

RESEARCH

Open Access



Genetic structure and demographic analysis of a true single-population species, *Camellia azalea*

Ruo-Qiu Wang^{1,2}, Ying Xu², Min Zhang^{3*}, Gang Hao⁴, Qiang-Min Zhao⁵, Xin-Kai Liu⁵, Xiao-Fei Liu⁶, Bo Yu⁶ and Wen-Ju Zhang^{2*}

Abstract

Single-population species (SPS) consist of only one natural population and often are at high risk of extinction. Although almost all species must go through this special stage in their evolutionary process, there is little understanding of how SPS survives. *Camellia azalea* C. F. Wei is a typical SPS, and has precious breeding values for its special flowering period. This study surveyed the age structure and spatial distribution of *C. azalea*, analyzed its genetic diversity and fine-scale spatial genetic structure (SGS) using microsatellite markers for 629 individuals, and estimated the effect of human disturbances on its population dynamics. Results showed that this species had a relatively moderate genetic diversity ($I = 0.989$, $He = 0.509$, and $Ho = 0.497$), high rate of sapling (~35%), and a narrow habitat (~6 km long, ~10 m wide). Although the construction of dams and roads did not lead to a significant loss of genetic diversity and genetic differentiation ($F_{ST} = 0.0096 \sim 0.0128$, $N_m = 19 \sim 26$), it limited *C. azalea*'s seed flow (adults, 95 m; juveniles, 60 m), which was a reason for juveniles having a stronger SGS than adults. These results indicate that as an SPS, *C. azalea* still possesses the potential capacity for self-evolution and regeneration, however, it is at risk of extinction due to its small range, narrow habitat, and human distances. Furthermore, the results are also of enlightening significance to the conservation of other SPS, especially those distributed along the riparian zone.

Keywords Genetic diversity, Fine-scale genetic structure, Habitat degradation, Conservation, SSR

*Correspondence:

Min Zhang

zhangmin@njfu.edu.cn

Wen-Ju Zhang

wjzhang@fudan.edu.cn

¹ Tech X Academy (Tech X Institute), Shenzhen Polytechnic University, Shenzhen, Guangdong 518055, China

² Ministry of Education Key Laboratory for Biodiversity Science and Ecological Engineering, School of Life Sciences, Fudan University, Shanghai 200438, China

³ Co-Innovation Center for Sustainable Forestry in Southern China, College of Life Sciences, Nanjing Forestry University, Nanjing 210037, Jiangsu, China

⁴ College of Life Sciences, South China Agricultural University, Guangzhou, Guangdong 510642, China

⁵ Palm Eco-Town Development Co, Ltd, Guangzhou, Guangdong 510627, China

⁶ Guangdong Key Lab of Ornamental Plant Germplasm Innovation and Utilization, Environmental Horticulture Research Institute, Guangdong Academy of Agricultural Sciences, Guangzhou 510640, China



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, 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 you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. 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-nc-nd/4.0/>.

Background

In the recent hundreds of years, with the continuous intensification of anthropogenic activities, some species have now become extinct [1], and more species are being threatened with extinction to become endangered plants [2, 3]. Current extinction rates are so high that biologists suggest that a sixth mass extinction may be underway, highlighting the need for effective conservation measures [4]. There is a special category of species, which consists of only one single population and faces the highest risk of extinction. We define these species as single-population species (SPS). In fact, the single population is a necessary stage in the evolutionary progress of most species. When a species is at its origin or near extinction, it is likely to be a single population. At this time, due to the lacking of gene flow, it can only rely on gene mutation to accumulate genetic variation, which is quite slow, and there is a greater risk of extinction at this stage [5, 6]. For SPS, the extinction of population means the extinction of species. Thus, the understanding of SPS is of special important significance for biological conservation.

The fate of SPS depends on the balance between its evolutionary ability and environmental changes, especially the changes caused by human activities. The genetic diversity within a species is the basis of evolutionary ability and an essential element in maintaining long-term survival of species [7]. Previous studies have suggested that small natural population species are at risk of extinction because of insufficient genetic diversity to adapt to environmental changes [8]. It is unclear whether SPS's genetic diversity is lower than multi-population species. Compared to genetic diversity, fine-scale spatial genetic structure (SGS) reveals the spatial distribution characteristics of genetic diversity and is more sensitive to habitat changes [9]. SGS is usually the result of multiple factors working together, including historical climate change, terrain conditions, population size, habitat heterogeneity, and gene flow [10], and it helps to predict the future natural selection direction of the population, and provides direct data reference and theoretical analysis basis for the development of natural population management and genetic resource protection strategies [11, 12]. A sufficient understanding of genetic diversity and SGS is the key to protecting endangered plants [13]. Furthermore, for the survival of SPS, a demographic census including the number of individuals and the area of distribution is necessary [14].

SPS should not be rare in the sixth mass extinction, however, in practice, they are difficult to find or confirm. There is little understanding of the characteristics of SPS. *Camellia azalea* C. F. Wei is a typical SPS endemic to Guangdong, China. It was first discovered and identified in 1985 [15–17]. According to several investigations, this

species contains only one single population and consists of about one thousand wild individuals, which grows wholly along one valley of the Ehuangzhang Nature Reserve, Guangdong Province [18]. In taxonomy, this species is a well-defined species and can be easily distinguished from other *Camellia* plants, since its leaves are non-serrated at the margin (Fig. 1C). It also has a characteristic different from other *Camellia* species, that is, its flowers not only have a long flowering period but also are in full bloom in summer. This is extremely rare in *Camellia* species and provides a possibility to change dramatically the flowering period of cultivated *Camellia* varieties and make them bloom in summer by inter-specific crossing [18]. With its potential value, *C. azalea* has been exploited by selective logging since it was discovered. According to the International Union for Conservation of Nature (IUCN Red-List, <https://www.iucnredlist.org/>) standards [19], it has been rated as a critically endangered (CR) species.

The genetic diversity and age structure of *C. azalea* were surveyed [20, 21]. However, only 60 individuals were collected in the former study and 174 individuals in the latter. For endangered species with self-incompatibility, all genetic variation in the population may be significant to the species' survival, so it is necessary to fully understand the genetic diversity of the population. In addition, the effects of recent environmental disturbances on the genetic diversity of woody plants may not be immediately apparent, but the effects on the spatial distribution pattern of genetic diversity often become detectable in the short term, which is beneficial for monitoring the species' threats in a timely manner and formulating reasonable protection policies. Some researchers found that *C. japonica*, a related species of *C. azalea*, had a significant SGS and revealed the strength of SGS changed with time [24, 25], which contributed to better understand the dynamics of spatial genetic structure within natural populations. However, previous studies on *C. azalea* have not addressed the spatial distribution pattern of genetic diversity in this species. In addition, as an SPS, the nature of *C. azalea*, the first or last population of a species, remains unclear, and the pressure from human activities also lacks evaluation. In this study, we aimed to estimate comprehensively the genetic diversity, spatial age structure, and fine-grained genetic structure of this species by surveying more individuals with precise spatial positions in its geographical range to assess the impact of human disturbances, including road construction, dam building, etc., and whether the population can maintain regeneration. The above information would be helpful to understand the stage of this SPS and propose appropriate sustainable utilization and protection strategies for the only extant population.

