



OPEN

Mating system and population structure in the natural distribution of *Toona ciliata* (Meliaceae) in South China

Wei Zhou^{1,2,3}, Xin-Xin Zhang^{1,2,3}, Ying Ren^{1,2}, Pei Li^{1,2}, Xiao-Yang Chen^{1,2}✉ & Xin-Sheng Hu^{1,2}

Most initially perfect flowers of *Toona ciliata* Roem subsequently develop into functionally unisexual flowers and their relative positions in the same inflorescence could enhance the outcrossing system in this species. Here we investigated the mating system of this species. We used eight nuclear microsatellite markers and investigated the progeny of 125 mother trees from six populations naturally distributed in South China, with sample sizes ranging from 64 to 300 seeds. The multilocus outcrossing rate was 0.970 ± 0.063 , and the single locus outcrossing rate was 0.859 ± 0.106 , indicating the pattern of predominant outcrossing. Selfing was present in one population, but biparental inbreeding occurred in five populations. Inbreeding was absent in maternal parents, and correlations of selfing among families or among loci were generally insignificant. Positive correlation of paternity at multiple loci was significant in four populations, but was not consistent with the results at single loci. Population substructure occurred in male similarity between outcrosses only in one population. Population genetic differentiation was significant ($F_{st} = 34.5\%$) and the effects of isolation-by-distance at the eight loci were significant among the six populations. These results provide evidence that self-compatibility and inbreeding naturally occur in *T. ciliata* and indicate that inbreeding avoidance is necessary during genetic improvement and breeding of this endangered tree species.

The flower structure of *Toona ciliata* Roem (Meliaceae) is characterized by flowers of about 5 mm long. There are short pedicels of 1–2 mm long, five short calyxes, five white and oblong petals each 4–5 mm long, five stamens that are about as long as the petals, filaments that are sparsely pilose, elliptic anthers, an ovary that is densely covered with long hard hairs and contains 8–10 ovules per locule, a glabrous style, and discoid stigmas with five fine lines¹. The flowers in the bud stage are perfect and hermaphrodite in the family of Meliaceae. Owing to the effects of unknown genetic or environmental factors, most floral development exhibits functionally male or female flowers although both anthers and pistils are present in each flower^{2,3}. Sterile ovaries in functionally male flowers are formed and would not develop further. Stamens with anthers in functionally female flowers seem not to produce fertile pollen. This is an alternative to those patterns that avoid selfing and inbreeding by forming sexually dioecious plants or employing self-incompatible genes. However, it is not clear whether such floral development is ubiquitous or not in genus *Toona* and other genera (e.g., *Cedrela* and *Swietenia*) of the Meliaceae², and the flowers that remain perfect cannot be excluded. The functionally male and female flowers are disproportionally mixed and distributed in different positions in the same inflorescence. The central flower in a cymule is female while the lateral flowers are male^{2,3}. Most individuals are sexually monoecious rather than dioecious. Thus, it is of interest to clarify the uncertainty of the mating system that *T. ciliata* possesses.

Preliminary inferences on the outcrossing rate of this species could be inferred from a recent study on population structure⁴. Based on the provenance populations of mixed seeds from multiple individuals, the estimate of inbreeding coefficients was generally negative ($F_{is} < 0$), implying an outcrossing system under an equilibrium between selfing and outcrossing where the outcrossing rate is $t = (1 - F_{is}) / (1 + F_{is})$ ⁵. However, more accurate estimate of the outcrossing rate or the selfing rate ($\alpha = 1 - t$) remains lacking based on the pedigree of family

¹College of Forestry and Landscape Architecture, South China Agricultural University, Guangdong 510642, China. ²Guangdong Key Laboratory for Innovative Development and Utilization of Forest Plant Germplasm, South China Agricultural University, Guangdong 510642, China. ³These authors contributed equally: Wei Zhou and Xin-Xin Zhang. ✉email: xychen@scau.edu.cn; xinsheng@scau.edu.cn



Figure 1. Locations of six populations of *Toona ciliata* for mating system analysis (Baoshan, Simao, Yongren, Tianlin, Guanshan and Nanping). The approximate natural geographical distribution of this species (in green) was plotted using Adobe Photoshop 2018 according to the distribution of 234 specimens available from <https://www.cvh.ac.cn/> and a reference to Chen et al.¹. The natural distribution of *T. ciliata* covers the regions of eastern, southern, central and southwest China.

structure, where the maximum likelihood estimate of outcrossing rate can be derived under disequilibrium between selfing and outcrossing⁶.

Scoring mating system of *T. ciliata* is of particular significance to its genetic conservation or sustainable management. This species is classified as an endangered species in China⁷ although it is not in the "Red lists" of the species at risk by the International Union for Conservation of Nature and Natural Resources (IUCN). The importance of this timber tree species is due to its straight stem form and quality red wood for furniture in China. However, the low natural regeneration and over-cutting enhanced a substantial decline in population density^{8,9}. This subsequently leads to the influences on mating system due to population demography and changes genetic variation within and between populations. So far, several programs have been set up to promote artificial plantations and genetic breeding, including provenance trials¹⁰ and the ongoing breeding program for insect-resistance. The critical issue is that the initial breeding programs were based on open-pollinated seeds derived from natural populations with unknown mating systems^{10,11}, which could yield biased estimates of genetic parameters and evaluation of progeny performance¹².

Here we investigated mating system of six natural populations distributed in South China. Previous studies with SRAP (sequence-related amplified polymorphism) and nuclear microsatellites (simple sequence repeats, SSRs) indicated that substantial genetic variation (about 79.26% with SRAP versus 32.4% with SSR) occurred among populations^{4,13}. The range-wide distribution of *T. ciliata* was grouped into two clusters, the western and eastern parts. To assess variation of mating system in different regions, six samples were collected from populations in both western and eastern parts of the species distribution (Fig. 1). We used nuclear SSRs because of the attributes of highly informative microsatellite markers for paternity/parentage and mating system analyses^{14,15}. SSR primers that were previously screened in *T. ciliata* were used for individual genotyping⁴. Besides, population genetic structure among the six populations was analyzed and then combined with the estimates of mating systems to inform strategies for managing this endangered forest species in China.

Results

Eight pairs of SSR primers were screened according to the most sequence variability observed in six populations (Table 1). Genotyping data at eight loci are shown in Supplementary Information Table 1. Analysis with Micro-Checker indicated that null alleles at locus TCR20 existed in all six populations, followed by locus TCR17 in three populations (Simao, Yongren and Tianlin), and loci TCR51 (Tianlin and Nanping), TCR83 (Simao and Yongren), and TCR81 (Yongren and Nanping) in two populations. Loci TCR26 (Guanshan) and TCR51 (Nanping) exhibited null alleles in one population. Three loci potentially had mis-scoring of stutters, including TCR18, TCR20 and TCR51 in Nanping (Table 2).

