J. Plant Biol. (2018) 61:159-168
DOI 10.1007/s12374-017-0373-8

ORIGINAL ARTICLE

Genetic Diversity, Population Genetic Structure and Protection Strategies for *Houpoëa officinalis* (Magnoliaceae), an Endangered Chinese Medical Plant

Yang Xu[1,2], Yang Zhiling[2] and Li Huogen[1,*]

[1]*Key Laboratory of Genetics & Biotechnology, Nanjing Forestry University, Ministry of Education, Nanjing 210037, Jiangsu, Chihna*
[2]*Research Institute of Subtropical Forestry, Chinese Academy of Forestry, Fuyang 311400, Zhejiang, China*

Received: September 27, 2017 / Accepted: March 11, 2018

Abstract *Houpoëa officinalis* is a traditional Chinese medical plant, which has significantly declined in the past decades because of human influence and habitat fragmentation. Twelve expressed sequence tag SSR (EST-SSR) markers developed from the EST sequence of *H. officinalis* were used to analyse the genetic diversity and structure of fourteen natural populations. The results indicated that moderate genetic diversity and high genetic differentiation existed in this plant (Ho = 0.600, Fst = 0.327). STRUCTURE and UPGMA analyses showed that *H. officinalis* populations could be divided into 3 different groups, and the west group had higher genetic diversity than the central and east groups. The historical migration rates among the groups were low and unsymmetrical, and there was no significant correlation between Nei's genetic distance and geographic distance. According to the genetic consequences, conservation strategies (in situ or ex situ, artificial pollination) should be carried out in all populations to preserve genetic diversity.

Keywords: EST-SSR, Genetic structrue, *Houpoëa officinalis*, Protection strategies

Introduction

Houpoëa officinalis, which belongs to the Magnoliaceae family (Xia et al. 2008), is an endangered medical plant in China. It is a hermaphroditic plant with obligate breeding system. There are dichogamy and herkogamy in the structure of the breeding systems to avoid self-fertilize. *H officinalis* is a entomophilous plant with beetle as its main pollination insect (Yang et al. 2012), which may be less efficient pollinators than bees (Ramsey 1988). The fruits are aggregated, with at most two red seeds per follicle. Like other Magnolia species, it may be dispersed by birds (Li and Yin 2004; Setsuko et al. 2008).

The dry bark of this plant is used to treat stomach problems, vomiting, coughing, and asthma (Committee of Pharmacopoeia 2010). According to historical records, this plant was relatively widely distributed throughout South China from 300-1500 m in hilly areas (Liu et al. 1996). *H. officinalis* population has been severely destroyed because of its high medical value. Only a few populations with a few thousand individuals of this species remain, and most subpopulations have become locally extinct in recent decades (Yang and Yang 2017). As a result, *H. officinalis* is now considered as endangered (EN) by IUCN's red list of 2011. These survival situations severely reduce the productive fitness and act as a mechanism that might reduce the gene flow between populations (Godt and Hamrick 2003). As a result, the rapid deterioration and loss of habitat caused by anthropic pressure on this species had further increased the fragmentation and caused a decline in the population.

In recent years, many molecular marker technologies have been widely used to analyse the genetic diversity or phylogeny of this species (Guo et al. 2001;Wang et al. 2001; Tong 2002; Liu et al. 2004; Yang 2007; He et al. 2009; Kong 2010; Yu 2010; Jiang 2010; Zheng 2010; Yu 2011; Weng 2014; Mai et al. 2015) such as RAPD, ISSR, ITS, SRAP and TRAP-PCR. But all these molecular markers are dominant genetic markers and can not distinguish homozygous and heterozygous easily. While the monogenetic CpDNA sequence is highly conserved and has less polymorphic site, and can not fully disclose the historical process of a group. On the other hand, since

*Corresponding author; Li Huogen
Tel : +86-25- 85428731
E-mail : hgli@njfu.edu.cn

only a part of the populations was collected, it may not have reflected the total genetic relationship among the species distributions (Kong 2010; Jiang 2010; Zheng 2010). Many artificial populations were considered to be local provenance in previous studies, while cultivation history and provenance sources of each place could not be determined, and many intermediate mutations made it difficult to exhaustively evaluate the genetic status of natural populations of *H. officinalis* (Yang and Yang 2017).

In contrast, microsatellites could be more adaptable for genetic analysis than other molecular markers due to their high abundance, reproducibility, degree of polymorphism and co-dominant nature (Sahu et al. 2012). Expressed sequence tags (ESTs) serve as the main resource for simple sequence repeats (SSRs). The computational approach to detect SSRs and develop SSR markers from EST-SSRs is preferred over the conventional methods, as it significantly reduces the time and cost (Sahu et al. 2012; Wu et al. 2014; Alisoltani et al. 2015). EST-SSR markers have been effectively used in different plant species, such as *Paris polyphylla* var. *yunnanensis* (Wang et al. 2016), *Dendrobium officinale* (Lu et al. 2013), *Vaccinium macrocarpon* (Schlautman et al. 2015), *Salvia miltiorrhiza* (Wang et al. 2011) and so on. As co-dominant molecular markers, EST-SSRs can reveal sufficient levels of variation in population genetic applications, which can be used in population genetic analyses (Ellis and Burke 2007; Levi 2011).

H. officinalis is assumed to have genetic drift resulting from human over-exploitation and habitat fragmentation (Yu et al. 2011). Hence, we developed EST-SSR markers for the species and selected twelve polymorphic markers to assess the genetic structure and diversity of *H. officinalis* in subtropical China. The study aimed to evaluate the following: (1) the genetic diversity and genetic structure in wild *H. officinalis* populations; (2) to develop genetic component of its conservation and restoration.

Results

Exact Test for HW for Each Population

Using an EM algorithm, we did find null alleles in the AHYX, HNXX, SCXH and ZJAJ population at locus HP390. After correcting with an IMA algorithm, the genetic diversity was estimated for further analysis. Likelihood ratio tests detected significant differences ($p < 0.05$) which showed a deviation from the Hardy-Weinberg equilibrium, indicating that probable inbreeding might have occurred in all of the populations. A linkage disequilibrium test for each population also showed that many combination loci deviated from the equilibrium, 2 combination loci in the SXAK population, 2 in the SCXH population, 1 in the HBES population, 4 in the HNXX population, 2 in the FJJN population, and 4 in the HNSM population. Thus, these loci were omitted from the populations presented above in the STRUCTURE analysis.

