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,**[†,‡] **Bruno Colas,**[§]
Hélène Fréville,[†,**] **Agnès Mignot**[†] **and Isabelle Olivieri**[†]

*Laboratoire de Génétique et Ecologie Végétales, Université Libre de Bruxelles, 1160 Brussels, Belgium, ‡Unidad de Genética Forestal, Centro de Investigación Forestal, Instituto Nacional de Investigación Agraria y Tecnología Alimentaria, 28040 Madrid, Spain, †Institut des Sciences de l'Evolution 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

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 EHRHARDT 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

[1]*Corresponding author:* Laboratoire d'Eco-éthologie Evolutive, Université Libre de Bruxelles, CP 160/12, 50 Av. F. Roosevelt, 1050 Brussels, Belgium. E-mail: ohardy@ulb.ac.be

(Nikkanen *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.*, Petit *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 also 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; Petit *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 gametic 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 3-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. Mignot, 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. Luijten, 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 12, 13, and 22 plants 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-matings 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 Φ_{FT}, 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 LOISELLE *et al.* (1995),

$$F_{ij} = \left[\sum_l \sum_a (p_{lai} - \overline{p}_{la})(p_{laj} - \overline{p}_{la}) / \sum_l \sum_a \overline{p}_{la}(1 - \overline{p}_{la}) \right] + 1/2(n-1),$$

where p_{lai}, p_{laj} are the frequencies of allele *a* at locus *l* in *i* and *j*, considering only the paternal genes, $\overline{p}_{la}$ being the average p_{lai} over all *n* offspring. When the paternal allele can be assessed unambiguously at a given locus, $p_{lai} = 1$ for *a* being the paternal allele and $p_{lai} = 0$ for the other alleles. However, when there is an ambiguity on the paternal allele, we set $p_{lai} = 0.5$ for each of the two possible alleles and $p_{lai} = 0$ 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, totalizing 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_{ij} 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 $f(\alpha, \beta; x, y) = [\beta/2\pi\Gamma(2/\beta)]\exp(-(r/\alpha)^\beta)$, where Γ is the gamma function, $r = (x^2 + y^2)^{1/2}$ is the distance crossed, and α, β 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*: $P(i|\alpha, \beta, j) = w_i f(\alpha, \beta; x_i, y_i) / \sum_k w_k f(\alpha, \beta; x_k, y_k)$, where x_i, y_i 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, P_{FS}, 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 the addition of a sporophytic self-incompatibility system, so that crossing among 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 *C. 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, $2\Phi_{FT}$ and $2F_S$ are all estimators of the proportion of full-sibs (P_{FS}) within progeny arrays (RITLAND 1988; AUSTERLITZ and SMOUSE 2001). Similarly, $2F_{ij}$ 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 α 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_{ij} between progenies of families separated by <3 m (F_{3m}). We also exported the genotypes obtained to compute r_p and Φ_{FT} with adequate software (see above). We then regressed r_p, $2\Phi_{FT}$, and $2F_S$ on P_{FS}, as well as $2F_{3m}$ on $P_{PS(3m)}$, and computed the mean square error (MSE; a combined measure of the bias and variance) for each estimator: $\text{MSE} = \Sigma(o_i - e_i)^2 / n$, where o_i is the r_p, $2\Phi_{FT}$, $2F_S$, or $2F_{3m}$ estimate of the ith replicate; e_i is the corresponding expected value [P_{FS} or $P_{PS(3m)}$]; and n is the number of replicates ($n = 1000$ for $2F_S$ and $2F_{3m}$ and $n = 200$ for r_p and $2\Phi_{FT}$). 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 paternal 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

TABLE 1

Average kinship coefficients between the paternal genes of maternal sibs according to the spatio-temporal proximity of fertilization events

| | Seed position effect | | | | |
| | Within capitulum | | | | Between capitula: |
	C-C	P-P	C-P	Average	Average
F_S	0.0976	0.1155	0.0926	0.0940	0.0841
SE	0.0122	0.0227	0.0200	0.0085	0.0184
No. of sib pairs	161	227	421	1152	413

| | Sampling date effect | | | | |
| | Within date | | | | Between dates: |
	June 1	June 6	June 22	Average	Average
F_S	0.0994	0.0892	0.0905	0.0936	0.0616
SE	0.0123	0.0230	0.0133	0.0108	0.0228
No. of sib pairs	538	212	721	1471	104

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.

found that the observed number of flowering capitula on a given plant and a given date closely followed a Poisson distribution with expectation equal to 12.5% the total number of capitula ($P = 0.912$).

