Genetic diversity and structure among natural populations of *Mytilaria laosensis* (Hamamelidaceae) revealed by microsatellite markers

Niu Yu[1], Jie Yuan[2], Guangtian Yin[1], Jinchang Yang[1*], Rongsheng Li[1], Wentao Zou[1]

[1] Key Laboratory of State Forestry Administration on Tropical Forestry Research, Research Institute of Tropical Forestry, Chinese Academy of Forestry, Guangzhou 510520, China
[2] Forestry Investigation and Design Team, Loudi, 417000, China

* Corresponding author: Jin-chang Yang. E-mail: fjyjc@126.com.

Abstract

Mytilaria laosensis is a fast-growing tropical broadleaf tree that is extensively used for wood production and has significant ecological benefits. To investigate the genetic diversity and population structure of *M. laosensis*, eight major natural populations in China were analyzed by using simple sequence repeat (SSR) markers. A total of 88 microsatellite-containing fragments were obtained by the method of magnetic bead enrichment, among which 26 pairs of SSR primers were screened out and used to generate a total of 190 alleles among 152 individuals. The average of observed number of alleles, Shannon's information index and polymorphism information content per locus were 18.3, 1.1577 and 0.7759, respectively, implying a high level of genetic diversity in *M. laosensis* populations. The variation within populations accounted for 81.74 % of total variation based on analysis of molecular variance. Cluster analysis divided the eight populations into four groups, among which five populations from the southern parts of Guangxi province were classified as one major group. Mantel test showed that there was highly significant correlation between Euclidean genetic distance and geographic distance, suggesting that geographic isolation contribute to the high genetic diversity of *M. laosensis*. Together, these could provide support for the feasibility of exploration and utilization of *M. laosensis* in subtropical areas of East Asia including Jiangxi, Hunan and Fujian province of China.

Keywords: : Mytilaria laosensis; SSR; Genetic diversity; Population structure

Introduction

Mytilaria laosensis Lecomte (Hamamelidaceae) is an important indigenous evergreen broad-leaf tree and has fast growth rate and ecological benefits. It is naturally distributed in southern China including Yunnan, Guangxi and Guangdong province, and northern Vietnam and Laos. *M. laosensis* has the advantages of fast growth and wood formation, wide adaptability, strong resistance, easy plantation and high survival rate and germination rate, and hence is beneficial for both economic and ecological development in many aspects including water conservation, soil amendment, fire prevention and environment protection (Guo et al. 2006; Huang et al. 2014). Its straight trunk and high quality of wood can be used for construction, furniture manufacture and paper-making. Considering the rapid growth rate and high quality of timber, *M. laosensis* has been widely used in the establishment of ecological benefit-oriented forests as well as fast-growing and high-yield plantations and is expected to become a major afforestation species in subtropical China and beyond (Huang et al. 2009). Furthermore, it is beneficial for the growth of the other species and improvement of soil quality when mixed afforestation with *Eucalyptus robusta, Fokienia hodginsii,* and *Choerospondias axillaris*. It is estimated that there are at least 2500 hectares of manmade forest of *M. laosensis* in Guangdong and Guangxi province of China (Ming et al. 2012; Ming et al. 2016). However, data on the germplasm evaluation and genetic test of *M. laosensis* used for plantation are still lacking. Due to the insufficient basic research, elite selection and germplasm enhancement of *M. laosensis* are lagging behind, which further limits the efficiency of resource cultivation and application. Furthermore, due to unreasonable exploitation of this species, the

DOI:10.2478/sg-2018-0013
edited by the Thünen Institute of Forest Genetics

natural stands of *M. laosensis* suffered serious destruction and large areas of natural stands are very scarce nowadays.

As a first step towards germplasm resources maintenance, elite selection and breeding, it is both necessary and urgent to investigate the genetic diversity of natural populations of *M. laosensis.* Molecular markers, which reveal the variation at the DNA level, are not affected by environmental factors or plant growth stage and therefore are very reliable and useful for analysis of genetic diversity. To date, simple sequence repeat (SSR) marker, which has been proven to be highly polymorphic, co-dominant, easily reproducible and highly reliable, is generally recognized as an ideal marker for breeding application (Olango et al. 2015). Until now, most of the research on *M. laosensis* has focused on introduction and cultivation (Chen et al. 2012). Information about genetic diversity and differentiation of *M. laosensis* is quite limited (Peng et al. 2012). In this paper, we investigated the genetic diversity and genetic structure among natural populations of *M. laosensis* using SSR markers. The results obtained in this study have implications for collection and maintenance of genetically diverse populations and implementation of further breeding programs.

Materials and Methods

Plant material

The collected eight populations represent the main natural distribution areas of *M. laosensis* in China. Location and sample size of populations were given in Figure 1 and Table 1. The distance between each sampling plant was at least 50 m and 3-5 young leaves from each plant were harvested, desiccated with silica gel and frozen at -20°C for further DNA extraction.

