CROP BREEDING AND APPLIED BIOTECHNOLOGY

ARTICLE

Genetic diversity of *Rhododendron henanense* subsp. *lingbaoense* revealed by whole-genome resequencing

Crop Breeding and Applied Biotechnology
23(1): e439023112, 2023
Brazilian Society of Plant Breeding.
Printed in Brazil
http://dx.doi.org/10.1590/1984-70332023v23n1a12

Xiaojun Zhou[1*], Chunrong Zhang[1] and Jianming Han[1]

Abstract: *Rhododendron henanense subsp. lingbaoense Fang (Rhl) is a subspecies endemic to China with important ornamental value. Recently, the Rhl population has been shrinking as a result of its limited distribution range, poor natural regeneration, and rising human exploitation. In this study, high-quality single nucleotide polymorphisms (SNPs) identified through whole-genome resequencing were examined in 50 individuals from four populations of Rhl with a depth of ~34× to evaluate genome-level diversity and population structure. Population structure profiling, principal component analysis, and phylogenetic analysis clustered the Rhl samples into two groups corresponding to their geographical distributions. Analysis of SNPs indicated that Rhl populations have high genetic diversity and little genetic differentiation. Additionally, demographic history indicated that all four populations of Rhl have experienced long-term population decline. The above findings showed that we should take action to protect this rare species.*

Keywords: *High-throughput sequencing, SNPs, population demography, conservation implications*

INTRODUCTION

Rhododendron is the collective name for *Rhododendron* (Ericaceae) plants, and of its flowers are among the ten most recognizable flowers in China. In addition to its high ornamental value, rhododendron also serves the purposes of ecological conservation, medical application, and scientific research (Li et al. 2018, Zhang et al. 2021). *Rhododendron* contains more than 1,000 species and has a global distribution. However, some unmanaged wild species with small populations have gone extinct or are on the verge of extinction due to narrow habitats or severe anthropogenic interference (Ma et al. 2014, Liu et al. 2020). *Rhododendron henanense lingbaoense* Fang (*Rhl*) is a perennial evergreen plant up to 3-6 m tall that belongs to subgen. *Hymenanthes* and subsect. *Campylocarpa* (Fang 1983). The main difference between *Rhl* and closely related species of *Rhododendron* is the white corolla without spots. Among the numerous *Rhododendron* taxa, *Rhl* has received less scientific attention. This may be because it is endemic to China, naturally distributed in the alpine region at altitudes above 2000 m on the border of Henan and Shaanxi provinces. *Rhl* has significant landscape usage value as a result of its extremely large blooms and clustered growth, as well as scientific research value and natural heritage value as an endemic species (Weng et al. 2012). The population of *Rhl* has

***Corresponding author:**
E-mail: lynubio@126.com
ORCID: 0000-0002-2234-0968

Received: 08 November 2022
Accepted: 29 March 2023
Published: 23 March 2023

[1] Luoyang Normal University, 6 Jiqing Rd, Luolong District, Luoyang, Henan, China, 471934

recently been declining due to its constrained geographic range, its weak natural regeneration potential, and growing human exploitation (Han et al. 2008, Ma et al. 2017, Zhou et al. 2022). As a result, it is essential to properly assess the natural resources of *Rhl* and take emergency conservation measures.

The core objective of species conservation is to protect their genetic diversity and genetic structure as much as possible (Hendricks et al. 2017). Studying the genetic variation of the target population can not only provide a basis for the development of artificial recovery techniques for wild populations of the species but also provide important information for the development of conservation strategies (Barbosa et al. 2018). With advances in high-throughput sequencing technology, population genomics can now support genetic analysis at a genomic scale in natural populations with thousands of genetic markers rather than a few genetic loci (Hohenlohe et al. 2010). Single nucleotide polymorphisms (SNPs) are the most abundant and universal sequence variations in all genomes, which makes them excellent markers for genetic studies (Wang et al. 2015). By using next-generation sequencing (NGS) platforms to resequence the genomes of representative individuals from different groups within a species, we can quickly and efficiently obtain genomic information from different groups and obtain a large number of SNP loci by comparing them with existing genomes, thus exploring the genetic, evolutionary and biological properties of living organisms at the genomic scale (Wang and Zhang 2011).