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 $\Phi_{FT} = 0.0974$. 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; number of sib pairs $N_p = 1575$), which is close to the global Φ_{FT} value and close to half the r_p value, as expected (see DISCUSSION). In the absence of biparental inbreeding, r_p, or twice Φ_{FT} or F_S, is expected to approximate the proportion of full-sibs, which might thus reach ~19%, in agreement with the results of the paternity analysis.

When maternal sib pairs were distinguished according to the relative positions of seeds within and among capitula, or according to the sampling date of seeds (and thus the phase of the flowering period when fertilization occurred), no significant effect was observed on F_S values (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
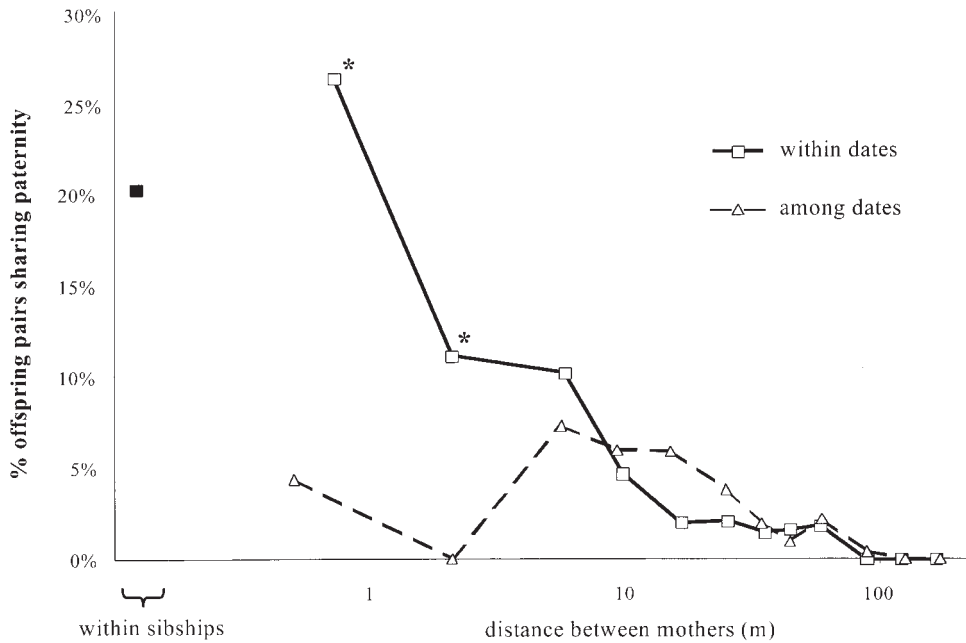
FIGURE 1.—Percentages of offspring pairs sharing paternity within maternal sibship (solid symbols) and among sibships according to the distance (log scale) separating the mothers (open symbols). We distinguish pairwise comparisons including offspring sampled the same date (squares, solid line) or at 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.