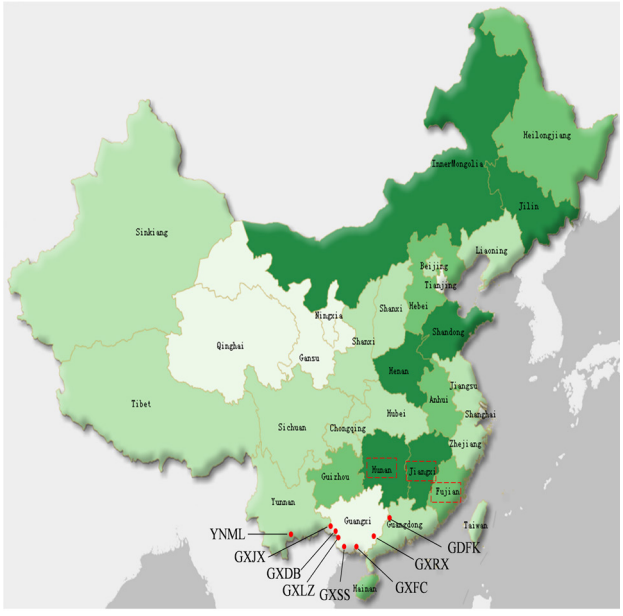

Figure 1
Sample collection sites of *M. laosensis.*

Table 1
Locations and sampling number of eight natural populations of *M. laosensis*

Code	Location	Province	Size	Longitude (E)	Latitude (N)	Altitude (m)
GDFK	Fengkai	Guangdong	6	111°54′	23°27′	442
GXRX	Rongxian	Guangxi	35	110°49′	22°41′	541
GXSS	Shangsi	Guangxi	36	107°54′	21°53′	412
GXFC	Fangcheng	Guangxi	8	107°33′	21°52′	483
GXLZ	Longzhou	Guangxi	17	106°43′	22°16′	526
GXDB	Debao	Guangxi	14	106°40′	23°03′	743
GXJX	Jingxi	Guangxi	27	106°36′	23°03′	761
YNML	Mali	Yunan	9	104°45′	22°54′	1120

Genomic DNA extraction and quantification

The genomic DNA was extracted using the Plant Genomic DNA extraction kit (Tiangen, Beijing, China) according to the manufacturer's instructions and then tested on 1.5 % agarose gel electrophoresis. The DNA concentration was quantified by spectrophotometer.

Development and screening of SSR markers

The SSR markers were developed according to Zane et al. (2002) with some modifications. The genomic DNA was digested with *Mse*I and ligated to *Mse*I adaptor (Forward: 5′-GACGATGAGTCCTGGC-3′; Reverse: 5′-TACTCAGGACCGCC-3′). The ligated DNA was pre-amplified with M_{00} primer (5′- GATGAGTCCTGAGCGG-3′) and purified with DNA agarose gel electrophoresis. The purified PCR products were hybridized with biotinylated probes $(AG)_{15}$ and then captured by streptavidin-coated paramagnetic beads and collected with a magnetic particle-collecting unit (Promega). The enriched DNA was amplified and purified again and ligated into pMD18-T vector (Takara) and cloned in DH5α competent cells. A total of 229 colonies were selected for PCR amplification with primers RV-M (5′-GAGCGGATAACAATTTCACACAGG-3′) and $(AG)_{9}$. Positive clones that produced single bands with the length of 200-1000 bp were subject to sequencing by Sangon Biotech (Shanghai Company, Ltd). These identified sequences were submitted to NCBI Genbank and the accession numbers are given in Table 2. SSR loci were screened with Chromas, and primers were designed with Primer5.0 (Clarke and Gorley, 2001).

A total of eight DNA samples were used as template and Polymerase Chain Reaction (PCR) amplification reactions were conducted with designed primers for preliminary screening. The amplified products were checked with 2.0 % agarose gel electrophoresis for specific bands and then double-checked with de-natured polyacrylamide gelelectrophoresis (Bassam et al. 1991).

SSR analysis

Based on the preliminary screening results, primer combinations which gave clearly distinguishable and consistent bands were used for SSR analysis (Li and Gan, 2011). The PCR program was as follows: pre-denaturation at 94°C for 4 min; 20 cycles each set at 94°C for 30 s, Tm+10°C for 30 s (the annealing temperature was lowered by 0.5°C after each cycle during the 20

Table 2
Details of 26 SSR primers of *M. laosensis*.

