

ORIGINAL ARTICLE

# Fine-scale spatial genetic structure and gene dispersal in *Silene latifolia*

M Barluenga<sup>1,2,3</sup>, F Austerlitz<sup>4</sup>, JA Elzinga<sup>1,5</sup>, S Teixeira<sup>1,6</sup>, J Goudet<sup>1</sup> and G Bernasconi<sup>1,7</sup>

<sup>1</sup>Department of Ecology and Evolution, University of Lausanne, Lausanne, Switzerland; <sup>2</sup>Institute of Ecology and Evolution, University of Bern, Bern, Switzerland; <sup>3</sup>EAWAG, Department of Fish Ecology and Evolution, Center of Ecology, Evolution and Biochemistry, Kastanienbaum, Switzerland; <sup>4</sup>Evolution et Systématique, Université Paris-Sud XI, Orsay Cedex, France; <sup>5</sup>Department of Biological and Environmental Sciences, University of Jyväskylä, Jyväskylä, Finland; <sup>6</sup>CCMAR, Grupo MAREE, Universidade do Algarve, Faro, Portugal and <sup>7</sup>Institute of Biology, University of Neuchâtel, Neuchâtel, Switzerland

Plants are sessile organisms, often characterized by limited dispersal. Seeds and pollen are the critical stages for gene flow. Here we investigate spatial genetic structure, gene dispersal and the relative contribution of pollen vs seed in the movement of genes in a stable metapopulation of the white campion *Silene latifolia* within its native range. This short-lived perennial plant is dioecious, has gravity-dispersed seeds and moth-mediated pollination. Direct measures of pollen dispersal suggested that large populations receive more pollen than small isolated populations and that most gene flow occurs within tens of meters. However, these studies were performed in the newly colonized range (North America) where the specialist pollinator is absent. In the native range (Europe), gene dispersal could fall on a different spatial scale. We genotyped 258 individuals from large and small (15) subpopulations along a 60 km, elongated meta-

population in Europe using six highly variable microsatellite markers, two X-linked and four autosomal. We found substantial genetic differentiation among subpopulations (global  $F_{ST} = 0.11$ ) and a general pattern of isolation by distance over the whole sampled area. Spatial autocorrelation revealed high relatedness among neighboring individuals over hundreds of meters. Estimates of gene dispersal revealed gene flow at the scale of tens of meters (5–30 m), similar to the newly colonized range. Contrary to expectations, estimates of dispersal based on X and autosomal markers showed very similar ranges, suggesting similar levels of pollen and seed dispersal. This may be explained by stochastic events of extensive seed dispersal in this area and limited pollen dispersal.

*Heredity* (2011) **106**, 13–24; doi:10.1038/hdy.2010.38; published online 14 April 2010

**Keywords:** isolation by distance; gene flow; microsatellites; seed dispersal; spatial autocorrelation; pollen dispersal

## Introduction

Plants are sessile organisms, often characterized by limited dispersal. Pollen and seed dispersal are the critical stages for gene flow, often showing quite different modes and distances (McCauley, 1994, 1997; Heuertz *et al.*, 2003; García *et al.*, 2007). Pollen-mediated gene movement is often important for genetic diversity within and among plant populations (Latta *et al.*, 1998; Streiff *et al.*, 1998; Imbert and Lefevre, 2003). Seed dispersal is an important determinant for colonization of new sites, extinction of local populations and migration among neighboring subpopulations (Ouborg *et al.*, 1999; Cain *et al.*, 2000). In spite of having profound demographic and genetic consequences for plant populations, the relative contribution of contemporary pollen and seed dispersal is still poorly investigated (Macdonald and Johnson, 2001) although some reliable data have recently

been accumulated (Ravigné *et al.*, 2006; García *et al.*, 2007).

Restricted gene flow should thus frequently result in spatial genetic structure of natural plant populations, that is, the nonrandom spatial distribution of genotypes (Wright, 1943, 1978; Turner *et al.*, 1982). Conversely, the study of fine-scale genetic structure within populations allows to gain insights into the dispersal behavior of genes and individuals (Queller and Goodnight, 1989; Smouse and Peakall, 1999; Rousset, 2000). Whereas wind-pollinated plants experience wide-range and un-directed pollen dispersal, in animal-pollinated plants gene flow can be more limited, depending on the behavior of the animal disperser (García *et al.*, 2007). Similarly, although in most species seeds are dispersed through gravity over very short distances (Howe and Smallwood, 1982; Ouborg *et al.*, 1999; Cain *et al.*, 2000), seed dispersal ranges can be substantially higher in anemochorous (Epperson and Allard, 1989; Tackenberg, 2003; Tackenberg *et al.*, 2003) or zoochorous species (Chung *et al.*, 2000). Self-fertilizing or clonal species are expected to have lower pollen movement and therefore higher genetic structure than obligatory out-crossing (for example, dioecious) species (Hamrick and Godt, 1996), as well as smaller effective populations sizes (Ingvarsson,

Correspondence: Professor G Bernasconi, Institute of Biology, University of Neuchâtel, Rue Emile Argand 11, Neuchâtel 2009, Switzerland.

E-mail: giorgina.bernasconi@unine.ch

Received 18 February 2009; revised 22 December 2009; accepted 15 January 2010; published online 14 April 2010

2002). Finally, rare or narrow endemic species might experience lower effective population sizes and higher population fragmentation than widespread ones, producing greater genetic differentiation between populations (Hamrick and Godt, 1996).

Here we investigate gene dispersal patterns in a metapopulation of the white campion *Silene latifolia* Poiret (Caryophyllaceae), within its native range in Eurasia. This short-lived perennial weed mainly grows in open and disturbed habitats, such as roadsides or cultivated areas (Baker, 1947) and thus is prone to experience frequent bouts of local extinctions and recolonizations. This weed has sex determination, the male being the heterogametic sex (Ironsides and Filatov, 2005). In this species, floral sex ratios in natural populations are often male biased, because of a larger floral display of males compared to females. *S. latifolia* is dioecious, and as such, obligatory out-crossing. Population persistence thus requires relatively large both effective population sizes and pollen movement. The white campion is pollinated by insects, mostly by the specialist pollinator/seed predator *Hadena bicruris*, Noctuidae (Blair and Wolfe, 2004; Bopp and Gottsberger, 2004; Jolivet and Bernasconi, 2006; Kephart *et al.*, 2006). Seeds are gravity dispersed (Bari, 1969; Moyle, 2006). However, this weed, which can thrive in a wide array of conditions, profits from human activity for long-distance seed dispersal. According to its life history, *S. latifolia* is thus expected to have poor seed dispersal at a local scale, and potentially greater pollen movement mediated by pollinator foraging behavior and flight distances (see McCauley, 1994; McCauley, 1997). Population size and isolation appear to influence pollination opportunities, as some pollinators behave differently in small plant populations than in large ones (Richards *et al.*, 1999). Pollinators in general show increased visitation of flowers with higher density of flowers, and with increasing population size (Fritz and Nilsson, 1994; Karron *et al.*, 1995; Mustajärvi *et al.*, 2001). This should result in low genetic variability in small and/or isolated populations and higher variability in larger and/or well-connected populations. Direct measures of pollen dispersal in experimental settings in *S. latifolia* (Richards *et al.*, 1999) and analysis of within-fruit paternity in natural populations (Teixeira and Bernasconi, 2007) suggest that large populations receive more pollen, and that this pollen has more diverse origins. However, we lack studies on spatial genetic structure of *S. latifolia* in its native range, with most of the population genetic studies to date focusing on North America (McCauley, 1994, 1997; Moyle, 2006; Taylor and Keller, 2007), a recently colonized area (ca 200 years) where the main pollinator

(the moth *H. bicruris*) is absent, and the species is pollinated by other insects (Young, 2002; Barthelmess *et al.*, 2006). In the presence of potentially coevolved pollinators in Europe, pollen dispersal could fall on a different spatial scale (see Wright and Meagher, 2004). Therefore, getting some knowledge about the spatial genetic structure and patterns of gene flow of *S. latifolia* in its native range is an important precondition to interpret the evolutionary changes that might have occurred in the newly colonized environments.

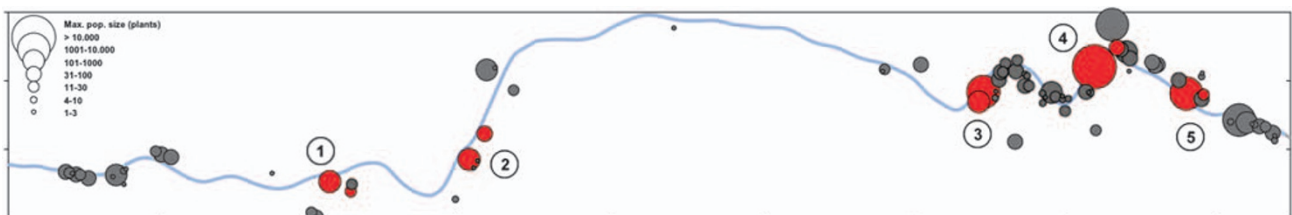
In this study, we estimated the fine-scale population structure along a stable *S. latifolia* metapopulation in Europe stretching over approximately 60 km, using six highly variable species-specific microsatellite markers (Teixeira and Bernasconi, 2007). First, we investigated the scale of gene flow and population differentiation within this large metapopulation and tested for isolation by distance (IBD; Slatkin, 1993). We addressed how the size and the degree of isolation of populations affect their genetic structure. Second, we investigated the patterns of gene dispersal at the individual level revealed by multivariate spatial autocorrelation analysis (Smouse and Peakall, 1999) and compared them with the available direct measures of pollen dispersal in North American populations. Finally, we estimated the relative contributions of pollen-mediated and seed-mediated gene flow by comparing the patterns revealed by autosomal vs X-linked microsatellite markers.

