Weinberg assumptions

From a methodological point of view, understanding populations as a community of loci varying across heterogeneous environments, it becomes attractive to adopt the statistical tools of community ecology in population genetics (Gram & Sork, 2001). Indeed, despite controversies over multivariate analyses in evolutionary genetics (Blows, 2007), genomic studies make great use of this statistical framework (Quackenbush, 2001) and the few population studies adopting it provided valuable insights about adaptation under natural conditions (Hamrick & Allard, 1972; Nevo *et al.*, 1988; Angers *et al.*, 1999; Gram & Sork, 2001; Manel *et al.*, 2003; Volis *et al.*, 2004). Amplified fragment length polymorphism markers (AFLPs) particularly suit the aims of population genomics and several statistical tools have been developed to deduce allele frequencies from dominant molecular markers in diploid organisms (Bonin *et al.*, 2007). However, Hardy–Weinberg assumptions on which these estimations rely are hardly verifiable and probably not met in several situations, in which no procedure seems yet available to deal with AFLPs. For example, autopolyploid lineages, which occur in many plant taxa (Soltis & Soltis, 2000; Soltis *et al.*, 2007), certainly violate Hardy–Weinberg expectations because they often present multisomic inheritance at certain loci, resulting in segregation complexities (Bever & Felber, 1992; Ronfort *et al.*, 1998). In order to investigate the evolutionary potential of all these insightful taxa, alternative statistical tools are needed.

Principal component analysis (PCA) aims at extracting major gradients of variation in data and the related between-group eigenanalysis (BPCA) ordinates specified groups as to maximize their separation in some space, allowing testing for the significance of the between-group variance (Doledec & Chessel, 1987). When used in population genetics, PCA and BPCA are band-based approaches that are independent of any theoretical genetic models and conveniently allow investigation of genetic structure (Patterson *et al.*, 2006), especially when Hardy-Weinberg equilibrium assumptions are a priori unexpected. Furthermore, when quantifying the effect of environmental heterogeneity on genetic structure and/or detecting candidate loci, spatial coincidence analyses are particularly suitable because they allow circumventing assumptions of neutrality. In that context, canonical correspondence analysis (CCA) suits the aims of population geneticists by attributing dependent variables (e.g. genetic data) to explanatory variables (e.g. ecological factors) and testing for their association. Canonical correspondence analysis integrates ordination and multiple regressions, taking all the AFLP loci and the environmental factors independently into account in a single analysis. It may represent an efficient method to explore the association between genetic differentiation and environmental heterogeneity (i.e. patterns of local adaptation; Manel *et al.*, 2003; Storfer *et al.*, 2007). Nevertheless, investigating the association between spatial patterns of genetic diversity and environmental heterogeneity is complicated by the presence of similar spatial autocorrelation in both independent datasets, which may artificially inflate their correlation (Volis *et al.*, 2004). Such a situation is particularly likely in continuous population at the local scale and inferences from spatial coincidence analyses have to be completed by a quantification of purely geographic effects, which may be achieved by partial Mantel tests comparing distances among multiple datasets (Riginos & Nachman, 2001; Bekkevold *et al.*, 2005).

Biscutella laevigata as a model species that does not follow Hardy-Weinberg expectations

Biscutella laevigata L. (Brassicaceae) is a spring blooming, long-lived perennial species with a sporophytic self-incompatibility

system (Olowokudejo & Heywood, 1984). Pollen dispersal is achieved by generalist Diptera and Lepidoptera, while seeds are passively dispersed by gravity and wind. Since the early studies of Manton (1937), *B. laevigata* is thought to have survived the Pleistocene vicissitudes as a diploid ($2n = 2x = 18$) in ice-free parts of the European continent and recolonized the Alps as an autotetraploid ($2n = 4x = 36$) after the ice ages. Using allozyme markers, Tremetsberger *et al.* (2002) showed multisomic segregation of alleles in the autopolyploids and recent plastid DNA phylogeography indicated that the species evolved through polytopic autopolyploidy before recolonizing previously glaciated areas with independent lineages (Parisod & Besnard, 2007). Whereas many studies focus on allopolyploidy (i.e. the merging and doubling of differentiated genomes), the evolutionary advantages of autopolyploid lineages, which do not combine the attributes of divergent parents, are neither straightforward nor well understood (Soltis & Soltis, 2000; Comai, 2005; Soltis *et al.*, 2007). Studies investigating multisomic taxa, such as *B. laevigata*, may thus provide valuable insights about the role of genome doubling under natural conditions.

Using a systematic sampling strategy in a continuous population of *Biscutella laevigata* (Brassicaceae) and AFLP on 124 individuals, this study aims at separating isolation-by-distance and isolation-by-time processes from those influenced by selection. Using appropriate statistical procedures, an association between the genome-wide AFLPs and the environmental variance was detected, suggesting restricted gene flow among habitats in relation to local adaptation and phenology.

Material and Methods

Sampling site and sampling strategy

The present study population (6°58′/46°26′; JAM in Parisod & Besnard, 2007) is located in an area with high level of plastid DNA haplotype diversity (Parisod & Besnard, 2007) and lies in putative refugia of the external area of the western Swiss Alps (Stehlik, 2000). It is therefore considered as a rear-edge population of one autopolypoid lineage. The altitude ranges between 1850 m and 2000 m, which represents the lower limit of the alpine belt. The bottom of the zone corresponds to subalpine habitats, while the plots located in higher elevation are in alpine conditions. Along this natural ecotone, contrasted habitats have a mosaic distribution according to the spatial repartition of ecological factors related to microtopography (C. Parisod *et al.*, unpublished).

Biscutella laevigata appeared in a linear, continuous population system over the drier areas of the study site. Following a systematic strategy, the species was censused in 2 × 2 m plots every 12.5 m and, when at least six individuals were present, leaves of each individual located nearest to each corner (total of four individuals) were collected and dried in silica gel (Fig. 1). Using this strategy, the resulting sampling is representative of the species natural distribution. The plot location was measured 100 times using a Trimble Geo-explorer 3 GPS (Sunnyvale, CA, USA) providing a precision of 1-m. *Biscutella laevigata* was present in 31 sampling plots, giving a dataset of 124 individuals for genetic analyses. The phenological stage of individuals was reported and coded as 1 for buds, 2 for buds and opened flowers, 3 for opened flowers, 4 for opened flowers and fruits, 5 for mature fruits and 6 for mature and dispersed fruits. Environmental values corresponding to the location of every plots were extracted from published GIS eco-climatic layers (Zimmermann & Kienast, 1999) and used as explanatory variables to explore their contribution to the phenology, but also their association with genetic markers (see later). Based on ecological modelling at the local and the regional scale (S. Joost and C. Parisod, unpublished), a few relevant environmental factors, which showed wide amplitude within the species habitat, were investigated. The following factors (presented with the range of values across the sampled plots and units between brackets) were taken into account (see the Supplementary material, Table S1): elevation (1851–1990 m), degree-days during the growing season (1124–1260 degree × days), slope (2–69°) and total solar radiation (18 057–75 418 kJ d^{-1}).