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_{ij} 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 Φ_{FT} between sibships shows that differentiation among pollen
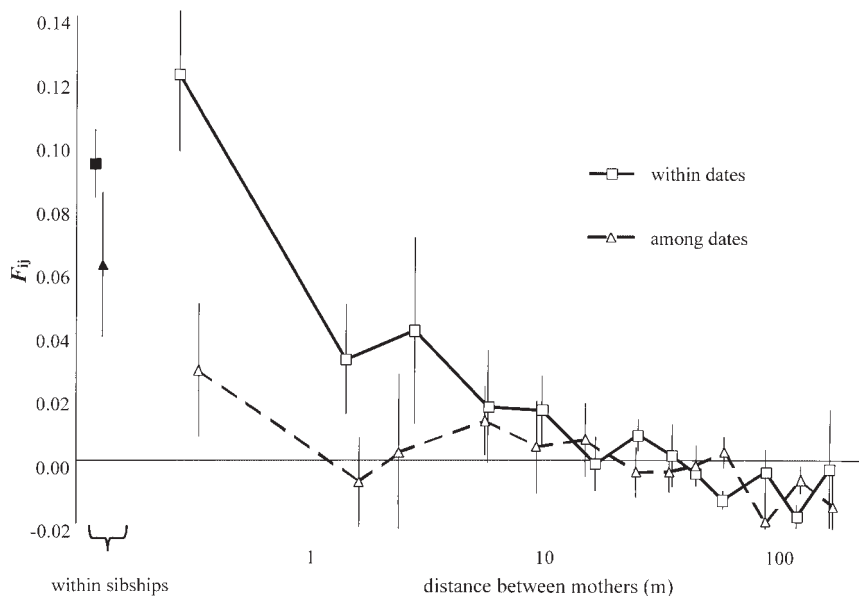

FIGURE 2.—Average kinship coefficient between the paternal genes of offspring pairs within maternal sibship (solid symbols) and among sibships according to the distance (log scale) separating the mothers (open symbols). We distinguish pairwise comparisons including offspring sampled the same date (squares, solid line) or at different dates (triangles, dashed line). Error bars indicate standard errors assessed by jackknifing loci.
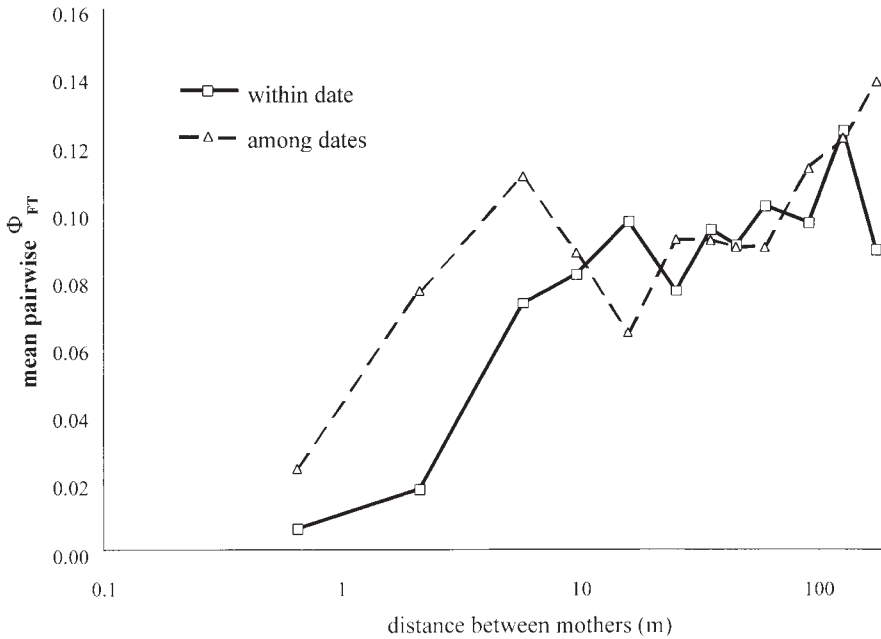
FIGURE 3.—Pairwise differentiation among pollen pools having fertilized different sibships (Φ_{FT}) according to the distance (log scale) between mothers. Average values per distance class are given. We distinguish pairwise comparisons including sibships sampled the same date (squares, solid line) or at different dates (triangles, dashed line).

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 ($\Phi_{FT} = 0.0037$).

