# High paternal diversity in the self-incompatible herb *Arabidopsis halleri* despite clonal reproduction and spatially restricted pollen dispersal

V. LLAURENS,* V. CASTRIC,* F. AUSTERLITZ† and X. VEKEMANS*

*Université des Sciences et Technologies de Lille 1, Laboratoire de Génétique et Evolution des Populations Végétales, CNRS UMR 8016, Bat. SN2, F-59655 Villeneuve d'Ascq Cedex, France, †CNRS, Laboratoire Ecologie Systématique et Evolution, UMR 8079, Orsay, F-91405; Université Paris-Sud, Orsay, F-91405; AgroParisTech, Paris, F-75231, France*

## Abstract

**The number of sires fertilizing a given dam is a key parameter of the mating system in species with spatially restricted offspring dispersal, since genetic relatedness among maternal sibs determines the intensity of sib competition. In flowering plants, the extent of multiple paternity is determined by factors such as floral biology, properties of the pollen vector, selfing rate, spatial organization of the population, and genetic compatibility between neighbours. To assess the extent of multiple paternity and identify ecological factors involved, we performed a detailed study of mating patterns in a small population of a self-incompatible clonal herb, *Arabidopsis halleri*. We mapped and genotyped 364 individuals and 256 of their offspring at 12 microsatellite loci and jointly analysed the level of multiple paternity, pollen and seed dispersal, and spatial genetic structure. We found very low levels of correlated paternity among sibs ($P_{\text{full-sib}}$ = 3.8%) indicating high multiple paternity. Our estimate of the outcrossing rate was 98.7%, suggesting functional self-incompatibility. The pollen dispersal distribution was significantly restricted (mean effective pollen dispersal distance: 4.42 m) but long-distance successful pollination occurred and immigrating pollen was at most 10% of all pollination events. Patterns of genetic structure indicated little extent of clonal reproduction, and a low but significant spatial genetic structure typical for a self-incompatible species. Overall, in spite of restricted pollen dispersal, the multiple paternity in this self-incompatible species was very high, a result that we interpret as a consequence of high plant density and high pollinator service in this population.**

*Keywords*: *Arabidopsis halleri*, mating patterns, multiple paternity, paternity analysis, pollen dispersal, self-incompatibility

*Received 11 August 2007; revision accepted 16 December 2007*

## Introduction

The extent of multiple paternity, i.e. the number of sires fertilizing a given dam, is a key parameter of the mating system in species with spatially restricted offspring dispersal (Ellstrand 1984). In such species, genetic relatedness among maternal sibs determines the intensity of sib competition (Yasui 1998). Hence, in plants, the extent of multiple paternity is determined by several factors, including spatial and temporal patterns of flower display, the characteristics of

the pollination vector and the ability of pistils to filter out individual pollen grains, including self-pollen, according to their origin (Hardy *et al*. 2004; Smouse & Sork 2004).

Correlated paternity among outcrossed maternal sibs can arise when fertilization events are not independent from each other. This may arise for instance when pollen carried by an animal pollinator leads to fertilization of all seeds of a fruit, as shown for example in species where pollen grains are carried together in pollinia (Broyles & Wyatt 1991). However, a pollinator may also simultaneously deposit a mixed pollen load in a single visit (Marshall & Ellstrand 1985; Campbell 1998). As pointed out by Hardy *et al*. (2004), correlated paternity also arises when the same

Correspondence: Xavier Vekemans, Fax: + 33 320 43 6979; E-mail: xavier.vekemans@univ-lille1.fr

pollen sources are used in independent pollination events, because the number of pollen donors for a given maternal plant is limited.

Despite the availability of several methodological approaches to estimate the extent of multiple paternity, data are still limited. A few trends in relation to life-history characteristics have, however, been deduced. Correlated paternity is generally higher in herbaceous than in tree species; and among tree species, correlated paternity is overall lower in wind-pollinated than in animal-pollinated species, and decreases with increasing conspecific density (Hardy *et al.* 2004; Smouse & Sork 2004). However, little is known on the factors influencing multiple paternity in flowering plants, particularly for herbaceous species.

In particular, the impact of self-incompatibility (SI) in this context has received only limited attention. Self-incompatibility systems are genetic systems that prevent selfing of hermaphrodite plants as well as cross-fertilization between individuals sharing identical SI phenotypes. The effect of SI is twofold: on the one hand, it prevents self-fertilization and thus enforces outcrossed pollen sources. On the other hand, SI filters out pollen sources that match the SI phenotype of the focal plant, and thus may also decrease the extent of multiple paternity. Hence, the net effect of SI on patterns of multiple paternity is difficult to predict. Several studies have provided evidence that in self-incompatible species, limited availability of compatible pollen sources may result in overall pollen limitation (e.g. Wagenius *et al.* 2007), suggesting that the sieving effect of SI may severely restrict multiple paternity. Theoretical investigations have shown that this process will depend on: (i) the type of SI; (ii) the number of SI alleles maintained within populations; and (iii) the genetic composition of the outcrossed pollen landing on a given pistil (Byers & Meagher 1992; Vekemans *et al.* 1998). The latter factor is determined by patterns of pollen dispersal, notably the extent of geitonogamous pollination, population density, spatial genetic structure, and population isolation (Robledo-Arnuncio & Austerlitz 2006).

Because clones typically occur close to each other, vegetative reproduction is another potentially important factor determining paternal diversity when pollen dispersal is restricted. In several species with SI systems, high level of vegetative propagation had been observed, causing the occurrence of patches of identical genotypes. In *Prunus avium* populations, Stoeckel *et al.* (2006) found several clone groups, with more than half of the trees belonging to one of these clone groups, with clone size up to 34 individuals on an overall sample of 247 adult trees. Because most pollination events occur over short spatial distance, this may severely decrease multiple paternity of plants within these patches. In species with reduced number of SI alleles, this reduction in the number of mates may ultimately cause a substantial decrease in their seed-set. Such low seed-set in clonal species with SI
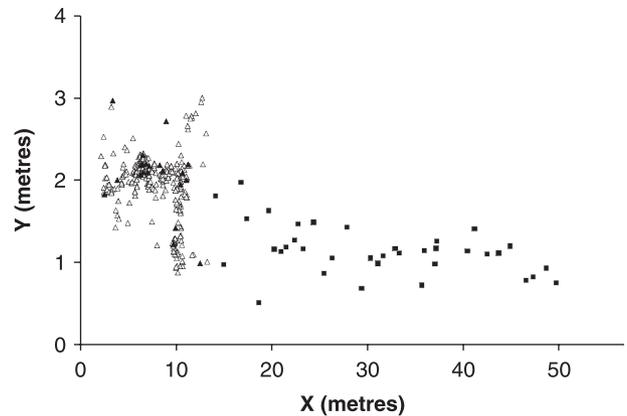


**Fig. 1** Spatial location of sampled individuals. Triangles represent individuals from area A (exhaustively sampled) and squares individuals from area B (nonexhaustively sampled). Solid triangles represent individuals from which seeds were collected for the paternity analysis.

has been reported in *Illicium floridanum* (Thien *et al.* 1983), in *Filipendula rubra* (Aspinwall & Christian 1992), *Hymenoxyx acaulis* (Demauro 1993), and the cactus *Stenocereus eruca* (Ricardo *et al.* 2006).

