Popul Ecol
DOI 10.1007/s10144-012-0332-7

ORIGINAL ARTICLE

Leading-edge populations do not show low genetic diversity or high differentiation in a wind-pollinated tree

Miao-Miao Shi · Xiao-Yong Chen

Received: 14 June 2011 / Accepted: 6 May 2012

Abstract Climate changes can shift species' ranges. Knowledge on genetic variation of the leading-edge populations provides critical information to understand responses and adaptation of plants to projected climate warming. To date, the research into genetic variation of leading-edge populations has been limited, particularly in the role of wind-mediated pollen flow in maintaining high genetic variation. *Castanopsis sclerophylla* (Fagaceae) is a wind-pollinated and gravity-dispersed tree. In the present study, we used seven polymorphic microsatellites to genotype 482 samples from five leading-edge and 12 non-edge populations. Significant effects of recent population bottleneck events were found in three of the five leading-edge populations, indicating that the leading-edge populations might have been recolonized after the Last Glacial Maximum. Genetic diversity was higher, though not significantly, in leading-edge than in non-edge populations. Relationship between genetic diversity and latitude indicated an increasing trend of genetic diversity towards leading-edge populations. No significant difference in genetic differentiation was found between leading-edge and non-edge populations. The inconsistence with the general predictions by leading-edge colonization model could be explained by high gene flow via pollen grains. Pollen-mediated gene flow could maintain high genetic diversity within and low differentiation among leading-edge populations. In response to climate warming, high genetic variation may provide leading-edge populations raw materials for evolutionary adaptation to future environmental conditions.

Keywords Bottleneck · *Castanopsis sclerophylla* · Gene flow · Genetic variation · Wind-direction

Introduction

It is well known that the Pleistocene climatic oscillations repeatedly influenced species' ranges (Hewitt 2000). During the glacial ages, species' range contracted towards the Equator, and recolonized polewards during the interglacial ages. Following the climate change, two types of edges of species' range can be identified: the rear edge, i.e., eroding range margin, and the leading edge, i.e., colonizing front (Hampe and Petit 2005). Rear edge populations are old and may maintain high genetic diversity, which will be gradually lost due to genetic drift and environmental stresses (Hampe and Petit 2005). During interglacial periods, polewards recolonizations were often controlled by rare long-distance dispersal events followed by exponential population growth (Hampe and Petit 2005), and later migrants would be very difficult to establish (Hewitt 1996). Therefore, leading-edge populations are often characterized by low genetic diversity and high differentiation. Such predictions have been confirmed by theoretical and empirical studies (Hewitt 1996; Hampe and Petit 2005; Excoffier et al. 2009), though controversy still exists, e.g.,

Electronic supplementary material The online version of this article (doi:10.1007/s10144-012-0332-7) contains supplementary material, which is available to authorized users.

M.-M. Shi · X.-Y. Chen (✉)
Department of Environmental Sciences, Tiantong National Observation Station for Forest Ecosystems, East China Normal University, Shanghai 200062, China
e-mail: xychen@des.ecnu.edu.cn

M.-M. Shi
Department of Community Ecology, Helmholtz Centre for Environmental Research, UFZ, Theodor-Lieser-Str. 4, 06120 Halle, Germany

Published online: 22 July 2012

Yakimowski and Eckert (2008). However, recent studies indicated that such a difference between edge and non-edge populations was rarely very large (Eckert et al. 2008), suggesting that knowledge on putative evolutionary processes affecting the leading-edge populations is still poor (Parisod and Bonvin 2008).

Gene dispersal is one of the most important processes dominating the leading-edge populations (Austerlitz and Garnier-Géré 2003). Theoretical studies indicated that frequent long distance dispersal can considerably increase genetic diversity within populations and decrease differentiation among populations (Bialozyt et al. 2006; Fayard et al. 2009; Ray and Excoffier 2010). In plants, besides seed dispersal, which involves colonization of new areas, pollen dispersal also contributes to the gene pool of colonizing populations (Chen et al. 2008; Wang et al. 2011). Especially in wind-pollinated species, pollen grains can be transported over long distances via the wind, and the relative contribution of pollen dispersal to gene pools was much larger than that of seed dispersal (Petit et al. 2005; Sutherland et al. 2010). Therefore, dominant wind directions during the flowering season may play a critical role in genetic variation of populations. If leading-edge populations locate downwind of non-edge populations during the flowering season, the non-edge populations may serve as pollen sources and the leading-edge populations are recipients (Duminil et al. 2007). Thus the leading-edge populations may harbor genetic variation comparable to non-edge populations (Alleaume-Benharira et al. 2006; Eckert et al. 2008; Bridle et al. 2010).

In central subtropical China (ca. 22°N–30/33°N), evergreen broadleaved forests are distributed, and dominated mainly by species of Fagaceae and Lauraceae. Although there was no large ice-covered area in South China during the Pleistocene (Hewitt 2000), climatic fluctuations have essentially affected distributions of plant species and evergreen broadleaved forests. During the Last Glacial Maximum (LGM: ca. 18,000 years BP), temperature in southern China was estimated 4–6 °C lower than that of the present (Zheng 2000), and evergreen broadleaved forests retracted southwards to the present tropical zone (Sun et al. 1999; Harrison et al. 2001) according to LGM biome map. Contrary to this expectation, several empirical studies revealed multiple isolated refugia located in main mountain ranges, e.g., Nanling Mountains (south), Wuyi Mountains (southeast) (Chen et al. 2012). From either hypothesis, a recolonization route from south to north can be expected. Thus, evergreen broadleaved forests with the leading edge lying around Yangtze River could be recognized.

Castanopsis sclerophylla (Lindley & Paxton) Schottky (Fagaceae) is a wind-pollinated monoecious tree. It is one of the dominant species of evergreen broadleaved forests in China. It mainly distributes in south of Yangtze River and north of Nanling Mountains, and is common in mountains below 1,000 m. *Castanopsis sclerophylla* is one of the most cold-tolerant evergreen broadleaved trees. Although palaeoecological information of this species is less available, a comparative phylogeographic study on *C. eyrei*, which has the similar distribution range, revealed two different clusters and south-north recolonization routes (M. Shi, unpublished data). Therefore, the leading edge of *C. sclerophylla* can be supposed similar to that of evergreen broadleaved forests. *Castanopsis sclerophylla* blooms in May and its fruits mature in October and November (http://www.efloras.org/florataxon.aspx?flora_id=2&taxon_id=200006250). It is gravity-dispersed, and rodents may serve as the secondary disperser. Maximum seed dispersal distance was observed to be less than 50 m (R. Wang, personal communication), and pollen-to-seed dispersal ratio (Ennos 1994) was over 150 in the scale of 20–30 km (Zhang et al. 2012), which corresponded to the value of oak species (pollen-to-seed dispersal ratio = 196) and species of Fagaceae (Petit et al. 2005). Previous studies indicated that, due to intensive pollen dispersal, populations do not become differentiated at scales up to tens of kilometers (Wang et al. 2011).

