

Genetic Diversity and Phylogeography of cultivated and wild *Ophiopogon japonicus* based on chloroplast DNA exploration for origin and sustainable cultivation

Luying Zhao

Zhejiang Chinese Medical University

Yi Shen

Zhejiang Chinese Medical University

Qiaoyan Zhang

Zhejiang Chinese Medical University

Sha Liu

Zhejiang Chinese Medical University

Qiuru Ren

Zhejiang Chinese Medical University

Luping Qin (✉ lpqin@zcmu.edu.cn)

Zhejiang Chinese Medical University

Yiqi Sun

Zhejiang Chinese Medical University

Research Article

Keywords: Genetic diversity, Origin, Domestication, *Ophiopogon japonicus*, Conservation

Posted Date: January 16th, 2023

DOI: <https://doi.org/10.21203/rs.3.rs-2229993/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at BMC Plant Biology on May 8th, 2023. See the published version at <https://doi.org/10.1186/s12870-023-04247-2>.

Abstract

Background

Ophipogon japonicus, planted in Sichuan (CMD) and Zhejiang (ZMD) Province in China, has a long cultivation history that can trace back to Ming dynasty (500 years ago) and Song dynasty (600 years ago) respectively. During the long period of domestication, the genetic diversity of cultivated *O. japonicus* has decreased a lot so that this species is facing with plant degradation. Therefore, it is urgent to clarify the genetic diversity and origin of cultivated *O. japonicus* to establish a theoretical basis for the utilization and conservation of the genetic resources of *O. japonicus*.

Result

We analyzed the genetic diversity and population structure of 266 *O. japonicus* individuals from 23 sampling sites based on 4 chloroplast DNA sequences (*atpB-rbcL*, *rpl16*, *psbA-trnH* and *rpl20-5' rps12*), identified the effects of domestication on genetic diversity of cultivars and determined the geographic origins of cultivars. The results showed that cultivated *O. japonicus* and wild *O. japonicus* had 4 haplotypes and 15 haplotypes respectively. The genetic diversity of two cultivars ($H_d = 0.35700$, $\pi = 0.06667$) was much lower than the wild populations ($H_d = 0.76200$, $\pi = 0.20378$), and the level of genetic diversity in CMD ($H_d = 0.01900$, $\pi = 0.00125$) was lower than that in ZMD ($H_d = 0.06900$, $\pi = 0.01096$). Genetic differentiation between the cultivated and the wild ($F_{ST} = 0.82044$) was significant, especially between the two cultivars ($F_{ST} = 0.98254$). This species showed a more pronounced phylogeographical structure ($N_{ST} > G_{ST}$, $P < 0.05$). The result of phylogenetic tree showed that demonstrated with the outgroup *O. amblyphyllus* Wang et Dai, the genetic difference between CMD and ZMD is not enough to distinguish the cultivars from two producing area. Besides, both CMD and ZMD have a closer relationship with wild populations in Sichuan than that in Zhejiang. In addition, according to the TCS network and MaxEnt suggested that the wild population TQ located in Sichuan Province should be served as the ancestor of cultivated *O. japonicus*, which is supported by RASP analysis.

Conclusion

These results suggest that cultivated *O. japonicus* has experienced dramatic loss of genetic diversity under anthropogenic influence. The genetic differentiation between CMD and ZMD was likely to be influenced by founder effect and strong artificial selection for plant traits. It appears that wild populations in Sichuan area were involved in the origin of not only CMD but also ZMD. Furthermore, some scientific strategies are suggested for resource conservation of *O. japonicus* based on its genetic diversity and population structure in this study.

Background

The use of medicinal plants is in an expanding commercial market over the past 50 years according to reports from China, Brazil and other countries [1–3]. The sharp increase in the usage and sales of herbal medicines worldwide, the habitat loss and other issues contribute to the reducing resource of wild medicinal plants. Consequently, several governments and organizations have advised that wild medicinal species be included in cultivation systems to cope with the resource shortage of wild medicinal plants [4]. However, previous studies on the genetic diversity of crops discovered that the domestication of wild plants was influenced by the founder effect and artificial selection to meet human needs, which may lead to a genetic bottleneck and eventually result in a decrease of genetic diversity [5]. As for medicinal plants, the impact of domestication for different medicinal plants is different due to different factors such as domestication time and propagation mode [6]. For instance, cultivated *Atractylodes macrocephala* had a higher genetic diversity than its wild populations [7], whereas cultivated *Angelica dahurica* had lower genetic diversity than its wild populations [8]. Accordingly, focusing on the effects of domestication on plants can assist to explore the changes in genetic diversity of medicinal plant during cultivation so as to provide essential information for the appropriate conservation and cultivation management strategies of medicinal plants.

Evaluation of genetic diversity is critical for the breeding processes and effective conservation strategies of genetic resources because genetic diversity provides the basis for evolutionary change, including acclimation and adaptive responses to changes in local environmental conditions [9]. Based on the assessment of genetic diversity, conservation strategies for medicinal plants like *Salvia officinalis* [10], *Vitex rotundifolia* [11] and *Gloriosa superba* [12] have been made. Nevertheless, the genetic diversity of many medicinal plants has not been studied, resulting in our ignorance of their gene pool and the extent of genetic structure changes. Therefore, a more comprehensive study on the genetic diversity of important medicinal plants can provide necessary information for understanding the impact of domestication on medicinal plants, formulating appropriate conservation strategies and making efforts for future breeding.

Genetic information gained from evaluation of genetic diversity also provides the basis for phylogeography. Phylogeography, combined genealogical and geographical information, has become a powerful tool for investigating the historical processes (fragmentation, expansion, and long-distance colonization) and gene flow in shaping geographical molecular variation [13, 14]. Phylogeographical approaches are well suitable to understand the cultivation origin and domestication history because cultivated species usually include many different degrees of divergence in the evolutionary lineages resulting from recent differentiation processes in certain populations [15, 16]. Accordingly, phylogeography can not only provide advice for species conservation, but also offer information for the breeding of cultivated plants. For example, using phylogeography, Wang found the ancestor of *Coptis chinensis* var. *chinensis* and *C. deltoidea* was first colonized Sichuan [17], while Jafari discovered that *Taraxacum officinale* was putatively originated in Europe [18]. Based on the exploration of the complex history of species dispersal and vicariance, relevant suggestions for the conservation of wild genetic resources and the improvement of cultivated medicinal plants are then proposed. Thus, it is appropriate to use phylogeographic approaches to understand the demographic and historical processes of medicinal plant and to further propose suitable conservation and sustainable cultivation practices for medicinal plants.

Ophiopogon japonicus (L. f.) Ker-Gawl., a typical domesticated medicinal plant whose cultivation history, is distributed in Sichuan, Zhejiang, Yunnan, Hebei, Henan, Shandong, Shaanxi, Jiangsu and other provinces in China. The tuberous root of *O. japonicus* has been used for the treatment of coughs, sore throats, thirstiness, insomnia and constipation in traditional Chinese medicine (TCM) over 2000 years due to the functions of nourishing Yin, promoting the production of body fluid, moistening lungs and clearing away internal heat [19]. Zhe maidong (ZMD) and Chuan maidong (CMD) are two genuine cultivars of *O. japonicus* in China planted in Zhejiang and Sichuan respectively. Zhejiang became a famous cultivation region of *O. japonicus* in the Song Dynasty, and was identified as a genuine cultivation region in the Ming Dynasty. While the cultivation history of CMD was first recorded in Ming Dynasty. It was not until Qing Dynasty that Sichuan became another genuine cultivation region of *O. japonicus* [20]. Different geographical location led to the different growing environment of the two cultivars [21]. Despite these external factors, CMD and ZMD have had stable mutation loci in chloroplast DNA (cpDNA) and nuclear genome [21, 22]. All these phenomena may cause different chemical constituents [23–25] and pharmacological activities [26–28] between ZMD and CMD, which would play pivotal roles in recognizing a superior clinical efficacy of ZMD over CMD. However, previous studies revealed that the genetic diversity of *O. japonicus* has decreased a lot which will affect the long-term survival and evolutionary potential of this species, and then affect the selection and breeding of high-quality medicinal herb [29, 30]. Hence, further study on evaluating the genetic diversity and locating the original site of cultivated *O. japonicus* for designing strategies of the protection of genetic resources and the breeding of new varieties is urgent.

Currently, we know little about the genetic variation background of *O. japonicus*, especially in studies concerning the population genetic structure and phylogeography. In order to provide practical information essential for sustainable cultivation of *O. japonicus*, we used 4 chloroplast DNA regions (*atpB-rbcL*, *rpl16*, *psbA-trnH* and *rpl20-5' rps12*) to figure out the population genetics and phylogeography of wild and cultivated *O. japonicus*, which can provide practical information necessary for the origin exploration and sustainable cultivation of *O. japonicus*. More specific objectives were as follows: (1) to evaluate the genetic diversity of cultivated populations under artificial selection and wild populations; (2) to clarify the genetic divergence between cultivated populations of *O. japonicus* from Sichuan Province and Zhejiang Province; (3) to reveal the geographical origins and ancestral populations of cultivated *O. japonicus*; and (4) to determine which wild populations can be introduced for the improvement of cultivation.

Results

Genetic Diversity and Differentiation of *O. japonicus*

One cpDNA introns (*rpl16*) and three cpDNA spacers (*atpB-rbcL*, *psbA-trnH* and *rpl20-5' rps12*) were chosen to analyze the genetic variation in *O. japonicus* from different geographical locations after preliminary screening for variation. Based on the 266 individuals representing 23 populations of *O. japonicus*, the nucleotide sequence lengths were 698 bp, 523 bp, 479 bp and 708 bp for the *atpB-rbcL*, *rpl16*, *psbA-trnH* and *rpl20-5' rps12* regions, respectively. The combined cpDNA sequence length after multiple alignments was 2,408bp. In total, 107 polymorphic sites and 19 haplotypes were detected from 10 cultivated populations and 13 wild populations of *O. japonicus*. The variation rate of the combined cpDNA sequence was 4.44%, and the aligned sequence variations are summarized in Table 1.

Genetic diversity of *O. japonicus* was estimated by the number of haplotypes (H), haplotype diversity (H_d) and nucleotide diversity (π) in *O. japonicus* (Table 2). 19 haplotypes were obtained from 23 populations, including 4 haplotypes obtained from cultivated *O. japonicus* and 15 haplotypes obtained from wild *O. japonicus* (Fig. 1). The H_d and π in all species of *O. japonicus* were 0.74500 and 0.00712 respectively. The H_d and π of CMD (0.01900; 0.00006) was lower than that of ZMD (0.06900; 0.00049) and wild populations (0.76200; 0.00905). Furthermore, the result of permutcpSSR found that the genetic diversity of all 13 wild populations ($h_T = 0.849$, $h_S = 0.341$, $V_T = 0.859$, $V_S = 0.204$) was much higher than that of two cultivars ($h_T = 0.521$, $h_S = 0.030$, $V_T = 0.842$, $V_S = 0.067$). These results indicated the genetic diversity of *O. japonicus* has declined a lot under anthropogenic influence during the cultivation.

According to the AMOVA analysis (Table 3), genetic differentiation of CMD and ZMD occurred mainly within populations (100.52% and 100.56%), which is different from wild *O. japonicus* whose variation mainly occurred among populations (70.09%). When comparing different groups, we found huge genetic variation between CMD and ZMD (98.37%), which is consistent with the low gene flow between the two cultivars ($N_m = 0.04$). The results means that there has been a high genetic differentiation between CMD and ZMD. What's more, the percentage of genetic variation between CMD and wild group (40.30%) was much higher than that between ZMD and wild group (8.64%). This is because the gene flow between ZMD and wild group ($N_m = 1.490$) can resist genetic differentiation caused by genetic drift among populations between ZMD and wild groups ($F_{ST} = 0.72934$), while the gene flow between CMD and wild group ($N_m = 0.830$) is insufficient to resist the genetic differentiation caused by migration or genetic drift between CMD and wild group ($F_{ST} = 0.82566$).