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 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_p, $2\Phi_{FT}$, 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_p and overestimated by $2\Phi_{FT}$ (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 $= 7 \times 10^{-4}$), $2\Phi_{FT} =$

$1.184P_{FS} - 0.0063$ ($R^2 = 0.983$; MSE $= 46 \times 10^{-4}$), $r_p = 1.018P_{FS} - 0.0633$ ($R^2 = 0.892$; MSE $= 107 \times 10^{-4}$). Under low actual P_{FS}, r_p performs particularly badly, giving often fairly negative values (Figure 4B). The lower performances of r_p 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 their actual values (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, $2F_{ij}$ somewhat underestimates P_{PS}: $2F_{3m} = 0.853P_{PS(3m)} - 0.0169$ ($R^2 = 0.907$; MSE $= 17 \times 10^{-4}$). When the actual paternal genes were used to assess F_S and F_{3m} (no ambiguities), it did not really improve F_S ($2F_S = 0.991P_{FS} - 0.0148$; $R^2 = 0.989$; MSE $= 9 \times 10^{-4}$) but improved a little F_{3m} ($2F_{3m} = 1.006P_{PS(3m)} - 0.0199$; $R^2 = 0.906$; MSE $= 10 \times 10^{-4}$), suggesting that ambiguous paternal gametic genotypes 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 $= 19 \times 10^{-4}$), $2\Phi_{FT} = 1.356 P_{FS} - 0.0081$ ($R^2 = 0.994$; MSE $= 99 \times 10^{-4}$), $r_p = 1.037P_{FS} - 0.0884$ ($R^2 = 0.936$; MSE $= 100 \times 10^{-4}$). Among progeny arrays, the impact of biparental inbreeding was more significant but the estimator still displayed fairly good statistical properties: $2F_{3m} = 1.191 P_{PS(3m)} - 0.0123$ ($R^2 = 0.940$; 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-

TABLE 2

Average kinship coefficients between the paternal genes of maternal sibs (F_S) and percentage of full-sib pairs (P_{FS}) obtained by simulating mating events under various assumptions

Model	F_S (95% C.I.)	P_{FS} (%)
Model 1		
Random mating	0.0093	1.05
	(−0.0019, 0.0208)	
Model 2		
Limited pollen dispersal	0.0363	6.38
Pollen production constant among adults	(0.0220, 0.0524)	
Pollen production constant through time		
Model 3		
Limited pollen dispersal	0.0431	9.19
Pollen production variable among adults	(0.0271, 0.0598)	
Pollen production constant through time		
Model 4		
Limited pollen dispersal	0.0582	15.09
Pollen production variable among adults	(0.0365, 0.0854)	
Pollen production variable through time		
Model 5		
Limited pollen dispersal	0.0659	17.20
Pollen production variable among adults	(0.0426, 0.0964)	
Pollen production variable through time		
Self-incompatibility system (20 alleles)		

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 $N_{ep} = 5.1$. This value lies in the upper range of other herbaceous species (Table 3), among which the highest N_{ep} 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_{ep}), 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-sib *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 *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 *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 *C. corymbosa* have not shown evidence of pollen limitation (S. LUIJTEN, unpublished results).

Factors determining correlated mating in *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 codispersal 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 *Rutidosis 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-
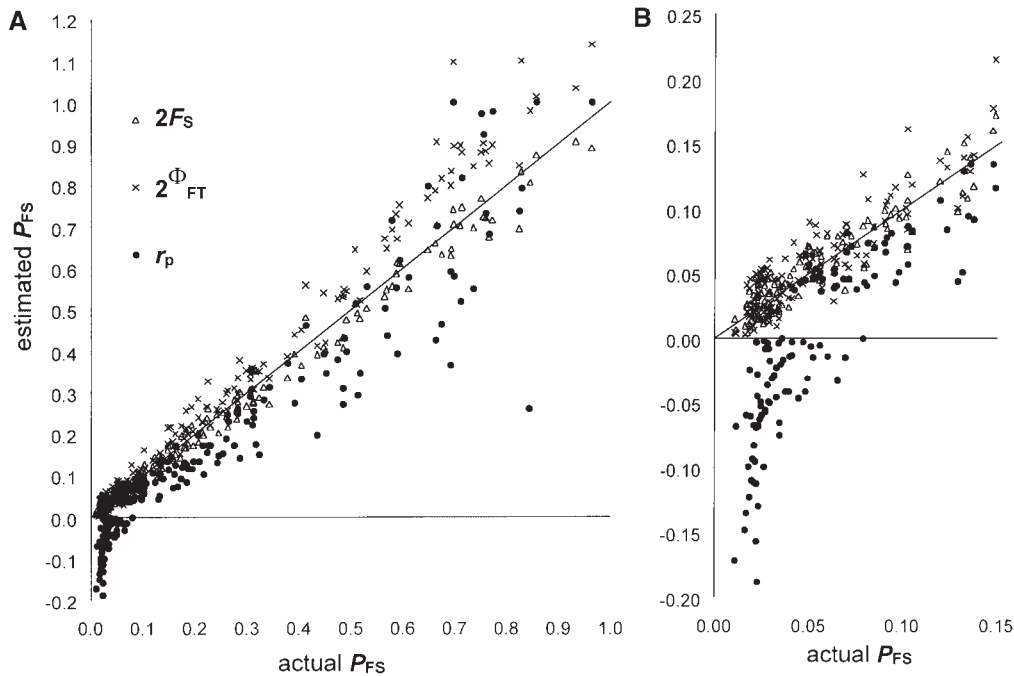
FIGURE 4.—Statistical performances of different estimators (r_p, $2\Phi_{FT}$, and $2F_S$) of the proportion of full-sibs within progeny arrays (P_{FS}). Values were computed on 200 simulated data sets (see text) where parental genotypes were reconstructed to ensure Hardy-Weinberg proportions and no biparental inbreeding. Results are given for (A) all runs and (B) runs with low P_{FS} (≤ 0.15). The diagonal line shows the expected relationship for an ideal estimator.