In this study, we investigated multiple paternity in a small isolated population of the herb *Arabidopsis halleri* (Brassicaceae), a species with sporophytic self-incompatibility. First, we estimated the extent of multiple paternity using adult and progeny data from 11 polymorphic microsatellite markers. Second, we analysed the following factors that influence patterns of multiple paternity within population: (i) the outcrossing rate; (ii) the distribution of pollen dispersal distances; (iii) the level of pollen immigration; (iv) the extent of clonality; and (v) the spatial genetic structure.

## Materials and methods

### Population sampling and DNA extraction

*Arabidopsis halleri* is a European insect-pollinated herbaceous species growing both in mountainous areas and on heavy-metal polluted sites, close to industrial areas, for example. It is a perennial species which survives in winter as a rosette and is able to develop stolons for asexual reproduction. *A. halleri* has been described as self-incompatible (Clauss & Koch 2006), but no study has been performed yet to estimate the functionality of this system in a natural population. We focused on a single isolated population located in Nivelle (France, 50°28′N, 3°27′E). With the aim of performing a paternity analysis, we intended to sample exhaustively all individuals from the population. Because of the very high plant density, an exhaustive sampling could be performed in only about one-half of the population (area A, Fig. 1), where 328 individuals were sampled by collecting leaves from

| Locus | Primer sequence (5′–3′) | Annealing temperature ($T_m$) |
|---|---|---|
| ATH | F: TCTATCAACAGAAACGCACCGAG | 56 |
| | R: CCACTTGTTTCTCTCTCTAG | |
| ELF3 | F: CGGAAGGACTGATATACAAGC | 50 |
| | R: GTTGGGTGTTCTGAAGAT | |
| GC16 | F: TTTTGGAGTTAGACACGGATCTG | 50 |
| | R: GTTGATCGCAGCTTGATAAGC | |
| LYR133 | F: GTTGCTGCTGCTGATGGTT | 56 |
| | R: CAAGGAAGGCAGCAAAGAAA | |
| LYR417 | F: AATCCCATCTCTTTCCGCTT | 56 |
| | R: GGAAGGAGAACCAACGATCA | |
| ATHZFPG | F: TTGCGTTTCCACATTTGTTT | 50 |
| | R: TGGGTCAATTCACATGTAGAGA | |
| ICE13 | F: GATCCTTCACCGGGTCTTG | 50 |
| | R: GTGGTGGAGACTCTTCGAGC | |
| NGA112 | F: TAATCACGTGTATGCAGCTGC | 50 |
| | R: CTCTCCACCTCCTCCAGTACC | |
| GC22 | F: GGTCTAATTGCCGTTGTTGC | 50 |
| | R: GAATTCTGTAACATCCCATTTCC | |
| H117 | F: CCGTCCTTGATCCTTGAGATTCTGAG | 50 |
| | R: CAATTCCGAAAATCATATTCATGCACC | |
| NGA361 | F: AGGGTTTTCCCAAAGAGATGA | 50 |
| | R: TCTTGTCCTTCGATTTTAGACCA | |
| MDC16 | F: GAGTGGCCTCGTGTAGAGAAAG | 50 |
| | R: TGTCACTCTTTTCCTCTGGTTTG | |

**Table 1** Primer sequences and annealing conditions used for amplifying microsatellite loci

each ramet that visually appeared as a distinct individual. In the remaining area of the population, we sampled leaves from 36 regularly spaced individuals (area B, Fig. 1) to investigate spatial genetic structure throughout the population. Each sampled individual was mapped. For each of 22 maternal plants randomly sampled from area A (Fig. 1), we sampled five fruits, collected the seeds, and grew the offspring for 8 weeks in a greenhouse. We thus analysed 26.7 (± 21.9 SD) seedlings per mother. All leaf samples (adult + offspring populations) were oven-dried at 55 °C, and DNA was extracted using the Nucleospin Multi-96 Plant extraction kit from Macherey-Nagel. Overall, we collected DNA from 364 adults and 256 offspring.

### Genotyping microsatellite markers

Based on a preliminary polymorphism survey using a small species-wide sample, 12 microsatellite loci were chosen to investigate genetic structure in this population. We amplified separately by polymerase chain reaction (PCR) five microsatellite loci: ATH, ELF3, GC16, LYR133 and LYR417 (see Table 1 for primer sequences). For these loci, each forward primer contained a 5′-tail of 19 bp homologous to the consensus M13 forward sequence (Oetting *et al.* 1995). The reaction mixture (15 μL) contained 20 ng DNA, 1× buffer (Applied Biosystems), 2 mM of MgCl$_2$, 200 μM of Fermentas

dNTP mix, 200 μg/mL of BSA, 0.2 μM of each microsatellite primer, 0.15 μM of M13 primer (fluorescence-labelled with either IRD-700 or IRD-800) and 0.025 U/μL of *Taq* polymerase (Ampli*Taq* DNA polymerase, Applied Biosystems). The amplification was carried out 5 min at 95 °C, eight cycles of 30 s at 95 °C, 45 s at 50 °C, 40 s at 72 °C, then 30 cycles of 30 s at 95 °C, 20 s at the annealing temperature (Table 1), 40 s at 72 °C, and one cycle of 7 min at 72 °C and performed in MJ Research PTC 200 thermocycler.

PCR products were separated on 6% polyacrylamide gels and visualized through fluorescence of M13 primers on a LI-COR sequencer. Size standards were run in every third lane to allow accurate band sizing.

We also amplified simultaneously by multiplex PCR seven additional loci (ATHZFPG, ICE13, NGA112, MDC16, GC22, H117, NGA361, Table 1) using primers labelled with Applied Biosystems dyes (VIC for MDC16 and GC22, 6FAM for NGA112 and NGA361, NED for H117 and ICE13, PET for ATHZFPG and LIZ for the sizing standard). The reaction mixture (15 μL) contained 20 ng DNA, 1× buffer (Applied Biosystems), 2 mM of MgCl$_2$, 0.6 mM of Fermentas dNTP mix, 200 μg/mL of BSA, 0.3 μL DMSO, 0.28 μM of each primer for H117, NGA361 and ATHZFPG, 0.1 μM for ICE13, 0.2 μM for GC22 and NGA112, 0.42 μM for MDC16 and 0.125 U/μL of *Taq* polymerase (Ampli*Taq* DNA polymerase, Applied Biosystems). Amplifications

were performed on an MJ Research PTC 200 thermocycler with the following conditions: 5 min at 95 °C, 35 cycles of 40 s at 95 °C, 40 s at 50 °C, 1 min at 72 °C, and one cycle of 7 min at 72 °C. Samples were loaded on a 16-capillary ABI 3100 sequencer, and genotypes were determined with the software GENEMAPPER (Applied Biosystems).

## *Analysis of multiple paternity*

We performed a paternity assignment analysis according to the maximum-likelihood method detailed in Marshall *et al*. (1998). Given the multilocus genotype of an offspring, of its mother, and of a set of potential fathers, the method assigns paternity to the father with the highest likelihood. Confidence in assignment was determined by computer simulations with the software CERVUS version 3.0 (Marshall *et al*. 1998). Briefly, assignment was considered unambiguous if the likelihood of the most likely father was significantly higher than that of the second most likely father. Significance was determined by simulating a population that included the genotyped individuals and a fixed proportion of unsampled individuals. These simulations were then used to compute for each offspring the log-likelihood difference between the two most likely potential fathers (Δ value) and determine the threshold value of Δ above which the most likely father was indeed the true father with a given confidence level. We chose the relaxed confidence level (i.e. 80%), which is commonly used in the literature when one of the two parents is already known and allowed for 5% genotyping error in likelihood calculations, which was a conservative hypothesis given that we had no precise estimation of our genotyping error rate. The average exclusion probability, given genetic information on the second parent, was calculated using the formula of Chakravarti & Li (1983).

