GMR

Genetic diversity and population structure of ancient *Platycladus orientalis* (Cupressaceae) in the Mausoleum of the Yellow Emperor revealed by SSR

Yao Xu[1,2], Jiantao Zhao[3], Sheng Zhang[1,2], Zhong Zhao[1,2*]

[1]College of Forestry, Northwest A&F University, Yangling, Shaanxi, P.R. China

[2]Key Comprehensive Laboratory of Forestry, Yangling, Shaanxi, P.R. China

[3]College of Horticulture, Northwest A&F University, Yangling, Shaanxi, P.R. China

Corresponding author: Zhong Zhao

E-mail: zhaozhlunwen2010@126.com

Genet. Mol. Res. 17 (4): gmr16039924

Received Apr 05, 2018

Accepted Sep 01, 2018

Published Nov 05, 2018

DOI: http://dx.doi.org/10.4238/gmr16039924

ABSTRACT. Ancient trees can serve as genetic reservoirs for special genes, and they are of significance in genetic applications dealing with forest fragmentation, climate change and drought. The Mausoleum of the Yellow Emperor in Shaanxi China has the largest number of ancient *Platycladus orientalis* trees. Here, Simple sequence repeat (SSR) technique was used to study the genetic diversity of ancient *P. orientalis*. Fifty-nine trees were sampled in four age groups and analyzed with 24 SSR markers. This set of SSRs had a moderate genetic diversity (mean He=0.327). Population structure and discriminant analysis of principal components (DAPC) both showed that these individuals were subdivided into four groups and the majority of the 19 oldest trees had a significant different genetic pool, compared to other age groups. This genetic differentiation could be caused by geographic or temporal originality. Our results show that ancient trees represent potential genetic resources for conservations. Better management of these ancient individuals and possible vegetative propagation could play an important role in the genetic conservation of these ancient trees.

Keywords: Genetic relationship; Population structure; Old-growth *Platycladus orientalis*; Simple sequence repeats; Discriminant analysis of principal components; The Mausoleum of the Yellow Emperor

INTRODUCTION

An ancient tree is one that has passed beyond maturity and is old compared to other trees of the same species (ATF, 2008a; Lonsdale, 2013). They are of particular significance: 1) to safeguard genetic resources; 2) to provide habitats for other organisms, such as plants, fungi and animals; 3) to perpetuate aesthetic values; 4) to maintain traditional practices; 5) for historical reasons, such as landmark trees and "Ancient and Famous Trees"; 6) to increase landscape value; and 7) to analyze the aging processes in trees and other organisms (Fay et al., 2000; ATF, 2008b; Read et al., 2013; Lonsdale, 2013).

Ancient trees are assumed to have a higher genetic diversity, compared with modern individuals, and they can serve as genetic reservoirs (ATF, 2008a; Short et al., 2011; Pauls et al., 2013; Zhu et al., 2013; FAO, 2014; Hu et al., 2015). They are commonly found in many historical places, including imperial parks, tombs, ancient temples and mausoleums (Zhu and Lou, 2013). To survive in a given habitat, they must endure climate changes and extreme environmental conditions. The high genetic diversity that characterizes tree populations and individuals and the associated stress tolerance and disease resistance mechanisms help explain their capacity to persist and thrive over long periods (Short et al., 2011; FAO, 2014; Zhang et al., 2016). These ancient trees are of particular significance in genetic applications in dealing with forest fragmentation, climate change, drought, and other disturbances (Lonsdale, 2013; Zhu and Lou, 2013).

Platycladus orientalis (L.), commonly known as Chinese arborvitae is a dominant forest species in China. It is a distinct tree species of the cypress family Cupressaceae with an extremely long lifespan. It is characterized with unisexual flowers, and disperses seeds and pollen mainly via the wind. It naturally occurs in China, Korea and eastern regions of Russia and was also introduced in Europe, North American, East Africa and some Asian countries (Hu et al., 2015; Li et al., 2016). It has been characterized as a salt-, cold-, drought-, and barren-tolerant species with wide adaptability and strong resistance (Hu et al., 2016; Jin et al., 2016; Zhang et al., 2016). Based on simulations, there are approximately 50 countries with a total area of 2.0×10^7 km^2 suitable for the introduction and cultivation of Chinese arborvitae (Li et al., 2016). Therefore, it has the potential for global wide forestation applications. There are many ancient Chinese arborvitaes in the Mausoleum of the Yellow Emperor. The origins of these most ancient trees have long been a mystery.

There are two main hypotheses:

1) They were brought from faraway places when ancient emperors and kings came to the Mausoleum of the Yellow Emperor to worship the Yellow Emperor;

2) They were from the local young Chinese arborvitae. Recently, these trees are under pressure due to climate change and population fragmentation, and their overall number and diversity are declining. Many of them are facing an increasing survival challenge. To protect these ancient Chinese arborvitaes, it is urgent to study their genetic diversity and population structure.

Microsatellites, or simple sequence repeats (SSRs), are ideal molecular markers for population genetic studies. SSRs are characterized by advantages including co-dominance, high polymorphism, reproducibility and cost-effective assessment (Agarwal et al., 2008; Guichoux et al., 2011; Porth et al., 2014). Despite growing competitions from new genetic markers such as SNPs (single nucleotide polymorphisms), SSRs are still very

convenient and cost-effective for evaluating the genetic diversity and population structures of many crops and trees, such as *Camellia japonica* (Zhao et al., 2017), *Genipa americana* (Ruzza et al., 2017), *Solanum lycopersicum* (Zhao et al., 2016), *Cunninghamia lanceolata* (Li et al., 2017) and *P. orientalis* (Zhu and Lou, 2013; Jin et al., 2016). Recently, high-throughput sequencing technologies bridged the gap to develop new polymorphic SSRs due to the exploration of transcriptome-based SSRs in non-model species (Guichoux et al., 2011; Tabbasam et al., 2014 Li et al., 2017). To date, only a limited number of SSRs are available for the genetic studies of *P. orientalis* (Zhu and Lou, 2012; Jin et al., 2016).

In this study, we first developed and evaluated a set of SSRs based on the transcriptome data for ancient *P. orientalis*. We then applied these SSRs to investigate the genetic diversity and population structure of ancient and young *P. orientalis* at four different age levels. Particular attention was paid to the 19 most ancient trees in the yard of the XuanYuan Temple. The objectives of this study were 1) to develop a set of informative SSRs for ancient *P. orientalis*; 2) to analyze the genetic diversity and population structure in relation to the age levels.

