



Conservation Genomics of Wild Red Sage (*Salvia miltiorrhiza*) and Its Endangered Relatives in China: Population Structure and Interspecific Relationships Revealed From 2b-RAD Data

OPEN ACCESS

Xuan Zhou^{1,2}, Zhi-Cheng Zhang^{1,2}, Yan-Bo Huang², Han-Wen Xiao², Jun-Jie Wu¹, Zhe-Chen Qi^{1,3*} and Yu-Kun Wei^{2*}

Edited by:

Jian-Feng Mao,
Beijing Forestry University, China

Reviewed by:

Detuan Liu,
Kunming Institute of Botany, Chinese
Academy of Sciences, China
Yuguo Wang,
Fudan University, China

*Correspondence:

Zhe-Chen Qi
zqi@zstu.edu.cn
Yu-Kun Wei
ykwei@aliyun.com

Specialty section:

This article was submitted to
Evolutionary and Population Genetics,
a section of the journal
Frontiers in Genetics

Received: 30 March 2021

Accepted: 13 April 2021

Published: 11 May 2021

Citation:

Zhou X, Zhang Z-C, Huang Y-B,
Xiao H-W, Wu J-J, Qi Z-C and
Wei Y-K (2021) Conservation
Genomics of Wild Red Sage (*Salvia
miltiorrhiza*) and Its Endangered
Relatives in China: Population
Structure and Interspecific
Relationships Revealed From 2b-RAD
Data. *Front. Genet.* 12:688323.
doi: 10.3389/fgene.2021.688323

¹ Zhejiang Province Key Laboratory of Plant Secondary Metabolism and Regulation, College of Life Sciences and Medicine, Zhejiang Sci-Tech University, Hangzhou, China, ² Shanghai Key Laboratory of Plant Functional Genomics and Resources and Eastern China Conservation Center for Wild Endangered Plant Resources, Shanghai Chenshan Botanical Garden, Shanghai, China, ³ Shaoxing Academy of Biomedicine of Zhejiang Sci-Tech University, Shaoxing, China

Red sage (*Salvia miltiorrhiza*) is a widely used medicinal plant for treatment of cardiovascular and cerebrovascular diseases. Because of excessive excavation by huge market demand and habitat loss by human activities, the wild population resources of *S. miltiorrhiza* have reduced drastically in recent years. Meanwhile, population status of two closely related species *S. bowleyana* and *S. paramiltiorrhiza* were in a trend of decreasing due to their potential replacement of *S. miltiorrhiza*. Particularly, *S. paramiltiorrhiza* was threatened and endemic to a small region in eastern China. However, to date there has been no conservation genetic research reported for wild *S. miltiorrhiza* population and its endangered relatives. Assess the wild germplasm diversity for *S. miltiorrhiza* and its related species would provide fundamental genetic background for cultivation and molecular breeding of this medicinally important species. In the present study, we investigated the genetic diversity, population structure, and intra/inter-specific differentiation of *S. miltiorrhiza* and above two relatives using 2b-RAD genome-wide genotyping method. By investigating 81 individuals of *S. miltiorrhiza*, 55 individuals of *S. bowleyana* and 15 individuals of *S. paramiltiorrhiza* from 23 locations in China, we obtained 23,928 SNPs in total. A comparatively high genetic diversity was observed in *S. miltiorrhiza* ($\pi = 0.0788$, $H_e = 0.0783 \pm 0.0007$). The observed and expected heterozygosity in populations of these three species ranged from 0.0297 to 0.1481 and 0.0251 to 0.831, respectively. Two major lineage groups were detected in the examined *S. miltiorrhiza* populations. The results indicated that Dabie Mountain as a genetic diversity center of *S. miltiorrhiza* and possible complex inter-specific genetic

exchange/hybridization occurred between *S. miltiorrhiza* and the two relatives. We suggest that strategic conservation and germplasm preservation should be considered not only for wild populations of *S. miltiorrhiza*, but also for its related *S. bowleyana* and *S. paramiltiorrhiza*.

Keywords: *Salvia miltiorrhiza*, 2b-RAD, population genomics, conservation, genetic diversity, *Salvia bowleyana*, *Salvia paramiltiorrhiza*, wild germplasm

INTRODUCTION

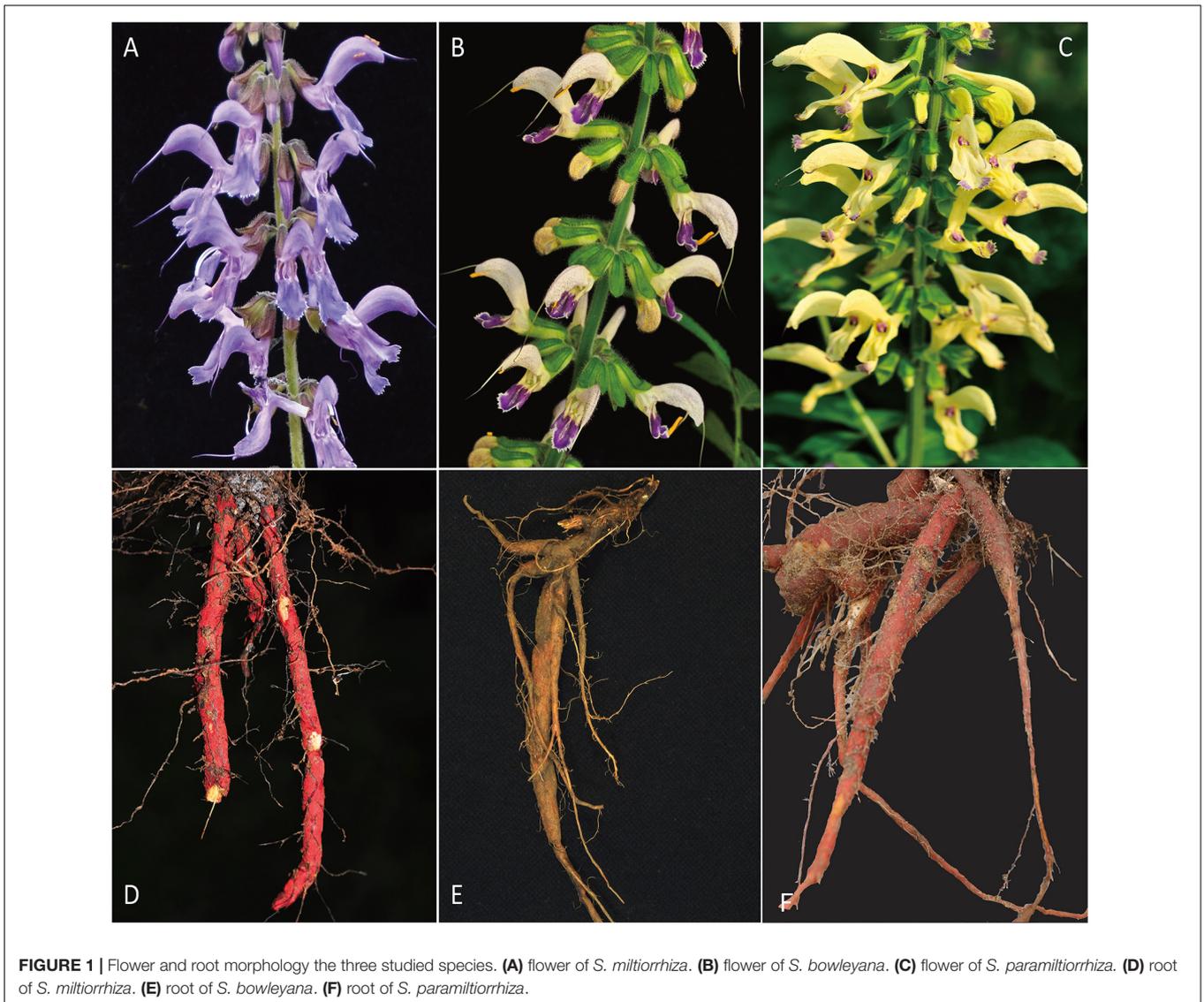
Medicinal plants are invaluable resource of maintaining health for the majority of world's population. At present, more than 80% of the world's population depends on herbal medicine for primary healthcare needs, while millions gain income from their wild harvest or cultivation, or are involved in their trading or processing (Hamilton, 2004). With the growing human populations, the demand for medicinal plants kept growing rapidly throughout the world for the past decades for their usage in herbal drugs and natural health products (Bentley, 2010; Jamshidi-Kia et al., 2018). According to IUCN, 50,000 and 80,000 flowering plant species are estimated to be used for pharmaceutical purposes around the world. For instance, China and India have the highest numbers of medicinal plants used, with 11,146 and 7,500 species, respectively. Among these numbers, 90% species were directly harvested from wild resources (Sarwat et al., 2011; Chen et al., 2016). Overexploitation and habitat destruction by human activity has led to an alarming number of medicinal plant species under threat of extinction. About 15,000 species are exposed to a risk of extinction due to high harvesting and destruction of habitats. A high proportion of medicinal plant species have become endangered in some species rich areas, threatening the regional biodiversity and ecosystem stability, especially in developing countries (Heywood and Iriondo, 2003; Hamilton, 2008; Bentley, 2010; Jamshidi-Kia et al., 2018).

One way to relief this situation is domestic cultivation. A number of governments and agencies are recommending that wild medicinal species should be brought into cultivation systems (WHO et al., 1993). Increased cultivation of medicinal plant is not only a means for meeting the huge market demands for drug and herbal remedies based on medicinal plant, but also a way to decrease harvest volume on wild populations, benefiting to the recovery of their wild resources (Schippmann et al., 2002; Canter et al., 2005). However, medicinal plant production through cultivation would probably result in genetic diversity reduce of cultivars due to founder effect and artificial selection, resembling the genetic bottleneck during the domestication of cereal species (Doebley et al., 2006). Decreased genetic diversity will lead them susceptible to pests, diseases, abiotic stress. Meanwhile, cultivation could also negatively affect the incentives to conserve wild populations (Schippmann et al., 2006). Wild population and relatives with genetic resources have proven to be very valuable in agricultural breeding practice (Tyack et al., 2020). The conservation of wild populations could maintain genetic diversity and allows

for the continuing evolution of the gene pool as a germplasm resource for future crop improvement (Tyack et al., 2020; Vining et al., 2020). In this regard, systematic collection and analyses of genetic diversity and relationships in wild medicinal plant populations can lay important foundations for the conservation of species as well as designing strategies for development of commercially relevant varieties (Govindaraj et al., 2015).