The proportion of confidently assigned paternity was 94% (241 out of 256 offspring). Based on the unambiguous paternity assignments, we computed for each sampled mother the total number of distinct successful fathers, and the average number of distinct fathers per fruit. To estimate the correlation of outcrossed paternity among maternal sibs, we also computed the proportion of full-sib pairs among pairs of maternal sibs, $P_{full-sibs}$, either for each maternal family or as an average per fruit (Hardy *et al*. 2004). To determine whether nearby mothers were sired by genetically similar sets of fathers, we assessed the impact of the spatial location of the sampled mothers on the paternal composition of the pollen load siring them. This was done with the method developed by Hardy *et al*. (2004), which consists in computing Loiselle *et al*.'s (1995) kinship coefficients between inferred paternal genes for pairs of maternal families and testing their correlation with the pairwise spatial distance between mothers using a Mantel test. The software SPAGEDI (Hardy & Vekemans 2002) was used to perform these computations.

## *Analysis of outcrossing rate*

We used the mixed-mating model based on maternal families analysis (Ritland & Jain 1981) to estimate the maximum likelihood multilocus outcrossing rate $t_m$, considered as the most reliable statistic, and an average single-locus outcrossing rate ($t_s$) using the software MLTR (Ritland 2002). We compared values of $t_m$ and $t_s$ to detect biparental inbreeding (Ritland & Jain 1981). Significance of these statistics were tested by performing 1000 bootstraps across maternal families using the method described in Fernandez & Sork (2005). Bootstrap estimates of $t_m$ were compared to 1, corresponding to strict outcrossing, $t_m - t_s$ estimates were compared to 0.

## *Estimates of pollen dispersal and immigration rate based on progeny genotypes*

We used results from the paternity assignment analysis described above to obtain an estimate of the distribution of effective pollen dispersal across one generation (i.e. successful pollen dispersal events). For each of the unambiguous paternity assignments, we computed the distance between the mother and the inferred father. We estimated the axial standard deviation of pollen dispersal by the formula: $(\sigma_{pollen})^2 = 1/2\,M$, where $M$ is the mean of squared father–mother distances. We tested the observed mean of effective pollen dispersal distances against a distribution of mean pollen dispersal distances obtained under random mating. To obtain this distribution, we simulated replicate progeny samples by randomly drawing a father from the overall adult population for each progeny from each mother sampled in the field in area A. For each of 1000 replicates, we computed the mean pollen dispersal distance. The hypothesis of random mating was then tested by comparing the observed mean pollen dispersal distance with the simulated distribution and computing a *P* value.

The effective pollen dispersal distances estimated above depend on both the intrinsic capacity of pollen to disperse (the probability that a pollen grain travels a given distance also called pollen dispersal curve or pollen dispersal 'kernel'), and on the relative spatial location of individuals. To disentangle these two factors, we used a maximum-likelihood method coined 'the spatially explicit mixed mating model' (SEMM, Oddou-Muratorio *et al*. 2005), to jointly estimate the pollen dispersal kernel, the selfing rate (*s*) and the proportion (*m*) of pollen from outside the sampled population (pollen immigration rate). This method, which stems from the neighbourhood model (Burczyk *et al*. 2002), allows estimation of the kernel by removing the effect of the position of adults on the landscape on the effective pollen dispersal curve. This hypothesis may not be fully relevant for an insect-pollinated species because insects may stay longer in denser patches. Nevertheless, in other insect-pollinated species like *Sorbus torminalis*

(Oddou-Muratorio *et al.* 2005) or partially insect-pollinated species like *Brassica napus* (Devaux *et al.* 2007), this model gave realistic estimates of pollen dispersal distance. The method computes maximum-likelihood values of the parameters of the pollen dispersal kernel according to each of five different families [normal, exponential, exponential power, geometric and Weibull; see Austerlitz *et al.* (2004)], given the spatial position and genotypes of mothers and potential fathers, and offspring genotypes. We used the Akaike information criterion (AIC) to estimate the best-fitting family. We allowed for the same genotyping error rate (5%) as above, assuming that an allele could be replaced by any other in case of genotyping error. For each family of dispersal kernels, we estimated from the values of the dispersal parameters (*a* for the one-parameter families; *a* and *b* for the two-parameter families) the mean distance of pollen dispersal ($\delta$), and the axial standard deviation of pollen dispersal ($\sigma$) using the formulas given in Austerlitz *et al.* (2004).

### Genetic structure in the adult population and indirect estimation of dispersal distances

The software FSTAT (Goudet 1995) was used to compute the observed and expected heterozygosities as well as the inbreeding coefficient $F_I$ on the overall sample of adult plants, and to test for departure from Hardy–Weinberg genotypic proportions.

The extent of vegetative propagation among the exhaustively sampled individuals (group A) was quantified by considering that plants from different ramets with identical multilocus genotypes at all microsatellite loci belonged to the same genet. To assess the risk of error, we estimated the probability of finding twice the same multilocus genotype at random within the population, i.e. probability of identity (Waits *et al.* 2001), using the software GIMLET (Valiere 2002). We estimated clone size as the maximal distance between two ramets with identical genotype. Note that rosettes presenting signs of physical attachment or situated very close to each other (< 0.5 cm) were not collected and considered as the same individual, thus biasing upwards the estimation of clone size. To avoid genetic replicates, and because clones were tightly clustered (see below), a single ramet was randomly chosen from each genet for further analyses.

Patterns of seed and pollen dispersal were characterized indirectly from the spatial genetic structure. Using a Mantel test (Mantel 1967), we tested for spatial genetic structure by comparing the matrix of pairwise spatial distances between individuals (on a log-scale) with that of pairwise genetic relatedness estimated using the kinship coefficient of Loiselle *et al.* (1995). Under a model of isolation by distance in two dimensions, a negative linear relationship is expected between spatial distances (log-scale) and kinship. Assuming migration-drift equilibrium, the slope of the linear regression is expected to be inversely proportional to the variance of gene dispersal, and directly proportional to the level of spatial genetic structure (Hardy & Vekemans 1999). Accordingly, we performed a linear regression between pairwise spatial distances and kinship coefficients and estimated the slope (*b*). We computed the axial variance of gene dispersal ($\sigma^2$) according to the following formula (Rousset 2000; Vekemans & Hardy 2004): $\sigma^2 = -(1 - F_N)/(4\pi D_e b)$, where $F_N$ is the average kinship coefficient between neighbouring individuals and $D_e$ is the effective population density. This axial variance of gene dispersal ($\sigma^2$) depends both on seed and pollen dispersal (Crawford 1982). The variation of kinship between individuals according to spatial distance, $F(d)$ was estimated as the average kinship coefficients among pairs of individuals in each of 10 distance classes. The upper boundary distances of the 10 distance classes were 0.63 m; 1.15 m; 1.85 m; 2.73 m; 3.51 m; 4.21 m; 5.38 m; 7.63 m; 20.66 m; 47.56 m, which were determined so as to include the same number of pairs of individuals in each distance class. The value of $F_N$ was taken as the average in the first distance class. In area A, where sampling was exhaustive, the observed density was $D_o = 14$ individuals/m$^2$. In the absence of information on effective population density ($D_e$), we computed $\sigma^2$ assuming a ratio of $D_e/D_o$ equal to 1/5 or 1/10. The kinship coefficient computations were performed using the software SPAGEDI (Hardy & Vekemans 2002).

