

RESEARCH

Open Access



# Genetic diversities in wild and cultivated populations of the two closely-related medical plants species, *Tripterygium Wilfordii* and *T. Hypoglaucom* (Celastraceae)

Chao Liu<sup>1†</sup>, Jingyi Wang<sup>1†</sup>, Ya-Zhu Ko<sup>2†</sup>, Meng-Shin Shiao<sup>3†</sup>, Yiheng Wang<sup>1</sup>, Jiahui Sun<sup>1</sup>, Qingjun Yuan<sup>1\*</sup>, Lisong Wang<sup>4\*</sup>, Yu-Chung Chiang<sup>2,5,6\*</sup> and Lanping Guo<sup>1\*</sup>

## Abstract

**Background** The sustainable supply of medicinal plants is important, and cultivating and domesticating them has been suggested as an optimal strategy. However, this can lead to a loss of genetic diversity. *Tripterygium wilfordii* Hook. f. is a medicinal plant commonly used in traditional Chinese medicine, but its wild populations are dwindling due to excessive harvesting. To protect the species and meet the increasing demand, it is urgent to cultivate it on a large scale. However, distinguishing between *T. wilfordii* and *T. hypoglaucom*, two similar species with different medicinal properties, is challenging. Therefore, it is crucial to understand the genetic diversity and population structure of these species for their sustainable utilization.

**Results** In this study, we investigated the genetic diversity and population structure of the two traditional medicinal semiwoody vines plant species, *Tripterygium wilfordii* and *T. hypoglaucom*, including wild and cultivated populations using chloroplast DNA (cpDNA) sequences and microsatellite loci. Our results indicated that the two species maintain a high level of genetic divergence, indicating possible genetic bases for the different contents of bioactive compounds of the two species. *T. wilfordii* showed lower genetic diversity and less subdivided population structures of both markers than *T. hypoglaucom*. The potential factors in shaping these interesting differences might be differentiated pollen-to-seed migration rates, interbreeding, and history of population divergence. Analyses of cpDNA and microsatellite loci supported that the two species are genetically distinct entities. In addition, a significant reduction of genetic diversity was observed for cultivated populations of the two species, which mainly resulted from the small initial population size and propagated vegetative practice during their cultivation.

<sup>†</sup>Chao Liu, Jingyi Wang, Ya-Zhu Ko and Meng-Shin Shiao contributed equally to this work.

\*Correspondence:

Qingjun Yuan

yuanqingjun@icmm.ac.cn

Lisong Wang

wanglisong@lsbg.cn

Yu-Chung Chiang

yuchung@mail.nsysu.edu.tw

Lanping Guo

glp01@126.com

Full list of author information is available at the end of the article



**Conclusion** Our findings indicate significant genetic divergence between *T. wilfordii* and *T. hypoglaucum*. The genetic diversity and population structure analyses provide important insights into the sustainable cultivation and utilization of these medicinal plants. Accurate identification and conservation efforts are necessary for both species to ensure the safety and effectiveness of crude drug use. Our study also highlighted the importance of combined analyses of different DNA markers in addressing population genetics of medicinal plants because of the contrasts of inheritance and rates of gene flow. Large-scale cultivation programs should consider preserving genetic diversity to enhance the long-term sustainability of *T. wilfordii* and *T. hypoglaucum*. Our study proposed that some populations showed higher genetic diversity and distinctness, which can be considered with priority for conservation and as the sources for future breeding and genetic improvement.

**Keywords** DNA sequence, Phylogeography, Population genetics, *Tripterygium*, Traditional chinese medicine

## Background

Medicinal plants play a central role in the health of humanity, but over-exploitation on wild resources led to biodiversity loss and caused global concerns [1–3]. Cultivation and domestication have been suggested as an optimal strategy for sustainable supply of medicinal plants [4–6]. However, the process of cultivation and domestication in most medicinal plants involves artificial selection and genetic bottleneck, which can lead to a serious loss of the genetic diversity found within the wild populations [7–9]. The loss of genetic diversity may have several severe impacts: it may lead to (1) decline of plasticity of secondary metabolism, which is the major source of medicinal compounds [10, 11], (2) progressive narrowing of the genetic diversity as the source for artificial selection to improve the yield of medicinal compounds [12–14], (3) lacking evolutionary potential to adapt to changing environments [15–17], and (4) negative impact on sustainable utilization. In particular, some promising medicinal plants for new drug development are being over-harvested, and their genetic diversity is rapidly declining. In order to protect the wild resources of medicinal plants and meet the increasing demand of human beings, there is an urgent need to introduce these medicinal plants into large-scale cultivation. Investigating genetic diversity patterns in wild and cultivated populations of these medicinal plants can provide guidelines for future cultivation programs.

*Tripterygium wilfordii* Hook. f. is a deciduous subshrub or scandent semiwoody vine of the family Celastraceae, the root and rhizome of which is a very important *ateria medica* in traditional Chinese medicine. During the last two decades, *T. wilfordii* has attracted much attention due to new discoveries about its pharmacological activities, such as immunosuppressive, anti-inflammatory, anticancer and anti-HIV activity [18–22]. Triptolide is the major bioactive component of *T. wilfordii* [23], which is a primary diterpene triepoxide with immunosuppressive and anti-inflammatory pharmacological effects [24]. Analogs of triptolide have been shown to be effective in

treating different diseases: Minnelide is a water-soluble triptolide analog, which has been shown to be highly effective in reducing pancreatic tumor and small cell lung carcinoma tumor growth and spread in animal models [25–27]; (5R)-5-Hydroxytriptolide (LLDT-8) is a new and optimized triptolide analog with lower cytotoxicity and relatively higher immunosuppressive activity [28]. These new pharmacological effects rapidly increased the demand for *T. wilfordii*. At present, the source of *T. wilfordii* medicinal materials mainly depends on the collection of wild resources, which results in the rapid decline in the sizes of the natural populations of *T. wilfordii* [22]. Thus, it is urgent to cultivate *T. wilfordii* artificially on a large scale.

In China, there are three species of *Tripterygium* Hook. f. in Flora Republicae Popularis Sinicae: *T. wilfordii* Hook. f., *T. hypoglaucum* (H. Lévy) Hutch, and *T. regelii* Sprague et Takeda [29]. The wild populations of *Tripterygium* species are restricted distribution: The wild *T. wilfordii* and *T. hypoglaucum* are native to eastern and southern China, and *T. regelii* is mainly distributed in northeast China. Wild *T. wilfordii* and *T. hypoglaucum* are plants growing naturally without human intervention. The wild *T. wilfordii* can be found in various habitats such as forests, hills, mountains, alpine regions, and other diverse terrain conditions [30]. Conversely, cultivated *T. wilfordii* and *T. hypoglaucum* are grown in controlled environments for selective planting and breeding. The cultivation of *T. wilfordii* mainly involves wild transplants, seedling cultivation, and cutting propagation, while cultivated *T. hypoglaucum* primarily relies on cutting propagation. *T. hypoglaucum* is also an important medicinal plant in traditional Chinese medicine, but its pharmacological effect differs from that of *T. wilfordii*. *Tripterygium wilfordii* and *T. hypoglaucum* have distinct chemical compositions. *Tripterygium hypoglaucum* consists of various unique secondary metabolites, and the content of its compounds differs from *T. wilfordii*. In particular, the sesquiterpene alkaloid content noticeably

varies between the two. Medicinally, *T. hypoglaucum* is believed to have fewer side effects and lower toxicity than *T. wilfordii* [18, 31]. However, there is no significant difference in morphological characteristics, geographical distribution and chemical composition between *T. wilfordii* and *T. hypoglaucum*, making identifying them difficult [18, 32, 33]. Therefore, some authors merged *T. wilfordii* and *T. hypoglaucum* into one species [22, 34]. *Tripterygium regelii* is an independent species due to significant differences from the other two species in morphological characteristics, geographical distribution and chemical composition [18, 22, 32–34]. In *Flora of China*, the three species of *Tripterygium* have been merged into one species [35]. However, *T. wilfordii* and *T. hypoglaucum* are two completely different types of materia medica in traditional Chinese medicine, the accurate identification of which is relevant to the safety of crude drug use and sustainable utilization of the two species.

Genetic markers with higher resolutions have been used independently to study further the divergence of *Tripterygium* species including random amplified polymorphic DNA (RAPD) [36], nuclear ribosomal internal transcribed spacer (ITS) and 5 S rDNA [37], chloroplast DNA (*psbA-trnH*, *rpl32-trnL*, and *trnL-trnF*), DNA barcodes (ITS2, *psbA-trnH*, *matK*, and *rbcL*) and simple-sequence repeats (also known as microsatellite markers, SSRs) [34, 38]. However, the relationship and genetic structure of *Tripterygium* species are inconclusive from these studies due to limited samples or insufficient genetic markers. Furthermore, the evolutionary rate, inheritance and dispersion mechanisms of different genetic markers are different [39, 40], which might be another reason for inconsistent results.

Chloroplast DNA (cpDNA) and microsatellite markers (SSRs) are commonly used genetic markers in studying population genetics. They have different mutation rates, inherited and dispersed mechanisms [39, 40]. Chloroplast DNA evolves relatively slowly and is maternal inheritance, reflecting the impacts of seed flow on genetic diversity. Microsatellite markers evolve relatively quickly and are biparental and codominant inheritance, reflecting the impacts of pollen flow on genetic diversity patterns [41, 42]. The cpDNA and SSRs markers have been successfully employed to detect genetic diversities in numerous plant species [43–49]. Nevertheless, few studies use both markers together to focus on the population genetics of medicinal plants. Here, we utilized two cpDNA regions and ten microsatellite loci to investigate the genetic diversity and genetic structure of *T. wilfordii* and *T. hypoglaucum* to (1) elucidate the phylogenetic relationships between *T. wilfordii* and *T. hypoglaucum*; (2) evaluate the level of genetic diversity and population

structure of the two species; (3) compare the amount of genetic variation in cultivated and wild populations of the two species and provide sustainable utilization strategy.

