



## **Genetic Diversity and Population Structure of Rhododendron longipedicellatum, an Endangered Species**

Authors: Cao, Yurong, Ma, Yongpeng, Li, Zhenghong, Liu, Xiongfang, Liu, Detuan, et al.

Source: Tropical Conservation Science, 15(1)

Published By: SAGE Publishing

URL: <https://doi.org/10.1177/19400829221078112>

---

BioOne Complete ([complete.BioOne.org](https://complete.BioOne.org)) is a full-text database of 200 subscribed and open-access titles in the biological, ecological, and environmental sciences published by nonprofit societies, associations, museums, institutions, and presses.


Your use of this PDF, the BioOne Complete website, and all posted and associated content indicates your acceptance of BioOne's Terms of Use, available at [www.bioone.org/terms-of-use](https://www.bioone.org/terms-of-use).


Usage of BioOne Complete content is strictly limited to personal, educational, and non - commercial use. Commercial inquiries or rights and permissions requests should be directed to the individual publisher as copyright holder.

---

BioOne sees sustainable scholarly publishing as an inherently collaborative enterprise connecting authors, nonprofit publishers, academic institutions, research libraries, and research funders in the common goal of maximizing access to critical research.

# Genetic Diversity and Population Structure of *Rhododendron longipedicellatum*, an Endangered Species

Tropical Conservation Science  
Volume 15: 1–10  
© The Author(s) 2022  
Article reuse guidelines:  
[sagepub.com/journals-permissions](https://sagepub.com/journals-permissions)  
DOI: 10.1177/19400829221078112  
[journals.sagepub.com/home/trc](https://journals.sagepub.com/home/trc)  


Yurong Cao<sup>1</sup> , Yongpeng Ma<sup>2</sup>, Zhenghong Li<sup>1</sup>, Xiongfang Liu<sup>1</sup>, Detuan Liu<sup>2</sup>, Suping Qu<sup>3</sup>, and Hong Ma<sup>1,4</sup>

## Abstract

*Rhododendron longipedicellatum* is an endangered species endemic to southeastern Yunnan, China. Assessment of genetic variation is critical for protecting endangered species. Therefore, we used EST-SSR markers to analyze the genetic characteristics of *R. longipedicellatum*. The results revealed high genetic diversity at the species level ( $H_e = 0.559$ ,  $N_A = 9.529$ ) and within populations ( $H_e = 0.507$ ,  $N_A = 5.910$ ) and moderate genetic differentiation between populations ( $F_{ST} = 0.083$ ). In addition, more genetic variation existed within populations (91.25%) compared with variation among populations (8.75%). The STRUCTURE analysis showed that 150 individuals from five existing populations could best be divided into two genetic groups. At the population level, the neighbor-joining (NJ) tree and principal coordinates analysis (PCoA) analyses also divided them into two groups. In addition, Bottleneck analyses using the Two-Phase Model (TPM) and Stepwise Mutation Model (SMM) as well as the Garza-Williamson Index revealed widespread signatures of bottleneck events. These results provide vital information for scientifically formulating conservation strategies for the endangered *R. longipedicellatum*.

## Keywords

conservation genetics, molecular markers, EST-SSR, endangered species, bottleneck effect

## Introduction

Genetic diversity is a key component of biodiversity. The study of genetic diversity and population structure of species is the theoretical basis for scientific and rational utilization of germplasm resources and improvement of genetic characteristics, as well as the prerequisite for effective conservation measures (Frankham, 2003; Mable, 2019). With the rapid development and continuous improvement of techniques in molecular biology, molecular markers have been widely used for population genetic diversity analysis, breeding programs, and evolutionary research (Ma et al., 2021; Yadav et al., 2019). Among the molecular marker types, simple sequence repeat (SSR) markers, including genomic SSR (gSSR) and expressed sequence tag SSR (EST-SSR), are highly regarded because of their rich content distribution, high polymorphism, ease of use, and repeatability (Liu et al., 2019; Wei et al., 2011). Compared with gSSR, EST-SSR have a higher level of transferability to related species, are more evolutionarily conserved, have lower development costs, and closer relationships with genes of known functions (Singh et al., 2013; Zhang et al., 2017).

*Rhododendron* is the largest genus in Ericaceae, with the widest distribution in the Northern Hemisphere, consisting of eight subgenera and more than 1025 species. China has about 581 species from six subgenera, of which 421 species are

<sup>1</sup>Institute of Highland Forest Science, Chinese Academy of Forestry, Kunming, China

<sup>2</sup>Yunnan Key Laboratory for Integrative Conservation of Plant Species with Extremely Small Populations, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming, China

<sup>3</sup>Flower Research Institute, Yunnan Academy of Agricultural Sciences, Kunming, China

<sup>4</sup>Key Laboratory of Breeding and Utilization of Resource Insects, National Forestry and Grassland Administration, Kunming, China

Received: 21 April 2021; accepted: 18 January 2022

### Corresponding Author:

Hong Ma, Institute of Highland Forest Science, Chinese Academy of Forestry, Kunming, China.

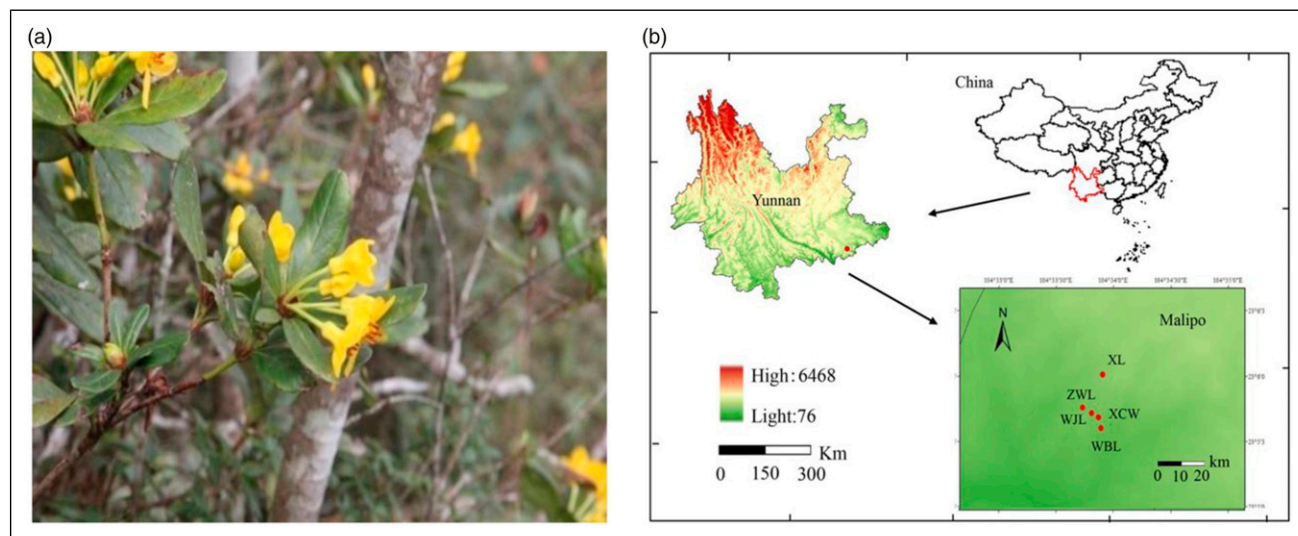
Email: [hortsience@163.com](mailto:hortsience@163.com); Suping Qu, Flower Research Institute, Yunnan Academy of Agricultural Sciences, Kunming, China.

