



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

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

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

Published By: The American Society of Plant Taxonomists

URL: <https://doi.org/10.1600/036364416X692343>

---

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.

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

Nan Jiang,<sup>1</sup> Li Man,<sup>1</sup> Wei Zhang,<sup>1</sup> Hai-Xiao Dong,<sup>2</sup> Hua-Ying Wang,<sup>1</sup> Ming-Rui Li,<sup>1</sup>  
Feng-Xue Shi,<sup>1</sup> and Ming-Zhou Sun<sup>1,3</sup>

<sup>1</sup>Key Laboratory of Vegetation Ecology of Ministry of Education, Institute of Grassland Science,  
Northeast Normal University, Changchun 130024, China

<sup>2</sup>College of Plant Science, Jilin University, Changchun 130062, China

<sup>3</sup>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-trnF*, *trnL-ndhJ*, *rpoC1-rpoB*, *trnS-trnG*, *trnH-psbA*, and *trnK-matK*). A total of 11 haplotypes were identified. The results showed a significant geographic distribution pattern ( $N_{ST} > G_{ST}$ ,  $p < 0.01$ ) 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 range 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 north temperate and boreal plants 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 southeast 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

Hook.) R. M. Tryon (Su et al. 2005), *Pinus kwangtungensis* Chun ex Tsiang (Tian et al. 2008), *Platycrater arguta* Siebold et Zucc. (Qiu et al. 2009a), *Kirengeshoma koreana* 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

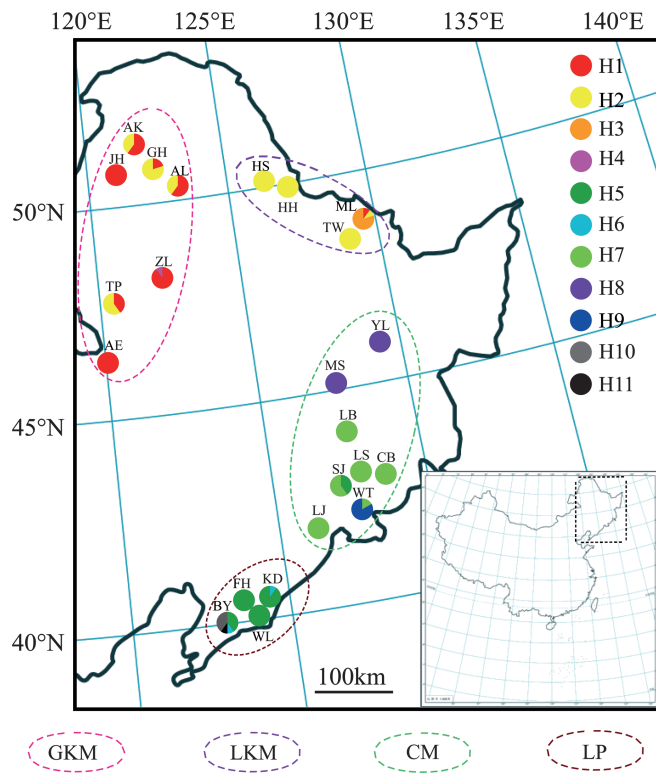


FIG. 1. The four geographic regions of the 23 populations and geographic distributions of 11 cpDNA haplotypes in *R. dauricum* identified with gaps set as missing. Section size of the pie reflects the frequency of haplotype occurrence in each population. Haplotype colors correspond to those shown in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

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 1.0% agarose gels, and qualified DNA samples were stored at  $-20^{\circ}\text{C}$  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; Demesure 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 SJ (from CM) were distant from each other, and accessions from these populations were used to perform the final 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 (*trnSf-trnMf*, *TabE-ndhJ*, *rpoC1f-rpoB*, *trnS<sup>(GCU)</sup>-trnG<sup>(UCC)</sup>*, *trnH-psbA* and *trnK-matK*), which were shown to be polymorphic among the examined accessions, were used to amplify the population samples. PCRs were performed in a 30  $\mu\text{L}$  total volume containing 20–50 ng template DNA, 1 × PCR buffer ( $\text{Mg}^{2+}$  plus), 0.6  $\mu\text{M}$  of both forward and reverse primers, 0.2 mM of dNTP Mixture, 3  $\mu\text{g}$  of

BSA, and 1 unit of rTaq polymerase (Takara, Dalian, Liaoning, China). The PCR parameters for all amplification programs of cpDNA were as follows: 5 min of predenaturation at  $95^{\circ}\text{C}$  followed by 35 cycles of 30 sec of denaturation at  $94^{\circ}\text{C}$ , 30 sec of annealing at the temperature in Appendix 2 for each chloroplast region, elongation (90 sec for *trnS-trnM*, *trnL-ndhJ* and *rpoC1-rpoB*, 1 min for *trnS-trnG* and *trnK-matK*, and 30 sec for *trnH-psbA*) at  $72^{\circ}\text{C}$ , and a final elongation step of 10 min at  $72^{\circ}\text{C}$ . 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 (Rozas et al. 2003), including number of polymorphic sites ( $S$ ), number of haplotypes ( $h$ ), haplotype diversity ( $H_d$ ), and nucleotide diversity ( $\pi$ ) 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 distance, 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 3.23 (Jensen et al. 2005). To test for the presence of phylogeographic structure,  $N_{ST}$  (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.pierroton.inra.fr/genetics/labo/Software/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 ( $H_d$ ) and nucleotide diversity ( $\pi$ ) ranged from 0–0.733 and from 0–0.00594, respectively. The genetic diversity parameters of population BY were the highest, with the highest values of both  $H_d$  and  $\pi$ . In all four geographic regions, the values of haplotype diversity ( $H_d$ ) and nucleotide diversity ( $\pi$ ) 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 non-significant or significant and positive (Table 1).

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

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

**Spatial Genetic Analysis, Phylogeographic Structure, and AMOVA**—The optimal number of groups indicated by SAMOVA was  $K = 3$  ( $F_{CT} = 0.82673$ ; 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 ( $r = 0.5084$ ,  $p < 0.0010$ ) as well as in each main SAMOVA group ( $r = 0.2974$ ,  $p = 0.0270$  for group 1 and  $r = 0.5752$ ,  $p < 0.0010$  for group 2).  $N_{ST}$  of *R. dauricum* (0.851) was significantly higher than  $G_{ST}$  (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.