## Results

### Higher number of haplotypes was observed in *T. hypoglaucum* while higher genetic diversity was observed in *T. wilfordii* using cpDNA sequences

We collected samples from 14 populations, 6 wild and 1 cultivated for both *T. wilfordii* and *T. hypoglaucum* (Table 1). Sequences of cpDNA (*psbA-trnH* and *trnL-trnF*) were obtained from a total of 179 individuals and submitted to GenBank (GenBank accession numbers: PP128529-PP128707 for *psbA-trnH*; PP128708-PP128886 for *trnL-trnF*). A total of 8 haplotypes (Hap1-8) were identified using combined sequences of the two cpDNA regions, where a total of 18 polymorphic sites were identified. Among those, one haplotype (Hap4) was shared in two populations of *T. wilfordii* (XYY and ZLX) and in one population of *T. hypoglaucum* (XLH), which indicates possible gene flows between the two species either naturally or artificially. Other than the shared haplotypes, two haplotypes (Hap1 and 2) were identified specifically in *T. wilfordii*, and five haplotypes (Hap3, 5, 6, 7 and 8) were identified specifically in *T. hypoglaucum* (Table 1; Fig. 1).

Overall, polymorphic sites ( $P$ ), nucleotide diversity ( $\pi$ ), nucleotide polymorphism ( $\theta_w$ ), and haplotype diversity ( $H_d$ ) were 18, 0.00650, 0.00264, and 0.836, respectively, from all the 179 individuals (Table 2). *T. wilfordii* exhibited more polymorphic sites, higher nucleotide diversity and nucleotide polymorphism ( $P=12$ ,  $\pi=0.00424$ ,  $\theta_w=0.00204$ ), than those of *T. hypoglaucum* ( $P=11$ ,  $\pi=0.00374$ ,  $\theta_w=0.00181$ ). However, *T. hypoglaucum* showed higher number of haplotypes and higher haplotype diversity ( $N_h=6$ ,  $H_d=0.828$ ) than those of *T. wilfordii* ( $N_h=3$ ,  $H_d=0.566$ ). In addition, the neutrality test (Tajima's  $D$  and Fu and Li's  $F^*$ ) for linkage disequilibrium showed significance ( $P<0.02$  or  $P<0.001$ ) in both species. The results of Tajima's  $D$  and Fu and Li's  $F^*$  tests show the significant positive value indicates that *T. wilfordii* and *T. hypoglaucum* may have experienced the effects of population bottlenecks or balancing selection. Consistent with this, the dynamic analysis of microsatellites also shows a decline in the current effective population size. This consistent result reveals that these species are facing the pressure of reduced genetic diversity.

### Geographic distribution and genealogy of haplotypes demonstrated that the populations of the two species have different evolutionary relations

The geographic distribution of haplotypes of *T. wilfordii* was less structured and differentiated compared to those of *T. hypoglaucum* (Table 1; Fig. 1a, b). By visual

**Table 1** Sampling information and haplotypes of *Tripterygium wilfordii* (*Tw*) and *Tripterygium hypoglaucum* (*Th*) for chloroplast DNA (cpDNA) and microsatellite loci (SSRs) analyses

Taxon	Location	Code	State	Longitude (E)	Latitude (N)	Sample size		cpDNA Haplotype (Sample numbers)
						SSR	cpDNA	
<i>Tripterygium wilfordii</i>	Tongcheng County, Hubei	ETC	Wild	113°53'43.880"	29°16'06.640"	20	17	Hap1 (17)
	Yueyang City, Hunan	YYY	Wild	113°18'54.690"	29°02'30.920"	20	13	Hap4 (13)
	Taining County, Fujian	FTN	Wild	117°21'35.690"	27°45'16.990"	20	13	Hap1 (13)
	Xinchang County, Zhejiang	ZXC	Cultivar	120°56'32.606"	29°21'36.332"	9	6	Hap1 (6)
	Lanxi City, Zhejiang	ZLX	Wild	119°45'16.600"	29°20'54.300"	22	13	Hap4 (13)
	Pingxiang City, Jiangxi	JPX	Wild	113°55'58.740"	27°42'55.147"	21	11	Hap1 (11)
	Qimen County, Anhui	WQM	Wild	117°23'19.250"	29°56'51.410"	10	9	Hap2 (9)
<i>Tripterygium hypoglaucum</i>	Longhui County, Hunan	XLH	Wild	110°44'40.730"	27°32'03.010"	22	14	Hap4 (14)
	Jianhe County, Guizhou	GJH	Cultivar	108°34'07.170"	26°39'27.020"	9	5	Hap5 (4), Hap6 (1)
	Leishan County, Guizhou	GLS	Wild	108°10'57.394"	26°22'36.866"	21	16	Hap5 (15), Hap6 (1)
	Suichuan County, Jiangxi	JSC	Wild	114°00'24.939"	26°14'33.489"	20	13	Hap3(13)
	Yuxi City, Yunnan	YYW	Wild	102°10'17.350"	23°56'06.560"	21	21	Hap6 (21)
	Jixi County, Anhui	WJX	Wild	118°45'19.500"	30°05'40.500"	9	9	Hap7 (9)
	Huangshan City, Anhui	WSMX	Wild	118°14'21.000"	30°09'02.100"	20	19	Hap8 (19)
Total						244	179	

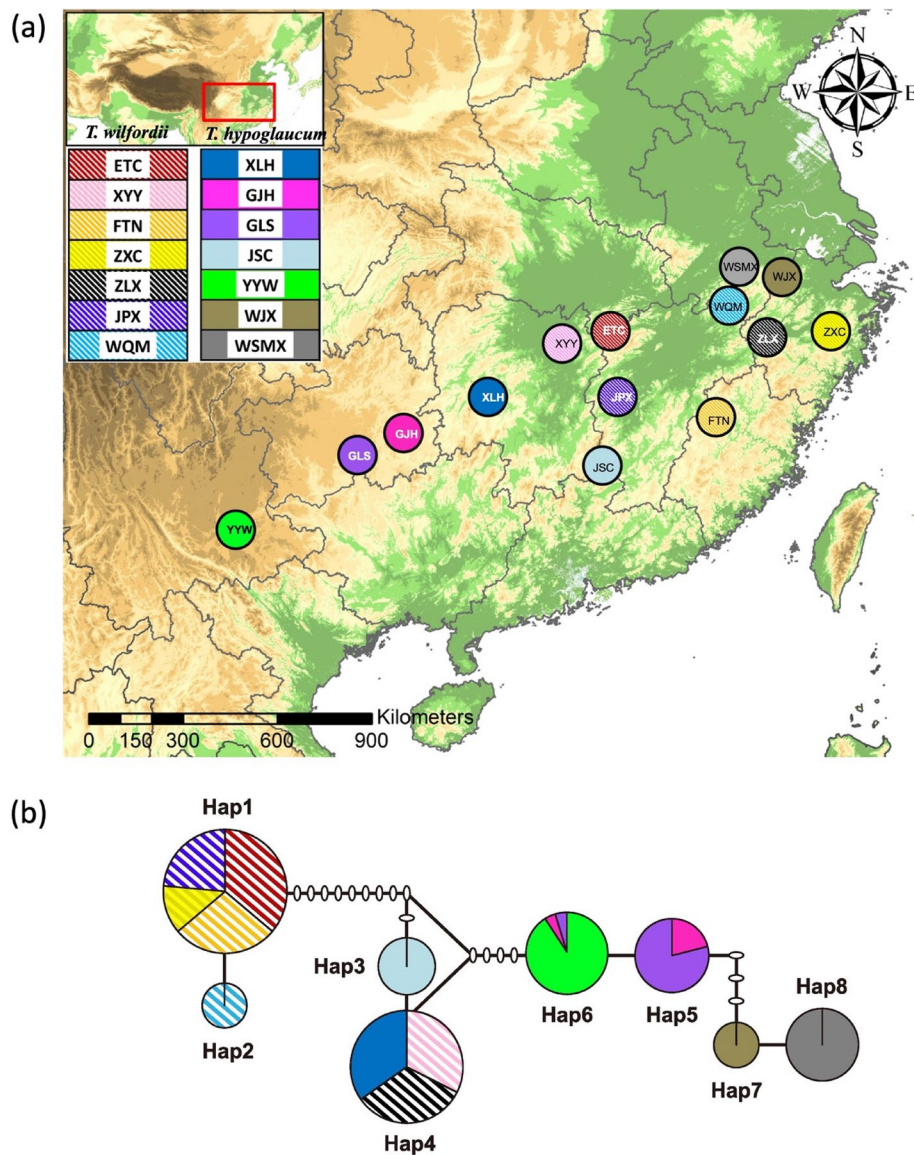
The abbreviations of code denote sample location. Professor Lisong Wang undertook the formal identification of the voucher specimens, which were deposited in herbaria of Institute of Chinese Materia Medica (CMMI), China Academy of Chinese Medical Sciences

inspection of Fig. 1a, the wild populations distributions of *T. wilfordii* and *T. hypoglaucum* had obviously different characteristics: *T. wilfordii* was mainly restricted to the central range of the two species distributions, whereas *T. hypoglaucum* was across the whole range, being divided into the northeastern group (WJX, WSMX) and the southwestern group (XLH, JSC, GLS, YYW).

Overall, Hap1 is the dominant haplotype of *T. wilfordii*: it was identified in the three wild populations (ETC, FTN, JPX) and the cultivated population (ZXC). Hap2, with only one mutation segregating from Hap1, was identified in an adjacent population (WQM), while a very distinct haplotype (Hap4), with at least 12 mutations segregating from Hap1, was identified in two wild populations (XXX and ZLX) of *T. wilfordii*. All the populations of *T. wilfordii* compose only one haplotype from the specimens we collected. In contrast, the wild populations of *T. hypoglaucum* were much more structured and differentiated: many populations were detected private haplotypes, i.e. the northeastern populations (WJX, WSMX) fixed Hap7 and 8, and the southwestern populations (JSC) fixed Hap3. Interestingly, a wild population of *T. hypoglaucum* (XLH) shared Hap4 with an adjacent wild population of *T. wilfordii* (XXX), which are all located in the central range of the geographic distribution. Due to the geographic relations, we may hypothesize that gene flows occurred more frequently between the two populations from two species.

The cultivated populations of *T. wilfordii* and *T. hypoglaucum* had completely different origins. The cultivated population of *T. wilfordii* (ZXC), located in the northeastern of the distribution, was detected with the same haplotype Hap1 as the wild populations (ETC, FTN, JPX) in the central range, which indicated that cultivated *T. wilfordii* may be originated from at least one of them *ex site* (Table 1; Fig. 1a, b). However, the cultivated population of *T. hypoglaucum* (GJH) was fixed at the haplotypes Hap6 with the local adjacent wild population (GLS) (Table 1; Fig. 1a, b), indicating the cultivated origin *in site* of *T. hypoglaucum*.