In this study, we sampled five leading-edge populations along the northern range limit and 12 central or subcentral populations. Because of potentially high pollen dispersal in *C. sclerophylla*, we hypothesized that leading-edge populations maintain high genetic diversity comparable to non-edge populations. We tested this hypothesis by (1) comparing genetic diversity of leading-edge with non-edge populations, (2) analyzing genetic differentiation among leading-edge and non-edge populations, and (3) inferring the potential causes in the genetic structure of *C. sclerophylla*.

Methods

Sampling

From 2005 to 2007, 482 individuals of *C. sclerophylla* were sampled from 17 populations located in Zhejiang, Anhui, Jiangsu, Fujian, Jiangxi provinces and Shanghai Municipality (Fig. 1). Wuyi Mountains might be the largest barrier in our examined area, though there are hills which may disturb the northward colonizations as well as pollen dispersal by wind. In southeastern China, mature natural forests are generally fragmented because of serious human disturbances. However, *C. sclerophylla* might have become more frequent in the naturally restored secondary forests, due to its strong ability of re-sprouting. Towards to the northern range limit, it distributes less frequently and continuously. Therefore, five populations (NS, SS, TZ, QL, and LC) located along the northern range limit were

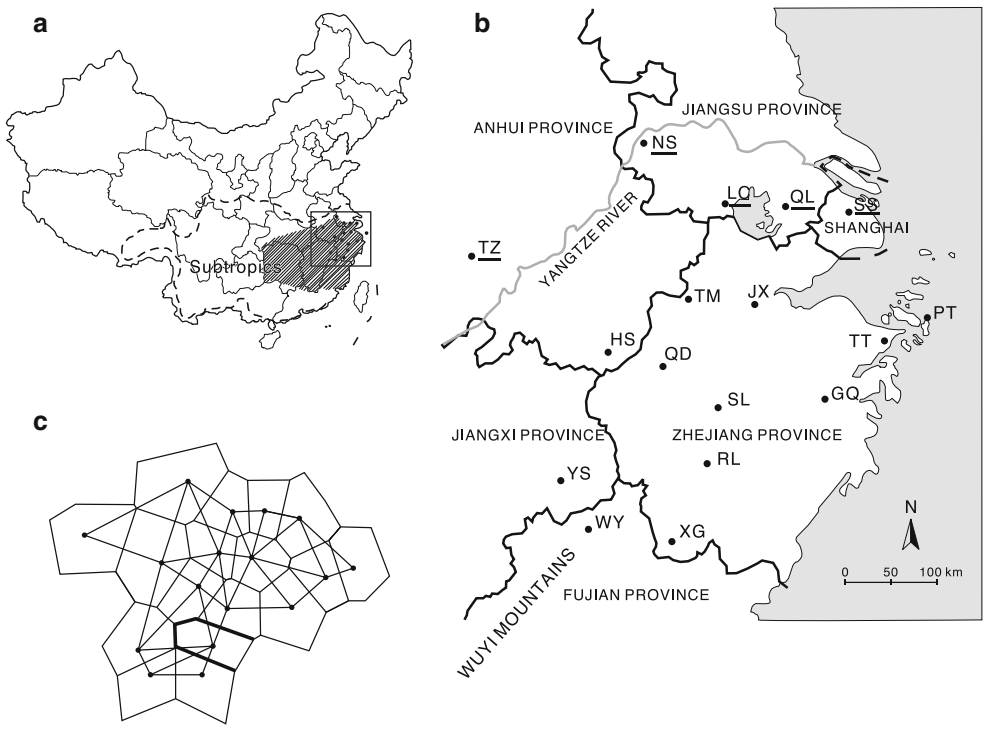

Fig. 1 **a** Map of China. *Shaded area* represents the main distribution of *Castanopsis sclerophylla*. The *dashed lines* indicate the region of subtropical China (modified after Fang et al. 2002). The *dots* in the *square* show the positions of sampled populations. **b** Locations of populations examined in this study (*dots*). Abbreviations for the sampled populations are the same as Table 1; the *underlined abbreviations* indicate leading populations. **c** Major genetic boundary (*bold line*, with bootstrap score 72 %) detected among the 17 populations of *C. sclerophylla* based on Monmonier's algorithm with significance tested by 1,000 bootstrap matrices of D_A genetic distance. The dots correspond to locations of sampled populations plotted on the map

considered as leading populations, while the others with more continuous and dense distribution were defined as non-edge populations (Fig. 1). In each population, leaves from 14 to 32 individuals were sampled, with at least 30 m distance between each other except population SS where we collected almost all trees with diameter at breast height (DBH) >10 cm (Table 1). Leaf tissue was dried by silica gel.

Microsatellite protocols

Genomic DNA was extracted from dried leaves (0.05 g) according to a modified CTAB protocol (Fan et al. 2004). Six microsatellite loci (Ccu16H15, Ccu33H25, Ccu62F15, Ccu87F23, Ccu93H17 and Ccu97H18) from *C. cuspidata* var. *sieboldii* Nakai (Ueno et al. 2000, 2003) and one locus (QM67-3M1) from *Quercus myrsinifolia* (Isagi and Suhandono 1997) were found to be polymorphic in *C. sclerophylla*. The microsatellite PCR reactions were performed in a 20 μL reaction mixture containing 50 ng genomic DNA, 0.2 μM of each primer, 2 μL of Tag DNA polymerase buffer [100 mM KCl, 80 mM $(NH_4)_2SO_4$, 100 mM Tris–HCl, NP-40, pH 9.0], 1 unit of Tag DNA polymerase, 0.2 mM of each dNTP and 1.5 mM $MgCl_2$. PCR reactions consisted of one cycle of 94 °C for 3 min, followed by 34 cycles of 30 s at 94 °C, 40 s at 45–66.5 °C depending on the primer used, and 30 s at 72 °C, with a final extension of 7 min at 72 °C. All reactions were run on a Mastercycler ep gradient PCR (Eppendorf, Hamburg, Germany). Amplification products were separated on 6 % denaturing polyacrylamide gels and detected by staining with silver nitrate. pUC 19 DNA/MspI (HpaII) marker 23 (Fermentas) was used as the reference of products' length.