In this study, SNP loci identified by genome resequencing were used to determine the genetic diversity, population structure, and demographic history of *Rhl*. We believe that the results presented here will be useful for the management and conservation of *Rhl* in the future and will also serve as an example for population genetic studies employing NGS for the preservation of species with restricted distributions.

MATERIAL AND METHODS

Plant sampling and DNA extraction

Rhl often grows in moist, slightly acidic (pH 6.2~7) mountain brown soil rich in organic matter. It is intolerant to high temperatures, with an optimum growth temperature of 10-25 □ (Han et al. 2008). Common companion species are *Betula platyphylla* Suk. and *Acer davidii* Franch. Rhl is mainly distributed in three populations (CA for Chang'an, LY for Laoya, and NC for Nanchuang) in the Xiaoqinling National Nature Reserve (lat 34° 25' 12" N, long 110° 28' 45" E, alt 2200 m asl) in Henan Province. While conducting fieldwork, we found another population showing characteristics consistent with those of *Rhl* in the Taibai Mountains (TB, lat 34° 00' 73" N, long 107° 48' 86" E, alt 2700 m asl) of Shaanxi Province, China. Therefore, we sampled the TB population as an outgroup for our study. Due to limitations imposed by field conditions, a total of 50 individuals from the four locations (CA 13, LY 18, NC 8, and TB 11) covering major habitats were sampled. Individuals from each population were chosen at random and spaced approximately 100 m apart. Fresh leaves were collected in self-sealing bags with dried silica gel and preserved until DNA extraction, and all plant materials were kept in the Plant Diversity Research Laboratory (PDRL) of Luoyang Normal University.

Genomic DNA was isolated using a modification of the CTAB method (Allen et al. 2006). The concentration and quality of DNA were measured using a NanoDrop™ 2000 spectrophotometer (IMPLEN, CA, USA) and the Qubit® DNA Assay Kit Fluorometer (Life Tech, CA, USA). DNA integrity was evaluated on a 1% (w/v) agarose gel.

Library construction and genome resequencing

The extracted DNA was enzymatically disrupted using the MGIEasy Enzymatic PCR-Free DNA Library Preparation Kit (MGI Tech, China), and the fragments (450~600 bp) were purified using two-step magnetic bead selection. The fragmented DNA was end-repaired, splice ligated, and then purified. The purified product was denatured using a PCR instrument to obtain a single-stranded product, which was then cyclized. Then, enzymatic digestion and purification were performed. The quality-checked products were subjected to DNBSEQ-T7 (MGI Tech, China) sequencing (Jeon et al. 2021).

SNP detection

Prefiltering of the reads on the basis of quality was performed using Cutadapt (version 1.16, https://pypi.org/project/cutadapt/1.2.1/) (-q 20 -e 0.1 -n 1 -m 20) (Martin 2011). BWA software (Li and Durbin 2009) was used to compare the reads with the reference genome sequence of *Rhl* (https://ncbi.nlm.nih.gov/assembly/GCA_020567845.1) and locate

the positions of clean reads in the reference genome (Zhou et al. 2022). GATK software was used to locally realign the reads near insertions/deletions (InDels) to obtain bam files and eliminate false-positive SNPs around the InDels (Narzisi et al. 2014). Then, Picard (version 1.102, https://sourceforge.net/projects/picard/files/picard-tools/1.102/) was used to mark duplicates and perform base recalibration to ensure the accuracy of the SNPs. The SNP calling and filtering software applied for each sample was GATK with the following parameters: "QD < 2.0 || MQ < 40.0 || FS > 60.0 || SOR > 3.0 || MQRankSum < -12.5 || ReadPosRankSum < -8.0" (Narzisi et al. 2014). Information on the sequencing depth and genome coverage of each sample was obtained, and the detection of variants was performed. SNPs of all samples were merged into a vcf file and filtered using the software VCFtools with the parameters "--maf 0.05 --minDP 4 --max-missing 0.1 --max-alleles 2 --min-alleles 2" (Danecek et al. 2011).

