Spatial heterogeneity in landscape structure influences dispersal and genetic structure: empirical evidence from a grasshopper in an agricultural landscape

BERTRAND GAUFFRE,* †‡ SOPHIE MALLEZ,† †‡ ¶ ** MARIE-PIERRE CHAPUIS,† †‡ RAPHAEL LEBLOIS,‡ †‡ ISABELLE LITRICO,§ †‡ SABRINA DELAUNAY§§ and ISABELLE BADENHAUSSER* †‡ *INRA, USC1339 (CEBC-CNRS), Villiers en Bois F-79360, France, †CNRS, UMR 7372 CEBC – Université de La Rochelle, Villiers en Bois F-79360, France, †‡LTER, ZA Plaine & Val de Sèvre, CNRS-CEBC, Villiers en Bois F-79360, France, §INRA, UMR1355 Institut Sophia Agrobiotech, Sophia Antipolis BP167-06903, France, ¶UMR Institute Sophia Agrobiotech, Université de Nice Sophia Antipolis, Sophia Antipolis BP 167-06903, France, **CNRS, UMR 7254 Institute Sophia Agrobiotech, Sophia Antipolis BP 167-06903, France, ††CIRAD, UMR CBGP, Montpellier F-34398, France, §§INRA, UMR CBGP, Montpellier F-34398, France, §§INRA, URP3F, site du chène, Lusignan 86600, France

Abstract

Dispersal may be strongly influenced by landscape and habitat characteristics that could either enhance or restrict movements of organisms. Therefore, spatial heterogeneity in landscape structure could influence gene flow and the spatial structure of populations. In the past decades, agricultural intensification has led to the reduction in grassland surfaces, their fragmentation and intensification. As these changes are not homogeneously distributed in landscapes, they have resulted in spatial heterogeneity with generally less intensified hedged farmland areas remaining alongside streams and rivers. In this study, we assessed spatial pattern of abundance and population genetic structure of a flightless grasshopper species, *Pezotettix giornae*, based on the surveys of 363 grasslands in a 430-km² agricultural landscape of western France. Data were analysed using geostatistics and landscape genetics based on microsatellites markers and computer simulations. Results suggested that small-scale intense dispersal allows this species to survive in intensive agricultural landscapes. A complex spatial genetic structure related to landscape and habitat characteristics was also detected. Two *P. giornae* genetic clusters bisected by a linear hedged farmland were inferred from clustering analyses. This linear hedged farmland was characterized by high hedgerow and grassland density as well as higher grassland temporal stability that were suspected to slow down dispersal. Computer simulations demonstrated that a linear-shaped landscape feature limiting dispersal could be detected as a barrier to gene flow and generate the observed genetic pattern. This study illustrates the relevance of using computer simulations to test hypotheses in landscape genetics studies.

Keywords: clustering methods, computer simulations, dispersal, landscape genetics, microsatellites, *Pezotettix giornae*

Received 15 October 2013; revision received 7 March 2015; accepted 10 March 2015

Introduction

Landscape and habitat characteristics can either restrict or enhance movements of individuals and consequently the extent of connectivity among populations (Taylor et al. 1993; Spear et al. 2010; Pflüger & Balkenhol 2014). Establishing how individuals are distributed over space and population spatial genetic structure in relation to the landscape and habitat characteristics may be critically important for management and conservation decisions (Palsboll et al. 2007). For instance, habitat loss and
fragmentation could lead to small and isolated populations that suffer from reduced genetic diversity (Frankham 2005) and are subject to higher extinction rates. Fragmentation is thus recognized to be one of the major threats to biodiversity (Fahrig 2003). This is particularly true in modern agricultural landscapes where intensification has caused loss of perennial habitats and their connection in space (Benton et al. 2003). Furthermore, farming practices have become increasingly intensive, even in perennial habitats, with an increasing proportion of artificial and temporary grasslands managed to ensure maximal yield of forage (Robinson & Sutherland 2002). This has resulted in strong declines of grassland invertebrates like grasshoppers (Barker 2004), leading to cascade effects at higher trophic levels among their predators (Vickery et al. 2001).

In agricultural landscapes, species persistence should be ensured by metapopulation dynamics in which processes of dispersal and colonization are crucial (Hanski 1998). Dispersal is a key factor of population dynamics and evolution in such disturbed environments, facilitating movement to favourable habitats and ensuring a global persistence despite local and recurrent extinctions (Ronc 2007). Theoretical and empirical studies demonstrated that less persistent, disturbed habitats promote animal movement capacities (Denno et al. 1996, 2001; Travis & Dytham 1999) and dispersal behaviour (Gauffre et al. 2009). Intense dispersal would result in genetic homogeneity at a large scale (Gauffre et al. 2008), but as dispersal is generally limited, genetic variation is expected to change gradually with distance (isolation by distance, IBD, Wright 1943; Rouset 1997) and animal distributions are expected to be spatially autocorrelated (Dormann et al. 2007). However, as landscape structure is rarely monotonous in space, gene flow is also modulated by physical and environmental heterogeneities in the landscape (isolation by resistance, IBR, McRae 2006).

The primary goal of landscape genetics is to relate landscape features directly to genetic structure and gene flow (Manel et al. 2003; Storfer et al. 2007; Holderegger & Wagner 2008). For example, identifying genetic disjunctions can elucidate the underlying landscape features and ecological processes promoting differentiation. Bayesian approaches of individual clustering (Guillot et al. 2005a; Francois et al. 2006; Chen et al. 2007) have proven to be efficient to detect genetic disjunctions (Blair et al. 2012), which are generally related to physical barriers such as roads or rivers (Holderegger & Di Guilio 2010; Pérez-Espona et al. 2012). However, in an IBR context, one could expect discontinuous features of landscape resistance to limit gene flow and promote genetic disjunctions, as recently evidenced in a long-winged grasshopper species (Blanchet et al. 2012). In complex landscape where IBD and IBR may occur simultaneously, computer simulations can be used to identify the processes driving population genetic patterns and investigate how reliably the effect of landscape resistance can be detected (Landguth et al. 2010). However, although the importance of using simulation in landscape genetics has been emphasized in a number of recent articles (Balkenhol et al. 2009; Epperson et al. 2010; Balkenhol & Landguth 2011), the effect of landscape resistance on Bayesian clustering methods has not been explored so far.