Analysis of genetic variation

As trans-species amplification of microsatellites often results in null alleles, we checked the data for the presence of null alleles under the assumption of Hardy–Weinberg equilibrium (van Oosterhout et al. 2004) and adjusted genotypes, if necessary, by adding an additional allele using MICRO-CHECKER. Null alleles were indicated in six out of seven loci with the frequency ranging from 0.028 to 0.313 (mean = 0.168). The average frequency of null alleles per population varied from 0.039 to 0.188 (Table 1). Corrected and uncorrected allele frequency is shown in Table S1 of Electronic Supplementary Material (ESM). However, overall the results were very similar with or

Table 1 Information of sampling sites and genetic diversity of leading-edge and non-edge populations of *Castanopsis sclerophylla*

Sampling site	ID	Location	Sampling size	*Null*	*A*	A_R	H_e	H_o	F_{IS}	*P*	*T* (95 % interval)	*Ancestry*
Leading-edge populations												
Tianzhushan, Anhui Province	TZ	E 116.47°, N 30.72°	30	0.153	6.6	5.9	0.70	0.37	0.034	0.34	–	20.25
Niushoushan, Jiangsu Province	NS	E 118.74°, N 31.91°	31	0.123	5.0	4.5	0.67	0.49	−0.096	**0.01**	12,361 (1,079, 40,599)	18.59
Qionglongshan, Jiangsu Province	QL	E 120.42°, N 31.26°	30	0.168	6.0	5.5	0.78	0.36	0.018	**0.01**	2,050 (315, 19,395)	19.39
Longchishan, Jiangsu Province	LC	E 119.73°, N 31.24°	30	0.173	6.3	5.7	0.73	0.38	−0.008	**0.02**	1,471 (98, 17,529)	10.79
Sheshan, Shanghai Municipality	SS	E 121.19°, N 31.10°	26	0.039	4.0	3.9	0.55	0.45	0.035	0.08	–	9.70
Mean					5.6	5.1	0.69	0.41	−0.006			15.74
Non-edge populations												
Ruoliaoxian, Zhejiang Province	RL	E 119.30°, N 28.27°	14	0.092	4.1	4.1	0.51	0.27	**0.130**	0.95	–	83.91
Shuanglongdong, Zhejiang Province	SL	E 119.62°, N 29.13°	30	0.188	5.7	5.3	0.67	0.29	0.006	0.06	–	82.73
Xingguang, Zhejiang Province	XG	E 119.05°, N 27.62°	20	0.131	3.9	3.6	0.56	0.27	0.043	0.53	–	89.82
Guoqingsi, Zhejiang Province	GQ	E 121.02°, N 29.15°	32	0.170	4.7	4.2	0.60	0.29	−0.056	0.53	–	97.28
Putuoshan, Zhejiang Province	PT	E 122.38°, N 30.00°	28	0.176	4.4	4.1	0.58	0.23	0.007	0.53	–	86.97
Jiuxi, Zhejiang Province	JX	E 120.13°, N 30.23°	31	0.177	5.7	5.1	0.65	0.29	−0.005	0.06	–	58.26
Tiantong, Zhejiang Province	TT	E 121.78°, N 29.80°	30	0.172	5.4	4.7	0.60	0.28	−0.041	0.41	–	12.17
Tianmushan, Zhejiang Province	TM	E 119.42°, N 30.35°	30	0.164	5.4	4.8	0.66	0.33	0.004	0.15	–	61.96
Qiandaohu, Zhejiang Province	QD	E 118.98°, N 29.60°	30	0.145	5.7	5.2	0.64	0.35	0.001	0.06	–	57.49
Wuyishan, Fujian Province	WY	E 117.92°, N 27.64°	30	0.125	4.1	3.7	0.55	0.32	−0.030	0.19	–	72.64
Yanshan, Jiangxi Province	YS	E 117.65°, N 28.21°	30	0.129	5.4	5.0	0.70	0.44	0.003	**0.03**	3,563 (285, 26,379)	27.34
Huangshan, Anhui Province	HS	E 118.17°, N 30.15°	30	0.111	6.0	5.3	0.68	0.41	0.063	0.47	–	41.03
Mean					5.1	4.6	0.62	0.31	0.003			64.30
Overall mean					5.2	4.8	0.60	0.34				
Total: calculate for all individuals as one population			482		6.5		0.71					
P value for difference between leading edge and non-edge populations					0.221		0.103		0.684			

Average frequency of null alleles per population (*Null*), number of alleles per locus (*A*), allelic richness per locus (A_R), expected heterozygosity (H_e), observed heterozygosity (H_o), inbreeding coefficient (F_{IS}) (the significant value in bold), *P* value of bottleneck result based on two-phase model (TPM) with 70 % stepwise mutation model (SMM) by Wilcoxon sign-rank test (significant values in bold), *T* estimation of time (in years) since population started to decline with the interval at 95 % and percentage ancestry derived from cluster A averaged across all individuals in the populations (*Ancestry*)

without taking null alleles into account. We presented the results using corrected data set hereafter.

Genetic diversity at species and population levels was characterized by calculating the number of alleles (A), allelic richness (A_R, calculated by a rarefaction method, corrected for sample size, El Mousadik and Petit 1996), expected heterozygosity (H_e) and inbreeding coefficient (F_{IS}) using the program FSTAT 2.9.3.2 (Goudet 1995). Two one-tailed t tests were performed in order to test the null hypothesis H_0: $G_{\text{leading-edge populations}} = G_{\text{non-edge populations}}$, with the alternative hypothesis H_1: $G_{\text{leading-edge populations}} < G_{\text{non-edge populations}}$ (Arnaud-Haond et al. 2006), where G is the allelic richness or expected heterozygosity. Also, relationship between genetic diversity and latitude was analyzed using a linear regression with R 2.8.1 (R Development Core Team 2010).

Overall population differentiation was assessed by F_{ST} based on Weir and Cockerham (1984) estimators using FSTAT (Goudet 1995). As F_{ST} is likely to underestimate genetic differentiation between populations for markers which show high levels of allelic variability, we calculated F'_{ST} value, a standardized parameter of genetic differentiation as $F'_{ST} = F_{ST}/F_{ST\,max}$(Hedrick 2005). $F_{ST\,max}$ was calculated after recoding the data using RECODEDATA (Meirmans 2006). To test whether genetic differentiation was higher among leading-edge populations than that in non-edge populations, differences among mean pairwise F_{ST} values were evaluated using a randomization procedure with 1,000 permutations in FSTAT (Goudet 1995). An analysis of molecular variation (AMOVA) was carried out using GenAlEx 6 (Peakall and Smouse 2006). In addition, we conducted barrier analysis (Manni et al. 2004) in Barrier 2.2 using Monmonier's (1973) algorithm to identify genetic barrier(s) among populations. The statistical confidence of predicted barrier(s) was tested using 1,000 bootstraps matrices of D_A genetic distances (Nei et al. 1983), which were computed using Microsatellite analyzer (Dieringer and Schlötterer 2003).