Genetic Diversity of *H. officinalis*

Genetic diversity among microsatellite loci (Table 3) and among populations (Table 4) showed that a total of 185 alleles were amplified across the twelve EST-SSR markers in 334 individuals. The number of alleles per locus ranged from 10 (HP390) to 20 (HP6984), with an average of 15.41. The number of effective alleles varied from 2.31 to 4.05, with an average of 3.10. Shannon's index varied from 0.75 to 1.53, with an average of 1.16. The observed heterozygosity varied from 0 to 0.48, with an average of 0.20; whereas the expected heterozygosity varied from 0.37 to 0.73, with an average of 0.60. Among the fourteen natural populations, Shannon's

Table 1. Information on fourteen natural populations of *H. officinalis*

Population	Location	Longitude°	Latitude°	Elevation/m	Sample size
AHYX	Yuexi, Anhui	111°21'E	30°50'N	824	4
FJJN	Jianning, Fujian	117°2'E	26°31'N	373	30
GXZY	Ziyuan, Guangxi	110°37'E	26°1'N	859	16
GZLP	Liping, Guizhou	119°8'E	26°14'N	896	30
HBES	Jianshi, Hubei	109°27'E	30°22'N	525	30
HNAH	Anhua, Hunan	111°12'E	28°22'N	1021	30
HNSM	Shimen,Hunan	110°19'E	30°1'N	1016	28
HNXX	Xinxian, Henan	114°52'E	31°37'N	715	4
JXLS	Lushan, Jiangxi	115°41'E	29°18'N	1128	30
SCXH	Xuanhan, Sichuan	108°5'E	31°29'N	1530	30
SXAK	Langao, Shanxi	108°31'E	32°12'N	798	30
YNTC	Tengchong, Yunnan	98°32'E	25°1'N	2252	24
ZJAJ	Anji Zhejiang	119°36'E	30°26' N	895	18
ZJSC	Suichang, Zhejiang	119°16'E	28°36' N	1045	30

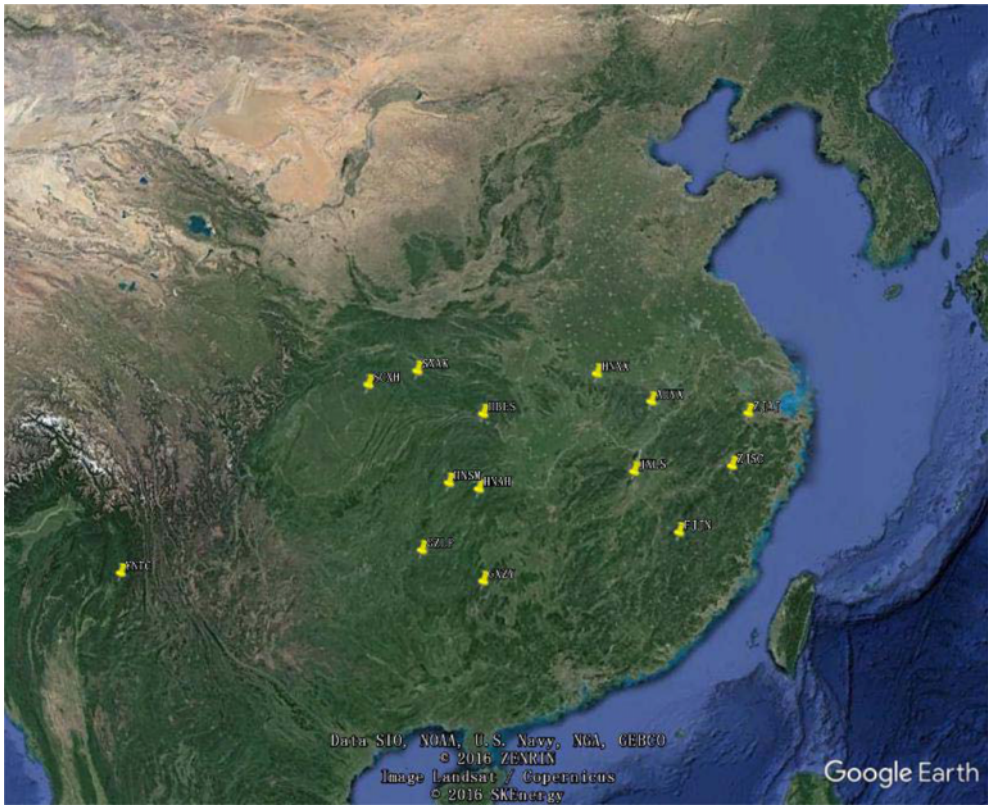

Fig. 1. The geographical locations of 14 *H.officinalis* populations sampled.

Table 2. Characteristics of the twelve EST-SSR primers analyzed in the fourteen *H.officinalis* populations included in this study

Name	GenBank accession	Forward Reverse	Repeat motif	Ta	Size range
HP386	JZ969904	CATCAACCTCCACCCTGTC CTCCCAATCATCCATCCC	(GA)8	59.7 52.7	314-386
HP387	JZ969905	CTTCTCAAGTCCGCTCCC GTCCCAGTTTCCTTCAGTCG	(CTG)5	59.6 59.8	147-197
HP390	JZ969906	CGTGGAGAAGAAAGGTGC GAAAGGGGCTGGAAAGTG	(CT)7	57.3 57.3	200-239
HP392	JZ969907	TGCCGTCAACCTCGTCGTA CCCTCATAGAAATCCCAGTC	(TGT)5	59.3 57.8	450-479
HP449	JZ969908	GCCACCACCCAATCTAAT CCATCACCCATCTTCCTT	(GA)6	55 55	203-222
HP501	JZ969909	GAACCAGAACATCACGAG TTGGGATAACTTGTAGCC	(AAT)5	57.3 57.3	140-179
HP508	JZ969910	GAATGAAAGCGAAAGAGG CCCGCAACAACGAGAACA	(GGA)7	55 55	180-233
HP630	JZ969911	CTGCCTTTCCTGGTTGAT AAATGTCTTGGTTGTGGG	(TG)8	55 52.7	159-203
HP693	JZ969912	CCGACCATCGTATCTCAT CCCTTCCATCCGCAACAC	(GA)8	55 59.6	405-453
HP696	JZ969913	TGACGATTTCCTCCTTTA GCTGAGATTGCTCCCTAT CCGCTGGTATGTGGGAAA	(CCT)6	55.4 52.7 57.3	341-361
HP699	JZ969914	GATGGAGGGCTTGAATGT TGATGGACGGCTGGGATA	(AT)6	56 56	140-199
HP6984	JZ969915	AAGGGTAGTGGATTGTGA CCGCTGGTATGTGGGAAA	(GA)6	52.7 57.3	153-200

