



# Phylogeography and Population Genetics of *Rosa chinensis* var. *spontanea* and *R. lucidissima* Complex, the Important Ancestor of Modern Roses

Hongying Jian<sup>1\*</sup>, Ling Zhao<sup>2</sup>, Hao Zhang<sup>1</sup>, Changle Ma<sup>2</sup>, Qigang Wang<sup>1</sup>, Huijun Yan<sup>1</sup>, Xianqin Qiu<sup>1</sup>, Ningning Zhou<sup>1</sup> and Ting Zhang<sup>1</sup>

<sup>1</sup> National Engineering Research Center for Ornamental Horticulture/Flower Research Institute, Yunnan Academy of Agricultural Sciences, Kunming, China, <sup>2</sup> School of Landscape Architecture and Horticulture Science, Southwest Forestry University, Kunming, China

## OPEN ACCESS

### Edited by:

Jordi López-Pujol,  
Spanish National Research Council  
(CSIC), Spain

### Reviewed by:

Yong Qi Zheng,  
Chinese Academy of Forestry, China  
Huaizhen Tian,  
East China Normal University, China

### \*Correspondence:

Hongying Jian  
ynwildflower@aliyun.com

### Specialty section:

This article was submitted to  
Plant Systematics and Evolution,  
a section of the journal  
Frontiers in Plant Science

Received: 09 January 2022

Accepted: 14 April 2022

Published: 20 May 2022

### Citation:

Jian HY, Zhao L, Zhang H, Ma CL,  
Wang QG, Yan HJ, Qiu XQ, Zhou NN  
and Zhang T (2022) Phylogeography  
and Population Genetics of *Rosa*  
*chinensis* var. *spontanea* and *R.*  
*lucidissima* Complex, the Important  
Ancestor of Modern Roses.  
*Front. Plant Sci.* 13:851396.  
doi: 10.3389/fpls.2022.851396

*Rosa chinensis* var. *spontanea* and *R. lucidissima* complex are the morphologically very similar key ancestors of modern roses with high importance in rose research and breeding. Although widely distributed in subtropical central and southwestern China, these two taxa are highly endangered. We sampled a total of 221 specimens and 330 DNA samples from 25 populations across the two taxa's whole range. Leaf morphological traits were compared. Two chloroplast DNA intergenic spacers (*trnG-trnS*, *petL-psbE*) and ITS were used for population genetics and phylogenetic study to delimit the boundary between the two taxa, assess the genetic variation, uncover the possible evolutionary mechanism responsible for the differentiation within the complex, and make the conservation recommendations. The complex exhibited high levels of genetic variation ( $h_T$  cpDNA = 0.768,  $h_T$  ITS = 0.726) and high population differentiation even over small geographic distance. We suggest *R. chinensis* var. *spontanea* and *R. lucidissima* be treated as independent taxa, and the northern populations around and within the Sichuan Basin being *R. chinensis* var. *spontanea*, having broader leaflets and paler full-blooming flowers, while those in the middle and southern Yunnan-Guizhou Plateau and the adjacent regions being *R. lucidissima*, having narrower leaflets and darker full-blooming flowers. Transitional areas between the southeastern Sichuan Basin and northeastern Guizhou are the contact or the hybridization zone of the two taxa. Ancestral haplotypes of the complex (*R. lucidissima*) evolved at about 1.21–0.86 Mya in southeastern Yunnan-Guizhou Plateau and its adjacent regions and survived there during the Quaternary Oscillation. Ancestral haplotypes of *R. chinensis* var. *spontanea* deviated from *R. lucidissima* at about 0.022–0.031 Mya at the transitional areas (Daloushan and Wulingshan Mountains) between the northeastern edge of Yunnan-Guizhou Plateau and the southeastern border of Sichuan Basin, where they survived the LGM. The evolution of the complex included spatial isolation and inter-species hybridization. The complex's endangered status might be the result of over-exploitation for its ornamental and medical value, or due to reforestation of some originally open habitats. We provide specific recommendations for the two taxa's *in situ* and *ex situ* conservation.

**Keywords:** conservation, genetic variation, population history, refugia, evolutionary mechanism, species boundary

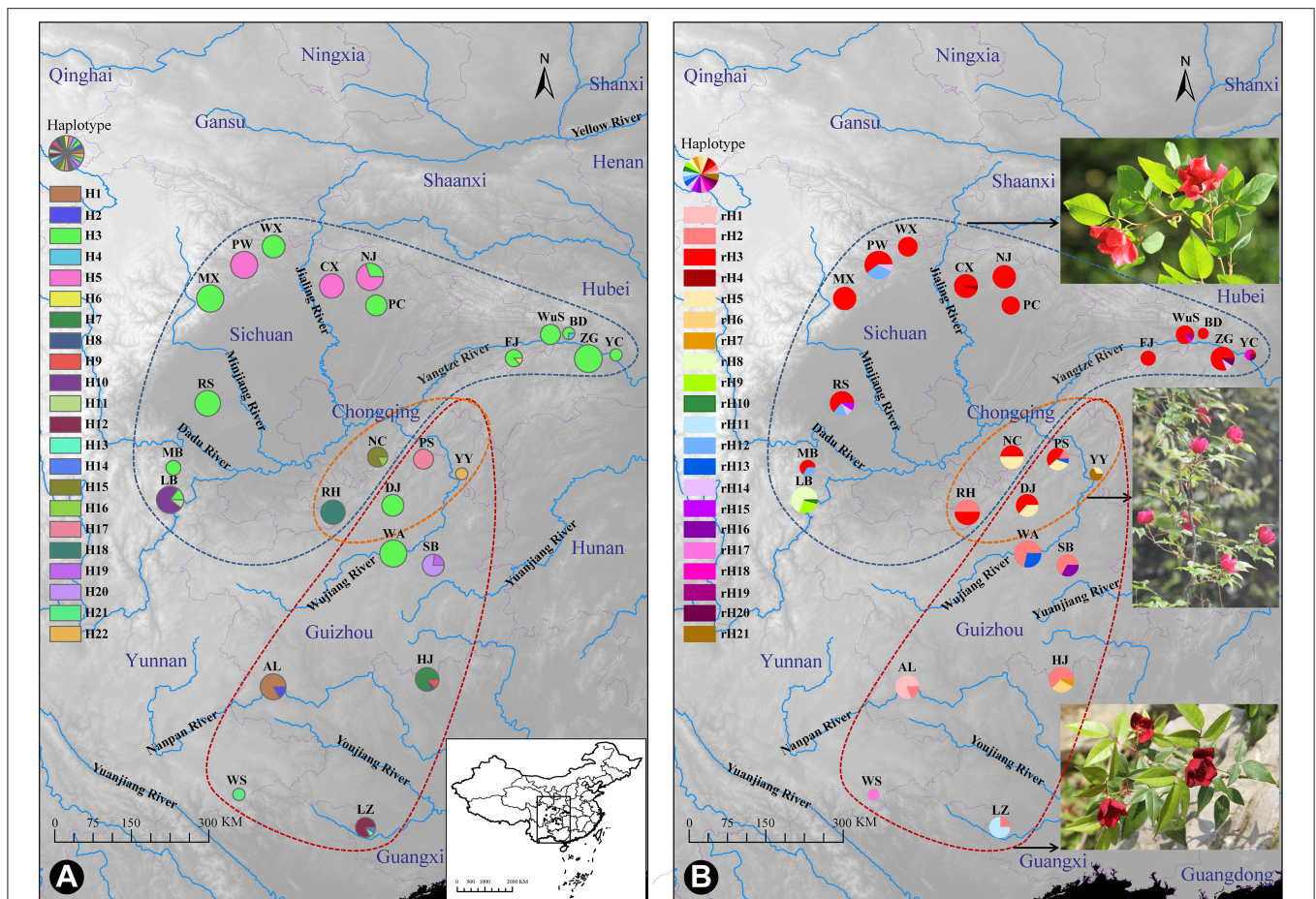
## INTRODUCTION

Phylogeographical studies utilizing genetic markers having phylogenetic signals such as cpDNAs and nrITS and analyzing the spatial distribution of haplotypes across species or species complex ranges, allow for verifying taxonomy at various taxonomic levels. Besides, such studies can unravel historical processes that led to the formation of species or affected their distribution (e.g., post-glacial migrations) and identify potential glacial refugia (Avice, 2000, 2004; Thaw et al., 2019). Refugia, i.e., locations where species have persisted during long-term climatic changes due to their eco-climatic stability, have high conservation value, so it is very important to identify precise refugial locations (Médail and Diadema, 2009; Hampe et al., 2013; Volis, 2019). In China, the area within  $\sim 34^{\circ}\text{N}$  to  $22^{\circ}\text{N}$  (Wu, 1980; Qiu et al., 2011) is considered a Pleistocene refugium for many ancient and relict species. However, within this area, multiple isolated refugia have been revealed, mostly in mountain regions (see Zhang et al., 2016; Deng et al., 2019 and references therein). Among them, the Sichuan Basin, which is part of the boundary between the Sino-Japanese Forest subkingdom and the Sino-Himalayan Forest subkingdom (Wu and Wu, 1996), played an important role in the formation and evolution of many plant lineages (Gao et al., 2007; Mitsui et al., 2008; Qiu et al., 2009; Guan et al., 2010; Ma et al., 2015).

Species are a fundamental unit of biology, but discerning species boundaries among closely related taxa is a difficult task. The genus *Rosa* L. (Rosaceae) comprises about 150–200 species widely distributed in the temperate and subtropical northern hemisphere, and about half of the rose species occur in Asia, especially in China (Ku, 1985; Ku and Robertson, 2003). Species of this genus are difficult to identify due to similar morphology because of naturally occurring hybridization (Fougère-Danezan et al., 2014) and polyploidization (De Riek et al., 2013). Chinese roses, especially natural or cultivated varieties of both *R. chinensis* Jacq. and *R. odorata* (Andr.) Sweet have been contributing characteristic of recurrent flowering, tea scent, and multiple floral colors to modern roses (Hurst, 1941; Wylie, 1954; Scalliet et al., 2008; Raymond et al., 2018). The wild type of *R. chinensis*, *R. chinensis* var. *spontanea* (Rehder et al. Wilson) Yü et Ku (Ku, 1985; Ku and Robertson, 2003) from sect. *Chinenses*, is the maternal parent of *R. chinensis* and the possible paternal parent of *R. odorata* (Meng et al., 2011). Thus, *R. chinensis* var. *spontanea* is one of the key ancestors of modern roses. *Rosa lucidissima* Lévl., also from sect. *Chinenses*, are so similar in morphology to *R. chinensis* var. *spontanea* that it was thought to be the state of *R. chinensis* var. *spontanea* by both Henry and Wilson (Rix, 2005). Both taxa are diploid (Akasaka et al., 2002; Cao et al., 2021), have shiny deep green leaves of 5–11 cm with 3–5 leaflets, solitary single flowers, and ovoid, obovoid, or pyriform glabrous hips fruiting from May to September. *Rosa chinensis* var. *spontanea* is an “erect shrub”. Its rachis and petiole are sparsely prickly and glandular-pubescent, leaflets are “broadly ovate or ovate-oblong” and flowers are red. By contrast, *R. lucidissima* is a “climbing shrub” having rachis and petiole that are shortly prickly and sparsely glandular-pubescent, leaflets are

“oblong-ovate or long elliptic” and pale green abaxially, and flowers are purple-red (Ku, 1985; Ku and Robertson, 2003; **Figure 1**). However, it is difficult to tell these two taxa apart in the wild, and it is common when specimens with the same collection ID are identified in one herbarium as *R. chinensis* var. *spontanea* and in another as *R. lucidissima*. In the field investigations, all the plants (and even in the place where the type specimen of *R. chinensis* var. *spontanea* was collected) are climbing shrubs. Plants flower from early March to early May with extraordinary petal color variation (Rix, 2005 and personal observations in the field). During flower opening, the petal color is changing from pink to purple (**Figures 2A–D**), and in populations in northern Sichuan, e.g., Pingwu county and Maoxian county, the flower colors of different plants varies from creamy white to dark crimson (**Figures 2E–H**). It is still not clear how much of this variation is due to phenotypic plasticity and how much is the result of genetic differentiation. Population level studies seem more effective for tracking evolutionary history and circumscribing close species (Joly et al., 2006; Liu et al., 2012; Sun et al., 2014). We treat *R. chinensis* var. *spontanea* and *R. lucidissima* as a complex and sampled according to populations for further study.