Table 1
Variable sites of the aligned sequences of 4 cpDNA in nineteen haplotypes (H1-H19) of *O. japonicus*

Haplotype		H20	H1	H2	H3	H4	H5	H6	H7	H8	H9	H10	H11	H12	H13	H14	H15	H16	H17	H18	H19
cpDNA	59	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	-
	60	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	-
	61	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	-
	62	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	-
	63	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	-
	192	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.
	204	T	C	C	.	.	C	C	C	.	.	.	.	.	.	.	.	.	.	.	.
	309	A	.	.	.	.	.	.	.	.	.	.	.	-	.	.	.	.	.	.	.
	404	G	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.
	545	G	.	.	A	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	666	T	.	.	.	.	.	.	-	.	.	.	.	.	.	.	.	.	.	.	.
	667	A	.	.	.	.	.	.	-	.	.	.	.	.	.	.	.	.	.	.	.
	668	T	.	.	.	.	.	.	-	.	.	.	.	.	.	.	.	.	.	.	.
	669	T	.	.	.	.	.	.	-	.	.	.	.	.	.	.	.	.	.	.	.
	670	T	.	.	.	.	.	.	-	.	.	.	.	.	.	.	.	.	.	.	.
	671	T	.	.	.	.	.	.	-	.	.	.	.	.	.	.	.	.	.	.	.
	672	C	.	.	.	.	.	.	-	.	.	.	.	.	.	.	.	.	.	.	.
	673	G	.	.	.	.	.	.	-	.	.	.	.	.	.	.	.	.	.	.	.
	674	A	.	.	.	.	.	.	-	.	.	.	.	.	.	.	.	.	.	.	.
	675	T	.	.	.	.	.	.	-	.	.	.	.	.	.	.	.	.	.	.	.
	676	T	.	.	.	.	.	.	-	.	.	.	.	.	.	.	.	.	.	.	.
	677	C	.	.	.	.	.	.	-	.	.	.	.	.	.	.	.	.	.	.	.
	678	A	.	.	.	.	.	.	-	.	.	.	.	.	.	.	.	.	.	.	.
	679	G	.	.	.	.	.	.	-	.	.	.	.	.	.	.	.	.	.	.	.
	723	A	.	-	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	724	G	.	-	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	725	C	.	-	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	726	A	.	-	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	727	G	.	-	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	728	G	.	-	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	729	A	.	-	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	756	C	T	T	T	T	T	T	T	T	T	.	.	.	T	T	T	T	T	.	.
	758	T	.	.	.	.	.	.	.	-	-	.	.	.	.	.	.	.	.	.	.
	759	A	.	.	.	.	.	.	.	-	-	.	.	.	.	.	.	.	.	.	.
	760	T	.	.	.	.	.	.	.	-	-	.	.	.	.	.	.	.	.	.	.
	761	T	.	.	.	.	.	.	.	-	-	.	.	.	.	.	.	.	.	.	.
	762	T	.	.	.	.	.	.	.	-	-	.	.	.	.	.	.	.	.	.	.
	763	A	.	.	.	.	.	.	.	-	-	.	.	.	.	.	.	.	.	.	.
	764	T	.	.	.	.	.	.	.	-	-	.	.	.	.	.	.	.	.	.	.
	765	T	.	.	.	.	.	.	.	-	-	.	.	.	.	.	.	.	.	.	.
	766	T	.	.	.	.	.	.	.	-	-	.	.	.	.	.	.	.	.	.	.

Haplotype	H20	H1	H2	H3	H4	H5	H6	H7	H8	H9	H10	H11	H12	H13	H14	H15	H16	H17	H18	H19
794	G	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A
907	T	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.
917	-	C	C	C	C	C	C	C	-	-	C	-	-	C	C	C	C	C	-	-
918	-	C	C	C	C	C	C	C	-	-	C	-	-	C	C	C	C	C	-	-
919	-	A	A	A	A	A	A	A	-	-	A	-	-	A	A	A	A	A	-	-
920	-	T	T	T	T	T	T	T	-	-	T	-	-	T	T	T	T	T	-	-
921	-	T	T	T	T	T	T	T	-	-	T	-	-	T	T	T	T	T	-	-
949	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	T	T	T	-	-
950	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	T	T	T	-	-
951	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	A	A	A	-	-
952	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	A	A	A	-	-
953	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	A	A	A	-	-
954	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	T	T	T	-	-
955	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	A	A	A	-	-
956	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	G	G	G	-	-
957	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	A	A	A	-	-
958	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	T	T	T	-	-
959	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	G	G	G	-	-
960	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	T	T	T	-	-
961	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	G	G	G	-	-
962	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	T	T	T	-	-
963	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	A	A	A	-	-
964	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	T	T	T	-	-
965	-	-	-	-	-	-	-	-	-	-	-	-	-	-	C	C	C	C	-	-
966	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	T	T	T	-	-
967	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	G	G	G	-	-
968	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	T	T	T	-	-
1055	G	A	A	.	.	.	A	.	.	.	A	.	.	A	.	.	.	.	.	.
1184	G	A	A	A	A	A	A	A	.	.	.	.	.	A	A	A	A	A	.	.
1678	C	T	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
1733	C	A	A	A	A	A	A	A	.	.	A	.	.	A	A	A	A	A	.	.
1789	G	.	.	.	.	.	.	.	.	.	.	.	.	.	A	A	A	A	.	.
1844	-	-	-	A	-	A	-	A	-	-	-	-	-	-	A	A	-	A	-	-
1845	-	-	-	T	-	T	-	T	-	-	-	-	-	-	T	T	-	T	-	-
1846	-	-	-	T	-	T	-	T	-	-	-	-	-	-	T	T	-	T	-	-
1847	-	-	-	T	-	T	-	T	-	-	-	-	-	-	T	T	-	T	-	-
1848	-	-	-	T	-	T	-	T	-	-	-	-	-	-	T	T	-	T	-	-
1849	-	-	-	A	-	A	-	A	-	-	-	-	-	-	A	A	-	A	-	-
1850	-	-	-	A	-	A	-	A	-	-	-	-	-	-	A	A	-	A	-	-
1851	-	-	-	A	-	A	-	A	-	-	-	-	-	-	A	A	-	A	-	-
1852	-	-	-	T	-	T	-	T	-	-	-	-	-	-	T	T	-	T	-	-
1853	-	-	-	T	-	C	-	C	-	-	-	-	-	-	C	C	-	C	-	-
1854	-	-	-	C	-	C	-	C	-	-	-	-	-	-	C	C	-	C	-	-

Haplotype	H20	H1	H2	H3	H4	H5	H6	H7	H8	H9	H10	H11	H12	H13	H14	H15	H16	H17	H18	H19
1855	-	-	-	T	-	T	-	T	-	-	-	-	-	-	T	T	-	T	-	-
1856	-	-	-	T	-	T	-	T	-	-	-	-	-	-	T	T	-	T	-	-
1857	-	-	-	C	-	C	-	C	-	-	-	-	-	-	C	C	-	C	-	-
1858	-	-	-	C	-	C	-	C	-	-	-	-	-	-	C	C	-	C	-	-
1859	-	-	-	A	-	A	-	A	-	-	-	-	-	-	A	A	-	A	-	-
1860	-	-	-	G	-	G	-	G	-	-	-	-	-	-	G	G	-	G	-	-
1866	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	C
1871	C	.	.	.	.	.	.	.	.	.	.	T	T	.	.	.	.	.	T	T
1909	A	-	-	-	-	-	-	-	.	.	-	.	.	-	-	-	-	-	.	.
1910	T	-	-	-	-	-	-	-	.	.	-	.	.	-	-	-	-	-	.	.
1911	A	-	-	-	-	-	-	-	.	.	-	.	.	-	-	-	-	-	.	.
1912	T	-	-	-	-	-	-	-	.	.	-	.	.	-	-	-	-	-	.	.
1913	T	-	-	-	-	-	-	-	.	.	-	.	.	-	-	-	-	-	.	.
1914	A	-	-	-	-	-	-	-	.	.	-	.	.	-	-	-	-	-	.	.
1951	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.
2169	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	-	.	.	.
2170	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	-	.	.	.
2171	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	-	.	.	.
2172	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	-	.	.	.
2173	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	-	.	.	.
2174	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	-	.	.	.
2175	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	-	.	.	.
2176	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	-	.	.	.
2177	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	-	.	.	.

Notes: All sequences are compared to the reference Hap20. "-" in sequences denote absence.

Table 2
Genetic diversity analysis based on cpDNA of *O. japonicus*

Population code	H	H_d	π	Tajima D	Fu's FS
Cultivated in Sichuan (CMD)					
YM	1	0	0	0	0
HY	1	0	0	0	0
W-HY	1	0	0	0	0
XD	1	0	0	0	0
QL	1	0	0	0	0
LF	2	0.125	0.00036	-2.06208	2.38195
				**	
LS	1	0	0	0	0
Species level	2	0.019	0.00125	-2.09486	-0.02002

Cultivated in Zhejiang (ZMD)					
FM	1	0	0	0	0
QS	1	0	0	0	0
SM	2	0.2	0.00141	-2.01902	5.88323

Species level	2	0.069	0.01096	-2.51175	3.84807

Wild					
YS	2	0.154	0.00211	-2.30289	8.81879

GY	2	0.566	0.00046	1.84427	2.42917
TQ	4	0.533	0.00232	-0.33388	3.34381
BZ	4	0.705	0.00488	0.52501	9.40284
NC	1	0	0	0	0
FLF	2	0.419	0.004	0.64702	12.53529
JH	2	0.429	0.00036	0.41421	1.65331
DZ	3	0.833	0.00568	-0.74309	3.15685
HC	4	0.803	0.01341	2.7618	14.2524
YH	1	0	0	0	0
SB	1	0	0	0	0
ZJJ	1	0	0	0	0
YC	1	0	0	0	0
Species level	17	0.762	0.20378	0.60829	20.39273

Note: H, the number of haplotypes; H_d , haplotype diversity; π , nucleotide diversity. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

Table 3
Analyses of molecular variance (AMOVA) based on the cpDNA sequences

Source of variation	d.f.	Sum of squares	Variance components	Percentage	Fixatio indices	P-value
of variation						
Wild						
Among populations	12	1016.788	8.09993 Va	70.09	$F_{ST}=0.70087$	0.00000
Within populations	119	411.378	3.45696 Vb	29.91		
Cultivated						
Among populations	9	452.518	3.75356 Va	95.51	$F_{ST}=0.95514$	0.00000
Within populations	124	21.863	0.17631 Vb	4.49		
Cultivated vs. Wild						
Among groups	1	369.238	2.19366 Va	22.09	$F_{SC}=0.76952$	0.00000
Among populations within groups	21	1469.307	5.95263 Vb	59.95	$F_{ST}=0.82044$	0.00000
Within populations	243	433.241	1.78288 Vc	17.96	$F_{CT}=0.22093$	0.00810
CMD						
Among populations	6	0.371	-0.00034 Va	-0.52	$F_{ST}=-0.00516$	1.00000
Within populations	98	6.562	0.06696 Vb	100.52		
ZMD						
Among populations	2	1.114	-0.00327 Va	-0.56	$F_{ST}=-0.00559$	1.00000
Within populations	26	15.300	0.58846 Vb	100.56		
CMD vs. ZMD						
Among groups	1	451.033	9.92018 Va	98.25	$F_{SC}=0.00383$	1.00000
Among populations within groups	8	1.485	0.00068 Vb	0.01	$F_{ST}=0.98254$	0.00000
Within populations	124	21.863	0.17631 Vc	1.75	$F_{CT}=0.98247$	0.00502
CMD vs. Wild						
Among groups	1	585.659	4.45178 Va	40.30	$F_{SC}=0.70800$	0.00000
Among populations within groups	18	1017.159	4.66986 Vb	42.27	$F_{ST}=0.82566$	0.00000
Within populations	217	417.94	1.92599 Vc	17.43	$F_{CT}=0.40296$	0.00000
ZMD vs. Wild						
Among groups	1	117.718	0.93887 Va	8.64	$F_{SC}=0.70376$	0.00000
Among populations within groups	14	1017.902	6.99048 Vb	64.30	$F_{ST}=0.72934$	0.00000
Within populations	145	426.678	2.94261 Vc	27.07	$F_{CT}=0.08636$	0.13675

Note: F_{ST} : Genetic differentiation among populations; F_{SC} : genetic differentiation among population within groups; F_{CT} : Genetic differentiation between groups.

