

ORIGINAL ARTICLE

Estimation of outcrossing rates at hierarchical levels of fruits, individuals, populations and species in *Magnolia stellata*

I Tamaki, S Setsuko and N Tomaru

Laboratory of Forest Ecology and Physiology, Graduate School of Bioagricultural Sciences, Nagoya University, Nagoya, Japan

In plant species with mixed mating systems, differences in diverse factors—including their pollination system, flowering phenology, life form and susceptibility to inbreeding depression—cause variation in outcrossing rates among fruits within individuals, among individuals within populations and among populations within species. To quantify this hierarchical variation, we examined outcrossing rates at the seed stage in five populations of *Magnolia stellata*, a self-compatible, insect-pollinated and protogynous tree species. For this purpose, we sampled 1498 seeds within 204 fruits obtained from 56 individuals of the five populations, determined genotypes of the sampled seeds and maternal individuals at six polymorphic microsatellite loci, then estimated outcrossing rates and their variance components at four hierarchical levels (fruits, individuals, populations and species) using a nested analysis of variance-type linear model with a Bayesian approach. The species-level outcrossing rate was 0.730

(95% credible interval, 0.595–0.842), indicating that this species has a mixed mating system. Outcrossing rates were not significantly different among populations, but were significantly different among individuals within populations. Variance components at the levels of individual and fruit were statistically supported and were highest for the former. Thus, factors influencing outcrossing rates at the individual level, such as differences in flowering phenology and early-stage inbreeding depression, appear to have important effects within these *M. stellata* populations, but not among them. The method of hierarchically estimating outcrossing rates using a Bayesian approach, as applied in this study, is compared with conventional methods for estimating outcrossing rates, and the statistical properties of the Bayesian approach are discussed.

Heredity (2009) **102**, 381–388; doi:10.1038/hdy.2008.128; published online 14 January 2009

Keywords: geitonogamy; hierarchical Bayesian model; insect pollination; mixed mating system; protogyny; variance component

Introduction

Since Darwin (1876), botanists have paid great attention to mixed mating systems of plants (Schemske and Lande, 1985; Goodwillie *et al.*, 2005; Jarne and David, 2008). Plants with such systems vary widely in terms of pollination system, flowering phenology, life form and susceptibility to inbreeding depression. Furthermore, variation in outcrossing rates, caused by these factors, has been widely observed at a range of hierarchical levels, including regions (Bower and Aitken, 2007; Mable and Adam, 2007), populations (Routley *et al.*, 1999; Bower and Aitken, 2007; Karasawa *et al.*, 2007; Michalski and Durka, 2007; Yates *et al.*, 2007; Ishida, 2008), individuals (El-Kassaby *et al.*, 1993; Lee, 2000; Bower and Aitken, 2007; Nassar *et al.*, 2007; de-Lucas *et al.*, 2008), parts within individual plants (El-Kassaby *et al.*, 1993; Patterson *et al.*, 2004) and fruits (Cruzan *et al.*, 1994). The effects of pollination systems can be illustrated by the differences between wind-pollinated and insect-

pollinated species. Generally, wind-pollinated species show either high or low outcrossing rate (Culley *et al.*, 2002), whereas insect-pollinated species show a wide range of outcrossing rates (Aide, 1986; Goodwillie *et al.*, 2005). Moreover, inter-population variation in outcrossing rates is higher for insect-pollinated species than for wind-pollinated species (Schemske and Lande, 1985; Aide, 1986). In addition, pollination by beetles and flies is generally less effective than pollination by bumble bees and other bees (Osada *et al.*, 2003), and the uncertainty of successful pollination seems to be generally higher among insect-pollinated species, especially those pollinated by insects other than bumble bees, than among wind-pollinated species. The effects of life form can be illustrated by findings that long-lived tree species generally have higher outcrossing rates than annuals (Hamrick and Godt, 1996; Petit and Hampe, 2006). Moreover, as members of tree species are of large plant sizes and environments are heterogeneous even within individuals' crowns, differences in outcrossing rates between upper and lower parts of their crowns have been reported (El-Kassaby *et al.*, 1993; Patterson *et al.*, 2004). The synchronization of flowering phenology also has a strong influence on the efficiency of pollination and, ultimately, on plants' mating systems (Hirao *et al.*, 2006). In dichogamous species, asynchronization of flowering phenology within individual plants increases

Correspondence: I Tamaki, c/o N Tomaru, Laboratory of Forest Ecology and Physiology, Graduate School of Bioagricultural Sciences, Nagoya University, Nagoya 464-8601, Japan.

E-mail: garageit@gmail.com

Received 29 June 2008; revised 29 October 2008; accepted 7 December 2008; published online 14 January 2009

the opportunity for geitonogamous selfing. Consequently, differences in the degrees of inbreeding depression during the development of embryos following fertilization can cause major variation in seeds' or later stages' outcrossing rates.

As many factors cause variation in outcrossing rates among components at hierarchical levels, when considering the ecological and/or evolutionary significance in the value of outcrossing rates, the hierarchical structure of data should be taken into account by estimating the components of variance in outcrossing rates at different hierarchical levels. Otherwise biased or inappropriate estimates of outcrossing rates may be obtained, and their significance may be misinterpreted. This problem is not typical for outcrossing rate estimation but does occur when estimating parameters associated with hierarchically nested binomial data. For example, the same problem occurs when estimating fruit set, seed set, germination rate or survival rate. To address this problem, a nested analysis of variance-type linear model seems to be appropriate, although there are some essential problems. At the lowest level, as obtainable data are finite, binomial sampling errors will inflate variance estimates, depending on sample sizes. Moreover, at the higher levels, the variances inherit sampling errors from the lower levels. This has been a statistical problem for decades. Maximum likelihood is hard to compute exactly as the number of hierarchical levels of data increases, so a Bayesian approach seems appropriate. Hierarchical Bayesian models have been used recently in a range of disciplines, including ecology, genetics, atmospheric sciences, finance and medicine (Clark, 2005), as they can provide more profound insights into the structure of complex, heterolytic data sets than conventional models (Clark, 2005, 2007). In the study presented here, a nested analysis of variance-type linear model with a Bayesian approach was applied using Markov Chain Monte Carlo methodology to estimate outcrossing rates at the seed stage and their components of variance, at multiple levels (fruits, individuals, populations and species) in five populations of *Magnolia stellata* (Magnoliaceae). The outcrossing rates obtained using the hierarchical Bayesian method are compared with estimates obtained by two conventional methods. The statistical properties of the Bayesian approach are then discussed.