MATERIALS AND METHODS

Plant materials

In the Mausoleum of the Yellow Emperor (109°15'E, 35°34'N), there is the largest number of ancient Chinese arborvitae, which have survived for hundreds or even thousands of years. Among them, there are 19 Chinese arborvitaes in the yard of the Xuan Yuan Temple that is oldest and most famous. Most of them are said to be older than 2000 years (hereafter named the HDL_XYM age group). For instance, No. 1 is called ‘Cypress Planted by Yellow Emperor’ (Figure 1).

Figure 1. Geographical location of the studied regions.

It is approximately 5,000 years old, 19 meters tall and 11 meters in circumference at the bottom. It has long been called the ‘No.1 Cypress in the world’. No. 2 is called ‘Han Emperor WuDi’s Cypress for Hanging Armor’ (Figure 2).

Figure 2. Cypress planted by Yellow Emperor (No. 1, or HDL_XYM-1).

Since the accurate ages were unknown, so the ages are roughly estimated based on flattering crown, hollowing trunk degree, staging headedness, breast diameter, etc. Another two ancient age groups were selected, including HDL_2000a (approximately 2000a, with 20 individuals) and HDL_1000a (approximately 1000a, with 10 individuals). In addition, one young age group HDL_50a (with 10 individuals) was also selected. The 2000a age group was randomly selected in the Front Mausoleum Region, which was the dominant region of ancient individuals aged around 2000 years. The 1000a age group was randomly selected in the Back Mausoleum Region, which was the dominant region of ancient individuals aged around 1000 years. The 50a age group was randomly selected in the Fengling Region, where many young individuals were planted (Figure 3).

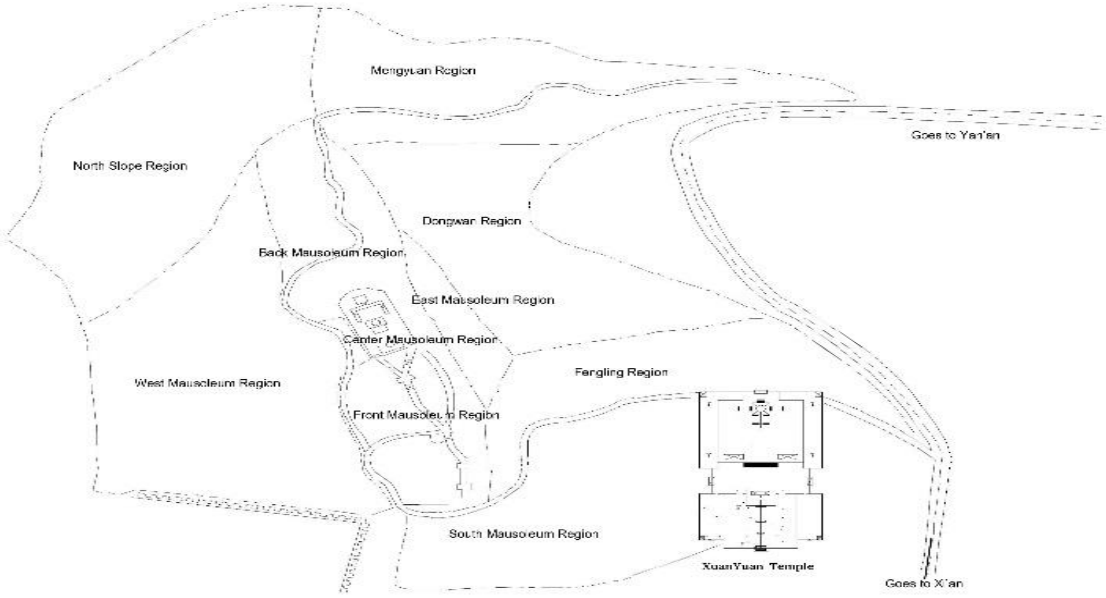

Figure 3. Han Emperor WuDi's Cypress for Hanging Armor (No. 2, or HDL_XYM-2)

SSR development and genotyping

Genomic DNA was extracted from young fresh leaves using the cetyltrimethylammonium bromide (CTAB) method (Doyle, 1973). Two hundred SSRs were selected from our previous transcriptomic data (Zhang et al., 2016). These SSRs represented the most common SSR repeat motifs. Sixteen individuals were randomly picked to select polymorphic SSRs, each age group with four individuals. PCR amplification was conducted using 2x Taq PCR Master Mix (Biomed Technologies, Beijing, China). PCR amplification conditions were performed as follows: 94°C for 5 min; 35 cycles at 94°C for 45 s, annealing at 55-60°C (adjusted according to each SSR) for 1 min and elongation at 72°C for 45 s, with a final extension at 72°C for 10 min. The PCR products were separated on 6% vertical polyacrylamide gels. Thirty-three SSRs with clear and common polymorphisms were selected from the 200 designed SSRs. All 59 studied individuals were genotyped using the same PCR amplification procedure. PCR products were also separated on 6% vertical polyacrylamide gels. There was no measurable difference between marker performances on young and/ or old trees. Genotyping calls were further checked for the presence of null alleles using Micro-Cheaker (Van Oosterhout et al., 2004). In total, 24 SSRs (minor allele frequencies >0.05) were selected and used in further analyses.

Data analysis

In a preliminary analysis, we analyzed the polymorphisms in the developed 24 SSR markers, and we compared the genetic diversity within the 59 individuals and the differentiation between age groups. We computed genetic diversity parameters and differentiation parameters using GenAlex v6.502 (Peakall et al., 2012), including allelic frequencies, the number of alleles (*Na*), the number of effective alleles (*Ne*), the observed *Ho*, the expected heterozygosity (*He*) and the unbiased expected heterozygosity (*uHe*). Fixation index (*F*) and F-statistics across different pre-defined age groups were also measured in GenAlEx v6.502. The average pairwise level of genetic differentiation (*Fst*) between four age groups was calculated using multi-locus comparisons (excluding null alleles) in GenAlEx based on 999 permutations. We also performed an AMOVA analysis in GenAlEx with 999 permutation tests.

In a second phase, we characterized the genetic structure of all individuals by using two different techniques: structure (Pritchard et al., 2000; Falush et al., 2003) and discriminant analysis of principal components (DAPC) (Jombart, 2008). The genetic structure was first evaluated using structure v2.3.4, which uses a Bayesian approach to infer different populations at Hardy-Weinberg equilibrium (hereafter called "genetic components") and assign individuals to them (Pritchard et al., 2000). The number of genetic components to infer *K* was set from 1 to 11, with a burn-in period of 100,000 steps and a Markov Chain Monte Carlo (MCMC) of 200,000 steps. Thirty separate runs were performed for each *K* evaluation. The most likely *K* value was detected using the Evanno transformation method (Evanno et al., 2005) and was processed with structure HARVESTER (Earl et al., 2012). To synthesize structure results, CLUMPP (Jakobsson and Rosenberg, 2007) with the Greedy algorithm was used to compute the average posterior assignment probability of each individual to each genetic component Population structure based on the mean values from CLUMPP was visualized using "ggplot2" package in R (Wickham, 2009). The assignment probability to each component at different ages was also tested using the t-test in "ggpubr" package in R. Dendrogram plot of the studied individuals were illustrated in "ggdendro" ("ave" method) and "ggplot2" in R using the mean value from CLUMPP.