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 Φ_{FT} values that were very low between progeny arrays sampled the same date (at <2 m, mean Φ_{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 Φ_{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 param-

eterization 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.*

<div align="center">

TABLE 3

Levels of correlated mating reported in natural populations of herbaceous and tree species.
The effective number of sires, N_{ep}, is estimated as $1/r_p$ or $1/(2\Phi_{FT})$

</div>

Species	N_{ep}	Reference
Herbs		
Centaurea corymbosa	5.1 (wC)	
C. solstialis	8.8 (1.2–21) (T)	Sun and Ritland (1998)
Crepis sancta	1.4–3.3 (wC)	Cheptou *et al.* (2001)
Eichhornia paniculata	3.1 (wF); 4.2 (aF)	Morgan and Barrett (1990)
Glycine argyrea	1.2 (wF); 1.7 (aF)	Brown (1989)
Mimulus guttatus	2.5 (wF); 4.7 (aF)	Ritland (1989); Dudash and Ritland (1991)
Rutidosis leptorrhynchoides	1.5–2.7 (T)	Wells and Young (2002)
Yucca filamentosa	1.6–3.1 (T)	Massey and Hamrick (1999)
Trees		
Acacia melanoxylon	1.3 (wF); 2.3 (wI); 30 (T)	Muona *et al.* (1991)
Albizia julibrissin	2.1 (T, 1 yr); 2.9 (T, 4 yr)	Irwin *et al.* (2003)
Caryocar brasiliense	4.9–11.6 (T)	Collevatti *et al.* (2001)
Dinizia excelsa	4.9 (T)	Dick *et al.* (2003)
Dryobalanops aromatica	12.5 (T)	Lee (2000)
Eucalyptus rameliana	3.8 (wF); 11 (aF)	Sampson (1998)
Larix occidentalis	9.1–16.1 (T)	El-Kassaby and Jaquish (1996)
Pinus sylvestris	83–125 (T)	Robledo-Arnuncio *et al.* (2004)
Quercus lobata	3.7 (T)	Sork *et al.* (2002)
Sorbus torminalis	4.9–14.7 (T)	S. Oddou-Muratorio (unpublished results)

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* Scopili) 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, Φ_{FT}, 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 Φ_{FT} 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; Malécot 1967). In the absence of inbreeding, in a diploid species, one expects $r_p = 2\Phi_{FT}$ (Cockerham 1969), and this is indeed what we found with our data. F_S and Φ_{FT} gave very similar values but there is, however,

a difference: Φ_{FT} 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 Φ_{FT} to be slightly higher than F_S, as observed with our data ($\Phi_{FT} = 0.0974$, $F_S = 0.0914$).