According to the specimen records, *R. chinensis* var. *spontanea* and *R. lucidissima* occur in subtropical central and Southwestern China, mainly on the borders of the Sichuan Basin, further north to the south of Gansu, further south to the east of Yunnan-Guizhou Plateau, Jiuwan Mountains and Shiwan Mountains in Guanxi Zhuang Autonomous Region and south of Nanling Mountains in Guangdong provinces (**Figure 1**). Plants have not only decorative but also medicinal value (Ye, 2015; Zhao et al., 2019) and demands for their roots are increasing with the development of public transportation and information allowing horticultural and pharmaceutical companies to reach the remote rural areas to buy the plants either for gardening or for medicine. Not surprisingly, the populations of these taxa are rapidly decreasing in size or completely disappearing. Meng et al. (2011) failed to collect any plants during their field expedition to the specimen-recorded locations. Only 14 individuals were found in Beichuan and Pingwu counties in NW Sichuan by Cui et al. (2016). *Rosa chinensis* var. *spontanea* and *R. lucidissima* have been listed as “endangered” and “critically endangered”, respectively, in a recent report on biodiversity (Qin et al., 2017). In the most recently published List of National Key Protected Wild Plants (<http://www.forestry.gov.cn/main/3954/20210908/163949170374051.html>), both taxa are evaluated as Rank II. The high importance of the above two taxa in rose research and breeding, and conservation concerns necessitate understanding the extent and structure of their genetic diversity. In our study, we sampled *R. chinensis* var. *spontanea* and *R. lucidissima* across their whole range, and used two chloroplast DNA regions and ITS to (1) delimit the boundary between the two taxa and uncover the possible evolutionary mechanism responsible for the differentiation within the complex; (2) analyze the extent and structure of the genetic variation of the complex; (3) identify the possible refugia during the glaciation period



**FIGURE 1 |** Distribution of cpDNA haplotypes **(A)** and ITS ribotypes **(B)** of *Rosa chinensis* var. *spontanea* (area encircled by the blue curve) and *R. lucidissima* (area encircled by the red curve). Pie charts show the frequency of haplotypes/ribotypes in each population. The approximate location of the complex's distribution was shown in **(A)**. The orange curve shows the transitional zones of the two taxa. The photos of each taxa and the plant in the transitional zones were shown close to their geographical distribution, respectively, in **(B)**.

and the population history of the complex; and (4) make the conservation recommendations.

## MATERIALS AND METHODS

### Plant Material Sampling

A total of 330 individuals (4 to 21 per population depending on the population size) were sampled from 25 populations, covering the whole geographical distribution of the complex (**Table 1**). Fresh young leaves were collected and dried in silica gel during field expeditions. A total of 221 specimen (1–20 per population) were collected for morphological trait study. The latitude, longitude, and altitude of each population were recorded. *Rosa rugosa* (sect. *Cinnamomeae*), *R. woodsii* (sect. *Carolinae*), *R. mairei* (sect. *Pimpinellifoliae*), *R. banksiae* (sect. *Banksianae*), and *R. bracteata* (sect. *Bracteatae*) were used as outgroups in the phylogenetic analysis. Fresh leaves of these species were collected from the Yunnan Provincial Flower Gemplasm Resources Nursery in the Flower

Research Institute, Yunnan Academy of Agricultural Sciences in Kunming, China.

### Molecular Techniques

Total genomic DNA was isolated from the leaves using Plant Axyprep Genomic DNA Kit (Axygen, CA, USA) with some modification to decrease the negative effect of secondary metabolites and epicuticular waxes. From a total of 10 cpDNA regions (*petL/psbE*, *trnL(UAG)/rpl32-F*, *trnH/psbA*, *trnDF/trnT*, *trnS/trnfM*, *rp120/rps12*, cpDNA-c/cpDNA-f, F71/R1516, *trnG(UUC)-trnS(GCU)*, *trnS/psbC*), we chose *trnG(UUC)-trnS(GCU)* and *petL-psbE* due to their relatively high variation and ease for amplification. The *trnG(UUC)-trnS(GCU)* (hereafter referred to as *trnG-trnS*) was amplified following the protocol of Shaw et al. (2007). The *petL-psbE* was amplified following that of Shaw et al. (2005). The internal transcribed spacer (ITS) was amplified using the two primers ITS1 and ITS4. The primer sequences of these three fragments were shown in **Table 2**. Then, the PCR products were purified and sequenced by TsingKe



**FIGURE 2 |** Change of petal color different opening stage and morphological variation within population MX. **(A)** Initial blooming stage. **(B)** Full-blooming stage. **(C)** Post-blooming stage. **(D)** Wither stage. **(E)** Individuals with pink flowers. **(F)** Individuals with crimson flowers. **(G)** Individuals with creamy white flowers. **(H)** Individuals with pinkish-white flowers.

**TABLE 1** | Details of the sample locations, sample sizes, haplotype frequencies, haplotype and nucleotide diversity based on cpDNA fragments and ITS.

Pop. ID	Province	LON*	LAT*	ALT (m)	Taxa#	Individuals	cpDNA ( <i>trnS-trnG</i> , <i>petL-psbE</i> )			ITS			
							Haplotype frequencies	Haplotype diversity, HD (SD)	Nucleotide diversity Pi (SD) × 10 <sup>-3</sup>	Number of sequence	Haplotype frequencies	Haplotype diversity, Hd (SD)	Nucleotide diversity, Pi (SD) × 10 <sup>-3</sup>
BD	Hubei	110.4	31.2	608	<i>R.C.S</i>	4	H3 (3), H4 (1)	0.50 (0.27)	0.20 (0.11)	4	rH3 (4)	0 (0)	0(0)
CX	Sichuan	106.3	32.0	494	<i>R.C.S</i>	16	H5(16)	0 (0)	0	21	rH3 (20), rH4 (1)	0.10 (0.08)	0.89 (0.79)
FJ	Chongqing	109.5	30.8	370	<i>R.C.S</i>	8	H3 (7), H6 (1)	0.25(0.18)	0.10 (0.07)	8	rH3 (8)	0 (0)	0 (0)
LB	Sichuan	103.4	28.3	1163	<i>R.C.S</i>	20	H3 (3), H10 (15), H11 (2)	0.43 (0.12)	0.19 (0.06)	26	rH8 (18), rH9 (7), rH10 (1)	0.47 (0.09)	1.05 (0.38)
MB	Sichuan	103.5	28.8	1046	<i>R.C.S</i>	6	H3 (6)	0 (0)	0	9	rH3 (6), rH12 (3)	0.50 (0.13)	5.46 (1.40)
MX	Sichuan	104.1	31.8	1134	<i>R.C.S</i>	20	H3 (20)	0 (0)	0	21	rH3 (20)	0	0
NC	Chongqing	107.1	29.0	300	<i>R.C.S</i>	11	H15 (9), H16 (2)	0.33 (0.15)	0.26 (0.12)	20	rH3 (10), rH5 (10)	0.53 (0.04)	5.75 (0.4)
NJ	Sichuan	106.9	22.5	585	<i>R.C.S</i>	20	H3 (6), H5 (14)	0.44 (0.09)	0.72 (0.14)	20	rH3 (20)	0	0
PC	Sichuan	107.0	31.7	390	<i>R.C.S</i>	12	H3 (12)	0 (0)	0	12	rH3 (12)	0	0
PW	Sichuan	104.7	32.4	850	<i>R.C.S</i>	20	H5 (20)	0 (0)	0	30	rH3 (18), rH12 (10), rH14 (2)	0.54 (0.06)	5.62 (0.51)
RH	Guizhou	106.3	28.1	591	<i>R.C.S</i>	16	H18 (16)	0 (0)	0	24	rH2 (12), rH3 (12)	0.52 (0.03)	6.51 (0.38)
RS	Chongqing	104.1	30.0	578	<i>R.C.S</i>	18	H3 (18)	0 (0)	0	22	rH3 (14), rH12 (4), rH14 (2), rH15 (2)	0.57 (0.11)	5.46 (0.79)
WuS	Chongqing	110.1	31.2	740	<i>R.C.S</i>	11	H3 (11)	0 (0)	0	12	rH3 (10), rH18 (1), rH19 (1)	0.32 (0.16)	0.52 (0.28)
WX	Gansu	105.2	32.7	930	<i>R.C.S</i>	14	H3 (14)	0 (0)	0	14	rH3 (14)	0	0
YC	Hubei	111.2	30.8	300	<i>R.C.S</i>	4	H3 (4)	0 (0)	0	4	rH3 (1), rH18 (2), rH20 (1)	0.83 (0.22)	1.56 (0.53)
ZG	Hubei	110.8	30.8	650	<i>R.C.S</i>	21	H3 (21)	0 (0)	0	21	rH3 (17), rH14 (2), rH20 (2)	0.34 (0.12)	2.7 (1.23)
						221		0.62 (0.03)	0.92 (0.06)	267		0.50 (0.04)	5.95 (0.45)
AL	Guizhou	105.2	25.0	1262	<i>R.L</i>	18	H1 (15), H2 (3)	0.29 (0.12)	0.12 (0.05)	21	rH1 (17), rH2 (4)	0.32 (0.11)	0.5 (0.17)
DJ	Guizhou	107.3	28.2	740	<i>R.L</i>	13	H3 (13)	0 (0)	0	19	rH3 (12), rH5 (7)	0.49 (0.07)	5.36 (0.75)
HJ	Guangxi	107.9	25.1	865	<i>R.L</i>	15	H7 (12), H8 (1), H9 (2)	0.36 (0.15)	0.86 (0.36)	24	rH2 (15), rH6(7), rH7(2)	0.54 (0.08)	10.91 (1.3)
LZ	Guangxi	106.9	22.5	423	<i>R.L</i>	11	H12 (9), H13 (1), H14 (1)	0.35 (0.17)	1.07 (0.63)	16	rH2 (4), rH11 (12)	0.40 (0.11)	0.62 (0.18)

(Continued)

TABLE 1 | Continued

Pop. ID	Province	LON*	LAT*	ALT (m)	Taxa#	Individuals	cpDNA ( <i>trnS-trnG</i> , <i>petL-psbE</i> )			ITS			
							Haplotype frequencies	Haplotype diversity, HD (SD)	Nucleotide diversity Pi (SD) × 10 <sup>-3</sup>	Number of sequence	Haplotype frequencies	Haplotype diversity, Hd (SD)	Nucleotide diversity, Pi (SD) × 10 <sup>-3</sup>
PS	Chongqing	107.9	29.0	756	<i>R.L</i>	11	H17 (11)	0 (0)	0	18	rH2 (3), rH3 (8), rH5 (6), rH13 (1)	0.70 (0.07)	6.35 (0.48)
SB	Guizhou	108.0	27.1	988	<i>R.L</i>	13	H19 (3), H20 (10)	0.39 (0.13)	0.16 (0.05)	18	rH2 (12), rH16 (6)	0.47 (0.08)	2.93 (0.51)
WA	Guizhou	107.3	27.3	868	<i>R.L</i>	20	H3 (20)	0 (0)	0	28	rH2 (20), rH13 (8)	0.42 (0.08)	4.62 (0.83)
WS	Yunnan	104.6	23.1	892	<i>R.L</i>	4	H21 (4)	0 (0)	0	4	rH17 (4)	0	0
YY	Chongqing	108.5	28.7	745	<i>R.L</i>	4	H22 (4)	0 (0)	0	6	rH5 (2), rH21 (4)	0.53 (0.17)	4.99 (1.61)
						109		0.89 (0.02)	1.68 (0.1)	154		0.81 (0.02)	7.39 (0.59)
Total						330		0.74 (0.02)	1.95 (0.11)	422		0.72 (0.02)	8.15 (0.25)

#Nominated according to both the specimen records and morphological characters. *R.C.S.*, *R. chinensis* var. *spontanea*; *R.L.*, *R. lucidissima*. \*Only with 1 decimal due to conservation concerns.

TABLE 2 | Primers for the amplification of selected chloroplast intergeneric spacers and ITS.

Fragment	Primer	Primer sequence	References
<i>trnG-trnS</i>	<i>trnG</i>	5'-GAATCGAACCCGCATCGTTAG-3'	Wakasugi et al., 1994
	<i>trnS</i>	5'-AACTCGTACAACGGATTAGCAATC-3'	
<i>petL-psbE</i>	<i>petL</i>	5'-AGTAGAAAACCGAAATAACTAGTTA-3'	Shaw et al., 2005
	<i>psbE</i>	5'-TATCGAATACTGGTAATAATATCAGC-3'	
ITS	ITS1	5'-TCCGTAGGTGAACCTGCGG-3'	White et al., 1990
	ITS4	5'-TCCTCCGCTTATTGATATGC-3'	

Biological Technology, Beijing, China. For sequence variants appearing in only one or two individuals, amplification and sequencing were repeated to avoid artifacts caused by PCR or sequencing errors. As for those putative hybrids which showed double-peak nucleotide signals in ITS by direct sequencing, the PCR products of ITS were purified with a Gel Extraction Kit (Tsingke, Beijing, China) and then cloned. Cloning of the ITS region was conducted using the pClone 007 versatile simple vector system kit and Trelief<sup>TM</sup> 5 $\alpha$  chemically competent cell (TsingKe, Beijing, China) following the manufacturer's instructions. Plasmids carrying the PCR fragment were then extracted and sequenced. For each PCR product, at least 2 clones were randomly sequenced.

## Leaf Morphological Traits Measurement

Because some specimens did not have blooming flowers, we focused on the leaflet shapes of both taxa. We randomly selected three mature odd-pinnately compound leaves in each specimen for study. Leaflet numbers per leaf were counted. The length

and the width of basal leaflets, middle leaflets (only those with more than 3 leaflets), and top leaflets of each compound leaf were measured with a caliper.