Locus	Forward/reverse primers (5'-3')	Repeat motifs	T_a (°C)	Allele sizes	Number of alleles
TCR17	F: GTGGCGTAACAGACCAAAAC R: CCAGAGATACTCCATCCAG	(GA) n	58	140–205	14
TCR18	F: GAAACCAGCAGGCAGAGC R: GAAGAAGGGTGAGCGAGA	(AG) n	58	107–210	44
TCR20	F: AAGCCAGTCAGCAACCTA R: GATTAAGTAATATTGGGTGGT	(GA) n	58	152–327	30
TCR26	F: ATGGATGAGTGTGCGATAGG R: TGTGATGTAGGAGTCTGAAC	(TC) n	58	233–271	30
TCR51	F: CAAAGACCAAGATTTGATGC R: ACTATGGGTGGCACAACTAC	(GA) n	56	108–140	5
TCR78	F: GATCTCACCCACTTGAAAAA R: GCTCATATTGAGAGGCATT	(GA) n (AG) n	58	163–255	24
TCR83	F: GAGATACAGTTGGTGGTTAGAGG R: TCTTCACCTGTTTGCCTCTC	(CT) n	58	183–287	47
TCR122	F: GTGCAGTGTCCATGTTGAAG R: GACATTTCTCGCAAGGTCA	(CA) n	56	158–298	14

Table 1. Primer sequences, repeat motifs, annealing temperature, allele sizes and the number of alleles at eight microsatellite loci for genotyping in six natural populations of *Toona ciliata*.

Populations	Latitude	Longitude	Altitude (m)	Samples	Stuttering loci	Large allele dropout	Loci with null alleles	Observed heterozygosity (H_o)	Expected heterozygosity with null alleles (H_e) ^a	Mean number of observed alleles per locus
Baoshan, Yunnan	N 24° 59'	E 99° 01'	1401	132 seeds (23 families, 4–6 seeds/family)	No	No	TCR20	0.505 ± 0.266	0.513 ± 0.219	12.750 ± 8.956
Simao, Yunnan	N 22° 46'	E 100° 58'	600	64 seeds (12 families, 4–6 seeds/family)	No	No	TCR17, TCR20, TCR83	0.227 ± 0.273	0.345 ± 0.249	5.500 ± 4.660
Yongren, Yunnan	N 26° 01'	E 101° 40'	1580	90 seeds (16 families, 4–6 seeds/family)	No	No	TCR17, TCR20, TCR83	0.266 ± 0.291	0.371 ± 0.245	7.000 ± 4.986
Tianlin, Guangxi	N 24° 17'	E 106° 13'	1200	119 seeds (20 families, 4–6 seeds/family)	No	No	TCR17, TCR20, TCR51	0.494 ± 0.310	0.603 ± 0.200	12.250 ± 9.285
Guanshan, Jiangxi	N 28° 32'	E 114° 33'	330	138 seeds (24 families, 4–6 seeds/family)	No	No	TCR18, TCR20, TCR26	0.332 ± 0.264	0.451 ± 0.281	7.250 ± 5.444
Nanping, Fujian	N 26° 38'	E 118° 10'	800	300 seeds (30 families, 10 seeds/family)	TCR18, TCR20, TCR51	No	TCR18, TCR20, TCR51, TCR78	0.256 ± 0.243	0.434 ± 0.286	2.875 ± 1.246

Table 2. Location, sample size, genotyping errors and genetic diversity in six natural populations of *Toona ciliata*. ^aThe expected heterozygosity was calculated according to the allele frequencies in SI Table 1 where null allele frequencies were estimated if present at a locus.

Allele frequencies were estimated using Genepop V4 for the loci without null alleles¹⁶, but with Micro-Checker (Oosterhout's method) for the loci with null alleles for each population (SI Table 2)¹⁷. Eight SSR loci exhibited a quite high level of variation among populations in heterozygosity and the number of alleles. Tianlin and Baoshan had the highest mean of observed heterozygosity over loci, followed by Guanshan, Yongren, Nanping and Simao. The expected heterozygosity was calculated according to the estimates of allele frequencies (SI Table 2). The expected heterozygosity (with the null alleles counted) was greater than the observed heterozygosity at each locus, but had the same ranking pattern as the observed heterozygosity among populations. This was slightly different from the pattern of the average number of observed alleles where Nanping (2.875 ± 1.246) and Simao (5.500 ± 4.660) had a fewer alleles (Table 2).

Analysis of population genetic differentiation with Eqs. (2) and (3) was summarized in Table 3. Generally, population genetic differentiation was significant at each locus except TCR51, ranging from $F_{st} = 0.005$ at locus TCR51 to 0.599 at locus TCR17. About 34.5% of SSR variation was accounted for by the differences among populations. Significant relationships between $F_{st}/(1 - F_{st})$ and the logarithmic geographical distance were present at all eight loci (Table 4). However, the analysis with multiple loci also showed a significant relationship between $F_{st}/(1 - F_{st})$ and the logarithmic geographical distance (p-value = 1.68×10^{-6} , Fig. 2), indicating the occurrence of isolation-by-distance (IBD) among the six natural populations. When the two populations from the eastern region (Guanshan and Nanping) were not included for analysis, IBD effects were insignificant (Fig. 2). Correlations between F_{st} and the altitude difference were not significant at both single and multiple loci (Table 3).

Locus	F_{st} (p-value)	a (p-value)	b (p-value)	R square	$r(F_{st}, \text{altitude})$
TCR17	0.599 (0.0000)	-4.202 (0.021)	3.155 (0.003)	0.647	0.370 (0.262)
TCR18	0.268 (0.0000)	-0.433 (0.010)	0.367 (8.81×10^{-5})	0.706	0.382 (0.160)
TCR20	0.302 (0.0000)	-0.403 (0.027)	0.394 (1.39×10^{-4})	0.685	0.347 (0.205)
TCR26	0.167 (0.0000)	-0.055 (0.358)	0.102 (0.002)	0.537	0.297 (0.283)
TCR51	0.005 (0.6215)	0.001 (0.767)	0.003 (0.015)	0.403	0.334 (0.244)
TCR78	0.622 (0.0000)	-4.635 (0.028)	3.497 (0.004)	0.620	0.404 (0.218)
TCR83	0.246 (0.0000)	-0.255 (0.047)	0.265 (2.44×10^{-4})	0.657	0.366 (0.179)
TCR122	0.207 (0.0000)	-0.090 (0.265)	0.141 (0.002)	0.550	0.287 (0.296)
Over loci	0.345 (0.0000)	-0.682 (0.0007)	0.576 (1.68×10^{-6})	0.826	0.349 (0.202)

Table 3. Population genetic differentiation (F_{st}), isolation-by-distance (IBD) and correlations between F_{st} and altitude differences between populations at eight polymorphic SSR loci in six populations IBD test was based on the regression of $F_{st}/(1 - F_{st}) = a + b \ln(\text{geographical distance})$, where a is the intercept and b is the regression coefficient.