In this study, we investigated spatial patterns of a small-sized grasshopper, *Pezotettix giornae* (Acrididae: Catantopinae), across a 430-km² agricultural landscape in western France where agricultural intensification substantially reduced the proportion and durability of grasslands in the past 50 years. This grasshopper species has a single generation each year (Richards et al. 1954) and is specialized on grassland and herbaceous habitats (Uvarov 1977). Moreover, this species has wings unfit for flight and thus, its ability to disperse long distances is supposed to be weak, with individuals moving only by walk and jumps. *P. giornae* dispersal behaviour is unknown and, more broadly, little knowledge is available on its biology (Baur et al. 2006). We hypothesize that, although dispersal distances must be short, *P. giornae* dispersal behaviour (i.e. its propensity to leave a habitat patch) should be intense to ensure its persistence in areas where agricultural intensification reduced grassland structure and durability. We predict that spatial autocorrelation in *P. giornae* distribution in the landscape results mainly from dispersal processes and is constrained by habitat distribution resulting in strong- and short-range spatial autocorrelation of abundance as well as a genetic pattern of IBD. In addition, we could expect dispersal and genetic structure to be influenced by spatial heterogeneity in landscape structure and habitat characteristics. More specifically, we suspect a linear hedged farmland and a motorway, both bisecting our study area to affect *P. giornae* movement and geneflow patterns. Consequently, we applied a geostatistical approach to provide quantitative statements about spatial autocorrelation patterns in abundance and allele frequencies in *P. giornae* and we used Bayesian clustering model and computer simulations to address the following questions: (i) Does spatial autocorrelation in abundance and IBD exist in *P. giornae*, what are their strengths and at what scales do they occur? (ii) Is *P. giornae* spatial genetic structure shaped by heterogeneities in landscape structure? (iii) Under which circumstances does spatial variation in landscape resistance and IBD lead to the detection of genetic disjunction by Bayesian approaches of individual clustering?
Materials and methods

Study area and landscape characteristics

The study area, the Long-Term Ecological Research Zone Atelier Plaine & Val de Sèvre, is located in central western France (46°23′N, 0°41′W; see Fig. 1). It is a farmland area of 430 km² containing more than 12 000 fields (mean area ± SD: 2.94 ha ± 3.15) mostly dedicated to cereal crop production. Since 1995, land use has been recorded annually for each field and mapped with a Geographical Information System (ArcGis 9.2 – ESRI Redlands, CA, USA). In 2010, grassland surfaces represented 12.85% of the study area and included alfalfa (3.12% of total land use) and meadows (temporary or permanent sown with grasses or resulting from Fig. 1 LTER ‘Plaine et Val de Sèvre’, located in the region Poitou Charentes, western France. Dark grey zones indicate permanent unfavourable habitats (woodland and build-up); light grey, annual crops; brown, woodlands; and green, grassland (meadows and alfalfa plots). The cross-hatched area indicates the spatial location of the linear hedged farmland, and the black line indicates the motorway path. Squares indicate the position of surveyed plots, and blacks squares indicate the 190 plots where the 377 individuals of the genetic survey were caught (colour online).
spontaneous flora, 9.73% of total land use). The study site is bordered by the Poitevin marshes, the city of Niort, and the Sèvre and Lambon rivers on its northern side and by the Belle and Boutonne rivers and a dense network of streams and hedged farmlands on its eastern and south-eastern sides (Fig. 1). Conversely, the landscape on the western side of the study site is a continuous and open agricultural area without any obvious physical boundary over dozens of km (Fig. 1). The study area is crossed by a motorway built in 1981 and by a linear hedged farmland associated with a temporary stream (La Guirande) (Fig. 1).

To identify landscape and habitat characteristics potentially influencing *P. giornae* dispersal, we calculated the proportion of meadows, the proportion of temporary (<2 years) and permanent grasslands (>5 years), the average distance between a grassland and its 10 closest grasslands and a hedgerow index (as the total length of hedges over the surface, m/km²) in the linear hedged farmland and in its western and eastern sides.

**Pezotettix giornae surveys and genetic samplings**

In 2010, 363 grasslands were sampled from July 19th to September 3rd (i.e. during the peak of adult density, Badenhausser et al. 2009) using four different sampling schemes that shared the same sampling method but differed in the type of grasslands targeted, the procedure used to select the fields, the number of sampled fields and the sampling date (Table S1, Supporting information). Sampling was realized by means of trapping with a 1-m² square cage sampler (Badenhausser et al. 2009). It was thrown haphazardly 10 times in each grassland field, and all individuals of *P. giornae* caught in the cage sampler were counted in the field. One adult individual of each sex was collected per field plot for genetic analyses. A total of 377 individuals were sampled in 190 field plots for genetic analyses (Fig. 1). Hence, in half of the surveyed fields, no grasshoppers were trapped (Table 1). Each collected individual was directly put in a tube containing alcohol for later DNA extraction. The coordinates of the grassland field centroid were attributed to the individuals.

**Analyses of Pezotettix giornae abundances and spatial autocorrelation patterns**

First, we tested the effect of grassland characteristics and landscape structure at large scale on *P. giornae* abundance, using generalized linear model (GLM) with Poisson distribution in R 3.0.2 (R Development Core Team 2013). The model included the location of the surveyed fields (in the linear hedged farmland, in its western side and in its eastern side), environmental variables which may impact grasshopper abundances such as the type of surveyed grassland (meadow or alfalfa) and its age (1, 2, 3, 4, 5 and >5 years), and sampling variables, that is the sampling scheme and date. Quantitative explanatory variables were standardized. Interaction terms between (i) the sampling scheme and date and (ii) the grassland type and age were included in the model. The response variable was the sum of *P. giornae* counts over the ten 1-m² replicates per grassland field. A backward stepwise selection procedure was used with type II analysis of variance and likelihood-ratio chi-square calculated using the package car (Fox 2008).

Second, we used a geostatistical approach to analyse the spatial autocorrelation structure of *P. giornae* abundance, based on the residuals of the previous GLM model. The semivariogram describes the spatial

---

**Table 1** Landscape characteristics and *Pezotettix giornae* abundance in (i) the western side, (ii) within and (iii) the eastern side of the linear hedged farmland. Total surface (*S*), proportion of meadow over total surface (*P. Meadow*), proportion of temporary and permanent grasslands (*P. grassland* <2 years and *P. grassland* >5 years, respectively), average distance (SD) of the 10 closest grasslands and hedgerow index are provided for each zone. Number of grassland fields sampled (*N*), frequency of occurrence as the percentage of sampled fields where *P. giornae* was recorded (*O*) and mean and standard error (SE) of *P. giornae* density (*D* = number of grasshoppers/10 m²) are reported for each zone.