Materials and methods

Study organism and sampling

M. stellata (Sieb. et Zucc.) Maxim. is a tree species of the Magnoliaceae, commonly known as 'star magnolia', that is endemic to the Tokai region in Japan and favored as an ornamental garden plant in Europe and North America (Callaway, 1994). It occurs at swampy sites such as the margins of small rivers or marshes (Ueda, 1988) and grows to 10 m in height and 20 cm in diameter at breast height. The main pollinators of the species are rove beetles and Thysanoptera, but rare visits by bumble bees and honey bees have also been observed (Setsuko *et al.*, 2008; I Tamaki, unpublished). *M. stellata* flowers in early spring, but strong fluctuations in the temperature and other climatic variables at this time, are thought to cause wide variation in its pollinators' activities and, conse-

quently, in *M. stellata*'s mating parameters. *M. stellata* has a mixed mating system (Hirayama *et al.*, 2005) and produces protogynous hermaphrodite flowers with self-compatibility. Therefore, autogamous selfing cannot occur (Saeda *et al.*, 2004), but geitonogamous selfing can as flowering is not fully synchronous within individual trees (Setsuko *et al.*, 2008). The fruits are aggregated and the numbers of seeds per fruit are highly variable, ranging from 1 to 40, within and among individual trees (I Tamaki, unpublished). Clearly, many factors affect the outcrossing rates of *M. stellata*, which thus provides ideal material for studying variation in mixed mating systems at hierarchical levels.

We collected leaves and fruits from individual trees of five populations (Kurogawa, Toshichibara, Ikawazu, Nagusa and Tabika). Detailed information regarding these populations, including their geographical positions and genetic variation, has been presented earlier (Tamaki *et al.*, 2008). Here, materials from individuals from which we collected more than two fruits were analyzed, including a total of 1498 seeds from 204 fruits obtained from 56 individuals representing the five populations (Table 1). Average numbers of sampled seeds per fruit, sampled fruits per individual and sampled individuals per population were 7.3, 3.6 and 11.2, respectively. The sampled leaves and the seeds extracted from sampled fruits were stored at -30 and 4°C , respectively, until DNA extraction.

DNA extraction and microsatellite analysis

Genomic DNA was extracted from the sampled leaves, and from the embryo and endosperm in sampled seeds, using the hexadecyltrimethylammonium bromide method (Murray and Thompson, 1980) with minor modifications or a DNeasy Plant Mini Kit (Qiagen, Heilden, Germany). The genotypes of the sampled materials were then determined using six microsatellite markers developed for *M. stellata* (Setsuko *et al.*, 2005): *stm0184*, *stm0222*, *stm0223*, *stm0251*, *stm0334* and *stm0423*. PCR fragments were separated using a 3100 Genetic Analyzer in conjunction with GeneScan software (Applied Biosystems, Carlsbad, CA, USA), see Setsuko *et al.* (2005) for details.

Data analysis

The existence of null alleles from maternal gametes (ovules) can be detected by analyzing several seeds per known maternal genotype. Null alleles at five loci were

Table 1 Population name, code and size, and the numbers of individuals, fruits and seeds sampled from each of the five studied *Magnolia stellata* populations

Name	Population		Number of samples		
	Code	Size	Individual	Fruit	Seed
Kurogawa	Kur	29	11	38	287
Toshichibara	Tos	130	12	41	347
Ikawazu	Ika	30	12	48	383
Nagusa	Nag	53	13	48	373
Tabika	Tab	51	8	29	108
Total			56	204	1498

Population sizes are the numbers of reproductive individuals in the populations.

recognized in five family arrays in our data sets in this manner, and the genotypes at those loci in those families were treated as missing data. It is not possible to infer null alleles originating from paternal gametes (pollen) in this way, which may lead to underestimation of outcrossing rates. However, as we estimated outcrossing rates using information from polymorphic multilocus data, any such underestimation is likely to have been minor.

In many studies, estimates of outcrossing rates have been obtained using mixed mating models implemented in the program MLTR (Ritland, 2002). However, seeds that originated from outcrossing events can be unambiguously identified by comparing the genotypes of seeds and their maternal trees, as alleles that are present in the seeds but not in the maternal tree must have come from other, paternal sources (Reusch, 2001; Obayashi *et al.*, 2002; Kenta *et al.*, 2004; Isagi *et al.*, 2007; Sampson and Byrne, 2008) apart from a very low frequency that may have originated from mutations. The accuracy of estimates obtained using this approach depends on the levels of polymorphism of the genetic markers used, as a seed that is derived from an outcrossing event will not be recognized as such if it has no alleles that differ from those of its maternal tree at the examined loci. Here, we minimized the probability of such failures by using multiple highly polymorphic microsatellite markers, and we calculated the probability that outcrossing would not be detected at locus *l* (the parameter β_l) from the following expression (Shaw *et al.*, 1981; Weir, 1996):

$$\beta_l = \sum_u P_{lu,lu} P_{lu} + \sum_u \sum_{u \neq v} P_{lu,lv} (P_{lu} + P_{lv}),$$

