

# SCIENTIFIC REPORTS

OPEN

## Inter-annual maintenance of the fine-scale genetic structure in a biennial plant

Javier Valverde<sup>1</sup>, José María Gómez<sup>1,2</sup>, Cristina García<sup>3</sup>, Timothy F. Sharbel<sup>4,5</sup>,  
María Noelia Jiménez<sup>6</sup> & Francisco Perfectti<sup>7</sup>

Received: 28 July 2016

Accepted: 02 November 2016

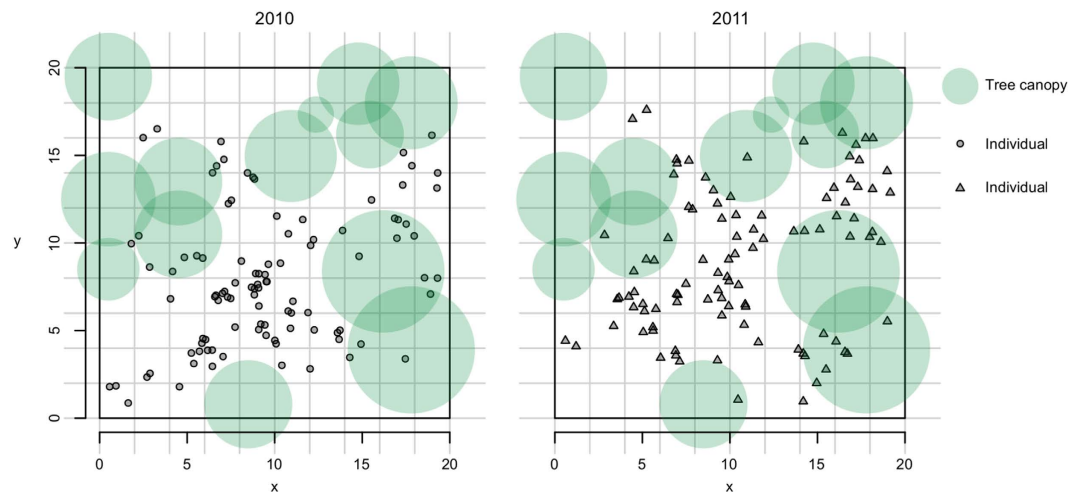
Published: 24 November 2016

Within plant populations, space-restricted gene movement, together with environmental heterogeneity, can result in a spatial variation in gene frequencies. In biennial plants, inter-annual flowering migrants can homogenize gene frequencies between consecutive cohorts. However, the actual impact of these migrants on spatial genetic variation remains unexplored. Here, we used 10 nuclear microsatellite and one plastid genetic marker to characterize the spatial genetic structure within two consecutive cohorts in a population of the biennial plant *Erysimum mediohispanicum* (Brassicaceae). We explored the maintenance of this structure between consecutive flowering cohorts at different levels of complexity, and investigated landscape effects on gene flow. We found that cohorts were not genetically differentiated and showed a spatial genetic structure defined by a negative genetic-spatial correlation at fine scale that varied in intensity with compass directions. This spatial genetic structure was maintained when comparing plants from different cohorts. Additionally, genotypes were consistently associated with environmental factors such as light availability and soil composition, but to a lesser extent compared with the spatial autocorrelation. We conclude that inter-annual migrants, in combination with limited seed dispersal and environmental heterogeneity, play a major role in shaping and maintaining the spatial genetic structure among cohorts in this biennial plant.

Spatial Genetic Structure (SGS) is the non-random spatial distribution of genotypes, a pattern occurring in most plant populations<sup>1</sup>. It results from an equilibrium among space-restricted gene movement, genetic drift and local selection<sup>2,3</sup>. SGS is most of the times the result of isolation by distance (IBD)<sup>2</sup>, in which local mating or short-distance seed dispersal generates a pattern such that kinship among individuals decreases with distance. The scale at which these processes act varies among species<sup>1</sup>, and recently, several studies have demonstrated that it can also happen at very fine scales<sup>4–7</sup>. The extent and strength of SGS depends on seed and pollen dispersal strategies as well as on mating system<sup>1,8,9</sup>. Moreover, dispersal can be directionally biased, resulting in differences in IBD intensity at different spatial directions<sup>10–13</sup> that promote spatial asymmetric SGS, a pattern also occurring at fine-scales<sup>5</sup>. This spatial asymmetry usually responds to environmental heterogeneity<sup>10,11</sup>. At-site variability in ecological conditions can create differences in habitat suitability and, therefore, affect genetic connectivity via seed or pollen dispersal<sup>12,14</sup>. Regarding the latter, an increasing number of studies are analysing environmental heterogeneity to understand fine-scale SGS and its effects on population dynamics and evolution.

Studies dealing with the temporal scale of SGS have been mostly aimed at disentangling how demography affects SGS by comparing among age-classes<sup>15–18</sup>. In perennials, age-classes coexist in the same spatio-temporal context, implying that individuals initially from different cohorts classes may interbreed. In contrast, in strict biennials, i.e., those plants completing their life cycle in two years, individuals plants from consecutive cohorts are reproductively isolated because they do not overlap in their flowering year<sup>19</sup>. However, this presumption is far from reality because only a few species are strictly biennial<sup>20–22</sup>. Several processes promote variation in biennial

<sup>1</sup>Departamento de Ecología, Universidad de Granada, E-18071, Granada, Spain. <sup>2</sup>Departamento de Ecología Funcional y Evolutiva, Estación Experimental de Zonas Áridas (EEZA-CSIC), E-04120, Almería, Spain. <sup>3</sup>Plant Biology, CIBIO/InBio, Centro de Investigação em Biodiversidade e Recursos Genéticos, Laboratório Associado, Universidade do Porto. Campus Agrário de Vairão, 4485-661, Vairão, Portugal. <sup>4</sup>Leibniz Institute for Plant Genetics and Crop Plant Research (IPK), 06466, Gatersleben, Germany. <sup>5</sup>Global Institute for Food Security (GIFS), University of Saskatchewan, Saskatoon, Canada. <sup>6</sup>Departamento de Botánica, Universidad de Granada, E-18071, Granada, Spain. <sup>7</sup>Departamento de Genética, Universidad de Granada, E-18071, Granada, Spain. Correspondence and requests for materials should be addressed to F.P. (email: fperfect@ugr.es)



**Figure 1. Spatial location of the marked plants.** Tree canopy is represented with green circles to evidence the fine-scale environmental heterogeneity. Scale is in meters.

life span, permitting the existence of inter-annual migrants that bridge years ('facultative biennials')<sup>23,24</sup>. These inter-annual migrants favour gene flow between cohorts, leading to homogenisation of their allelic frequencies<sup>25</sup>. However, little is known regarding how inter-annual migrants affect the similarities in SGS among cohorts of a biennial, particularly for small plant populations where temporal changes in population size could dramatically vary their genetic composition<sup>26</sup>. We think that the spatial location at which inter-annual migrants arrive to the next cohort may affect SGS. Particularly, in plants presenting limited seed dispersal, the localized arrival of inter-annual migrants may have predictable effects on the SGS, equalizing the spatial variation in allelic frequencies between consecutive cohorts and balancing population size fluctuations.

We have studied two consecutive cohorts of the biennial plant *Erysimum mediohispanicum* (Brassicaceae) to explore fine-scale SGS and to investigate the temporal maintenance of SGS between cohorts. *E. mediohispanicum* is endemic to the Iberian Peninsula<sup>27</sup>, where a highly diverse assemblage of pollinators visits their flowers. Although slightly self-compatible, this species needs pollinators for a complete seed set<sup>28</sup>, whose dispersal occurs at very short distances due to its barochorous dispersal strategy<sup>29</sup>. Populations of this species are patchily distributed and formed by tens to several hundreds of individuals<sup>30</sup>. These populations vary largely in size due to climate fluctuations, especially linked to drought years.

We have analysed the temporal maintenance of SGS between cohorts of *E. mediohispanicum* using 10 nuclear microsatellite and one plastid DNA marker. We also assess to what extent spatial environmental variation shapes the observed SGS. We first explore if cohorts were genetically differentiated, followed by characterization and comparison of SGS between cohorts at different levels: (1) by using spatial principal component analysis (sPCA) to assess global genetic structure; (2) by testing for isotropic SGS using kinship-distance distograms; (3) by using bearing correlograms to visualize SGS asymmetry (anisotropy). Finally, we assess associations among the spatial variation in genotypes and environmental factors using autoregressive spatial models.

## Results

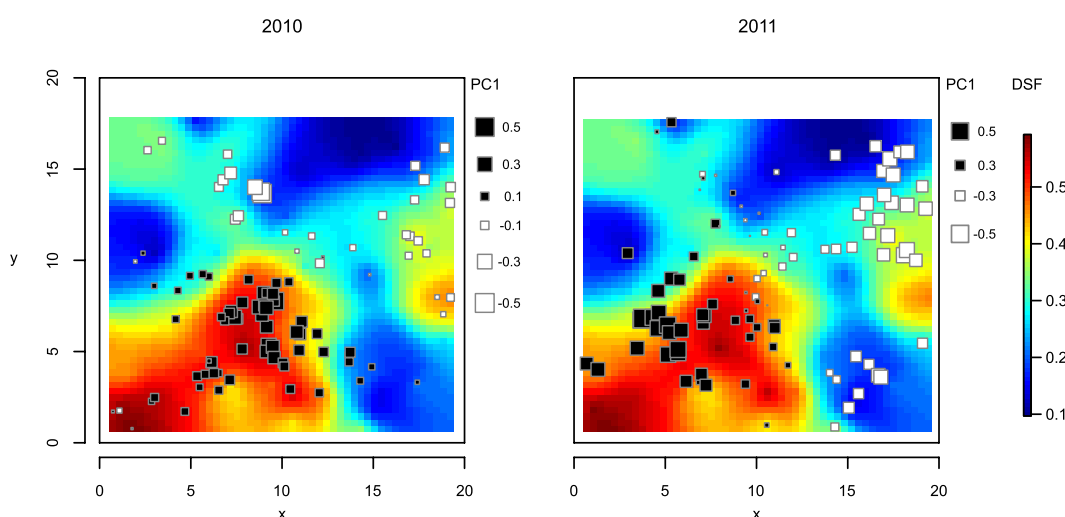
**Genetic characterization.** We successfully characterized 200 sampled individuals (100 per cohort, Fig. 1). The highest proportion of missing data (3%) was found in the locus E3 in 2011. Allelic richness was 11.7, ranging between four and 22 per locus (Table 1). Most loci were at Hardy-Weinberg equilibrium, with only one locus in 2010 and two in 2011 presenting a deficit of heterozygotes. The analysis of the plastid *trnL-trnF* spacer yielded five haplotypes with an uneven occurrence across the population, with 95.5% of individuals bearing one of the two most frequent haplotypes.

**Inter-annual genetic structure.** All loci exhibited extremely low values of the genetic variance component related to inter-annual differences ( $F_{ST} < 0.006$ ,  $p > 0.05$  for all loci), similar to what we found for the multilocus value ( $F_{ST} < 0.001$ ,  $p = 0.226$ ). These low values were consistent with the lack of significance found after an AMOVA ( $\sigma^2 = 0$ ;  $p = 0.765$ ). Moreover, the proportion of inferred migrants between cohorts was high (0.272).