Table 3. Genetic diversity of twelve EST-SSR markers within fourteen natural populations of *H.officinalis*

	Na	Ne	I	Ho	He	Fst	Gst
HP 386	17	3.14	1.243	0	0.632	0.471	0.476
HP 387	14	3.048	1.146	0	0.604	0.312	0.503
HP 390	10	2.455	0.75	0.133	0.371	0.336	0.640
HP 392	18	3.379	1.208	0.012	0.635	0.310	0.473
HP 449	11	2.306	0.883	0	0.51	0.432	0.604
HP 501	19	3.738	1.348	0.439	0.674	0.274	0.430
HP 508	15	3.182	1.249	0.478	0.652	0.286	0.445
HP 630	17	3.157	1.198	0.312	0.607	0.313	0.477
HP 693	13	2.901	1.147	0.044	0.627	0.285	0.444
HP 696	14	2.728	1.078	0.09	0.569	0.378	0.548
HP 699	17	3.068	1.11	0.341	0.598	0.189	0.505
HP 6984	20	4.046	1.532	0.523	0.725	0.338	0.318
Average	15.417	3.096	1.158	0.197	0.600	0.327	0.489

Na observed alleles number; *Ne* effective number of alleles; *I* Shannon's index; *Ho* observed heterozygosity; *He* average expected heterozygosity; *Fst* the total fixation index; *Gst* the total genetic differentiation coefficient

Table 4. Genetic diversity among fourteen natural population of *H.officinalis*

Pop	Na	Ne	I	Ho	He	Nei	F	Fst
AHYX	26	1.91	0.636	0.313	0.414	0.271	0.281	N.A
FJJN	56	3.014	1.176	0.129	0.593	0.577	0.834	0.244
GXZY	69	4.024	1.442	0.440	0.709	0.702	0.42	0.166
GZLP	56	3.355	1.234	0.094	0.638	0.621	0.858	0.181
HBES	62	2.98	1.209	0.234	0.618	0.557	0.666	0.186
HNAH	68	3.72	1.368	0.13	0.665	0.660	0.821	0.190
HNSM	65	3.512	1.34	0.022	0.666	0.627	0.969	0.183
HNXX	30	2.352	0.803	0.229	0.497	0.545	0.562	N.A
JXLS	61	3.286	1.224	0.333	0.615	0.575	0.544	0.191
SCXH	52	2.522	0.979	0.132	0.506	0.505	0.803	0.224
SXAK	62	3.274	1.314	0.171	0.660	0.629	0.757	0.189
YNTC	69	4.229	1.471	0.134	0.721	0.709	0.810	0.167
ZJAJ	44	2.596	0.987	0.313	0.552	0.572	0.437	0.210
ZJSC	49	2.564	1.024	0.094	0.551	0.406	0.857	0.226
Average	54.929	3.096	1.157	0.197	0.600	0.569	0.687	0.197

Na observed alleles number; *Ne* effective number of alleles; *I* Shannon's index; *Ho* observed heterozygosity; *He* average expected heterozygosity; *Nei* Nei's expected heterozygosity *F* Fix index; *Fst* the total fixation index N.A not analyzed

index values ranged from 1.47 (YNTC) to 0.64 (AHYX); Nei's genetic diversity values varied from 0.27 (AHYX) to 0.71 (YNTC); the observed heterozygosity varied from 0.02 (HNSM) to 0.44 (GXZY); and the expected heterozygosity varied from 0.41 (AHYX) to 0.72 (YNTC). The fixation index of all fourteen populations was greater than 0 and varied from 0.28 (AHYX) to 0.97 (HNSM), and most of them significantly deviated from equilibrium. Population differentiation was significant at twelve loci ($P < 0.05$, Table 3), with an average Fst value of 0.327 (range: 0.189-0.471), which indicated a certain level of genetic differentiation among the populations. The Mantel test indicated no significant correlation between Nei's genetic distance (Table 5) and the geographical distance among the populations (Fig. 2).

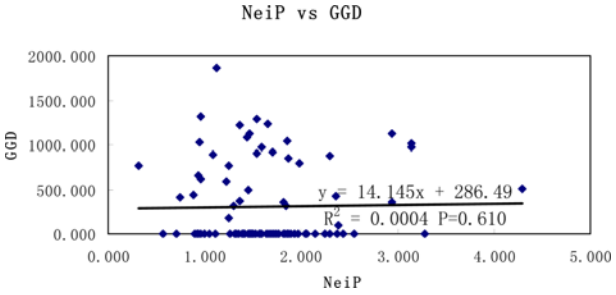

Fig. 2. Mantel test between geographic distance and genetic distance.