In addition to the structure approach, we performed a discriminant analysis of principal components (DAPC) to cluster individuals based on their genetic similarity using the "adegenet" package in R (Jombart, 2008). This method partitioned the variance within and among groups without assumptions on linkage disequilibrium or Hardy-Weinberg equilibrium (Jombart et al., 2010). The optimal number of clusters was determined by Bayesian Information Criteria (BIC) with the lowest value. All PCs and all discriminant functions were retained to find the optimal number of clusters. In the following DAPC analyses, all discriminant functions and the first 19 principal components were retained. To obtain a robust population structure from the DAPC analyses, 20 independent runs were performed and the results were further synthesized using CLUMPP with the greedy algorithm. Mean assignment probabilities were used to plot a population structure-like result

using the “barplot” function in R. Dendrogram plot of the studied individuals was illustrated in “ggdendro” (“ave” method) and “ggplot2” in R, using the mean value from CLUMPP. Since the accurate age of the ancient individuals was unknown, we redefined the population into slightly different clusters based on the results of structure and DAPC in the following analyses. Significant tests across ages/groups/clusters were calculated using the t-test in “ggpubr” package in R.

RESULTS

Genetic diversity

A set of 24 SSRs was developed in this study (Table 1).

Table 1. Detailed information for 24 SSRs used in this study.

SSR	Motif	Forward primer (5'-3')	Reverse primer (5'-3')	Product (bp)
M-01	(TGATA)5	ATCCCACCATGAAGCTGTTC	TTTACCCCCTACAGCCACTG	159
M-02	(AT)6	AAACGAATGAGGCTGAATGG	GGATGCACGCAATTTTCTTT	166
M-03	(GT)7	ACGGCCTTTGTTTTCTCTCA	AAACCGCCAACACAGGTAAT	265
M-05	(TCT)5	AGTGAGAGCACCTGCTGGAT	AGCAGTGGGCTTTACCCTTT	236
M-08	(AT)6	CATGAATGCATGTGTGTCTCA	TTGGAAATGGCACTGTGGTA	231
M-10	(AT)8	CATGGGACCTTCCATTCATC	AAATGTGGCACCAAATGCTT	198
M-11	(CAG)6	CTTCGTCCCCGATACAAGAG	CATCATGCCCGATATCATCA	251
M-12	(CT)7	GAGTCCAACAAGCTGCATCA	CTGGAGAAAAGGCGTGGTTA	237
M-14	(CTC)5	TGGGGATATAATGCCAAGGA	GCGACGATGAAAAGAATGGT	242
M-15	(TG)6	TAACATCCATCCCTCCATCC	AGCATATCGCTGTTGGCAAT	182
M-18	(CCT)5	CCGGTGCTCCACAAATGTAT	ATCAAGCAAAATTGAGCGGT	183
M-20	(CCT)5	CGAATTGTTTGCCCTGTTCT	CTGATTTTGACTGCTGTGGTT	231
M-21	(AC)6	TCCATTAATAATCGCATACGTCC	GTAGTTGCCATGAGCCTGGT	183
M-22	(AC)6	TGCATTCTATGCGCTTGTTC	GAATGGCTTGCATGCATCTT	277
M-23	(AT)6	CCTACCTTTTGCTACCACGG	CTAGGGTGAATCGCCATGTT	255
M-24	(GA)6	GGGTTTTGCAAAGTATTGCTG	TAGAGGCGCCAATCTCTCAT	270
M-25	(TG)6	AGTGCATGCGTTCATCTCAG	GCCATCAAACAATCAAGCCT	219
M-27	(CA)6	ACATTGATTTGCATTGGGGT	AGAGCACATTCCGGTACCAC	213
M-28	(CTG)5	GATGGCTTTCGCTTGGATTA	GATGAGCACTCAGGATGCAA	140
M-29	(CTG)6	CCTCTGCTCGCAGTCTCTCT	GATGCGCTTGTTTCCGTTAT	189
M-30	(TG)6	GTTGGGCCCCTACTATGGTT	TCCCTTCACTGTCCACTCCT	279
M-31	(GA)6	GGGATATGGGGACCAAGAGT	TGCCGGAGTGTACAACAAAG	204
M-32	(CT)7	TTTCCCTGCCTTCAATCATC	CCTGTCCGGAAAAAGAATGA	263
M-33	(GTA)5	CTCAGCTTCTTGGTTGGAGG	CGCAGATAATGCAAACCTGA	250

They covered the most common SSR motifs in the transcriptomic data, with an average product size ranging from 140 to 279 bp. The average number of alleles (*Na*) was 2.389, ranging from 1.667 to 4.333 among SSRs. The observed heterozygosity (*Ho*) had an average of 0.350 and the highest value was detected at locus M-14 (0.593). In contrast, the expected heterozygosity (*He*) ranged from 0.057 at locus M-31 to 0.589 at locus M-22, with an average value of 0.327. The Shannon's Information index had an average of 0.538 (Table 2).

Table 2. Characterization of 24 microsatellite loci isolated for ancient *P. orientalis*.

Locus	*Na*	*Ne*	*I*	*Ho*	*He*	*uHe*	*F*
M-01	2.333	1.836	0.655	0.496	0.431	0.443	-0.147
M-02	3.333	1.942	0.806	0.466	0.465	0.479	0.044
M-03	2.667	2.119	0.817	0.506	0.522	0.540	0.042
M-05	2.000	1.900	0.664	0.588	0.471	0.488	-0.233
M-08	1.667	1.613	0.448	0.269	0.319	0.326	0.157
M-10	2.333	1.606	0.513	0.342	0.317	0.324	-0.081
M-11	2.000	1.405	0.441	0.279	0.276	0.286	-0.040
M-12	2.333	1.756	0.645	0.517	0.417	0.430	-0.230
M-14	3.667	2.047	0.892	0.593	0.501	0.518	-0.199
M-15	3.000	1.435	0.563	0.351	0.301	0.310	-0.168
M-18	4.333	1.995	0.940	0.527	0.499	0.515	-0.059
M-20	2.000	1.554	0.540	0.442	0.355	0.366	-0.242
M-21	2.000	1.557	0.518	0.342	0.340	0.353	0.033
M-22	3.000	2.447	0.964	0.438	0.589	0.609	0.240
M-23	2.000	1.753	0.609	0.411	0.421	0.433	-0.002
M-24	2.000	1.323	0.376	0.280	0.227	0.236	-0.174
M-25	3.000	1.643	0.676	0.469	0.375	0.387	-0.228
M-27	2.000	1.145	0.243	0.113	0.125	0.128	0.050
M-28	2.000	1.179	0.284	0.165	0.151	0.157	-0.090
M-29	2.000	1.210	0.274	0.184	0.155	0.162	-0.110
M-30	2.000	1.257	0.328	0.226	0.191	0.199	-0.134
M-31	1.667	1.060	0.117	0.057	0.055	0.056	-0.045
M-32	2.000	1.158	0.255	0.146	0.134	0.137	-0.080
M-33	2.000	1.254	0.349	0.192	0.200	0.206	0.002
Mean	2.389	1.591	0.538	0.350	0.327	0.337	-0.074

On average, ancient Chinese arborvitae (HDL_XYM and HDL_2000a) had similar genetic diversity compared to the young age group (not significant) (Table 3).

