plant biology

Plant Biology ISSN 1435-8603

RESEARCH PAPER

# Historical habitat connectivity affects current genetic structure in a grassland species

Z. Münzbergová[1,2], S. A. O. Cousins[3], T. Herben[1,2], I. Plačková[2], M. Mildén[3] & J. Ehrlén[3]

1 Institute of Botany, Academy of Sciences of the Czech Republic, Průhonice, Czech Republic

2 Department of Botany, Faculty of Science, Charles University, Praha 2, Czech Republic

3 Department of Botany, Stockholm University, Stockholm, Sweden

**Keywords**

Beals index; habitat isolation; habitat suitability; heterozygote excess; landscape fragmentation; landscape history; population size; *Succisa pratensis*.

**Correspondence**

Z. Münzbergová, Institute of Botany, Academy of Sciences of the Czech Republic, CZ-252 43 Průhonice, Czech Republic.
E-mail: zuzmun@natur.cuni.cz

**Editor**

F. Roux

Received: 14 February 2012; Accepted: 9 March 2012

doi:10.1111/j.1438-8677.2012.00601.x

## ABSTRACT

**Many recent studies have explored the effects of present and past landscape structure on species distribution and diversity. However, we know little about the effects of past landscape structure on distribution of genetic diversity within and between populations of a single species. Here we describe the relationship between present and past landscape structure (landscape connectivity and habitat size estimated from historical maps) and current genetic structure in a perennial herb, *Succisa pratensis*. We used allozymes as co-dominant markers to estimate genetic diversity and deviation from Hardy–Weinberg equilibrium in 31 populations distributed within a 5 km$^2$ agricultural landscape. The results showed that current genetic diversity of populations was related to habitat suitability, habitat age, habitat size and habitat connectivity in the past. The effects of habitat age and past connectivity on genetic diversity were in most cases also significant after taking the current landscape structure into account. Moreover, current genetic similarity between populations was affected by past connectivity after accounting for current landscape structure. In both cases, the oldest time layer (1850) was the most informative. Most populations showed heterozygote excess, indicating disequilibrium due to recent gene flow or selection against homozygotes. These results suggest that habitat age and past connectivity are important determinants of distribution of genetic diversity between populations at a scale of a few kilometres. Landscape history may significantly contribute to our understanding of distribution of current genetic structure within species and the genetic structure may be used to better understand landscape history, even at a small scale.**

## INTRODUCTION

In the past two decades, the discipline of phylogeography has been established by exploring the relationship between past habitat connectivity and current distribution of genetic markers over large geographic scales (Avise *et al.* 1987; Schönswetter *et al.* 2005; Provan & Bennett 2008). Such relationships arise because genetic equilibrium is reached only after a considerable time (*e.g.* Koch & Kiefer 2006; Soltis *et al.* 2006). Studies indicate that signatures of large-scale past landscape structure can be seen in genetic characteristics of populations much longer than in, for example, species composition of communities (but see Svenning *et al.* 2004; Svenning & Skov 2005).

Surprisingly, much less is known about the effects of past habitat connectivity on genetic structure over smaller spatial and temporal scales. Nevertheless, indirect evidence suggests that such effects may be important. First, many studies have demonstrated significant effects of current habitat connectivity on genetic similarity and diversity of populations (*e.g.* Lienert *et al.* 2002; Lopez–Pujol *et al.* 2003; Hooftman *et al.* 2004; Fér & Hroudová 2009). Second, several recent studies have indicated that habitat age can have a significant effect on genetic diversity of populations (Jacquemyn *et al.* 2004; Prentice *et al.* 2006). Lastly, past habitat connectivity has been shown to influence species distribution and diversity in fragmented landscapes (Lindborg & Eriksson 2004; Herben *et al.* 2006; Lindborg 2007; Reitalu *et al.* 2010).

Based on such findings, it is reasonable to assume that past landscape structure can also leave a signature in the distribution of genetic diversity of species across a landscape. Such a signature may be evident as a correlation between past landscape structure and genetic diversity of individual subpopulations, or between past landscape structure and deviation from Hardy–Weinberg equilibrium. In particular, genetic diversity is likely to increase with increasing connectivity of a population in the past, also when accounting for present-day landscape structure. Further, populations that were connected in the past should be genetically more similar than populations that were always separated. Finally, populations with evidence of recent contact with other habitats should show heterozygote excess, while small populations isolated for a long time should show evidence of inbreeding.

Until now, only two studies have explicitly explored the relationship between past landscape structure and distribution of genetic diversity within a species (Jacquemyn *et al.* 2004; Prentice *et al.* 2006). They have shown legacies of past

habitat sizes, but have not used any information on past landscape connectivity.

In the present study, we address effects of the past landscape structure on genetic diversity and structure, but explicitly include a much larger variety of past landscape indicators, specifically including habitat connectivity in the past. We work with a set of 31 local populations of a perennial herb, *Succisa pratensis*, in a small (2.0 × 2.5 km) region in southern Sweden, which have been the subject of earlier studies (Münzbergová *et al.* 2005; Herben *et al.* 2006; Mildén *et al.* 2006). These studies have provided detailed information on the past habitat structure (four time layers beginning with 1850). They have also shown that past landscape structure has important effects on the current distribution of the species, indicating that the distribution is not in full equilibrium with the landscape.

Specifically, we link the existing detailed data on current and past landscape structure derived from historical maps with data on genetic diversity and structure of these populations. We use eight allozyme loci that are known to be variable in these populations. We take advantage of the fact that allozyme markers are co-dominant and can be allelically interpreted in diploids such as *Succisa pratensis*. These can thus be used to estimate whether and how populations deviate from Hardy–Weinberg equilibrium, in addition to information on their allelic diversity.

To test existence of landscape legacies on genetic structure, we used regression modelling to identify sources of variation in allelic diversity (measured as Nei's diversity) and in genetic structure (measured as the inbreeding coefficient). We partialled out information on current habitat characteristics and current population parameters (such as population size) and assessed whether habitat size and connectivity in the past significantly contributed to explaining genetic diversity and structure of the populations. Furthermore, we examined the relationship between genetic similarity of populations and their connectivity in the past. A significant contribution of past landscape structure indicators, after the information on current landscape and population structure had been removed, was interpreted as evidence of current distribution of genes within populations not being in equilibrium with the present-day landscape and partly reflecting landscape structure in the past.