To infer the relative distances of pollen and seed dispersal from the spatial genetic structure, we applied the method developed by Heuertz *et al.* (2003) to identify the most likely combinations of axial standard deviations of pollen ($\sigma_{pollen}$) and seeds ($\sigma_{seed}$). To that end, we simulated spatial genetic structure in a theoretical population resembling as closely as possible to area A with a program kindly provided by O. Hardy which had already been used in Heuertz *et al.* (2003). We considered a strictly outcrossing two-dimensional population with the same density as in the observed population (14 individuals/m$^2$). We assumed a perennial life cycle with 80% annual survival probability, as estimated from greenhouse observations, in the absence of field data. In the simulations, we considered the same number of microsatellite loci and the same number of alleles per locus as observed in our study population. Simulations were started with identical population frequencies for all alleles at each locus. Pollen and seed dispersal followed a normal distribution with parameters $\sigma_{pollen}$ and $\sigma_{seed}$, respectively. Simulations of gene dispersal were made with values of $\sigma_{pollen}$ and $\sigma_{seed}$ ranging from 0 to 3 m, and the $\sigma_{pollen} - \sigma_{seed}$ combinations providing the best fit to the observed slope were retained as the most likely. We performed 1000 simulations for each [$\sigma_{pollen}$; $\sigma_{seed}$] combination. The slope observed in Nivelle was compared to the 1000 ranked slopes independently simulated for each [$\sigma_{pollen}$; $\sigma_{seed}$] combination. The parameter *P*, described by Sokal & Rohlf (1995) was used to determine which [$\sigma_{pollen}$; $\sigma_{seed}$] pair gives

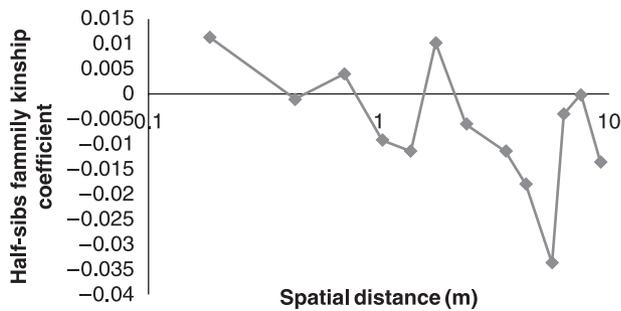**Fig. 2** Correlation between pairwise kinship coefficients between inferred paternal genes of maternal families and spatial distance (log scale).



**Fig. 3** Distribution of effective pollen dispersal distances obtained by paternity analysis: frequency of mating events for each class of distance between mates. The broken line represents the distribution of pollen dispersal distances expected under random mating.

the best fit to the data. $P = (n + 1)/(N + 1)$ if $n < N/2$ and $P = 1 - (n + 1)/(N + 1)$ if $n > N/2$ because it is a bilateral test; where $n$ is the rank of the observed slope among the ordered simulated slopes, and $N$ the number of replicates for each parameter set (here $N = 1000$).

## Results

### Patterns of multiple paternity

The cumulative exclusion probability over all loci in the paternity analysis was 0.961. The proportion of confidently assigned paternity was exceptionally high (94%), thus providing solid ground for a detailed analysis of patterns of multiple paternity. Analysis of the diversity of paternal parents siring offspring from a given maternal individual showed a strikingly high diversity. Of the 26.7 (± 21.9 SD) offspring harvested per maternal plant on average, the estimated number of distinct fathers was 22.4 (± 16.4). On average, the estimate of the fraction of maternal sibs that share the same father was very low, $P_{full-sib} = 3.8\% \pm 5.6\%$ (SD), hence correlated paternity among outcrossed maternal sibs is very low. High level of multiple paternity was also observed within single fruits, since the mean number of seeds per fruit observed was 2.7 (± 1.8 SD) and the mean number of distinct fathers per fruit was 2.6 (± 1.7). The fraction of maternal sibs per fruit that share the same father was surprisingly low, $P_{full-sib} = 1.5\% \pm 3.9\%$ (SD), indicating that correlated paternity among outcrossed maternal sibs is similar within and among fruits. We observed a decrease of the average pairwise kinship coefficients between inferred paternal genes of pairs of maternal families with spatial distance among mothers (Fig. 2). The Mantel test on this correlation was marginally significant ($P = 0.061$), indicating that plants located closer to each other were pollinated by more similar fathers than distant plants. Thus, correlated outcrossed paternity occurs not only at low level both within maternal families but also among neighbouring maternal families.
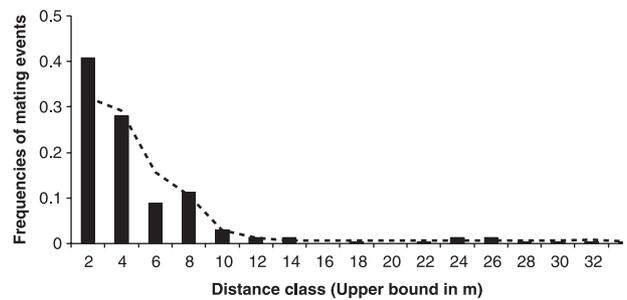
### Outcrossing rate

The maximum-likelihood multilocus estimate of outcrossing rate obtained under the mixed-mating model (Ritland & Jain 1981) was very high, $t_m = 0.987 \pm 0.052$ (SE). Bootstrap analysis showed that the $t_m$ statistic was not significantly different from 1, suggesting that selfing may be strictly avoided in this population. The comparison of single and multilocus outcrossing rate, $t_m - t_s = 0.047 \pm 0.052$ departed significantly from 0, thus indicating a significant amount of biparental inbreeding.

### Analyses of pollen dispersal based on progeny genotypes

From the results of the paternity analysis, we computed the distribution of realized pollen dispersal distances (Fig. 3). The observed mean pollen dispersal distance was significantly smaller than the mean pollen dispersal distance obtained under the hypothesis of random mating ($P$ value = 0.024), thus indicating that pollen dispersal was spatially restricted. Most successful pollination events occurred over limited distances, but successful population-wide pollen dispersal events were also observed occasionally. The mean effective pollen dispersal distance was 4.42 ± 6.3 m (SD), giving an estimate of the axial standard deviation of pollen dispersal, $\sigma_{pollen} = 5.44$ m.