Genetic analysis

Total DNA was extracted from dried leaf material using the FastDNA kit (Q-Biogen) in a room devoted to DNA extraction. The manufacturer instructions were followed except that extracted DNA was washed twice. DNA quality and concentration were checked on agarose gels.

Dominant AFLP markers were generated for the whole population (124 profiles) following Vos *et al.* (1995) with minor modifications. Recommendations of Bonin *et al.* (2004) were followed in order to reduce genotyping errors. Restrictions and ligations were carried in a Biometra thermocycler (Gottingen, Germany) at 37°C for 2 h each. Restrictions were performed on 150 ng of genomic DNA in 20-µl reactions of 10 × buffer for *Eco*RI with 2 U *Mse*I, 5 U *Eco*RI and 0.3 µg bovine serum albumin (BSA). Then, 0.7 µM double-stranded *Eco*RI and *Mse*I adapters were ligated in further 20 µl of 10 × T4 buffer with 1 Weiss U of T4 DNA ligase giving a 40 µl ligation mix. Two microlitres of ligation solution were used for a nonselective amplification in 20 µl by 28 preselective cycles (120 s at 94°C, 28 cycles for 45 s at 94°C, 45 s at 56°C and 60 s at 72°C and a final elongation of 600 s at 72°C). Three microlitres of 20-times diluted preamplification solution were then used for a selective amplification by 36 cycles of a touch-down polymerase chain reaction (PCR) (60 s at 94°C, 13 cycles of 30 s at 94°C, 30 s at 65°C to 56°C and 60 s at 72°C followed by 23 cycles of 30 s at 94°C, 30 s at 56°C, 60 s at 72°C with a final elongation of 300 s at 72°C). Based on a preliminary assessment, where eight selective primer combinations were

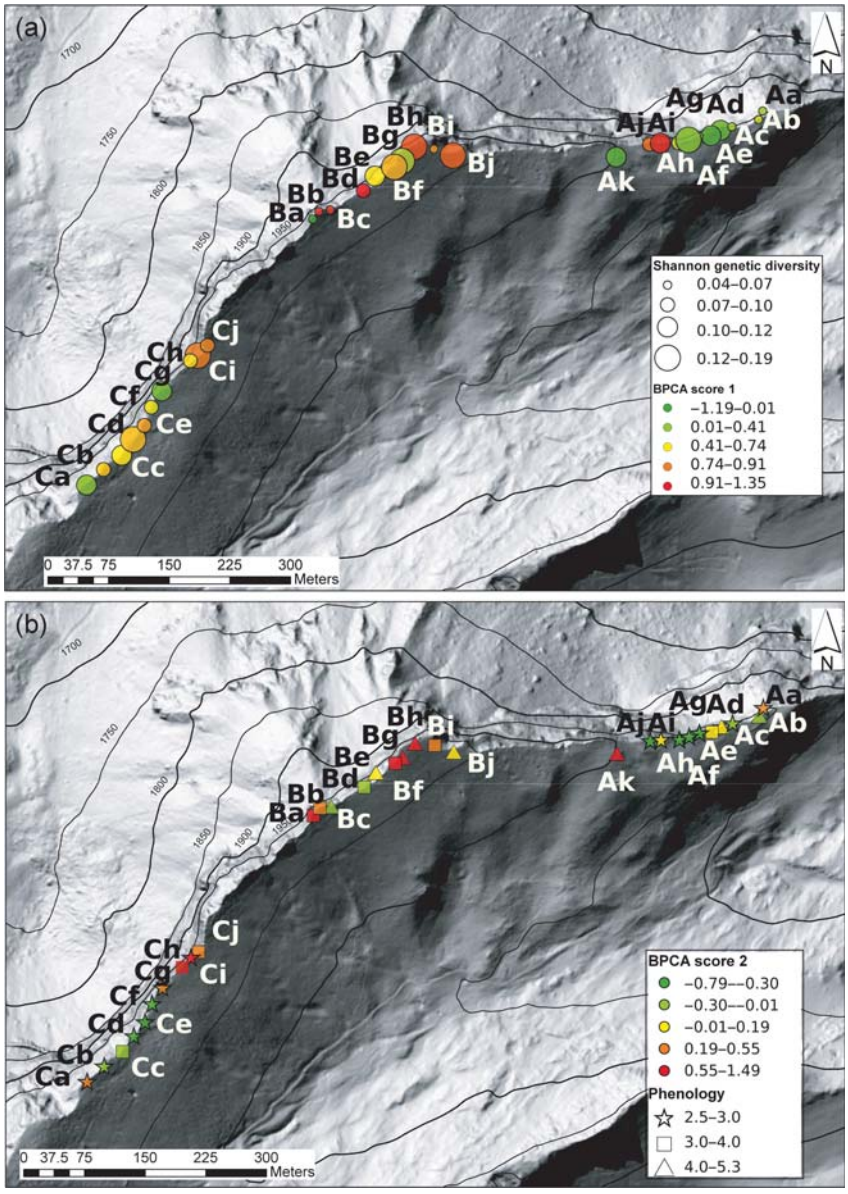

Fig. 1 Distribution of the sampled plots located within the continuous population of *Biscutella laevigata* and spatial representation of the genetic structure. Basemap presents the topography of the site with conventional northwest lighting and isoelevation lines every 50 m. (a) Circles represent the plots sampled, and colour indicates their quantiled score along the between-group eigenanalysis (BPCA) first axis and whose diameter is correlated with the genetic diversity, according to the panel on the right. (b) Plots are represented by star, square or triangle indicating the mean phenological stage and coloured according to their quantiled score along the second BPCA axis, according to the panel on the right.

tested three times independently to check for polymorphisms and reproducibility, the complete sample set was processed using two selective primer combinations: M-CAG/E-ACA and M-CTC/E-AGG with the fluorescent E primer labelled by FAM and JOE, respectively. PCR-products were visualized using an ABI-PRISM 377 (Applied Biosystems, Foster City, CA, USA) sequencer on a 6% Long Ranger denaturing gel for 5 h (2000 V, 50 A). Fragment sizes were estimated using the 500-ROX standard and scored manually as dominant markers using Genescan 3.1.2 (Applied Biosystems). Only clearly identifiable bands with high peaks were scored as present or absent and weak profiles were discarded. According to Vekemans *et al.* (2002), AFLP fragments inferior to 50 bp, which are reputed to be highly homoplasic, were not taken into account. Since no significant correlation between band size and frequency was detected ($P = 0.237$), homoplasy was considered negligible here. The whole AFLP procedure was replicated twice for 20 individuals and an error rate was calculated as the number of band differences between the two profiles of the same individual divided by the total number of fragments scored. The estimated error rate was 1.8%, which is lower than commonly reported with similar techniques (Bonin *et al.*, 2004).