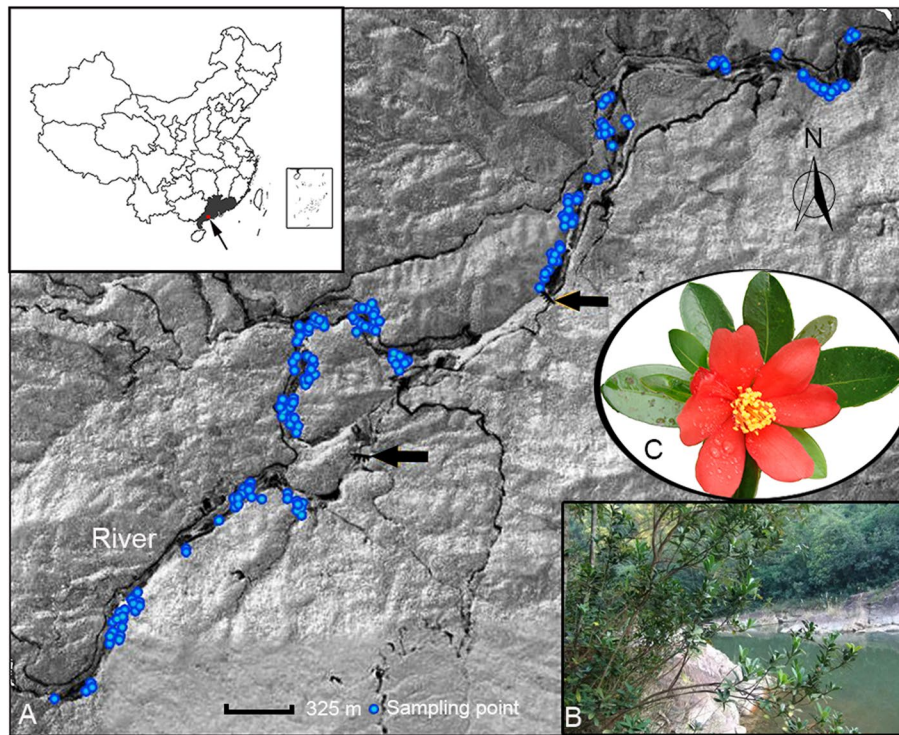


Fig. 1 Locations of sampled sites. **A** Circles represent distribution of sampling point. The direction of the river is from southwest (bottom left corner) to northeast (top right corner). Arrows represent dams (upstream dam (Dam1) were built in1986, downstream dam (Dam2) were built in1998). The population was divided into three subpopulations: SUB-U (individuals upstream from Dam1), SUB-M (individuals between Dam1 and Dam2), and SUB-D (individuals downstream from Dam2). **B** Habitat of *Camellia azalea*. **C** Unique flower and leaf morphology

Results

Geographic pattern and diameter structures of *C. azalea*

In total, 697 individuals of *C. azalea* were found in our survey, and among them, 629 individuals were positioned and measured accurately, and their precise spatial

positions were mapped (Fig. 1). As shown in Fig. 2, all individuals grew at two banks of a river with ~6000 m length and 5~91 m width; 96.6% of them (608 individuals) were within 10 m vertical to the edge of the river, 2.5% (16 individuals) grew in riffles or small islands of

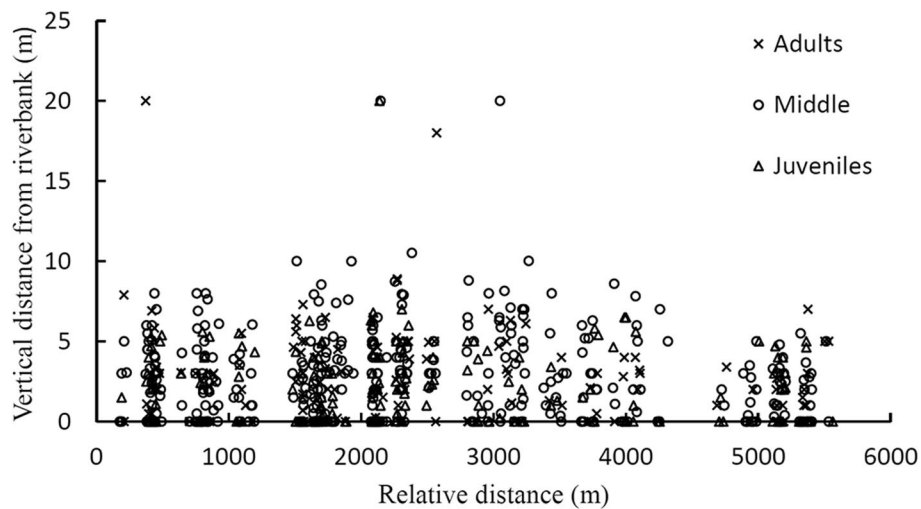


Fig. 2 Spatial distribution of *Camellia azalea* within the study plot. Crosses represent adult individuals (232 individuals); Triangles represent juveniles (203 individuals); Circles represent middle age individuals (262 individuals)

the river. Only four individuals (0.6%) were more than 20 m away from the edge vertically. The population was divided into three subpopulations by two dams, which from the upper to the down contained 180, 308, and 141 individuals, respectively. No individuals were distributed around the two dams. Based on their basal diameters, the diameter structures were shown in Fig. 3, and the number of middle individuals (2–3 cm in basal diameter) was accounted for ~40%, and adults and juveniles accounted for ~26% and ~34%, respectively.

Genetic diversity and differentiation within population

A set of nine SSR loci were used in this research, all the loci showed a high level of diversity. The average number of alleles per locus was 4.556~5.111, distributed among most sampled plants (Table 1). The percentage of null alleles was 0.15%~2.51%. Four loci showed deviation from Hardy–Weinberg equilibrium, and five loci did not deviate from Hardy–Weinberg equilibrium (Table S1). The observed heterozygosity ($H_o=0.509$) and expected heterozygosity ($H_e=0.497$) for total samples was similar to that for juveniles ($H_o=0.493$; $H_e=0.511$) and adults ($H_o=0.503$; $H_e=0.513$). The

mean fixation index (F_{IS}) of total samples was not significantly different from zero (0.0236). The values of Wright statistics (F_{ST}) between three subpopulations (from the upper to the down: SUB-U, SUB-M, and SUB-D) were from 0.0096 to 0.0128, and gene flow (N_m) was 19.33 to 25.85. The result of the principal component analysis (PCA) of total samples was shown in Fig. 4, in which the first principal component (PC) axes was 19.30%, the second was 19.11%, for a total of 38.41%. The first principal component (horizontal axis) divided all individuals into two groups with a significant genetic differentiation ($F_{ST}=0.077$), but three subpopulations according to spatial position (SUB-U, SUB-M, and SUB-D) did not gather, respectively.

All individuals of *C. azalea* were analyzed with STRU CTURE to infer their population genetic structure, and $K=2$ or 4 was the most plausible clustering solution of genetic population clusters (Fig. 5). There was no significant differentiation between the three subpopulations, and the genetic composition of all individuals was highly mixed. Mantel test showed a very weak correlation ($r=0.0058$, $P=0.002$) between distance and differentiation (Fig. S1).