**Haplotype and Phylogenetic Analyses**—The topology obtained from TCS was used to infer relationships among the 11 cpDNA haplotypes (Fig. 2a; Fig. 1), of which H1 (observed 39 times, in 23.49% of individuals), H2 (observed 36 times, 21.69%), H5 (observed 30 times, 18.07%), and H7 (observed 26 times, 15.66%) were the most frequent haplotypes. In addition, the haplotypes H10, H2, H5, and H7 were located in the center of the haplotype network. Moreover, our results also showed that all haplotypes from CM (H5, H7, H8, and H9) and haplotype H6 from LP were clustered together as one clade (Clade I). Similarly, Clade II consisted of H1, H2, H3, and H4, which only occurred in GKM and LKM. The remaining haplotypes (H10 and H11) occurred in LP and were clustered as Clade III. More than half of the populations were fixed for a single haplotype (13 of the 23 populations). Haplotypes H3, H4, and H9 were found to be private alleles of populations ML, ZL, and WT, respectively, whereas haplotypes H10 and H11 were found to be private alleles in population BY. When gaps were set as fifth informative sites (Fig. 2b), there were no significant changes in the haplotype network. Although more haplotypes were identified, they were merely further subdivisions of the haplotypes with gaps regarded as missing.

Phylogenetic relationships of the 11 cpDNA haplotypes were assessed under NJ analyses. All haplotypes were clustered into three clades (Fig. 2c), which was consistent with the haplotype network (Fig. 2a). Haplotypes H5–H9 were clustered as Clade I; Clade II contained H1–H4; H10 and

TABLE 2. Results of the analysis of molecular variance (AMOVA) and the test of phylogeographic structure ( $N_{ST} > G_{ST}$ ) 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  $K = 3$ ).  $F_{CT}$  = differentiation among groups of populations,  $F_{SC}$  = differentiation among populations within groups,  $F_{ST}$  = differentiation among individuals within populations.

Group	d.f.	Fixation indices	% of variation	<i>P</i>	$N_{ST}$	$G_{ST}$	<i>P</i>
(a)	3	$F_{CT} = 0.72268$	72.27	<0.01	0.721	0.512	<0.01
	19	$F_{SC} = 0.47232$	13.10	<0.01	-	-	-
	143	$F_{ST} = 0.85366$	14.63	<0.01	-	-	-
(b)	2	$F_{CT} = 0.82673$	82.67	<0.01	0.594	0.286	<0.01
	20	$F_{SC} = 0.33991$	5.89	<0.01	-	-	-
	143	$F_{ST} = 0.88562$	11.44	<0.01	-	-	-



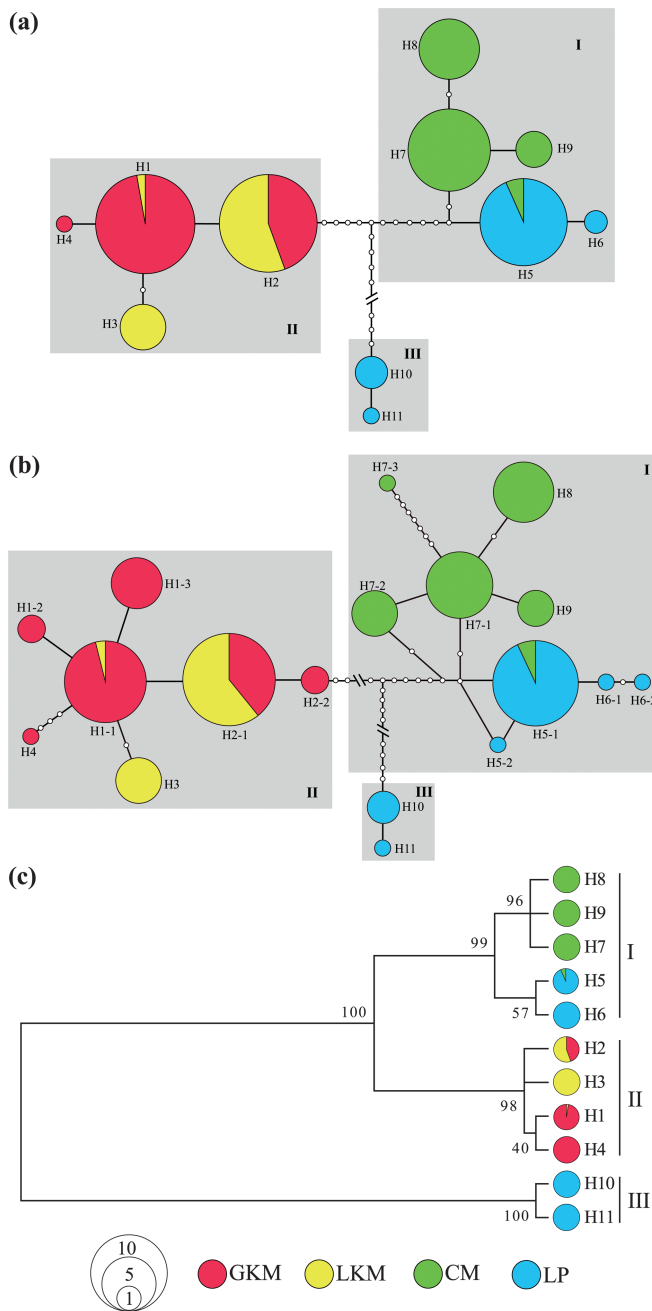


FIG. 2. (a) Statistical parsimony network of 11 cpDNA haplotypes recovered from 166 accessions of *R. dauricum* with gaps set as missing. Haplotype circle area represents relative frequency. Haplotypes were arbitrarily assigned numbers 1–11 and color-coded to indicate different geographic regions. The dots on branches indicate mutational steps, and the double hatch marks represent the omission of dots. (b) Statistical parsimony network with gaps set as a fifth state. Haplotype numbers were based on haplotypes in (a). (c) NJ phylogenetic trees for the 11 cpDNA haplotypes. Numbers above branches are bootstrap support values based on 1,000 replicates. Section size of the pie reflects the frequency of individuals from each geographic region.

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 the 11 haplotypes, attempting to discover the pattern of nucleotide variation among the three clades (Fig. 3). In all, 88 polymorphisms (including 57 SNPs and 31 indels) were observed, and we identified the fixation sites among these clades (six for Clade I, 10 for Clade II,

and 56 for Clade III, including indels), which 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 ( $H_d = 0.834$ ) of *R. dauricum* was higher than that of other temperate shrub or small-sized tree species, such as *R. delavayi* (C. B. Clarke) Ridley ( $H_d = 0.634$ ; Sharma et al. 2014), wild *Prunus pseudocerasus* Lindl. ( $H_d = 0.557$ ; Chen et al. 2015), *Zygophyllum xanthoxylon* (Bunge) Maxim. ( $H_d = 0.534$ ; Shi and Zhang 2015), and the *Rosa sericea* Wall. ex Lindl. complex ( $H_d = 0.522$ – $0.779$ ; 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 ( $H_d = 0.670$ ).