Email: [qsp@yaas.org.cn](mailto:qsp@yaas.org.cn)



Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (<https://creativecommons.org/licenses/by-nc/4.0/>) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE

and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).



**Figure 1.** *R. longipedicellatum* (A) and the geographical distribution of the five extant populations of *R. longipedicellatum* (B). Note: ZWL, XL, WJL, XCW, and WBL are population codes.

endemic to China (Cai et al., 2016; Ma et al., 2017). China is the most important modern distribution and evolution center of *Rhododendron* species, but the loss and destruction of wild resources has become a serious threat in recent years, and many endemic and rare endangered species are extinct or on the verge of extinction (Ma et al., 2014). Although much research has been conducted on the genetic structure and population dynamics of *Rhododendron*, few studies have been conducted on endangered species whose habitats are seriously threatened. Thus, conservation research on endemic and endangered *Rhododendron* species is urgently needed.

*Rhododendron longipedicellatum* is endemic to Southeast Yunnan, China (Cai et al., 2016). Its beautiful tree shape, bright pure yellow flowers, and unusual flowering time give this species a high value for development and utilization in landscaping. Since 2014, our study group has found only five wild populations with less than 1000 plants remaining in limestone habitats in Malipo County, Yunnan Province at an altitude of approximately 1183–1316 m. Due to the proximity to human settlements and high ornamental value, additional habitat loss and removal of wild plants is a relatively serious threat to conserving these remaining populations. Furthermore, within the narrow geographic distribution, the number of populations and individual plants is small, few true seedlings are present within each population, and natural regeneration is weak. At present, the wild, undeveloped habitat available to sustain populations such as ZWL and XL is continuously decreasing (Supplementary Figure S1). According to the IUCN *Red List Categories and Criteria* (IUCN, 2001), *R. longipedicellatum* is already extremely endangered (CR B1ab [i, iii, v]) and requires urgent protection. Thus, this study used 17 previously developed high-polymorphism EST-SSR markers (Li, Liu, Li, Wan,

et al., 2018) to study the genetic characteristics of 150 individuals from five residual populations of *R. longipedicellatum*, attempting to assess the following: (1) genetic diversity and structure; (2) gene flow between populations; (3) correlation between genetic distance and geographical distance; and (4) bottleneck effects. Findings are expected to illustrate the possible genetic consequences of residual small populations in the limestone habitat so as to provide a reference for effective conservation and germplasm resource management of *R. longipedicellatum*.

## Methods

### Experimental Materials

Young leaves of 150 healthy mature plants (30 samples from each population) were collected from five residual populations. Locality and voucher information for samples are provided in Supplementary Table S1 and Figure 1. Samples were collected at a distance of at least 2.5 m to avoid collecting clone-mates. The collected leaf tissues were quickly dried with silica gel and stored at  $-20^{\circ}\text{C}$  until further use.

### DNA Extraction and EST-SSR Genotyping

A modified cetyltrimethylammonium bromide (CTAB) method was used to extract genomic DNA (Doyle & Doyle, 1987). In brief, gel-dried fragments of leaves (0.1 g) were placed in liquid nitrogen, quickly ground to a powder, and then mixed with 1 mL of CTAB buffer (100 mM Tris-HCl, pH 8.0, 20 mM EDTA, 2% CTAB) with 0.1%  $\beta$ -mercaptoethanol added immediately prior to use. The mixture was incubated at  $65^{\circ}\text{C}$  for 60 min and centrifuged for 2 min at 13,000 rpm, after which 600  $\mu\text{L}$  of the supernatant was transferred to a 1.5 mL microcentrifuge tube. The supernatant was mixed twice with

**Table 1.** Genetic Diversity Indices for Five Populations of *R. longipedicellatum*.

| Pop              | $N_A$ | $N_E$ | $H_o$ | $H_e$ | $I$   | $F_{IS}^a$ | Null              | $N_m$ |
|------------------|-------|-------|-------|-------|-------|------------|-------------------|-------|
| WBL              | 4.647 | 2.146 | 0.448 | 0.432 | 0.860 | -0.037     | 0.01              | -     |
| WJL              | 5.235 | 2.817 | 0.581 | 0.545 | 1.099 | -0.066     | 0.00              | -     |
| XCW              | 6.177 | 2.844 | 0.631 | 0.596 | 1.207 | -0.059     | 0.01 <sup>†</sup> | -     |
| XL               | 5.647 | 2.610 | 0.626 | 0.568 | 1.123 | -0.102*    | 0.00              | -     |
| ZWL              | 5.118 | 2.177 | 0.465 | 0.461 | 0.917 | -0.009     | 0.02 <sup>†</sup> | -     |
| Population level | 5.910 | 2.744 | 0.515 | 0.507 | 1.101 | -0.016     | 0.01              | 3.373 |
| Species level    | 9.529 | 2.897 | 0.550 | 0.529 | 1.244 | -0.040     | 0.00              | 2.762 |

$N_A$ , average number of alleles;  $N_E$ , effective number of alleles;  $H_o$ , observed heterozygosity;  $H_e$ , expected heterozygosity;  $I$ , Shannon's diversity index;  $F_{IS}$ , inbreeding coefficient; Null, null allele frequency averaged across the 17 loci,

† indicates the significance of null alleles in the full model, which accounts for null alleles (n), inbreeding (f) and genotyping failures (b) simultaneously;  $N_m$ , gene flow.

<sup>a</sup>significant deviation from Hardy-Weinberg equilibrium after sequential Bonferroni corrections: \* $p < 0.05$ .

phenol/chloroform/isoamyl alcohol (25:24:1) and mixed for 10 min, followed by centrifugation for 10 min at 12,000 rpm. DNA was precipitated from the aqueous phase by adding 0.1 mL of 3 M sodium acetate and 2.5 mL of ethanol. The DNA precipitate was washed twice with 70% ethanol and dissolved in water. DNA quantity was assessed using 0.8% agarose gel electrophoresis (Supplementary Figure S2).