where $P_{lu,lu}$ is the frequency of a homozygous genotype of allele *u*, $P_{lu,lv}$ is the frequency of a heterozygous genotype of alleles *u* and *v*, and p_{lu} and p_{lv} are the frequencies of alleles *u* and *v*, respectively. The estimated frequencies of genotypes and alleles each of the populations used in this analysis were derived in an earlier study (Tamaki *et al.*, 2008), and genotype frequencies were weighted according to the numbers of seeds contributed by maternal individuals with the respective genotypes. The multilocus probability of not detecting outcrossing, α , is the product of β_l over all loci:

$$\alpha = \prod_l \beta_l.$$

To estimate outcrossing rates and their components of variance at the levels of fruits, individuals, populations and species, a nested analysis of variance-type linear model with a hierarchical Bayesian approach was used. The number of outcrossed seeds in fruit *i* within individual *j* within population *k* (oc_{ijk}) was assumed to follow a binomial (Bin) distribution with the expected outcrossing rate (θ_{ijk}^{fruit}) and the number of total seeds (n_{ijk}):

$$oc_{ijk} \sim \text{Bin}(\theta_{ijk}^{fruit}, n_{ijk}).$$

Outcrossing rates depend hierarchically on the effects of fruits (β_{ijk}^{fruit}), individuals ($\beta_{jk}^{individual}$), populations ($\beta_k^{population}$) and species (a common, suprapopulation component; $\beta^{species}$) levels:

$$\text{logit}(\theta_{ijk}^{fruit}) = \beta_{ijk}^{fruit} + \beta_{jk}^{individual} + \beta_k^{population} + \beta^{species}.$$

Where these parameters are assumed to follow a weakly informative normal (N) prior distribution:

$$\begin{aligned} \beta_{ijk}^{fruit} &\sim N(0, \sigma^{2fruit}), \\ \beta_{jk}^{individual} &\sim N(0, \sigma^{2individual}), \\ \beta_k^{population} &\sim N(0, \sigma^{2population}), \\ \beta^{species} &\sim N(0, 1000), \end{aligned}$$

and variances of β_{ijk}^{fruit} , $\beta_{jk}^{individual}$ and $\beta_k^{population}$ are assumed to follow a weakly informative inverse gamma (IG) hyperprior distribution:

$$\begin{aligned} \sigma^{2fruit} &\sim \text{IG}(0.001, 0.001), \\ \sigma^{2individual} &\sim \text{IG}(0.001, 0.001), \\ \sigma^{2population} &\sim \text{IG}(0.001, 0.001). \end{aligned}$$

In this model, outcrossing rates at the levels of fruits (θ_{ijk}^{fruit}), individuals ($\theta_{jk}^{individual}$), populations ($\theta_k^{population}$) and species ($\theta^{species}$) are estimated by logistic ($\beta_{ijk}^{fruit} + \beta_{jk}^{individual} + \beta_k^{population} + \beta^{species}$), logistic ($\beta_{jk}^{individual} + \beta_k^{population} + \beta^{species}$), logistic ($\beta_k^{population} + \beta^{species}$) and logistic ($\beta^{species}$), respectively. To improve the convergence properties of Markov Chain Monte Carlo simulation in the above model scheme, hierarchical centering reparameterization was conducted following Gelfand *et al.* (1995):

$$\begin{aligned} \text{logit}(\theta_{ijk}^{fruit}) &\sim N(\text{logit}(\theta_{jk}^{individual}), \sigma^{2fruit}), \\ \text{logit}(\theta_{jk}^{individual}) &\sim N(\text{logit}(\theta_k^{population}), \sigma^{2individual}), \\ \text{logit}(\theta_k^{population}) &\sim N(\text{logit}(\theta^{species}), \sigma^{2population}), \\ \text{logit}(\theta^{species}) &\sim N(0, 1000). \end{aligned}$$

We implemented this model using WinBUGS 1.4.3 (Spiegelhalter *et al.*, 2003) and R2WinBUGS package in R 2.7.2 (Sturtz *et al.*, 2005; R Development Core Team, 2008), running Markov Chain Monte Carlo simulations with three chains (5000 iterations after 5000 burn-in iterations, during which convergence was reached) and kept every tenth sample from each chain to provide 1500 independent sets of samples from the posterior distribution (see the data and scripts in the Supplementary information). A deviance information criterion (DIC) analysis was conducted to examine the statistical support for each variance component. DIC is intended to serve as a generalization of Akaike's information criterion in hierarchical models; models with smaller values of DIC are regarded as being better (Spiegelhalter *et al.*, 2002). DIC is calculated by the following equation:

$$\text{DIC} = \bar{D}(\phi) - 7pD,$$

where $\bar{D}(\phi)$ is the posterior mean of the deviance that is -2 times log-likelihood of the model with the parameter vector ϕ , pD is the effective number of parameters and is given by $\bar{D}(\phi) - D(\bar{\phi})$, and $D(\bar{\phi})$ is a point estimate of the deviance obtained using the posterior mean of ϕ . When the difference between DICs is larger than 10, the model with the lower DIC value is statistically supported, but when this difference is lower than 5, there is little difference between the models. We compared the DIC value for the full model with those for the models in which each hierarchical component was eliminated from the full model.

In addition, we estimated outcrossing rates at each hierarchical level using two alternative methods that disregard the number of samples and the hierarchical structure of the data, and compared them with those estimated by the hierarchical Bayesian model. First, the iterative averaging method was used, in which outcrossing rates at higher levels were estimated by averaging those at lower levels. Second, we used the pooling method, estimating outcrossing rates at a given level by pooling data for all seeds within each component (that is, individuals, populations or species) at that level. Note that outcrossing rates at the lowest level, that is, the fruit level, were estimated (in both cases) by dividing the number of outcrossed seeds within a fruit by the total number of seeds within it.