## MATERIAL AND METHODS

### Studied species and area

*Succisa pratensis* Moench is a polycarpic herb with a maximum life span of at least 25 years (Hooftman *et al.* 2004). It flowers in August and September, with one to six, sometimes more, flower heads, on one to three stems of 20–80 cm high. Each flower head contains up to 100 small pale violet flowers. The flowers are self-compatible, but outcrossing enhances seed set considerably. Vergeer *et al.* (2003) showed that inbreeding affects small and isolated populations negatively. Clonal propagation is rare in the study area (Mildén *et al.* 2006). In Sweden, *S. pratensis* is most commonly found in dry to mesic semi-natural grasslands. It benefits from grazing and the present distribution in the landscape largely depends on the management history, *i.e.* is closely correlated with former grazing and mowing (Mildén *et al.* 2006). The species forms a persistent seed bank (Mildén *et al.* 2006) and is diploid (2n = 2x = 20).

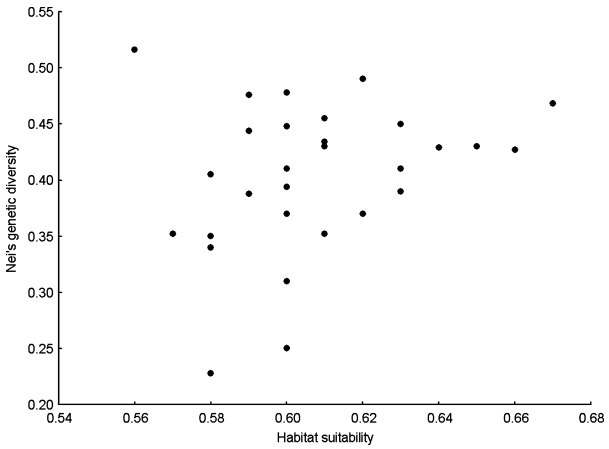


**Fig. 1.** Relationship between Nei's genetic diversity and habitat suitability measured using Beals index. In the full model, the effect of habitat is significant (P = 0.005). The pair-wise relationship shown here is not significant (P = 0.19).

The species is dispersed by exozoochory (Münzbergová *et al.* 2005) of both domestic animals (in this region represented by cattle) and wild animals (mainly roe deer). The main pollinators of *S. pratensis* are bees and bumblebees (Z. Münzbergová, unpublished observation), with dispersal usually within a maximum 100 m. We thus assume that exozoochorous dispersal on grazing animals is the main dispersal type.

The study area, Nynäs (58°50′ N, 17°24′ E), encompasses 5 $km^2$ and is situated on the Baltic Sea coast. It is a traditional rural landscape with a grazing history of at least 4000 years (Cousins *et al.* 2002) and with many semi-natural habitats, fairly small agricultural fields but also a lot of forest. During the last 100 years many changes have occurred. Fields have become larger and more open, wetlands have been drained, forestry intensified, dry grasslands abandoned and fertile moist grasslands have been cultivated (Cousins 2001).

In this study we used the same set of habitat patches as in Herben *et al.* (2006), *i.e.* all habitat patches identified as suitable for *S. pratensis* within the study area. There are 36 current populations of *S. pratensis* and 31 suitable unoccupied habitat patches. Herben *et al.* (2006) provide maps of the location of these patches in their Fig. 1.

### Data on genetic diversity

We used allozyme markers to estimate genetic diversity of the studied populations. We sampled 31 populations (all populations for which complete habitat and historical information was available) from the region for the genetic analyses. We randomly selected ten adult individuals, no matter whether vegetative or flowering, in each population. We assumed that large vegetative plants are individuals that have reproduced in the past or will reproduce in the near future, thus being valid members of the adult population. This assumption is supported by frequent observations of transitions between flowering and vegetative stages in permanent plots (Mildén *et al.* 2006). In some cases, a population consisted of two or more separate patches of individuals (separated

by 10–30 m). In that case, we selected several individuals from each patch to represent the population. It should be noted that units referred to as populations in this study may not necessarily be populations in a genetic sense. Rather, the populations in this study are patches of *S. pratensis* separated by more than 100 m from another patch of *S. pratensis*.

The sample size of ten individuals was selected based on the size of the smallest populations. To explore if this relatively small sample size could affect the results, we used a rarefaction technique to estimate the effect of sample size on number of alleles sampled from the population. The rarefaction method compares the number of species/alleles found in different samples with different sampling effort (Krebs 1989). We used the implementation of J. Brzustowski (Rarefaction Calculator; http://www.biology.ualberta.ca/jbrzusto/rarefact.php). The results suggest that the sample size of ten covers the genetic diversity of the populations relatively well, and the limited sample size thus is not a major problem in our data. We assume that any negative effects of the small sample size are likely to act in a conservative manner as our data are less precise and we are thus less likely to find any patterns.

From each selected individual, we took a sample of a fresh leaf of approximately 10 $cm^2$. The samples were kept on ice after sampling and the isozymes were extracted within 3 days after collection. Electrophoresis was performed on crude protein extracts of leaf material. All enzymes were resolved on polyacrylamide gels using 8.16% separating gel and 4% stacking gel. Eight enzyme systems were investigated in the first step; four of them yielded clear patterns and were analysed further: phosphoglucomutase (PGM, EC 5.4.2.2), leucine aminopeptidase (LAP, EC 3.4.11.1), alcohol dehydrogenase (ADH, EC 1.1.1.1) and aspartate aminotransferase (AAT, EC 2.6.1.1). For more details on the isozyme analyses see the protocol in Münzbergová & Plačková (2010). One locus for PGM, two loci for LAP and ADH, and three loci for AAT were allelically interpreted, yielding eight independent allelically interpreted polymorphic loci. We used the data on genetic constitution of each population to calculate Nei's (1973) measure of the average gene diversity per locus, observed and expected heterozygosity and inbreeding coefficient using POPGENE version 1.31 (Yeh et al. 1997).