Population structure and population genetics analysis

After filtering, high-quality SNPs were employed in the population genetics analysis. A phylogenetic tree is a diagram representing evolutionary relationships among organisms derived from a common ancestral form. We constructed neighbor-joining (NJ) trees for 50 individuals using FastTreeMP software (version 2.1.11 SSE3) with 200 bootstrap values (Price et al. 2010). To examine population stratification, structure analysis was carried out by Admixture software (v1.3.0) with K values from 1 to 4 (Alexander et al. 2009). Principal component analysis (PCA) was performed with the EIG-6.1.4 program package (Patterson et al. 2006). Nucleotide diversity (Pi) and fixation statistic (Fst) analyses were applied to estimate the degree of variation among groups and to explain population differentiation based on the variance between the two groups, respectively. Both Pi and Fst were calculated using the "population" command in Stacks software (version 2.54) (Catchen et al. 2013).

Historical population size estimates

The size and evolutionary history of each of the four populations were inferred using MSMC2 software (v2.3.1, https://github.com/stschiff/msmc2) with the following parameters: "1*2+25*1+1*2+1*3" (Schiffels and Wang 2020). The mutation rate used to calculate the population size was 4.15×10^{-9} (Yang et al. 2020).

RESULTS AND DISCUSSION

Identification of SNPs in *Rhl* germplasms

In this study, the genomes of 50 *Rhl* individuals were resequenced to explore genomic diversity with a mean sequencing depth of ~ 34× and 95% coverage. In total, ~ 114 G bp of sequences or 7.60 billion reads were generated. Our sequencing data showed high Phred quality (Q20 > 96%, Q30 > 88%), with a stable GC content ranging from 40.91% to 42.01%. A total of 316,555,560 SNPs were obtained in the 50 individuals by comparison with the reference genome. After stringent alignment and filtering, 249 million high-quality SNPs were discovered and used for subsequent analysis. The SNP data have been deposited in the National Genomics Data Center (NGDC) under accession number GVM000370.

Population genetic structure of *Rhl* germplasms

Using the high-quality SNPs, a phylogenetic relationship analysis of 50 *Rhl* individuals was conducted to explore the degree of admixture in the populations. It is clear from Figure 1a that the 50 individuals of *Rhl* were divided into Shaanxi and Henan populations when K=2. When K=3 and 4, two different subpopulations emerged in the Shaanxi population, and there was no significant genetic structure between populations. In addition, at the junction of different populations, the genetic structures were intermixed with each other. Therefore, the optimal K value is two, at which the four populations can be divided into two groups based on geographic distance: Henan and Shaanxi. Moreover, the NJ tree was reconstructed. The samples were clearly divided between two branches, where the red section showed individuals from the Shaanxi TB population and the blue section showed individuals from Henan Province (Figure 2). This result is consistent with the genetic structure revealed by admixture. PCA was further performed based on the high-quality SNPs to study the genetic background similarity and clustering relationships among the 50 samples. As shown in Figure 3a, PC1 and PC2 accurately divided the individuals into two groups in Henan and Shaanxi, which is fully consistent with the geographical information of the source of *Rhl*. PC1 and PC2 explained 2.14% and 1.57% of the total variation, respectively.

