Patterns of Pollen Flow in a Dense Population of the Insect-Pollinated Canopy Tree Species *Castanopsis sieboldii*

ATSUSHI NAKANISHI, HIROSHI YOSHIMARU, NOBUHIRO TOMARU, MASAHIRO MIURA, TOHRU MANABE, AND SHIN-ICHI YAMAMOTO

Nagoya 465-0042, Japan (Nakanishi); Tama Forest Science Garden, Forestry and Forest Products Research Institute, Tokyo 193-0843, Japan (Yoshimaru); Graduate School of Bioagicultural Sciences, Nagoya University, Nagoya 464-8601, Japan (Tomaru); Okayama University, Okayama 700-8530, Japan (Yamamoto); the Forest Tree Breeding Center, Forestry and Forest Products Research Institute, Hitachi, Ibaraki 319-1301, Japan (Miura); and Kitakyushu Museum of Natural History and Human History, Kitakyushu 805-0071, Japan (Manabe).

Address correspondence to Atsushi Nakanishi at the address above, or e-mail: atsunakanishi@yahoo.co.jp.

Abstract

Insect pollinations of tree species with high-density populations have rarely been studied. Since the density of adults can affect effective pollen dispersal, short-distance pollination, even by insects, may frequently occur in high-density populations. To test this prediction, we investigated pollination patterns in a high-density population of the insect-pollinated canopy tree species *Castanopsis sieboldii* by paternity analysis using genotypes at 8 microsatellite loci of 145 adult trees and 439 seeds from 11 seed parents in a 4-ha plot. We then explored their genetic effects on the population by calculating other population genetics parameters. Although *C. sieboldii* has high potential for long-distance dispersal of pollen (as indicated by a fat-tailed dispersal kernel), the cumulative pollination at the local scale was spatially limited and strongly dependent on the distance between parents due to the high density of adults. Genetic diversity estimates for pollen pools accepted by each seed parent converged on a maximum as the effective number of pollen parents increased. The genetic diversity of pollen pool bulked over all the seed parents from inside the plot did not differ from that of the total pollen pools. Therefore, although pollen flow from distant pollen parents may help to maintain the genetic diversity of offspring, pollen parents neighboring seed parents may be the main contributors to the genetic diversity of the offspring at the seed stage.

Key words: dispersal kernel, gene flow, genetic diversity, genetic structure, paternity analysis, pollen dispersal

Gene movement within and between populations determines their spatial and temporal genetic variation and structure, which in turn influences their evolutionary potential. Populations constitute the actual evolving units of a species (Hartl and Clark 1997). Therefore, to understand the evolutionary dynamics of a species, it is important to study gene movement within and between its populations. In seed plants, the historical levels of pollen flows are generally at least an order of magnitude larger than the levels of seed flow at the range-wide scale (Petit et al. 2005). In forest tree species as well as other plant species, pollen dispersal may thus be the most important component of the gene flow that maintains genetic diversity within their populations. Gene movement via seeds and pollen in forest tree species has been directly studied by parentage analysis using microsatellite markers, which are suited to such studies as they have very high levels of polymorphism and codominant expression (Chase et al. 1996; Dow and Ashley 1996, 1998; Streiff et al. 1999; Isagi et al. 2000; Bacles et al. 2006).

Forest tree species that naturally occur in low densities are also suitable for parentage analysis because all of the potential parents in populations of such species can be relatively easily found and genotyped (Hardy 2009; Kamm et al. 2009). Consequently, most studies that have used parentage analysis have focused on low-density species (Chase et al. 1996; Isagi et al. 2000, 2007; Sezen et al. 2005; Kamm et al. 2009). In contrast, there has been very little
Materials and Methods

Study Species

Castanopsis sieboldii (Makino) Hatus. (synonym, C. cuspidata var. sieboldii [Makino] Nakai), an evergreen broad-leaved tree of the Fagaceae, is distributed in Japan and Korea (Horikawa 1972). The species generally reaches heights up to 15–25 m, dominantly forms a canopy layer in warm-temperate evergreen broad-leaved forests in Japan, and blooms from May to June. According to observations by Yumoto (1987), it usually flowers only in the canopy (seldom in the understory) and individual trees produce large number of flowers that project outward simultaneously. Thus, the flowering trees are conspicuous, even from a long distance. Yumoto (1987) also noted that individuals in the same region tightly synchronize their flowering times, and it is pollinated by insects, including bees, wasps, flies and beetles. The seeds are dispersed primarily by gravity and sometimes by hoarding behavior of rodents (Shimada 2001) and birds (Higuchi 1977).