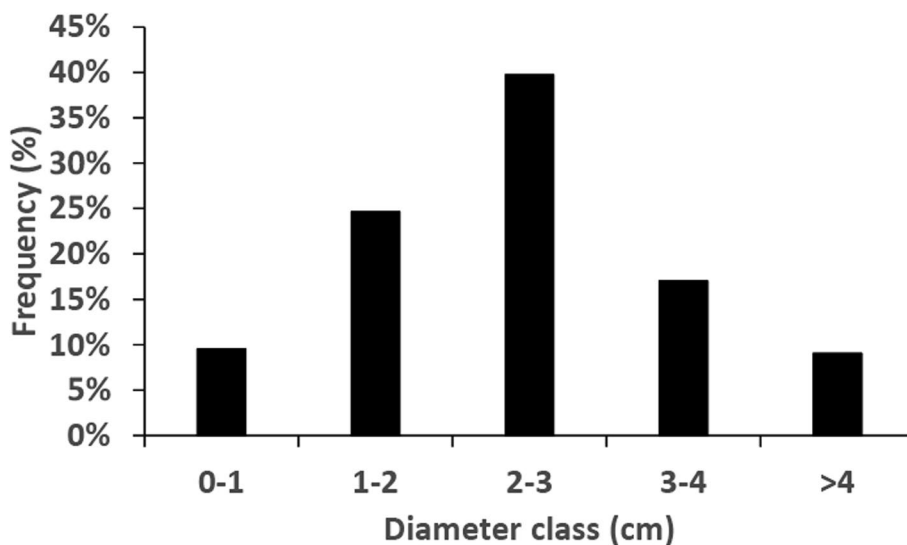


Fig. 3 Diameter class of *Camellia azalea*. The proportion of individuals with different diameter classes to the total number of individuals

Table 1 Genetic diversity parameters in *Camellia azalea* population

Sample	N	Na	Ne	Ho	He	I	F_{IS}	S_p
Juveniles	203	5.111	2.422	0.493	0.511	0.984	0.0126	0.0057
Adults	232	5.000	2.429	0.503	0.513	0.998	0.0082	0.0018
Total samples	677	5.111	2.416	0.497	0.509	0.989	0.0236	0.0059

N sample size, Na number of alleles, Ne effective number of alleles, Ho observed heterozygosity, He expected heterozygosity, I Shannon information index, F_{IS} fixation index, S_p the statistic that measures the extent of spatial genetic structure

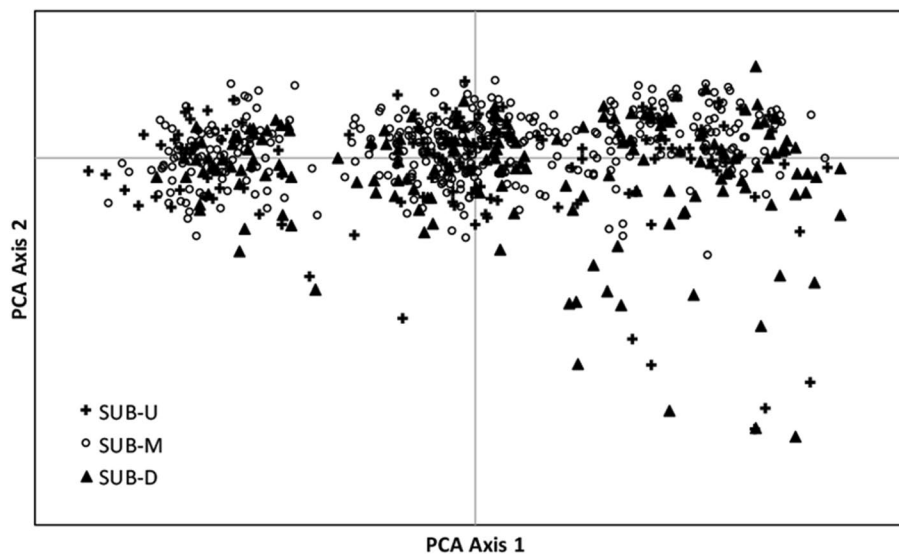


Fig. 4 Two-dimensional plot of the principal coordinates analysis (PCoA) of 629 *Camellia azalea* individuals. Trees with different distributions are represented by different symbols in the figure: Crosses represent SUB-U individuals (180 individuals); Circles represent SUB-M individuals (308 individuals); Triangles represent SUB-D individuals (141 individuals)

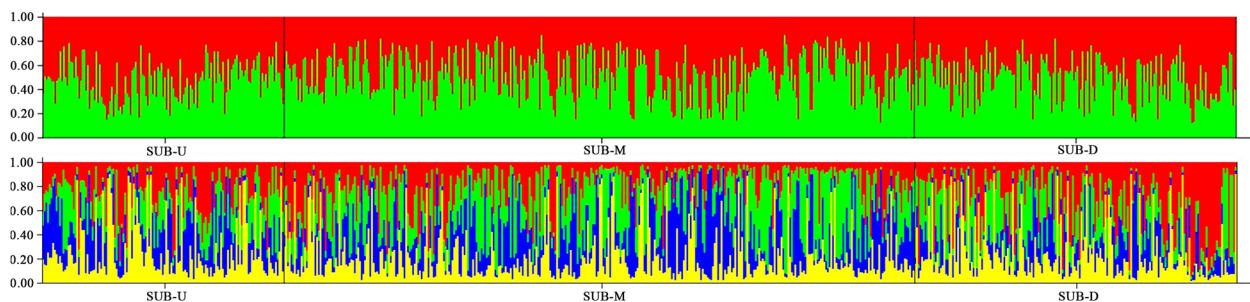


Fig. 5 Genetic structure bar plots of three *Camellia azalea* subpopulations (SUB-U, SUB-M and SUB-D). STRUCTURE for the optimal value of K estimated (Upper: K=2, Lower: K=4). Each solid bar represents a single individual, while colored areas correspond to distinct genetic clusters

Fine-scale spatial genetic structure

The results of the spatial autocorrelation analysis of all individuals were shown in Fig. 6, and the dotted line in the figure was the range of the 95% confidence interval. The significant positive correlation distance was detected within 75 m, and the intercept with the X-axis was 205 m. There is no significant spatial genetic structure over a further distance. Besides this, we conducted spatial autocorrelation analysis for the upper, middle, and down individuals, respectively, and the significant positive correlation distances of the three subpopulations (SUB-U, SUB-M, and SUB-D) are 10 m, 35 m, and 20 m, respectively. For different age groups, the significant positive correlation distance of the adults was 25 m, and the juveniles was 10 m (Fig. 6). The X-axis intercept also reflected the same trend.

Parentage analysis

Four individuals in the adult and juvenile groups found their parents with a confidence level of 80%, respectively. We estimated the most likely parents for all individuals, and a total of eight individuals found their parents (Table 2). In the adult group, the pollen dispersal distance ranged from 20 to 1137 m (445 m average), and the seed flow ranged from 1.5 to 200 m (95 m average). In the juvenile, the pollen flow ranged from 817 to 1687 m (1019 m average), and seed flow ranged from 1.2 to 134 m (60 m average).