Population Structure And Phylogeographical Analysis

A permutation test showed that N_{ST} was significantly higher than G_{ST} ($N_{ST}/G_{ST} = 0.920/0.753$, $P < 0.01$), which can be indicated the haplotypes geographically close were distributed in the same or contiguous population, and the genetic differentiation among population increased with the increase of geographical distance. It indicated that there is a significant phylogeographic structure in *O. japonicus* cpDNA haplotypes. Mantel test also supported the result that the genetic distances between *O. japonicus* populations were significantly and positively correlated to their geographical distances ($P = 0.01 < 0.05$, $R^2 = 0.111$). With the increase of geographical distance, the genetic distance of *O. japonicus* also showed a trend of increase (Fig. 2). Combined with the low gene flow among populations of *O. japonicus* ($N_m = 0.17$), it can be concluded this species showed a pattern of restricted gene flow with isolated by distance.

The phylogenetic relationships between the 19 cpDNA haplotypes were assessed under Bayesian inferences drawn using *O. amblyphyllus* Wang et Dai (H20) as an outgroup. The result showed that the phylogenetic tree of the 19 cpDNA haplotypes had a comb-like structure (Fig. 3). Haplotypes (H1, H2, H3 and H4) from cultivated populations occurred in different clades, but still grouped together compared with the outgroup. As a result, we could not distinguish CMD from ZMD based on the 4 cpDNA sequences.

Furthermore, the rooted TCS network of cpDNA haplotypes revealed the relationship of the interior (ancestral) and the tip (derived) haplotypes (Fig. 4). H13, the haplotype only found in the TQ population, was inferred as the ancestral haplotype because it lies at the center of TCS network and was connected with other derivative haplotypes. 4 cpDNA haplotypes (H1, H2, H3 and H4) found within cultivated populations, all of which comes from H13. H1, the most frequent and widely distributed haplotype, was shared between CMD and wild *O. japonicus*. H3 was shared between ZMD and wild *O. japonicus*. H2 was shared only by CMD. H4 was shared only by ZMD. The other 15 haplotypes were found only in wild populations. The haplotype distribution also confirmed the low genetic diversity of cultivated *O. japonicus*.

Demographic History

The mismatch distribution for *O. japonicus* was clearly multimodal and inconsistent with the bell-shaped curve expected for an expanding population at specie level (SSD = 0.10627, $P = 0.000 < 0.05$), suggesting that *O. japonicus* did not underwent a recent expansion in the face of favorable climatic and environmental conditions (Fig. 5). Neither did the wild population of *O. japonicus* expand recently (SSD = 0.08589, $P = 0.000 < 0.05$). However, CMD whose SSD-value was 0.00028 ($P = 0.058 > 0.05$) and ZMD whose SSD-value was 0.00401 ($P = 0.160 > 0.05$) had a relatively smooth unimodal distribution, suggesting CMD and ZMD assumed as an expansion model.

At species level, both Tajima's D (-0.03985, $P = 0.799$) and Fu's F_s (20.36184, $P = 0.926$) indicated no significant difference from zero, indicating *O. japonicus* has not passed through a recent demographic expansion. Neither did wild populations expand recently (Tajima's $D = 0.60829$, $P = 0.732$; Fu's $F_s = 20.39273$, $P = 1.000$). However, the value of Tajima's D and Fu's F_s (Table 2) implied that demographic expansion existed in certain populations from CMD and ZMD, which is consistent with the result of mismatch distribution.

Historical Biogeography

The results of S-DIVA ancestral area reconstruction based on RASP showed that *O. japonicus* may originated in Sichuan (A) with subsequent dispersal and vicariance events between the ancestral and areas of present distribution (Fig. 6). These facts support that Sichuan may be an ancestral area of *O. japonicus*.

Species Distribution Modeling

Based on the geographical distribution of *O. japonicus*, MaxEnt was used to simulate the potential distribution of *O. japonicus* during LIG, LGM, MID and present, as shown in Fig. 7. The AUC scores that were the average of 10 runs were higher than 0.8 for the four models built, which supported their predictive power [31].

Simulation results of species distribution showed that *O. japonicus* distributed in most areas of China during LIG period (Fig. 7-A). But the cold climate deprived the previously hospitable environment of this species, leading to a large-scale migration. *O. japonicus* migrated to the east of Hengduan Mountains and south of Qinling Mountains (Fig. 7-B), suggesting that these mountains may have been the refuge of this species during LGM. After LGM period, the distribution ranges of the species had a slight expansion or recolonization in the geographical range (Fig. 7-C). While the distribution range of *O. japonicus* at present is similar to that of MID (Fig. 7-D), indicating that the migration and expansion of *O. japonicus* at MID tends to be basically stable. In addition, the glacial refugium hypothesis (GRH) proposes that glaciers promoted differentiation and generation of intraspecific diversity by isolating populations in ice-free refugia [32], which was also supported by the results of haplotype diversity, geographical analysis and historical biogeography.

Discussion

Genetic diversity and differentiation

For many cultivated crops, the consequences of artificial selection, founder effects and genetic bottlenecks result in considerable decline in genetic diversity [33–35], which may reduce their capacity for long-term survival and evolution in dynamic environments. Due to the comparatively brief domestication period and the exchange of genes with wild populations during extensive cultivation, the genetic diversity of medicinal plants typically declined much milder than that of crops. In our fieldwork, we investigated 13 sites of wild *O. japonicus*, 7 CMD cultivation bases in the villages along the Fujiang River in Sichuan Province, 3 ZMD cultivation bases, 2 of which were situated in traditional genuine producing area Cixi and one of which located in new producing area Santai. The present study indicated that cultivated *O. japonicus* exhibited a striking decrease in genetic diversity ($h_T = 0.521$, $h_S = 0.030$, $V_T = 0.842$, $V_S = 0.067$) compared to wild populations ($h_T = 0.849$, $h_S = 0.341$, $V_T = 0.859$, $V_S = 0.204$) after a few hundred years of domestication history. Out of 19 cpDNA haplotypes in the species, only 4 haplotypes were detected in cultivated populations. All these domestication result of decreased genetic diversity is similar to that of *Scrophularia ningpoensis* [36] and *Angelica dahurica* [8]. These dramatic changes in the genetic diversity of *O. japonicus* during such a short domestication history are most likely caused by founder effect and the clonal mode of reproduction. Firstly, the founder effect is that when a population is established and developed by a small number of individuals, the genetic information carried by these few individuals does not completely reflect the genetic information of their source population, resulting in a low genetic diversity of the resulting new population. The loss of genetic diversity for medical plants like *Allium mongolicum* [37], *Leonurus cardiaca* [38] and *Lycium barbarum* [39] are more or less related to the founder effect. In this study, due to the founder effect, only

2 haplotypes of cultivated *O. japonicus* was gathered from haplotypes of wild populations, meaning that the founders contained only a few samples of the genetic diversity of wild populations. Secondly, while the domestication history of *O. japonicus* (600 years) was shorter than other medical plants (2000 years), the genetic diversity of cultivate *O. Japonicus* dramatically decreased. It is probably attributed to the fact that wild *O. Japonicus* may expand via fruits, whereas the cultivated *O. Japonicus* were influenced by asexual reproduction and artificial selection. In addition, previous researches demonstrated that genetic bottlenecks caused by domestication in annual and perennial fruit crops. It concluded that perennial crops preserved a higher percentage of the genetic variation present in their wild progenitors than annual crops [40]. And juvenile phase length is a principal difference between domesticated annual and domesticated perennial fruit crops. In this study, the cultivation time before harvest was 2 or 3 years for ZMD and 1 year for CMD. ZMD ($H_d = 0.06900$; $\pi = 0.00049$) showed a higher genetic diversity than CMD ($H_d = 0.06900$; $\pi = 0.00049$). And gene flow of ZMD-wild ($N_m=1.490$) is higher than that of CMD-wild ($N_m=0.830$). Thus, it can be inferred that the relatively lengthy juvenile phases coupled with ongoing cultivar–wild gene flow due to the change of habitat may contribute to milder genetic bottlenecks in perennial ZMD than annual CMD, resulting in relatively high genetic diversity in ZMD.

In addition to reducing the genetic diversity of medical plants, domestication also affects the structural patterns of genetic variation [41]. For instance, notable genetic divergence was found between wild and cultivated *Angelica dahurica* and between its two cultivars ($F_{ct} = 0.148$, $P = 0.000$) [8]. Our result of AMOVA analysis indicated that 21.8% of the total genetic variation occurred between cultivated and wild *O. japonicus*, while 98.37% of the total genetic variation occurred between CMD and ZMD (Table 3). Further analysis revealed that the genetic differentiation between ZMD and wild *O. japonicus* ($F_{ST} = 0.72934$) was lower than that between CMD and wild *O. japonicus* ($F_{ST} = 0.82566$). The results of mismatch distribution and neutrality test indicated that both ZMD and CMD had undergone a recent population expansion. Frequent extinctions and recolonizations of local populations can also be an important source of gene flow (Slatkin 1987). It has been reported that ZMD production areas have experienced significant changes in recent years [42]. Due to urbanization and regional economic reasons, most farmers in the traditional production area of ZMD (Hangzhou) no longer plant ZMD. Sanmen County, located in Taizhou City, Zhejiang Province, has gradually become a new producing region of ZMD. As a result, it is probable that one of the factors contributing to the higher gene flow between ZMD and wild populations ($N_m = 1.490$) than between CMD and wild populations ($N_m = 0.830$) is the shrinkage of traditional production regions and the creation of new production areas. In a group of completely isolated populations, genetic drift tends to fix different alleles in different local populations. Genetic drift is the unpredictable change in gene frequency due to finite population size. Gene flow between any populations will prevent complete fixation, but gene flow must exceed a certain level ($N_m > 1$) to prevent substantial genetic differentiation due to genetic drift [43]. This also clarified why the gene flow between ZMD and wild populations ($N_m = 1.490$) is able to withstand the genetic differentiation caused by genetic drift while the gene flow between CMD and wild populations ($N_m = 0.830$) is unable.

As for the genetic differentiation between CMD and ZMD, on the one hand, we believe that it is the result of artificial selection and the isolation by distance. Our surveys conducted in cultivated regions indicated that seedlings of CMD are prostrate CMKY-2, whereas the majority of seedlings of ZMD are erect *O. japonicus* and a small number are prostrate *O. japonicus* [44]. This combination of artificial selection and clonal propagation of *O. japonicus* rapidly fixes the alleles carried by selected individuals, which means that one or two haplotypes are fixed in each cultivar population, leading to large genetic differences between ZMD and CMD. Moreover, the herbal textual researches and the description of the local planters exhibited there was no record about the introduction of one cultivar of *O. japonicus* into another cultivar's growing region, which may have prevented gene flow and then leading to genetic differentiation between CMD and ZMD. Finally, population differentiation is also influenced by geographical distance. Isolation by distance is a model of population structure in which genetic differences between populations increase with geographic scale [45]. In other words, populations that are adjacent to each other are more similar than that are far apart, because they are linked by larger gene exchange, unless some impassable isolation breaks this mechanism and reduces communication between some populations and their neighbors. In our result, a greater N_{ST} (0.920) than G_{ST} (0.753) indicated the presence of phylogeographical structure of *O. japonicus*. Combined with the result of mantel test, it can be inferred that the restricted gene flow with isolated by distance contributed to the population structure of *O. japonicus* [46]. Geographical isolation can contribute to the lack of gene flow among populations, and then leading to the genetic differentiation between two cultivars.

Origin of cultivated *O. japonicus*

Designing strategies for the protection of genetic resources and the breeding of new varieties can be assisted by identifying the original site of domestication and evaluating the evolution history of wild-domesticated species [47]. Before Tang Dynasty (1200 years ago), it was believed that *O. japonicus* came from the wild and mainly distributed in Henan, Jiangsu, Zhejiang and Anhui provinces. On cultivation history, ZMD can trace back to Song dynasty (600 year ago) while CMD can trace back to Ming Dynasty (500 years ago) [20]. As the first study exploring the origin of *O. japonicus*, we explored the ancestral distribution of *O. japonicus* and its historical changes based on 4 chloroplast DNA data. ML phylogenetic tree (Fig. 3) showed that compared with the outgroup, ZMD and CMD grouped as a cluster. Further studies found ZMD and some individuals in the wild population BZ located in Sichuan grouped as a cluster, while CMD and the wild population NC located in Sichuan grouped as another cluster. These results suggested that the cultivated *O. japonicus* is more closely linked to wild populations in Sichuan than the wild *O. japonicus* in Zhejiang, which is similar to study on the genetic diversity of cultivated *O. japonicus* based on ISSR markers [29]. Therefore, the wild populations NC and BZ located in the Sichuan province may have played an important role in the origin of cultivated *O. japonicus*.