Results

The multilocus probability that an outcrossing event had occurred but was not discernible, α , was lower than 0.0001 for each of the five populations, and the numbers of seeds sampled from each of the populations ranged from 108 to 383 (Table 1), thus the number of outcrossed seeds from each population that was erroneously designated selfed is likely to have been less than one.

The posterior mean outcrossing rate at the species level obtained by applying the hierarchical Bayesian method to our samples was 0.730 (95% credible interval (CI), 0.595–0.842; Figure 1a), whereas the outcrossing rates at the species level obtained by the iterative averaging and pooling methods were 0.652 and 0.666, respectively (Figures 2a and b). Posterior mean values of the outcrossing rates at the population level ranged from 0.671 to 0.769, and their CIs overlapped (Figure 1b). In contrast, outcrossing rates at the individual level varied widely and their CIs did not overlap between some pairs of individuals in all of the populations except the Tabika population (for example, between individuals Kur08 and

Kur09 of the Kurogawa population or individuals Tos01 and Tos14 of the Toshichibara population; Figure 3). Outcrossing rates at the fruit level were generally very similar within individuals; only a pair of fruits from individual Tos10 of the Toshichibara population showed no overlapping of CIs (Figure 4).

Posterior mean values of the components of variance in outcrossing rates among fruits within individuals, among individuals within populations and among populations were 1.121 (CI, 0.599–1.880), 1.691 (0.856–2.872) and 0.252 (0.001–1.682), respectively (Table 2). Their CIs overlapped, although the variance components at the individual and the population levels showed the highest and lowest values, respectively. On the basis of the DIC analysis, the statistical significances of variance components at the fruit and the individual levels were supported (Δ DIC = 79.6 and 12.3, respectively), whereas those at the population level were neither supported nor rejected (Δ DIC = 0.3; Table 3). However, differences in the values of pD were very small when comparing model 0 ($pD = 117.5$) and model 3 ($pD = 117.4$). This may suggest that $\beta^{\text{population}}$ makes a small contribution to the model.

Similar trends were observed in the relationships between the outcrossing rates estimated using the hierarchical Bayesian model and those obtained using both the iterative averaging and pooling methods (Figures 2a and b). When the outcrossing rates estimated using either the iterative averaging or pooling method were lower than 0.6 and higher than 0.8, they were lower and higher than those estimated using the hierarchical Bayesian model, respectively. These trends were also observed in the estimated fruit-level outcrossing rates, and especially prominent when sample sizes were small (Figure 2c).

Discussion

Mating system and its variation in *M. stellata*

The 95% CI of species-level outcrossing rates in *M. stellata* ranged from 0.595 to 0.842, showing that the species has a mixed mating system that is slightly biased toward outcrossing. The estimate of the outcrossing rate of *M. stellata* obtained in this study (0.730) is similar to, or slightly higher than, estimates obtained for two congeneric species: *M. obovata* (0.10–0.66; Ishida, 2008) and *M. salicifolia* (0.650; K Takabayashi and N Tomaru, unpublished). The similarity of those values may result from the shared mating characteristics of members of the Magnoliaceae, such as insect pollination and protogyny (Bernhardt and Thien, 1987; Callaway, 1994).

Population-level outcrossing rates obtained for the five populations (0.671–0.769) did not significantly differ from one another, as indicated by their overlapping CIs, and were consistent with those obtained for two other populations with 10 and 20 individuals (0.612 and 0.739, respectively) examined in an earlier study of the species (Hirayama *et al.*, 2007). In the cited study, the outcrossing rate obtained for the small population was lower than that of the large population, but in this study no positive relationships between population size and population-level outcrossing rates were detected ($r = -0.8$, $P = 0.09$). Therefore, the population-level outcrossing rate may be quite constant, regardless of

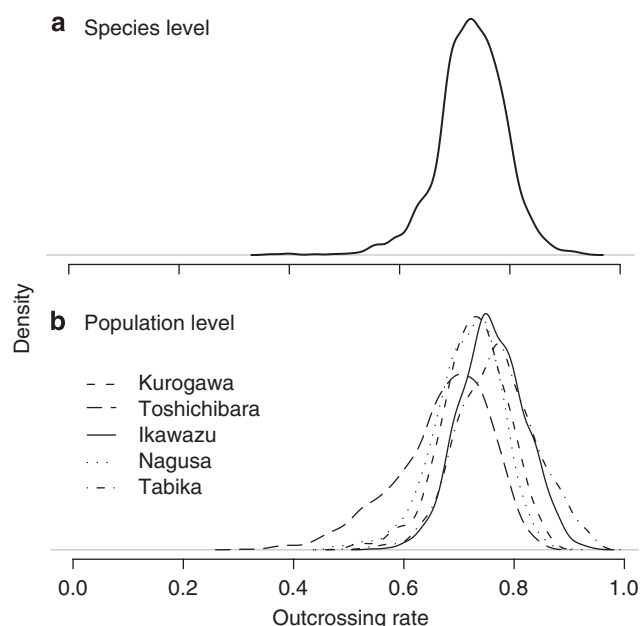


Figure 1 Posterior distributions of outcrossing rates at the levels of species (a) and populations (b) in *Magnolia stellata*.