<table>
<thead>
<tr>
<th></th>
<th>Western side</th>
<th>Linear hedged farmland</th>
<th>Eastern side</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>S</em> (km²)</td>
<td>271.39</td>
<td>48.71</td>
<td>107.37</td>
</tr>
<tr>
<td><em>P. meadow</em> (%)</td>
<td>11.57%</td>
<td>24.44%</td>
<td>8.68%</td>
</tr>
<tr>
<td><em>P. grassland</em> &lt;2 years</td>
<td>42.61</td>
<td>32.57</td>
<td>52.1</td>
</tr>
<tr>
<td><em>P. grassland</em> &gt;5 years</td>
<td>31.7</td>
<td>45.04</td>
<td>22.9</td>
</tr>
<tr>
<td>Distance (SD) 10 closest grassland (m)</td>
<td>338.9 (171.5)</td>
<td>242.2 (95.5)</td>
<td>443.5 (226.7)</td>
</tr>
<tr>
<td>Hedgerow index</td>
<td>2931.4</td>
<td>9303.8</td>
<td>2014.9</td>
</tr>
<tr>
<td><em>N</em></td>
<td>253</td>
<td>54</td>
<td>56</td>
</tr>
<tr>
<td><em>O</em></td>
<td>49.8</td>
<td>50</td>
<td>46.4</td>
</tr>
<tr>
<td><em>D</em> (SE)</td>
<td>3.92 (0.54)</td>
<td>2.79 (0.60)</td>
<td>3.32 (1.02)</td>
</tr>
</tbody>
</table>

© 2015 John Wiley & Sons Ltd
contiguity of this variable and was calculated by averaging the empirical variogram values $\gamma (h)$ for a range of distance interval $h$ (Wackernagel 1995):

$$\gamma (h) = \frac{1}{2N(h)} \sum_{x=1}^{N(h)} [z(x) - z(x + h)]^2,$$

where $z(x)$ is the value of the variable located at point $x$. After examination of the empirical variograms, a spherical model appeared the most appropriate to model it. Following Maestre et al. (2005), the parameters were estimated by minimizing the mean squared error between the variogram model and the sample variogram data weighted by the number of pairs per distance lag (minimum number of pairs set to 30). The parameters of interest for the spherical model are the range (distance beyond which the semivariance reaches a plateau and samples become spatially independent), the nugget ($C_0$, the variogram intercept, usually interpreted as local random effects and measurement error) and the sill ($C$, the value of the plateau). The ratio $C_0/C$ can be calculated to indicate the strength of spatial autocorrelation in the data (Rong et al. 2007). The smaller this ratio is, the more autocorrelated the data tend to be at short range. In the case of a constant model (pure nugget effect), this ratio is 1. Both $C_0$ and $C$ were taken from the variogram model parameters. Distance lags started at 0 with 200-m increments. Analyses were conducted in R 3.0.2 (R Development Core Team 2013) using the gstat 1.7-4 package (Ribeiro & Diggle 2001).

**DNA extraction and genotyping**

DNA extraction was performed from a piece of the posterior leg femur (2–4 mm), using the chloroform–octanol method developed by the CIMMYT laboratory (CIMMYT 2005). To improve the lysis, we added 20 mg/mL of proteinase K in the buffer at the second stage.

A set of *P. giornae*-specific microsatellite markers was developed for this study. The library was carried out from a pool of ten individuals by Genoscreen (Lille, France) using a recently developed procedure (Malausa et al. 2011) combining DNA enrichment and high-throughput pyrosequencing (GS-FLX, Roche Diagnostics). We chose 40 markers, based on the number of repeated motifs and the size of amplified fragments, from a total of 271 microsatellite marker sequences. The variability of these 40 loci was screened on 16 individuals using an automatic sequencer (Li-Cor, global IR2 system double laser sequencing). Eleven loci were selected (Table 2) for this study, the rest were unsuitable due to apparently high incidence of null alleles, ambiguous PCR products and/or failed reactions.

**Table 2 Characteristics of the 11 microsatellite loci isolated in *Pezotettix. giornae*. Raw reads from the GS-FLX Roche sequencing have been posted on Dryad (doi:10.5061/dryad.88b12)**

<table>
<thead>
<tr>
<th>Locus</th>
<th>Primer sequence (5’–3’)</th>
<th>Repeat motif</th>
<th>Size of cloned allele (bp)</th>
</tr>
</thead>
</table>
| Pezo_3  | F: AGACGCGATTTTGGAGGCGACAT (TTG)$_{13}$
R: CTCGAATGAACCCAGCTCTTG |
| Pezo_8  | F: TTCTCTTCACATAGAATTCAGC
R: TGTCACACTCTCCTG |
| Pezo_9  | F: TTATAGCCACCTCCGTGAC
R: CTGATCACAAGAACCTCTG |
| Pezo_13 | F: CAGCTGGAAGGGCTTTCTCTT
R: TTACCTGATGGAGAGGAGAC |
| Pezo_19 | F: GCCAGCTTGCTATCCATTTA
R: GGCATCAACCCCTCCAAGA |
| Pezo_24 | F: GGCTGTCACAGGATAGG\textsuperscript{AGAGG}
R: TTCAAAATTCCTGAGTCCG |
| Pezo_27 | F: GGAGGGCGGAGCCACGTCTCTG
R: TTCGGTCTAGATTAGTGGTTG |
| Pezo_29 | F: TTATCACTCACAGAAGGAAAGCA
R: GGCTTACAGAGCAGTTAGG |
| Pezo_31 | F: TTGTTTATGCGGCTCGGCG
R: TCTTCAATATTTGTACGGTC |
| Pezo_32 | F: AACCCGTTAGCCGCTTGAG
R: CCTCGGGTCTGAAAACAAAA |
| Pezo_37 | F: CCGCTAAATTTGCTCGAAGATC
R: AGATAAAATGGCGCTTG |

© 2015 John Wiley & Sons Ltd
crosatellite markers were amplified by PCR in 10 μL of a reaction solution containing 2.5 μL of DNA, 5 μL of Master Mix (Ampli Taq Gold 360 Master Mix), 0.5 μL of Primer Mix (10 μM of forward primer; 1 μM of M13-labelled reverse primer), 0.05 μL of M13-labelled reverse primer (100 μm) and 1.95 μL of distilled water. The PCR was performed using a denaturation step of 5 min at 95 °C, followed by 7 cycles of 30 s at 95 °C, 30 s at 62 °C with a decrease of 1 °C per cycle and 30 s at 72 °C, then by 30 cycles of 30 s at 95 °C, 30 s at 55 °C and 30 s at 72 °C and by eight cycles of 30 s at 95 °C, 30 s at 56 °C and 30 s at 72 °C and finally by 5 min at 72 °C. The PCR products were genotyped in a genotyper ABI Prism® 3730 Genetic Analyser (Applied Biosystems), and the genotypes were determined using GENEMAPPER™ software (version 3.7, Applied Biosystems).