Population demographics and effective population size

The Infinite Allelic Model (IAM), the Stepwise Mutation Model (SMM), and the Two-Phase Model (TPM) were used to conduct two tests on total samples: a sign

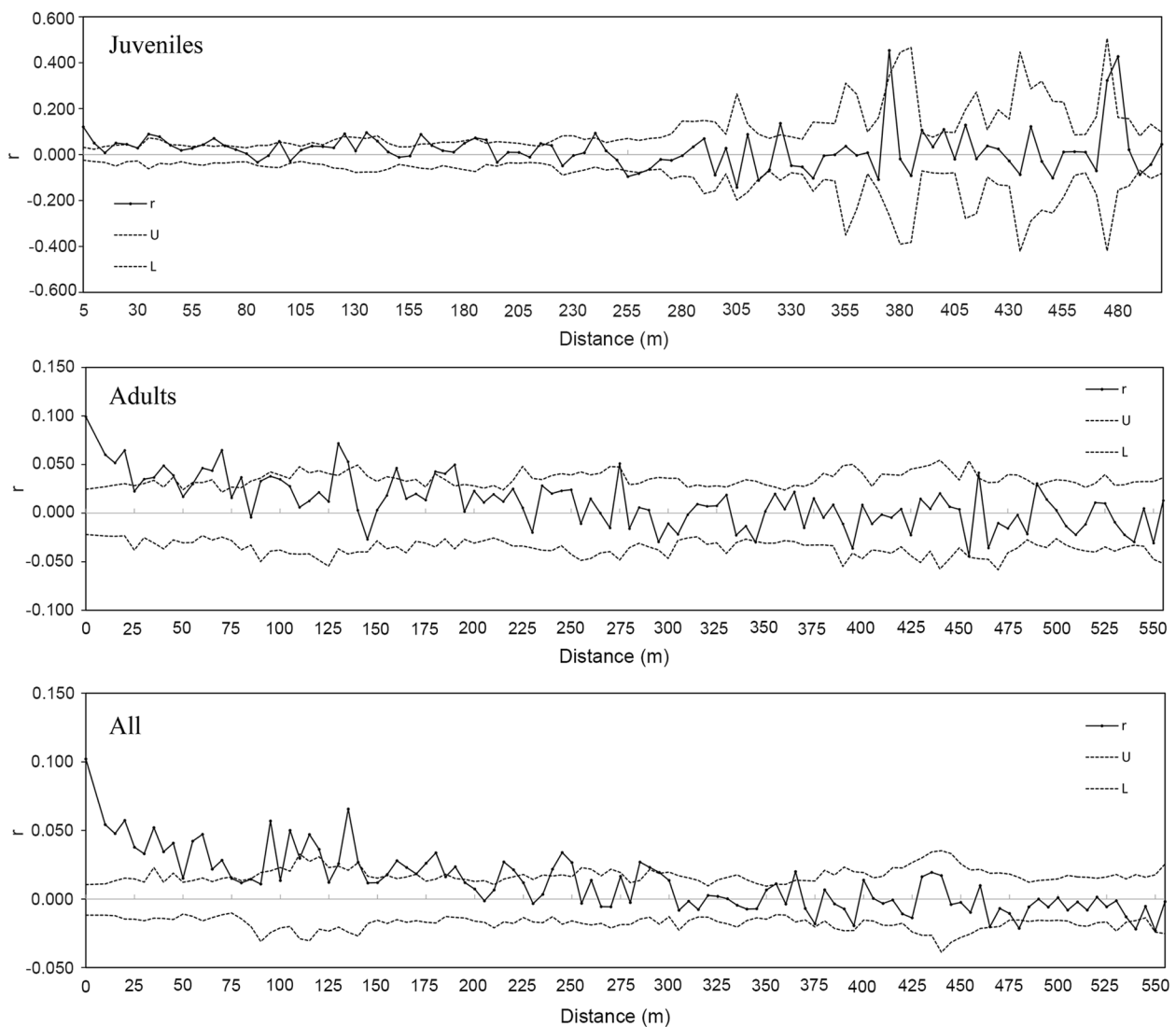


Fig. 6 Spatial autocorrelagrams of genetic variation obtained in adults (upper), juveniles (middle) and all individuals (lower). The dotted lines give the range of the 95% confidence interval and the solid line represents the estimate of average correlation coefficient. r: correlation coefficient

Table 2 Estimation of pollen and seed dispersal distances^a

	Individuals	Seed flow (m)		Pollen flow (m)	
		Max	Mean	Max	Mean
Juveniles	4	134	60	1687	1019
Adults	4	200	95	1137	445
Total samples	6	-	80	-	998

^a -80% credible interval for the estimates

test and the Wilcoxon signed-rank test. The *P* values were all higher than 0.05. That is the *C. azalea* population did not deviate from the mutation-drift balance and had not experienced the bottleneck effect recently.

Besides this, we tested the bottleneck effect of adults and juveniles, and the results were similar to those of the total population. All the tests under the above models did not detect the recent bottleneck effect. The results of mode-shift analysis showed that the allele frequency distributions of the whole population, the adult sub-population, and the juvenile sub-population were normal "L" shape distribution.

Since most of the individuals were distributed within 10 m of the river bank, we multiplied the length of the river (6000 m) by 16.5 m (both sides of the river) as its distribution area, $S = 99,000 \text{ m}^2$, dividing the number of all individuals by the area to obtain the density of the *C. azalea* population ($D = 0.007$). Table 3 exhibited the

Table 3 Neighborhood size and gene dispersion values at different effective population densities

	De									Average value
	D/2	D/3	D/4	D/5	D/6	D/7	D/8	D/9	D/10	
N_b	110	128	298	271	263	304	-	-	-	214
SE	58	343	155	785	1038	-	-	-	-	-
σ_g	50	67	116	124	134	156	-	-	-	98
SE	12	59	31	118	159	-	-	-	-	-

De effective population density, N_b neighborhood size, σ_g gene dispersal, SE Standard Error

- indicates that the value does not converge

neighborhood size and gene dispersal (σ_g) for different effective population densities of the population. In the first five gradients of effective population density (D_e), the neighborhood size and gene dispersal of *C. azalea* population all converge, the analyses entirely converged when D/6. Based on this, the effective population size was 176.

Discussion

Geographical distribution, age structure and population regeneration of *C. azalea*

For a SPS, the information on the habitat of every individual is important. SPS is the most likely to occupy a narrow habitat, otherwise, it should have more populations. In this study, we measured the spatial positions of 629 individuals and found that all individuals of this species grow in a valley of ~6000 m in length. This range has only an altitude difference of <120 m, and most of individuals found (99.1%) grow on opposite riparian zones of the river and are not far from 10 m to one of two river edges (Figs. 1 and 2). These results show this species indeed has a narrow niche, implying it is susceptible to environmental changes. In fact, this study found that this species has been affected by human activities to some extent, especially buildings of water dams and roads. Individuals grown in the riparian zone with roads were much fewer than those in the riparian zone without roads, and there were no individuals in a stretch of riverbank in front of each dam because after the dams were built, some suitable habitats were inundated by rising waters. Also, these disturbances affected the genetic structure of this species, and this effect will be discussed later.

