Chloroplast View of the Population Genetics and Phylogeography of a Widely Distributed Shrub Species, Rhododendron dauricum (Ericaceae)

Authors: Nan Jiang, Li Man, Wei Zhang, Hai-Xiao Dong, Hua-Ying Wang, et. al.

Source: Systematic Botany, 41(3) : 626-633

Published By: The American Society of Plant Taxonomists

URL: https://doi.org/10.1600/036364416X692343
Chloroplast View of the Population Genetics and Phylogeography of a Widely Distributed Shrub Species, *Rhododendron dauricum* (Ericaceae)

Nan Jiang, Li Man, Wei Zhang, Hai-Xiao Dong, Hua-Ying Wang, Ming-Rui Li, Feng-Xue Shi, and Ming-Zhou Sun

Key Laboratory of Vegetation Ecology of Ministry of Education, Institute of Grassland Science, Northeast Normal University, Changchun 130024, China

College of Plant Science, Jilin University, Changchun 130062, China

Author for correspondence (sunmz100@nenu.edu.cn)

Communicating Editor: Michael J. Moore

**Abstract**—Climatic oscillations during the last glacial maximum (LGM) have played important roles in the distributions and genetic structures of extant species. Here, we investigated the population structure of a deciduous shrub species, *Rhododendron dauricum* L., across its current geographic distribution in northeastern China based on six chloroplast regions (*trnS-trnM, trnL-ndhJ, rps16-rpoB, trnS-trnG, trnH-psbA*, and *trnK-matK*). A total of 11 haplotypes were identified. The results showed a significant geographic distribution pattern (*NOD > GST, p < 0.01*) in northeastern China. In addition, our results showed that there were high levels of genetic diversity in *R. dauricum* and that more than 70% of the genetic variability was distributed among groups. Furthermore, results from the neutrality test failed to identify population expansion at the whole-species level. These findings together suggested that *R. dauricum* shows spatial genetic structure and that the LGM has had profound impacts on the geographic distribution and genetic variability of this species.

**Keywords**—Genetic diversity, haplotype, last glacial maximum, spatial genetic structure.

The climatic oscillations of the last glacial maximum (LGM, 21–18 ka) have resulted in repeated drastic environmental changes that have profoundly shaped the current distributions and genetic structures of extant species (Hewitt 2000, 2003). Based on population genetics, fossil pollen, and paleoecological evidence, a series of recent studies from diverse plant species have revealed some basic phylogeographic patterns and concepts that have together shed light on the Quaternary history of temperate plants (Kamei 1981; Huntley and Birks 1983; Tsukada 1988; Webb and Bartlein 1992; Lang 1994). One classic view is the southern refugia model (Bennett et al. 1991). According to this model, extensive latitudinal range shifts occurred during the LGM. For example, southward retreats during glacials were followed by rapid expansions northward during inter-/postglacials (Fujii et al. 2002; Griffin and Barrett 2004; Tsuda and Ide 2005; Bartish et al. 2006; de Lafontaine et al. 2010). The repeated contractions and expansions of plant species have shaped plant distribution patterns in the Northern Hemisphere (Hewitt 1996, 1999). Examples from Europe and North America have demonstrated the roles of southern glacial refugia in the formation of current ecosystems (Taberlet et al. 1998; Soltis et al. 2006). However, recent studies of widespread boreal and boreal forests have revealed a different viewpoint, namely that some species survived the LGM in northern refugia near the ice margin (McLachlan et al. 2005; Wall et al. 2010; Beatty and Provan 2011; Nakamura et al. 2014). These studies indicated that in the East Asian temperate forests, phylogeographic patterns cannot be simply explained by southern glacial refugia (Qiu et al. 2011). For example, Opgenoorth et al. (2010) inferred that the *Juniperus tibetica* (Kom.) Kom. complex survived the LGM in situ on the southeastern edge of the Qinghai-Tibetan Plateau, which had multiple refugia across its range. In addition, Wang et al. (2008) revealed that *Primula secundiflora* Franch. showed multiple separated refugia within a large geographic region without range expansion. Similarly, studies in Ginkgo biloba L. (Gong et al. 2008), *Croomia japonica* Miq., and *C. heterosepala* (Bak.) Oku. (Li et al. 2008) also showed only limited (the former) or even no (the latter two) range expansion; the range expansions during the inter-/postglacials of *Alsophila spinulosa* (Wall. ex R. M. Tryon (Su et al. 2005), *Pinus kwangtungensis* Chun ex Tseng (Tian et al. 2008), *Platycerat argula* Siebold et Zucc. (Qiu et al. 2009a), *Kirengeshoma korajana* Nakai, and *K. palmata* Yatabe (Qiu et al. 2009b) were localized. Those studies indicated that some (primarily dominant) species in the temperate mixed conifer and broadleaved forests of East Asia might have survived the glaciations in situ and displayed a limited population expansion during post-glaciations. These species might not have experienced drastic cycles of southward range contraction and northward expansion.

