



# Insights Into the Significance of the Chinese Loess Plateau for Preserving Biodiversity From the Phylogeography of *Speranskia tuberculata* (Euphorbiaceae)

Jun-Wei Ye<sup>1,2</sup>, Hai-Yang Wu<sup>1</sup>, Meng-Jiao Fu<sup>1</sup>, Pei Zhang<sup>1</sup> and Bin Tian<sup>1\*</sup>

<sup>1</sup> Key Laboratory for Forest Resources Conservation and Utilization in the Southwest Mountains of China, Ministry of Education, Southwest Forestry University, Kunming, China, <sup>2</sup> Germplasm Bank of Wild Species, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming, China

## OPEN ACCESS

### Edited by:

Gerald Matthias Schneeweiss,  
University of Vienna, Austria

### Reviewed by:

Hong-Hu Meng,  
Chinese Academy of Sciences, China  
Hongxiang Zhang,  
Chinese Academy of Sciences, China

### \*Correspondence:

Bin Tian  
tianbin@swfu.edu.cn

### Specialty section:

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

**Received:** 09 September 2020

**Accepted:** 12 January 2021

**Published:** 04 February 2021

### Citation:

Ye J-W, Wu H-Y, Fu M-J, Zhang P  
and Tian B (2021) Insights Into  
the Significance of the Chinese  
Loess Plateau for Preserving  
Biodiversity From the Phylogeography  
of *Speranskia tuberculata*  
(Euphorbiaceae).  
*Front. Plant Sci.* 12:604251.  
doi: 10.3389/fpls.2021.604251

The significance of the Chinese Loess Plateau (CLP) in maintaining biodiversity for northern China has rarely been shown, as previous phylogeographic studies are mostly woody species and they have revealed that Quaternary refugia are mainly located in mountain regions. We selected a drought-enduring endemic herb, *Speranskia tuberculata* (Euphorbiaceae), to determine its glacial refugia and postglacial demographic history. To this end, we sampled 423 individuals from 38 populations covering its entire geographic distribution. Three chloroplast DNA (cpDNA) fragments, two low-copy nuclear genes, and six nuclear microsatellites (nSSRs) were used and supplemented with ecological niche modeling (ENM) to infer the phylogeographic history of this species. Populations with private haplotypes and high haplotype diversity of cpDNA are mainly located in the CLP or scattered around northeastern China and the coastal region. Spatial expansion, detected using a neutrality test and mismatch distribution, may have resulted in a widely distributed ancestral cpDNA haplotype, especially outside of the CLP. For nuclear DNA, private haplotypes are also distributed mainly in the CLP. In nSSRs, STRUCTURE clustering identified two genetic clusters, which are distributed in the west (western cluster) and east (eastern cluster), respectively. Many populations belonged, with little to no admixture, to the western cluster while (hardly) pure populations of the eastern cluster were barely found. Genetic differentiation is significantly correlated with geographic distance, although genetic diversity is uniformly distributed. ENM suggests that the distribution of *S. tuberculata* has recently expanded northwards from the southern CLP, whereas it has experienced habitat loss in the south. Thus, *S. tuberculata* populations probably survived the last glacial maximum (LGM) in the southern CLP and experienced post-glacial expansion. Wind-dispersed pollen could bring the majority of genotypes to the front during spatial expansion, resulting in uniformly distributed genetic diversity. Based on evidence from molecular data and vegetation and climate changes since the LGM, we conclude that drought-enduring species, especially herbaceous species, are likely to have persisted in the CLP during the LGM and to have experienced expansion to other regions in northern China.

**Keywords:** Chinese Loess Plateau, refugia, nuclear microsatellites, Quaternary, *Speranskia tuberculata*

## INTRODUCTION

The Chinese Loess Plateau (CLP), which spans from 34 – 40°N and 101 – 113°E in northern China, contains heterogeneous topography and climate (Zhang et al., 2002) and is at the intersection of multiple floristic regions (Wang, 1999; Wu et al., 2010). A number of relict taxa and a diverse drought-enduring plants indicate that the CLP could have acted as a macrorefugium during Quaternary climate fluctuations (Harrison et al., 2001; Zhang et al., 2002; Ni et al., 2010). However, phylogeographic investigations of plants in northern China are limited (Qiu et al., 2011) as previous studies are focused on northwestern China (Meng et al., 2015), subtropical China (Ye et al., 2017a), northeastern China (Ye et al., 2017b) or southwestern China (Liu et al., 2017). Thus, the significance of the CLP for preserving the biodiversity of adjacent regions may be underestimated.

Paleovegetation reconstruction for northern China indicates that steppe and even desert vegetation covered this area during the last glacial maximum (LGM, ca. 26,000 – 18,000 years ago), while the present coniferous and deciduous forests experienced southward retreat to below 30°N (Harrison et al., 2001; Ni et al., 2010). Phylogeographic studies, however, challenged this hypothesis by discovering multiple LGM refugia, which have allowed species to persist across northern China (Chen et al., 2008; Tian et al., 2009; Zeng et al., 2011). In an endemic and dominant coniferous tree, *Pinus tabulaeformis*, Chen et al. (2008) suggested that there are multiple geographically isolated refugia based on restrictedly distributed mitotypes, although the plastid haplotypes were almost homogeneously distributed, which may be due to greater mobility of pollen relative to seed. Multiple localized refugia patterns were also discovered in a tree, *Quercus liaotungensis* (Zeng et al., 2011) and a shrub, *Ostryopsis davidiana* (Tian et al., 2009), as different plastid haplotypes were geographically restricted. Other studies on widely distributed East Asian species have shown that northern China is likely a contact zone between southern subtropical and northern cold temperate populations (Guo et al., 2014; Bai et al., 2016).

Previous studies have largely focused on woody plants (Chen et al., 2008; Tian et al., 2009; Zeng et al., 2011); therefore, these results may not be representative of the biome as a whole. Herbaceous species, especially drought-enduring plants that are mainly distributed in the CLP (Zhang et al., 2002) with much shorter generation times compared to woody species, may have responded differently to past climate changes (Wang et al., 2015). Second, only haplotypes of cytoplasmic markers are used to infer population demographic histories (Chen et al., 2008; Tian et al., 2009; Zeng et al., 2011), whereas the importance of pollen-mediated gene flow has been largely ignored except for *P. tabulaeformis* (Chen et al., 2008). As a result, many aspects of past population and range dynamics in northern China still need further exploration to provide insights into the preservation of biodiversity.

To assess the significance of the CLP for preserving biodiversity, *Speranskia tuberculata* (Bunge) Baillon (Euphorbiaceae), belonging to an East Asian endemic genus with two species only, was selected. This species is a drought-enduring endemic herb in northern China that is mainly distributed

in the CLP and adjacent areas, where it usually occurs in dry habitats such as xerothermophilous grasslands, sparse thickets, and forest edges. Its pollen is dispersed by wind, while seeds are dispersed by gravity (personal observations). Its generation time is about one year base on personal observations. Genetic variation estimated from chloroplast (cp) and low-copy nuclear (n) DNA as well as nuclear microsatellites (nSSRs) was combined with ecological niche modeling (ENM) to investigate (a) the population genetic structure of *S. tuberculata* and to test (b) whether any refugia were located in the CLP.

## MATERIALS AND METHODS

### Sampling and DNA Extraction

A total of 423 *S. tuberculata* individuals representing 38 populations were sampled through its entire geographic distribution (Table 1). We avoided collecting clones by sampling individuals more than 10 m apart. Genomic DNA was extracted from silica gel-dried leaves using the Ezup DNA Extraction Kit (Sangon Biotech, Shanghai, China). Voucher specimens were deposited in the herbarium SWFC (Southwest Forestry University), Kunming, China.