Study Site and Field Methods

The study site was in the Tatera Forest Reserve, on the South Island of Tsushima, which is located between the Japanese Archipelago and the Korean Peninsula. The reserve, which is protected as a National Natural Monument, has an area of approximately 100 ha and is situated on the north-facing slope of Mt Tatera. There has been no human interference in the reserve for several centuries, hence the vegetation within it is a well-developed primary evergreen broad-leaved forest (Itow 1991).

A 4-ha permanent plot (200 × 200 m) was established in 1990 (Manabe et al. 2000), from 150 to 190 m above sea level within a forest in the lower part of the reserve, which is identified phytosociologically as Distylium–Quercetum salicinae (Itow et al. 1993). The forest has a general canopy 20–30 m tall, dominated by evergreen broad-leaved trees with a diameter at breast height (d.b.h.) exceeding 1 m. These trees include Castanopsis sieboldii, Distylium racemosum, and Quercus salicina, although the forest in the upper part of the reserve (to the south of the lower part) is scrub dominated by Quercus acuta growing to heights of just 6–7 m, due to strong winds and rocky conditions (Manabe et al. 2000). The forest in the lower part of the reserve covers approximately 40 ha and is adjacent to a coniferous plantation to the east and secondary forests to the north and west, respectively. The plot was located approximately 600, 250, 100, and 100 m away from the edges of the adjacent forests to the south, east, west, and north, respectively. Tree censuses were performed in 1990, 1992, 1997, 2002, and 2007 of all stems ≥5 cm in d.b.h. The plot contained a total of 45 species and 4570 living stems ≥5 cm d.b.h., with a total basal area of 63.9 m²/ha. Castanopsis sieboldii was the dominant species in the plot, having the largest basal area of all species (24.9 m²/ha), largest average d.b.h. (75.9 cm; max. 209.2 cm), and fifth highest density (41.0 stems/ha) (Manabe et al. 2000). One hundred and forty-six individual C. sieboldii trees were found with living stems ≥5 cm d.b.h. in 1997, which were defined as adult trees in this study. This was an operational definition because (as mentioned) C. sieboldii flowers are usually found only in the canopy, and the minimum d.b.h. of its canopy trees in the plot was 26.9 cm in 1997. In 1996 and 1997, leaf samples of 145 of the adult trees (all except one, which died during the course of the study) were collected, and in 2000, seeds that had fallen under the canopies of 11 adult C. sieboldii trees (Figure 1) were collected within 3 m
from the trunk of each seed parent to avoid collecting seeds from other seed parents due to overlapping seed shadows.

**DNA Extraction and Microsatellite Genotyping**

Genomic DNA was extracted from leaves and seeds using the hexadecyltrimethylammonium bromide (CTAB) method (Murray and Thompson 1980) with minor modifications. Eight polymorphic microsatellite loci were selected for genotyping *Castanopsis sieboldii* adults and seeds: *Ccu22F30, Ccu16H13, Ccu5F43, Ccu17F15, Ccu9T20, and Ccu33H25*, which were developed for *C. sieboldii* (Ueno et al. 2000); and *QpZAG16* and *QpZAG9*, which were originally developed for *Quercus petraea* (Steinkellner et al. 1997). Genotyping at the microsatellite loci was conducted using capillary electrophoresis by a 3100 Genetic Analyzer and GeneScan software (Applied Biosystems).

**Statistical Analyses**

To estimate the genetic diversity of *C. sieboldii* adults, we used the following standard population genetic parameters, calculated by GENEPOP ver. 3.4 (Raymond and Rousset 1995): the number of alleles (*A*); observed heterozygosity (*H*o); gene diversity (*H*_0); and inbreeding coefficient (*F*_is) for each locus and over all loci. Deviations from Hardy-Weinberg equilibrium at each locus were tested using the Markov-chain method.

