Fine-Scale Genetic Structure and Gene Dispersal in *Centaurea corymbosa* (Asteraceae). II. Correlated Paternity Within and Among Sibships

Olivier J. Hardy,†,1 Santiago C. González-Martínez,‡,2 Bruno Colas,§ Hélène Fréville,†,** Agnès Mignot† and Isabelle Olivieri†

*Laboratoire de Génétique et Ecologie Végétales, Université Libre de Bruxelles, 1160 Brussels, Belgium, 1Unidad de Genética Forestal, Centro de Investigación Forestal, Instituto Nacional de Investigación Agraria y Tecnología Alimentaria, 28040 Madrid, Spain, †Institut des Sciences de l’Évolution de Montpellier, Université de Montpellier 2, 34035 Montpellier, France, ‡Laboratoire Fonctionnement et Evolution des Systèmes Ecologiques, Université de Paris VI, 75252 Paris, France and **Département Ecologie et Gestion de la Biodiversité, UMR 5176 CNRS-MNHN, 91800 Brunoy, France

Manuscript received February 17, 2004
Accepted for publication July 5, 2004

ABSTRACT

The fine-scale pattern of correlated paternity was characterized within a population of the narrow-endemic model plant species, *Centaurea corymbosa*, using microsatellites and natural progeny arrays. We used classical approaches to assess correlated mating within sibships and developed a new method based on pairwise kinship coefficients to assess correlated paternity within and among sibships in a spatio-temporal perspective. We also performed numerical simulations to assess the relative significance of different mechanisms promoting correlated paternity and to compare the statistical properties of different estimators of correlated paternity. Our new approach proved very informative to assess which factors contributed most to correlated paternity and presented good statistical properties. Within progeny arrays, we found that about one-fifth of offspring pairs were full-sibs. This level of correlated mating did not result from correlated pollen dispersal events (i.e., pollen codispersion) but rather from limited mate availability, the latter being due to limited pollen dispersal distances, the heterogeneity of pollen production among plants, phenological heterogeneity and, according to simulations, the self-incompatibility system. We point out the close connection between correlated paternity and the “TwoGener” approach recently developed to infer pollen dispersal and discuss the conditions to be met when applying the latter.

CORRELATED paternity refers to the fact that different offspring may be sired by the same father. Within maternal progeny arrays it is often referred to as “correlated mating” and can be expressed by the fraction of full-sib pairs (e.g., Ritland 1989; El-Kassaby and Jaquish 1996) or by the number of different fathers involved (e.g., Campbell 1998). In this context, pure half-sib and pure full-sib families represent the extreme alternatives of a continuum from uncorrelated to totally correlated mating events (e.g., polyads of mimosoid legumes and tropical figs; Nason et al. 1998). Correlated paternity can also be considered between maternal progeny arrays, where it can be expressed by the relative proportions of (paternal) half-sibs and non-sibs.

In plant populations, correlated paternity is important for several reasons. First, it is of evolutionary significance as it affects the genetic relatedness between maternal sibs and hence the response to selection when sibs are competing for maternal resources during seed maturation, affecting the effectiveness of selective fruit or embryo abortion, as well as resource allocation to each sex (Charnov 1982). Under limited seed dispersal, where interacting individuals are likely sibs, it may also act on the type of competitive interactions involved (e.g., kin selection; Hamilton 1964; Schuster and Mitton 1991; Rousset and Billiard 2000), the average fitness of competing siblings (Young 1981; Schmitt and Ehnhardt 1987; Karron and Marshall 1990, 1993), or the success of mating events between nearby individuals when inbreeding depression or self-incompatibility occurs. Second, together with the outcrossing rate, the pattern of correlated mating is a key parameter of the mating system (Ritland 1988, 2002) and can provide valuable information on pollination biology because it depends on a set of biological factors related in particular to floral biology, phenology, flowering intensity, pollinator behavior, pollen dispersal distances, and pollen competition. For example, several studies investigating correlated mating in a hierarchical fashion, distinguishing maternal sibs within and among flowers or inflorescences, gathered insightful information on the way pollen was brought on flowers (e.g., Schoen 1985; Ritland 1989; Morgan and Barrett 1990; Muona et al. 1991; Sampson 1998). Variation in pollen viability among pollen donors may result in male gametophyte competition.
(Nikkannen et al. 2000) and, consequently, affects patterns of correlated mating in successfully developed seeds. Finally, the level of correlated paternity is a key parameter on which many methods rely; for instance, as is detailed in the discussion, the so-called “TwoGener” approach recently developed to assess pollen dispersal distances (Austerlitz and Smouse 2001, 2002; Smouse et al. 2001; Irwin et al. 2003; Austerlitz et al. 2004) is fundamentally based on the level of correlated paternity. Quantifying correlated mating can also be important in quantitative genetic studies that make use of natural sib families, where half-sibs are often assumed (e.g., Pettit et al. 2001).

In plant populations, we can distinguish two types of mechanisms causing correlated paternity between outcrossing events: (1) correlated dispersal events, i.e., the codispersion of pollen from the same origin (nonindependent dispersal events), and (2) limited mate availability, i.e., the low number of pollen donors (independent dispersal events). The first mechanism occurs typically within multiple-seeded fruits when ovules are fertilized by a single pollen load (i.e., brought during a single pollinator visit) from few origins. Such correlated dispersal events could also occur among nearby flowers (e.g., within inflorescences) that are simultaneously receptive. Hence, the extent of pollen carryover and the spatio-temporal proximity of fertilization events are key determinants of the extent of correlated paternity due to correlated dispersal events. When dispersal events are independent, the pattern of pollen dispersal (dispersal distances and degree of anisotropy), the density of pollen producers, the heterogeneity of their flowering intensity, and the existence of a self-incompatibility system are key parameters determining the extent of correlated paternity due to limited mate availability. The heterogeneity of flowering phenology among plants is also important because even when there are many potential mating partners over the whole flowering period, mate availability can be limited on a given day if each plant releases most of its pollen over a short period. Hence, the temporal proximity of fertilization events can also be a key determinant of the extent of correlated paternity due to limited mate availability.

Different methods have been developed to quantify correlated paternity using genetic markers (e.g., Schoen 1985; Brown et al. 1986; Ritland 1988, 2002; Dewoody et al. 2000). The most precise consists in identifying the father of each offspring through a paternity analysis, so that the whole information regarding correlated paternity is available. However, paternity analyses require (1) an exhaustive sampling of the potential fathers and (2) sufficiently informative markers (highly polymorphic and/or many loci available) to identify unambiguously the true father. Alternative approaches are based on genotyped progeny arrays, using models to deduce the number of sires per family (e.g., Ellstrand 1984; Campbell 1998) or the fraction of sib-pairs sharing the same father (e.g., Ritland’s correlated-matings model; Ritland 1989, 2002). Although less precise than paternity analyses, these approaches are much more convenient to apply (nonexhaustive sampling) and can be used with less-polymorphic markers. Maximum-likelihood approaches have also been developed to assess whether a particular pair of individuals are half- or full-sibs (e.g., González-Martínez et al. 2003).