### Chloroplast and Nuclear DNA Sequence Analyses

Three chloroplast fragments, *psbJ-petA*, *psbB-psbF*, and *trnL-trnF*, and two species-specific low-copy nuclear genes (6146 and 38274) developed using transcriptome data (Fu et al., 2016) following the procedure described by Ye et al. (2017c) were used to detect potential genetic variation in *S. tuberculata* (Supplementary Table 1). The PCRs were performed following the protocol of Ma et al. (2019), except for different annealing temperatures (Supplementary Table 1), and the products were sequenced by Sangon Biotech (Shanghai, China).

cpDNA and nDNA chromatograms were read, edited, and aligned in MEGA 6.0 using the ClustalW algorithm (Tamura et al., 2013) and the alignments were carefully checked manually. For cpDNA, we deleted a region from *psbJ-petA* because it contained complex indels information and was too divergent to be aligned. The three plastid fragments were concatenated after the indels were recoded as single mutations. For nDNA, heterozygote sites were resolved using the coalescent-based Bayesian method of the phase function in DnaSP 5.10 (Librado and Rozas, 2009). Then, for both cpDNA and nDNA, the number of variable sites, number of haplotypes, haplotype diversity ( $H_d$ ), and nucleotide diversity ( $\pi$ ) were calculated in DnaSP. The most parsimonious networks for both plastid and nuclear data were inferred using Network 5.0 (Bandelt et al., 1999). Haplotype distributions were displayed in ArcGis 10.2 (ESRI, Inc.).

Potential population spatial expansion was detected using mismatch distribution analysis and a neutrality test, Tajima's  $D$  (Tajima, 1989) and Fu's  $F_s$  (Fu, 1997), with default settings for all *S. tuberculata* populations using cpDNA genetic variations in Arlequin 3.5.1.2 (Excoffier and Lischer, 2010). Neither the absolute expansion time nor a Bayesian skyline plot

**TABLE 1** | Details of sample location, sample size, and genetic diversity of chloroplast DNA and nuclear microsatellites in 38 *Speranskia tuberculata* populations.

Population	Location	Lat	Long	Ele (m)	cpDNA				Nuclear microsatellites						
					n	$H_d$	$\pi$ ( $\times 10^{-4}$ )	Haplotype (n)	n	$A_O$	$H_E$	$H_O$	$R_S$	$P_{AR}$	$F_{IS}$
AHQ	Aohanqi, Inner Mongolia	42.53	120.58	670	15	0.42	1.6	H1(11), H2(4)	15	28	0.57	0.56	2.97	0.00	0.06
BJ	Mentonggou, Beijing	39.98	116.03	400	10	0	0	H1(10)	10	28	0.54	0.43	3.29	0.07	0.25
CF	Wengniuteqi, Inner Mongolia	42.97	118.98	631	12	0	0	H1(12)	12	26	0.52	0.50	3.05	0.05	0.08
CFS	Chifeng, Inner Mongolia	42.42	118.92	1500	14	0	0	H1(14)	12	26	0.53	0.60	3.2	0.07	-0.08
DF	Dengfeng, Henan	34.52	112.93	650	16	0.49	3.3	H1(11), <b>H5(4), H6(1)</b>	16	22	0.45	0.42	2.6	0.02	0.11
DL	Dalian, Liaoning	38.92	121.62	130	10	0	0	<b>H26(10)</b>	-	-	-	-	-	-	-
FF	Fufeng, Shaanxi	34.28	107.90	530	8	0.71	4.9	H10(4), <b>H9(2), H21(2)</b>	6	23	0.59	0.47	-	-	0.31
GG	Gangu, Gansu	34.82	105.30	1325	5	0	0	H1(5)	5	12	0.32	0.37	-	-	-0.02
GH	Guanghe, Gansu	35.50	103.57	2178	17	0	0	<b>H12(17)</b>	17	25	0.54	0.46	2.96	0.04	0.19
GS	Huixian, Henan	35.57	113.52	980	14	0.58	5.2	H1(5), <b>H15(1), H16(8)</b>	14	27	0.62	0.63	3.41	0.00	0.02
GY	Guyuan, Ningxia	36.27	106.40	1927	3	0	0	H1(3)	3	15	0.44	0.42	-	-	0.29*
HC	Hancheng, Shaanxi	35.40	110.38	470	4	0	0	<b>H8(4)</b>	4	17	0.46	0.38	-	-	0.32*
HM	Houma, Shanxi	35.58	111.40	530	8	0.43	1.6	H1(6), H7(2)	8	24	0.63	0.49	3.46	0.00	0.30*
HT	Huhehaote, Inner Mongolia	40.87	111.60	1147	14	0	0	H1(14)	14	23	0.54	0.61	2.94	0.00	-0.10
HY	Heyang, Shaanxi	35.02	110.22	475	9	0.22	0.9	<b>H20(8), H1(1)</b>	9	19	0.52	0.59	2.7	0.00	-0.07
JB	Jingbian, Shaanxi	37.45	108.57	1689	1	0	0	H7(1)	-	-	-	-	-	-	-
JNS	Jinan, Shandong	36.62	117.07	250	14	0	0	H1(14)	14	21	0.55	0.39	2.74	0.03	0.32*
KLQ	Kulunqi, Inner Mongolia	42.75	121.38	400	8	0	0	H1(8)	8	25	0.59	0.69	3.26	0.08	-0.10
KQ	Keqi, Inner Mongolia	43.27	117.55	2067	13	0	0	H1(13)	13	21	0.47	0.39	2.76	0.09	0.22
LL	Luliang, Shanxi	37.12	111.15	1587	13	0	0	H7(13)	13	26	0.60	0.46	3.2	0.10	0.26
LX	Longxian, Gansu	34.92	104.68	1850	14	0.14	0.5	<b>H11(1), H7(13)</b>	13	22	0.55	0.47	2.9	0.00	0.19
MZ	Mizhi, Shaanxi	37.78	110.12	1050	16	0.23	0.9	H7(14), <b>H17(2)</b>	16	29	0.65	0.76	3.37	0.00	-0.13
PC	Pucheng, Shaanxi	34.95	109.58	488	13	0	0	H1(13)	13	21	0.55	0.36	2.9	0.03	0.40
QD	Qingdao, Shandong	36.13	120.72	170	10	0	0	H1(10)	12	20	0.53	0.38	2.81	0.00	0.32
QY	Qingyang, Gansu	35.72	107.65	1408	11	0.18	0.7	H1(10), H7(1)	11	27	0.61	0.65	3.31	0.00	-0.01
QYS	Qinyang, Henan	35.22	112.80	787	12	0.3	4.7	H1(10), <b>H18(2)</b>	-	-	-	-	-	-	-
SH	Shahe, Hebei	36.88	114.15	339	9	0	0	H1(9)	9	22	0.51	0.46	2.87	0.00	0.16
SJZ	Shijiazhuang, Hebei	38.07	114.27	200	15	0.13	0.5	H1(14), H3(1)	15	28	0.64	0.55	3.34	0.00	0.17
SLQ	Salaqi, Inner Mongolia	40.60	110.58	1200	8	0	0	H1(8)	8	23	0.48	0.54	2.97	0.13	-0.07
TB	Heshui, Gansu	36.13	108.65	1166	12	0.49	3.7	H10(8), H7(4)	12	27	0.65	0.50	3.54	0.03	0.28*
TC	Tongchuan, Shaanxi	34.92	108.98	623	14	0.26	1	H10(12), <b>H19(2)</b>	14	30	0.69	0.53	3.71	0.04	0.27*
WC	Wenchuan, Sichuan	31.48	103.58	1390	10	0.47	1.8	<b>H13(3), H14(7)</b>	10	12	0.26	0.26	1.77	0.00	0.04
XX	Xiaoxian, Anhui	34.18	116.90	179	17	0.22	0.8	H1(15), <b>H4(2)</b>	17	31	0.56	0.57	3.26	0.13	0.02
XZ	Xinzhou, Shanxi	39.32	113.58	1160	14	0	0	H3(14)	14	28	0.57	0.58	3.33	0.14	0.02
YA	Yanan, Shaanxi	36.58	109.48	1061	13	0.39	10.4	H1(3), <b>H22(10)</b>	13	33	0.63	0.62	3.7	0.06	0.07
YJ	Yongji, Shanxi	34.87	110.45	365	9	0.50	5.8	H1(6), <b>H23(3)</b>	9	21	0.61	0.61	2.96	0.00	0.05
YT	Yantai, Shandong	37.28	121.73	120	15	0	0	<b>H24(15)</b>	15	24	0.55	0.48	2.97	0.00	0.17
ZS	Zishan, Shanxi	35.60	110.97	1300	3	0	0	<b>H25(3)</b>	3	8	0.12	0.17	-	-	-0.20