In this study, we investigated the genetic structure of *Rhododendron dauricum* L. (Ericaceae) across its current geographic distribution in northeastern China. *R. dauricum* is a temperate-deciduous Northeast Asia endemic shrub species widely distributed in northeastern China, the Korean Peninsula, and the Russian Far East. In northeastern China, *R. dauricum* occurs primarily in the Greater Khingan Mountains (GKM), Lesser Khingan Mountains (LKM), and Changbai Mountains (CM). Characterized by both a temperate climate and a complex topography, these areas have never been directly impacted by extensive and unified ice sheets (Shi et al. 1986; Liu 1988). The Quaternary history of species distributed in this region may display divergent characteristics.

To this end, we used cpDNA sequences to describe the population structure and investigate the phylogeographic history of *R. dauricum* in northeastern China. The aims of this study are as follows: (1) examine the genetic diversity and genetic differentiation among different geographic and spatial genetic groups of *R. dauricum*, (2) identify the intra-species phylogenetic relationships of chloroplast haplotypes and (3) test phylogeographic and demographic patterns in *R. dauricum*. We expect our findings to provide novel evidence for the hypothesis that some taxa persisted in situ in northeastern China during the LGM.

**Materials and Methods**

**Sampling and DNA Extraction**—A total of 166 *R. dauricum* individuals were collected from 23 populations that covered its current geographic range in northeastern China (Appendix 1; Fig. 1). In brief, a total of
is also distributed in the —

1. The four geographic regions of the 23 populations and geographic distributions of 11 cpDNA haplotypes in *R. dauricum* identified with gaps set as missing. Section size of the pie reflects the frequency of haplotype occurrence in each population. Haplotype colors correspond to those shown in panel.

55 accessions from seven populations were sampled from the GKM; 29 individuals of four populations were collected from the LKM; eight populations sampled from the CM consisted of 47 accessions; and 35 accessions from four populations were collected from the Liaodong Peninsula (LP). It is known that *R. dauricum* is also distributed in the Russian Far East and the Korean Peninsula. However, detailed information on these geographic distributions is not yet known. No samples were obtained from these areas, primarily because it is infeasible to collect samples from these regions. Mature leaves or flowers were sampled from each individual and preserved in silica gel. Sample vouchers have been deposited at the Northeast Normal University Herbarium (NENU).

Total DNA was extracted from the collected samples using a modified 4 × CTAB procedure (Doyle and Doyle 1987). DNA quality was checked by electrophoresis on 1% agarose gels, and qualified DNA samples were stored at −20°C for polymerase chain reaction (PCR) amplification.