To test the pattern of isolation by distance, the association between genetic differentiation (F_{ST}) and geographic distances was assessed using a Mantel test in R 2.8.1 (R Development Core Team 2010). Because pollen dispersal plays a critical role in gene flow, we also tested the relationship between differentiation and geographical distance of pairwise populations in or off predominant wind direction during flowering periods, i.e., northwestwards. If direction of two populations was within 10 degrees biased from southeast–northwest, they were in the wind direction. Otherwise, they were off the wind direction.

Population clusters and ancestry were detected by Bayesian clustering using STRUCTURE 2.3.3 (Pritchard et al. 2000). An admixture model was run with correlated allele frequencies. Each run was pursued for 50,000 MCMC interactions, with an initial burn-in of 10,000. To estimate the K number of ancestral genetic populations and the ancestry membership coefficients of each individual in these clusters, the algorithm was run 10 times for each user-defined K value from 1 to 10. The mean log-likelihood at each value of K [ln Pr($X|K$)] was plotted, and an ad-hoc statistics ΔK, as suggested by Evanno et al. (2005), was used to estimate the most likely number of clusters.

The software BOTTLENECK (Cornuet and Luikart 1996) was used to check a recent reduction of effective population size under the assumption that alleles are generally lost faster than heterozygosity, and recently bottlenecked populations will therefore display an excess of heterozygosity relative to that expected from the equilibrium between mutation and drift (Piry et al. 1999). Three models there are optional: stepwise mutation model (SMM), infinite allele model (IAM) and two-phase model (TPM). Microsatellite loci with dinucleotide repeats or compound motifs are thought to follow a mutational model deviating from the SMM towards the IAM (Straub and Doyle 2009), thus in this study the TPM was the most appropriate and consequently the TPM with 70 % SMM character was employed. Significance was evaluated by Wilcoxon sign-rank test. In addition, to determine the time of bottlenecks, we used MSVAR program (Beaumont 1999), which implements a coalescent simulation-based Bayesian likelihood analysis, assumes a strict SMM, and estimates the posterior probability distribution of population parameters using Markov Chain Monte Carlo simulations, based on the observed distribution of microsatellite alleles and their repeat numbers. We set the generation time of *C. sclerophylla* as 15 years (T. Fang, personal communication). In each simulation, we ran each chain with 20,000 thinned updates and a thinning interval of 10,000 steps, leading to a total number of 2×10^8 updates. The first 10 % of updates were discarded to avoid bias in parameter estimation at starting conditions, and the remaining data were used to obtain the lower (5 %), the median (50 %), and the upper (95 %) quantiles of the posterior distributions.

Results

A total of 57 alleles were detected at seven loci among 482 samples from 17 populations. The number of alleles per locus ranged from 4 (Ccu97H18 and QM67-3M1) to 15 (Ccu93H17), with a mean of 8. At the species level, *C. sclerophylla* had a high value of expected heterozygosity (H_e) of 0.71.

The mean number of alleles per locus (A) for non-edge populations ranged from 3.9 (population XG) to 6.0

(population HS), while that of leading-edge populations ranged from 4.0 (SS) to 6.6 (TZ). Considering allelic richness (A_R) which is used to correct the different sample size, it ranged from 3.9 to 5.9 (mean = 5.1) in leading-edge populations, while from 3.6 to 5.3 (mean = 4.6) in non-edge populations. The expected heterozygosity (H_e) varied from 0.55 to 0.78 (mean = 0.69) in leading-edge populations and from 0.51 to 0.70 (mean = 0.62) in non-edge populations. Except in population RL, inbreeding coefficients were not biased from 0 in leading-edge and non-edge populations using corrected data set (Table 1). Leading-edge populations had higher A_R and H_e values, but not significantly (t test: $P = 0.139$ and 0.077, respectively). However, when excluding population SS whose upwind only one population with low genetic diversity (population PT) was laid, t test showed significantly higher A_R and H_e in leading populations than in non-edge populations ($P = 0.036$ and 0.005, respectively). Number of alleles per locus (A) (adjusted $R^2 = 0.202$, $P = 0.040$), allelic richness per locus (A_R) (adjusted $R^2 = 0.199$, $P = 0.042$) and expected heterozygosity (H_e) (adjusted $R^2 = 0.254$, $P = 0.023$) were positively related with latitude (Fig. 2), indicating that genetic diversity increased towards northern margin limit.

An analysis of molecular variance (AMOVA) revealed that the majority of genetic variation (80 %) resided within populations, among groups only 2 %, while 18 % of the total genetic variation was detected among populations within group (Table 2). Overall F_{ST} over seven loci was 0.116 ($P < 0.01$), and the standardized F'_{ST} was 0.327 indicating moderate differentiation. Genetic differentiation of leading-edge populations ($F_{ST} = 0.081$) was lower but not significantly compared to non-edge populations ($F_{ST} = 0.117$) ($P = 0.691$). Pairwise F_{ST} values varied from 0.022 (YS and TT) to 0.373 (PT and RL), and were marginally significantly related to geographic distances according to Mantel test ($r = 0.212$, $P = 0.071$). However, when pairwise populations were grouped into in or off wind direction, significant isolation-by-distance pattern was observed in population pairs both in ($r = 0.423$, $P = 0.024$) and off wind direction ($r = 0.180$, $P = 0.032$) (Fig. 3). F_{ST} values of pairwise populations in wind direction (mean = 0.113) were smaller, but not significantly, than those of off wind direction (mean = 0.121). Barrier analysis revealed a single genetic boundary in central populations, with 72 % bootstrap support, separating population RL from others (Fig. 1).

An estimate of number of clusters was obtained by ΔK statistic, producing a distinct peak at $K = 2$ (Fig. 4) which would correlate to the most likely value of K. All the 17 populations can be assigned to two clusters (A and B) (Fig. 4). The majority of individuals from leading-edge populations were assigned into one substructure together with non-edge populations TT, YS and HS. The percentage contribution that each cluster makes per population is shown in Table 1.