We show in this article that estimating the kinship coefficient between the paternal genes of any two given progenies (or equivalently between the fertilizing pollen grains) provides a convenient way to characterize the pattern of correlated paternity within as well as among progeny arrays. This approach is closely related to both Ritland’s correlated-matings model and the TwoGener method, but has the advantage of being more appropriate (1) to identify the factors causing correlated mating and (2) to characterize correlated paternity among maternal sibships in a spatial perspective. Simulations suggest that this method benefits better statistical properties.

We use microsatellite markers to characterize correlated paternity in Centaurea corymbosa Pourret (Asteraceae), a well-established model for narrow-endemic endangered species (Colas et al. 1997, 2001; Fréville et al. 1998, 2001, 2004; Pettit et al. 2001). Common garden experiments have provided some evidence of inbreeding depression in this species (H. Fréville and A. Mignot, unpublished results). In 2000, adults and maternal progenies from one population were genotyped to assess the selfing rate and pollen dispersal distances by paternity analysis (Hardy et al. 2004). In this article we use the same data to address the following questions:

1. What is the level of correlated mating within sibships and what is the level of correlated paternity among sibships?
2. To what extent is correlated mating due to correlated pollen dispersal events vs. limited mate availability?
3. Does correlated mating vary along the flowering period or among families according to the degree of isolation of the mother plant?
4. To what extent are correlated mating and correlated paternity accounted for by (a) limited pollen dispersal, (b) the heterogeneity of flowering intensity and flowering phenology among plants, and (c) the self-incompatibility system?

To answer these questions, we (1) use the results of the paternity analysis carried out previously (Hardy et al. 2004), (2) use standard approaches (Ritland 2002; Smouse et al. 2001) to assess correlated mating within progeny arrays, (3) develop a new approach based on pairwise kinship coefficients between the paternal genetic genotypes of offspring pairs to assess correlated paternity within as well as among progeny arrays, and (4) compare the level of correlated mating observed with the level expected under various hypotheses (ran-
dom mating vs. limited pollen dispersal, equal vs. unequal pollen release by the potential fathers, and presence of a self-incompatibility system), using numerical simulations. We sampled outer seeds and inner seeds within capitula on different dates. As is shown, this original sampling addresses the subject of the temporal proximity of fertilization events, which is of high importance but quite neglected in the literature. Finally, we also discuss the relationships between the different estimators used to assess correlated mating and we compare their statistical properties.

MATERIALS AND METHODS

Study organism: C. corymbosa is a monocarpic, self-incompatible perennial herb known from only six populations located within a 5-km² area along the French Mediterranean coast (Colas et al. 1997). It has specialized into an extreme habitat, the top of limestone cliffs where few other plant species survive. Its life cycle begins with a rosette stage during at least 2 years (5.5 years on average according to demographic data; Fréville et al. 2004). The year of flowering (the ultimate year as it is a monocarpic species), a mean of 30 capitula (range 1–200) are produced on several stems and mature sequentially. Capitula include on average 17 peripheral sterile ray flowers and 43 central hermaphroditic disc flowers (Colas et al. 2001). Each of the latter can produce only one seed and, on average, 40–60% of the flowers produce viable seeds (Colas et al. 2001). Flowers are protandrous and maturation is centripetal within capitula (i.e., outer flowers open first), so that 3–5 days are necessary for the opening of all flowers of a capitulum (H. Fréville and A. Miélot, unpublished data). Flowering begins in early May and extends through mid-July with a peak in June. Pollination is insect mediated, the main foragers observed in the field being small Coleoptera, Hymenoptera, Lepidoptera, Diptera, and Thysanoptera (S. Luijtjen, personal communication). C. corymbosa has a self-incompatible system. Paternity analyses made on the data set described below showed that (1) selfing is rare (1.6%) and (2) pollen dispersal is limited (50% of the fertilizing pollen moved by <11 m) and can be modeled by a bivariate exponential power function (dispersal kernel) more leptokurtic than the bivariate normal or exponential functions (i.e., fat-tailed dispersal distribution; Hardy et al. 2004).

Individual sampling: We studied one medium-size population (population “A” in Colas et al. 1997), considering here only its “main part” described in Hardy et al. (2004). We visited this population on June 1, 6, and 22, 2000, and mapped, in three dimensions, all individuals that were found to reproduce that year, totaling 96 plants. Leaf samples for DNA extraction were collected on 85 reproducing plants, the remaining 11 ones being unreachable (Hardy et al. 2004). To assess roughly the phenology and the pollen production per plant, the number of capitula at each stage (bud, flowering, senesced) was recorded for each accessible plant (55, 29, and 1 plants on our first, second and third visit, respectively).

Maternal sibships were obtained from 47 reproducing plants, collecting usually one capitulum per plant to limit human impact on this rare endangered species, but two or three capitula were collected on a few plants. Capitula from three dates were sampled on our first, second, and third visit, respectively. Seeds sampled on June 1 were among the very first ripe ones and were thus fertilized during the very early flowering period. Seeds collected on June 6 likely still belong to the early-flowering period. On the contrary, seeds collected on June 22 were probably fertilized during the mid-flowering period. No samplings were made during the late-flowering period; actually, the latter contributes little to plant reproduction because of the low fertilization rate and the high ovule abortion rate observed at this stage (Colas et al. 2001). Because capitula mature in a centripetal fashion, their peripheral seeds (outer ring) are fertilized 2 or 3 days before their central seeds (inner ring). Hence, to assess if correlated dispersal events (which could occur only between seeds from the same ring) contributed to correlated mating, we attempted to sample in each capitulum four central and four peripheral seeds. We succeeded to do so for 32 families; for the 15 remaining ones, we had to select eight seeds at random. DNA was extracted from young leaves from plantlets after seed germination, so that DNA material was obtained for 372 offspring from 47 families.

Genetic markers: Genotypes were assessed for 85 reproducing individuals and 372 offspring using nine microsatellite loci described in Fréville et al. (2000) and Hardy et al. (2004). These loci had 3–11 alleles and gene diversities ranged from 0.122 to 0.776 (average = 0.500). Comparison of offspring genotypes with maternal ones showed that genotyping error rate was <1% per locus (Hardy et al. 2004).

DATA ANALYSIS

Characterizing correlated paternity: A first characterization of correlated paternity was based on a paternity analysis using an exclusion approach described in Hardy et al. (2004): considering offspring having only one candidate father with a genotype fully compatible at all loci, paternity could be assessed for 123 offspring. From these data (excluding two selfing events), we computed the proportion of pairs of progenies sharing the same father (1) within progeny arrays (full-sibs) and (2) among progeny arrays, and according to the spatial distance separating the respective mothers, distinguishing pairs of offspring sampled on the same date or on different dates.

A second set of analyses used three closely related approaches to characterize correlated mating using the genotypes of all 372 progenies and their respective mothers: (1) Ritland’s (1989, 2002) correlated-mating model, using his software MLTR 2.4, which estimates in particular \( r_{P} \), the correlation of outcrossed paternity within progeny arrays (multilocus estimator); (2) the TwoGener approach (Austerlitz and Smouse 2001; Smouse et al. 2001) whereby global and pairwise \( \Phi_{T} \), expressing the differentiation among pollen pools sampled on different mothers, were estimated using a software provided by F. Austerlitz; and (3) a new approach described hereafter whereby the relatedness between the paternal genes of each pair of offspring was estimated using pairwise kinship coefficients.