Parameters	Baoshan	Simao	Yongren	Tianlin	Guanshan	Nanping
t_m	1.166 (0.067)	0.992 (0.088)	1.082 (0.089)	1.194 (0.031)	0.980 (0.225)	0.843 (0.033)
α_m	-0.166 (0.067)	0.008 (0.088)	-0.082 (0.090)	-0.194 (0.031)	0.020 (0.212)	0.157 (0.033)
t_s	0.992 (0.038)	0.844 (0.044)	0.899 (0.046)	0.897 (0.022)	0.850 (0.050)	0.672 (0.039)
α_s	0.008 (0.038)	0.156 (0.044)	0.100 (0.047)	0.103 (0.022)	0.150 (0.050)	0.338 (0.039)
$t_m - t_s$	0.174 (0.085)	0.148 (0.081)	0.183 (0.089)	0.297 (0.041)	0.136 (0.180)	0.171 (0.022)
F (maternal)	-0.187 (0.029)	-0.178 (0.052)	-0.141 (0.082)	-0.184 (0.031)	-0.027 (0.091)	-0.200 (0.001)
r_s	-0.765 (0.630)	-0.546 (0.674)	-0.292 (0.765)	-0.919 (0.391)	-0.261 (0.789)	0.016 (0.048)
$r_{p(m)}$	0.109 (0.037)	-0.133 (0.274)	0.080 (0.109)	0.208 (0.055)	0.304 (0.120)	0.312 (0.060)
$r_{p(s)}$	0.069 (0.054)	-0.083 (0.180)	0.116 (0.156)	0.281 (0.121)	0.350 (0.184)	0.452 (0.087)
$r_{p(s)} - r_{p(m)}$	-0.039 (0.045)	0.050 (0.282)	0.036 (0.157)	0.073 (0.104)	0.046 (0.187)	0.140 (0.043)
r_{loci}	0.293 (0.572)	-0.022 (0.238)	-0.054 (0.283)	-0.086 (0.107)	0.017 (0.172)	-0.007 (0.108)

Table 4. Estimates of parameters for mating system in six natural populations of *Toona ciliata*. t_m : the multilocus outcrossing rate; α_m : the multilocus selfing rate ($= 1 - t_m$); t_s : the single locus outcrossing rate; α_s : the single locus selfing rate; F : the single locus inbreeding coefficient of maternal parents; r_s : the correlation of selfing among families; r_p : the correlation of paternity among siblings; r_{loci} : the correlation of selfing among loci.

According to the instruction provided in MLTR¹⁸, null alleles at relevant loci in different populations were counted in estimating the outcrossing rate (or the selfing rate). Results showed that all populations exhibited a predominant outcrossing system (Table 4). Estimates of the multilocus outcrossing rates ranged from 0.843 (Nanping) to 1.194 (Tianlin), with a mean of 0.970 (0.063) over the six populations [t_m in Baoshan, Yongren and Tianlin was set as 1.0 in calculating the mean and standard deviation]. Estimates of the single outcrossing rates ranged from 0.672 (Nanping) to 0.992 (Baoshan), with the mean of 0.859 (0.106). From the relation of selfing and outcrossing rates, $\alpha = 1 - t$ and their equal variances between α and t , significant selfing rates were present in Nanping ($\alpha = 0.157$, Se = 0.033) but not in the remaining five populations for the multilocus estimates. However, significant selfing rates at single loci were present in all populations except Baoshan. Significant differences between t_m and t_s occurred in Nanping and Simao, indicating that biparental inbreeding in addition to the occurrence of selfing in these populations⁶. Inbreeding coefficient of maternal parents (F) was not significant from zero in Yongren and Guanshan but were significantly negative in the remaining four populations, indicating no inbreeding in female parents.

The correlation of selfing among families was not significant in all populations, indicating independence among families in selfing. The positive correlation of paternity at multiple loci was significant in Baoshan ($r_{p(m)} = 0.109$, Se = 0.037), Tianlin ($r_{p(m)} = 0.208$, Se = 0.055), Guanshan ($r_{p(m)} = 0.304$, Se = 0.120), and Nanping ($r_{p(m)} = 0.312$, Se = 0.06), indicating that the same fathers in pollination were partially shared in these populations but not in Yongren ($r_{p(m)} = 0.080$, Se = 0.109) and Simao ($r_{p(m)} = -0.133$, Se = 0.274). The results were different from single-locus estimates where a significant correlation of paternity was found only in Nanping ($r_{p(s)} = 0.452$, Se = 0.087). Single locus estimates of paternity were significantly greater than multilocus estimate only in Nanping ($r_{p(s)} - r_{p(m)} = 0.140$, Se = 0.043), indicating that population substructure occurred in male similarity between outcrosses in Nanping but not in other populations. Significant correlation of selfing among loci (r_{loci}) was not found in all populations, indicating the independence among the eight SSR loci in paternity correlation.