All samples were genotyped using 17 EST-SSR markers previously developed for this species (Supplementary Table S1) (Li, Liu, Li, Wan, et al., 2018). For all markers, PCR amplification (three replicates of each sample) was performed on an ABI 3730 Thermal Cycler (Thermo Fisher Scientific, Waltham, Massachusetts, USA). PCR amplifications were performed in a final volume of 10  $\mu$ L, containing 1  $\mu$ L (10–30 ng) of template DNA, 5  $\mu$ L of 0.7 $\times$  Multiplex PCR Master Mix (QIAGEN, Hilden, Germany), 0.5  $\mu$ L (10 pM) of each primer, and 3  $\mu$ L of RNase-free water. The PCR program was as follows: 5 min at 94°C; 30 cycles of 30 s at 94°C, annealing at 58°C for 30 s, and 30 s at 72°C; and a final extension step at 72°C for 5 min. The PCR products were pooled and purified using the Wizard SV Gel and PCR Clean-Up System (Promega), following the manufacturer's instructions. The 5' end of each forward primer for the 17 markers was tagged with one of three fluorescent dyes (FAM, HEX, or ROX), and multiplex PCR amplifications were performed for the individuals of *R. longipedicellatum* using the PCR conditions described above. The fluorescence signal detected by capillary electrophoresis was interpreted using the GeneMapper v4.1. All individual-specific bands corresponding to each primer were read out to establish the original data matrix. Data editing and format conversion were carried out in Convert v1.31 and Genalex v6.2.

## Data Analysis

The frequency of null alleles was estimated for each locus in each population using INEST 2.2 (Chybicki & Burczyk, 2009). Genepop v4.2 was used to calculate Linkage disequilibrium (LD) for all locus pairs and Hardy-Weinberg equilibrium (HWE) in each population, with deviations from expected proportions evaluated with a Bonferroni corrected  $p$ -value threshold of 0.05

(Rice, 1989). The genetic diversity indices were calculated in PopGene v1.31 (Yeh et al., 1999), including the average number of alleles ( $N_A$ ), the number of effective alleles ( $N_E$ ), Shannon information index ( $I$ ), observed heterozygosity ( $H_o$ ), expected heterozygosity ( $H_e$ ), inbreeding coefficient ( $F_{IS}$ ).

We adopted the Markov Chain Monte Carlo (MCMC) method and used STRUCTURE v2.3.4 (Pritchard et al., 2000) to evaluate genetic structure and shared ancestry across the five putative population groups. The analysis was performed using the admixture model of ancestry, which assumes independent origins of each allele, with operating parameter settings "Length of Burnin period" and "Number of MCMC Reps after Burnin" in the running length set to 30,000 and 100,000, respectively. The  $K$  value range was set at 1–20, and 20 independent runs were conducted for each  $K$  value. The most likely number of clusters was evaluated using both  $\Delta K$  and the log-likelihood value with STRUCTURE HARVESTER v0.6.94 (Earl & VonHoldt, 2012; Evanno et al., 2005). Following this, cluster analysis based on shared allele distance (DAS) was performed using Population v1.2.30 (Langella, 2007) with the neighbor-joining (NJ) method with 1000 bootstraps. The clustering tree was edited using the FigTree v1.4.0. Based on Nei's genetic distance (Nei, 1973), principal coordinates analysis (PCoA) of populations was conducted with ape package v5.4 (Paradis & Schliep, 2019). The genetic distance and geographic distance matrices between populations were computed using PopGene v1.31 and Geographic Distance Matrix Generator v1.2.2, respectively, and the Mantel test was carried out with PASSaGE v2 (999 permutations were calculated).

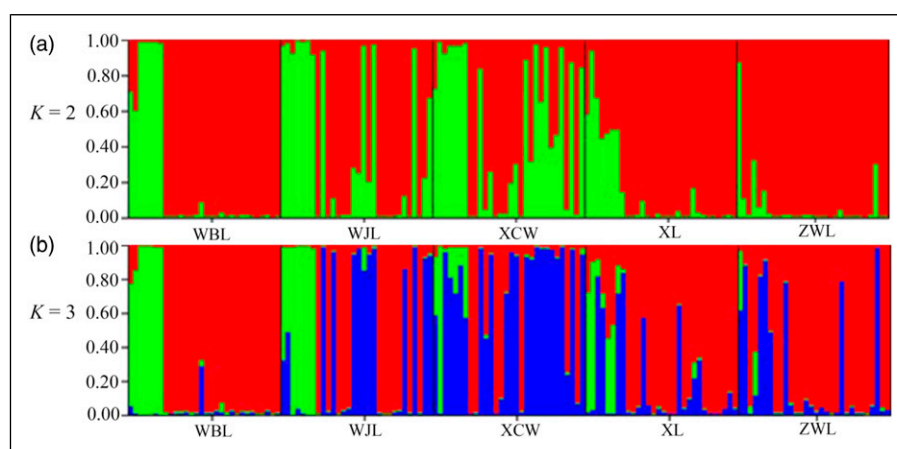
Analysis of molecular variance (AMOVA) was carried out to estimate the partitioning of genetic variation between and within the populations using Arlequin v3.5 (Excoffier & Lischer, 2010). The genetic differentiation between populations was evaluated by calculating the genetic differentiation index ( $F_{ST}$ ; Weir & Cockerham, 1984) using FSTAT v2.9.3 (Goudet, 2001). Gene flow ( $N_m$ ) was calculated by the formula  $N_m = ([1/F_{ST}] - 1)/4$  (Slatkin, 1985). The Garza-Williamson index (G-W) was calculated using Arlequin v3.5 (Excoffier & Lischer, 2010). The data were assessed for a population genetic bottleneck signature using

**Table 2.** Analysis of Molecular Variance (AMOVA) based on 17 EST-SSR Markers for Populations of *R. longipedicellatum*.

| Groups                  | Source of Variation | d.f | Sum of Squares | Variance Components | Percentage of Variation | F        |
|-------------------------|---------------------|-----|----------------|---------------------|-------------------------|----------|
| Whole                   | Among populations   | 4   | 120.617        | 0.428Va             | 8.75                    | 0.0875** |
|                         | Within populations  | 295 | 1317.217       | 4.465Vb             | 91.25                   | -        |
|                         | Total               | 299 | 1437.833       | 4.893               | -                       | -        |
| Group 1<br>WBL, ZWL, XL | Among populations   | 2   | 64.022         | 0.464Va             | 9.95                    | 0.099**  |
|                         | Within populations  | 177 | 742.533        | 4.195Vb             | 90.05                   | -        |
|                         | Total               | 179 | 806.556        | 4.659               | -                       | -        |
| Group 2<br>WJL, XCW     | Among populations   | 1   | 19.533         | 0.244Va             | 4.78                    | 0.048**  |
|                         | Within populations  | 118 | 574.683        | 4.870Vb             | 95.22                   | -        |
|                         | Total               | 119 | 594.217        | 5.115               | -                       | -        |

F, fixation Index.