To estimate the proportion of full-sibs within progeny arrays in an outcrossed population, r_p, $2\Phi_{FT}$, 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 Φ_{FT} (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_i, can be assessed for a random set of N plants: $d_e = d(\Sigma w_i)^2/(N\Sigma w_i^2)$, where summations hold for $i = 1–N$. 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/m², $d_e = 0.100$ 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 Φ_{FT} (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 $\delta_k = 9.8$ m from the global Φ_{FT}, whereas the best fit to the 121 observed mating events gave $\delta_k = 37$ m ($\delta_k = 91$ m assuming an exponential power kernel; HARDY *et al.* 2004). It must be noted here that δ_k is not an estimate of the mean distance between mates (which was $\delta_e = 21.6$ 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 $(1/3)^2 =$ one-ninth 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.

LITERATURE CITED

AUSTERLITZ, F., and P. E. SMOUSE, 2001 Two-generation analysis of pollen flow across a landscape. II. Relationship between Φ_{ft}, pollen dispersal and interfemale distance. Genetics **157:** 851–857.

AUSTERLITZ, F., and P. E. SMOUSE, 2002 Two-generation analysis of pollen flow across a landscape. IV. Estimating the dispersal parameter. Genetics **161:** 355–363.

AUSTERLITZ, F., C. W. DICK, C. DUTECH, E. K. KLEIN, S. ODDOU-MURATORIO *et al.*, 2004 Using genetic markers to estimate the pollen dispersal curve. Mol. Ecol. **13:** 937–955.

BROWN, A. H. D., J. E. GRANT and R. PULLEN, 1986 Outcrossing and paternity in *Glycine argyrea* by paired fruit analysis. Biol. J. Linn. Soc. **29:** 283–294.

BROYLES, S. B., and R. WYATT, 1991 Effective pollen dispersal in a natural population of *Asclepias exaltata*: the influence of pollinator behavior, genetic similarity, and mating success. Am. Nat. **138:** 1239–1249.

CAMPBELL, D. R., 1998 Multiple paternity in fruits of *Ipomopsis aggregata* (Polemoniaceae). Am. J. Bot. **85:** 1022–1027.

CHARNOV, E., 1982 *The Theory of Sex Allocation.* Princeton University Press, Princeton, NJ.

CHEPTOU, P.-O., J. LEPART and J. ESCARRÉ, 2001 Differential outcrossing rates in dispersing and non-dispersing achenes in the heterocarpic plant *Crepis sancta* (Asteraceae). Evol. Ecol. **15:** 1–13.

COCKERHAM, C. C., 1969 Variance in gene frequencies. Evolution **23:** 72–84.

COLAS, B., I. OLIVIERI and M. RIBA, 1997 *Centaurea corymbosa*, a cliff-dwelling species tottering on the brink of extinction: a demographic and genetic study. Proc. Natl. Acad. Sci. USA **94:** 3471–3476.

COLAS, B., I. OLIVIERI and M. RIBA, 2001 Spatio-temporal variation

of reproductive success and conservation of the narrow-endemic *Centaurea corymbosa* (Asteraceae). Biol. Conserv. **99:** 375–386.

COLLEVATTI, R. G., D. GRATTAPAGLIA and J. D. HAY, 2001 High resolution microsatellite based analysis of the mating system allows the detection of significant biparental inbreeding in *Caryocar brasiliense*, an endangered tropical tree species. Heredity **86:** 60–67.

DEWOODY, J. A., Y. D. DEWOODY, A. C. FIUMERA and J. C. AVISE, 2000 On the number of reproductives contributing to a half-sib progeny array. Genet. Res. **75:** 95–105.

DICK, C. W., G. ETCHELECU and F. AUSTERLITZ, 2003 Pollen dispersal of tropical trees (*Dinizia excelsa*: Fabaceae) by native insects and African honeybees in pristine and fragmented Amazonian rainforest. Mol. Ecol. **12:** 753–764.

DUDASH, M. R., and K. RITLAND, 1991 Multiple paternity and self-fertilization in relation to floral age in *Mimulus guttatus* (Scrophulariaceae). Am. J. Bot. **78:** 1746–1753.

EL-KASSABY, Y. A., and B. JAQUISH, 1996 Population density and mating pattern in western larch. J. Hered. **87:** 438–443.