Name	Repeat motif	Primer sequence (5'-3')	Ta (°C)	Size (bp)	Accession no
MLP3	$(GA)_{13}$	F: GGACTCACCTCCACTGAGC R: GTTTGAACCCTTCCCTTTC	55	144	KJ796804
MLP6	$(GA)_{14}$	F: CGCCATTTAGATTAGACAGG R: GTCCCAGCATTCTCAACCAC	56	212	KJ796806
MLP7	$(GA)_{14}$	F: CAAACACATAGCAAACATCC R: ATACACCAGTCCTACCACCA	56	125	KJ796807
MLP8	$(CT)_{21}$	F: GCTATGGTGTGAACAAGGA R: AGCAAAAACCCAGAGTAAT	56	216	KJ796808
MLP12	$(CT)_{10}$	F: TCTTTCTTCACTCACTCTGCAT R: CTCTATCAACACCCACCACA	58	182	KJ796811
MLP16	$(CT)_{17}$	F: GTTAGTTGCAGCATAGAATG R: ATCTACTTTTCCCTATCAGC	56	282	KJ796813
MLP22	$(TC)_{20}$	F: CAGGACGGAATAGTTAGGAC R: CCAGGTGAAAGATGGGCAAA	56	218	KJ796816
MLP23	$(CT)_{12}$	F: ACCTGGTTTTGCGTGTATCA R: CACTGGCCTTTCTTCTTGGA	56	294	KJ796817
MLP25	$(CT)_{25}$	F: GTGACTTAGGGGCAAAACT R: CCTCCAAAATCCAAGAATA	56	251	KJ796819
MLP27	$(TC)_{19}$	F: TAGTGGCAAGCGGAAGAATC R: CAGCCACAACCTGGGAAAAT	56	292	KJ796821
MLP30	$(CT)_{15}$	F: AAGTTTGTTTGGGTGAGTGAT R: ACCAAGAAGACACCTTTAGCA	58	225	KJ796822
MLP31	$(TC)_3TT(CT)_{14}$	F: AATCAAAGGTCACTGTCGTGC R: AAAGGGGTTACATGGAAAATC	56	210	KJ796823
MLP32	$(GA)_{16}$	F: CATCATCTTCGCTTGATTTG R: TTCGTCTATGGCATTCTTGT	56	251	KJ796824
MLP33	$(TC)_{18}$	F: AGTCCGTATTGCTTGTTCATT R: CACTGGCTGTTTGCTCTTCTA	56	270	KJ796825
MLP34	$(CT)_{20}$	F: ACTCCGATCCTTTCCTTTCA R: GTCCGACTTGTTATGTTGTTTCC	56	217	KJ796826
MLP35	$(CT)_{17}$	F: TTCACTGCCTTCCCTGTAT R: GCAGGAGAAGAAGCGACCA	55	129	KJ796827
MLP36	$(TC)_{12}$	F: TTCTTTTGCTCCGCCTTCTA R: GTCTGCATGACTGCATGTGC	55	123	KJ796828
MLP37	$(TC)_{21}$	F: CTGATAATGGCGTTCTTACT R: TCTTCTGCTGCTCAATACAA	53	213	KJ796829
MLP39	$(CT)_{14}$	F: AGGCGTATCGGACACTTCG R: CAAACCCTAGAGCAAATGAAAA	55	243	KJ796830
MLP40	$(TC)_{10}$	F: AGTAGTGTTGGGGAGTAGGA R: GAAAGGTTATCGGGTGTTAG	56	297	KJ796831
MLP41	$(TC)_{16}$	F: CTCCTCCTTCTATGTATGC R: CAAAAGGGAATGAAGTAAC	55	188	KJ796832
MLP43	$(CT)_{17}$	F: CTTCTTTCCATCCTCATCG R: TATAAAACTACCTCGCAGA	55	167	KJ796834
MLP46	$(CT)_{19}$	F: ATTTTGGATACAATGACTCG R: TAATGAGAAACAATCAACCC	55	226	KJ796835
MLP48	$(CT)_{11}$	F: GAAATCAGTGAAGAACATAC R: GCTTCTCTACCATAAATCTA	51	173	KJ796836
MLP49	$(CT)_9CA(CT)_{12}$	F: CGCTTCAGAACTCGCTTTCAC R: GACGGTCGGTGCCGATTGATA	60	276	KJ796837
MLP50	$(CT)_{12}$	F: TTGTGTGGGTCATTGCACCTT R: TTCAAAATCATCATCCCGCTG	55	120	KJ796838

Ta, annealing temperature.

cycles); 30 cycles each set at 94°C for 30 s, Tm for 30 s and 72°C for 1 min; and final extension at 72°C for 10 min. Finally, 2 µL of PCR products were mixed with 7.82 µL of formamide and 0.18 µL of GneScan 500 LIZ size standard and analyzed on ABI 3730xl DNA analyser (Applied Biosystems). Data scoring and collection were carried out by GeneMapper 4.1.