Salvia miltiorrhiza Bunge (Lamiaceae), the red sage, is an indigenous medicinal herb in China. Wild populations of *S. miltiorrhiza* usually grows on the hillsides, stream banks and understories in broad-leaved deciduous forests in central and eastern China. Its dried roots, commonly known as Danshen, have long been used in traditional Chinese medicine for treatment of cardiovascular and cerebrovascular diseases (Wang, 2010; Zhang et al., 2014). The production of *S. miltiorrhiza* in China achieved a total revenue of around 75 million USD in 2015 (Kum, 2019). Due to the huge market demand, although mostly harvested from cultivation, the wild populations of *S. miltiorrhiza* reduced drastically in recent years by excessive excavation and habitat destruction (Yang et al., 2017). Moreover, germplasm depression was frequently observed in cultivated *S. miltiorrhiza*. Meanwhile, two closely related species *S. bowleyana* Dunn and *S. paramiltiorrhiza* Li and Huang are also under a pressure of population decrease as frequent substitutes to *S. miltiorrhiza* (Field observation in the past decade). Particularly, populations of *S. paramiltiorrhiza* was critically threatened and can only be found in restricted areas in Hubei and Anhui, China (Huang and Li, 1981; Hao et al., 2015). Morphologically, these three species can be distinguished from each other mainly by their flower and root colors (Figure 1). However, previous phylogenetic studies based on plastid DNA and a few nuclear loci showed that their evolutionary relationships remain unresolved (Li et al., 2013; Hu et al., 2018). As wild populations and relatives as germplasm resources may provide infinite potentials for superior breeding varieties to improve the quality of cultivars (Vining et al., 2020; Wang et al., 2020), it is essential to conserve and assess the genetic diversity of *S. miltiorrhiza* and its relatives in natural populations.

Molecular markers are playing a significant role in studying population genetics of medicinal plants (Sarwat et al., 2011; Wang et al., 2020). With suitable genetic makers, we can understand the germplasm diversity, such as single nucleotide polymorphism (SNP), allele frequencies, extend and distribution of genetic diversity, and population structure, under a rigorous mathematical framework. Such information gained can be then referred to devise a proper conservation strategy and genebank management. With its ultra-high-throughput, next-generation sequencing (NGS) technology could provide an



efficient toolkit in producing large amount of SNP markers for organisms through simplified genome sequencing, for example, restriction site-associated DNA sequencing (RADSeq), genotyping-by-sequencing (GBS) or other related techniques (Davey et al., 2011; Peterson et al., 2012; Andrews et al., 2016). 2b-RAD, characterized by producing uniform fragments generated by type IIB restriction endonuclease, is a flexible and cost-effective method outperform other RAD-seq methods with higher markers numbers, repeatability, and genomic coverage density (Wang et al., 2012).

In present study, we assessed the genetic diversity of *S. miltiorrhiza* and two closely related species (*S. bowleyana* and *S. paramiltiorrhiza*) using 2b-RAD sequencing. We investigated the phylogenetic relationships and several aspects of the population genetics of focusing on wild *S. miltiorrhiza* populations including genetic diversity, population structure, and differentiation, as well as estimation of demographic history. We hope the methods and results presented here could be useful

in guidance of wild germplasm conservation and management for *S. miltiorrhiza* and other medicinal plants in the future.

MATERIALS AND METHODS

Sampling and DNA Extraction

A total of 151 individuals from 23 sites in central and eastern China were sampled and analyzed in this study, including *S. miltiorrhiza* (81 individuals from 13 wild populations), *S. bowleyana* (55 individuals from 9 wild populations) and *S. paramiltiorrhiza* (15 individuals from 1 wild populations) (**Supplementary Table 1**). Our sampling focused on an area where populations of all three species occur, in order to assess their potential gene flows. Leaves collected from each individual were dried in silica gel. Genomic DNA was extracted from dried leaf powder using a modified CTAB method (Allen et al., 2006). First, gel-dried leaf samples were put into liquid nitrogen and

quickly ground to powder; the powder was then immediately transferred into a clean 1.5 ml microcentrifuge tube and mixed immediately with 600 μ l 2 \times CTAB extracting solution and incubated in a water bath at 65°C for 45 min. The samples were then centrifuged twice at 12,000 rpm for 10 min, after which the supernatant was transferred to another clean 1.5 ml microcentrifuge tube. Then 5 μ l DNase free RNase (10 μ g/ μ l) was added and incubate at 37°C for 15 min, intermittently mixed up for 2–3 times. The same volume of chloroform-isoamyl alcohol (24:1) was then added and mixed up the phases by gentle inversion. The sample was centrifuged at 12,000 rpm for 5 min, after which the supernatant was transferred to a new 1.5 ml centrifuge tube and mixed with the same volume of isopropanol, then incubated at –20°C for about 1 h. The solution was centrifuged at 12,000 rpm for 10 min, then the supernatant liquor was carefully discarded. The DNA was then cleaned twice with 75% ethanol and air dried at room temperature. The cleaned DNA was then dissolved in 50 μ l TE. DNA quantity was assessed with 1% agarose gel electrophoresis, and Qubit 3.0 fluorescent quantitative assay and NanoVue were used to test purity and density.

Multi-IsoRAD Sequencing and SNP Calling

Library construction and sequencing were conducted following Wang et al. (2016), an updated 2b-RAD approach that allows the preparation of five concatenated isoRAD tags for Illumina paired-end sequencing. After DNA extraction and quality control, the high-quality genomic DNA (600 ng) was digested with restriction enzyme of IIB type (*Bsa*XI) at 37° for 1 h, followed by enzyme heat inactivation at 65° for 20 min. The ligation reaction was performed by combining 5 μ l of digested DNA with 20 μ l of a ligation master mix containing 0.4 μ M each of library-specific adaptors with fully degenerate cohesive ends (5' -NNN- 3'), 0.2 mM adenosine 5'-triphosphate (New England Biolabs), and 1000 U T4 DNA ligase. Then PCR amplification was performed. Ligation was carried out at 16° for 3 h, with subsequent heat inactivation for 10 min at 65°. In Multi-isoRAD protocol, the single-tag constructs can be further digested by *Sap*I to generate distinct cohesive ends and then ligated in a predefined order to produce five concatenated tags for Illumina paired-end sequencing using modified adaptors and biotin-labeled primers. After individual libraries were pooled into equimolar amounts, the pool quality was verified on Agilent 2100 Bioanalyzer. The paired-end RAD-seq were generated using Illumina HiSeq X ten platform (Illumina, San Diego, CA, United States) using 150 bp pair-end sequencing.

Quality control were performed with Trimmomatic (Bolger et al., 2014) according to the following criteria: (a) any reads with adapter; (b) any reads with > 50% bases having phred quality <15; (c) any reads with >8% unidentified nucleotides (N). After that, the high-quality reads of each sample were aligned to a reference genome of *S. miltiorrhiza* (Xu et al., 2016) using the SOAP2 (version 2.21) program following the protocols by Li et al. (2009). A maximum of two mismatches (–v 2) was allowed for each read, and those that mapped onto more than one position

in the genomic reference sequence were excluded (–r 0). Reads with depths <3 were not counted. The match mode was set to “find the best hits” (–M 4). and the SNPs were obtained based on a maximum likelihood algorithm and filtered by the RADtyping program (Fu et al., 2013). Finally, Quality control was performed using VCF tools (Danecek et al., 2011) to remove SNPs showing a call rate of less than 90%, a minor allele frequency (MAF) of less than 0.05 or significant deviation from Hardy-Weinberg equilibrium ($p < 0.05$).

Phylogenetic Reconstruction

To infer relationships among samples, we performed phylogenetic analyses with concatenated SNP dataset using maximum likelihood (ML) method. The optimal substitution models for the ML phylogenetic analyses was determined by ModelFinder program (Kalyaanamoorthy et al., 2017) using the Bayesian information criterion (BIC), as implemented in IQ-TREE v2 (Minh et al., 2020). The ML analysis was conducted with IQ-TREE. Five searches were compared to ensure identical topologies. Statistical supports were analyzed using 10,000 replicates of ultrafast bootstrapping (UFBoot: Minh et al., 2013) and 10,000 bootstrap replicates of the Shimodaira/Hasegawa approximate likelihood ratio test (SH-aLRT: Guindon et al., 2010).

Analyses of Genetic Diversity and Structure

Population genetic statistics, including observed and expected heterozygosity (H_e , H_o), nucleotide diversity (π), and Wright's F statistics of genetic differentiation (F_{st}) and inbreeding coefficient (F_{IS}) were assessed using the “populations” program in Stacks v2.2 (Rochette et al., 2019), and the pairwise statistics between populations were performed using non-parametric tests. Principal component analysis (PCA) and population structure analysis with all SNPs were performed using Plink v1.9 (Purcell et al., 2007) and ADMIXTURE v1.3.0 (Alexander et al., 2009), respectively. Population genetic variation was calculated using an analysis of molecular variance (AMOVA) with packages *adgenet* v2.1 (Jombart and Ahmed, 2011) and *poppr* v2.9 (Kamvar et al., 2015) in R v4.0.3 (R Development Core Team, 2018). Pairwise comparisons of F_{st} among and within three studied species were estimated. To investigate the correlation between genetic distance and geographic distances, Mantel test was performed using GenALEX v6.5 (Peakall and Smouse, 2012). Geographic distance between field sampling sites were calculated using *geosphere* v1.5 package, and genetic distances were obtained with *nei.dist* function of *poppr* package.