Of all 23 populations, population BY had the highest genetic diversity ( $H_d = 0.733$ ,  $\pi = 0.00594$ ). Of all four geographic regions, CM had the highest haplotype diversity ( $H_d = 0.605$ ) and LP had the highest nucleotide diversity ( $\pi = 0.00269$ ). 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 ( $F_{CT}$ ), populations ( $F_{SC}$ ), and individuals ( $F_{ST}$ ) 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 ( $\pi = 0.00238$ ) 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 ( $\pi = 0.00304$ ; Chiang et al. 2001), *Fatsia japonica* (Thunb.) Decne. et Planch. ( $\pi = 0.00285$ ; Chen 2004), and wild *Prunus pseudocerasus* ( $\pi = 0.00209$ ; 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,  $N_{ST}$  was significantly

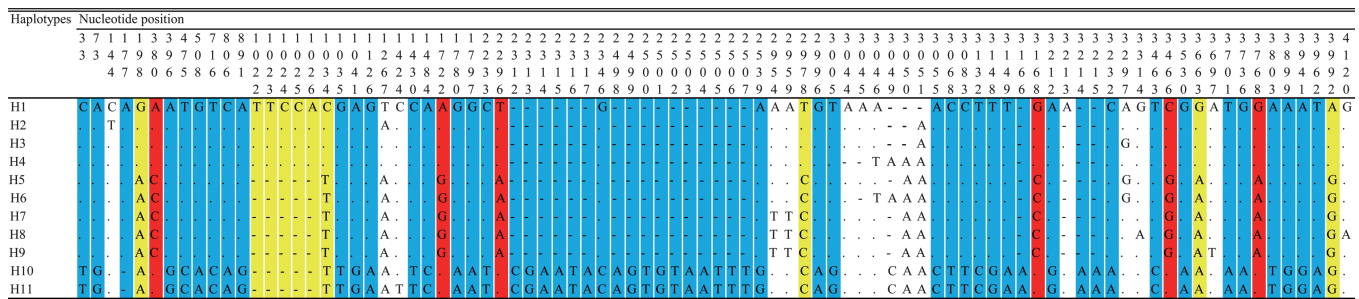


FIG. 3. Segregating sites of the 11 cpDNA haplotypes. All sequences are compared to the reference haplotype H1. The fixation sites in Clade I (H5–H9), Clade II (H1–H4) and Clade III (H10 and H11) are expressed in red, yellow, and blue.

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 *Phyllodoce 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

- Bai, W. N., W. T. Wang, and D. Y. Zhang. 2014. Contrasts between the phylogeographic patterns of chloroplast and nuclear DNA highlight a role for pollen-mediated gene flow in preventing population divergence in an East Asian temperate tree. *Molecular Phylogenetics and Evolution* 81: 37–48.
- Bartish, I. V., J. W. Kadereit, and H. P. Comes. 2006. Late Quaternary history of *Hippophae rhamnoides* L. (Elaeagnaceae) inferred from chalcone synthase intron (Chsi) sequences and chloroplast DNA variation. *Molecular Ecology* 15: 4065–4083.
- Beatty, G. E. and J. Provan. 2011. Phylogeographic analysis of North American populations of the parasitic herbaceous plant *Monotropa hypopitys* L. reveals a complex history of range expansion from multiple late glacial refugia. *Journal of Biogeography* 38: 1585–1599.
- Bennett, K., P. Tzedakis, and K. Willis. 1991. Quaternary refugia of north European trees. *Journal of Biogeography* 18: 103–115.
- Chen, D., X. Zhang, H. Kang, X. Sun, S. Yin, H. Du, N. Yamanaka, W. Gapare, H. X. Wu, and C. Liu. 2012a. Phylogeography of *Quercus variabilis* based on chloroplast DNA sequence in East Asia: Multiple glacial refugia and mainland-migrated island populations. *PLoS One* 7: e47268.
- Chen, K., R. J. Abbott, R. I. Milne, X. M. Tian, and J. Liu. 2008. Phylogeography of *Pinus tabulaeformis* Carr. (Pinaceae), a dominant