**PCR and DNA Sequencing** — A total of 73 primer pairs collected from previous studies were examined in our study (Appendix 2; Taberlet et al. 1991; Johnson and Soltis 1994; Johnson and Soltis 1995; Desmear et al. 1995; Oxelman et al. 1997; Hamilton 1999; Grivet et al. 2001; Shaw et al. 2005; Shaw et al. 2007; Kress and Erickson 2007; Lahaye et al. 2008; Fior et al. 2013). We used two filtering steps to obtain markers showing sufficient polymorphism. First, three randomly selected individuals were amplified by each of the 73 primer pairs. After adjusting the annealing temperature, primers with nonspecific amplifications or no obvious electrophoretic bands were regarded as having failed to amplify and were then discarded. Populations AK (from GKM), BY (from LP) and H (from CM) were distant from each other, and accessions from these populations were used to perform an initial filtering step. Two individuals from each of the three populations were chosen on a random basis. PCR amplification, DNA sequencing, and alignment were used to test for polymorphism, and this step excluded markers that revealed neither SNPs nor indels. Finally, six primer pairs (trnS-trnM, Tab-e ndh1, rpoC1-rpoB, trnS-trnG, trnH-psbA, and tmk-matK), which were shown to be polymorphic among the examined accessions, were used to amplify the population samples. PCRs were performed in a 30 μl total volume containing 20–50 ng template DNA, 1 × PCR buffer (Mg²⁺ plus), 0.6 μM of both forward and reverse primers, 0.2 mM of dNTP Mixture, 3 μg of BSA, and 1 unit of tRaq polymerase (Jakara, Dalian, Liaoning, China). The PCR parameters for all amplification programs of cpDNA were as follows: 5 min of predenaturation at 95°C followed by 35 cycles of 30 sec of denaturation at 94°C, 30 sec of annealing at the temperature in Appendix 2 for each chloroplast region, elongation (90 sec for *trnS-trnM*, *trnL-ndh1* and rpoC1-rpoB, 1 min for *trnS-trnG* and *tmk-matK*), and 30 sec for *Has-H* (72°C), and a final elongation step of 5 min at 72°C. All amplified products were separated by electrophoresis on 1.5% agarose gels and sequenced with the ABI3730 sequencer (Haerbin Boshi Biotechnology Co., Ltd., Haerbin, Heilongjiang, China).

**Data Analyses** — **Population Analyses** — Initial sequence editing and assembling were performed using ContigExpress (Informax Inc., North Bethesda, MD, USA, 2000). DNA sequence alignment was performed using ClustalX 2.0 (Larkin et al. 2007) and alignments were edited manually in BioEdit 7.0.1 (Hall 1999). Genetic diversity parameters were calculated using DnaSP 5.10 (Roza 2003), including number of polymorphic sites (*S*), number of haplotypes (*h*), haplotype diversity (*H*), and nucleotide diversity (*π*) for each population and every geographic region (GKM, LKM, CM, and LP). To further detect if the studied populations have experienced expansion, Fu and Li's D and F (Fu and Li 1993) were calculated for the total population and the four geographic regions using DnaSP. A spatial analysis of molecular variance (SAMOVA) was used to analyze the spatial genetic structure of populations using SAMOVA 1.0 (Dupanloup et al. 2002), with the K varying from 2–20 and each simulation starting from 100 random initial conditions. To further assess the partitioning of genetic variation, an analysis of molecular variance (AMOVA; Excoffier et al. 1992) was performed using Arlequin 3.11 with 1,000 random permutations to test for the significance of partitions, grouping populations by geographic region.

**Phylogenetic and Phylogeographic Analyses** — Haplotype identification and network reconstruction were performed using TCS 1.21 (Clement et al. 2000). The fix connection limit was set as 41 steps when gaps were treated as missing and 65 steps when gaps were considered a fifth state to enable all haplotypes to be connected. Phylogenetic relationships among haplotypes were inferred from tests of the Neighbor-Joining (NJ) method, which were calculated using MEGA 5 (Tamura et al. 2011) with the Maximum Composite Likelihood method as a substitution model. We chose 1,000 bootstrap replicates (Felsenstein 1985) to evaluate clade support. To assess the correlation between genetic and geographic distances, an isolation-by-distance (IBD) analysis, with 1,000 randomizations on matrices of pairwise population *F*~ST~ values and the geographic distances, was performed using Isolation By Distance Web Service Version 3.32 (jensen et al. 2005). To test for the presence of phylogeographic structure, *F*~ST~ (a measure of genetic differentiation that incorporates phylogenetic distance) was compared to permuted values of *G*~ST~ (an unordered measure of genetic differentiation that does not incorporate phylogenetic distance) with the program PERMUT (http://www.pierrrono.inra.fr/genetics/labO/Softw are/PermutCpSSR). Populations containing fewer than three individuals cannot be incorporated in PERMUT; accordingly, these populations were not included.

**Results**

**cpDNA Diversity and Demographic Analyses** — The concatenated cpDNA sequences (including indels) had an aligned length of 4,130 bp, in which 57 SNPs and 11 haplotypes were identified (Table 1). All new sequence data generated from this study were submitted to GenBank under the accession numbers KT696608 – KT697615. In all 22 populations analyzed, the haplotype diversity (*H*~S~) and nucleotide diversity (*π*) ranged from 0.00594, respectively. The genetic diversity parameters of population BY were the highest, with the highest values of both *H*~S~ and *π*. In all four geographic regions, the values of haplotype diversity (*H*~S~) and nucleotide diversity (*π*) ranged from 0.353–0.605 and from 0.00011–0.00269, respectively. LP had the highest genetic diversity.