### Habitat data

Habitat suitability for *S. pratensis* in the present landscape was assessed using data on species composition. Specifically, we used Beals index (Münzbergová & Herben 2004) to calculate the probability of occurrence of *S. pratensis* based on presence of other plant species on the habitat. Beals index uses information on co-occurrence of the target species with all the other species in the community to calculate the probability that the target species can occur in each habitat. The probability of occurrence of species *j* at a habitat *i* is defined as

$$p_{ij=}(1/S_i)\sum_k N_{jk}/N_k \quad (1)$$

where $p_{ij}$ is probability of finding species j at habitat i, $S_i$ is number of species at habitat i (minus 1 if species j is present), $N_{jk}$ is the number of joint occurrences of species j and k in the reference database, $N_k$ is the number of occurrences of species k in the reference database. Since we did not have an independent reference data set, we used the data set collected at the study sites as both the reference database and the database for which the Beals index was calculated. A seed addition experiment for a subset of these habitats demonstrated a strong correlation between Beals index and recruitment success in this system (see Mildén *et al.* 2006 for details). It is thus reasonable to assume that habitat suitability as defined by Beals index captures important variation on the present growing conditions of the populations. As the Beals index indicates a probability that a species will occur at a given habitat, we suggest that habitats with higher probability of occurrence are more suitable for *S. pratensis,* in terms of individual growth and potential population size.

Habitat suitability in the past (suitable or non-suitable) was assessed using information on habitat types from old land-use maps from 1850 and 1900 and aerial photographs from 1945 derived from Cousins (2001); see Herben *et al.* 2006 for details of the habitat classification. It is clear that the classification in the past is much rougher. However, the results of the estimation of habitat suitability using Beals index largely correspond to the major habitat types in the area and thus the two classifications should be comparable.

Habitat size in each time layer was calculated as the sum of sizes of all habitats at least partly contributing to the current habitat in the respective time layer. Habitat age was calculated as the weighted mean of the time over which the different parts of the current habitat patch were continuously suitable for *S. pratensis* (weights represented the proportion of the locality with the given age). We expected this value to be correlated with the age of the population. Another alternative would be to use a maximum age of the habitat, *i.e.* the age of the oldest part of the habitat. However, in our data set, the maximum age of the habitats was 175 years in all but three habitats. The maximum age was thus not used in the further analyses.

Current habitat connectivity was calculated as hypothetical number of seeds expected to arrive at each habitat from all surrounding habitats, given a mean dispersal distance of the seed (equal to 0.94 m; Herben *et al.* 2006). Hypothetical number of seeds arriving at a habitat was thus defined as

$$n = \sum_{ijk} I_{jk}\exp(-\alpha d_{ij}) \quad (2)$$

where $\alpha$ is dispersal shape parameter derived using mean dispersal distance of the seeds under the assumption of exponential dispersal, $d_{ij}$ is the distance between centres of cells i and j (the whole region was divided into 5 m × 5 m grid cells), $I_{jk}$ is 1 if cell j lies within the habitat k, and 0 otherwise; i represents indexing over all cells of the focus habitat, j represents indexing over all cells in the grid, and k represents indexing over all source habitats. Using this approach, the connectivity measure incorporates the strength of all the surrounding potential sources of seeds and distance to these sources.

For measures of past connectivity, *i.e.* connectivity in 1850, 1900, 1945, number of seeds expected to arrive from all the surrounding populations to all patches at least partly contributing to a current habitat patch in the respective time layer in the past were summed. In all cases, the habitat connectivity was calculated as connectivity between all suitable habitats, *i.e.* not only the habitats currently occupied by the

species because we had no information on the actual species distribution in the past. Habitat suitability and habitat age were uncorrelated with the other independent variables (not shown). Habitat connectivity was also only weakly correlated to habitat size. Habitat connectivities from the different time layers were significantly correlated to each other, but the correlations were rather weak. Habitat sizes from the different time layers were, however, quite closely correlated to each other (Appendix S1).

Finally, we used the total population size and the number of flowering individuals (averaged over the period 2000–2002; M. Mildén, unpublished observation) at each occupied habitat as another predictor of genetic diversity of the populations, as population size is considered the key determinant of genetic status of populations (*e.g.* van Treuren *et al.* 1991; Prentice *et al.* 2006; Dostálek *et al.* 2010; Šmídová *et al.* 2011).

### Data analysis: genetic diversity

We examined the effects of habitat and population characteristics on Nei's (1973) estimate of average gene diversity over loci, as a measure of genetic diversity of the population, and on the inbreeding coefficient as a measure of the deviation of the population from Hardy–Weinberg equilibrium. As predictor variables in regression models we used habitat suitability, mean age of the habitat, current population size, number of flowering individuals, habitat connectivity and habitat size in the past (*i.e.* in 1945, 1900 and 1850). No transformation of the dependent variables was necessary. We, however, transformed values of some of the predictors to remove skew in their values and to achieve reasonably uniform distribution. Data on mean habitat age were log-transformed after being subtracted from 200 years and data on habitat size and habitat connectivity were log-transformed.

To account for the correlation between independent variables and to select the best subset of independent variables explaining variation in genetic diversity and structure we used bi-directional step-wise linear regression, as implemented by the function step in (S-PLUS 2000; forward step-wise regression combined with backward regression in each step). Inclusion or exclusion of each term was decided using Akaike information criterion (AIC), a form of penalised log-likelihood analysis (Crawley 2005).

We used a hierarchical series of four step-wise procedures for conservative testing of the potential effects of past habitat size and connectivity. First, we ran a step-wise analysis as described above with all predictors available for selection. Second, we forced population size and number of flowering individuals into the regression, and made only the remaining predictors available for selection. In the third model, we forced present-day habitat sizes and connectivities into the model, with remaining predictors available for selection. Finally, we forced both population size and number of flowering individuals, and present-day habitat sizes and connectivities into the model before the selection proper.