## Materials and methods

### Sample collection

In the summer of 2006, we sampled leaves from 258 *S. latifolia* plants in a large linear metapopulation in an area of river plains and grassland along the River Waal, a tributary of the River Rhine, in the Netherlands. Over 3 consecutive years (2001–2003) we mapped all occurrences of *S. latifolia* in this region, defining as patches all groups of plants separated by  $\geq 100$  m, and estimating the number of flowering plants in each patch (Elzinga, 2005; Elzinga *et al.*, 2005, 2007).

We designed a sampling scheme allowing the investigation of genetic structure at multiple scales within this metapopulation. We collected samples in five distinct regions spanning 60 km (from west to east: 1, Zalt; 2, Tiel; 3, Nij; 4, Mil and 5, Lob; Figure 1). Within each region we selected two plant patches, one larger (min = 130 and max = 13 300 flowering individuals), and one smaller (min = 15 and max = 135 flowering individuals), separated on average by 1.5 km, whereby within regions the 'large' patch had always more flowering individuals



**Figure 1** Map of the *S. latifolia* metapopulation along the River Waal (blue line) in the Netherlands. Circles represent individual populations separated by 100 m or more. Size of circles resembles the size of the populations. Red circles denote sampled populations (numbers represent the five sampled regions: 1, Zalt; 2, Tiel; 3, Nij; 4, Mil; 5, Lob) and gray circles other *S. latifolia* populations not included in this study.

than the 'small' one. We will further refer to these as 'populations within regions'. Within each large population (L) we sampled 15–20 plants in two independent transects (L1 and L2) separated from each other by a maximum of 500 m. Within each small population (S) we sampled 15–20 plants on a single transect. We will further refer to these transects as 'subpopulations'. Distances among subpopulations within and among regions are given in Supplementary Table 2. Transects were linear and 30–250 m long, depending on plant density. We recorded the relative position of each sampled plant as polar coordinates (distance and angle) relative to the starting point of the transect. When plants were not on the transect but close to them, we determined their relative positions by additionally recording the normal distance to the transect with a measuring tape. Average interplant distances within transects ranged from 2 to 12 m. We further recorded the geographical coordinates of the starting point of each transect with a GPS (GISbackpack; Cometius, Leiderdorp, the Netherlands). This allowed converting the relative position of each plant into exact coordinates. If the plant was flowering, we also recorded the sex. Leaves of individual plants were preserved in bags with silica gel for genetic analyses.

#### DNA extraction and microsatellite amplification

Total DNA was extracted from leaf tissue using the Qiagen BioSprint DNA Plant kit (Valencia, CA, USA). Six nuclear microsatellite loci designed for *S. latifolia* were genotyped, four autosomal (SI1, SI4, SI6 and SI8) and two X-linked (SI14 and SI15). Primer sequences and amplification conditions followed (Teixeira and Bernasconi, 2007). Microsatellites were amplified with fluorescently labeled forward primers (FAM, HEX and TAMRA dyes) and fragment length was analyzed with the internal size marker Genescan-500 ROX (Applied Biosystems, Foster City, CA, USA) with an ABI 3100 Automated Sequencer (Applied Biosystems), and scored with Genescan version 3.7 and Genotyper version 3.7 (Applied Biosystems) software packages.

#### Population genetic analyses

We used FSTAT version 2.9.3 (Goudet, 2002) to determine the number of alleles and fixation index (F) and whether these significantly differed from 0, and to calculate allelic richness and gene diversity, averaged across loci for each subpopulation, as well as genetic differentiation among them. We compared large (L) vs small (S) subpopulations to address the effect of population size on genetic diversity. *S. latifolia* patches are not uniformly distributed over the study area, with more and larger populations concentrated on the eastern compared to the western zone (Figure 1). Therefore, we compared differentiation of the genetic diversity indices between the western (Zalt and Tiel) and eastern (Nij, Mil and Lob) zones with FSTAT. In addition, for each subpopulation we plotted the saturation curves of absolute allele number per locus as a function of number of individuals, to test for effects of sample size on allele detection. Samples were added following the position in the transect. Linkage disequilibrium was tested for all possible pairs of loci in each population and globally for each pair of loci across populations with Arlequin version 3.0 (Excoffier *et al.*, 2005).

Genetic differences between populations were estimated with pairwise F-statistics (Weir and Cockerham, 1984) based on differences in allele frequencies, and their significance levels were estimated with 10 000 permutations as implemented in Arlequin. To determine how neutral genetic variation was partitioned among and within regions, we estimated a hierarchical nesting of genetic diversity using the analysis of molecular variance approach of Excoffier *et al.* (1992) with Arlequin. Nesting levels considered were (1) among regions, (2) among subpopulations within regions and (3) within subpopulations.

Critical significance levels in all analyses were corrected for multiple testing following the sequential Bonferroni procedure (Rice, 1989).

A principal coordinate analysis (PCoA) of individual diploid genotypes was performed based on pairwise  $F_{ST}$  between subpopulations to examine their genetic clustering using Genalex version 6.1 (<http://www.anu.edu.au/BoZo/GenALEX/>). Tests of the significance of total inertia as well as individual axes inertia was calculated with 1000 permutations with PCAGEN version 1.2 (Goudet, 2005).

#### Isolation by distance, spatial autocorrelation and rate of pollen vs seed dispersal

Spatial genetic structure was assessed both at the population level and at the individual level. At the population level, we tested for IBD by estimating the correlation between the matrix of genetic distance between pairs of populations ( $F_{ST}/(1-F_{ST})$ ) and the matrix of geographic distances, and tested the significance of the correlation using a Mantel test with 10 000 permutations (Smouse *et al.*, 1986) as implemented in Arlequin. Geographic distances between subpopulations were computed as the minimum distance between the calculated coordinates for each subpopulation. We also obtained the neighborhood size ( $N_b$ ) at among-population level for both autosomal and X-linked markers as the inverse of the slope of the regression of ( $F_{ST}/(1-F_{ST})$ ) over the logarithm of distance between populations (Rousset, 1997). This neighborhood size is  $N_b^A = 4\pi N_e \sigma_A^2$  for autosomes and  $N_b^X = 3\pi N_e \sigma_X^2$  for X-linked markers (because there are only 1.5 X-linked genes for two autosomal genes), where  $N_e$  is the average effective population size across populations and  $\sigma_A^2$  and  $\sigma_X^2$  are the axial variances of gene dispersal for autosomal and X-linked markers, respectively. We assumed  $N_e$  values of 100 or 1000 and obtained thus estimates of  $\sigma_A^2$  and  $\sigma_X^2$  at the among-population level.

To investigate spatial genetic structure at the individual level, we used a spatial autocorrelation approach for multiallelic codominant markers (Smouse and Peakall, 1999; Peakall *et al.*, 2003; Double *et al.*, 2005). These analyses were performed using both Genalex and SPAGeDI version 1.2 (Hardy and Vekemans, 2002). Genalex calculates an autocorrelation coefficient  $r$  using the matrices containing pairwise geographic distances and pairwise squared genetic distances for codominant data (following Peakall *et al.*, 1995; Smouse and Peakall, 1999). The autocorrelation coefficient is calculated for a specified number of distance classes, and provides a measure of the genetic similarity between pairs of individuals whose geographic separation falls within the specified distance class. Pairwise

individual-by-individual geographic distances were calculated from the individual coordinates. We defined distance classes containing even sample size (a minimum of 250 pairwise comparisons) to increase the robustness of results. We performed an analysis with geographic distances not exceeding 1.5 km to measure genetic structure within regions. We then plotted the calculated autocorrelation coefficients  $r$  as a function of distance to produce spatial genetic autocorrelograms. Tests for statistical significance were performed using two methods: random permutations of individuals and bootstrapping across loci estimates of  $r$ , with the number of permutations and bootstraps set to 1000.

Spatial genetic structure was also assessed following the procedure described by Vekemans and Hardy (2004) based on pairwise kinship coefficients between individuals (Loiselle *et al.*, 1995) using the software *SPAGeDI*. To visualize the spatial genetic structure, we averaged pairwise kinship coefficients over a set of distance classes, and plotted against the distance. We calculated 95% confidence intervals by 1000 permutations over loci. This analysis was performed for all microsatellite loci together, and for X-linked and autosomal markers separately. To estimate gene dispersal and the rate of seed and pollen dispersal, we regressed pairwise kinship coefficients on the logarithm of the pairwise distances to provide the regression slope ( $b$ ), and Wright's neighborhood size was estimated as  $N_b = -(1-F)/b_r$ , where  $b_r$  is the restricted regression slope on the range within regions and  $F$  is the inbreeding coefficient.