Due to no seed immigration from other populations, population regeneration is greatly important for SPS's survival, which can be estimated by population age distribution. This index reflects the history of survival, reproduction, and potential for future growth of a population. Li et al. estimated the age structure of *C. azalea*, and concluded that the population was mainly composed of adults and was declining due to a serious lack of seedlings

[21]. The result was based on the data of 167 individuals. In this study, we surveyed the whole distribution of this species and measured nearly all the individuals that we could find (629 individuals), and a similar pattern was obtained (Fig. 3). However, our research indicated this species still maintains the ability to boost natural regeneration, because although the rate of juvenile individuals (0–1 cm in basal diameter) is only 10%, near 50% of individuals are in a period of vigorous reproduction. In our field investigation, we found many individuals can produce normal seeds, and those juvenile individuals (0–1 cm in basal diameter) should keep the viability of this population. Another reason is that the real rate of juvenile individuals was likely to be underestimated. There are ~1000 individuals in the population of *C. azalea*, but we only found 629 individuals in this study. We guess that most of the residual ~300 individuals should be juvenile because it was more difficult to find smaller individuals in a bushy forest than to find larger individuals in our field investigation. Thus, the real rate of juvenile individuals is more than 10%. In fact, in many forest communities, some species' seedlings are not at the highest rate, but it does not mean they begin to decline, and the appearance of a larger of seedlings needs to wait for some opportunities, such as canopy gaps, soil moisture, etc. [22, 23]. Meanwhile, according to the analysis of parent–child relationships, four individuals identified their parents. Although the detectable rate is low, it is still a sign that the population regeneration is continuing. Theoretically, due to the absence of gene flow with other populations and the large-scale felling of adult individuals, most seedlings in studied population should find their corresponding parents. However, only very few individual parents of seedlings have been found. There may be two reasons: 1) too few markers were used, so judgment cannot be made with high execution; 2) In the habitat shown in Fig. 1, in some regions before two dams and on the left bank of the river, we could not find any individuals due to constructing dams and roads. In addition, some adult plants were cut down before the establishment of the Natural Reserve

due to other reasons. If it is necessary to determine the parents of most juveniles, more markers will be needed to accurately calibrate the genotype of every individual. Based on these reasons, if the environment does not change dramatically, *C. azalea* could survive in the current habitat.

Fine-scale spatial genetic structure

The change of fine-scale SGS preceded the decrease of genetic diversity and supplied an early warning of conservation. *C. azalea* has been disturbed due to human construction of dams and roads, and its SGS information is necessary to evaluate its living conditions. Similar to other *Camellia* species, a significant spatial genetic structure was detected at the population level. The spatial distribution pattern of genetic variation is closely related to inbreeding, self-crossing, and cloning reproduction and limited gene flow [26–28], and for plants that can only reproduce sexually, the non-random distribution of gene flow is the main factor of the spatial genetic structure of population [29]. The direct effect of habitat fragmentation is the change in the effective dispersal of pollen or seeds in the population. Although the pollen dispersal mode of *C. azalea* is similar to *C. japonica*, animal-dispersed, different from the latter whose seeds were dispersed mainly by gravity, the former's seeds were also spread by the river, meaning they could be carried farther by water. This inference is supported by the fact that the average seed spread distance of *C. azalea* population (up to 80 m) is much higher than that of other *Camellia* species.

For species with overlapping generations, the spatial genetic structure should focus on age class to reveal the whole picture accurately [24]. For example, Ueno et al. revealed a strong SGS for juveniles in *C. japonica* and no significant genetic structure for adults in *C. japonica* with no human interference [25]. Chung et al. demonstrated the internal spatial genetic structure found in younger classes but dissipated in older age classes [24]. In this study, the SGS showed a significant positive correlation distance of 25 m for adults but 10 m for juveniles, and the intense of fine-scale spatial genetic structure (measured by S_p statistics) revealed a slightly stronger fine-scale spatial genetic structure in juveniles ($S_p=0.0057$) compared to adults ($S_p=0.0018$).

There are several reasons for this result, such as density, breeding system, cloning reproduction, and limited gene flow. The density effect can be excluded in this species since the distribution of juveniles and adults was relatively even. In the absence of clonal reproduction and selfing, limited gene flow may be a reasonable explanation for the formation of fine-scale spatial genetic structure. Pollen flow of the juvenile stage (1019 m) was much

greater than that of the adult stage (445 m), but the seed spread distance based on the juvenile stage individuals (60 m) was smaller than that based on the adult stage individuals (95 m), probably due to the limitation of seed propagation by dams. The stronger SGS in juveniles revealed that dams affect the gene flow to some extent. However, the difference in SGS between the young group and the adult group (S_p value of the former is greater than that of the latter) may be due to natural factors such as self-thinning. Considering the influence of many factors that weaken the SGS intensity, SGS, as a sensitive marker of genetic consequences, should be paid close attention to.

Genetic diversity and effective population size

Genetic diversity is both the result of long-term evolution and represents the potential for species evolution [7, 13, 30]. The Nei's indices of the *C. azalea* population ($H_o=0.493$; $H_e=0.511$; $I=0.989$) indicated it possessed a moderate to high level of genetic diversity. Unexpectedly, the genetic diversity of a single population of *C. azalea* was not significantly deviation from other *Camellia* species [31, 32] and even more than that of endemic species of flowering plants ($H_e=0.42$) [33]. Previous studies have suggested that a common feature of rare species is a lower level of genetic diversity [34]. However, recent studies have found that some endangered or narrowly distributed species also have higher levels of genetic variation [34, 35]. High genetic diversity at the species level is often consistent with its self-incompatibility breeding system [36]. Its fixation index at the species level ($F_{IS}=0.0236$) was not significantly different from zero, indicating *C. azalea*, like other *Camellia* species, is highly outcrossing [37]. Especially, when there is a lack of gene flow between populations, its outcrossing mating system can reserve rare alleles in the population and consequently keep a high genetic diversity. The rich genetic diversity of *C. azalea* might be sufficient to withstand a certain extent of external interference.

Effective population size is one of the important factors affecting genetic and population variation [38]. Populations with an effective population size of 50 individuals have been considered the minimum limit for maintaining sufficient allele richness; while the population with an effective population size greater than 500 individuals can maintain genetic variations in quantitative traits within the population and their adaptability to future environmental changes [39]. The effective population size of *C. azalea* was 176, and this value is larger than the minimum limit but less than 500. Therefore, although *C. azalea* is a SPS and has suffered some disturbances, it will not go extinct in a very short time if its habitat and all individuals can be conserved strictly from now.

Conservation strategy and fate of SPS

As a SPS, the extant population is unlikely the starting population of *C. azalea*, because no closely related species were found in its sympatric or neighboring region and it was different from any known *Camellia* species in morphology. It was implausible that the ancestral species immediately become extinct when a new species establishes its first population. According to previous studies, *C. azalea* has a long evolutionary history, its DNA sequence has a large divergence with other *Camellia* species [40, 41].

The fate of this SPS in the future has several possibilities (Fig. 7). Based on the genetic characteristics of *C. azalea*, it is foreseeable that with sufficient saplings and rich genetic diversity, this species will continue to live as a single population. Under effective protection measures, this SPS meets the conditions for the establishment of multi-population. If the environment changes dramatically and there is a lack of effective protection measures, in addition to the narrow habitat of *C. azalea*, the population may decline inevitably. So, the effective conservation measures (*ex-situ* conservation) are critical to *C. azalea*'s fate.