To test whether the selected models were significant, we compared them with the appropriate baseline models (null model or model with predictors forced into the equation) using F statistics. The significance of the single terms in the model was assessed by comparing models with and without the tested term using F statistics.

### Data analysis: genetic similarity

To explore the effect of current and past landscape structure on genetic similarity of the populations, we calculated pairwise genetic distances between all populations using Nei's unbiased genetic distance (Nei 1978) with POPGENE version 1.31 (Yeh *et al.* 1997). We then expressed geographic distance between each pair of populations by calculating Euclidean distance between the centroids of the current populations. We also calculated edge-to-edge distances between populations. The Euclidean distances between the centroids, however, turned out to be better predictors of genetic similarity of the populations and were thus used in the subsequent analyses.

To examine if genetic similarity was affected by physical connections between past habitats, we created a matrix for each time layer describing whether each pair of current populations was a part of the same habitat patch at a given time step in the past. This resulted into a square habitat × habitat matrix of 0 and 1 for each time layer in the past (1850, 1900 and 1945). We refer to this matrix as information on common history of the sites in a given time period. We also created a matrix of times since disconnection for each pair of populations. We assumed that the populations separated midway between the two time layers in which they were connected and disconnected, respectively. For habitats not connected in 1850 (the time of the oldest available map) we assumed that they were connected up to 1800. This assumption was based on the fact that the habitats before 1800 occurred in a mosaic where livestock grazed fallow fields every second year and all fields after harvest in the whole area, resulting in high functional connectivity between patches (Cousins 2006; Dahlström *et al.* 2006).

We used Mantel tests to examine relationships between the different dissimilarity matrices. Specifically, we tested for a relationship between genetic and geographic distance, between genetic similarity and the matrices describing if they were part of the same habitat in the different time layers (referred to as common history of the sites), and between the genetic similarity and the time since disconnection. We then used a partial Mantel test to estimate the relationship between genetic similarity and the different measures of past connectivity after removing the effect of current geographic distance. Significance of the relationships was assessed using 9999 permutations of the data using the Vegan package in R (R Development Core Team 2003).

## RESULTS

There was large variation in the allelic diversity among populations and strong differences between individual populations (mean Nei's diversity was 0.41 ± 0.06 SD, mean expected heterozygosity 0.43 ± 0.07, n = 31, mean $F_{st}$ among populations = 0.14). Allelic diversity (measured as Nei's diversity) was not significantly correlated with the population size ($r = 0.16$, $P = 0.39$, $n = 31$), but was negatively correlated with the proportion of individuals flowering ($r = -0.35$, $P = 0.05$, $n = 31$).

Step-wise regression modelling showed that allelic diversity of a population increased with increasing habitat suitability and decreased with increasing habitat size in 2000 (Table 1, Fig. 1). It was also influenced by historical factors. It increased with both increasing habitat age and with increasing habitat connectivity in 1850 (Fig. 2). After accounting for habitat size,

**Table 1.** Effect of current and past landscape structure on genetic diversity of *S. pratensis* populations (Nei's genetic diversity and inbreeding coefficient). Models were identified using bi-directional step-wise regressions using penalised log-likelihood analysis (AIC criterion).

| | | all effects | effects of 2000 removed | effects of population size removed | effects of 2000 and population size removed |
|---|---|---|---|---|---|
| Nei's genetic diversity | | | | | |
| Population size | | 0.01 n.s. (+) | 0.01 n.s. (+) | Inc. | Inc. |
| Number of flowering plants | | 0.12 **(–) | 0.12 **(–) | Inc. | Inc. |
| Habitat suitability | | 0.06 *(+) | 0.06 *(+) | 0.06 *(+) | 0.06 †(+) |
| Habitat age | | 0.13 **(+) | 0.09 *(+) | 0.13 **(+) | 0.09 *(+) |
| Habitat connectivity | 1850 | 0.03 *(+) | 0.03 †(+) | 0.03 *(+) | 0.03 n.s.(+) |
| | 1900 | | | | |
| | 1945 | | | | |
| | 2000 | | Inc. | | Inc. |
| Habitat size | 1850 | | 0.01(–) | | |
| | 1900 | | | | |
| | 1945 | | | | |
| | 2000 | 0.17 ** (–) | Inc. | 0.17 ** (–) | Inc. |
| | Model | | | | |
| | $R^2$ | 0.52 | 0.32 | 0.39 | 0.18 |
| | Model P | 0.019 | 0.049 | 0.033 | 0.147 |
| Inbreeding coefficient | | | | | |
| Population size | | 0.02 n.s. (+) | 0.01 n.s. (+) | Inc. | Inc. |
| Number of flowering plants | | 0.14 * (–) | 0.14 * (–) | Inc. | Inc. |
| Habitat suitability | | | | | |
| Habitat age | | | | | |
| Habitat connectivity | 1850 | 0.05 * (+) | 0.05 * (+) | 0.05 * (+) | 0.05 * (+) |
| | 1900 | 0.18 ** (–) | 0.18 ** (–) | 0.18 ** (–) | 0.18 ** (–) |
| | 1945 | 0.09 * (–) | 0.09 * (–) | 0.09 * (–) | 0.09 * (–) |
| | 2000 | | Inc. | | Inc. |
| Habitat size | 1850 | | | | |
| | 1900 | | | | |
| | 1945 | | | | |
| | 2000 | | Inc. | | Inc. |
| | Model | | | | |
| | $R^2$ | 0.48 | 0.47 | 0.32 | 0.32 |
| | Model P | 0.015 | 0.014 | 0.019 | 0.018 |

The values in the table are $R^2$ values for the single terms, followed by indication of significance of the single term (n.s. for $P > 0.10$, † for $P \leq 0.1$, * for $P \leq 0.05$ and ** for $P \leq 0.01$) and the signs of direction of the effect for the terms in brackets. The significance of the whole models was determined by comparing a model with the respective baseline model (with only forced variables included). N = 31. Inc. indicates that the parameter was included in the baseline model (to remove the effects of conditions in 2000 and population size).

connectivity and suitability in 2000, habitat connectivity in 1850 was still included in the model selected using step-wise selection, but its individual effect was not significant (Table 1).