We calculated *F*_is (coancestry; Loiselle et al. 1995) kinship coefficients in order to evaluate the spatial genetic structure of all the adults in the 4-ha plot. The average *F*_is value was calculated for each of 10 continuous distance classes at 20-m intervals from 0–20 to 180–200 m. The significance of the average *F*_is values was then assessed by permutation tests (with 1000 permutations), in which spatial coordinates were permuted randomly among adults. These calculations were performed by SPAGeDi ver. 1.2 (Hardy and Vekemans 2002). Paternity analysis was performed using the maximum-likelihood method based on the multilocus genotypes (Meagher and Thompson 1986; Gerber et al. 2000) of the adults and seeds, using FaMoz (Gerber et al. 2003). We assumed that each of the 11 adult trees whose crowns covered the seed-sampling site was a putative seed parent of the seeds collected in that area. The most likely pollen parents were determined using “log of the odds ratios” (LOD) scores, based on reference allele frequencies, which were calculated for the adult population within the 4-ha plot. The test thresholds of the LOD scores for rejecting a candidate parent as a true pollen parent (TF) were determined using a simulation procedure (for further details, see Gerber et al. 2000, 2003). At the same time as the simulation procedure, we also calculated the following probabilities for the TF: correct classification (i.e., the most likely pollen parent being the true pollen parent); type I errors (erroneous rejection of the null hypothesis that the pollen parents are present inside the plot); and type II errors (erroneous acceptance of the null hypothesis). We incorporated an error rate of 1.0% into genotyping in the simulations. Morrissey and Wilson (2005) postulated that the most powerful approach for handling genotype errors in parentage analyses might be to apply likelihood equations with error rates set to values substantially lower than the rates at which genotype errors occur. To test this hypothesis, we applied 2 error rates (0.1% and 1.0%) for the LOD calculations in the simulations. If a candidate parent had an LOD score exceeding TF, it was considered a true potential parent. If a seed had no true potential parent, it was regarded as having no pollen parent within the plot. If a seed had only one true potential parent, the potential parent was regarded as its true pollen parent. If a seed had multiple true potential parents, it was regarded as having a pollen parent within the plot, but its paternity could not be assigned.

Based on the results of the paternity analysis, the pollinations within the plot were analyzed using 4 methods. First, all matings were divided into classes of 20 m between-parent distance intervals, and the frequency of matings relative to all considered matings was calculated for each distance class for each seed parent (Streiff et al. 1999). The frequency distribution was then plotted as a function of the distance between the parents. After data were averaged over all seed parents, an exponential function was fitted to the data and a regression coefficient (*R*_2) was determined. Second, to test whether the average pollination distance within the plot deviated significantly from that expected under random mating, it was compared with the distribution of the average values generated by a randomization procedure, repeated 1000 times, as follows. If the *i*th seed parent had *n*_i seeds whose pollen parents could be assigned by paternity analysis (so the total number of seeds, *N*, was the sum of *n*_i over all 11 seed parents), a sample from the seed parent was generated with *n*_i randomly chosen adult trees from the 145 potential pollen parents. The distances...
between the seed parent and the \( n_i \) extracted adult trees were then estimated, and an average value over all 11 seed parents (average values for \( N \) distances) was calculated. Third, to test whether the average \( F_{st} \) between pairs of mating parents, which reflects the degree of inbreeding, within the plot differed significantly from that expected under random mating, we compared the actual data with the distribution of the average values generated by the same randomization procedure used to test the average distance of pollen flow within the plot. The analysis was conducted for both cases including and excluding self-pollination. Fourth, we estimated the effective number of pollen parents (\( N_{ep} \)) within the plot (Smouse and Robledo-Arnuncio 2005). In this procedure, \( N_{cp} \) was calculated by using the inverse of the probabilities for paternal identity (PPI, defined as the probability that two different offspring, seeds, were sired by the same pollen parent). We estimated \( N_{cp} \) for each seed parent, for an average seed parent and the global \( N_{cp} \) for all 11 seed parents. These values were calculated using the estimated values of PPI for each seed parent (\( r_{gs} \)), for an average seed parent (\( \bar{r}_0 \)), and the global PPI (\( R_0 \)), respectively, as follows:

\[
R_0 = \frac{\binom{2}{2} + \binom{X_2}{2} + \cdots + \binom{X_K}{2}}{\binom{N}{2}},
\]

\[
r_{gs} = \frac{\binom{X_{g1}}{2} + \binom{X_{g2}}{2} + \cdots + \binom{X_{gK}}{2}}{\binom{n_g}{2}},
\]

\[
\bar{r}_0 = \frac{\sum_{g=1}^{G} \frac{\binom{X_g}{2}}{\binom{n_g}{2}}}{\sum_{g=1}^{G} \binom{n_g}{2}},
\]

where \( X_{gk} \) is the number of offspring from the \( g \)th seed parent for whom the \( k \)th adult is the designated pollen parent, \( K \) is the total number of pollen parents, \( n_g \) is the number of offspring from the \( g \)th seed parent, \( X_g \) is the total number of offspring from the \( g \)th pollen parent over all seed parents, and \( N \) is the number of offspring over all seed parents. These calculations were conducted for the seeds for which the pollen parents were inside the plot and assigned unambiguously.

We calculated inbreeding coefficient (\( F_{is} \)) for adults and seeds based on the reference allele frequencies of the adult population by SPAGeDi ver. 1.2 (Hardy and Vekemans 2002) and compared its values between adults and seeds by Mann–Whitney U-test to examine inbreeding depression. The inbreeding coefficient was computed as the kinship coefficient between genes at a locus within an individual (Hardy and Vekemans 2002), and its average value for adults should be equal to the average value of \( F_{is} \) calculated by comparing the frequency of heterozygotes in the population with the frequency expected under random mating (we calculated the value by using Genepop ver. 3.4, Raymond and Rousset 1995, as mentioned above).