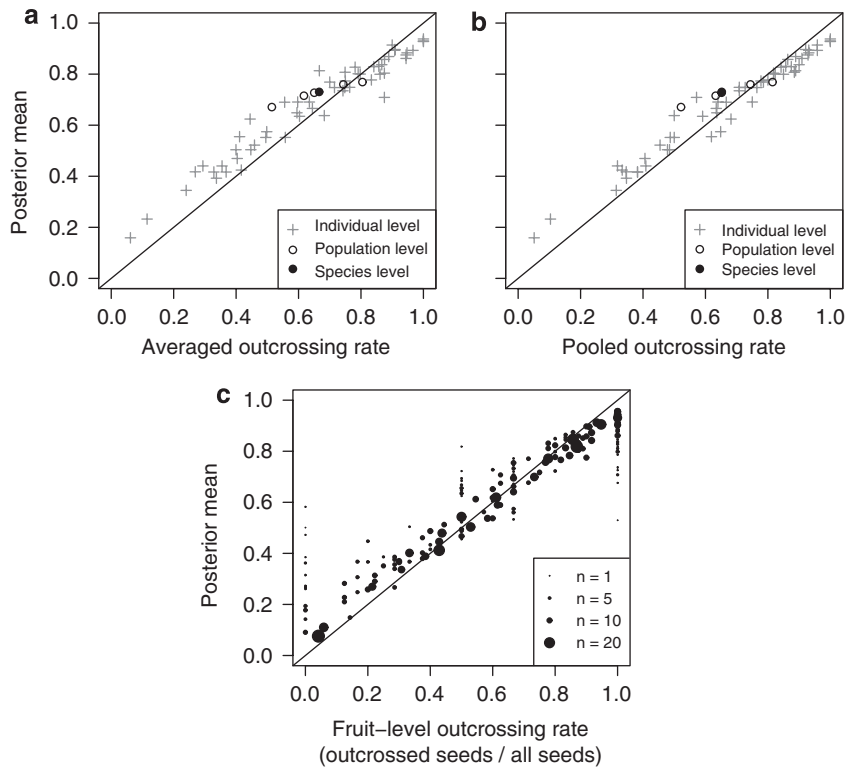


Figure 2 Relationships of posterior mean values of outcrossing rates obtained using the hierarchical Bayesian model and outcrossing rates obtained using the iterative averaging (a) and pooling (b) methods, and outcrossing rates at the fruit level (c). Lines indicate that outcrossing rates calculated by each method are identical to posterior mean values.

population size. However, the five populations investigated in this study and the small population (Yokkaichi) examined by Hirayama *et al.* (2007) were distributed at the edge of the species' range (Tamaki *et al.*, 2008). Therefore, additional analyses of populations distributed in the central part of the species' range are required to test the hypothesis more rigorously. The CIs of the outcrossing rates at the population level overlapped, whereas the variance component at the population level was neither supported nor rejected from DIC analysis. This may be due to the small numbers of study populations, and further research, based on a greater number of populations, is needed. It should also be noted that Setsuko *et al.* (2007) obtained a much higher population-level estimate of outcrossing rate of *M. stellata* (0.943) at the seedling stage in a paternity analysis (Setsuko *et al.*, 2007), possibly due to selection against selfed seeds before and during germination.

In contrast to the quite constant population- and fruit-level outcrossing rates we obtained, the individual-level outcrossing rates were highly varied; the variance component in outcrossing rates among individuals within populations was higher than that at any of the other examined hierarchical levels. Thus, in this species, outcrossing rates seem to vary mostly among individuals within populations. Similarly, de-Lucas *et al.* (2008) examined the mating system of Mediterranean maritime pine at the hierarchical levels, in a nested analysis of 61 individuals, in 24 plots, of three populations, and found that most (99%) of the variation in outcrossing rates of the species resided among individuals. The cited authors attributed some of this variation to differences in tree size and crown width. Unfortunately, we did not

measure these variables, so we were unable to assess their effects. In an *M. stellata* population, synchronous, but less than perfectly synchronous, flowering among individual trees has been reported (Setsuko *et al.*, 2008). As trees that flower early in a population are relatively unlikely to be pollinated by pollen from other trees, outcrossing rates of these individuals may be low. In addition, in two populations of this species, differences in early-life-stage inbreeding depression (inbreeding depression during the development of embryo following fertilization) among individual trees have been reported (Hirayama *et al.*, 2007). Thus, among-individual differences in early-life-stage inbreeding depression are likely to provide a further source of variation in outcrossing rates among individuals. The variation of individual-level outcrossing rates observed in this study may result from the combined effects of these two phenological and genetic factors. There are some differences in flowering phenology among flowers within an individual tree (Setsuko *et al.*, 2008) and the variance component was statistically supported; however, there were few differences among fruits within individual trees as indicated by their overlapping CIs. Thus, the individual-level factors influencing outcrossing rates discussed above may be stronger than fruit-level factors.

Comparisons between the hierarchical Bayesian model and two alternative methods, and the statistical properties of the model

The outcrossing rates estimated using both the iterative averaging and pooling methods were lower than those obtained using the hierarchical Bayesian model when