**Basic population genetics analyses**

First, on the whole data set, the number of alleles per locus and the frequency of the major allele were calculated. We also tested for deviation from Hardy–Weinberg equilibrium (HWE) using the exact test implemented in GENEPOP 4.1 (Rousset 2008) for each locus and globally. Observed and expected heterozygosities (Ho and He, Nei 1987), and Weir & Cockerham’s estimate of Fs (Weir & Cockerham 1984) were also calculated using GENEPOP 4.1. We tested genotypic linkage disequilibrium (LD) for each locus and the frequency of the major allele were calculated allele frequency model, included the presence of null alleles at microsatellite loci (Guillot et al. 2008) into the genetic distances (Nei et al. 2005a) into the genetic distances (Nei et al. 2005a, b). Ten independent runs with 1 000 000 MCMC iterations (thinning = 100) were performed allowing K (the number of genetic clusters) to vary from 1 to 10. The potential error for spatial coordinates was fixed at 50 m. The number of clusters inferred, K*, was determined from the modal value of the run with the highest posterior probability. The β-prior distribution included in the estimation of the drift coefficient was set to the default value [β(2,20)] corresponding to a medium-sized differentiation. However, because the inference of K using the model of correlated allele frequencies is sensitive to the choice of the β-distribution parameters (Guillot 2008), we also carried out inferences with 2 other settings for the β-distribution: β (1,1) (uninformative) and β(1,100) (low differentiation) (Guillot 2008). For each run, we reported K*, the proportion of saved iterations along the MCMC chain with K = K*, the mean log-posterior density over saved iterations and, when K* ≥ 2, the proportion of individuals with a membership probability 1.5 times greater than the equiprobability (i.e. 1/K*). We finally checked visually for the consistency of results across the 10 runs. In a second step, for each set of β-prior setting, a long run was performed with K fixed to K* and 2 000 000 MCMC iterations (thinning = 100) to determine thoroughly the posterior probability of cluster membership of each individual. If clusters were detected, genetic differentiation among them was tested using G tests and quantified by calculating FST following Weir & Cockerham (1984) and using the ENA correction method for data set harbouring null alleles implemented in FreeNA (Chapuis & Estoup 2007).

**Isolation by distance and spatial autocorrelation**

We determined whether dispersal was limited by distance by testing the correlation between the pairwise genetic distances (r, Watts et al. 2007) and the logarithm of the geographical distances using GENEPOP 4.1. Pairs of individuals from the same field plot were excluded, and significance of the correlation was tested using Mantel test. First, the analysis was performed on the whole data set. Second, to investigate the linear hedged farmland influence on IBD, we also analysed IBD according to cluster membership. Each inferred cluster was identified as an independent spatial category, and computations were then restricted to the pairs of individuals belonging to a same cluster only or, alternatively, to different clusters only.

Finally, the spatial pattern of genetic variation was investigated using spatial autocorrelation analyses that assess the genetic similarity between pairs of individuals at different distance classes, thus providing results on the scale at which spatial patterns occur. Using SPAGEDI
Assessing the effect of spatial heterogeneity in dispersal on the clustering analysis

We evaluated the effect of spatial heterogeneity in dispersal distribution on our inferences with a focus on the GENELAND clustering analysis. We simulated a single continuous population under isolation by distance using a generation-by-generation coalescent algorithm (Leblois et al. 2003) implemented in IBDSIM v1.3 (Leblois et al. 2009; Leblois et al. in prep). Simulated data sets had 200 diploid individuals genotyped at 10 independent microsatellite loci, reflecting the sample values in our real P. giornae data set. Mutations of the microsatellite loci followed a symmetric generalized stepwise-mutation model with a variance equal to 0.36 (Estoup et al. 2001) and a maximum range of allelic states of 40. We fixed the mutation rate to $10^{-8}$, which resulted in ranges of expected heterozygosity values from 0.25 to 0.51 and of allelic diversity values from 2.6 to 4.1 that both comprise the values observed at our microsatellite markers (see Tables 3 and 4).

We mimicked the spatial context of our real data set by simulating a $40 \times 10$ lattice with five individuals per node and subsampling a single individual per node for the right half of the lattice only (i.e. from the 21th node) (Fig. S1, Supporting information). With this spatial model, we assumed that rivers that border all sides but the western side of our real study area correspond to the geographical limits of the population (Fig. 1). To create dispersal heterogeneity in space, the lattice was bisected at its 31th node from north to south by a $5 \times 10$ zone, which corresponds to the thin hedged farmland located in the western side of our sampling area (see Figs 1 and 2). This zone was parameterized with more stringent values for IBD model parameters (hereafter referred to as lower dispersal zone). We considered a geometric dispersal distribution, with absorbing boundaries and a maximal dispersal distance of 5 nodes, and contrasted values of the emigration rates and shape parameters ($\epsilon$ and $g$), in and out of the lower dispersal zone. Two levels of contrast in dispersal parameters were simulated (see sim1 and sim2 in Table 3 and Fig. S2, Supporting information). The strongest contrast (i.e. sim2) generated 100% of significant correlations between individual pairwise genetic distances $c_e$ (Watts et al. 2007) and the logarithm of geographical distances, with regression slopes ranging from 0.001 to 0.042, which is compatible with the value of 0.015 inferred from the real data set (Table 3). Keeping the IBD model parameter values of sim2, we then simulated a lower dispersal zone with either a larger width (i.e. $10 \times 10$ nodes in sim3) or a central position in the sampling area (in sim4) (Table 3 and Fig. S1, Supporting information). These two sets of simulations may generate different results from the realistic simulation scenario sim2 that featured a linear hedged farmland located asymmetrically in the eastern half of the sampling area.

For each of the four parameter sets, we simulated 30 data sets for each of which we performed five runs of GENELAND using the same parameters settings as for the real data set and a $\beta$-prior distribution set to $\beta(2,20)$. For each data set, we reported the inferred number of clusters (K*), the percentage of individuals that had a membership probability 1.5 times greater than the equiprobability to belong to a given cluster and the values for the estimators of $F_{IS}$ within clusters and $F_{ST}$ between clusters when applicable.