Genetic Structure

When using the Bayesian clustering approach, the true number

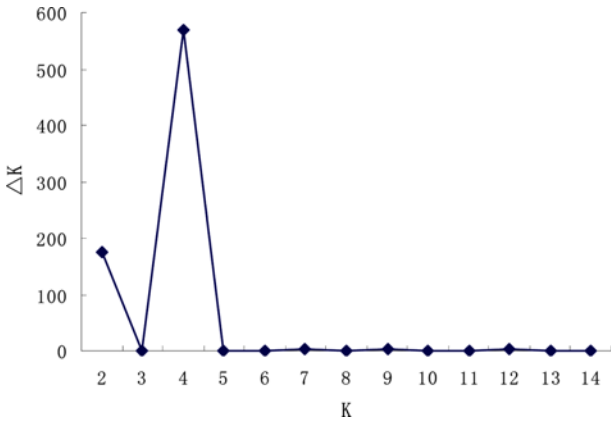

Fig. 3. Model choice criterion ΔK of the structure analysis for K value.

of gene pools K could not to determined by a straightforward approach because ln Pr(K) increased progressively with K, Moreover, the ΔK statistic allowed the detection of a rate change in ln Pr(K) corresponding to K = 4. Therefore, the individuals were classified into 4 different groups (Fig. 3). The SXAK, GZLP, YNTC, GXZY and HNSM populations were classified into a group; the AHYX, HNXX, FJJN, ZJAJ and HNAH populations were classified into a group; the HBES, JXLS and SCHX populations were classified into a group, and the ZJSC populations constitute its own group (Fig. 4). The clustering of individuals in a UPGMA dendrogram based on Nei's genetic distance was very similar to that in STRUCTURE, except that the ZJSC population was merged into a group with the AHYX, HNXX, FJJN, ZJAJ and HNAH populations, whereas the HNXX population was closer to the JXLS, HBES and SCXH population (Fig. 5). According to the genetic structure results, we classified the fourteen populations into three group: the west group containing the population of SXAK, GZLP, YNTC, GXZY and HNSM populations, the central group containing the HBES, JXLS and SCHX population, and the east group containing the AHYX, HNXX, FJJN, ZJAJ, ZJSC and HNAH population. The measures of genetic diversity in the west group were higher than those in the central and east groups (Table 6).

Historical Gene Flow between Populations

Due to the small number of individuals in the AHYX and HNXX populations, these two populations were eliminated from the statistics of gene flow. STRUCTURE analyses showed that only a few individuals were responsible for having mixed genealogies, which indicated low migration rates between the populations. MIGRATE analyses indicated that the migration rates between groups were low and unsymmetrical. The highest level of migration occurred from the west group to the central group and the lowest occurred from the east group to the west group. Unsymmetrical gene

Table 5. Pairwise Population Matrix of Nei Genetic Distance

	FJJN	GXZY	GZLP	HBES	HNAH	HNSM	JXLS	SCXH	SXAK	YNTC	ZJAJ	ZJSC
FJJN												
GXZY	1.486											
GZLP	1.983	1.764										
HBES	1.578	1.818	1.249									
HNAH	1.357	1.267	2.142	1.966								
HNSM	1.313	1.004	1.352	1.562	1.456							
JXLS	1.858	1.667	1.221	0.311	2.245	1.563						
SCXH	1.545	1.524	0.937	0.747	2.047	1.396	0.886					
SXAK	1.564	1.647	1.470	1.739	1.111	1.596	1.812	1.820				
YNTC	1.860	1.393	0.707	1.482	1.834	1.054	1.696	1.413	1.499			
ZJAJ	1.838	0.936	0.942	1.366	2.379	0.963	1.245	0.961	2.434	1.114		
ZJSC	1.814	1.722	0.956	1.656	1.875	1.581	1.444	1.364	2.295	1.927	1.125	

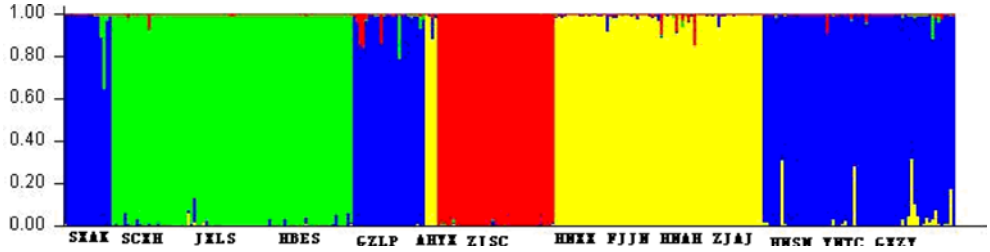

Fig. 4. Estimated population structure for 14 populations of *H. officinalis*.

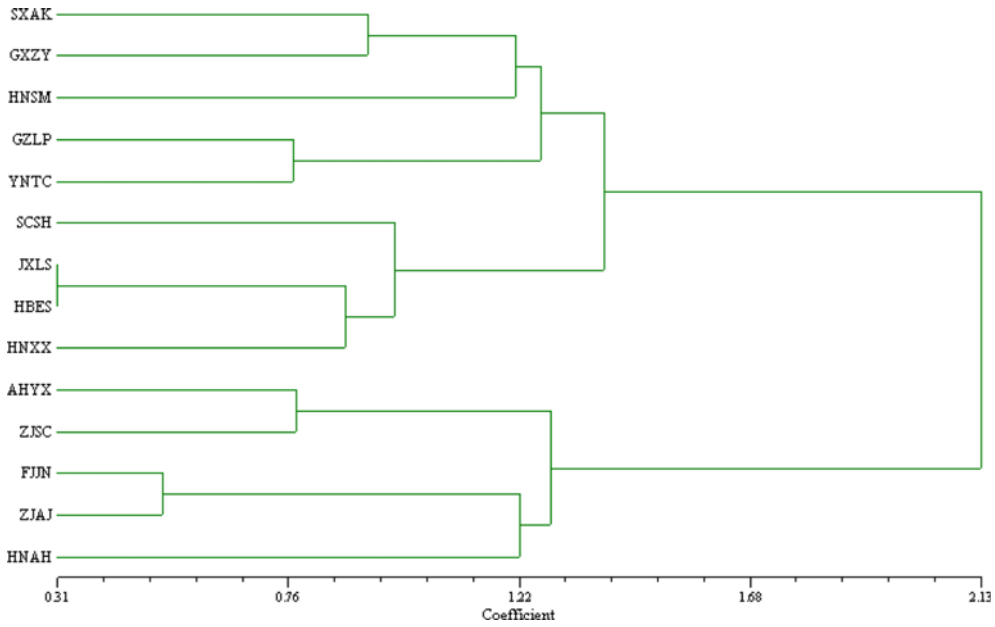

Fig. 5. Dendrogram of fourteen natural populations of *H. officinalis.*

Table 6. Difference between three groups of population genetic diversity

Group	Ao	Ne	I	Ho	He	Nei	F	FIS
Central group	155	6.886	2.117	0.274	0.837	0.537	0.685	0.67264
East group	157	7.406	2.149	0.138	0.839	0.548	0.835	0.835518
West group	159	8.067	2.240	0.167	0.871	0.627	0.808	0.808266