The novel approach developed here requires the computation of a pairwise kinship coefficient between the paternal gametic genotypes of offspring pairs. To this end, the first step was to find out the paternal gametic genotype of each offspring by subtracting, for each locus, the maternal gametic genotype from the
offspring genotype (Smouse et al. 2001). This can be done unambiguously at a locus unless both the offspring and its mother are heterozygous for the same pair of alleles. For example, if an offspring carries alleles A and B and its mother alleles B and C, it means the father brought allele A. However, if the mother carried alleles A and B, the father could have brought allele A or allele B. The second step consists in computing, for each pair of offspring i and j, the kinship coefficients, $F_{ij}$, between the paternal gametic genotypes of i and j (with $n = 372$ offspring, there were 69,006 pairwise comparisons). The expectation of $F_{ij}$ is 0.5 if i and j have the same noninbred father, and 0 if they have different (and unrelated) fathers, so that twice an average $F_{ij}$ over many pairs of progenies should provide the proportion of these pairs being paternal sibs. We used J. Nason’s multilocus $F_{ij}$ estimator described in Loisel et al. (1995),

$$F_{ij} = \frac{1}{2} \sum_{k=1}^{n} \left( p_{ia} - \bar{p}_a \right) \left( p_{ja} - \bar{p}_a \right) / \sqrt{\sum_{k=1}^{n} \left( 1 - \bar{p}_a \right)},$$

where $p_{ia}$, $p_{ja}$ are the frequencies of allele a at locus l in i and j, considering only the paternal genes, $\bar{p}_a$ being the average $p_a$ over all n offspring. When the paternal allele can be assessed unambiguously at a given locus, $p_{ia} = 1$ for a being the paternal allele and $p_{ia} = 0$ for the other alleles. However, when there is an ambiguity on the paternal allele, we set $p_{ia} = 0.5$ for each of the two possible alleles and $p_{ia} = 0$ for the other ones. Consequently, if the paternal alleles of i and j at a locus are ambiguous and unambiguous, respectively, the single locus $F_{ij}$ is an average estimate of the kinship coefficients between the fathers of i and j, and between the mother of i and the father of j, potentially causing an estimation bias (see discussion). The correct computations were carried out using the software SPAGeDi (Hardy and Vekemans 2002) by encoding paternal gametic genotypes as diploid genotypes, using a homozygote when the paternal contribution was unambiguous and a heterozygote when it was ambiguous (i.e., recalling the example given above), pollen genotypes would be considered as AA and AB in the unambiguous and ambiguous cases, respectively.

The third step is to compute various averages from the matrix of pairwise $F_{ij}$ coefficients. First, averages were made over maternal sib pairs (i.e., within sibships, totaling 1575 sib-pairs), giving $F_S$ values. To assess the impact of the spatio-temporal proximity of fertilization events on correlated mating, we contrasted $F_S$ estimates for the following kinds of pairwise comparisons: (1) seeds sampled on the same date vs. on different dates, (2) seeds sampled within capitulum vs. among capitula, and (3) seeds sampled within ring (i.e., fertilized on the same date) vs. among rings (i.e., fertilized on different dates) within a capitulum. To assess if the extent of correlated mating varied in the course of the flowering period, we compared $F_S$ among the three different sampling dates. Moreover, to assess if there was an effect of the degree of plant isolation or of the heterogeneity of the proximity of the nearest neighbors, the $F_S$ values within each family were regressed on the number of flowering plants situated within 20 m of the mother plant (20 m being approximately the mean distance between mates; Hardy et al. 2004) as well as on the variance of the distances to the five closest flowering plants. Second, $F_S$ values between offspring from different mothers were averaged according to the distance separating the mother plants, so that correlated paternity among progeny arrays could be examined in a spatial perspective. Here again we distinguished pairs of offspring sampled on the same date or different dates. A jackknife procedure over loci provided approximate standard errors for all these estimates. Randomization procedures whereby the sampling dates or the spatial positions were reshuffled allowed us to test for temporal and spatial effects.

Simulations: Simulations of mating events in the studied population were carried out to investigate which mechanisms were necessary to account for the observed level of correlated mating. Following previous results (Hardy et al. 2004), we assumed that pollen dispersed according to an exponential power function (the kernel): the probability that a pollen grain moves by x and y along two orthogonal axes is $f(\alpha; \beta; x, y) = [\beta / 2\pi \Gamma(2/\beta)] \exp \left(-x^2 + y^2 \right)^{1/\beta}$, where $\Gamma$ is the gamma function, $r = (x^2 + y^2)^{1/\beta}$ is the distance crossed, and $\alpha$, $\beta$ are the dispersal parameters, controlling the extent and shape of the distribution, respectively. The probability that adult i is the father of a progeny collected on a given mother j is the proportion of the pollen received by j that originates from i: $P(i | \alpha, \beta, j) = w_i f(\alpha; \beta; x_i, y_j) / \sum_i w_i f(\alpha; \beta; x_i, y_j)$, where $x_i, y_j$ is the position of i relative to j, $w_i$ represents the relative amount of pollen produced by individual i, and the summation over k concerns all potential fathers.

To simulate mating events, all individuals found in the population but the mother were considered as potential fathers (i.e., self-fertilization not allowed), and pollen flow from other populations was assumed negligible (Colas et al. 1997; Fréville et al. 2001; Hardy et al. 2004). Each simulation run produced 372 offspring on the same set of mothers and with the same number of offspring per mother as in the real sample. For each offspring of mother j, we selected randomly a father according to $P(i | \alpha, \beta, j)$ values. The genotype of the progeny was then constructed by picking up randomly, for each locus, one gene from its mother and one gene from its father. To this end, we used the actual genotypes of the 85 adults that were genotyped and, for the 11 nonsampled adults, we constructed virtual genotypes by picking up alleles randomly according to their observed frequencies in the 85 genotyped adults. Once all 372 offspring were defined, the actual proportion of full-sibs within progeny arrays, $F_S$, was determined, and the mean relatedness between paternal genes within
sibships, $F_S$, was estimated following the approach described previously in this section. Using 1000 simulation replicates, we assessed the 95% confidence intervals on the $F_S$ values.

Simulations were run under various hypotheses, starting from panmixia and implementing step-by-step realistic processes expected to affect correlated mating:

1. **Random dispersal:** We approximated panmixia by setting $\alpha = 10,000$ and $w_i = 1$ for any $i$.

2. **Limited pollen dispersal, constant pollen production among adults:** Here we set $w_i = 1$ for any $i$, and $\alpha = 2.30$ and $\beta = 0.44$, the parameters fitting best with the pollen dispersal events inferred by paternity analysis (Hardy et al. 2004).

3. **Limited pollen dispersal, pollen production variable among individuals but constant through time:** The difference from simulation 2 is that, for each individual, $w_i$ was equal to the total number of capitula counted on $i$.

4. **Limited pollen dispersal, pollen production variable among individuals and through time:** The difference from simulation 2 is that, for each individual, $w_i$ was a random variable following a Poisson distribution with expectation equal to one-eighth the total number of capitula counted on $i$ (in accordance with observations, see results), and three different $w_i$ corresponding to the three sampling dates of sibs were drawn randomly (we assumed in the simulations that progeny arrays for a given family were fertilized on the same day).