Almost all populations showed an excess of heterozygotes, as indicated by negative values of the inbreeding coefficient, but the degree of heterozygote excess varied among populations (Fig. 3). Heterozygote excess was marginally significantly correlated with allelic diversity ($r = 0.33$, $P = 0.07$, $n = 31$) and with the proportion of individuals flowering ($r = -0.32$, $P = 0.08$, $n = 31$), but not with population size ($r = 0.15$, $P = 0.43$, $n = 31$). Results of step-wise regression indicated that the excess of heterozygotes decreased with current connectivity. It increased with the habitat connectivity in 1900 and 1945, but decreased with the connectivity in 1850 (Table 1). These effects were significant even after accounting for habitat size, connectivity and suitability in 2000.

There was significant negative correlation between genetic similarity and geographic distance between populations (Mantel test, $z = -0.17$, $P = 0.017$). There was also a signifi-

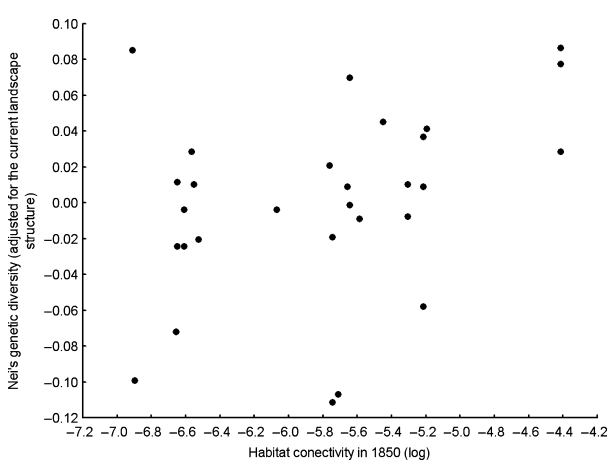


**Fig. 2.** Relationship between habitat connectivity in 1850 and genetic diversity. Genetic diversity is plotted as residuals after removing the effect of habitat size, connectivity and suitability in 2000 (P = 0.050).

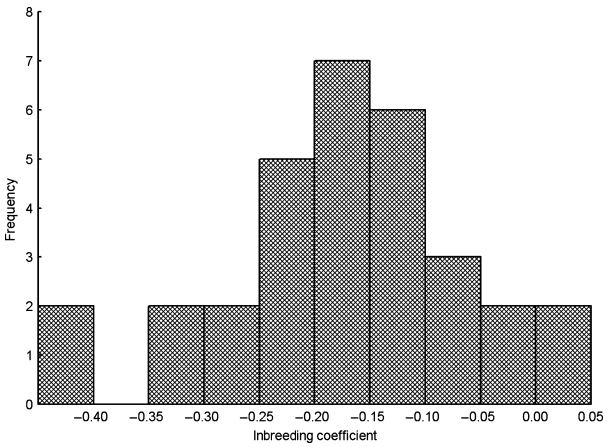


**Fig. 3.** Histogram of values of the inbreeding coefficient for each population.

cant positive correlation between genetic similarity and the common history of the populations in the three time layers in the past ($z = 0.15$, $P = 0.004$, $z = 0.13$, $P = 0.004$ and $z = 0.13$, $P = 0.006$ for connection in 1850, 1900 and 1945, respectively). Relationships between genetic similarity and common history were marginally significant also after removing the effect of current geographic distance ($z = 0.09$, $P = 0.06$, $z = 0.07$, $P = 0.10$ and $z = 0.07$, $P = 0.10$ for common history in 1850, 1900 and 1945, respectively). There was a significant negative correlation between genetic similarity and time since disconnection ($z = -0.14$, $P = 0.005$). This relationship was also marginally significant after removing the effect of geographic distance ($z = -0.08$, $P = 0.08$).

## DISCUSSION

Our results suggest that the allelic diversity and inbreeding coefficient of *S. pratensis* populations and their genetic similarity are influenced not only by the contemporary landscape structure but also by past landscape structure, namely habitat connectivity. This indicates that the population structure of *S. pratensis* in the study landscape is far from equilibrium, both in terms of spatial distribution of individual alleles and in terms of Hardy–Weinberg equilibrium of individual populations.

For the rest of the discussion, it is important to mention that the differences in genetic diversity and inbreeding are interpreted (with one exception below) assuming that allozymes are neutral markers. Some of the patterns could, however, also be due to different selection pressure at different localities in cases where the marker was under selection. We consider strong divergent selection at the nearby localities in such a small landscape fragment unlikely and thus mostly retain the neutrality assumption. However, confirming or rejecting neutrality of the marker is not possible based on our data, and there is also no general consensus on this issue in the literature (*e.g.* Hartl & Dykhuizen 1981; Hamrick & Godt 1989; Skibinski & Ward 1998; Volis *et al.* 2003).

The genetic structure of populations is due to interplay of both present-day and past influences. However, as *S. pratensis* is a rather long-lived species with overlapping generations (mean life span 22.2 years, mean age to first reproduction 5.0 years; estimated based on transition matrix models used in Mildén *et al.* 2006), a persistent seed bank and not very frequent seedling reproduction (Mildén *et al.* 2006), the populations are likely to change their genetic structure rather slowly (see *e.g.* Emigh & Pollak 1979). This may result in effects of past landscape structure that persist for many years, which in our study is also supported by the generally negative values of the inbreeding coefficient. No local population was isolated long enough to show a detectable effect of inbreeding; in contrast, many of the populations showed disequilibrium due to gene flow from other, genetically different, populations.

The potential for resilience in the genetic structure implies that the effects of the present-day landscape indicators must be interpreted with caution and must take into account possible time lags in the population genetic response to landscape structure.