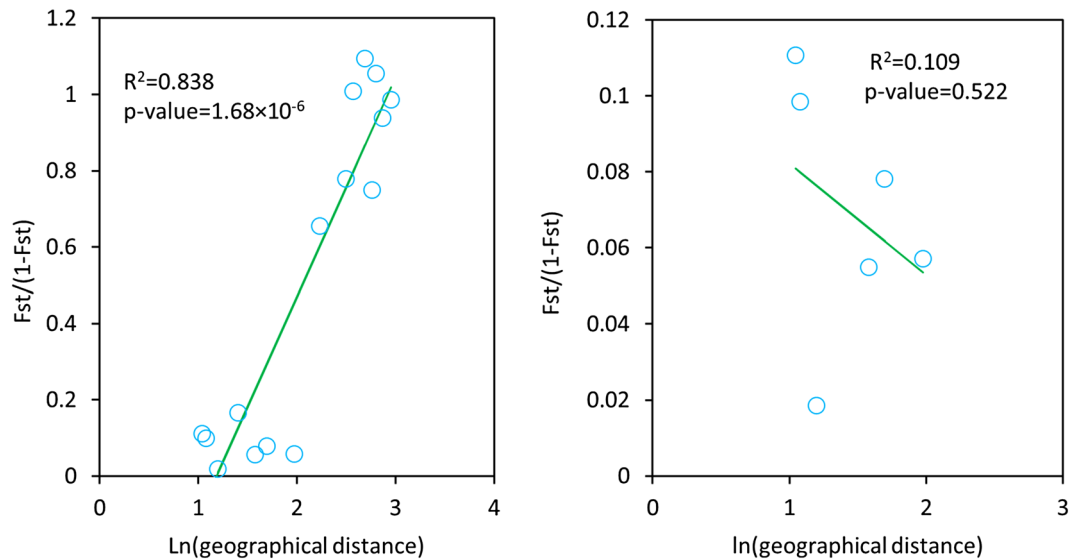


Figure 2. Test of isolation-by-distance (IBD) effects based on multilocus estimates of pairwise F_{st} 's. (A) Significant IBD effects existed among the six natural populations, $\frac{F_{st}}{1-F_{st}} = -0.752 + 0.7 \ln(\text{geographical distance})$; (B) Insignificant IBD effects occurred among four populations in the western region (without Guanshan and Nanping).

Discussion

Mating system. In this study, we confirmed that *T. ciliata* generally exhibited a predominant outcrossing system, which is consistent with most flower structure showing functionally unisexual flowers^{2,3}. However, we observed diverse rates of outcrossing in different regions of natural distribution of the species in South China. These estimates of multilocus outcrossing rates reflect the attributes of reproductive ecology of this species. Pollen dispersal could be driven mainly by insects², with an expectation of predominant outcrossing system. However, significant selfing for single locus estimates in Simao, Yongren, Guanshan and Nanping populations or for multilocus estimates in Nanping suggests that this species is self-compatible when there is limitation of pollen within and among populations (Table 4). This study provides evidence of occurrence of selfing and inbreeding despite the floral development that could facilitate cross-breeding. The mixed distribution of functionally male and female flowers on the same inflorescence enhances self-fertilization.

Estimates of paternity correlations indicated that a certain proportion of full-sibship in the same half-sib progeny occurred in four populations except Simao and Yongren that possessed multiple paternity for pollination. Further, Nanping and Guanshan populations exhibited population substructure in male similarity among outcrosses ($r_{p(s)} > r_{p(m)}$). This strongly suggests the occurrence of limited pollen dispersal both within and among populations in different regions of the species distribution. Although maternal parents showed lack of inbreeding ($F = 0$), the number of paternal parents was limited in four out of six populations investigated. This could be partially attributed to the limited numbers of pollen donors within or between populations or to the over exploitation that led to low population density. Reports of low natural regeneration was probably associated with the low quality of seeds due to inbreeding depression⁹.

Biparental inbreeding inferred from a comparison of t_m and t_s suggests the presence of cross-fertilization between genetically close relatives in six populations. This may occur in structured populations or in populations of low density where inbreeding takes place within populations. Restriction of pollen flow in *T. ciliata* reinforces biparental inbreeding. Further, distinct pollen pools in allele frequency existed among maternal parents and the progeny so produced came from a few individuals, which could also be inferred from the inverse values of estimates ($1/r_{p(m)}$) that ranged from 3 to 12 individuals¹⁹.

The mating system of *T. ciliata* is essentially different from *T. sinensis*, one genetically related species in monophyly genus *Toona* in Meliaceae^{1,20}. The differences between *T. ciliata* and *T. sinensis* are subtle in flower morphological traits. The flower structure of *T. sinensis* is characterized by flowers of 4–5 mm long, short pedicels, 5-dentate or shallowly undulate calyxes, five white, oblong and glabrous petals each 4–5 mm long and 2–3 mm wide, ten stamens (five fertile and five degenerate), eight ovules per locule, the style longer than the ovary, and the discoid stigmas with five fine lines. *T. sinensis* is sexually diecious and has the system of complete outcrossing. *T. ciliata* retains perfect flowers, and its mating system is evidently predominant outcrossing from this study. Their qualitative difference in mating system is likely related to the speciation time between *T. ciliata* and *T. sinensis*^{21,22}, and distinct evolutionary history modulates each species (inbreeding depression vs. sexual polymorphism).

Sample sizes and the number of loci. Sampling of seeds is difficult in this species because most mature trees are tall. Rate of seed germination was low in either our green house (South China Agricultural University, Guangzhou) or in the field²³, resulting in different sizes of progeny populations available for analysis. Although

the sample sizes per progeny (4–10 seeds per family) were relatively small in comparison with some reports in the literature, these samples were appropriate for scoring this predominant outcrossing system. These are based on two reasons. One is that microsatellite markers are highly polymorphic and the number of eight loci is sufficient to identify cross individuals. The maternal genotypes can be directly observed (see “Supplementary materials” for genotyping data) without sophisticated statistical inferences.

The second reason is based on a theoretical study on the array of open-pollinated progeny genotypes descended from maternal parents of unknown genotypes. Simulation studies showed that the optimal number of progeny per family was from 4 to 8 zygotes for estimating fixation index and selfing parameters²⁴. This range of sample sizes provides appropriate estimates of variances. The sampling errors are calculated from the binomial distribution by $\sqrt{\hat{t}(1-\hat{t})/n}$ ¹⁸. Our estimates of standard errors for the outcrossing rate or the selfing rate were less than 10% except Guanshan for the estimate of multilocus outcrossing rate (Table 4).

The number of eight SSR loci was sufficient to identify heterozygotes. This can be viewed from a simple inference. Suppose that all alleles at a locus follow uniform distribution. According to the number of observed alleles at each locus, the joint probability that two individuals had the exact same genotypes at eight loci was infinitely small by chance ($= \frac{1}{14} \cdot \frac{1}{44} \cdot \frac{1}{30} \cdot \frac{1}{30} \cdot \frac{1}{5} \cdot \frac{1}{24} \cdot \frac{1}{47} \cdot \frac{1}{14}$). Although this probability varied with populations, the power for fingerprinting individuals or identifying heterozygotes was large with the eight polymorphic loci.