Ao observed alleles number; *Ne* effective number of alleles; *I* Shannon's index; *Ho* observed heterozygosity; *He* average expected heterozygosity; *Nei* Nei's expected heterozygosity *F* Fix index; *FIS* inbreeding coefficient

Table 7. Estimates of gene flow among three population groups of *H. officinalis*

Regions	Θ*	West West→i	Central Central→i	East East→i
West	0.615 (0.603-0.649)		0.356 (0.349-0.663)	0.230 (0.228-0.234)
Central	0.487 (0.468-0.511)	0.797 (0.782-0.806)		0.723 (0.719-0.727)
East	0.426 (0.405-0.456)	0.644 (0.636-0.655)	0.587 (0.583-0.592)	

*The range estimates given below each value are the 95% confidence limits.

flow occurred not only between the west and the central groups, but also between the central and the east groups (Table 7).

Discussion

H. officinalis has long been known for its high medical value, which has led to the overharvesting during the last century. Intensive logging often results in the fragmentation and reduction of the effective population size in many Magnoliaece species such as *Kmeria septentrionalis* (Zhao et al. 2010), *Manglietiastrum sinicum* (Tian et al. 2003) and *M. sargentiana* (Xie et al. 2014).This factor has a direct or indirect impact on the evolutionary processes, causing a loss of genetic diversity, changes in intra- and inter-population genetic structures and genetic drift (Soldati et al. 2013). Our study of *H. officinalis* regarding severe historical anthropogenic pressures on populations could be considered representative of the current status of this medicinal resource.

Genetic Diversity and Differentiation in *H. officinalis* Populations

As a type of widespred species, the moderate genetic diversity and high genetic differentiation observed in *H. officinalis* (He = 0.600, Fst = 0.327) might reflect the situation of this plant more accurately than other out-crossing species determined based on SSR markers (He = 0.650 Fst = 0.22, Nybom, 2004). The results might be relative to the biological characteristics of this plant and the external factors described as follows: First of all, the complex topography and climates of southern China have caused genetically initiated differentiation among *H. officinalis* geographical populations. PCR amplification and sequencing performed with two chloroplast intergenic spacers, psbA-trnH and trnL-tmF haplotypes showed that the differentiation of haplotype frequencies among *H. officinalis* populations was significant (Yu et al. 2010); Second, *H. officinalis* is a hermaphroditic plant with beetles as its pollinator (Yang et al. 2012). Thus, habitat fragmentation cause the reduction of pollen capacity and pollen quality (González-Varo et al. 2009; Rosas et al. 2011) and restricts the rate and efficacy of beetles visiting behavior (Moniem et al. 2013) which will increase the selfing rates and inbreeding depression on the population.; Third, like other Magnolia species, *H. officinalis* seeds cannot germinate spontaneously, and depend on digestion by birds. Habitat fragmention cause the extinction or decrease of the distributing animal species and reduce seed's diffusion ability (Bai et al. 2011; Jones et al. 2017), creating regeneration barriers for this plant.

Genetic Relationship of *H. officinalis*

STRUCTURE and UPGMA analyses showed that *H. officinalis* populations can be divided into 3 different groups. The boundaries were exactly on the edge of "Tanaka line" (Sino-Himalayan and Sino-Japanese subregion transitional area), which divideds the west and central groups, while the Dabie, Tianmu, Wuyi-Donggong mountains divided the east and central groups. The genetic diversity in the west group was higher than that in the central and east groups. This difference is probably because these two areas are surrounded by long-term agricultural practice (Thiago et al. 2008), and isolation might have occurred earlier than in the west region. The isolation and the relatively small population size decreased the genetic diversity. For instance, AHYX and HNXX populations maintained less than 30 individuals and possessed the lowest levels of genetic diversity among the species (He = 0.414 and 0.497 respectively), ZJSC and FJJN populations showed high genetic distance in our study (Fst = 0.244 and 0.226 respectively).

Southwest China is considered one of the most important global refuges during the Last Glacial Maximum (Huang et al. 2010; Yong et al. 2012), and is the distribution centre for many Magnoliaceae plants (Liu et al. 1995). Since continuous core populations had comparable higher levels of genetic diversity than peripheral and disjunctive populations (Gapare et al. 2005), we considered these areas to be the distribution centres of *H. officinalis.*

Historical Gene Flow of *H. officinalis*

STRUCTURE and MIGRATE analyses showed the historical migration rates among the groups were low and unsymmetrical. Meanwhile, great differentiations were observed among the *H. officinalis* groups. The leaf of the west groups is apex acute with short and upright hairs; the leaf of the central groups is shortly acuminate or obtuse, with long and upright hairs; the apex concave leaf of east groups is the most differentiated one with long, bent or twisted hairs (Si et al. 2000; Yang et al. 2016). These may suggest limitations of gene flow, not only because of its reproductive characteristics (Yang et al. 2012), but also because the mountains existed between groups form the natural barriers, often have important effects on intraspecific genetic structure through restraining gene flow and enhancing differentiation among populations (Yang et al. 2015). However, reduction of effective population size and reproductive fitness (Yang et al. 2012) will diminish contemporary gene flow significantly (Hoebee et al. 2007; Nagamitsu et al. 2014).

Conservation Implications

Genetic diversity is critically important for a species to maintain it's evolution in order to cope with an ever-changing environment (Hsu et al. 2015). The information obtained in this study should be helpful to provide a clear framework for developing conservation programmes for the endangered species *H. officinalis.* Since excessive exploitation is the most significant cause for the endangerment of *H. officinalis* populations, cutting adult trees and digging seedlings must be prohibited to maintain the effective population size. Moreover, anthropogenic habitat destruction should be prevented to encourage seedling establishment and the recruitment of new individuals, which will increase the population size through natural regeneration.