The pollen dispersal kernel was examined using the neighborhood model (Adams and Birkes 1991) based on the multilocus genotypes of adults and seeds, by NM+ software (Chybicki and Burczyk 2010). We applied the following exponential–power function characterized by the parameters \( a \) and \( b \) (Clark 1998):

\[
P(r) = \frac{b}{2\pi a\Gamma(\frac{b}{a})}\exp\{-\left(\frac{r}{a}\right)^b\},
\]

where \( r \) is pollination distance, \( \Gamma \) is the gamma function, \( a \) is a scale parameter, and \( b \) is the shape parameter affecting the “fatness” of the tail of the dispersal distribution (Austerlitz et al. 2004). When \( b = 1 \), the above equation degenerates to the exponential distribution. When \( b < 1 \) the dispersal kernel is fat-tailed (Clark 1998), that is, the long-range decay of probability is slower than for the exponential distribution. Conversely, when \( b > 1 \), the dispersal is thin tailed, with a rapid decrease of the dispersal function, implying few long-distance dispersal events (Austerlitz et al. 2004). Treating all the adults within the plot as the neighborhood, we estimated the parameters \( a \) and \( b \), pollen immigration rate, average distance of pollination, and self-pollination rate.

To quantify genetic heterogeneity in the pollen pools among the seed parents, AMOVA (Excoffier et al. 1992) of pollen haplotypes (TwoGen) was performed following the method described by Smouse et al. (2001). AMOVA allowed us to calculate both within and among variance components for the pollen pools and “\( \Phi_{PT} \)” values (analogs of Wright’s \( F_{ST} \) values). The significance of the \( \Phi_{PT} \) values was tested by 1000 randomizations (Excoffier et al. 1992). Pairwise \( \Phi_{PT} \) values were calculated and used as a measure of the genetic distance between pollen pools of seed parents. These calculations and tests were conducted by GenAlEx ver. 6 (Peakall and Smouse 2006). To analyze the spatial genetic structure of pollen pools, the relationship between the pairwise \( \Phi_{PT} \) and spatial distance between the seed parents was examined by Spearman’s rank correlation analysis.

Finally, we quantified the genetic diversity of pollen pools for each seed parent. The paternal haplotypes were estimated by the methods described in Hardy et al. (2004). The paternal alleles of each seed were identified by comparing its genotype with that of its seed parent, then the paternal haplotypes were converted into diploid homozygous genotypes with the alleles. In ambiguous cases, in which both the seed and its seed parent were heterozygotes with the same alleles, the paternal haplotypes were converted into the corresponding homozygous genotypes. Then, average \( H_i \) and allelic richness (\( R \); El Mousadik and Petit 1996) were calculated over all loci for the pollen pools from inside and outside the plot and for the total pollen pool using the converted genotypes for each seed parent. The two parameters were also calculated for the converted genotype data pooled over all the seed parents. In this study, \( R \) was calculated as the expected number of different alleles in a sample of 6 gene copies; the smallest sample size for seeds (3) from a given seed parent with pollen parents from inside and outside the plot. We tested
the significance of the relationships between $N_{cp}$ and the 2 genetic diversity parameters ($H_E$ and $R$) for the pollen pool of each seed parent by examining the distributions of $H_E$ and $R$ as linear functions of inverse of $N_{cp}$ using the data for seeds with unambiguously assigned pollen parents.

**Results**

**Genetic Diversity and Structure in Adult Trees**

The 145 adult trees were genotyped with the 8 microsatellite loci. The 8 loci were highly polymorphic in adults, with 10–40 alleles per locus and a mean $H_E$ value of 0.856 (Table 1). $F_{IS}$ values for each locus ranged from −0.081 to 0.073, and the $F_{IS}$ value across all loci was 0.014. Deviations from Hardy–Weinberg equilibrium were not significant at any of the loci after Bonferroni correction. In the correlogram of Hardy–Weinberg equilibrium were not significant at any of the sampled under its canopy may have null alleles at the parents. One seed parent (SP3242) and 17 of 48 seeds and were compared with those of their putative seed parents. The 486 seeds were genotyped with the 8 microsatellite loci, Paternity and Pollen Flow

...unambiguously assigned pollen parents, 5 were produced by the following analyses, a dummy maternal allele was added for this seed parent and the 17 seeds at this locus. In addition to these 17 seeds, the haplotypes of 47 seeds under the canopies of eight putative seed parents did not match the genotypes of their putative seed parents; either the seeds or putative seed parents were heterozygous at the mismatched locus in all cases. Therefore, those 47 seeds may have originated from other adult trees. The haplotypes of the remaining 422 seeds matched the genotypes of their putative seed parents. Paternity analysis was conducted using data for the 439 seeds with haplotypes matching those of their respective putative seed parents.