Table 3 Genetic measures for *P. orientalis* divided into four age groups

Pop	*Na*	*Ne*	*I*	*Ho*	*He*	*uHe*	*F*
HDL_50a	2.167 ± 0.702	1.543 ± 0.421	0.501 ± 0.290	0.357 ± 0.251	0.307 ± 0.179	0.324 ± 0.189	-0.158 ± 0.316
HDL_1000a	2.167 ± 0.482	1.555 ± 0.401	0.508 ± 0.257	0.338 ± 0.239	0.315 ± 0.172	0.332 ± 0.181	-0.025 ± 0.358
HDL_2000a	2.542 ± 0.833	1.654 ± 0.474	0.579 ± 0.293	0.361 ± 0.205	0.347 ± 0.181	0.356 ± 0.186	-0.048 ± 0.256
HDL_XYM	2.458 ± 0.779	1.593 ± 0.426	0.538 ± 0.268	0.339 ± 0.195	0.327 ± 0.178	0.336 ± 0.183	-0.046 ± 0.176
Mean	2.389 ± 0.074	1.591 ± 0.044	0.538 ± 0.028	0.350 ± 0.023	0.327 ± 0.018	0.337 ± 0.019	-0.074 ± 0.029

Genetic differentiation among age groups across the 24 SSR loci was low, ranging from 0.025 to 0.064. The *Fst* increased gradually with increasing age (Table 4).

Table 4. Pairwise levels of genetic differentiation *Fst* values (below diagonal) and *Nei* genetic distance (upper diagonal) across age groups.

Variables	HDL_50a	HDL_1000a	HDL_2000a	HDL_XYM
HDL_50a	-	0.069	0.054	0.075
HDL_1000a	0.060	-	0.030	0.053
HDL_2000a	0.051	0.025	-	0.018
HDL_XYM	0.064	0.041	0.015	-

Similar trends were also found for the *Nei* genetic distance (Table 4).

Population structure

Population structure was first evaluated in structure v2.3.4 with *K* ranging from 1 to 11 with 30 replicates each. The most likely *K* (*K* =4) value was determined based on both the rate of changes of the likelihood distribution (Figure 4A) and the Evanno transformation method (Figure 4B). Four genetic components (Q1, Q2, Q3 and Q4) were found according to the optimal *K* (Figure 4B). Individuals in the most ancient age group (HDL_XYM) had a high significant different genetic component (Q1) ($p<0.001$), compared with other age groups (Figure 4C). Dendrogram tree revealed that these 59 studied individuals could be subdivided into three clusters (Figure 4D): 1) Cluster 1 consisted of all individuals with dominant genetic component Q3; 2) Cluster 2 consisted of all individuals with dominant genetic component Q1; 3) Cluster 3 included the remaining individuals who are admixed with individuals from all age groups (Q2 and Q4 accounted for most genetic components). Notably, only ancient Chinese arborvitae appeared in Cluster 1, including several of the most ancient trees in HDL_XYM and some ancient trees in HDL_2000a and HDL_1000a. Cluster 2 consisted of only the most ancient trees in HDL_XYM (Figure 4E). Population structure was partly consistent with the age subdivision (Figure 4D).