Results from the analysis of the SEMM are summarized in Table 2 for different pollen dispersal kernels. For different shapes of the pollen dispersal kernel, we consistently estimated that about 10% of pollen came from outside area A from the sampled population ($m = 0.099–0.128$, Table 2). Consistent with the estimate from the mixed-mating model analysis, the SEMM model estimate of selfing rate (*s*) converged to 0% for each model of dispersal kernel. The random mating model (corresponding to uniform dispersal distribution) had an AIC much higher than all other dispersal kernels, indicating that pollen dispersal distance distribution was significantly different from random mating

**Table 2** Log-likelihood and parameter estimates from the spatially explicit mixed mating model for each of four shapes of the pollen dispersal kernel

| Curve | – log likelihood | $P$ | AIC | $m$ | $s$ | $a$ | $b$ | $\delta$ | $\sigma$ |
|---|---|---|---|---|---|---|---|---|---|
| Uniform (random mating) | −2498.0 | 2 | 5000.0 | 0.064 | 0 | | | | |
| Gaussian | −2486.4 | 3 | 4978.7 | 0.125 | 0 | 2.774 | 2 | 2.458 | 1.961 |
| Exponential | −2482.9 | 3 | 4971.8 | 0.128 | 0 | 1.474 | 1 | 2.948 | 2.553 |
| Exponential power | −2478.1 | 4 | 4964.2 | 0.103 | 0 | 0.001 | 0.209 | 142.10 | 233.168 |
| Geometric | −2477.6 | 4 | 4963.2 | 0.099 | 0 | 0.098 | 0.989 | ∞ | |
| Weibull | −2478.9 | 4 | 4965.9 | 0.114 | 0 | 5.144 | 1.421 | 4.678 | 4.064 |

$P$, number of parameters estimated; AIC, Akaike information criterion; $m$, immigration rate; $s$, selfing rate; $1-m-s$, outcrossing rate in the population; $a$, scale parameter; $b$, shape parameter; $\delta$, mean distance of pollen dispersal; $\sigma$, axial standard deviation of pollen dispersal.

**Table 3** Single-locus and multilocus statistics describing genetic diversity, departure from Hardy and Weinberg genotypic proportions, and spatial genetic structure in the adult population of *Arabidopsis halleri*

| Locus | Allele size range (in bp) | No. of alleles | $H_O$ | $H_E$ | $F_I$ | Slope ($b$) |
|---|---|---|---|---|---|---|
| AthZFPG | 132–170 | 3 | 0.3571 | 0.4831 | 0.279* | −0.0091 |
| GC22 | 184–188 | 2 | 0.5287 | 0.4905 | −0.074 | −0.0012 |
| H117 | 124–132 | 3 | 0.3291 | 0.3327 | 0.018 | −0.0030 |
| MDC16 | 113–115 | 2 | 0.3656 | 0.4162 | 0.124* | −0.0024 |
| nga112 | 176–202 | 4 | 0.4217 | 0.4772 | 0.137* | −0.0040 |
| nga361 | 123–127 | 3 | 0.2104 | 0.3099 | 0.296* | −0.0052 |
| GC16 | 162–174 | 4 | 0.5774 | 0.6299 | 0.085* | −0.0156 |
| ATH | 169–206 | 7 | 0.5194 | 0.5602 | 0.079* | −0.0062 |
| Lyr133 | 163–169 | 4 | 0.4889 | 0.5455 | 0.121* | −0.0049 |
| lyr417 | 203–247 | 4 | 0.5845 | 0.5661 | −0.028 | −0.0092 |
| Elf3 | 320–350 | 5 | 0.5472 | 0.6394 | 0.166* | −0.0288 |
| Mean | — | 3.7 | 0.4108 | 0.4542 | 0.106* | −0.0099 |

No. of alleles, observed ($H_O$) and expected ($H_E$) heterozygosities, inbreeding coefficient ($F_I$) and exact test of departure from Hardy–Weinberg genotypic proportions (*, $P < 0.05$), slope ($b$) of the regression of pairwise kinship coefficients on the logarithm of geographical distance.

in this population. The pollen dispersal kernel providing the best fit to the data was obtained with the geometric distribution. This distribution has a fat tail, meaning that pollen dispersal events occurred over a large range of distances, including long distances relative to the population size. The estimated value of the $b$ parameter for the geometric distribution was below 3, meaning that the estimated mean distance of pollen dispersal and axial standard deviation ($\sigma_{pollen}$) were infinite (Austerlitz *et al*. 2004). This result suggests that the scale of the population is too small for a proper analysis of the tail of the pollen dispersal curve, as observed in other studies (e.g. Devaux *et al*. 2007). However, the fact that the second best-fitting dispersal curve, the exponential power, has a very low estimated $b$ confirms the conclusion of fat-tailed pollen dispersal.

### Genetic structure in the adult population

Statistics describing genetic diversity in the total adult population (areas A and B) are given in Table 3. Locus Ice13 appeared monomorphic in Nivelle, and was thus discarded from the analyses. For the other loci, we observed from two to seven alleles per locus, with observed ($H_O$) and expected ($H_E$) heterozygosities ranging from 0.210 to 0.584, and from 0.310 to 0.639, respectively. The average inbreeding coefficient ($F_I$) in the adult population was 0.106 [95% CI: (0.081; 0.117)], showing an overall deficit in heterozygotes. This overall departure from Hardy–Weinberg genotypic proportions was significant at eight of the 11 loci analysed. The high number of moderately to highly variable microsatellite loci ensured that our analysis had a high power to detect clonal reproduction: the probability of identity was $2.5 \times 10^{-6}$. Overall, we found multiple ramets for only eight genets out of 355 distinct genotypes and the same genotype was never found in more than two ramets. Among those pairs, the mean distance between ramets was 0.10 m with a range of (0.004 m – 0.389 m). Hence, the level of vegetative reproduction was rather limited in this population.
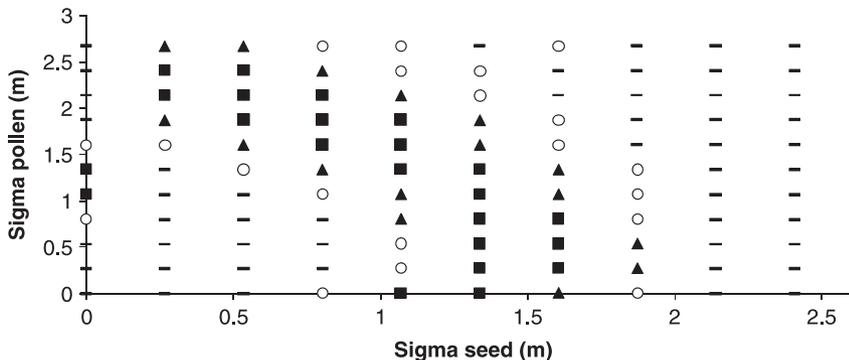
We found a significant correlation between pairwise multilocus kinship coefficients and spatial distance between individuals (Mantel test, $P < 0.001$), thus indicating a pattern of isolation by distance within the adult population. The
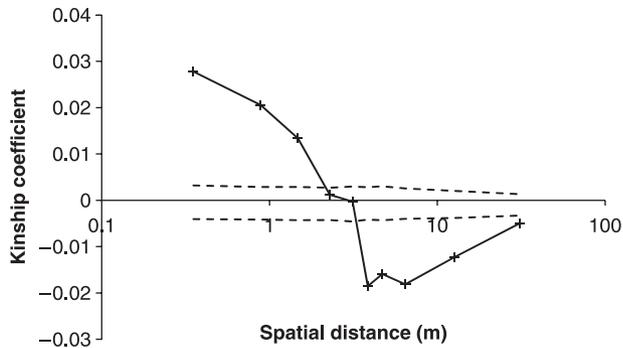
**Fig. 4** Correlogram showing spatial genetic structure. The y-axis represents the mean multilocus pairwise kinship coefficient between individuals; x-axis represents spatial distance intervals in meters (log-scale). The solid line represents the correlogram obtained on the population data and the broken line the 95% confidence interval.

average kinship showed an overall linear decrease up to about 8 m (Fig. 4). For eight class intervals out of 10, the average kinship coefficient between individuals was significantly different from that of random pairs of individuals obtained by permutations with a confidence interval of 95%. The slope of the linear regression (*b* value) was −0.010. To quantify the spatial genetic structure, we computed the *Sp* statistic defined by Vekemans & Hardy (2004) as $Sp = -b/(1 - F_I)$, where *b*, is the slope of the linear regression and $F_I$ is the mean kinship coefficient between individuals belonging to a first distance interval. The value of the *Sp* statistic obtained was 0.010.