Results

Landscape characteristics and demographic population structuring

The linear hedged farmland appeared the more favourable zone for the grasshopper with a higher proportion of land surface covered by meadows and a larger proportion of permanent grasslands (Table 1). On the other hand, the higher quantity of hedgerow in the linear hedged farmland (Table 1) could limit dispersal. The western side appeared slightly more favourable than the eastern side (Table 1). P. giornae was present in 49.3% of sampled grasslands, and overall, P. giornae abundance was low with an average $3.66 \pm 0.42$ (SE) individuals/10 m$^2$ (details per zone in Table 1). No significant differences in grasshopper density (location effect in GLM: chi-square = 5.40, $P$\text{-}value = 0.06, Table S2, Supporting information) was detected between the grasslands within the linear hedged farmland and the grasslands from the two other zones. Grassland age had a positive
Table 3 Effects of spatial heterogeneity in dispersal on the Bayesian clustering method. For each set of parameters are indicated the parameter settings, average summary statistics over the simulated data sets and GENELAND results including mean values of \( F \)-indices and the proportion of individuals assigned in each cluster according to their position in the sampling area. \( c \): parameter controlling total emigration; \( g \): shape parameter of the geometric distribution; \( A \): allelic diversity; \( H_E \): expected heterozygosity; \( K^* \): modal estimates of the number of clusters; and \( % K = K^* \): proportion of simulated data sets with \( K = K^* \). For convenience, cluster A always refers to the left side, cluster B to the right side and cluster C to the centre of the sampling area.

<table>
<thead>
<tr>
<th>Parameter settings</th>
<th>sim1</th>
<th>sim2</th>
<th>sim3</th>
<th>sim4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lattice size</td>
<td>40 x 10</td>
<td>40 x 10</td>
<td>45 x 10</td>
<td>45 x 10</td>
</tr>
<tr>
<td>Sampling area size</td>
<td>20 x 10*</td>
<td>20 x 10*</td>
<td>25 x 10*</td>
<td>25 x 10*</td>
</tr>
<tr>
<td>LDZ size</td>
<td>5 x 10</td>
<td>5 x 10</td>
<td>10 x 10</td>
<td>5 x 10</td>
</tr>
<tr>
<td>LDZ position</td>
<td>Nodes 31 to 35</td>
<td>Nodes 31 to 35</td>
<td>Nodes 31 to 40</td>
<td>Nodes 31 to 35</td>
</tr>
<tr>
<td>( e/g ) in LDZ</td>
<td>0.4/0.675</td>
<td>0.2/0.5</td>
<td>0.2/0.5</td>
<td>0.2/0.5</td>
</tr>
<tr>
<td>( e/g ) in HDZ</td>
<td>0.6/0.75</td>
<td>0.8/0.75</td>
<td>0.8/0.75</td>
<td>0.8/0.75</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Summary statistics</th>
<th>sim1</th>
<th>sim2</th>
<th>sim3</th>
<th>sim4</th>
</tr>
</thead>
<tbody>
<tr>
<td>( A )</td>
<td>3.5</td>
<td>3.3</td>
<td>3.4</td>
<td>3.4</td>
</tr>
<tr>
<td>( H_E )</td>
<td>0.38</td>
<td>0.37</td>
<td>0.39</td>
<td>0.37</td>
</tr>
<tr>
<td>IBD slope</td>
<td>0.006</td>
<td>0.012</td>
<td>0.045</td>
<td>0.019</td>
</tr>
<tr>
<td>% Significant IBD tests</td>
<td>89</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>GENELAND results†</th>
<th>sim1</th>
<th>sim2</th>
<th>sim3</th>
<th>sim4</th>
</tr>
</thead>
<tbody>
<tr>
<td>( K^* )</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>( % K = K^* )</td>
<td>81</td>
<td>82</td>
<td>54</td>
<td>85</td>
</tr>
<tr>
<td>( F_{ST} A-B )</td>
<td>–</td>
<td>0.024</td>
<td>0.097</td>
<td>0.031</td>
</tr>
<tr>
<td>( F_{ST} A-C )</td>
<td>–</td>
<td>–</td>
<td>0.045</td>
<td>–</td>
</tr>
<tr>
<td>( F_{ST} B-C )</td>
<td>–</td>
<td>–</td>
<td>0.05</td>
<td>–</td>
</tr>
<tr>
<td>( F_{IS} ) cluster A</td>
<td>–0.001</td>
<td>0</td>
<td>0.007</td>
<td>0.014</td>
</tr>
<tr>
<td>( F_{IS} ) cluster B</td>
<td>–</td>
<td>0.015</td>
<td>0.027</td>
<td>0.01</td>
</tr>
<tr>
<td>( F_{IS} ) cluster C</td>
<td>–</td>
<td>–</td>
<td>0.072</td>
<td>–</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Zone</th>
<th>HDZ (left)</th>
<th>HDZ (right)</th>
<th>HDZ (left)</th>
<th>HDZ (right)</th>
<th>HDZ (left)</th>
<th>HDZ (right)</th>
<th>HDZ (left)</th>
<th>HDZ (right)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LDZ (left)</td>
<td>LDZ (right)</td>
<td>LDZ (left)</td>
<td>LDZ (right)</td>
<td>LDZ (left)</td>
<td>LDZ (right)</td>
<td>LDZ (left)</td>
<td>LDZ (right)</td>
</tr>
<tr>
<td>% Inds in cluster A</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>97.5</td>
<td>38.4</td>
<td>0.5</td>
<td>100</td>
<td>37.5</td>
</tr>
<tr>
<td>% Inds in cluster B</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>2.5</td>
<td>61.6</td>
<td>99.5</td>
<td>0</td>
<td>37.8</td>
</tr>
<tr>
<td>% Inds in cluster C</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>0</td>
<td>100</td>
<td>24.8</td>
</tr>
</tbody>
</table>

LDZ, lower dispersal zone; HDZ, higher dispersal zone (i.e. the rest of the lattice).

*Right half of the lattice.
†Results based on datasets with \( K = K^* \).
and significant linear effect on *P. giornae* abundance (chi-square = 24.65, \( P (>\text{chi-square}) < 0.001; \) estimate = 0.49 ± 0.07 (SE); Table S2, Supporting information). Consistently, *P. giornae* was detected in only 31.1% of 1-year grasslands, while it was detected in 61.5% of permanent grasslands (5 years and more). Grassland type (meadow or alfalfa) had no effect on *P. giornae* abundances when corrected by grassland age (chi-square = 0.44, \( P (>\text{chi-square}) = 0.51; \) Table S2, Supporting information).