The TCS network claimed that 2 haplotypes (H1 and H2) found in CMD and 2 haplotypes (H2 and H4) found in ZMD were derived from H6 and H13 respectively. The 2 haplotypes (H6 and H13) both found only in the wild population BZ and TQ located in Sichuan province. All of the aforementioned haplotypes were derived from H13, which lied at the center of the haplotype network and only found in the wild population TQ located in Sichuan. It may be the ancestor of the two cultivars in this study. Furthermore, as the most frequent and widely distributed haplotype, H1 was found not only in CMD but also in the wild population NC, while H3, the one occurred in most ZMD also found in the wild population BZ, showing that the wild populations in Sichuan made a great contribution to the origin of *O. japonicus*.

Result of historical biogeography and the model of species distribution indicated that the ancestor of *O. japonicus* was located in Sichuan, and then experienced multiple vicariance-dispersal between Sichuan and Zhejiang due to emergence suitable climatic conditions after LGM. This may explain why haplotypes H5 appeared in different populations located in Sichuan, Zhejiang, Hubei, Hunan, Guizhou and other provinces. In addition, simulation results of species distribution also suggested that Sichuan province should be one of the glacial refugia for *O. japonicus*.

In summary, it can be inferred that *O. japonicus* was widely distributed in China during LIG period. With the passage of time, the cold climate caused *O. japonicus* to migrate east of Hengduan Mountains and south of Qinling Mountains during the LGM period. The present distribution was formed after several vicariance-dispersal events on the basis. More importantly, wild populations similar to the extant TQ located in Sichuan Province may be the ancestor of cultivated *O. japonicus*.

Sustainable cultivation of *O. japonicus*

Wild populations have been considered as a critical resource for breeding efforts in medicinal plants because their genetic information are used infrequently and generally exhibit no reproductively isolation from cultivated populations [48]. Previous study comparing cultivated and wild populations of *O. japonicus* indicated that cultivated *O. japonicus* has much lower genetic diversity than that of wild populations (Fig. 1, Table 1). The large-scale cultivation of *O. japonicus* has affected the genetic diversity of wild populations. For example, wild population NC showed a single, identical haplotypes (H1) to CMD. It may result in reduction in plasticity of cultivated *O. japonicus* to respond to changes in climate, pathogen populations, agricultural practices, or quality requirements. Hence, it is necessary to improve the genetic diversity of cultivated *O. japonicus* through sexual reproduction by using wild resources.

Accumulation evidence indicated that a proper assessment of genetic resources of medicinal plant is useful information for the development of conservation plans to protect genetic diversity. For example, protection of medicinal plants like *Angelica dahurica* [8], *Aspidopterys obcordata* var. *obcordata* [49] and *Vitex rotundifolia* [11] were mostly based on the conservation of genetic diversity. Our study exhibited wild populations DZ, HC and BZ had more cpDNA haplotypes and higher genetic diversity ($H_d = 0.83300$, $\pi = 0.00568$; $H_d = 0.80300$, $\pi = 0.01341$; $H_d = 0.70500$, $\pi = 0.00488$) than other populations (Fig. 1, Table 2). In addition, populations located in glacial refugia should be considered as priority areas for conservation because they not only preserve the genetic differentiation of species during the period of climate upheaval, but also serve as the starting point for post-glacial re-dispersal of species which can continue to maintain the continuity of genetic diversity [50–52]. In our study the ENM results suggested that east of Hengduan Mountains and south of Qinling Mountains was the glacial refugia of *O. japonicus* during LIG, while RASP analysis indicated that the ancestor of *O. japonicus* was located in Sichuan. These results demonstrated that Sichuan should be the starting point for post-glacial and re-dispersal of *O. japonicus*. In summary, populations DZ, HC and BZ that showed high genetic diversity as well as populations TQ and BZ that located in Sichuan with high genetic diversity should be served as key objects for biodiversity protection of *O. japonicus* and preferentially selected as the gene pool for improving the genotypes of cultivated *O. japonicus*.

Conclusion

In this study, we accessed the level of genetic structure of *O. japonicus* from 23 sampling sites by using 4 maternally inherited cpDNA primers and investigated the origin of *O. japonicus* for the first time. In our study, it was the anthropogenic influence that led to the low genetic diversity and a high degree of genetic differentiation of *O. japonicus* from two producing areas as well as the high degree of genetic differentiation between cultivated populations and wild populations. In terms of the origin, Sichuan may be one of the origins of *O. japonicus* and wild populations TQ in Sichuan area were recognized as the ancestor of cultivated *O. japonicus*. In addition, we postulate the populations containing high genetic diversity (DZ, HC and BZ) and populations distribution in glacial refugia with high genetic diversity (TQ and BZ) should be protected for the conservation of genetic resource and improvement of genetic diversity for *O. japonicus*. Our findings might contribute to a better understanding of effects of domestication on genetic diversity of medicinal plants and providing genetic insight to improve conservation and sustainable cultivation for *O. japonicus*.

Method

Plant material and DNA extraction

A total of 266 individuals of *O. japonicus* from 23 sampling sites for cpDNA were sampled from distribution areas in China (Fig. 8 and Table 4). Leaf samples were collected from individuals separated by more than 20 m to avoid collecting the same clones (genotypes). The leaf samples were dried in silica gel and then stored in a freezer (-20°C) for use. A sample of *O. amblyphyllus* Wang et Dai from Sichuan was also collected as the outgroup. Genomic DNA was extracted using the Plant Genomic DNA kit (Tiangen Biotech, Beijing, China). For the wild and cultivated plants we used in this study, we obtained relevant permission to collect. All voucher specimens were morphologically identified by professor Luping Qin from the College of Pharmaceutical Sciences, Zhejiang Chinese Medical University (ZCMU) and deposited at the herbarium of ZCMU (voucher ID numbers: 510000190328LY001-510000190328LY013; 330000190630LY001-330000190630LY009; 520000200602LY001; 430000200527LY001; 420000200530LY001). Our field study and experimental research complied with local legislation, national and international guidelines.

Table 4
Details of sample locations, sample sizes

Population code	Locality	Longitude (E)	Latitude (N)	Altitude (m)	Sample size
Cultivated in Sichuan (CMD)					
YM	Tianping, Yongming, SiChuan, China	104°51'58.40"	31°21'38.04"	417	15
HY	Xinglou, Huayuan, SiChuan, China	104°55'05.10"	31°16'08.28"	411	15
W-HY	Qinshan, Huayuan, SiChuan, China	104°53'36.44"	31°17'19.59"	417	15
XD	Maji,Xinde, SiChuan, China	105°05'19.91"	31°07'50.30"	367	15
QL	Wenfeng, Qinglian, SiChuan, China	104°40'40.27"	31°40'20.50"	496	15
LF	Shunjiang, Longfeng, SiChuan, China	104°42'40.61"	31°37'08.97"	480	16
LS	Luocheng, Luoshui, SiChuan, China	104°02'36.99"	31°14'58.92"	675	14
Cultivated in Zhejiang (ZMD)					
FM	Fumin, Andong, ZheJiang, China	121°13'48.21"	30°18'00.94"	3	10
QS	Fujia road, Chongshou, ZheJiang, China	121°15'18.54"	30°15'49.99"	4	10
SM	Nanlin, Hengdu ZheJiang, China	121°26'33.72"	29°00'33.34"	60	10
Wild					
YS	Luocheng, Luoshui, SiChuan, China	104°01'45.66"	31°15'10.62"	817	13
GY	Yueba,Baichao, SiChuan, China	105°25'40.45"	32°22'58.33"	1470	10
TQ	Hongling mountain, Labahe, SiChuan, China	102°34'17.99"	30°02'27.81"	1663	10
BZ	Fangdou mountain, Qunle, SiChuan, China	106°36'28.89"	31°31'57.93"	585	15
NC	Liujiawan, Fengzhan, SiChuan, China	106°19'00.48"	31°48'48.82"	456	11
FLF	Feilai peak, Xihu, ZheJiang, China	120°06'09.84"	30°14'17.92"	90	15
JH	Daxikou, Xilian, ZheJiang, China	120°22'35.34"	28°53'47.66"	374	8
DZ	Danzhu, Danzhu, ZheJiang, China	120°31'02.29"	28°31'13.41"	402	4
HC	Huang, Longnan, ZheJiang, China	119°14'38.54"	27°55'00.70"	1430	12
YH	Meijiujian, Chongtuo, ZheJiang, China	119°27'11.09"	28°01'05.23"	1306	2
SB	Wengtang, Liutang, Guizhou, China	108°12'27.39"	26°58'43.27"	762	13
ZJJ	Xiangshuijian, Sifangxi, Hunan, China	110°10'30.07"	29°39'16.85"	446	7
YC	Mawang mountain, Nanqiao, Hubei, China	111°20'42.22"	30°31'18.84"	204	12
Outgroup					
DY	Hongling mountain, Labahe, SiChuan, China	102°34'07.26"	30°02'56.08"	1787	1

Sequencing Of cpDNA

Polymerase chain reaction (PCR) was performed in an Veriti96 thermocycler (Thermofisher). The reaction mixture (50 µL) contained 20µL of total DNA, 1 µL of each primer, 21 µL of ddH₂O, and 25 µL of 2× Trans Taq High Fidelity (HiFi) PCR SuperMix (Transgen Biotech, Beijing, China). Double-stranded DNA was amplified after 4-min incubation at 94°C; followed by 30 cycles at 94 °C for 50 s, 54°C for 50 s, and 72 °C for 90 s; with a final extension at 72 °C for 7 min. PCR products were identified on 1.0% agarose gels, visualized with ultraviolet light and photographed. DNA Marker DL2000 (Takara Bio Inc.) was used. The PCR products with a single band were cleaned by a EasyPure PCR Purification Kit (TransGen Biotech. Beijing, China) for direct sequencing. Sequencing was conducted from both ends.

Phylogenetic Analyses

Geneious Pro V4.8.3 [53] was used to trim, assemble and align the DNA sequences. Bayesian phylogenetic analyses was performed on MrBayes 3.2.4 [54] to construct the phylogenetic tree based on the combined cpDNA data. The sequence evolution model T92 + G was evaluated by the Akaike Information Criterion (AIC) in Modeltest 3.7.0 [55]. Two independent analyses with four simultaneous Monte Carlo Markov Chains (MCMC) analyses were run for 10 million generations, with sampling done every 100 generations. The first 50,000 sampled trees were discarded as burn-in samples. A statistical parsimony haplotype network estimated the genetic relationships among haplotypes was built using TCS 1.21 [56].

Population Genetic Diversity And Genetic Differentiation

The number of haplotypes (H), haplotype diversity (H_d) and nucleotide diversity (π) of *O. japonicus* were computed by DnaSP v5.10 [57]. The program PermutCpSSR_1.2.1 [58] was used to calculate the average within-population gene diversity (h_S), total gene diversity (h_T) and the level of population differentiation at the species level (G_{ST}), as well as the geographical total haplotype diversity (V_T), geographical average haplotype diversity (V_S) and an estimate of population subdivisions for phylogenetically ordered alleles (N_{ST}) obtained by taking into account the similarities between haplotypes. G_{ST} and N_{ST} are used to assess the geographical structure which affects population differentiation. To evaluate the existence of correlation between the geographic and genetic distances, Mantel test was performed using GenAlEx v.6.502 [59]. The geographic distance matrix was constructed in Geographic Distance Matrix Generator v.1.2.3 [60], while the genetic distance matrix was constructed in Mega v.6.06 [61].

Demographic History

Mismatch distributions analysis was performed using the DnaSP v5.10 to develop figures and Arlequin 3.5 to calculate the values of the sum of squared deviations (SSD). The lack of significant differences in the SSD values indicated that the population assumed an expansion model.

The neutrality tests (Tajima's D and Fu's F_s) were carried out on Arlequin 3.5 [62]. Negative Tajima's D values suggest that a population has recently increased in size, whose results can tell the bottlenecks, selective effects, population expansion, or heterogeneity of mutation rates of *O. japonicus*.

Biogeographic Analyses

The statistical dispersal-vicariance analysis (S-DIVA) was implemented to infer ancestral distribution using RASP 3.2 [63]. The trees previously constructed by MrBayes were used as the input data. The geographical distribution of the species was categorized into five areas: (A) Sichuan province, China; (B) Zhejiang province, China; (C) Hubei province, China; (D) Hunan province, China; (E) Guizhou province, China.