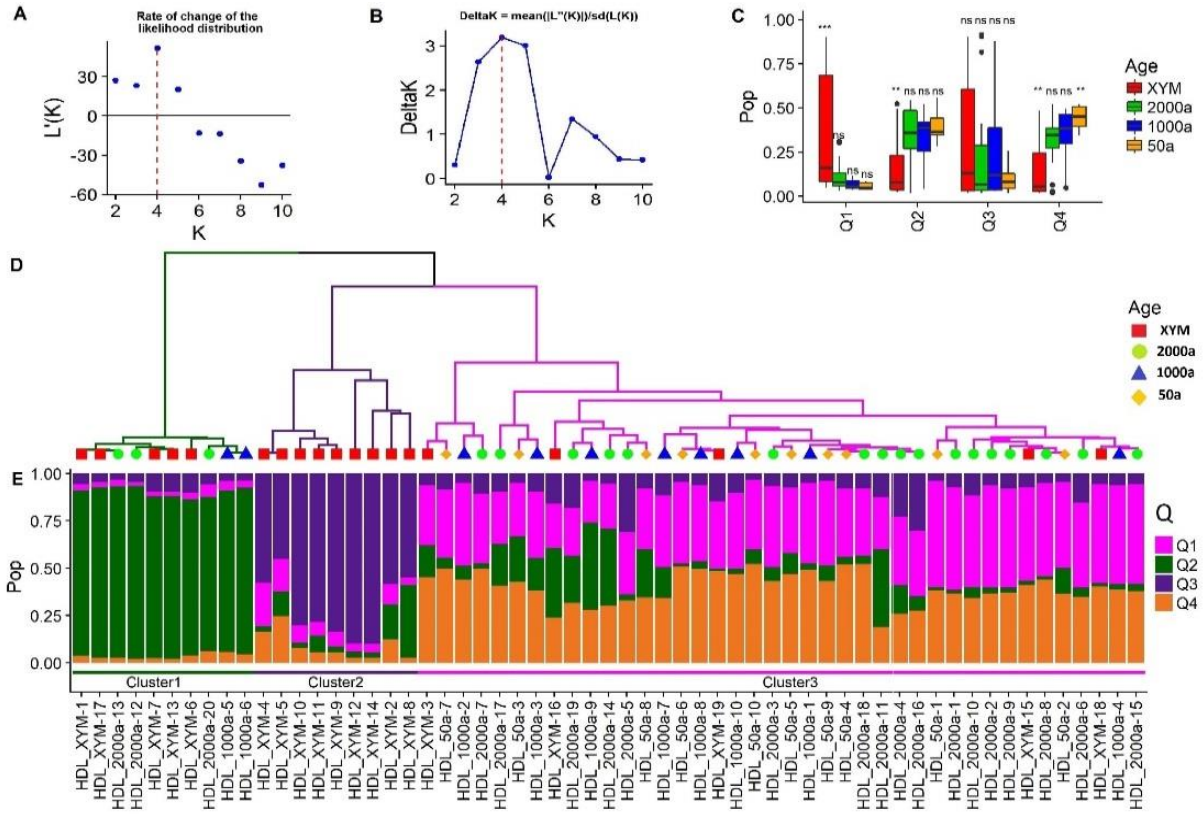

Figure 4. Population structure of 59 *Platycladus orientalis* based on 24 SSRs. (A) Rate of change of the likelihood distribution. The red vertical line indicates the optimal number of population structure. (B) Optimal K evaluation based on the Evanno method. The optimal number (*K*=4) was indicated by the vertical red line. (C) Significant t-tests among four studied age groups. Pop represents the corresponding mean assignment probability values from CLUMPP. *, $p<0.05$; *, $p<0.01$; ***, $p<0.005$; ns, not significant. (D) Dendrogram of 59 *P. orientalis* based on the population structure from the mean value of 30 independent runs in structure. Individuals in HDL_XYM were labeled with a red square; individuals in HDL_2000a were labeled with a green circle; individuals in HDL_1000a were labeled with a blue triangle; individuals in HDL_50a were labeled with an orange diamond. (E) Population structure based on 30 independent runs. Each bar represents a single sample and based on the dominant genetic component (Q), all individuals could be mainly subdivided into three clusters.

AMOVA analysis revealed that most of the genetic variance was occurred within populations (Table 5). Together, these results showed that there were weak population stratifications for main individuals apart from the oldest age group (HDL_XYM).

Table 5. Analysis of molecular variance (AMOVA) within/among different age groups of ancient Chinese arborvitae.

Source	df	SS	MS	Estimated Variance	Percentage
Among Pops	3	30.336	10.112	0.206	5%
Within Pops	114	483.961	4.245	4.245	95%
Total	117	514.297	--	4.451	100%

Since the 19 oldest individuals had a high significant genetic component (Figure 4C), we provided further analyses of them (Figure 5).

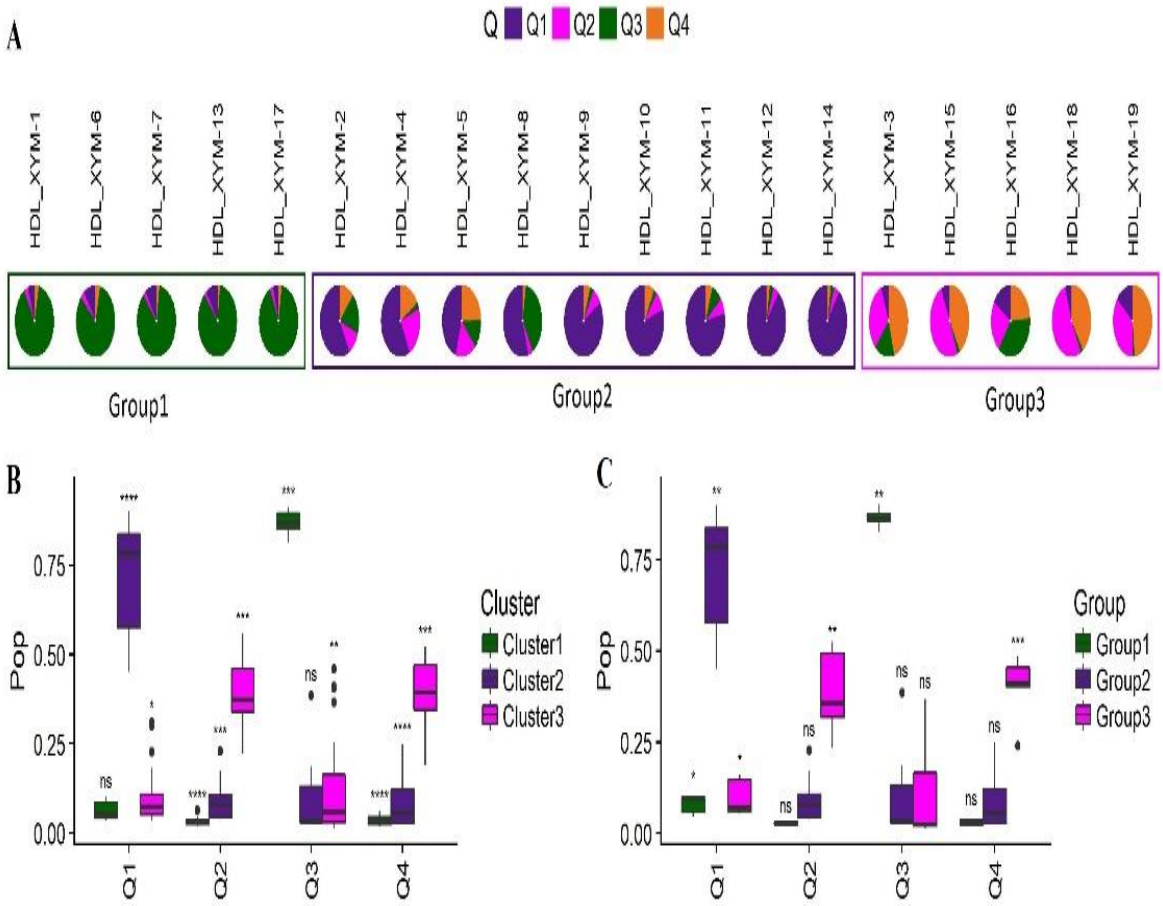

Figure 5. Sub-group assignment of the 19 oldest *Platycladus orientalis* and significant tests based on the mean results of structure. (A) Group division of the 19 most ancient individuals in HDL_XYM. The five individuals appeared in Cluster 1 were grouped as Group 1. The nine individuals in Cluster 2 were grouped as Group 2. The remaining five individuals were grouped as Group 3. (B) Significant t-tests of the three main clusters revealed in the phylogenetic relationship. *, p<0.05; *, p<0.01; ***, p<0.005; ****, p<0.001; ns, not significant. (C) Significant t-tests of the three groups in HDL_XYM. *, p<0.05; *, p<0.01; ***, p<0.005; ns, not significant.

For the 19 most ancient individuals in HDL_XYM, 14 of them were clearly assigned to two dominant components (Q1 and Q3) (Table 6).

Table 6. Detailed information for the division of groups, clusters and components for structure analysis.