**Spatial variation of allele frequencies.** Tests for spatial structure indicated the existence of global spatial structures within cohorts ( $\max(t) = 0.019$  for both cohorts,  $p = 0.016$  and  $0.036$  for 2010 and 2011 respectively), but no local structures, e.g., genetic differentiation among close neighbours, ( $\max(t) = 0.020$  for both cohorts,  $p = 0.460$  and  $0.126$  for 2010 and 2011). The sPCA analyses showed that most of the data structure was explained by the first three principal components (PC) in 2010 and by the first two in 2011 (Figure S1.A). The eigenvalues of these PCs were 0.045, 0.030 and 0.025 in 2010, and 0.040 and 0.027 in 2011, with the remaining eigenvalues showing lower values (Figure S1.B). For environment-genotype association studies, we retained the first two PCs, as these stand out in terms of genetic variance and spatial autocorrelation (Figure S1.A). The global structure

Locus	2010						2011					
	$R_S$	% MD	$F_{IS}$	$H_O$	$H_O/H_E$	$\chi^2$	$R_S$	% MD	$F_{IS}$	$H_O$	$H_O/H_E$	$\chi^2$
C5	10	0	-0.024	0.756	1.024	21.670	10	0	0.043	0.766	0.957	25.373
D4	8	0	0.027	0.558	0.973	11.625	8	0	-0.025	0.588	1.025	45.185
E4	14	1	0.053	0.794	0.946	148.649	15	0	0.064	0.816	0.936	<b>92.432</b>
D2	22	0	0.053	0.912	0.947	200.625	21	0	<b>0.076</b>	0.925	0.924	211.690
E3	4	0	0.094	0.244	0.906	7.963	6	3	0.062	0.287	0.938	14.710
E6	13	0	-0.057	0.694	1.057	47.827	12	0	0.030	0.684	0.969	40.239
E8	15	0	<b>0.192</b>	0.758	0.808	<b>204.151</b>	16	0	<b>0.170</b>	0.751	0.830	166.314
E5	16	0	0.052	0.869	0.948	117.717	14	1	<b>0.123</b>	0.857	0.877	<b>143.413</b>
D10	5	0	-0.022	0.531	1.022	5.928	5	1	-0.043	0.506	1.043	13.180
D11	6	0	-0.043	0.520	1.043	9.327	5	1	0.079	0.529	0.921	7.506

**Table 1. Population genetic parameters per locus and year (cohort).** For each locus and cohort: allelic richness ( $R_S$ ), percentage of missing data (%MD), inbreeding coefficient ( $F_{IS}$ ), observed heterozygosity ( $H_O$ ), departure from the expected heterozygosity under Hardy-Weinberg equilibrium ( $H_O/H_E$ ), and their corresponding chi-squared values are shown. Significant values are indicated in bold.



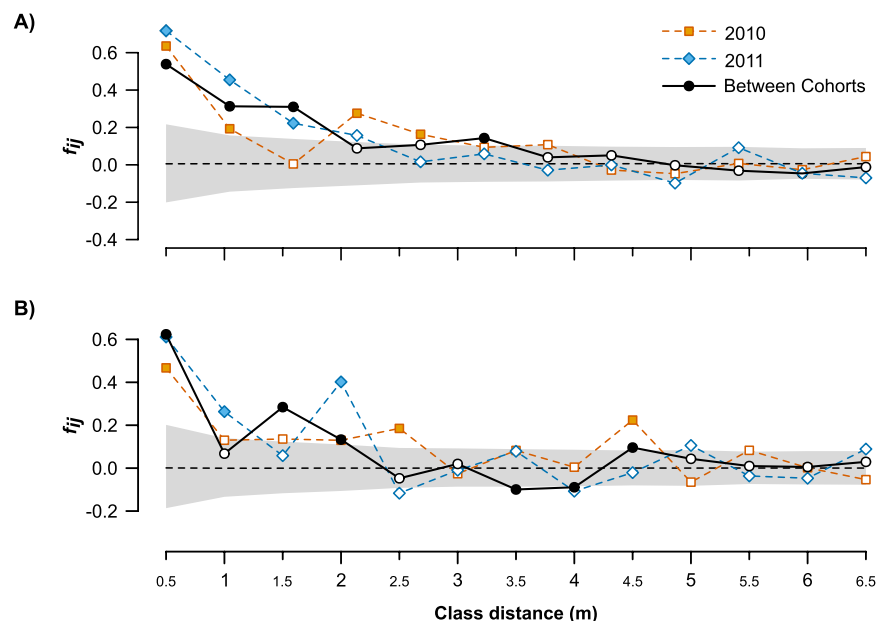
**Figure 2. Spatial representation of the first vector from the sPCA for both cohorts and light availability (DSF).** Individuals tend to be surrounded by other individuals with similar score values, indicating local aggregation of related genotypes. Scores are represented with solid black squared symbols when positive and empty when negative. Square size is relative to the score value.

represented by the first PC (PC1) from the sPCA, revealed similar spatial genetic patterns for both cohorts, extending as a cline along the X-axis of our study plot (Fig. 2), while the second (PC2) did not revealed a clear spatial pattern (Figure S2.B).

**Isotropic spatial genetic structure.** Kinship-distance distograms performed using nuclear markers showed positive significant values only at short distances ( $<2.5$  m; Fig. 3A). The significant positive values for the first distance class ( $0.063 \pm 0.013$  in 2010;  $0.071 \pm 0.020$  in 2011) and the negative slope of the relationship between kinship and the natural logarithm of spatial distance ( $-0.011$  for both cohorts) lead to the same significant value of intensity of SGS ( $Sp = 0.012 \pm 0.001$ ; Table 2). The analysis of the plastid haplotypes showed similar results, with positive and significant values below 0.5 and 1 m in 2010 and 2011 respectively, in addition to occasional significant values at further distances (Fig. 3B). The  $F_I$  and  $Sp$  coefficients depicting SGS intensity showed significant values for both cohorts ( $F_I = 0.467$  and  $0.627$ ;  $Sp = 0.207$  and  $0.019$  in 2010 and 2011 respectively;  $p < 0.05$ ).

We found similar patterns for the isotropic SGS between cohorts. When plotting the kinship coefficients of plants belonging to different cohorts against distance we observed that the pattern was maintained for both types of marker. We obtained significant positive values of average kinship coefficient below 1.5 m for the nuclear markers (Fig. 3A) and below 2 m for the plastid marker (Fig. 3B). Moreover, there were non-significant between-cohort differences in  $F_I$  and  $Sp$  (Table 2).

**Anisotropic spatial genetic structure.** The correlation between kinship and spatial distance varied with compass directions for each cohort and marker, indicating the occurrence of anisotropy in SGS (Fig. 4). This



**Figure 3. Isotropic distograms.** For nuclear (A) and plastid (B) markers, the average kinship at each distance class are plotted with differing symbols depending on the comparison: orange squares (2010), blue diamonds (2011) and black circles (comparisons between cohorts). Filled symbols denote significance of the value when compared with the null hypothesis of no isolation by distance. Grey filled areas represents 95% confidence intervals for the null hypothesis of no spatial structure in the between-cohort comparisons.

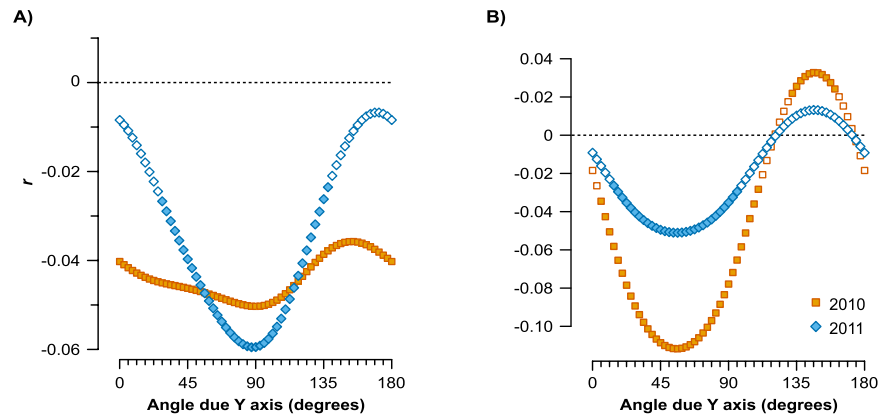
Marker	Cohort	$b_F$	$F_1$	$Sp$
SSR	2010	$-0.011 \pm 0.001$	$0.064 \pm 0.013$	$0.012 \pm 0.001$
	2011	$-0.011 \pm 0.001$	$0.072 \pm 0.020$	$0.012 \pm 0.001$
	Between	$-0.011 \pm 0.001$	$0.063 \pm 0.016$	$0.012 \pm 0.001$
t-test between cohorts		1	0.751	1
cpDNA	2010	$-0.110$	$0.467$	$0.207$
	2011	$-0.007$	$0.627$	$0.019$
	Between	$-0.083$	$0.624$	$0.220$

**Table 2. Isotropic SGS strength within and between cohorts.**  $b_F$ : kinship-log spatial distance regression slope within an optimized range of distances.  $F_1$ : average kinship for the first distance class.  $Sp$ : slope of the regression of kinship on the logarithm of the spatial distance, it equals to  $-b_F/[1 - F_1]$ . Between cohorts parameters were obtained restricting the comparisons to individuals belonging to different cohorts. Significant values are indicated in bold. P-values of the t-tests comparing between cohorts are also shown.

variation coincided between cohorts such that there was congruence in the bearing angles with the strongest and weakest SGS for each marker. For the nuclear markers, both cohorts presented the strongest correlation along the X-axis of the study plot ( $r = -0.050$ ,  $\theta_s = 90^\circ$  in 2010;  $r = -0.059$ ,  $\theta_s = 87^\circ$  in 2011; Table 3; Fig. 4A), and the weakest correlation at relatively close angles ( $r = -0.036$ ,  $\theta_w = 154^\circ$  in 2010;  $r = -0.007$ ,  $\theta_w = 171^\circ$  in 2011). Plastid markers showed the same bearing angle with the strongest correlation ( $\theta_s = 56^\circ$ ,  $r = -0.112$  and  $-0.051$  in 2010 and 2011; Table 3; Fig. 4B) similar to the bearing angle with the weakest correlation ( $\theta_w = 146^\circ$ ,  $r = 0.033$  and  $0.013$  in 2010 and 2011 respectively).

For the nuclear markers the strength of the correlation at  $\theta_s$  showed the same values ( $Sp = 0.013$  in 2010 and 2011), while at  $\theta_w$  these were different ( $Sp = 0.020$  and  $-0.007$ ). The plastid marker showed differing values of strength at both angles ( $\theta_s$ :  $Sp = 0.107$  and  $0.032$ ;  $\theta_w$ :  $Sp = 0.058$  and  $-0.011$ ).

**Environment-genotype correlations.** The lagged autoregressive models performed on PC1 and PC2 presented low values of Moran's I index in the residuals ( $<0.099$ ), indicating a good performance of these models (Supplementary Table S1), while  $\rho$ , the parameter associated with the inherent spatial autocorrelation term, was the most important parameter with values consistently above 0.8 (Table 4). The averaged parameters from the selected models indicated some environment-genotype associations. Light availability (DSF) was strongly associated in both cohorts with PC1 ( $w+ = 1$  and  $0.585$  for 2010 and 2011; Table 4; Fig. 2) and in a lesser extent with PC2. Anions were strongly associated with PC1 in 2011 ( $w+ = 0.681$ ; Figure S2.A) and PC2 in 2010 ( $w+ = 0.479$ ), whereas cations had the strongest effect on PC2 both years ( $w+ = 0.721$  and  $0.438$  for 2010 and 2011; Figure S2.B).