Compared to other populations, protection is urgently needed for populations such as the AHYX and HNXX populations because of their small population size and relatively low genetic diversity. Any loss of individuals may result in decreased potential for the species to adapt to environmental changes in the long term (Tremblay and Ackerman 2001). On the other hand, widespread species always consist of geographically differentiated populations, which contain many excellent properties and specific genes

from isolation and local adaptation. The populations used for the establishment of ex situ gene conservation should be from different geographical areas to maximize genetic diversity.

Finally, appropriate strategies to increase gene flow are required for *H. officinalis* in order to maintain the largest amount of genetic resources and promote long-term survival. Promoting gene flow among populations may be necessary for small isolated populations. However, since a mixture of individuals from genetically distinct populations may result in outbreeding depression (Hufford and Mazer 2003; Edmands and Timmerman 2003), hybridization between regions cannot be currently recommended. Artificial pollination with caution were proven to be efficient in enhancing gene exchange among subpopulations and lead to increased viability of *H. officinalis* subpopulations in the study area over the long term (Mai et al. 2015).

Materials and Methods

Plant Materials

First of all, a comprehensive literature textual research on *H officinalis* was carried out. We consulted more than 1000 specimens in main Chinese herbariums including EP, HKSA, NAS, HZU, ZFJC etc, then assemble the geographic distribution and location information of this species synthetically. Through extensive field investigation on full distribution area and carefully interview with forestry bureau workers and the farmers, we selected fourteen natural populations which can be affirmed as wild provenance (Table 1 and Fig. 1). We collected leaf samples from 334 adult individuals from the fourteen natural populations. The samples were collected from 4 to 30 individuals (mean = 21.14). Detailed information on these populations and the geographical distribution were recorded with GPS tools (Table 1).

Total DNA Extraction and SSR Amplification

Since a few EST sequences of *H. officinalis* were obtained from NCBI, which suggested that only a few molecular studies had been carried out on this endemic species. We used the EST library which was developed from *H offcinalis* transcriptome sequencing by the research institute of subtropical forestry, Chinese academy of forestry. A total of 180 SSR markers were randomly selected for validation, and twelve polymorphic markers (seven di-nucleotide repeats and five tri-nucleotide repeats) with the GenBank accession JZ969904-JZ969915 were selected to estimate the genetic diversity and structure of the *H. officinalis* populations (Table 2).

Total genomic DNA was extracted from young leaves using a DN0702 EASYspin Plus kit (Aillab). DNA concentrations were evaluated using a NanoDrop™ 2000 spectrophotometer (Thermo Scientific, Wilmington, DE, USA). All genomic DNA samples were amplified and analysed with the twelve EST-SSR markers. PCR amplification of the EST-SSR markers was performed in 20-uL reactions containing 10 μL of 2×Taq PCR MasterMix, 10 pmol of each primer, 20 ng of template DNA and 7.2 μL of deionized water using the following settings: 1 cycle at 94°C for 5 min and 94°C for 15 s, 1 cycle of 61-52°C for 15 s, 72°C for 1 min, and 94°C for 15 s, and 23 cycles of 52°C for 15 s, followed by 72°C for 10 min. The PCR products were analysed on a Qsep100 Genetic Analyzer with GeneScan software (Applied Biosystems, Mendota Heights, MN, USA).

Data Analysis

The co-dominant data were obtained and the frequency of null alleles was estimated by a maximum-likelihood estimator using an EM algorithm (Chapuis and Estoup 2007). The Hardy-Weinberg equilibrium was estimated from a Markov chain (Yuan and Bonney 2003) and the linkage disequilibrium was estimated using a maximum-likelihood-binomial (MLB) method (Cobat et al. 2011). The parameters were as follows: 10000 Dememorization, 100 batches, and 5000 Iterations per batch.

Genetic diversity was assessed by calculating the observed number of alleles (Na), the mean effective number of alleles per locus (Ne), the effective number of alleles (Ne), the observed heterozygosity (Ho), the average expected heterozygosity (He), Nei's genetic diversity (Nei), Shannon's information index (I) for each microsatellite locus and each population using GenAlEx 6.501 software (Peakall and Smouse 2012). Mantel test for matrix correlation between Nei's genetic distance and geographic distance for fourteen H. officinalis populations using GenAlEx 6.501 software (Peakall and Smouse 2012). Genetic differentiation among populations was assessed using θ (Fst) (Weir and Cokerham 1984) and the standardized genetic differentiation Gst (Hedrick 2005).

UPGMA cluster trees were constructed to define the population relationships based on Nei's pairwise genetic distance and the actual genetic distances among the population with MEGA5.0. The Landscape Shape interpolation procedure and Monmonier's Maximum Difference Algorithm analysis were used to visualize the genetic distance and geographical distance patterns among the populations, with 10000 random permutations (Rufai et al 2013).

The population structure of the fourteen populations was determined with a Bayesian Clustering analysis in STRUCTURE v2.3.4 (Pritchard et al. 2000). Ten independent simulations were run for each K(1-19), with 50000 burn-in steps followed by 200000 MCMC steps with the LOCPRIOR model. We adopted two criteria with the ln Pr(K) and delta K(ΔK) values to indentify the suitable values of K for the population structure (Evanno et al. 2005). The number of clusters (populations) was estimated using the delta K method developed by Evanno et al. (2005), and calculated using STRUCTURE HARVESTER 0.6.8 to visualize the STRUCTURE output.

To estimate the historical gene flow between the populations, MIGRATE version 3.6.11 (Beerli 2006) was used to calculate maximum-likelihood (ML) estimates of the effective number of migrants between pairs of populations under a Bayesian coalescent framework. We relied on the maximum likelihood estimation and used ten short chains (1000 trees) and three long chains (100000 trees) with 10000 trees discarded as initial "burn-in", replicates = YES, longChains, randomtree = YES, heating = ADAPTIVE:1(1.0,1.2,1.5,3.0) (Beerli and Felsenstein 1999).

Acknowledgements

The research has been financially supported by the National Natural Science Foundation of China (31270585).

Author's Contributions

YX conducted the experiment, analyzed the data and wrote the manuscript; LH design the study, control the quality and proposal for the manuscript; YZ is the person who gain the foundation. All authors

read and approved the manuscript.