Data analysis

Polymorphism information content (PIC) and Euclidean genetic distance were calculated using PowerMarker v3.25 (Liu and Muse, 2005). The genetic distances were subjected to cluster analysis by Unweighted pair group method with arithmetic average (UPGMA) method using MEGA v4 (Tamura et al. 2007). Bayesian clustering method implemented in STRUCTURE v2.3.4 (Pritchard et al. 2000) was used to investigated population genetic structure. The *K* value was set between 2 and 7. The most likely numbers of clusters (*K*) were estimated using STRUCTURE HARVESTER (Evanno et al. 2005). The clustering result was displayed by the program distruct 1.1 (Rosenberg 2004). GenAlEx v6 (Peakall and Smouse, 2006) was utilized for calculations of genetic diversity parameters including the observed number of alleles (N_a), observed heterozygosity (H_o), expected heterozygosity (H_e) and Shannon's information index (I); the observed number of alleles per population (N_a) and the effective number of alleles (N_e); genetic differentiation parameters among populations including genetic differentiation coefficient (F_{st}), gene flow (N_m), inbreeding coefficient (F_{is}) and analysis of molecular variance (AMOVA) (Nei, 1978; Excoffier et al. 2005). TFPGA was used for Mantel correlation analysis (Sokal, 1979; Miller, 1999).

Results and discussion

Development and evaluation of SSR markers

Fast isolation by AFLP of sequences containing repeats (FIASCO) method was employed to develop SSR primers for *M. laosensis*. A total of 95 positive colonies that produced specific bands with the length of 200-1000 bp were identified by PCR screening. Then these colonies were subject to sequencing, among which 88 microsatellite-containing fragments were characterized. According to primer design standards and subsequent electrophoresis test, 37 out of 51 pairs of designed SSR primers gave clear and polymorphic bands. Among these primers, 26 pairs were found to produce consistent and clearly distinguishable bands and then were used for subsequent genetic analysis of *M. laosensis*. The characteristics of the 26 SSR primers are given in Table 2. These developed SSR markers were very effective in characterizing the genetic diversity and genetic relationships within the *M. laosensis* populations. Hence these SSR markers would have great significance as they would be further utilized for genetic study, evaluation of germplasm resources and molecular-assisted breeding of *M. laosensis* in the future.

Analysis of genetic diversity

In total, 190 alleles were generated among 152 individuals in eight populations of *M. laosensis*. Table 3 presents results of genetic diversity parameters of *M. laosensis*. The number of alleles per locus (N_a) ranged from 6 (MLP31) to 50 (MLP22) with an average of 18.3. The average of Shannon's information index (I) value was 1.1577. The highest observed heterozygosity (H_o) was 0.6715 (MLP12), while the lowest H_o was 0.0882 (MLP50) and the average of H_o was 0.4114. The expected heterozygosity (H_e) ranged from 0.2972 (MLP50) to 0.8609 (MLP22) with an average of 0.5663. Polymorphism information content (PIC) ranged from 0.5383 (MLP50) to 0.9389 (MLP22) with an average of 0.7759. Together, these parameters revealed a high level of genetic diversity among the natural populations of *M. laosensis*. The average of genetic differentiation coefficient (F_{st}) was 0.3031, indicating middle level of differentiation among populations of *M. laosensis* (30.31 %). The F_{st} was relatively higher than other tree species distributed in subtropical China, such as *Populas adenopoda* (F_{st}=0.1558) (Fan et al. 2018), *Loropetalum chinense* (0.2255) (Gong et al. 2016) and *Castanopsis eyrei* (F_{st}=0.097) (Shi et al. 2014). The inbreeding coefficient (F_{is}) value at each locus were all above zero except locus MLP7 and MLP12, suggesting high extent of inbreeding and large amount of homozygotes in *M. laosensis*.

Within eight populations of *M. laosensis*, the number of alleles (N_a) in each population ranged from 2.1 (GDFK) to 8.6