A total of 14 mutations were identified segregating the two major haplotype clusters: Hap1-2 (dominant in *T. wilfordii*) and Hap5-8 (dominant in *T. hypoglaucum*), which were linked by a shared haplotype (Hap4) (Fig. 1b). The results suggested that the diverging between *T. wilfordii* and *T. hypoglaucum* was significant, while the shared and unresolved haplotypes may need to be further examined by phylogenetic analysis. We, thus, reconstructed gene genealogy using different approaches, i.e. Bayesian Inferences, maximum parsimony and neighbor joining, with cpDNA sequences to elucidate the relations between haplotype clusters (Figure S1). All analyses indicated a consistent phylogeny showing that the eight haplotypes were grouped as two paraphyly lineages (Fig. 2): the basal of the tree was resolved by Hap1 and Hap2 while Hap3, 5–8 were further grouped together as descendant lineages. This result further strengthened the previous



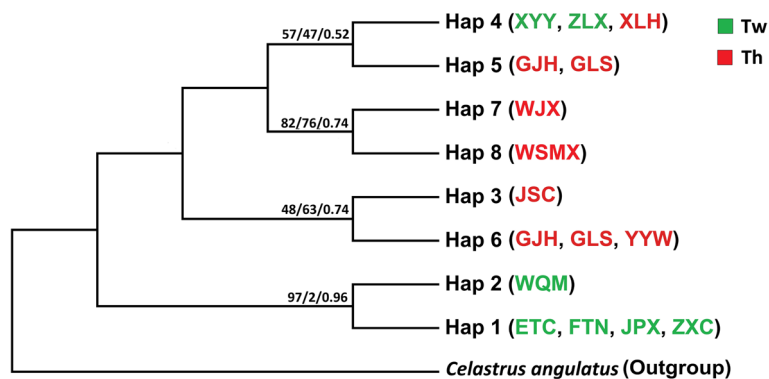
**Fig. 1** **a** Geographic distribution of *Tripterygium wilfordii* and *Tripterygium hypoglaucum* populations. **b** Median-joining networks of eight cpDNA haplotypes (Hap1-8) identified in this study. The sizes of the circles are proportional to the overall frequency of the haplotypes in the entire sample of all population. The white oval are the missing or inferred haplotypes and each line represents one mutational step

**Table 2** Genetic diversity estimated from cpDNA for all populations of *Tripterygium wilfordii* and *Tripterygium hypoglaucum*

	<i>N</i>	<i>P</i>	$\pi$	$\theta_w$	<i>Nh</i>	<i>Hd</i>	<i>D</i>	<i>D</i> *	<i>F</i> *
All species	179	18	0.00650	0.00264	8	0.836	3.89391***	1.68314*	3.04834**
<i>Tw</i>	82	12	0.00424	0.00204	3	0.566	2.95790**	1.47296*	2.36612**
<i>Th</i>	97	11	0.00374	0.00181	6	0.828	2.80140***	1.42315*	2.26723**

*N* sample size, *P* polymorphic sites,  $\pi$  nucleotide diversity,  $\theta_w$  nucleotide polymorphism, *Nh* No. of haplotype, *Hd* Haplotype diversity, *D* Tajima's *D*, *D*\* Fu and Li's *D*\*, *F*\* Fu and Li's *F*\*

Significance level: \* *P* < 0.05, \*\* *P* < 0.02, \*\*\* *P* < 0.001



**Fig. 2** A summarized evolutionary relationship among cpDNA haplotypes of *Tripterygium hypoglaucum* (Th) and *Tripterygium wilfordii* (Tw) inferred from neighbor joining (NJ), maximum parsimony (MP) and Bayesian Inference (BI) methods. The values above the branches indicate bootstrap support values of NJ and MP trees, and Bayesian posterior probabilities (PPs) of BI tree (NJ/MP/BI)

finding of genetic divergence between the two species (Fig. 1b). The phylogeny also suggested that *T. wilfordii* might have the ancestral state among all the haplotypes, as Hap1 and 2 were exclusive in the species. Hap4, a shared haplotype of the two species, also suggested its transition state during the evolution of the two species based on the genealogy results.

#### Genetic diversities and population differentiations were identified using microsatellite loci

In general, microsatellite loci provide more detailed population structures than cpDNA sequences. Therefore, we further analyzed polymorphisms of 10 microsatellite loci to elucidate the population structures of the two species, including 244 individuals. We then estimated 4 different parameters representing population diversities: average numbers of alleles ( $N_a$ ), number of effective alleles ( $N_e$ ), observed heterozygosity ( $H_o$ ) and expected heterozygosity ( $H_e$ ) (Table 3, Table S1 and S2). Among 12 wild and 2 cultivated populations, the average number of alleles ( $N_a$ ) is 6.0 for *T. wilfordii* and 12.5 for *T. hypoglaucum*. Across all the populations, the number of alleles ( $N_a$ ) per locus ranges from 1.4 to 5.0, the number of effective alleles ( $N_e$ ) ranges from 1.40 to 3.52, observed heterozygosity ( $H_o$ ) range from 0.25 to 0.57, and expected heterozygosity ( $H_e$ ) range from 0.20 to 0.58. Overall, populations of *T. hypoglaucum* showed higher genetic diversities (Th:  $N_a = 12.5$ ,  $N_e = 4.52$ ) than those of *T. wilfordii* (Tw:  $N_a = 6.0$ ,  $N_e = 2.32$ ). Nearly all indices indicated that the lowest value was observed in the population of *T. wilfordii*, while the highest value was observed in those of *T. hypoglaucum*, with the exception of the value of  $H_o$  (both lowest and highest values were found in populations of *T. wilfordii*). However, *T. wilfordii* was found to have higher observed heterozygosity ( $H_o = 0.49$ ) than that of *T. hypoglaucum* ( $H_o = 0.43$ ), which was almost equal

**Table 3** Average number of alleles ( $N_a$ ), number of effective alleles ( $N_e$ ), observed heterozygosity ( $H_o$ ) and expected heterozygosity ( $H_e$ ) of microsatellite loci for two species of *Tripterygium*. <sup>a</sup>indicates cultivated populations of the two species. Detailed results of all the parameters are shown in Table S1 and S2

Species	Population	$N_a$	$N_e$	$H_o$	$H_e$
<i>T. wilfordii</i>	All populations	6.0	2.32	0.49	0.51
	ETC	3.1	1.86	0.45	0.37
	XYY	3.6	2.30	0.37	0.42
	FTN	1.8	1.79	0.74	0.40
	ZXC <sup>a</sup>	1.4	1.40	0.40	0.20
	ZLX	2.3	1.67	0.57	0.33
	JPX	2.0	1.61	0.50	0.30
	WQM	2.5	1.59	0.25	0.32
<i>T. hypoglaucum</i>	All populations	12.5	4.52	0.43	0.73
	XLH	3.8	2.21	0.49	0.46
	GJH <sup>a</sup>	3.1	1.82	0.32	0.38
	GLS	3.8	2.26	0.42	0.44
	JSC	2.8	1.80	0.36	0.39
	YYW	5.0	2.57	0.38	0.47
	WJX	3.9	2.72	0.44	0.50
	WSMX	5.0	3.52	0.53	0.58

to its expected heterozygosity ( $H_e = 0.51$ ). This indicates that *T. wilfordii* experienced strong gene exchange and significant random mating among populations, but *T. hypoglaucum* had more intense inbreeding and isolation among populations.

As expected, the two cultivated populations (ZXC, GJH) significantly reduced genetic diversity compared with the wild populations (Table 3). The lowest numbers of alleles, effective alleles and expected heterozygosity were observed from ZXC (*T. wilfordii*) among all 14

populations. In addition, all the values of genetic diversity ( $N_a$ ,  $N_e$ ,  $H_o$  and  $H_e$ ) in GJH (*T. hypoglaucum*) were also lower than those of the 6 wild populations of the species.

#### Species divergence and population subdivision, particularly in *T. hypoglaucum*, were revealed using both cpDNA and microsatellite loci

As there is no significant difference in morphological characteristics, geographical distribution, and chemical composition between *T. wilfordii* and *T. hypoglaucum*, it is sometimes difficult to distinguish the two species. Here, we combined all the individuals from the two species and performed AMOVA to show the genetic discrepancies between the two species using both chloroplast and nucleus markers. Interestingly, we discovered very different percentages of variations (% variation in Table 4) using different markers.

Overall, the variations of cpDNA sequences revealed that only 0.2% of the genetic variations were partitioned within populations, while 51.33% was among populations within each species, and 48.47% was among species (Table 4). This indicates that the populations differ significantly, likely due to genetic divergence of the two species. In addition, we identified very high fixation indices between populations ( $F_{ST} = 0.9980$ ) using cpDNA sequences. Unlike cpDNA sequences, polymorphisms in microsatellite loci indicated that 50.71% of the genetic variation was partitioned within populations, 29.81% was among populations within each species, and 19.48% was among species (Table 4). In line with cpDNA sequences, fixation indices estimated by microsatellite loci also suggest the highest inbreeding coefficient between populations in the two species ( $F_{ST} = 0.4929$ ). Taken together,

the results showed the genetic discrepancies among populations, which indicates the differences in the genetics of the two species.

Genetic structure at a finer resolution, i.e. genetic clustering, was analyzed using STRUCTURE based on both markers (Table S3). Different numbers of clusters (K: estimated number of groups/clusters) were estimated from samples based on the variations in cpDNA sequences or microsatellite loci (Fig. 3). For cpDNA, the best and second best-fit numbers of K values were K=2 and K=3 based on  $\Delta K$ , which indicated that the two species were clustered into two or three groups ( $\Delta K = 993.57$  for K=2 and  $\Delta K = 479.28$  for K=3 based on Bayesian clustering analyses) (Fig. 3a). With an estimation of 2 clusters (K=2), *T. hypoglaucum* (mostly in blue bars) and *T. wilfordii* (mostly red bars) were divided into two differentiated clusters except for ZLX and XYY populations of *T. wilfordii*. The two populations showed closer relations with *T. hypoglaucum*, which is in line with the haplotype analyses: they shared haplotype (Hap4) with XLH of *T. hypoglaucum*. With an estimation of 3 clusters (K=3), *T. hypoglaucum* was further divided into two clusters corresponding to the geographic subdivisions: populations in the northeast (bars in green) and populations in the southwest (bars in blue).