Analyses of a neutrality test of the complete data set did not support the hypothesis that *R. dauricum* had undergone a historical population expansion. Neutrality test statistics, both overall and in the four geographic regions, were nonsignificant or significant and positive (Table 1).

**Fig. 1.** The four geographic regions of the 23 populations and geographic distributions of 11 cpDNA haplotypes in *R. dauricum* identified with gaps set as missing. Section size of the pie reflects the frequency of haplotype occurrence in each population. Haplotype colors correspond to those shown in panel.
Spatial Genetic Analysis, Phylogeographic Structure, and AMOVA—The optimal number of groups indicated by SAMOVA was K = 3 (FCT = 0.82673; Appendix 3). Thus, all 23 populations were divided into three groups. Group 1 contained all 11 populations from GKM and LKM, group 2 was composed of populations from CM and LP except for population BY, and group 3 is a single-population group consisting of population BY.

The IBD test revealed a significant correlation between genetic and geographic distances of cpDNA haplotypes over all populations (r = 0.5084, p = 0.00010) as well as in each main SAMOVA group (r = 0.2974, p = 0.0270 for group 1 and r = 0.5752, p = 0.00010 for group 2). NQS of R. dauricum (0.851) was significantly higher than GST (0.749) at a global level (p < 0.01) and the two grouping schemes tested (Table 2), which indicated significant phylogeographic structure in the data set regardless of the grouping scheme used. AMOVA analysis revealed that the main source of variation was found among the groups (Table 2). For example, 72.27% and 82.67% (p < 0.01) of the total variation were derived from the geographic and SAMOVA groups, respectively. These findings suggested that R. dauricum showed obvious spatial genetic structure in its current distribution range in northeastern China.

**Table 1.** Chloroplast genetic diversity and results of neutrality tests for R. dauricum. S = number of polymorphic sites (SNPs); h = number of haplotypes, $H_d$ = haplotype diversity; $\pi$ = nucleotide diversity; NC, not calculated; * = p < 0.02.