We used here the availability of both autosomal and X-linked markers (Teixeira and Bernasconi, 2007) to estimate gene dispersal parameters for both types of markers and a ratio of pollen vs seed dispersal. This stems from the fact that the neighborhood size ( $N_b$ ) is related for any genetic system to the axial standard deviation of gene dispersal ( $\sigma^2$ ) through the relation  $N_b = 2\pi D_{eg}\sigma^2$ , where  $D_{eg}$  is the effective number of gene copies in the population for the system considered. For an autosomal locus in a diploid species with balanced sex ratio, this number is  $D_{eg}^A = 2D_e$ , where  $D_e$  is the effective density of individuals in the population. For an X-linked locus, the relation becomes  $D_{eg}^X = \frac{3}{2}D_e$ , as males carry only one copy of the X-linked gene.

We estimated gene dispersal using the iterative procedure provided by *SPAGeDI* (Vekemans and Hardy, 2004; Hardy *et al.*, 2006), but it did not always converge here, as it occurs in some cases according to the *SPAGeDI* manual. Thus, we simply used the relation  $\sigma^2 = N_b / (4\pi D_e)$ . Owing to the patchy distribution of this weed, we used two different values of density representing the upper and lower boundaries of the observed patterns (0.007–0.25 individuals per m<sup>2</sup>, corresponding to the density estimates over the entire area vs within patches, respectively), which were corrected for X-linked markers (three-fourths of the effective population size, as there is only 1.5 gene copies for an X-linked gene vs 2 for an autosomal gene).

To calculate the ration of pollen vs seed dispersal, we considered the possibility that sex ratios are unbalanced. Then the effective number of gene copies in the population becomes

$$D_{eg}^A = \frac{8D_f D_m}{D_f + D_m} \quad (1)$$

and

$$D_{eg}^X = \frac{9D_f D_m}{D_f + 2D_m} \quad (2)$$

where  $D_f$  is the effective density of female individuals in the population and  $D_m$  the effective density of male individuals. These formulas are the direct transposition to densities of the formulas for effective population sizes given in Wright (1939) and recalled in Segurel *et al.* (2008).

As  $N_b^A = 2\pi D_{eg}^A \sigma_A^2$  for autosomal markers, and  $N_b^X = 2\pi D_{eg}^X \sigma_X^2$  for X-linked markers, where  $\sigma_A^2$  and  $\sigma_X^2$  are the axial variances of gene dispersal for autosomal and X-linked markers, respectively, we can write the ratio between these two quantities as

$$\frac{N_b^X}{N_b^A} = \frac{D_{eg}^X \sigma_X^2}{D_{eg}^A \sigma_A^2} = \frac{9(D_f + D_m)}{8(D_f + 2D_m)} \frac{\sigma_X^2}{\sigma_A^2} \quad (3)$$

(3) can be transformed into

$$\frac{N_b^X}{N_b^A} = \frac{D_{eg}^X \sigma_X^2}{D_{eg}^A \sigma_A^2} = \frac{9(1 + \beta)}{8(2 + \beta)} \frac{\sigma_X^2}{\sigma_A^2} \quad (4)$$

where  $\beta = D_f/D_m$  is the sex ratio. Because the dispersal parameter ( $\sigma^2$ ) is composed of a seed and a pollen dispersal component, with respective variances  $\sigma_s^2$  and  $\sigma_p^2$ , the dispersal variance for autosomal markers is known to translate into  $\sigma_A^2 = \sigma_s^2 + (\sigma_p^2/2)$  (Crawford, 1984), because all genes disperse through seeds, but only half of them disperse through pollen. For X-linked markers, we can write the relation as  $\sigma_X^2 = \sigma_s^2 + (\sigma_p^2/3)$ , because all genes disperse through seeds, but only one third of them (those in males) will also disperse through pollen. Thus, Eq. (4) can be further transformed into

$$\frac{N_b^X}{N_b^A} = \frac{9(1 + \beta)}{8(2 + \beta)} \frac{(6\sigma_s^2 + 2\sigma_p^2)}{(6\sigma_s^2 + 3\sigma_p^2)} \quad (5)$$

Finally, by denoting  $\alpha = \sigma_p^2/\sigma_s^2$  the ratio of seed over pollen dispersal, we have

$$\frac{N_b^X}{N_b^A} = \frac{3(1 + \beta)}{4(2 + \beta)} \frac{(3 + \alpha)}{(2 + \alpha)} \quad (6)$$

This relation can also be expressed by solving for  $\alpha$  as

$$\alpha = \frac{9(1 + \beta) - 8(2 + \beta)(N_b^X/N_b^A)}{(-3(1 + \beta) + 4(2 + \beta))(N_b^X/N_b^A)} \quad (7)$$

Thus, in theory, assuming a given value for the sex ratio ( $\beta$ ) and estimating  $N_b^X$  and  $N_b^A$  from genetic data, Eq. (5) provides an estimate from  $\alpha$ . One difficulty, however, is that  $\alpha$  is by definition positive (ratio of two positive quantities), whereas Eq. (5) can yield negative values. For the right hand side of Eq. (5) to be positive, the ratio  $N_b^X/N_b^A$  has to be bounded between two values  $r_{\min}$  and  $r_{\max}$ , where

$$r_{\min} = \frac{3(1 + \beta)}{4(2 + \beta)} \quad (8)$$

and

$$r_{\max} = \frac{9(1 + \beta)}{8(2 + \beta)} \quad (9)$$

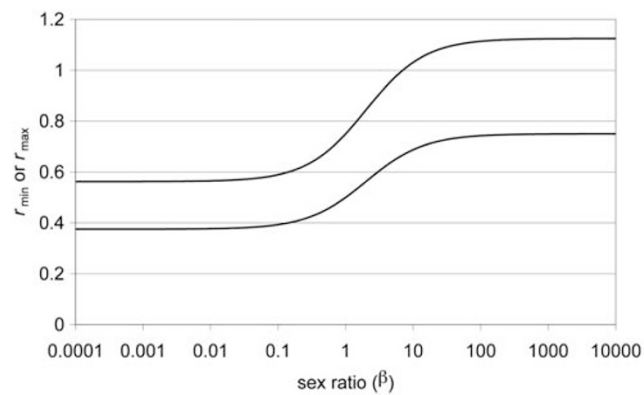
These boundaries are plotted in Figure 2 as a function of  $\beta$ . They appear as quite narrow, and their value clearly

depends on  $\beta$ . Here, we estimated the  $N_b^X$  and  $N_b^A$  values with SPAGeDI as explained above. We also estimated the standard error of these estimates ( $s_{eX}$  and  $s_{eA}$ ) using the jackknifing procedure over loci proposed in SPAGeDI. This provided us with approximate 95% confidence intervals ( $N_{b,min}^X, N_{b,max}^X$ ) and ( $N_{b,min}^A, N_{b,max}^A$ ), by adding or subtracting  $1.96 \sigma_e$  to the estimated values. We applied Eq. (7) to all possible combinations of  $N_b^X, N_{b,min}^X$  and  $N_{b,max}^X$  on one side, and  $N_b^A, N_{b,min}^A$  and  $N_{b,max}^A$  on the other side. This was done for three values of the sex ratio ( $\beta$ ): 1.0, 1.6 and 2.1, which correspond respectively to a balanced sex ratio, the estimated sex ratio over the whole set of populations, and the maximum value for the estimated sex ratio per populations. Note also that because the populations are rather elongated, we also performed all the above analyses under an assumption of a linear model, in which the kinship coefficients are regressed against the linear distance and not its logarithm (Rousset, 2000; Vekemans and Hardy, 2004).

## Results

### Microsatellite diversity

As already reported in other *S. latifolia* natural populations (Jolivet and Bernasconi, 2006, 2007; Teixeira and



**Figure 2** Minimal ( $r_{min}$ ) and maximal ( $r_{max}$ ) values that the ratio  $N_b^X/N_b^A$  can take as a function of the sex ratio ( $\beta = N_f/N_m$ ). Note that the x axis is in log scale.