\*\* $p < 0.01$ , extremely significant difference.



**Figure 2.** Results of Bayesian clustering of 150 individuals of *R. longipedicellatum*. Note: (A) Assignment of 150 individuals into  $K = 2$  genetically distinguishable groups. (B) Assignment of 150 individuals into  $K = 3$  genetically distinguishable groups. Each individual is represented by a vertical bar colored according to the assigned group(s).

the Sign test and Wilcoxon test under the Stepwise Mutation Model (SMM) and Two-Phase Model (TPM) in Bottleneck v1.2 (Cornuet & Luikart, 1997; Piry et al., 1999). The variance for TPM was set to 30, and the proportion of SMM in TPM was 70%, with 1000 iterations. The remaining relevant parameters were set as the software default values.

## Results

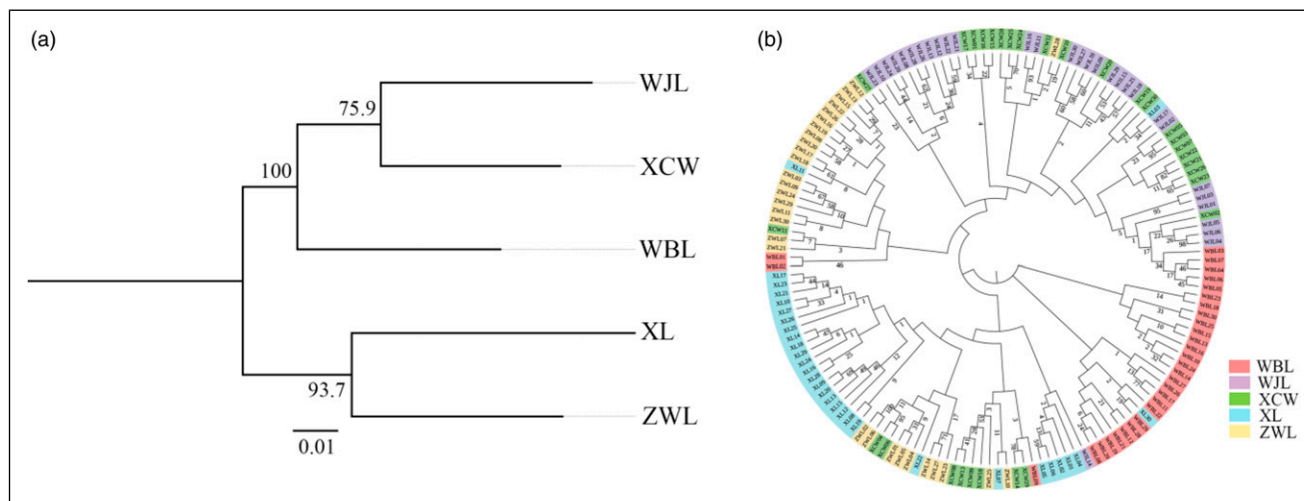
### Genetic Diversity

Using INEST, the frequency of null alleles was estimated to be lower than the threshold of 0.05 at each of the 17 markers across populations (Table 1) (Dakin & Avise, 2004; Y. Li et al., 2019). No consistent genotypic disequilibrium was found between any two loci across all populations; therefore, all loci were used for further analyses. Significant deviation from HWE was only detected in one of the five populations ( $p < 0.05$ , after Bonferroni correction; Table 1). At the population level, the five *R. longipedicellatum* populations showed high genetic diversity (Table 1). The average number of alleles per locus was high (9.529) across all populations, while individual population values ranged from 4.647 to 6.177.

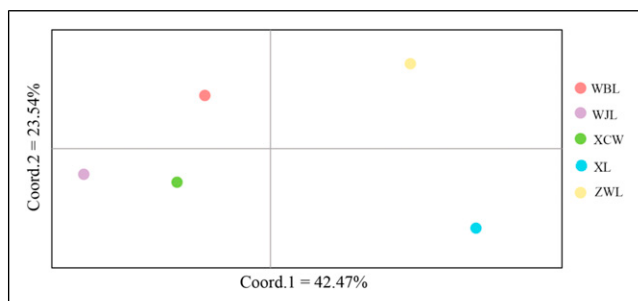
The average number of alleles was higher in the pooled population data set because unique alleles were present in some of the populations. Shannon's information index as a measure of gene diversity was similar across populations, ranging from 1.207 to 0.860. In general, the population XCW contained the greatest level of genetic diversity, and WBL the least.

### Genetic Structure

The results of the AMOVA analysis of the five populations are presented in Table 2. Of the total genetic variation, 91.25% occurred within populations, and only 8.75% occurred between populations. The genetic differentiation value of 0.083 was within the range of moderate differentiation. The STRUCTURE analysis showed that  $\Delta K$  reached its maximum value ( $\Delta K = 310.98$ ) at  $K = 2$  (Supplementary Figure S3), suggesting that the 150 individuals from the five populations most likely belong to two principal genetic clusters (Figure 2(A)). Cluster one mainly included individuals from the WBL, WJL, and XCW populations, whereas cluster two mainly included individuals from the XL and ZWL populations. The STRUCTURE results at  $K = 3$  were consistent with those of



**Figure 3.** Neighbor-Joining (NJ) dendrogram of five extant populations and 150 individuals of *R. longipedicellatum* based on DAS genetic distance. The bootstrap values are indicated at the nodes in each cluster.



**Figure 4.** Principal coordinates analysis (PCoA) of five *R. longipedicellatum* populations based on a genetic distance matrix of 17 EST-SSR loci. Note: The first and second axes extracted 42.47% and 23.54% of the total genetic variance, respectively.

the NJ tree (Figure 2(B)), and there were mixed individuals among the populations, indicating some gene flow or differentiation.

Based on the DAS genetic distance, NJ trees of populations and individuals were constructed, as shown in Figure 3. The genetic relationship between the WJL and XCW populations was close, clustering as a clade, and the XL and ZWL populations clustered into another clade. The WBL population formed a separate clade and belonged to a large branch along with the WJL and XCW populations (Figure 3(A)). The NJ tree of the 150 individuals (Figure 3(B)) revealed some cross-mixing among the three groups and among the populations. The PCoA analysis of the five populations is shown in Figure 4. Approximately 66.01% of the variation occurred between principal coordinates one and two, and their contribution rates were 42.47% and 23.54%, respectively. The results were similar to the NJ tree, and the five populations were roughly divided into two groups. The genetic distance and geographical distance among the five populations were positively correlated

(Mantel test,  $r = 0.77$ ,  $t = 2.95$ ,  $p < 0.01$ ). This result is consistent with the NJ tree and two-dimensional PCoA map.