## Data Analysis

Sequences of ITS, *trnG-trnS*, and *petL-psbE* were assembled with DNASTar (Gene Codes Cooperation, USA), aligned with Clustal X version 1.81 (igbmc u-strasbg, France; Thompson et al., 1997), and manually adjusted in BioEdit version 7.0.4.1 (Tom Hall, Ibis Therapeutics, USA; Hall, 1999). Indels in the cpDNA dataset were treated as single mutation and coded as substitution (A or T). Diversity and differentiation parameters (within-population diversity,  $h_S$ ; total diversity,  $h_T$ ; differentiation for unordered and ordered alleles,  $G_{ST}$  and  $N_{ST}$ , respectively) were calculated with HAPLONST (Pons and Petit, 1996). This program was also used to test whether  $N_{ST}$  is larger than  $G_{ST}$  (Pons and Petit, 1996). The significance of the difference between  $N_{ST}$  and  $G_{ST}$  was assessed with one thousand random permutations following Burban et al. (1999). An analysis of molecular variance (AMOVA) was performed using Arlequin version 3 (CMPG, University of Berne, Switzerland; Excoffier et al., 2005). The 95% CIs for  $F_{ST}$  values were obtained using 10,000 permutations. Gene flow among populations was calculated by the formula:  $Nm = 0.25 \times (1 - G_{ST})/G_{ST}$  (Wright, 1931).

Spatial genetic structure was analyzed with SAMOVA 1.0 (Dupanloup et al., 2002). The  $F_{CT}$  index of genetic differentiation among K groups was computed to obtain the best configuration of groups. In this study, the tested K values ranged from 2 to 20 for chloroplast haplotypes and 2–24 for ITS haplotypes, with each simulation starting from 100 random initial conditions and 1,000 permutations. Tajima's D (Tajima, 1989), Fu and Li's F (Fu, 1997) statistics, and mismatch distribution analysis were also calculated

using Arlequin version 3 (Excoffier et al., 2005) with 1,000 parametric bootstrap replicates to test for selective neutrality and for evidence of range expansion. Observed and inferred haplotypes were nested according to Posada and Crandall (2001). Network reconstruction and Nested Clade Phylogeographic Analysis (NCPA) were carried out using ANeCA software (University of Vigo, Spain; Panchal, 2007). The closed loops were resolved using the rules and predictions based on coalescence theory (Crandall and Templeton, 1993; Templeton and Sing, 1993). Finally, a list of inference keys for the origin of the nested clades was generated.

Phylogenetic relationships among the haplotypes of the complex and the outgroups were reconstructed using both maximum likelihood (ML) and Bayesian inference (BI). Prior to phylogenetic analyses, sequences were aligned using MAFFT (Osaka University, Japan; Katoh and Standley, 2013). Based on the AIC calculated from the MrModetest 2.3 (Uppsala University, Sweden; Nylander, 2004), the GTR+I was the best-fitting substitution model for both the ML and BI analyses. The ML analysis was performed using RAxML 8.2.11 (Heidelberg Institute for Theoretical Studies, Germany) with 1,000 bootstrap replicates (Stamatakis, 2014). The BI analysis was conducted using MrBayes 3.2.6 (University of Rochester, USA and Uppsala University, Sweden; Huelsenbeck and Ronquist, 2001). The Markov chain Monte Carlo algorithm was run for 1,000,000 generations with four incrementally heated chains, starting from random trees and sampling every 1,000 generations. The first 25% of the trees were discarded as burn-in and the remaining trees were used to construct a 50% majority-rule consensus tree. Internodes with posterior probabilities > 70% were considered statistically significant. A log-likelihood ratio test (LRT) was first performed using MEGA 10.2.2 (Tokyo Metropolitan University, Japan; Kumar et al., 2018) to estimate divergence times and to test the existence of a constant substitution rate among the cpDNA haplotypes by comparing the log L scores of ML trees with or without a molecular clock enforced, following the test statistic:  $-2(\log L_{\text{clock}} - \log L_{\text{noclock}})$ , which should be distributed as  $\chi^2$  with  $(N - 2)$  degrees of freedom ( $df$ ), where  $N$  is the number of sequences in the tree (Posada and Crandall, 1998). Then, as no fossils are available to calibrate the intergenic spacers substitution rate for genus *Rosa*, branch lengths of the clock-constrained ML tree of cpDNA were transformed into absolute time by assuming the substitution rates of these spacers to be  $(1.2-1.7) \times 10^{-9}$  substitutions per site per year (s/s/y) (Graur and Li, 2000).

For leaf morphology analysis, the ratio of length to width of a leaflet was calculated to indicate its shape. One-way ANOVA was used to test for statistical differences between the two taxa. Statistical analyses were done using Excel 2007 (Microsoft Corporation, USA).

## Population Prioritization for Conservation

In order to rank each population by its conservation priority, we used Contrib ver.1.02 (Petit et al., 1998) to estimate the contribution of each population to total diversity ( $CT$ ) and to total allelic differentiation ( $CrT$ ). The data of cpDNA haplotypes and their distribution among populations were used as input data for this analysis. Negative values indicate that the diversity or

the differentiation of a population is lower than the mean of the whole dataset (Petit et al., 1998; Swatdipong et al., 2009). The higher population of  $CT$  or  $CrT$  is, the higher the conservation priority this population has.

## RESULTS

### Leaf Morphology Comparison

The number and size of the leaflets of *R. chinensis* var. *spontanea* and *R. lucidissima* were shown in **Table 3**. About 56.2% individuals of *R. chinensis* var. *spontanea* has 5 leaflets per compound leaf while 67.9% individuals of *R. lucidissima* has only 3 leaflets. The top leaflets of both taxa are larger than the middle and the basal ones. Averagely, the leaflets at the same position of *R. lucidissima* are longer and narrower than those of *R. chinensis* var. *spontanea*, despite that the difference is not significant. However, the ratio of length to width of all the leaflets of *R. chinensis* var. *spontanea* is significantly less than that of *R. lucidissima*, indicating the leaflets of *R. chinensis* var. *spontanea* are much broader than those of *R. lucidissima*.

### Haplotypes, Genetic Diversity, and Structure Based on CpDNA Fragments

The aligned sequences of *trnG-trnS* and *petL-psbE* of the *R. chinensis* var. *spontanea* and *R. lucidissima* complex (hereafter *R. chinensis* var. *spontanea* - *R. lucidissima* complex) was 1,359 and 1,213 bp in length, respectively. A total of 25 polymorphic sites and 10 indels were detected in the concatenated dataset. In total, there were 22 haplotypes (H1-H22) and their GenBank accession numbers were shown (**Supplementary Table S1**). *Rosa chinensis* var. *spontanea* had 9 haplotypes and *R. lucidissima* had 14 haplotypes. The only haplotype shared by both taxa was H3.

The spatial distribution and relative frequencies of the haplotypes are shown in **Figure 1A** and **Table 1**. Haplotypes H4, H6, H8, H13, and H14 all were represented by a single extant individual. The haplotype H6 was found in *R. chinensis* var. *spontanea* and the others in *R. lucidissima*. Haplotype H3 (158/330 individuals) was present in 12 out of the 16 *R. chinensis* var. *spontanea* populations and in 2 out of the 9 *R. lucidissima* populations. All the 14 populations with H3 were on the boundary of the Sichuan Basin, such as WuS and YC, etc. east in the Three Gorges Mountain Region (TGMR), NJ and PC north in the southern Qingling-Dabashan Mountains, WX and MX northwest in the Longmenshan Mountains, MB and LB southwest in the Xiaoliangling Mountains. Although not detected in the south edge of Sichuan Basin, H3 existed more south to DJ and WA where the south edge of Sichuan Basin transited to the northeastern Yunnan-Guizhou Plateau. Haplotype H5 was only found in northern populations of *R. chinensis* var. *spontanea* (PW, CX, and NJ). Except for H3 and H5, the other seven haplotypes were all “private” to a single population of *R. chinensis* var. *spontanea*. Among the 14 haplotypes detected in *R. lucidissima*, except for the shared haplotype H3, all others were private ones.

**TABLE 3** | Comparison of compound leaves between *Rosa chinensis* var. *spontanea* and *Rosa lucidissima*.

Taxa	Leaflet number	Top leaflet (cm)			Middle leaflet (cm)			Basal leaflet (cm)		
		Length	Width	Length/width	Length	Width	Length/width	Length	Width	Length/width
<i>R. chinensis</i> var. <i>spontanea</i>	4.64 ± 0.65	4.96 ± 0.83	2.62 ± 0.61	1.93 ± 0.17	3.89 ± 0.46	2.04 ± 0.36	1.95 ± 0.23	3.38 ± 0.71	1.85 ± 0.45	1.85 ± 0.16
<i>R. lucidissima</i>	3.52 ± 0.72	5.19 ± 0.97	2.30 ± 0.39	2.27 ± 0.29	4.62 ± 0.37	2.08 ± 0.56	2.24 ± 0.12	3.65 ± 0.75	1.73 ± 0.38	2.16 ± 0.25
Significance (P)	0.050	0.549	0.219	0.0015**	0.106	0.898	0.035**	0.414	0.509	0.006**

Data were Means ± SE, \*\*means significant difference between the two taxa.

Three populations had three haplotypes. Most individuals in population HJ were haplotype H7, two individuals were haplotype H2 and the rest had haplotype H8. Besides the widespread haplotype H3, population LB also had two more haplotypes. Most individuals of population LZ in southwestern Guangxi were haplotype H12, and the other two individuals belonged to H13 and H14, respectively. Five populations had two haplotypes. Populations BD and FJ in the TGMR, and population NJ in the Qingling-Dabashan Mountains, had another private haplotype (H4, H6, and H5, respectively) besides H3. Most individuals in population AL had haplotype H1, while three individuals were H2. Most individuals in population NC were haplotype H9 and two individuals were H16. The other populations had only one haplotype.

The chloroplast genetic diversity ( $h_T$ ) of the complex was  $0.768 \pm 0.085$ , that of *R. chinensis* var. *spontanea* was  $0.596 \pm 0.12$  and that of *R. lucidissima* was  $0.972 \pm 0.04$ . The genetic diversity within the population ( $h_S$ ) of the complex was  $0.133 \pm 0.037$ , that of *chinensis* var. *spontanea* was  $0.122 \pm 0.048$  and that of *R. lucidissima* was  $0.154 \pm 0.061$ . The total haplotype diversity ( $H_d$ ) and the nucleotide ( $P_i$ ,  $\pi$ ) of the *R. chinensis* var. *spontanea* - *lucidissima* complex and each taxa are shown in **Table 1**. For the complex, AMOVA revealed that most of the variation (92.29%) is among populations (**Supplementary Table S3**). In *R. chinensis* var. *spontanea*, the among population component was 89.54%, and in *R. lucidissima*, it was 92.44%. The  $N_m$  calculated for the complex was 0.053, and for *R. chinensis* var. *spontanea* and *R. lucidissima* it was 0.064 and 0.047, respectively, meaning extremely low gene flow among populations. The genetic diversity analysis revealed that genetic differentiation among populations of the complex ( $N_{ST} = 0.93 \pm 0.02$ ,  $G_{ST} = 0.83 \pm 0.04$ ), *R. chinensis* var. *spontanea* ( $N_{ST} = 0.89 \pm 0.06$ ,  $G_{ST} = 0.8 \pm 0.08$ ) and *R. lucidissima* ( $N_{ST} = 0.93 \pm 0.03$ ,  $G_{ST} = 0.84 \pm 0.06$ ) all was high. The comparison of  $N_{ST}$  and  $G_{ST}$  indicated a lack of phylogeographic structure either in the complex ( $p = 0.245$ ), *R. chinensis* var. *spontanea* ( $p = 0.152$ ), or in *R. lucidissima* ( $p = 0.999$ ).

The SAMOVA revealed, especially when the two taxa were analyzed separately, a strong structure of genetic variation. For the complex,  $F_{CT}$  varied irregularly making it impossible detection of the optimal K value (**Supplementary Figure S1A**). For *R. chinensis* var. *spontanea*, the optimal  $F_{CT}$  was found with  $K=5$ , indicating populations could be divided into 5 groups.

**TABLE 4** | The results of Tajima's test and Fu and Li's test of *R. chinensis* var. *spontanea*-*lucidissima* complex and each taxon.

Taxa	Tajima's test		Fu and Li's test
	D	D*	F*
<i>R. chinensis</i> var. <i>spontanea</i> - <i>lucidissima</i> complex	0.86016 ( $P > 0.10$ )	1.90954 ( $P < 0.02$ )	1.77600 ( $P < 0.05$ )
<i>R. chinensis</i> var. <i>spontanea</i>	1.40005 ( $P > 0.10$ )	1.55297 ( $P < 0.02$ )	1.79885 ( $P < 0.05$ )
<i>R. lucidissima</i>	0.90757 ( $P > 0.10$ )	1.74811 ( $P < 0.02$ )	1.70509 (0.10 > $P > 0.05$ )

The single population LB of southwestern Sichuan was Group I. Group II included populations CX, NJ, and PW of northern Sichuan. The single population RH in northern Guizhou was Group IV and the single population NC on the southeastern border of Sichuan Basin was Group V. Group III included all the other populations. The optimal  $F_{CT}$  of *R. lucidissima* was found with  $K = 4$ , implying the population could be divided into 4 groups. Group I included population HJ, LZ, and WS; population DJ, WA, and YY formed Group II; the single population PS was Group III; and populations AL and SB formed Group IV.