For microsatellite loci, the optimal estimated number of clusters were 2, 3 and 7, with highest  $\Delta K$  values 9.28, 5.42 and 9.84, respectively (Table S1). Different from cpDNA, *T. hypoglaucum* and *T. wilfordii* were divided into two distinct groups without exception, in contrast to cpDNA analysis with K=2 (Fig. 3b). With an estimation of 3 clusters (K=3), the northeast populations (bars in green) and southwest populations (bars in blue) were

**Table 4** Analysis of Molecular Variance (AMOVA) and  $F$  test of fixation indices based on cpDNA and microsatellite loci for *T. wilfordii* and *T. hypoglaucum*

	Source of variation	d.f.	Sum of squares	Variance components	% variation
cpDNA	Among groups	1	269.129	2.59380	48.47
	Among populations within groups	12	413.86	2.74674	51.33
	Within populations	165	1.738	0.01053	0.20
	Total	178	684.726	5.35107	
Microsatellites	Among groups	1	242.112	0.79625	19.48
	Among populations within groups	12	525.638	1.21859	29.81
	Within populations	474	979.896	2.07286	50.71
	Total	487	1747.464	4.08769	

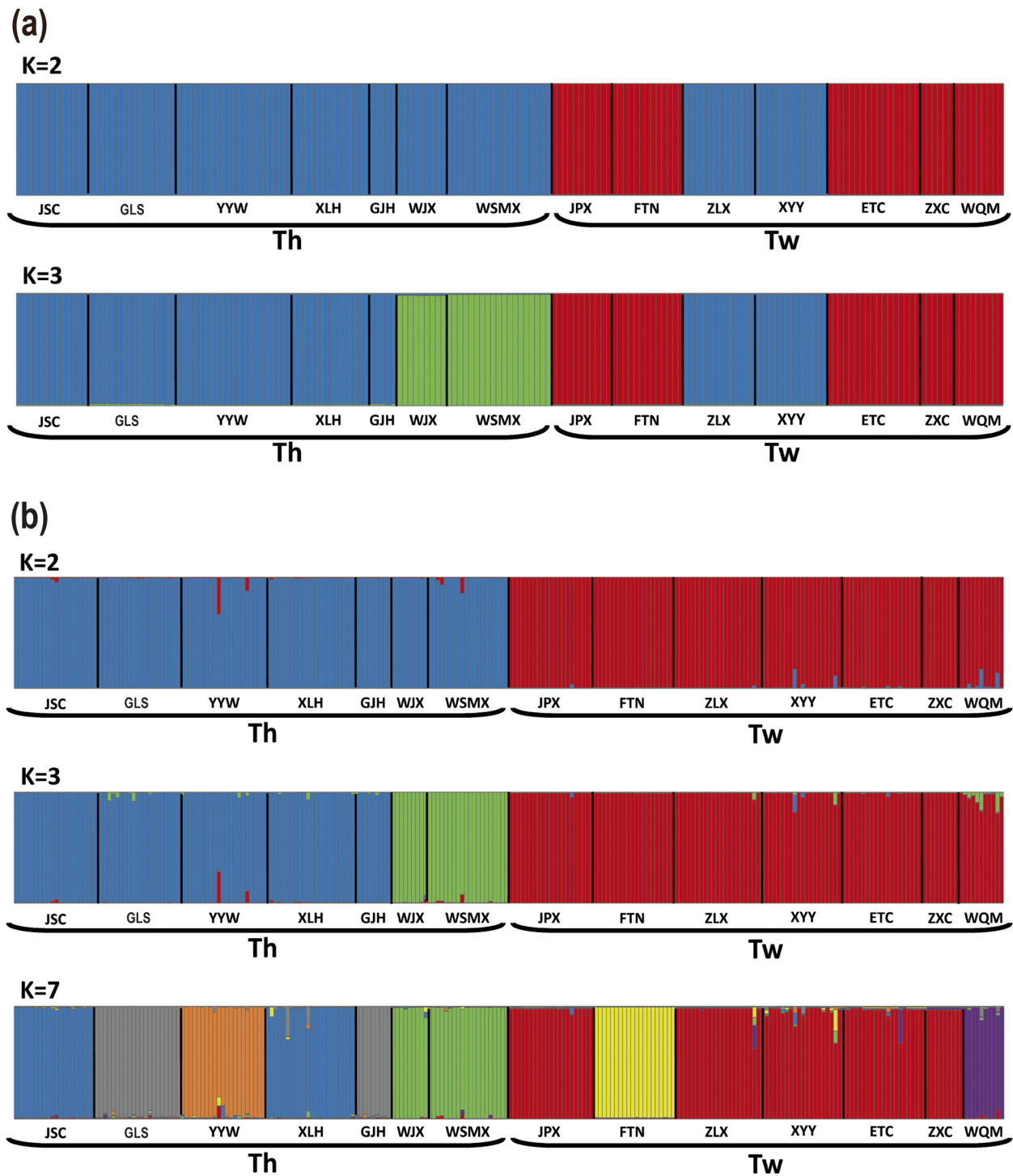
Fixation indices of cpDNA:  $F_{IT} = 0.4847$ ;  $F_{ST} = 0.9980$ ;  $F_{IS} = 0.9962$

Fixation indices of microsatellites:  $F_{IT} = 0.1948$ ;  $F_{ST} = 0.4929$ ;  $F_{IS} = 0.3702$

$F_{IT}$ : overall

$F_{ST}$ : inbreeding coefficient between populations

$F_{IS}$ : inbreeding coefficient within populations



**Fig. 3** Clustering analysis of (a) cpDNA and (b) microsatellite loci for *Tripterygium hypoglaucum* (Th) and *Tripterygium wilfordii* (Tw) populations using STRUCTURE. Bold lines distinguish populations, and different colors within each population represent the probability of assigning the individuals to each cluster. The alphabet below the bar indicates the population code

able to be distinguished as seen in cpDNA. With an estimated number of 7 clusters showed that more sub-clusters can be identified within each species, i.e. 4 and

3 clusters can be identified in *T. hypoglaucum* and *T. wilfordii*, respectively. Among 4 clusters in *T. hypoglaucum*, the southwest populations (bars in blue when K=3)



were further divided into central-range sub-clusters (bars in blue) and southwest sub-clusters (bars in grey and orange). However, 3 clusters in *T. wilfordii* did not correspond to the geographical structure. This also indicates that a finer scale and resolution of population structures can be identified using microsatellite loci compared to cpDNA haplotypes.

The genetic homogeneity of population samples between *T. wilfordii* and *T. hypoglaucum* was further tested by principal coordinate analysis (PCoA) based on microsatellite data (Fig. 4). At the species level, the genetic compositions of *T. wilfordii* and *T. hypoglaucum* can be obviously distinguished into two clusters at the first axis (explained 37.92% variations) and the second principle (explained 19.33% variations). At the population level, northeast, central-range, and southwest population sub-clusters in *T. hypoglaucum* observed by STRU CTURE analysis were also separated into three groups on the second axis. However, the populations in *T. wilfordii* were not further subdivided into different groups. For cultivated and wild populations of both species, the genetic compositions showed no difference from that of their wild populations at either of the two principles.

Taken together, we demonstrated that different genetic markers and analysis methods all pointed out the genetic divergence of the two species even though they have indistinguishable morphologies and are without geographic isolations.

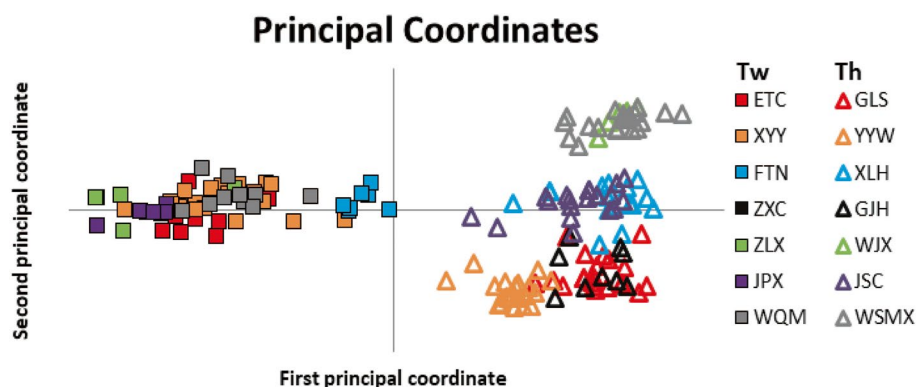
#### Demographic dynamics and historical gene flows between species

As both genetic markers confirmed the species divergence, we would like to ask how the population size fluctuates in the two species and whether there were gene flows between the two species historically.

The marginal distribution of posterior probabilities for demographic parameters was estimated by IMA analyses based on both markers, including the highest posterior parameter estimate (HiPt) and lower to upper bounds of 95% highest posterior densities (HPD) intervals (Table S4, Figure S2). The effective population sizes of *T. hypoglaucum* ( $N_1=0.605$  and 0.195 from cpDNA and microsatellite markers, respectively) were larger than that of *T. wilfordii* ( $N_2=0.195$  and 0.054 from cpDNA and microsatellite markers, respectively) (Table S4). However, the effective population sizes of both species are significantly lower than those of their ancestral populations ( $N_A = 8.461$  and 421.445 from cpDNA and microsatellite markers, respectively). This indicates that both species may have undergone population reduction and recently experienced a change in population size again.

The migration estimates between *T. hypoglaucum* and *T. wilfordii* were asymmetric and different using cpDNA and microsatellite loci (Table S4). The cpDNA gene flow from *T. hypoglaucum* to *T. wilfordii* ( $M_{1\rightarrow 2} = 0.015$ , 95% HPD interval: 0.015–5.595) was lesser than that in the opposite direction ( $M_{2\rightarrow 1} = 0.075$ , 95% HPD interval: 0.025–12.025), while gene flow estimated by microsatellite loci showed different results: significantly larger gene flow from *T. hypoglaucum* to *T. wilfordii* ( $M_{1\rightarrow 2} = 6.125$ , 95% HPD interval: 1.625–18.575;  $M_{2\rightarrow 1} = 0.813$ , 95% HPD interval: 1.188–62.813) was detected.

The results showed that shallow genetic exchange could be identified between the two species, which is in line with the results of haplotype diversity: only one haplotype (Hap4) was shared between the two species based on cpDNA sequences. The exchange might be due to closer geographic relations as the two populations (XXX of *T. wilfordii* and XLH of *T. hypoglaucum*) with Hap4 are both in Hunan Province.