Effects of IBD on mating system. Mating system of a natural plant population can be shaped by many factors, including local population density, pollen dispersal within and between populations, plant height, phenological period, and global changes (rapid changes in climate, habitat fragmentation and land usage)^{25,26}. With a reference to the information derived from the present analysis of population genetic structure, we here discuss the effects of IBD and gene flow on mating system. Significant IBD effects could be consistent with the entomophilous feature of the species. The relationship between IBD effects within a continuous population and mating system has been long appreciated in theory²⁷, and was recently assessed with genome-wide single nucleotide polymorphisms²⁸. A relatively higher level of selfing and inbreeding occurred in the eastern part of species distribution, while a lower level of selfing or outcrossing was present in the western part. Pollen flow of *T. ciliata* could be mainly entomophilous, and seed dispersal is mainly mediated by wind together with birds⁹, with an anticipation of a certain level of population genetic differentiation. Previous studies demonstrated that population differentiation of *T. ciliata* was substantially large, with $F_{st}=0.792$ among range-wide populations assayed with nuclear SRAP¹³ and 0.324 with nuclear SSR loci among 29 populations⁴. The present study also indicated significant genetic differentiation ($F_{st}=0.345$) among the six populations, and detected significant IBD effects among the six populations. Zhan et al.⁴ showed that IBD effects were absent among populations in the western part but present in the eastern part. IBD effects among populations could likely affect mating system since mating systems differed in different regions under significant IBD effects²⁹.

The relationship between IBD between populations and mating system could be complicated^{30,31}. From Zhang et al.³¹, an analytical relationship among IBD effects, gene flow and mating system can be roughly inferred by synthesizing the results of Rousset³² ($F_{st}/(1-F_{st})=a+b \ln(\text{geographical distance})$), Caballero and Hill³³, and Hu and Ennos³⁴ under classical island model³⁵, i.e.

$$\frac{1}{N(1-\alpha/2)(m_s + (1-\alpha)m_p/2)} = a + b \ln(\text{geographical distance}) \quad (1)$$

where $N(1-\alpha/2)$ is the effective population size (N_e) in which N is the population size, α is the selfing rate, and m_s and m_p are the migration rates of seeds and pollen, respectively; a is the intercept of regression analysis, and b is the regression coefficient. A positive correlation is possible between selfing rate (α) and geographical distance under significant IBD ($b > 0$) effects.

Thus, two processes could relate IBD effects to mating system. One is the demographic change owing to the reduction of effective population size by selfing/inbreeding, $N(1-\alpha/2)$. This change comes from alternation of gene flow within rather than between populations. IBD at the population level reinforces this process by increasing the chance of interbreeding of relatives, which indirectly decreases genetic diversity within populations (e.g., a high level of inbreeding in Nanping population). Reduction in effective population size was reported in an isolated population of *Myracrodruon urundeuva* although pollen dispersal of this species is mediated by wind and bees³⁶. The second process is that IBD impedes both seed and pollen flow between rather than within populations, $m_s + (1-\alpha)m_p/2$. Selfing only interacts with pollen flow that is restricted by IBD through decreasing population connectivity or increasing physical distances due to clear-cutting. Empirical studies support the impediment to gene flow by habitat fragmentation³⁷ or IBD at the population level, with a few exceptions of case studies^{38,39}. Taken together, the interaction of the two processes could reinforce the effects of IBD on mating system.

Concerned with our results (Table 4), the relatively higher rate of selfing or inbreeding in Nanping ($\alpha_m > 0$ and $t_m - t_s > 0$) and Guanshan ($t_m - t_s > 0$) could be associated with Wuyi (2158 m above sea level at the highest peak) and Xuefeng (a part of Yunnan-Guizhou Plateau) Mountains that naturally impeded seed and pollen flow between western and eastern parts of the species distribution (Fig. 1). Parents were mainly confined to local neighbor relatives. Seeds of *T. ciliata* are mainly dispersed by gravity and less often by rodents and birds⁹, and natural regeneration is derived from seeds distributed around mother trees, increasing the rate of both selfing and biparental inbreeding. The average germination rate was about 38% for families²³, and might reflect inbreeding depression in this species⁸.

Management implications. Practices of forest management frequently affect population genetic structure within and/or between populations. For instance, selective logging, such as removal of individuals in adults or

in the reproductive stage, results in low population density⁴⁰. Other practices of forest management, such as clear-cutting and habitat fragmentation, more or less create population structure⁴¹. This was the same situation in *T. ciliata* in China where adult trees of straight stem were cut for furniture and natural regeneration by seedlings was low⁸. Carneiro et al.¹⁴ showed that selective logging in a population of *Hymenaea courbaril* reduced the migration (m_p) of pollen from outside populations but increased pollen or seed dispersal distance within the population. They also showed that selective logging reduced the reproductive population size, and increased bottleneck effects (related to N_e) and the rate of selfing or inbreeding among relatives. Conclusions similar to Carneiro et al.¹⁴ were also drawn in *Baillonella toxisperma*⁴² and other species⁴¹. Thus, population genetic structure could be affected by selective logging in natural forests of *T. ciliata*^{8,9}.

The existence of self-fertilization and inbreeding in *T. ciliata* could produce substantial population genetic differentiation. Management, such as selective logging and clear-cutting followed by natural ecological restoration could yield interbreeding among relatives^{14,43}. These practices of conventional forest management erode genetic diversity and destroy habitats. Barriers to the inter-population gene flow enhance population genetic differentiation and reduce genetic variation within populations, leading to high homozygosity in local habitats. This may temporally improve adaptation of subdivided populations to local habitats (high frequency of adaptive homozygotes), but reduce the potential for long-term evolutionary response to globally changing environments. Further, when locally adapted populations hybridize, progeny so produced becomes less fit to local habitats, analogous to the outbreeding depression. Therefore, silvicultural management, such as assisted migration of seeds from distinct provenances^{44,45} and maintenance of non-isolated populations in multiple preserved areas, is the key to the viability of *T. ciliata* in China. Since *T. ciliata* is evidently self-compatible, practices that promote extensive pollen dispersal within and between populations, such as improving population density by planting seedlings in logging gaps^{42,46} and/or plantation with genotypes differing from neighborhood trees, can enhance outcrossing and reduce inbreeding depression.

Conclusions

T. ciliata is an important timber species that has been over exploited owing to its high quality of wood for furniture in China. Scoring its mating system helps to understand the reproductive ecology of this species and hence to imply its natural regeneration. This study investigated mating system of six natural populations in different regions and showed occurrence of selfing or inbreeding in different geographical regions. *T. ciliata* exhibited a predominant outcrossing, but allowed selfing and inbreeding. Analysis of population genetic structure showed that substantial genetic variation existed among the six populations. Isolation-by-distance effects were significant among the six populations but insignificant among the four populations in the western part of the distribution of this species. The results suggest that one practice of management is to remove barriers to gene flow in this species, including assisted migration and planting seedlings in logging gaps. The results also indicate the needs of avoiding inbreeding in tree breeding programmes.