**Figure 4. Bearing correlograms.** For a series of bearing angles from 0 to 180° due the Y-axis of the plot, the Mantel correlation coefficient between genetic similarity and transformed distance matrix is plotted for both cohorts and for the nuclear (A) and plastidial (B) markers. Orange squares denote values for the cohort in 2010; blue diamonds values for the cohort in 2011. Significance after permutation is represented with filled symbols.

## Discussion

We did not find genetic differentiation between two consecutive cohorts of the biennial plant *E. mediohispanicum*, a result indicating that these cohorts belong to the same gene pool and behave cohesively. Furthermore, fine scale SGS was congruent between cohorts, indicating recurrent seed dispersal limitation among years. Additionally, we found an association between the microenvironment and the spatial genetic variation. Below, we discuss the SGS found within the studied population and the causes that may be inducing it. We end by discussing how the absence of genetic differentiation and the maintenance of the SGS may be driven by inter-annual migrants between cohorts.

The spatial distribution of genotypes deviated each studied cohort from what it is expected under a random distribution of alleles, indicating the existence of SGS. Particularly, the genotypes were more similar at short distances, fitting a pattern of isolation by distance even at this fine spatial scale. These results are supported by the global structure found using the sPCA and by the Mantel correlograms for both plastid and nuclear markers. This pattern emerges from restricted gene flow driven by the local dispersal of seeds or pollen, whose contributions to the final SGS may differ<sup>31</sup>. Limited seed dispersal promotes siblings to germinate and establish locally, creating aggregations of relatives and therefore structuring genotypes at the fine-scale<sup>6,7</sup>. This is likely to have happened in our study population because the distances at which the significant positive correlation vanished (<1–2.5 meters) are consistent with the reported distances of seed dispersal in *E. mediohispanicum* (0–0.38 m)<sup>29</sup>. However, limited seed dispersal may not be acting alone. The congruent SGS pattern showed by the nuclear and plastidial markers suggest that pollinators moving pollen at long distances are not blurring the SGS. In addition, we think that the spatial aggregation of individuals (Figure S3) together with the prevalence of floral visitors that tend to forage at short distances<sup>32,33</sup>, may facilitate mating events among nearby relatives, causing biparental inbreeding, and reinforcing the observed SGS. Nevertheless, even under random mating, SGS can occur as a byproduct of limited seed dispersal<sup>8,15</sup>, and therefore further analyses are needed to disentangle the relative role of pollen and seeds dispersal as drivers of the observed pattern.

We found asymmetry in SGS for both cohorts, indicating directionally biased gene flow. Kinship-distance correlation varied periodically along the spatial directions, revealing higher resistances to gene flow at 90 and 87° due the Y-axis of the plot (Fig. 4). Directionally biased gene flow has been related to spatial gradients on environmental factors<sup>13</sup>, and has been reported in plant species where topography or environmental factors reinforce SGS in particular directions<sup>10,11,13,34</sup>. If these gradients are maintained over years, the predominance of gene flow along some axes may leave their signature in the SGS. However, we did not identify any environmental gradients because the analysed environmental factors showed a marked patchy heterogeneity. Several factors acting locally could produce the observed SGS asymmetry. For example, the non-random movement of pollinators driven by preferences for plant traits, or micro-environmental factors, may produce heterogeneous patterns of gene flow<sup>12</sup>. These environmental factors could also affect seed germination and seedling establishment, contributing to directionally biased SGS<sup>35</sup>. Additionally, directional genetic flow from other populations could contribute to the observed pattern, but at the present we have no evidences for this to be considered as an important factor shaping SGS asymmetry.

Site-specific ecological conditions mediate genetic connectivity<sup>36</sup>. In our study, the spatial variation in allele frequencies showed by the PC1 matched the spatial variation in understory light availability (Fig. 2). This pattern could be produced by light modulating the suitability of spaces for seed germination and seedling establishment<sup>37</sup>, and therefore contributing to a kind of fine-scale isolation-by-environment<sup>38</sup>. In addition, this pattern could be reinforced by the influence of light on the identity and behaviour of the pollinators visiting the plants<sup>39</sup>. Under this scenario, plants sharing the same light conditions would be visited by similar pollinators and therefore may exhibit higher genetic connectivity. This hypothesis, however, needs to be tested by further analyses of paternity on the offspring to relate pollen flow with understory light conditions and pollinators.



Marker	Cohort	$\theta_s$				$\theta_w$			
		$\theta$	$b_F$	$F_I$	$Sp$	$\theta$	$b_F$	$F_I$	$Sp$
SSR	2010	90°	−0.013 ± 0.003	−0.001 ± 0.001	0.013 ± 0.003	154° 42′	−0.020 ± 0.006	0.000 ± 0.001	0.020 ± 0.006
	2011	87° 11′	−0.013 ± 0.005	0.000	0.013 ± 0.005	171° 34′	0.007 ± 0.003	−0.001 ± 0.001	−0.007 ± 0.003
t-test between cohorts			0.962	0.782	0.962		0.000	0.426	0.000
cpDNA	2010	56° 15′	−0.104	0.029	0.107	146° 15′	−0.059	−0.020	0.058
	2011	56° 15′	−0.032	0.005	0.032	146° 15′	0.011	−0.011	−0.011

**Table 3. SGS strength within cohorts at the bearing angles with the strongest and weakest kinship-distance correlations.** For the compass directions with the strongest ( $\theta_s$ ) and weakest ( $\theta_w$ ) kinship-distance correlations, we show their bearing angles ( $\theta$ ) and SGS strength parameters  $b_F$ ,  $F_I$  and  $Sp$ . P-values of the t-tests comparing between cohorts are also shown.

		2010			2011		
		Estimate	SE	w+	Estimate	SE	w+
PC1	$\rho$	0.883	0.078	1	0.874	0.087	1
	Light availability (DSF)	0.158	0.041	1	0.036	0.042	0.585
	Anions (N and P)	−0.004	0.019	0.135	−0.060	0.059	0.681
	Cations 1 (Mg <sup>2+</sup> and K <sup>+</sup> )	−0.049	0.050	0.632	0.046	0.063	0.498
	Cations 2 (Na <sup>+</sup> )	0.003	0.013	0.140	−0.001	0.014	0.078
	Field capacity	−0.054	0.042	0.807	−0.001	0.021	0.079
PC2	$\rho$	0.899	0.068	1	0.927	0.051	1
	Light availability (DSF)	−0.035	0.040	0.596	−0.004	0.018	0.217
	Anions (N and P)	−0.034	0.045	0.479	−0.001	0.021	0.158
	Cations 1 (Mg <sup>2+</sup> and K <sup>+</sup> )	−0.053	0.049	0.721	−0.022	0.041	0.438
	Cations 2 (Na <sup>+</sup> )	−0.125	0.028	1	−0.008	0.026	0.244
	Field capacity	0.008	0.020	0.268	0.014	0.038	0.326

**Table 4. Model-averaged parameter estimates after model selection.** Regression model parameters with estimates, standard errors, and relative importance values (w+) resulting from the model selection.  $\rho$  denotes the parameter associated with the inherent spatial autocorrelation term. The w+ values of  $\rho$  equals to 1 because this is a structural parameter of the lagged simultaneous autoregressive models.

We also found that SGS was associated with some soil variables. Effects of soil heterogeneity on the pattern of genetic variation have been reported in several plant species. For example, Segarra-Moragues *et al.*<sup>35</sup> have recently shown significant genetic differentiation between rosemary plants (*Rosmarinus officinalis*) growing in siliceous vs. calcareous soils. In our case, the observed fine-scale soil-SGS association could be caused by the differential performance (establishment, survival, growth) of some genotypes in slightly different edaphic conditions. However, it is noteworthy to highlight the high values of  $\rho$  (>0.867) in our models, indicating that, at this scale, spatially limited dispersion prevails over these site-specific environmental factors. In this sense, more evidence is necessary to disentangle the importance of the genotype-environment interactions in shaping SGS at this fine scale.

Prudence is cautioned in interpreting spatially lagged regressions because little is known about their performance when applied to landscape genetic analyses<sup>40</sup>. However, we are confident of our results for a number of reasons. It has been shown that these models are robust to type I errors<sup>41,42</sup> and do not fail when using sPCA principal components as response variables<sup>40,43</sup>. In addition,  $\rho$  was consistently high in our models, and the residuals showed little to no spatial autocorrelation (Table S1), indicating that these models accounted adequately for the spatial autocorrelation. Nevertheless we are aware that the associations found here have to be interpreted cautiously when inferring the causal mechanisms leading to these environment-genotype associations.

Plants flowering in different years did not differ in their allelic frequencies, indicating that the two studied cohorts belong to the same gene pool. This is confirmed by the low values of  $F_{ST}$  and the non-significance of the AMOVA. In small populations of biennial plants, such as *E. mediohispanicum*, recurring and catastrophic declines of population size could conduce to genetic differentiation even among cohorts, since these fluctuations in population size are expected to produce intense genetic drift in small populations<sup>26</sup>. But even in the case of large demographic events, the occurrence of gene flow between cohorts may counterbalance genetic differentiation. Gene flow is surely due to the high amount of inter-annual migrants, up to 17% according to our Bayesass estimates, a high value when compared to most studies using this same algorithm<sup>44</sup>. Several factors may promote between-cohort gene flow via different pathways in *E. mediohispanicum* (Figure S4). Seed dormancy can be activated by unfavourable weather conditions<sup>45</sup> and leads to mating between the descendants of different cohorts<sup>25</sup>. In addition, the arrival of inter-annual migrants can take other different paths. Firstly, via the repeated flowering of some individuals in the oncoming year. This could be caused by delayed reproduction or via a reduced reproductive success during a year followed by subsequent resource allocation, permitting a new reproductive event

in the following year<sup>46</sup>. In our study species, several factors such as herbivory<sup>47</sup> or pollen limitation<sup>48</sup> may trigger the existence of this type of inter-annual migrants. Additionally, delayed or precocious flowering can occur as a response to climatic suitability to plant growth<sup>46</sup>, a likely candidate factor considering the climatic unpredictability of Mediterranean ecosystems<sup>49</sup>. Whatever the reasons, our results demonstrate that inter-annual migrants prevent the genetic differentiation among cohorts, contributing to population cohesion.

Between-cohort gene flow, besides homogenizing allelic frequencies, also reinforced SGS. Firstly, kinship-distance correlations showed a similar pattern within and between cohorts, with the same significant positive values spanning up to 1.5 and 2 m for nuclear and plastid markers respectively. Along with this, SGS intensity did not differ significantly between cohorts, emphasizing the similarities in isotropic SGS. These findings confirm that plants growing nearby are genetically more similar independently of whether they belong to the same flowering cohort or not. In addition, SGSs were similar in their spatial asymmetry, with coinciding bearing angles for the strongest and weakest correlations. The maintenance of the SGS between cohorts is likely due to the limited dispersal of seeds. The integration of all individual seed shadows determines the starting template for the spatial genetic patterns of future flowering plants<sup>16,50</sup>. Therefore, an inter-annual migrant stands as a representative of this spatial template in the new cohort. From this localized migration spot, the oncoming processes of gene flow through pollination and seed dispersal will spread the migrant genes in the new cohort similarly to what would occur in its parental cohort. In this sense, the environmental factors will similarly favour or restrict gene flow in some directions, and therefore migrant alleles will flow asymmetrically. These recurrent migration events will, in the long term, match the spatial genetic variation among cohorts, leading to the observed congruence in SGS. Furthermore, between-populations migration could also contribute to produce the observed SGS. However, biased and recurrent migration, probably mediated by long-distance pollination events, would be needed to explain the observed pattern. As these pollination events do not contribute to the plastidial SGS, the importance of recurrent spatial migration in the temporal maintenance of the SGS is, at least, doubtful.