The total exclusion probability for the first and second parent, calculated from the allele frequencies at the 8 loci for the 145 adult trees, was 0.999544 and 0.999989, respectively. For all loci, frequencies of null alleles were lower than 0.05. From the results of the simulation procedures with the error rates of 0.1% and 1.0% in the LOD calculation, the calculated TF values were 7.21 and 4.67, respectively. The estimated probability of correct paternity classification with the error rates of 0.1% and 1.0% was 99.1% and 95.0%, respectively. The estimated probabilities of type I and type II errors were <0.02 and <0.02 for the TF value of 7.21 and <0.11 and <0.12 for the TF value of 4.67, respectively. These results indicate that use of substantially lower error rates than those at which genotype errors occur in the LOD calculations increased the accuracy of the paternity analysis, as found by Morrissey and Wilson (2005). Therefore, we used the error rate of 0.1% in the LOD score calculations and the TF value of 7.21 in further analyses. Of the 439 seeds, 154 (35.1%) had no true potential pollen parent in the plot (Table 2). The other 285 (64.9%) seeds had at least one true potential pollen parent in the plot and 247 of those only had one true potential pollen parent, which could thus be defined as the true pollen parent (with unambiguous assignment), while 38 had multiple true potential pollen parents (ambiguous assignment). Of the 247 seeds with unambiguously assigned pollen parents, 5 were produced by

**Table 1** Genetic diversity at the 8 polymorphic microsatellite loci in adult *Castanopsis sieboldii* individuals in the 4-ha plot

<table>
<thead>
<tr>
<th>Locus</th>
<th>$A$</th>
<th>$H_O$</th>
<th>$H_E$</th>
<th>$F_{IS}$</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>CusF45</em></td>
<td>30</td>
<td>0.938</td>
<td>0.942</td>
<td>0.005</td>
</tr>
<tr>
<td><em>Cus17F15</em></td>
<td>40</td>
<td>0.938</td>
<td>0.949</td>
<td>0.012</td>
</tr>
<tr>
<td><em>Cus22F30</em></td>
<td>21</td>
<td>0.786</td>
<td>0.848</td>
<td>0.073</td>
</tr>
<tr>
<td><em>Cus16H15</em></td>
<td>10</td>
<td>0.738</td>
<td>0.734</td>
<td>−0.005</td>
</tr>
<tr>
<td><em>Cus33H25</em></td>
<td>16</td>
<td>0.883</td>
<td>0.878</td>
<td>−0.005</td>
</tr>
<tr>
<td><em>Cus2T20</em></td>
<td>21</td>
<td>0.848</td>
<td>0.889</td>
<td>0.046</td>
</tr>
<tr>
<td><em>QpAG9</em></td>
<td>15</td>
<td>0.779</td>
<td>0.826</td>
<td>0.057</td>
</tr>
<tr>
<td><em>QpAG16</em></td>
<td>11</td>
<td>0.848</td>
<td>0.785</td>
<td>−0.081</td>
</tr>
<tr>
<td>Mean</td>
<td>20.5</td>
<td>0.845</td>
<td>0.856</td>
<td>0.014</td>
</tr>
<tr>
<td>Standard error</td>
<td>3.6</td>
<td>0.026</td>
<td>0.026</td>
<td></td>
</tr>
</tbody>
</table>

$A$, number of alleles detected; $H_O$, observed heterozygosity; $H_E$, gene diversity; $F_{IS}$, inbreeding coefficient.

**Figure 2.** Correlogram of the average $F_{ij}$ values for *Castanopsis sieboldii* adults. Distance classes were defined at continuous 20-m intervals from 0–20 to 180–200 m. The dashed lines represent the 95% (two-tailed) confidence intervals for the average $F_{ij}$ distribution calculated by 1000 permutations of spatial distances among pairs of adults.
self-pollination and thus the rate of self-pollination was estimated to be 1.2% (the number of self-pollinated seeds [5] divided by the total number of seeds analyzed [439]) excluding those with multiple potential parents [38]). The analyses of the 247 seeds indicated that they originated from the seed parents mating, on average, with 11.8 ± 3.3 (mean ± standard deviation) pollen parents and that 65 of the 145 potential parents within the plot were pollen parents of at least one seed, whereas the other 80 potential pollen parents made no contribution to the seeds. The 65 pollen parents all had a d.b.h. ≥36.2 cm and reached the canopy layer.