Table 3
Genetic diversity of *M. laosensis* in SSR loci

Locus	N_a	I	H_o	H_e	F_{st}	F_{is}	PIC
MLP6	30	1.3322	0.5487	0.5668	0.1738	0.0290	0.7777
MLP7	12	1.1102	0.5988	0.5584	0.2815	-0.1076	0.8044
MLP8	32	2.0729	0.6465	0.8333	0.1154	0.2247	0.9303
MLP12	15	1.2215	0.6715	0.6221	0.2510	-0.1266	0.7972
MLP22	50	2.2833	0.6630	0.8609	0.0927	0.2330	0.9389
MLP23	7	0.6893	0.1740	0.4087	0.4340	0.5887	0.7341
MLP27	29	1.3780	0.4428	0.6728	0.2296	0.3830	0.8625
MLP31	6	0.5841	0.2935	0.3273	0.3356	0.1371	0.4732
MLP32	22	0.9906	0.4439	0.4607	0.4564	0.0981	0.7417
MLP33	13	0.9471	0.4602	0.5307	0.3332	0.0548	0.6878
MLP41	18	1.1733	0.3323	0.5789	0.2577	0.4034	0.8391
MLP35	16	1.2046	0.3216	0.6123	0.2763	0.5011	0.8775
MLP37	21	1.1031	0.2944	0.5246	0.3393	0.4374	0.8451
MLP48	11	0.8808	0.2634	0.4943	0.2776	0.5084	0.6694
MLP16	24	1.4867	0.6006	0.7018	0.2160	0.0913	0.8824
MLP34	20	1.4796	0.4580	0.6993	0.2408	0.3076	0.9197
MLP30	11	0.8474	0.3838	0.4716	0.4585	0.1682	0.7976
MLP40	7	0.6691	0.3932	0.3851	0.4256	0.0536	0.7009
MLP43	16	1.2697	0.2729	0.6448	0.2032	0.5355	0.8038
MLP46	12	0.9083	0.3017	0.4774	0.4513	0.3922	0.7569
MLP49	20	0.9477	0.0917	0.4825	0.4571	0.8258	0.5927
MLP3	23	1.1735	0.5151	0.6135	0.3428	0.0499	0.7248
MLP36	12	1.1506	0.4650	0.6063	0.2418	0.2401	0.7750
MLP25	15	1.2350	0.5836	0.6357	0.2250	0.0515	0.8205
MLP39	22	1.4041	0.3877	0.6562	0.2663	0.3974	0.8810
MLP50	11	0.5566	0.0882	0.2972	0.4983	0.6598	0.5383
Mean	18.3	1.1577	0.4114	0.5663	0.3031	0.2658	0.7759
(S.E.)	(1.9)	(0.0450)	(0.0192)	(0.0175)	(0.0218)	(0.0286)	(0.023)

N_a, observed number of alleles; I, Shannon's information index; H_o, observed heterozygosity; H_e, expected heterozygosity; F_{st}, genetic differentiation coefficient of populations; F_{is}, inbreeding coefficient; PIC, polymorphism information content. S.E., standard error of mean.

(GXSS) with an average of 5.3 (Table 4). The number of expected alleles (N_e) ranged from 1.7 (GDFK) to 4.7 (GXJX) with an average of 3.3. The highest I value was 1.6492 (GXJX), while the lowest I was 0.5220 (GDFK). The population with the highest H_o was GXJX (0.5197) and the population with the lowest H_o was YNML (0.2698). The population with the highest H_e was GXJX (0.7376) and the population with the lowest H_e was GDFK (0.3072). GXSS and GDFK population had the highest and the lowest F_{is} value, respectively. Overall, GXJX population had the highest level of genetic diversity, while the GDFK population had the lowest level of genetic diversity. SSR markers could display genetic information about types of homozygote and heterozygote (Wu et al. 1993; Powell et al. 1996). The values of H_o (0.4114), H_e (0.5663) and F_{is} (0.2658) further indicated high extent of inbreeding and excessive homozygote in *M. laosensis*. Therefore, it would be of great importance to select elite individuals from distant populations for hybrid breeding in future.

Table 4
Genetic diversity of *M. laosensis* within eight populations

Code	N_a (S.E.)	N_e (S.E.)	I (S.E.)	H_o (S.E.)	H_e (S.E.)	F_{is} (S.E.)
YNML	2.8 (0.4)	2.1 (0.2)	0.7374 (0.1050)	0.2698 (0.0526)	0.4174 (0.0515)	0.3590 (0.0942)
GDFK	2.1 (0.3)	1.7 (0.3)	0.5220 (0.1023)	0.3282 (0.0708)	0.3072 (0.0529)	-0.0559 (0.1128)
GXDB	4.8 (0.4)	3.0 (0.2)	1.2347 (0.0656)	0.4713 (0.0470)	0.6444 (0.0187)	0.2585 (0.0771)
GXFC	3.8 (0.5)	2.9 (0.4)	0.9792 (0.1301)	0.3557 (0.0595)	0.5013 (0.0556)	0.3021 (0.0910)
GXRX	5.8 (0.7)	3.0 (0.4)	1.1300 (0.1152)	0.4479 (0.0494)	0.5519 (0.0439)	0.2133 (0.0720)
GXLZ	6.0 (0.5)	4.2 (0.4)	1.4645 (0.0927)	0.4890 (0.0516)	0.7023 (0.0262)	0.2966 (0.0720)
GXJX	8.3 (0.8)	4.7 (0.5)	1.6492 (0.0928)	0.5197 (0.0421)	0.7376 (0.0219)	0.2905 (0.0585)
GXSS	8.6 (0.9)	4.5 (0.5)	1.5443 (0.1302)	0.4093 (0.0446)	0.6681 (0.0445)	0.3740 (0.0614)
Mean	5.3 (0.3)	3.3 (0.2)	1.1577 (0.0450)	0.4114 (0.0192)	0.5663 (0.0175)	0.2658 (0.0286)

N_a, observed number of alleles per population; N_e, effective number of alleles; I, Shannon's information index; H_o, observed heterozygosity; H_e, expected heterozygosity; F_{is}, inbreeding coefficient; S.E., standard error of mean.