Bernasconi, 2007), very high level of polymorphism was detected for all microsatellite loci used in this study, as well as significant linkage disequilibrium between SI1, SI14 and SI15. Although the genomic position of the marker SI1 is unknown, this pattern may possibly suggest that it is a pseudoautosomal region, inherited as autosomal, but homologous to sequences in the X chromosome. The number of alleles per locus ranged between 34 (SI8) and 57 (SI6), and the number of alleles per subpopulation and locus ranged from 4 (MilS, SI8) to 24 (LobL2, SI6). Average values across loci per subpopulation of  $F_{IS}$ , allelic richness and gene diversity are given in Table 1 (see Supplementary Table 3 for values of  $F_{IS}$  per locus and subpopulation). Contrary to our expectations, statistical comparisons of genetic diversity indexes between large (L) and small (S) populations were nonsignificant ( $P > 0.05$ , FSTAT). Eastern subpopulations (more frequent and forming larger patches) had significantly higher allelic richness (East, 11.547; West, 8.872;  $P = 0.012$ ), and marginally significant higher gene diversity (East, 0.905; West, 0.857;  $P = 0.09$ ) than western subpopulations (fewer, scattered and forming smaller patches). The subpopulation MilS consistently behaved as an outlier within the eastern regions with a surprisingly small allelic richness (Table 1). Repeating the analysis without this subpopulation confirmed the differences between the two geographic areas in both indices (allelic richness: East, 12.141; West, 8.872;  $P = 0.003$  and gene diversity: East, 0.924; West, 0.857;  $P = 0.016$ ). Saturation curves corroborated these results showing earlier saturation for all loci (but SI8) in all western subpopulations, whereas eastern subpopulations showed later saturation or even no saturation in some loci (Supplementary Figure 1). The outlier subpopulation MilS showed faster saturation than any other subpopulation (excluding locus SI1), confirming its unusually low allelic richness.

### Population genetic structure

We found surprisingly large genetic differentiation among subpopulations of *S. latifolia* along a 60 km transect. All pairwise  $F_{ST}$  estimates were significant with the exception of few comparisons between adjacent subpopulations within regions (see Supplementary Table 2).

**Table 1** List of subpopulations, ordered by region (see Figure 1), with sample sizes ( $n$ ), summary of basic descriptive statistics averaged across all six loci and populations sizes measured in 2003

Region	Population	n	$F_{IS}$	Allelic richness	Gene diversity	Population size
West (1)	Zalt L1	15	0.227	8.3	0.77	130
	Zalt L2	15	0.329	8.4	0.86	
	Zalt S	23	0.440	8.4	0.84	
West (2)	Tiel L1	15	0.368	8.6	0.88	784
	Tiel L2	25	0.382	10.1	0.90	
	Tiel S	20	0.256	9.5	0.87	
East (3)	Nij L1	15	0.390	12.1	0.92	1019
	Nij L2	15	0.267	10.2	0.90	
	Nij S	20	0.371	11.0	0.91	
East (4)	Mil L1	15	0.451	12.7	0.93	13 273
	Mil L2	15	0.434	12.2	0.93	
	Mil S	20	0.418	6.8	0.78	
East (5)	Lob L1	15	0.303	13.0	0.93	2783
	Lob L2	15	0.362	13.9	0.94	
	Lob S	15	0.296	12.2	0.93	

All  $F_{IS}$  values were significantly different from 0,  $P < 0.001$ .

**Table 2** Results of the hierarchical analysis of molecular variance comparing the neutral genetic variation among regions, among subpopulations within regions and within subpopulations

Source of variation	d.f.	Sum of squares	Variation (%)
Among regions	4	46.02	2.75
Among subpopulations within regions	10	86.56	6.25
Within subpopulations	501	1003.10	91.00

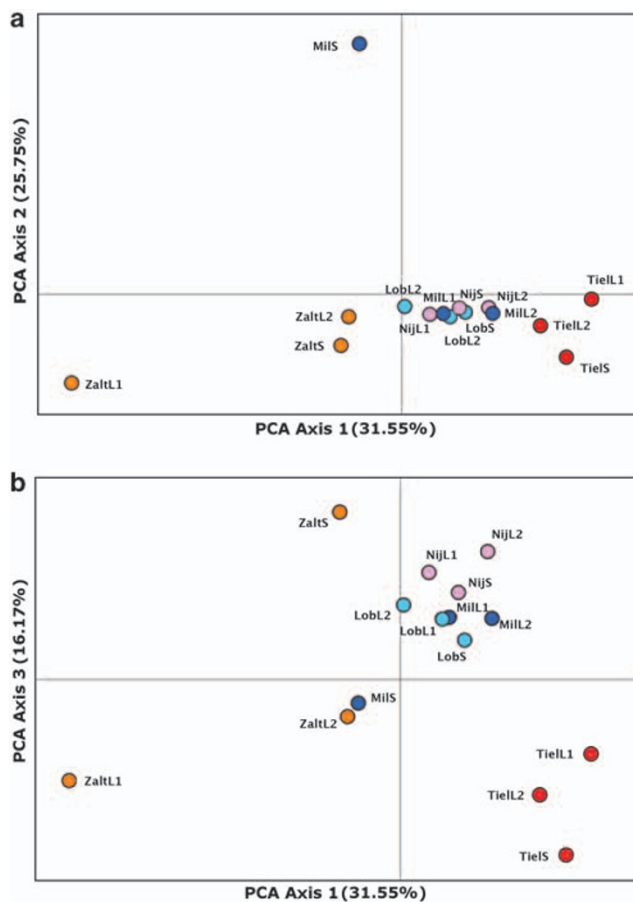
The highest differentiation was found between distant subpopulations. However, surprisingly high differentiation was also found between subpopulations within the Zalt region, as well as between MilS and the two subpopulations in population MilL. These results were consistent for all microsatellites used separately. The hierarchical analysis of molecular variance revealed that 91% of neutral genetic variation resides within subpopulations, 6.25% is distributed within regions and an additional 2.75% of the variance can be explained by differentiation among regions (Table 2). Global  $F_{ST}$  was 0.11.

#### Clustering of populations

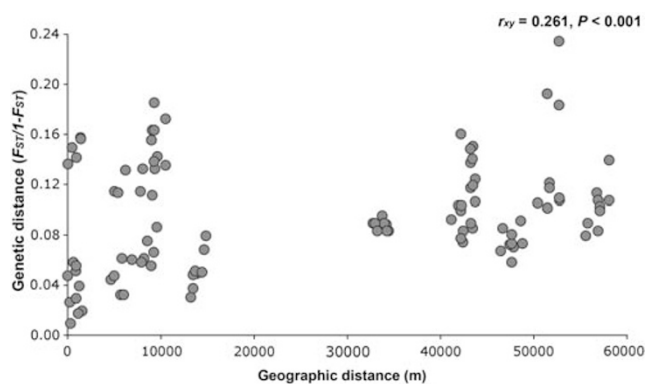
We used a principal coordinate analysis to investigate the relative position of subpopulations and regions in a multidimensional space (Figure 3). The first three principal component (PC) axes explained 31.55, 25.75 and 16.17% of the neutral genetic variation among subpopulations (cumulative inertia: 73.47%). PC1 revealed a cluster of all subpopulations from eastern regions (with the exception of the outlier subpopulation MilS), clearly separated from the western subpopulations, which surprisingly did not cluster close to each other. Subpopulations from the Tiel region came together in all axes, but did not group together with subpopulations from the Zalt region, which appeared scattered on the multidimensional space. PC2 discriminated between the outlier subpopulation MilS and all other subpopulations, corroborating previous analyses. PC3 corroborated the clustering of eastern populations and evidenced the diversity within the western Zalt population. Tests of statistical significance of both total inertia (global  $F_{ST}=0.107$ ) and the first three individual axes inertia (per Axis  $F_{ST}=0.021$ , 0.016 and 0.013) were highly significant ( $P<0.01$ ).

#### Spatial genetic structure

There were significant associations between genetic differences and geographic distances at the subpopulation level (Figure 4), although only a small proportion of the genetic variance was explained by geography ( $r^2=0.07$ ;  $P<0.001$ ). Although there was a general pattern of IBD in *S. latifolia* in the studied area, the region Zalt clearly deviated from this trend (Supplementary Table 2), with the two subpopulations ZaltL1 and ZaltL2, which were only 80 meters away from each other but genetically distinct ( $F_{ST}=0.136$ ). Similarly, the outlier subpopulation MilS was genetically very differentiated ( $F_{ST}=0.140$ ) from the geographically close MilL subpopulations (1000 m). When the region Zalt and the outlier subpopulation MilS were removed, a much



**Figure 3** Plot of the first three axes of the principal coordinate analysis (a, axes 1 and 2; b axes 1 and 3) based on pairwise  $F_{ST}$  in *S. latifolia* subpopulations. Colors represent different regions. The proportions (in percent) of the variance explained by each axis are given in parenthesis.



**Figure 4** Correlation between pairwise genetic differences (linearized  $F_{ST}$  values) and pairwise geographic distances of *S. latifolia* subpopulations.  $r_{xy}$  represents the correlation coefficient and  $P$  the probability value.