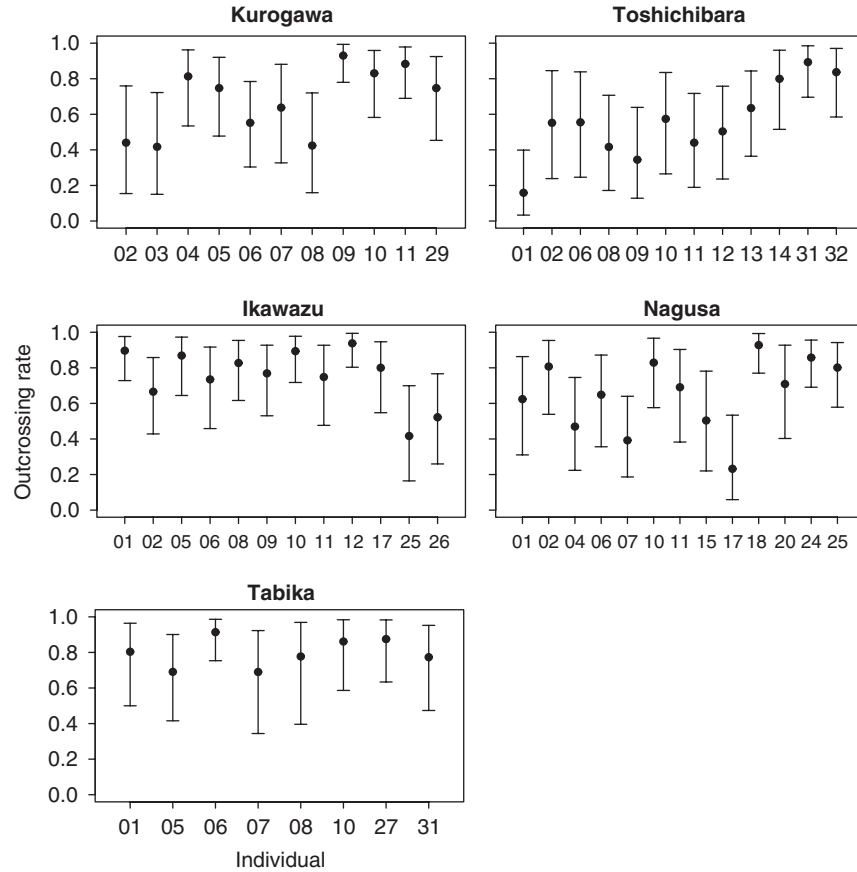


Figure 3 Posterior distributions of outcrossing rates at the individual level in five *Magnolia stellata* populations. Population's names are indicated above each graph, whereas individuals are designated by their ID codes below each graph. Dots and bars show posterior mean values and their 95% credible intervals, respectively.

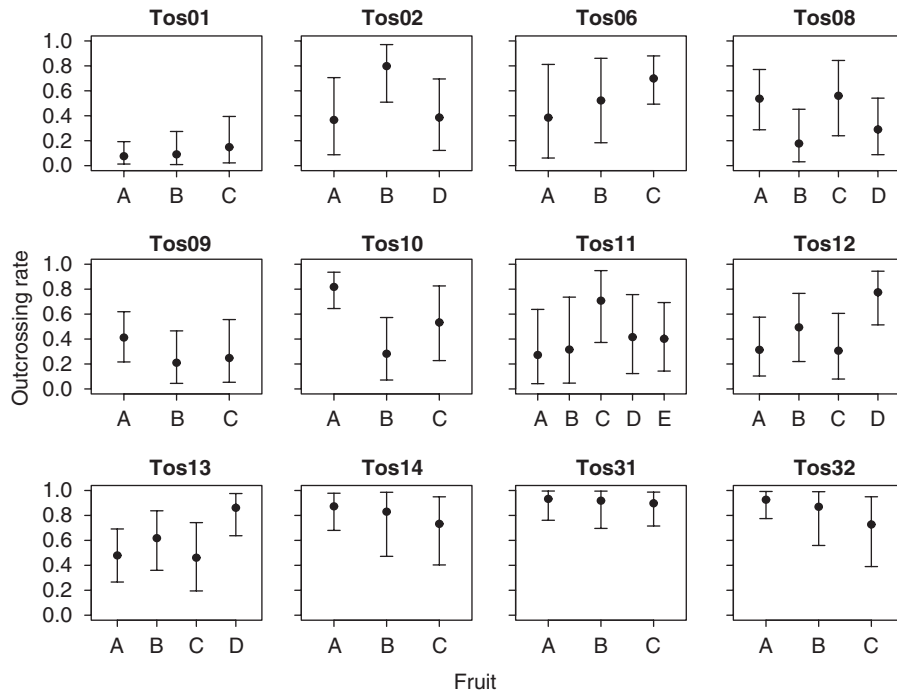


Figure 4 Posterior distributions of outcrossing rates at the fruit level in the Toshichibara population of *Magnolia stellata*. Codes above each graph indicate ID codes of individuals, whereas fruits are designated by their ID codes below each graph. Dots and bars show posterior mean values and their 95% credible intervals, respectively.

Table 2 Posterior distributions of the components of variance in outcrossing rates at each hierarchical level in *Magnolia stellata*

Source of variation	Mean	s.d.	Credible interval	
			2.5%	97.5%
Among fruits within individuals	1.121	0.324	0.599	1.880
Among individuals within populations	1.691	0.519	0.856	2.872
Among populations	0.252	0.521	0.001	1.682

Table 3 Posterior distribution of deviance, effective number of parameters (pD), deviance information criterion (DIC) and the difference in DIC from model 0 (ΔDIC) in the full model and the models in which each hierarchical component is eliminated from the full model

Model	No.	Description	Deviance		pD	DIC	ΔDIC		
			Mean	s.d.					
			Credible interval						
			2.5%	97.5%					
0		Full model	510.7	20.7	470.1	551.8	117.5	627.7	—
1		Full model— β^{fruit}	659.2	10.8	640.2	682.6	48.1	707.3	79.6
2		Full model— $\beta^{\text{individual}}$	502.3	20.9	462.0	544.0	137.6	640.0	12.3
3		Full model— $\beta^{\text{population}}$	510.6	20.5	470.5	551.1	117.4	628.0	0.3

outcrossing rates were low and higher when outcrossing rates were high. When outcrossing rates were extreme (0 and 1) or the number of sampled seeds was small, some fruit-level outcrossing rates differed markedly from those derived using the hierarchical Bayesian model. These results may depend on two properties of the model. First, the logit link function does not allow rates of just 0 and 1. Second, in the nested design, the predicted outcrossing rate for a particular fruit is closer to the mean value at the individual level than the actual proportion of outcrossed seeds observed in that particular fruit, especially when the number of sampled seeds per fruit is small. This is because of binomial sampling error. It could be suggested that analyses of outcrossing rates should be conducted using data sets that consist of fruits containing sufficient seeds. However, fruits with small numbers of seeds due to a shortage of pollen occur commonly in the field (Hirayama et al., 2005), so omitting such fruits from the analyses may lead to biased estimates of outcrossing rates and variance components.