Individual	Age group	Cluster	Group
HDL_XYM-1	HDL_XYM	Cluster 1	Group 1
HDL_XYM-2	HDL_XYM	Cluster 2	Group 2
HDL_XYM-3	HDL_XYM	Cluster 3	Group 3
HDL_XYM-4	HDL_XYM	Cluster 2	Group 2
HDL_XYM-5	HDL_XYM	Cluster 2	Group 2
HDL_XYM-6	HDL_XYM	Cluster 1	Group 1
HDL_XYM-7	HDL_XYM	Cluster 1	Group 1
HDL_XYM-8	HDL_XYM	Cluster 2	Group 2
HDL_XYM-9	HDL_XYM	Cluster 2	Group 2

HDL_XYM-10	HDL_XYM	Cluster 2	Group 2
HDL_XYM-11	HDL_XYM	Cluster 2	Group 2
HDL_XYM-12	HDL_XYM	Cluster 2	Group 2
HDL_XYM-13	HDL_XYM	Cluster 1	Group 1
HDL_XYM-14	HDL_XYM	Cluster 2	Group 2
HDL_XYM-15	HDL_XYM	Cluster 3	Group 3
HDL_XYM-16	HDL_XYM	Cluster 3	Group 3
HDL_XYM-17	HDL_XYM	Cluster 1	Group 1
HDL_XYM-18	HDL_XYM	Cluster 3	Group 3
HDL_XYM-19	HDL_XYM	Cluster 3	Group 3
HDL_2000a-1	Around 2000a	Cluster 3	
HDL_2000a-2	Around 2000a	Cluster 3	
HDL_2000a-3	Around 2000a	Cluster 3	
HDL_2000a-4	Around 2000a	Cluster 3	
HDL_2000a-5	Around 2000a	Cluster 3	
HDL_2000a-6	Around 2000a	Cluster 3	
HDL_2000a-7	Around 2000a	Cluster 3	
HDL_2000a-8	Around 2000a	Cluster 3	
HDL_2000a-9	Around 2000a	Cluster 3	
HDL_2000a-10	Around 2000a	Cluster 3	
HDL_2000a-11	Around 2000a	Cluster 3	
HDL_2000a-12	Around 2000a	Cluster 1	
HDL_2000a-13	Around 2000a	Cluster 1	
HDL_2000a-14	Around 2000a	Cluster 3	
HDL_2000a-15	Around 2000a	Cluster 3	
HDL_2000a-16	Around 2000a	Cluster 3	
HDL_2000a-17	Around 2000a	Cluster 3	
HDL_2000a-18	Around 2000a	Cluster 3	
HDL_2000a-19	Around 2000a	Cluster 3	
HDL_2000a-20	Around 2000a	Cluster 1	
HDL_1000a-1	Around 1000a	Cluster 3	
HDL_1000a-2	Around 1000a	Cluster 3	
HDL_1000a-3	Around 1000a	Cluster 3	
HDL_1000a-4	Around 1000a	Cluster 3	
HDL_1000a-5	Around 1000a	Cluster 1	
HDL_1000a-6	Around 1000a	Cluster 1	
HDL_1000a-7	Around 1000a	Cluster 3	
HDL_1000a-8	Around 1000a	Cluster 3	
HDL_1000a-9	Around 1000a	Cluster 3	
HDL_1000a-10	Around 1000a	Cluster 3	
HDL_50a-1	Around 50a	Cluster 3	
HDL_50a-2	Around 50a	Cluster 3	
HDL_50a-3	Around 50a	Cluster 3	
HDL_50a-4	Around 50a	Cluster 3	

HDL_50a-5	Around 50a	Cluster 3
HDL_50a-6	Around 50a	Cluster 3
HDL_50a-7	Around 50a	Cluster 3
HDL_50a-8	Around 50a	Cluster 3
HDL_50a-9	Around 50a	Cluster 3
HDL_50a-10	Around 50a	Cluster 3

For the 19 individuals we subdivided those into three groups (Group 1, 2 and 3) with 5, 9 and 5 individuals, respectively (Figure 2A, Table 6). For cluster 1, the dominant component Q3 was high significantly (p<0.001) different from the other two clusters. The dominant Q1 in Cluster 2 was also significantly different from all the other clusters (Figure 5B). For the three groups in HDL_XYM, Q1 in Group 2 was significantly different from the other groups; Q3 in Group 1 was also significantly different from the other groups (Figure 5C). Similarly, the genetic component Q3, mainly in Group 1, was significantly different (p<0.001) from the other groups. The admixed Q2/Q4 represented the main genetic pool for Group 3, which was significantly different (p<0.001) from the other groups (Figure 5C).

Discriminant analysis of principal components

Discriminant analysis of principal components (DAPC) was performed to further investigate the genetic structure. Four clusters were found in DAPC, which was consistent with structure (Figure 6).