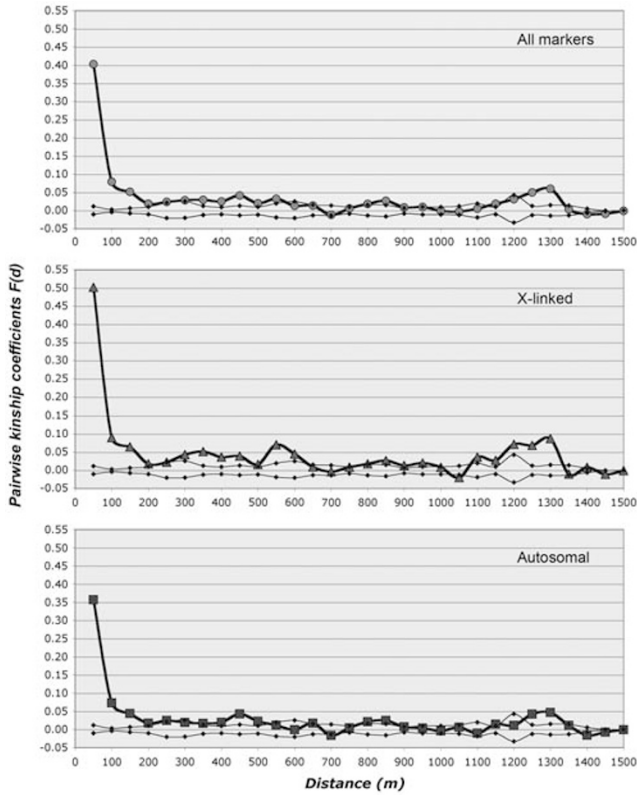
stronger overall pattern of IBD was detected ( $r_{xy}=0.737$ ,  $r^2=0.544$ ,  $P<0.001$ ). When performing a regression of  $F_{ST}/(1-F_{ST})$  against the logarithm of distance, we obtain a neighborhood size  $N_b$  of 437.90 for autosomes and 144.87 for X-linked markers. Assuming an effective population size  $N_e$  of 1000 per population, this translates

into  $\sigma$  values:  $\sigma_A = 0.18$  m for autosomes and  $\sigma_x = 0.107$  m. Assuming an  $N_e$  value of 100 gives  $\sigma_A = 0.590$  m for autosomes and  $\sigma_x = 0.339$  m.

The characterization of spatial genetic structure at the individual level as calculated with *SPAGeDI* revealed a pattern consistent with IBD, where pairwise kinship coefficients decrease with distance. Positive kinship

coefficients were found at the subpopulation (160 m), and population (500 m) levels, only intercepting the  $x$  axis at about 600 m (Figure 5). Kinship coefficients were close to the 95% confidence intervals probably due to the lower power of bootstrapping over loci with a small number of loci, but mostly over this range until 600 m. This result was consistent for X-linked and autosomal markers. Thus, within both subpopulations and populations, neighboring individuals are increasingly related, and this trend mostly disappears over distances of 600 m. The kinship coefficients were highest at the first distance class and rapidly decreased. The autocorrelation analysis calculated with Genalex revealed a very similar pattern (Supplementary Figure 2). Relatedness decreased with distance, and genetic autocorrelation coefficients only intercepted the  $x$  axis (that is, the null hypothesis of no genetic correlation) over 600 m.

The neighborhood sizes estimated from autosomal and X-linked markers were respectively  $N_b^A = 45.4$  (95% CI: 42.18–49.10) and  $N_b^X = 42.27$  (95% CI: 35.40–52.45). Using Eq. (7), this translates into a negative estimate for the ratio of seed to pollen flow ( $\alpha$ ) of  $-0.841$  if a balanced sex ratio ( $\beta = 1$ ) is assumed (Table 3). This value remains negative whatever the sex ratio assumed. A negative true value of  $\alpha$  is obviously mathematically impossible. This discrepancy stems probably from the fact that the estimates of  $N_b^A$  and  $N_b^X$  are point estimates, which deviate to some extent from the true values of this parameter in a way that they do not fulfill the conditions given in Eqs. (8) and (9). However, using Eq. (7) on the minimal or maximal value of the confidence interval of  $N_b^A$  and  $N_b^X$  provided in some cases positive values for  $\alpha$  (Table 3), these values being in general around 1.0, whatever the sex ratio assumed. This indicates a clear tendency of similar level of seed and pollen flow in this area. Repeating this analysis assuming that the population is linear gives very similar results as the analysis assuming a two-dimensional population (Table 3; see also Rousset, 1997). Finally, for the dispersal parameters, we obtained relatively low values in all cases: for an assumed density of  $D_e = 0.007$ ,  $\sigma_A = 22.70$  m for the autosomal markers and  $\sigma_x = 25.31$  m for the X-linked markers; and for an assumed density of  $D_e = 0.25$ ,  $\sigma_A = 3.80$  m and  $\sigma_x = 4.23$  m. The similar values obtained



**Figure 5** Spatial autocorrelograms plotting pairwise kinship coefficients  $F(d)$  against distance for all microsatellite loci together, and autosomal and X-linked markers separately. Pairwise kinship coefficients were calculated with the software *SPAGeDI*, and were averaged over geographic distances every 50 meters. Ninety-five percent confidence intervals calculated with 1000 permutations of loci among all individuals are shown.

**Table 3** Estimated ratio of pollen over seed dispersal ( $\alpha$ ) for all combinations of the point estimates of the neighborhood sizes for X-linked ( $N_b^X$ ) and autosomal ( $N_b^A$ ) markers, and the limits of their confidence intervals ( $N_{b,min}^X$ ,  $N_{b,max}^X$ ) and ( $N_{b,min}^A$ ,  $N_{b,max}^A$ ), for three possible values of the sex ratio ( $\beta$ ), assuming a quadratic (a) or linear (b) model

(a)	$\beta = 1.0$			$\beta = 1.6$			$\beta = 2.21$		
	$N_{b,min}^A$	$N_b^A$	$N_{b,max}^X$	$N_{b,min}^A$	$N_b^A$	$N_{b,max}^X$	$N_{b,min}^A$	$N_b^A$	$N_{b,max}^X$
$N_{b,min}^X$	-0.526	-0.214	0.264	-0.179	0.273	1.023	0.139	0.748	1.839
$N_b^X$	-1.004	-0.841	-0.614	-0.823	-0.610	-0.303	-0.671	-0.410	-0.021
$N_{b,max}^X$	-1.328	-1.238	-1.120	-1.228	-1.118	-0.971	-1.149	-1.021	-0.848

(b)	$\beta = 1.0$			$\beta = 1.6$			$\beta = 2.21$		
	$N_{b,min}^A$	$N_b^A$	$N_{b,max}^X$	$N_{b,min}^A$	$N_b^A$	$N_{b,max}^X$	$N_{b,min}^A$	$N_b^A$	$N_{b,max}^X$
$N_{b,min}^X$	-0.191	0.126	0.980	0.307	0.800	2.294	0.795	1.502	3.959
$N_b^X$	-1.111	-1.011	-0.791	-0.960	-0.833	-0.544	-0.834	-0.682	-0.327
$N_{b,max}^X$	-1.577	-1.542	-1.471	-1.525	-1.484	-1.401	-1.485	-1.439	-1.346

$\beta = 1.6$  was the overall sex ratio observed in our metapopulation and 2.21 the maximum observed among populations.

for the X-linked and autosomal markers point again toward similar levels of seed and pollen flow.

## Discussion

### Metapopulation genetic structure and isolation by distance

The white campion *S. latifolia* in the River Waal metapopulation shows high levels of genetic diversity and subpopulation structure (global  $F_{ST}=0.11$ ) even between relatively nearby subpopulations. However, despite pronounced subpopulation differentiation, we found a general pattern of IBD at the scale of tens of kilometers, providing either evidence of ongoing gene flow or alternatively a recent colonization process (Slatkin, 1993). Substantial among-population structure has been reported across the genus *Silene* (Moyle, 2006), however these studies used molecular markers other than microsatellites, which limits the possibilities for comparison with our results. In general, analyses based on chloroplast DNA result in higher estimates of population differentiation than those published based on allozymes and random nuclear markers (McCauley, 1994, 1998; Olson and McCauley, 2002; Petit *et al.*, 2005). Interestingly, it has been proposed that the chloroplast of dioecious *Silene* has undergone a selective sweep, which would be reflected in reduced genetic diversity on its genes (Muir and Filatov, 2007). Similarly, Y-linked markers appear to have low genetic diversity due to Y-chromosome degeneration (but see Ironside and Filatov, 2005; Laporte *et al.*, 2005). Comparatively, X-linked and autosomal markers could harbor more genetic diversity (Laporte *et al.*, 2005). We found similar levels of genetic diversity in both X-linked and autosomal markers. Owing to their high mutation rates microsatellites typically show high polymorphism and low population differentiation (Balloux *et al.*, 2000). Therefore, the high level of differentiation observed in this study is a very strong sign of genetic structure. Genetic IBD also appears to be common in *Silene* species. However, population genetic studies of *Silene*, and particularly with *S. latifolia*, have been mainly performed in North America, a recently colonized area (McCauley, 1994, 1998; Richards *et al.*, 1999). North American populations of *S. latifolia* appear to be the result of multiple invasions of diverse origins (Taylor and Keller, 2007) and, therefore, population differentiation might be the consequence of contemporary demographic processes as well as ancestral genetic differentiation of source populations.