Based on GLM residuals, empirical and modelled variograms revealed a range of spatial autocorrelations of 1034 m (Fig. 3a), indicating that *P. giornae* densities were spatially positively correlated up to 1034 m. Strength of spatial autocorrelation at small scales was very high as shown by the small value of 0.059 of the normalized

### Table 4

<table>
<thead>
<tr>
<th>Locus</th>
<th>A</th>
<th>FreqA</th>
<th>He</th>
<th>Ho</th>
<th>( F_{IS} ) (sign)</th>
<th>Pmg</th>
<th>Na</th>
<th>Freq Na</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pezo_3</td>
<td>3</td>
<td>0.92</td>
<td>0.152</td>
<td>0.085</td>
<td>0.439*</td>
<td>0</td>
<td>+</td>
<td>0.12</td>
</tr>
<tr>
<td>Pezo_8</td>
<td>2</td>
<td>0.7</td>
<td>0.417</td>
<td>0.592</td>
<td>-0.419*</td>
<td>0</td>
<td>-</td>
<td>0</td>
</tr>
<tr>
<td>Pezo_9</td>
<td>3</td>
<td>0.86</td>
<td>0.241</td>
<td>0.111</td>
<td>0.538*</td>
<td>0</td>
<td>+</td>
<td>0.13</td>
</tr>
<tr>
<td>Pezo_13</td>
<td>2</td>
<td>0.62</td>
<td>0.473</td>
<td>0.366</td>
<td>0.227*</td>
<td>0.013</td>
<td>+</td>
<td>0.12</td>
</tr>
<tr>
<td>Pezo_19</td>
<td>2</td>
<td>0.87</td>
<td>0.223</td>
<td>0.207</td>
<td>0.07**</td>
<td>0</td>
<td>-</td>
<td>0.046</td>
</tr>
<tr>
<td>Pezo_24</td>
<td>3</td>
<td>0.44</td>
<td>0.264</td>
<td>0.043</td>
<td>0.839*</td>
<td>0.003</td>
<td>+</td>
<td>0.221</td>
</tr>
<tr>
<td>Pezo_27</td>
<td>2</td>
<td>0.81</td>
<td>0.307</td>
<td>0.271</td>
<td>0.116*</td>
<td>0.003</td>
<td>+</td>
<td>0.055</td>
</tr>
<tr>
<td>Pezo_29</td>
<td>6</td>
<td>0.82</td>
<td>0.304</td>
<td>0.311</td>
<td>-0.024**</td>
<td>0.003</td>
<td>-</td>
<td>0.024</td>
</tr>
<tr>
<td>Pezo_31</td>
<td>3</td>
<td>0.6</td>
<td>0.482</td>
<td>0.264</td>
<td>0.453*</td>
<td>0.032</td>
<td>+</td>
<td>0.219</td>
</tr>
<tr>
<td>Pezo_32</td>
<td>3</td>
<td>0.56</td>
<td>0.494</td>
<td>0.501</td>
<td>-0.016**</td>
<td>0</td>
<td>-</td>
<td>0</td>
</tr>
<tr>
<td>Pezo_37</td>
<td>2</td>
<td>0.54</td>
<td>0.497</td>
<td>0.449</td>
<td>0.096**</td>
<td>0</td>
<td>-</td>
<td>0.051</td>
</tr>
</tbody>
</table>

© 2015 John Wiley & Sons Ltd
expression $C_0/C$ which indicated that a large fraction of spatial dependence was due to spatial autocorrelation over a short range.

**Microsatellite characteristics**

Genotypic data revealed low levels of genetic diversity at all 11 microsatellite loci (Table 4). The number of alleles per locus ranged between two and six. Expected heterozygosity for each locus ranged from 0.15 to 0.49, with an average value of 0.35 for all loci. Observed heterozygosity ranged from 0.04 to 0.59, with an average value of 0.29 for all loci. There was a significant overall heterozygosity deficit in the total sample ($P < 0.001$) with a high (0.169) and significant global $F_{IS}$ value ($P < 0.001$) that could be related to the presence of spatial structure as well as null alleles in our data. After FDR correction, we found a single pair of loci showing significant LD (Pezo_8 and Pezo_24). Null alleles were suspected to occur at six loci with estimated frequency ranging from 0 to 0.21 (Table 4). We took the problem of null alleles into account in our analyses of genetic population structure as described above. Locus Pezo_13 appeared homozygous for all males. Because grasshopper females are XX and males X0 (Hake & O’Connor 2008), Pezo_13 was probably located on the sexual chromosome. Consequently, this locus was excluded in the following analyses.

**Delimitation of genetic clusters**

With a medium-sized differentiation prior [i.e. β(2,20)], the correlated allele frequencies model implemented in **GENELAND** inferred two genetic clusters consistently across all ten runs with an average 96.66% of individuals assigned unambiguously to either cluster (Table S3, Supporting information). Results when setting an uninformative prior [i.e. β(1,1)] were very similar (Table S3, Supporting information), and assignment of individuals to one of the two clusters was fully identical among both sets of β-prior setting. The first cluster, hereafter named Western cluster, occupied more than half of the study area on the western side of the linear hedged farmland and had 280 individuals. The second cluster, named Eastern cluster, covered the eastern side of the study area and had 97 individuals (Fig. 4). A more complex picture emerged in the western side when inferences were performed with a low differentiation prior to drift parameters [i.e. β(1,100)]. Whereas the Eastern cluster remained almost identical to the one inferred with previous prior settings, the former Western cluster was split into two distinct clusters by the motorway consistently across all ten runs. However, only individuals belonging to the Eastern cluster were unambiguously assigned (membership probability >0.5). Hence, the subdivision in the western side was poorly supported and not considered in the following analyses.
The Eastern cluster covered a large section of the linear hedged farmland, with the assignment of 39 of the 56 caught individuals (i.e. 69.6%). However, only 83.9% of the individuals from the hedged farmland were unambiguously assigned to their cluster (membership probability > 0.75), while the proportion of unambiguous assignments was 96.6% and 98.1% for individuals originating from the eastern and western sides of the linear hedged farmland, respectively (all were assigned to the corresponding cluster). Therefore, the linear hedged farmland corresponds to the delineation between the two main inferred genetic clusters.

The level of genetic differentiation between the two clusters inferred from the GENELAND analysis was relatively high ($F_{ST} = 0.045, P < 0.001$). We found specific alleles in the two clusters at six of the microsatellite markers (Table S5, Supporting information). Both clusters were still characterized by a significant deficit in heterozygotes: $H_o = 0.305, H_e = 0.351, F_{IS} = 0.131$ (exact test, $P < 0.001$) in the Western cluster and $H_o = 0.223, H_e = 0.277, F_{IS} = 0.195$ (exact test, $P < 0.001$) in the Eastern cluster (see Table S5, Supporting information, for locus by locus estimates). These values of expected and observed heterozygosities did not differ significantly between clusters (Wilcoxon test, $P = 0.32$ and $P = 0.17$ for $H_o$ and $H_e$, respectively). After FDR correction, we found no significant LD in the two inferred clusters.