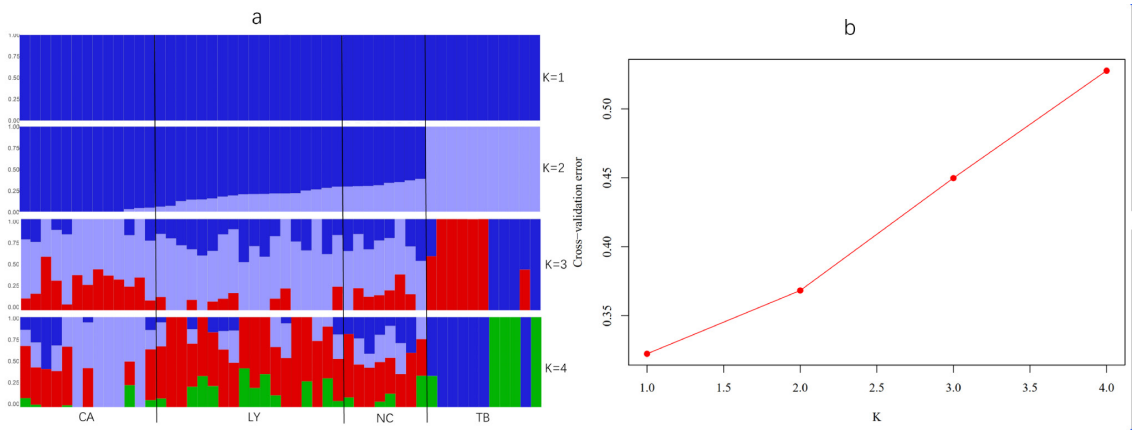

Figure 1. Genetic structure analysis of four populations (CA, LY, NC, and TB) of *R. henanense* subsp. *lingbaoense* based on SNPs. Population assignment by admixture analysis for K = 1-4 (a). The cross-validation (CV) error rate of K value (b).

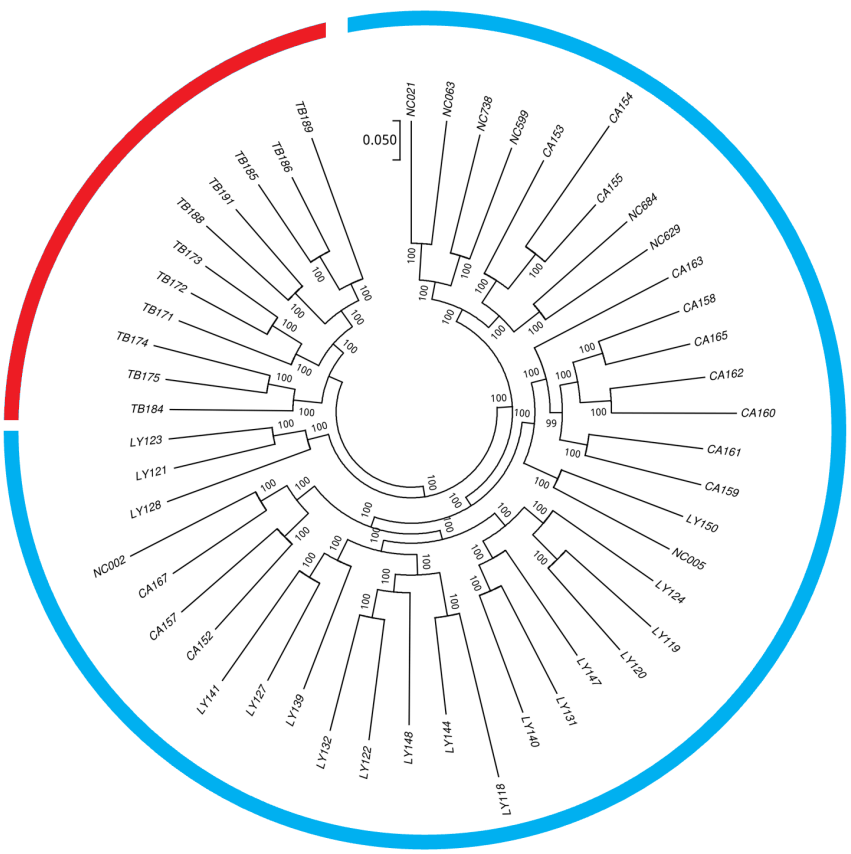

Figure 2. Neighbor-joining (NJ) phylogenetic tree of 50 *R. henanense* subsp. *lingbaoense* samples. The red section indicates individuals from the Shaanxi TB population, and the blue section indicates individuals from Henan Province.