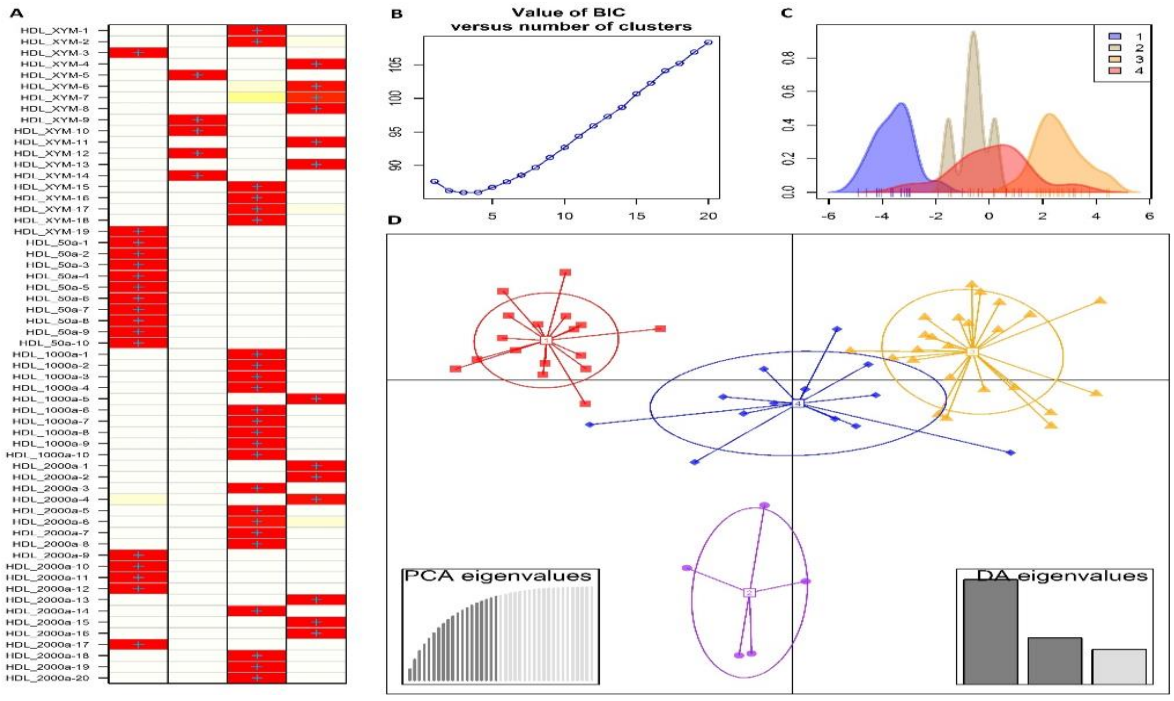

Figure 6. Discriminant analysis of principal components (DAPC) of the 59 *Platycladus orientalis*. (A) Proportions of successful reassignment of individuals to their original clusters (based on the discriminant functions). Heat colors represent membership probabilities (red=1, white=0); blue crosses represent the prior cluster provided to DAPC. (B) Value of Bayesian Information Criterion (BIC) versus the number of clusters. The lowest BIC value indicates the optimal number of clusters (K=4). (C) Densities of individuals on the first discriminant function. Each color represents one cluster. (D) DAPC results of the 59 *P. orientalis* with the first 19 principal components and all three discriminant function retained.

Most of the individuals had a high proportion of successful assignment (Figure 6A). The optimal number of clusters (*K*=4) was determined using Bayesian Information Criteria (BIC) with the lowest value (Figure 6B). The first discriminant function could separate cluster 1 and cluster 3 clearly, while cluster 2 and 4 were admixed (Figure 6C). Comparing the correspondence between the three clusters in structure with the four clusters in DAPC, Cluster 2 in DAPC only contained Cluster 3 in structure; Cluster 4 in DAPC was dominated by Cluster 2 in structure; Cluster 1 and 3 in DAPC contained the three structure clusters in various proportions (Figure 6A).

Dendrogram analyses using the mean values of 20 replicate DAPC analyses revealed four clusters (Figure 7A). Cluster 1 consisted of ancient individuals who were mainly from HDL_XYM and HDL_2000a. Cluster 2 was dominated by the individuals from the modern young age group HDL_50a. Cluster 3 was the admixture of HDL_2000a and HDL_1000a. Cluster 4 was dominated by individuals from HDL_XYM (Figure 7B). DAPC had a better performance in assigning the individuals in HDL_1000a and HDL_50a compared to the results from structure (Figure 7B).

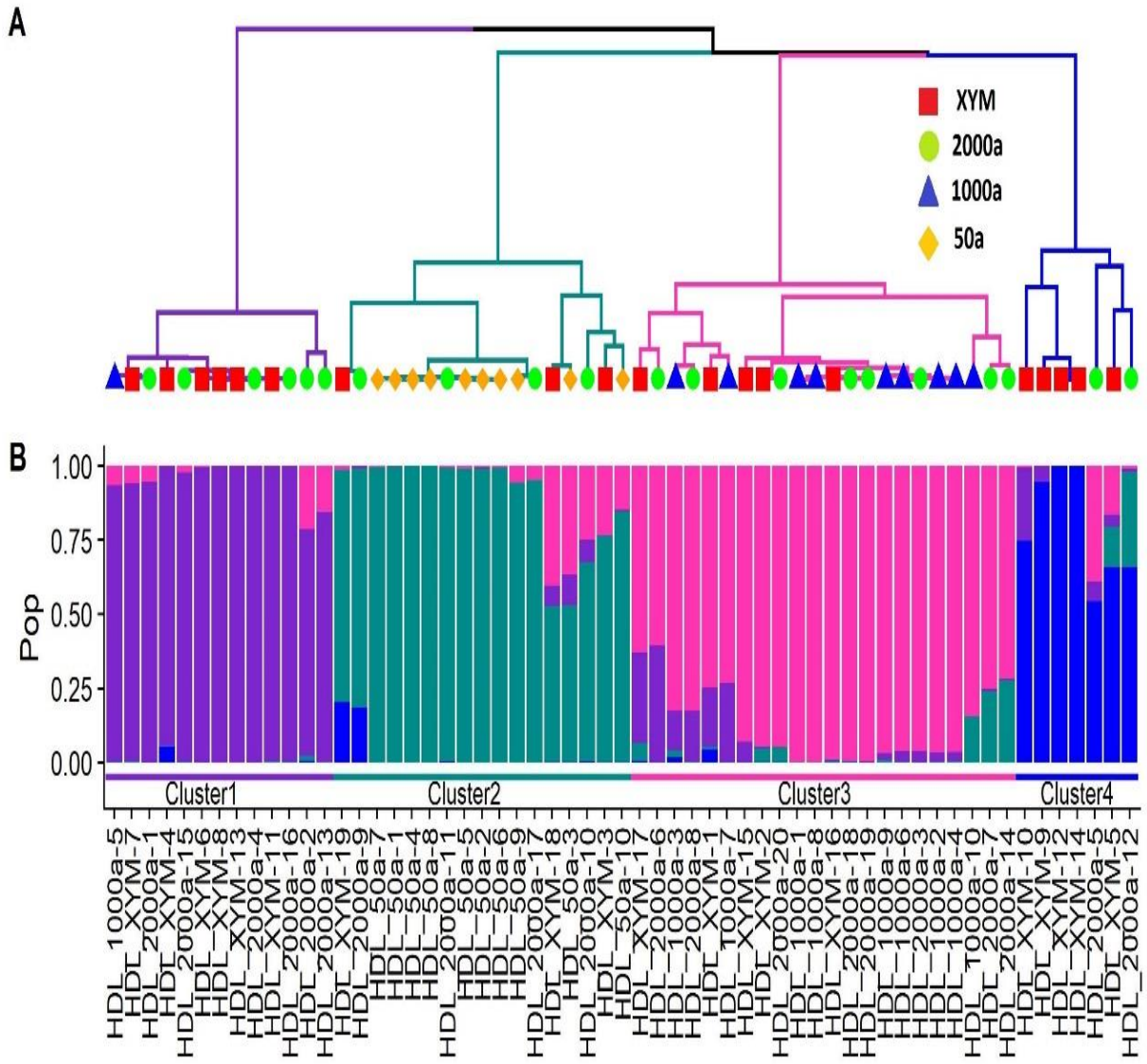

Figure 7. Phylogenetic relationship and significant tests of the 59 *Platycladus orientalis* based on DAPC analyses. (A) Phylogenetic relationship of the 59 *Platycladus orientalis* based on the mean values of 20 independent runs with the first 19 principal components and all discriminant functions retained. (B) structure-like distribution based on the mean values of 20 independent runs

By contrast, the genetic structure of age group HDL_XYM was more complex, indicating a diverse genetic background for the 19 oldest Chinese arborvitaes. They could also be subdivided into four groups, corresponding to which cluster they belonged (Figure 8 and Table 7).