Lat, latitude; Long, longitude; Ele, elevation;  $H_d$ , haplotype diversity;  $A_O$ , total observed allele number of six loci;  $H_O$ , observed heterozygosity;  $H_E$ , expected heterozygosity;  $R_S$ , allelic richness;  $P_{AR}$ , private allelic richness;  $F_{IS}$ , inbreeding coefficient.

\*significant deviation from zero; -, not genotyped or calculated; haplotypes that are private to only one population are in bold.

(BSP) (Heled and Drummond, 2008) that can detect changes in effective population size through time were estimated, as no reasonable substitution rate of cpDNA is provided in *S. tuberculata*.

## Nuclear Microsatellites Analyses

Six of the 18 nSSR markers developed for *Speranskia* species (Fu et al., 2016), which produced reproducible, scorable, and

polymorphic products, were used to detect genetic variation in *S. tuberculata* (Table 2). The six nSSRs were amplified (Fu et al., 2016) and then genotyped using a 3730xl automated Genetic Analyzer (Applied Biosystems, Foster City, CA, United States). Allele sizes were determined using GENEMAPPER 3.7 (Applied Biosystems). Null alleles were checked using Micro-checker 2.2 (Oosterhout et al., 2004). Linkage disequilibrium was tested using sequential Bonferroni corrections, and the deviation of  $F_{IS}$

**TABLE 2** | Measures of genetic diversity and genetic differentiation at six nuclear microsatellite loci in *Speranskia tuberculata*.

Locus	$A_O$	$H_O$	$H_S$	$H_T$	$F_{ST}$
9441	9	0.476	0.505	0.731	0.308
25439	9	0.719	0.623	0.745	0.164
26221	14	0.461	0.735	0.855	0.141
10809	9	0.493	0.568	0.679	0.163
16226	6	0.336	0.352	0.497	0.292
24490	10	0.489	0.645	0.832	0.224
Mean	9.5	0.496	0.571	0.723	0.215

$A_O$ , observed allele number;  $H_O$ , observed heterozygosity;  $H_S$ , gene diversity;  $H_T$ , overall gene diversity;  $F_{ST}$ , among-population differentiation.

(fixation index) from zero was used to test for Hardy-Weinberg equilibrium using FSTAT 2.9.3 (Goudet, 2001).

For all six loci, the genetic diversity index of observed number of alleles ( $A_O$ ), observed heterozygosity ( $H_O$ ), gene diversity ( $H_S$ ), overall gene diversity ( $H_T$ ), and among-population differentiation ( $F_{ST}$ ) were calculated using FSTAT. For all populations,  $A_O$ ,  $H_O$ , and expected heterozygosity ( $H_E$ ) were calculated in GenALEX 6.5 (Smouse et al., 2017). The  $R_S$  (allele richness) and  $P_{AR}$  (private allele richness) in populations with more than eight samples were calculated in hp-rare 1.0 using rarefaction with a sample number of eight (Kalinowski, 2005). To test isolation by distance (IBD), a Mantel test with 10,000 random permutations was performed between the matrix of pairwise  $F_{ST}$  and that of the natural logarithm of the geographic distances using GenALEX.

The genetic structure was inferred using the Bayesian clustering method in STRUCTURE 2.3 (Falush et al., 2007) with an admixture model and assuming allele frequencies to be correlated among populations. Of the 38 populations sampled, populations DL, JB, and QYS were excluded due to limited sample size or a high proportion of missing data. Ten independent runs were performed for each  $K$  from 1 to 20 with  $1 \times 10^5$  burn-in cycles followed by  $1 \times 10^6$  MCMC cycles. Potential clusters ( $K$ ) were determined by  $LnP(D)$ , the change in log-likelihood of the data for each run (Pritchard et al., 2000), and  $\Delta K$ , the second-order rate of change of  $LnP(D)$  between successive  $K$ -values (Evanno et al., 2005).

## Ecological Niche Modeling

To explore the historical distribution range shifts of *S. tuberculata*, its potential distribution at present and during the LGM were modeled using MAXENT (Phillips et al., 2006). A total of 94 occurrence points were collected from our field records (35) and herbarium records (59). Eight low-correlated ( $r < 0.75$ ) climate variables (Supplementary Table 2, available at<sup>1</sup>) with 2.5 arc-min resolution were used to model the niche at present (Hijmans et al., 2005). The established model was projected onto the LGM climatic conditions reconstructed using the Community Climate System Model (CCSM) (Collins et al., 2006) at 2.5 arc-min resolution. We used 80% of the species records for training and 20% for model testing with 20 replicates.

<sup>1</sup> <http://www.worldclim.org/>

The accuracy of the model's performance was evaluated based on the area under the receiver operating characteristic curve (AUC) (Fawcett, 2006).

## RESULTS

### cpDNA Analyses

Overall, 34 variable sites including 15 indels were detected in 423 *S. tuberculata* individuals, and 26 haplotypes were identified with a total of 2,600 bp using three chloroplast fragments (Supplementary Table 3). The most parsimonious network showed that most haplotypes were closely related, except H22 and H24. The most abundant and centrally located haplotype H1 showed the closest relationship with outgroup *S. cantonensis* (Figure 1a). Haplotype H1 is widely distributed across the entire range of *S. tuberculata*, especially outside of the CLP (Figure 1b). Two distantly related haplotypes (H15 and H16) were both distributed in population GS, and three relatively closely related haplotypes (H22–H24) showed disjunct distributions between the CLP (YA and YJ population) and coastal regions (YT population) (Figure 1b). Populations that have private haplotypes were mainly distributed in the CLP except two found in the coastal region (DL and YT population) (Figure 1b). Populations with high haplotype diversity were mainly located in the southern CLP and adjacent regions except the AHQ population in northeastern China (Figure 2a).

Both mismatch analysis (sum of squares deviation = 0.01,  $P = 0.327$ , Harpending's Raggedness index = 0.04,  $P = 0.845$ , Supplementary Figure 1) and the neutrality tests (Tajima's  $D = -1.72$ ,  $P = 0.01$ , Fu's  $F_S = -12.29$ ,  $P = 0$ ) indicated significant spatial expansion for all populations.