The following results were derived from the analyses of the 247 seeds with unambiguously assigned pollen parents. The average pollination distance within the plot was 35.1 ± 28.4 m. A negative exponential function of the distance significantly explained the averages of the frequencies of matings over all seed parents at each distance class ($R^2 = 0.972$, $P < 0.001$; Figure 3). The actual average distance (35.1 m) of pollination within the plot was significantly shorter than that (81.7 ± 2.3 m) generated by the permutation procedures ($R = 0.863 ± 0.026$ and $4.4 ± 0.2$ for inside the plot, and the total pollen pools were significantly lower than those for the total pollen pools from inside and outside the plot, and the total pollen pools were 3.9 ± 0.1, 4.0 ± 0.2 and 4.2 ± 0.1, respectively. The average R values for the 11 seed parents for pollen pools from inside the plot were significantly lower than those for the total pollen pool ($t$-test for paired comparison, $P < 0.01$). However, when pollen pools were pooled over all 11 seed parents, the average values of both $H_E$ and $R$ for the pollen pools from inside the plot were not significantly different from those for the total pollen pool (0.859 ± 0.024 and 4.4 ± 0.2 vs. 0.863 ± 0.026 and 4.4 ± 0.2 for $H_E$ and $R$, respectively; $t$-test for paired comparison). Linear functions of inverse of $N_{cp}$ for each seed parent significantly explained the $H_E$ and $R$ values for the accepted pollen pool for each seed parent ($R^2 = 0.460$, $P < 0.05$ and $R^2 = 0.441$, $P < 0.05$, respectively; Figure 4). Both the $H_E$ and $R$ values increased as $N_{cp}$ increased, but the values seemed to converge at a maximum.

### Table 2  Paternity of seeds from 11 Seed parents of Castanopsis sieboldii

<table>
<thead>
<tr>
<th>Seed parent</th>
<th>N</th>
<th>Total</th>
<th>One candidate</th>
<th>Multiple candidates</th>
<th>No candidate</th>
</tr>
</thead>
<tbody>
<tr>
<td>SP1613</td>
<td>21</td>
<td>18</td>
<td>(0.857)</td>
<td>13</td>
<td>5</td>
</tr>
<tr>
<td>SP1628</td>
<td>47</td>
<td>36</td>
<td>(0.766)</td>
<td>29</td>
<td>7</td>
</tr>
<tr>
<td>SP2254</td>
<td>45</td>
<td>22</td>
<td>(0.489)</td>
<td>19</td>
<td>3</td>
</tr>
<tr>
<td>SP2308</td>
<td>46</td>
<td>30</td>
<td>(0.652)</td>
<td>26</td>
<td>4</td>
</tr>
<tr>
<td>SP2466</td>
<td>29</td>
<td>14</td>
<td>(0.483)</td>
<td>12</td>
<td>2</td>
</tr>
<tr>
<td>SP2552</td>
<td>40</td>
<td>25</td>
<td>(0.625)</td>
<td>22</td>
<td>3</td>
</tr>
<tr>
<td>SP2982</td>
<td>48</td>
<td>38</td>
<td>(0.792)</td>
<td>24</td>
<td>3</td>
</tr>
<tr>
<td>SP3037</td>
<td>43</td>
<td>27</td>
<td>(0.628)</td>
<td>24</td>
<td>3</td>
</tr>
<tr>
<td>SP3050</td>
<td>44</td>
<td>26</td>
<td>(0.591)</td>
<td>26</td>
<td>0</td>
</tr>
<tr>
<td>SP3063</td>
<td>34</td>
<td>22</td>
<td>(0.647)</td>
<td>22</td>
<td>0</td>
</tr>
<tr>
<td>SP3242</td>
<td>42</td>
<td>27</td>
<td>(0.643)</td>
<td>21</td>
<td>6</td>
</tr>
<tr>
<td>Total</td>
<td>439</td>
<td>285</td>
<td>(0.649)</td>
<td>247</td>
<td>154 (0.351)</td>
</tr>
</tbody>
</table>

$N$, the total number of seeds analyzed. One, multiple, and no candidates indicate the number of seeds having one, multiple, or no candidate pollen parents, respectively.