Ecological Niche Modeling (ENM)

19 bioclimatic variables at a resolution of 2.5 arc minutes were obtained from the WorldClim website (available at <http://www.worldclim.org/download>), and community climate system model version 4 (CCSM4) model simulation was selected in Last inter-glacial (LIG), Last Glacial Maximum (LGM) and Mid Holocene (MID) periods. The current distribution information for *O. japonicus* was obtained from major databases and information systems, including the Global Biodiversity Information Facility (GBIF; <http://www.gbif.org/>) and the Chinese Virtual Herbarium (<http://www.cvh.ac.cn>). After a pairwise correlation analysis using SPSS, we selected a smaller set of nine relatively uncorrelated variables ($r \leq |0.8|$): mean diurnal range (bio2), isothermality (bio3), temperature Seasonality (bio4), mean temperature of the wettest quarter (bio8), mean temperature of warmest quarter (bio10), mean temperature of coldest quarter (bio11), precipitation seasonality (bio15), precipitation of wettest quarter (bio16), and precipitation of driest quarter (bio17). Replicate runs (10) of MaxEnt using the "bootstrap" method were performed to ensure reliable results. Model performance was assessed using the area under the curve (AUC) of the receiver operating characteristic plot, with 25% of the localities randomly selected to test the model. Graphics for each predicted species distribution model were drawn using S-DIVA.

List Of Abbreviations

cpDNA
chloroplast DNA
 H
the number of haplotypes
 H_d
haplotype diversity
 π
nucleotide diversity
 h_S
the average within-population gene diversity
 h_T
total gene diversity
 G_{ST}
the level of population differentiation at the species level
 V_T
the geographical total haplotype diversity
 V_S
geographical average haplotype diversity
 N_{ST}
an estimate of population subdivisions for phylogenetically ordered alleles
SSD

the sum of squared derivations
S-DIVA
the statistical dispersal-vicariance analysis
CCSM4
community climate system model version 4
LIG
last inter-glacial
LGM
last glacial maximum
MID
mid Holocene
GBIF
global biodiversity information facility
AUC
the area under the curve
 N_m
gene flow
AMOVA
analyses of molecular variance
RASP
reconstruct ancestral state in phylogenies
GRH
glacial refugium hypothesis
 F_{ST}
Genetic differentiation among populations
 F_{SC}
genetic differentiation among population within groups
 F_{CT}
Genetic differentiation between groups.

Declarations

Ethics approval and consent to participate

QLP undertook the formal identification of the plant material used in our study. The authors have promised that relevant permission to collect the plant used were obtained from all the owners of the private land during the collection of the plants. All local, national or international guidelines and legislation were adhered to in the production of this study.

Consent for publication

Not applicable.

Availability of data and materials

The datasets generated and analysed during the current study are available in GenBank (accession numbers: OP785779-OP786045; OP786046-OP786312; OP786313-OP786579; OP786580 - OP786846).

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by Research Project of Zhejiang Chinese Medical University [grant number: 2021JKJNTZ007B]

Authors' contributions

ZLY, ZQY and SYQ designed the research; ZLY and LS collected the samples; ZLY, SY and RQR performed the experiment, analysis and wrote the manuscript; all authors revised the manuscript. The author(s) read and approved the final manuscript.

Acknowledgements

I am grateful to related local institutions and staff in Sichuan and Zhejiang for helping to collect experimental materials. I am thankful to Kaohua Liu and Xiaoqing Jin for their assistance in the collection of samples.

References

1. Dang H, Wang Q, Wang H, Yan M, and Liu X. The Integration of Chinese Material Medica into the Chinese Health Care Delivery System, an Update. *Phytother Res.* 2016; 30: 292-7.
2. Carvalho ACB, Lana TN, Perfeito JPS, and Silveira D. The Brazilian Market of Herbal Medicinal Products and the Impacts of the New Legislation on Traditional Medicines. *J Ethnopharmacol.* 2018; 212: 29-35.
3. Abdullahi AA. Trends and Challenges of Traditional Medicine in Africa. *Afr J Tradit Complement Altern Med.* 2011; 8: 115-23.
4. WHO, IUCN, and WWF. Guidelines on the Conservation of Medicinal Plants. Switzerland: International Union for Conservation of Nature and Natural Resources; 1993.
5. Doebley JF, Gaut BS, and Smith BD. The Molecular Genetics of Crop Domestication. *Cell.* 2006; 127: 1309-21.
6. Olsen K and Gross B. Detecting Multiple Origins of Domesticated Crops. *Proc Natl Acad Sci U S A.* 2008; 105: 13701-2.
7. Chen C, Zheng L, Ma Q, Zhou WB, Lu Y, Zhao YP, and Fu CX. Impacts of Domestication on Population Genetics of a Traditional Chinese Medicinal Herb, *Atractylodes Macrocephala* (Asteraceae). *J Syst Evol.* 2019; 57: 222-33.
8. Huang R, Liu Y, Chen J, Lu Z, Wang J, He W, Chao Z, and Tian E. Limited Genetic Diversity and High Differentiation in *Angelica Dahurica* Resulted from Domestication: Insights to Breeding and Conservation. *BMC Plant Biol.* 2022; 22: 141.
9. Pauls SU, Nowak C, Bálint M, and Pfenninger M. The Impact of Global Climate Change on Genetic Diversity within Populations and Species. *Mol Ecol.* 2013; 22: 925-46.
10. Rešetnik I, Baričević D, Batir Rusu D, Carović-Stanko K, Chatzopoulou P, Dajić-Stevanović Z, Goncariuc M, Grdiša M, Greguraš D, Ibraliu A, Jug-Dujaković M, Krasniqi E, Liber Z, Murtić S, Pečanac D, Radosavljević I, Stefkov G, Stešević D, Šošarić I, and Šatović Z. Genetic Diversity and Demographic History of Wild and Cultivated/Naturalised Plant Populations: Evidence from Dalmatian Sage (*Salvia Officinalis* L., Lamiaceae). *PLoS ONE.* 2016; 11: e0159545.
11. Sun YQ, Yang H, Zhang QY, Qin LP, Li P, Lee JK, Chen SC, Rahman K, Kang TG, and Jia M. Genetic Diversity and Its Conservation Implications of *Vitex Rotundifolia* (Lamiaceae) Populations in East Asia. *PeerJ.* 2019; 7: e6194.
12. Tilwari A and Sharma R. Random Amplified Polymorphic DNA and Inter Simple Sequence Repeat Markers Reveals Genetic Diversity between Micro Propagated, Wild and Field Cultivated Genotypes of *Gloriosa Superba*: An Endangered Medicinal Plant. *Mol Biol Rep.* 2021; 48: 2437-52.
13. Schaal BA, Hayworth DA, Oseni KM, Rauscher JT, and Smith WA. Phylogeographic Studies in Plants: Problems and Prospects. *Mol. Ecol.* 1998; 7: 465–74.
14. Avise JC. *Phylogeography: The History and Formation of Species.* Cambridge: Harvard University Press; 2000.
15. Rogers AR and Harpending H. Population Growth Makes Waves in the Distribution of Pairwise Genetic Differences. *Mol Biol Evol.* 1992; 9: 552-69.
16. Aguirre-Dugua X and González-Rodríguez A. Phylogeographical Approaches to the Study of Plant Domestication, with Special Emphasis on Perennial Plants, In: Lira R., Editor. *Ethnobotany of Mexico.* New York: Springer; 2016. p. 319-66.
17. Wang X, Liu XQ, Ko YZ, Jin XL, Sun JH, Zhao ZY, Yuan QJ, Chiang YC, and Huang LQ. Genetic Diversity and Phylogeography of the Important Medical Herb, Cultivated Huang-Lian Populations, and the Wild Relatives *Coptis* Species in China. *Front Genet.* 2020; 11: 708.
18. Jafari M, Akram W, Pang YJ, Ahmad A, Ahmed S, Yasin NA, Anjum T, Ali B, Hu XD, Li XH, Dong SA, Cai Q, Ciprian M, Bielec M, Hu S, Sefidkon F, and Hu XB. Genetic Diversity and Biogeography of *T. Officinale* Inferred from Multi Locus Sequence Typing Approach. *PLoS ONE.* 2018; 13: e0203275.
19. Author. *Pharmacopoeia of the People's Republic of China.* Beijing, China: Press of Chinese Medical Science and Technology. 2020.
20. Zhao JC, Wang YH, Weng QQ, Jin Y, and Zhang W. Herbal Textual Research on "Ophiopogonis Radix" in Chinese Classical Prescriptions. *Mod Chin Med.* 2020; 22: 1381-92.
21. Liu X, Lin YH, Xie CX, Zeng FL, MA XC, Sun W, Tu PF, and Chen SL. Ecogenetic and Differentiation of Genuine Medicinal Materials: *Ophiopogonis Radix* in Sichuan and Zhejiang Province. *Chinese Journal of Experimental Traditional Medical Formulae.* 2017; 23: 27-33.
22. Huang YJ, Chen JY, Hai-Lan SU, Huang YZ, and Wan XF. Rdnal-Its of *Radix Ophiopogonis* and *Radix Liriodopsis* from Different Regions. *Fujian Journal of Agricultural Sciences.* 2009; 24: 508-12.
23. Li X, Wang Y, Sun P, and Liao D. Determination of Saponin Content in Hang Maidong and Chuan Maidong Via Hplc-Elcd Analysis. *Journal of analytical methods in chemistry.* 2016; 2016: 7214607.
24. Ge YH, Chen XJ, Godevac D, Bueno PCP, Abarca LFS, Jang YP, Wang M, and Choi YH. Metabolic Profiling of Saponin-Rich *Ophiopogon Japonicus* Roots Based on (1) H Nmr and Hptlc Platforms 1. *Planta Med.* 2019; 85: 917-24.
25. Tan MX, Chen JL, Wang CC, Zou LS, Chen SY, Shi JJ, Mei YQ, Wei LF, and Liu XH. Quality Evaluation of *Ophiopogonis Radix* from Two Different Producing Areas. *Molecules.* 2019; 24: 3220.
26. Lin Y, Zhu D, Qi J, Qin M, and Yu B. Characterization of Homoisoflavonoids in Different Cultivation Regions of *Ophiopogon Japonicus* and Related Antioxidant Activity. *J Pharm Biomed Anal.* 2010; 52: 757-62.
27. Lu XY, Tong W, Wang SF, Li JH, Zheng J, Fan XH, and Liu L. Comparison of the Chemical Constituents and Immunomodulatory Activity of *Ophiopogonis Radix* from Two Different Producing Areas. *J Pharm Biomed Anal.* 2017; 134: 60-70.
28. Zhao M, Xu WF, Shen HY, Shen PQ, Zhang J, Wang DD, Xu H, Wang H, Yan TT, and Wang L. Comparison of Bioactive Components and Pharmacological Activities of *Ophiopogon Japonicus* Extracts from Different Geographical Origins. *J Pharm Biomed Anal.* 2017; 138: 134-41.
29. Cen XX, Sun J, Shen XX, and An. WZ. Study on Origin and Genetic Diversity of Cultivated *Ophiopogon Japonicus* Based on Srap and Issr Markers. *Journal of Chinese Medicinal Materials.* 2021; 44: 555-61.