- species of coniferous forest in northern China. *Molecular Ecology* 17: 4276–4288.
- Chen, S. 2004. *Phylogeography of Fatsia (Araliaceae): an endemic genus in East Asia*. M.S. thesis. Tainan, Taiwan: Cheng Kung University.
- Chen, S., Y. Xing, T. Su, Z. Zhou, E. D. L. Dilcher, and D. E. Soltis. 2012b. Phylogeographic analysis reveals significant spatial genetic structure of *Incarvillea sinensis* as a product of mountain building. *BMC Plant Biology* 12: 58.
- Chen, T., Q. Chen, Y. Luo, Z. L. Huang, J. Zhang, H. R. Tang, D. M. Pan, and X. R. Wang. 2015. Phylogeography of Chinese cherry (*Prunus pseudocerasus* Lindl.) inferred from chloroplast and nuclear DNA: insights into evolutionary patterns and demographic history. *Plant Biology* 17: 787–797.
- Chiang, T. Y., Y. C. Chiang, Y. J. Chen, C. H. Chou, S. Havanond, T. N. Hong, and S. Huang. 2001. Phylogeography of *Kandelia candel* in East Asiatic mangroves based on nucleotide variation of chloroplast and mitochondrial DNAs. *Molecular Ecology* 10: 2697–2710.
- Clement, M., D. Posada, and K. A. Crandall. 2000. TCS: a computer program to estimate gene genealogies. *Molecular Ecology* 9: 1657–1659.
- de Lafontaine, G., J. Turgeon, and S. Payette. 2010. Phylogeography of white spruce (*Picea glauca*) in eastern North America reveals contrasting ecological trajectories. *Journal of Biogeography* 37: 741–751.
- DeChaine, E. G. and A. P. Martin. 2004. Historic cycles of fragmentation and expansion in *Parnassius smintheus* (Papilionidae) inferred using mitochondrial DNA. *Evolution* 58: 113–127.
- Demesure, B., N. Sodji, and R. J. Petit. 1995. A set of universal primers for amplification of polymorphic non-coding regions of mitochondrial and chloroplast DNA in plants. *Molecular Ecology* 4: 129–131.
- Doyle, J. J. and J. L. Doyle. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin* 19: 11–15.
- Dupanloup, I., S. Schneider, and L. Excoffier. 2002. A simulated annealing approach to define the genetic structure of populations. *Molecular Ecology* 11: 2571–2581.
- Excoffier, L., P. E. Smouse, and J. M. Quattro. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* 131: 479–491.
- Felsenstein, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39: 783–791.
- Fior, S., M. Li, B. Oxelman, R. Viola, S. A. Hodges, L. Ometto, and C. Varotto. 2013. Spatiotemporal reconstruction of the *Aquilegia* rapid radiation through next-generation sequencing of rapidly evolving cpDNA regions. *The New Phytologist* 198: 579–592.
- Fu, Y. X. and W. H. Li. 1993. Statistical tests of neutrality of mutations. *Genetics* 133: 693–709.
- Fujii, N., N. Tomaru, K. Okuyama, T. Koike, T. Mikami, and K. Ueda. 2002. Chloroplast DNA phylogeography of *Fagus crenata* (Fagaceae) in Japan. *Plant Systematics and Evolution* 232: 21–33.
- Gao, Y. D., Y. Zhang, X. F. Gao, and Z. M. Zhu. 2015. Pleistocene glaciations, demographic expansion and subsequent isolation promoted morphological heterogeneity: A phylogeographic study of the alpine *Rosa sericea* complex (Rosaceae). *Scientific Reports* 5: 11698.
- Gong, W., C. Chen, C. Dobes, C. X. Fu, and M. A. Koch. 2008. Phylogeography of a living fossil: Pleistocene glaciations forced *Ginkgo biloba* L. (Ginkgoaceae) into two refuge areas in China with limited subsequent postglacial expansion. *Molecular Phylogenetics and Evolution* 48: 1094–1105.
- Griffin, S. R. and S. C. H. Barrett. 2004. Post-glacial history of *Trillium grandiflorum* (Melanthiaceae) in Eastern North America: Inferences from phylogeography. *American Journal of Botany* 91: 465–473.
- Grivet, D., B. Heinze, G. Vendramin, and R. Petit. 2001. Genome walking with consensus primers: application to the large single copy region of chloroplast DNA. *Molecular Ecology Notes* 1: 345–349.
- Hall, T. A. 1999. BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* 41: 95–98.
- Hamilton, M. 1999. Four primer pairs for the amplification of chloroplast intergenic regions with intraspecific variation. *Molecular Ecology* 8: 521–523.
- Harrison, S. P., G. Yu, H. Takahara, and I. C. Prentice. 2001. Palaeovegetation — Diversity of temperate plants in east Asia. *Nature* 413: 129–130.
- Hewitt, G. M. 1996. Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society. Linnean Society of London* 58: 247–276.
- Hewitt, G. M. 1999. Post-glacial re-colonization of European biota. *Biological Journal of the Linnean Society. Linnean Society of London* 68: 87–112.
- Hewitt, G. M. 2000. The genetic legacy of the Quaternary ice ages. *Nature* 405: 907–913.
- Hewitt, G. M. 2003. Ice ages, species distributions, and evolution. Evolution on planet earth. Pp. 339–361 in *Evolution on Planet Earth: Impact of the Physical Environment*, ed. L. Rothschild and A. Lister. London: Academic Press.
- Huntley, B. and H. J. B. Birks. 1983. *An atlas of past and present pollen maps for Europe: 0–13,000 years ago*. Cambridge: Cambridge University Press.
- Ikeda, H. and H. Setoguchi. 2013. A multilocus sequencing approach reveals the cryptic phylogeographical history of *Phyllocladus nipponica* Makino (Ericaceae). *Biological Journal of the Linnean Society. Linnean Society of London* 110: 214–226.
- Jensen, J. L., A. J. Bohonak, and S. T. Kelley. 2005. Isolation by distance, web service. *BMC Genetics* 6: 13.
- Johnson, L. A. and D. E. Soltis. 1994. *matK* DNA sequences and phylogenetic reconstruction in Saxifragaceae s. str. *Systematic Botany* 19: 143–156.
- Johnson, L. A. and D. E. Soltis. 1995. Phylogenetic inference in Saxifragaceae sensu stricto and *Gilia* (Polemoniaceae) using *matK* sequences. *Annals of the Missouri Botanical Garden* 82: 149–175.
- Kamei, T. 1981. Research Group for the Biogeography from Würm Glacial. 1981. Fauna and flora of the Japanese Islands in the last glacial. *Quaternary Research* 20: 191–205.
- Kress, W. J. and D. L. Erickson. 2007. A two-locus global DNA barcode for land plants: the coding *rbcl* gene complements the non-coding *trnH-psbA* spacer region. *PLoS One* 2: e508.
- Lahaye, R., V. Savolainen, S. Duthoit, O. Maurin, and M. van der Bank. 2008. A test of *psbK-psbI* and *atpF-atpH* as potential plant DNA barcodes using the flora of the Kruger National Park as a model system (South Africa). *Nature Precedings* hdl:10101/npre.2008.1896.1 : 1.
- Lang, G. 1994. *Quartäre Vegetationsgeschichte Europas: Methoden und Ergebnisse*. Jena: Spektrum Akademischer Verlag.
- Larkin, M. A., G. Blackshields, N. P. Brown, R. Chenna, P. A. McGettigan, H. McWilliam, F. Valentin, I. M. Wallace, A. Wilm, R. Lopez, J. D. Thompson, T. J. Gibson, and D. G. Higgins. 2007. Clustal W and clustal X version 2.0. *Bioinformatics* 23: 2947–2948.
- Lei, M., Q. Wang, Z. J. Wu, J. Lopez-Pujol, D. Z. Li, and Z. Y. Zhang. 2012. Molecular phylogeography of *Fagus engleriana* (Fagaceae) in subtropical China: limited admixture among multiple refugia. *Tree Genetics & Genomes* 8: 1203–1212.
- Li, E. X., S. Yi, Y. X. Qiu, J. T. Guo, H. P. Comes, and C. X. Fu. 2008. Phylogeography of two East Asian species in *Crotonia* (Stemonaceae) inferred from chloroplast DNA and ISSR fingerprinting variation. *Molecular Phylogenetics and Evolution* 49: 702–714.
- Li, Z., G. Yu, D. Rao, and J. Yang. 2012. Phylogeography and Demographic History of *Babina pleuraden* (Anura, Ranidae) in Southwestern China. *PLoS One* 7: e34013.
- Liu, K. 1988. Quaternary history of the temperate forests of China. *Quaternary Science Reviews* 7: 1–20.
- McLachlan, J. S., J. S. Clark, and P. S. Manos. 2005. Molecular indicators of tree migration capacity under rapid climate change. *Ecology* 86: 2088–2098.
- Nakamura, K., G. Kokubugata, R. R. Rubite, C. J. Huang, Y. Kono, H. A. Yang, A. Lopez-Feliciano, M. L. Labuguen, M. Yokota, and C. I. Peng. 2014. In situ glacial survival at the northern limit of tropical insular Asia by a lowland herb *Begonia fenicis* (Begoniaceae). *Botanical Journal of the Linnean Society* 174: 305–325.
- Opgenoorth, L., G. G. Vendramin, K. Mao, G. Miehle, S. Miehle, L. J. Liu, and B. Ziegenhagen. 2010. Tree endurance on the Tibetan Plateau marks the world's highest known tree line of the Last Glacial Maximum. *The New Phytologist* 185: 332–342.
- Oxelman, B., M. Lidén, and D. Berglund. 1997. Chloroplast rps16 intron phylogeny of the tribe Sileneae (Caryophyllaceae). *Plant Systematics and Evolution* 206: 393–410.
- Petit, R. J., J. Duminil, S. Fineschi, A. Hampe, D. Salvini, and G. G. Vendramin. 2005. Comparative organization of chloroplast, mitochondrial and nuclear diversity in plant populations. *Molecular Ecology* 14: 689–701.
- Qian, H. and R. E. Ricklefs. 2000. Large-scale processes and the Asian bias in species diversity of temperate plants. *Nature* 407: 180–182.
- Qiu, Y. X., C. X. Fu, and H. P. Comes. 2011. Plant molecular phylogeography in China and adjacent regions: Tracing the genetic imprints of Quaternary climate and environmental change in the world's most diverse temperate flora. *Molecular Phylogenetics and Evolution* 59: 225–244.
- Qiu, Y. X., X. S. Qi, X. F. Jin, X. Y. Tao, C. X. Fu, A. Naiki, and H. P. Comes. 2009a. Population genetic structure, phylogeography, and demographic history of *Platycodon arguta* (Hydrangeaceae) endemic