### Background: effects of present-day parameters of populations and habitats

Genetic diversity of the populations was quite independent of population size. The relationship between population size and genetic diversity has been repeatedly demonstrated in other systems (*e.g.* van Treuren *et al.* 1991; Raijmann *et al.* 1994; Prentice *et al.* 2006; Dostálek *et al.* 2010; Šmídová *et al.* 2011; but see *e.g.* Jacquemyn *et al.* 2004) and is due to higher genetic drift, and thus loss of genetic diversity, in populations that have been small for a long time. The absence of such a relationship in our system suggests that landscape structure and population sizes have changed rapidly relative to the generation time of the study species. This is confirmed in the transition matrix analyses of the population dynamics of the species, showing that *S. pratensis* is able to expand rather fast if a habitat is suitable (Mildén *et al.* 2006). Current population size thus need not be very informative for the average population sizes in the past. Larger populations are also typically closer to Hardy–Weinberg equilibrium (effect of population size on inbreeding coefficient is positive), indicating less (relative) influence of outside migrants and more time to reach equilibrium.

In contrast to the lack of effects of population size, genetic diversity was negatively correlated with the proportion of flowering plants. This is likely related to a third factor, most likely habitat suitability, which influences both genetic diversity and the probability of flowering. In less suitable habitats only a few really vigorous plants (which tend to flower) remain in the population, but the overall allelic diversity is low (for a similar pattern, see *e.g.* Luijten *et al.* 2000). These populations also have larger heterozygote excess, most likely due to large proportion of pollination from outside the population or to selection against homozygotes. In contrast, the large populations often host a large proportion of vegetative individuals and only a few plants flower each year.

Among the characteristics of current landscape structure, both habitat suitability and habitat size affected the allelic diversity, while present-day connectivity had no significant effect. The positive effect of habitat suitability on genetic diversity could possibly be attributed to the fact that more suitable habitats support higher sexual reproduction in *S. pratensis* (Mildén *et al.* 2006), as in other species (Münzbergová 2004). More frequent sexual reproduction has also been shown to result in maintenance of high genetic diversity in other systems (*e.g.* Alberto *et al.* 2006; Kjolner *et al.* 2006; Mandák *et al.* 2006; Scarcelli *et al.* 2006; Tsyusko *et al.* 2006),

although the data from our system do not fully support this contention (see above). We assume that the key effect of habitat suitability in our system is related to the fact that it is estimated from total species composition of the habitats, which may also reflect past habitat characteristics (*e.g.* Peterken & Game 1984; Bellemare *et al.* 2002; Jacquemyn *et al.* 2003; Ehrlén *et al.* 2006; Chýlová & Münzbergová 2008; Tájek *et al.* 2011). It may thus be an indicator of the population size in the past, or of the duration of the population at the site. Present-day habitat size, in contrast, had a negative effect. This was most likely caused by the structure of the landscape, where larger habitats generally host less dense populations, possibly leading to lower outcrossing rates and lower genetic diversity. This is supported by numerically lower inbreeding coefficients in larger habitats, indicating that populations in such habitats are closer to equilibrium.

### Effects of past landscape structure

Genetic diversity of populations increased with increasing habitat connectivity in 1850, *i.e.* in the oldest time layer. This suggests that spatial structure of the allelic diversity in *S. pratensis* is still partly a result of landscape structure 150 years ago. Both the effect of habitat age and of habitat connectivity in 1850 were also included in the final model after taking the current landscape structure into account (although the separate effect of connectivity in 1850 was not significant in the case of genetic diversity), suggesting that past landscape structure partly influences genetic diversity of the current populations.

The finding that genetic diversity of the populations is partly determined by landscape structure in the past was also supported by the comparison of genetic similarity between populations. After accounting for current geographic distance between populations, the effects of past connectivity of the habitats were only marginally significant. Similar to the data exploring the effect on genetic diversity of populations, this analysis also suggested that the effect of connectivity in 1850 was stronger than the effect of connectivity in 1900 and 1945.

The good predictive power of the landscape structure in 1850 as compared to subsequent time layers may be linked to the fact that the landscape structure occurring in 1850 is likely to have been fairly static since AD 1000–1200 (Cousins *et al.* 2002). This long-standing landscape structure was thus the main driver, over time, of the genetic structure of plant populations. The results thus suggest that this past landscape structure could still be seen in the current distribution of genetic diversity, despite the large changes in the landscape over the last century (Cousins 2001; Cousins *et al.* 2007). The effect of landscape structure in 1850 also corresponds to results of our previous studies suggesting that colonisation–extinction dynamics of *S. pratensis* are very slow (due to its limited dispersal capacity and long life-span) and that past landscape structure has strong effects on distribution of *S. pratensis* (Münzbergová *et al.* 2005; Herben *et al.* 2006).

The signal of past landscape connectivity on spatial structure of allelic diversity goes back to 1850, with later changes having less effect. In contrast, the inbreeding coefficient is, rather as expected, more affected by more recent landscape parameters. In particular, habitat connectivity in the past (namely in the 1900 and 1945 layers) had a positive effect on the absolute value of the inbreeding coefficient (*i.e.* heterozygote excess). This means that habitats that were recently connected to other habitats are further from the equilibrium than habitats with no history of such contacts, and that the effect could persist in local populations for 50–100 years.

In summary, this study provides the first direct demonstration of effects of past landscape connectivity on spatial structure of genetic diversity and on local genetic structure. This adds to earlier studies that showed effects of habitat age on distribution of genetic diversity (Jacquemyn *et al.* 2004; Prentice *et al.* 2006). Here we were able to show the absolute time scale over which such effects persist. Our results suggest that landscape structure in 1850 is a stronger determinant of the distribution of genes within the landscape than landscape structure in the subsequent time periods. This indicates that genetic properties of the populations are quite conserved and that we cannot fully understand them without taking past landscape structure into account.

These findings extend the list of ecological legacies of past landscape structure (*e.g.* Peterken & Game 1984; Bellamare *et al.* 2002; Jacquemyn *et al.* 2003; Lindborg & Eriksson 2004; Dahlström *et al.* 2006). All of these legacies (including the genetic ones) are due to transient state/disequilibrium of rapidly changing agricultural landscapes to which many species of conservation interest are bound. On the other hand, this disequilibrium due to landscape history makes phylogeographic reasoning applicable at small spatial and temporal scales (see *e.g.* Kolseth & Lonn 2005; Fér & Hroudová 2009). We suggest that our results may also hold for many other poorly dispersed long-lived perennial herbs of habitats with similar temporal dynamics.