Based on the value of *b* and assuming an isolation-by-distance model, the indirect estimate of the average axial parent–offspring spatial distance was estimated to be σ = 1.67 or 2.36 m, assuming a ratio of $D_o/D_e$ is equal to 1/5 or 1/10, respectively. The numerical simulations aimed at determining combinations of pollen and seed dispersal that best fit the observed pattern of isolation by distance suggested that the average axial pollen dispersal did not exceed 2.5 m, whereas the maximum value fitting the data for seed dispersal was 1.6 m (Fig. 5).

## Discussion

The level of multiple paternity per mother plant observed in our study was remarkably high, even within single fruits. Overall, the fraction of siblings sharing the same father (0.04) obtained in *Arabidopsis halleri* was one of the lowest reported for an herbaceous outcrossing species: values obtained in *Centaurea corymbosa* (Hardy *et al.* 2004), *Rutidosis leptorrhychoides* (Wells & Young 2002), *Grevillea iaspicula* (Hoebee & Young 2001), and *Mimulus guttatus* (Ritland 1989) were, respectively, 0.20, 0.37–0.65, 0.31–0.54, and 0.2. The observed level of multiple paternity in the Nivelle population was instead comparable to that typically observed for wind-pollinated plants (Smouse & Sork 2004). Our observation that multiple paternity is also high within single fruits suggests the occurrence of multiple pollinator visits to each flower, or extensive pollen carryover. However, the occurrence of correlated paternity among neighbouring maternal families (Fig. 2) indicates nonrandom patterns of pollen dispersal. We discuss these results in the light of our inferences on mating patterns and spatial genetic structure in this population.

### Self-incompatibility and multiple paternity

*Arabidopsis halleri* has been described as a self-incompatible species, but formal investigation was lacking (Clauss & Koch 2006). Functional haplotypes have recently been identified and sequenced in *A. halleri* at a locus considered to be orthologous to the S locus of the sister species *Arabidopsis lyrata* (Bechsgaard *et al.* 2006; Castric & Vekemans 2007), but the level of outcrossing in a natural *A. halleri* population had never been studied. Our direct estimate of population outcrossing rate based on either the maximum-likelihood method using maternal families (Ritland 2002) or the spatially explicit mixed mating model using both maternal families and potential paternal genotypes (Oddou-Muratorio *et al.* 2005) are consistent and indicate strict outcrossing. As we genotyped offspring from seedlings which had germinated and grown enough to produce leaves, the hypothesis that some selfing actually occurred



**Fig. 5** Fit of the observed slope of kinship–distance plot for the statistic *P* with simulations under various combinations of pollen and seed dispersal standard deviations ($\sigma_{seed}$ and $\sigma_{pollen}$) to the observed slope of kinship–distance plot in the population of *Arabidopsis halleri*. Hyphens, $P < 0.05$; circles, $0.05 < P < 0.2$; triangles, $0.2 < P < 0.5$; and squares, $P > 0.5$.

but remained masked by early acting inbreeding depression remains possible. But since selfed progeny were successfully obtained in controlled conditions (V. Llaurens, unpublished data), the hypothesis of early acting inbreeding depression alone could not explain the observed outcrossing rate. Our results are thus compatible with a fully functional self-incompatibility system in this population of *A. halleri*. The small but significant level of inbreeding detected in the population ($F_I = 0.11$) is most probably due to biparental inbreeding, as suggested by the observation of a significant difference in single-locus vs. multilocus estimates of selfing rate using Ritland's method (Ritland 2002). In the closely related species *A. lyrata*, several populations with completely self-compatible individuals have been described recently, with population outcrossing rate as low as 0.2, which reveals the occurrence of substantial variation in the functionality of the SI system (Mable *et al*. 2005). It will therefore be important to extend the present mating system analysis to additional natural populations to determine whether such variation in SI also occurs in *A. halleri*.

A fully functional SI system will exclude correlated paternity due to self-fertilization. This is certainly one factor that contributes to the low correlated paternity observed in *A. halleri*. However, SI may also restrict the diversity of sires, by filtering out outcrossed pollen donors sharing the same SI phenotype. This effect is expected to be substantial if the number of SI alleles is low and if the alleles have mostly codominant expression (Vekemans *et al*. 1998). Our results suggest that this situation does not occur in the Nivelle population, and investigations aiming at determining allelic diversity at the SI locus and patterns of dominance relationships among alleles are still ongoing.

*Vegetative propagation*

Because clones typically occur close to each other, vegetative reproduction is a potentially important factor determining paternal diversity when pollen dispersal is restricted. Excluding the individuals presenting signs of physical attachment, the extent of vegetative propagation seemed to be very limited in this population. Indeed, we found that only 2% of the multilocus genotypes exhibit more than a single ramet, with a distance between ramets not exceeding 0.39 m. In another population of *A. halleri* (Auby, France), about 22% of multilocus genotypes showed multiple ramets, with a maximum distance among ramets of a given clone below 1 m (Van Rossum *et al*. 2004). This suggests that clonal reproduction is particularly limited in the Nivelle population, but the discrepancy between these results could be due to the following three reasons: (i) genotypic diversity in the Auby population may have been underestimated because of the low number of microsatellite loci used in that study (five loci as compared to 11 in the present study); (ii) the sampling protocols differed slightly, i.e. rosettes were collected every 10 cm in the previous study, irrespective of whether traces of physical attachment were present or not, while we deliberately discarded such samples; (ii) plant density was very contrasted in the two populations (higher in Nivelle than in Auby on average), a factor that was indeed reported to negatively affect the extent of clonal reproduction in *A. halleri* (Van Rossum *et al*. 2004). Our results indicate that in spite of restricted pollen dispersal, the level of clonal reproduction observed in this *A. halleri* population is unlikely to negatively affect seed set.

*Shape of pollen dispersal*

The shape of the pollen dispersal distribution has a great influence on the extent of multiple paternity because it determines the number of different pollen sources potentially pollinating a given plant in the population.

Analysis using the SEMM model suggests that the observed pollen dispersal distribution shape was different from that expected under random mating and was highly leptokurtic, suggesting an excess of pollination at short distances, but also an excess of population-wide pollination events.

The average estimate of effective pollen dispersal distances (i.e. the distance observed between the two parents of each seed) by the direct method of paternity analysis ($\sigma_{pollen} = 4.72$ m) is consistent with the $\sigma_{pollen}$ found for actual pollen dispersal distances (pollen dispersal kernel) estimated by the SEMM model. Given the high density of the population, this would lead to a large neighbourhood size, suggesting a high number of potential partners available for each plant in this population.