5. **Limited pollen dispersal, pollen production variable among individuals and through time, self-incompatibility system:** The difference from simulation 4 is that, across individuals sharing at least one allele at the self-incompatibility locus was not allowed. Therefore, we considered 20 equi-frequent alleles randomly distributed among the 96 individuals. This number of alleles is close to the one expected at drift-selection-mutation equilibrium for an effective population size $N_e = 1000$ (a realistic value for Centaurea corymbosa at the species level; O. J. Hardy, unpublished results) and a mutation rate $\mu = 10^{-5}$ (Schierup et al. 1997). Note that self-incompatibility alleles being subject to strong frequency-dependent selection, the number of alleles at drift-mutation equilibrium within a population belonging to a metapopulation is essentially a function of the metapopulation effective size (Schierup et al. 1997).

**Statistical properties:** A thorough analysis of the statistical properties of the estimators used is beyond the scope of this article but, as we introduced a new approach to assess correlated paternity, it is useful to compare the statistical properties of the new estimator with classical approaches, at least for the sampling scheme adopted and the genetic markers used. In the absence of selfing and biparental inbreeding, $r_p$, $2\Phi_{3m}$ and $2\Phi_{5s}$ are all estimators of the proportion of full-sibs ($P_{fs}$) within progeny arrays (Ritland 1988; Austerlitz and Smouse 2001). Similarly, $2\Phi_{5s}$ estimates the proportion of progeny pairs from different families being paternal sibs ($P_{ps}$). The statistical properties were thus assessed by comparing the estimates obtained with the actual parameters using simulated data sets. To this end, we carried out the same kind of simulations presented in the previous section but, for each run, the $\alpha$ and $w_i$ parameters were taken from log-normal distributions to get a wide range of realized levels of correlated paternity (self-incompatibility was not simulated). For each run, we recorded $P_{fs}$ as well as $P_{ps}$ between progenies of families separated by $<3$ m [P_{ps(3m)}] and computed $F_S$ and the average $F_i$ between progenies of families separated by $<3$ m ($F_{3m}$). We also exported the genotypes obtained to compute $r_p$ and $\Phi_{3m}$ with adequate software (see above). We then regressed $r_p$, $2\Phi_{3m}$, and $2\Phi_{5m}$ on $P_{fs(3m)}$, and computed the mean square error (MSE; a combined measure of the bias and variance) for each estimator: $\text{MSE} = \Sigma (\hat{e} - e)^2/n$, where $e$ is the $r_p$, $2\Phi_{3m}$, $2\Phi_{5m}$ or $2\Phi_{3m}$ estimate of the $i$th replicate; $\hat{e}$ is the corresponding expected value [P_{fs(3m)}] or [P_{ps(3m)}]; and $n$ is the number of replicates ($n = 1000$ for $2\Phi_{5m}$ and $2\Phi_{3m}$ and $n = 200$ for $r_p$ and $2\Phi_{3m}$). Note that the sampling scheme of the artificial data sets is identical to the one of the experimental study.

We tried to identify two possible causes of estimator bias: (1) the presence of a spatial genetic structure in the parental generation (O. J. Hardy, unpublished results) and (2) the ambiguity regarding the paternal genotype when an offspring and its mother have the same heterozygote genotype. To investigate the impact of a spatial genetic structure, in a first set of simulations the parental genotypes were reconstructed by assigning alleles randomly according to the observed allele frequencies, whereas a second set of simulations used the original parental genotypes. To investigate the impact of the ambiguous parental genotypes, $F_S$ and $F_{3m}$ were also computed using the actual paternal genes (no ambiguities).

**RESULTS**

**Phenological data:** The total number of capitula produced per flowering plant, estimated on 85 accessible individuals, ranged from 3 to 140 (mean = 32.2, SD = 26.4). On our first visit (June 1), all 55 reproducing plants then characterized had at least one capitulum being in flower or having flowered and, considering all the capitula recorded that date, 23% of them had flowered and 14% were flowering. On June 6, 46% of the capitula from 29 other plants had flowered and 9% were flowering. On average, 12.5% of the total number of capitula per plant were flowering at any date, a percentage uncorrelated with the total number of capitula per plant ($r = −0.012, N = 87, P = 0.92$). We also
Average kinship coefficients between the paternal genes of maternal sibs according to the spatio-temporal proximity of fertilization events

**Seed position effect**

<table>
<thead>
<tr>
<th></th>
<th>Within capitulum</th>
<th>Between capitula:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>C-C</td>
<td>P-P</td>
</tr>
<tr>
<td>$F_s$</td>
<td>0.0976</td>
<td>0.1155</td>
</tr>
<tr>
<td>SE</td>
<td>0.0122</td>
<td>0.0227</td>
</tr>
<tr>
<td>No. of sib pairs</td>
<td>161</td>
<td>227</td>
</tr>
</tbody>
</table>

**Sampling date effect**

<table>
<thead>
<tr>
<th></th>
<th>Within date</th>
<th>Between dates:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>June 1</td>
<td>June 6</td>
</tr>
<tr>
<td>$F_s$</td>
<td>0.0994</td>
<td>0.0892</td>
</tr>
<tr>
<td>SE</td>
<td>0.0123</td>
<td>0.0230</td>
</tr>
<tr>
<td>No. of sib pairs</td>
<td>538</td>
<td>212</td>
</tr>
</tbody>
</table>

Standard errors (SE) are obtained by jackknifing over loci. Different kinds of sib pairs are considered according to their respective origins [from a same or different capitula and, in the former case, from the central (C) or peripheral (P) part of a capitulum] or the sampling date.

Assessment of paternal gametic genotypes: The percentage of offspring for which the paternal allele could not be determined unambiguously ranged from 2.4 to 19.4% among loci, the lowest values corresponding to the two least polymorphic loci because heterozygosity was low for these. For the seven other loci (which had 3–10 alleles and gene diversities between 0.39 and 0.83), these percentages ranged between 14.8 and 19.4% and were not clearly related to their level of polymorphism.

Correlated mating within progeny arrays: The paternity analysis detailed in Hardy et al. (2004) identified the fathers of 121 outcrossed progenies, making 121(121 – 1)/2 = 7260 offspring pairs, 148 of these being maternal sib pairs (offspring sharing the same mother). Twenty-nine of these sib pairs were full-sibs, so that maternal progeny arrays contained on average 19.6% of full-sib pairs. Interestingly, the mean distance traveled by pollen was 20.5 m for the 121 outcrossing events, whereas it was only 10.5 m when we considered solely offspring having at least one full-sib, suggesting that limited pollen dispersal is an important cause of correlated mating.