- to East China and South Japan, inferred from chloroplast DNA sequence variation. *Taxon* 58: 1226–1241.
- Qiu, Y. X., Y. Sun, X. P. Zhang, J. Lee, C. X. Fu, and H. P. Comes. 2009b. Molecular phylogeography of East Asian *Kirengeshoma* (Hydrangeaceae) in relation to Quaternary climate change and landbridge configurations. *The New Phytologist* 183: 480–495.
- Rozas, J., J. C. Sanchez-DelBarrio, X. Messeguer, and R. Rozas. 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* 19: 2496–2497.
- Sharma, A., R. C. Poudel, A. Li, J. Xu, and K. Guan. 2014. Genetic diversity of *Rhododendron delavayi* var. *delavayi* (C. B. Clarke) Ridley inferred from nuclear and chloroplast DNA: implications for the conservation of fragmented populations. *Plant Systematics and Evolution* 300: 1853–1866.
- Shaw, J., E. B. Lickey, J. T. Beck, S. B. Farmer, W. Liu, J. Miller, K. C. Siripun, C. T. Winder, E. E. Schilling, and R. L. Small. 2005. The tortoise and the hare II: relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. *American Journal of Botany* 92: 142–166.
- Shaw, J., E. B. Lickey, E. E. Schilling, and R. L. Small. 2007. Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: The tortoise and the hare III. *American Journal of Botany* 94: 275–288.
- Shi, X. J. and M. L. Zhang. 2015. Phylogeographical structure inferred from cpDNA sequence variation of *Zygophyllum xanthoxylon* across north-west China. *Journal of Plant Research* 128: 269–282.
- Shi, Y., B. Ren, J. Wang, and E. Derbyshire. 1986. Quaternary glaciation in China. *Quaternary Science Reviews* 5: 503–507.
- Soltis, D. E., A. B. Morris, J. S. McLachlan, P. S. Manos, and P. S. Soltis. 2006. Comparative phylogeography of unglaciated eastern North America. *Molecular Ecology* 15: 4261–4293.
- Su, Y. J., T. Wang, B. Zheng, Y. Jiang, G. P. Chen, P. Y. Ouyang, and Y. F. Sun. 2005. Genetic differentiation of relict populations of *Alsophila spinulosa* in southern China inferred from cpDNA *trnL-F* noncoding sequences. *Molecular Phylogenetics and Evolution* 34: 323–333.
- Taberlet, P., L. Gelly, G. Pautou, and J. Bouvet. 1991. Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Molecular Biology* 17: 1105–1109.
- Taberlet, P., L. Fumagalli, A. G. Wust-Saucy, and J. F. Cosson. 1998. Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology* 7: 453–464.
- Tamura, K., D. Peterson, N. Peterson, G. Stecher, M. Nei, and S. Kumar. 2011. MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Molecular Biology and Evolution* 28: 2731–2739.
- Tian, B., R. Liu, L. Wang, Q. Qiu, K. Chen, and J. Liu. 2009. Phylogeographic analyses suggest that a deciduous species (*Ostryopsis davidiana* Decne., Betulaceae) survived in northern China during the last glacial maximum. *Journal of Biogeography* 36: 2148–2155.
- Tian, S., L. C. Luo, S. Ge, and Z. Y. Zhang. 2008. Clear genetic structure of *Pinus kwangtungensis* (Pinaceae) revealed by a plastid DNA fragment with a novel minisatellite. *Annals of Botany* 102: 69–78.
- Tsuda, Y. and Y. Ide. 2005. Wide-range analysis of genetic structure of *Betula maximowicziana*, a long-lived pioneer tree species and noble hardwood in the cool temperate zone of Japan. *Molecular Ecology* 14: 3929–3941.
- Tsukada, M. 1988. Japan. Pp. 459–518 in *Vegetation History*, eds. B. Huntley and T. Webb III. Dordrecht: Kluwer Academic Publishers.
- Wall, W. A., N. A. Douglas, Q. Y. J. Xiang, W. A. Hoffmann, T. R. Wentworth, and M. G. Hohmann. 2010. Evidence for range stasis during the latter Pleistocene for the Atlantic Coastal Plain endemic genus, *Pyxidanthra* Michaux. *Molecular Ecology* 19: 4302–4314.
- Wang, F. Y., X. Gong, C. M. Hu, and G. Hao. 2008. Phylogeography of an alpine species *Primula secundiflora* inferred from the chloroplast DNA sequence variation. *Journal of Systematics and Evolution* 46: 13–22.
- Webb, T. and P. Bartlein. 1992. Global changes during the last 3 million years: climatic controls and biotic responses. *Annual Review of Ecology and Systematics* 23: 141–173.
- Wright, S. 1978. *Evolution and the Genetics of Populations*, vol. 4, Variability Within and Among Natural Populations. Chicago: University of Chicago Press.
- AL, Ali River, Inner Mongolia, GKM, 50°33.387'N, 123°32.587'E, 518, 5, 06/23/2011, 58–62; AE Arshaan, Inner Mongolia, GKM, 46°40.410'N, 120°03.100'E, 965, 10, 06/26/2011, 74–83; AK, Aokelidui, Inner Mongolia, GKM, 51°50.307'N, 122°02.222'E, 1,310, 10, 06/24/2011, 84–93; JH, Jinhe, Inner Mongolia, GKM, 51°02.479'N, 121°17.613'E, 1,026, 5, 06/23/2011, 94–98; ZL, Zhilantun, Inner Mongolia, GKM, 48°05.141'N, 122°42.339'E, 332, 10, 05/02/2011, 21–25, 112–116; TP, Taipingling, Inner Mongolia, GKM, 47°29.302'N, 120°29.430'E, 1,595, 10, 06/03/2012, 122–131; GH, Genhe, Inner Mongolia, GKM, 50°46.498'N, 121°29.496'E, 865, 5, 05/07/2012, 132–136; HS, Daheishan, Heilongjiang, LKM, 50°14.468'N, 126°26.905'E, 518, 5, 05/06/2012, 52–56; HH, Wudaohuoluo, Heilongjiang, LKM, 50°14.718'N, 127°31.708'E, 121, 5, 05/28/2012, 117–121; ML, Maolangou, Heilongjiang, LKM, 49°05.893'N, 129°46.257'E, 166, 10, 08/03/2014, 146–155; TW, Tangwanghe, Heilongjiang, LKM, 49°05.591'N, 129°45.527'E, 444, 9, 08/03/2014, 156–164; MS, Maershan, Heilongjiang, CM, 45°15.532'N, 127°26.948'E, 424, 5, 07/22/2011, 38, 40–43; YL, Yilan, Heilongjiang, CM, 46°19.471'N, 129°34.079'E, 119, 9, 04/25/2012, 137–145; LJ, Linjiang, Jilin, CM, 41°53.678'N, 126°45.195'E, 673, 11, 08/02/2011, 26–30, 32–37; CB, CM, Jilin, CM, 42°27.387'N, 128°08.639'E, 682, 2, 07/08/2011, 44–45; WT, Wangtiane, Jilin, CM, 41°32.821'N, 127°56.597'E, 725, 6, 08/03/2011, 46–51; LB, Laobaishan, Jilin, CM, 44°14.547'N, 127°55.155'E, 450, 1, 07/29/2010, 57; LS, Lushuihe, Jilin, CM, 42°32.338'N, 127°53.024'E, 767, 8, 08/04/2011 99–106; SJ, Sanjiaolongwan, Jilin, CM, 42°21.973'N, 126°26.362'E, 785, 5, 07/26/2011, 107–111; KD, Kuandian, Liaoning, LP, 40°30.412'N, 124°29.533'E, 216, 10, 04/26/2011, 1–10; FH, Fenghuangshan, Liaoning, LP, 40°24.790'N, 124°04.686'E, 363, 5, 04/26/2011, 11–14, 16; BY, Bingyugou, Liaoning, LP, 39°59.484'N, 122°57.802'E, 230, 10, 06/12/2011, 63–66, 68–73; WL, Wulongshan, Liaoning, LP, 40°15.303'N, 124°20.130'E, 135, 10, 08/03/2014, 165–174.