Analysis of genetic differentiation

Genetic variation and differentiation level are important index for measurement of genetic diversity of populations, which are affected by factors such as microenvironment, reproductive system and life cycle (Adams et al. 1990). In general, populations from significantly different microenvironment are greatly influenced by selection pressure exerted by the external environment, which in turn enhance the extent of differentiation among populations. Second, perennial species usually display lower levels of genetic differentiation than that of annual species. Third, the outcrossing dominant species can produce continuous gene flow among populations and thus display lower level of genetic differentiation than that of selfing dominant species. *M. laosensis* has bisexual flowers that can be fertilized by either self-pollination or cross-pollination. Analysis of molecular variance (AMOVA) revealed that variation within populations account for 81.74 % of total variation and variation among populations account for 18.26 % of total variation, which was consistent with the result that 82.43 % of genetic variation existed within populations as shown by ISSR markers (Peng et al. 2012). The two different markers revealed that variation within populations were the main source of genetic variation in *M. laosensis*.

Pairwise F_{st} and gene flow (N_m) value between populations are given in Table 5. The N_m value varied from 0.316 to 3.626, indicating limited gene flow. The lowest value of pairwise F_{st} was 0.064, while the highest value of pairwise F_{st} was 0.442, suggesting relatively high genetic differentiation in *M. laosensis* populations. Furthermore, the highest N_m value and the lowest pairwise F_{st} value was observed between GXDB and GXJX population, suggesting relatively close relationship between the two populations, whereas the lowest N_m value and the highest pairwise F_{st} value was observed between GDFK and YNML population, suggesting relatively far relationship between them.

Bayesian cluster analysis using STRUCTURE revealed that the genetic structure of *M. laosensis* populations can best be explained by three clusters (Figure 2). Cluster 1 was composed of population YNML, GXDB, GXLZ and GXJX, all of which were in south-western parts of the distribution range. Cluster 2 was composed of population GDFK and GXRX, both of which were in south-central parts of the distribution range. Cluster 3 was composed of population GXFC and GXSS, both of which were in south-eastern part of the distribution range. The three clusters of *M. laosensis* populations based on Bayesian analysis were generally consistent with the geographical distribution patterns of populations.

Cluster analysis

Geographic isolation may affect genetic differentiation of some species in certain extent (Bacilieri et al. 2013; Zehdi-Azouzi et al. 2015). Among the eight natural populations of *M. laosensis*, the genetic distance between GXFC and GXSS population was the smallest (0.3334), while the genetic distance between GXRX and YNML was the largest (0.7126). The genetic distance can directly reflect the geographic distribution of populations. The GXFC population, geographically located in

Table 5
F_{st} (below diagonal) and N_m (above diagonal) values among populations of *M. laosensis*

Code	YNML	GDFK	GXDB	GXFC	GXRX	GXLZ	GXJX	GXSS
YNML		0.316	0.974	0.641	0.567	0.940	1.122	0.881
GDFK	0.442		0.568	0.378	0.734	0.581	0.592	0.562
GXDB	0.204	0.306		1.218	1.031	2.059	3.626	1.566
GXFC	0.281	0.398	0.170		0.774	1.068	1.525	2.442
GXRX	0.306	0.254	0.195	0.244		1.240	1.344	1.498
GXLZ	0.210	0.301	0.108	0.190	0.168		2.693	1.893
GXJX	0.182	0.297	0.064	0.141	0.157	0.085		2.487
GXSS	0.221	0.308	0.138	0.093	0.143	0.117	0.091	

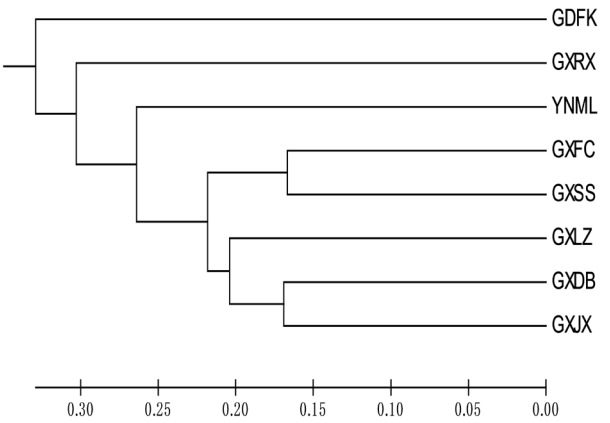

Figure 3
UPGMA dendrogram among eight populations of *M. laosensis*. Population code followed those given in Table 1.