The MLTR program provided the following estimates when allele frequencies in pollen and seeds were assumed equal: multilocus estimate of the outcrossing rate $t = 0.972$ (SE = 0.011), parental inbreeding coefficient $F = 0.046$ (SE = 0.028), and multilocus estimate of the correlation of outcrossed paternity $r_p = 0.196$ (SE = 0.031). Following the TwoGener approach, the global differentiation among pollen pools having fertilized different mothers is $\Phi_{ST} = 0.0974$. The new procedure developed in this article gave an average kinship coefficient between the paternal genes of maternal sibs equal to $F_s = 0.0914$ (SE = 0.0094 by jackknifing over loci; 148 of these sib pairs (offspring sharing the same mother). Twenty-nine of these sib pairs were full-sibs, so that maternal (Table 1). Hence, the spatio-temporal proximity of fertilization events does not seem to affect significantly the extent of correlated mating for sibs, suggesting correlated dispersal events is not a significant mechanism promoting correlated mating in *C. corymbosa*. However, sibs collected at different phases of the flowering period seemed to show less correlated mating ($F_s = 0.0616$) than sibs collected during the same phase ($F_s = 0.0936$), but precision was lacking because few pairs of maternal sibs were sampled at different dates (Table 1) so that the difference was not statistically significant. Average $F_s$’s for all sib pairs per family were not significantly correlated with the number of adults found within 20
m around the mother plant ($r = -0.064; N = 46; P = 0.67$) or with the variance of distances to the five nearest flowering plants ($r = 0.028; N = 46; P = 0.80$).

**Correlated paternity among progeny arrays:** On the basis of the 121 outcrossing events identified by paternity analysis, the percentage of offspring pairs sharing the same father decreased with the spatial distance among mothers and, at short distances, it was much higher between offspring sampled on the same date than on different dates (Figure 1). Chi-square tests show that the difference between offspring sampled on the same date and offspring sampled on different dates is significant at distances $<1.5$ m ($P = 0.025$) and nearly so between 1.5 and 3 m ($P = 0.05$), but not significant at larger distances ($P > 0.05$).

The kinship coefficient $F_t$ between the paternal genes of offspring from different mothers decreased with the distance between mothers (Mantel test, $P < 0.001$), indicating again that limited pollen dispersal was a key determinant of correlated paternity (Figure 2). In agreement with the paternity analysis, for pairs of offspring sampled on mothers separated by $<10$ m, the extent of correlated paternity was much more pronounced for offspring pairs sampled on the same date than on different dates (Figure 2). Permutation tests showed that the difference is significant for distances $<3$ m ($P = 0.01$). This confirms the trend already observed within sibships that correlated paternity is more likely for progeny fertilized during the same period. Accordingly, pairwise $\Phi_{tt}$ between sibships shows that differentiation among pollen

**Figure 1.—** Percentages of offspring pairs sharing paternity within maternal sibship (solid symbols) and among sibships according to the distance (log scale) separating the mothers (open symbols). We distinguish pairwise comparisons including offspring sampled the same date (squares, solid line) or at different dates (triangles, dashed line). For pairwise comparisons within sibships and among sampling dates, no value is given because there were only five pairs. The asterisks show estimates within sampling dates that are significantly higher than estimates between sampling dates. These results are based on 121 outcrossing events identified by paternity analysis.

**Figure 2.—** Average kinship coefficient between the paternal genes of offspring pairs within maternal sibship (solid symbols) and among sibships according to the distance (log scale) separating the mothers (open symbols). We distinguish pairwise comparisons including offspring sampled the same date (squares, solid line) or at different dates (triangles, dashed line). Error bars indicate standard errors assessed by jackknifing loci.
pools increases with the distance between the mother and, at short distances, is higher between sibships sampled at different dates (Figure 3). For sibships separated by <2 m and sampled on the same date, there is nearly no differentiation ($\Phi_{TT} = 0.0037$).

**Simulation results:** The proportion of full-sibs ($P_{FS}$) and the relatedness between the paternal genes of sibs ($F_{S}$) increased progressively when the following features were successively added in the simulations: (1) panmixia, (2) limited pollen dispersal, (3) pollen production per individual variable among individuals, (4) pollen production per individual variable through time, and (5) self-incompatibility system (Table 2). The observed $F_{S}$ value (0.0914) was included into the 95% confidence interval obtained by simulations only in model 5 when all these features were implemented (0.0426–0.0964), in which case the percentage of full-sibs reached $P_{FS} = 17.2\%$ on average. It is noteworthy that limited dispersal taken alone accounted for only one-third of the effect ($P_{FS} = 6.4\%$), the heterogeneity of pollen production among plants and through time accounting for as much as limited dispersal. Finally, the self-incompatibility system also contributed substantially to correlated mating.

**Statistical properties:** Using simulated data sets, estimated $P_{FS}$ ($r_{ip}$, $2\Phi_{TT}$, or $2F_{S}$) were regressed on actual $P_{FS}$ so that estimator accuracy (bias) was assessed by the departure of the regression slope from unity and of the intercept from zero, and estimator precision (variance) was assessed by $R^2$. The MSE provides a combined measure of the impact of accuracy and precision. When there is no spatial genetic structure and no inbreeding in the parental population, Figure 4 shows that $P_{FS}$ tends to be underestimated by $r_{ip}$ and overestimated by $2\Phi_{TT}$ (at least for large $P_{FS}$), whereas $2F_{S}$ presents the best statistical properties (low bias and variance): $2F_{S} = 0.957P_{FS} + 0.0021$ ($R^2 = 0.990$; MSE = $7 \times 10^{-4}$), $2\Phi_{TT} = 1.184P_{FS} - 0.0063$ ($R^2 = 0.983$; MSE = $46 \times 10^{-4}$), $r_{ip} = 1.018P_{FS} - 0.0633$ ($R^2 = 0.892$; MSE = $107 \times 10^{-4}$).

Under low actual $P_{FS}$, $r_{ip}$ performs particularly badly, giving often fairly negative values (Figure 4B). The lower performances of $r_{ip}$ might be due to the fact that the MLTR software estimates simultaneously several statistics. Hence, we also used the MLTR software after setting the inbreeding coefficient and the outcrossing rate to (existence of a spatial genetic structure), all the estimates show better performances (0 and 1, respectively) and got better estimates (MSE = $39 \times 10^{-4}$) but still not as good as for $2F_{S}$. Among progeny arrays, $2F_{ip}$ somewhat underestimated $P_{FS}$: $2F_{FS} = 0.853P_{FS(3m)} - 0.0169$ ($R^2 = 0.907$; MSE = $17 \times 10^{-4}$). When the actual paternal genes were used to assess $F_{S}$ and $F_{fs}$ (no ambiguities), it did not really improve $F_{S}$ ($2F_{FS} = 0.991P_{FS} - 0.0148$; $R^2 = 0.989$; MSE = $9 \times 10^{-4}$) but improved a little $F_{FS}$ ($2F_{FS(3m)} = 1.006P_{FS(3m)} - 0.0199$; $R^2 = 0.906$; MSE = $10 \times 10^{-4}$), suggesting that ambiguous paternal genotypic causes only limited bias.

Using the original parental genotypes in the simulations (existence of a spatial genetic structure), all the estimators showed higher bias but their performances were not much affected within progeny arrays: $2F_{S} = 1.024P_{FS} + 0.0240$ ($R^2 = 0.989$; MSE = $19 \times 10^{-4}$), $2\Phi_{TT} = 1.356P_{FS} - 0.0081$ ($R^2 = 0.994$; MSE = $99 \times 10^{-4}$), $r_{ip} = 1.037P_{FS} - 0.0884$ ($R^2 = 0.936$; MSE = $100 \times 10^{-4}$). Among progeny arrays, the impact of biparental in-breeding was more significant but the estimator still displayed fairly good statistical properties: $2F_{FS} = 1.191P_{FS(3m)} - 0.0123$ ($R^2 = 0.940$; MSE = $14 \times 10^{-4}$).