The paternity analysis suggested that the population was moderately isolated, as at most 10% of pollen came from outside the population. This estimate is conservative, however, as only half of the population was exhaustively sampled. This value is rather low as compared to the literature: pollen immigration rate in tree species is generally much higher, with 40%, 65%, and 69% estimates in *Sorbus terminalis* (Oddou-Muratorio *et al*. 2005), *Quercus robur* and *Quercus petraea* (Streiff *et al*. 1999), respectively, although a lower value has been reported in a small isolated population of the tree species *Pinus sylvestris* (4.3% of all observed matings, Robledo-Arnuncio & Gil 2005). Intermediate values have typically been reported in herbs: from 0% in *Heloniopsis orientalis* (Miyazaki & Isagi 2000), 8.2–17.9% depending on the year of sampling in *Raphanus sativus* (Ellstrand & Marshall 1985), 38.1% in the herbaceous self-incompatible *Beta vulgaris* (Fénart *et al*. 2007), and up to 54.3% in *Iris fulva* (Cornman *et al*. 2004). Within species, variation in pollen immigration rate as a function of the degree of population isolation has been reported: Hoebee *et al*. (2007) observed that in *Sorbus torminalis* immigrant pollen was about 30% for continuous populations, whereas it was only

4% in a small isolated population. Altogether, our results are compatible with levels of immigration typically observed in moderately isolated populations of an herbaceous species, and suggest that immigration slightly contributes to increase the level of multiple paternity.

*Spatial genetic structure*

More than any other factor, we identified the spatial location of individuals in the population as the main determinant of the composition of pollen siring a given maternal plant. Spatial genetic structure analysis in the adult population revealed a typical pattern of isolation by distance due to restricted dispersal. The extent of within-population spatial genetic structure as estimated by the slope of the regression of pairwise kinship coefficients against the log of spatial distance between individuals ($Sp = 0.010$) was lower than observed in the other *A. halleri* population studied by Van Rossum *et al*. (2004) ($Sp = 0.003$), but was indeed strikingly similar to the average value obtained for 17 self-incompatible species, $Sp = 0.013 \pm 0.08$ (Vekemans & Hardy 2004). This is thus consistent with the general pattern of low extent of spatial genetic structure in strictly outcrossing species, in comparison with selfing species (mean $Sp = 0.143$) and species with a mixed mating system (mean $Sp = 0.037$).

Despite a low extent of spatial genetic structure, the indirect estimate of gene dispersal is rather low, $\sigma = 2.36$ m, with a maximum value of $\sigma_{pollen} \approx 2.5$ m and of $\sigma_{seed} \approx 1.5$ m. The estimation of $\sigma_{pollen}$ and $\sigma_{seed}$ could be biased to some extent by the normal distribution chosen for the pollen and seed dispersal curve in the model of Heuertz *et al*. (2003), which may appear unrealistic because it provided low fit for the pollen dispersal kernel in the mixed-mating model. They could also be biased by nonequilibrium conditions. The direct estimate of effective pollen dispersal distances (paternity analysis, $\sigma_{pollen} = 4.72$ m) was indeed higher than the maximum estimate based on spatial genetic structure ($\sigma_{pollen} < 2.5$ m), which assumed drift/mutation equilibrium. The latter estimate thus integrates pollen dispersal over the long-term, whereas the paternity estimate depends only on the last generation. We suggest, however, that the apparent contradiction between restricted dispersal distance and low extent of spatial genetic structure could be due to the high density of this population ($\approx 14$ ind/m²). Indeed, it has been shown that within species, plant density has a paramount influence on patterns of spatial genetic structure and pollen dispersal (Oddou-Muratorio *et al*. 2004; Vekemans & Hardy 2004). This is because although pollen grains travel short distances, it represents a large diversity of potential paternal individuals: for instance, assuming $\sigma_{pollen} = 2.5$ m and density of 14 ind/m², an individual is surrounded on average by 275 potential fathers at a distance within a circle of radius $\sigma_{pollen}$, thereby decreasing the probability that the father and mother would be genetically related. Altogether,

these results indicate that pollen and seed dispersal were restricted, but as the density found in this population was high, a large number of potential mates are available over short distance, allowing multiple paternity to be high.

## Conclusion

Using genotypes of adult plants and maternal progenies at 11 microsatellite loci, we obtained very high estimates of paternal diversity within and among fruits of maternal individuals in a self-incompatible species, despite spatially restricted pollen dispersal, and vegetative reproduction. We suggest that this result may be due to the very high plant density and/or high pollinator activity in this population. This result suggests that the constraint on mating patterns due to self-incompatibility must be weak in this population, and more generally depends on the ecological context.

## Acknowledgements

## References

Aspinwall N, Christian T (1992) Clonal structure, genotypic diversity, and seed production in populations of *Filipendula rubra* (Rosaceae) from the northcentral United States. *American Journal of Botany*, **79**, 294–299.

Austerlitz F, Dick CW, Dutech C *et al*. (2004) Using genetic markers to estimate the pollen dispersal curve. *Molecular Ecology*, **13**, 937–954.

Bechsgaard JS, Castric V, Charlesworth D, Vekemans X, Schierup MH (2006) The transition to self-compatibility in *Arabidopsis thaliana* and evolution within S-haplotypes over 10 Myr. *Molecular Biology and Evolution*, **23**, 1741–1750.

Broyles SB, Wyatt R (1991) Effective pollen dispersal in a natural population of *Asclepias exaltata*: the influence of pollinator behavior, genetic similarity, and mating success. *American Naturalist*, **138**, 1239–1249.

Burczyk J, Adams WT, Moran GF, Griffin AR (2002) Complex patterns of mating revealed in a Eucalyptus regnans seed orchard using allozyme markers and the neighbourhood model. *Molecular Ecology*, **11**, 2379–2391.

Byers DL, Meagher TR (1992) Mate availability in small populations of plant species with homomorphic sporophytic self-incompatibility. *Heredity*, **68**, 353–359.

Campbell DR (1998) Multiple paternity in fruits of *Ipomopsis aggregata* (Polemoniaceae). *American Journal of Botany*, **85**, 1022–1027.

Castric V, Vekemans X (2007) Evolution under strong balancing

selection: how many codons determine specificity at the female self-incompatibility gene *SRK* in Brassicaceae? *BMC Evolutionary Biology*, **7**, 132.

Chakravarti A, Li CC (1983) The effect of linkage on paternity calculations. In: *Inclusion Probabilities in Parentage Testing* (ed. Walker R), pp. 411–422. American Association of Blood Banks, Arlington, Virginia.

Clauss MJ, Koch MA (2006) Poorly known relatives of Arabidopsis thaliana. *Trends in Plant Science*, **11**, 449–459.

Cornman RS, Burke JM, Wesselingh RA, LAM (2004) Contrasting genetic structure of adults and progeny in a Louisiana iris hybrid population. *Evolution*, **58**, 2669–2681.

Crawford TJ (1982) What is a population? In: *The 23rd Symposium of the British Ecological Society* (ed. Shorrocks B). Blackwell Scientific Publications, Oxford.

Demauro MM (1993) Relationship of breeding system to rarity in the lakeside daisy (*Hymenoxys acaulis* var. *glabra*). *Conservation Biology*, **7**, 542–550.

Devaux C, Lavigne C, Austerlitz F, Klein EK (2007) Modelling and estimating pollen movement in oilseed rape (*Brassica napus*) at the landscape scale using genetic markers. *Molecular Ecology*, **16**, 487–499.