In the population neutrality test applied to either the complex or each of the two taxa, the D value did not deviate significantly from zero (**Table 4**). The test of mismatch distribution of the number of pairwise nucleotide differences for cpDNA haplotypes showed that the observed value departed from the expectation with the multimodal distribution (**Supplementary Figure S2A**), indicating a lack of recent range expansion within each taxon and the whole complex.

## Ribosome Haplotypes, Genetic Diversity, and Structure Based on ITS

The aligned 422 ITS sequences had 642 nucleotide positions of which 26 were polymorphic. In total, 21 ITS haplotypes (ribotypes) were identified and sequence variation for ribotypes can be seen in **Supplementary Table S2**. There were 13 and 11 ribotypes specific to *R. chinensis* var. *spontanea* and



*R. lucidissima*, respectively, while rH2, rH3, and rH5 were shared by the two taxa. The spatial distribution and relative frequencies of the ribotypes are presented in **Table 1** and **Figure 1B**.

The ribotype rH3 (206 of 422 sequences) predominated in the complex, especially in *R. chinensis* var. *spontanea*. Ribotype rH3 was present in almost all populations of *R. chinensis* var. *spontanea* and in two populations (PS and DJ) of *R. lucidissima*. These populations were around the Sichuan Basin. The second most popular ribotype was rH2, present in the majority of the populations of *R. lucidissima*, but in only one population of *R. chinensis* var. *spontanea* population (RH). Ribotype rH5 was found in populations of both taxa at the conjunction of Sichuan Basin and Yunnan-Guizhou Plateau, namely in populations NC, PS, YY, and DJ. Ribotype rH12 was found in *R. chinensis* var. *spontanea* populations of MB, RS, and PW on the western border of the Sichuan Basin. rH20 was in ZG and YC of *R. chinensis* var. *spontanea*, while rH13 was found in *R. lucidissima* populations PS and WA. The other ribotypes were all private. Among them, ribotypes rH4, rH10, and rH19 were represented by a single individual each.

The ITS genetic diversity ( $h_T$ ) of the complex was  $0.726 \pm 0.075$ , that of *R. chinensis* var. *spontanea* was  $0.464 \pm 0.107$  and that of *R. lucidissima* was  $0.907 \pm 0.031$ . The genetic diversity within the population ( $h_S$ ) of the complex was  $0.344 \pm 0.051$ , that of *chinensis* var. *spontanea* was  $0.295 \pm 0.07$  and that of *R. lucidissima* was  $0.431 \pm 0.064$ . The total haplotype diversity ( $H_d$ ) and the nucleotide ( $P_i$ ,  $\pi$ ) of the complex and each taxa based on ITS are shown in **Table 1**. For the complex, AMOVA revealed over half of the variation (58.90%) is among populations (**Supplementary Table S3**). In *R. chinensis* var. *spontanea*, the among population component was 55.72%, and in *R. lucidissima*, it was 40.22%. The  $N_m$  calculated for the complex was 0.225, and for *R. chinensis* var. *spontanea* and *R. lucidissima* it was 0.434 and 0.226, respectively. The genetic differentiation of the whole complex ( $G_{ST} = 0.53 \pm 0.06$ ,  $N_{ST} = 0.64 \pm 0.08$ ) and *R. lucidissima* ( $G_{ST} = 0.53 \pm 0.08$ ,  $N_{ST} = 0.52 \pm 0.13$ ) was higher than *R. chinensis* var. *spontanea* ( $G_{ST} = 0.37 \pm 0.07$ ,  $N_{ST} = 0.54 \pm 0.15$ ).

In SAMOVA, the optimal  $F_{CT}$  value for *R. chinensis* var. *spontanea* - *R. lucidissima* complex was found for  $K = 3$  (**Supplementary Figure S1B**), with Group I comprising almost all populations around the Sichuan Basin; Group II including populations of Yunnan-Guizhou Plateau and adjacent areas; and Group III comprising population LB in southwestern Sichuan and WS in southeastern Yunnan. For *R. chinensis* var. *spontanea*, SAMOVA identified two groups: Group I with only population LB, and Group II with all the remained populations. For *R. lucidissima*,  $F_{CT}$  was changing irregularly with an increase in  $K$  value, indicating no group structure.

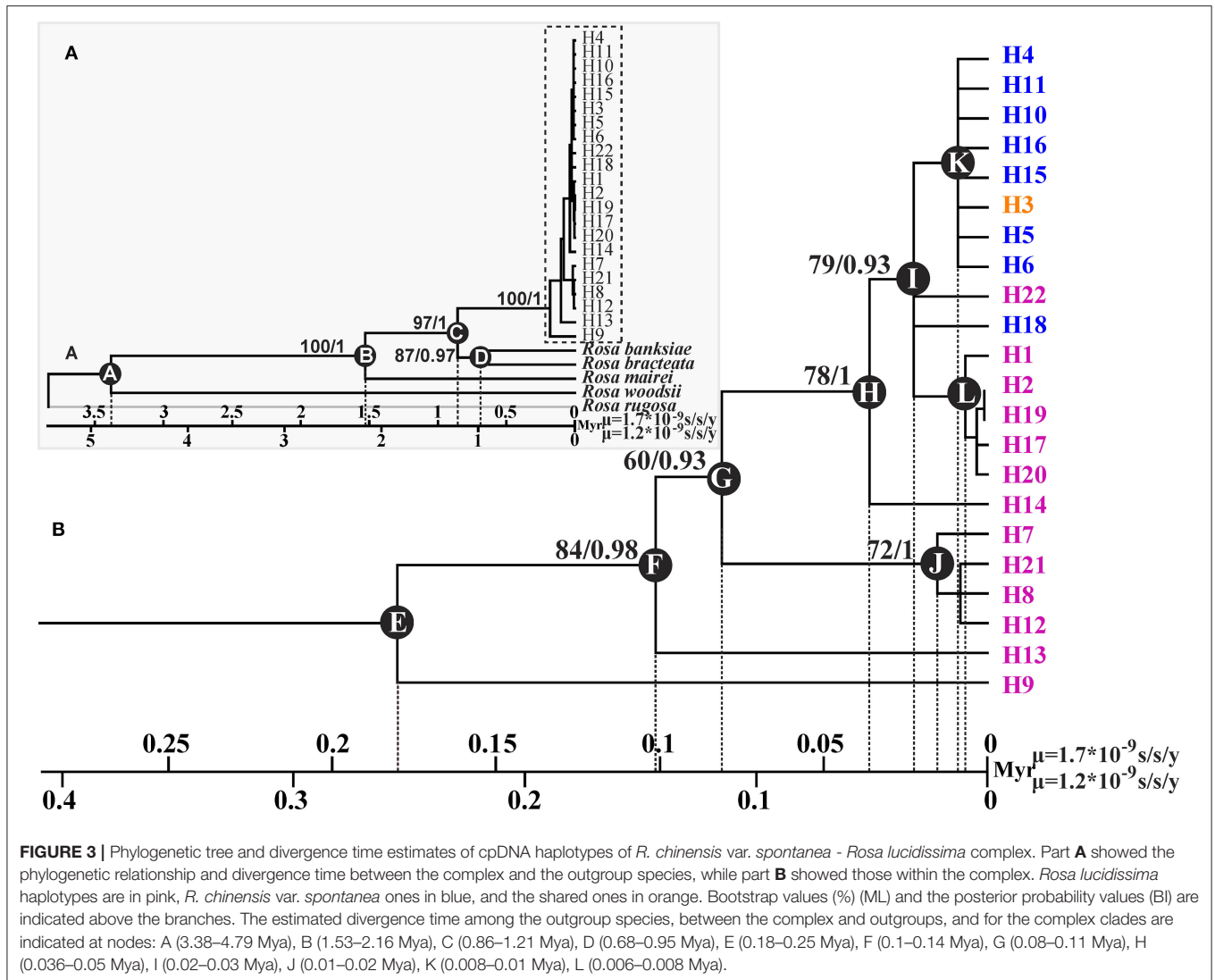
While Tajima's  $D$  values were not significant,  $D^*$  and  $F^*$  in Fu and Li's test significantly deviated from zero (**Table 4**), indicating a bottleneck effect and possible balancing selection, and the test of mismatch distribution revealed departure from the expectation with bimodal distribution

(**Supplementary Figure S2B**), indicating that populations in the entire geographical region lack of recent range expansion. This was also found both in the whole complex and for the two taxa analyzed separately.

## Genetic Relationship of Populations, Phylogenetic and Nested Clade Analysis

GenBank accession numbers from OL774542 to OL774546 were for the *trnG-trnS* sequences of *R. banksiae*, *R. bracteata*, *R. mairei*, *R. rugosa*, and *R. woodsii*, respectively, and those from OL774562 to OL774566 were for their *petL-psbE* sequences, respectively. The ML and BI applied to cpDNA data produced trees with identical topology (**Figure 3**). The 22 haplotypes identified in *R. chinensis* var. *spontanea* - *R. lucidissima* complex formed one clade that diverged from other congeneric species at about 1.21–0.86 Mya (node C). The majority of the haplotypes of *R. lucidissima* evolved much earlier than those of *R. chinensis* var. *spontanea*. Within *R. lucidissima*, haplotypes H9 and H13 are the most ancestral that diverged about 0.25–0.18 and 0.14–0.1 Mya (node E and F, respectively). The group of haplotypes H7, H21, and H8 diverged the next at about 0.08–0.11 Mya (node G) followed by haplotype H14 at about 0.036–0.051 Mya (node H). An ancestor that evolved at that time (node H), gave rise to a large number of haplotypes of both *R. lucidissima* and *R. chinensis* var. *spontanea*. Within this clade, the ancestral haplotype H18 of *R. chinensis* var. *spontanea* diverged from *R. lucidissima* haplotypes about 0.022–0.031 Mya (node I), and a group of *R. lucidissima* haplotypes H1, H2, H17, H19, H20 diverged about 0.006–0.008 Mya (node L). In *R. chinensis* var. *spontanea*, except for the most ancestral haplotype H18, other haplotypes diverged very recently at about 0.0086–0.012 Mya (node K).

GenBank accession numbers from OL719054 to OL719058 were ITS sequences of *R. bracteata*, *R. banksiae*, *R. rugosa*, *R. woodsii*, and *R. mairei*, respectively. The ITS ribotypes identified in *R. chinensis* var. *spontanea* - *R. lucidissima* complex also formed with high support of a separate clade in the tree (**Figure 4**). This clade had two distinct subclades. The first subclade included ribotypes rH18, rH19, rH20 of *R. chinensis* var. *spontanea* from the TGMR areas, the widely distributed rH3 shared by both *R. chinensis* var. *spontanea* and *R. lucidissima* around the Sichuan Basin, rH21 of *R. lucidissima* in PS and YY at the southeastern Sichuan basin near the Wulingshan Mountains, and rH6 and rH7 of *R. lucidissima* in HJ in northern Guangxi. The second one was further divided into two paraphyletic branches. The first contained rH4, rH14, rH12, and rH15 of *R. chinensis* var. *spontanea* from the northern and northwestern border of the Sichuan Basin. The second branch included rH1, rH11, and rH16 of *R. lucidissima* from AL, LZ, and SB, rH2 and rH5 shared by both taxa from NC, DJ, YY, and SB in the conjunction of Yunnan-Guizhou Plateau and southeastern Sichuan Basin, rH8 and rH9 of *R. chinensis* var. *spontanea* from LB in southwestern Sichuan and rH17 of *R. lucidissima* from WS in southeastern Yunnan.