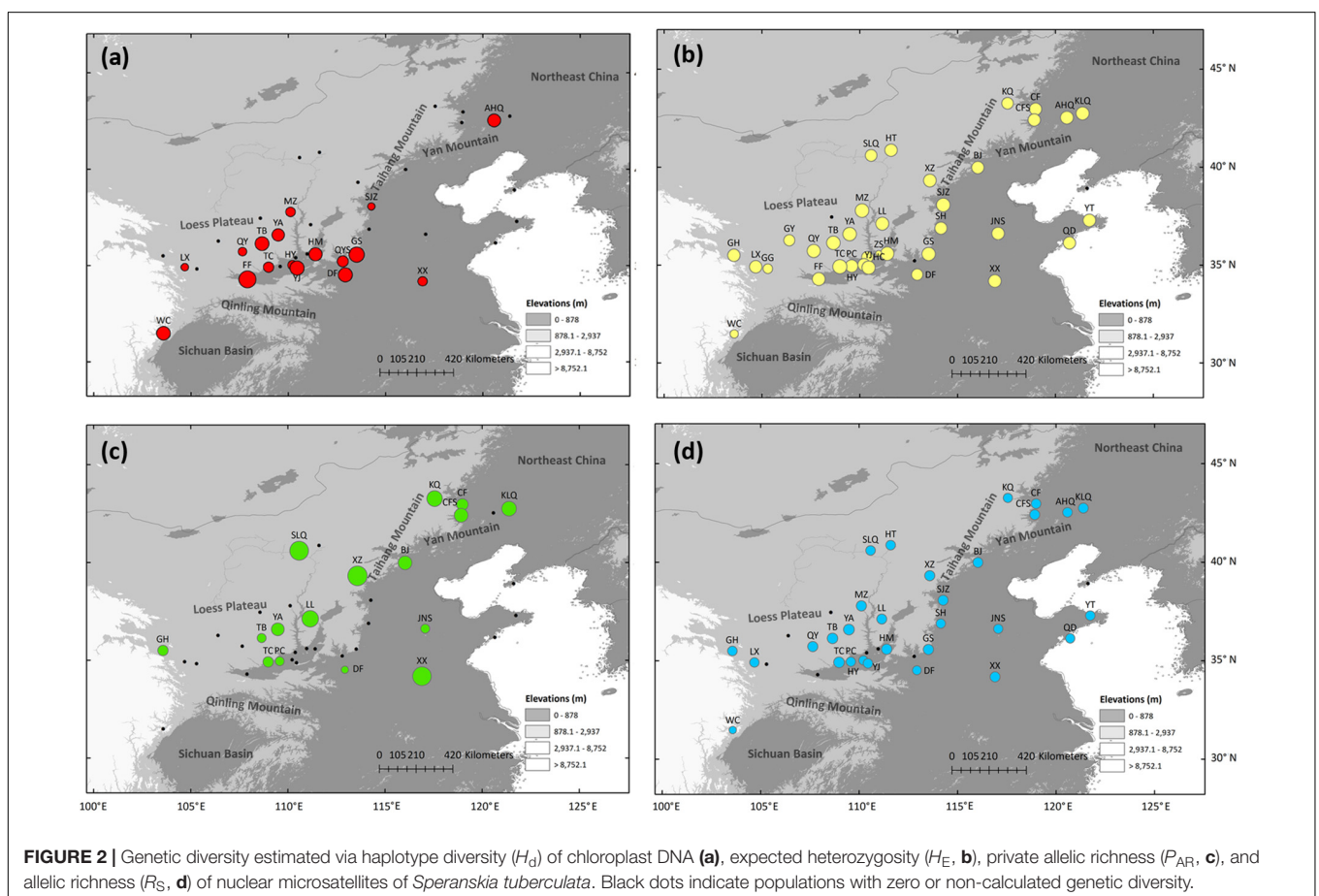
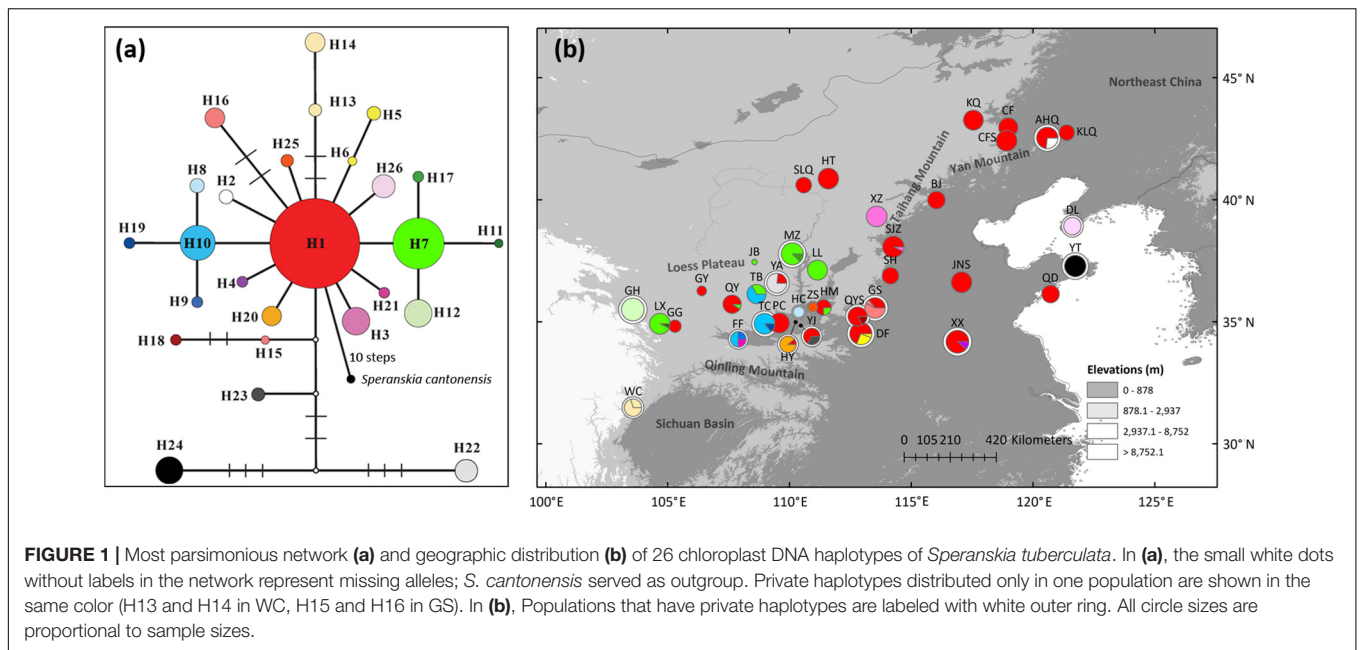
### nDNA Analyses

Low genetic variation was found in both nuclear loci (6146 and 38274, Supplementary Table 4). In locus 38274, eight variable sites in 405 bp of 181 individuals resulted in nine haplotypes. Haplotypes H2–H9 were all one mutation away from the centrally located haplotype H1 (Figure 3a). Private haplotypes were distributed in populations south of the Qinling Mountains (WC population), the southern CLP (GH, FF, TC, LL, and YJ population), and coastal regions (DL population) (Figure 3b). In locus 6146, six haplotypes were detected through five variable sites along 294 bp in 184 individuals. H1 was separated from H2–H5 by only one mutation site (Figure 3c). H1 was widely distributed, and three populations that have private haplotypes (H3 in GH, H2 in FF, and H4 in the HM population) were distributed in the southern CLP (Figure 3d).