Table 1. Genetic diversity analysis of four *R. henanense* subsp. *lingbaoense* populations performed by using Stacks software

Pop ID	%Poly Loci	Num Indv	Ho	He	P	Pi	Fis
CA	78.2649	9.9270	0.5958	0.3502	0.7021	0.3709	-0.4466
LY	78.0040	13.1429	0.6019	0.3541	0.6990	0.3698	-0.4677
NC	80.4387	5.8319	0.6342	0.3629	0.6829	0.4017	-0.4356
TB	77.8868	8.5374	0.5819	0.3412	0.7091	0.3648	-0.4261

Notes: Pop, populations; % Poly Loci, percentage of sites found to be polymorphic; Num Indv, mean number of individuals per locus in this population. P, average frequency of major allele; Ho, average observed heterozygosity; He, average expected heterozygosity; Pi, mean nucleotide diversity; Fis, average Wright's inbreeding coefficient.

To further analyze the relationship between the *Rhl* materials in different populations, the expected heterozygosity (He), observed heterozygosity (Ho), and nucleotide diversity (Pi) of each population were quantified. Analysis of SNP datasets for all populations revealed high genetic diversity (Pi>0.3648, He >0.3412) in *Rhl* (Table 1). The results of Fst analysis indicated little genetic differentiation among these populations (Fst = 0.0110-0.0240) (Table 2).

Table 2. Genetic differentiation among populations of *R. henanense* subsp. *lingbaoense* based on SNPs

	LY	NC	TB
CA	0.0110	0.0149	0.0228
LY		0.0117	0.0194
NC			0.0240

It is widely recognized that characterizing population structure and genetic diversity is essential for the effective conservation of threatened species. Compared to widespread plant species, narrowly distributed species may have lower genetic diversity due to genetic drift and inbreeding depression (Gibson et al. 2008). However, resequencing indicated that the nucleotide diversity (0.3648-0.4017) of *Rhl* is higher than that of other endangered plant species, such as *Geodorum densiflorum* (0.0359), *Rhododendron meddianum* (0.0741), and *Pinus bungeana* (0.2427-0.2842) (Tian et al. 2022, Roy et al. 2016, Zhang et al. 2021, respectively). Additionally, *Rhl* has higher genetic diversity than other widely distributed species, e.g., *Physalis philadelphica* (0.127-0.206) and *Phaseolus vulgaris* (0.31) (Delfini et al. 2021, Alcalá-Gómez et al. 2022, respectively). These findings are largely consistent with the hypothesis that some rare and endangered species can maintain high levels of genetic diversity even at small population sizes (Zhao et al. 2012, Wu et al. 2015, Stone et al. 2019).

The genetic diversity of a species is generally influenced by its breeding system, life span, seed dispersal mechanism, distribution range, and evolutionary history. Additionally, outcrossing organisms typically exhibit higher genetic variety than selfing species (Nybom 2004, Zhang et al. 2021). Previous research revealed that *Rhododendron* requires pollinators and produces adhesive pollen, suggesting that relies mostly on outcrossing (Ng and Corlett 2000, Huang et al. 2017, Li et al. 2018). The high genetic diversity of the four populations of *Rhl* might also be derived from their ancestral populations, as observed for other threatened *Rhododendron* plants, e.g., *R. protistum* var. *gigantum* and *R. meddianum* (Wu et al. 2015, Zhang et al. 2021).

The population genetic structure, PCA results, and NJ trees demonstrated that populations (LY, NC, and CA) located in Henan Province were clustered together, while population TB from Shaanxi Province was separated. This is due to the geographical proximity of the three populations in Henan Province, whereas the TB population in Shaanxi Province is distant from them. The results of the population genetic analysis were consistent with the geographical distributions of the populations. Fst is a classical measure of population genetic differentiation and population genetic outcome (Holsinger and Weir 2009). Fst values range from 0-1, with a maximum value of 1 indicating complete differentiation between two populations and a minimum value of 0 indicating no differentiation between populations. The lowest peak was discovered to be at K = 1 according to the cross-validation errors of K values (Figure 1b). In light of this finding and the Fst values (0.0110-0.0240), it is likely that the four populations of *Rhl* used in this study all descended from the same ancestral population.