<table>
<thead>
<tr>
<th>Group</th>
<th>S</th>
<th>h</th>
<th>$H_d$</th>
<th>$\pi$</th>
<th>Haplotype</th>
<th>Fu and Li’s D</th>
<th>Fu and Li’s F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>57</td>
<td>11</td>
<td>0.834</td>
<td>0.00238</td>
<td>H1 – H11</td>
<td>2.12684*</td>
<td>1.39211</td>
</tr>
<tr>
<td>AL</td>
<td>1</td>
<td>2</td>
<td>0.600</td>
<td>0.00015</td>
<td>H1, H2</td>
<td>NC</td>
<td>NC</td>
</tr>
<tr>
<td>AE</td>
<td>0</td>
<td>1</td>
<td>0.000</td>
<td>0.00000</td>
<td>H1</td>
<td>NC</td>
<td>NC</td>
</tr>
<tr>
<td>AK</td>
<td>1</td>
<td>2</td>
<td>0.533</td>
<td>0.00013</td>
<td>H1, H2</td>
<td>NC</td>
<td>NC</td>
</tr>
<tr>
<td>JH</td>
<td>0</td>
<td>1</td>
<td>0.000</td>
<td>0.00000</td>
<td>H1</td>
<td>NC</td>
<td>NC</td>
</tr>
<tr>
<td>ZL</td>
<td>1</td>
<td>2</td>
<td>0.200</td>
<td>0.00005</td>
<td>H1, H4</td>
<td>NC</td>
<td>NC</td>
</tr>
<tr>
<td>TP</td>
<td>1</td>
<td>2</td>
<td>0.533</td>
<td>0.00013</td>
<td>H1, H2</td>
<td>NC</td>
<td>NC</td>
</tr>
<tr>
<td>GH</td>
<td>1</td>
<td>2</td>
<td>0.400</td>
<td>0.00010</td>
<td>H1, H2</td>
<td>NC</td>
<td>NC</td>
</tr>
<tr>
<td>HS</td>
<td>0</td>
<td>1</td>
<td>0.000</td>
<td>0.00000</td>
<td>H2</td>
<td>NC</td>
<td>NC</td>
</tr>
<tr>
<td>HH</td>
<td>0</td>
<td>1</td>
<td>0.000</td>
<td>0.00000</td>
<td>H2</td>
<td>NC</td>
<td>NC</td>
</tr>
<tr>
<td>ML</td>
<td>3</td>
<td>3</td>
<td>0.378</td>
<td>0.00022</td>
<td>H1, H2, H3</td>
<td>NC</td>
<td>NC</td>
</tr>
<tr>
<td>TW</td>
<td>0</td>
<td>1</td>
<td>0.000</td>
<td>0.00000</td>
<td>H2</td>
<td>NC</td>
<td>NC</td>
</tr>
<tr>
<td>MS</td>
<td>0</td>
<td>1</td>
<td>0.000</td>
<td>0.00000</td>
<td>H8</td>
<td>NC</td>
<td>NC</td>
</tr>
<tr>
<td>YL</td>
<td>0</td>
<td>1</td>
<td>0.000</td>
<td>0.00000</td>
<td>H8</td>
<td>NC</td>
<td>NC</td>
</tr>
<tr>
<td>LJ</td>
<td>0</td>
<td>1</td>
<td>0.000</td>
<td>0.00000</td>
<td>H7</td>
<td>NC</td>
<td>NC</td>
</tr>
<tr>
<td>CB</td>
<td>0</td>
<td>1</td>
<td>0.000</td>
<td>0.00000</td>
<td>H7</td>
<td>NC</td>
<td>NC</td>
</tr>
<tr>
<td>WT</td>
<td>1</td>
<td>2</td>
<td>0.333</td>
<td>0.00008</td>
<td>H7, H9</td>
<td>NC</td>
<td>NC</td>
</tr>
<tr>
<td>LS</td>
<td>0</td>
<td>1</td>
<td>0.000</td>
<td>0.00000</td>
<td>H7</td>
<td>NC</td>
<td>NC</td>
</tr>
<tr>
<td>SJ</td>
<td>3</td>
<td>2</td>
<td>0.600</td>
<td>0.00044</td>
<td>H5, H7</td>
<td>NC</td>
<td>NC</td>
</tr>
<tr>
<td>KD</td>
<td>1</td>
<td>2</td>
<td>0.200</td>
<td>0.00005</td>
<td>H5, H6</td>
<td>NC</td>
<td>NC</td>
</tr>
<tr>
<td>FH</td>
<td>0</td>
<td>1</td>
<td>0.000</td>
<td>0.00000</td>
<td>H5</td>
<td>NC</td>
<td>NC</td>
</tr>
<tr>
<td>BY</td>
<td>45</td>
<td>4</td>
<td>0.733</td>
<td>0.00594</td>
<td>H5, H6, H10, H11</td>
<td>NC</td>
<td>NC</td>
</tr>
<tr>
<td>WL</td>
<td>0</td>
<td>1</td>
<td>0.000</td>
<td>0.00000</td>
<td>H5</td>
<td>NC</td>
<td>NC</td>
</tr>
<tr>
<td>GKM</td>
<td>2</td>
<td>3</td>
<td>0.446</td>
<td>0.00011</td>
<td>H1, H2, H4</td>
<td>-0.91934</td>
<td>-0.72396</td>
</tr>
<tr>
<td>LKM</td>
<td>3</td>
<td>3</td>
<td>0.463</td>
<td>0.00031</td>
<td>H1, H2, H3</td>
<td>0.95411</td>
<td>1.31094</td>
</tr>
<tr>
<td>CM</td>
<td>6</td>
<td>4</td>
<td>0.605</td>
<td>0.00032</td>
<td>H5, H7, H8, H9</td>
<td>1.17948</td>
<td>0.90646</td>
</tr>
<tr>
<td>LP</td>
<td>45</td>
<td>4</td>
<td>0.353</td>
<td>0.00269</td>
<td>H5, H6, H10, H11</td>
<td>1.68697*</td>
<td>1.33221</td>
</tr>
</tbody>
</table>

**Table 2.** Results of the analysis of molecular variance (AMOVA) and the test of phylogeographic structure (NQS > GST) of cpDNA sequence data from R. dauricum populations. Groupings were as follows: (a) by geography with four regions and (b) by groupings from SAMOVA (for K = 3).