The clustering analysis (principal coordinate analysis) revealed that the studied metapopulation along 60 km of the River Waal is not a panmictic unit, but is instead composed by an eastern metapopulation (with the exception of the outlier subpopulation MilS) and distinct populations in the western area (see Figure 3). This suggests higher effective population sizes and gene flow in the more continuous eastern metapopulation (see Supplementary Figure 1) also supported by generally lower genetic differentiation among subpopulations in this area (Supplementary Table 2), and higher effects of restricted gene flow and genetic drift in the other four genetically distinct populations. This could indicate an older age and a higher stability of the eastern population, whereas the western populations could be the

consequence of more recent invasions from diverse origins (see, for a similar case in *S. dioica*, Giles and Goudet, 1997) or remnants from larger populations. Testing this idea would require to investigate the more western populations not included in this study (gray circles in Figure 1).

### Effects of population demography: size and distance of populations

Theory predicts that small isolated populations should show less genetic variability and more inbreeding than large and well-connected ones, and therefore, undergo higher risk of extinction (Holsinger, 2000; Richards, 2000). Gene flow operates against genetic drift and inbreeding by homogenizing populations. Distance between populations will constrain the possibility of gene exchange between populations. However, the scale at which populations can be considered isolated is not clear. In the case of *S. latifolia*, with insect-mediated pollination, flight distances and pattern of pollen uptake and deposition will determine the potential for gene flow through pollen. *S. latifolia* has white flowers, nocturnal anthesis and produces a scent that attracts nocturnal moths (Dötterl *et al.*, 2006). In agreement with this, experimental approaches have shown that nocturnal pollinators are responsible for most gene flow between populations, diurnal pollinators having only a small role for gene flow (Barthelmess *et al.*, 2006). In its native range the main pollinator of *S. latifolia* is the moth *Hadena bicuris*. Moths are often able to fly over long distances (Linhart and Mendenhall, 1977; Elzinga *et al.*, 2007), which would promote gene exchange between isolated subpopulations (Altizer *et al.*, 1998; Barthelmess *et al.*, 2006). Based on this, potential pollen flow in this species could be on the range of hundreds of meters. However, the pollen load that moths of this species receive from visiting one male flower is essentially deposited on the next three female flowers visited (A-M Labouche and G Bernasconi, unpublished), suggesting that most pollen may in fact be dispersed within patches of neighboring plants (see also Wright and Meagher, 2004).

In addition, the size and isolation of populations appear to affect the potential for gene flow, as some pollinators visit most often, and proportionally more flowers in larger patches. Experimental studies of *S. latifolia* in North America suggest that plant patches (consisting of 10 females) separated over distances of 80 meters can be considered as isolated (Richards *et al.*, 1999). We sampled small plant patches in all regions separated by over 1 km (see population distances in Supplementary Table 2) and compared their genetic diversity with larger and nearby plant patches. All these populations persisted over 6 years of monitoring (2001–2006; Elzinga *et al.*, 2005), which is longer than the reproductive life of single individuals (2–3 years). Contrary to our expectations, we found no evidence of reduced genetic diversity in the small populations compared with the nearby larger ones. This pattern could be caused by extensive pollen flow mediated by pollinators or by punctual large distance events of seed dispersal. Alternatively, small populations could be recent remnants from larger populations.

One population, MilS, deviated in all analyses from all other surrounding sites, showing consistently less



genetic diversity and different genetic composition as revealed by the clustering analysis. The size and distance of this subpopulation was within the range of the other sampled populations. One explanation for the particular genetic structure of this site could be that it is located on the north shore of the river, and the water body could constitute a barrier for seed or pollinator movement. Alternatively, this population may be related to other *S. latifolia* populations that were not sampled (gray circles Figure 1). Knowledge on the genetics of these other populations would give more information on the potential of rivers to prevent dispersal on this plant species. Because *S. latifolia* hybridizes naturally with *S. dioica* (Baker, 1947; Goulson and Jerrim, 1997; Minder *et al.*, 2007), hybridization could be an alternative explanation to the existence of a genetically very divergent population.

#### Spatial scale of gene dispersal

Spatial autocorrelation analyses (for example, García *et al.*, 2007; Hardy *et al.*, 2006; Jones *et al.*, 2007; Peakall and Smouse, 2006; Temple *et al.*, 2006; Valbuena-Carabaña *et al.*, 2007) are a widespread tool for estimating local genetic structure and patterns of gene flow. From this approach indirect estimates of gene dispersal can be obtained (see for example, Austerlitz *et al.*, 2004; Hardy *et al.*, 2006). Using this methodology we found positive kinship coefficients between individuals even at the level of hundreds of meters (Figure 5), only disrupted over distances of about 600 m, after which individuals showed no evidence of genetic autocorrelation. This trend suggests potential extensive gene flow in this species. There is a debate on whether direct conclusions about historical gene dispersal can be directly drawn from spatial autocorrelation patterns, as in those, equilibrium isolation-by-distance patterns are assumed (see, for example, Hardy *et al.*, 2006). Instead, an estimation of gene dispersal  $\sigma_g^2$  is recommended. In fact, our estimates of gene dispersal revealed flow of genes only at the level of a few tens of meters (5–25 m). These results are in accordance to indirect estimates of decay of pollen dispersal with increasing distance at the scale of 50 or even 20 meters in Spanish populations of *S. latifolia* based on maximum likelihood paternity analyses (Smouse *et al.*, 1999) on allozymes. Young (2002) measured direct pollen dispersal through diurnal and nocturnal pollinators using fluorescent dyes in North American populations (patch distances 18–95 m), and detected fluorescent pollen on a scale of centimeters to a few meters. Using experimental populations consisting of genotypes homozygous at a diagnostic locus, and interpopulations distances of 20–80 m, Richards *et al.* (1999) and Barthelmess *et al.* (2006) experimentally proved movement of pollen on a scale of tens of meters also in North American populations. With an indirect method, in this study we also found evidence of pollen dispersal at the scale of tens of meters for natural plant stands from the native range of *S. latifolia*. The study at among population level indicated even lower estimated values of gene dispersal (up to ~0.6 m), confirming the restricted level of dispersal for this species in this area. The fact that the main pollinator in the native range of this species, the moth *H. bicurris*, is absent in North

America suggests that the two distribution areas might differ in their biotic interactions, and therefore in the potential for gene flow. Pollinator role in North America is taken over by other insects, mostly by other moths (Young, 2002). Despite this, our study results do not indicate a general congruence between the ranges of gene flow in both areas. However, different approaches have been taken, and additional local differences may subsist in the presence or absence of adjoining sites to the study populations, so that definitive support for greater pollen flow distances in the presence than in the absence of the specialist moth *H. bicurris* would only be gained by a parallel study including samples from both distribution areas.

Genetic studies on related *Silene* species (*S. dioica*, Giles *et al.*, 1998; *S. acaulis*, Gehring and Delph, 1999; *S. tatarica*, Tero *et al.*, 2005) also suggest gene dispersal in the range of tens of meters. However, some of these species have diurnal 'pollination syndrome' with pink flowers and are partly self-compatible (Aspi *et al.*, 2003; Marr, 2006), and/or are mostly pollinated by bees (Kephart *et al.*, 2006) or bumblebees (Shykoff, 1988; Marr, 2006). Diurnal insect pollinators tend to forage primarily within rather than between flower patches (Altizer *et al.*, 1998). Therefore, gene flow distance in these species would be expected to be more restricted than in *S. latifolia*. As pointed out above, better conclusions could be drawn from an integrative study on different species where similar methodology would be applied.

Spatial autocorrelation analyses on other plant species revealed very diverse patterns of gene flow depending on mating system, life form and population density (Vekemans and Hardy, 2004). The smallest gene flow values of just a few meters were found in herbaceous annual and perennial herbs (Godt and Hamrick, 1993; Miyazaki and Isagi, 2000; Silvertown and Charlesworth, 2001; Fenster *et al.*, 2003). Insect-pollinated plants with gravity seed dispersal appear to have the shortest dispersal ranges, followed by wind-pollinated and dispersed species. The largest dispersal ranges appear to be achieved by birds as pollinators and seed dispersers. The gene dispersal role of nocturnal insects would *a priori* be expected to be intermediate between diurnal insects and birds, although to the best of our knowledge there is to date no evidence in favor of this, including this study.

Probably one of the most surprising results of this study concerns the estimate of the relative dispersal rate of the two critical stages of gene flow, seeds and pollen. Although in invasive populations in North America the comparison of cpDNA and allozyme variation indicated that *S. latifolia* has higher pollen than seed movement (McCauley, 1994, 1997), our study found similar levels of pollen and seed dispersal (Table 3) whatever the assumed sex ratio (up to 2.21, the maximal sex ratio observed in our populations). Several explanations may account for this. Stochastic events can produce extensive seed dispersal in a weed like *S. latifolia* and could potentially explain the pattern found in our analysis, whereas pollen dispersal appears limited (Young, 2002). Finally, our estimate of the ratio of pollen to seed dispersal may improve with more markers. With the increasing availability of X- and Y-linked genes for this species (McCauley, 1997; Ironside and Filatov, 2005; Bergero *et al.*, 2007), it should be possible for future

studies to perform a similar analysis with a higher resolution.

