Flora 233 (2017) 1–6

Contents lists available at ScienceDirect

Flora

journal homepage: www.elsevier.com/locate/flora

Inferences of genetic structure and demographic history of *Rhododendron protistum* var. *giganteum*—The world's largest Rhododendron using microsatellite markers

Fuqin Wu[a,b], Shikang Shen[a,*], Xue Zhang[a], Guansong Yang[a], Yuehua Wang[a]

[a] School of Life Sciences, Yunnan University, Kunming, 650091, China
[b] Yunnan Research and Monitoring Center of Nature Reserve, Yunnan Institute for Forest Inventory and Planning, Kunming, 650051, China

ARTICLE INFO

Article history:
Received 31 July 2016
Received in revised form 11 April 2017
Accepted 18 April 2017
Edited by Dr. W Durka
Available online 24 April 2017

Keywords:
Gene flow
Endangered plant
Population demography
Small population
Conservation strategies

ABSTRACT

Big tree rhododendron (*Rhododendron protistum* var. *giganteum*) is an extremely endangered plant with only two remnant populations distributed in the Gaoligong Mountains of northwestern Yunnan Province, China. However, the genetic structure, demographic history, and the population dynamics of this species have been rarely described. In the present study, data from 14 polymorphic nuclear microsatellite loci were used to assess the levels of genetic diversity, genetic structure, and demographic history of big tree rhododendron remnant populations. Results suggested that renmants maintained a moderate level of genetic diversity (N_E = 2.86, A_R = 5.726, H_o = 0.510, H_e = 0.602, I = 1.174) despite their small population. Bayesian clustering and principal coordinate analysis indicated that sampled individuals were clustered into two distinct genetic groups. Demographic history analysis revealed high historical and low contemporary gene flow between two remnant populations. BOTTLENECK analysis under a two-phase mutation model showed that the populations could have experienced a recent historical bottleneck. The effective population sizes of the two populations were smaller than 30. Therefore, *in situ* conservation of this species should be prioritized. Furtheremore, *ex situ* cultivation of field collected seeds and introduction of saplings into the source site might be considered for the future conservation and restoration of big tree rhododendron.

1. Introduction

Inference of historical demographic and population dynamics and their complex interactions are of practical importance to conservation biologists in formulating conservation strategies for endangered plants (Kang et al., 2008). Population dynamics inference can help elucidate the current status of a species experiencing habitat degradation, varying landscapes, and climate changes (Selwood et al., 2014). Historical and contemporary microevolutionary processes, such as drift and gene flow, can contribute to current patterns of genetic variation within populations (Adams et al., 2006; Yao et al., 2007; Shaddick et al., 2011). Recent reductions in population size and population isolation can decrease genetic diversity and promote genetic divergence within populations, especially in small populations (Consuegra et al., 2005; Whiteley et al., 2010). Therefore, genetic diversity, population structure, and historical population dynamics should be evaluated to improve our understanding of the protection of endangered, threatened, and endemic plants with conservation value.

Big tree rhododendron, *Rhododendron protistum* var. *giganteum* (Forrest) D. F. Chamberlain, is a perennial woody plant belonging to Ericaceae. The species is honored as "*the king of Rhododendron*" in China (Shen et al., 2015). It is a critically endangered species listed on the China Plant Red Data Book (Fu and Jin, 1992). The taxon has only two wild populations with less than 1500 individuals remaining in the wild (Ma et al., 2012; Wu et al., 2015). Although amplified fragment length polymorphism markers (AFLP) have been used to investigate the genetic diversity of big tree rhododendron remnant populations (Wu et al., 2015), microsatellites, due to their high polymorphism and codominant inheritance would be more appropriate to assess inbreeding and bottlenecks and indirect estimate population demography (Nybom, 2004). Thus, we employed microsatellite markers to address the following questions: (1) what is the degree of genetic diversity within and gene flow between the remnant big tree rhododendron populations? (2)

* Corresponding author at: School of life sciences, Yunnan University, No. 2 Green lake North road Kunming, Yunnan, 650091, China.
E-mail addresses: shikang168@yahoo.com, yunda123456@126.com (S. Shen).

http://dx.doi.org/10.1016/j.flora.2017.04.009
0367-2530/

how is demographic history reflected in the current population genetic structure? This information is used to develop optimum management strategies for big tree rhododendron conservation.

2. Materials and methods

2.1. Study species and plant sampling

Big tree rhododendron is distributed in the Gaoligong Mountains National Nature Reserve, Tengchong County of Yunnan Province, China (Ma et al., 2012). The species is one of the tallest rhododendron trees and reaches 30 m in height and 1 m in basal diameter. Big tree rhododendron is evergreen and has large purple-red flowers from January to March. Fruits ripen between October and December producing numerous small seeds with thin membranous wings (Shen et al., 2015; Wu et al., 2015).

In November 2013, 60 individuals were collected from the two natural populations. Of these individuals, 30 were obtained from the Cizhuhe (CZH, N: 25°39′47″; E: 98°43′44″) and Dahetou (DHT, N: 25°46′42″; E: 98°42′29″) population. The distance between the collected individual samples was at least 15 m. Fresh young leaves were removed from shoots, dried in silica gel, and stored at −20 °C until DNA was extracted.

2.2. DNA extraction and microsatellite genotyping

Genomic DNA was extracted from dried leaves by using the modified CTAB method (Doyle and Doyle, 1987). Total purified DNA was stored at −20 °C until use.