Genetic structure with AFLPs

At the individual level, the pattern of spatial genetic structure was explored by spatial autocorrelation methods (Vekemans & Hardy, 2004) on the 124 individuals over the continuous population. The association between pairwise relationship coefficients calculated on AFLP markers (Hardy, 2003) in

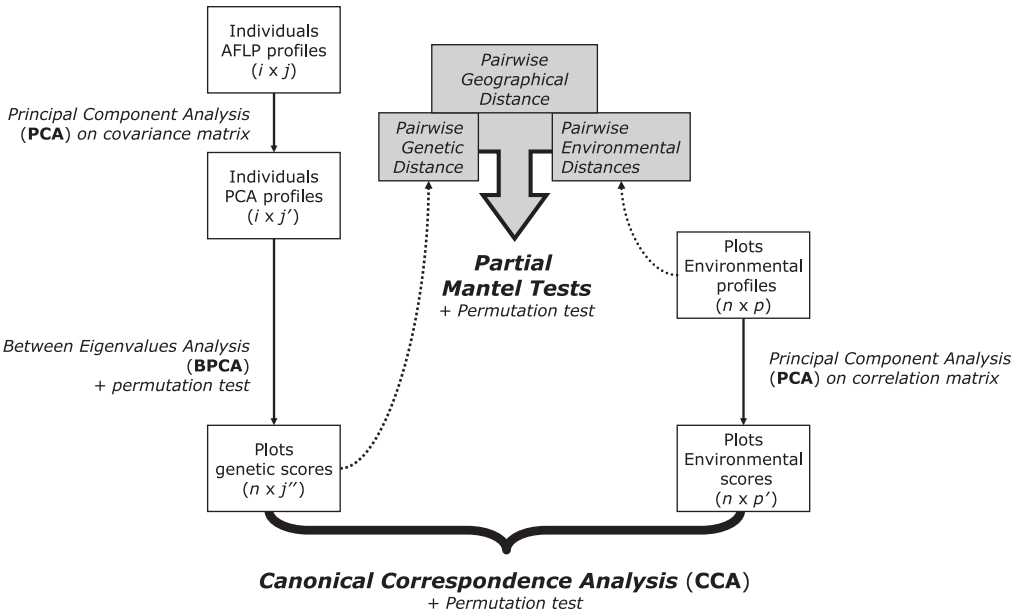

Fig. 2 Conceptual diagram of the multivariate statistical procedure used to investigate the association of individual genetic amplified fragment length polymorphism (AFLP) profiles and environmental factors. Variance partitioning by between-group eigenanalyses (BPCA) is analogous to F_{ST} and can be used to summarize most of the individual genetic information in a subspace at the plot level. Similarly, most of the environmental information is combined into an environmental subspace at the plot level. *i* and *n* represent the number of individuals and plots, respectively. *j* and *p* are the number of AFLP loci and environmental factors, respectively. After multivariate procedures (presented along the arrows), the column of the matrix contains independent information (scores, *j'* or *j''* and *p'*) that can be further investigated by univariate and/or spatial analyses. The canonical correspondence analysis (CCA, at the bottom) provides the framework to regress the multivariate environmental subspace on the multivariate genetic subspace and test for their association by permutations. Euclidian distances between plots multidimensional scores (*j''* and *p'*) can be computed in each subspace (dashed arrows) and used to investigate the influence of the geographical distance between plots with partial Mantel tests (at the top).

SPAGEDI (Hardy & Vekemans, 2002) and the logarithm of the geographic distance was tested by Mantel tests with 9999 permutations, using FSTAT (Goudet, 1995). An autocorrelogram was then constructed to assess correlations among genotypes at increasing geographical distance intervals with balanced number of individuals. Significance of distance-class mean relationship coefficients was assessed with 999 permutations and Bonferroni corrected (Vekemans & Hardy, 2004).

Given the lack of knowledge concerning the segregation of dominant AFLP markers in the autopolyploid *B. laevigata* (i.e. disomic vs multisomic inheritance, rate of double reduction), the genetic structure of the continuous population was investigated by a multivariate analysis (Fig. 2). Since it is a band-based rather than allele frequency-based approach, this procedure does not assume Hardy–Weinberg equilibrium and is therefore convenient to explore the genetic structure of polyploid populations with multisomic inheritance. Using ADE-4 (Thioulouse *et al.*, 1997), PCA on the covariance matrix was computed on AFLP profiles at the individual level and then at the plot level with BPCA in order to subsequently correlate genetic data with environmental factors. Principal component analysis conserves Euclidian distances and decomposes the covariance of all descriptors (here, loci) into components for each object (here, individuals or plots) along each of the full-ranked eigenvectors derived from general singular value decomposition (Doledec & Chessel, 1987; Patterson *et al.*, 2006). Principal component analysis thus summarizes a maximum of variance into fewer, interpretable dimensions. Thereafter, BPCA (i.e. PCA between plots based on PCA among individuals) was performed. This analysis groups individual PCA profiles into sampling plots in order to maximize the between-group genetic variance. The plot-centroids are then projected in a new reduced space, along the full-ranked BPCA eigenvectors. Significance of the between-group variance was estimated by 9999 permutations using ADE-4 (Thioulouse *et al.*, 1997). The mathematical details of the BPCA can be found elsewhere (Culhane *et al.*, 2002; Pavoine *et al.*, 2004), but variance partitioning by BPCA is an Euclidian discriminant approach and can be safely used with any combination of plots and loci. One interesting use of BPCA is to produce a set of univariate genotypic variables for each plot (BPCA scores) that can be further analysed by univariate and/or spatial statistics in order to help the interpretation of CCA (see later). Since BPCA is based on Euclidian distances, it can be considered as analogous to *F*-statistics (Parisod *et al.*, 2005). However, those estimators, named β_{ST}, are not equivalent to *F*-statistics and β_{ST} values may be overestimated because BPCA maximizes the between-group

variance. Nevertheless, relative β_{ST} represented the genetic structure of the population well. Indeed, pairwise β_{ST}, which were computed here as multidimensional Euclidian distance between the multidimensional BPCA scores of plot centroids, were slightly inflated but highly correlated with other traditional estimators of genetic differentiation (Supplementary material, Fig. S1), such as the G_{ST} calculated on Shannon diversity (Mantel test, $r = 0.81$, P-value < 0.001; Bussell, 1999) and the Φ_{ST} obtained from the AMOVA implemented in the ARLEQUIN software (Mantel test, $r = 0.70$, P-value < 0.001; Excoffier *et al.*, 1992).