Demographic History and Gene Flow

Recent demographic dynamics of *S. miltiorrhiza* was investigated following Liu et al. (2020). A python script easySFS¹ was used to generate the folded site frequency spectrum (SFS) formatted file with VCF files and sample list file, and demographic history was then inferred using the program Stairway plot v.0.2

¹<https://github.com/isaacovercast/easySFS>

(Liu and Fu, 2015). The mutation rate of *S. miltiorrhiza* was set to 2.67×10^{-8} per site per generation (unpublished data, based on genomic scale mutation and divergence between *S. miltiorrhiza* and *S. bowleyana*). The generation time was set to 1 year per generation. The program BAYESASS version 3.0.3 (Wilson and Rannala, 2003) was used to examine contemporary gene flow (over the last few generations) and migration rates. Analyses were performed inter-specific and inter-cluster levels separately. We conducted five independent runs using 10^7 iterations, a burn-in length of 10^6 and a sampling interval of 10^3 steps. To reach the recommended acceptance rates, Markov chain Monte Carlo mixing parameter values for migration rates (gene flow), allele frequencies, and inbreeding coefficients were adjusted to 0.50, 0.95, and 0.50, respectively. To confirm convergence, each run was executed with different seed numbers, and trace files were examined for consistent oscillations using Tracer 1.7 (Rambaut et al., 2018).

RESULTS

Summary Statistics

A total of 1,097,351,870 raw reads were generated from 151 individuals. After all quality filter, we retained a total of 941,122,629 reads from the initial raw reads, with an average of 6,232,600 reads per sample. The average number of unique tags was 50,031 after removing the unique tags whose sequencing depth was less than 3. Mean sequence depths across all samples was $32\times$, ranging from $17.77\times$ to $58.41\times$. After filtering, we obtained 23,928 SNPs in all samples for further analyses.

Phylogenetic Relationships

The ML phylogeny inferred from TVM + F + R2 model identified four well-supported clades (A–D), with SH-aLRT/UFBoot values displayed above branches (Figure 2). Except several subclades in clade C, all other nodes received robust statistical supports. Clade A includes individuals from 7 populations of *S. miltiorrhiza* and one population of *S. paramiltiorrhiza*. *Salvia paramiltiorrhiza* population is well-embedded in *S. miltiorrhiza*, forming a monophyly with population SMHX. Clade B includes individuals from a single *S. bowleyana* population (SBZL) of south Zhejiang province. In clade C, SBAQ and one SBAH individual of *S. bowleyana* formed the basal clade C1 (SH-aLRT/UFBoot: 99.6/87). Clade C2 (SBHL and SBHY) is sister to clade C3 + C4. Clade C3 includes SBAH and SBZS of *S. bowleyana* populations is in turn sister to three *S. miltiorrhiza* populations (SMAT, SMAH and SMAX) with weak support (SH-aLRT/UFBoot: 32.4/54). SMHS, SMAJ and SMHB of *S. miltiorrhiza* and SBAY, SBZT and SBJY of *S. bowleyana* showed inter-mixed relationships in clade D.

Genetic Diversity and Population Structure

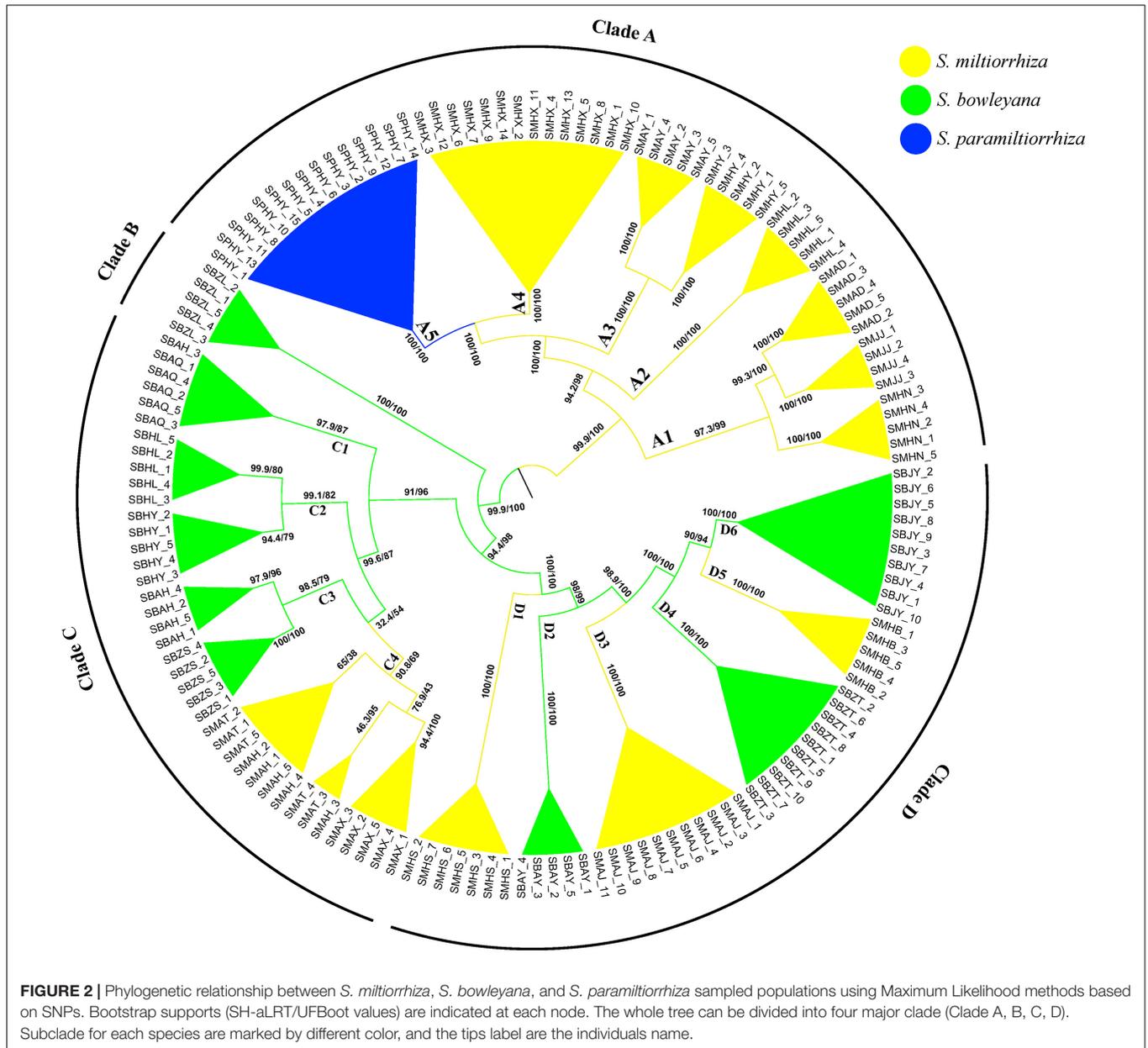
In *S. miltiorrhiza* populations, the observed and expected heterozygosity range from 0.0297 to 0.0647 and 0.0251 to 0.0571, respectively. The mean values of π and F_{IS} are 0.0282 to 0.0641

and -0.0066 to 0.0095, respectively. Significant difference of the genetic parameters were detected among the populations (H_o : $p = 0.000$; H_e : $p = 0.000$; π : $p = 0.000$; F_{IS} : $p = 0.000$). In *S. bowleyana* populations, H_o and H_e range from 0.0353 to 0.1481 and 0.0322 to 0.0831, respectively. The mean values of π and F_{IS} are 0.0371 to 0.0890 and -0.1129 to 0.0078. Except F_{IS} , significant difference of the genetic parameters were also detected among the *S. bowleyana* populations (H_o : $p = 0.000$; H_e : $p = 0.000$; π : $p = 0.000$; F_{IS} : $p = 0.028$). The mean values of H_o , H_e , π and F_{IS} in *S. paramiltiorrhiza* are 0.0464, 0.0445, 0.0462, and 0.0018, respectively. High values of private alleles are observed in some populations sampled in the present study (Table 1). Average private alleles in a population are 239 (*S. miltiorrhiza*, highest = 995 in SMHB), 787 (*S. bowleyana*, highest = 4790 in SBZT) and 353 (*S. paramiltiorrhiza*), respectively. The summary statistics of genetic diversity of all three species are given in Table 1.

In the Admixture analysis, the CV errors continued to decrease as K increased in value, giving no clear indication of the appropriate K for the study populations. As such, conclusions on population substructuring were based on visual inspection of the admixture plots (Figure 3) and PCA plots (Figure 4). A graphic representation of cluster structure analysis is depicted in Figure 3 with K from 3 to 5. The species and the genotype boundary between groups is clearer at $K = 5$. *Salvia paramiltiorrhiza* is dominated by cluster 1. Four major clusters were observed in *S. miltiorrhiza* (cluster 2–5). It also indicated that cluster 4 and 5 were shared by six populations of *S. miltiorrhiza* (SMAX, SMAH, SMAT, SMAJ, SMHS, and SMHB) and *S. bowleyana* (Figure 3, $K = 5$). Different degrees of inter/intra-specific admixtures were observed in the studied populations. Further insights into population structure were obtained by plotting the five genetic clusters on the map (Figure 5). Geographical distribution of the genetic clusters at population level did not show conspicuous genetic structure.

For PCA plots, PC1, PC2, and PC3 explained approximately 33.43% of the total variation. When we grouped the samples according to the five clusters in Admixture analysis, cluster 1 and 2 overlapped to each other in PC1-PC2 plot (Figure 4A), while showed a distant relationship in PC1-PC3 plot (Figure 4B). Cluster 3, a genetic group from *S. miltiorrhiza* showed close relationship with cluster 4 in both plots. In cluster 4, individuals of *S. miltiorrhiza* and *S. bowleyana* mixed together (Figures 4A,B). Cluster 5 showed isolated relationships with other clusters. It is also noted that while share the same cluster 5, individuals of *S. miltiorrhiza* (SMHB) show a clear distance with *S. bowleyana* populations (SBZT and SBJY).