For any SPS, strict conservation measures are necessary. However, specific measures need to be formulated according to the genetic diversity, niche, and age structure of the species. As a SPS, *C. azalea* had a relatively high genetic diversity and regeneration capacity. Thus, it can survive and evolve if its habitat can be conserved. However, its small geographical range, fewer wild individuals, and narrow niche make it very sensitive to human interference and environmental changes. The adoption of strict protective measures to conserve its habitat and field individuals is the most important for its survival, and constructing new populations by transporting seeds or seedlings to other similar valleys in the

Nature Reserve is also another way to save this species from extinction. Information on the SGS of this species can help us to select appropriate individuals to obtain seeds in constructing new populations with high genetic diversity. To limit consanguinity, the seeds should be obtained from trees at least 80 m apart. It is very important to investigate the niche of SPS. A narrow niche may be a common feature of SPS, but it needs a lot of research to confirm. This information is critical to assess the survival status of SPS.

In the same Nature Reserve, except *C. azalea*, there are other SPS and more endangered plant species (eg. *Euryodendron excelsum*, *Paraboeica chunii* and *Apteroperma oplate* [31]; and the finding on *C. azalea* should have important implications for conserving these species. In addition, the streamside or riparian zone in valley is a special category of habitats, and is characterized by a high level of species-richness. Many endangered *Camellia* plants also grow along streams or riparian zone, such as *C. euphlebica*, *C. pyxidicea*, *C. fluviatilis*, *C. huana*, etc. [42]. Although these species are not SPS, their ranges are very narrow and their natural populations are declining. Results in this study show the river plays an important role in the survival of *C. azalea*. Although the water dams in the river did not cause the intraspecific genetic differentiation, but they destroyed some suitable habitats of *C. azalea* and changed its SGS by limiting seed flows. For perennial woody plants, the influence of environmental disturbance on their genetic diversity often lags, and it takes several generations to show up. Therefore, in addition to the comprehensive investigation of genetic diversity and habitat changes, attention should also be paid to the spatial genetic structure after habitat fragmentation to formulate the most reasonable protection strategy. New conservation strategies should be put forward to ensure that rivers can spread seeds smoothly, and

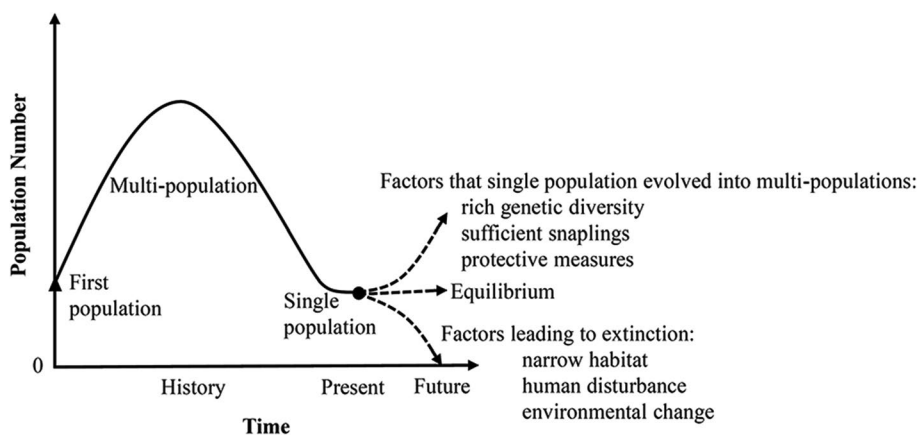


Fig. 7 Dynamics of population number. Possible evolutionary direction and corresponding influencing factors of single population species

the knowledge of *C. azalea* will help to conserve similar species.

Conclusion

Single population is a necessary stage in the evolutionary progress of most species, the first or the last population. This study investigated the spatial distribution, age structure, genetic diversity, fine spatial genetic segments, population historical dynamics, and the effects of human activities on a true single-population species, *C. azalea*. This species has a long evolutionary history, and the extant population is most likely the last population rather than the first. We found that this species possessed a moderate to high level of genetic diversity, however, its current ecological niche is very narrow, with most individuals distributed on the bank of a stream. Recent human activities have changed gene flow of this species and thus probably contributed to the strengthening of SGS in juveniles. Considering that this species has maintained a moderate level of genetic diversity, it could face near-term extinction if effective conservation measures continue to be implemented, especially to protect its riverine habitat and avoid human exploitation. These results are of enlightening significance to the conservation of other SPS, especially those distributed along the riparian zone.

Materials and methods

Study sites and sampling distribution

C. azalea is a small tree (shrub like small tree), most individuals are less than 3 m in height and 3 cm in diameter. It has red flowers and an out-crossing breeding system, and is a typical insect-pollinated and self-incompatible [20]. Its seeds spread by water flow mainly. The research site is part of the Ehuangzhang Nature Reserve (21°50′36″~21°58′40″ N, 111°21′29″~111°36′03″ E) in Yangchun, Guangdong Province. The species is distributed in the range of 6 km along a river (about 20 m in width, Fig. 1B) in an altitudinal range from 50 to 175 m. There are two dams for hydropower stations in the river. The two dams were built in 1986 (upstream, dam1 in Fig. 1) and 1998 (downstream, dam2), respectively. We first identified one individual's geographical coordinates using GPS (Global Positional System) as benchmark (21°54′40″ N, 111°30′46″ E) and then measured other individuals' relative positions to the benchmark. We surveyed carefully the range of the species along the banks of the river, and measured every individual we could find, including the basal diameter (BD), plant height, and distance from the river bank.

For sample collection, fresh leaves of 697 individuals of wild *C. azalea* were collected and dried in silica, respectively (Table S2). The samples were identified by Wen-Ju

Zhang, and the voucher specimens were preserved in the Institute of Biodiversity Science of Fudan University. Field study and sample collection were conducted with the permission of the Ehuangzhang Nature Reserve, and complied with the institutional and national guidelines. To explore the impact of the dams on *C. azalea* population, *C. azalea* individuals were divided into adults (SUB-A) and juveniles (SUB-J) groups. We sampled 232 adults (pre-fragmentation, BD > 3 cm, 4.25 cm average) and 203 juveniles (post-fragmentation, BD < 2 cm, 1.52 cm average) individuals (Table S3), respectively, simulating the groups formed before dam construction and the groups formed after dam construction. To reduce the error caused by the age division, only individuals with larger and smaller basal diameter were taken, and the intermediate individuals (262 individuals) were discarded. In addition, for testing the effect of the dams to the genetic structure, we spatially divided the population into three subpopulations: the upstream population (SUB-U), which includes all individuals above dam1, and the mid-stream population (SUB-M) which includes individuals between the two dams, and downstream group (SUB-D) which includes all individuals below dam2.

DNA extraction and microsatellite genotyping

Genomic DNA was extracted by the cetyltrimethylammonium bromide (CTAB) method [43]. According to the polymorphism records, 69 pairs of primers with polymorphism in other *Camellia* species were selected, and 12 pairs of primers were selected for PCR amplification and fluorescence capillary electrophoresis, and finally a total of nine pairs of SSR primers with polymorphism in *C. azalea* were selected and fluorescently labeled at the 5'-end. The primer sequences and other information are shown in Supplementary Material 1.

Genetic diversity analysis

Mean number of expected heterozygosity (H_e) and observed heterozygosity (H_o) per locus were calculated by the Nei's method. Gene flow was estimated using Genetix 4.05 software [44]. FSTAT 2.9.3.2 [45] software and Weir & Cockerham's calculation standard were used to estimate Wright's (1965) F statistics (F_{IS} , F_{ST}). The effective alleles and Shannon's information index were estimated using PopGen 32 [46]. Genepop 4.1.4 [47] was used to assess whether per SSR locus and the population as a whole deviated significantly from the Hardy–Weinberg equilibrium. The null allele frequency was estimated by FreeNA [48].