**Fig. 4** Principal coordinate analysis (PCoA) based on microsatellite data for all populations of *Tripterygium wilfordii* (Tw) and *Tripterygium hypoglaucum* (Th). (The first and second coordinate described 37.92% and 19.33% of total variations, respectively)

## Discussion

The sustainable use of a species relies on properly managing the source plant to avoid over-gathering of the wild population, which may result in severe depletion of the number of individuals and genetic pools [50]. In addition, understanding a species' genetic compositions and population structures provides the most fundamental knowledge in maintaining their genetic diversities for sustainable use, particularly medicinal plants. In this study, we investigated genetic variations and population differentiations of two widely used medicinal semiwoody vines plants in traditional Chinese medicine, *T. wilfordii* and *T. hypoglaucum*, by using cpDNA and microsatellite loci as genetic markers. From both markers, all the analyses pointed out that the populations of the two species were well grouped into two major clusters according to their species. *T. wilfordii* showed lower genetic diversity of both markers than *T. hypoglaucum*, but it was found to have higher nucleotide divergence of cpDNA and observed heterozygosity of microsatellite loci than those of *T. hypoglaucum*. Further clustering analyses showed that subdivided population structures can be identified in *T. hypoglaucum* across their geographic distribution, i.e. northeast and southwest sub-clusters. In line with our expectation, cultivated populations (ZXC of *T. wilfordii* and GJH of *T. hypoglaucum*) showed a reduction in genetic diversities compared to wild populations, which may result in the loss of the ability to select the strains with high content of medicinal compounds in the long run. This suggests that an urgent effort is required to select individuals from both species to preserve the genetic diversity for future preservation purposes.

The taxonomic investigation suggested one single species in *Tripterygium* because of the indistinguishable morphological differences [30, 35]. DNA barcoding [22] and phytogeographic investigation [34] showed that *T. wilfordii* and *T. hypoglaucum* could be conspecific, but *T. regelii* should be treated separately. However, *T. wilfordii* and *T. hypoglaucum* are two different kinds of materia medica due to different pharmacological activities in traditional Chinese medicine [18, 19, 21], although their identification was difficult by morphological and pharmacognostic methods [18, 37, 51]. In this study, the phylogenetic trees (NJ, MP, and BI) and median-joining network of cpDNA indicated that *T. wilfordii* and *T. hypoglaucum* were clustered into two significant divergent clades except for few shared and unresolved haplotypes (Figs. 1 and 2). Similar results were shown by STRUCTURE analyses based on the cpDNA (Fig. 3). However, significant genetic differentiation between *T. wilfordii* and *T. hypoglaucum* was further revealed by STRUCTURE and PCoA analyses based on microsatellite loci (Figs. 3 and 4). These results suggest that the

taxonomic treatments of two species *T. wilfordii* and *T. hypoglaucum* made by Cheng (1999) were justified, which suggested that genetic bases for the content of bioactive compounds of the two species should be different [18, 29]. Phenotypic similarities and shared lesser haplotypes both show a certain degree of similarity between *T. wilfordii* and *T. hypoglaucum*. However, the more refined distinction of microsatellite marker between *T. wilfordii* and *T. hypoglaucum* has verified the differentiation of the two species. As for the phenotypic similarities between *T. wilfordii* and *T. hypoglaucum*, the most likely explanation may be the retention of ancestral polymorphisms due to incomplete lineage sorting [52]. Furthermore, the similarities may also be ascribed to pollen-mediated gene flow among the two species and/or phenotypic convergence under similar selection schemes in the same geographical distribution region.

The analysis of migration estimates revealed inconsistent results between *T. hypoglaucum* and *T. wilfordii* when using cpDNA and microsatellites. The results from cpDNA showed that the gene flow intensity from *T. wilfordii* to *T. hypoglaucum* was greater than vice versa. However, microsatellite loci analysis reveals the opposite, indicating a stronger gene flow from *T. hypoglaucum* to *T. wilfordii*. The discrepancies between cpDNA and microsatellite results can be attributed to their different transmission modes and evolutionary rates. Chloroplast DNA, a primarily maternally inherited organelle marker, is dispersed by seeds and evolves at a relatively slow rate, reflecting more historical, long-term gene flow [53, 54]. The phylogenetic trees derived from the study (Fig. 2) further substantiate this inference. In contrast, microsatellite is biparentally inherited, transferred via pollen flow, and evolves faster, reflecting more recent genetic exchanges [53–55]. On the other hand, the difference in effective population sizes can also indirectly provide an explanation for the inconsistency in migration estimates. Evaluations based on effective population sizes reveal that, regardless of cpDNA and microsatellite, the effective population sizes of *T. hypoglaucum* were larger than those of *T. wilfordii*. This characteristic could enable greater gene flow for *T. hypoglaucum*, thus explaining the microsatellite analysis results. Additionally, the cultivation method of *T. wilfordii* is mainly wild transplantation, seedling cultivation, and branch-cutting seedling transplantation. Unlike *T. wilfordii*, *T. hypoglaucum* is mainly propagated by cutting seedlings. These distinct reproduction methods may also lead to different gene flow patterns.

From the literature, the mean value of cpDNA haplotype diversity ( $H_d$ ) detected by various cpDNA markers is 0.67 in 170 species of plants compiled by Petit et al. (2005) [53]. The cpDNA haplotype diversity in *T. wilfordii*

( $Hd=0.566$ ) was low relative to those plants and was significantly lower than that in *T. hypoglaucum* ( $Hd=0.828$ ). *Tripterygium wilfordii* and *T. hypoglaucum* are plants in an archaic taxon, Celastraceae, with a long evolutionary history. Unlike other species in Celastraceae, e.g. *Dipentodon sinicus*, showing very high haplotype diversity ( $Hd=0.902$ ) [56], *T. wilfordii* showed unexpectedly low haplotype diversity in this study. The possible reason might be the over-exploitation due to the heavy usage of the species for medicinal purposes. Demographic dynamics also showed that *T. wilfordii* has undergone the recent contraction of effective population sizes (Table S5, Fig. 2). Therefore, there is an urgent need to develop conservation and cultivation practices for *T. wilfordii*.

Cultivation of wild plants always produces genetic bottlenecks and thus results in loss of genetic diversity due to founder effects and unconscious or conscious selections [7]. The cultivation of *Tripterygium* plants originated from local farmers' practice in traditional agricultural habitats with possibly less than fifty years of history [57]. Compared with long-term and large-scale cultivated medicinal plants in China (e.g., *Coptis chinensis* var. *chinensis*; *Scrophularia ningpoensis*) [58, 59], recently cultivated medicinal plants used to have larger population size and thus likely be able to maintain greater genetic diversity due to multiple origins and shorter artificial selection time (e.g., *Scutellaria baicalensis*; *Magnolia officinalis* var. *biloba*) [60, 61]. However, in the study, significantly lower genetic diversity of cultivated *T. wilfordii* and *T. hypoglaucum* was revealed: the main values of genetic diversity of microsatellite loci ( $N_a$ ,  $N_e$  and  $H_e$ ) in cultivated *T. wilfordii* were the lowest among all the 14 populations, and these values of cultivated *T. hypoglaucum* were also lower than those of its 6 wild populations (Table 3). These results may be explained by the small initial population size and vegetative propagation during the cultivation of the two species. First, The geographical distribution of cpDNA haplotypes indicates that only one haplotype (Hap1) and two haplotypes (Hap5 and 6) were brought into cultivated populations of *T. wilfordii* and *T. hypoglaucum*, respectively, from at least one wild population. (Fig. 1; Table 1). The evidence supports that cultivated *T. wilfordii* or *T. hypoglaucum* could have originated from one place for a single time, resulting in a small initial population size during cultivation. Second, cultivation of *Tripterygium* plants was mainly from cuttage because of several issues such as the cost of seed gathering, lower germination rate, and investment in factory equipment [57]. This cultivation practice inevitably led to a shift from sexually reproducing wild populations to vegetatively propagating cultivated populations [62, 63] and a significant reduction of genetic diversity in the cultivated populations [12, 63].

The decline of genetic diversity in cultivation not only results in a narrow genetic range of material to meet breeding objectives such as selections for high-yielding or pharmacological standard individuals but also might cause the loss of evolutionary potential of cultivated species due to inbreeding. These experiences were revealed by many crop species [64, 65] and a few medicinal plants, such as *Scrophularia ningpoensis* [58]. Therefore, the integrity of genetic diversity is urgently needed to be conserved during the cultivation of *Tripterygium* plants.

Facing the rapidly growing demands for materia medica of *T. wilfordii* for its new pharmacological activities, the sustainable utilization of *T. wilfordii* resources has become an urgent problem that needs to be solved. The study on the genetic diversity pattern of *T. wilfordii* and its sibling species *T. hypoglaucum* is the precondition for the sustainable utilization of *T. wilfordii* resources. Our results suggest that the taxonomic treatments of two species *T. wilfordii* and *T. hypoglaucum* made by Cheng (1999) are justified [29], which provides genetic bases for the contents of bioactive compounds of the two species being different. Over-exploitation might result in significantly lower genetic diversity and less subdivided genetic structure of *T. wilfordii* than those of *T. hypoglaucum*. During the course of cultivation of the two species, genetic diversity should be increased and the introduction and cultivation *ex site* should be avoided. This study not only provides baseline data for protecting and utilizing the genetic resource of *T. wilfordii* through a conservation-by-cultivation approach but also represents a paradigm for instructing future cultivation projects of important medicinal plants with new pharmacological activities.

## Materials and methods

### Sampling, DNA extraction and microsatellite genotyping

Field investigations were conducted throughout the distribution range of *T. wilfordii* and *T. hypoglaucum*. A total of 244 individuals representing 7 *T. wilfordii* and 7 *T. hypoglaucum* populations were collected (Table 1; Fig. 1a). Five to twenty-two individuals were sampled for each population, and the distances of each population were more than 50 km. Fresh leaves of each individual were collected and dried with silica gel for DNA extraction. Dr. Lisong Wang conducted the formal identification of the voucher specimens, and voucher specimens have been deposited in the herbaria of the Institute of Chinese Materia Medica (CMMI), China Academy of Chinese Medical Sciences. The precise geographic location of each sampled population was determined using a Garmin GPS unit. The distribution map was generated using Esri's ArcGIS v9.3 software, incorporating the ESRI

World Elevation GMTED2010 dataset and source from USGS.