**Figure 3.** Distributions of the frequencies of matings over all the matings at each distance class for each seed parent (open circles) and of the averages over all seed parents (filled circles) as functions of the distance between the parents. The frequency of matings was calculated as the number of matings at each distance class divided by the total number of matings for the mated seed parent. The averages as a function of the distance were fitted to an exponential function.
shorter than that expected under random mating. In other studies of insect-pollinated tree species, high proportions of long-distance pollen flow have been identified and the cumulative frequency of pollination has been found to depend less strongly on the distances between parents (Chase et al. 1996; Isagi et al. 2000; Isagi et al. 2007; Kamm et al. 2009). In this study, the proportions of outcrossing occurring at the distances within 50 and 100 m relative to all the outcrossing were at least 41.7% and 54.1%, respectively, which were higher than those in the other studies (Chase et al. 1996; Isagi et al. 2000; Isagi et al. 2007; Kamm et al. 2009). Although 35.1% of pollen flow came from outside the plot, the pollination occurred at short distance with a high frequency, suggesting that the pollen dispersal at the local scale is spatially limited in this species. However, in contrast to the results of the paternity analysis, the shape parameter of the dispersal kernel we calculated even based on the pollinations within the plot (0.537) represents fat-tailed distribution, indicating that *C. sieboldii* has high potential for long-distance dispersal, like other insect-pollinated tree species examined in previous studies (Austerlitz et al. 2004, Oddou-Muratorio et al. 2005). A dispersal kernel of pollen dispersal represents the probability density that a pollen lands at a given position away from the source individual, and reflects, therefore, the pattern of primary dispersal (Hardy 2009). However, effective dispersal of pollen depends on both the kernel and the spatial arrangement of suitable sites (Hardy 2009) (in our case, the spatial positions of seed parents capable of accepting pollen). Hardy (2009) suggested that population density can affect dispersal in two ways. First, birds or insects may forage more locally in search of nectar (or pollen, for species such as *C. sieboldii* that do not produce nectar) in high-density populations. Second, pollen dispersal is “effective” only if pollen reaches a conspecific individual. Therefore, the density and spatial distribution of conspecifics also condition effective pollen dispersal, and mean effective pollen dispersal is expected to be lower under high density and/or when individuals are aggregated (Hardy 2009). The abovementioned strong dependence of the cumulative pollination frequency on distance and the resulting short average distance of pollination at local scale we found may be due to the high density of reproductive *C. sieboldii* trees (>30 individuals ha⁻¹). Although the pollen-dispersal kernel showed the potential for long-distance dispersal, the probability of pollination for adults at the local scale rapidly decreased as the distance from seed parents increased. For example, the probability of an adult 30 m away from a seed parent pollinating that seed parent was less than 5% of the probability of an adult 0 m away from the seed parent (i.e., immediately adjacent to the seed parent) pollinating it at the dispersal kernel. Therefore, the presence of adults near seed parents strongly affected the average distance of pollination at the local scale. Distance dependence of effective pollen dispersal has been reported in other studies of relatively high-density species that rely on both wind and insect pollination (Streiff et al. 1999; Setsuko et al. 2007). The pollen dispersal in *C. sieboldii* may potentially occur over long distance, as indicated by the fat-tailed dispersal kernel, but cumulative

**Figure 4.** Distributions of gene diversity ($H_e$) and allelic richness ($R$), defined as the expected number of different alleles in a sample of 6 genes, of pollen pools accepted by each seed parent as a function of the $N_{cp}$ for seed parents of *Castanopsis sieboldii*. Linear functions of inverse of the $N_{cp}$ were fitted to the data.

**Discussion**

**Patterns of Pollen Flow**

The paternity analysis indicated that all the pollen parents that had at least one offspring, had a d.b.h. ≥ 36.2 cm and reached the canopy layer. *Castanopsis sieboldii* is categorized as a canopy-flowering tree species, that is, its flowers are usually found only in the canopy, seldom in the understory (Yumoto 1987). The plot in this study was established in the site with the most fully developed canopy in the old-growth evergreen broad-leaved forest (Manabe et al. 2000). Therefore, only the canopy trees can presumably flower in the plot. We were unable to sample the leaves of one adult, which had a d.b.h. of 5.4 cm, did not reach the canopy layer, and died during the course of the study. The contribution of the adult that could not be genotyped was presumably null. Therefore, we were able to sample virtually all of the reproductive individuals within the 4-ha plot that contributed to the seeds sampled in 2000, taking into account the size of reproductive individuals (d.b.h. ≥ 36.2 cm, reaching the canopy layer). The density of reproductive individuals may be at least 30 individuals ha⁻¹, based on the number of individuals (120) having a d.b.h. ≥ 36.2 cm and reaching the canopy layer.