**DISCUSSION**

**Consequences of correlated mating in C. corymbosa:** All analyses (Ritland’s correlated-matings model, Two-Gener approach, the pairwise kinship approach devel-
oped in this article, and the paternity analysis) provided a consistent result: within progeny arrays, about one-fifth of sib pairs were full-sibs. One over this proportion gives the “effective number of sires” (i.e., the number of fathers necessary to cause the observed level of correlated paternity assuming equal mating chances and independent mating events; Ritland 1989), which is $N_p = 5.1$. This value lies in the upper range of other herbaceous species (Table 3), among which the highest $N_p$ was observed for another Centaurea species (C. solstitialis; Table 3). Tree species tend to show higher values than herbs (Table 3).

\[ C. corymbosa \] is self-incompatible with very restricted seed dispersal (Colas et al. 1997) so that the extent of correlated mating within sibships determines to a large measure the degree of relatedness between competing seedlings. Compared to other herbaceous species, correlated paternity appears fairly low (higher $N_p$), so that we expect a fairly high level of genetic diversity among competing sib seedlings. It has been suggested that such diversity might favor the maternal fitness by reducing the competition among sibs through resource partitioning or, in a heterogeneous environment, by increasing the chance that some progenies are well adapted to some microsites (Young 1981). However, experimental works in other species failed to support these hypotheses, at least when comparing half-sibs vs. full-sib progeny arrays (e.g., Schmitt and Ehrhardt 1987; Karron and Marshall 1990, 1993). Another potential advantage of the low level of correlated paternity concerns colonization ability. With a sporophytic incompatibility system, typical in Asteraceae (Richards 1997), and assuming codominance among incompatibility alleles in the style and in the pollen, sib mating compatibility is at most 50% for half-sibs and 25% for full-sibs. Hence, a low level of correlated paternity should enhance the chance of establishment of a new population in a site recently colonized by a single plant or by a progeny array (e.g., if a whole ripe capitulum was transported, for example, by a bird using \textit{C. corymbosa} twigs for its nest).

It is still unclear whether pollen limitation (i.e., reduced seed set due to insufficient pollen fertilization) occurs in \textit{C. corymbosa}. Such pollen limitation was suggested by the reduced seed set of isolated plants, i.e., plants without flowering neighbors at < 4 m (Colas et al. 1997), but it appears that most pollen disperse over distances > 4 m (Hardy et al. 2004) and capitula are regularly visited by pollinators (S. Luijten, personal communication, see below). Our present results suggest that mate availability is limited, notably by the self-incompatibility system, conditions favoring pollen limitation. In animal-pollinated angiosperm species, pollen limitation is a common phenomenon as shown by experimental data comparing the percentages of fruit set in open-pollinated controls vs. in plants that received supplemental cross pollen (see review in Larson and Barrett 2000). However, such experiments performed on \textit{C. corymbosa} have not shown evidence of pollen limitation (S. Luijten, unpublished results).

Factors determining correlated mating in \textit{C. corymbosa}: For seeds sampled within capitula, the proportion of full-sibs was not lower when ovules were fertilized on different days (seeds from different rings) than on the same day (seeds from the same ring). Hence, correlated mating was not contributed by correlated dispersal events (i.e., the codispersion of pollen from the same origin by a single pollinator movement). We can thus conclude that correlated mating resulted essentially from limited mate availability, just as was observed recently by Wells and Young (2002) in the Australian endemic shrub \textit{Rutidosia leptorrhychoides} (Asteraceae).

Two lines of evidence indicate that limited pollen dispersal was the main factor limiting the diversity of sires. First, the paternity analysis showed that the mean distance between parents was lower for full-sibs (~ 10 m) than for random offspring (~ 20 m). Second, correlated paternity for pairs of offspring from different families was a decreasing function of the spatial distance between the mothers. Interestingly, for seeds collected the same date, the extent of correlated paternity among progeny arrays separated by < 1.5 m (i.e., the percentage of off-

\[ \begin{array}{c|c|c}
\text{Model} & P_{FS} (95\% \text{ C.I.}) & P_{FS} (%) \\
\hline
\text{Model 1} & \text{Random mating} & 0.0093 \quad 1.05 \\
 & & (0.0019, 0.0208) \\
\text{Model 2} & \text{Limited pollen dispersal} & 0.0363 \quad 6.38 \\
 & \text{Pollen production} & (0.0220, 0.0524) \\
 & \text{constant among adults} & \\
 & \text{Pollen production} & \\
 & \text{constant through time} & \\
\text{Model 3} & \text{Limited pollen dispersal} & 0.0431 \quad 9.19 \\
 & \text{Pollen production variable} & (0.0271, 0.0598) \\
 & \text{among adults} & \\
 & \text{Pollen production} & \\
 & \text{constant through time} & \\
\text{Model 4} & \text{Limited pollen dispersal} & 0.0582 \quad 15.09 \\
 & \text{Pollen production variable} & (0.0365, 0.0854) \\
 & \text{among adults} & \\
 & \text{Pollen production variable} & \\
 & \text{through time} & \\
\text{Model 5} & \text{Limited pollen dispersal} & 0.0659 \quad 17.20 \\
 & \text{Pollen production variable} & (0.0426, 0.0964) \\
 & \text{among adults} & \\
 & \text{Pollen production variable} & \\
 & \text{through time} & \\
 & \text{Self-incompatibility system} & \\
 & (20 alleles) & \\
\end{array} \]
spring pairs sharing the same father) was similar to the extent of correlated mating within progeny arrays (Figures 1 and 2). Hence, the fertilizing pollen loads arriving on adjacent plants had very similar compositions, at least within a few days. This was confirmed by pairwise $\Phi_{FT}$ values that were very low between progeny arrays sampled the same date (at $<2$ m, mean $\Phi_{FT} = 0.0037$, SE = 0.0160; Figure 3).

A second factor limiting significantly mate availability was the phenological heterogeneity among plants. This was clearly evident from the levels of correlated paternity for pairs of offspring collected on nearby mothers the same date vs. at different dates (Figures 1 and 2), a trend also observed within sibship. Accordingly, when progeny arrays sampled on different dates were considered, the pollen pools on nearby mothers were now differentiated (mean pairwise $\Phi_{FT}$ at $<2$ m = 0.0387, SE = 0.0187). Thus, the composition of the pollen load arriving on a given location varied in the course of the flowering period because the relative flowering intensities of the neighboring plants changed. Consequently, the diversity of sires at a given location and a given date was considerably lower than when averaged over the whole flowering period. Such an effect was also observed among years in the tree *Albizia julibrissin* (Irwin et al. 2003). As seeds from most progeny arrays were collected on a single date (different dates for only two arrays), our estimate of correlated mating is substantially higher than the average level over the whole flowering period.