In this study, GenAlEx 6.41 [49] was used to perform PCA (Principal Component Analysis) cluster analysis on the genetic distance data of *C. azalea* to reveal the relationship between different individuals in the population

and whether these relationships are related to geographical distribution or other factors.

To verify whether a genetic bottleneck, three microsatellite mutation models in Bottleneck [50] software were used for bottleneck analysis and mode-shift test: the infinite allele model (IAM), the stepwise mutation model (SMM), and the two-phase model (TPM). Two different tests, the sign test and the Wilcoxon signed rank test, were employed.

Analyses of genetic structure

The most likely parental detection of all *C. azalea* individuals was analyzed by Cervus 3.0 software [51]. With the help of the allele frequency of the observed site, 10,000 cycles of simulations were performed to estimate the likelihood value ratio (LOD value) of each candidate parent of the offspring. The one with the highest LOD value was the “real” parent. The difference between the largest and second largest LOD values is defined as the D value, and the significance of the D value is tested by comparing the D value with the critical value. The significance level of the D value is expressed by confidence level (CL). There are generally two types of credibility, namely 95% (strict CL, equivalent to the significance level of $P=0.05$) and 80% (relaxed CL) [52]. Among the most likely parents, the individual closer to the seed was considered to be the female parent, while the farther plant was considered to be the male parent. The distance between offspring and the female parent is the seed flow, and the distance between parents is the pollen flow. The phylogenetic correlation coefficient (F_{ij}), the population gene spread value σ_g , and the neighborhood size (N_b) were calculated using the software SPAGeDi 1.3 [53]. Effective population size calculation formula: $N_e = N_b * S / Sg = N_b * S / (4\pi\sigma_g^2)$, where S is the area of the population distribution.

STRUCTURE 2.2 software [54, 55] was used to estimate the Bayesian genetic clustering algorithm. The parameters set are as follows: admixture model and correlated allele frequencies, burn-in length is 10,000 times, Markov's chain Monte Carlo (MCMC) times are set as 100,000 times, $K=1-12$, and iteration is 10 Times. The ΔK value is then based on the ad hoc statistics presented by Evanno et al. [56], choose the most realistic genetic type, calculated as follows: $\Delta K = m [|L(K+1) - 2L(K) + L(K-1)|] / S[L(K)]$. Where m is the average of $\ln P(D)$ in repeated runs, and S is the variance of $\ln P(D)$; the K value corresponding to the highest ΔK value is the true number of genetic types. Finally, based on the calculated K values, a bar plot was obtained. GenALEX (version 6.41) was used to conduct Mantel test based on 629 individuals [49]. S_p statistics were calculated using the formula:

$$S_p = -b_{\log} / (1 - F1)$$

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12870-024-05886-9>.

Supplementary Material 1.

Supplementary Material 2.

Acknowledgements

Ziyi Chen, Yi Xu, and Yongju Chen helped to collect some materials, Jing Xu and Helian Song conducted some works in SSR analysis.

Authors' contributions

WRQ, ZWJ and XY initiated and designed the research; ZWJ obtained funding for the study; XY, HG, ZQM, LXX collected materials and performed the experiments, WRQ, ZM, XY, YB, LXF analyzed the data and wrote the paper. All authors read, edited and agreed to submit the manuscript.

Funding

This study was financially supported by the Key Research and Development Program of Guangdong Province (Grant Number: 2018B0202020002), the National Natural Science Foundation of China (Grant Number: 31670223), and SZPU Research Project-6024330001 K.

Data availability

The SSR genotype and phenotype data for the accessions included in this study are provided in the supplementary materials (Table S2 and S3).

Declarations

Ethics approval and consent to participate

All methods for the use of plants comply with relevant institutional, national, and international guidelines and legislation.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 13 August 2024 Accepted: 27 November 2024

Published online: 30 December 2024

References

1. Tilman D, Clark M, Williams DR, Kimmel K, Polasky S, Packer C. Future threats to biodiversity and pathways to their prevention. *Nature*. 2017;546(7656):73–81.
2. Chen LL, Hui C, Lin ZS. Habitat destruction and the extinction debt revisited: the Allee effect. *Math Biosci*. 2009;221(1):26–32.
3. Pimm SL, Jenkins CN, Abell R, Brooks TM, Gittleman JL, Joppa LN, et al. The biodiversity of species and their rates of extinction, distribution, and protection. *Science*. 2014;344(6187):1246752–1246752.
4. Barnosky AD, Matzke N, Tomiya S, Wogan GOU, Swartz B, Quental TB, et al. Has the earth's sixth mass extinction already arrived? *Nature*. 2011;471(7336):51–7.
5. Thomas CD, Cameron A, Green RE, Bakkenes M, Beaumont LJ, Collingham YC, et al. Extinction risk from climate change. *Nature*. 2004;427(6970):145–8.
6. Caballero A, Garcia-Dorado A. Allelic diversity and its implications for the rate of adaptation. *Genetics*. 2013;195(4):1373–84.