We have demonstrated the existence of congruent fine-scale spatial genetic variation in two consecutive cohorts of a natural population of the biennial plant *E. mediohispanicum*. This SGS was consistently associated with environmental factors, such as light availability and soil composition. More importantly, there was always a strong spatial autocorrelation in SGS, suggesting that SGS was mostly caused by the spatial pattern of limited seed dispersal in distance rather than by any environment-genotype association. The two studied cohorts had homogenized allelic frequencies, indicating the absence of genetic isolation and the existence of inter-annual migrants. These migrants, due to spatial limited seed dispersal exhibited by *E. mediohispanicum*, imprint a spatially localized genetic signature on the new cohort, which ultimately results in a between-cohort matching of SGS.

## Materials and Methods

**Sampling design.** In years 2010 and 2011 we set up a 20 × 20 m plot delimiting a population of *E. mediohispanicum* situated at 1723 m.a.s.l. at the Sierra Nevada Protected Area (SE Spain; 37° 8.07' N 3° 21.71' W). Each year we randomly selected 100 flowering plants and mapped their spatial location at the centimetre level (Fig. 1). These plants represented ~ 85% of the total individuals comprising a reproductive cohort. We characterized light and soil conditions for each individual plant. Light availability, measured as Direct Site Factor (DSF), was derived from hemispherical photographs taken at the ground level using a Nikon coolpix 4500 and subsequently analysed by applying the Hemiview software (Delta-T Devices, Cambridge, UK). Soil was characterized by means of cations (Na<sup>+</sup>, K<sup>+</sup> and Mg<sup>2+</sup>) and anions (total nitrogen and phosphorous), and moisture content at field capacity as a reliable measure of the soil's ability to retain water (see Supplementary Information S1). To decrease variable dimensionality we ran a PCA over cations and anions independently. We found two components explaining cation variability: the first (cations 1) explaining 58% and was associated to Mg<sup>2+</sup> and K<sup>+</sup> and the second (cations 2) explaining 33% and was associated to Na<sup>+</sup> (Figure S5). Anions were synthesised in a variable explaining 70% of the variability and equally associated to nitrogen and phosphorous.

**Genetic characterization.** We genotyped each plant using nuclear and plastid molecular markers. We isolated DNA using the GenElute Plant Genomic DNA Miniprep Kit (Sigma-Aldrich) on 60 mg of plant material previously crunched in liquid nitrogen.

We amplified the plastidial *trnL-trnF* spacer (~1300 bp) using the tabC and tabF primers (see Abdelaziz *et al.*<sup>51</sup> and Supplementary Information S2). PCR products were mixed with 0.15 volumes of 3 M sodium acetate, pH 4.6, and 3 volumes 95% (v/v) ethanol and subsequently precipitated after centrifuging at 4 °C. Amplicons were then sent to Macrogen (Geumchun-gu, Seoul, Korea) for sequencing in both directions, using the respective PCR primers. Chromatograms were reviewed and contigs were produced using Geneious v.7<sup>52</sup> (Biomatters, <http://www.geneious.com/>). Sequences were uploaded to Genbank (accession numbers KX641272 to KX641276).

From each plant we amplified 10 unlinked nuclear microsatellites loci (SSR) described in Muñoz-Pajares *et al.*<sup>53</sup> (Supplementary Information S2). We additionally genotyped 500 offspring in order to detect and correct genotyping errors (e.g., allele dropout or null alleles) in the parental plant genotypes. Electropherograms were analysed and genotypes called using PeakScanner v.2 (Applied Biosystems) and exhaustive eye-inspection. We characterized each microsatellite locus within cohorts by computing observed and expected heterozygosity, number of alleles, and inbreeding coefficient. For this later we tested if deviated significantly from zero using 999 bootstrap replicates. We tested for Hardy-Weinberg equilibrium using an exact test based on Monte Carlo permutations for each locus and for all loci together. These tests were performed using the 'pegas'<sup>54</sup> and 'diveRsity'<sup>55</sup> packages in the open source software R v. 3.2.2.

**Inter-annual genetic structure.** We evaluated the amount of genetic differentiation between cohorts using Wright's  $F_{ST}$  statistic. The significance of this statistic was obtained by comparing with a null distribution obtained after 500 random permutations of genotypes. We also performed an AMOVA between cohorts using Chord's

distances among genotypes<sup>56</sup>. Moreover, we used Bayesass v.1.3<sup>57</sup> to infer rates of recent inter-annual migrants. This software uses Markov chain Monte Carlo techniques to estimate the posterior probabilities of recent individual immigrants. We ran three independent MCMC runs for  $10^7$  iterations with a thinning of 2000 and a burn-in of the first 10% of the iterations. Delta parameters of allele frequency, migration, and inbreeding were set as 0.15.

**Spatial variation of allele frequencies.** We checked for the occurrence of fine-scale SGS within each cohort by conducting a spatial principal component analysis (sPCA)<sup>58</sup> on the nuclear markers. This multivariate ordination technique tests for the existence of global structures (such as spatial clines) or local structures (genetic differences between neighbours). sPCA integrates principal component analysis and Moran's autocorrelation index to reduce the multidimensional nature of genotype data into a set of highly informative orthogonal vectors that differentiate spatial patterns of genetic variation. To accomplish this analysis, a connection network depicting spatial weights among individuals has to be previously defined. We used the inverse of the spatial distances among pairs so that all plants were considered neighbours while accounting for the spatial cost distance. With the resulting sPCA principal components, we assessed spatial structures by representing the scores of the two main vectors in space. We only used the two first components (PC1 and PC2) to reduce the burden of the computation to the most informative sPCA principal components. We evaluated the significance of global and local structures using the test proposed by Jombart *et al.*<sup>58</sup>.

**Isotropic spatial genetic structure.** We explored the extent and intensity of SGS to evaluate if relatives were spatially clumped as a result of isolation by distance. For each cohort and for each nuclear (SSR) and plastid marker, we calculated Nason's kinship coefficients<sup>59</sup>, which are standardized to a population's allele frequencies and are highly robust to HWE deviations<sup>1</sup>. When calculating this coefficient using the plastid marker, we treated haplotypes as alleles from a haploid organism. To explore the spatial extent of SGS we calculated the average kinship coefficient for a set of distance classes ( $F_D$ ) spanning 0.5 meters, and plotted them against distance<sup>60</sup>. We subsequently measured the intensity of SGS by calculating two gene dispersal parameters:  $F_I$ , the average kinship coefficient for the first distance class (<0.5 meters), and the  $Sp$  coefficient<sup>1</sup>, defined as  $-b_F/(1-F_I)$ , where  $b_F$  is the slope of the regression of kinship coefficients on the natural logarithm of the spatial distance among individuals within an optimized range of distances. The range of distances for which  $b_F$  is calculated is related to the average parent-offspring distance and its approximation follows an iterative procedure<sup>1</sup> implemented in the software SPAGeDi 1.5a<sup>61</sup>. Standard errors for all parameters were calculated jackknifing over SSR loci<sup>1,9</sup>, except in the case of cpDNA haplotypes, where only a marker was available. For SSR  $Sp$ , and each SSR  $F_D$  we obtained a null distribution of expected values under the hypothesis of no SGS by permuting genotypes among individuals 9999 times, and compared them with the observed values to obtain a significance value.

We investigated the maintenance of the isotropic SGS between cohorts by plotting kinship coefficients against spatial distance as explained before, but restricting the comparisons to individuals belonging to different cohorts. Then, we compared the intensity of the SGS between cohorts by a t-test using the average and standard errors obtained by jackknifing over loci. We also calculated  $b_F$ ,  $F_I$  and  $Sp$  dispersal parameters restricting the comparisons to individuals belonging to different cohorts.

**Anisotropic spatial genetic structure.** We explored the anisotropy of SGS by testing whether its strength varied in different spatial directions using Rosenberg's<sup>62</sup> bearing correlogram, an analysis that uses a Mantel test to correlate a genetic similarity matrix with a transformed distance matrix for a set of spatial directions, measured as the angles formed with the Y-axis ( $\theta$ ). For a given direction, the transformed distance matrix is obtained by first calculating the natural logarithm of the original distance matrix, followed by weighting its values by the squared cosine of the clockwise bearing angle depicted by each pair of individuals and the fixed spatial direction. For this analysis we used the obtained kinship matrix and a set of 128 equidistant bearing angles. The significance of the correlations was obtained using a permutation test (999 permutations). Through this analysis we were able to find the bearing angle denoting the strongest ( $\theta_s$ , minimal Mantel correlation) and weakest SGS directions ( $\theta_w$ , maximal Mantel correlation). Next, we calculated the strength of SGS at  $\theta_s$  and  $\theta_w$  for each cohort. To perform this, and to augment the number of paired kinship coefficients, we included the plants inside a range of  $30^\circ$  around  $\theta_s$  and  $\theta_w$ <sup>10</sup>. For each cohort and  $\theta_s$  and  $\theta_w$ , we computed  $Sp$  and  $F_I$  and obtained their significance as explained before. These analyses were performed using SPAGeDi 1.5a<sup>61</sup> and personalized scripts in R (Supplementary Information S3). As the temporal maintenance of anisotropy is determined by the congruence and strength of  $\theta_s$  and  $\theta_w$  between cohorts, we therefore compared the gene dispersal parameters  $Sp$  and  $F_I$  by means of a t-test.

**Environment-genotype correlations.** We evaluated the contribution of environmental factors to the spatial genotypic variation beyond isolation by distance. We used lagged simultaneous autoregressive models using the first two components of the sPCA as response variables<sup>43</sup>. These models assume the inherent spatial autocorrelation occurring in the response variable and are fairly robust to type I errors<sup>41,42</sup>. The models take the form:

$$G_i = \rho W_{ij} G_j + \beta X + \varepsilon$$

In this equation, the typical ordinary least squares regression ( $G_i = \beta X + \varepsilon$ ) is modified by adding a term ( $\rho W_{ij} G_j$ ) which controls for the inherent spatial autocorrelation which is assumed to occur in the response variable. The PC scores of all other individuals ( $G_j$ ) are weighted by the parameter  $\rho$ , which accounts for the lack of independence among individuals, and by the cost-distance-weighting matrix ( $W_{ij}$ ). We computed  $W_{ij}$  using the inverse spatial distances among individuals because it approximately emulates a spatial autocorrelation under IBD<sup>40,43</sup>. By using site-based measures we were able to determine the environmental factors associated with genetic structure<sup>63</sup>.



From all possible combinations of models we selected those with  $\Delta AIC < 2$  from the best model. Following, we model-averaged parameter estimates and calculated their relative importance following Burnham and Anderson<sup>64</sup>. Moreover, the selected models were checked for remaining autocorrelation in residuals by using a Moran's I test. Model performances and selection were calculated using the R packages 'spdep'<sup>65</sup> and 'AICcmodavg'<sup>66</sup>.