## ACKNOWLEDGEMENTS

We would like to thank to the anonymous reviewers for comments on the manuscript. This study was supported by grant GACR P504/10/0456 and partly by AV0Z60050516.

## SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

**Appendix S1.** Pair wise correlations among the independent variables used in the study.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

## REFERENCES

Alberto F., Arnaud-Haond S., Duarte C.M., Serrao E.A. (2006) Genetic diversity of a clonal angiosperm near its range limit: the case of *Cymodocea nodosa* at the Canary Islands. *Marine Ecology-Progress Series*, **309**, 117–129.

Avise J.C., Arnold J., Ball R.M., Bermingham E., Lamb T., Neigel J.E., Reeb C.A., Saunders N.C. (1987) Intraspecific phylogeography – the mitochondrial-DNA bridge between population genetics and systematics. *Annual Review of Ecology and Systematics*, **18**, 489–522.

Bellemare J., Motzkin G., Foster D.R. (2002) Legacies of the agricultural past in the forested present: an assessment of historical land-use effects on rich mesic forests. *Journal of Biogeography*, **29**, 1401–1420.

Chýlová T., Münzbergová Z. (2008) Past land use co-determines present distribution of dry grassland plant species. *Preslia*, **80**, 183–198.

Cousins S.A.O. (2001) Analysis of land-cover transitions based on 17th and 18th century cadastral maps and aerial photographs. *Landscape Ecology*, **16**, 41–54.

Cousins S.A.O. (2006) Plant species richness in mid-field islets and road verges – the effect of landscape fragmentation. *Biological Conservation*, **127**, 500–509.

Cousins S.A.O., Eriksson Å., Franzén D. (2002) Reconstructing past land use and vegetation patterns using palaeogeographical and archaeological data A focus on grasslands in Nynäs by the Baltic Sea in south-eastern Sweden. *Landscape and Urban Planning*, **61**, 1–18.

Cousins S.A.O., Ohlson H., Eriksson O. (2007) Effects of historical and present fragmentation on plant species diversity in semi-natural grasslands in Swedish rural landscapes. *Landscape Ecology*, **22**, 723–730.

Crawley M.J. (2005) Statistical computing. An introduction to data analysis using S-Plus. Wiley, Chichester.

Dahlström A., Cousins S.A.O., Eriksson O. (2006) The history (1620–2003) of land use, people and livestock, and the relationship to present plant species diversity in a rural landscape in Sweden. *Environment and History*, **12**, 191–212.

Dostálek T., Münzbergová Z., Plačková I. (2010) Genetic diversity and its effect on fitness in an endangered plant species, *Dracocephalum austriacum* L. *Conservation Genetics*, **11**, 773–783.

Ehrlén J., Münzbergová Z., Diekmann M., Eriksson O. (2006) Long-term assessment of seed limitation in plants: results from an 11-year experiment. *Journal of Ecology*, **94**, 1224–1232.

Emigh T.H., Pollak E. (1979) Fixation probabilities and effective population numbers in diploid populations with overlapping generations. *Theoretical Population Biology*, **15**, 86–107.

Fér T., Hroudová Z. (2009) Genetic diversity and dispersal of *Phragmites australis* in a small river system. *Aquatic Botany*, **90**, 165–171.

Hamrick J.L., Godt M.J. (1989) Allozyme diversity in plant species. In: Brown A.H.D., Clegg M.T., Kahler A.L., Weir B.S. (Eds), *Plant population genetics, breeding and germplasm resources*. Sinauer, Sunderland, MA. pp 43–63.

Hartl D.L., Dykhuizen D.E. (1981) Potential for selection among nearly neutral allozymes of 6-phosphogluconate dehydrogenase in *Escherichia coli*. *Proceedings of the National Academy of Sciences USA*, **78**, 6344–6348.

Herben T., Münzbergová Z., Mildén M., Ehrlén J., Cousins S.O.A., Eriksson O. (2006) Long-term spatial dynamics of *Succisa pratensis* in a changing rural landscape: linking dynamical modeling with historical maps. *Journal of Ecology*, **94**, 131–143.

Hooftman D.A.P., Billeter R.C., Schmid B., Diemer M. (2004) Genetic effects of habitat fragmentation on common species of Swiss fen meadows. *Conservation Biology*, **18**, 1043–1051.

Jacquemyn H., Butaye J., Hermy M. (2003) Influence of environmental and spatial variables on regional distribution of forest plant species in a fragmented and changing landscape. *Ecography*, **26**, 768–776.

Jacquemyn H., Honnay O., Galbusera P., Roldán-Ruiz I. (2004) Genetic structure of the forest herb *Primula elatior* in a changing landscape. *Molecular Ecology*, **13**, 211–219.

Kjolner S., Sastad S.M., Brochmann C. (2006) Clonality and recombination in the arctic plant *Saxifraga cernua*. *Botanical Journal of the Linnean Society*, **152**, 209–217.

Koch M.A., Kiefer C. (2006) Molecules and migration: biogeographical studies in cruciferous plants *Plant Systematics and Evolution*, **259**, 121–142.

Kolseth A.K., Lonn M. (2005) Genetic structure of *Euphrasia stricta* on the Baltic island of Gotland, Sweden. *Ecography*, **28**, 443–452.

Krebs C.J. (1989) *Ecological methodology*. Harper and Row, New York.

Lienert J., Fischer M., Schneller J., Diemer M. (2002) Isozyme variability of the wetland specialist *Swertia perennis* (Gentianaceae) in relation to habitat size, isolation, and plant fitness. *American Journal of Botany*, **89**, 801–811.

Lindborg R. (2007) Evaluating the distribution of plant life-history traits in relation to current and historical landscape configurations. *Journal of Ecology*, **95**, 555–564.

Lindborg R., Eriksson O. (2004) Historical landscape connectivity affects present plant species diversity *Ecology*, **85**, 1840–1845.