Table 2 Analysis of molecular variance (AMOVA) displaying the genetic variation between the two groups of leading-edge and non-edge populations, among populations within the groups and individuals of *Castanopsis sclerophylla*

Source of variation	*df*	Sum of squares	Percentage of variation
Among groups	1	66.5	2**
Among populations within groups	15	594.4	18**
Within populations	465	2,669.3	80**
Total	481	3,330.2	

** $P < 0.01$

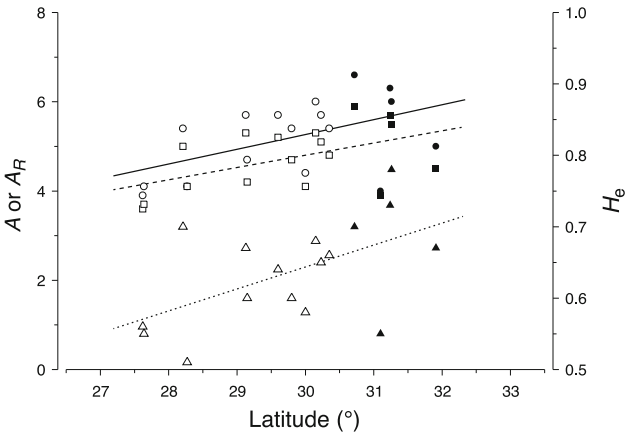

Fig. 2 Changes of number of alleles per locus (A) (*circles*, *solid line*), allelic richness per locus (A_R) (*squares*, *dashed line*) and expected heterozygosity (H_e) (*triangles*, *dotted line*) along latitude. *Filled labels* were data of leading-edge populations

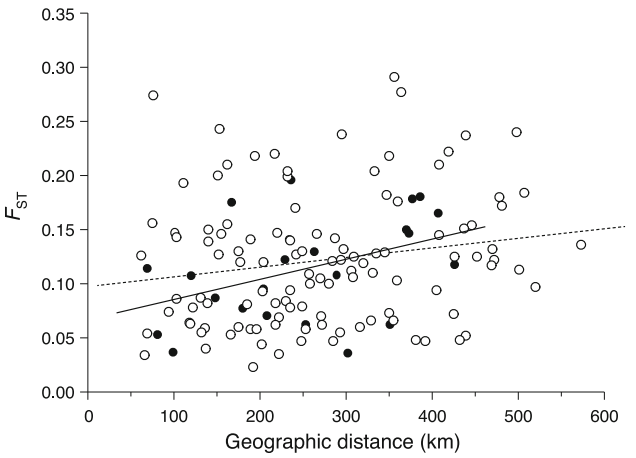

Fig. 3 Relationship between differentiation (F_{ST}) and geographic distances of population pairs in (*filled circles*, *solid line*; $r = 0.423$, $P = 0.020$) and off (*empty circles*, *dashed line*; $r = 0.180$, $P = 0.031$) dominant wind direction during flowering season by the Mantel test with 1,000 randomizations

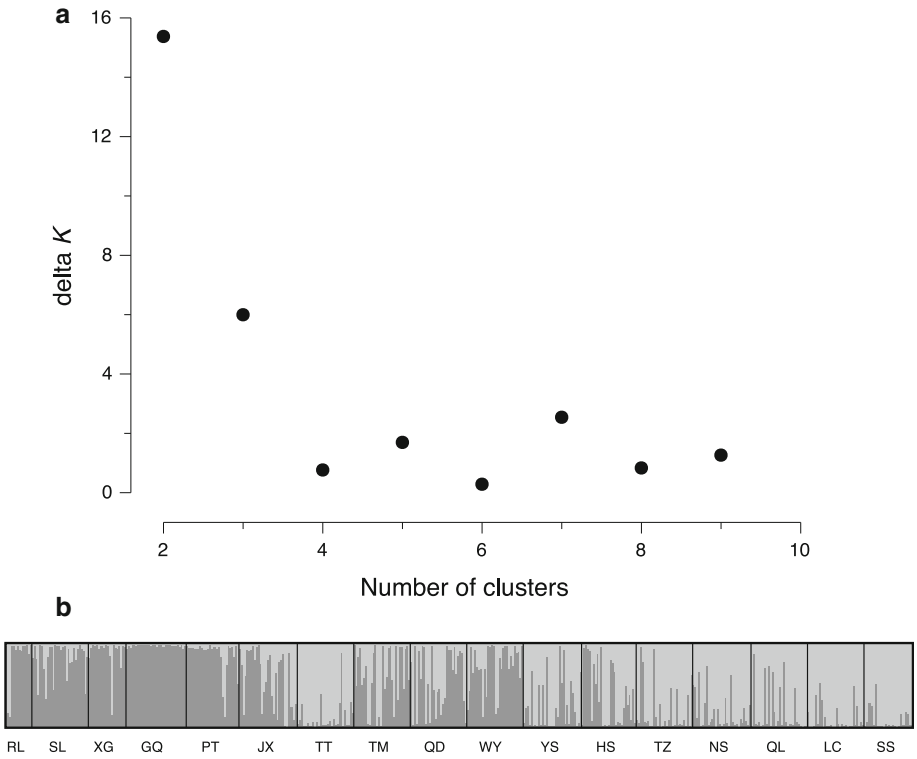

Fig. 4 **a** Values of ΔK (delta K) based on the rate of the change of ln Pr($X|K$) as a function of the number of clusters K. **b** Distribution of cluster membership at individual and population levels when $K = 2$ (*grey* cluster A, *light grey* cluster B)

The analysis of recent bottleneck effects indicated that almost all of the non-edge populations except YS (8.3 % of non-edge populations) did not suffer from population bottlenecks. However, three of the five leading-edge populations (NS, QL, LC) (60 % of leading populations) did experience significant recent bottlenecks ($P < 0.05$) (Table 1). The estimated time since decline (T) ranged from 1,500 to 13,000 years ago according to MSVAR.

Discussion

Genetic diversity of *Castanopsis sclerophylla* populations

As predicted by leading-edge colonization model, significant signs of recent bottlenecks due to founder events were found in three of the five leading-edge populations, while only one non-edge populations experienced recent bottleneck. However, genetic variation in leading-edge populations was higher, though not significantly than that in non-edge populations. Furthermore, there was a significant increasing trend of genetic diversity towards leading edge. These observations of genetic variation were not consistent with the general predictions by leading-edge colonization model. If heterosis is more advantageous at the range edge because of more extreme environmental conditions (Hansson and Westerberg 2002), higher heterozygosity can be observed in the leading edge populations. However, similar levels of inbreeding coefficients (F_{IS}) and heterozygosity were observed in the leading and non-edge populations (Table 1). Furthermore, microsatellites, in general, are neutral and not impacted by selection (Ellegren 2004).