7. Agashe D, Falk JJ, Bolnick DI. Effects of founding genetic variation on adaptation to a novel resource. *Evolution*. 2011;65(9):2481–91.
8. Jump AS, Marchant R, Penuelas J. Environmental change and the option value of genetic diversity. *Trends Plant Sci*. 2009;14(1):51–8.
9. Epperson KB. Estimating dispersion from short distance spatial autocorrelation. *Heredity*. 2005;95(1):7–15.
10. Ewédjè EBK, Ahanchédé A, Hardy OJ. Breeding system, gene dispersal and small-scale spatial genetic structure of a threatened food tree species, *Pentadesma butyracea* (Clusiaceae) in Benin. *Conserv Genet*. 2017;18(4):799–811.
11. Kalisz S, Nason JD, Hanzawa FM, Tonsor SJ. Spatial population genetic structure in *Trillium grandiflorum*: the roles of dispersal, mating, history, and selection. *Evolution*. 2001;55(8):1560–8.
12. De-Lucas AI, González-Martínez SC, Vendramin GG, Hidalgo E, Heuertz M. Spatial genetic structure in continuous and fragmented populations of *Pinus pinaster* Aiton. *Mol Ecol*. 2009;18(22):4564–76.
13. Evans SM, Sinclair EA, Poore AG, Bain KF, Vergés A. Genotypic richness predicts phenotypic variation in an endangered clonal plant. *Peer J*. 2016;4:e1633.
14. Volis S, Deng T. Importance of a single population demographic census as a first step of threatened species conservation planning. *Biodivers Conserv*. 2020;29(2):527–43.
15. Ye CX. A new species of *Camellia*. In: The Collection of Paper Abstracts of 1985 Annual Academic Congress of Botanical society in Guangdong province. *Bull Bot Soc Guangdong Prov*. 1986;3:25
16. Wei CF. A new species of *Camellia* from China. *Bull Bot Res*. 1986;6:141–4.
17. Li KK, Li LL, Jennifer T, Li CR, Dai SP, Ye CX, et al. *Camellia changii*, a legitimate and correct name with priority. *Ecol Sci*. 2009;28(05):424–7
18. Ren H, Jian S, Chen YJ, Liu H, Zhang Q, Liu N, et al. Distribution, status, and conservation of *Camellia changii* Ye (Theaceae), a critically endangered plant endemic to southern China. *Oryx*. 2014;48(3):358–60.
19. Rodriguez JP, Rodriguez-Clark KM, Baillie JE, Ash N, Benson J, Boucher T, et al. Establishing IUCN red list criteria for threatened ecosystems. *Conserv Biol*. 2001;25(1):21–9.
20. Luo XY, Zhuang XY, Yang YS. Genetic diversity of *Camellia changii* Ye (Theaceae) using ISSR markers. *J Trop Subtrop Bot*. 2007;15(2):93–100.
21. Li XL, Sun ZY, Li JY, Yin HF, Fan ZQ, Xu C, et al. Population structure and dynamic change of endangered plant *Camellia azalea*. *J Plant Res Environ*. 2018;27(02):17–23.
22. Hubbell SP, Foster RB, O'Brien ST, Harm KE, Condit R, Wechsler B, et al. Light-gap disturbances, recruitment limitation, and tree diversity in a neotropical forest. *Science*. 1999;283(5401):554–7.
23. Norghauer JM, Newbery DM. Seed fate and seedling dynamics after masting in two African rain forest trees. *Ecol Mongr*. 2011;81(3):443–69.
24. Chung MY, Epperson BK, Chung MG. Genetic structure of age classes in *Camellia japonica* (Theaceae). *Evolution*. 2003;57(1):62–73.
25. Ueno S, Tomaru N, Yoshimaru H, Manabe T, Yamamoto T. Size-class differences in genetic structure and individual distribution of *Camellia japonica* L. in a Japanese old-growth evergreen forest. *Heredity*. 2002;89(2):120–6.
26. Sokal RR, Jacquez GM, Wooten MC. Spatial autocorrelation analysis of migration and selection. *Genetics*. 1989;121(4):845–55.
27. Shapcott A. The spatial genetic structure in natural populations of the Australian temperate rainforest tree *Atherosperma moschatum* (Labill.) (Monimiaceae). *Heredity*. 1995;74(1):28–38.
28. Reusch TBH, Hukriede W, Stam WT, Olsen JL. Differentiation between clonal growth and limited gene flow using spatial autocorrelation of microsatellites. *Heredity*. 1999;83(2):120–6.
29. Ennos RA. Inferences about spatial processes in plant populations from the analysis of molecular markers. In: Silvertown J, Antonovics J, editors. *Integrating ecology and evolution in a spatial context*. Oxford: Blackwell Scientific; 2001. p. 45–71.
30. Agbani PO, Kafoutchoni KM, Salako KV, Gbedomon RC, Kégbé AM, Karen H, et al. Traditional ecological knowledge-based assessment of threatened woody species and their potential substitutes in the Atakora mountain chain, a threatened hotspot of biodiversity in Northwestern Benin, West Africa. *J Ethnobiol Ethnomed*. 2018;14:21.
31. Xu W, Dong LJ, Duan JH, Li SJ, Zhang SG. Genetic diversity and relationship of 84 tea cultivars (*Camellia sinensis* (L.) O. Kuntze) by SSR markers. *J Hunan Agric Univ*. 2011;37(3):260–6.
32. Jia BG, Lin Q, Zhang L, Tan X, Lei X, Hu X, et al. Development of 15 genic-SSR markers in oil-tea tree (*Camellia oleifera*) based on transcriptome sequencing. *Genetika*. 2014;46(3):789–97.
33. Nybom H. Comparison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants. *Mol Ecol*. 2004;13(5):1143–55.
34. Gitzendanner MA, Soltis PS. Patterns of genetic variation in rare and widespread plant congeners. *Am J Bot*. 2000;87(6):783–92.
35. Gonzalez-Astorga J, Castillo-Campos G. Genetic variability of the narrow endemic tree *Antirhea aromatica* Castillo-Campos Lorence, (Rubiaceae, Guettardeae) in a tropical forest of Mexico. *Ann Bot*. 2004;93(5):521–8.
36. Vekemans X, Hardy OJ. New insights from fine-scale spatial genetic structure analyses in plant populations. *Mol Ecol*. 2004;13(4):921–35.
37. Neel MC, Ross-Ibarra J, Ellstrand NC. Implications of mating patterns for conservation of the endangered plant *Eriogonum ovalifolium* var. *vineum* (Polygonaceae). *Am J Bot*. 2001;88(7):1214–22.
38. Crow JF, Kimura M. An introduction to population genetics theory. Caldwell: Blackburn Press; 1970.
39. Franklin IR. Evolutionary change in small populations. In: M. E. Soule ME, Wilcox BA, Editors. *Conservation Biology: An Evolutionary-Ecological Perspective*. Sunderland: Sinauer Associates Inc.; 1980. p. 135–149.
40. Vijayan K, Zhang WJ, Tsou CH. Molecular taxonomy of *Camellia* (Theaceae) inferred from nrITS sequences. *Am J Bot*. 2009;96(7):1348–60.
41. Zan T, He YT, Zhang M, Yonezawa T, Ma H, Zhao QM, et al. Phylogenomic analyses of *Camellia* support reticulate evolution among major clades. *Mol Phylogenet Evol*. 2023;182: 107744.
42. Ming TL. A systematic synopsis of the genus *Camellia*. *Acta Bot Yunnanica*. 1998;21:149–59.
43. Doyle JJ, Doyle JL. A rapid DNA isolation for small quantities of fresh tissue. *Phytochem Bull*. 1987;19:11–5.
44. Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F. GENETIX, Version 4.04 logiciel sous Windows TM pour la Génétique des populations. In: *Laboratoire Génome, Populations, Interactions CNRS UMR 5000. Université Montpellier II, Montpellier*; 2002.
45. Goudet J. FSTAT (version 1.2): a computer program to calculate F-statistics. *J Hered*. 1995;86(6):485–6.
46. Yeh FC, Yang RC. Popgene, the user-friendly shareware for population genetic analysis. *Molecular Biology and Biotechnology Centre, University of Alberta, Canada*. 1997;10:295–301.
47. Raymond M, Rousset F. GENEPOP: population genetics software for exact tests and ecumenicism. *J Hered*. 1995;86:248–9.
48. Chapuis MP, Estoup A. Microsatellite null alleles and estimation of population differentiation. *Mol Bio Evol*. 2007;24(3):621–31.
49. Peakall R, Smouse PE. GENALEX 6: genetic analysis in excel. Population genetic software for teaching and research. *Mol Ecol Notes*. 2006;6(1):288–95.
50. Piry S, Luikart G, Cornuet JM. Bottleneck: a computer program for detecting recent reductions in the effective population size using allele frequency data. *J Hered*. 1999;90:502–3.
51. Kalinowski ST, Taper ML, Marshall TC. Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Mol Ecol*. 2007;16:1099–106.
52. Marshall TC, Slate J, Kruuk LEB, Pemberton JM. Statistical confidence for likelihood-based paternity inference in natural populations. *Mol Ecol*. 1998;7(5):639–55.
53. Hardy OJ, Vekemans X. SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Mol Ecol Notes*. 2002;2(4):618–20.
54. Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. *Genetics*. 2000;155(2):945–59.
55. Falush D, Stephens M, Pritchard JK. Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Mol Ecol Notes*. 2007;7(4):574–8.
56. Evanno G, Regnaut S, Goudet J. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol*. 2005;14(8):2611–20.

Publisher's Note

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