Isolation by distance and spatial autocorrelation

A positive and significant linear relationship between $e_r$ and the logarithm of spatial distance was found in the whole data set (slope = 0.015, $P \leq 0.001$). IBD analysis considering all pairs of individuals belonging to the same cluster only (i.e. distances estimated within each cluster were merged in a single data set) revealed a positive and significant correlation, with a lower slope value (slope = 0.006, $P = 0.004$). The linear IBD relationship was greater when considering only the pairs of individuals belonging to distinct clusters (slope = 0.016, $P = 0.018$), suggesting a higher resistance to gene flow associated with the linear hedged farmland.

Finally, the spatial genetic autocorrelation analysis performed on pairs of individuals belonging to a same cluster only displayed significant and positive kinship values for the two first distance classes (zero and 500 m) (Fig. 3b). Genetic similarities between individuals
decreased sharply between distance classes 0 and 1.5 km. For distances >3.5 km, we observed a tendency for a slight decrease in similarity with distance.

**Simulation results**

We used computer simulations to investigate whether a spatial heterogeneity in dispersal distribution could explain the inference of two genetic clusters with GENELAND. Results are summarized in Table 3 and Fig. 2, and detailed in supporting information (Table S4, Supporting information). We first used a spatial model that reflects the real north–south bisection of the eastern part of our study area by the linear hedged farmland with two levels of contrast in dispersal parameters (sim1 and sim2). When simulating a moderate contrast between the lower dispersal zone and the rest of the lattice, GENELAND detected a single genetic cluster in 80% of the simulated data sets (sim1). By contrast, two genetic clusters were detected in 82% of the simulated data sets with a stronger contrast in dispersal parameters between the lower dispersal zone and the rest of the lattice (sim2). The averages of population genetic summary statistics over sim2 replicates ($F_{ST} = 0.024$, IBD slope = 0.012) were close to the values obtained in the real data set ($F_{ST} = 0.045$, IBD slope = 0.015), with the exception of $F_{IS}$ statistics that was much lower in sim2 (0 and 0.015 for clusters A and B, respectively) than in the real data set (0.131 and 0.195 for the Western and Eastern clusters, respectively). Larger values of $F$-statistics in the real data set may result from the presence of null alleles in the *P. giorni* microsatellite loci (Chapuis & Estoup 2007). In this asymmetrical spatial model, 62% of the nodes from the lower dispersal zone were assigned to the Eastern cluster, which is close to the 69.6% in our real data set. On the contrary, when the lower dispersal zone was placed in the centre of the sampled area, its nodes were assigned equally to the Western and Eastern clusters (sim4). Finally, when the lower dispersal zone was two times wider, GENELAND inferred in 85% of the simulated data sets a third genetic cluster, which was rigorously located within this heterogeneous zone (sim3).

**Discussion**

**Small population size**

The genetic diversity characterizing the studied population (from two to six alleles per locus) was much lower than usually reported in Orthopteran insects (Chapuis et al. 2012). For example, six to 25 alleles were found in a *Locusta migratoria* population (Chapuis et al. 2005), 15–62 in a German *Chorthippus parallelus* population (Wiesner et al. 2011) and six to 29 alleles at six microsatellite loci in a small endemic population of *Prionotropis rhodanica* which is considered at the edge of extinction (Streiff et al. 2002). At all 11 microsatellites loci presented here, the frequency of the most frequent allele was > 0.5. Overall, these characteristics suggested very small effective population size, congruent with observed low densities (0.37 inds per m²) and the small proportion of habitats (i.e. grasslands) in the landscape (12.85%). In addition, we could hypothesize that only few individuals succeeded in reproduction given the level of spatio-temporal instability characterizing the agricultural mosaic. Indeed, the average duration of grasslands does not exceed three years in the study area that results in high local extinction rates. Alternatively, the low genetic diversity could be the consequence of a founder effect as the study site is close to the distribution limit of this Mediterranean species (Kruseman 1982). The comparison of genetic diversity with populations from the Mediterranean region would allow testing this hypothesis.

**Dispersal syndrome in intensive agricultural land**

The heterozygosity deficit measured in the whole data set could not be fully attributed to the presence of genetic clusters as there was also a significant heterozygosis deficit in the inferred clusters. This should be partly due to the presence of null alleles (Dakin & Avise 2004), known to be frequent in Orthoptera (Chapuis et al. 2005) and detected at six of the ten loci used in this study. IBD in the inferred clusters could also account for the observed deviation from Hardy–Weinberg expectation. The autocorrelation pattern, with pairs of individuals more related than expected under random distribution of genotypes until 500 m, confirmed the limited dispersal capacities of this flightless grasshopper species. Consistently, previous studies on flightless grasshoppers reported strong rates of genetic differentiation (Streiff et al. 2006) and limited movement distances of adults (Weyer et al. 2012). In phytophagous insects, a general relationship between mobility and gene flow has been documented with gene flow usually extensive and weakly declining with distance in highly mobile species while it was declining rapidly with distance to produce IBD in less mobile species (Peterson & Denno 1998). Interestingly, the geostatistical analysis provided results similar to the genetic autocorrelation (Fig. 3). The strength of spatial autocorrelation in *P. giorni* densities at short range was substantial, and densities were spatially correlated up to 1034 m. Altogether, results suggest intense and small-scale dispersal allowing the colonization and connection of habitat patches to ensure the persistence of *P. giorni* in highly disturbed environments. Densities and occurrence did not differ between grasslands in and out of the linear hedged farmland, while
grassland age had a positive effect on \textit{P. giornae} density (Badenhauser & Cordeau 2012).

Similar ranges of spatial autocorrelation observed in the same study area for other grasshoppers species (750 m for \textit{Calliptamus italicus}; and 1500 m for \textit{Gomphocerinae} subfamily; Badenhauser \textit{et al.} 2012) and of genetic autocorrelation in the common vole (Gaufrisse \textit{et al.} 2008) suggest that similarities in landscape influence co-occurring species in managed grassland habitats. This could reflect that a general set of rules may determine how different organisms respond to landscape (With 1994). We could thus postulate that in intensive agricultural areas, \textit{P. giornae} dispersal is enhanced by disturbance due to increasing grasslands turnover. As a long-term consequence, one could expect the selection of increasing dispersal capacities. Both theoretical (Olivieri \textit{et al.} 1995; Travis & Dytham 1999) and experimental (Friedenberg 2003) studies showed that habitat instability as well as spatial heterogeneity induces evolution towards more dispersal. Empirical studies comparing movement-related morphology among populations demonstrated that higher degree of habitat fragmentation and/or disturbance could select for more mobile ‘morphologies’ in insects (Denno et al. 2001; San Martin y Gomez & Van Dyck 2012).