References

Alisoltani A, Ebrahimi S, Azarian S, Hematyar M, Shiran B, Jahanbazi H (2015) Parallel consideration of SSRs and differentially expressed genes under abiotic stress for targeted development of functional markers in almond and related Prunus species. Scientia Hortic 198:462–472

Bai B, Li N, Lu CH (2011) Influence of habitat fragmentation on frugivores and their seed dispersal. Chin J Ecol 30:2613–2620

Beerli P (2006) Comparison of Bayesian and maximum-likelihood inference of population genetic parameters. Bioinformatics 22: 341–345

Beerli P, Felsenstein J (1999) Maximum-likelihood estimation of migration rates and effective population numbers in two populations using a coalescent approach. Genetics 152:763–773

Chapuis MP, Estoup A (2007) Microsatellite null alleles and estimation of population differentiation. Mol Biol Evol 24:621–631

Cobat A, Abel L, Alcaïs A (2011) The Maximum-Likelihood-Binomial method revisited: a robust approach for model-free linkage analysis of quantitative traits in large sibships. Genetic Epidemiology 35:46–56

Committee of Pharmacopoeia (2010) Chinese Pharmacopoeia. Beijing: chemical industry publishers. 275

Edmands S, Timmerman CC (2003) Modeling Factors Affecting the Severity of Outbreeding Depression. Conserv Biol 17:883–892

Ellis JR, Burke JM (2007) EST-SSRs as a resource for population genetic analyses. Heredity 99:125–132

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 14:2611–2620

Gapare WJ, Aitken SN, Ritland CE (2005) Genetic diversity of core and peripheral Sitka spruce (*Picea sitchensis*, (Bong.) Carr) populations: implications for conservation of widespread species. Biol Conserv 123:113–123

Godt MJW, Hamrick JL (2003) Consequences of landscape patterns on the genetic composition of remnant hardwood stands in the Southeast: A pilot study. Off Sci & Tech InformTech Rep 2–48

González-Varo JP, Arroyo J, Aparicio A (2009) Effects of fragmentation on pollinator assemblage, pollen limitation and seed production of Mediterranean myrtle (Myrtus communis). Biol Conserv 142:1058–1065

Guo BL, Wu M, Si JP, Li JS, Xiao PG (2001) Research on DNA molecular marker of *Magnolia officinalis* Rehd. et Wils----RAPD study on certified species. Acta Pharmacol Sin 36:386–389

Hedrick PW (2005) A standardized genetic differentiation measure. Evolution 59:1633–1638

He JS, Ling C, Ying S, Bo H, Ban XQ, Wang YW (2009) Population structure and genetic diversity distribution in wild and cultivated populations of the traditional Chinese medicinal plant *Magnolia officinalis* subsp. *biloba* (Magnoliaceae) Genetica 135:233

Hoebee SE, Arnold U, Düggelin C, Brodbeck S, Rotach P, Holderegger R, Hoebee SE, Amold U (2007) Mating patterns and contemporary gene flow by pollen in a large continuous and a small isolated population of the scattered forest tree *Sorbus torminalis*. Heredity 99:47–55

Hsu TH, Takata K, Onozato H et al (2015) Causes of the drastic loss of genetic variation in the Critically Endangered Formosa landlocked salmon of Taiwan. Endange Species Res 27:277–287

Huang Z, Liu N, Liang W, Zhang Y, Liao X, Ruan L (2010). Phylogeography of Chinese bamboo partridge, *Bambusicola thoracica* thoracica, (Aves: Galliformes) in south China: Inference from mitochondrial DNA control-region sequences. Mol Phylogenet Evol 56:273–280

Hufford KM, Mazer SJ (2003) Plant ecotypes: genetic differentiation in the age of ecological restoration. Trends Ecol Evol 18:147–155

Jiang YF (2010) Study on variation of genetic diversity of *Magnolia officinalis* at different levels. MsD Hangzhou: Zhejiang Agriculture and Forest University

Jones LR, Dukesylvester SM, Leberg PL, Johnso DM (2017) Closing the gaps for animal seed dispersal: Separating the effects of habitat loss on dispersal distances and seed aggregation. Ecol Evol 7:5410

Kong L (2010) Study on variety selection of *Magnolia officinalis* and its identification of provenance of RAPD. MsD. Fuzhou: Fujian Agriculture and forestr University

Levi A (2011) EST-SSRs of watermelon (Citrullus sp.) Useful in assessing genetic diversity among *Lagenaria Siceraria* accessions. Int J Wildland Fire 20:776–791

Li Y, Yan HF, Ge XJ (2012) Phylogeographic analysis and environmental niche modeling of widespread shrub *Rhododendron simsii* in China reveals multiple glacial refugia during the last glacial maximum. Acta Phytotaxon Sin 50:362–373

Li X, Yin X (2004) Seed dispersal by birds in Nanjing Botanical Garden Mem. Sun Yat.-Sen in spring and summer. Act Ecol Sin 24:1452–1458

Liu YH, Xia NH, Yang HQ (1995). The origin, evolution, and phytogeography of Magnoliaceae. J Tropic & Subtropic Bot 34:1–12

Liu YH, Luo XR, Wu RF (1996) Flora of China. Beijing: Science Press 30:221–222

Liu WS, Zhu JM, He B, Su Y (2004) Studies on random amplified polymorphic DNA fingerprinting of cortex *Magnoliae officinalis*. J Tradit Chin Med 27:164–169

Lu JJ, Gao L, Kang JY, Feng SG, He RF, Wang HZ (2013) Thirteen novel polymorphic microsatellite markers for endangered Chinese endemic herb *Dendrobium officinale*. Conservation Genetics Resources 5:359–361

Mai J, Yang ZL, Yang X, Wang L (2015) Analysis and genetic diversity of population of parents and their progenies of *Magnolia officinalis* based on SSR markers. J Plant Resour &Envir 24:10–17

Nagamitsu T, Kikuchi S, Hotta M, Kenta T, Hiura T (2014) Effects of Population Size, Forest Fragmentation, and Urbanization on Seed Production and Gene Flow in an Endangered Maple (*Acer miyabei*). Am Midl Nat 172:303–316

Moniem HEMA, Holland JD (2013) Habitat connectivity for pollinator beetles using surface metrics. Landscape Ecol 28:1251–1267