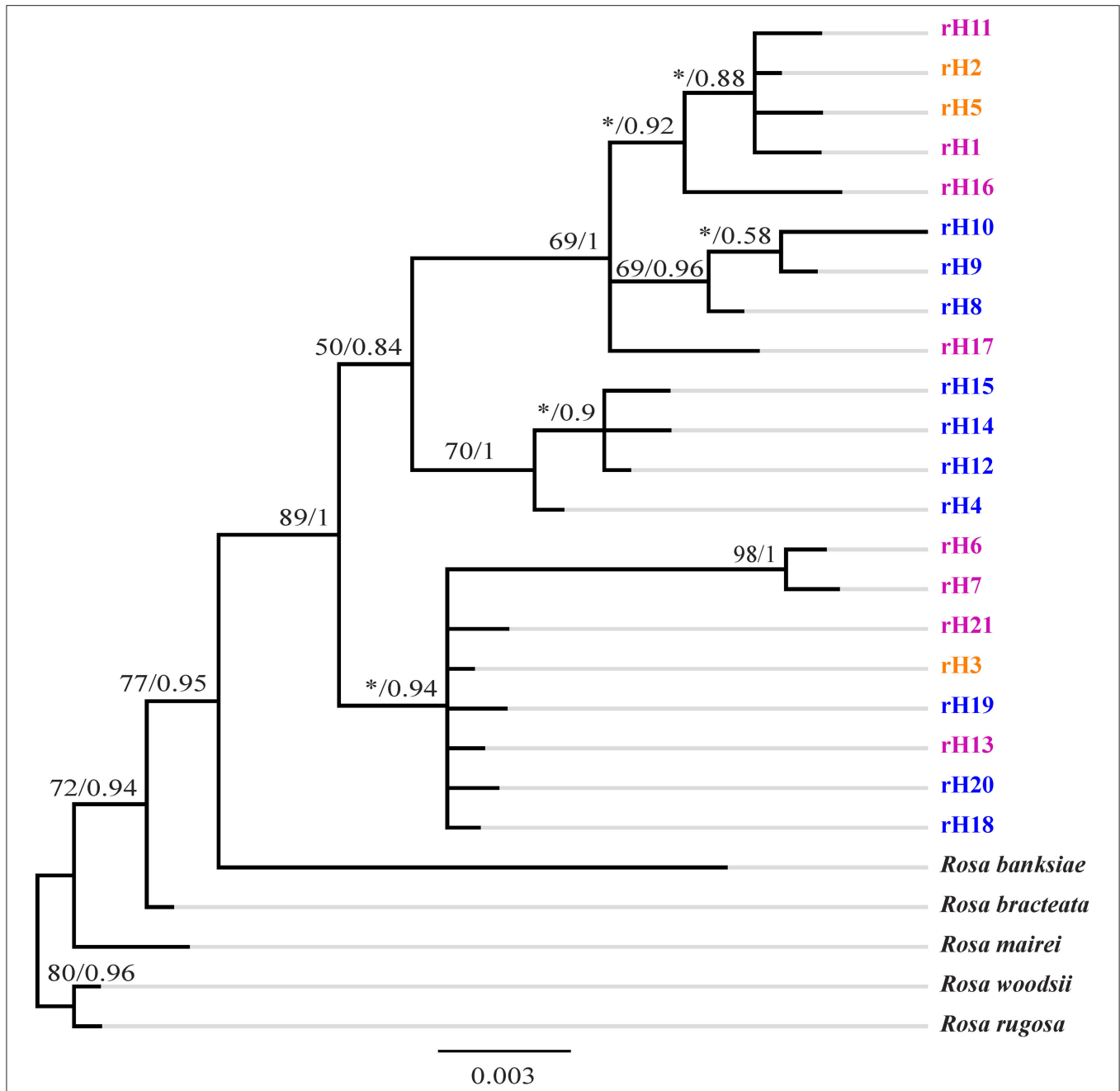


The NCPA applied to the chloroplast haplotype data produced a network of seven main clades (Figure 5). For only a few of them, the inferences of historical events that could shape their present-day spatial distribution can be made, with no overall conclusive inference for the complex evolution (Supplementary Table S4). Both the interior-clade 1-1 is formed by H20 (SB) and H17 (PS), and the subclade 2-12 is formed by interior-clade 1-4 (H5, NJ, and PW) and interior-clade 1-17 (H16, NC) could result from “allopatric fragmentation”. The clade 3-7 formed by subclade 2-12 and subclade 2-2 (H18, RH) could result from “allopatric fragmentation”, too. The interior-clade 1-5 including widely distributed H3, H4 from BD, H6 from FJ and H15 from NC could result from “restricted gene flow with isolation by distance”. The NCPA for ribotypes also produced a network of seven main clades (Supplementary Figure S3). At the complex level, no conclusive inference could be made for the historical events that shaped the present-day ribotype spatial distribution (Supplementary Table S4). Only subclade

1-3 including rH1 and rH2 from AL, rH11 from LZ was inferred as a result of “restricted gene flow/dispersal but with some long-distance dispersal over intermediate areas not occupied by the species; or past gene flow followed by extinction of intermediate populations”.

### Contribution of the Extant Populations to CpDNA Variation

The results of the analysis of population contribution to genetic diversity by Contrib are shown in Figure 6. The population RH, NC, PW, LB, CX, and NJ made the highest contribution to the total cpDNA genetic diversity and allelic diversity of *R. chinensis* var. *spontanea*. Among them, population RH, PW and CX contributed due to their high genetic differentiation (Cd) from others, even though they had low genetic diversity (Cs), while NC and LB had both high genetic differentiation and high genetic diversity. Contributions of the population of *R. lucidissima* (except for WA and DJ having low



**FIGURE 4** | Phylogenetic tree of ribotypes of *R. chinensis* var. *spontanea* and *R. lucidissima* complex. Bootstrap values (%) (ML) and the posterior probability values (BI) are indicated above the branches. *R. lucidissima* haplotypes are in pink, *R. chinensis* var. *spontanea* ones in blue, and the shared ones in orange. \*Means the bootstrap value is less than 50%.

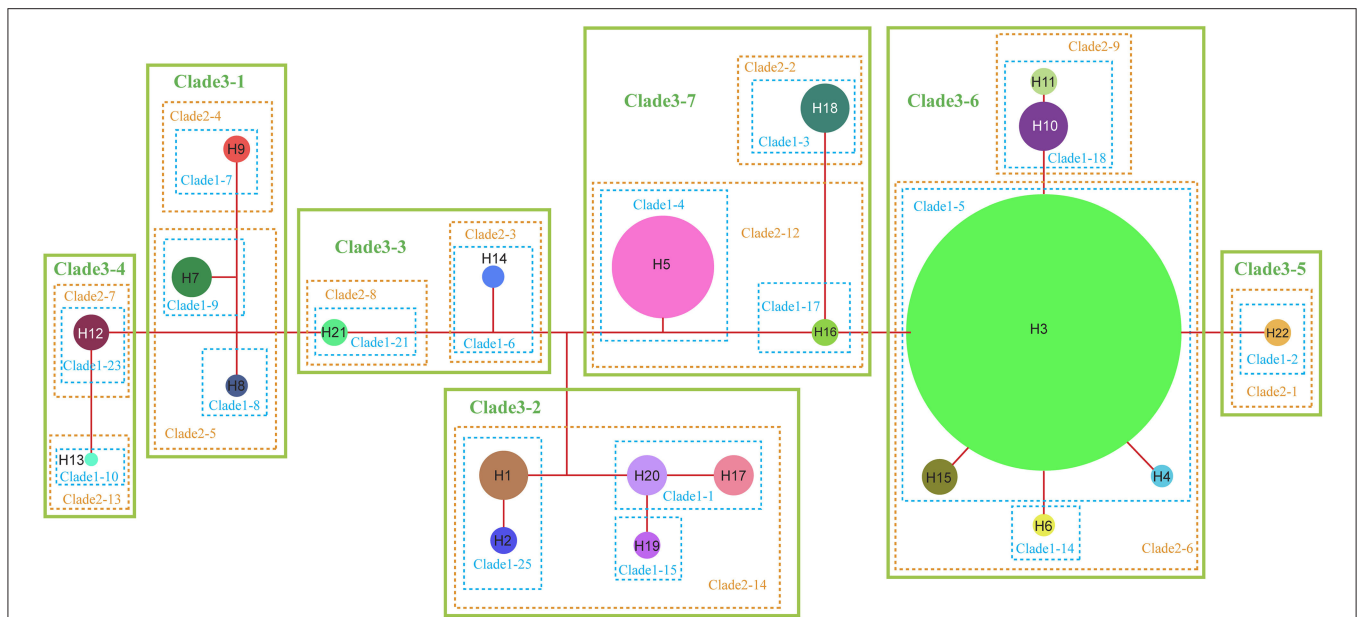
contribution) to both genetic diversity and allelic diversity was similarly high.

## DISCUSSION

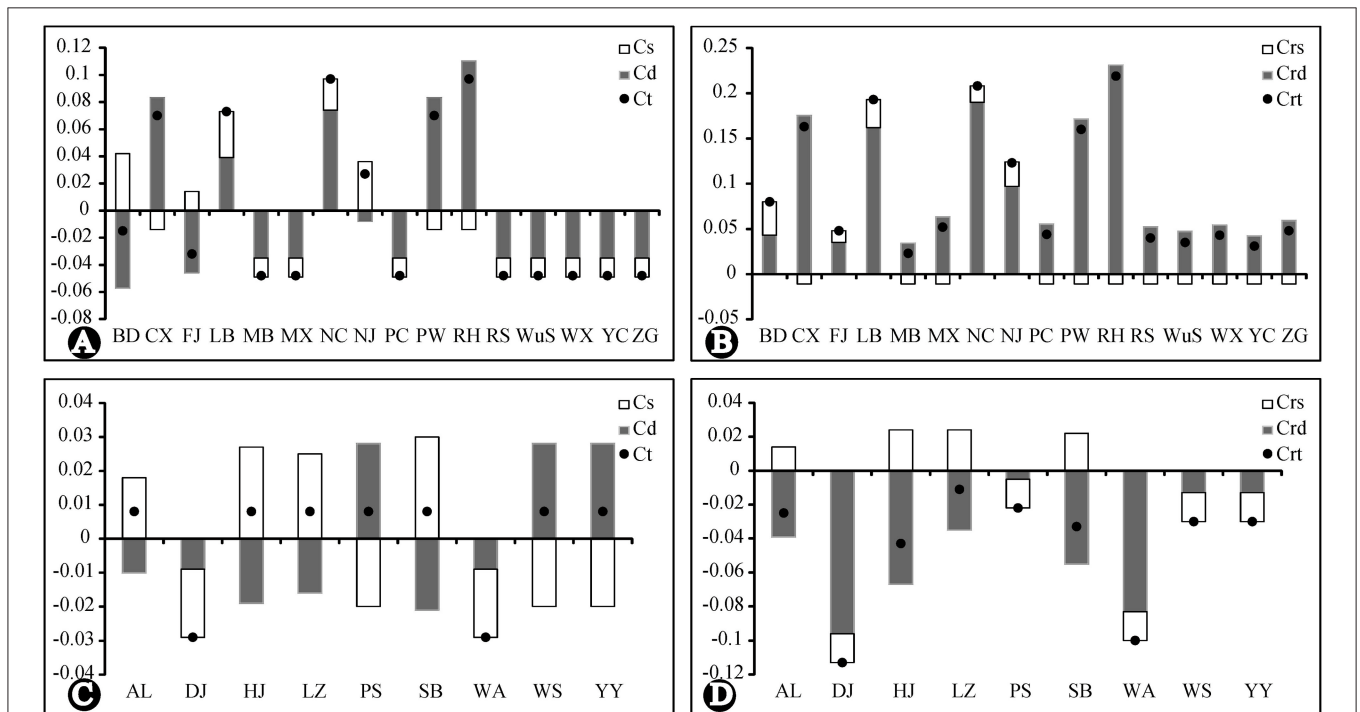
### Species Boundary Between *R. chinensis* var. *spontanea* and *R. lucidissima*

Wild rose species are characterized by extensive morphological variation, which has resulted in a notoriously complex taxonomy

(Wissemann, 2003; Joly et al., 2006), e.g., *R. caninae* complex (De Cock et al., 2008), *R. soulieana* (Jian et al., 2015), *R. sericea* complex (Gao et al., 2015, 2019) and *R. chinensis* var. *spontanea* and *R. lucidissima* complex. *Rosa chinensis* var. *spontanea* and *R. lucidissima* are very similar morphologically to each other (Henry, 1902; Léveillé, 1911; Ku, 1985; Ku and Robertson, 2003; Rix, 2005), and taxonomic circumscription between the two taxa is confusing both in herbarium sheets and in the field. Studies focusing on natural population are critical for understanding not



**FIGURE 5** | Nested cladogram of the chloroplast haplotypes (H1–H22) of *R. chinensis* var. *spontanea* and *R. lucidissima* complex. Circles with numbers denote haplotypes and their sizes are proportional to the observed frequencies of the haplotypes. Dots represent putative haplotypes. Each branch represents one mutation.



**FIGURE 6** | Contribution of each population to the genetic and allelic diversity of *R. chinensis* var. *spontanea* and *R. lucidissima* complex. *Cs*, the contribution of genetic diversity. *Cd*, the contribution of genetic differentiation. *Ct*, the contribution of total diversity. *Crs*, the contribution of allelic diversity. *Crd*, the contribution of allelic differentiation. *Crt*, the contribution of total allelic richness. **(A,B)**: *R. chinensis* var. *spontanea*. **(C,D)**: *R. lucidissima*.

only *Rosa* sections *Chinenses* and *Synstylae* but even the whole genus since many closely related species resulted from a fast diversification, incomplete lineage sorting or reticulate evolution

(Zhu et al., 2015). In our study, *R. chinensis* var. *spontanea* and *R. lucidissima* shared one cpDNA haplotype and three ribotypes, mainly in a transitional area between southeastern Sichuan Basin

and Yunnan-Guizhou Plateau, where populations of the two taxa meet. Secondly, although both taxa have many haplotypes limited to a small region or occurring in a single population, spatial distribution of the haplotypes roughly corresponds to the taxa boundaries based on morphology. Thirdly, populations of the whole complex were divided into two main clades based on the genetic distance among populations based on both ITS and cpDNA (**Supplementary Figure S4**). Fourthly, the ecological niches of *R. chinensis* var. *spontanea* and *R. lucidissima* only overlap at the transition zone between Yunnan-Guizhou Plateau and southeastern border of Sichuan Basin. So, we suggest *R. chinensis* var. *spontanea* and *R. lucidissima* be treated as independent taxa, according to morpho-geographical species concept (Mitaka, 2003). The northern populations of the complex, mainly those around and within the Sichuan Basin were *R. chinensis* var. *spontanea*, which has broader leaflets and paler full-blooming flowers in general, and those of the southern populations, mainly those distribute in Guangxi, Yunnan, Middle and Southern Guizhou, with narrower leaflets and darker flowers, were *R. lucidissima* (**Supplementary Figure S5**). The southeastern Sichuan Basin and the northeastern Guizhou are the contact or the hybridization zone of the two taxa. A large panel of SNPs at a genome-level should be necessary for a more precisely delimiting of both taxa in these transitional areas.