Demographic history of *Rhl* populations

MSMC2 software was used to assess changes in the effective population size of *Rhl* over time, and the results are shown in Figure 3b. Long-term shrinkage was observed in all four populations, and the trends in effective population size were

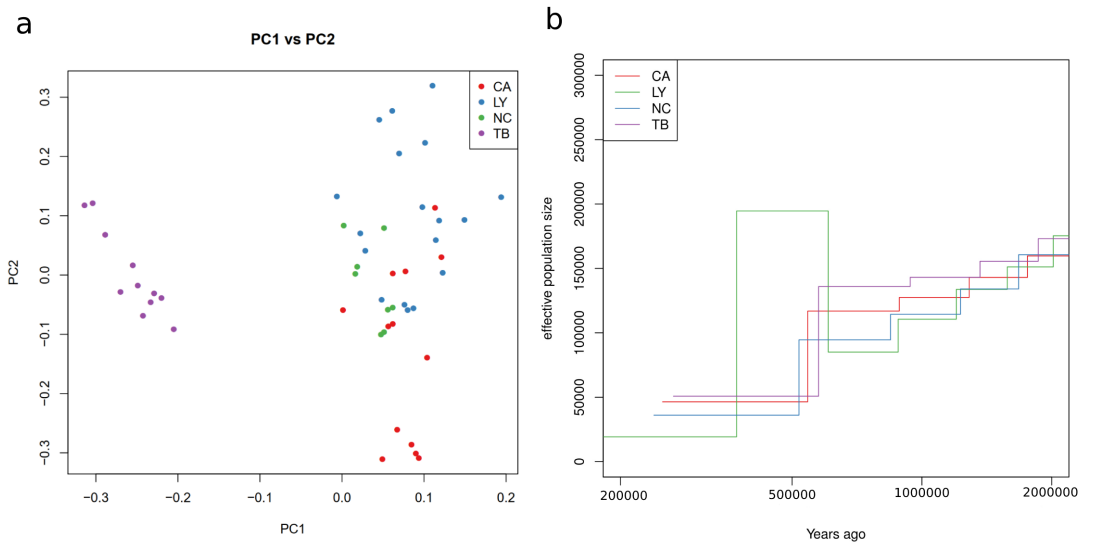

Figure 3. Principal component analysis (PCA) of all 50 *R. henanense* subsp. *lingbaoense* individuals, with the proportion of the variance explained being 2.14% for PC1 and 1.57% for PC2 (a). The effective population size of four *R. henanense* subsp. *lingbaoense* populations estimated using MSMC2 software (b).

generally consistent. After a long period of bottleneck effects, the LY population expanded significantly approximately 700,000 years ago. By approximately 400,000 years ago, the LY population had undergone another significant contraction, after which its changes remained largely consistent with those of other populations.

The key to conserving a species is preserving its genetic diversity (Rauch and Bar-Yam 2005). Our study on the genetic analysis of *Rhl* has significant conservation implications for this rare species. All populations of *Rhl* maintain a high level of genetic diversity, so it is important to enhance the restoration and conservation of its native habitat. In wild populations of *Rhl*, seedlings and saplings are rarely found, and its habitat is disturbed and destroyed by mining and other forms of anthropogenic interference. Therefore, it is extremely urgent to establish conservation plots to protect its natural habitat. In addition, the results of the population history dynamics study revealed long-term population decline in all four *Rhl* populations accompanied by Pleistocene climatic oscillations. Therefore, seed collection from all four *Rhl* populations should be conducted as soon as possible so that they can be used for ex situ conservation and storage of the germplasm.

ACKNOWLEDGMENTS

This study was funded by the National Natural Science Foundation of China (31870697).

REFERENCES

Alcalá-Gómez G, Pérez-Alquicira J, Cabrera-Toledo D, Cortés-Cruz M, del Pilar Zamora-Tavares M and Vargas-Ponce O (2022) Genetic diversity and structure in husk tomato (*Physalis philadelphica* Lam.) based on SNPs: a case of diffuse domestication. **Genetic Resources and Crop Evolution 69**: 443-459.

Alexander DH, Novembre J and Lange K (2009) Fast model-based estimation of ancestry in unrelated individuals. **Genome Research 19**: 1655-1664.