30. Liu YY, Lu JJ, Wang HZ, Wang ZA, Chen JG, and Min H. Evaluation of Genetic Diversity of Germplasm from *Ophiopogon Japonicus* Planted in Zhejiang. *Journal of Zhejiang Agricultural Sciences*. 2017; 58: 2146-9.
31. Fawcett T. An Introduction to Roc Analysis. *Pattern Recognition Letters*. 2005; 27: 861-74.
32. Holder K, Montgomerie R, and Friesen V. A Test of the Glacial Refugium Hypothesis Using Patterns of Mitochondrial and Nuclear DNA Sequence Variation in Rock Ptarmigan (*Lagopus Mutus*). *Evolution; international journal of organic evolution*. 1999; 53: 1936-50.
33. Hyten D, Song QJ, Zhu YL, Choi I, Nelson R, Costa J, Specht J, Shoemaker R, and Cregan P. Impacts of Genetic Bottlenecks on Soybean Genome Diversity. *Proc Natl Acad Sci U S A*. 2006; 103: 16666–71.
34. Yamasaki M, Wright S, and McMullen M. Genomic Screening for Artificial Selection During Domestication and Improvement in Maize. *Annals of botany*. 2007; 100: 967-73.
35. Kovach M and McCouch S. Leveraging Natural Diversity: Back through the Bottleneck. *Curr Opin Plant Biol*. 2008; 11: 193-200.
36. Chen C, Li P, Wang RH, Schaal BA, and Fu CX. The Population Genetics of Cultivation: Domestication of a Traditional Chinese Medicine, *Scrophularia Ningpoensis* Hemsl. (*Scrophulariaceae*). *PLoS ONE*. 2014; 9: e105064.
37. Zhang YH, Yu QS, Zhang Q, Hu XK, Hu J, and Fan BL. Regional-Scale Differentiation and Phylogeography of a Desert Plant *Allium Mongolicum* (*Liliaceae*) Inferred from Chloroplast DNA Sequence Variation. *Plant Syst Evol*. 2017; 303: 451-66.
38. Sun JH, Wang YH, Garran TA, Qiao P, Wang ML, Yuan QJ, Guo LP, and Huang LQ. Heterogeneous Genetic Diversity Estimation of a Promising Domestication Medicinal Motherwort *Leonurus Cardiaca* Based on Chloroplast Genome Resources. *Front Genet*. 2021; 12: 721022.
39. Yu Y, Wang L, Sun JH, Zhong WH, Zhang T, Guo LP, and Yuan QJ. Genetic Diversity and Genetic Structure Evaluation of Cultivated Chinese Wolfberry (*Lycium Barbarum*) Based on Microsatellite Population Genetics. *Zhongguo Zhong Yao Za Zhi*. 2020; 45: 838-45.
40. Miller A and Gross B. From Forest to Field: Perennial Fruit Crop Domestication. *Am J Bot*. 2011; 98: 1389-414.
41. Miller A and Schaal B. Domestication and the Distribution of Genetic Variation in Wild and Cultivated Populations of the Mesoamerican Fruit Tree *Spondias Purpurea* L. (*Anacardiaceae*). *Mol Ecol*. 2006; 15: 1467-80.
42. Wang H, Li ZF, Chen J, Zeng FL, and Xie CX. Suitability Evaluation and Ecological Characteristics of Traditional Producing Areas of Zhejiang *Ophiopogon Japonicas*. *Journal of Henan Agricultural Sciences*. 2016; 45: 89-95.
43. Slatkin M. Gene Flow and the Geographic Structure of Natural Populations, In: Real LA, Editor. *Ecological Genetics*. Princeton: Princeton University Press; 1987. p. 3-17.
44. Liu J, Liu S, Chen XF, Zhang SP, and Zou YF. Comprehensive Comparison between Erective and Creepy *Ophiopogon Japonicus* of Sichuan. *China Journal of Chinese Materia Medica*. 2009; 34: 2449-53.
45. Wright S. Isolation by Distance. *Genetics*. 1943; 28: 114-38.
46. Templeton A, Routman E, and Phillips C. Separating Population Structure from Population History: A Cladistic Analysis of the Geographical Distribution of Mitochondrial DNA Haplotypes in the Tiger Salamander, *Ambystoma Tigrinum*. *Genetics*. 1995; 140: 767-82.
47. Frankel OH. *The Conservation of Plant Biodiversity*. Cambridge: Cambridge University Press; 1995.
48. Kassa M, Penmetza R, Carrasquilla-Garcia N, Sarma B, Datta S, Upadhyaya H, Varshney R, von Wettberg E, and Cook D. Genetic Patterns of Domestication in Pigeonpea (*Cajanus Cajan* (L.) Millsp.) and Wild *Cajanus* Relatives. *PLoS ONE*. 2012; 7: e39563.
49. Song MF, Guan YH, Zhang Y, Li HT, Zhang LX, and Zhang ZL. Genetic Diversity Assessment of a Chinese Medicinal Endemic Species, *Aspidopterys Obcordata* Var. *Obcordata*, by Combined Molecular Marker Methods (Issr & Srap). *Biochem Genet*. 2021; 59: 283-99.
50. Tzedakis P. C. Buffered Tree Population Changes in a Quaternary Refugium: Evolutionary Implications. *Science*. 2003; 299: 825.
51. Sommer RS and Zachos FE. Fossil Evidence and Phylogeography of Temperate Species: 'Glacial Refugia' and Post-Glacial Recolonization. *J Biogeogr*. 2010; 36: 2013–20.
52. Chung MY, Vu SH, López-Pujol J, Herrando-Moraira S, Son S, Suh GU, Le HTQ, and Chung MG. Comparison of Genetic Variation between Northern and Southern Populations of *Lilium Cernuum* (*Liliaceae*): Implications for Pleistocene Refugia. *PLoS ONE*. 2018; 13: e0190520.
53. AJ D, B A, S B, M C, A C, C D, M F, J H, M K, S M, R M, S S-H, S S, and TWA T. Geneious V4.8.5. 2017. Available from <http://www.geneious.com>. Accessed 2022-6-6.
54. Ronquist F TM, Mark P Van Der, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP. Mrbayes 3.2: Efficient Bayesian Phylogenetic Inference and Model Choice across a Large Model Space. *Systematic Biology* 61:539–542 DOI 10.1093/sysbio/sys029. 2012.
55. Posada D CK. Modeltest: Testing the Model of DNA Substitution. *Bioinformatics* 14:817–818 DOI 10.1093/bioinformatics/14.9.817. 1998.
56. Clement M PD, Crandall KA. Tcs: A Computer Program to Estimate Gene Genealogies. *Molecular Ecology* 9:1657–1659 DOI 10.1046/j.1365-294x.2000.01020.x. 2000.
57. Librado P RJ. Dnasp V5: A Software for Comprehensive Analysis of DNA Polymorphism Data. *Bioinformatics* 25:1451–1452 DOI 10.1093/bioinformatics/btp187. 2009.
58. Pons O PR. Measwring and Testing Genetic Differentiation with Ordered Versus Unordered Alleles. *Genet Breed*. 1996; 144: 1237–45.
59. Peakall R and Smouse P. Genalex 6.5: Genetic Analysis in Excel. Population Genetic Software for Teaching and Research—an Update. *Bioinformatics* (Oxford, England). 2012; 28: 2537-9.
60. Ersts P. Geographic Distance Matrix Generator (Version 1.2.3). 2012. Available from http://biodiversityinformatics.amnh.org/open_source/gdmg. Accessed 2022-6-19.
61. Tamura K, Stecher G, Peterson D, Filipski A, and Kumar S. Mega6: Molecular Evolutionary Genetics Analysis Version 6.0. *Mol Biol Evol*. 2013; 30: 2725-9.

62. L E and HE L. Arlequin Suite Ver 3.5: A New Series of Programs to Perform Population Genetics Analyses under Linux and Windows. Molecular Ecology Resources. 2010; 10: 564–7

63. Yu Y, Harris AJ, Blair C, and He X. Rasp (Reconstruct Ancestral State in Phylogenies): A Tool for Historical Biogeography. Mol Phylogenet Evol. 2015; 87: 46-9.

Figures

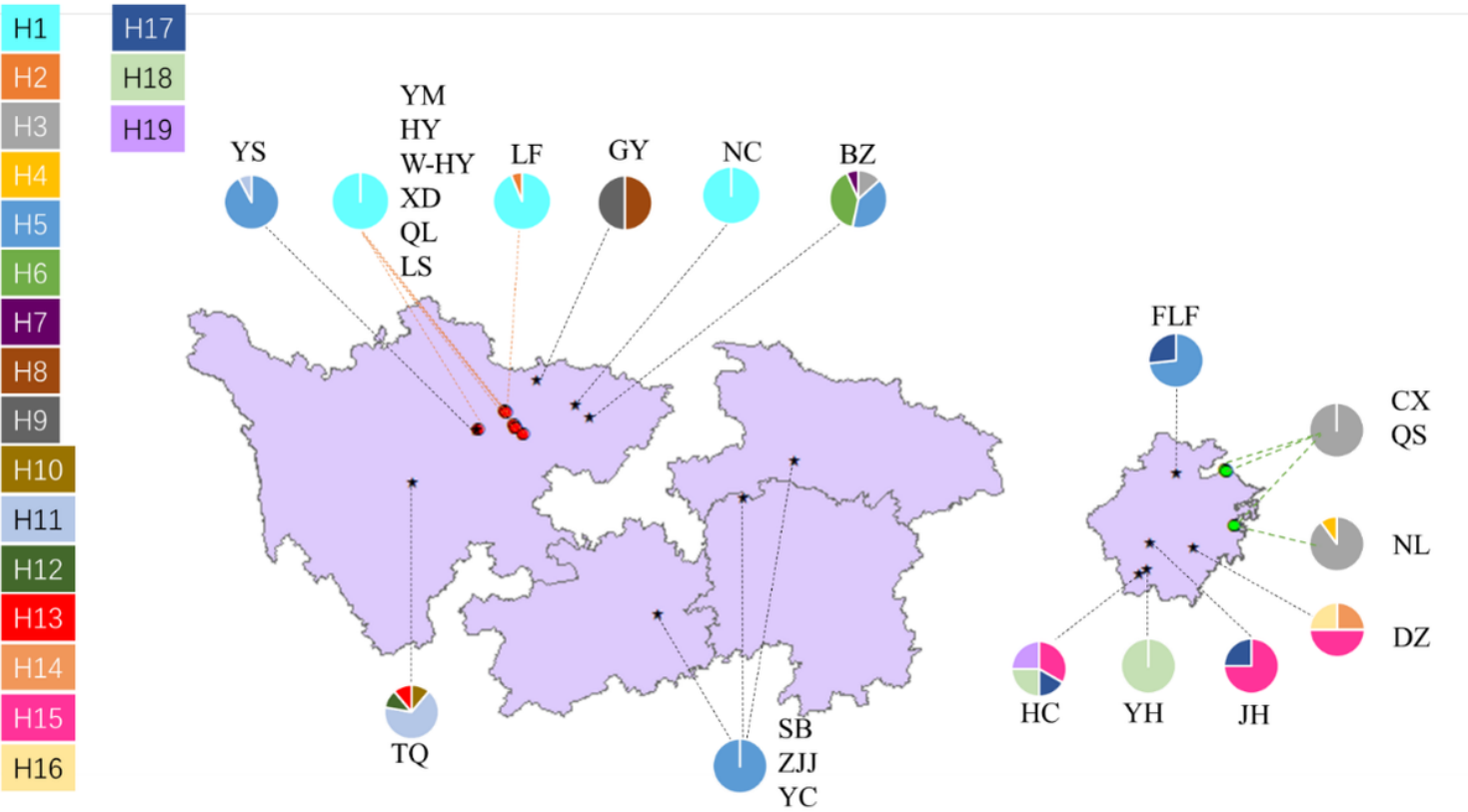


Figure 1

A geographic distribution of 19 cpDNA haplotypes (H1-H19) detected in *O. japonicus*.