Grouping individuals into sampling plots, the Shannon diversity index was calculated following Bussell (1999) for each AFLP locus and averaged in each plot, as

$$H = -\frac{1}{n}\sum p_i \cdot \log_2(p_i)$$

The relationships between the Shannon diversity index, plot size and density, as well as ecological factors were explored by stepwise multiple linear regressions and robust regressions for each factor.

Association between the genetic and environmental heterogeneity

The ecogenetic structure of the population was explored by CCA and partial Mantel tests (Fig. 2). Both procedures were computed in parallel because CCA cannot properly deal with geographical distance (but see Borcard *et al.*, 1992) and has been claimed to possibly raise misleading positive correlations because of shared spatial autocorrelation between genetic and ecological traits (Volis *et al.*, 2004). Conversely, partial Mantel tests account for geography but are limited to distance type of relationship among descriptors.

Canonical correspondence analysis explored the association between the AFLP data summarized by the BPCA at the plot level and the environmental factors summarized by PCA on a correlation matrix (Angers *et al.*, 1999; Gram & Sork, 2001; Volis *et al.*, 2004; Blows, 2007). Canonical correspondence analysis integrates ordination and multiple regression, resulting here in ordination axes ordered according to the genetic variance explained by linear combination of the independent ecological variables (Angers *et al.*, 1999). The CCA was performed with ADE-4 (Thioulouse *et al.*, 1997). Global significance of the association was tested by 9999 Monte Carlo permutations, which is permissive concerning the data distribution. In a single analysis, CCA produced a reduced space that maximized the joint structure between the genetic data (i.e. BPCA matrix derived from the individual AFLP bands occurrences and maximizing the genetic structure between plots) and the environmental data (PCA on correlation matrix of ecological factors). Therefore, CCA summarized the pattern of genetic variation that was best explained by the environment within the population (Gram & Sork, 2001) and highlighted the contribution of single ecological factor as well as AFLP loci to the CCA axes by producing an ordination diagram showing sampling plots as well as vectors presenting their AFLP composition and ecological factors acting on plots.

Patterns of genetic isolation-by-distance and spatial autocorrelation with ecological factors were investigated by Mantel tests with 9999 permutations, using FSTAT (Goudet, 1995). The correlation between pairwise β_{ST} was first tested separately with pairwise Euclidian distances among plots for each ecological factor and for the mean phenological stage and the logarithm of geographic distances among plots. All the environmental factors and the logarithmic geographic distance were then added as explanatory variables in a partial Mantel test. The best model was determined through a stepwise procedure, with the less significant explanatory variable removed until all the remaining factors were significant. This was repeated several times with a randomisation of the explanatory factors order to verify that the results were consistent. The same procedure was applied to determine the best models for the pairwise distances on the first and second axis of the BPCA.

Results

Population AFLP structure

One-hundred and two reliable AFLP bands were generated. A Mantel test between pairwise relationship coefficients and logarithm of the geographic distances among individuals revealed significant isolation by distance ($r = 0.097$, $P < 0.001$) and spatial autocorrelogram further showed fine-scale genetic structure (Fig. 3). It showed that geographically nearby individuals were genetically related as only the first three distance classes presented a mean relationship coefficient significantly greater than zero (0.234 between 1 and 21.3 m; 0.047 between 21.3 and 41.1 m and 0.036 between 41.1 and 60.5, respectively). The first significant distance-class with a negative relationship coefficient (i.e. trough) was composed of individuals located *c.* 200 m apart and the relationship coefficient reached zero below 100 m. Congruently, individuals located further than 400 m were all negatively related. Although predominantly negative relationship coefficients were significant after conservative Bonferroni correction, consecutive peaks and troughs of similar amplitude succeeded on the correlogram with a period of *c.* 100 m (Fig. 3).

The first axis of the PCA on individual AFLP profiles summarized only 6.71% of the total inertia and the 40 first PCA axes, which together represented 89.79% of the information contained in the genetic dataset, were kept for further analyses. The BPCA showed that most of the genetic variation was between plots ($\beta_{ST} = 0.485$, $P < 0.001$). The two first ordination axes showed roughly even contribution of most AFLP loci and summarized 10.4% and 9.1%, respectively (Fig. 4). Amplified fragment length polymorphism loci

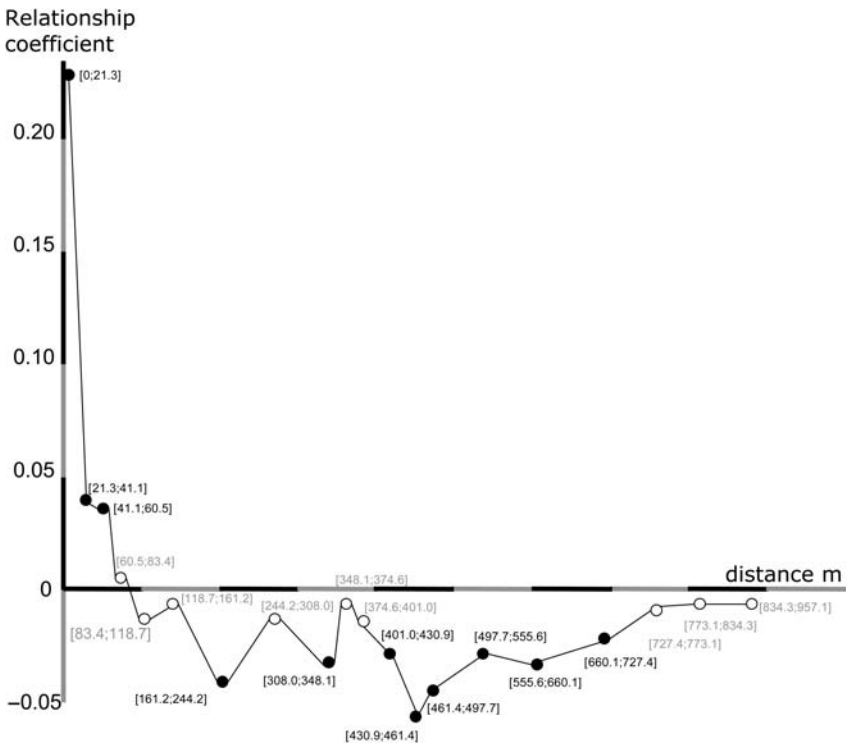

Fig. 3 Fine-scale spatial genetic structure in the continuous population of *Biscutella laevigata* located at the rear edge of the species range. Autocorrelogram presenting the mean coefficient of relationship within 20 balanced log-distance classes. The distance scale is linear and the intervals of distance classes are presented between brackets. Closed circles represent significant relationship coefficient for the distance class after Bonferroni correction; open circles represent nonsignificant values. Similarly, intervals of significant distance classes are presented in black and non significant in grey.