## References

- Vekemans, X. & Hardy, O. J. New insights from fine-scale spatial genetic structure analyses in plant populations. *Mol. Ecol.* **13**, 921–935 (2004).
- Wright, S. Isolation by Distance. *Genetics* **28**, 114–138 (1943).
- Epperson, B. K. Recent advances in correlation analysis of spatial patterns of genetic variation. *Evol. Biol.* **27**, 95–155 (1993).
- Hardy, O. J. *et al.* Fine-scale genetic structure and gene dispersal in *Centaurea corymbosa* (Asteraceae). II. Correlated paternity within and among sibships. *Genetics* **168**, 1601–1614 (2004).
- Rouger, R. & Jump, A. S. Fine-scale spatial genetic structure across a strong environmental gradient in the saltmarsh plant *Puccinellia maritima*. *Evol. Ecol.* **29**, 609–623 (2015).
- Lara-Romero, C. *et al.* Individual spatial aggregation correlates with between-population variation in fine-scale genetic structure of *Silene ciliata* (Caryophyllaceae). *Heredity (Edinb)* **116**, 417–423 (2016).
- Volis, S., Zaretsky, M. & Shulgina, I. Fine-scale spatial genetic structure in a predominantly selfing plant: role of seed and pollen dispersal. *Heredity (Edinb)* **105**, 384–393 (2010).
- Hamrick, J. L., Murawski, D. A. & Nason, J. D. The influence of seed dispersal mechanisms on the genetic structure of tropical tree populations. *Vegetatio* **107/108**, 281–297 (1993).
- Hardy, O. J. *et al.* Fine-scale genetic structure and gene dispersal inferences in 10 neotropical tree species. *Mol. Ecol.* **15**, 559–571 (2006).
- Born, C., Le Roux, P. C., Spohr, C., McGeoch, M. A. & van Vuuren, B. J. Plant dispersal in the sub-Antarctic inferred from anisotropic genetic structure. *Mol. Ecol.* **21**, 184–194 (2011).
- Rhodes, M. K., Fant, J. B. & Skogen, K. A. Local topography shapes fine-scale spatial genetic structure in the Arkansas Valley evening primrose, *Oenothera harringtonii* (Onagraceae). *J. Hered.* **4**, 1–10 (2014).
- DiLeo, M. F. *et al.* The gravity of pollination: integrating at-site features into spatial analysis of contemporary pollen movement. *Mol. Ecol.* **23**, 3973–3982 (2014).
- Wang, Z.-F. *et al.* Pollen and seed flow under different predominant winds in wind-pollinated and wind-dispersed species *Engelhardtia roxburghiana*. *Tree Genet. Genomes* **12**, 19 (2016).
- Holderegger, R., Buehler, D., Gugerli, F. & Manel, S. Landscape genetics of plants. *Trends Plant Sci.* **15**, 675–683 (2010).
- Kalisz, S., Nason, J. D., Hanzawa, F. M. & Tonsor, S. J. Spatial population genetic structure in *Trillium grandiflorum*: the roles of dispersal, mating, history, and selection. *Evolution* **55**, 1560–1568 (2001).
- Jones, F. A. & Hubbell, S. P. Demographic spatial genetic structure of the Neotropical tree, *Jacaranda copaia*. *Mol. Ecol.* **15**, 3205–3217 (2006).
- Zhou, H. P. & Chen, J. Spatial genetic structure in an understory dioecious fig species: The roles of seed rain, seed and pollen-mediated gene flow, and local selection. *J. Ecol.* **98**, 1168–1177 (2010).
- Berens, D. G. *et al.* Fine-scale spatial genetic dynamics over the life cycle of the tropical tree *Prunus africana*. *Heredity (Edinb)* **113**, 401–407 (2014).
- Kelly, D. Demography of short-lived plants in chalk grassland. I. Life cycle variation in annuals and strict biennials. *J. Ecol.* **77**, 747–769 (1989).
- Hart, R. Why are biennials so few? *Am. Nat.* **111**, 792–799 (1977).
- Silvertown, J. Death of the elusive biennial. *Nature* **310**, 271 (1984).
- Klinkhamer, P. G. L., de Jong, T. J. & Meelis, E. Delay of flowering in the 'biennial' *Cirsium vulgare*: size effects and devernization. *Oikos* **49**, 303–308 (1987).
- de Jong, T. J., Klinkhamer, P. G. L. & Metz, J. A. J. Selection for biennial life histories in plants. *Vegetatio* **70**, 149–156 (1987).
- de Jong, T. J. & Klinkhamer, P. G. L. Population ecology of the biennials *Cirsium vulgare* and *Cynoglossum officinale*. *J. Ecol.* **76**, 366–382 (1988).
- Königer, J., Rebernik, C. a., Brabec, J., Kiehl, K. & Greimler, J. Spatial and temporal determinants of genetic structure in *Gentianaella bohemica*. *Ecol. Evol.* **2**, 636–648 (2012).
- Ellstrand, N. C. & Elam, D. R. Population genetic consequences of small population size: Implications for plant conservation. *Annu. Rev. Ecol. Syst.* **24**, 217–242 (1993).
- Nieto-Feliner, G. In *Flora Iberica vol. 4* (ed. Castroviejo, S.) (Consejo Superior de Investigaciones Científicas, 2003). at [http://www.floraiberica.es/eng/PHP/direcciones\\_todas2.php?autor=Nieto+Feliner,+G](http://www.floraiberica.es/eng/PHP/direcciones_todas2.php?autor=Nieto+Feliner,+G).
- Gómez, J. M. *et al.* Association between floral traits and rewards in *Erysimum mediohispanicum* (Brassicaceae). *Ann. Bot.* **101**, 1413–1420 (2008).
- Gómez, J. M. Dispersal-mediated selection on plant height in an autochorously dispersed herb. *Plant Syst. Evol.* **268**, 119–130 (2007).
- Gómez, J. M. & Perfectti, F. Evolution of Complex Traits: The Case of *Erysimum* Corolla Shape. *Int. J. Plant Sci.* **171**, 987–998 (2010).
- Freeland, J. R., Biss, P. & Silvertown, J. Contrasting patterns of pollen and seed flow influence the spatial genetic structure of sweet vernal grass (*Anthoxanthum odoratum*) populations. *J. Hered.* **103**, 28–35 (2012).
- Pérez-Collazos, E., Segarra-Moragues, J. G., Villar, L. & Catalán, P. Ant pollination promotes spatial genetic structure in the long-lived plant *Borderea pyrenaica* (Dioscoreaceae). *Biol. J. Linn. Soc.* **116**, 144–155 (2015).
- Valverde, J., Gómez, J. M. & Perfectti, F. The temporal dimension in individual-based plant pollination networks. *Oikos* **125**, 468–479 (2016).
- Austerlitz, F., Dutech, C., Smouse, P. E., Davis, F. & Sork, V. L. Estimating anisotropic pollen dispersal: a case study in *Quercus lobata*. *Heredity (Edinb)* **99**, 193–204 (2007).
- Segarra-Moragues, J. G., Carrión Marco, Y., Castellanos, M. C., Molina, M. J. & García-Fayos, P. Ecological and historical determinants of population genetic structure and diversity in the Mediterranean shrub *Rosmarinus officinalis*. *Bot. J. Linn. Soc.* **180**, 50–63 (2016).
- Dyer, R. J. & Sork, V. L. Pollen pool heterogeneity in shortleaf pine, *Pinus echinata* Mill. *Mol. Ecol.* **10**, 859–866 (2001).
- Isselstein, J., Tallwin, J. R. B. & Smith, R. E. N. Factors affecting seed germination and seedling establishment of fen-meadow species. *Restor. Ecol.* **10**, 173–184 (2002).
- Wang, I. J. & Bradburd, G. S. Isolation by environment. *Molecular Ecology* **23**, 5649–5662 (2014).
- Herrera, C. M. Microclimate and individual variation in pollinators: flowering plants are more than their flowers. *Ecology* **76**, 1516–1524 (1995).
- Kierepka, E. M. & Latch, E. K. Fine-scale landscape genetics of the American badger (*Taxidea taxus*): disentangling landscape effects and sampling artifacts in a poorly understood species. *Heredity (Edinb)* **116**, 33–43 (2016).
- Kissling, W. D. & Carl, G. Spatial autocorrelation and the selection of simultaneous autoregressive models. *Glob. Ecol. Biogeogr.* **17**, 59–71 (2007).

42. Beale, C. M., Lennon, J. J., Yearsley, J. M., Brewer, M. J. & Elston, D. A. Regression analysis of spatial data. *Ecol. Lett.* **13**, 246–264 (2010).
43. Robinson, S. J., Samuel, M. D., Lopez, D. L. & Shelton, P. The walk is never random: Subtle landscape effects shape gene flow in a continuous white-tailed deer population in the Midwestern United States. *Mol. Ecol.* **21**, 4190–4205 (2012).
44. Meirmans, P. G. Nonconvergence in Bayesian estimation of migration rates. *Mol. Ecol. Resour.* **14**, 726–733 (2014).
45. Baskin, C. C. & Baskin, J. M. In *Seeds: Ecology, biogeography, and, evolution of dormancy and germination* 37–77 (Elsevier Inc.), doi: 10.1016/B978-0-12-416677-6.00003-2 (2014).
46. Klinkhamer, P. G. L., de Jong, T. J. & Meelis, E. Life-history variation and the control of flowering in short-lived monocarps. *Oikos* **49**, 309–314 (1987).
47. Gómez, J. M. Herbivory Reduces the Strength of Pollinator-Mediated Selection in the Mediterranean Herb. *Am. Nat.* **162**, 242–256 (2003).
48. Gómez, J. M., Abdelaziz, M., Lorite, J., Jesús Muñoz-Pajares, a. & Perfectti, F. Changes in pollinator fauna cause spatial variation in pollen limitation. *J. Ecol.* **98**, 1243–1252 (2010).
49. Gómez, J. M., Valladares, F. & Puerta-Piñero, C. Differences between structural and functional environmental heterogeneity caused by seed dispersal. *Funct. Ecol.* **18**, 787–792 (2004).
50. Shimono, A., Ueno, S., Tsumura, Y. & Washitani, I. Spatial genetic structure links between soil seed banks and above-ground populations of *Primula modesta* in subalpine grassland. *J. Ecol.* **94**, 77–86 (2006).
51. Abdelaziz, M. *et al.* Using complementary techniques to distinguish cryptic species: a new *Erysimum* (Brassicaceae) species from North Africa. *Am. J. Bot.* **98**, 1049–1060 (2011).
52. Kearse, M. *et al.* Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28**, 1647–1649 (2012).
53. Muñoz-Pajares, A. J. *et al.* Characterization of microsatellite loci in *Erysimum mediohispanicum* (Brassicaceae) and cross-amplification in related species. *Am. J. Bot.* **98**, e287–e289 (2011).
54. Paradis, E. pegas: an R package for population genetics with an integrated-modular approach. *Bioinformatics* **26**, 419–420 (2010).
55. Keenan, K., McGinnity, P., Cross, T. F., Crozier, W. W. & Prodöhl, P. A. diveRsity: An R package for the estimation and exploration of population genetics parameters and their associated errors. *Methods Ecol. Evol.* **4**, 782–788 (2013).
56. Excoffier, L. & Smouse, P. E. Analysis of Molecular Variance Inferred From Metric Distances Among DNA Haplotypes: Application. *Genetics* **131**, 479–491 (1992).
57. Wilson, G. A. & Rannala, B. Bayesian inference of recent migration rates using multilocus genotypes. *Genetics* **163**, 1177–1191 (2003).
58. Jombart, T., Devillard, S., Dufour, A.-B. & Pontier, D. Revealing cryptic spatial patterns in genetic variability by a new multivariate method. *Heredity (Edinb.)* **101**, 92–103 (2008).
59. Loiselle, B. A., Sork, V. L., Nason, J. & Graham, C. Spatial genetic structure of a tropical understory shrub, *Psychotria officinalis* (Rubiaceae). *Am. J. Bot.* **82**, 1420–1425 (1995).
60. Hardy, O. J. & Vekemans, X. Isolation by distance in a continuous population: reconciliation between spatial autocorrelation analysis and population genetics models. *Heredity (Edinb.)* **83**, 145–154 (1999).
61. Hardy, O. J. & Vekemans, X. spagedi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Mol. Ecol. Notes* **2**, 618–620 (2002).
62. Rosenberg, M. S. The bearing correlogram: a new method of analyzing directional spatial autocorrelation. *Geogr. Anal.* **32**, 267–278 (2000).
63. Jombart, T., Pontier, D. & Dufour, A.-B. Genetic markers in the playground of multivariate analysis. *Heredity (Edinb.)* **102**, 330–341 (2009).
64. Burnham, K. P. & Anderson, D. R. *Model selection and multimodel inference: a practical information-theoretic approach*. (Springer-Verlag, 2002). at <http://www.springer.com/us/book/9780387953649>.
65. Bivand, R. *et al.* *spdep: spatial dependence: weighting schemes, statistics and models*. *spdep R package version 0.5-56*. (2011). at <http://cran.r-project.org/package=spdep>.
66. Mazerolle, M. J. *AICcmodavg: model selection and multimodel inference based on (Q)AIC(c)*. *R package version 2.0-4*. (2016). at <http://cran.r-project.org/package=AICcmodavg>.