APPENDIX 2. Information about the 73 primer pairs tested in this study. Data are in the following order: chloroplast regions, primer names and sequences, product size (bp), annealing temperature (°C), literature source, and result of PCR amplification. Primers in bold were ultimately chosen.

**trnH-psbA**, **trnH**: ACTGCCTTGATCCACTTGGC, **psbA**: CGAAGCTCCATCTACAAATGG, ~400, 47, Hamilton 1999, Amplified; *rpl20-5'* *rps12*, *rpl20*: TTTGTCTACGTCTCCGAGC, 5'-*rps12*: GTCGAGGAACATGTACTAGC, ~1000, 48, Hamilton 1999, Amplified; *psbB-psbF*, *psbB*: GTTACTTTTGGGCATGCTTCG, *psbF*: CGCAGTTCGCTTGACCAG, -, -, Hamilton 1999, Failed to amplify; *trnS-trnG*, **trnS<sup>(GCU)</sup>**: GCCGCTTTAGTCCACTCAGC, **trnG<sup>(UCC)</sup>**: GAACGAATCACACTTTTACCAC, ~750, 48, Hamilton 1999, Amplified; *trnL-ndhF*, **TabE**: GGTTCAAGTCCCTCTATCCC, **ndhJ**: ATGCCYGAAGTTGGATAGG, ~1100, 48, Taberlet et al. 1991, Shaw et al. 1991, Amplified; *trnL<sup>(UAA)</sup>*, **trnL-C**: GGGGATAGAGGGACTTGAAC, **trnLd**: CGAAATCGGTAGACGCTACG, ~600, 48, Taberlet et al. 1991, Amplified; *trnL<sup>(UAA)</sup>-trnF<sup>(GAA)</sup>*, **TrnL-F(f)**: ATTTGAATCGGTGACACGAG, **TrnL-F(c)**: CGAAATCGGTAGACGCTACG, -, -, Taberlet et al. 1991, Failed to amplify; *trnT<sup>(UGU)</sup>-trnL<sup>(UAA)</sup>*, **trnTa**: CATTACAAATGCGATGCTCT, **trnLb**: TCTACCGATTTCCGCTATATC, -, -, Taberlet et al. 1991, Failed to amplify; *trnT<sup>(UGU)</sup>-trnF<sup>(GAA)</sup>*, **trnTf**: CATTACAAATGCGATGCTCT, **trnFr**: ATTTGAATCGGTGACACGAG, -, -, Taberlet et al. 1991, Failed to amplify; *rps16* intron, **rps16R2**: TCGGGATCGAACATCAATTGCAAC, **rps16F**: GTGGTAGAAAGCAACGTGCGACTT, ~900, 48, Oxelman et al. 1997, Amplified; *rbcL*, **rbcLr**: CTCTGCTACAAATAAGAATCGATCTC, **rbcLf**: ATGTACCACAAA CAGAGCTACAAAGC, -, -, Kress and Erickson 2007, Failed to amplify; *atpF-atpH*, **atpF**: ACTCGCACACACTCCCTTTCC, **atpH**: GCTTTTATGGAAGCTTTAACAAT, ~600, 48, Lahaye et al. 2008, Amplified; *psbK-psbI*, **psbK**: TTAGCCTTTGTTTGCAAG, **psbI**: AGAGTTTGAGAGTAAAGCAT, ~500, 47, Lahaye et al. 2008, Amplified; *trnS-trnM*, **trnSf**: GAGAGAGAGGGATTGCAACC, **trnMf**: CATAACCTTGAGGTCA CGGG, ~1100, 48, Demesure et al. 1995, Amplified; *rpoC2*, **rpoC2-f**: CGAGCAGTTTCTTGAAACTCGC, **rpoC2-r**: GAGGATTAATGKRCGATCCWCAAGG, -, -, Grivet et al. 2001, Failed to amplify; *trnK*, **trnK1**: GGGTTGCCCGGGACTCGAAC, **trnK2r**: CAACGGTAGAGTACTCGGCTTTTA, -, -, Grivet et al. 2001, Failed to amplify; *trnK-trnQ*, **trnK2f**: TAAAGCCGAGTACTCTACCGTTG, **trnQr**: CTATTGGAGGTTTCGAA TCCCTGCC, -, -, Grivet et al. 2001, Failed to amplify; *trnQ-trnK*, **trnQf**: GGGACGGAAGGATTGCAACC, **trnS**: ATTGCGTCCAATAGGATTGAA, -, -, Grivet et al. 2001, Failed to amplify; *rpoC1-trnK*, **rpoC1**: GCACAAATTCCTTTTATRGG, **trnCr**: CGACACCCRGATTGAAC TGG, -, -, Grivet et al. 2001, Failed to amplify; *trnC-trnD*, **trnCf**: CCAGTTCAAATCTGGGTGTC, **trnDr**: GGGATTGTAGTTCAATTGGT, -, -, Grivet et al. 2001, Failed to amplify; *psbA-trnS*, **psaAf**: ACTTCTGGTTCGGCGCAACGAA, **trnSr**: AACCACTCGGCCATCTCT

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.