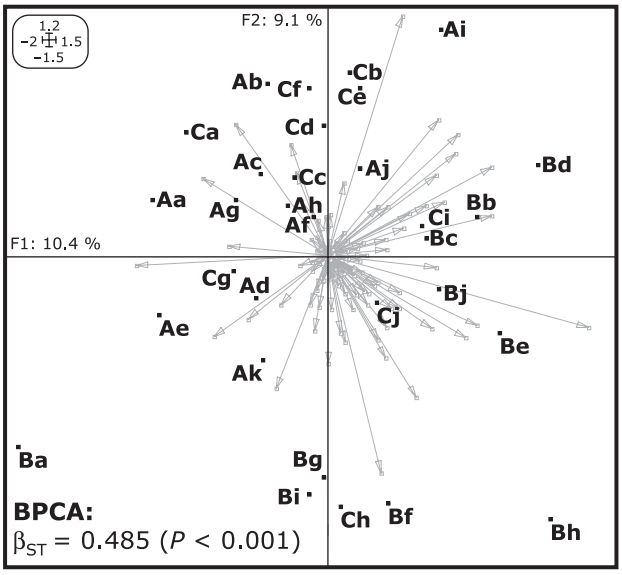

Fig. 4 Representation of the genetic structure among plots based on the between-group eigenanalysis (BPCA) of the individual amplified fragment length polymorphism (AFLP) profiles in a continuous population of *Biscutella laevigata*. The contribution of AFLP loci to the two first axes of the BPCA (F1 and F2, summarizing 10.4% and 9.1% of the genetic information, respectively) is presented as grey arrows. The scale of this ordination space is shown in the panel at the upper left. BPCA shows that 48.5% of the genetic variance is between plots ($P < 0.001$). See Fig. 1 for a spatial representation.

showed a broadly even contribution to the first BPCA axes, without strict covariation (i.e. most arrows had similar length and were distributed around 360°), and no particular loci appeared as mainly responsible for the observed structure. The ordination of plot centroids did not show obvious clustering and, although plots from field-patches A and B were genetically differentiated from one another, those belonging to the patch C were related to both (Fig. 4).

The CCA indicated that a great part of the genetic structure among plots was correlated with the structure of the environmental factors (25.1%, $P < 0.001$). The two first CCA axes summarized 27.9% and 25.2% of the inertia contained in the multiple regression of the environmental factors on the genetic structure, respectively. In contrast to BPCA, plots belonging to a different field-patch clustered with the CCA and plots of patch A especially were sorted from those of patches B and C (Fig. 5a). Although a few AFLP loci showed a great contribution to the first CCA axes (Fig. 5b), none is likely to be mainly responsible for the observed association between genetic and environmental heterogeneity. Total radiation and degree-days during growing season were shown to be responsible for most of the ecological differentiation among plots across the population (Fig. 5c).

The mean Shannon diversity index within the 31 plots was 0.101 (± 0.033). Within-plot genetic diversity did not significantly correlate with the environmental factors tested (stepwise regression, see the Supplementary Material, Table S1): neither abiotic factors (elevation, $P = 0.251$; degree-days during the growing season, $P = 0.276$; slope,

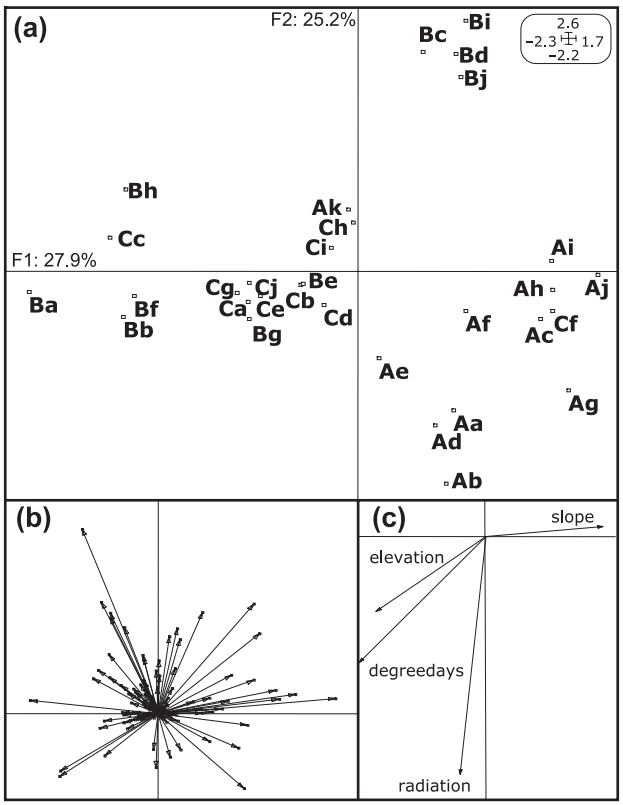

Fig. 5 Representation of the canonical correspondence analysis (CCA) showing the association between the environmental factors and the genetic structure among plots in a continuous population of *Biscutella laevigata*. (a) CCA ordination maximizing the genetic structure based on a model of environmental variance ($P < 0.001$). The two first axes show 27.9% and 25.2% of the common variance between the genetic and ecological datasets, respectively. The scale of this ordination space is shown in the panel at the above right. (b,c) Contribution of the amplified fragment length polymorphism (AFLP) loci and the environmental factors to the common reduced space presented in (a) respectively.

$P = 0.832$; and total solar radiation, $P = 0.077$), nor plot density ($P = 0.665$) or size ($P = 0.472$).

The influence of each environmental factor on the genetic distance was investigated when controlled for geographical distance by partial Mantel tests and indicated that, once Bonferroni corrected, only radiation distances had a significant influence on the genetic distance between plots (Table 1). Interestingly, phenology had none. Although a few ecological factors were significantly regressed on the first BPCA axis (data not shown), partial Mantel test showed that only radiation was significantly correlated with the first BPCA axis once the geographical distance is taken into account (Table 1). The mean phenological stage of plots, which was not spatially autocorrelated (Mantel test; $P = 0.125$) and not related to environmental factors (number of degree-days, $P = 0.535$; total radiation, $P = 0.148$), did not associate to the score along the first axis ($P = 0.904$), but correlate to the score along the second axis of the BPCA ($R^2 = 0.172$, $P = 0.023$, Fig. 1b; see the Supplementary Material, Table S1). This correlation cannot be attributed to isolation by distance since it remained significant even once the geographical distance is taken into account (Table 1).