### nSSRs Analyses

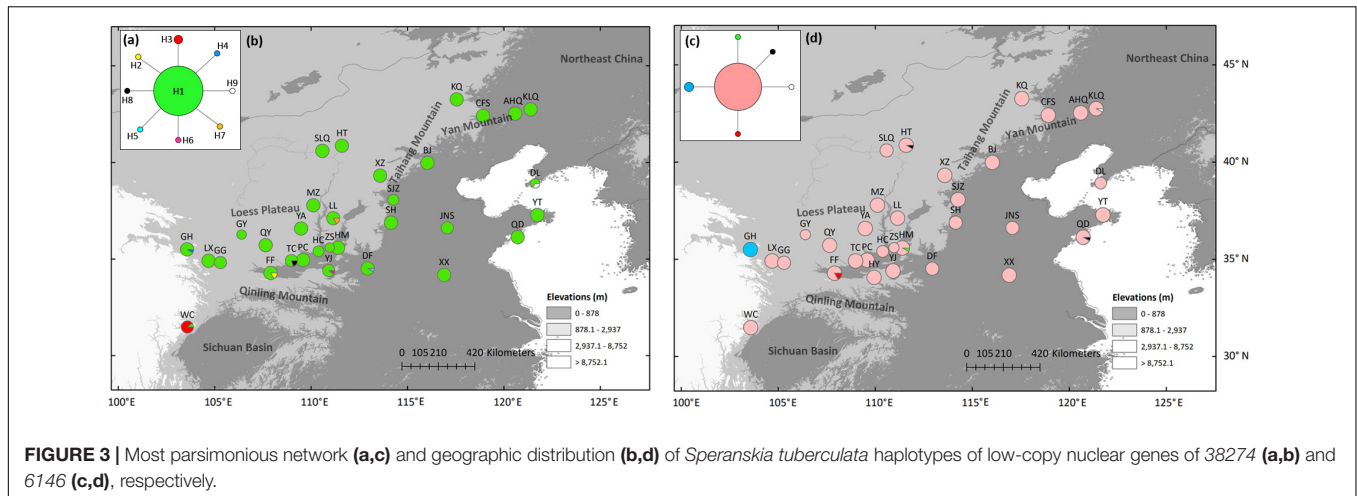
Null alleles were not found in the six nSSR loci, and no significant genotypic disequilibrium was observed among the 15 locus pairs in any populations. Several populations had a fixation index ( $F_{is}$ ) that is significantly deviated from zero, suggesting that these populations are not under Hardy-Weinberg equilibrium (Table 1). The genetic diversity indices of the 35





sampled populations and six loci are listed in **Tables 1** and **2**, respectively. Within-population genetic diversities,  $H_e$ ,  $P_{AR}$ , and  $R_S$ , are all uniformly distributed (**Figures 2b–d**). A Mantel

test ( $r = 0.171$ ,  $P = 0.037$ ) indicated a significant correlation between genetic differentiation and geographic distance among all sampled populations.



**FIGURE 3** | Most parsimonious network (a,c) and geographic distribution (b,d) of *Speranskia tuberculata* haplotypes of low-copy nuclear genes of 38274 (a,b) and 6146 (c,d), respectively.

Using STRUCTURE, the highest  $\Delta K$  is observed at  $K = 2$ , while the  $LnP(D)$  gradually increased from  $K = 2$  to  $K = 11$  (Supplementary Figure 2). We chose  $K = 2$  as the most possible clustering (Figure 4a), and the histogram of the STRUCTURE assignment test for all populations from  $K = 2$  is shown in Figure 4b. The western green cluster is mainly distributed in the CLP and adjacent regions, while the eastern red cluster is in the eastern part of the present *S. tuberculata* distribution range (Figure 4a). Pure western cluster populations were observed in many populations, while pure eastern populations were barely found.

## Ecological Niche Modeling

High ROC values ( $0.968 \pm 0.012$ ) indicate good accuracy of model predictions. At present, *S. tuberculata* is distributed mainly in the CLP and adjacent regions with extensions from Taihang-Yan mountains to northeastern China and from the North China Plain to the coastal region (Figure 5a). During the LGM, the most suitable habitat is predicted to have been located in the southern CLP and the Qinling Mountains with extension to the Sichuan Basin (Figure 5b). Comparison of absolute changes in habitat suitability showed that the northern distribution of *S. tuberculata* was reached after the LGM from the southern CLP and that the species has experienced habitat loss in the Sichuan Basin (Figure 5c).

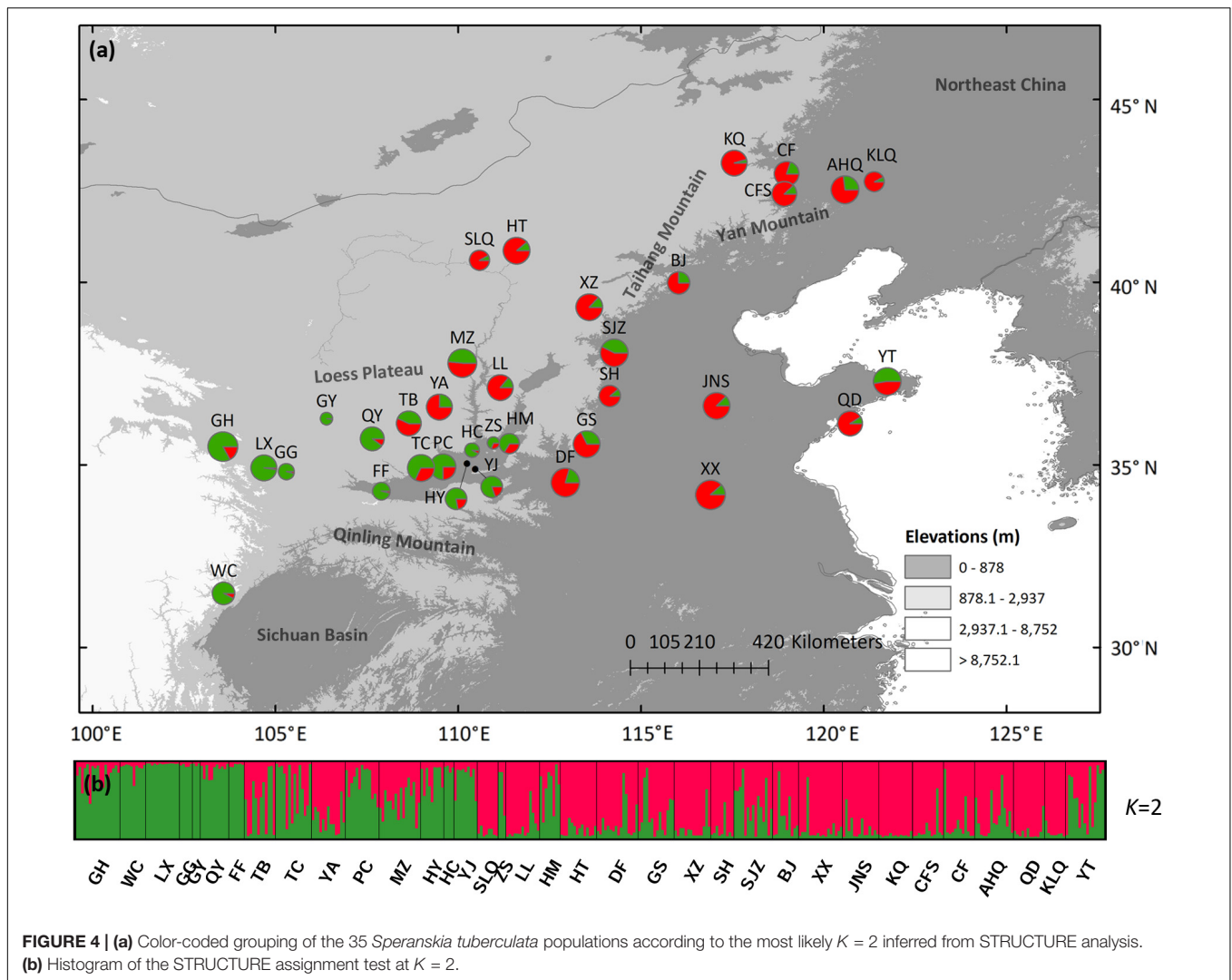
## DISCUSSION

### Southern Macrorefugium and Multiple Distant Microrefugia

In the present study, the phylogeography history of *S. tuberculata* was investigated using multiple genetic markers and potential habitat modeling. Chloroplast and nuclear DNA as well as ENM consistently support the hypothesis that *S. tuberculata* survived the unsuitable climate in the macrorefugium in the southern CLP and in multiple distant microrefugia. In contrast, nSSRs are ambiguous in inferring refugia and postglacial demographic history.