AMOVA analyses was performed among species and within each species separately. For within-species analysis, we grouped the populations according to the five ancestry clusters recovered by Admixture analysis. The results revealed that inter-specific differences account for 13.12% of the genetic variation. Most of genetic variation is due to the differences within populations (49.78%) (Table 2). The same pattern was observed in *S. miltiorrhiza* analyses, that high percentage of the genetic variance (65.52%) came from the differences within populations. In *S. bowleyana*, most of the genetic variation (56.36%) is due to



differences between populations within clusters. Pairwise genetic differentiation (F_{st}) between samples from sampling sites were different significantly, as shown in **Supplementary Table 2**, and the results are consistent with the genetic variation between populations by AMOVA. Mantel test failed to detect a significant isolation by distance (IBD) based on the geographical distance and genetic divergence matrices in both species (*S. miltiorrhiza*: $R^2 = 0.0013$, $p = 0.460$; *S. bowleyana*: $R^2 = 0.0174$, $p = 0.450$).

Migration Rates and Gene Flow Analysis

For inter-specific migrations, we detected significant gene flow from *S. bowleyana* to *S. miltiorrhiza* (21.45%) estimated in BAYESASS (**Table 3**). It is consistent with the Admixture results that several *S. miltiorrhiza* populations shared clusters with

S. bowleyana. The second highest but not significant migration was inferred from *S. paramiltiorrhiza* to *S. miltiorrhiza*. For inter-cluster runs, overall low levels of contemporary gene flow were revealed. The highest migration rates were inferred from the cluster 5 to 4 (6.44%), cluster 1 to 2 (3.23%), and cluster 3 to 4 (2.56%). Cluster 4 received comparatively higher gene flows from other clusters. This is also observed in Admixture analyses that most individuals from cluster 4 showed differential admixture levels with other clusters.

Effective Population Size and Demographic History

Effective population size and demographic history of three studied species were inferred based on SNP frequency spectra

TABLE 1 | Population genetic statistics of *S. miltiorrhiza*, *S. bowleyana*, and *S. paramiltiorrhiza* in this study: observed heterozygosity (H_o), expected heterozygosity (H_e), genetic diversity (π), inbreeding coefficients (F_{IS}).

Taxon	Pop ID	Private alleles	H_o	H_e	π	F_{IS}
<i>S. bowleyana</i>	SBAH	16	0.0591 ± 0.0010	0.0547 ± 0.0009	0.0614 ± 0.0010	0.0058 ± 0.0029
	SBAQ	141	0.0353 ± 0.0009	0.0322 ± 0.0007	0.0371 ± 0.0008	0.0041 ± 0.0048
	SBAY	146	0.0484 ± 0.0010	0.0416 ± 0.0008	0.0471 ± 0.0009	-0.0018 ± 0.0041
	SBHL	13	0.0581 ± 0.0010	0.0545 ± 0.0008	0.0612 ± 0.0010	0.0073 ± 0.0030
	SBHY	11	0.0614 ± 0.0011	0.0552 ± 0.0009	0.0619 ± 0.0010	0.0020 ± 0.0025
	SBJY	1918	0.0508 ± 0.0012	0.0405 ± 0.0009	0.0430 ± 0.0009	-0.0142 ± 0.0105
	SBZL	29	0.0553 ± 0.0010	0.0521 ± 0.0008	0.0587 ± 0.0010	0.0077 ± 0.0035
	SBZS	22	0.0551 ± 0.0010	0.0519 ± 0.0009	0.0587 ± 0.0010	0.0078 ± 0.0035
	SBZT	4790	0.1481 ± 0.0027	0.0831 ± 0.0015	0.0890 ± 0.0016	-0.1129 ± 0.0154
	<i>S. miltiorrhiza</i>	SMAD	97	0.0408 ± 0.0009	0.0398 ± 0.0008	0.0449 ± 0.0009
SMAH		8	0.0489 ± 0.0010	0.0452 ± 0.0008	0.0526 ± 0.0009	0.0083 ± 0.0054
SMAJ		516	0.0563 ± 0.0010	0.0535 ± 0.0009	0.0566 ± 0.0009	0.0025 ± 0.0085
SMAT		12	0.0647 ± 0.0011	0.0571 ± 0.0009	0.0641 ± 0.0010	-0.0003 ± 0.0028
SMAX		39	0.0548 ± 0.0010	0.0493 ± 0.0008	0.0560 ± 0.0010	0.0031 ± 0.0039
SMAY		55	0.0469 ± 0.0010	0.0408 ± 0.0008	0.0462 ± 0.0009	-0.0009 ± 0.0036
SMHB		995	0.0456 ± 0.0011	0.0401 ± 0.0008	0.0452 ± 0.0009	-0.0002 ± 0.0053
SMHL		185	0.0297 ± 0.0009	0.0251 ± 0.0007	0.0282 ± 0.0007	-0.0022 ± 0.0041
SMHN		77	0.0520 ± 0.0011	0.0450 ± 0.0008	0.0505 ± 0.0009	-0.0027 ± 0.0030
SMHS		500	0.0430 ± 0.0010	0.0373 ± 0.0008	0.0407 ± 0.0009	-0.0041 ± 0.0054
SMHX		298	0.0443 ± 0.0009	0.0425 ± 0.0008	0.0443 ± 0.0008	0.0019 ± 0.0090
SMHY		130	0.0378 ± 0.0010	0.0321 ± 0.0007	0.0362 ± 0.0008	-0.0022 ± 0.0038
SMJJ		195	0.0370 ± 0.0010	0.0284 ± 0.0007	0.0332 ± 0.0008	-0.0066 ± 0.0030
<i>S. paramiltiorrhiza</i>		SPHY	353	0.0464 ± 0.0009	0.0445 ± 0.0008	0.0462 ± 0.0008

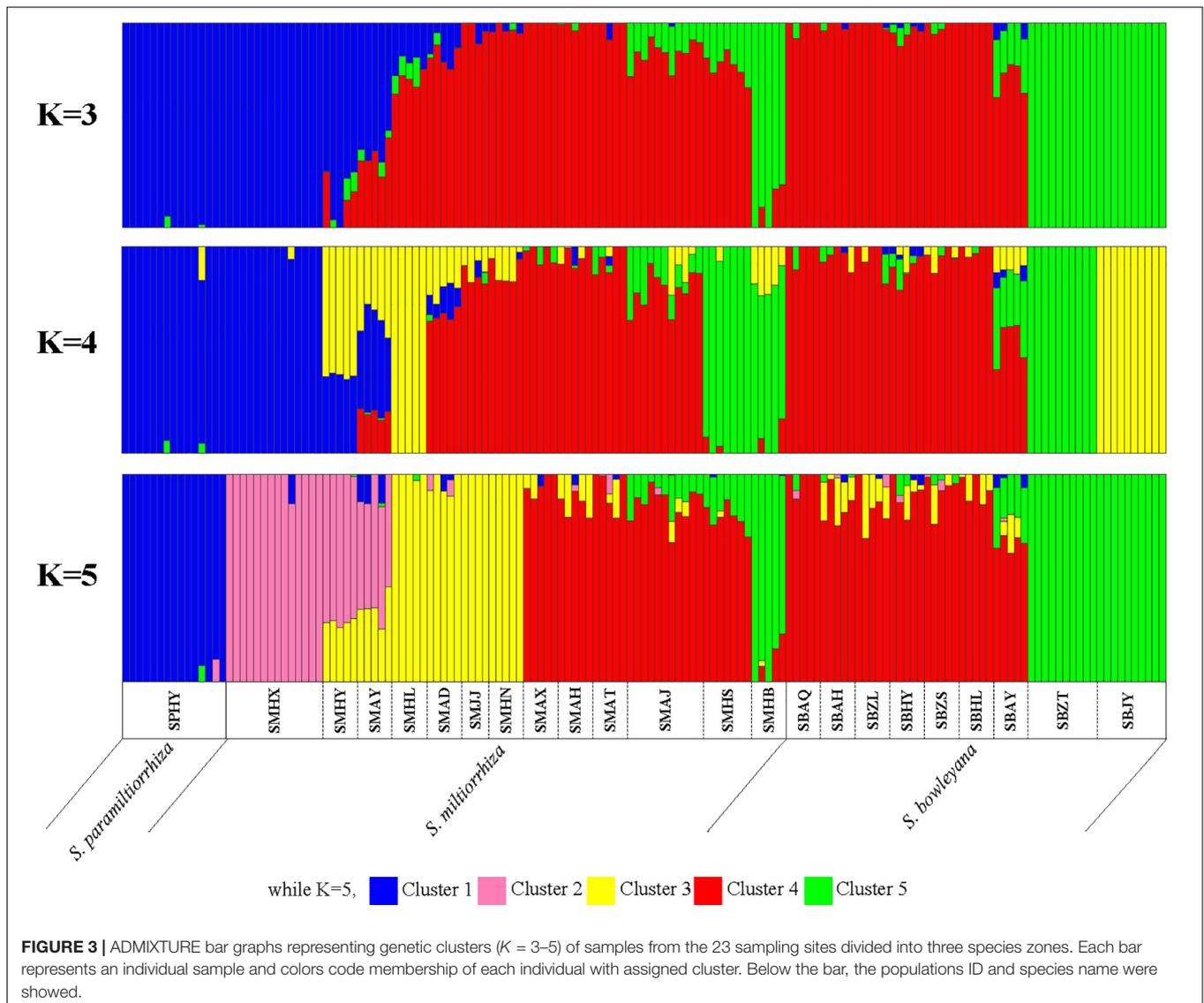
and displayed on stairway plots (Figure 6). For *S. miltiorrhiza*, an initial shallow population contraction occurred at ca. 60 kya (thousand years ago), after which followed a population expansion and then a flat period until 8 kya. Then, a second more severe population bottleneck occurred during ca. 8–6 kya, which followed another population expansion and then a flat growth period maintained to present. For *S. bowleyana* populations, a deep population bottleneck was also detected around 8–6 kya, which followed a similar population dynamics as *S. miltiorrhiza*. *Salvia paramiltiorrhiza* experienced an early population contraction at ca. 104 kya, and then rapidly expanded and maintained its population size to present.