Discussion

Genetic structure within the continuous population

The investigated population is nearly continuous and formed of hundreds of individuals located < 800 m apart (Fig. 1), so that panmixia through gene dispersal could be a priori expected. However, isolation by distance was detected and only the relationship coefficients among nearby individuals (i.e. located < 50 m apart) were significantly positive (Fig. 3), pointing to fine-scale genetic structure owing to restricted gene flow. Although a precise quantification of isolation by distance and neighbourhood size in autotetraploid individuals is not achievable using dominant markers (Vekemans & Hardy, 2004), the correlogram showed a declining wave-like pattern, indicating that individuals at 200 m are genetically consistently differentiated and suggesting that the neighbourhood is between 50 m and 100 m. Periodicity of the interindividual relationship coefficient along geographical distance pointed towards a mosaic distribution of AFLP genotypes. Congruently, high genetic structure was observed with the BPCA ($\beta_{ST} = 0.485$, $P < 0.001$). Since no particular AFLP loci disproportionably contributed to the reduced space (Fig. 4), this ordination suggests a genome-wide structure. Such genetic differentiation could be explained by genetic drift alone, but the genetic architecture of the population pointed to other evolutionary factors as well. Indeed, genetic discontinuities did not map to any obvious landscape feature that could have hindered recurrent gene flow (Fig. 1). Furthermore, plots did not cluster into obvious higher-level geographical units and field patches that were distinguished from the sampling strategy presented low genetic structure (data not shown), indicating that isolation by distance was not entirely responsible for the genetic structure at the plot level.

The distribution of genotypes across the continuous population of *B. laevigata* strongly suggests that the environment played a major role in shaping the genetic structure along this natural ecotone between subalpine and alpine habitats. Canonical correspondence analysis showed that a great part of the genetic structure was significantly associated with the environmental structure (Fig. 5). Including ecological factors into the analysis of population structure apparently increased the organization of the genetic data in relation to the landscape and, compared with the ordination based on genetic similarity alone (BPCA, Fig. 4), the ecogenetic ordination (CCA, Fig. 5) showed a simplified structure with conspicuous clustering of plots into field patches. However, notable exceptions to this clear-cut pattern (i.e. Cf clustered with Ac and Ah, Ak

Table 1 Mantel tests for the association between genetic, geographic and environmental distances among plots in the studied population of *Biscutella laevigata*

	β_{ST}	BPCA, first axis	BPCA, second axis
R-squared of best model	0.056	0.018	0.083
Log_e (geographic distance)	0.0962 (*P* = 0.039)	0.0810 (*P* = 0.082)	0.1181 (*P* = 0.012)
Degree day distance	–0.0369 (*P* = 0.428)	0.0756 (*P* = 0.101)	–0.0814 (*P* = 0.080)
Radiation distance	**0.2366*** **(*P* 0.001)**	**0.1338*** **(*P* 0.004)**	**0.1503*** **(*P* 0.002)**
Slope distance	0.0055 (*P* = 0.909)	0.0308 (*P* = 0.512)	**0.1589*** **(*P* 0.001)**
Elevation distance	–0.0394 (*P* = 0.396)	0.0595 (*P* = 0.200)	–0.0665 (*P* = 0.152)
Phenology distance	0.0289 (*P* = 0.541)	0.0344 (*P* = 0.454)	**0.2333*** **(*P* 0.001)**

Genetic distances have been computed for amplified fragment length polymorphisms (AFLPs) from between-group eigenanalysis (BPCA), as β_{ST} for all axes, first axis and second axis independently. Slopes of the explanatory factors that were significant after Bonferroni sequential correction are presented in bold type. The asterisks mark factors that participate to the best model, whose *R*-squared is given.

clustered with Ch and Ci, Cc clustered with Bf and Bh) further suggested that environmental rather than geographical conditions were responsible for the genetic similarity among plots (Fig. 5). The distribution of environmental factors matched that of the genotypes, suggesting restricted gene flow among habitats and a repartition of the continuous population into subpopulations adapted to different micro-environmental conditions. Partial Mantel test showed that the pairwise differences in total radiation between plots significantly explained the genetic structure (Table 1), whereas the effect of the logarithmic distance was not significant when the radiation factor was taken into account. The consistent association between genetic and environmental heterogeneity therefore indicates that evolutionary processes other than drift have shaped the structure (Lenormand, 2002) and reduced gene flow between habitats in the continuous population of *B. laevigata*.

Genome-wide association between markers and ecological factors

Since AFLPs are reputed to be mainly neutral markers (Bensch & Akesson, 2005), the observed structure is not prone to be mainly shaped by direct selection on all AFLP loci. The genome-wide association between genetic and environmental differentiations could result either from habitat-directed selection acting on several loci across the genome resulting in a genome-wide 'hitch-hiking' phenomenon or from temporally restricted gene flow because of contrasting phenology of individuals living in different habitats (Fox, 2003). The CCA provides independent information along its orthogonal axes (Gram & Sork, 2001) and, even if this property is increasing the difficulty of interpretation, it may be useful to disentangle multivariate evolutionary processes (Houle, 2007; Walsh, 2007). Here, the two first axes of the CCA indicated that different processes are apparently shaping the structure of the continuous population.

On one hand, most of the genetic architecture (i.e. BPCA plot scores along the first axis) was not explained by the phenological stage ($P = 0.904$), but rather by ecological factors (Figs 1a and 5). Partial Mantel test (Table 1) showed that total radiation apparently determined most of the genetic structure of the population. Since this ecological factor influences microclimate and apparently controls the fine-scale mosaic distribution of subalpine vs alpine habitats over the area studied (C. Parisod *et al.*, unpublished), its consistent association with particular AFLP genotypes (Table 1) suggests that homogenization of genetic variance is prevented by habitat-related thermal conditions. On the other hand, the phenology, which was not spatially autocorrelated (Mantel test, $P = 0.125$), apparently also shapes the genetic structure within the continuous population since the score along the second BPCA axis was significantly explained by the mean phenological stage of plots ($P = 0.020$, $R^2 = 0.144$). This relation holds when controlled for the geographical distance (Table 1) and indicates that plots with individuals flowering synchronously were genetically more similar (Fig. 1b). Therefore, it seems that part of the genetic structure in the continuous population results from flowering time differences (Hendry & Day, 2005).

Based on AFLPs, local adaptation and phenology seem to be complementary forces shaping the genetic structure of the continuous population and, together with isolation by distance, explain different parts of the total genetic variance. It must nevertheless be stressed out that any association between genetic and environmental heterogeneity remains correlational, which does not necessarily imply causality (Kawecki & Ebert,

2004), and that the adaptive value of genotypes predicted that way still has to be addressed by molecular dissection of the phenotypes and manipulative experiments. However, the procedure adopted here (Fig. 2) provides preliminary tools guiding the construction of a stronger experimental design.