Allen GC, Flores-Vergara MA, Krasynanski S, Kumar S and Thompson WF (2006) A modified protocol for rapid DNA isolation from plant tissues using cetyltrimethylammonium bromide. **Nature Protocols 1**: 2320-2325.

Barbosa S, Mestre F, White TA, Pauperio J, Alves PC and Searle JB (2018) Integrative approaches to guide conservation decisions: Using genomics to define conservation units and functional corridors. **Molecular Ecology 27**: 3452-3465.

Catchen J, Hohenlohe PA, Bassham S, Amores A and Cresko WA (2013) Stacks: an analysis tool set for population genomics. **Molecular Ecology 22**: 3124-3140.

Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, McVean G, Durbin R

and 1000 Genomes Project Analysis Group (2011) The variant call format and VCFtools. **Bioinformatics 27**: 2156-2158.

Delfini J, Moda-Cirino V, Santos Neto J, Ruas PM, Sant'Ana GC, Gepts P and Gonçalves LSA (2021) Population structure, genetic diversity and genomic selection signatures among a Brazilian common bean germplasm. **Scientific Reports 11**: 1-12.

Fang WP (1983) New taxa of the genus *Rhododendron* from China. **Acta Phytotaxon Sin 21**: 457-459.

Gibson JP, Rice SA and Stucke CM (2008) Comparison of population genetic diversity between a rare, narrowly distributed species and a common, widespread species of *Alnus* (Betulaceae). **American Journal of Botany 95**: 588-596.

Han JW, Zhang Y, Yuan ZL and Ye YZ (2008) Study on the introduction and development of wild *Rhododendron lingbaoense* in Xiaoqinling National Nature Reserve of Henan. **Journal of Anhui Agriculture 36**: 14954-14955.

Hendricks S, Epstein B, Schönfeld B, Wiench C, Hamede R, Jones M, Storfer A and Hohenlohe P (2017) Conservation implications of limited genetic diversity and population structure in Tasmanian devils (*Sarcophilus harrisii*). **Conservation Genetics 18**: 977-982.

Hohenlohe PA, Bassham S, Etter PD, Stiffler N, Johnson EA and Cresko WA (2010) Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. **Plos Genetics 6**: e1000862.

Holsinger KE and Weir BS (2009) Genetics in geographically structured populations: defining, estimating and interpreting F_{ST}. **Nature Reviews Genetics 10**: 639-650.

Huang ZH, Song YP and Huang SQ (2017) Evidence for passerine bird pollination in *Rhododendron* species. **AoB Plants 9**: plx62.

Jeon SA, Park JL, Park SJ, Kim JH, Goh SH, Han JY and Kim SY (2021) Comparison between MGI and Illumina sequencing platforms for whole genome sequencing. **Genes & Genomics 43**: 713-724.

Li H and Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler Transform. **Bioinformatics 25**: 1754-1760.

Li TQ, Liu XF, Li ZH, Ma H, Wan YM, Liu XX and Fu LY (2018) Study on reproductive biology of *Rhododendron longipedicellatum*: a newly discovered and special threatened plant surviving in limestone habitat in southeast Yunnan. **Frontiers in Plant Science 9**: 33.

Liu D, Zhang L, Wang J and Ma Y (2020) Conservation genomics of a threatened *Rhododendron*: contrasting patterns of population structure revealed from neutral and selected SNPs. **Frontiers in Genetics 11**: 757.

Ma H, Li T, Liu X, Wan Y, Li Y, Liu X and Li Z (2017) Research progress in conservation biology of *Rhododendron*. **World Forestry Research 30**: 13-17.

Ma YP, Nielsen J, Chamberlain DF, Li XY and Sun WB (2014) The conservation of *Rhododendrons* is of greater urgency than has been previously acknowledged in China. **Biodiversity and Conservation 23**: 3149-3154.

Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. **EMBnet 17**: 10-12.

Narzisi G, O'rawe JA, Iossifov I, Fang H, Lee YH, Wang Z, Wu Y, Lyon GJ, Wigler M and Schatz MC (2014) Accurate de novo and transmitted indel detection in exome-capture data using microassembly. **Nature Methods 11**: 1033-1036.