Given colonization by long-distance dispersers and difficulty for later migrants to establish, leading-edge populations were characterized by low genetic diversity (Hewitt 1996). However, life-history traits may determine genetic structure among populations (Hamrick and Godt 1996; Pannell and Dorken 2006; Duminil et al. 2007), particularly mating system (Duminil et al. 2007). The species with outcrossing mating system are considered capable to harbor high genetic diversity (van Rossum et al. 1997), which has been evidenced by many studies (Shi et al. 2011). In addition, although the role of seed and pollen dispersal type was indicated not as significant as mating system for nuclear markers (Duminil et al. 2007), in a view of long-distance dispersal, diversity levels between different populations should be more similar (Cottrell et al. 2003). Seeds of Fagaceae species are gravity-dispersed, and rodents may serve as the secondary disperser. In these species, seed dispersal plays a minor role in gene flow compared to pollen dispersal, especially in a large scale. Ratios of pollen-to-seed dispersal of species of Fagaceae are among the highest values of studied species (Petit et al. 2005). Therefore, pollen dispersal plays a key role in shaping genetic structure of Fagaceae species.

Secondly, for wind-pollinated species, wind is supposed to have a great contribution to their pollen dispersal. Morphometric data on oak pollen grains have suggested that they could be wind-transported at distances exceeding hundreds of kilometers (Stanley and Linskens 1974). Streiff et al. (1999) also demonstrated a high proportion (67 % on average) of pollen in *Quercus petraea* and *Q. robur* originating from outside the study stand of 5.76 ha. Furthermore, direction of dominant winds determines the direction of pollen dispersal of wind-pollinated species, and thus source and recipient populations. *Castanopsis sclerophylla* flowers in May, when dominant wind currents are northwestwards (Yu et al. 2009). Therefore, leading-edge populations are recipients, which may receive pollen grains from multiple non-edge populations (i.e., source populations), resulting in higher genetic variation in leading-edge compared to non-edge populations (Table 1). The conclusion that wind facilitates maintaining high genetic diversity in the leading-edge populations was also supported by the increasing trend in genetic variation to the northern leading edge and by more significant isolation by distance pattern along wind directions (Fig. 2). Similarly, Sutherland et al. (2010) investigated 42 British populations of *Fraxinus excelsior* L., a wind pollinated tree also, and found higher genetic variation in eastern Britain, which could be attributed to the prevailing wind direction from the west.

Thirdly, combination of multiple sources may contribute to high genetic diversity in leading-edge populations, such as *Himantoglossum hircinum* (Preifer et al. 2010). In our study, two clusters were detected in the 17 populations in the eastern distribution range. The leading-edge populations more or less received genes from two clusters with 15.7 % genes contributed by cluster A, on average. Such a mixing of genes can maintain high genetic diversity in populations during their rapidly expanding (Pluess 2011). Lastly, population size is thought to be one of the factors affecting the level of genetic diversity in populations (Hensen and Oberprieler 2005). As we know, population size can influence the dynamics of plant populations (Li et al. 2012). Genetic drift and inbreeding depression often threaten small populations (Chen et al. 2003), resulting in low level of genetic variation. However, we did not observe significant signs of small size/fragmentation in leading-edge populations, except population SS. Population SS is located in Shanghai Municipality and has been affected by highly urbanization with remnant individuals <50. The hypothesis that small populations harbor low levels of genetic diversity was supported.

Genetic differentiation of *C. sclerophylla* populations

In this study, genetic differentiation among leading-edge populations was lower, though not significantly, than that of non-edge populations. AMOVA showed that only 2 % variation existed among groups. These results were also inconsistent with the general prediction that edge populations are more differentiated compared to non-edge populations because edge populations are generally isolated and of small size (Eckert et al. 2008).

As suggested by Duminil et al. (2007), mating system was one of the only few traits related with genetic structure. Outcrossed species are not expected to show significant population structure. Additionally, wind-pollination manner could also explain such a discrepancy. Generally, genetic differentiation of wind-pollinated species is lower than that of insect-pollinated and self-fertilized species (Hamrick and Godt 1989; Duminil et al. 2007; Puşcaş et al. 2008), though not significant under phylogenetic independent contrast (Duminil et al. 2007). Potentially high gene flow in *C. sclerophylla* is expected to even out genetic differentiation among leading-edge and non-edge populations. As mentioned above, leading-edge populations may act as receipts, which can receive pollen from multiple non-edge populations during flowering periods. Actually, Bayesian clustering supplied some clues (Fig. 4). STRUCTURE results showed that leading-edge populations exhibited majority genes from cluster B, like non-edge populations YS and TT, meanwhile, as well as minority from cluster A, characterized by most of non-edge populations. Sharing same gene pools tone down differentiation between leading-edge and non-edge populations.

Low differentiation by high pollen dispersal in *C. sclerophylla* is also supported by the isolation-by-distance pattern observed in pairwise populations in and off wind direction (Fig. 3). Since microsatellites are generally neutral and selection is absent or weak, isolation-by-distance pattern indicated a balance between gene flow and drift (Hutchison and Templeton 1999). However, more significant patterns of isolation by distance ($r = 0.423$) revealed by population pairs in wind direction during flowering season state clearly that gene flow predominates in wind direction, whereas more genetic drift off wind direction. Geographic barrier plays an important role in shaping population structure. However, in the region we studied, even in the eastern distribution range of *C. sclerophylla*, mountains over than 2,000 m altitude are sporadically distributed in Wuyi Mountain ranges where population WY is located. Although the leading-edge populations are distributed locally less continuously than non-edge populations, they are still connected to other populations, because there were generally scattered *C. sclerophylla* trees around or in local villages (X. Chen, personal observation), serving as stepping stones. In fact, the only significant genetic barrier was detected in surrounding of RL, where the northern extension of Wuyi Mountains is located, thus supporting genetic continuities among leading-edge and non-edge populations.

Climate change and leading-edge populations

Since the LGM about 18,000 years ago, evergreen broad-leaved forest recolonized the subtropical areas of China, though there might experience several contractions (Member of China Quaternary Pollen Data Base 2000). In spite of lack of pollen or phylogeographic evidence on *C. sclerophylla*, leading edge was most likely to be recolonized in recent several thousands of years (Wang 1992). Therefore, these populations might experience founder events. In fact, we found that three leading-edge populations (60 %) (NS, QL, LC) had suffered from significant population bottlenecks and population declines from 1,500 to 13,000 years ago. Although founder events during colonization should result in stochastic loss of genetic variation and divergence among populations, high gene flow may have compensated for these processes (Muir et al. 2004).