ELLSTRAND, N. C., 1984 Multiple paternity within the fruits of the wild radish, *Raphnus sativus*. Am. Nat. **123:** 819–828.

FRÉVILLE, H., B. COLAS, J. RONFORT, M. RIBA and I. OLIVIERI, 1998 Predicting endemism from population structure of a widespread species: case study in *Centaurea maculosa* Lam. (Asteraceae). Conserv. Biol. **12:** 1269–1278.

FRÉVILLE, H., E. IMBERT, F. JUSTY, R. VITALIS and I. OLIVIERI, 2000 Isolation and characterization of microsatellites in the endemic species *Centaurea corymbosa* Pourret (Asteraceae) and other related species. Mol. Ecol. **9:** 1671–1672.

FRÉVILLE, H., F. JUSTY and I. OLIVIERI, 2001 Comparative allozyme and microsatellite population structure in a narrow endemic plant species, *Centaurea corymbosa* Pourret (Asteraceae). Mol. Ecol. **10:** 878–889.

FRÉVILLE, H., B. COLAS, M. RIBA, H. CASWELL, A. MIGNOT *et al.*, 2004 Spatial and temporal demographic variability in small populations of the endemic plant species, *Centaurea corymbosa* (Asteraceae). Ecology **85:** 694–703.

GOLD, J. J., and J. S. SHORE, 1995 Multiple paternity in *Asclepias syriaca* using a paired-fruit analysis. Can. J. Bot. **73:** 1212–1216.

GONZÁLEZ-MARTÍNEZ, S. C., S. GERBER, M. T. CERVERA, J. M. MARTINEZ-ZAPATER, R. ALIA *et al.*, 2003 Selfing and sibship structure in a two-cohort stand of maritime pine (*Pinus pinaster* Ait.) using nuclear SSR markers. Ann. For. Sci. **60:** 115–121.

HAMILTON, W. D., 1964 The genetical evolution of social behaviour. J. Theor. Biol. **7:** 1–16.

HARDY, O. J., and X. VEKEMANS, 2002 SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. Mol. Ecol. Notes **2:** 618–620.

HARDY, O. J., S. C. GONZÁLEZ-MARTÍNEZ, H. FRÉVILLE, G. BOQUIEN, A. MIGNOT *et al.*, 2004 Fine-scale genetic structure and gene dispersal in *Centaurea corymbosa* (Asteraceae). I. Pattern of pollen dispersal. J. Evol. Biol. **17:** 795–806.

IRWIN, A. J., J. L. HAMRICK, M. J. W. GODT and P. E. SMOUSE, 2003 A multiyear estimate of the effective pollen donor pool for *Albizia julibrissin*. Heredity **90:** 187–194.

KARRON, J. D., and D. L. MARSHALL, 1990 Fitness consequences of multiple paternity in wild radish, *Raphanus sativus*. Evolution **44:** 260–268.

KARRON, J. D., and D. L. MARSHALL, 1993 Effects of environmental variation on fitness of singly and multiply sired progenies of *Raphanus sativus* (Brassicaceae). Am. J. Bot. **80:** 1407–1412.

LARSON, B. M. H., and S. C. H. BARRETT, 2000 A comparative analysis of pollen limitation in flowering plants. Biol. J. Linn. Soc. **69:** 503–520.

LEE, S. L., 2000 Mating system parameters of *Dryobalanops aromatica* Gaertn. f. (Dipterocarpaceae) in three different forest types and seed orchards. Heredity **85:** 338–345.

LOISELLE, B. A., V. L. SORK, J. NASON and C. GRAHAM, 1995 Spatial genetic structure of a tropical understory shrub, *Psychotria officinalis* (Rubiaceae). Am. J. Bot. **82:** 1420–1425.

LYNCH, M., and K. RITLAND, 1999 Estimation of pairwise relatedness with molecular markers. Genetics **152:** 1753–1766.

MALÉCOT, G., 1967 Identical loci and relationship. Proc. Fifth Berk. Symp. Math. Stat. Prob. **4:** 317–332.