Microsatellite primers were selected from recently developed nuclear microsatellites in *Rhododendron* subg. *Hymenanthes* (Wang et al., 2009; Wang et al., 2010; Li et al., 2011). Fourteen high-polymorphism microsatellite loci were selected from the published 58 microsatellite loci. PCR amplification was performed in a 25 μL reaction, with the forward primers labeled with fluorescent dye (FAM, TAMRA or HEX) and visualized on an ABI 3730xl Capillary DNA Analyzer by Sangon Biotech (Shanghai) Co., Ltd. Fragment sizes were assessed using GeneMapper version 4.0.

2.3. Data analysis

2.3.1. Genetic diversity

The dataset editing and transformation was performed in GenAlEx, version 6.5 (Peakall and Smouse, 2012). Hardy-Weinberg equilibrium (HWE) was tested with default parameters using Genepop version 4.3 (Raymond and Rousset, 1995). Linkage disequilibrium (LD) was investigated at the 5% statistical significance level among loci pairs with 1000 permutations using FSTAT v2.9.3 (Goudet, 2001) and then corrected for multiple tests using the sequential Bonferroni method (Rice, 1989). The number of average alleles (N_a), effective number of alleles (N_E), allelic richness (A_R), expected heterozygosity (H_e), observed heterozygosity (H_O), information index (I) and fixation index (F_{IS}) were calculated using GenAlEx, version 6.5 (Peakall and Smouse, 2012), and FSTAT, v.2.9.3 (Goudet, 2001).

2.3.2. Genetic structure

The genetic differentiation between populations was evaluated by calculating the genetic differentiation index (F_{ST}; Weir and Cockerham, 1984) and standardized G_{ST} index (G'_{ST}; Hedrick, 2005) using FSTAT v.2.9.3 (Goudet, 2001), and SMOGD v. 1.2.5 (Crawford, 2010), respectively. An analysis of molecular variance (AMOVA) using the pairwise difference was performed in Arlequin version 3.5 (Excoffier et al., 2005). Bayesian clustering analysis implemented in Structure v2.3.1 (Pritchard et al., 2000) was used to assign the 60 individuals to genetic clusters (K) and to estimate admixture proportions for each individual. The analyses were run using an admixture model with correlated allele-frequency models. The simulation was run with values of K from 1 to 20 and repeated 20 times for each set. Each run included a burn-in of 1×10^5 iterations and 1×10^5 subsequent Markov Chain Monte Carlo (MCMC) steps. The best-fit number of groupings was evaluated using ΔK by STRUCTURE HARVESTER v0.6.8 (Evanno et al., 2005; Earl and vonHoldt, 2012). Furthermore, an individual-based principal coordinate analysis (PCoA) visualized inter-individual genetic distances (Nei, 1983) by the program MVSP, v 3.1.2 (Kovach, 1999).

2.3.3. Historical and contemporary gene flow

We used Migrate-n version 3.6.4 (Beerli, 2006) to estimate historical gene flow. MIGRATE applies coalescent theory to estimate jointly effective population size θ ($4Ne\mu$) and asymmetric gene flow M (m/μ) between pairs of populations over a long period of time (~4Ne generations). By applying a Brownian motion model in the maximum likelihood runs of Migrate-n, we sampled one of every 20 reconstructed genealogies for each locus for three long and ten short chains. In the recorded 10,000 and 1000 genealogies for short and long chains, respectively, the first 2000 and 200 genealogies were discarded as burn-in. In addition, the amount of long-term gene flow (Nm) between populations was indirectly estimated by a traditional genetic differentiation method based on F_{ST} value ($Nm = (1-F_{ST})/4F_{ST}$; Slatkin and Barton, 1989; Hedrick, 2005).

To estimate contemporary migration patterns, we used the software GeneClass v2.0 (Piry et al., 2004) and BayesAss v3.0 (Wilson and Rannala, 2003). GeneClass was used to detect putative first generation migrants basing on Bayesian method (Rannala and Mountain, 1997). Assignment probabilities were calculated using Monte Carlo resampling with 10,000 permutations and using a threshold probability of 0.01 (Piry et al., 2004). Contemporary inter-population migrations were estimated following a Bayesian approach implemented in BayesAss version 3.0 with default parameters (Wilson and Rannala, 2003).

2.3.4. Population bottlenecks and effective population size

We tested for a recent bottleneck using BOTTLENECK v.1.2.02 (Piry et al., 1999) applying both the Sign test and the Wilcoxon test as suggested in the manual. We assumed a two-phase mutation model (TPM) because it is more suitable for microsatellite markers compared to the infinite allele (IAM) and one-step stepwise mutation model (SMM) (Di Rienzo et al., 1994; Spencer et al., 2000).

We calculated the effective population sizes of two populations to establish the degree of endangerment of the species. We used the program LDNeInterface (Waples and Do, 2008) using levels of the lowest allele frequency (=0.01, 0.02, 0.05) at a 95% confidence interval.

3. Results

3.1. Equilibrium test and genetic diversity

Most loci and two populations showed significant deviation from Hardy-Weinberg equilibrium (HWE) ($P < 0.01$) except for locus 6. Of the 91 locus pairwise comparisons, 28 showed significant LD ($P < 0.05$) after Bonferroni correction.

A total of 81 alleles were detected for the 14 SSR loci analyzed. At the species level, A_R, H_o, H_e and I were 5.726, 0.510, 0.602 and 1.174, respectively. At the population level, the CZH population showed a lower genetic diversity level than DHT. F_{IS} was significantly positive ($p < 0.001$) at the species level (0.161) and within the DHT population (0.122) (Table 1).