DISCUSSION

Preserving the biological and genetic diversity from wild plant species were considered as an essential and effective way of conservation and sustainable use of medicinal plants (Hamilton, 2008; Chen et al., 2016; Vining et al., 2020). Therefore, it is important to characterize genetic diversity and population structure of targeted species and its relatives, especially threatened ones (Lee et al., 2018, 2020). The genetic diversity in three studied species are relatively high compare to other medicinal herbs, for example, *Coptis* species (Wang et al., 2020), *Valeriana officinalis* (Boczkowska et al., 2020), *Aconitum austrokoreense* (Lee et al., 2018). Especially for *S. paramiltiorrhiza*, characterized by yellow flower and longer tube without hair annulus, high genetic diversity was

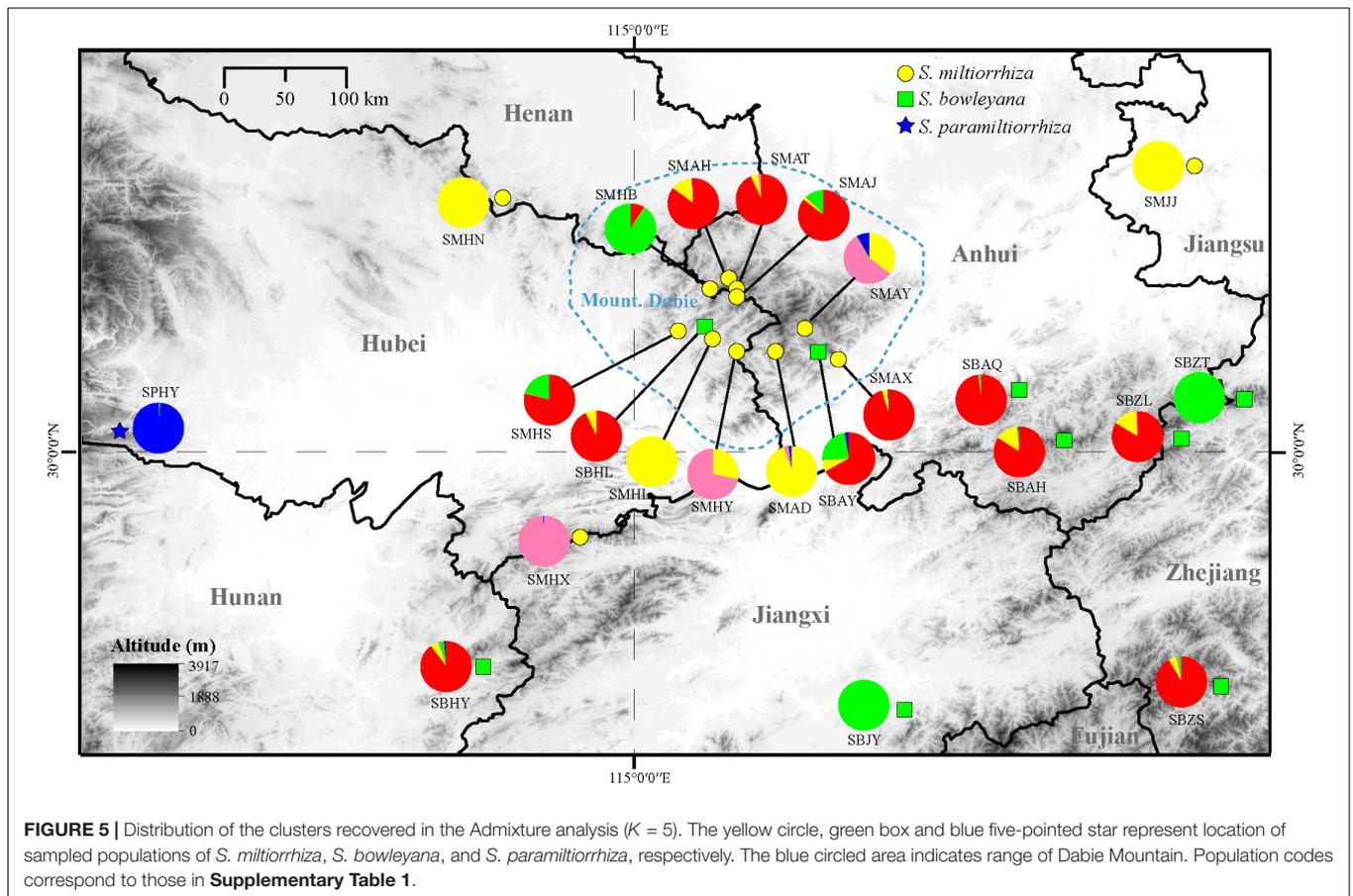
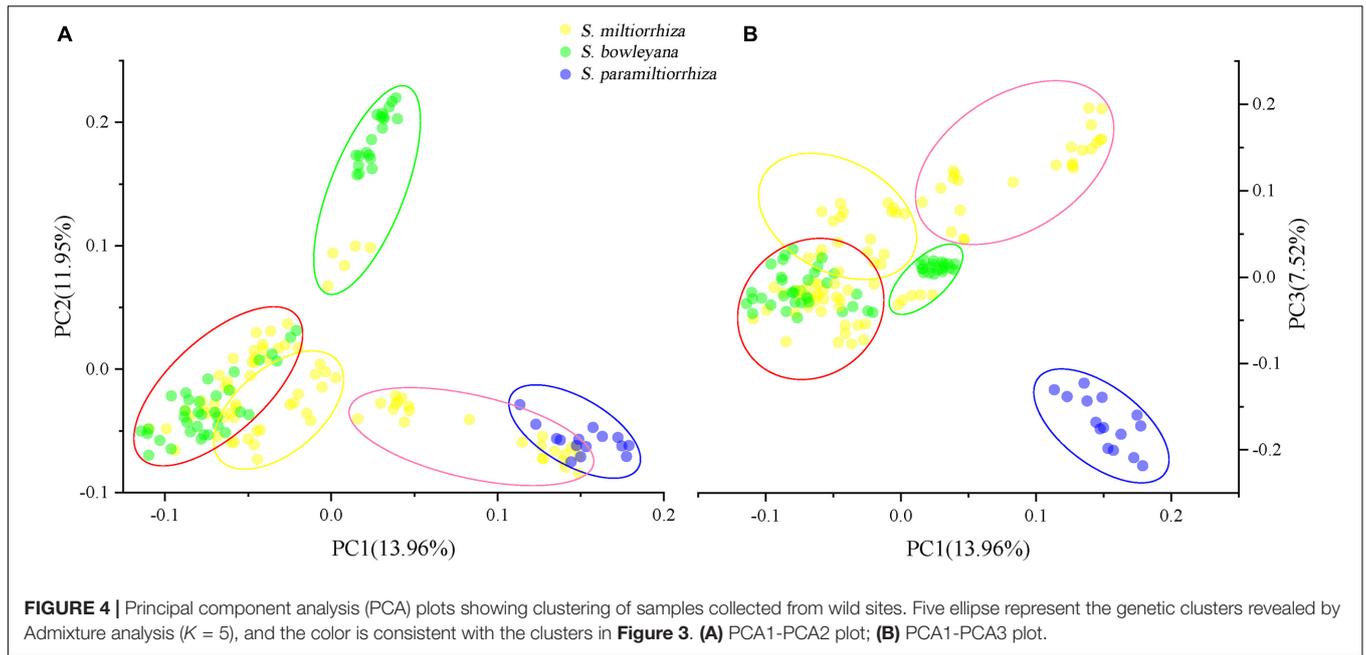
observed (0.0462) despite its small population size and restricted distribution. This could be due to its outcrossing behavior and possible historical introgression with *S. miltiorrhiza*, which is supported by PCA plot (Figure 4A). Although phylogenetic analysis indicated that *S. paramiltiorrhiza* is nested within *S. miltiorrhiza* populations (Figure 2), Admixture analysis (Figure 3, $K = 5$) and PCA plot (Figure 4B) revealed that it's a distinct cluster to *S. miltiorrhiza* and *S. bowleyana* populations. Pairwise F_{st} also indicates that SPHY, the *S. paramiltiorrhiza* population, is significantly differentiated from all other studied populations of *S. miltiorrhiza* and *S. bowleyana* (Supplementary Table 2). Geographically, *S. paramiltiorrhiza* is isolated from other sampled populations (Figure 5). However, admixture from *S. paramiltiorrhiza* to *S. miltiorrhiza* were detected in population SMHX, SMAY, and SMAD. A more extensive field examination is needed to confirm whether coexistence of *S. miltiorrhiza* is occurred in its distribution range.