### Bottleneck Effect Analysis

The results of the bottleneck analysis for the five residual populations based on the SMM and TPM are shown in Table 3. Assuming that all of the EST-SSR loci conformed to the SMM model, the Sign test and Wilcoxon test detection methods showed that all five populations experienced heterozygote deficit or excess and a certain degree of genetic bottleneck. Based on the TPM model, five populations all experienced heterozygote excess, and the ZWL population showed a significant deviation from mutation-drift equilibrium in Wilcoxon test ( $p < 0.01$ , after Bonferroni correction). The G-W values of the five populations ranged from 0.22 to 0.26 with an overall value of 0.26 (Table 3), which was lower than the critical value of 0.68 (Garza & Williamson, 2001) and consistent with a population decline in all five populations. At the species level, the results of the different detection methods consistently indicated a genetic bottleneck.

## Discussion

### Genetic Diversity

The evolutionary potential of a species and its ability to withstand adverse environments depend on both the genetic diversity of the species and its population genetic structure (Stebbins, 1950). Different levels of genetic variation and genetic structure of populations can be attributed to multiple factors, such as breeding systems, biological characteristics, evolutionary history, and natural selection (Hamrick & Godt, 1996; Oyundelger et al., 2021; Zhou et al., 2020). The results

**Table 3.** Bottleneck Analysis for Five Populations of *R. longipedicellatum*.

| Pop           | Mutation Model          |                       |                         |                       | G-W  |
|---------------|-------------------------|-----------------------|-------------------------|-----------------------|------|
|               | SMM                     |                       | TPM                     |                       |      |
|               | Sign Test ( $H_E/H_D$ ) | Wilcoxon Test ( $p$ ) | Sign Test ( $H_E/H_D$ ) | Wilcoxon Test ( $p$ ) |      |
| WBL           | 6/1                     | 0.003*                | 7/5                     | 0.278                 | 0.24 |
| WJL           | 3/9                     | 0.177                 | 7/5                     | 0.529                 | 0.26 |
| XCW           | 7/5                     | 0.060                 | 9/3                     | 0.582                 | 0.23 |
| XL            | 4/8                     | 0.001*                | 8/4                     | 0.436                 | 0.25 |
| ZWL           | 2/10                    | 0.000*                | 10/2                    | 0.009*                | 0.22 |
| Species level | 8/4                     | 0.000*                | 6/1                     | 0.005*                | 0.26 |

SMM, Stepwise Mutation Model; TPM, Two-Phase Model; G-W, Garza-Williamson Indices;  $H_E/H_D$ , number of loci with heterozygosity excess/number of loci with heterozygosity deficiency.

\* indicates  $p < 0.01$  after Bonferroni correction.

of this EST-SSR study showed that *R. longipedicellatum* has high genetic diversity at the species level ( $H_e = 0.559$ ,  $N_A = 9.529$ ), and the mean genetic diversity of the five populations investigated ( $H_e = 0.507$ ,  $N_A = 5.910$ ) was higher than that of *Rhododendron oldhamii* ( $H_e = 0.284$ , ranged 0.240–0.336) (Hsieh et al., 2013) and the *Rhododendron pseudochrysanthum* complex (average  $H_e = 0.424$ ) (Chen et al., 2014), but lower than the genetic diversity of *Rhododendron jinggangshanicum* ( $H_e = 0.642$ ) (M. Li et al., 2015) and *Rhododendron decorum* ( $H_e = 0.758$ ) (Wang et al., 2013) based on gSSR markers. Although gSSR markers vary more, possibly due to a higher mutation rate, the variation observed in EST-SSR markers may be more functionally significant and represent responses to selection more effectively (Ellis & Burke, 2007; Kane & Rieseberg, 2007).

Despite the restricted distribution and small population sizes of *R. longipedicellatum*, this species has retained high levels of genetic diversity, possibly due to the following reasons. The extensive temporal and spatial heterogeneity of the limestone habitat in which these plants occur may contribute to their genetic diversity. The limestone habitat includes variation in light, temperature, rock type, degree of differentiation, fissure soil, karst water, humus layer thickness, and the species composition and density of the vegetation. Some *R. longipedicellatum* plants occurring in the limestone habitat are attached to tree trunks, providing a stable yet polymorphic microenvironment that reduces the impact of natural and human-induced disturbance. Second, outcrossing plants tend to have higher genetic variation than selfing plants (Huang et al., 2021; Nybom, 2004). The cross-pollination reproductive system of *R. longipedicellatum* reported by Li, Liu, Li, Ma, et al. (2018) may, therefore, also support the relatively high genetic diversity in this species. The high genetic variability of *R. longipedicellatum* may also be attributed to abundant diversity in the ancestral population (Zhang et al., 2021) that may have existed as a large, continuous, and widespread species, although little is

known about the origin and evolutionary history of *R. longipedicellatum*. The current restricted distribution and small population sizes of the species result from habitat deterioration and loss, as well as direct human interference. *R. longipedicellatum* is a perennial woody plant whose biological characteristics may also play a role in the maintenance of genetic diversity. In the five populations of *R. longipedicellatum* studied, few or no seedlings were observed and the plants usually propagated vegetatively, which tends to maintain genetic diversity.

### Genetic Structure

According to AMOVA, 91.25% of the total genetic variation occurred within populations, and only 8.75% was inter-population variation, which is consistent with the observation that most of the genetic variation in woody plants occurs within populations (Liu et al., 2012; X. Liu et al., 2020). In this study, the overall  $F_{ST}$  was 0.083, indicating moderate differentiation among populations (Jakobsson et al., 2013) due to drift or possibly divergent selection among patchily-distributed populations caused by human disturbance or habitat heterogeneity (Liu et al., 2019; Willi et al., 2006). The STRUCTURE analysis divided 150 individuals from the five existing populations into two genetic populations, and the NJ tree and PCoA analyses also divided the populations into two groups. This clustering of populations may be due to local environmental conditions, which often shape patterns of genetic structure, especially in heterogeneous and fragmented habitats (Young et al., 1996; Zhu et al., 2016). Both at the species and the population levels,  $F_{IS}$  was less than 0 and the cross-pollination rate was greater than 1, indicating that outbreeding was present in the population, contrary to the high inbreeding coefficient expected in a small population of plants occurring in a unique habitat and subjected to an increasingly shrinking geographical distribution. The obvious homozygote deficiency or heterozygote excess may be associated with population regeneration in some plants

(Stoeckel et al., 2010). In addition, the increased heterozygosity may be due to the small number of populations with few individual plants in each population or to the presence of a high allelic richness in the parent plants (Huang et al., 2021; Pudovkin et al., 1996).