Origin and maintenance of local adaptation patterns

The rear edge population presently investigated only shows one plastid DNA haplotype (Eb; Parisod & Besnard, 2007) in a regional hotspot of diversity and the whole population probably has a common origin. Therefore, the different genotypes may have evolved either in sympatry or in parapatry, before having been clumped by climate changes (Ackerley, 2003). Several lines of evidence indicate that the observed subpopulations evolved in parapatric habitats, before they merged at the top of local mountains (i.e. secondary sympatry). First, locally adapted plots were not genetically depauperate and evenly distributed as a mosaic across the landscape (Fig. 1), suggesting that intense *in situ* directional selection was not entirely responsible for the genome-wide structure observed across the continuous population. Furthermore, plots from ecologically contrasted areas were genetically well differentiated and rather homogeneous, while plots lying in mosaic habitats were related to those extremes in association with the fine-scale ecological conditions (Figs 1, 4). Finally, tetraploid populations of *B. laevigata* across elevation gradients have already been demonstrated to be locally adapted to abiotic factors (Gasser, 1986). Although polytopic autopolyploidy resulted in genetic diversity that probably account for the evolutionary success of *B. laevigata*, allopatric lineages independently expanded along the whole altitudinal gradient, which further suggests adaptive radiation (Parisod & Besnard, 2007). These evidence and those reported here further suggest adaptive ability of autopolyploid genomes within lineages.

Although the observed association between the genome-wide structure and the environment may be of historical origin, evolutionary mechanisms explaining its maintenance against admixture in sympatry are crucial. In natural populations, local adaptation to abiotic factors can be maintained by an immigrant inviability barrier (i.e. genetically based differences between individuals growing in contrasted habitats restricting gene flow between them and lowering survivorship as well as reproduction of immigrants; Nosil *et al.*, 2005). Such a reproductive barrier may explain observed patterns of local adaptation affecting the structure of most AFLP loci. However, *B. laevigata* presents a sporophytic self-incompatibility system enforcing admixture and promoting kin avoidance (Olowokudejo & Heywood, 1984). Phenological divergence is thus the most straightforward counteracting mechanism that could promote assortative mating in continuous plant populations (Fox, 2003). Since the AFLP genetic structure was significantly explained by both environmental and phenological distances (Table 1), habitat and flowering time differences apparently work in concert to maintain the observed genome-wide spatial structure. Nevertheless, flowering time divergence was shown here to weakly influence genetic differentiation at the local scale, suggesting that this prezygotic mechanism may have only slightly reinforced the observed genome-wide structure. These findings suggest that the postzygotic component of the immigrant inviability barrier is important in maintaining the genome-wide integrity of locally adapted subpopulations in this continuous population of *B. laevigata*. However, the relative contribution of different reproductive barriers in limiting gene flow remains to be elucidated.

Acknowledgements

We would like to thank G. Besnard, L. Buechi, N. Galland, R. Holderegger, P. Taberlet as well as anonymous reviewers for valuable comments. We would like to thank A. Pasche and M. Antonin for their help in the field as well as O. Broennimann for help with the maps.

References

Ackerley D. 2003. Community assembly, niche conservatism, and adaptive evolution in changing environments. *International Journal of Plant Science* **164**: S165–S184.

Angers B, Magnan P, Plante M, Bernatchez L. 1999. Canonical correspondence analysis for estimating spatial and environmental effects on microsatellite gene diversity in brook charr (*Salvelinus fontinalis*). *Molecular Ecology* **8**: 1043–1053.

Bekkevold D, André C, Dahlgren TG, Clausen LAW, Torstensen E, Mosegaard H, Carvalho GR, Christensen TB, Norlinder E, Ruzzante DE. 2005. Environmental correlates of population differentiation in Atlantic herring. *Evolution* **59**: 2656–2668.

Bensch S, Akesson M. 2005. Ten years of AFLP in ecology and evolution: why so few animals? *Molecular Ecology* **14**: 2899–2914.

Bever JD, Felber F. 1992. The theoretical population genetics of autopolyploidy. *Oxford Surveys in Evolutionary Biology* **8**: 185–217.

Blows MW. 2007. A tale of two matrices: multivariate approaches in evolutionary biology. *Journal of Evolutionary Biology* **20**: 1–8.

Bonin A, Bellmain E, Bronken-Eidesen P, Pompanon F, Brochmann C, Taberlet P. 2004. How to track and assess genotyping errors in population genetics studies. *Molecular Ecology* **13**: 3261–3273.

Bonin A, Ehrich D, Manel S. 2007. Statistical analysis of amplified fragment length polymorphism data: a toolbox for molecular ecologists and evolutionists. *Molecular Ecology* **16**: 3737–3758.

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

Bussell JD. 1999. The distribution of random amplified polymorphic DNA (RAPD) diversity amongst populations of *Isotoma petrea* (Lobeliaceae). *Molecular Ecology* **8**: 775–789.

Comai L. 2005. The advantages and disadvantages of being polyploid. *Nature Reviews Genetics* **6**: 836–846.

Culhane AC, Perriere G, Considine EC, Cotter TG, Higgins DG. 2002. Between-group analysis of microarray data. *Bioinformatics* **18**: 1600–1608.

Doledec S, Chessel D. 1987. Rythmes saisonniers et composantes stationnelles en milieu aquatique I – Description d'un plan d'observations complet par projection de variables. *Acta Œcologica, Œcologia Generalis* **8**: 403–426.

Escudero A, Iriondo JM, Torres ME. 2003. Spatial analysis of genetic diversity as a tool for plant conservation. *Biological Conservation* **113**: 351–365.

Excoffier L, Smouse PE, Quattro JM. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* **131**: 479–491.

Fox GA. 2003. Assortative mating and plant phenology: evolutionary and practical consequences. *Evolutionary Ecology Research* **5**: 1–18.

Gasser M. 1986. Genetic–ecological investigations in *Biscutella levigata* L. *Veröffentlichungen des Geobotanischen Institutes des ETH* **86**: 7–86.

Goudet J. 1995. FSTAT (version 1.2): a computer program to calculate F-statistics. *Journal of Heredity* **86**: 485–486.

Gram WK, Sork VL. 2001. Association between environmental and genetic heterogeneity in forest tree populations. *Ecology* **82**: 2012–2021.

Hamrick JL, Allard RW. 1972. Microgeographical variation in allozyme frequencies in *Avena barbata*. *Proceedings of the National Academy of Sciences, USA* **69**: 2100–2104.

Hardy OJ. 2003. Estimation of pairwise relatedness between individuals and characterization of isolation-by-distance processes using dominant genetic markers. *Molecular Ecology* **12**: 1577–1588.

Hardy OJ, Vekemans X. 2002. SPAGEDI: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Molecular Ecology Notes* **2**: 618–620.

Hendry AP, Day T. 2005. Population structure attributable to reproductive time: isolation by time and adaptation by time. *Molecular Ecology* **14**: 901–916.

Houle D. 2007. A dispatch from the multivariate frontier. *Journal of Evolutionary Biology* **20**: 22–23.