MARSHALL, D. L., and N. C. ELLSTRAND, 1985 Proximal causes of

multiple paternity in wild radish, *Raphanus sativus*. Am. Nat. **126:** 596–605.

MASSEY, L. K., and J. L. HAMRICK, 1999 Breeding structure of a *Yucca filamentosa* (Agavaceae) population. Evolution **53:** 1293–1298.

MORGAN, M. T., and S. C. H. BARRETT, 1990 Outcrossing rates and correlated mating within a population of *Eichhornia paniculata* (Pontederiaceae). Heredity **64:** 271–280.

MUONA, O., G. F. MORAN and J. C. BELL, 1991 Hierarchical patterns of correlated mating in *Acacia melanoxylon*. Genetics **127:** 619–626.

NASON, J. D., E. A. HERRE and J. L. HAMRICK, 1998 The breeding structure of a tropical keystone plant resource. Nature **391:** 685–687.

NIKKANEN, T., T. ARONEN, H. HÄGGMAN and M. VENÄLÄINEN, 2000 Variation in pollen viability among *Picea abies* genotypes—potential for unequal paternal success. Theor. Appl. Genet. **101:** 511–518.

PETIT, C., H. FRÉVILLE, A. MIGNOT, B. COLAS, M. RIBA *et al.*, 2001 Gene flow and local adaptation in two endemic plant species. Biol. Conserv. **100:** 21–34.

RICHARDS, A. J., 1997 *Plant Breeding System*. Chapman & Hall, London.

RITLAND, K., 1988 The genetic mating structure of subdivided populations. II. Correlated mating models. Theor. Popul. Biol. **34:** 320–346.

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.

ROBLEDO-ARNUNCIO, J. J., P. E. SMOUSE, L. GIL and R. ALIA, 2004 Pollen movement under alternative silvicultural practices in native populations of Scots pine (*Pinus sylvestris* L) in Central Spain. For. Ecol. Manage. **197:** 245–255.

ROUSSET, F., and S. BILLIARD, 2000 A theoretical basis for measures of kin selection in subdivided populations: finite populations and localized dispersal. J. Evol. Biol. **13:** 814–825.

SAMPSON, J. F., 1998 Multiple paternity in *Eucalyptus rameliana* (Myrtaceae). Heredity **81:** 349–355.

SCHIERUP, M. H., X. VEKEMANS and F. B. CHRISTIANSEN, 1997 Evolutionary dynamics of sporophytic self-incompatibility alleles in plants. Genetics **147:** 835–846.

SCHMITT, J., and D. W. EHRHARDT, 1987 A test of sib-competition hypothesis for outcrossing advantage in *Impatiens capensis*. Evolution **41:** 579–590.

SCHOEN, D. J., 1985 Correlation between classes of mating events in two experimental plant populations. Heredity **55:** 381–385.

SCHUSTER, W. S. F., and J. B. MITTON, 1991 Relatedness within clusters of bird-dispersed pine and the potential for kin interactions. Heredity **67:** 41–48.

SMOUSE, P. E., R. J. DYER, R. D. WESTFALL and V. L. SORK, 2001 Two-generation analysis of pollen flow across a landscape. I. Male gamete heterogeneity among females. Evolution **55:** 260–271.

SORK, V. L., F. W. DAVIS, P. E. SMOUSE, V. J. APSIT, R. J. DYER *et al.*, 2002 Pollen movement in declining populations of California Valley oak, *Quercus lobata*: Where have all the fathers gone? Mol. Ecol. **11:** 1657–1668.

SUN, M., and K. RITLAND, 1998 Mating system of yellow starthistle (*Centaurea solstitialis*), a successful colonizer in North America. Heredity **80:** 225–232.

WELLS, G. P., and A. G. YOUNG, 2002 Effect of seed dispersal on genetic structure in populations of *Rutidosis leptorrhychoides* with different levels of correlated paternity. Genet. Res. **79:** 219–226.

WRIGHT, S., 1922 Coefficients of inbreeding and relationship. Am. Nat. **56:** 330–338.

YOUNG, J. P., 1981 Sib competition can favor sex in two ways. J. Theor. Biol. **88:** 755–756.