A negative exponential function could explain the cumulative frequencies of pollinations dependent on the distance between parents within the plot, and the actual average distance of pollination within the plot was significantly
pollen pools from inside the plot. Therefore, the level of genetic diversity appeared to be constrained, possibly due to the genetic structure of the adult trees and frequent short distance pollination. Adult trees near a seed parent may contribute frequently to offspring, but the adults also tend to be genetically related to each other due to the genetic structure of the population, thus the genetic diversity of pollen pools accepted by each seed parent could not be higher than a certain level. The genetic diversity of the pollen pool bulked over all seed parents from inside the plot did not differ from that for the total pollen pool. Therefore, long-distance pollen flow does not seem to enhance the genetic diversity of the offspring significantly. In the correlogram of the average $F_{ij}$ values of adults, indicating weak genetic structure at a local scale, the values were significantly positive at short distance classes and decreased as distances increased up to (but not beyond) 120 m. Therefore, the level of genetic differentiation between adults at a certain scale (longer than 120 m) may not increase as the distance increases, and thus long-distance pollen flow from outside the plot may not substantially increase the level of genetic diversity among the offspring. The pollen parents near seed parents may be the main contributors to the genetic diversity of offspring at the seed stage because of their frequent pollination. However, as discussed above, the long-distance pollination with a relatively low frequency may have a selective advantage at the later stages and thus the contribution of pollen parents located far from seed parents may be important for sustaining genetic diversity of populations in the species.

The observed pattern of pollen dispersal in *C. sieboldii* indicates high potential for long-distance dispersal (as indicated by the fat-tailed dispersal kernel), in accordance with previous findings for other insect-pollinated tree species. However, the cumulative pollination is limited and strongly dependent on the distance between parents at the local scale due to the high density of reproductive trees. Due to the distance-dependent pollination, we found that: 1) neighboring trees that are genetically related to their seed parents may have frequently mated; 2) there were substantial differences in the contributions of pollen parents to the offspring among the seed parents, probably due to the abovementioned dependence of pollination on the between-parent distance. Different seed parents located near to and far from each other may be more and less frequently pollinated by the same pollen parents, respectively, due to the limited pollen dispersal. Hence, the genetic similarity of the pollen pools accepted by seed parents may be inversely related to the distance between them. Hardy et al. (2004) discussed different mechanisms causing differentiation among pollen pools and suggested limited pollen dispersal as the main factor and phenological heterogeneity among plants as the second factor.

Linear functions of inverse of the $N_{ep}$ for each seed parent significantly explained the genetic diversity of the pollen pools for each seed parent. The estimated genetic diversity of the pollen pools increased as $N_{ep}$ increased, but the estimates seemed to converge at a maximum. This result was determined from the analyses of $N_{ep}$ for pollen parents within the plot and the pollen pools from trees inside the plot using data for seeds with unambiguously assigned pollen parents. However, the average values for the total pollen pool over all 11 seed parents ($0.839 \pm 0.008$ and $4.2 \pm 0.1$ for $H_k$ and $R$, respectively) were lower than or similar to the maximum values ($0.861$ and $4.2$ for $H_k$ and $R$, respectively; Figure 4) estimated based on the functions for

Heterogeneity of Pollen Parents and Pollen Pools

The $N_{ep}$ estimated for each seed parent and for an average seed parent within the plot were lower than the estimated global $N_{ep}$ value. The genetic composition of the total pollen pool significantly differed among the 11 seed parents, and the genetic differentiation was correlated with the distance between them. These findings indicate that there were substantial differences in the contributions of pollen parents to the offspring among the seed parents, probably due to the abovementioned dependence of pollination on the between-parent distance. Different seed parents located near to and far from each other may be more and less frequently pollinated by the same pollen parents, respectively, due to the limited pollen dispersal. Hence, the genetic similarity of the pollen pools accepted by seed parents may be inversely related to the distance between them. Hardy et al. (2004) discussed different mechanisms causing differentiation among pollen pools and suggested limited pollen dispersal as the main factor and phenological heterogeneity among plants as the second factor.

The actual average $F_{ij}$ between mating pairs of parents within the plot including and excluding self-pollination were significantly higher than those expected under random mating. The high average $F_{ij}$ values may be caused by the genetic structure of the adult trees and the abovementioned relationship between spatial distance and pollination (i.e., that adult trees near seed parents tend to be genetically related to the seed parents due to the genetic structure, and may often be pollinated by the seed parents because of the short distances between them). Consequently, neighboring trees that are genetically related to their seed parents may have frequently mated. However, the self-pollination rate was very low, although the distance-dependent pollination should be expected to cause a high frequency of self-pollination. Therefore, postpollination mechanism may favor outcrossing as discussed in the studies of other fagaceous species (Lumaret et al. 1991). Furthermore, because the average value of $F_{ij}$ for seeds was significantly higher than that for adults, inbreeding depression may occur and the long-distance pollination with a relatively low frequency may have a selective advantage at the later stages.

Funding

Grants-in-Aid for Scientific Research (Nos. 11460069 and 14206017) from the Japan Society for the Promotion of Science.

Acknowledgments

We are grateful to T. Fujita and other members of the Laboratory of Forest Ecology and Physiology of Nagoya University, who provided field and laboratory assistance. We thank the Tsushima District Forest Office for permitting this study.


Received November 8, 2010; Revised November 23, 2011; Accepted March 14, 2012

Corresponding Editor: Scott Hodges