Joost S, Bonin A, Bruford MW, Despres L, Conord C, Erhardt G, Taberlet P. 2007. A spatial analysis method (SAM) to detect candidate loci for selection: towards a landscape genomics approach to adaptation. *Molecular Ecology* **16**: 3955–3969.

Kawecki TJ, Ebert D. 2004. Conceptual issues in local adaptation. *Ecology Letters* 7: 1225–1241.

Latta RG. 2003. Gene flow, adaptive population divergence and comparative population structure across loci. *New Phytologist* **161**: 51–58.

Lenormand T. 2002. Gene flow and the limits to natural selection. *Trends in Ecology & Evolution* **17**: 183–189.

Levin DA. 2000. *The origin, expansion, and demise of plant species*. New York, NY, USA: Oxford University Press.

Linhart YB, Grant MC. 1996. Evolutionary significance of local genetic differentiation in plants. *Annual Reviews in Ecology and Systematics* **27**: 237–277.

Luikart G, England PR, Tallmon D, Jordan S, Taberlet P. 2003. The power and promise of population genomics: from genotyping to genome typing. *Nature Reviews Genetics* **4**: 981–994.

Manel S, Schwartz MK, Luikart G, Taberlet P. 2003. Landscape genetics: combining landscape ecology and population genetics. *Trends in Ecology & Evolution* **18**: 189–197.

Manton I. 1937. The problem of *Biscutella laevigata* L. II. The evidence from meiosis. *Annals of Botany* **51**: 439–465.

Nevo E, Beiles A, Krugman T. 1988. Natural selection of allozyme polymorphisms: a microgeographic climatic differentiation in wild emmer wheat, *Triticum dicoccoides*. *Theoretical and Applied Genetics* **75**: 529–538.

Nosil P, Vines TH, Funk D. 2005. Reproductive isolation caused by natural selection against immigrants from divergent habitats. *Evolution* **59**: 705–719.

Olowokudejo JD, Heywood VH. 1984. Cytotaxonomy and breeding system on the genus *Biscutella* (Cruciferae). *Plant Systematics and Evolution* **145**: 291–309.

Parisod C, Besnard G. 2007. Glacial *in situ* survival in the Western Alps and polytopic autopolyploidy in *Biscutella laevigata* L. (Brassicaceae). *Molecular Ecology* **16**: 2755–2767.

Parisod C, Trippi C, Galland N. 2005. Genetic variability and founder effect in the pitcher plant *Sarracenia purpurea* (Sarraceniaceae) in populations introduced into Switzerland: from inbreeding to invasion. *Annals of Botany* **95**: 277–286.

Patterson N, Price AL, Reich D. 2006. Population structure and Eigenanalysis. *PLoS Genetics* **2**: e190.

Pavoine S, Dufour AB, Chessel D. 2004. From dissimilarities among species to dissimilarities among communities: a double principal coordinate analysis. *Journal of Theoretical Biology* **228**: 523–537.

Quackenbush J. 2001. Computational analysis of microarray data. *Nature Reviews Genetics* **2**: 418–427.

Riginos C, Nachman MW. 2001. Population subdivision in marine environments: the contributions of biogeography, geographical distance and discontinuous habitat to genetic differentiation in a blennioid fish, *Axoclinus nigricaudus*. *Molecular Ecology* **10**: 1439–1453.

Ronfort J, Jenczewski E, Bataillon T, Rousset F. 1998. Analysis of population structure in autotetraploid species. *Genetics* **150**: 921–930.

Slatkin M. 1993. Isolation by distance in equilibrium and nonequilibrium populations. *Evolution* **47**: 264–279.

Soltis PS, Soltis DE. 2000. The role of genetic and genomic attributes in the success of polyploids. *Proceedings of the National Academy of Sciences, USA* **97**: 7051–7057.

Soltis DE, Soltis PS, Schemske DW, Hancock JF, Thompson JN, Husband BC, Judd WS. 2007. Autopolyploidy in angiosperms: have we grossly underestimated the number of species? *Taxon* **56**: 13–30.

Sork VL, Nason J, Campbell DR, Fernandez JF. 1999. Landscape approaches to historical and contemporary gene flow in plants. *Trends in Ecology & Evolution* **14**: 219–224.

Stehlik I. 2000. Nunataks and peripheral refugia for alpine plants during quaternary glaciation in the middle part of the Alps. *Botanica Helvetica* **110**: 25–30.

Storfer A, Murphy M, Evans J, Goldberg C, Robinson S, Spear S, Dezzani R, Delmelle E, Vierling L, Waits L. 2007. Putting the 'landscape' in landscape genetics. *Heredity* **98**: 128–142.

Thioulouse J, Chessel D, Doledec S, Olivier JM. 1997. ADE-4: a multivariate analysis and graphical display software. *Statistics and Computing* **7**: 75–83.

Tremetsberger K, König C, Samuel R, Pinsker W, Stuessy TF. 2002. Infraspecific genetic variation in *Biscutella laevigata* (Brassicaceae): new focus on Irene Manton's hypothesis. *Plant Systematics and Evolution* **233**: 163–181.

Vekemans X, Beauwens T, Lemaire M, Roldan-Ruiz I. 2002. Data from amplified fragment length polymorphism (AFLP) markers show indication of size homoplasy and of relationship between degree of homoplasy and fragment size. *Molecular Ecology* **11**: 139–151.

Vekemans X, Hardy OJ. 2004. New insights from fine-scale spatial genetic structure analyses in plant populations. *Molecular Ecology* **13**: 921–935.

Volis S, Anikster Y, Olsvig-Whittaker L, Mendlinger S. 2004. The influence of space in genetic-environmental relationships when environmental heterogeneity and seed dispersal occur at similar scale. *American Naturalist* **163**: 312–327.

Vos P, Hogers R, Bleeker M, Reijans M, VanDelee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M *et al.* 1995. AFLP – a new technique for DNA-fingerprinting. *Nucleic Acid Research* **23**: 4407–4414.

Walsh B. 2007. Escape from flatland. *Journal of Evolutionary Biology* **20**: 36–38.

Zimmermann NE, Kienast F. 1999. Predictive mapping of alpine grasslands in Switzerland: species versus community approach. *Journal of Vegetation Science* **10**: 469–482.

Supplementary Material

The following supplementary material is available for this article online:

Fig. S1 Comparisons of genetic differentiation estimators among plots in the continuous population of *Biscutella laevigata*.

Table S1 Detailed information on sampled plots within the continuous population of *Biscutella laevigata*

This material is available as part of the online article from: http://www.blackwell-synergy.com/doi/abs/10.1111/j.1469-8137.2007.02361.x
(This link will take you to the article abstract).

Please note: Blackwell Publishing are not responsible for the content or functionality of any supplementary materials supplied by the authors. Any queries (other than missing material) should be directed to the journal at *New Phytologist* Central Office.