The magnitude of gene flow between populations has a significant influence on the genetic structure of populations and is a key factor that predicts the effect of changes in the environment, such as human interference and population isolation (Wu et al., 2017). In this study, the estimated number of migrants per generation among populations of *R. longipedicellatum* was  $>1$ , indicating that gene flow was high enough to overcome differentiation among populations by genetic drift (Wright, 1949). The diffusion and transmission of pollen and seeds are the two major forms of gene exchange between populations. Similar to the majority of *Rhododendron* species (Ng & Corlett, 2000; Yang et al., 2020), our field surveys found that the seed transmission capability of *R. longipedicellatum* was very limited, seedlings in the populations were rare, and the pollen dispersal depended mainly on male *Bombus braccatus* (Li, Liu, Li, Ma, et al., 2018), with a propagation distance of 3–10 km (Ng & Corlett, 2000). All of the study populations of *R. longipedicellatum* were geographically close, and even the relatively distant populations, WBL and ZWL, were only about 9 km apart. Thus, pollen flow by pollinators is expected.

### Population Bottlenecks

Bottleneck v1.2 identified genetic bottlenecks in *R. longipedicellatum* using different models and detection methods. The results showed that at the species level, under the assumptions of the SMM and TPM models, *R. longipedicellatum* significantly deviated from mutation-drift equilibrium, and at the population level, the ZWL population also deviated from equilibrium ( $p < 0.01$ , after Bonferroni correction). Recent damage caused by human activities, such as grazing and indiscriminate felling, has increased, especially in the ZWL population. The *M*-ratio test reflected a loss of allelic diversity in all five populations, consistent with population contraction and supporting the Bottleneck results. As bottlenecks tend to cause reduction the quality of the species gene pool and genetic variability of populations and consequently, their fitness. The high genetic diversity observed in *R. longipedicellatum* in this study may be due to the occurrence of the genetic bottleneck being relatively recent.

### Implications for Conservation

These results indicate that genetic diversity is still present, to a certain degree, in *R. longipedicellatum*. To preserve this species, an effective conservation measure could be to strengthen the habitat protection of *R. longipedicellatum*, strictly forbid indiscriminate felling, and protect it from man-made interference. All five populations contained a high level

of genetic diversity, with the XCW population having the richest genetic diversity and the highest number of effective alleles. The connectivity and mixing between XCW and other populations were higher than those between XCW and other populations, which indicated that XCW is of great significance in maintaining genetic polymorphism. This population is geographically located in the middle of the other populations and thus beneficial for facilitating dispersal and connecting the other populations as a corridor. Therefore, we suggest that additional conservation efforts be focused on the XCW population. In addition, protection efforts can be enhanced by appealing to the related local authorities, setting up conservation warning signs, and publicizing the value of, and conservation strategy for, this endangered species to local residents.

### Acknowledgments

We would like to thank all the funding for their support. We would also like to thank Editage ([www.editage.com](http://www.editage.com)) for English language editing. Finally, we are particularly grateful to the anonymous referees for their time and their valuable comments.

### Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

### Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Fundamental Research Funds for the Central Non-profit Research Institution of the Chinese Academy of Forestry (CAFYBB2019ZB007); the Biodiversity Survey and Assessment Project of the Ministry of Ecology and Environment, China (2019HJ2096001006); the Ten Thousand Talent Program of Yunnan Province (YNWR-QNBJ-2019-010; YNWR-QNBJ-2018-174) and the Young and middle-aged academic and technical leaders in Yunnan Province reserve talents (2018HB066).

### ORCID iD

Yurong Cao  <https://orcid.org/0000-0003-3448-697X>

### Supplemental material

Supplemental material for this article is available online.

### References

- Cai, L., Neilsen, J., Dao, Z. L., & Ma, Y. P. (2016). *Rhododendron longipedicellatum* (Ericaceae), a new species from South-eastern Yunnan, China. *Phytotaxa*, 282(4), 296–300. DOI:10.11646/phytotaxa.282.4.7. <https://doi.org/10.11646/phytotaxa.282.4.7>
- Chen, C. Y., Liang, B. K., Chung, J. D., Chang, C. T., Hsieh, Y. C., Lin, T. C., & Hwang, S. Y. (2014). Demography of the upward-shifting temperate woody species of the *Rhododendron*