Whether or not the outcrossing rates estimated by the model in this study are better than those estimated using conventional methods depends on the assumption within the model that outcrossing rates are normally distributed on the logit scale. This should be remembered when applying the model. Moreover, two aspects should be noted. First, genotype data for both maternal trees and their seeds are needed. Second, multiple highly polymorphic markers, such as nuclear microsatellite markers, are needed and the multilocus probability of failure to detect outcrossing events must be calculated. When the status of seeds (outcrossed or selfed) is ambiguous, the model cannot be applied in a straightforward way and a process that estimates the status of individual seeds stochastically from allele frequencies, such as the multi-locus mixed mating model (Ritland and Jain, 1981), will need to be embedded to the model.

In current population genetics, many microsatellite markers have been developed; applying microsatellite markers developed for related species should address the problem easily and directly, avoiding the need to develop a much more complex algorithm. Extension of the Bayesian model by adding the effects of different years would enable the differences and variances in outcrossing rates between years to be considered too. This model could also be fitted to data with similar hierarchical structures for estimating fruit set, seed set or germination rate. As these data do not have any uncertainty associated with recording their status, applying the model to these data would be easier than its use in the estimation of outcrossing rates. If the results of mating system studies using this model are archived, they could be very useful in any attempts to elucidate poorly understood aspects of the evolutionary and ecological implications of mixed mating systems in plants.

Acknowledgements

We are grateful to the members of the laboratory of Forest Ecology and Physiology of Nagoya University for useful discussions. We also thank the Education Offices of Tahara and Komono Cities for permitting us to collect the materials, T Kubo for his helpful suggestion about the hierarchical Bayesian model and two anonymous reviewers for their variable comments on a previous draft of the paper. Financial support was provided by a grant-in-aid for Scientific Research (no. 16380100) from the Japan Society for the Promotion of Science.

References

- Aide TM (1986). The influence of wind and animal pollination on variation in outcrossing rates. *Evolution* **40**: 434–435.
- Bernhardt P, Thien LB (1987). Self-isolation and insect pollination in the primitive angiosperms: new evaluations of older hypotheses. *Plant Syst Evol* **156**: 159–176.
- Bower AD, Aitken SN (2007). Mating system and inbreeding depression in whitebark pine (*Pinus albicaulis* Engelm.). *Tree Genet Genomes* **3**: 379–388.
- Callaway DJ (1994). *The World of Magnolias*. Timber Press: Oregon, USA, 260pp.
- Clark JS (2005). Why environmental scientists are becoming Bayesians. *Ecol Lett* **8**: 2–14.
- Clark JS (2007). *Models for Ecological Data*. Princeton University Press: New Jersey, USA.
- Cruzan MB, Hamrick JL, Arnold ML, Bennett BD (1994). Mating system variation in hybridizing irises: effects of phenology and floral densities on family outcrossing rates. *Heredity* **72**: 95–105.
- Culley TM, Weller SG, Sakai AK (2002). The evolution of wind pollination in angiosperms. *Trends Ecol Evol* **17**: 361–369.
- Darwin C (1876). *The Effects of Cross and Self Fertilisation in the Vegetable Kingdom*. John Murray: London, UK.
- de-Lucas AI, Robledo-Arnuncio JJ, Hidalgo E, Gonzalez-Martinez SC (2008). Mating system and pollen gene flow in Mediterranean maritime pine. *Heredity* **100**: 390–399.
- El-Kassaby YA, Meagher MD, Davidson R (1993). Temporal variation in the outcrossing rate in a natural stand of western white pine. *Silvae Genet* **42**: 131–135.
- Gelfand AE, Sahu SK, Carlin BP (1995). Efficient parametrisations for normal linear mixed models. *Biometrika* **82**: 479–488.
- Goodwillie C, Kalisz S, Eckert CG (2005). The evolutionary enigma of mixed mating systems in plants: occurrence,