Nybom H (2004) Comparison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants. Mol Ecol 13: 1143–1155

Peakall R, Smouse PE (2012) GenAlEx 6.5. Bioinformatics 28:2537–2539

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–959

Rosas F, Quesada M, Lobo JA, Sork VL (2011) Effects of habitat fragmentation on pollen flow and genetic diversity of the endangered tropical tree *Swietenia humilis* (meliaceae). Biol Conserv 144:3082–3088

Rufai S, Hanafi MM, Rafii MY, Ahmad S, Arolu IW, Ferdous J (2013) Genetic dissection of new genotypes of drumstick tree (*Moringa oleifera* Lam.) using random amplified polymorphic DNA marker. Biomed Res Int 2013:604598

Sahu J, Sen P, Choudhury MD, Modi MK, Talukdar AD (2012) Towards an efficient computational mining approach to identify EST-SSR markers. Bioinformation 8:201–202

Setsuko S, Nishimura N, Tomaru N (2008) Analysis of seed dispersal pattern of Magnolia stellata by using microsatellite markers.

Japanese forestry society congress
Schlautman B, Fajardo D, Bougie T, Wiesman E, Polashock J, Vorsa N (2015) Development and validation of 697 novel polymorphic genomic and EST-SSR markers in the American cranberry (*Vaccinium macrocarpon* Ait.). Molecules 20:2001–2013
Si JP, Tong ZK, Guo BL (2000) Discussion on current classification of *Magnolia officinalis*. J Pharm Practi 18:301–302
Soldati MC, Fornes L, Zonneveld MV, Thomas E, Zelener N (2013) An assessment of the genetic diversity of *Cedrela balansae*, C. DC. (Meliaceae) in Northwestern Argentina by means of combined use of SSR and AFLP molecular markers. Biochem Syst Ecol 47:45–55
Tian K, Zhang G, Cheng X, He S, Yang Y (2003) The habitat fragility of *Manglietiastrum sinicum*. Acta Bot Yunn 2003
Thiago A, Maristerrar L, James G, Rogério G (2008) Post-logging loss of genetic diversity in a mahogany (*Swietenia macrophylla* King, Meliaceae) population in Brazilian Amazonia. Forest Ecol Manag 255:340–345
Tong ZK (2002) Studies on genetic improvement of medical plant *Magnoliae officinalis*. PhD. Nanjing: Nanjing Forestry University
Tremblay RL, Ackerman JD (2001) Gene flow and effective population size in Lepanthes (Orchidaceae): a case for genetic drift. Biol J Linn Soc 72:47–62
Wang L, Yang Y, Zhao Y, Yang S, Udikeri S, Liu T (2016) De novo characterization of the root transcriptome and development of EST-SSR markers in *Paris polyphylla* Smith var. yunnanensis, an endangered medical plant. J Agric Sci Tech-Iran 18:437–452
Wang X, Zhou X, Gao W, Cui G, Huang L, Liu C (2011). New analysis of EST-SSR distribution and development of EST-SSR markers in *Salvia miltiorrhiza*. China J Chin mater med 36:289–93
Wang T, Sun YJ, Zhu J et al (2001) Studies on DNA amplification fingerprinting of cortex *Magnoliae officinalis*. J Tradit Chin Med 24:710–715
Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. Evolution 38:1358–1370
Weng J (2014) TRAP analysis and seedling preliminary provenance selection on different provenances of *Magnolia officinalis* subsp. Biloba. MsD. Fuzhou: Fujian Agriculture and forestry university
Wu HB, Gong H, Liu P, He XL, Luo SB, Zheng XM (2014) Large-scale development of EST-SSR markers in sponge gourd via transcriptome sequencing. Mol Breeding 34:1903–1915
Xia N H, Liu Y H, Nooteboom HP (2008) Magnoliaceae. Flora of China. Beijing:Science Press 7:64
Xiong M, Tian S, Zhang Z, Fan D, Zhang Z (2014) Population genetic structure and conservation units of *Sinomanglietia glauca* (Magnoliaceae). Biodiver Sci 22:476
Yang HB (2007) Quality of Cortex *Magnolia officinalis* in Hubei Enshi. MsD. Wuhang: Hubei university of Chinese medicine
Yang L, Liu ZL, Li J, Dyer RJ (2015) Genetic structure of *Pinus henryi*, and *Pinus tabuliformis* : Natural landscapes as significant barriers to gene flow among populations. Biochem Syst Ecol 61:124–132
Yang X, Yang ZL, Wang J, Tan GY, He ZS (2012) Floral syndrome and breeding system of endangered species *Magnolia officinalis* subsp. Biloba. Chin J Ecol 31:551–556
Yang X, Yang ZL, Wang J, Tan GY (2012) Variation of fruit and seed characters of endangered species *Magnolia officinalis* subsp. Biloba. Chin J Ecol 31:1077–1081
Yang X, Yang Z L, Mai J, Pan WT, Li HG (2016) Comparison of characters of leaf trichomes in *Houpoëa officinalis* and their taxonomical significances, Guihaia 36:1335–1343
Yang ZL, Yang X (2017) Conservation biology of *Houpoëa officinalis*. Beijing: Science Press 36–37
Yu SX, Yuan QJ, Yang B, Zhang C, Chen C (2010) Population genetic study of *Magnolia officinalis* and *Magnolia officinalis* var biloba. Zhongguo Zhong Yao Za Zhi 35:2129–2132
Yuan A, Bonney GE (2003) Exact test of Hardy-Weinberg equilibrium by Markov chain Monte Carlo. Math Med Biol 20:327-340
Yu HH, Yang ZL, Sun B, Liu RN, Yang X (2011) Genetic diversity and relationship of the endangered plant *Magnolia officinalis* (Magnoliaceae) assessed with ISSR polymorphisms. Biochem Syst Ecol 39:71–78
Zhao JL, He TH, Li QM (2010) Contrasting coarse and fine scale genetic structure among isolated relic populations of *Kmeria septentrionalis*. Genetica 138:939–944
Zheng ZL (2010). Studies on genetic diversity and construction of fingerprinting of *Magnolia officinalis*. MsD. Fuzhou:Fujian Agriculture and forestry university