- pseudochrysanthum* complex and ecologically relevant adaptive divergence in its trailing edge populations. *Tijdschrift voor Gerontologie en Geriatrie*, 10(1), 111–126. <https://doi.org/10.1007/s11295-013-0669-x>.
- Chybicki, I. J., & Burczyk, J. (2009). Simultaneous estimation of null alleles and inbreeding coefficients. *Journal of Heredity*, 100(1), 106–113. <https://doi.org/10.1093/jhered/esn088>.
- Cornuet, J. M., & Luikart, G. (1997). Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics*, 144(4), 2001–2014. <https://doi.org/10.1093/genetics/144.4.2001>.
- Dakin, E. E., & Avise, J. C. (2004). Microsatellite null alleles in parentage analysis. *Heredity*, 93(5), 504–509. <https://doi.org/10.1038/sj.hdy.6800545>.
- Doyle, J., & Doyle, J. (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem Bull*, 19(1), 11–15. <https://doi.org/10.12691/ajmr-3-1-7>.
- Earl, D. A., & VonHoldt, B. M. (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, 4(2), 359–361. <https://doi.org/10.1007/s12686-011-9548-7>.
- Ellis, J. R., & Burke, J. M. (2007). EST-SSRs as a resource for population genetic analyses. *Heredity*, 99(2), 125–132. <https://doi.org/10.1038/sj.hdy.6801001>.
- Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software structure: a simulation study. *Molecular Ecology*, 14(8), 2611–2620. <https://doi.org/10.1111/j.1365-294X.2005.02553.x>.
- Excoffier, L., & Lischer, H. E. (2010). Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, 10(3), 564–567. <https://doi.org/10.1111/j.1755-0998.2010.02847.x>.
- Frankham, R. (2003). Genetics and conservation biology. *Comptes Rendus Biologies*, 326(1), 22–29. [https://doi.org/10.1016/S1631-0691\(03\)00023-4](https://doi.org/10.1016/S1631-0691(03)00023-4).
- Garza, J. C., & Williamson, E. G. (2001). Detection of reduction in population size using data from microsatellite loci. *Molecular Ecology*, 10(2), 305–318. <https://doi.org/10.1046/j.1365-294X.2001.01190.x>.
- Goudet, J. (2001). FSTAT, a program to estimate and test gene diversities and fixation indices. <http://www.unil.ch/izea/software/FSTat.html>. Version 2.9.3
- Hamrick, J. L., & Godt, M. J. W. (1996). Effects of life history traits on genetic diversity in plant species. *Philosophical Transactions of the Royal Society B-Biological Sciences*, 351, 1291–1298. <http://www.jstor.org/stable/56204>.
- Hsieh, Y. C., Chung, J. D., Wang, C. N., Chang, C. T., Chen, C. Y., & Hwang, S. Y. (2013). Historical connectivity, contemporary isolation and local adaptation in a widespread but discontinuously distributed species endemic to Taiwan, *Rhododendron oldhamii* (Ericaceae). *Heredity*, 111(2), 147–156. <https://doi.org/10.1038/hdy.2013.31>.
- Huang, R., Wang, Y., Li, K., & Wang, Y. Q. (2021). Genetic variation and population structure of clonal *zingiber zerumbet* at a fine geographic scale: a comparison with two closely related selfing and outcrossing *zingiber* species. *BMC Ecology Evolution*, 21(1), 116. <https://doi.org/10.1186/s12862-021-01853-2>.
- IUCN (2001) *IUCN red list Categories and Criteria: version 3.1* (30). IUCN Species Survival Commission.
- Jakobsson, M., Edge, M. D., & Rosenberg, N. A. (2013). The relationship between  $F_{ST}$  and the frequency of the most frequent allele. *Genetics*, 193(2), 515–528. <https://doi.org/10.1534/genetics.112.144758>.
- Kane, N. C., & Rieseberg, L. H. (2007). Selective sweeps reveal candidate genes for adaptation to drought and salt tolerance in common sunflower, *Helianthus annuus*. *Genetics*, 175(4), 1823–1834. <https://doi.org/10.1534/genetics.106.067728>.
- Langella, O. (2007). *Population genetic software (individuals or populations distances, phylogenetic trees)*. Populations 1.2.30. Available online at: <http://bioinformatics.org/~tryphon/populations/>.
- Li, M. W., Chen, S. F., Shi, S., Zhang, Z., Liao, W. B., Wu, W., Zhou, R. C., & Fan, Q. (2015). High genetic diversity and weak population structure of *Rhododendron jinggangshanicum*, a threatened endemic species in Mount Jinggangshan of China. *Biochemical Systematics & Ecology*, 58, 178–186. <https://doi.org/10.1016/j.bse.2014.12.008>.
- Li, T. Q., Liu, X. F., Li, Z. H., Ma, H., Wan, Y. M., Liu, X. X., & Fu, L. Y. (2018). Study on reproductive biology of *Rhododendron longipedicellatum*: A newly discovered and special threatened plant surviving in limestone habitat in Southeast Yunnan, China. *Facial Plastic Surgery: FPS*, 9, 33. <https://doi.org/10.3389/fpls.2018.00033>.
- Li, T. Q., Liu, X. F., Li, Z. H., Wan, Y. M., Liu, X. X., & Ma, H. (2018). Development of novel EST-SSR markers for *Rhododendron longipedicellatum* (Ericaceae) and cross-amplification in two congeners. *Applications in Plant Sciences*, 6(6), e01162. <https://doi.org/10.1002/aps3.1162>.
- Li, Y., Zhang, X. W., & Fang, Y. M. (2019). Landscape features and climatic forces shape the genetic structure and evolutionary history of an oak species (*quercus chenii*) in east china. *Frontiers in Plant Science*, 10, 1060. <https://doi.org/10.3389/fpls.2019.01060>.
- Liu, X. F., Ma, Y. P., Wan, Y. M., Li, Z. H., & Ma, H. (2020). Genetic diversity of *Phyllanthus emblica* from two different climate type areas. *Facial Plastic Surgery: FPS*, 11, 580812. <https://doi.org/10.3389/fpls.2020.580812>.
- Liu, Y. L., Geng, Y. P., Song, M. L., Zhang, P. F., Hou, J. L., & Wang, W. Q. (2019). Genetic structure and diversity of *Glycyrrhiza* populations based on transcriptome SSR markers. *Plant Molecular Biology Reporter*, 37(5–6), 401–412. <https://doi.org/10.1007/s11105-019-01165-2>.
- Liu, Y. F., Xing, M., Zhao, W., Fan, R. J., Luo, S., & Chen, X. (2012). Genetic diversity analysis of *Rhododendron aureum* Georgi (Ericaceae) located on Changbai Mountain using ISSR and RAPD markers. *Plant Systematics & Evolution*, 298(5), 921–930. <https://doi.org/10.1007/s00606-012-0601-0>.
- Ma, H., Li, T. Q., Liu, X. F., Wan, Y. M., Li, Y. Y., Liu, X. X., & Li, Z. H. (2017). Research progress in conservation biology of