Relationships between *S. miltiorrhiza* and *S. bowleyana* indicated a possible complex history. Morphologically, *S. bowleyana* can be distinguished from *S. miltiorrhiza* by its brownish root and yellow to pink upper corolla lips (Figure 1). Genetic analyses indicated that six populations (SMAX, SMAH, SMAT, SMAJ, SMHS, SMHB) of *S. miltiorrhiza* shared cluster 4 and 5 with *S. bowleyana* populations in Admixture analysis, which is supported by phylogenetic analysis (clade C and D, Figure 2). All of these six populations distributed in the Dabie Mountain area, which overlaps the northern distribution range of *S. bowleyana* (Figure 5). On one hand, we assumed that the frequent gene flows and possible hybridization might



contributed to the mixture between these two species. The species used in this study were carefully identified, nonetheless, one could actually be hybrids with morphological similarity to one of the parent species. Although unidirectional migration of gene flow is detected from *S. bowleyana* to *S. miltiorrhiza* in BAYESASS analyses, we observed reciprocal admixtures between populations of these two species (Figure 3), which is also indicated by the low inter-specific differentiation (0.1312) by AMOVA (Table 2). The weakly supported subclades in clade C might indicate the influence of partially contradict informative SNPs introduced by the gene flow between populations of these two species. Moreover, *S. miltiorrhiza* and *S. bowleyana* blossom at the same time and share pollinators (field observation), which indicate that the reproductive barriers between them might be weak to some extent, leading to inter-specific gene flow. On the other hand, since these two species occur in partial allopatry, the undifferentiated genetic divergence could also be a result of retention of ancestral polymorphism, as incomplete lineage

sorting could lead to a similar phenomenon which was observed in many other closely related taxa (Alexander et al., 2017; Goetze et al., 2017; Cerca et al., 2021). Pairwise F_{st} estimates between SMAJ, SMAH, SMAT, SMHB from *S. miltiorrhiza* and population SBHL, SBHY, SBZS, SBZL, SBAH, SBAQ of *S. bowleyana* were not statistically significant (Supplementary Table 2). In phylogenetic analysis, clade B & C, dominated by populations from cluster 4, is sister to clade D, which is dominated by populations from cluster 5. Nonetheless, PCA analysis suggested a closer distance of the mixed cluster 4 and cluster 3 (*S. miltiorrhiza* populations in clade A). This indicated that both species could potentially contribute to the ancestry of cluster 4. It is also noteworthy that populations of *S. miltiorrhiza* (SMHB, SMHS, SMAJ) and *S. bowleyana* (SBZT, SBJY, SBAY) in cluster 5 did not mix to each other as shown in PCA plot (Figure 4), despite sharing the same genetic ancestry revealed by Admixture analysis. Geographically, population SMHB is distantly located to SBZT/SBJY, indicating low recent gene flow



may occur between them. It suggests a retention of ancestral polymorphism of cluster 5 in both species. The AMOVA revealed different genetic differentiation scenarios of *S. miltiorrhiza* and

S. bowleyana. Differentiation among populations within clusters of *S. bowleyana* is high, and composed most of the genetic variation (56.36%). However, in *S. miltiorrhiza*, most of variation

TABLE 2 | Analysis of molecular variance (AMOVA) showing genetic variation among and within three studied species from collection sites.

Source of variation	Sum Sq	Variance components	% of variation	P-value	F-Statistics
Interspecific comparison					
Among three species	108340.00	704.14	13.12	0.076	0.1312
Among populations within species	295574.20	1989.97	37.09	0.001	0.4270
Within populations	341878.20	2670.92	49.78	0.001	0.5022
Salvia miltiorrhiza					
Among Clusters	52783.73	566.43	15.35	0.001	0.1535
Among populations within Clusters	62582.16	706.21	19.13	0.001	0.2260
Within populations	162008.66	2418.04	65.52	0.001	0.3448
Salvia bowleyana					
Among Clusters	74212.78	1346.46	16.46	0.197	0.1646
Among populations within Clusters	199950.89	4609.83	56.36	0.001	0.6747
Within populations	102235.60	2222.51	27.17	0.001	0.7283

were detected within populations, and low genetic differentiation were observed in all group settings. This might imply that in spite of frequent gene flow, these two species might experienced differential evolutionary history. Further studies on *S. miltiorrhiza* and *S. bowleyana* at a larger population scale will be necessary to clarify the boundary between these two species.

Our analyses provided insights into the distribution of genetic diversity in *S. miltiorrhiza*, *S. bowleyana*, and *S. paramiltiorrhiza*, revealing that populations located in the Dabie Mountain range harbor the greatest amount of diversity (Figure 5). This area could potentially serve as a connection between *S. miltiorrhiza* from the north and *S. bowleyana* from the south. They might have experienced continuous gene flow over time, increasing toward the present day. According to the demographic history of *S. miltiorrhiza* and *S. bowleyana* illustrated by stairway plot, they experienced a contemporary bottleneck occurred at 8–6 kya (Figure 6), which was consistent with the 8.2 k event of global cooling and population reduction in the Holocene epoch (Alley and Agustsdottir, 2005; Wicks and Mithen, 2014), which imply that *S. miltiorrhiza* and *S. bowleyana* may have shared refugia during climatic oscillations, where hybridization occurred. *Salvia miltiorrhiza* experienced an additional population contraction occurred at ca. 60 kya, which was consistent with the abrupt climate change and ecological turnover in late Pleistocene glaciation period (Finlayson and Carrión, 2007; Astakhov, 2008). *Salvia bowleyana* and *S. paramiltiorrhiza* might avoid the contraction during the glaciation due to their further south distribution range compare to *S. miltiorrhiza*. Therefore, genetic exchange could also occur among them when *S. miltiorrhiza* retreated

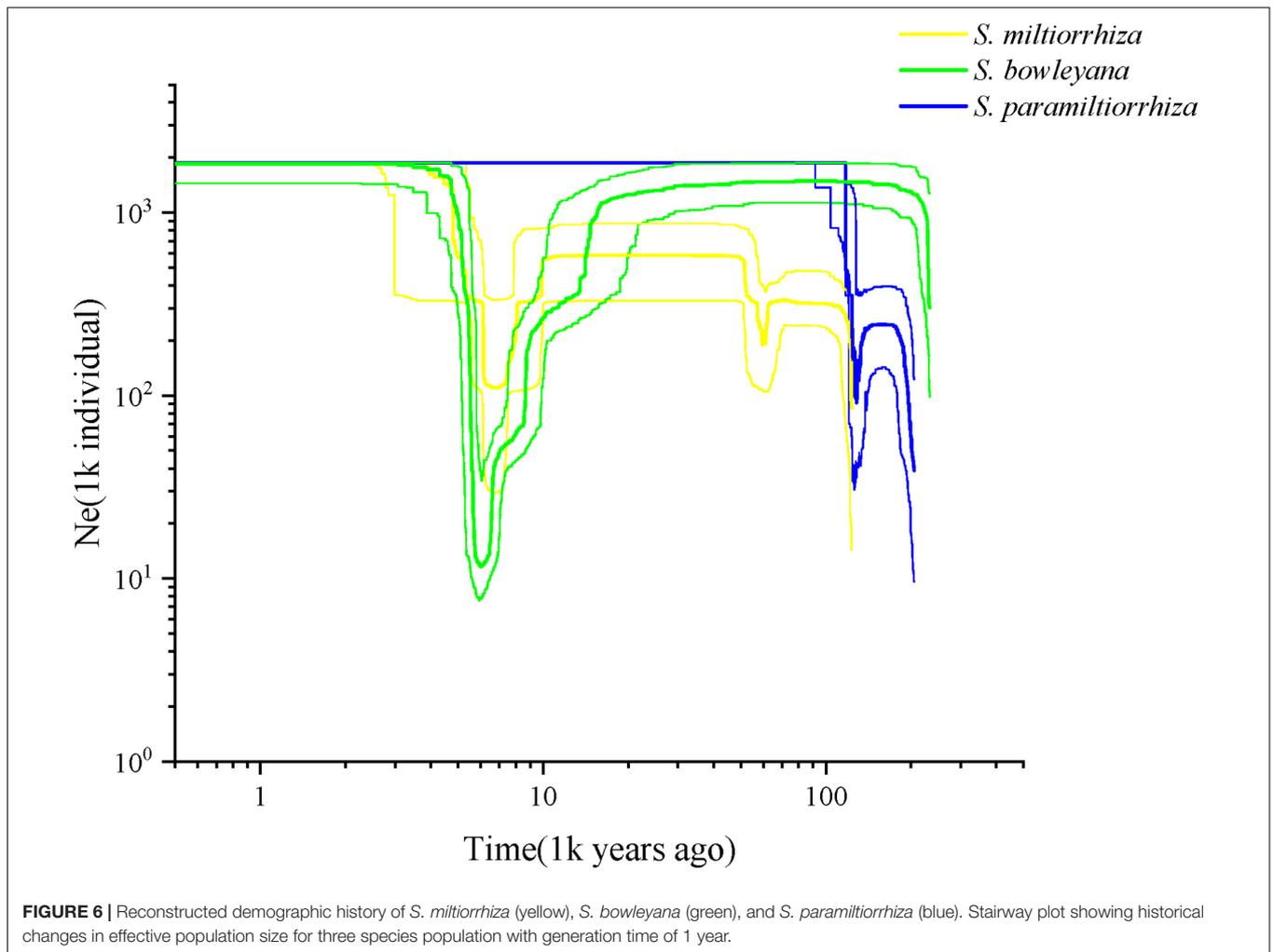
TABLE 3 | Posterior means of contemporary migration rates among species and genetic clusters estimated by BAYESASS (values in brackets are standard deviations, and mean values more than 0.1 are considered significant).

Populations	Migration rates
SB→SM	0.2145 (0.0219)
SB→SP	0.0060 (0.0058)
SM→SB	0.0156 (0.0076)
SM→SP	0.0040 (0.0040)
SP→SB	0.0175 (0.0165)
SP→SM	0.0250 (0.0217)
cluster1→cluster4	0.0158 (0.0152)
cluster1→cluster2	0.0323 (0.0211)
cluster1→cluster3	0.0165 (0.0157)
cluster1→cluster5	0.0158 (0.0152)
cluster2→cluster4	0.0227 (0.0152)
cluster2→cluster3	0.0115 (0.0110)
cluster2→cluster5	0.0115 (0.0113)
cluster2→cluster1	0.0118 (0.0115)
cluster3→cluster4	0.0256 (0.0175)
cluster3→cluster2	0.0134 (0.0131)
cluster3→cluster5	0.0134 (0.0127)
cluster3→cluster1	0.0134 (0.0128)
cluster4→cluster2	0.0046 (0.0045)
cluster4→cluster3	0.0191 (0.0092)
cluster4→cluster5	0.0238 (0.0101)
cluster4→cluster1	0.0048 (0.0047)
cluster5→cluster4	0.0644 (0.0235)
cluster5→cluster2	0.0109 (0.0105)
cluster5→cluster3	0.0106 (0.0105)
cluster5→cluster1	0.0107 (0.0102)

SM: *S. miltiorrhiza*; SB: *S. bowleyana*; SP: *S. paramiltiorrhiza*. The bold value indicates strong signal of contemporary migration.

further south during that period. Population expansion of both *S. miltiorrhiza* and *S. bowleyana* took place during 6–4 kya, after which the population size of both species reached a stable status to present. No recent reduction in population size was observed based on genetic data. However, due to recent excessive excavation and habitat destruction, the wild populations of *S. miltiorrhiza* and its relatives reduced drastically. This could be explained by that the populations we sampled in the study were mostly collected in the nature reserves, where human activities are rare and restricted. Moreover, the population destruction started just decades ago, as such the role of genetic drift will be limited over such a short period of time (Johnson et al., 2018). Otherwise, the frequent gene flow among the species could also contributed to their maintenance of genetic diversity.