Total genomic DNA was extracted using a modified cetyltrimethyl ammonium bromide (CTAB) protocol [66]. The relative purity and concentration of extracted DNA were estimated by ethidium bromide staining on agarose gels compared with known DNA concentration markers. Two cpDNA regions *psbA-trnH* and *trnL-trnF* were chosen for amplification and sequencing using the primers indicated in Taberlet et al. (1991) and Sang et al. (1997) [67, 68], respectively (Table S5). PCR amplifications were carried out in a volume of 20  $\mu$ L using 1  $\mu$ L of template DNA (50–100 ng), 2  $\mu$ L of 10 $\times$ reaction buffer, 1.6  $\mu$ L dNTP mix (2.5 mM), 1.25  $\mu$ L 10  $\mu$ M of each primer, 0.2  $\mu$ L Ex-Taq DNA polymerase (Takara Shuzo Co., Ltd., Otsu, Japan), 12.7  $\mu$ L sterile distilled water. Reactions run on a Veriti thermocycler (Applied Biosystems, USA). The protocol for *psbA-trnH* consisted of one cycle of denaturation at 95  $^{\circ}$ C for 5 min; 35 cycles of 1 min denaturation at 94  $^{\circ}$ C, 1 min of annealing at 55  $^{\circ}$ C and 1 min 30 s of extension at 72  $^{\circ}$ C; followed by an 8 min extension at 72 $^{\circ}$ C. The protocol for *trnL-trnF* consisted of one cycle of denaturation at 94  $^{\circ}$ C for 4 min; 35 cycles of 45 s denaturation at 94  $^{\circ}$ C, 30 s annealing at 58  $^{\circ}$ C, and 1 min 30 s extension at 72  $^{\circ}$ C; then 7 min extension at 72  $^{\circ}$ C. All the purified PCR products were sequenced directly in both directions on an ABI 3730XL automated sequencer (Applied Biosystem, Foster City, CA).

A total of 10 microsatellite primer pairs derived from Novy & Jones (2011) [69] were chosen for this study (Table S6). Amplification of DNA was carried out in 25  $\mu$ L reactions consisting of 1  $\mu$ L DNA, 2.5  $\mu$ L of 10 $\times$ reaction buffer, 2  $\mu$ L dNTP (2.5mM), 0.4  $\mu$ L each primer, and 0.25  $\mu$ L rTaq DNA polymerase (5U/ $\mu$ L), 18.45  $\mu$ L sterile distilled water. The reaction was optimized and programmed on a PCR System 9700 (Applied Biosystems GeneAmp, USA) for one cycle of denaturation at 94  $^{\circ}$ C for 5 min, 40 cycles of 30 s denaturation at 94  $^{\circ}$ C, 30 s at proper annealing temperatures, and 30 s extension at 72  $^{\circ}$ C; then 10 min extension at 72  $^{\circ}$ C. The PCR amplified products were scored using an ABI 3730XL automated sequencer using a 50-cm capillary, polymerPOP-7 and ROX 500 (both Applied Biosystems) as an internal standard. Fragment sizes were assessed using GeneMarker software version 1.5. Allele size is indicated twice manually to reduce scoring error.

#### Sequence alignment, population genetics statistics, and genealogy of cpDNA

All 244 samples were used in analyses of microsatellite markers (SSRs), and only 179 were used in cpDNA sequencing analyses because partial samples could not be sequenced due to degradation (Table 1). All sequences

were edited and assembled by SeqMan 7.1.0 [70] and BioEdit ver. 7.0.9 [71] with manual correction. Sequence alignments were performed with CLUSTALX version 2.0 [72] and then manually modified. Two cpDNA regions were combined and treated as a single marker. Summary statistics, including the number of haplotypes ( $N_h$ ) and haplotype diversity ( $H_d$ ) were determined. The level of genetic diversity within populations was calculated by measures of nucleotide divergence,  $\theta_W$  [73] and  $\pi$  [74, 75], using DnaSP Version 5.10.1 [76]. The signatures of demographic change were examined by neutrality tests using DnaSP, including Tajima's  $D$  statistic [77], Fu and Li's  $D^*$  statistic [78], Fu and Li's  $F^*$  tests. An analysis of molecular variance (AMOVA) [79] was conducted to estimate the genetic variation within and among populations grouped by species by using the program Arlequin v.3.5.1.2 [80].

Phylogenetic relationships among cpDNA haplotypes of *T. wilfordii* and *T. hypoglaucum* were inferred using maximum parsimony (MP) in PAUP\* 4.0b10 [81] and Bayesian Inference (BI) implemented in MrBayes v. 3.1.2 [82, 83] with *Celastrus angulatus* (Celastraceae) as the outgroup. For BI inference, the normal prior probability distribution was chosen as prior. All parameters once every 1000 steps from  $10^7$  Markov coupled Markov chain (MCMC) steps, with the first 10% of samples discarded as burn-in. The TRACER program v.1.6 [84] was used to examine the convergence of chains to the stationary distribution. Trees were then compiled as maximum clade credibility trees by TREEANNOTATOR v.1.7.5 [85] and visualized by FIGTREE v.1.4 [86] to display the highest posterior density (HPD) intervals at 95% (upper and lower) for each node. Neighbor joining (NJ) phylogeny relationships of cpDNA haplotypes that was based on the neighbor-joining method were constructed using MEGA v. 7.02 [26]. The degree of relatedness among cpDNA haplotypes was also estimated using Network v. 4.2.0.1 [87]. with indels that were treated as single mutational events.

#### Population genetic analysis and gene flow

For each microsatellite locus, genetic diversity was assessed by the average number of alleles ( $N_a$ ), the number of effective alleles ( $N_e$ ), the observed heterozygosity ( $H_o$ ) and the expected heterozygosity ( $H_e$ ), which were estimated by using GenAlEx v.6.5 [88]. The population structure was evaluated to estimate the distinction between the two species by calculating the population differentiation level (Weir and Cockerham's  $F_{ST}$ ) [89] in AMOVA analysis implemented in Arlequin v.3.5.1.2., partitioning the genetic diversity into three levels: among species (groups), among populations within species (groups) and within populations.

Significance tests were conducted using 10,000 permutations. Pairwise  $F_{ST}$  among populations and species were also estimated hierarchically in the program DnaSP. The principal coordinate analysis (PCoA) was also conducted using GenAlEx 6.5. Genetic groups were inferred using a Bayesian clustering method implemented in STRUCTURE v.2.3.4 [90]. by using both cpDNA and microsatellite loci. A total of 20 independent simulations were run for each ( $K = 1-15$ ) with 500 000 burn-in steps followed by 500 000 MCMC steps using the admixture model with correlated allelic frequencies [91] because hybridization events are common in plants [92]. The true number of  $K$  was determined using the method of Pritchard et al. (2000) [90] and Evanno et al. (2005) [93]. Structure Harvester v.0.6.8 [94] was used to visualize the STRUCTURE output.

The rates of gene flow between *T. wilfordii* and *T. hypoglaucum* ( $M$ ) and effective population size ( $N$ ) based on cpDNA sequences and microsatellite loci were estimated by using IMA2 software [95]. The IMA analyses were performed using > 3 million burn-in steps per run to ensure all parameters were well mixed, and additional genealogies were saved until all effective sample sizes of parameters were > 100. Three independent runs with random seeds were performed to check the consistency between results. The final results were calculated by averaging the values estimated in the three runs. A swapping procedure was also used to enhance the mixing of chains using a geometric heating scheme with 15 parallel chains. All parameters were scaled by the mutation rate of  $1.52 \times 10^{-9}$  ( $1.0 \times 10^{-9} - 3.0 \times 10^{-9}$ ) substitutions per year for cpDNA by Wolfe et al. (1987) and Richardson & Stojiljkovic (2001) and  $4.76 \times 10^{-3}$  ( $1.2 \times 10^{-4} - 2.7 \times 10^{-2}$ ) per allele per generation for microsatellite loci by Cieslarová et al. (2011) [42, 96, 97].

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12870-024-04826-x>.

**Supplementary Material 1.**

**Supplementary Material 2.**

## Acknowledgments

Not applicable.

## Authors' contributions

Conceptualization, C.L., Q.-J.Y., L.-S.W., Y.-C.C. and L.-P.G.; methodology, J.-Y.W., Y.-Z.K., M.-S.S., Q.-J.Y., and Y.C.C.; software, J.-Y.W., Y.-Z.K. and Y.-C.C.; validation, Y.-H.W., J.-H.S., Q.-J.Y. and Y.C.C.; formal analysis, Y.-Z.K., L.-S.W., Y.-C.C. and L.-P.G.; investigation, C.L., J.-Y.W., Y.-H.W., J.-H.S., Q.-J.Y., L.-S.W. and L.-P.G.; resources, C.L., Q.-J.Y., L.-S.W., Y.-C.C. and L.-P.G.; data curation, C.L., J.-Y.W., Y.-Z.K., Q.-J.Y. and Y.-C.C.; writing-original draft preparation, C.L., J.-Y.W., Y.-Z.K., M.-S.S., Q.-J.Y., L.-S.W. and Y.-C.C.; writing-review and editing, Y.-Z.K., M.-S.S., Q.-J.Y. and Y.-C.C.; visualization, J.-Y.W. and Y.-Z.K.; supervision, Q.-J.Y., Y.-C.C. and L.-P.G.; project

administration, Q.-J.Y., L.-S.W., Y.-C.C. and L.-P.G.; funding acquisition, Q.-J.Y., L.-S.W., Y.-C.C. and L.-P.G. All authors have read and agreed to the published version of the manuscript.

## Funding

This study was funded by the National Key Research and Development Program of China (2017YFC1703700: 2017YFC1703704), Key Project at Central Government Level: The Ability Establishment of Sustainable Use for Valuable Chinese Medicine Resources (2060302) to Q-JY and L-PG, and by Lushan Talent Special Project (2021ZWZX06) to L-SW and by partial financing of NSYSU to Y-CC.

## Availability of data and materials

All sequences for this study have been uploaded to the NCBI GenBank as follows: Sequences of cpDNA (*psbA-trnH* and *trnL-trnF*) were obtained from a total of 179 individuals and submitted to GenBank with the following accession numbers: PP128529-PP128707 (<https://www.ncbi.nlm.nih.gov/nuccore/PP128529> to <https://www.ncbi.nlm.nih.gov/nuccore/PP128707>) for *psbA-trnH* and PP128708-PP128886 for *trnL-trnF*. (<https://www.ncbi.nlm.nih.gov/nuccore/PP128708> to <https://www.ncbi.nlm.nih.gov/nuccore/PP128886>)

## Declarations

### Ethics approval and consent to participate

*Tripterygium wilfordii* and *T. hypoglaucum* are not endangered species and do not require permissions in this research.