## Conflict of interest

The authors declare no conflict of interest.

## Acknowledgements

We thank Arjen Biere, Gabriela Gleiser, Gerald Heckel, John Pannell and the reviewers for valuable comments on earlier stages of this study. Arjen Biere and the NIOO Wageningen provided logistical support in the collection of samples. This study was supported by Swiss National Science Foundation (Grant nos. 3100A0-122004/1 to GB; PIOIA-119443 to FA and GB) and the Fondation Pierre Mercier pour la Science (to GB).

## References

- Altizer SM, Thrall PH, Antonovics J (1998). The role of pollinators as vectors for anther-smut infection in *Silene alba*. *Am Midl Nat* **139**: 147–163.
- Aspi J, Jäkäläniemi A, Tuomi J, Siikamäki P (2003). Multilevel phenotypic selection on morphological characters in a metapopulation of *Silene tatarica*. *Evolution* **57**: 509–517.
- Austerlitz F, Dick CW, Dutech C, Klein EK, Oudou-Muratorio S, Smouse P et al. (2004). Using genetic markers to estimate the pollen dispersal curve. *Mol Ecol* **13**: 937–954.
- Baker HG (1947). *Melandrium album* (Mill.) Garcke. *J Ecol* **35**: 274–282.
- Balloux F, Brunner H, Lugon-Moulin N, Hausser J, Goudet J (2000). Microsatellites can be misleading: an empirical and simulation study. *Evolution* **54**: 1414–1422.
- Bari EA (1969). *Experimental Taxonomy of Annual European Species of Silene L.* PhD thesis, University of Cambridge.
- Barthelmeß EL, Richards CM, McCauley DE (2006). Relative effects of nocturnal vs diurnal pollinators and distance on gene flow in small *Silene alba* populations. *New Phytol* **169**: 689–698.
- Bergero R, Forrest A, Kamau E, Charlesworth D (2007). Evolutionary strata on the X chromosomes of the dioecious plant *Silene latifolia*: evidence from new sex-linked genes. *Genetics* **175**: 1945–1954.
- Blair AC, Wolfe LM (2004). The evolution of an invasive plant: an experimental study with *Silene latifolia*. *Am Nat* **85**: 3035–3042.
- Bopp S, Gottsberger G (2004). Importance of *Silene latifolia* ssp. *alba* and *S. dioica* (Caryophyllaceae) as a host plants of the parasitic pollinator *Hadena bicurvis* (Lepidoptera, Noctuidae). *Oikos* **105**: 221–228.
- Cain ML, Milligan BG, Strand AE (2000). Long-distance seed dispersal in plant populations. *Am J Bot* **87**: 1217–1227.
- Chung MG, Chung MY, Oh GS, Epperson BK (2000). Spatial genetic structure in a *Neolitsea sericea* population (Lauraceae). *Heredity* **85**: 490–497.
- Crawford TJ (1984). The estimation of neighborhood parameters for plant populations. *Heredity* **52**: 273–283.
- Dötterl S, Jürgens A, Seifert K, Laube T, Weissbecker B, Schütz S (2006). Nursery pollination by a moth in *Silene latifolia*: the role of odours in eliciting antennal and behavioral responses. *New Phytol* **169**: 707–718.
- Double MC, Peakall R, Beck NR, Cockburn A (2005). Dispersal, phylopatry, and infidelity: dissecting local genetic structure in superb fairy-wrens (*Malurus cyaneus*). *Evolution* **59**: 625–635.
- Elzinga JA (2005). Effects of habitat fragmentation on a tri-trophic system: *Silene latifolia*, *Hadena bicurvis* and its parasitoids. PhD thesis, University of Utrecht.
- Elzinga JA, Turin H, van Damme JMM, Biere A (2005). Plant population size and isolation affect herbivory of *Silene latifolia* by the specialist herbivore *Hadena bicurvis* and parasitism of the herbivore by parasitoids. *Oecologia* **144**: 416–426.
- Elzinga JA, van Nouhuys S, van Leeuwen D-J, Biere A (2007). Distribution and colonisation ability of three parasitoids and their herbivorous host in a fragmented landscape. *Basic Appl Ecol* **8**: 75–88.
- Epperson BK, Allard RW (1989). Spatial autocorrelation analysis of the distribution of genotypes within populations of lodge-pole pine. *Genetics* **121**: 369–377.
- Excoffier L, Laval G, Schneider S (2005). Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evol Bioinform Online* **1**: 47–50.
- Excoffier L, Smouse PE, Quattro J (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* **131**: 479–491.
- Fenster CB, Vekemans X, Hardy OJ (2003). Quantifying gene flow from spatial genetic structure data in a metapopulation of *Chamaecrista fasciculata* (Leguminosae). *Evolution* **57**: 995–1007.
- Fritz AL, Nilsson LA (1994). How pollinator-mediated mating varies with population size in plants. *Oecologia* **100**: 451–462.
- García C, Jordano P, Godoy JA (2007). Contemporary pollen and seed dispersal in a *Prunus mahaleb* population: patterns in distance and direction. *Mol Ecol* **16**: 1947–1955.
- Gehring JL, Delph LF (1999). Fine-scale genetic structure and clinal variation in *Silene acaulis* despite of high gene flow. *Heredity* **82**: 628–637.
- Giles BE, Goudet J (1997). Genetic differentiation in *Silene dioica* metapopulations: estimation of spatiotemporal effects in a successional plant species. *Am Nat* **149**: 507–526.
- Giles BE, Lundqvist E, Goudet J (1998). Restricted gene flow and subpopulation differentiation in *Silene dioica*. *Heredity* **80**: 715–723.
- Godt MJW, Hamrick JL (1993). Patterns and levels of pollen-mediated gene flow in *Lathyrus latifolius*. *Evolution* **47**: 98–110.
- Goudet J (2002). FSTAT: a program to estimate and test gene diversities and fixation indices. v. 2.9.3.2. (Online) Available at <http://www.unil.ch/izea/software/fstat.html>.
- Goudet J (2005). PCAGEN 1.2. Available at <http://www2.unil.ch/popgen/software/pcagen.htm>.
- Goulson D, Jerrim K (1997). Maintenance of the species boundary between *Silene dioica* and *S. latifolia*. *Oikos* **79**: 115–126.
- Hamrick JL, Godt MJW (1996). Effects of life history traits on genetic diversity in plant species. *Philos Trans R Soc London B* **351**: 1291–1298.
- Hardy OJ, Maggia L, Bandou E, Breyne P, Caron H, Chevallier M-H et al. (2006). Fine-scale genetic structure and gene dispersal inferences in 10 neotropical tree species. *Mol Ecol* **15**: 559–571.
- Hardy OJ, Vekemans X (2002). SPAGEDI: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Mol Ecol Notes* **2**: 618–620.
- Heuertz M, Vekemans X, Hausman J-F, Paladas M, Hardy OJ (2003). Estimating seed vs. pollen dispersal from spatial genetic structure in the common ash. *Mol Ecol* **12**: 2483–2495.
- Holsinger KE (2000). Demography and extinction in small populations. In: Young AG CG (ed). *Genetics, Demography and Viability of Fragmented Populations*. Cambridge University Press: Cambridge. pp 54–74.
- Howe HF, Smallwood J (1982). Ecology of seed dispersal. *Annu Rev Ecol Syst* **13**: 201–228.
- Imbert E, Lefevre F (2003). Dispersal and gene flow of *Populus nigra* (Salicaceae) along a dynamic river. *J Ecol* **91**: 447–456.