Our study suggests a complex scenario of inter/intra-specific evolutionary history of among *S. miltiorrhiza*, *S. bowleyana*, and *S. paramiltiorrhiza*. Incomplete lineage sorting, repeated gene flow, as well as localized events of hybridization together might have shaped the current genetic structure and intermixed relationship among these three closely related species. The molecular data indicates that *S. paramiltiorrhiza* exhibits a unique lineage after comparing with *S. miltiorrhiza* and *S. bowleyana* populations. Additionally, several novel compounds has been



discovered from the roots of *S. paramiltiorrhiza* (Sun et al., 1991, 1992; Lin and Chang, 2000). For *S. bowleyana*, a complex interspecific interaction history between *S. miltiorrhiza* was revealed. It leads to a better understanding of the evolutionary history and dynamics of this species complex group. As relatives with *S. miltiorrhiza*, *S. paramiltiorrhiza* and *S. bowleyana* could serve as valuable wild germplasm resources to improve the quality of *S. miltiorrhiza* or develop novel breedings. Protection of the wild relatives of the targeted plant species will greatly enhance the scale of the genetic diversity and increase resilience to the agricultural system through breeding practice (Ford-Lloyd et al., 2011; Tyack et al., 2020). Therefore, we suggest that wild populations of *S. paramiltiorrhiza* and *S. bowleyana* should be considered together as whole with *S. miltiorrhiza* in wild germplasm conservation practice. For *S. paramiltiorrhiza*, on the basis of its restricted geographical distribution and threatened condition, we suggested that *in situ* conservation efforts should be made such as establishing protected areas and population recovering in natural habitat. Our study also revealed Dabie Mountain as a genetic diversity center and corridor of these three species. This region should be prioritized as a potential area of conservation, which would ensure the survival and maintain

the evolutionary potential of the group. In the meantime, seed bank construction and *ex situ* conservation through botanical gardens are also recommended, as it also effectively preserve the biological and genetic diversity of *S. miltiorrhiza*, *S. bowleyana* and *S. paramiltiorrhiza* for sustainable usage. We hope this study would provide a guidance to the conservation of other medicinal plants when species history are complex and wild relatives are involved.

DATA AVAILABILITY STATEMENT

The raw sequencing data has been successfully uploaded to SRA in NCBI, <https://www.ncbi.nlm.nih.gov/Traces/study/?acc=PRJNA716793>.

AUTHOR CONTRIBUTIONS

Y-KW and Z-CQ conceived and designed the experiments and revised the manuscript. Y-BH and H-WX carried out

investigation and collected materials. XZ, Z-CZ, J-JW, Z-CQ, and Y-KW analyzed the data and wrote the manuscript. Y-KW and Z-CQ acquired the funding. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the Natural Science Foundation of Zhejiang Province (LY21C030008), the Specific Project for Strategic Biological Resources and Technology Supporting System from the Chinese Academy of Sciences (ZSZY-001), the Science and Technology Program of Shanghai Science and Technology Committee (20392000600), the Chenshan Special Foundations from Shanghai Municipal Administration of Forestation and City Appearances (G102404 and F132423), Open Fund of Shanghai Key Laboratory of Plant Functional Genomics and Resources (PFGR202002 and PFGR202104), and the Open

Fund of Shaoxing Academy of Biomedicine of Zhejiang Sci-Tech University (SXAB202020).

ACKNOWLEDGMENTS

We would like to thank Gui-Bin Li for help with data analysis. We also thank the editor and two reviewers for their insightful comments and suggestions on the early version of the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2021.688323/full#supplementary-material>

REFERENCES

- Alexander, A. M., Su, Y., Oliveros, C. H., Olson, K. V., Travers, S. L., and Brown, R. M. (2017). Genomic data reveals potential for hybridization, introgression, and incomplete lineage sorting to confound phylogenetic relationships in an adaptive radiation of narrow-mouth frogs. *Evolution* 71, 475–488. doi: 10.2307/48575966
- Alexander, D. H., Novembre, J., and Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* 19, 1655–1664. doi: 10.1101/gr.094052.109
- Allen, G. C., Flores-Vergara, M. A., Krasynanski, S., Kumar, S., and Thompson, W. F. (2006). A modified protocol for rapid DNA isolation from plant tissues using cetyltrimethylammonium bromide. *Nat. Protoc.* 1, 2320–2325. doi: 10.1038/nprot.2006.384
- Alley, R., and Agustsdottir, A. (2005). The 8k event: cause and consequences of a major Holocene abrupt climate change. *Q. Sci. Rev.* 24, 1123–1149. doi: 10.1016/j.quascirev.2004.12.004
- Andrews, K. R., Good, J. M., Miller, M. R., Luikart, G., and Hohenlohe, P. A. (2016). Harnessing the power of RADseq for ecological and evolutionary genomics. *Nat. Rev. Genet.* 17, 81–92. doi: 10.1038/nrg.2015.28
- Astakhov, V. (2008). Geographical extremes in the glacial history of northern Eurasia: post-QUEEN considerations. *Polar Res.* 27, 280–288. doi: 10.1111/j.1751-8369.2008.00069.x
- Bentley, R. (2010). *Medicinal plants*. London: Domville-Fife Press.
- Boczkowska, M., Bączek, K., Kosakowska, O., Rucińska, A., Podyma, W., and Węglarz, Z. (2020). Genome-wide diversity analysis of *Valeriana officinalis* L. using DArT-seq derived SNP markers. *Agronomy* 10:1346. doi: 10.3390/agronomy10091346
- Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. doi: 10.1093/bioinformatics/btu170
- Canter, P. H., Thomas, H., and Ernst, E. (2005). Bringing medicinal plants into cultivation: opportunities and challenges for biotechnology. *Trends Biotechnol.* 23, 180–185. doi: 10.1016/j.tibtech.2005.02.002
- Cerca, J., Rivera-Colón, A. G., Ferreira, M. S., Ravinet, M., Nowak, M. D., Catchen, J. M., et al. (2021). Incomplete lineage sorting and ancient admixture, and speciation without morphological change in ghost-worm cryptic species. *PeerJ*. 9:e10896. doi: 10.7717/peerj.10896
- Chen, S., Yu, H., Luo, H., Wu, Q., Li, C., and Steinmetz, A. (2016). Conservation and sustainable use of medicinal plants: problems, progress, and prospects. *Chin. Med.* 11:37. doi: 10.1186/s13020-016-0108-7
- Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., et al. (2011). The variant call format and VCFtools. *Bioinformatics* 27, 2156–2158. doi: 10.1093/bioinformatics/btr330
- Davey, J. W., Hohenlohe, P. A., Etter, P. D., Boone, J. Q., Catchen, J. M., and Blaxter, M. L. (2011). Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nat. Rev. Genet.* 12, 499–510. doi: 10.1038/nrg3012
- Doebley, J. F., Gaut, B. S., and Smith, B. D. (2006). The Molecular genetics of crop domestication. *Cell* 127, 1309–1321. doi: 10.1016/j.cell.2006.12.006
- Finlayson, C., and Carrión, J. S. (2007). Rapid ecological turnover and its impact on Neanderthal and other human populations. *Trends Ecol. Evol.* 22, 213–222. doi: 10.1016/j.tree.2007.02.001
- Ford-Lloyd, B. V., Schmidt, M., Armstrong, S. J., Barazani, O., Engels, J., Hadas, R., et al. (2011). Crop Wild Relatives—Undervalued, Underutilized and under Threat? *BioScience* 61, 559–565. doi: 10.1525/bio.2011.61.7.10
- Fu, X., Dou, J., Mao, J., Su, H., Jiao, W., Zhang, L., et al. (2013). RADtyping: an integrated package for accurate de novo codominant and dominant RAD genotyping in mapping populations. *PLoS One* 8:e79960. doi: 10.1371/journal.pone.0079960
- Goetze, M., Zanella, C. M., Palma-Silva, C., Büttow, M. V., and Bered, F. (2017). Incomplete lineage sorting and hybridization in the evolutionary history of closely related, endemic yellow-flowered *Aechmea* species of subgenus *Ortgiesia* (Bromeliaceae). *Am. J. Bot.* 104, 1073–1087. doi: 10.3732/ajb.1701013
- Govindaraj, M., Vetriventhan, M., and Srinivasan, M. (2015). Importance of genetic diversity assessment in crop plants and its recent advances: An overview of its analytical perspectives. *Genet. Res. Int.* 2015, 1–14. doi: 10.1155/2015/431487
- Guindon, S., Dufayard, J., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Syst. Biol.* 59, 307–321. doi: 10.1093/sysbio/syq010
- Hamilton, A. C. (2004). Medicinal plants, conservation and livelihoods. *Biodiv. Conserv.* 13, 1477–1517. doi: 10.1023/B:BIOC.0000021333.23413.42
- Hamilton, A. C. (2008). *Medicinal plants in conservation and development: case studies and lessons learnt*. Salisbury, UK: Plantlife International.
- Hao, D. C., Gu, X., and Xiao, P. G. (2015). “14 - Phytochemical and biological research of *Salvia* medicinal resources,” in *Medicinal Plants* (587–639), eds D. C. Hao, X. Gu, and P. G. Xiao (Sawston: Woodhead Publishing).
- Heywood, V. H., and Iriondo, J. M. (2003). Plant conservation: old problems, new perspectives. *Biol. Conserv.* 113, 321–335. doi: 10.1016/S0006-3207(03)00121-6
- Hu, G., Takano, A., Drew, B. T., Liu, E., Soltis, D. E., Soltis, P. S., et al. (2018). Phylogeny and staminal evolution of *Salvia* (Lamiaceae, Nepetoideae) in East Asia. *Ann. Bot.* 2018:5045153. doi: 10.1093/aob/mcy104/5045153
- Huang, X., and Li, H. (1981). Notes on *Salvia* Linn. in China. *Acta Phytotaxonom. Sin.* 19, 245–249.