## Acknowledgements

We are grateful to Belén Herrador and José María Corral for help and advice on DNA extraction and genotyping. We thank M. Abdelaziz for discussion on the previous version on this work and Susana Hitos for helping with soil analyses. M. Berbel, A. L. Caravantes and J. Lorite helped with data collection. P. Jordano provided us the photographic equipment used to take the hemispherical photographs. Our work in the study area was approved by the Sierra Nevada Natural and National Park headquarters. Grants from the Spanish Ministerio de Economía y Competitividad (CGL2009-07015, CGL2012-34736, CGL2013-47558-P), including EU-FEDER funds, and Plan Andaluz de Investigación (P11-RNM-7676) supported this study. JV was supported by a fellowship (BES-2010-030067) from MINECO. CG was supported by the FCT Investigador Programme and co-funded by the European Program COMPETE: FCOMP-01-0124-FEDER-019772 and FCT-ANR/BIA-BIC/0010/2013.

## Author Contributions

J.V., J.M.G. and F.P. performed fieldwork and wrote the manuscript. J.V. performed lab-work and analysed the results with the contribution of F.P. and C.G., M.N.J. performed soil analyses. T.F.S. provided laboratory facilities and SSR guidance. All authors reviewed the manuscript.

## Additional Information

**Supplementary information** accompanies this paper at <http://www.nature.com/srep>

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Valverde, J. *et al.* Inter-annual maintenance of the fine-scale genetic structure in a biennial plant. *Sci. Rep.* **6**, 37712; doi: 10.1038/srep37712 (2016).

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2016

## **Supplementary Information**

### **Inter-annual maintenance of the fine-scale genetic structure in a biennial plant**

Javier Valverde<sup>1</sup>, José María Gómez<sup>1,2</sup>, Cristina García<sup>3</sup>, Timothy F. Sharbel<sup>4</sup>, M<sup>a</sup> Noelia Jimenez<sup>5</sup>,

Francisco Perfectti<sup>6</sup>

<sup>1</sup> Dpto de Ecología, Universidad de Granada, E-18071 Granada, Spain

<sup>2</sup> Dpto de Ecología Funcional y Evolutiva, Estación Experimental de Zonas Áridas (EEZA-CSIC), E-04120 Almería, Spain

<sup>3</sup> Plant Biology, CIBIO/InBio, Centro de Investigação em Biodiversidade e Recursos Genéticos, Laboratório Associado, Universidade do Porto. Campus Agrário de Vairão, 4485-661 Vairão, Portugal

<sup>4</sup> Leibniz Institute for Plant Genetics and Crop Plant Research (IPK), 06466 Gatersleben, Germany. Present address: Global Institute for Food Security (GIFS), University of Saskatchewan, Saskatoon, Canada

<sup>5</sup> Dpto de Botánica, Universidad de Granada, E-18071 Granada, Spain

<sup>6</sup> Dpto de Genética, Universidad de Granada, E-18071 Granada, Spain

## **Supplementary Methods S1. Soil analyses**

During 2010 we collected soil samples under each individual plant to a depth of 0-10 cm. We measured moisture content at field capacity extracted in a pressure plate at -33 kPa according to Cassel and Nielsen<sup>1</sup>. We estimated the amount of exchangeable cations (Na, K<sup>+</sup>, Ca<sup>2+</sup>, Mg<sup>2+</sup>) by extracting with NH<sub>4</sub>OAc 1N. Ca<sup>2+</sup> and Mg<sup>2+</sup> were measured by atomic-absorption spectroscopy and Na and K<sup>+</sup> by flame photometry. Finally, total nitrogen was inferred by the Kjeldahl method<sup>2</sup>, and available P using the classical Olsen method<sup>3</sup>. Soil variables for plant in 2011 were estimated using kriging interpolation. Before PCA analysis, all variables were log-transformed in order to normalize and afterwards centered and scaled by their standard deviations.

## **Supplementary Methods S2. DNA amplifications**

### **trnL-trn-F**

PCR reactions for the chloroplastidial trnL-trnF intergenic spacer was performed using the tabC and tabF primers<sup>4</sup>. Following Abdelaziz et al.<sup>5</sup>, we performed these amplifications in 25 µL reaction volume, including 0.5 units of Taq DNA polymerase, 2,5 µL 10x reaction buffer containing MgCl<sub>2</sub> at 1.5 mM, 0.2 mM each dNTP (all product from New England BioLabs), 0.2 µM each primer, and 5 ng of template DNA. Amplification conditions included 35 cycles with 94 °C (15 s) denaturation, 58 °C (30 s) annealing, 72 °C (90 s) extension, and 72 °C (3 min) final extension.

### **Microsatellites**

The PCR reactions were performed in 14 µL reactions including 0.3 units of Taq polymerase (New England BioLabs), 1.4 µL 10× buffer containing MgCl<sub>2</sub> at 1.5 mM (New England BioLabs), 0.16 mM of each dNTP, 0.33 µM for each primer and 2.5 ng of DNA template. We used a Gradient Master Cycler Pro S (Eppendorf) to perform PCRs. The initial denaturing step consisted in 3 min at 94 °C, followed by 35 to 40 extension cycles consisting in 15 s at 94 °C, 30 s at a specific annealing



temperature (see table below) and 90 s at 72 °C, with a final extension step of 3 min at 72 °C.

**Table** Characteristics of the 10 microsatellite markers. Adapted from Muñoz-Pajares et al.<sup>6</sup>.

Locus	Primer sequence (5'- 3')	Repeat motif	Product size (bp)	T <sub>a</sub> (°C)
C5 (JF766210)	F: TCTTTCTTCTGCGGTTTATTC R: CGTTTTTGTGTGTTCTGG	CCA <sub>8</sub>	164 – 182	56
D2 (JF766211)	F: ACGGAAGATGACGATGATCGACTG R: CAATGTCCCTAATTGGTCAATGG	CAT <sub>23</sub>	117 – 189	54
D4 (JF766212)	F: TAAGGTGTTACCGGATTGTC R: GTGACGATTCGCTCCTTG	ATC <sub>7</sub>	200 – 215	57
E4 (JF766213)	F: CCTTCCTCCGACTACTCTCC R: TGAGCGACTGATGATGATTC	CT <sub>20</sub>	145 – 178	57
E8 (JF766214)	F: AGCTCACAGCCGTCGATGTTTGC R: GAGGTGAAATACACGTAGAACCT	CT <sub>50</sub>	157 – 229	50
D11 (JF766215)	F: TCCAGGGTCTGAGTCAATATG R: TTACCACTCCTTGCTTCTGAA	TCA <sub>14</sub>	179 – 197	53
E6 (JF766216)	F: CTTGTAACCGAGCCACTCA R: ATACGGAGAAGAAAGCGAATC	TC <sub>14</sub>	131 – 159	53
D10 (JF766217)	F: ACTGCCATCAAACGACCTC R: TTGGTTGGAAAAGGGATTG	TCA <sub>12</sub>	166 – 185	53
E5 (JF766218)	F: TCCATTTACACAATCCGTTTCAT R: CCAACCTGACATCTTGCTTC	GA <sub>13</sub>	167 – 195	50
E3 (JF766219)	F: TTCCTCCAGATGAAACTACACAGG R: ACTTACATCGGATCGGTTGAG	GA <sub>17</sub>	215 – 253	56

For each locus the GenBank accession number is given in parenthesis, forward (F) and reverse (R) primers, repeat motif, allele size ranges, and optimal annealing temperatures (T<sub>a</sub>) are indicated.

### Supplementary Methods S3. R code for bearing correlograms

The following script performs a bearing correlogram, an analytical tool for determining the directions of maximum and minimum spatial-genetic correlation<sup>7</sup>. This method weights the pairwise distance matrix for a set of directional angles. For a given directional angle ( $\theta$ ) the log-transformed distance between two points ( $\log r_{ij}$ ) is weighted by the squared cosine of the angle formed between their associated vector and the directional angle of study ( $\cos^2 (\theta - \alpha_{ij})$ ).

$$d_{ij} = \log r_{ij} \cos^2 (\theta - \alpha_{ij})$$

The correlation between the kinship and the resulting weighted distances is then tested via Mantel test, and the statistical significance obtained by through a permutational procedure. The script uses functions included in the packages 'geosphere'<sup>8</sup>, 'aspace'<sup>9</sup>, 'RandomFields'<sup>10</sup> and 'ecodist'<sup>11</sup>.

Usage:

bearing.analysis (K, coord, divisions = 64, perms=1000)

### Arguments:

K            a numeric matrix containing all paired kinship values  
coord       data frame with each individual coordinates  
divisions   a numeric variable denoting the number of angle divisions between the 0 and the 180  
             degrees from the (0,1) directional  
perms       a numeric variable defining the number of permutations used to obtain the statistical  
             significance of the Mantel test

### Code:

```
bearing.analysis <- function(Kij, coord, divisions=64, perms=1000)
{
  Kij <- as.matrix(Kij)

  # required libraries
  require(geosphere)
  require(aspace)
  require(ecodist)

  # Obtain the logarithm of the distance matrix
  log.Rij <- log(dist(coord))

  # set divisions in radians
  angle <- seq(0,360,360/divisions)
  angle <- angle[-length(angle)]

  # data frame to store for each angle mantel test data:
  main <- data.frame("angle"=angle, "radians"=NA, "r.Mantel"=NA, "p.val1"=NA, "p.val2"=NA, "p.val3"=NA,"color"=NA)

  # set angles to radians
  angle <- as_radians(angle)

  # set matrix to store the bearing angles for each pair of points
  alphas <- Kij; alphas[] <- NA

  # Loop for setting bearing angles in the M matrix
  for(i in 1:nrow(coord))
  {
    for(j in 1:nrow(coord))
    {
```

```

        # this returns angles in degrees
        alphas[i,j]<-bearingRhumb(coord[i,],coord[j,])
    }
}

# transform bearing degrees to radians.
alphas <- as_radians(alphas)

# Loop for each angle
for (t in 1:length(angle))
{
    print(paste("angle:",t))

    # Theta. Reference angle (in radians)
    main[t,"radians"] <- theta <- angle[t]

    # Transformed distance matrix
    cos2.ang <- cos(theta-alphas)^2
    Dij <- as.matrix(log.Rij)*cos2.ang

    # Mantel test using the kinship matrix (Kij) and the weighted spatial distance matrix (Rij).
    # Permutations (user defined in 'perms' argument) are used for significance.
    main[t,3:6] <- mantel(as.dist(Kij)~ as.dist(Dij), nperm = perms)[1:4]

    # Define colours for significance levels
    if(any(main[t,4:6] <= 0.05)) {main[t,7] <- "black"}
    else {main[t,7] <- "white"}

    print(paste("Mantel r:",round(m[1],4),"p-values:",m[2],m[3],m[4]))
}

# Plot bearing correlogram
if(max(main$r.Mantel)<0){m <- 0}
else {m <- max(main$r.Mantel)}
plot(main[,1],main[,3], axes=F,type="n", ylim=c(min(main$r.Mantel),m), ylab="Mantel r", xlab="angle due X-axis (degrees)")
lines(main$angle, main$r.Mantel, lty=3, col="gray50")
points(main[,1], main[,3], pch=22, col="black", bg=main$color, cex=.6)
axis(1, labels=seq(0,max(main$angle),22.5), at=seq(0,max(main$angle),22.5), cex.axis=.6)
axis(2, cex.axis=.6)
abline(h=0, lty=3, col="gray50")

return(main)
}

```

## SUPPLEMENTARY REFERENCES:

1. Cassel, D. K. & Nielsen, D. R. in *Methods of soil analysis. Part. 1: physical and mineralogical methods* (ed. Klute, A.) 901–926 (ASA-SSSA, 1986).