Fangcheng of Guangxi province, was close to the GXSS population located in Shangsi of Guangxi province, whereas the GXRX population located in Rongxian of Guangxi province was quite far from the YNML population, which was located in Mali of Yunnan province. Mantel test provided evidence that there was highly significant correlation between genetic distance and geographic distance (r = 0.8389, p=0.001), indicating that the genetic diversity of *M. laosensis* populations were affected by geographic distance. Furthermore, based on the Euclidean genetic distance of eight populations of *M. laosensis*, UPGMA cluster analysis showed that the eight populations were classified into one major group and three individual groups at the level of 0.25 (Figure 3). One cluster consisted of five populati

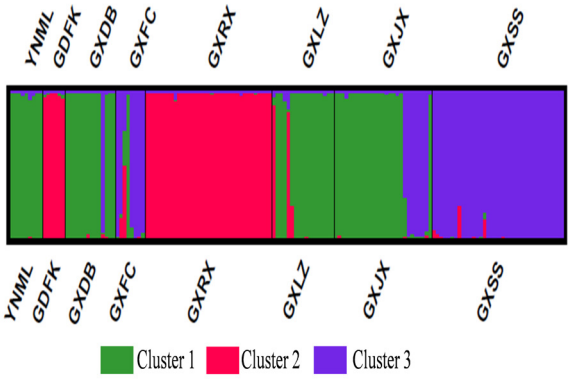

Figure 2
Bayesian cluster analysis of *M. laosensis* population with STRUCTURE.

ons GXFC, GXSS, GXLZ, GXDB and GXJX, all of which were located in southern parts of Guangxi province. The other three populations GDFK (Guangdong), GXRX (southern east of Guangxi) and YNML (Yunnan) were clustered separately. In addition, the GXDB and GXJX populations were closely clustered, while GDFK and YNML population were classified into different groups, which were also consistent with the N_m and F_{st} value of populations (Table 5). Successful genetic improvement programs should start with a breeding population of broad genetic base. It would be sound strategy to import new germplasm from faraway natural distribution areas such as Vietnam and Laos for germplasm resource conservation of *M. laosensis*. Furthermore, selection of elite individuals from populations of far relationship for hybridization and molecular-assisted breeding are important for future genetic improvement programs. Moreover, *M. laosensis* trees were recently successfully introduced to Fujian province of China, which is north of the native distribution areas (Figure 1), and were found to grow fast and well, and generate significant ecological and economic benefits (Huang et al. 2009). The introduction and cultivation of *M. laosensis* in Jiangxi and Hunan province of China were under progress. Here the genetic information could provide support for good adaptability of *M. laosensis*, therefore indicated the feasibility of exploration and development of *M. laosensis* in tropical areas of East Asia.

Conclusions

In this study, twenty-six pairs of SSR primers were developed to investigate the genetic diversity and population structure of 152 individuals from eight natural populations of *M. laosensis*. The genetic parameters suggested high diversity of *M. laosensis*. UPGMA cluster analysis and Mantel test suggested that geographic isolation contribute to genetic differentiation of *M. laosensis* populations. This study suggests a reasonable strategy for germplasm resource conservation and genetic breeding of *M. laosensis*, and provides support for the feasibility of development and promotion of *M. laosensis* fast-growing and high-yield plantations in subtropical areas of East Asia besides the native South Asia.

Acknowledgements

This work was supported by the Fundamental Research Funds of Chinese Academy of Forestry under Grant [number CAFYBB2016SZ002 and CAFYBB2018QB003], Forestry

Scientific and Technological Innovation Funds of Guangdong Province (2016KJCX004), and National Natural Science Foundation of China under Grant [number 31600538].

References

Adams WT, Strauss, SH, Copes DL, Griffin AR (1990) Population Genetics of Forest Trees. Proceedings of the International Symposium on Population Genetics of Forest Trees Corvallis, Oregon, U.S.A.

Bacilieri R, Lacombe T, Le Cunff L, Di Vecchi-Staraz M, Laucou V, Genna B, Péros JP, This P, Boursiquot JM (2013) Genetic structure in cultivated grapevines is linked to geography and human selection. BMC Plant Biol. 13: 25. https://doi.org/10.1186/1471-2229-13-25

Bassam BJ, Caetano-Anollés G, Gresshoff PM (1991) Fast and sensitive silver staining of DNA in polyacrylamide gels. Anal Biochem. 196: 80-83. https://doi.org/10.1016/0003-2697(91)90120-i

Chen L, Zeng J, Jia HY, Zeng J, Guo WF, Cai DX (2012) Growth and nutrient uptake dynamics of Mytilaria laosensis seedlings under exponential and conventional fertilizations. Soil Sci Plant Nutr. 58: 618-626. https://doi.org/10.1080/00380768.2012.708879

Clarke KR, Gorley RN (2001) PRIMER v5: User Manual/Tutorial. PRIMER-E Ltd, Plymouth.