CCTA, -, -, Grivet et al. 2001, Failed to amplify; *trnV-rbcL*, *trnVf*: CGAACCGTAGACCTTCTCGG, *rbcL*: GCTTTAGTCTCTGTTTGTGG, ~500, 44.5, Grivet et al. 2001, Amplified; *orf512-orf184*, *orf512*: AGTATGGGATCCGTAGTMGG, *orf184r*: GGCCYCGGATTTCATAT AAAG, -, -, Grivet et al. 2001, Failed to amplify; *petA-psbE*, *petAf*: GCATCTGTTATTTTGGCACA, *petEr*: TACCTTCCCTATTCATTCGCG, -, -, Grivet et al. 2001, Failed to amplify; *psbB1-psbB2*, *psbB1*: TGCCTT GGTATCGTGTTCATAC, *psbB2r*: CYGTCTTCTTGTAGTTGGAT, ~1000, 48, Grivet et al. 2001, Amplified; *petB-petD*, *petB3f*: CTATCGTCCRACCG TTACWAGGCT, *petDr*: CAAAYGGATAYGCAGGTTACCC, -, -, Grivet et al. 2001, Failed to amplify; *trnS-trnR*, *trnS*: CGCCGCTTTAG TCCACTCA, *trnRr*: ATTGCGTCCAATAGGATTGAA, -, -, Grivet et al. 2001, Failed to amplify; *rps18-clpp*, *rps18f*: GCTCGTATTTTATC TTGTGTACC, *clppr*: AACCTGCTAGTCTTWTAT, -, -, Grivet et al. 2001, Failed to amplify; *rps8-rpl16*, *rps8f*: TGAACAATATTTTCGGTAAT, *rpl16r*: AACCAGATTTCGTAAACAAC, -, -, Grivet et al. 2001, Failed to amplify; *atpH-atpI*, *atpHf*: CCAGCAGCAATAACGGAAGC, *atpIr*: ATA GGTGAATCCATGGAGGG, ~900, 48, Grivet et al. 2001, Amplified; *atpI-rpoC2*, *atpIf*: GATGRCCCTCCATGGATTACCC, *rpoC2r*: GCGAGTTTT CAAGAACTGCTCG, -, -, Grivet et al. 2001, Failed to amplify; *clpp-psbB*, *clppf*: ATAAWAAGAACTAGCAGGTT, *psbBr*: ATAYACCAATGCCAR TAG, -, -, Shaw et al. 2007, Failed to amplify; *rpl23-psbA*, *rpl23f*: GCCCTTTTAACTCAGTGGTA, *psbCr*: GAGCTTGAGAAGCTTCTGCT, -, -, Grivet et al. 2001, Failed to amplify; *psbD-trnT*, *psbD*: CTCCGTARCCAGTCATCCATA, *trnT<sup>(GCU)</sup>*: CCCTTTTAACTCAGTGG TAG, -, -, Shaw et al. 2007, Failed to amplify; *rpl23-psbA*, *rpl23f*: TAAGACAGAAATAAAGCATTGCGTCGAAC, *psbA3'*: CTAGCACTG AAAACCGTCTT, -, -, Grivet et al. 2001, Failed to amplify; *trnD-trnT*, *trnDf*: ACCAATTGAAGTACAATCCC, *trnTr*: CTACCACTGAGTTAAAA GGG, -, -, Grivet et al. 2001, Failed to amplify; *rpl23-psbA*, *rpl23f*: GGTTCGTGACCAAGAAACCAC, *trnSr*: GGTTCAATCCCTCTCTCTC, ~1100, 49, Grivet et al. 2001, Amplified; *trnM-psaA*, *trnMf*: GAA CCCGTGACCTCAAGGTTATG, *psaAr*: ATTCGTTCCCGGAACCAGAGA AGT, -, -, Grivet et al. 2001, Failed to amplify; *rpl23-psbA*, *rpl23f*: CGAGGGTTTCAATCCCTCTC, *trnTr*: AGAGCATCGCATTTGTAATG, -, -, Grivet et al. 2001, Failed to amplify; *trnF-trnV*, *trnFf*: CTCGTGTCACC AGTTCAAAT, *trnVr*: CCGAGAAGGTCTACGGTTCCG, -, -, Grivet et al. 2001, Failed to amplify; *rbcLf-orf106*, *rbcLf*: ATGTCCACAAACAGAGAAA CTAAGCAAGT, *orf106*: ACTACAGATCTCATACTACCCC, -, -, Grivet et al. 2001, Failed to amplify; *orf184-petA*, *orf184f*: GGCCYCGGATTTC ATATAAG, *petAr*: CCCTCKGAAACAAGAAGTT, -, -, Grivet et al. 2001, Failed to amplify; *petA-psbE*, *petAf*: GCATCTGTTATTTTGGCACA, *psbEr*: TACCTTCCCTATTCATTCGCG, -, -, Grivet et al. 2001, Failed to amplify; *psbL-rps18*, *psbL*: GAAAATAAAACAGCAAGTAC, *rps18r*: GGTAAC AAAGATAAAATACGAGC, -, -, Grivet et al. 2001, Failed to amplify; *psbB2-petB3*, *psbB2f*: CAGAAGCTTGGTCTAAAATTCC, *petB3r*: GRITCCC AAGGGAARGAATAACCAGT, -, -, Grivet et al. 2001, Failed to amplify; *petD-rps8*, *petDf*: TCCAGTAGTAATTCTAGGTA, *rps8r*: TAAAATAC TTTTACGAGAAAG, -, -, Grivet et al. 2001, Failed to amplify; *rpoC1-rpoB*, *rpoC1f*: CTGGTAATARRCACAACCATCC, *rpoB*: GTTACACAAC AACCCCTTAGAGG, ~1100, 48, Grivet et al. 2001, Amplified; *trnD<sup>(GUC)</sup>*-*psbM*, *trnD<sup>(GUC)</sup>*RAq: GGGACTGTAGTCAATCGGT, *psbM*-FAq651: ACTCCCTCACTAAGAAAAG, -, -, Shaw et al. 2005, Fior et al. 2013, Failed to amplify; *rps4-trnE<sup>(GAA)</sup>*, *rps4R2*: CTGTSAGACCATAAT GAAAMCG, *trnT<sup>(GCU)</sup>*: AGGTTAGAGCATCGCATTG, -, -, Shaw et al. 2005, Failed to amplify; *trnD<sup>(GUC)</sup>*-*trnT<sup>(GGU)</sup>*, *trnT<sup>(GCU)</sup>*: CTACCACTGAG TTAAGAGGG, *trnE<sup>(UAC)</sup>*: TCCTTGAAAGAGAGATGTCCT, -, -, Shaw et al. 2005, Failed to amplify; *trnS<sup>(UGA)</sup>*-*trnG<sup>(GCC)</sup>*, *trnS<sup>(UGA)</sup>*Aq: GAGAGA