When mating events were simulated accounting for limited pollen dispersal and the heterogeneity of pollen production and phenology among plants, the level of correlated mating was still inferior to the one observed. This discrepancy might result from an inaccurate parameterization of the simulated model, especially regarding the phenology of flowering for which precise information was lacking. However, even when phenology was modeled in a way to enhance correlated mating by amplifying the temporal heterogeneity of pollen production among plants, we hardly reached the level observed in reality (results not shown), until we implemented a self-incompatibility system in the simulations. The parameterization of such a system is difficult because information on the number and frequencies of the self-incompatibility alleles and their dominant/codominant effects in the pollen and stigma is lacking, so that we had to rely on theoretical expectations. With 20 self-incompatibility alleles, simulated levels of correlated mating were consistent with the observed value, and this was also the case using 10 or 30 alleles (results not shown). Note that the impact of the self-incompatibility system might be somewhat underestimated because we did not consider the spatial autocorrelation of self-incompatibility alleles that should arise under limited gene dispersal.

We cannot ensure that our simulation model contained all the essential ingredients determining correlated mating in the natural population, nature being often complex. However, these simulations give an idea of the relative impact of different factors on correlated paternity, showing in particular that limited pollen dispersal does not necessarily constitute the major cause.

In many species, full-sibs occur in much higher proportion within fruits than among fruits (Table 3), inducing a hierarchical pattern of correlated matings (*e.g.*, Schoen 1985; Brown et al. 1986; Ritland 1989; Morgan and Barrett 1990; Muona et al. 1991; Sampson 1998). This effect is particularly strong in species where pollen is carried in clusters [*e.g.*, Acacia (Muona et al. 1991)].
TABLE 3

Levels of correlated mating reported in natural populations of herbaceous and tree species.

The effective number of sires, $N_s$, is estimated as $1/r_p$ or $1/(2\Phi_{RT})$

<table>
<thead>
<tr>
<th>Species</th>
<th>$N_p$</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Herbs</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Centaurea corymbosa</td>
<td>5.1 (wC)</td>
<td>Sun and Ritland (1998)</td>
</tr>
<tr>
<td>C. solstitialis</td>
<td>8.8 (1.2–21) (T)</td>
<td>Cheptou et al. (2001)</td>
</tr>
<tr>
<td>Crepis sancta</td>
<td>1.4–3.5 (wC)</td>
<td>Ritland (1991)</td>
</tr>
<tr>
<td>Eichhornia paniculata</td>
<td>3.1 (wF); 4.2 (aF)</td>
<td>Morgan and Barrett (1990)</td>
</tr>
<tr>
<td>Glycine argyrea</td>
<td>1.2 (wF); 1.7 (aF)</td>
<td>Brown (1989)</td>
</tr>
<tr>
<td>Mimulus guttatus</td>
<td>2.5 (wF); 4.7 (aF)</td>
<td>Ritland (1989); Dudash and Ritland (1991)</td>
</tr>
<tr>
<td>Ruitidosis leptorrhynchoides</td>
<td>1.5–2.7 (T)</td>
<td>Wells and Young (2002)</td>
</tr>
<tr>
<td>Yucca filamentosa</td>
<td>1.6–3.1 (T)</td>
<td>Massey and Hamrick (1999)</td>
</tr>
<tr>
<td><strong>Trees</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acacia melanoxylon</td>
<td>1.3 (wF); 2.5 (wI); 30 (T)</td>
<td>Muona et al. (1991)</td>
</tr>
<tr>
<td>Albizia julibrissin</td>
<td>2.1 (T, 1 yr); 2.9 (T, 4 yr)</td>
<td>Irwin et al. (2003)</td>
</tr>
<tr>
<td>Caryocar brasiliense</td>
<td>4.9–11.6 (T)</td>
<td>Collevatti et al. (2001)</td>
</tr>
<tr>
<td>Dinizia excelsa</td>
<td>4.9 (T)</td>
<td>Dick et al. (2005)</td>
</tr>
<tr>
<td>Dryobalanops aromatica</td>
<td>12.5 (T)</td>
<td>Sampson (1998)</td>
</tr>
<tr>
<td>Eucalyptus rameliana</td>
<td>3.8 (wF); 11 (aF)</td>
<td>El-Kassaby and Jaquish (1996)</td>
</tr>
<tr>
<td>Pinus sylvestris</td>
<td>83–125 (T)</td>
<td>Sork et al. (2002)</td>
</tr>
<tr>
<td>Quercus lobata</td>
<td>3.7 (T)</td>
<td></td>
</tr>
<tr>
<td>Sorbus torminalis</td>
<td>4.9–14.7 (T)</td>
<td>S. Oddou-Muratorio (unpublished results)</td>
</tr>
</tbody>
</table>

Seed sampling terms are as follows: wC, within capitula (Asteraceae); wF, within flowers; aF, among flowers; wI, within inflorescences; T, over the whole plant; 1 yr, for 1 year; 4 yr, over 4 years.

1991) and Asclepias (Broyles and Wyatt 1991; Gold and Shore 1995)] or in some species with a particular type of plant-pollinator interactions [e.g., figs and fig wasps (Nason et al. 1998)], where multiple paternity within fruits is the exception rather than the rule. A significant impact of the timing of sib fertilization has also been reported recently in A. julibrissin (Irwin et al. 2003). In animal-pollinated species, correlated paternity within fruits is reduced when pollen loads from different fathers are deposited sequentially (e.g., Dudash and Ritland 1991) and/or when a single pollen load is a mixture of different fathers, implying extensive pollen carryover (e.g., Marshall and Ellstrand 1985; Campbell 1998). In Asteraceae, such as C. corymbosa, fruits are one-seeded, so that correlated paternity cannot result from the deposition of a pollen load on a given stigma, but stigma from adjacent flowers within capitula might be fertilized by the same pollen load, capitula in Asteraceae having the same functional role to attract pollinators as typical flowers do in many other families. When there is a delayed, centripetal maturation of flowers within capitulum, correlated paternity through this process is limited to adjacent flowers maturing simultaneously. In C. corymbosa, the observation that correlated paternity within sibship was not contributed by such a mechanism implies that each flowering capitulum got several pollen loads each day (and was thus visited several times) and/or that pollen carryover was extensive. This is consistent with field observations carried out in 2002 in population A, showing that capitula were visited mainly by Coleoptera (Oedemera nobilis Scopilli) at a rate of 3.6 visits per hour (S. Luijten, personal communication).

**Some unifications:** The TwoGener approach was devised to infer pollen dispersal distances on the basis of the pattern of differentiation among pollen pools having fertilized different maternal sibships (Austerlitz and Smouse 2001, 2002; Smouse et al. 2001; Austerlitz et al. 2004). This approach is fundamentally based on the pattern of correlated paternity because differentiation among pollen pools implies that paternity is correlated within sibships. Actually, $\Phi_{RT}$, the measure of genetic differentiation among pollen pools having fertilized different mothers in the TwoGener approach; $r_p$, the correlation of outcrossed paternity within sibships following Ritland’s (1989) approach; and $F_s$, the mean kinship coefficient between the paternal gametic genotypes of maternal sibs following our approach, are all estimators of the average genetic relatedness among fathers having fertilized maternal sibships. However, they are not exactly equal because $r_p$ is a function of the probability that an allele in one father is identical to one of the alleles of the other father ("relationship" coefficient; Wright 1922), whereas $\Phi_{RT}$ and $F_s$ are functions of the probability that an allele in one father is identical to a random allele of the other father ("kinship" coefficient; Malecót 1967). In the absence of inbreeding, in a diploid species, one expects $r_p = 2\Phi_{RT}$ (Cockerham 1969), and this is indeed what we found with our data. $F_s$ and $\Phi_{RT}$ gave very similar values but there is, however,
a difference: $\Phi_{TT}$ measures the paternal genetic resemblance within sibships relative to the one among sibships, whereas $F_s$ measures the paternal genetic resemblance within sibships relative to the average found in the sample, both within and among sibships. Hence, one expects $\Phi_{TT}$ to be slightly higher than $F_s$ as observed with our data ($\Phi_{TT} = 0.0974, F_s = 0.0914$).