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. https://doi.org/10.1111/j.1365-294x.2005.02553.x

Excoffier L, Laval G, Schneider S (2005) Arlequin (version 3.0): an integrated software package for population genetics data analysis. Evol Bioinform. 1: 47-50. https://doi.org/10.1177/117693430500100003

Fan L, Zheng H, Milne RI, Zhang L, Mao K (2018) Strong population bottleneck and repeated demographic expansions of Populus adenopoda (Salicaceae) in subtropical China. Ann Bot. 121(4): 665-679. https://doi.org/10.1093/aob/mcx198

Gong W, Liu WZ, Gu L, Kaneko S, Koch M. A, Zhang DX (2016) From glacial refugia to wide distribution range: demographic expansion of Loropetalum chinense (Hamamelidaceae) in Chinese subtropical evergreen broadleaved forest. Org Divers Evol. 16(1): 23-38. https://doi.org/10.1007/s13127-015-0252-4

Guo WF, Cai DX, Jia HY, Li YX, Lu ZF (2006) Growth laws of Mytilaria laosensis plantation. For Res. 19: 585–589.

Huang ZQ, Yu ZP, Wang MH (2014) Environmental controls and the influence of tree species on temporal variation in soil respiration in subtropical China. Plant Soil. 382: 75-87. https://doi.org/10.1007/s11104-014-2141-6

Huang Z T, Wang S F, Jiang Y M, Mo J X (2009) Exploitation and utilization prospects of eximious native tree species Mytilaria laosensis. J Guangxi Agric Sci. 40: 1220–1223.

Li F, Gan S (2011) An optimised protocol for fluorescent-dUTP based SSR genotyping and its application to genetic mapping in Eucalyptus. Silvae Genet. 60(1): 18-25. https://doi.org/10.1515/sg-2011-0003

Liu K, Muse SV (2005) PowerMarker: an integrated analysis environment for genetic marker analysis. Bioinformatics. 21: 2128-2129. https://doi.org/10.1093/bioinformatics/bti282

Miller MP (1999) Tools for population genetic analyses (TFPGA): a windows ™ program for the analysis of allozyme and molecular population genetic data. http://www.marksgeneticsoftware.net/tfpga.htm.

Ming AG, Jia HY, Tao Y, Lu LH, Su JM (2012) Biomass and its allocation in 28-year-old Mytilaria laosensis plantation in southwest Guangxi. Chinese J Ecol. 31: 1050–1056.

Ming AG, Jia HY, Zhao JL, Tao Y, Li YF (2016) Above- and below-ground carbon stocks in an indigenous tree (Mytilaria laosensis) plantation chronosequence in subtropical China. Plos One 9: e109730. https://doi.org/10.1371/journal.pone.0109730

Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics. 89: 583-90.

Olango TM, Tesfaye B, Pagnotta MA, Pè ME, Catellani M (2015) Development of SSR markers and genetic diversity analysis in enset (Ensete ventricosum (Welw.) Cheesman), an orphan food security crop from Southern Ethiopia. BMC Genet. 16: 98. https://doi.org/10.1186/s12863-015-0250-8

Peakall ROD, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol ecol notes. 6: 288-295. https://doi.org/10.1111/j.1471-8286.2005.01155.x

Peng JQ, Cao FX, Xu RX (2012) Genetic diversity of the Mytilaria laosensis in Guangxi detected by ISSR markers. J Nat Sci Hunan Norm Uni. 35: 61-65.

Powell W, Morgante M, Andre C, Hanafey M, Vogel J, Tingey S, Rafalski A (1996) The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. Mol Breed. 2(3): 225-238. https://doi.org/10.1007/bf00564200

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

Rosenberg NA (2004) Distruct: a program for the graphical display of population structure. Mol Ecol Notes. 4: 137–138. https://doi.org/10.1046/j.1471-8286.2003.00566.x

Shi MM, Michalski SG, Welk E, Chen XY, Durka W (2014) Phylogeography of a widespread Asian subtropical tree: genetic east–west differentiation and climate envelope modelling suggest multiple glacial refugia. J Biogeogr. 41: 1710-1720. https://doi.org/10.1111/jbi.12322

Sokal R (1979) Testing statistical significance of geographic variation patterns. Syst Zool. 28: 227-232. https://doi.org/10.1093/sysbio/28.2.227a

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol Biol Evol. 24: 1596-1599. https://doi.org/10.1093/molbev/msm092

Wu KS, Tanksley SD (1993) Abundance, polymorphism and genetic mapping of microsatellites in rice. Mol Gen Genet. 241(1-2):225-35. https://doi.org/10.1007/bf00280220

Zane L, Bargelloni L, Patarnello T (2002) Strategies for microsatellite isolation: a review. Molecular Ecology, 11(1): 1-16 https://doi.org/10.1046/j.0962-1083.2001.01418.x

Zehdi-Azouzi S, Cherif E, Moussouni S, et al (2015) Genetic structure of the date palm (Phoenix dactylifera) in the Old World reveals a strong differentiation between eastern and western populations. Ann Bot. 116: 101-112. https://doi.org/10.1093/aob/mcv068