In chloroplast and nuclear sequences, populations with high haplotype diversity and private haplotypes are mainly distributed in the southern CLP, indicating that the macrorefugium of *S. tuberculata* is located there (Hewitt, 1996, 2004). Disjointly distributed closely related or co-distributed distant chloroplast haplotypes are evidence of population loss during past climate changes (Qiu et al., 2011), likely due to the cold and dry climate during the LGM. The mismatch distribution and the neutrality tests of chloroplast sequences support spatial expansion, which results in a wide distribution of haplotype H1 (Figure 1b). ENM also shows that most of the present *S. tuberculata* distribution north of the Qinling Mountains has been gained after the LGM (Figure 5). Post-glacial colonizations to the eastern part of the present distribution are likely through the Taihang-Yan Mountains and the North China Plain. Private chloroplast or nuclear haplotypes in northeastern China (KLQ and AHQ population) or in coastal regions (DL and YT population) may indicate multiple restricted microrefugia (Provan and Bennett, 2008). Therefore, *S. tuberculata* is likely to have persisted in the southern CLP macrorefugium and in multiple microrefugia in northeastern China and the coastal region (Ye et al., 2017b).

Seemingly, the nSSRs data tell a different story as two different scenarios can both generate genetic patterns displayed by nSSRs. First, expansions from an eastern and a western macrorefugium can result in genetically mixed populations in the central contact zone (Figure 4), similar to some widely distributed species (e.g., *Acer mono* and *Juglans* spp.) that show a genetic mixture in northern China (Guo et al., 2014; Bai et al., 2016). Such populations in contact zones are expected to have higher genetic diversity (Petit et al., 2003). However, this is not the case for *S. tuberculata* (Figures 3b–d), additionally pure populations belonging to the eastern cluster are barely found. Besides, uniformly distributed genetic diversity is probably attributed to massive gene flow via wind-dispersed pollen, which can reduce the surfing effect during spatial expansion and bring the majority of genetic composition to the expansion front (Wang et al., 2016). This scenario could explain that genetic structure changes gradually from west to east in agreement with a significant IBS



**FIGURE 4 |** (a) Color-coded grouping of the 35 *Speranskia tuberculata* populations according to the most likely  $K = 2$  inferred from STRUCTURE analysis. (b) Histogram of the STRUCTURE assignment test at  $K = 2$ .

pattern (Figure 4). Thus, nSSR data are compatible with the southern macrorefugium scenario.

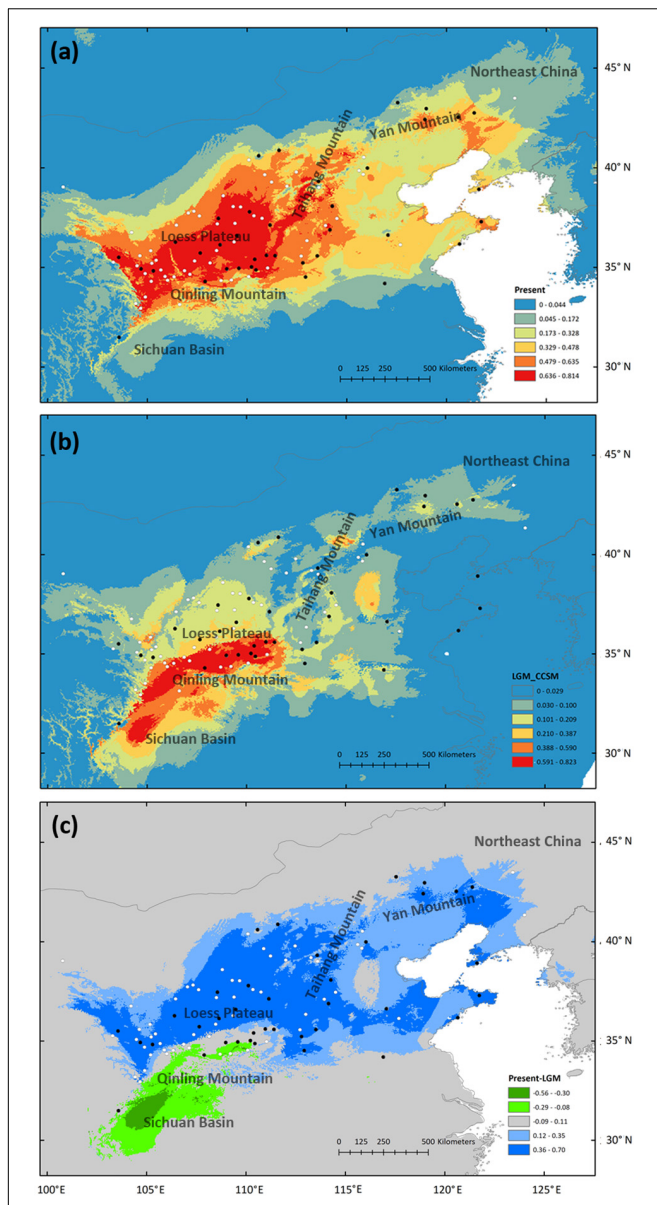
## The Significance of the CLP in Preserving Biodiversity

In previous phylogeographic studies in northern China, detected refugia were mainly located in mountain regions, such as the Qinling Mountains (Tian et al., 2009) or the Yan-Taihang Mountains (Zeng et al., 2011), but not in the CLP. In some other species mainly distributed in arid northwestern China, the CLP acts as a dispersal corridor for colonization (*Lagochilus ilicifolius*, Meng and Zhang, 2011; *Ribes meyeri*, Xie and Zhang, 2013). In this study, using the drought-enduring herb *S. tuberculata* as a case, multiple genetic data and potential habitat modeling uncovered the role of the CLP as a macrorefugium, indicating its importance in preserving biodiversity in northern China.

The CLP is mainly influenced by the Asian winter monsoon, which brings dust from the deserts of the Asian interior. The initial eolian dust deposits could date back to the Eocene-Oligocene (Licht et al., 2014; Clift et al., 2015).

Such deposits experienced several intensifications up to the Quaternary (Wang et al., 2007). Precipitation, which is higher in the southeast and lower in the northwest, has a direct influence on the vegetation of the CLP as it is located in the dry interior of Asia (Zhang et al., 2012), and this gradient was also present in the interglacial periods (Jiang and Ding, 2005). During the LGM, as the CLP is located in the marginal zone of the Asian summer monsoon, the climate is prone to be the same compared to prominent north-south and west-east gradients (warmer and wetter conditions in the south/east) during the interglacial period (Jiang et al., 2014). Higher precipitation is needed to develop wood forest while lower precipitation is needed for grass forest (Lv et al., 1999), and herbaceous species are thought to be more likely to find refugia in the CLP than tree species that survived the LGM in mountain regions (Tian et al., 2009; Zeng et al., 2011). In the CLP flora, many genera that are adapted to arid environments (Wang, 1999), such as *Caragana*, *Ephedra*, and *Tamarix*, are highly diversified, and majority of the CLP endemic species show tolerance to aridity (Zhang et al., 2002). In addition to the herb *S. tuberculata*, other drought-enduring species are also likely to have persisted





**FIGURE 5 |** Potential species distribution of *Speranskia tuberculata* estimated by ecological niche modeling at present (a) and during the last glacial maximum (LGM) using the CCSM3 model (b). Absolute changes in habitat suitability are compared between present and LGM (c); Blue represents suitability gains and green represents losses in suitability. Black dots represent locations of sampled populations, while white dots represent occurrence points collected from herbarium records.

in the CLP macrorefugium and to have experienced post-glacial expansions.