### Genetic Differentiation Within and Between *R. chinensis* var. *spontanea* and *R. lucidissima*

Southwestern China was weakly but widely affected by the Quaternary Glacials (Committee of Chinese Academy of Sciences for Physical Geography of China, 1984; Liu, 1991). The mountain boundaries between the Second and Third ladders in China were not only the obstacles that retarded gene flow between both sides of the mountains but also the corridors of plant dispersal (Hoorn et al., 2013). The Nanling Mountains (Wang et al., 2009; Kou et al., 2016), Daloushan Mountains (Ma et al., 2015), Wushan and Dabashan Mountains (Liu et al., 2013) all have played such a role in the Sino-Japanese Flora (reviewed by Ye et al., 2017b and references therein). The subtropical flora of China has been traditionally thought to be originated from southwestern China and expanded eastward through such mountain corridors (Wang, 1992a,b). As parts of this flora, *R. chinensis* var. *spontanea* – *R. lucidissima* complex originated and migrated like this. Haplotypes of *R. lucidissima* diverged from other roses at about 0.86–1.21 Mya presumably in the Günz glacial substage of the Quaternary Oscillation mainly in mountain areas of southeastern Yunnan-Guizhou Plateau and adjacent regions. During the third glacial advance in the Quaternary Oscillation (0.37–0.24 Mya), the ancestors of the southern haplotypes apparently survived on the Jiuwan Mountain in Huanjiang (HJ), where there exist many relic plant species of the Quaternary Glacials (Compilation Committee of Local Chronicles of Guangxi Zhuang Autonomous Region, 1994). The ancestral haplotypes of the middle ones dispersed north to the Daloushan and then east to the Wulingshan Mountains at the northeastern edge of Yunnan-Guizhou Plateau and southeastern of Sichuan Basin,

where the ancestor of *R. chinensis* var. *spontanea* evolved from *R. lucidissima*, and then both taxa survived the Last Glacial Maximum (LGM) there. So, the Daloushan and Wulingshan Mountains were the complex's refugia in LGM. With the ending of the Quaternary glacials and the coming of the postglacial period, the ancestral haplotypes of the haplotypes in the refugia dispersed northeast along the Wushan Mountains to the Three Gorges Mountains Region and dispersed north along the Ridge and Valleys of the east Sichuan to the Dabashan Mountains, from where the most widely distributed haplotypes (H3) originated and then widely distributed the mountains around the Basin. The southern haplotypes and northern ones within the complex kept being isolated by the environment (IBE) due to the combination of altitude, climate and soil (Bai et al., 2016; Ye et al., 2017a). Then they differentiated further and evolved into the southern *R. lucidissima* and the northern *R. chinensis* var. *spontanea*.

Within *R. chinensis* var. *spontanea*, almost all haplotypes were very recently diverged at about 0.0086–0.012 Mya. For *R. lucidissima*, most haplotypes originated during or even later than this period. On the other hand, almost all cpDNA haplotypes were private to some single population. The reason for this recent and fast diversification of haplotypes might firstly be due to the wide and complex geographical range covering a mosaic of plateaus, mountains, basins, and gorges which could act as geographic barriers of gene flow and thus provided ample opportunity for isolation, drift and mutation (Wang et al., 2009; Ma et al., 2015). Secondly, hybridization between the two taxa, or gene introgression from other rose species could be another reason. It is well known that inter-species hybridization is common in genus *Rosa* (Ritz et al., 2005; Joly and Bruneau, 2006; Schanzer and Kutlunina, 2010; Ritz and Wissemann, 2011; Kellner et al., 2012; Fougère-Danezan et al., 2014). Populations in the contact areas of the two taxa showed richer ribotypes in this study, and we found that populations with higher  $H_d$  value often have individuals with more or less phenotypic traits significantly alike other sympatric rose, such as *R. banksiae* f. *normalis* and *R. rubus* with same blooming period.

### Population History and Refugia of *R. chinensis* var. *spontanea* and *R. lucidissima*

Our demographic analyses indicated that there was no range expansion of either *R. chinensis* var. *spontanea* or *R. lucidissima*, which could be due to their narrow habitat requirements. Both taxa are highly heliophilous and readily colonize deforested locations but can not compete with the evergreen forest vegetation covering the Sichuan Basin and East China at lower altitudes (Yu et al., 2000; Harrison et al., 2001).

Refugial populations are expected to have higher genetic diversity and longer demographic history than populations that evolved later (Tribsch and Schönswetter, 2003; Lowe et al., 2005; Provan and Bennett, 2008). The Yunnan-Guizhou Plateau and its adjacent regions have long been believed to be an important center of origin for the East Asiatic flora (Wang, 1992a,b) and a refugial area for many species in China (Ying, 2001). Populations of *R. lucidissima* in the southeastern Yunnan-Guizhou Plateau

and its adjacent regions, especially populations HJ and LZ had much higher genetic diversity and harbored much more ancient haplotypes (both cpDNA and ITS) than in the other parts of the taxa range, supporting the above hypothesis. In *R. chinensis* var. *spontanea*, population BD, FJ, NJ, LB, and NC all had high genetic diversity, but only RH within Daloushan in the northeastern Yunnan-Guizhou Plateau adjacent to the Sichuan Basin harbored the ancient haplotype H18. Thus, these areas could be possible refugia for *R. chinensis* var. *spontanea*. The population LB in the southwestern edge of Sichuan Basin (Xiaoxiangling Mountains), and NC in southeastern Sichuan Basin having high diversity but young haplotypes, indicated that these areas might be the contact zone or areas of high immigration from diverse sources of *R. chinensis* var. *spontanea* - *R. lucidissima* complex because of their consistent and favorable local conditions.

## The Genetic Diversity and Genetic Structure of the Complex

Despite with a relatively short evolutionary history (1.21–0.86 Mya), *R. chinensis* var. *spontanea* - *R. lucidissima* had a relatively high level of chloroplast genetic diversity ( $h_T$ , 0.768), comparable to that of *R. praelucens*, endemic to Shangrila of Yunnan Plateau (0.720, Jian et al., 2018), and the widely distributed *R. omeiensis* - *R. sericea* complex (0.790, Gao et al., 2019), and even much higher than that of another widely distributed wild rose, *R. soulieana* (0.316, Jian et al., 2015). The complex is also characterized by very high genetic differentiation among populations and extremely low genetic variation within populations ( $G_{ST} = 0.83$ ). This extent and structure of genetic diversity appears to be a result of, on one hand, a wide geographical range, covering a mosaic of plateaus, mountains, basins, and gorges in Southwest China, and, on the other hand, numerous geographic barriers within this range limiting seed dispersal and promoting genetic drift. The comprising complex plants are long-living woody shrubs with hermaphroditic flowers pollinated mainly by moths, bees, and other small insects. The migration distance of these insects is limited to several kilometers (Rands and Whitney, 2011). The seed dispersal, unlike many other species, is by gravity and small mammals. In such rose species as *R. canina* (Herrera, 1984), *R. arvensis* (Vander Mijnsbrugge et al., 2010), *R. multiflora* (Ghosh, 2009), and *R. soulieana* (Jian et al., 2015), the hips are red, often < 1 cm in diameter and persist on the plants for a long time and therefore often dispersed by birds. In contrast, the hips of *R. chinensis* var. *spontanea* and *R. lucidissima* are greenish yellow and usually about 2–3 cm in diameter. The hips easily fall off the plants as soon as or even before they ripe. The hips only fall around the mother plants and the achenes (seeds) are dispersed mainly by the vertebrates not the birds, and thus within a rather limited distance. Also, because it is mainly distributed in the middle and lower parts of the mountains and along the valleys, the huge mountains act as the barriers to the pollen and seeds dispersal of *R. chinensis* var. *spontanea* and *R. lucidissima*. All these factors may limit inter-population gene flow and contribute to isolation and genetic differentiation of populations.

However, compared to that of cpDNA, genetic diversity based on nuclear DNA (ITS) within the populations ( $h_S$ ) was relatively higher, and resulting in less percent of genetic variation existing among the populations and a higher gene flow among populations. These discordant between cpDNA and ITS may result from different dispersal vectors for seeds and pollen, as gene flows due to pollination might be easier and further than those due to seeds dispersal. And also, hybridization within the complex might decrease the genetic differentiation among populations.

## Conservation Implications

Wild plant resources provide infinite potential for superior breeding to improve the genetic diversity of cultivated plants (Long et al., 2018; Wang et al., 2020). Rose is the world's most important ornamental plant, with economic, cultural, and symbolic value (Gudin, 1999; Hibrand Saint-Oyant et al., 2018; Raymond et al., 2018). It is known that 8 to 15 species contributed to the modern rose cultivars (Leus et al., 2018), and unfortunately, this important genetic pool of modern roses is decaying due to intensive breeding (Matsumoto et al., 1998; Scariot et al., 2006; Leus et al., 2018). Wild rose species germplasm is necessary for resistance to introgression (Young et al., 2019), flower color breeding (Zlesak, 2015), and introduction of specific traits (Leus et al., 2018), and therefore it is necessary to protect *R. chinensis* var. *spontanea* - *R. lucidissima* complex.

During our 2 years of outdoor investigation, we did not find any individuals in Yanhe and Xingren in Guizhou, Lingchuan, Fushui, and Cangwu in Guangxi, where there once had specimens records and local people used to dig the roots of *R. lucidissima* to cure stomach diseases. Local people near several other northern populations told us some companies collected old plants of *R. chinensis* var. *spontanea* as the tree rose rootstocks. Also, populations in the dense forest usually had very slim and weak individuals, which were gradually replaced by other plants. So the endangered status of the whole complex might be the result of the over-exploitation of its medical and ornamental value or reforestation of many originally open habitats. We recommend, based on the results of genetic analysis and the contribution of each population, focusing conservation efforts on the populations RH, NC, PW, LB, CX, and NJ of *R. chinensis* var. *spontanea* having the highest contribution to the total cpDNA genetic diversity and allelic richness. Also, populations BD, FJ, RS, WuS, and YC having private haplo- and ribotypes must be given priority in conservation planning and implementation. In *R. lucidissima*, all populations (except for WA and DJ) have high conservation value according to their genetic contribution and haplotype distribution.

The most appropriate and easy to implement an approach for the conservation of these populations *in situ* are plant micro-reserves (Laguna et al., 2004) allowing legal protection, long-term monitoring, and necessary management actions (Volis, 2016). In addition, living collections preserving the existing species' genetic diversity or special phenotypes in a form of seed orchards or quasi *in situ* collections should be established. Given a low number of the extant populations

and the high value of this gene pool for breeding, germplasm nurseries such as, for example, those created in Belgium for autochthonous populations of woody species of Flanders (Vander Mijnsbrugge, 2014), seem to be the especially suited for *Rosa ex situ* approach.

## DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Material**.

## AUTHOR CONTRIBUTIONS

HYJ conceived and designed the article and wrote the manuscript. LZ, HJY, XQQ, and TZ performed the experiment. HYJ, LZ, HZ, and QGW collected the plant materials. HYJ, CLM, and NNZ analyzed the data. All authors contributed to the article and approved the submitted version.