- theoretical explanations, and empirical evidence. *Annu Rev Ecol Syst* **36**: 47–79.
- Hamrick JL, Godt MJW (1996). Effects of life history traits on genetic diversity in plant species. *Philos Trans: Biol Sci* **351**: 1291–1298.
- Hirao AS, Kameyama Y, Ohara M, Isagi Y, Kudo G (2006). Seasonal changes in pollinator activity influence pollen dispersal and seed production of the alpine shrub *Rhododendron aureum* (Ericaceae). *Mol Ecol* **15**: 1165–1173.
- Hirayama K, Ishida K, Setsuko S, Tomaru N (2007). Reduced seed production, inbreeding, and pollen shortage in a small population of a threatened tree, *Magnolia stellata*. *Biol Conserv* **136**: 315–323.
- Hirayama K, Ishida K, Tomaru N (2005). Effects of pollen shortage and self-pollination on seed production of an endangered tree, *Magnolia stellata*. *Ann Bot* **95**: 1009–1015.
- Isagi Y, Tateno R, Matsuki Y, Hirao A, Watanabe S, Shibata M (2007). Genetic and reproductive consequences of forest fragmentation for populations of *Magnolia obovata*. *Ecol Res* **22**: 382–389.
- Ishida K (2008). Effects of inbreeding on the magnitude of inbreeding depression in a highly self-fertilizing tree, *Magnolia obovata*. *Ecol Res* **23**: 995–1003.
- Jarne P, David P (2008). Quantifying inbreeding in natural populations of hermaphroditic organisms. *Heredity* **100**: 431–439.
- Karasawa MMG, Vencovsky R, Silva CM, Zucchi MI, Oliveira GCX, Veasey EA (2007). Mating system of Brazilian *Oryza glumaepatula* populations studied with microsatellite markers. *Ann Bot* **99**: 245–253.
- Kenta T, Isagi Y, Nakagawa M, Yamashita M, Nakashizuka T (2004). Variation in pollen dispersal between years with different population conditions in a tropical emergent tree. *Mol Ecol* **13**: 3575–3584.
- Lee SL (2000). Mating system parameters in tropical tree species, *Shorea leprosula* Miq. (Dipterocarpaceae), from Malaysian lowland Dipterocarp forest. *Biotropica* **32**: 693–702.
- Mable BK, Adam A (2007). Patterns of genetic diversity in outcrossing and selfing populations of *Arabidopsis lyrata*. *Mol Ecol* **16**: 3565–3580.
- Michalski SG, Durka W (2007). High selfing and high inbreeding depression in peripheral populations of *Juncus atratus*. *Mol Ecol* **16**: 4715–4727.
- Murray MG, Thompson WF (1980). Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Res* **8**: 4321–4325.
- Nassar JM, Ramirez N, Lampo M, Gonzalez JA, Casado R, Nava F (2007). Reproductive biology and mating system estimates of two Andean melocacti *Melocactus schatzlii* and *M. andinus* (Cactaceae). *Ann Bot* **99**: 29–38.
- Obayashi K, Tsumura Y, Ihara-Ujino T, Niiyama K, Tanouchi H, Suyama Y et al. (2002). Genetic diversity and outcrossing rate between undisturbed and selectively logged forests of *Shorea curtisii* (Dipterocarpaceae) using microsatellite DNA analysis. *Int J Plant Sci* **163**: 151–158.
- Osada N, Sugiura S, Kawamura K, Cho M, Takeda H (2003). Community-level flowering phenology and fruit set: comparative study of 25 woody species in a secondary forest in Japan. *Ecol Res* **18**: 711–723.
- Patterson B, Vaillancourt RE, Pilbeam DJ, Potts BM (2004). Factors affecting variation in outcrossing rate in *Eucalyptus globulus*. *Aust J Bot* **52**: 773–780.
- Petit RJ, Hampe A (2006). Some evolutionary consequences of being a tree. *Annu Rev Ecol Syst* **37**: 187–214.
- R Development Core Team (2008). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing: Vienna, Austria.
- Reusch TBH (2001). Fitness-consequences of geitonogamous selfing in a clonal marine angiosperm (*Zostera marina*). *J Evol Biol* **14**: 129–138.
- Ritland K (2002). Extensions of models for the estimation of mating systems using n independent loci. *Heredity* **88**: 221–228.
- Ritland K, Jain S (1981). A model for the estimation of outcrossing rate and gene frequencies using n independent loci. *Heredity* **47**: 35–52.
- Routley MB, Mavranganis K, Eckert CG (1999). Effect of population size on the mating system in a self-compatible, autogamous plant, *Aquilegia canadensis* (Ranunculaceae). *Heredity* **82**: 518–528.
- Saeda T, Nakashima M, Sakai Y (2004). Crossing and self-compatibility between *Magnolia stellata* and *M. salicifolia*. *Report of Gifu Prefectural Research Institute for Forests* **33**: 27–32 (in Japanese).
- Sampson J, Byrne M (2008). Outcrossing between an agroforestry plantation and remnant native populations of *Eucalyptus loxophleba*. *Mol Ecol* **17**: 2769–2781.
- Schemske DW, Lande R (1985). The evolution of self-fertilization and inbreeding depression in plants. II. Empirical observations. *Evolution* **39**: 41–52.
- Setsuko S, Ishida K, Ueno S, Tsumura Y, Tomaru N (2007). Population differentiation and gene flow within a metapopulation of a threatened tree, *Magnolia stellata* (Magnoliaceae). *Am J Bot* **94**: 128–136.
- Setsuko S, Tamaki I, Ishida K, Tomaru N (2008). Relationships between flowering phenology and female reproductive success in the Japanese species *Magnolia stellata*. *Botany* **86**: 248–258.
- Setsuko S, Ueno S, Tsumura Y, Tomaru N (2005). Development of microsatellite markers in *Magnolia stellata* (Magnoliaceae), a threatened Japanese tree. *Conserv Genet* **6**: 317–320.
- Shaw DV, Kahler AL, Allard RW (1981). A multilocus estimator of mating system parameters in plant-populations. *Proc Natl Acad Sci USA-Biol Sci* **78**: 1298–1302.
- Spiegelhalter DJ, Best NG, Carlin BP, Van der Linde A (2002). Bayesian measure of model complexity and fit. *J R Stat Soc Ser B* **64**: 583–616.
- Spiegelhalter DJ, Thomas A, Best N, Lunn D (2003). *WinBUGS. Ver1.4 User Manual*. MRC Biostatistics Unit: Cambridge, UK.
- Sturtz S, Ligges U, Gelman A (2005). R2WinBUGS: a package for running WinBUGS from R. *J Stat Softw* **12**: 1–16.
- Tamaki I, Setsuko S, Tomaru N (2008). Genetic variation and differentiation in populations of a threatened tree, *Magnolia stellata*: factors influencing the level of within-population genetic variation. *Heredity* **100**: 415–423.
- Ueda K (1988). Star magnolia (*Magnolia tomentosa*) an indigenous Japanese plant. *J Arnold Arbor* **69**: 281–288.
- Weir BS (1996). *Genetic Data Analysis II*. Sinauer Associates: Sunderland, MA, USA.
- Yates CJ, Coates DJ, Elliott C, Byrne M (2007). Composition of the pollinator community, pollination and the mating system for a shrub in fragments of species rich kwongan in south-west Western Australia. *Biodivers Conserv* **16**: 1379–1395.

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