<table>
<thead>
<tr>
<th>Group</th>
<th>d.f.</th>
<th>Fixation indices</th>
<th>% of variation</th>
<th>P</th>
<th>NQS</th>
<th>GST</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>FCT</td>
<td>0.72268</td>
<td>72.27</td>
<td>&lt;0.01</td>
<td>0.721</td>
<td>0.512</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>19</td>
<td>FSC</td>
<td>0.47232</td>
<td>13.10</td>
<td>&lt;0.01</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>143</td>
<td>FST</td>
<td>0.85566</td>
<td>14.63</td>
<td>&lt;0.01</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>(b)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>FCT</td>
<td>0.82673</td>
<td>82.67</td>
<td>&lt;0.01</td>
<td>0.594</td>
<td>0.286</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>20</td>
<td>FSC</td>
<td>0.33991</td>
<td>5.89</td>
<td>&lt;0.01</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>143</td>
<td>FST</td>
<td>0.88852</td>
<td>11.44</td>
<td>&lt;0.01</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
H11 were clustered as Clade III. The support bootstrap values for internal nodes of the tree were generally high. We checked the segregating sites of these haplotypes. Only 57 SNPs were observed, and the three clades showed quite different nucleotide variation patterns among the three clades.

**Discussion**

**Genetic Diversity and Population Structure**—Of the 73 primer pairs tested in this study, six were amplified successfully and shown to be polymorphic. When confronted with chloroplast marker choice in *Rhododendron*, Ericaceae, or other related families, these six primer pairs and associated chloroplast regions may be priorities. From the sequences amplified by those primer pairs, our results showed that the haplotype diversity (*H*<sub>d</sub> = 0.834) of *R. dauricum* was higher than that of other temperate shrub or small-sized tree species, such as *R. delavayi* (C. B. Clarke) Ridley (*H*<sub>d</sub> = 0.634; Sharma et al. 2014), wild *Prunus pseudocerasus* Lindl. (*H*<sub>d</sub> = 0.557; Shi and Zhang 2015), and the *Rosa sericea* Wall. ex Lindl. complex (*H*<sub>d</sub> = 0.522–0.779; Gao et al. 2015), and it was also higher than the mean for cpDNA as reported by Petit et al. (2005) in 170 species of angiosperms (*H*<sub>d</sub> = 0.670).

Of all 23 populations, population BY had the highest genetic diversity (*H*<sub>d</sub> = 0.733, *π* = 0.00594). Of all four geographic regions, CM had the highest haplotype diversity (*H*<sub>d</sub> = 0.605) and LP had the highest nucleotide diversity (*π* = 0.00269). Therefore, the two regions were identified as potential genetic diversity centers of *R. dauricum*. The results of AMOVA showed that the source of most genetic variability was among groups. According to the standard of Wright (1978), the differentiation values calculated among groups (*F*<sub>CT</sub>), populations (*F*<sub>SC</sub>), and individuals (*F*<sub>ST</sub>) were all more than 0.25, indicating a great genetic differentiation of groups on different levels. Similar results were obtained from the SAMOVA cluster, showing that the genetic structures of different clusters were significantly different, while the origins of *R. dauricum* from different groups were diverse.

The observation of high genetic diversity indicated that *R. dauricum* might not have undergone a severe genetic bottleneck in the evolutionary process. In contrast, the nucleotide diversity (*π* = 0.00238) of *R. dauricum* does not show a high level compared with the haplotype diversity. Our results were similar to previous studies, such as *Kandelia candel* (L.) Druce (*π* = 0.00304; Chiang et al. 2001), *Fatsia japonica* (Thunb.) Decne. et Planch. (*π* = 0.00285; Chen 2004), and wild *Prunus pseudocerasus* (*π* = 0.00209; Chen et al. 2015). To test the factors that have led to the contrasting observation of the high haplotype diversity and the low nucleotide diversity of *R. dauricum*, we checked the segregating sites of these haplotypes. Only 57 SNPs were observed, and the three clades showed quite different nucleotide variation patterns. Accordingly, we consider that the small number of SNPs within each population might be one factor that led to the contrasting observations. Moreover, the deep phylogeographic splits between the geographic or SAMOVA groups are, most likely, the other reason for high haplotype diversity but low nucleotide diversity, a pattern that is considered to result from isolation among different refugia.