## REFERENCES

- Akasaka, M., Ueda, Y., and Koba, T. (2002). Karyotype analyses of five wild rose species belonging to septet A by fluorescence in situ hybridization. *Chromosome Sci.* 6, 17–26.
- Avise, J. C. (2000). *Phylogeography: The History and Formation of Species*. Cambridge: Harvard University Press.
- Avise, J. C. (2004). *Molecular Markers, Natural History, and Evolution*, 2nd edn. Sunderland: Sinauer Associates.
- Bai, W. N., Wang, W. T., and Zhang, D. Y. (2016). Phylogeographic breaks within Asian butternuts indicate the existence of a phytogeographic divide in East Asia. *New Phytol.* 209, 1757–1772. doi: 10.1111/nph.13711
- Burban, C., Petit, R. J., Carcreff, E., and Jactel, H. (1999). Rangewide variation of the maritime pine bast scale *Matsucoccus feytaudi* Duc. (Homoptera: Matsucoccidae) in relation to the genetic structure of its host. *Mol. Ecol.* 8, 1593–1602. doi: 10.1046/j.1365-294x.1999.00739.x
- Cao, S. R., Zhang, T., Wang, Q. G., Qiu, X. Q., Ma, C. L., and Jian, H. Y. (2021). Karyological analysis of eight rose germplasm resources. *North. Hort.* 12, 85–90. doi: 10.11937/bfyf.20204689
- Committee of Chinese Academy of Sciences for Physical Geography of China (1984). *Physical Geography of China: Palaeogeography*. Beijing: Science Press.
- Compilation Committee of Local Chronicles of Guangxi Zhuang Autonomous Region (1994). *General Records of Guangxi: Natural Geography*. Nanling: Guangxi People's Publishing House.
- Crandall, K. A., and Templeton, A. R. (1993). Empirical tests of some predictions from coalescent theory with applications to intraspecific phylogeny reconstruction. *Genetics* 134, 959–969. doi: 10.1093/genetics/134.3.959
- Cui, J. P., Zhao, S. W., Dong, Z. Y., and Liu, H. X. (2016). "Investigation and collection of *Rosa chinensis* var. *spontanea* in some areas of Northern Sichuan," in *Proceedings of 2016 WFRS Regional Convention Jointly With the 14th International Heritage Rose Conference*, ed The Chinese Rose Society, Organizing Committee of 2016 WFRS Regional Convention (Beijing), 149–158.
- De Cock, K., Vander Mijnsbrugge, K., Breyne, P., Van Bockstaele, E., and Van Slycken, J. (2008). Morphological and AFLP-based differentiation within the taxonomical complex Section *Caninae* (subgenus *Rosa*). *Ann. Bot.* 102, 685–697. doi: 10.1093/aob/mcn151

## FUNDING

This work was financially supported by the National Key R&D Program of China (2019YFD1000400) and the National Natural Science Foundation of China (31760087), the Major Program for Science and Technology of Yunnan Province (202102AE090052 and 202002AA100007).

## ACKNOWLEDGMENTS

We sincerely thank Dr. Sergei Volis for their comments and suggestions to improve the manuscript. We thank Xiaoyun Zhang and Xingjia Tang for their help in the sample collection. We thank Dr. Yonghong Zhang for his help in data analysis.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2022.851396/full#supplementary-material>

- De Riek, J., De Cock, K., Smulders, M. J. M., and Nybom, H. (2013). AFLP-based population structure analysis as a means to validate the complex taxonomy of dogroses (*Rosa* section *Caninae*). *Mol. Phylogenet. Evol.* 67, 547–559. doi: 10.1016/j.ympev.2013.02.024
- Deng, T., Abbott, R. J., Li, W., Sun, H., and Volis, S. (2019). Genetic diversity hotspots and refugia identified by mapping multi-plant species haplotype diversity in China. *Isr. J. Plant Sci.* 66, 136–151. doi: 10.1163/22238980-20191083
- Dupanloup, I., Schneider, S., and Excoffier, L. (2002). A simulated annealing approach to define the genetic structure of populations. *Mol. Ecol.* 11, 2571–2581. doi: 10.1046/j.1365-294X.2002.01650.x
- Excoffier, L., Laval, G., and Schneider, S. (2005). Arlequin ver. 3.0: an integrated software package for population genetics data analysis. *Evol. Bioinform. Online* 1, 47–50. doi: 10.1177/117693430500100003
- Fougère-Danezan, M., Joly, S., Bruneau, A., Gao, X. F., and Zhang, L. B. (2014). Phylogeny and biogeography of wild roses with specific attention to polyploids. *Ann. Bot.* 115, 275–291. doi: 10.1093/aob/mcu245
- Fu, Y. X. (1997). Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* 147, 915–925. doi: 10.1101/gad.7.3.517
- Gao, L. M., Moeller, M., Zhang, X. M., Hollingsworth, M. L., Liu, J., Mill, R. R., et al. (2007). High variation and strong phylogeographic pattern among cpDNA haplotypes in *Taxus wallichiana* (Taxaceae) in China and North Vietnam. *Mol. Ecol.* 16, 4684–4698. doi: 10.1111/j.1365-294X.2007.03537.x
- Gao, Y. D., Gao, X. F., and Harris, A. (2019). Species boundaries and parapatric speciation in the complex of alpine shrubs, *Rosa sericea* (Rosaceae), based on population genetics and ecological tolerances. *Front. Plant Sci.* 10, 321. doi: 10.3389/fpls.2019.00321
- Gao, Y. D., Zhang, Y., Gao, X. F., and Zhu, Z. M. (2015). Pleistocene glaciations, demographic expansion and subsequent isolation promoted morphological heterogeneity: a phylogeographic study of the alpine *Rosa sericea* complex (Rosaceae). *Sci. Rep.* 5, 11698. doi: 10.1038/srep11698
- Ghosh, R. (2009). *Assessment of genetic variation and population differentiation in invasive multiflora rose, Rosa multiflora Thunberg (Rosaceae) in Northeastern Ohio*. (Master's thesis) Kent State University.
- Graur, D., and Li, W. H. (2000). *Fundamentals of Molecular Evolution*, 2nd Edn. Sunderland: Sinauer Associates.
- Guan, B. C., Fu, C. X., Qiu, Y. X., Zhou, S. L., and Comes, H. P. (2010). Genetic structure and breeding system of a rare understory herb, *Dysosma versipellis*

- (Berberidaceae), from temperate deciduous forests in China. *Am. J. Bot.* 97, 111–122. doi: 10.3732/ajb.0900160
- Gudin, S. (1999). Improvement of rose varietal creation in the world. *Acta Hortic.* 495, 283–291. doi: 10.17660/ActaHortic.1999.495.14
- Hall, T. A. (1999). BioEdit: a user-friendly sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.* 41, 95–98.
- Hampe, A., Rodríguez-Sánchez, F., Dobrowski, S., Hu, F. S., and Gavin, D. G. (2013). Climate refugia: from the Last Glacial Maximum to the twenty-first century. *New Phytol.* 197, 16–18. doi: 10.1111/nph.12059
- Harrison, S. P., Yu, G., Takahara, H., and Prentice, I. C. (2001). Palaeovegetation: diversity of temperate plants in East Asia. *Nature* 413, 129–130. doi: 10.1038/35093166
- Henry, A. (1902). Wild Chinese roses. *Gardeners Chron.* 809, 438–439.
- Herrera, C. M. (1984). Seed dispersal and fitness determinants in wild rose: Combined effects of hawthorn, birds, mice and browsing ungulates. *Oecologia.* 63, 386–393. doi: 10.1007/BF00390670
- Hibrand Saint-Oyant, L., Ruttink, T., Hamama, L., Kirov, I., Lakhwani, D., Zhou, N. N., et al. (2018). A high-quality genome sequence of *Rosa chinensis* to elucidate ornamental traits. *Nat. Plants* 4, 473–484. doi: 10.1038/s41477-018-0166-1
- Hoorn, C., Mosbrugger, V., Mulch, A., and Antonelli, A. (2013). Biodiversity from mountain building. *Nat. Geosci.* 6, 154. doi: 10.1038/ngeo1742
- Huelsenbeck, J. P., and Ronquist, F. (2001). MRBAYES: Bayesian inference of phylogeny. *Bioinformatics* 17, 754–755. doi: 10.1093/bioinformatics/17.8.754
- Hurst, C. C. (1941). Notes on the origin and evolution of our garden roses. *J. R. Hort. Soc.* 66, 73–82.
- Jian, H. Y., Li, S. F., Guo, J. L., Li, S. B., Wang, Q. G., Yan, H. J., et al. (2018). High genetic diversity and differentiation of an extremely narrowly distributed and critically endangered decaploid rose (*Rosa praelucens*): implications for its conservation. *Conserv. Genet.* 19, 761–776. doi: 10.1007/s10592-018-1052-0
- Jian, H. Y., Tang, K. X., and Sun, H. (2015). Phylogeography of *Rosa soulieana* (Rosaceae) in the Hengduan Mountains: refugia and ‘melting’ pots in the Quaternary climate oscillations. *Plant Syst. Evol.* 301, 1819–1830. doi: 10.1007/s00606-015-1195-0
- Joly, S., and Bruneau, A. (2006). Incorporating allelic variation for reconstructing the evolutionary history of organisms from multiple genes: an example from *Rosa* in North America. *Syst. Biol.* 55, 623–636. doi: 10.1080/10635150600863109
- Joly, S., Starr, J. R., Lewis, W. H., and Bruneau, A. (2006). Polyploid and hybrid evolution in roses east of the Rocky Mountains. *Am. J. Bot.* 93, 412–425. doi: 10.3732/ajb.93.3.412
- Katoh, K., and Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780. doi: 10.1093/molbev/mst010
- Kellner, A., Ritz, C. M., and Wissemann, V. (2012). Hybridization with invasive *Rosa rugosa* threatens the genetic integrity of native *Rosa mollis*. *Bot. J. Linn. Soc.* 170, 472–484. doi: 10.1111/j.1095-8339.2012.01298.x
- Kou, Y. X., Cheng, S. M., Tian, S., Li, B., Fan, D. M., Chen, Y. J., et al. (2016). The antiquity of *Cyclocarya paliurus* (Juglandaceae) provides new insights into the evolution of relict plants in subtropical China since the late Early Miocene. *J. Biogeogr.* 43, 351–360. doi: 10.1111/jbi.12635
- Ku, T. C. (1985). “*Rosa*,” in *Flora Reipublicae Popularis Sinicae*, Vol. 37. eds T. T. Yu and T. C. Ku (Beijing: Science Press), 360–455.
- Ku, T. C., and Robertson, K. R. (2003). “*Rosa* (Rosaceae),” in *Flora of China*, Vol. 9. eds Z. Y. Wu and P. H. Raven (Beijing: Science Press; St. Louis: Missouri Botanical Garden Press), 339–381.
- Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018). MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* 35, 1547–1549. doi: 10.1093/molbev/msy096
- Laguna, E., Deltoro, V. I., Perez-Botella, J., Perez-Rovira, P., Olivares, A., and Fabregat, C. (2004). The role of small reserves in plant conservation in a region of high diversity in eastern Spain. *Biol. Conserv.* 119, 421–426. doi: 10.1016/j.biocon.2004.01.001
- Leus, L., Van Laere, K., De Riek, J., and Van Huylenbroeck, J. (2018). “Rose,” in *Ornamental Crops, Handbook of Plant Breeding*, Vol. 11. ed J. Van Huylenbroeck (Cham: Springer International Publishing), 719–768.
- Léveillé, H. (1911). *Repertorium specierum Novarum Regni Vegetabilis*. 9, 444.
- Liu, G. D. (1991). “Climate changes,” in *Climate of Southwest China*, ed Y. H. Xu. (Beijing: China Meteorological Press), 281–297.
- Liu, J. Q., Sun, Y. S., Ge, X. J., Gao, L. M., and Qiu, Y. X. (2012). Phylogeographic studies of plants in China: advances in the past and directions in the future. *J. Syst. Evol.* 50, 267–275. doi: 10.1111/j.1759-6831.2012.00214.x
- Liu, W., Kang, M., Tian, H., and Huang, H. (2013). A range wide geographic pattern of genetic diversity and population structure of *Castanea mollissima* populations inferred from nuclear and chloroplast microsatellites. *Tree Genet. Genom.* 9, 975–987. doi: 10.1007/s11295-013-0610-3
- Long, C. L., Cheng, Z., Zhou, Y., and Long, B. (2018). “The role of biodiversity and plant conservation for ornamental breeding,” in *Ornamental Crops, Handbook of Plant Breeding*, Vol. 11. ed J. Van Huylenbroeck (Cham: Springer International Publishing), 1–12.
- Lowe, A. J., Boshier, D., Ward, M., Bacles, C. E. F., and Navarro, C. (2005). Genetic resource impacts of habitat loss and degradation: reconciling empirical evidence and predicted theory for Neotropical trees. *Heredity* 95, 255–273. doi: 10.1038/sj.hdy.6800725
- Ma, Q., Du, Y. J., Chen, N., Zhang, L. Y., Li, J. H., and Fu, C. X. (2015). Phylogeography of *Davidia involucrata* (Davidiaceae) inferred from cpDNA haplotypes and nSSR Data. *Syst. Bot.* 40, 796–810. doi: 10.1600/036364415X689267
- Matsumoto, S., Kouchi, M., Yabuki, J., Kusunoki, M., Ueda, Y., and Fukui, H. (1998). Phylogenetic analysis of the genus *Rosa* using MatK sequence: molecular evidence for the narrow genetic background of modern roses. *Sci. Hortic.* 77, 73–82. doi: 10.1016/S0304-4238(98)00169-1
- Médail, F., and Diadema, K. (2009). Glacial refugia influence plant diversity patterns in the mediterranean basin. *J. Biogeogr.* 36, 1333–1345. doi: 10.1111/j.1365-2699.2008.02051.x
- Meng, J., Fougère-Danezan, M., Zhang, L. B., Li, D. Z., and Yi, T. S. (2011). Untangling the hybrid origin of the Chinese tea roses: evidence from DNA sequences of single-copy nuclear and chloroplast genes. *Plant Syst. Evol.* 297:157–170. doi: 10.1007/s00606-011-0504-5
- Mitaka, J. (2003). *The Genus Aconitum L. (Ranunculaceae) in Poland and Adjacent Countries. A Phonetic-Geographic Study*. Krakow: The Institute of Botany of the Jagiellonian University.
- Mitsui, Y., Chen, S. T., Zhou, Z. K., Peng, C. I., Deng, Y. F., and Setoguchi, H. (2008). Phylogeny and biogeography of the genus *Ainsliaea* (Asteraceae) in the Sino-Japanese region based on nuclear rDNA and plastid DNA sequence data. *Ann. Bot.* 101, 111–124. doi: 10.1093/aob/mcm267
- Nylander, J. A. A. (2004). *MrModeltest v2*. Program distributed by the author. Evolutionary Biology Centre, Uppsala University.
- Panchal, M. (2007). The automation of nested clade phylogeographic analysis. *Bioinformatics* 23, 509–510. doi: 10.1093/bioinformatics/btl614
- Petit, R. J., el Mousadik, A., and Pons, O. (1998). Identifying populations for conservation on the basis of genetic markers. *Conserv. Biol.* 12, 844–855. doi: 10.1111/j.1523-1739.1998.96489.x
- Pons, O., and Petit, R. J. (1996). Measuring and testing genetic differentiation with ordered versus unordered alleles. *Genetics* 144, 1237–1245. doi: 10.1093/genetics/144.3.1237
- Posada, D., and Crandall, K. A. (1998). Modeltest: testing the model of DNA substitution. *Bioinformatics* 14, 817–818. doi: 10.1093/bioinformatics/14.9.817
- Posada, D., and Crandall, K. A. (2001). Intraspecific gene genealogies: trees grafting into networks. *Trends Ecol. Evol.* 16, 37–45. doi: 10.1016/S0169-5347(00)02026-7
- Provan, J., and Bennett, K. (2008). Phylogeographic insights into cryptic glacial refugia. *Trends Ecol. Evol.* 23, 564–571. doi: 10.1016/j.tree.2008.06.010
- Qin, H. N., Yang, Y., Dong, S. Y., He, Q., Jia, Y., Zhao, L. N., et al. (2017). Threatened species list of China’s higher plants. *Biodiv. Sci.* 25, 696–744. doi: 10.17520/biodis.2017144
- Qiu, Y. X., Fu, C. X., and Comes, H. P. (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. *Mol. Phylogenet. Evol.* 59, 225–244. doi: 10.1016/j.ympev.2011.01.012
- Qiu, Y. X., Guan, B. C., Fu, C. X., and Comes, H. P. (2009). Did glacials and/or interglacials promote allopatric incipient speciation in East Asian