To estimate the proportion of full-sibs within progeny arrays in an outcrossed population, $r_p, 2\Phi_{TT}$, and $2F_s$ are all fairly good estimators. However, simulations showed that, for the sampling scheme and the genetic markers we considered, $2F_s$ was the best estimator and $r_p$ the worst one. Assessing the statistical properties of these estimators in a more general way (i.e., for different types of sampling schemes and allele frequencies) is beyond the scope of this article but it would be worth doing in a future study.

**Implications of our results for the TwoGener approach:** The TwoGener approach relies on a model predicting how $\Phi_{TT}$ (global or pairwise values) varies with the extent of pollen dispersal (Austerlitz and Smouse 2001, 2002; Austerlitz et al. 2004), given the density of pollen donors. It assumes implicitly that limited mate availability due to limited pollen dispersal is the major factor determining the extent of correlated mating. It implies among other things that pollen dispersal events must be independent. Independent dispersal events seem to hold for C. corymbosa but, in other species, especially when sibs come from multiseeded fruits, the extent of correlated mating may be essentially due to correlated dispersal events. In C. corymbosa, the heterogeneity in pollen production and phenology among plants and (probably) the presence of a self-incompatibility system contributed substantially to correlated mating, so that limited pollen dispersal accounted for less than half the observed level. Under such a condition, the TwoGener approach is expected to underestimate significantly pollen dispersal distances.

Hence, care must be taken when applying the TwoGener approach because the sampling scheme can have considerable impacts on the output (see also Irwin et al. 2003 for the impact of temporal sampling). To avoid bias due to correlated dispersal, sibs should be sampled on distant flowers. To account for the heterogeneity of pollen production among plants, the density of pollen donors, $d$ (census density), should be corrected to provide an effective density, $d_e$. This can be done if the relative flowering intensity regarding pollen production, $w$, can be assessed for a random set of $N$ plants: $d_e = (\sum w^2)/(N \sum w^2)$, where summations hold for $i = 1$ to $N$. To account for the phenological heterogeneity, if all progenies were fertilized during a given period, $w_i$ should be assessed during this period. Applying this formula in our population and assuming that pollen production was proportional to the number of capitula, $d = 0.0167$ flowering plants/m$^2$, $d_e = 0.100$ when $w_i$ is the total number of capitula per plant (effective density over the whole flowering period), and $d_e = 0.0085$ when $w_i$ is the number of flowering capitula per plant at the date of visit (instantaneous effective density); the effective density was thus half the census one. Self-incompatibility also reduces the effective density; as a first approximation, $d_e$ should be multiplied by the proportion of compatible mating, but this information is rarely available. Note that $d_e$ can also be estimated simultaneously with the dispersal parameters using pairwise $\Phi_{TT}$ (Austerlitz and Smouse 2002; Austerlitz et al. 2004).

Finally, it should be noted that the TwoGener approach assumes that flowering individuals are homogeneously spread over a two-dimensional landscape larger than the scale of pollen dispersal. The studied population was an elongated one (200-m length and 20- to 60-m width) with clumped individuals (Hardy et al. 2004), a geometry that enhances correlated mating. Consequently, even when the density was corrected ($d_e = 0.0085$), the TwoGener approach did not succeed in providing pollen dispersal parameters consistent with paternity analyses in our population: assuming an exponential dispersal kernel, the estimated mean pollen dispersal distance was $d_e = 9.8$ m from the global $\Phi_{TT}$, whereas the best fit to the 121 observed mating events gave $d_e = 37$ m ($d_e = 91$ m assuming an exponential power kernel; Hardy et al. 2004). It must be noted here that $d_e$ is not an estimate of the mean distance between mates (which was $d_e = 21.6$ m), unless individuals are distributed randomly over an area much larger than the scale of pollen dispersal (Hardy et al. 2004). In conclusion, although the TwoGener approach provides a convenient way to assess pollen dispersal in natural populations, care must be taken in the interpretation of the results because reliable estimates require adequate sampling schemes and population geometry.

**Comparing the methods to assess correlated mating:** The identification of mating events by a paternity analysis is *a priori* the most accurate way of characterizing correlated mating, but successful paternity analyses are often difficult to obtain. In our population, paternity was assessed for about one-third of the offspring sampled, which is quite good, but implies that correlated paternity could be assessed for only $(1/3)^2 = 1/9$ of the total number of offspring pairs. On the contrary, the approach based on pairwise kinship coefficients considers all pairs of offspring and can be applied even if potential fathers have not been genotyped. Nevertheless, pairwise kinship coefficients typically suffer high sampling variance (Lynch and Ritland 1999) so that they are not powerful to characterize correlated mating for a single pair of offspring, and they do not exploit fully the multilocus information (they are averages of single-locus estimates). Ritland’s correlated-matings model uses maximum-likelihood methods to estimate mating system parameters and therefore better exploits multilocus information. However, the statistical properties of the $r_p$ estimator were not so good and, as designed
nowadays in Ritland’s MLTR software (Ritland 2002), correlated paternity assessment is restricted within sibships. As we have seen, much information can be obtained by assessing correlated paternity among sibships in a spatial perspective.

In conclusion, we showed that pairwise kinship coefficients are convenient and powerful to characterize correlated paternity on a fine scale and to identify its causal mechanisms because (1) they can assess correlated paternity both within and among sibships, (2) they can be averaged over particular types of pairs (e.g., within or among sampling dates) or regressed on explanatory variables (e.g., spatial distance), and (3) their statistical properties are good. Further methodological progress could be achieved by designing methods to assess correlated paternity, exploiting more fully the multilocus information.

We thank Eric Imbert and three anonymous reviewers for their comments on a previous draft and Frédéric Austerlitz for providing software to compute TwoGener parameters as well as for helpful discussions. This work was supported by the Ministère de l’Aménagement du Territoire et de l’Environnement (MATE; no. 98153 to I. Olivieri) through the National Programme Diversitas, Fragmented Populations network. The molecular work was also funded by the European Union program “Plant Dispersal” headed by Ben Vosman. The Belgian National Fund for Scientific Research and the Commune du Territoire et de l’Environnement (MATE; no. 98153 to I. Olivieri) through the National Programme Diversitas, Fragmented Populations network. The molecular work was also funded by the European Union program “Plant Dispersal” headed by Ben Vosman. The Belgian National Fund for Scientific Research and the Commune du Territoire et de l’Environnement (MATE; no. 98153 to I. Olivieri) through the National Programme Diversitas, Fragmented Populations network. The molecular work was also funded by the European Union program “Plant Dispersal” headed by Ben Vosman.

LITERATURE CITED


Marshall, D. L., and N. C. Ellstrand, 1985 Proximal causes of


Communicating editor: O. Savolainen