As climate warms, it is likely that *C. sclerophylla* will shift to higher latitudes in response to climate changes. This implies that the leading-edge populations will move northwards while the central and subcentral populations will occupy the areas currently occupied by the leading-edge populations. Though microsatellites are thought to be neutral, they may indicate variation at selective loci (Shi et al. 2011). High genetic diversity not only provides an adaptation to less favorable habitats for leading-edge populations (Honjo et al. 2008), but may also provide leading-edge populations raw materials for evolutionary adaptation to future environmental conditions.

Acknowledgments We appreciate Walter Durka to give constructive suggestions and supply some statistical helps. We thank Naoki Tani and two anonymous reviewers for helpful comments and suggestes and Shuo Yu, Mei-Hua Liu, Dawei Gao, Bang-Quan Gao and Xiao-Yan Wang for helps in sample collection. This work was supported by the National Natural Science Foundation of China (30970430, 30470287) and the Fundamental Research Funds for the Central Universities (78220028).

References

Alleaume-Benharira M, Pen IR, Ronce O (2006) Geographical patterns of adaptation within a species' range: interactions between drift and gene flow. J Evol Biol 19:203–215

Arnaud-Haond S, Teixeira S, Massa SI, Billot C, Saenger P, Coupland G, Duarte CM, Serrao EA (2006) Genetic structure at range edge: low diversity and high inbreeding in Southeast Asian mangrove (*Avicennia marina*) populations. Mol Ecol 15:3515–3525

Austerlitz F, Garnier-Géré PH (2003) Modelling the impact of colonisation on genetic diversity and differentiation of forest trees: interaction of life cycle, pollen flow and seed long-distance dispersal. Heredity 90:282–290

Beaumont MA (1999) Detecting population expansion and decline using microsatellites. Genetics 153:2013–2029

Bialozyt R, Ziegenhagen B, Petit RJ (2006) Contrasting effects of long distance seed dispersal on genetic diversity during range expansion. J Evol Biol 19:12–20

Bridle JR, Polechov J, Kawata M, Butlin RK (2010) Why is adaptation prevented at ecological margins? New insights from individual-based simulations. Ecol Lett 13:485–494

Chen X-Y, Li YY, Wu TY, Zhang X, Lu HP (2003) Size-class differences in genetic structure of *Metasequoia glyptostroboides* Hu et Cheng (Taxodiaceae) plantations in Shanghai. Silvae Genet 52:107–109

Chen X-Y, Fan X-X, Hu X-S (2008) Roles of seed and pollen dispersal in natural regeneration of *Castanopsis fargesii* (Fagaceae): implications for forest management. For Ecol Manag 256:1143–1150

Chen Y, Compton SG, Liu M, Chen X-Y (2012) Fig trees at the northern limit of their range: the distributions of cryptic pollinators indicate multiple glacial refugia. Mol Ecol 21: 1687–1701

Cornuet JM, Luikart G (1996) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. Genetics 144:2001–2014

Cottrell JE, Munro RC, Tabbener HE, Milner AD, Forrest GI, Lowe AJ (2003) Comparison of fine-scale genetic structure using nuclear microsatellites within two British oakwoods differing in population history. For Ecol Manag 176:287–303

Dieringer D, Schlötterer C (2003) Microsatellite analyser (MSA): a platform independent analysis tool for large microsatellite data sets. Mol Ecol Notes 3:167–169

Duminil J, Fineschi S, Hampe A, Jordano P, Salvini D, Vendramin GG, Petit RJ (2007) Can population genetic structure be predicted from life-history traits? Am Nat 169:662–672

Eckert CG, Samis KE, Lougheed SC (2008) Genetic variation across species' geographical ranges: the central-marginal hypothesis and beyond. Mol Ecol 17:1170–1188

El Mousadik A, Petit RJ (1996) Chloroplast DNA phylogeography of the argan tree of Morocco. Mol Ecol 5:547–555

Ellegren H (2004) Microsatellites: simple sequences with complex evolution. Nat Rev Genet 5:435–445

Ennos RA (1994) Estimating the relative rates of pollen and seed migration among plants populations. Heredity 72:250–259

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 14:2611–2620

Excoffier L, Foll M, Petit RJ (2009) Genetic consequences of range expansions. Annu Rev Ecol Evol Syst 40:481–501

Fan XX, Shen L, Zhang X, Chen X-Y, Fu C (2004) Assessing genetic diversity of *Ginkgo biloba* L. (Ginkgoaceae) populations from China by RAPD markers. Biochem Genet 42:269–278

Fang JY, Song YC, Liu HY, Piao SL (2002) Vegetation-climate relationship and its application in the division of vegetation zone in China. Acta Bot Sin 44:1105–1122

Fayard J, Klein EK, Lefevrè F (2009) Long distance dispersal and the fate of a gene from the colonization front. J Evol Biol 22: 2171–2182

Goudet J (1995) FSTAT (version 1.2): a computer program to calculate F-statistics. J Hered 86:485–486

Hampe A, Petit RJ (2005) Conserving biodiversity under climate change: the rear edge matters. Ecol Lett 8:461–467

Hamrick JL, Godt MJW (1989) Allozyme diversity in plant species. In: Brown AHD, Clegg MT, Kahler AL, Weir BS (eds) Plant population genetics, breeding, and genetic resources. Sinauer, Sunderland, pp 43–63

Hamrick JL, Godt MJW (1996) Effects of life history traits on genetic diversity in plant species. Philos Trans R Soc B Biol Sci 351: 1291–1298

Hansson B, Westerberg L (2002) On the correlation between heterozygosity and fitness in natural populations. Mol Ecol 11:2467–2474

Harrison SP, Yu G, Takahara H, Prentice IC (2001) Palaeovegetation: diversity of temperate plants in East Asia. Nature 413:129–130

Hedrick PW (2005) A standardized genetic differentiation measure. Evolution 59:1633–1638

Hensen I, Oberprieler C (2005) Effects of population size on genetic diversity and seed production in the rare *Dictamnus albus* (Rutaceae) in central Germany. Conserv Genet 6:63–73

Hewitt G (1996) Some genetic consequences of ice ages, and their role in divergence and speciation. Biol J Linnean Soc 58:247–276

Hewitt G (2000) The genetic legacy of the Quaternary ice ages. Nature 405:907–913