Ng S and Corlett T (2000) Comparative reproductive biology of the six species of *Rhododendron* (Ericaceae) in Hong Kong, South China. **Canadian Journal of Botany 78**: 221-229.

Nybom H (2004) Comparison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants. **Molecular Ecology 13**: 1143-1155.

Patterson N, Price AL and Reich D (2006) Population structure and eigenanalysis. **Plos Genetics 2**: e190.

Price MN, Dehal PS and Arkin AP (2010) FastTree 2-approximately maximum-likelihood trees for large alignments. **Plos One 5**: e9490.

Rauch EM and Bar-Yam Y (2005) Estimating the total genetic diversity of a spatial field population from a sample and implications of its dependence on habitat area. **Proceedings of the National Academy of Sciences 102**: 9826-9829.

Roy SC, Moitra K and De Sarker D (2016) Assessment of genetic diversity among four orchids based on ddRAD sequencing data for conservation purposes. **Physiology and Molecular Biology of Plants 23**: 169-183.

Schiffels S and Wang K (2020) MSMC and MSMC2: the multiple sequentially markovian coalescent. In Dutheil JY (ed.) **Statistical population genomics**. Humana/Springer Nature, New York, p. 147-166.

Stone BW, Ward A, Farenwald M, Lutz AW and Wolfe AD (2019) Genetic diversity and population structure in Cary's Beardtongue *Penstemon caryi* (Plantaginaceae), a rare plant endemic to the eastern Rocky Mountains of Wyoming and Montana. **Conservation Genetics 20**: 1149-1161.

Tian Q, El-Kassaby YA and Li W (2022) Revealing the genetic structure and differentiation in endangered *Pinus bungeana* by genome-wide SNP markers. **Forests 13**: 326.

Wang C and Zhang D (2011) A novel compression tool for efficient storage of genome re-sequencing data. **Nucleic Acids Research 39**: e45-e45.

Wang S, Chen JD, Zhang WP, Hu Y, Chang LJ, Fang L, Wang Q, Lv F, Wu H, Si Z, Chen S, Cai C, Zhu X, Zhou B, Guo W and Zhang T (2015) Sequence-based ultra-dense genetic and physical maps reveal structural variations of allopolyploid cotton genomes. **Genome Biology 16**: 108.

Weng M, Sun J and Ye YZ (2012) Study of conservation priority classes of key protected plant in Henan. **Journal of Henan Agricultural Sciences 41**: 129-133.

Wu FQ, Shen SK, Zhang XJ, Wang YH and Sun WB (2015) Genetic diversity and population structure of an extremely endangered species: the world's largest Rhododendron. **AoB Plants 7**: plu082.

Yang F S, Nie S, Liu H, Shi TL, Tian XC, Zhou SS, Bao YT, Jia KH, Guo JF, Zhao W, An N, Zhang RG, Yun QZ, Wang XZ, Mannapperuma C, Porth I, El-Kassaby YA, Street NR, Wang XR, Van de Peer Y and Mao JF (2020) Chromosome-level genome assembly of a parent species of widely cultivated azaleas. **Nature Communications 11**: 1-13.

Zhang XJ, Liu XF, Liu DT, Cao YR, Li ZH, Ma YP and Ma H (2021) Genetic diversity and structure of *Rhododendron meddianum*, a plant species with extremely small populations. **Plant Diversity 43**: 472-479.

Zhao X, Ma Y, SunW, Wen X and Milne R (2012) High genetic diversity and low differentiation of *Michelia coriacea* (Magnoliaceae), a critically endangered endemic in southeast Yunnan, China. **International Journal of Molecular Sciences 13**: 4396-4411.

Zhou XJ, Li JT, Wang HL, Han JW, Zhang K, Dong SW, Zhang YZ, Ya HY, Cheng YW and Sun SS (2022) The chromosome-scale genome assembly, annotation and evolution of *Rhododendron henanense* subsp. *lingbaoense*. **Molecular Ecology Resources 22**: 988-1001.