GAGGATACGAACC, *psbZ-AqR*: ACTAGACCAACCATCAGAAGAAG, ~600, 54, Shaw et al. 2005, Fior et al. 2013, Amplified; *rps12-rpl20*, *5'rps12-Aq*: AATAGAAACACAAGACAGCCAAAT, *rpl20-Aq515*: ACGTCTCT CTATACACCG, ~500, 52, Shaw et al. 2005, Fior et al. 2013, Amplified; *ycf6-psbM*, *psbM*-RAq: ATGGAAGTAAATATTGTCCG, *ycf6*-FAq558: TCCATTGATTGATTGTTTCG, -, -, Shaw et al. 2005, Fior et al. 2013, Failed to amplify; *ndhA* intron, *ndhAx1*: GCYCAATCWATTAGTTAT GAAATACC, *ndhAx2*: GGTGACGCCACAGATTCCA, ~1000, 51, Shaw et al. 2007, Amplified; *3'trnV-ndhC*, *trnV<sup>(UAC)</sup>*x2: GTCTACGGTTCGART CCGTA, *ndhC*: TATTATTAGAAATGYCCARAAAATATCATATTC, ~500, 48, Shaw et al. 2007, Amplified; *rpl32-trnL*, *trnL<sup>(UAG)</sup>*: rpl32-F: CAGTT CCAAAAAAACGTACTTC, CTGCTTCCTAAGAGCAGCGT, ~700, 48, Shaw et al. 2007, Amplified; *ndhJ-trnT<sup>(GAA)</sup>*, *ndhJ*: ATGCCYGAAGTT GGATAGG, *TabFR*: CTCGTGTCACCAGTTCAAAT, ~700, 51, Shaw et al. 2007, Amplified; *rpl14-rps8-infA-rpl36*, *rpl14*: AAGGAAATCCAAAAGG AACTCG, *rpl36*: GGRITGGAACAAATTAATAATTCG, -, -, Shaw et al. 2007, Failed to amplify; *psbJ-petA*, *psbJ*: ATAGGTACTGTAR CYGGTATT, *petA*: AACARTTYGARAAGGTTCAATT, -, -, Shaw et al. 2007, Failed to amplify; *ndhF-rpl32*, *ndhF*: GAAAGGTATKATCCAY GMATATT, *rpl32-R*: CCAATATCCCTTYTTTTCCTCAA, -, -, Shaw et al. 2007, Failed to amplify; *ndhC-trnV<sup>(UAC)</sup>*, *ndhC*: TATTATTAGAAA TGYCCARAAAATATCATATTC, *trnV<sup>(UAC)</sup>*x2-Aq605: TCATATACTAAAA CCCGATTGC, -, -, Shaw et al. 2007, Fior et al. 2013, Failed to amplify; *trnK<sup>(UUU)</sup>*-*rps16*, *trnK<sup>(UUU)</sup>*x1: TTAAAGCCGAGTACTCTACC, *rps16x2F2-Aq517*: GGGTGGTTTTTGGTGAGACT, -, -, Shaw et al. 2007, Fior et al. 2013, Failed to amplify; *trnT<sup>(GGU)</sup>*-*psbD*, *trnT<sup>(GGU)</sup>*R: CCCTTTAACT CAGTGGTAG *psbD*-Aq536: CTCTACCGAATTTGAATATC, -, -, Shaw et al. 2007, Fior et al. 2013, Failed to amplify; *psbE-petL*, *psbE*: TATCGAATACTGGTAATAATATCAGC, *petL*-Aq664: ATAGCGCGGA TGCAACAATG, -, -, Shaw et al. 2007, Fior et al. 2013, Failed to amplify; *atpH-atpI*, *atpH*: CCAAYCCAGCAGCAATAAC, *atpI*-Aq551: GTATATGG AATTTAATGCCTAGAAG, -, -, Shaw et al. 2007, Fior et al. 2013, Failed to amplify; *psaI-accD*, *psaI*-75R: AGAAGCCATTGCAATTGCCGGA AAA, *accD*: AATYGTACCACGTAATCYTTAAA, -, -, Shaw et al. 2007, Failed to amplify; *rps16-trnQ<sup>(UUG)</sup>*, *rps16x1*: GTTGCTTTTACCACATCGTTT, *trnQ<sup>(UUG)</sup>*Aq597: TGTCAGTAGATCGGAGTTAAGC, -, -, Shaw et al. 2007, Fior et al. 2013, Failed to amplify; *ndhA* intron, *ndhAx2-Aq*: GGTGACGCCACAGATTCCA, *ndhAx1-Aq597*: GAATCCTTATGGGG CTTTACG, -, -, Shaw et al. 2007, Fior et al. 2013, Failed to amplify; *trnC<sup>(GCA)</sup>*-*ycf6*, *trnC<sup>(GCA)</sup>*FAq520: TACGAGTGCTCAGACATTCA, *ycf6*-RAq: GCCCCGAGCAAGACTTACTATATCCAT, -, -, Shaw et al. 2007, Fior et al. 2013, Failed to amplify; *trnC<sup>(GCA)</sup>*-*rpoB*, *trnC<sup>(GCA)</sup>*RAq: CACAC RGATTTGAACTGGGG, *rpoB*-Aq573: TTGAATGAAAAATGGAAGC, -, -, Shaw et al. 2005, Fior et al. 2013, Failed to amplify; *matK*, *matK900*: CGTTTAATCACTCGAATGTATCA, *matK1470*: AACATGTTGATCGT AAATAA, -, -, Fior et al. 2013, Johnson and Soltis 1994, Failed to amplify; *trnK-matK*, *trnK3914F*: AAGATGTTGATWGTAATGA, *matK1470R*: TGGGTTGCTAACTCAATGG, -, -, Johnson and Soltis 1995, Failed to amplify; *trnK-matK*, *trnK*: AGAAGATTGGTTGCGGAGAA, *matK*: AAAAGAGCGATTGGGTTGAA, ~900, 48, This study, Amplified.

APPENDIX 3.  $K$  value and  $F_{CT}$  value pairs observed in SAMOVA with the  $K$  values varying from 2 to 20.

2, 0.75011; 3, 0.82673; 4, 0.83737; 5, 0.84424; 6, 0.84910; 7, 0.84968; 8, 0.85014; 9, 0.85345; 10, 0.85498; 11, 0.85523; 12, 0.85582; 13, 0.85475; 14, 0.85514; 15, 0.85542; 16, 0.85631; 17, 0.85686; 18, 0.85790; 19, 0.85901; 20, 0.86137.