Materials and methods

Plant material. Six natural populations were selected in 2017 and 2018 (Fig. 1). Three populations were in Yunnan Province (Baoshan, Simao and Yongren), one population in Guangxi Province (Tianlin), one in Jiangxi Province (Guanshan), and one in Fujian Province (Nanping). Samples from Yunnan Province were collected at the altitude of 600–1580 m above sea level in the western part of the species distribution. Samples from Tianlin was collected at 1200 m above sea level in the central region. Samples in Guanshan and Nanping were collected respectively at the altitude of 300 m and 800 m in the eastern part (Table 2). Note that these six populations were resampled and different from those populations published by Zhan et al.⁴ where seed samples were mixed and the pedigree information was not recorded.

Effective sample sizes ranged from 64 in Simao (12 open-pollinated families) to 300 seeds in Nanping (30 open-pollinated families), and an average of 140 seedlings per population (4–10 seeds per family) were germinated for mating system analysis. Note that the sample sizes were appropriate when highly polymorphic markers were employed (say, 50 seeds)^{2,23} and maternal genotypes were observable from progeny genotypes. Seeds were sowed in small containers with mixture of matrix and perlite sand in October of 2017 and 2018 in green house at South China Agricultural University, Guangzhou. Seedlings were grown for 90 days to a height of 2–3 cm.

DNA extraction and SSR-PCR amplification. DNA extraction was conducted by using 100 mg fresh leaves and following manufacturer's instructions of the DNAsure Plant Kit (TIANGEN Biotech Beijing Co., Ltd.). Agarose gel electrophoresis (1.0% w/v) was applied to check the quality of DNA extraction. The DNA concentration was checked with a Thermo Scientific NanoDrop 1000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA), and adjusted to 50 ng μL^{-1} . Quantified DNA samples were stored at -20°C for PCR amplification.

Selection of SSR markers was based on a previous study of *T. ciliata*⁴. Eight pairs of SSR primers were screened according to the most sequence variability observed in six populations. These SSR markers are nuclear genomes-based, and the resultant amplicons follow biparental Mendelian inheritance. Table 1 shows the SSR primer sequences, repeat motifs, annealing temperature (T_a), and the observed allele sizes for each SSR locus.

The same procedure as our previous study on *Machilus pauhoi*⁴⁷ was employed. The PCR amplification was carried out in a 25 μL reaction volume containing 1U Taq DNA polymerase (TaKaRa), 1 μL 20 ng μL^{-1} template DNA, 2 μL 25 mmol L^{-1} MgCl₂, 0.5 μL 10 mmol L^{-1} dNTP, 0.5 μL 10 mmol L^{-1} of each forward and reverse fluorescent primer pair, 2.5 μL 10 \times PCR buffer (without MgCl₂, 100 mmol L^{-1} Tris-HCl, pH 8.8 at 25 $^\circ\text{C}$, 500 mmol L^{-1} KCl). PCR amplification was conducted using following cycle procedure: an initial 3 min of denaturing at 95 $^\circ\text{C}$, followed by ten cycles of three steps (30 s of denaturing at 95 $^\circ\text{C}$, 30 s of annealing at 60 $^\circ\text{C}$, and 30 s of elongation at 72 $^\circ\text{C}$). The next 20 cycles were then used with the same steps (30 s of denaturing at 95 $^\circ\text{C}$, 30 s of annealing

at 55 °C, and 30 s of elongation at 72 °C). A final elongation step was at 72 °C for 6 min. The PCR products were detected by capillary electrophoresis using an ABI 3730XL DNA analyzer after confirmation of PCR amplification on a 2% agarose gel. Genemapper 4.0 software (Applied Biosystems) was used to estimate molecular sizes of SSR-PCR amplifications.

Data analysis. To check genotyping errors, we applied Micro-Checker¹⁷ to discriminate errors arising from short allele dominance (large allele dropout), stuttering and null alleles. A pooled population consisting of all families was used to check genotyping errors. In this program, presence of null alleles was suggested when homozygotes were overall excess and evenly distributed across the homozygote-classes. Mis-scoring of stutters was suggested when heterozygotes were deficit with alleles differing in size by a single repeat and large homozygotes were in excess. Short allele dominance was suggested when homozygotes were biased towards either extreme of the allele size-distribution. Null allele frequencies were estimated using the method of Oosterrhout et al.¹⁷.

Let p_{ijk} be the frequency of allele k of locus j at population i . The expected heterozygosity (H_e) at a locus in a population was calculated by $1 - \sum_k p_{ijk}^2$, where the null allele frequency was included. Estimates of allele frequencies were shown in SI Table 2. Following Wright³⁵ and Weir⁴⁸, population genetic differentiation at a single locus was calculated by

$$F_{st(s)} = \frac{\sum_i \sum_k n_i (p_{ijk} - \bar{p}_{jk})^2 / (r - 1) \bar{n}}{\sum_k \bar{p}_{jk} (1 - \bar{p}_{jk})}, \quad (2)$$

where \bar{p}_{jk} be the average frequency of allele k at locus j over r populations, and $\bar{n} = \sum_i n_i / r$. Test of the null hypothesis $H_0: F_{st(s)} = 0$ can be approximated by Chi-square statistic, $\chi^2 = (r - 1) \bar{n} F_{st(s)}$, with $df = r - 1$ ⁴⁸. For multilocus estimate,

$$F_{st(m)} = \frac{\sum_i \sum_j \sum_k n_i (p_{ijk} - \bar{p}_{ijk})^2 / (r - 1) \bar{n}}{\sum_j \sum_k \bar{p}_{jk} (1 - \bar{p}_{jk})}. \quad (3)$$

A program in R was provided to calculate population genetic differentiation based on allele frequencies (“SI Program”). Since geographical coordinates of mother trees were not recorded in natural populations, further estimation of contemporary pollen dispersal was not analyzed using POLDISP 1.0⁴⁹.

MLTR was used to handle the presence of null alleles with SSR markers^{6,24} where each allele status at each locus (null allele present with code 1, or absent with code 0) was indicated in preparing dataset. Estimates (expectation-maximization method) of mating system parameters included the multilocus population outcrossing rate (t_m), the average single-locus population outcrossing rate (t_s), the inbreeding coefficient of maternal parents (F), the correlation of paternity at multilocus ($r_{p(m)}$) and single locus ($r_{p(s)}$), the correlated selfing rate (r_s), the difference between t_m and t_s , and the difference between $r_{p(s)}$ and $r_{p(m)}$. Because estimates of allele frequencies between pollen and ovules were heterogenous in some half-sib families, their inequality was set in running the program (not pooled), which yielded greater maximum likelihoods with our datasets. Standard errors of population estimates were calculated using 1000 bootstraps that were based on resampling of progeny array.