The CLP contributes not only to the biodiversity of northern China, but also to other regions in East Asia. First, Meng et al. (2015) have suggested the CLP to be an important refugium for the dry-adapted flora of northwestern China. Furthermore, disjoint distribution of same haplotype between CLP and northeastern China may result from long distance dispersal from the CLP (Tian et al., 2009; Song et al., 2020).

Finally, haplotypes of *S. tuberculata* in the Sichuan Basin (H13–14) are derived from the ancestral haplotype (H1), indicating the CLP may be the source of genotypes in southwestern China.

## CONCLUSION

The present phylogeographic work is an important supplement to previous studies in East Asia, which are much more numerous for northwestern China (Meng et al., 2015), subtropical China (Ye et al., 2017a), northeastern China (Ye et al., 2017b) or southwestern China (Liu et al., 2017). However, further phylogeographic studies on various species are still needed to depict a comprehensive response of the flora from northern China to Quaternary climate changes (Zhang et al., 2002).

## 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

J-WY and BT conceived the ideas. H-YW, M-JF, and PZ contributed to the sampling and the data analyses. J-WY and BT wrote the manuscript. All authors have read and approved the final manuscript.

## FUNDING

This study was supported by the National Natural Science Foundation of China (NSFC) (31260050 and 41861008) and the Ten-thousand Talents Program of Yunnan Province (YNWR-QNBJ-2020).

## ACKNOWLEDGMENTS

We thank Mr. Yi Fu for help during the field survey. We are grateful to Dr. Kang-Shan Mao for very helpful comments on earlier drafts of this manuscript.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2021.604251/full#supplementary-material>

**Supplementary Figure 1 |** Mismatch distribution of simulation under spatial expansion assuming constant deme size for all *Speranskia tuberculata* populations.

**Supplementary Figure 2 |** Delta- $K$  (A) and  $\text{Ln}P(D)$  (B) in STRUCTURE analysis on all *Speranskia tuberculata* populations are shown for predefined group number  $K = 1-20$ . Standard deviations of  $\text{Ln}P(D)$  obtained from 10 independent runs for each predefined group size are also shown.



## REFERENCES

- Bai, W. N., Wang, W. T., and Zhang, D. Y. (2016). Phylogeographic breaks within Asian butternuts indicate the existence of a phylogeographic divide in East Asia. *New Phytol.* 209, 1757–1772. doi: 10.1111/nph.13711
- Bandelt, H. J., Forster, P., and Röhl, A. (1999). Median-joining networks for inferring intraspecific phylogenies. *Mol. Biol. Evol.* 16, 37–48. doi: 10.1093/oxfordjournals.molbev.a026036
- Chen, K. M., Abbott, R. J., Milne, R. I., Tian, X. M., and Liu, J. Q. (2008). Phylogeography of *Pinus tabulaeformis* Carr. (Pinaceae), a dominant species of coniferous forest in northern China. *Mol. Ecol.* 17, 4276–4288. doi: 10.1111/j.1365-294X.2008.03911.x
- Clift, P. D., Wan, S., and Blusztajn, J. (2015). Reconstructing chemical weathering, physical erosion and monsoon intensity since 25 Ma in the northern South China Sea: a review of competing proxies. *Earth Sci. Rev.* 130, 86–102. doi: 10.1016/j.earscirev.2014.01.002
- Collins, W. D., Bitz, C. M., Blackmon, M. L., Bonan, G. B., Bretherton, C. S., Carton, J. A., et al. (2006). The community climate system model version 3 (CCSM3). *J. Climate* 19:2122. doi: 10.1175/JCLI3761.1
- Evanno, G., Regnaut, S., and Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol. Ecol.* 14, 2611–2620. doi: 10.1111/j.1365-294X.2005.02553.x
- Excoffier, L., and Lischer, H. E. L. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.* 10, 564–567. doi: 10.1111/j.1755-0998.2010.02847.x
- Falush, D., Stephens, M., and Pritchard, J. K. (2007). Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Mol. Ecol. Notes* 7, 574–578. doi: 10.1111/j.1471-8286.2007.01758.x
- Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recogn. Lett.* 27, 861–874. doi: 10.1016/j.patrec.2005.10.010
- Fu, Y., Ju, M. M., Ma, H. C., Xin, P. Y., He, C. Z., Jia, D. R., et al. (2016). Development and characterization of novel EST-SSR markers for *Speranskia tuberculata* (Euphorbiaceae). *Appl. Plant Sci.* 4:1600067. doi: 10.3732/apps.1600067
- Fu, Y. X. (1997). Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* 147, 915–925. doi: 10.1093/genetics/147.2.915
- Goudet, J. (2001). *FSTAT, A Program To Estimate and Test Gene Diversities And Fixation Indices (version 2.9.3)*. (accessed Sep 1, 2014). Available online at: <https://www2.unil.ch/popgen/softwares/fstat.htm>
- Guo, X. D., Wang, H. F., Bao, L., Wang, T. M., Bai, W. N., Ye, J. W., et al. (2014). Evolutionary history of a widespread tree species *Acer mono* in East Asia. *Ecol. Evol.* 4, 4332–4345. doi: 10.1002/ece3.1278
- Harrison, S. P., Yu, G., Takahara, H., and Prentice, I. C. (2001). Palaeovegetation (Communications arising): diversity of temperate plants in East Asia. *Nature* 413, 129–130. doi: 10.1038/35093166
- Heled, J., and Drummond, A. J. (2008). Bayesian inference of population size history from multiple loci. *BMC Evol. Biol.* 8:289. doi: 10.1186/1471-2148-8-289
- Hewitt, G. M. (1996). Some genetic consequences of ice ages, and their role in divergence and speciation. *Biol. J. Linn. Soc.* 58, 247–276. doi: 10.1006/bijl.1996.0035
- Hewitt, G. M. (2004). Genetic consequences of climatic oscillations in the Quaternary. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 359, 183–195. doi: 10.1098/rstb.2003.1388
- Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., and Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *Int. J. Climatol.* 25, 1965–1978. doi: 10.1002/joc.1276
- Jiang, H. C., and Ding, Z. L. (2005). Temporal and spatial changes of vegetation cover on the Chinese Loess Plateau through the last glacial cycle: evidence from spore-pollen records. *Rev. Palaeobot. Palynol.* 133, 23–37. doi: 10.1016/j.revpalbo.2004.08.003
- Jiang, W. Y., Cheng, Y. F., and Yang, X. X. (2014). Spatial patterns of vegetation and climate on the Chinese Loess Plateau since the last glacial maximum. *Quat. Int.* 33, 52–60. doi: 10.1016/j.quaint.2013.10.039
- Kalinowski, S. T. (2005). hp-rare 1.0: a computer program for performing rarefaction on measures of allelic richness. *Mol. Ecol. Notes* 5, 187–189. doi: 10.1111/j.1471-8286.2004.00845.x
- Librado, P., and Rozas, J. (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25, 1451–1452. doi: 10.1093/bioinformatics/btp187
- Licht, A., Van Cappelle, M., Abels, H. A., Ladant, J. B., Trabuco-Alexandre, J., France-Lanord, C., et al. (2014). Asian monsoons in a late Eocene greenhouse world. *Nature* 513, 501–506. doi: 10.1038/nature13704
- Liu, J., Luo, Y. H., Li, D. Z., and Gao, L. M. (2017). Evolution and maintenance mechanisms of plant diversity in the Qinghai-Tibet Plateau and adjacent regions: retrospect and prospect. *Biodiv. Sci.* 25, 163–174. doi: 10.17520/biods.2016293
- Lv, H. Y., Liu, T. S., Wu, N. Q., Han, J. M., and Guo, Z. T. (1999). Phytolith record of vegetation succession in the southern Loess Plateau since late Pleistocene. *Quat. Sci.* 4, 336–349.
- Ma, X. G., Wang, Z. W., Tian, B., and Sun, H. (2019). Phylogeographic analyses of the east asian endemic genus *Prinsepia* and the role of the East Asian monsoon system in shaping a north-south divergence pattern in China. *Front. Genet.* 10:128. doi: 10.3389/fgene.2019.00128
- Meng, H. H., Gao, X. Y., Huang, J. F., and Zhang, M. L. (2015). Plant phylogeography in arid Northwest China: retrospectives and perspectives. *J. Syst. Evol.* 53, 33–46. doi: 10.1111/jse.12088
- Meng, H. H., and Zhang, M. L. (2011). Phylogeography of *Lagochilus ilicifolius* (Lamiaceae) in relation to Quaternary climatic oscillation and aridification in northern China. *Biochem. Syst. Ecol.* 39, 787–796. doi: 10.1016/j.bse.2011.07.015
- Ni, J., Yu, G., Harrison, S. P., and Prentice, I. C. (2010). Palaeovegetation in China during the late Quaternary: biome reconstructions based on a global scheme of plant functional types. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 289, 44–61. doi: 10.1016/j.palaeo.2010.02.008
- Oosterhout, C. V., Hutchinson, W. F., Wills, D. P. M., and Shipley, P. (2004). micro-checker: software for identifying and correcting genotyping errors in microsatellite data. *Mol. Ecol. Resour.* 4, 535–538. doi: 10.1111/j.1471-8286.2004.00684.x
- Petit, R., Aguinagalde, I., De Beaulieu, J. L., Bittkau, C., Brewer, S., Cheddadi, R., et al. (2003). Glacial refugia: hotspots but not melting pots of genetic diversity. *Science* 300, 1563–1565. doi: 10.1126/science.1083264
- Phillips, S. J., Anderson, R. P., and Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecol. Model.* 190, 231–259. doi: 10.1016/j.ecolmodel.2005.03.026
- Pritchard, J. K., Stephens, M., and Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics* 155, 945–959.
- Provan, J., and Bennett, K. D. (2008). Phylogeographic insights into cryptic glacial refugia. *Trends Ecol. Evol.* 23, 564–571. doi: 10.1016/j.tree.2008.06.010
- 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
- Smouse, P. E., Banks, S. C., and Peakall, R. (2017). Converting quadratic entropy to diversity: both animals and alleles are diverse, but some are more diverse than others. *PLoS One* 12:e0185499. doi: 10.1371/journal.pone.0185499
- Song, X. Y., Milne, R. I., Fan, X. Y., Xie, S. Y., Zhang, L., Zheng, H. L., et al. (2020). Blow to the Northeast? Intraspecific differentiation of *Populus davidiana* suggests a north-eastward skew of a phylogeographic break in East Asia. *J. Biogeogr.* 48, 187–201. doi: 10.1111/jbi.13992
- Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123, 585–595. doi: 10.1093/genetics/123.3.585
- Tamura, K., Stecher, G., Peterson, D., Filipiński, A., and Kumar, S. (2013). MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* 30, 2725–2729. doi: 10.1093/molbev/mst197
- Tian, B., Liu, R. R., Wang, L. Y., Qiu, Q., Chen, K. M., and Liu, J. Q. (2009). Phylogeographic analyses suggest that a deciduous species (*Ostryopsis davidiana* Decne., Betulaceae) survived in northern China during the last glacial maximum. *J. Biogeogr.* 36, 2148–2155. doi: 10.1111/j.1365-2699.2009.02157.x
- Wang, H. S. (1999). The evolution and sources of North China flora. *Acta Geogr. Sin.* 54, 213–223.
- Wang, W. T., Xu, B., Zhang, D. Y., and Bai, W. N. (2016). Phylogeography of postglacial range expansion in *Juglans mandshurica* (Juglandaceae) reveals no evidence of bottleneck, loss of genetic diversity, or isolation by distance in the