Lopez-Pujol J., Orellana M.R., Bosch M., Simon J., Blanche C. (2003) Effects of habitat fragmentation on allozyme diversity and conservation status of the coastal sand dune plant *Stachys maritima* (Lamiaceae) in the Iberian Peninsula. *Plant Biology*, **5**, 504–512.

Luijten S.H., Dierick A., Oostermeijer J.G.B., Raijmann L.E.L., Den Nijs H.C.M. (2000) Population size, genetic variation, and reproductive success in a rapidly declining, self-incompatible perennial (*Arnica montana*) in The Netherlands. *Conservation Biology*, **14**, 1776–1787.

Mandák B., Bímová K., Mahelka V., Plačková I. (2006) How much genetic variation is stored in the seed bank? A study of Atriplex tatarica (Chenopodiaceae). *Molecular Ecology*, **15**, 2653–2663.

Mildén M., Münzbergová Z., Herben T., Ehrlén J. (2006) Metapopulation dynamics of a perennial plant, *Succisa pratensis*, in an agricultural landscape. *Ecological Modelling*, **199**, 464–475.

Münzbergová Z. (2004) Effect of spatial scale on factors limiting species distributions in dry grassland fragments. *Journal of Ecology*, **92**, 854–867.

Münzbergová Z., Herben T. (2004) Identification of suitable unoccupied habitats in metapopulation studies using co-occurrence of species. *Oikos*, **105**, 408–414.

Münzbergová Z., Plačková I. (2010) Seed mass and population characteristics interact to determine performance of *Scorzonera hispanica* under common garden conditions. *Flora*, **205**, 552–559.

Münzbergová Z., Mildén M., Ehrlén J., Herben T. (2005) Population viability and reintroduction strategies: a spatially explicit landscape-level approach. *Ecological Applications*, **15**, 1377–1786.

Nei M. (1973) Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences USA*, **70**, 3321–3323.

Nei M. (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*, **89**, 583–590.

Peterken G.F., Game M. (1984) Historical factors affecting the number and distribution of vascular plant species in the woodlands of central Lincolnshire. *Journal of Ecology*, **72**, 155–182.

Prentice H.C., Lonn M., Rosquist G., Ihse M., Kindstrom M. (2006) Gene diversity in a fragmented population of Briza media: grassland continuity in a landscape context. *Journal of Ecology*, **94**, 87–97.

Provan J., Bennett K.D. (2008) Phylogeographic insights into cryptic glacial refugia. *Trends in Ecology and Evolution*, **23**, 564–571.

R Development Core Team (2003) *R: a Language and Environment for Statistical Computing*. http://cran.rproject.org.

Raijmann L.L., van Leeuwen N.C., Kersten R., Oostermeijer J.G.B., Dennijs H.C.M., Menken S.B.J. (1994) Genetic variation and outcrossing rate in relation to population size in *Gentiana pneumonanthe* L. *Conservation Biology*, **8**, 1014–1026.

Reitalu T., Johansson L.J., Sykes M.T., Hall K., Prentice H.C. (2010) History matters: village distances, grazing and grassland species diversity. *Journal of Applied Ecology*, **47**, 1216–1224.

Scarcelli N., Tostain S., Vigouroux Y., Agbangla C., Dainou O., Pham J.L. (2006) Farmers' use of wild relatives and sexual reproduction in a vegetatively propagated crop. *The case of yam in Benin. Molecular Ecology*, **15**, 2421–2431.

Schönswetter P., Stehlik I., Holderegge R., Tribsch A. (2005) Molecular evidence for glacial refugia of mountain plants in the European Alps. *Molecular Ecology*, **14**, 3547–3555.

Skibinski D.O.F., Ward R.D. (1998) Are polymorphism and evolutionary rate of allozyme proteins limited by mutation or selection? *Heredity*, **81**, 692–702.

Šmídová A., Münzbergová Z., Plačková I. (2011) Genetic diversity of a relict plant species, *Ligularia sibirica* (L.) Cass (Asteraceae). *Flora*, **206**, 151–157.

Soltis D.E., Morris A.B., McLachlan J.S., Manos P.S., Soltis P.S. (2006) Comparative phylogeography of unglaciated eastern North America. *Molecular Ecology*, **15**, 4261–4293.

S-PLUS (2000) *Professional edition for Windows, Release 2*. MathSoft, Inc., Cambridge, MA

Svenning J.C., Skov F. (2005) The relative roles of environment and history as controls of tree species composition and richness in Europe. *Journal of Biogeography*, **32**, 1019–1033.

Svenning J.C, Kinner D.A., Stallard R.F., Engelbrecht B.M.J., Wright S.J. (2004) Ecological determinism in plant community structure across a tropical forest landscape. *Ecology*, **85**, 2526–2538.

Tájek P., Bucharová A., Münzbergová Z. (2011) Limitation of distribution of two rare ferns in fragmented landscape. *Acta Oecologica*, **37**, 495–502.

van Treuren R., Bijlsma R., van Delden W., Ouborg N.J. (1991) The significance of genetic erosion in the process of extinction.1 Genetic differentiation in *Salvia pratensis* and *Scabiosa columbaria* in relation to population size. *Heredity*, **66**, 181–189.

Tsyusko O.V., Smith M.H., Oleksyk T.K., Goryanaya J., Glenn T.C. (2006) Genetics of cattails in radioactively contaminated areas around Chernobyl. *Molecular Ecology*, **15**, 2611–2625.

Vergeer P., Rengelink R., Copal A., Ouborg J. (2003) The interacting effects of genetic variation, habitat quality and population size on performance of *Succisa pratensis*. *Journal of Ecology*, **91**, 18–26.

Volis S., Shulgina I., Ward D., Mendlinger S. (2003) Regional subdivision in wild barley – allozyme variation: adaptive or neutral? *Journal of Heredity*, **94**, 341–351.

Yeh F. C., Yang R.C., Boyle T.B.J., Ye Z.H., Mao J.X. (1997) *POPGENE, the user-friendly shareware for population genetic analysis*. Molecular Biology and Biotechnology Centre, University of Alberta, Edmonton, AB, Canada.