Received: 24 November 2019; Accepted: 23 September 2020

Published online: 12 October 2020

References

- Chen, S. K., Li, H. & Chen, B. Y. *Meliaceae in Flora of China* Vol. 43, 37–104 (Science Press, Beijing, 1997) (in Chinese).
- Styles, B. T. The flower biology of Meliaceae and its bearing on tree breeding. *Silvae Genet.* **21**, 175–182 (1972).
- Gouvea, C. F., Dornelas, M. C. & Rodriguez, A. P. M. Floral development in the tribe Cedreleae (Meliaceae, sub-family Swietenioideae): *Cedreleae* and *Toona*. *Ann. Bot.* **101**, 39–48. <https://doi.org/10.1093/aob/mcm279> (2008).
- Zhan, X. et al. Genetic diversity and population structure of *Toona ciliata* revealed by simple sequence repeat markers. *Biotechnol. Biotech. Eq.* **5**, 1–9. <https://doi.org/10.1080/13102818.2018.1561210> (2019).
- Fyfe, J. L. & Bailey, N. T. J. Plant breeding studies in leguminous forage crops. I. Natural cross-breeding in winter beans. *J. Agric. Sci.* **41**, 371–378. <https://doi.org/10.1017/S0021859600049558> (1951).
- Ritland, K. Extensions of models for the estimation of mating systems using n independent loci. *Heredity* **88**, 221–228. <https://doi.org/10.1038/sj.hdy.6800029> (2002).
- Fu, L. G. *Red Paper on Chinese Plants—Rare And Endangered Plants* Vol. 1, 458–459 (Science Press, Beijing, 1992) (in Chinese).
- Liang, R. L., Liao, R. Y. & Dai, J. Endangered causes and protection strategy of *Toona ciliata*. *Guangxi Forest Sci.* **40**, 201–203. <https://doi.org/10.19692/j.cnki.gfs.2011.03.011> (2011) (in Chinese).
- Huang, H. L., Zhang, L. & Liao, C. K. Seed rain, soil seed bank, and natural regeneration of natural *Toona ciliata* var. *pubescens* forest. *Chin. J. Appl. Ecol.* **23**, 972–978. <https://doi.org/10.13287/j.1001-9332.2012.0133> (2012) (in Chinese).
- Li, P., Que, Q. M., Wu, L. Y., Zhu, Q. & Chen, X. Y. Growth rhythms of *Toona ciliata* seedlings from different provenances. *J. South China Agric. Univ.* **38**, 96–102. <https://doi.org/10.7671/j.issn.1001-411x.2017.01.016> (2017) (in Chinese).
- Xu, X. H., Peng, C., Zhang, Y. D., Wang, R. W. & Chen, H. L. Seedling growing traits of different provenances of *Toona ciliata* in Wuhan. *Chin. Agric. Sci. Bull.* **33**, 54–58 (2017).
- Mandal, A. K. & Ennos, R. A. Mating system analysis in a natural population of *Acacia nilotica* subspecies *kraussiana*. *Forest Ecol. Manag.* **79**, 235–240. <https://doi.org/10.1007/BF00224520> (1995).