- Ingvarsson PK (2002). A metapopulation perspective on genetic diversity and differentiation in partially self-fertilizing plants. *Evolution* **56**: 2368–2373.
- Ironside JE, Filatov DA (2005). Extreme population structure and high interspecific divergence of the *Silene* Y chromosome. *Genetics* **171**: 705–713.
- Jolivet C, Bernasconi G (2006). Experimental analysis of constitutive and induced defense in a plant–seed predator system. *Funct Ecol* **20**: 966–972.
- Jolivet C, Bernasconi G (2007). Molecular and quantitative genetic differentiation in European populations of *Silene latifolia* (Caryophyllaceae). *Ann Bot* **100**: 119–127.
- Jones TH, Vaillancourt RE, Potts BM (2007). Detection and visualization of spatial genetic structure in continuous *Eucalyptus globulus* forest. *Mol Ecol* **16**: 697–707.
- Karron J, Tucker R, Thumser N, Reinartz J (1995). Comparison of pollinator flight movements and gene dispersal patterns in *Mimulus ringens*. *Heredity* **75**: 612–617.
- Kephart S, Reynolds RJ, Rutter MT, Fenster CB, Dudash MR (2006). Pollination and seed predation by moths on *Silene* and allied Caryophyllaceae: evaluating a model system to study the evolution of mutualisms. *New Phytol* **169**: 667–680.
- Laporte V, Filatov DA, Kamau E, Charlesworth B (2005). Indirect evidence from DNA sequence diversity for genetic degeneration of the Y-chromosome in dioecious species of the plant *Silene*: the SIY4/SIX4 and DD44-X/DD44-Y gene pairs. *J Evol Biol* **18**: 337–347.
- Latta RG, Linhart YB, Fleck D, Elliot M (1998). Direct and indirect estimates of seed versus pollen movement within a population of ponderosa pine. *Evolution* **52**: 61–67.
- Linhart YB, Mendenhall JA (1977). Pollen dispersal by hawkmoths in a *Lindenia rivalis* Benth. population in Belize. *Biotropica* **9**: 143.
- Loiselle BA, Sork VL, Nason J, Graham C (1995). Spatial genetic structure of a tropical understory shrub, *Psychotria officinalis* (Rubiaceae). *Am J Bot* **82**: 1420–1425.
- Macdonald DW, Johnson DDP (2001). Dispersal in theory and practice: consequences for conservation biology. In: Clobert J, Danchin E, Dhondt AA and Nichols JD (eds). *Dispersal*. Oxford University Press: New York, pp 358–372.
- Marr D (2006). Seed fitness of hermaphrodites in areas with females and anther-smut disease: *Silene acaulis* and *Microbotryum violaceum*. *New Phytol* **169**: 741–751.
- McCauley DE (1994). Contrasting the distribution of chloroplast DNA and allozyme polymorphism among local populations of *Silene alba*: implications for studies of gene flow in plants. *Proc Natl Acad Sci USA* **91**: 8127–8131.
- McCauley DE (1997). The relative contribution of seed and pollen movement to local genetic structure of *Silene alba*. *J Heredity* **88**: 257–263.
- McCauley DE (1998). The genetic structure of a gynodioecious plant: nuclear and cytoplasmic genes. *Evolution* **52**: 255–260.
- Minder AM, Rothenbuehler C, Widmer A (2007). Genetic structure of hybrid zones between *Silene latifolia* and *Silene dioica* (Caryophyllaceae): evidence for introgressive hybridization. *Mol Ecol* **16**: 2504–2516.
- Miyazaki Y, Isagi Y (2000). Pollen flow and the intrapopulation genetic structure of *Helionopsis orientalis* on forest floor as determined using microsatellite markers. *Theor Appl Genet* **101**: 718–723.
- Moyle LC (2006). Correlates of genetic differentiation and isolation by distance in 17 congeneric *Silene* species. *Mol Ecol* **15**: 1067–1081.
- Muir G, Filatov DA (2007). A selective sweep in the chloroplast DNA of dioecious silene (Section *Elisanthe*). *Genetics* **177**: 1239–1247.
- Mustajärvi K, Siikamäki P, Rytönen S, Lammi A (2001). Consequences of plant population size and density for plant–pollinator interactions and plant performance. *J Ecol* **89**: 80–87.
- Olson MS, McCauley DE (2002). Mitochondrial diversity, population structure, and gender association in the gynodioecious plant *Silene vulgaris*. *Evolution* **56**: 253–262.
- Ouborg NJ, Piquot Y, van Groenendael JM (1999). Population genetics, molecular markers and the study of dispersal in plants. *J Ecol* **87**: 551–568.
- Peakall R, Ruibal M, Lindenmayer DB (2003). Spatial autocorrelation analysis offers new insights into gene flow in the Australian bush rat, *Rattus fuscipes*. *Evolution* **57**: 1182–1195.
- Peakall R, Smouse PE (2006). GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Mol Ecol Notes* **6**: 288–295.
- Peakall R, Smouse PE, Huff DR (1995). Evolutionary implications of allozyme and RAPD variation in diploid populations of Buffalograss *Buchloe dactyloides*. *Mol Ecol* **4**: 135–147.
- Petit RJ, Duminil J, Fineschi S, Hampe A, Salvini D, Vendramin GG (2005). Comparative organization of chloroplast, mitochondrial and nuclear diversity in plant populations. *Mol Ecol* **14**: 689–701.
- Queller DC, Goodnight KF (1989). Estimating relatedness using genetic markers. *Evolution* **43**: 258–275.
- Ravigné V, Olivieri I, González-Martínez SC, Rousset F (2006). Selective interactions between short-distance pollen and seed dispersal in self-compatible species. *Evolution* **60**: 2257–2271.
- Rice WR (1989). Analyzing tables of statistical tests. *Evolution* **43**: 223–225.
- Richards CM (2000). Inbreeding depression and genetic rescue in a plant metapopulation. *Am Nat* **155**: 383–394.
- Richards CM, Church S, McCauley DE (1999). The influence of population size and isolation on gene flow by pollen in *Silene alba*. *Evolution* **53**: 63–73.
- Rousset F (1997). Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics* **145**: 1219–1228.
- Rousset F (2000). Genetic differentiation between individuals. *J Evol Biol* **13**: 58–62.
- Segurel L, Martínez-Cruz B, Quintana-Murci L, Balaesque P, Georges M, Hegay T et al. (2008). Sex-specific genetic structure and social organization in Central Asia: insights from a multi-locus study. *PLoS Genet* **4**: e1000200.
- Shykoff JA (1988). Maintenance of gynodioecy in *Silene acaulis* (Caryophyllaceae): stage-specific fecundity and viability selection. *Am J Bot* **75**: 844–850.
- Silvertown J, Charlesworth D (2001). *Introduction to Plant Population Biology*. Blackwell: USA.
- Slatkin M (1993). Isolation by distance in equilibrium and non-equilibrium populations. *Evolution* **47**: 264–279.
- Smouse PE, Long JC, Sokal RR (1986). Multiple regression and correlation extensions of the Mantel test of matrix correspondence. *Syst Biol* **35**: 627–632.
- Smouse PE, Meagher TR, Kobak CJ (1999). Parentage analysis in *Chamaelirium luteum* (L.) Gray (Liliaceae): why do some males have higher reproductive contributions? *J Evol Biol* **12**: 1069–1077.
- Smouse PE, Peakall R (1999). Spatial autocorrelation analysis of individual multiallele and multilocus genetic structure. *Heredity* **82**: 561–573.
- Streiff R, Labbe T, Bacilieri R, Steinkellner H, Glossl J, Kremer A (1998). Within-population genetic structure in *Quercus robur* L. & *Quercus petraea* (Matt.) Liebl. assessed with isozymes and microsatellites. *Mol Ecol* **7**: 317–328.
- Tackenberg O (2003). Modelling long-distance dispersal of plant diaspores by wind. *Ecol Monogr* **73**: 173–189.
- Tackenberg O, Poschlod P, Bonn S (2003). Assessment of wind dispersal potential in plant species. *Ecol Monogr* **73**: 191–205.
- Taylor DR, Keller SR (2007). Historical range expansion determines the phylogenetic diversity introduced during contemporary species invasion. *Evolution* **61**: 334–345.
- Teixeira S, Bernasconi G (2007). High prevalence of multiple paternity within fruits in natural populations of *Silene*

- latifolia*, as revealed by microsatellite DNA analysis. *Mol Ecol* **16**: 4370–4379.
- Temple HJ, Hoffman I, Amos W (2006). Dispersal, philopatry and intergroup relatedness: fine-scale genetic structure in the white-breasted thrasher, *Ramphocinclus brachyurus*. *Mol Ecol* **15**: 3449–3458.
- Tero N, Aspi J, Siikamäki P, Jäkäläniemi A (2005). Local genetic population structure in an endangered plant species, *Silene tatarica* (Caryophyllaceae). *Heredity* **94**: 478–487.
- Turner ME, Stephens JC, Anderson WW (1982). Homozygosity and patch structure in plant populations as a result of nearest-neighbor pollination. *Proc Natl Acad Sci USA* **79**: 203–207.
- Valbuena-Carabaña M, González-Martínez SC, Hardy OJ, Gil L (2007). Fine-scale spatial genetic structure in mixed oak stands with different levels of hybridization. *Mol Ecol* **16**: 1207–1219.
- Vekemans X, Hardy OJ (2004). New insights from fine-scale spatial genetic structure analyses in plant populations. *Mol Ecol* **13**: 921–935.
- Weir BS, Cockerham CC (1984). Estimating F-statistics for the analysis of population structure. *Evolution* **38**: 1358–1370.
- Wright JW, Meagher TR (2004). Selection on floral characters in natural Spanish populations of *Silene latifolia*. *J Evol Biol* **17**: 382–395.
- Wright S (1939). *Statistical genetics in relation to evolution*, Vol. 802. Hermann et Cie.: Paris.
- Wright S (1943). Isolation by distance. *Genetics* **28**: 114–138.
- Wright S (1978). *Evolution and the Genetics of Populations*, Vol. 4. University of Chicago Press: Chicago.
- Young HJ (2002). Diurnal and nocturnal pollination of *Silene alba* (Caryophyllaceae). *Am J Bot* **89**: 433–440.

Supplementary Information accompanies the paper on Heredity website (<http://www.nature.com/hdy>)