\textbf{Consequences of landscape heterogeneity on dispersal and gene flow}

We found little support for genetic subdivision associated with the motorway bisecting the study area. However, in a previous study, computer simulations demonstrated that the motorway was too recent to be detected, at least for a species with large effective population sizes (see Gaufrisse \textit{et al.} 2008). By contrast, we detected unambiguously two genetic clusters separated by the linear hedged farmland crossing the study area. Because \textit{P. giornae} densities did not significantly differ between the linear hedged farmland and its western and eastern sides, this subdivision cannot result from density variations between the zones. By contrast, comparison of landscape characteristics revealed strong differences regarding two landscape elements: (i) hedgerows, where density was more than three times higher in the linear hedged farmland, and (ii) grasslands, which were two times more abundant in the linear hedged farmland. In addition, grasslands were dominated by permanent meadows in the linear hedged farmland, while they were more intensively managed in the rest of the study site with a more frequent turnover. Hedgerows may be difficult to cross for an apterous grasshopper and thus increase landscape resistance to dispersal. This has also been found in other nonflying arthropods inhabiting agricultural landscapes (e.g. carabid beetles, Mauremooto \textit{et al.} 1995) or even in flying insects such as Syrphidae (Wratten \textit{et al.} 2003). Second, we could postulate that the largest proportion of grasslands and their greater temporal stability in the linear hedged farmland could also increase landscape resistance and reduce gene flow. Such hypothesis follows the Circe principle recently proposed to explain how areas with more attractive habitats can waylay pollinators rather than facilitating their movement in fragmented landscapes (Bartomeus & Winfree 2011; Lander \textit{et al.} 2011). Because hedgerows and grasslands are strongly correlated in our landscape, we cannot separate the respective effect of hedgerows and grasslands on the limitation of \textit{P. giornae} dispersal in the linear hedged farmland.

Computer simulations clearly supported the hypothesis that a higher landscape resistance in the linear hedged farmland bisecting our sampling area could result in the inference of two distinct genetic clusters by Bayesian clustering approaches. Note that these results were obtained in the specific simulated landscape of an open western space and a closed eastern space. Interestingly, simulating the same spatial asymmetry as in our real landscape (sim2) led to results mirroring our observed data (e.g. a higher \textit{F}_{st} in the Eastern cluster and the assignation of individuals from the linear hedged farmland biased towards the Eastern cluster). The inference of two genetic clusters was associated with an equal assignation of individuals from the lower dispersal zone to the two peripheral clusters when border effects were minimized (sim4; i.e. by moving the lower dispersal zone from the eastern side to the centre of the sampling area). We also showed that if nonlinear (i.e. larger), the lower dispersal zone would have led to the detection of further genetic disjunctions that rigorously delineate the breaks in landscape resistance (sim3). In contrast, we found a single genetic cluster when simulating low contrast in dispersal parameters between the lower and higher dispersal zones (sim1). This is in agreement with previous simulation studies on the effect of landscape heterogeneity on dispersal which showed that the contrast between the resistance of the different landscape elements was required to be strong to be detected (Jaquiéry \textit{et al.} 2011; Cushman \textit{et al.} 2013). To conclude, an important point that arises from our simulation data is that strong spatial heterogeneity in isolation by distance (e.g. related to heterogeneous landscape features that act as resistance to movement of genes) results in signatures similar to discrete physical barriers (i.e. sharp genetic discontinuities).

\textbf{Conclusion}

In this study, we detected a complex spatial genetic pattern in a flightless grasshopper (\textit{P. giornae}) inhabiting
patchily distributed semipermanent habitats in an agricultural landscape likely resulting from spatially structured variations in landscape resistance. From analysis of empirical data, we anticipated that distribution of hedgerow barriers and favourable permanent grasslands modulated spatial variations in landscape resistance in this species. Computer simulations allowed us to confirm this hypothesis and evaluate the conditions under which spatial heterogeneity in dispersal correctly re-created the observed population genetic substructure (Landguth et al. 2010). We demonstrated that reduced dispersal in a thin zone in the middle of sampling area (i.e. the hedged farmland in our real landscape) created a genetic discontinuity that separated the two peripheral higher dispersal zones. This study is thus a stark illustration of a spatial heterogeneity in landscape resistance that acts similarly to a discrete barrier to gene flow. Our study emphasizes the importance of addressing landscape genetics processes in terms of isolation-by-differential resistance across continuous space (Cushman & Landguth 2010), rather than limiting investigation to the identification of discrete barrier features.

Acknowledgements

We thank Sylvain Piry, Gael Caro, Pascal Monestiez and Kevin Lerest for constructive comments on the analyses and Marilyn Roncoroni for help in trapping grasshoppers. Microsatellite genotyping was performed at GENTYANE platform. The project was supported by INRA-SPE.

References


CIMMYT (2005) Laboratory Protocols: CIMMYT Applied Molecular Genetics Laboratory, 3rd edn. CIMMYT, Mexico, D.F.


© 2015 John Wiley & Sons Ltd
Landscape influences grasshopper genetic patterns

San Martín y Gomez G, Van Dyck H (2012) Ecotypic differentiation between urban and rural populations of the grasshop-
per *Chorthippus brunneus* relative to climate and habitat fragmentation. *Oecologia*, 169, 125–133.


Data accessibility

Georeferenced abundance and microsatellite data and raw reads from the GS-FLX Roche sequencing have been posted on Dryad (doi:10.5061/dryad.88b12).

Supporting information

Additional supporting information may be found in the online version of this article.

Table S1 Details of the four sampling designs

Table S2 Summary statistics of the selected GLM (Generalized Linear Model) testing the effect of the location of the surveyed grasslands (in the linear hedged farmland, in its western side and in its eastern side), the grassland type (meadow or alfalfa) and age (1, 2, 3, 4, 5 and ≥ 5 years), the sampling scheme, the sampling date and interaction terms (I) on *P. giornae* abundance (number per 10 m²)

Table S3 Postprocess parameters and estimated number of clusters (K) from the *GENELAND* inferences

Table S4 Summary statistics for each simulated data set analyzed by *GENELAND*

Table S5 Number of allele (A), unbiased expected heterozygosity (Hₑ) (Nei 1987), observed heterozygosity (Hₒ), Fₛ value and departure from HWE (sign: <0.05∗, <0.01** and <0.001***)) computed for each locus in the two inferred clusters using a medium differentiation prior or no prior for drift parameters

Fig. S1 Representation of the lattice and sampling area in sim1 and sim2 (a), sim3 (b) and sim4 (c).

Fig. S2 Shape of the dispersal function implemented in the lower dispersal zone (top line) and the rest of the lattice (bottom line) in sim1 (e = 0.4, g = 0.675) (a) and e = 0.6, g = 0.75 (b), respectively) and in sim2, sim3 and sim4 (e = 0.2, g = 0.5 (c) and e = 0.8, g = 0.75 (d), respectively).