### Consent for publication

Not applicable.

### Competing interests

The authors declare no competing interests.

### Author details

<sup>1</sup>State Key Laboratory Breeding Base of Dao-di Herbs, National Resource Center for Chinese Materia Medica, China Academy of Chinese Medical Sciences, Beijing 100700, China. <sup>2</sup>Department of Biological Sciences, National Sun Yat-sen University, Kaohsiung City, Taiwan. <sup>3</sup>Research Center, Faculty of Medicine, Ramathibodi Hospital, Mahidol University, Bangkok 10400, Thailand. <sup>4</sup>Lushan Botanical Garden, Chinese Academy of Sciences, Jiujiang, Jiangxi 332900, China. <sup>5</sup>Department of Biomedical Science and Environmental Biology, Kaohsiung Medical University, Kaohsiung City, Taiwan. <sup>6</sup>The Multi-disciplinary and Data Science Research Center (MDSRC), National Sun Yat-sen University, Kaohsiung 804, Taiwan.

Received: 2 October 2023 Accepted: 15 February 2024

Published online: 16 March 2024

## References

- Howes MJR, Quave CL, Collemare J, Tatsis EC, Twilley D, Lulekal E, Farlow A, Li L, Cazar ME, Leaman DJ. Molecules from nature: reconciling biodiversity conservation and global healthcare imperatives for sustainable use of medicinal plants and fungi. *Plants People Planet*. 2020;2(5):463–81.
- Ramawat KG, Arora J. Medicinal plants domestication, cultivation, improvement, and alternative technologies for the production of high value therapeutics: an overview. In: Ekiert HM, Ramawat KG, Arora J, editors. *Medicinal plants: domestication, biotechnology and Regional Importance Cham*. Switzerland: Springer; 2021. p. 1–29.
- Schippmann U, Leaman DJ, Cunningham AB. Impact of cultivation and gathering of medicinal plants on biodiversity: global trends and issues. *Biodiversity and the ecosystem approach in agriculture, forestry and fisheries*. Rome: FAO; 2002.
- Worlds Health Organization. Guidelines on the conservation of medicinal plants. Gland: The International Union for Conservation of Nature and Natural Resources (IUCN); 1993.

5. Lambert J, Srivastava J, Vietmeyer N. Medicinal plants: rescuing a global Heritage. Washington D.C.: World Bank; 1997.
6. Schippmann UWE, Leaman D, Cunningham AB. A comparison of cultivation and wild collection of medicinal and aromatic plants under sustainability aspects. In: Bogers RJ, Craker LE, Lange D, editors. Medicinal and aromatic plants: agricultural, commercial, ecological, legal, pharmacological and social aspects, vol. 17. Wageningen, The Netherlands: Springer; 2006. p. 75–95.
7. Doebley JF, Gaut BS, Smith BD. The molecular genetics of crop domestication. *Cell*. 2006;127(7):1309–21.
8. Hyten DL, Song Q, Zhu Y, Choi I-Y, Nelson RL, Costa JM, Specht JE, Shoemaker RC, Cregan PB. Impacts of genetic bottlenecks on soybean genome diversity. *Proc Natl Acad Sci USA*. 2006;103(45):16666–71.
9. Kovach MJ, McCouch SR. Leveraging natural diversity: back through the bottleneck. *Curr Opin Plant Biol*. 2008;11(2):193–200.
10. Hartmann T. From waste products to ecochemicals: fifty years research of plant secondary metabolism. *Phytochem*. 2007;68(22–24):2831–46.
11. Jan R, Asaf S, Numan M, Lubna, Kim KM. Plant secondary metabolite biosynthesis and transcriptional regulation in response to biotic and abiotic stress conditions. *Agronomy*. 2021;11(5):968.
12. Abbo S, Berger J, Turner NC. Evolution of cultivated chickpea: four bottlenecks limit diversity and constrain adaptation. *Funct Plant Biol*. 2003;30(10):1081–7.
13. Guo J, Wang Y, Song C, Zhou J, Qiu L, Huang H, Wang Y. A single origin and moderate bottleneck during domestication of soybean (*Glycine max*): implications from microsatellites and nucleotide sequences. *Ann Bot*. 2010;106(3):505–14.
14. Tanksley SD, McCouch SR. Seed banks and molecular maps: unlocking genetic potential from the wild. *Science*. 1997;277(5329):1063–6.
15. Larson G, Piperno DR, Allaby RG, Purugganan MD, Andersson L, Arroyo-Kalin M, Barton L, Climer Vigueira C, Denham T, Dobney K. Current perspectives and the future of domestication studies. *Proc Natl Acad Sci USA*. 2014;111(17):6139–46.
16. Purugganan MD. Evolutionary insights into the nature of plant domestication. *Curr Biol*. 2019;29(14):R705–714.
17. Ross-Ibarra J, Morrell PL, Gaut BS. Plant domestication, a unique opportunity to identify the genetic basis of adaptation. *Proc Natl Acad Sci USA*. 2007;104:8641–8.
18. Guo L, Duan L, Liu K, Liu EH, Li P. Chemical comparison of *Tripterygium Wilfordii* and *Tripterygium hypoglaucum* based on quantitative analysis and chemometrics methods. *J Pharm Biomed Anal*. 2014;95:220–8.
19. Li XJ, Jiang ZZ, Zhang LY. Triptolide: progress on research in pharmacodynamics and toxicology. *J Ethnopharmacol*. 2014;155(1):67–79.
20. Lv H, Jiang L, Zhu M, Li Y, Luo M, Jiang P, Tong S, Zhang H, Yan J. The genus *Tripterygium*: a phytochemistry and pharmacological review. *Fitoterapia*. 2019;137:104190.
21. Tu L, Su P, Zhang Z, Gao L, Wang J, Hu T, Zhou J, Zhang Y, Zhao Y, Liu Y. Genome of *Tripterygium Wilfordii* and identification of cytochrome P450 involved in triptolide biosynthesis. *Nat Commun*. 2020;11(1):971.
22. Zhang X, Li N, Yao Y, Liang X, Qu X, Liu X, Zhu Y, Yang D, Sun W. Identification of species in *Tripterygium* (Celastraceae) based on DNA barcoding. *Biol Pharm Bull*. 2016;39(11):1760–6.
23. Wong KF, Yuan Y, Luk JM. *Tripterygium Wilfordii* bioactive compounds as anticancer and anti-inflammatory agents. *Clin Exp Pharmacol Physiol*. 2012;39(3):311–20.
24. Liu Q. Triptolide and its expanding multiple pharmacological functions. *Int Immunopharmacol*. 2011;11(3):377–83.
25. Chugh R, Sangwan V, Patil SP, Dudeja V, Dawra RK, Banerjee S, Schumacher RJ, Blazar BR, Georg GI, Vickers SM. A preclinical evaluation of Minnelide as a therapeutic agent against pancreatic cancer. *Sci Transl Med*. 2012;4(156):4085–94.
26. Kumar S, Stecher G, Tamura K. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol*. 2016;33(7):1870–4.
27. Rousalova I, Banerjee S, Sangwan V, Evenson K, McCauley JA, Kratzke R, Vickers SM, Saluja A, D'Cunha J. Minnelide: a novel therapeutic that promotes apoptosis in non-small cell lung carcinoma in vivo. *PLoS One*. 2013;8(10):e77411.
28. Shen Y, Jiang T, Wang R, He S, Guo M, Zuo J, He D. (5R)-5-Hydroxytriptolide (LLDT-8) inhibits osteoclastogenesis via RANKL/RANK/OPG signaling pathway. *BMC Complement Altern Med*. 2015;15(77):1–10.
29. Cheng CY. *Tripterygium*. In: Cheng CY, Huang PH, editors. Flora Republicae Popularis Sinicae. Beijing: Science; 1999. p. 178–81.
30. Jin-Shuang M, Brach A, Liu QR. A revision of the genus *Tripterygium* (Celastraceae). *Edinb J Bot*. 1999;56(1):33–46.
31. Zhao J, Zhang F, Xiao X, Wu Z, Hu Q, Jiang Y, Zhang W, Wei S, Ma X, Zhang X. *Tripterygium Hypoglaucum* (Lévl.) Hutch and its main bioactive components: recent advances in pharmacological activity, pharmacokinetics and potential toxicity. *Front Pharmacol*. 2021;12:715359.
32. Chen YL, Liu X, Qu XY, Yao YY, Li N, Liang XM, Zhang YY, Yang DJ, Zhang XM. Studies on difference of chemical compositions in plant species of *Tripterygium* Genus. *Zhongguo Zhong Yao Za Zhi*. 2017;42(2):319–25.
33. Huang W, Guo B, Si J. Triptolide in different populations and individuals of three species of *Tripterygium* Hook. F. *Zhong Cao Yao*. 2005;36(7):1065–8.
34. Ma B, Hu T, Li P, Yuan Q, Lin Z, Tu Y, Li J, Zhang X, Wu X, Wang X. Phylogeographic and phylogenetic analysis for *Tripterygium* species delimitation. *Ecol Evol*. 2017;7(20):8612–23.
35. Ma J, Funston A. *Tripterygium* Hook. f. (Celastraceae). In: Wu ZY, Raven PH, editors. Flora of China, vol. 11. Beijing: Science; 2008. p. 286–7.
36. Liu WS, Guo BL, Huang WH, Si JP. RAPD analysis for genetic relationship and diversity of three species of genus *Tripterygium*. *Zhongguo Zhong Yao Za Zhi*. 2007;32(16):1615–21.
37. Law SK, Simmons MP, Techen N, Khan IA, He M-F, Shaw PC, But PP. Molecular analyses of the Chinese herb Leigongteng (*Tripterygium Wilfordii* Hook. F.). *Phytochem*. 2011;72(1):21–6.
38. Li N, Yao YY, Chen YL, Liang XM, Zhang DD, Zhang XM, Yang DJ. Evaluation of genetic diversity and population structure of genus *Tripterygium* based on SSR markers. *Acta Pharm Sin*. 2017;52(1):153–61.
39. Nadeem MA, Nawaz MA, Shahid MQ, Doğan Y, Comertpay G, Yıldız M, Hatipoğlu R, Ahmad F, Alsaleh A, Labhane N. DNA molecular markers in plant breeding: current status and recent advancements in genomic selection and genome editing. *Biotechnol Biotechnol Equip*. 2018;32(2):261–85.
40. Ouborg NJ, Piquot Y, Van Groenendael JM. Population genetics, molecular markers and the study of dispersal in plants. *J Ecol*. 1999;87(4):551–68.
41. Ellis JR, Bentley KE, McCauley DE. Detection of rare paternal chloroplast inheritance in controlled crosses of the endangered sunflower *Helianthus verticillatus*. *Hered*. 2008;100(6):574–80.
42. Wolfe KH, Li WH, Sharp PM. Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. *Proc Natl Acad Sci USA*. 1987;84(24):9054–8.
43. Chen J, Hao Z, Xu H, Yang L, Liu G, Sheng Y, Zheng C, Zheng W, Cheng T, Shi J. The complete chloroplast genome sequence of the relict woody plant *Metasequoia glyptostroboides* Hu et Cheng. *Front Plant Sci*. 2015;6:447.
44. Feng X, Wang Y, Gong X. Genetic diversity, genetic structure and demographic history of *Cycas simplicipinna* (Cycadaceae) assessed by DNA sequences and SSR markers. *BMC Plant Biol*. 2014;14(1):1–16.
45. Ge XJ, Hsu TW, Hung K-H, Lin CJ, Huang CC, Huang CC, Chiang YC, Chiang TY. Inferring multiple refugia and phylogeographical patterns in *Pinus massoniana* based on nucleotide sequence variation and DNA fingerprinting. *PLoS One*. 2012;7(8):e43717.
46. Ge XJ, Hung KH, Ko YZ, Hsu TW, Gong X, Chiang TY, Chiang YC. Genetic divergence and biogeographical patterns in *Amentotaxus argotaenia* species complex. *Plant Mol Biol Rep*. 2014;33:264–80.
47. Gong YQ, Zhan QQ, Nguyen KS, Nguyen HT, Wang YH, Gong X. The historical demography and genetic variation of the endangered *Cycas multipinnata* (Cycadaceae) in the red river region, examined by chloroplast DNA sequences and microsatellite markers. *PLoS One*. 2015;10(2):e0117719.
48. Liao PC, Tsai CC, Chou CH, Chiang YC. Introgression between cultivars and wild populations of *Momordica charantia* L. (Cucurbitaceae) in Taiwan. *Int J Mol Sci*. 2012;13(5):6469–91.
49. Zhang X, Su H, Yang J, Feng L, Li Z, Zhao G. Population genetic structure, migration, and polyploidy origin of a medicinal species *Gynostemma pentaphyllum* (Cucurbitaceae). *Ecol Evol*. 2019;9(19):11145–70.
50. Zhao-Seiler N. Sustainability of Chinese medicinal herbs: a discussion. *J Chin Med*. 2013;101:52–6.
51. Dai KM, Qin WZ. Identification of Leigongteng and Kunming Shanhaitang with scanning electron microscope. Shanghai: The Fourth National Symposium on Leigongteng; 2004. p. 134–6.
52. Avise JC. Molecular markers, natural history, and evolution. 2nd ed. Sunderland: Springer Science & Business Media; 2012.