- Jamshidi-Kia, F., Lorigooini, Z., and Amini-Khoei, H. (2018). Medicinal plants: Past history and future perspective. *J. Herbm. Pharm.* 7, 1–7. doi: 10.15171/jhp.2018.01
- Johnson, M. T. J., Prashad, C. M., Lavoignat, M., and Saini, H. S. (2018). Contrasting the effects of natural selection, genetic drift and gene flow on urban evolution in white clover (*Trifolium repens*). *Proc. R. Soc. B: Biol. Sci.* 285:20181019. doi: 10.1098/rspb.2018.1019
- Jombart, T., and Ahmed, I. (2011). adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics* 27, 3070–3071. doi: 10.1093/bioinformatics/btr521
- Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K., von Haeseler, A., and Jermini, L. S. (2017). ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods* 14:587. doi: 10.1038/nmeth.4285
- Kamvar, Z. N., Brooks, J. C., and Grünwald, N. J. (2015). Novel R tools for analysis of genome-wide population genetic data with emphasis on clonality. *Front. Genet.* 6:208. doi: 10.3389/fgene.2015.00208
- Kum, K. Y. (2019). *Herbal products' quality control using 1H-NMR metabolomics approach – An example from Danshen (Salvia miltiorrhiza Bunge)*. Ph.D., London: University College London.
- Lee, K. J., Lee, J., Sebastin, R., Cho, G., and Hyun, D. Y. (2020). Molecular genetic diversity and population structure of ginseng germplasm in RDA-Genebank: Implications for breeding and conservation. *Agronomy* 10:68. doi: 10.3390/agronomy10010068
- Lee, S., Choi, J., Lee, B., Yu, J., and Lim, C. E. (2018). Genetic diversity and structure of an endangered medicinal herb: implications for conservation. *AoB Plants* 10:y21. doi: 10.1093/aobpla/ply021
- Li, Q., Li, M., Yuan, Q., Cui, Z., Huang, L., and Xiao, P. (2013). Phylogenetic relationships of *Salvia* (Lamiaceae) in China: Evidence from DNA sequence datasets. *J. Syst. Evol.* 51, 184–195. doi: 10.1111/j.1759-6831.2012.00232.x
- Li, R., Yu, C., Li, Y., Lam, T., Yiu, S., Kristiansen, K., et al. (2009). SOAP2: an improved ultrafast tool for short read alignment. *Bioinformatics* 25, 1966–1967. doi: 10.1093/bioinformatics/btp336
- Lin, H., and Chang, W. (2000). Diterpenoids from *Salvia miltiorrhiza*. *Phytochemistry* 53, 951–953. doi: 10.1016/S0031-9422(99)00433-1
- Liu, D., Zhang, L., Wang, J., and Ma, Y. (2020). Conservation genomics of a threatened *Rhododendron*: Contrasting patterns of population structure revealed from neutral and selected SNPs. *Front. Genet.* 11:00757. doi: 10.3389/fgene.2020.00757
- Liu, X., and Fu, Y. (2015). Exploring population size changes using SNP frequency spectra. *Nat. Genet.* 47, 555–559. doi: 10.1038/ng.3254
- Minh, B. Q., Nguyen, M. A. T., and von Haeseler, A. (2013). Ultrafast Approximation for Phylogenetic Bootstrap. *Mol. Biol. Evol.* 30, 1188–1195. doi: 10.1093/molbev/mst024
- Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., von Haeseler, A., et al. (2020). IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Mol. Biol. Evol.* 37, 1530–1534. doi: 10.1093/molbev/msaa015
- Peakall, R., and Smouse, P. E. (2012). GenA1Ex 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics* 28, 2537–2539. doi: 10.1093/bioinformatics/bts460
- Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S., and Hoekstra, H. E. (2012). Double digest RADseq: An inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS One*. 7:e37135. doi: 10.1371/journal.pone.0037135
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., et al. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 81, 559–575. doi: 10.1086/519795
- R Development Core Team (2018). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.
- Rambaut, A., Drummond, A. J., Xie, D., Baele, G., and Suchard, M. A. (2018). Posterior summarisation in Bayesian phylogenetics using Tracer 1.7. *Syst. Biol.* 67, 901–904.
- Rochette, N. C., Rivera Colón, A. G., and Catchen, J. M. (2019). Stacks 2: Analytical methods for paired-end sequencing improve RADseq-based population genomics. *Mol. Ecol.* 28, 4737–4754. doi: 10.1111/mec.15253
- Sarwat, M., Nabi, G., Das, S., and Srivastava, P. S. (2011). Molecular markers in medicinal plant biotechnology: past and present. *Crit. Rev. Biotechnol.* 32, 74–92. doi: 10.3109/07388551.2011.551872
- Schippmann, U., Leaman, D., and Cunningham, A. (2002). *Impact of Cultivation and Gathering of Medicinal Plants on Biodiversity: Global Trends and Issues Biodiversity and the ecosystem approach in agriculture, forestry and fisheries (142–167)*. Rome: Inter-Departmental Working Group on Biological Diversity for Food and Agriculture.
- Schippmann, U., Leaman, D., and Cunningham, A. (2006). A comparison of cultivation and wild collection of medicinal and aromatic plants under sustainability aspects. *Med. Aromat. Plants* 17:6. doi: 10.1007/1-4020-5449-1_6
- Sun, X., Luo, H., and Sakai, T. (1992). Pentacyclic Triterpenes from *Salvia paramiltiorrhiza*. *Japan. Soc. Pharm.* 46, 202–205.
- Sun, X., Luo, H., Sakai, T., and Niwa, M. (1991). Paramiltioic acid, a novel rearranged norabietanoid from *Salvia paramiltiorrhiza*. *Tetrahedron Lett.* 32, 5797–5800. doi: 10.1016/S0040-4039(00)93558-5
- Tyack, N., Dempewolf, H., and Khoury, C. K. (2020). The potential of payment for ecosystem services for crop wild relative conservation. *Plants* 9:101305. doi: 10.3390/plants9101305
- Vining, K. J., Hummer, K. E., Bassil, N. V., Lange, B. M., Khoury, C. K., and Carver, D. (2020). Crop wild relatives as germplasm resource for cultivar improvement in mint (*Mentha L.*). *Front. Plant Sci.* 11:01217. doi: 10.3389/fpls.2020.01217
- Wang, B. (2010). *Salvia miltiorrhiza*: Chemical and pharmacological review of a medicinal plant. *J. Med. Plants Res.* 4, 2813–2820. doi: 10.1590/S0102-695X2011005000024
- Wang, S., Liu, P., Lv, J., Li, Y., Cheng, T., Zhang, L., et al. (2016). Serial sequencing of isologous RAD tags for cost-efficient genome-wide profiling of genetic and epigenetic variations. *Nat. Protoc.* 11, 2189–2200. doi: 10.1038/nprot.2016.133
- Wang, S., Meyer, E., McKay, J. K., and Matz, M. V. (2012). 2b-RAD: a simple and flexible method for genome-wide genotyping. *Nat. Methods* 9, 808–810. doi: 10.1038/nmeth.2023
- Wang, X., Liu, X., Ko, Y., Jin, X., Sun, J., Zhao, Z., et al. (2020). Genetic diversity and phylogeography of the important medicinal herb, cultivated Huang-Lian populations, and the wild relatives *Coptis* species in China. *Front. Genet.* 11:00708. doi: 10.3389/fgene.2020.00708
- WHO, IUCN, and WWF (1993). *Guidelines on the Conservation of Medicinal Plants*. Gland: International Union for Conservation of Nature and Natural Resources.
- Wicks, K., and Mithen, S. (2014). The impact of the abrupt 8.2 ka cold event on the Mesolithic population of western Scotland: a Bayesian chronological analysis using 'activity events' as a population proxy. *J. Archaeol. Sci.* 45, 240–269. doi: 10.1016/j.jas.2014.02.003
- Wilson, G. A., and Rannala, B. (2003). Bayesian inference of recent migration rates using multilocus genotypes. *Genetics* 163, 1177–1191. doi: 10.1093/eurpub/13.1.11
- Xu, H., Song, J., Luo, H., Zhang, Y., and Li, Q. (2016). Analysis of the genome sequence of the medicinal plant *Salvia miltiorrhiza*. *Mol. Plant.* 6, 949–952. doi: 10.1016/j.molp.2016.03.010
- Yang, A., Wu, B., Shen, C., Zhang, G., Liu, H., Guo, H., et al. (2017). Microsatellite records for volume 9, issue 3. *Conserv. Genet. Resour.* 9, 507–511. doi: 10.1007/s12686-017-0806-1
- Zhang, Y., Yan, Y., Wu, Y., Hua, W., Chen, C., Ge, Q., et al. (2014). Pathway engineering for phenolic acid accumulations in *Salvia miltiorrhiza* by combinational genetic manipulation. *Metab. Eng.* 21, 71–80. doi: 10.1016/j.mben.2013.10.009

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 Zhou, Zhang, Huang, Xiao, Wu, Qi and Wei. 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.