Ellstrand NC (1984) Multiple paternity within the fruits of the wild radish, *Raphanus sativus*. *American Naturalist*, **123**, 819–828.

Ellstrand NC, Marshall DL (1985) Interpopulation gene flow by pollen in wild radish, *Raphanus sativus*. *American Naturalist*, **126**, 606–616.

Fénart S, Austerlitz F, Cuguen J, Arnaud J-F (2007) Long distance pollen-mediated gene flow at a landscape level: the weed beet as a case study. *Molecular Ecology*, **16**, 3801–3813.

Fernandez MJF, Sork VL (2005) Mating patterns of subdivided population of the Andean oak (*Quercus humboldtii* Bonpl., Fagaceae). *Journal of Heredity*, **96**, 635–643.

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

Hardy OJ, González-Martínez SC, Colas B *et al.* (2004) Fine-scale genetic structure and gene dispersal in *Centaurea corymbosa* (Asteraceae). II. Correlated paternity within and among sibships. *Genetics*, **168**, 1601–1614.

Hardy OJ, Vekemans X (1999) Isolation by distance in a continuous population: reconciliation between spatial autocorrelation analysis and population genetics models. *Heredity*, **83**, 145–154.

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

Heuertz M, Vekemans X, Hausman J-F, Palada M, Hardy OJ (2003) Estimating seed vs. pollen dispersal from spatial genetic structure in the common ash. *Molecular Ecology*, **12**, 2483–2495.

Hoebee SE, Arnold U, Duggelin C *et al.* (2007) Mating patterns and contemporary gene flow by pollen in a large continuous and a small isolated population of the scattered forest tree *Sorbus torminalis*. *Heredity*, **99**, 47–55.

Hoebee SE, Young AG (2001) Low neighbourhood size and high interpopulation differentiation in the endangered shrub *Grevillea iaspicula* McGill (Proteaceae). *Heredity*, **86**, 489–496.

Loiselle BA, Sork VL, Nason J, Graham C (1995) Spatial genetic structure of a tropical understory shrub, *Psychotria officinalis* (Rubiaceae). *American Journal of Botany*, **82**, 1420–1425.

Mable BK, Robertson AV, Dart S, Di Berardo C, Witham L (2005) Breakdown of self-incompatibility in the perennial Arabidopsis lyrata (Brassicaceae) and its genetic consequences. *Evolution*, **59**, 1437–1448.

Mantel N (1967) The detection of disease clustering and generalized regression approach. *Journal of Cancer Research*, **27**, 209–220.

Marshall DL, Ellstrand NC (1985) Proximal causes of multiple paternity in wild radish, *Raphanus sativus*. *American Naturalist*, **126**, 596–605.

Marshall TC, Slate J, Kruuk L, Pemberton JM (1998) Statistical confidence for likelihood-based paternity inference in natural populations. *Molecular Ecology*, **7**, 639–655.

Miyazaki Y, Isagi Y (2000) Pollen flow and the intrapopulation genetic structure of *Heloniopsis orientalis* on the forest floor as determined using microsatellite markers. *Theoretical and Applied Genetics*, **101**, 718–723.

Oddou-Muratorio S, Demesure-Musch B, Pelissier R, Gouyon P-H (2004) Impacts of gene flow and logging history on the local genetic structure of a scattered tree species, *Sorbus torminalis* L. Crantz. *Molecular Ecology*, **13**, 3689–3702.

Oddou-Muratorio S, Klein EK, Austerlitz F (2005) Pollen flow in the wildservice tree, *Sorbus torminalis* (L.) Crantz. II. Pollen dispersal and heterogeneity in mating success inferred from parent–offspring analysis. *Molecular Ecology*, **14**, 4441–4452.

Oetting WS, Lee HK, Landers DJ (1995) Linkage analysis with mutiplexed short tandem polymorphism using infrared fluorescence and M13 tailed primers. *Genomics*, **30**, 450–458.

Ricardo CT, Corrado CA, Mandujano MC, Molina-Freaner F (2006) Reproductive consequences of clonal growth in *Stenocereus eruca*, a rare clonal cactus of the Sonoran desert. *Evolutionary Ecology*, **20**, 131–142.

Ritland K (1989) Correlated matings in the partial selfer *Mimulus guttatus*. *Evolution*, **43**, 848–859.

Ritland K (2002) Extensions of models for the estimation of mating systems using *n* independent loci. *Heredity*, **88**, 221–228.

Ritland K, Jain S (1981) A model for the estimation of outcrossing rate and gene frequencies using *n* independent loci. *Heredity*, **47**, 35–52.

Robledo-Arnuncio JJ, Austerlitz F (2006) Pollen dispersal in spatially aggregated populations. *American Naturalist*, **168**, 500–511.

Robledo-Arnuncio JJ, Gil L (2005) Patterns of pollen dispersal in a small population of *Pinus sylvestris* L. revealed by total-exclusion paternity analysis. *Heredity*, **94**, 13–22.

Rousset F (2000) Genetic differentiation between individuals. *Journal of Evolutionary Biology*, **13**, 58–62.

Smouse PE, Sork VL (2004) Measuring pollen flow in forest trees: an exposition of alternative approaches. *Forest Ecology and Management*, **191**, 21–38.

Sokal RR, Rohlf FJ (1995) *Biometry: the Principles and Practice of Statistics in Biological Research*, 3rd edn. W.H. Freeman, New York.

Stoeckel S, Grange J, Fernandez-Manjarres JF *et al.* (2006) Heterozygote excess in a self-incompatible and partially clonal forest tree species — *Prunus avium* L. *Molecular Ecology*, **15**, 2109–2118.

Streiff R, Ducousso A, Lexer C *et al.* (1999) Pollen dispersal inferred from paternity analysis in a mixed oak stand of *Quercus robur* L. and *Q. petraea* (Matt.) Liebl. *Molecular Ecology*, **8**, 831–841.

Thien LB, White DA, Yatsu LY (1983) The reproductive biology of a relict — *Illicium floridanum* Ellis. *American Journal of Botany*, **70**, 719–727.

Valiere N (2002) GIMLET: a computer program for analysing genetic individual identification data. *Molecular Ecology Notes*, **2**, 377–379.

Van Rossum F, Bonnin I, Fénart S *et al.* (2004) Spatial genetic structure within a metallicolous population of *Arabidopsis halleri*, a clonal, self-incompatible and heavy-metal-tolerant species. *Molecular Ecology*, **13**, 2959–2967.

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

Vekemans X, Schierup MH, Christiansen FB (1998) Mate availability and fecundity selection in multi-allelic self-incompatibility systems in plants. *Evolution*, **52**, 19–29.

Wagenius S, Lonsdorf E, Neuhauser C (2007) Patch aging and the *S*-allee effect: breeding system effects on the demographic response of plants to habitat fragmentation. *American Naturalist*, **169**, 383–397.

Waits LP, Luikart G, Taberlet P (2001) Estimating the probability of identity among genotypes in natural populations: cautions and guidelines. *Molecular Ecology*, **10**, 249–256.

Wells GP, Young AG (2002) Effetcs of seed dispersal on spatial genetic structure in populations of *Rutidosis leptorrhychoides* with different levels of correlated paternity. *Genetical Research Cambridge*, **79**, 219–226.

Yasui Y (1998) The 'genetic benefits' of female multiple mating reconsidered. *Trends in Ecology & Evolution*, **13**, 246–250.