- temperate plants? Phylogeographic and coalescent analyses on refugial isolation and divergence in *Dyosma versipellis*. *Mol. Phylogenet. Evol.* 49, 702–714. doi: 10.1016/j.ympev.2009.01.016
- Rands, S. A., and Whitney, H. M. (2011). Field margins, foraging distances and their impacts on nesting pollinator success. *PLoS ONE* 6, e25971. doi: 10.1371/journal.pone.0025971
- Raymond, O., Gouzy, J., Just, J., Badouin, H., Verdenaud, M., Lemainque, A., et al. (2018). The *Rosa* genome provides new insights into the domestication of modern roses. *Nat. Genet.* 50, 772–777. doi: 10.1038/s41588-018-0110-3
- Ritz, C. M., Schmutz, H., and Wissemann, V. (2005). Evolution by reticulation: European dogroses originated by multiple hybridization across the genus *Rosa*. *J. Hered.* 96, 4–14. doi: 10.1093/jhered/esi011
- Ritz, C. M., and Wissemann, V. (2011). Microsatellites analyses of artificial and spontaneous dogrose hybrids reveal the hybridogenic origin of *Rosa micrantha* by the contribution of unreduced gametes. *J. Hered.* 102, 217–227. doi: 10.1093/jhered/esq124
- Rix, M. (2005). *Rosa chinensis*. f. *spontanea*. *Curtiss Bot. Mag.* 22, 214–219. doi: 10.1111/j.1355-4905.2005.00494.x
- Scalliet, G., Piola, F., Douady, C. J., Réty, S., Raymond, O., Baudino, S., et al. (2008). Scent evolution in Chinese roses. *Proc. Natl. Acad. Sci.* 105, 5927–5932. doi: 10.1073/pnas.0711551105
- Scariot, V., Akkac, A., and Botta, R. (2006). Characterization and genetic relationships of wild species and old garden roses based on microsatellite analysis. *J. Am. Soc. Hortic. Sci.* 131, 66–73. doi: 10.1007/s10658-005-2501-4
- Schanzer, I. A., and Kutlunina, N. A. (2010). Interspecific hybridization in wild roses (*Rosa* L. sect. *Caninae* DC.). *Biol. Bull.* 37, 480–488. doi: 10.1134/S1062359010050079
- Shaw, J., Lickey, E. B., Beck, J. T., Farmer, S. B., Liu, W., Miller, J., et al. (2005). The tortoise and the hare II: relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. *Am. J. Bot.* 92, 142–166. doi: 10.3732/ajb.92.1.142
- Shaw, J., Lickey, E. B., Schilling, E. E., and Small, R. L. (2007). Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in Angiosperms: the tortoise and the hare III. *Am. J. Bot.* 94, 75–288. doi: 10.3732/ajb.94.3.275
- Stamatakis, A. (2014). RAXML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30, 1312–1313. doi: 10.1093/bioinformatics/btu033
- Sun, Y. S., Abbott, R. J., Li, L. L., Li, L., Zou, J. B., and Liu, J. Q. (2014). Evolutionary history of purple cone spruce (*Picea purpurea*) in the Qinghai-Tibet Plateau: homoploid hybrid origin and Pleistocene expansion. *Mol. Ecol.* 23, 343–359. doi: 10.1111/mec.12599
- Swatdipong, A., Vasemägi, A., Koskinen, M. T., Piironen, J., and Primmer, C. R. (2009). Unanticipated population structure of European grayling in its northern distribution: implications for conservation prioritization. *Front. Zool.* 2009, 6:6. doi: 10.1186/1742-9994-6-6
- Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123, 585–595. doi: 10.1101/gad.3.11.1801
- Templeton, A. R., and Sing, C. F. (1993). A cladistics analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping. IV. Nested analyses with cladogram uncertainty and recombination. *Genetics* 134, 659–669.
- Thaw, S., White, T. A., Bannikova, A. A., and Searle, J. B. (2019). “Phylogeography,” in *Shrews, Chromosomes and Speciation*, eds J. Searle, P. Polly, and J. Zima, J. (Cambridge: Cambridge University Press), 112–133.
- Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F., and Higgins, D. G. (1997). The CLUSTAL X windows interface: flexible strategies for multiple sequences alignment aided by quality analysis tools. *Nucleic Acids Res.* 25, 4876–4882. doi: 10.1093/nar/25.24.4876
- Tribisch, A., and Schönswetter, P. (2003). Patterns of endemism and comparative phylogeography confirm palaeo-environmental evidence for Pleistocene refugia in the Eastern Alps. *Taxon* 52, 477–497. doi: 10.2307/3647447
- Vander Mijnsbrugge, K. (2014). “Continuity of local genetic diversity as an alternative to importing foreign provenances,” in *Genetic Considerations in Ecosystem Restoration Using Native Tree Species*. *State of the World's Forest Genetic Resources – Thematic Study*, eds M. Bozzano, R. Jalonen, E. Thomas, D. Boshier, L. Gallo, and S. Cavers (Rome: FAO and Bioversity International), 38–46.
- Vander Mijnsbrugge, K., De Cock, K., Cox, K., and Breynne, P. (2010). Conservation measures for *Rosa arvensis* Huds. in Flanders (Belgium) based on congruent genetic and phenotypic population differentiation. *Conserv. Genet.* 11, 2243–2253. doi: 10.1007/s10592-010-0109-5
- Volis, S. (2016). How to conserve threatened Chinese plant species with extremely small populations? *Plant Diver.* 1, 1–11. doi: 10.1016/j.pld.2016.05.003
- Volis, S. (2019). *Plant Conservation: The Role of Habitat Restoration*. Cambridge: Cambridge University Press.
- Wakasugi, T., Tsudzuki, J., Ito, S., Nakashima, K., Tsudzuki, T., and Sugiura, M. (1994). Loss of all *ndh* genes as determined by sequencing the entire chloroplast genome of the black pine *Pinus thunbergii*. *Proc. Natl. Acad. Sci. U. S. A.* 91, 9794–9798. doi: 10.1073/pnas.91.21.9794
- Wang, J., Gao, P., Kang, M., Lowe, A. J., and Huang, H. (2009). Refugia within refugia: the case study of a canopy tree (*Eurycorymbus cavaleriei*) in subtropical China. *J. Biogeogr.* 36, 2156–2164. doi: 10.1111/j.1365-2699.2009.02165.x
- Wang, W. T. (1992a). On some distribution patterns and some migration routes found in the eastern Asiatic region. *Acta Phytotaxon. Sin.* 30, 1–24. (in Chinese)
- Wang, W. T. (1992b). On some distribution patterns and some migration routes found in the eastern Asiatic region (Cont.). *Acta Phytotaxon. Sin.* 30, 97–117. (in Chinese)
- Wang, X., Liu, X. Q., Ko, Y. Z., Jin, X. L., Sun, J. H., and Zhao, Z. Y. (2020). Genetic diversity and phylogeography of the important medical herb, cultivated Huang-Lian populations, and the wild relatives *Coptis* Species in China. *Front. Genet.* 11, 708. doi: 10.3389/fgene.2020.00708
- White, T. J., Bruns, T. D., Lee, S. B., and Taylor, J. (1990). “Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics,” in *PCR Protocols: A Guide to Methods and Applications*, eds M. A. Innis, D. H. Gelfand, and J. J. Sninsky (New York, NY: Academic), 315–321.
- Wissemann, V. (2003). “Conventional taxonomy (wild roses),” in *Encyclopedia of Rose Science*, eds A. V. Roberts, T. Debener, and S. Gudín (Amsterdam: Elsevier), 111–117.
- Wright, S. (1931). Evolution in Mendelian populations. *Genetics* 16, 97–159.
- Wu, Z. Y. (1980). *Vegetation of China*. Beijing: Science Press.
- Wu, Z. Y., and Wu, S. G. (1996). “A proposal for a new floristic kingdom (realm): The E. Asiatic kingdom, its delineation and characteristics,” in *Floristic characteristics and diversity of East Asian plants*, eds A. L. Zhang, and S. G. Wu (New York, NY: Springer), 3–4.
- Wylie, A. (1954). The history of garden roses. *J. R. Hort. Soc.* 79, 555–571.
- Ye, J. Q. (2015). *Modern Practical Herb*. Beijing: China Press of Traditional Chinese Medicine, 129–130.
- Ye, J. W., Bai, W. N., Bao, L., Wang, T. M., Wang, H. F., and Ge, J. P. (2017a). Sharp genetic discontinuity in the aridity-sensitive species *Lindera obtusiloba* (Lauraceae): solid evidence supporting the Tertiary floral subdivision in East Asia. *J. Biogeogr.* 44, 2082–2095. doi: 10.1111/jbi.13020
- Ye, J. W., Zhang, Y., and Wang, X. J. (2017b). Phylogeographic breaks and the mechanisms of their formation in the Sino-Japanese floristic region. *Chin. J. Plant Ecol.* 41, 1003–1019. doi: 10.17521/cjpe.2016.0388
- Ying, T. S. (2001). Species diversity and distribution pattern of seed plants in China. *Biodiv. Sci.* 9, 393–398. doi: 10.17520/biods.2001058
- Young, E. L., Anderson, N. A., Bedard, C., Scheiber, M., and Byrne, D. H. (2019). *Rosa palustris* and *rosa setigera*: breeding challenges. *Acta Hort.* 1232, 41–46. doi: 10.17660/ActaHortic.2019.1232.7
- Yu, G., Chen, X., Ni, J., Cheddadi, R., Guiot, J., and Han, H. (2000). Palaeovegetation of China: a pollen data-based synthesis for the mid-Holocene and last glacial maximum. *J. Biogeogr.* 27, 635–664. doi: 10.1046/j.1365-2699.2000.00431.x
- Zhang, M. G., Slik, J. W. F., and Ma, K. P. (2016). Priority areas for the conservation of perennial plants in China. *Biol. Conserv.* 210, 56–63. doi: 10.1016/j.biocon.2016.06.007
- Zhao, L., Zhang, H., Wang, Q. G., Ma, C. L., and Jian, H. Y. (2019). The complete chloroplast genome of *Rosa lucidissima*, a critically endangered

- wild rose endemic to China. *Mitochondrial DNA B*. 4, 1826–1827. doi: 10.1080/23802359.2019.1613198
- Zhu, Z. M., Gao, X. F., and Fougère-Danezan, M. (2015). Phylogeny of *Rosa* sections *Chinenses* and *Synstylae* (Rosaceae) based on chloroplast and nuclear markers. *Mol. Phylogenet. Evol.* 87, 50–64. doi: 10.1016/j.ympev.2015.03.014
- Zlesak, D. C. (2015). “Rose – *Rosa* x hybrid,” in *Flower Breeding and Genetics: Issue, Challenges and Opportunities for the 21st Century*, ed N. O. Anderson (Dordrecht: Springer).

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher’s Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Jian, Zhao, Zhang, Ma, Wang, Yan, Qiu, Zhou and Zhang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.