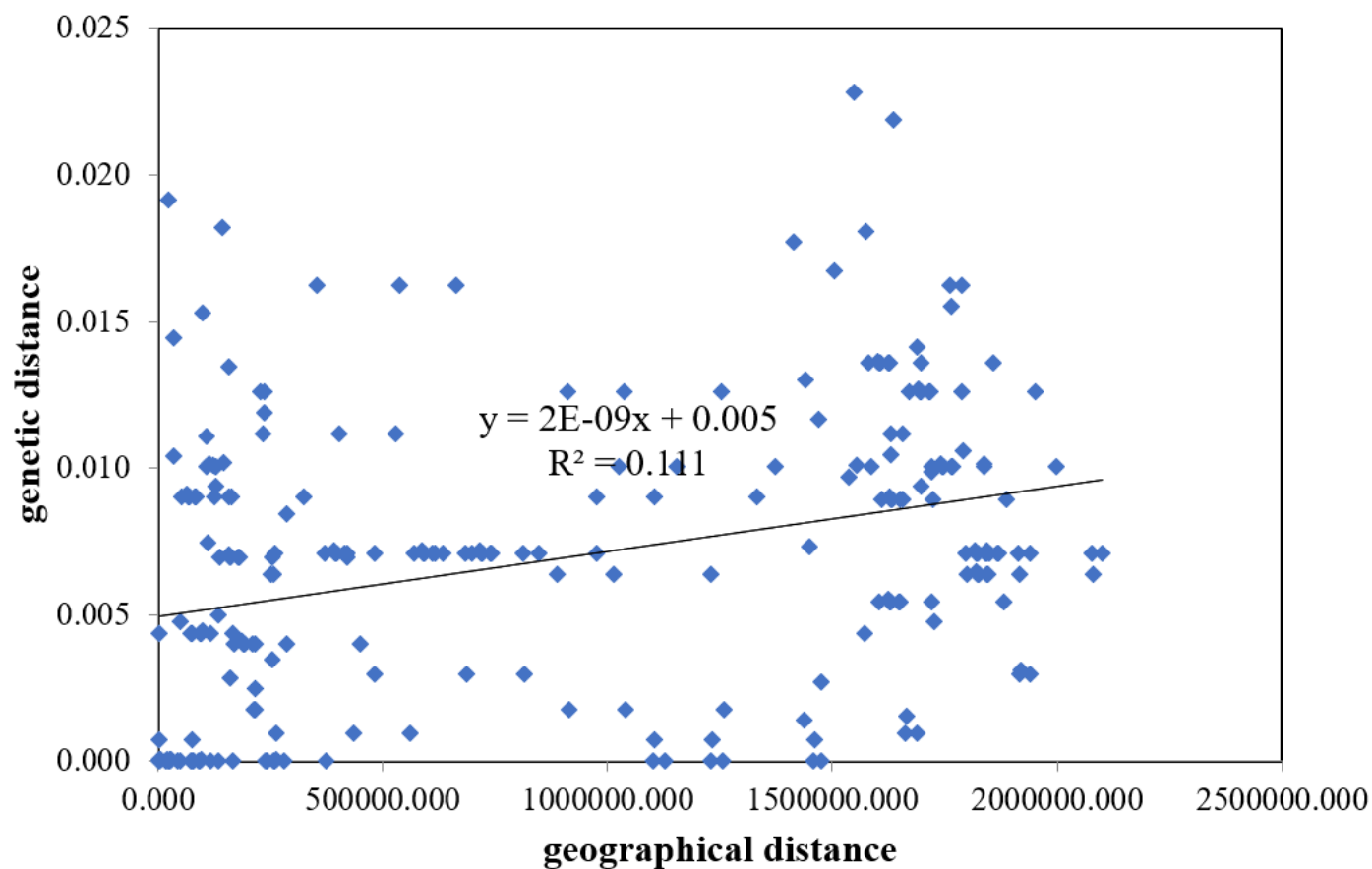


Figure 2

Mantel test of relationship between genetic distance and geographical distance among populations of *O. japonicus*

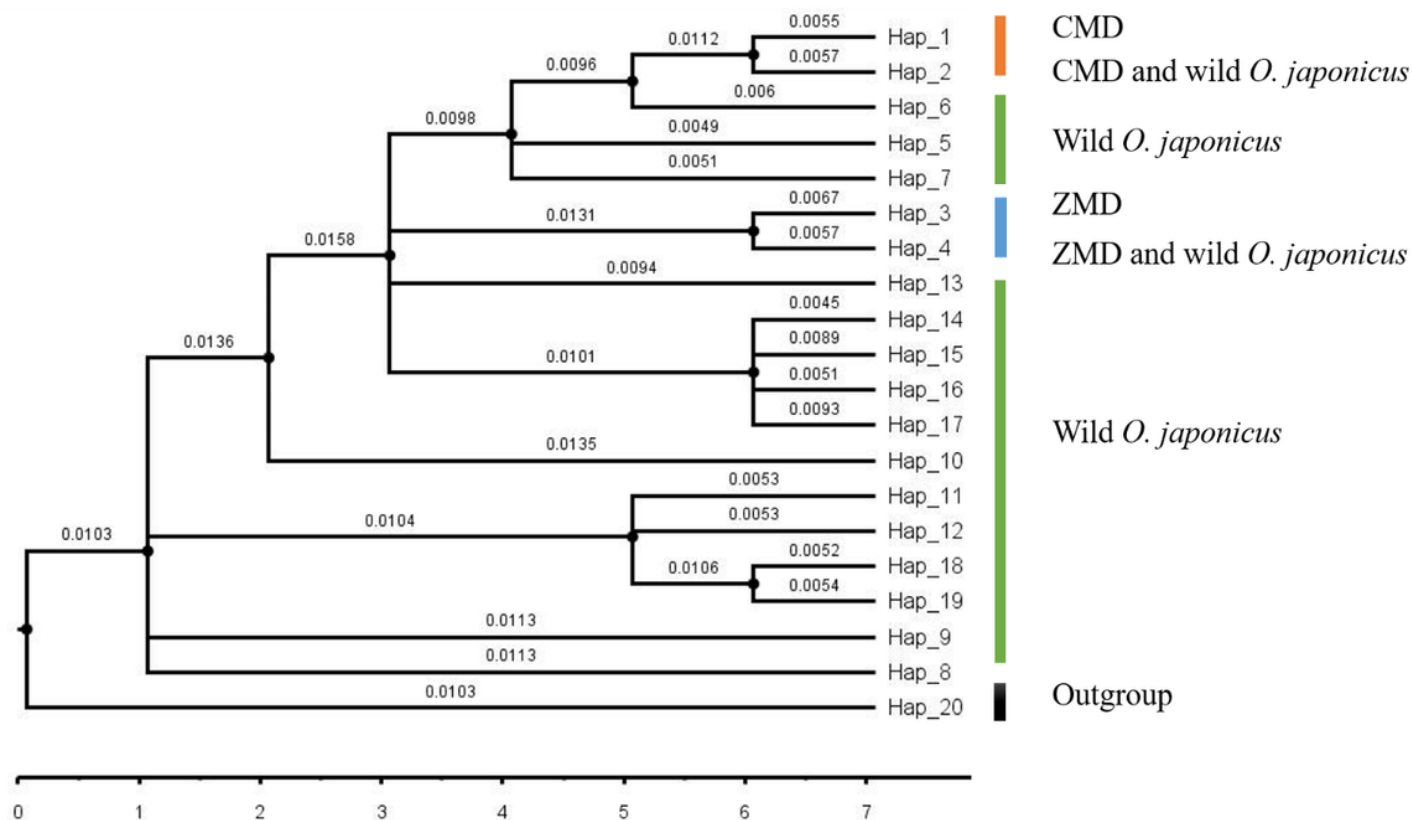


Figure 3

Maximum-likelihood (ML) phylogenetic tree: the number showed on node stand for bootstrap value on 1000 replicates; *O. amblyphyllus* Wang et Dai is set as an outgroup.

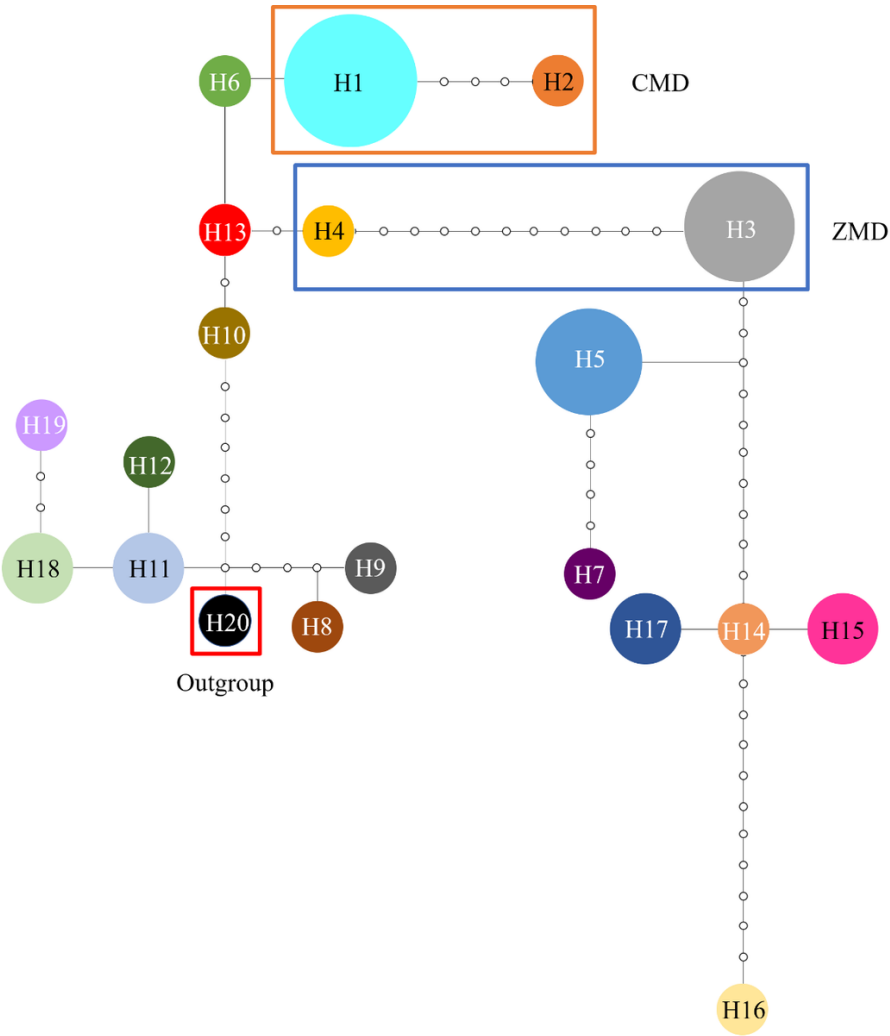


Figure 4

Minimum spanning network of 19 cpDNA haplotypes in *O. japonicus*. The network was rooted at the *O. amblyphyllus* Wang et Dai. Circle sizes are proportional to the number of samples per haplotypes. Hollow dots indicate the number of mutational steps.

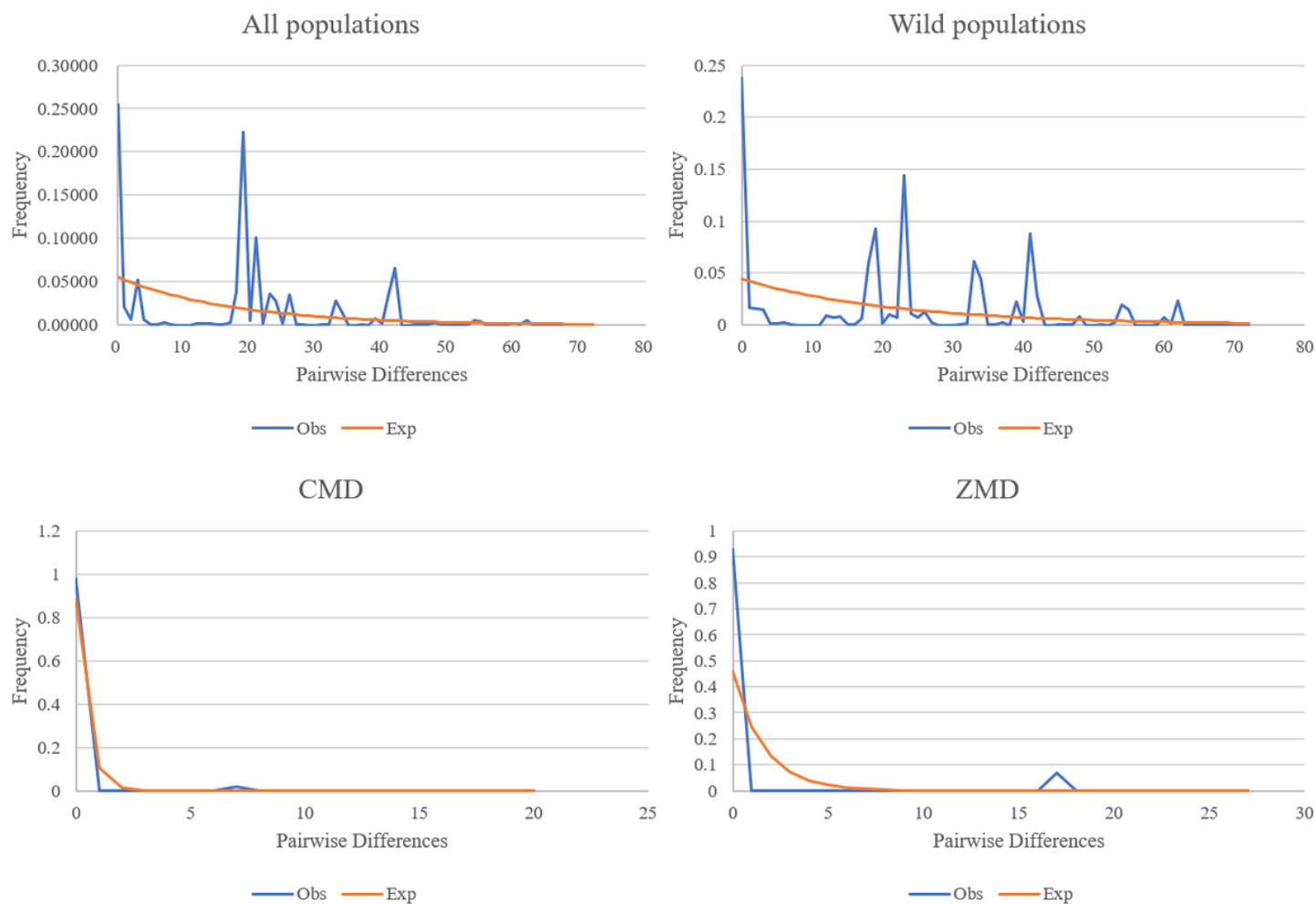


Figure 5

Mismatch distribution for cpDNA sequence data in *O. japonicus*. The red line represents expected values and blue line represents observed values.