2. Bremner, J. M. in *Methods of soil analysis. Part. 2. Chemical and microbiological properties* (eds. Black, C. A., Evans, D. D., Esminger, L. E. & Clark, F. E.) 1324–1345 (American Society of Agronomy, 1965).
3. Olsen, S. R. & Sommers, L. E. in *Methods of Soil Analysis. Chemical and Microbiological Properties* (ed. Page, D. L.) 403–430 (American Society of Agronomy and Soil Science Society of America, 1982).
4. Taberlet, P., Gielly, L., Pautou, G. & Bouvet, J. Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Mol. Biol.* **17**, 1105–1109 (1991).
5. Abdelaziz, M. *et al.* Using complementary techniques to distinguish cryptic species: a new *Erysimum* (Brassicaceae) species from North Africa. *Am. J. Bot.* **98**, 1049–60 (2011).
6. Muñoz-Pajares, a J. *et al.* Characterization of microsatellite loci in *Erysimum mediohispanicum* (Brassicaceae) and cross-amplification in related species. *Am. J. Bot.* **98**, e287–9 (2011).
7. Rosenberg, M. S. The bearing correlogram: a new method of analyzing directional spatial autocorrelation. *Geogr. Anal.* **32**, 267–278 (2000).
8. Hijmans, R. J. *geosphere: Spherical Trigonometry. R package version 1.4-3.* (2015). at <<http://cran.r-project.org/package=geosphere>>
9. Bui, R., Buliung, R. N. & Rummel, T. K. *aspace: A collection of functions for estimating centrographic statistics and computational geometries for spatial point patterns. R package version 3.2.* (2012). at <<http://cran.r-project.org/package=aspace>>
10. Schlather, M., Malinowski, A., Menck, P. J., Oestin, M. & Stokor, K. Analysis, simulation and prediction of multivariate random fields with package randomfields. *J. Stat. Softw.* **63**, 1–25 (2015).
11. Goslee, S. C. & Urban, D. L. The ecodist package for dissimilarity-based analysis of ecological data. *J. Stat. Softw.* **22**, 1–19 (2007).

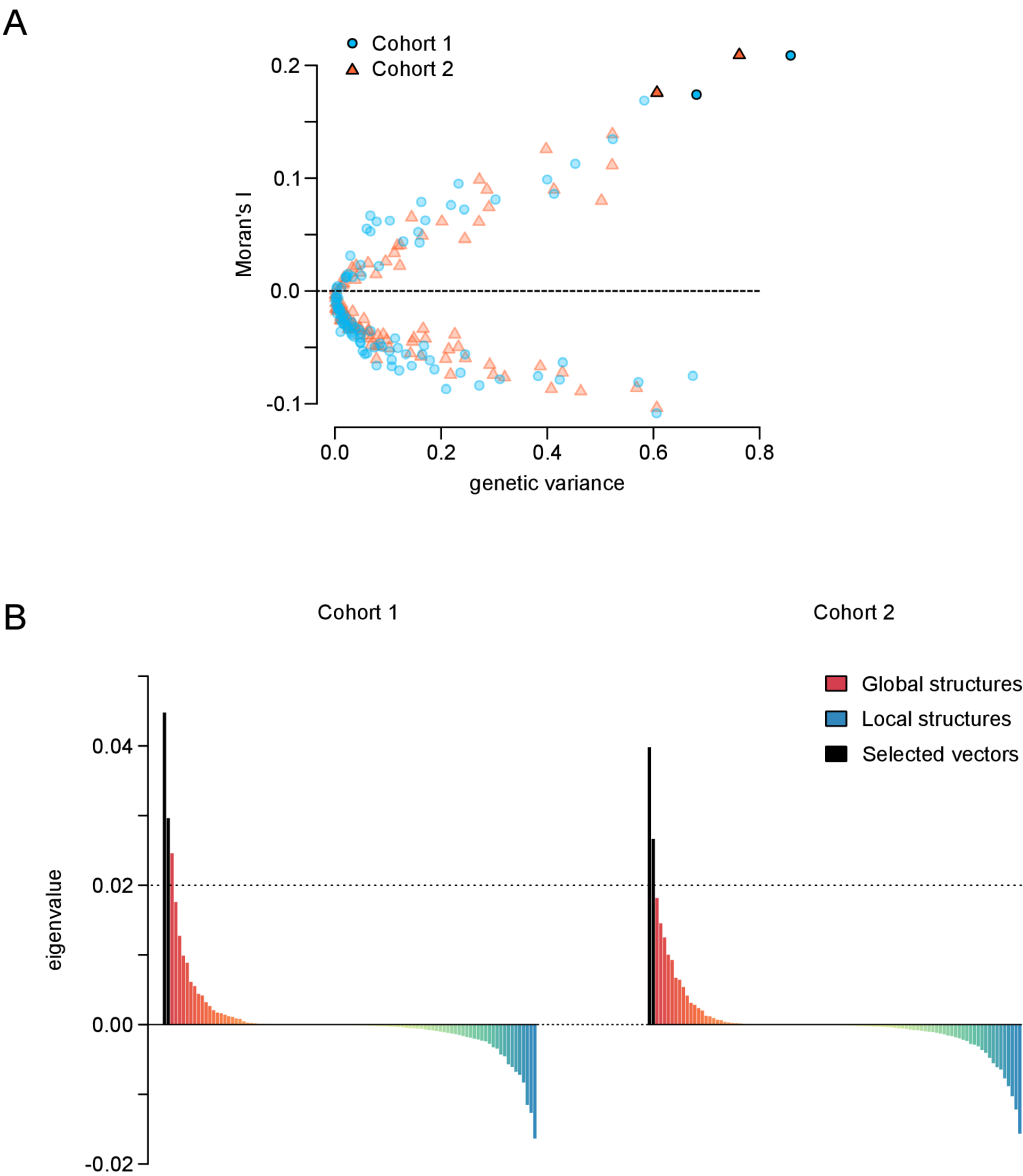
**Table S1** Model selection for spatially lagged autoregressive models for sPC1 and sPC2. Best-ranked models are listed in ascendant AIC showing parameter estimates and values of Moran's I test over the residuals.

PC1	Light availability	Anions (N, P)	Cations 1 (Mg <sup>2+</sup> , K <sup>+</sup> )	Cations 2 (Na <sup>+</sup> )	Field capacity	$\rho$	Moran's I	AICc	LogLikelihood	w <sub>i</sub>
2010	<b>0.165</b>		<b>-0.075</b>		<b>-0.075</b>	<b>0.888</b>	0.025	91.27	-40.316	0.357
	<b>0.169</b>		-0.074	0.019	<b>-0.074</b>	<b>0.883</b>	0.023	93.144	-40.12	0.14
	<b>0.144</b>					<b>0.88</b>	<b>0.03</b>	92.509	-43.129	0.192
	<b>0.143</b>				-0.042	<b>0.875</b>	<b>0.027</b>	92.693	-42.136	0.175
	<b>0.173</b>	-0.026	-0.089		<b>-0.07</b>	<b>0.886</b>	<b>0.025</b>	93.222	-40.159	0.135
2011	0.069	<b>-0.1</b>				<b>0.869</b>	<b>0.026</b>	86.77	-39.175	0.232
			<b>0.117</b>			<b>0.878</b>	<b>0.032</b>	86.996	-40.373	0.207
	0.038		<b>0.101</b>			<b>0.866</b>	<b>0.028</b>	88.219	-39.899	0.112
		<b>-0.099</b>				<b>0.905</b>	<b>0.044</b>	88.202	-40.976	0.113
		-0.044	0.082			<b>0.879</b>	<b>0.031</b>	88.56	-40.07	0.095
	0.059	-0.079	0.032			<b>0.861</b>	<b>0.024</b>	88.788	-39.075	0.084
	0.068	-0.089			-0.019	<b>0.865</b>	<b>0.025</b>	88.922	-39.142	0.079
	0.068	<b>-0.102</b>		-0.009		<b>0.868</b>	<b>0.026</b>	88.951	-39.156	0.078
PC2	Light availability	Anions (N, P)	Cations 1 (Mg <sup>2+</sup> , K <sup>+</sup> )	Cations 2 (Na <sup>+</sup> )	Field capacity	$\rho$	Moran's I	AICc	LogLikelihood	w <sub>i</sub>
2010		-0.072	<b>-0.098</b>	<b>-0.127</b>		<b>0.898</b>	<b>0.045</b>	70.497	-29.93	0.2
	<b>-0.07</b>			<b>-0.125</b>		<b>0.904</b>	<b>0.06</b>	70.81	-31.195	0.171
	-0.057		-0.041	<b>-0.125</b>		<b>0.893</b>	<b>0.052</b>	71.235	-30.298	0.138
	-0.04	-0.057	-0.079	<b>-0.131</b>		<b>0.893</b>	<b>0.047</b>	71.541	-29.319	0.119
		<b>-0.083</b>	<b>-0.089</b>	<b>-0.127</b>	0.03	<b>0.901</b>	<b>0.049</b>	71.888	-29.492	0.1
	<b>-0.069</b>			<b>-0.123</b>	0.03	<b>0.903</b>	<b>0.061</b>	71.735	-30.548	0.108
			-0.055	<b>-0.117</b>		<b>0.901</b>	<b>0.052</b>	71.803	-31.691	0.104
	-0.041	-0.068	-0.069	<b>-0.13</b>	0.031	<b>0.896</b>	<b>0.052</b>	72.915	-28.849	0.06
2011			-0.049			<b>0.927</b>	<b>0.098</b>	74.035	-33.893	0.173
					0.052	<b>0.928</b>	<b>0.095</b>	74.354	-34.052	0.147
				-0.037		<b>0.927</b>	<b>0.093</b>	74.954	-34.352	0.109
	-0.024					<b>0.927</b>	<b>0.092</b>	75.17	-34.46	0.098
		0.018				<b>0.927</b>	<b>0.093</b>	75.431	-34.59	0.086
			-0.046	-0.031		<b>0.927</b>	<b>0.097</b>	75.726	-33.653	0.074
		-0.035	-0.076			<b>0.929</b>	<b>0.101</b>	75.775	-33.677	0.072
				-0.029	0.047	<b>0.927</b>	<b>0.094</b>	76.111	-33.845	0.061
	-0.009		-0.045			<b>0.927</b>	<b>0.097</b>	76.15	-33.864	0.06
			-0.039		0.014	<b>0.927</b>	<b>0.097</b>	76.174	-33.876	0.059
	-0.019				0.049	<b>0.927</b>	<b>0.095</b>	76.199	-33.889	0.059