- Rhododendron. *World Forestry Research*, 30(4), 13–17. <https://doi.org/10.13348/j.cnki.sjlyyj.2017.0037.y>.
- Ma, H., Liu, Y. B., Liu, D. T., Sun, W. B., Liu, X. F., Wan, Y. M., Zhang, X. J., Zhang, R. G., Yun, Q. Z., Wang, J. H., Li, Z. H., & Ma, Y. P. (2021). Chromosome-level genome assembly and population genetic analysis of a critically endangered rhododendron provide insights into its conservation. *The Plant Journal*, 107(5), 1533–1545. <https://doi.org/10.1111/tpj.15399>.
- Ma, Y. P., Nielsen, J., Chamberlain, D. F., Li, X. Y., & Sun, W. B. (2014). The conservation of Rhododendrons is of greater urgency than has been previously acknowledged in China. *Biodiversity & Conservation*, 23(12), 3149–3154. <https://doi.org/10.1007/s10531-014-0764-9>.
- Mable, B. K. (2019). Conservation of adaptive potential and functional diversity: integrating old and new approaches. *Conservation Genetics*, 20(1), 89–100. <https://doi.org/10.1007/s10592-018-1129-9>.
- Nei, M. (1973). Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences of the United States of America*, 70(12), 3321–3323. <https://doi.org/10.1073/pnas.70.12.3321>.
- Ng, S. C., & Corlett, R. T. (2000). Comparative reproductive biology of the six species of *Rhododendron* (Ericaceae) in Hong Kong, South China. *Industrial Crops & Products*, 78(2), 221–229. <https://doi.org/10.1139/b99-181>.
- Nybom, H. (2004). Comparison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants. *Molecular Ecology*, 13(5), 1143–1155. <https://doi.org/10.1111/j.1365-294X.2004.02141.x>.
- Oyundelger, K., Herklotz, V., Harpke, D., Oyuntsetseg, B., & Ritz, C. M. (2021). Contrasting effects of local environment and grazing pressure on the genetic diversity and structure of *Artemisia frigida*. *Conservation Genetics*, 39(2), 364–375. <https://doi.org/10.1007/s10592-021-01375-w>.
- Paradis, E., & Schliep, K. (2019). ape 5.0: An environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*, 35(3), 526–528. <https://doi.org/10.1093/bioinformatics/bty633>.
- Piry, S., Luikart, G., & Cornuet, J.-M. (1999). Bottleneck: A computer program for detecting recent reductions in the effective population size using allele frequency data. *Journal of Heredity*, 90(4), 502–503. <https://doi.org/10.1093/jhered/90.4.502>.
- Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, 155(2), 945–959. <https://pubmed.ncbi.nlm.nih.gov/10835412/>.
- Pudovkin, A. I., Zaykin, D. V., & Hedgecock, D. (1996). On the potential for estimating the effective number of breeders from heterozygote-excess in progeny. *Genetics*, 144(1), 383–387. <https://doi.org/10.1093/genetics/144.1.383>.
- Rice, W. R. (1989). Analyzing tables of statistical tests. *Evolution*, 43(1), 223–225. <https://doi.org/10.1111/j.1558-5646.1989.tb04220.x>.
- Singh, R. K., Jena, S. N., Khan, S., Yadav, S., Banarjee, N., Raghuvanshi, S., Bhardwaj, V., Dattamajumder, S. K., Kapur, R., Solomon, S., Swapna, M., Srivastava, S., Tyagi, A. K., & Tyagi, A. K. (2013). Development, cross-species/genera transferability of novel EST-SSR markers and their utility in revealing population structure and genetic diversity in *sugarcane*. *Genes*, 524(2), 309–329. <https://doi.org/10.1016/j.gene.2013.03.125>.
- Slatkin, M. (1985). Gene flow in natural populations. *Annual Review of Ecology Evolution & Systematics*, 16, 393–430. <https://doi.org/10.1146/annurev.es.16.110185.002141>.
- Stebbins, C. L. J. (1950). *Variation & evolution in plants*. Oxford University Press.
- Stoeckel, S., Grange, J., Fernández-Manjarres, J. F., Bilger, I., Frascaria-Lacoste, N., & Mariette, S. (2010). Heterozygote excess in a self-incompatible and partially clonal forest tree species - *Prunus avium* L. *Molecular Ecology*, 15(8), 2109–2118. <https://doi.org/10.1111/j.1365-294X.2006.02926.x>.
- Wang, X. Q., Huang, Y., & Long, C. L. (2013). Assessing the genetic consequences of flower-harvesting in *Rhododendron decorum* Franchet (Ericaceae) using microsatellite markers. *Biochemical Systematics & Ecology*, 50, 296–303. <https://doi.org/10.1016/j.bse.2013.04.009>.
- Wei, W. L., Qi, X. Q., Wang, L. H., Zhang, Y. X., Hua, W., Li, D. H., Lv, H. X., & Zhang, X. R. (2011). Characterization of the sesame (*Sesamum indicum* L.) global transcriptome using illumina paired-end sequencing and development of est-ssr markers. *BMC Genomics*, 12(1), 451. <https://doi.org/10.1186/1471-2164-12-451>.
- Weir, B. S., & Cockerham, C. C. (1984). Estimating *F*-statistics for the analysis of population structure. *Evolution*, 38, 1358–1370. <https://doi.org/10.1111/j.1558-5646.1984.tb05657.x>.
- Willi, Y., Buskirk, J. V., & Hoffmann, A. A. (2006). Limits to the adaptive potential of small populations. *Annual Review of Ecology Evolution & Systematics*, 37(1), 433–458. <https://doi.org/10.2307/30033839>.
- Wright, S. (1949). The genetic structure of populations. *Annals of Human Genetics*, 15(1), 323–354. <https://doi.org/10.1111/j.1469-1809.1949.tb02451.x>.
- Wu, F. Q., Shen, S. K., Zhang, X., Yang, G., & Wang, Y. H. (2017). Inferences of genetic structure and demographic history of *Rhododendron protistum* var. *giganteum*-The world's largest Rhododendron using microsatellite markers. *Flora*, 233, 1–6. <https://doi.org/10.1016/j.flora.2017.04.009>.
- Yadav, L. K., Mcassey, E. V., & Wilde, H. D. (2019). Genetic diversity and population structure analysis of *Rhododendron canescens*, a Native Azalea for urban landscaping. *HortScience*, 54(4), 647–651. <https://doi.org/10.21273/HORTSCI13840-18>.
- Yang, Y., Wang, H. Y., & Ma, L. H. (2020). Characteristics of fruiting and seed germination of endangered plant, *Rhododendron changii*. *Chinese Journal of Scientia Silvae Sinicae*, 56(10), 173–183. <https://doi.org/10.11707/j.1001-7488.20201019>.
- Yeh, F. C., Yang, R. C., & Boyle, T. (1999). *POPGENE VERSION 1.31: Microsoft Window-based Freeware for Population Genetic Analysis*. <ftp://ftp.microsoft.com/Softlib/HPGL>.
- Young, A., Boyle, T., & Brown, T. (1996). The population genetic consequences of habitat fragmentation for plants. *Trends in Ecology & Evolution*, 11, 413–418. [https://doi.org/10.1016/0169-5347\(96\)10045-8](https://doi.org/10.1016/0169-5347(96)10045-8).

- Zhang, X. J., Liu, X. F., Liu, D. T., Cao, Y. R., Li, Z. H., Ma, Y. P., & Ma, H. (2021). Genetic diversity and structure of *Rhododendron meddianum*, a plant species with extremely small populations. *Plant Diversity*, 43(6), 472–479. <https://doi.org/10.1016/j.pld.2021.05.005>.
- Zhang, Y., Zhang, X., Wang, Y. H., & Shen, S. K. (2017). *De Novo* Assembly of transcriptome and development of novel EST-SSR markers in *Rhododendron rex* Lévl. through Illumina sequencing. *Facial Plastic Surgery: FPS*, 8, 1664. <https://doi.org/10.3389/fpls.2017.01664>.
- Zhou, Q., Mu, K., Ni, Z. X., Liu, X. H., Li, Y. A., & Xu, L. A. (2020). Analysis of genetic diversity of ancient *Ginkgo* populations using SSR markers. *Industrial Crops & Products*, 145, e111942. <https://doi.org/10.1016/j.indcrop.2019.111942>.
- Zhu, Z., Zhang, L., Gao, L., Tang, S., Zhao, Y., & Yang, J. (2016). Local habitat condition rather than geographic distance determines the genetic structure of *Tamarix chinensis* populations in Yellow River Delta, China. *Tijdschrift Voor Gerontologie en Geriatrie*, 12(1), 14. <https://doi.org/10.1007/s11295-016-0971-5>.