**Phylogeographic Structure and Inference of Demographic Patterns**—Our results on the phylogeographic structure of *R. dauricum* were consistent with previous studies of temperate montane flora in northeastern China, which were strongly structured phylogeographically (Chen et al. 2008; Tian et al. 2009; Chen et al. 2012b). In this study, *N*<sub>ST</sub> was significantly
higher than $G_{ST}$ at the global, geographic, and grouping levels. Moreover, the network and NJ tree of haplotypes showed clear geographic structure. The amazing uniformity of SAMOVA clustering ($K = 3$) and geography further confirmed that populations from different origins had a distinct genetic makeup. Therefore, we propose that the current distribution of R. dauricum might result from occupying different refugia during the LGM. For example, populations from the GKM and CM might have had distinct refugia during the LGM. Similar phenomena have also been observed in Ostryopsis davidiana Decne. (Tian et al. 2009), Fagus engleriana Seemen ex Diels (Lei et al. 2012), Quercus variabilis Blume (Chen et al. 2012a) and Juglans cathayensis Dode (Bai et al. 2014). In addition, the results of the neutrality test suggested stable population dynamics recently, implying no spatial expansions in the post-glaciations. Regarding Fu and Li’s $D$, we found that it shows a significant positive value at the species level, possibly due to the population structure observed within this species. Similarly, when all 23 populations or populations within each main SAMOVA group were pooled together, the IBD test showed a significant correlation between genetic and geographic distance. It also indicates that these populations might not have undergone recent population expansion. Ikeda and Setoguchi (2013) used an isolation-with-migration (IM) model to estimate demographic history of Phyllocladus nipponica Makino. Their results showed that mixing of populations and gene exchange did not emerge in the post-glaciations, so the differences between isolated refugia have not been eliminated. Likewise, our results proved that the genetic imprint from different refugia still deeply influences the genetic structure of R. dauricum in northeastern China.

Qian and Ricklefs (2000) and Harrison et al. (2001) proposed two different hypotheses for Asian species diversity. Harrison et al. (2001) suggested that temperate forests would have retreated southward to approximately 30°N during the LGM, so a genetic signature of ‘southern richness’ and ‘northern purity’ in many species should be detected. On the contrary, Qian and Ricklefs (2000) observed that there have been multiple isolated refugia for northern China plants possibly due to the great physiographical heterogeneity in this region, providing abundant opportunities for allopatric speciation in the Quaternary. Our results are consistent with the latter view. Habitat fragmentation formed barriers against gene flow, further leading to genetic diversification, accelerating variation and even speciation (DeChaine and Martin 2004; Li et al. 2012; Gao et al. 2015). Under this hypothesis, the long-term isolation among multiple refugia and little admixture among populations of different origins have caused substantial genetic diversification in R. dauricum. Given that this may be an example of cryptic speciation, we therefore checked the nucleotide variation patterns of R. dauricum and we found that there exist some fixation sites among the three clades. The deep divergence of these clades suggested the possibility of cryptic speciation within this species. However, only a few selected loci were used in this study, which largely limited our inference on this topic. To this end, further investigations based on genome-wide markers might provide novel insights into this issue.

In conclusion, cpDNA haplotypes have delineated the genetic diversity and population genetic structure of R. dauricum, providing insights into the phylogeographic patterns. Our results showed high levels of genetic diversity and great genetic differentiation among different groups. A significant phylogeographic structure was also detected in this species. Moreover, the phylogeographic and demographic patterns of multiple localized refugia and no postglacial range expansion have been proved. Our results are consistent with the hypothesis that the complex landforms of East Asia have provided local species with opportunities to survive the last glaciation in situ.

ACKNOWLEDGMENTS. We thank Dr. Hong-Xing Xiao for help during sampling and Dr. Lin-Feng Li for his preview of this manuscript. We also thank the editors and two anonymous reviewers for their valuable comments and suggestions, which have greatly improved the manuscript. This study was supported by the National Natural Science Foundation of China (Nos. J1210070, 31170429 and 31570194).

LITERATURE CITED


Plants On Earth? e340313.


to East China and South Japan, inferred from chloroplast DNA sequence variation. Taxon 58: 1226–1241.


APPENDIX 1. Sampling information for populations used in this study. Data are presented in the following order: population, location, geographic regions, latitude, longitude, elevation (m), sample size, sampling date, and individual numbers.
Appendix 3. K value and FST value pairs observed in SAMOVA with the K values varying from 2 to 20.

| K | FST value pairs observed in SAMOVA |