- leading-edge populations. *Mol. Phylogenet. Evol.* 102, 255–264. doi: 10.1016/j.ympev.2016.06.005
- Wang, Y. H., Jiang, W. M., Comes, H. P., Hu, F. S., Qiu, Y. X., and Fu, C. X. (2015). Molecular phylogeography and ecological niche modelling of a widespread herbaceous climber, *Tetragymma hemsleyanum* (Vitaceae): insights into Plio-Pleistocene range dynamics of evergreen forest in subtropical China. *New Phytol.* 206, 852–867. doi: 10.1111/nph.13261
- Wang, Y. X., Yang, J. D., Chen, J., Zhang, K. J., and Rao, W. B. (2007). The Sr and Nd isotopic variations of the Chinese Loess Plateau during the past 7 Ma: implications for the East Asian winter monsoon and source areas of loess. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 249, 351–361. doi: 10.1016/j.palaeo.2007.02.010
- Wu, Z. Y., Sun, H., Zhou, Z. K., Li, D. Z., and Peng, H. (2010). *Floristics of Seed Plants From China*. Beijing: Science Press.
- Xie, K. Q., and Zhang, M. L. (2013). The effect of Quaternary climatic oscillations on *Ribes meyeri* (Saxifragaceae) in northwestern China. *Biochem. Syst. Ecol.* 50, 39–47. doi: 10.1016/j.bse.2013.03.031
- Ye, J. W., Li, Q., Tian, X. Y., Wang, H. F., Bao, L., and Ge, J. P. (2017c). Twenty-seven low-copy nuclear gene primers for *Lindera obtusiloba* (Lauraceae): a Tertiary relict species in East Asia. *Appl. Plant Sci.* 5:1700120. doi: 10.3732/apps.1700120
- Ye, J. W., Yuan, Y. G., Cai, L., and Wang, X. J. (2017b). Research progress of phylogeographic studies of plant species in temperate coniferous and broadleaf mixed forests in Northeastern China. *Biodiv. Sci.* 25, 1339–1349. doi: 10.17520/biods.2017265
- Ye, J. W., Zhang, Y., and Wang, X. J. (2017a). Phylogeographic history of broad-leaved forest plants in subtropical China. *Acta Ecol. Sin.* 37, 5894–5904. doi: 10.5846/stxb201606031072
- Zeng, Y. F., Liao, W. J., Petit, R. J., and Zhang, D. Y. (2011). Geographic variation in the structure of oak hybrid zones provides insights into the dynamics of speciation. *Mol. Ecol.* 20, 4995–5011. doi: 10.1111/j.1365-294X.2011.05354.x
- Zhang, C., Ren, Z. Y., and Li, X. Y. (2012). Research on vegetation response to temperature and precipitation in Loess Plateau. *Sci. Agric. Sin.* 45, 4205–4215.
- Zhang, W. H., Li, D. W., Liu, G. B., and Xu, X. H. (2002). The characteristics of the seed plant flora in Loess Plateau. *Bull. Bot. Res.* 22, 373–379.

**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.

Copyright © 2021 Ye, Wu, Fu, Zhang and Tian. 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.