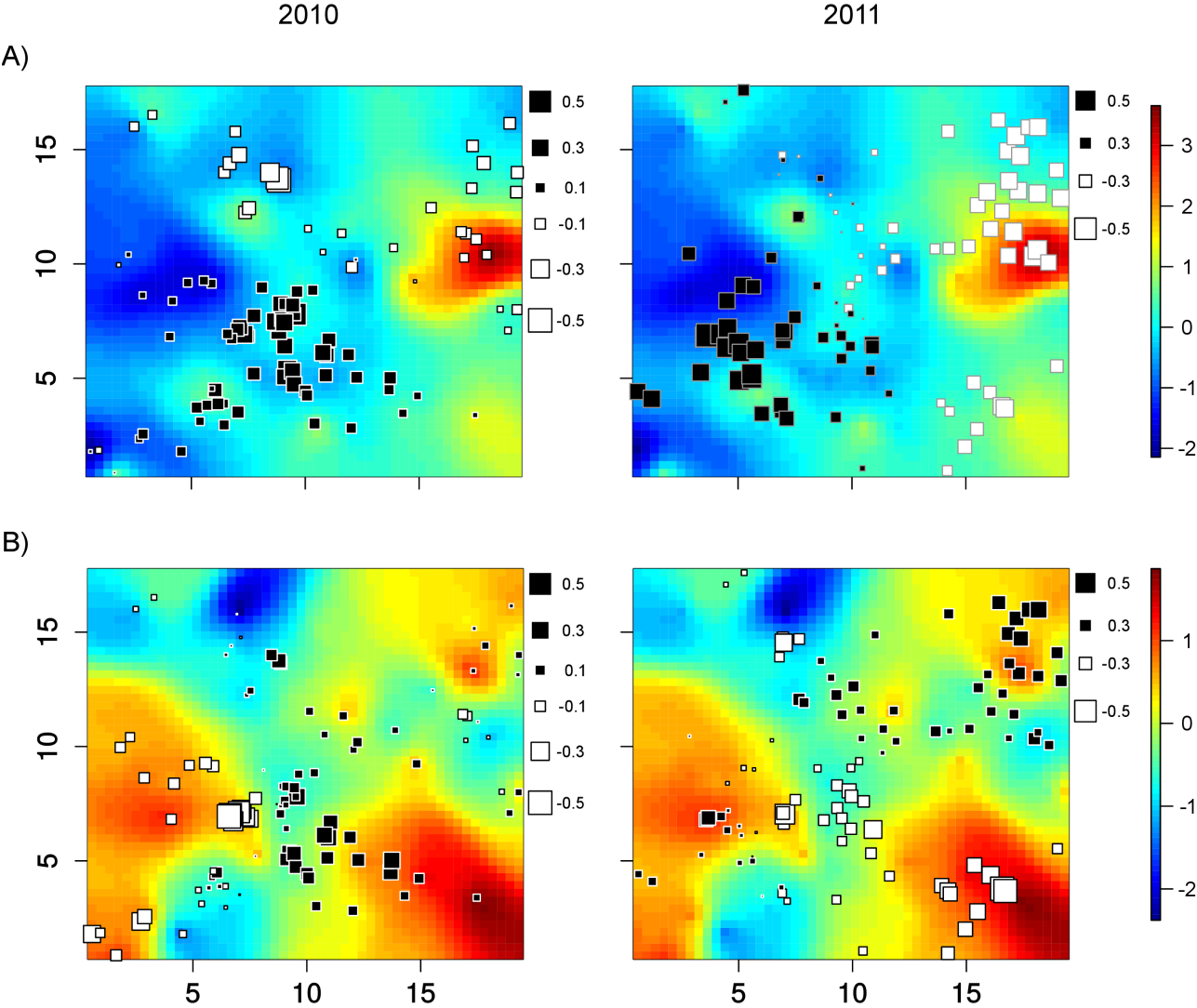
Abbreviations: AICc, second order AIC; w<sub>i</sub>, Akaike weight of each model. Significant values of parameters estimates and of Moran's I are in bold.



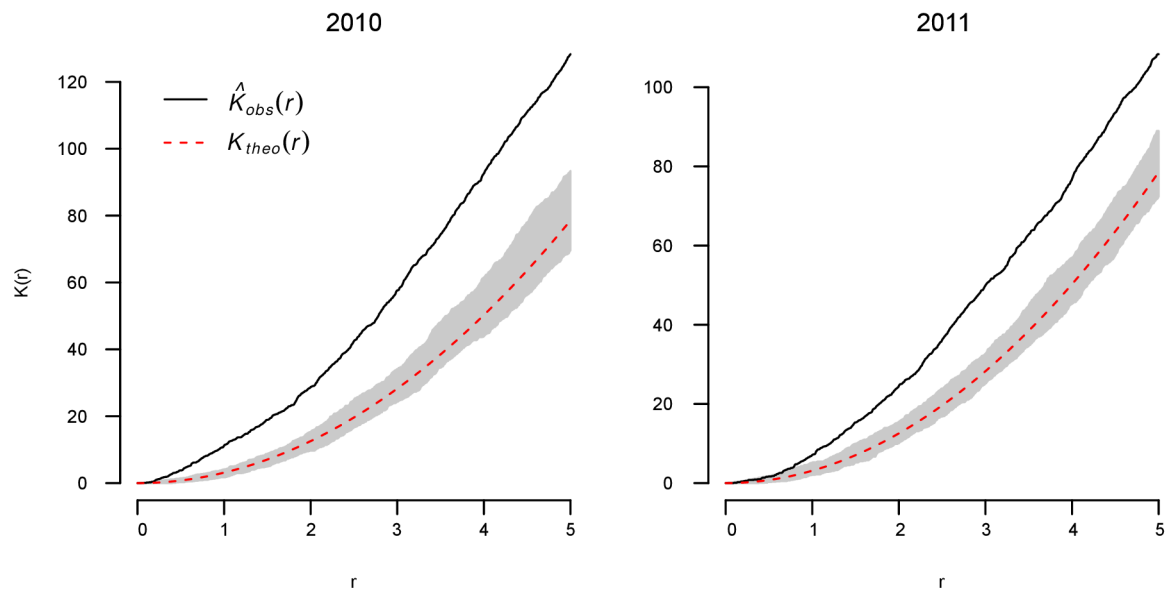
**Figure S1** Eigenvalues of sPCA for cohorts 1 and 2. A) Eigenvalues decomposition into the genetic variance explained and Moran's I values. B) Barplot of eigenvalues showing positive (global structures) and negative (local structures) axes. In both, the selected sPCs for the spatial analyses are highlighted.



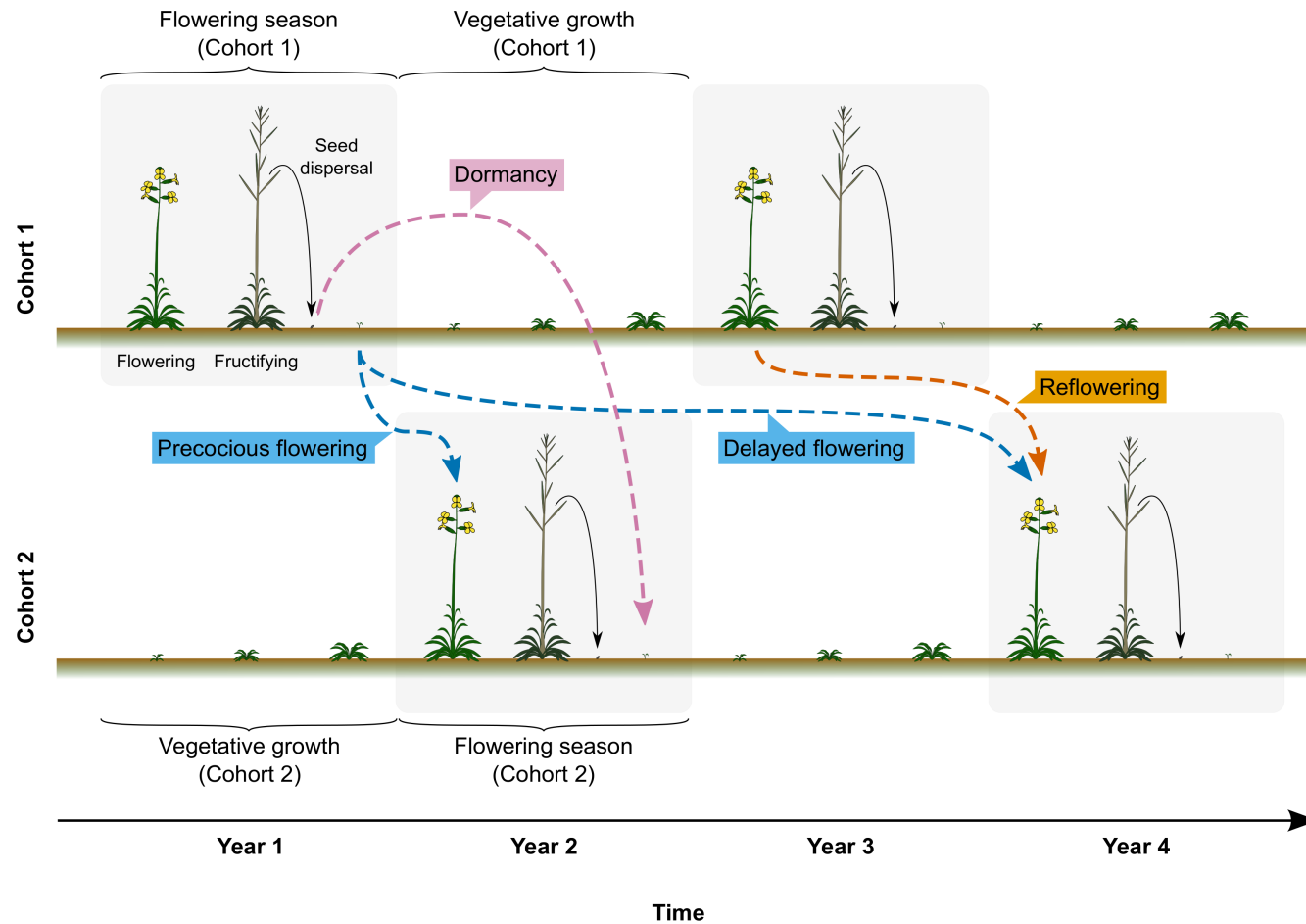
**Figure S2** Lagged principal scores and inferred values of environmental variables. A) PC1 and anions. B) PC2 and cations 1.



**Figure S3** Ripley's K functions for both cohorts



**Figure S4** Pathways of inter-annual migration. Life cycle of two flowering cohorts (cohort 1 and cohort 2) are shown along four years (year 1 to 4). Dashed lines define the different hypothetical pathways taken by an inter-annual migrant from the flowering season of the cohort 1 to that of cohort 2.



**Figure S5** Biplot of the PCA performed over the soil cations. Circles and triangles depict the scores of the observations on the principal components 1 and 2 (cations 1 and cations 2). The coefficients of the variables are marked with arrows.