13. Li, P. *et al.* Genetic diversity and population structure of *Toona ciliata* Roem. based on sequence-related amplified polymorphism (SRAP) markers. *Forests* **6**(12), 1094–1106. <https://doi.org/10.3390/f6041094> (2015).
14. Carneiro, F. S. *et al.* Effects of selective logging on the mating system and pollen dispersal of *Hymenaea courbaril* L. (Leguminosae) in the Eastern Brazilian Amazon as revealed by microsatellite analysis. *Forest Ecol. Manag.* **262**, 1758–1765. <https://doi.org/10.1016/j.foreco.2011.07.023> (2011).
15. Vinson, C. C., Mangaravite, E., Sebbenn, A. M. & Lander, T. A. Using molecular markers to investigate genetic diversity, mating system and gene flow of Neotropical trees. *Braz. J. Bot.* **41**, 481–496. <https://doi.org/10.1007/s40415-018-0472-x> (2018).
16. Raymond, M. & Rousset, F. GENEPOP (version 1.2): Population genetics software for exact tests and ecumenicism. *J. Heredity* **86**, 248–249. <https://doi.org/10.1093/oxfordjournals.jhered.a111573> (1995).
17. Oosterhout, C. V., Hutchinson, W. F., Wills, D. P. M. & Shipley, P. MICRO-CHECKER: Software for identifying and correcting genotyping errors in microsatellite data. *Mol. Ecol. Notes* **4**, 535–538. <https://doi.org/10.1111/j.1471-8286.2004.00684.x> (2004).
18. Ritland, K. Multilocus mating system program: MLTR. <http://kermitzii.com/softwares/>. Accessed 29 Sep 2020.
19. Sun, M. & Ritland, K. Mating system of yellow starthistle (*Centaurea solstitialis*), a successful colonizer in North America. *Heredity* **80**, 225–232. <https://doi.org/10.1046/j.1365-2540.1998.00290.x> (1998).
20. Sun, Y. L. *et al.* Genetic diversity of *Toona sinensis* in Korea and the phylogenetic relationship of this species based on chloroplast DNA and ribosomal DNA sequences. *Res. J. Biotechnol.* **9**, 1–14 (2014).
21. Charlesworth, D. Evolution of plant breeding systems. *Curr. Biol.* **16**, R726–R735. <https://doi.org/10.1016/j.cub.2006.07.068> (2006).
22. Barrett, S. C. H. Evolution of mating systems: outcrossing versus selfing. In *The Princeton Guide to Evolution* (ed. Losos, J.) 356–362 (Princeton Univ. Press, Princeton, 2014).
23. Wang, R. W., Huang, G. W., Chen, H. L. & Zhang, X. Y. Seed germination rate and seeding characteristics of *Toona ciliata* in Enshi. *Chin. Agric. Sci. Bull.* **34**, 39–43 (2018).
24. Ritland, K. Joint maximum likelihood estimation of genetic and mating structure using open-pollinated progenies. *Biometrics* **42**, 25–43. <https://doi.org/10.2307/2531241> (1986).
25. Eckert, C. G. *et al.* Plant mating systems in a changing world. *Trends Ecol. Evol.* **25**, 35–43. <https://doi.org/10.1016/j.tree.2009.06.013> (2010).
26. Levin, D. A. Mating system shifts on trailing edge. *Ann. Bot.* **109**, 613–620. <https://doi.org/10.1093/aob/mcr159> (2012).
27. Wright, S. Isolation by distance under diverse systems of mating. *Genetics* **31**, 39–59 (1946).
28. Aguillon, S. M., Fitzpatrick, J. W., Bowman, R., Schoech, S. J. & Chen, N. Deconstructing isolation-by-distance: The genomic consequences of limited dispersal. *PLoS Genet.* **13**, e1006911. <https://doi.org/10.1371/journal.pgen.1006911> (2017).
29. Ennos, R. A. & Clegg, M. T. Effect of population substructure on estimates of outcrossing rate in plant populations. *Heredity* **48**, 283–292. <https://doi.org/10.1038/hdy.1982.33> (1982).
30. Gaiotto, F. A., Grattapaglia, D. & Vencovsky, R. Genetic structure, mating system, and long-distance gene flow in heart of palm (*Euterpe edulis* Mart.). *J. Heredity* **94**, 399–406. <https://doi.org/10.1093/jhered/esg087> (2003).
31. Zhang, X. X. *et al.* Advances in the study of population genetic diversity at plant species' margins. *Chin. J. Plant Ecol.* **43**, 383–395. <https://doi.org/10.17521/cjpe.2018.0252> (2019) (in Chinese).
32. Rousset, F. Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics* **145**, 1219–1228. <https://doi.org/10.1017/S0016672397002656> (1997).
33. Caballero, A. & Hill, W. G. Effective size of non-random mating populations. *Genetics* **130**, 909–916 (1992).
34. Hu, X. S. & Ennos, R. A. Impacts of seed and pollen flow on population genetic structure for plant genomes with three contrasting modes of inheritance. *Genetics* **152**, 441–450. <https://doi.org/10.1046/j.1365-2443.1999.00261.x> (1999).
35. Wright, S. *Evolution And The Genetics of Populations: The Theory of Gene Frequencies*. Vol. 2, 291–293 (University Chicago Press, Chicago, 1969).
36. Gaino, A. P. S. C. *et al.* Understanding the effect of isolation on seed and pollen flow, spatial genetic structure and effective population size of the dioecious tropical tree species *Myracrodruon urundeuva*. *Conserv. Genet.* **11**, 1631–1643. <https://doi.org/10.1007/s10592-010-0046-3> (2010).
37. Sebbenn, A. M. *et al.* Low levels of realized seed and pollen gene flow and strong spatial genetic structure in a small, isolated and fragmented population of the tropical tree *Copaifera langsdorffii* Desf. *Heredity* **106**, 134–145. <https://doi.org/10.1038/hdy.2010.33> (2011).
38. White, G. M., Boshier, D. H. & Powell, W. Increased pollen flow counteracts fragmentation in a tropical dry forest: An example from *Swietenia humilis* Zuccarini. *Proc. Natl. Acad. Sci. USA* **90**, 2038–2042. <https://doi.org/10.1073/pnas.042649999> (2002).
39. O'Connell, L. M., Mosseler, A. & Rajora, O. P. Impacts of forest fragmentation on the mating system and genetic diversity of white spruce (*Picea glauca*) at the landscape level. *Heredity* **97**, 418–426. <https://doi.org/10.1038/sj.hdy.6800886> (2006).
40. Medjibe, V. P., Putz, F. E., Starkey, M. P., Ndouna, A. A. & Memiaghe, H. R. Impacts of selective logging on above-ground forest biomass in the Monts de Cristal in Gabon. *Forest Ecol. Manag.* **262**, 1799–1806. <https://doi.org/10.1016/j.foreco.2011.07.014> (2011).
41. Ratnam, W. *et al.* Genetic effects of forest management practices: Global synthesis and perspectives. *Forest Ecol. Manag.* **333**, 52–65. <https://doi.org/10.1016/j.foreco.2014.06.008> (2014).
42. Duminil, J. *et al.* High selfing rate, limited pollen dispersal and inbreeding depression in the emblematic African rain forest tree *aillonella toxisperma*—Management implications. *Forest Ecol. Manag.* **379**, 20–29. <https://doi.org/10.1016/j.foreco.2016.08.003> (2016).
43. Donna, B. & Krauss, S. L. Limited impact of fragmentation and disturbance on the mating system of tuart (*Eucalyptus gomphocephala*, Myrtaceae): Implications for seed-source quality in ecological restoration. *Aust. J. Bot.* **61**(2), 148–160. <https://doi.org/10.1071/BT13032> (2013).
44. Smulders, M. J. M., Cobben, M. M. P., Arens, P. & Verboom, J. Landscape genetics of fragmented forests: Anticipating climate change by facilitating migration. *Iforest* **2**, 128–132. <https://doi.org/10.3832/ifor0505-002> (2009).
45. Krutovsky, K. V. *et al.* Gene flow, spatial structure, local adaptation, and assisted migration in trees. In *Genomics of Tree Crops* (eds Schnell, R. & Priyadarshan, P.) 71–116 (Springer, New York, 2012).
46. Doucet, J. L., Kouadio, Y. L., Monticelli, D. & Lejeune, P. Enrichment of logging gaps with moabi (*Baillonella toxisperma* Pierre) in a Central African rain forest. *Forest Ecol. Manag.* **258**, 2407–2415. <https://doi.org/10.1016/j.foreco.2009.08.018> (2009).
47. Zhu, Q. *et al.* Phylogeographic pattern suggests a major eastward dispersal in the distribution of *Machilus pauhoi* in South China. *PLoS ONE* **12**, e0184456. <https://doi.org/10.1371/journal.pone.0184456> (2017).
48. Weir, B. S. *Genetic Data Analysis II: Methods for Discrete Population Genetic Data* (Sinauer Associates, Inc. Publishers, Sunderland, 1996).
49. Robledo-Arnuncio, J. J., Austerlitz, F. & Smouse, P. E. POLDISP: A software package for indirect estimation of contemporary pollen dispersal. *Mol. Ecol. Notes* **7**, 763–766. <https://doi.org/10.1111/j.1471-8286.2007.01706.x> (2007).

Acknowledgements

We sincerely appreciate editors and two anonymous reviewers for comments that significantly improved the presentation of this article. We are grateful to Richard A Ennos for proof reading this manuscript. The authors acknowledge the funding by Science and Technology Department of Guangdong Province (2017B020201008), National Key Research Projects, Forestry Resources Cultivation and Utilization Technology Innovation

(2016YFD0600606), People's Republic of China, and the funding from South China Agricultural University (4400-K16013).

Author contributions

X.S.H. and X.Y.C. conceived and designed the study; W.Z. and X.X.Z. conducted the experiments and drafted the manuscript; Y.R. and P.L. participated in sampling; X.S.H. analyzed data and finalized the manuscript. All authors read and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-020-74123-8>.

Correspondence and requests for materials should be addressed to X.-Y.C. or X.-S.H.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020