53. Petit RJ, Duménil J, Fineschi S, Hampe A, Salvini D, Vendramin GG. Comparative organization of chloroplast, mitochondrial and nuclear diversity in plant populations. *Mol Ecol*. 2005;14:689–701.
54. Ennos R. Estimating the relative rates of pollen and seed migration among plant populations. *Heredity*. 1994;72(3):250–9.
55. Harrison HB, Saenz-Agudelo P, Planes S, Jones GP, Berumen ML. Relative accuracy of three common methods of parentage analysis in natural populations. *Mol Ecol*. 2013;22(4):1158–70.
56. Yuan QJ, Zhang ZY, Peng H, Ge S. Chloroplast phylogeography of Dipentodon (Dipentodontaceae) in southwest China and northern Vietnam. *Mol Ecol*. 2008;17(4):1054–65.
57. Hong W, Li J, Wu C, Hong T, Fan H, Tang J, Chen C. Research summary of cultivation and utilization of *Tripterygium Wilfordii*. *J Fujian Coll Forest*. 2007;27(1):92–6.
58. Chen C, Li P, Wang RH, Schaal BA, Fu CX. The population genetics of cultivation: domestication of a traditional Chinese medicine, *Scrophularia Ningpoensis* Hemsl.(Scrophulariaceae). *PLoS One*. 2014;9(8):e105064.
59. Wang X, Liu XQ, Ko YZ, Jin XL, Sun JH, Zhao ZY, Yuan QJ, Chiang YC, Huang LQ. Genetic diversity and phylogeography of the important medicinal herb, cultivated Huang-Lian populations, and the wild relatives *Coptis* species in China. *Front Genet*. 2020;11:708.
60. He J, Chen L, Si Y, Huang B, Ban X, Wang Y. Population structure and genetic diversity distribution in wild and cultivated populations of the traditional Chinese medicinal plant *Magnolia officinalis* subsp. *biloba* (Magnoliaceae). *Genetica*. 2009;135:233–43.
61. Yuan Q-J, Zhang ZY, Hu J, Guo LP, Shao AJ, Huang LQ. Impacts of recent cultivation on genetic diversity pattern of a medicinal plant, *Scutellaria baicalensis* (Lamiaceae). *BMC Genet*. 2010;11(1):1–13.
62. Zohary D. Unconscious selection and the evolution of domesticated plants. *Econ Bot*. 2004;58(1):5–10.
63. Zohary D, Spiegel-Roy P. Beginnings of fruit growing in the old world. *Science*. 1975;187(4174):319–27.
64. Baute GJ, Kane NC, Grassa CJ, Lai Z, Rieseberg LH. Genome scans reveal candidate domestication and improvement genes in cultivated sunflower, as well as post-domestication introgression with wild relatives. *New Phytol*. 2015;206(2):830–8.
65. Kantar MB, Nashoba AR, Anderson JE, Blackman BK, Rieseberg LH. The genetics and genomics of plant domestication. *Bioscience*. 2017;67(11):971–82.
66. Doyle JJ, Doyle JL. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem Bull*. 1987;19:11–5.
67. Sang T, Crawford DJ, Stuessy TF. Chloroplast DNA phylogeny, reticulate evolution, and biogeography of *Paeonia* (Paeoniaceae). *Am J Bot*. 1997;84(8):1120–36.
68. Taberlet P, Gielly L, Pautou G, Bouvet J. Universal primers for amplification of 3 noncoding regions of chloroplast DNA. *PI Molec Biol*. 1991;17:1105–9.
69. Novy A, Jones KC. Characterization of polymorphic microsatellites for *Tripterygium* (Celastraceae), a monospecific genus of medicinal importance. *Am J Bot*. 2011;98(10):e280–281.
70. Burland TG. DNASTAR's lasergene sequence analysis software. *Methods Mol Biol*. 2000;132:71–91.
71. Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser*. 1999;41:95–8.
72. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R. Clustal W and Clustal X version 2.0. *Bioinformatics*. 2007;23(21):2947–8.
73. Watterson GA. On the number of segregating sites in genetical models without recombination. *Theor Popul Biol*. 1975;7(2):256–76.
74. Nei M. *Molecular evolutionary genetics*. New York, NY: Columbia University; 1987.
75. Tajima F. Evolutionary relationship of DNA sequences in finite populations. *Genetics*. 1983;105(2):437–60.
76. Librado P, Rozas J. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*. 2009;25(11):1451–2.
77. Tajima F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*. 1989;123(3):585–95.
78. Fu YX, Li WH. Statistical tests of neutrality of mutations. *Genetics*. 1993;133(3):693–709.
79. Excoffier L, Smouse PE, Quattro J. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*. 1992;131(2):479–91.
80. Excoffier L, Lischer HE. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol Ecol Resour*. 2010;10(3):564–7.
81. Swofford D. PAUP\*: Phylogeny analysis using parsimony (\* and other methods), version 4. Sunderland: Sinauer Association, Inc; 2002.
82. Huelsenbeck JP, Ronquist F. MRBAYES: bayesian inference of phylogenetic trees. *Bioinformatics*. 2001;17(8):754–5.
83. Ronquist F, Huelsenbeck JP. MrBayes 3: bayesian phylogenetic inference under mixed models. *Bioinformatics*. 2003;19(12):1572–4.
84. Rambaut A, Suchard M, Xie D, Drummond A. Tracer version 1.6. Auckland: The University of Auckland; 2014. <http://beast.bio.ed.ac.uk/Tracer>.
85. Drummond AJ, Suchard MA, Xie D, Rambaut A. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol Biol Evol*. 2012;29(8):1969–73.
86. Rambaut A, Drummond A. FigTree version 1.4.0. Scotland, UK: The University of Edinburgh; 2012. <http://tree.bio.ed.ac.uk/software/figtree>.
87. Bandelt HJ, Forster P, Röhl A. Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol*. 1999;16(1):37–48.
88. Peakall R, Smouse PE. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics*. 2012;28(19):2537–9.
89. Weir BS, Cockerham CC. Estimating F-statistics for the analysis of population structure. *Evolution*. 1984;38(6):1358–70.
90. Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. *Genetics*. 2000;155(2):945–59.
91. Falush D, Stephens M, Pritchard JK. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics*. 2003;164(4):1567–87.
92. Taylor SA, Larson EL. Insights from genomes into the evolutionary importance and prevalence of hybridization in nature. *Nat Ecol Evol*. 2019;3(2):170–7.
93. Evanno G, Regnaut S, Goudet J. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol*. 2005;14(8):2611–20.
94. Earl DA, VonHoldt BM. Structure harvester: a website and program for visualizing Structure output and implementing the Evanno method. *Conserv Genet Res*. 2012;4(2):359–61.
95. Hey J. Isolation with migration models for more than two populations. *Mol Biol Evol*. 2010;27(4):905–20.
96. Cieslarová J, Hanáček P, Fialová E, Hýbl M, Smýkal P. Estimation of pea (*Pisum sativum* L.) microsatellite mutation rate based on pedigree and single-seed descent analyses. *J Appl Genet*. 2011;52:391–401.
97. Richardson AR, Stojiljkovic I. Mismatch repair and the regulation of phase variation in *Neisseria meningitidis*. *Mol Microbiol*. 2001;40(3):645–55.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.