Honjo M, Ueno S, Tsumura Y, Handa T, Washitani I, Ohsawa R (2008) Tracing the origins of stocks of the endangered species *Primula sieboldii* using nuclear microsatellites and chloroplast DNA. Conserv Genet 9:1139–1147

Hutchison DW, Templeton AR (1999) Correlation of pairwise genetic and geographic distance measures: inferring the relative influences of gene flow and drift on the distribution of genetic variability. Evolution 53:1898–1914

Isagi Y, Suhandono S (1997) PCR primers amplifying microsatellite loci of *Quercus myrsinifolia* Blume and their conservation between oak species. Mol Ecol 6:897–899

Li YY, Guan SM, Yang SZ, Luo Y, Chen X-Y (2012) Genetic decline and inbreeding depression in the extremely rare tree, *Ostrya rehderiana* Chun (Betulaceae). Conserv Genet 13:343–347

Manni F, Guerard E, Heyer E (2004) Geographic patterns of (genetic, morphologic, linguistic) variation: how barriers can be detected by "Monmonier's algorithm". Hum Biol 76:173–190

Meirmans PG (2006) Using the AMOVA framework to estimate a standardized genetic differentiation measure. Evolution 60:2399–2402

Member of China Quaternary Pollen Data Base (2000) Pollen-based biome reconstruction at Middle Holocene (6 Ka BP) and Last Galcial Maximum (18 ka BP) in China. Acta Bot Sin 42: 1201–1209

Monmonier M (1973) Maximum-difference barriers: an alternative numerical regionalization method. Geogr Anal 5:245–261

Muir G, Lowe AJ, Fleming CC, Vogl C (2004) High nuclear genetic diversity, high levels of outcrossing and low differentiation among remnant populations of *Quercus petraea* at the margin of its range in Ireland. Ann Bot London 93:691–697

Nei M, Tajima F, Tateno Y (1983) Accuracy of estimated phylogenetic trees from molecular data. J Mol Evol 19:153–170

Pannell JR, Dorken ME (2006) Colonisation as a common denominator in plant metapopulations and range expansions: effects on genetic diversity and sexual systems. Landsc Ecol 21:837–848

Parisod C, Bonvin G (2008) Fine-scale genetic structure and marginal processes in an expanding population of *Biscutella laevigata* L. (Brassicaceae). Heredity 101:536–542

Peakall R, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol Ecol Notes 6:288–295

Petit RJ, Duminil J, Fineschi S, Hampe A, Salvini D, Vendramin GG (2005) Comparative organization of chloroplast, mitochondrial and nuclear diversity in plant populations. Mol Ecol 14:689–701

Piry S, Luikart G, Cornuet J-M (1999) BOTTLENECK: a computer program for detecting recent reductions in the effective size using allele frequency data. J Hered 90:502–503

Pluess AR (2011) Pursuing glacier retreat: genetic structure of a rapidly expanding *Larix decidua* population. Mol Ecol 20:473–485

Preifer M, Passalacqua NG, Bartram S, Schatz B, Croce A, Carey PD, Kraudelt H, Jeltsch F (2010) Conservation priorities differ at opposing species borders of a European orchid. Biol Conserv 143:2207–2220

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–959

Puşcaş M, Taberlet P, Choler P (2008) No positive correlation between species and genetic diversity in European alpine grasslands dominated by *Carex curvula*. Divers Distrib 14:852–861

R Development Core Team (2010) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna

Ray N, Excoffier L (2010) A first step towards inferring levels of long-distance dispersal during past expansions. Mol Ecol Resour 10:902–914

Shi MM, Michalski SG, Chen X-Y, Durka W (2011) Isolation by elevation: genetic structure at neutral and putatively non-neutral loci in a dominant tree of subtropical forests, *Castanopsis eyrei*. PLoS One 6:e21302

Stanley R, Linskens H (1974) Pollen: biology, biochemistry and management. Springer, Berlin

Straub SCK, Doyle JJ (2009) Conservation genetics of *Amorpha georgiana* (Fabaceae), an endangered legume of the Southeastern United States. Mol Ecol 18:4349–4365

Streiff R, Ducousso A, Lexer C, Steinkellner H, Gloessl J, Kremer A (1999) Pollen dispersal inferred from paternity analysis in a mixed oak stand of *Quercus robur* L. and *Q. petraea* (Matt.) Liebl. Mol Ecol 8:831–841

Sun XJ, Song CQ, Chen XD (1999) China Quaternary pollen database (CPD) and "Biome 6000" project. Adv Earth Sci 14:407–411

Sutherland BG, Belaj A, Nier S, Cottrell JE, Vaughan SP, Hubert J, Russell K (2010) Molecular biodiversity and population structure in common ash (*Fraxinus excelsior* L.) in Britain: implications for conservation. Mol Ecol 19:2196–2211

Ueno S, Yoshimaru H, Kawahara T, Yamamoto S (2000) Isolation of microsatellite markers in *Castanopsis cuspidata* var. *sieboldii* Nakai from an enriched library. Mol Ecol 9:1188–1190

Ueno S, Yoshimaru H, Kawahara T, Yamamoto S (2003) A further six microsatellite markers for *Castanopsis cuspidata* var. *sieboldii* Nakai. Conserv Genet 4:813–815

van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Mol Ecol Notes 4:535–538

van Rossum F, Vekemans X, Meerts P, Gratia E, Lefebvre C (1997) Allozyme variation in relation to ecotypic differentiation and population size in marginal populations of *Silene nutans*. Heredity 78:552–560

Wang W-T (1992) On some distribution patterns and some migration routes found in the eastern Asiatic region. Acta Phytotaxon Sin 30:1–24

Wang R, Compton SG, Chen X-Y (2011) Fragmentation can increase spatial genetic structure without decreasing pollen-mediated gene flow in a wind-pollinated tree. Mol Ecol 20:4421–4432

Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. Evolution 38:1358–1370

Yakimowski SB, Eckert CG (2008) Populations do not become less genetically diverse or more differentiated towards the northern limit of the geographical range in clonal *Vaccinium stamineum* (Ericaceae). New Phytol 180:534–544

Yu J, Li J, Chen H (2009) Diurnal variation of surface wind over central eastern China. Climate Dynam 33:1089–1097

Zhang X, Shi M–M, Shen D-W, Chen X-Y (2012) Habitat loss other than fragmentation per se decreased nuclear and chloroplast genetic diversity in a monoecious tree. PLoS One 7:e39146

Zheng Z (2000) Late Quaternary vegetational and climatic changes in the tropical and subtropical areas of China. Acta Micropalae Sin 17:125–146