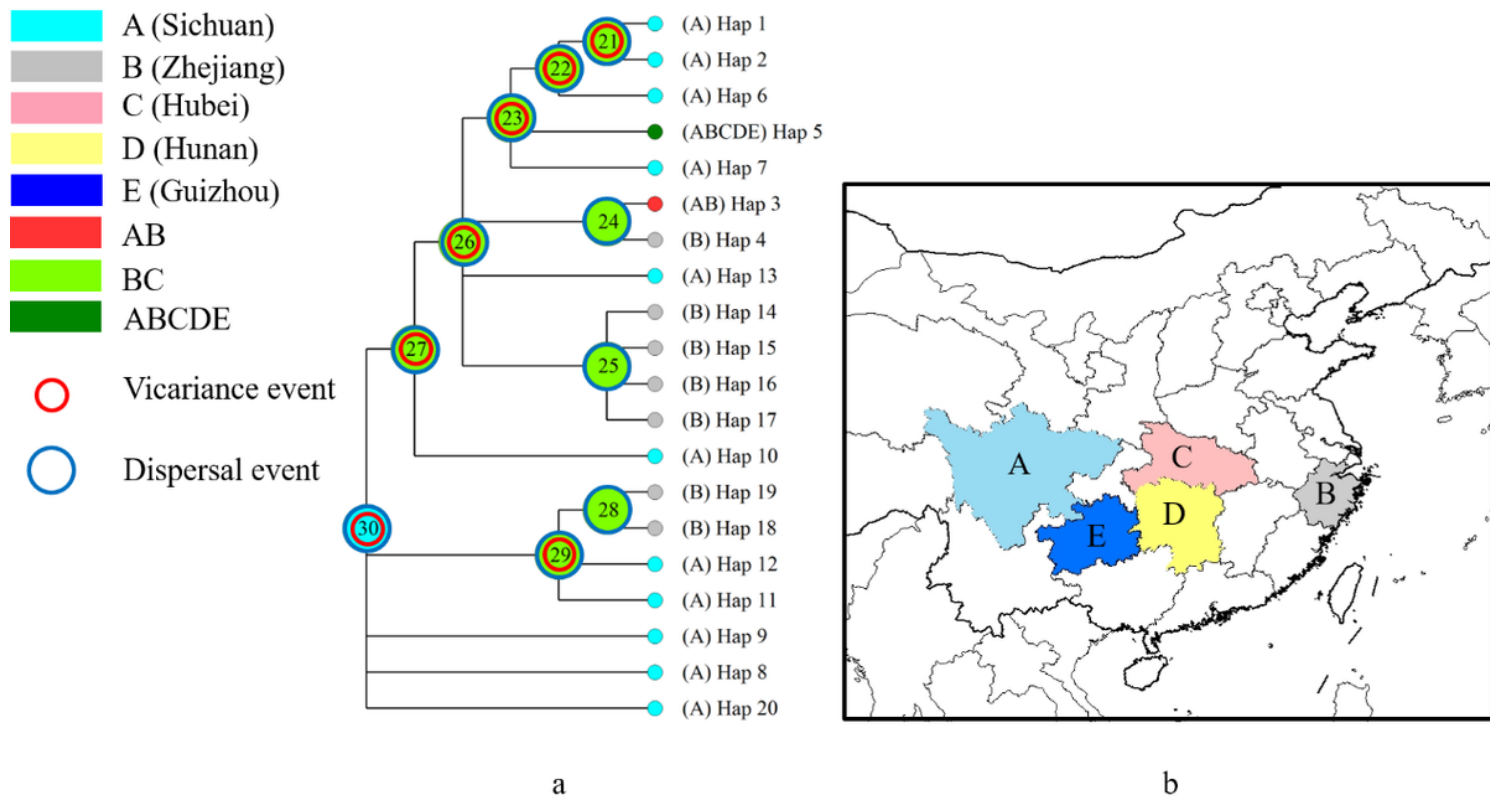


Figure 6

Historical biogeography of *O. japonicus* and ancestral area reconstruction. Map showing five geographical regions in colors as defined in RASP analysis. Pie charts on nodes display the relative probabilities of possible ancestral ranges.

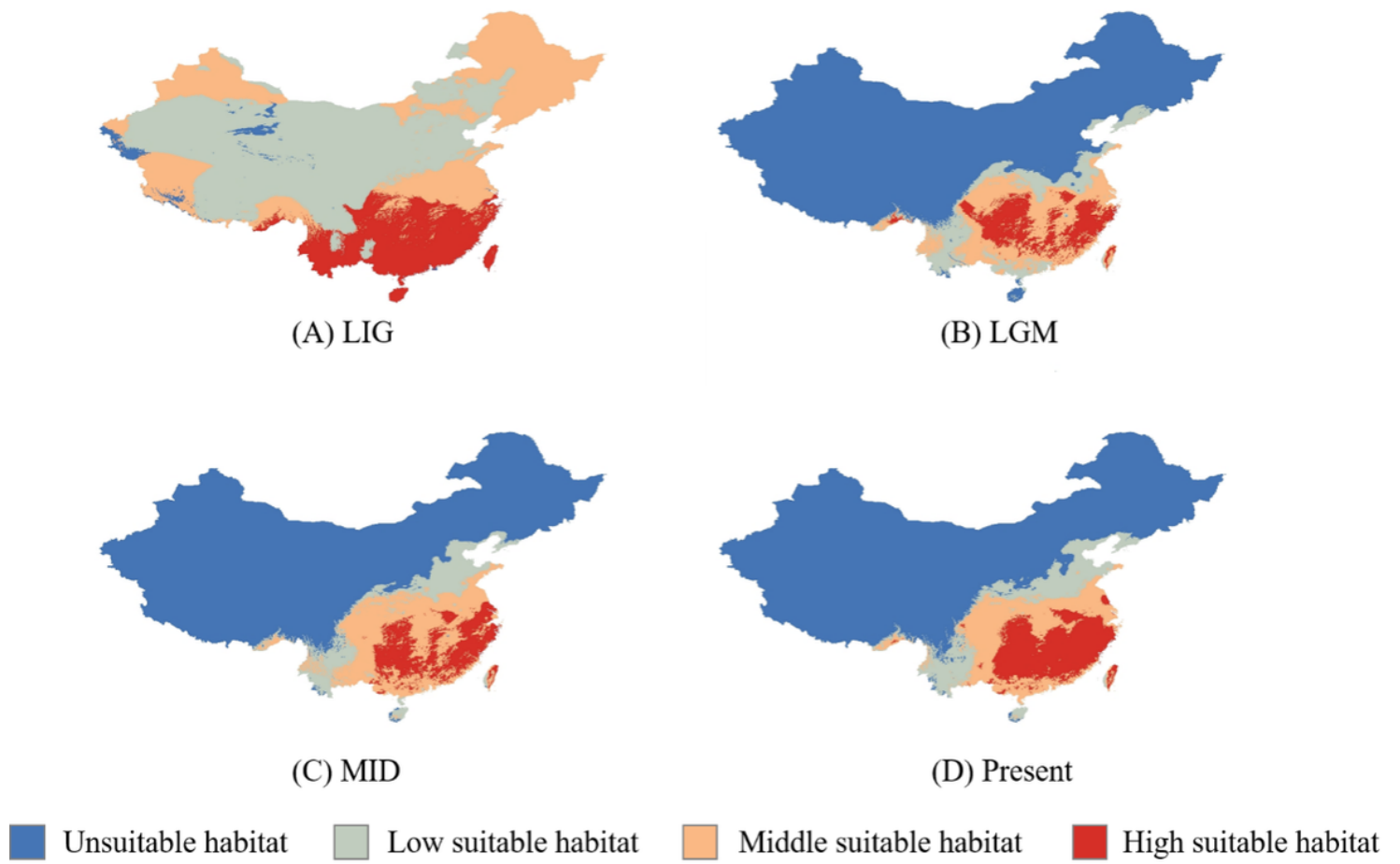


Figure 7

Simulation of the distribution area of *O. japonicus* in different climatic conditions in China.

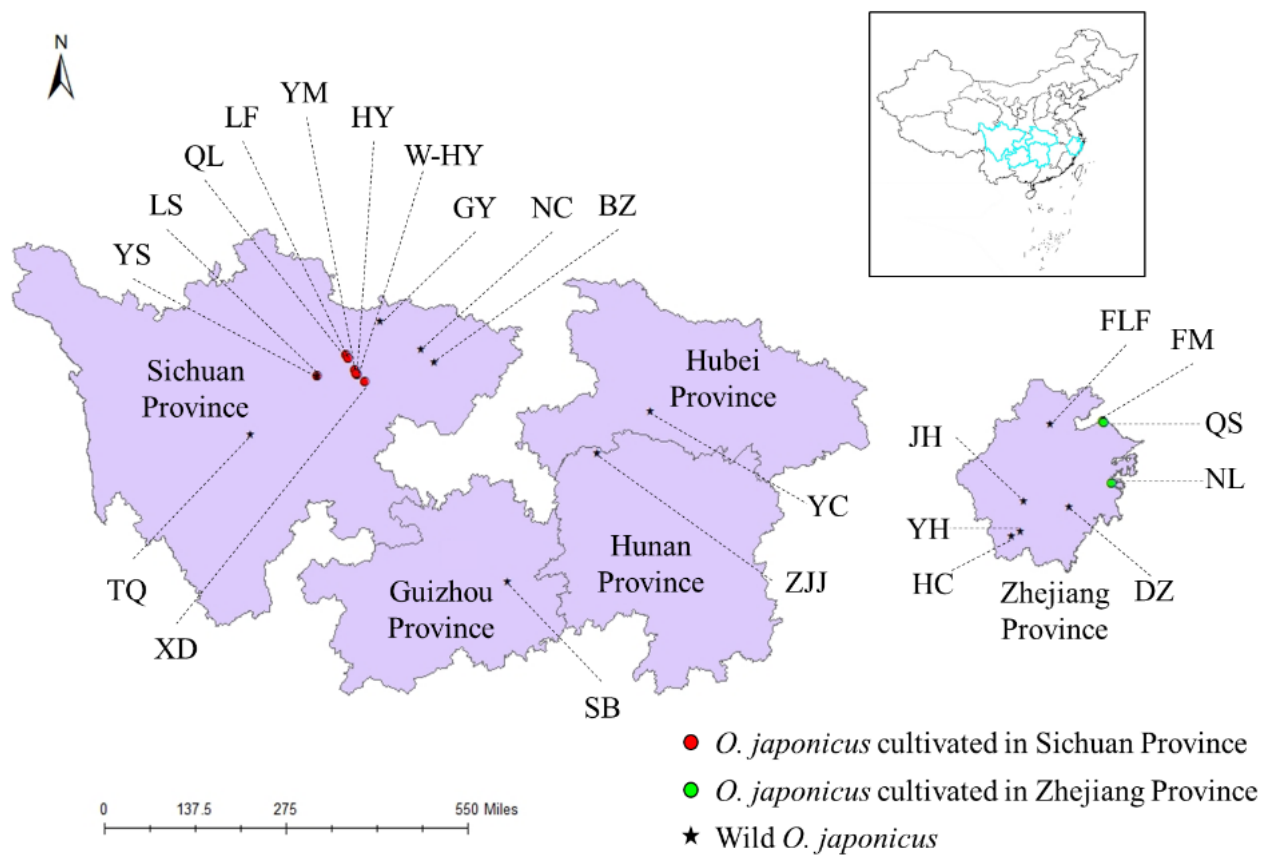


Figure 8

The sample locations of *O. japonicus* used in this study.