Table 7. Detailed information for the division of groups, clusters and components for discriminant analysis of principal components analysis.

Individual	Age group	Cluster	Group
HDL_XYM-1	HDL_XYM	Cluster 3	Group 3
HDL_XYM-2	HDL_XYM	Cluster 3	Group 3
HDL_XYM-3	HDL_XYM	Cluster 2	Group 2
HDL_XYM-4	HDL_XYM	Cluster 1	Group 1
HDL_XYM-5	HDL_XYM	Cluster 4	Group 4
HDL_XYM-6	HDL_XYM	Cluster 1	Group 1
HDL_XYM-7	HDL_XYM	Cluster 1	Group 1
HDL_XYM-8	HDL_XYM	Cluster 1	Group 1
HDL_XYM-9	HDL_XYM	Cluster 4	Group 4
HDL_XYM-10	HDL_XYM	Cluster 4	Group 4
HDL_XYM-11	HDL_XYM	Cluster 1	Group 1
HDL_XYM-12	HDL_XYM	Cluster 4	Group 4
HDL_XYM-13	HDL_XYM	Cluster 1	Group 1
HDL_XYM-14	HDL_XYM	Cluster 4	Group 4
HDL_XYM-15	HDL_XYM	Cluster 3	Group 3
HDL_XYM-16	HDL_XYM	Cluster 3	Group 3
HDL_XYM-17	HDL_XYM	Cluster 3	Group 3
HDL_XYM-18	HDL_XYM	Cluster 2	Group 2
HDL_XYM-19	HDL_XYM	Cluster 2	Group 2
HDL_2000a-1	Around 2000a	Cluster 1	
HDL_2000a-2	Around 2000a	Cluster 1	
HDL_2000a-3	Around 2000a	Cluster 3	
HDL_2000a-4	Around 2000a	Cluster 1	
HDL_2000a-5	Around 2000a	Cluster 4	
HDL_2000a-6	Around 2000a	Cluster 3	
HDL_2000a-7	Around 2000a	Cluster 3	
HDL_2000a-8	Around 2000a	Cluster 3	
HDL_2000a-9	Around 2000a	Cluster 2	
HDL_2000a-10	Around 2000a	Cluster 2	
HDL_2000a-11	Around 2000a	Cluster 2	
HDL_2000a-12	Around 2000a	Cluster 4	
HDL_2000a-13	Around 2000a	Cluster 1	
HDL_2000a-14	Around 2000a	Cluster 3	
HDL_2000a-15	Around 2000a	Cluster 1	

HDL_2000a-16	Around 2000a	Cluster 1
HDL_2000a-17	Around 2000a	Cluster 2
HDL_2000a-18	Around 2000a	Cluster 3
HDL_2000a-19	Around 2000a	Cluster 3
HDL_2000a-20	Around 2000a	Cluster 3
HDL_1000a-1	Around 1000a	Cluster 3
HDL_1000a-2	Around 1000a	Cluster 3
HDL_1000a-3	Around 1000a	Cluster 3
HDL_1000a-4	Around 1000a	Cluster 3
HDL_1000a-5	Around 1000a	Cluster 1
HDL_1000a-6	Around 1000a	Cluster 3
HDL_1000a-7	Around 1000a	Cluster 3
HDL_1000a-8	Around 1000a	Cluster 3
HDL_1000a-9	Around 1000a	Cluster 3
HDL_1000a-10	Around 1000a	Cluster 3
HDL_50a-1	Around 50a	Cluster 2
HDL_50a-2	Around 50a	Cluster 2
HDL_50a-3	Around 50a	Cluster 2
HDL_50a-4	Around 50a	Cluster 2
HDL_50a-5	Around 50a	Cluster 2
HDL_50a-6	Around 50a	Cluster 2
HDL_50a-7	Around 50a	Cluster 2
HDL_50a-8	Around 50a	Cluster 2
HDL_50a-9	Around 50a	Cluster 2
HDL_50a-10	Around 50a	Cluster 2

Significant differences were commonly detected across four clusters for each genetic component (Figure 8B). Notably, Q4 was significantly (p<0.001) different from all the others. It was clearly defined in cluster 4, which was dominated by the most ancient individuals from HDL_XYM (Figure 7C).

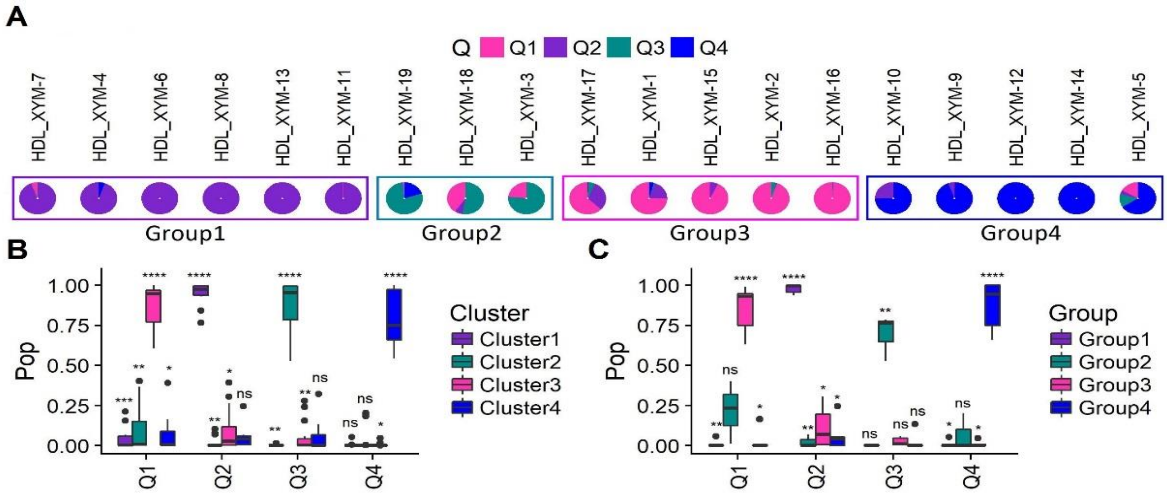

Figure 8. Subgroup assignment of the 19 oldest *Platycladus orientalis* and significant tests based on the mean results of DAPC. (A) Group division of the 19 most ancient individuals in HDL_XYM. (B) Significant t-tests of clusters revealed in the phylogenetic relationship. *, p<0.05; *, p<0.01; ***, p<0.005; ****, p<0.001; ns, not significant. (C) Significant t-tests of groups in HDL_XYM. *, p<0.05; *, p<0.01; ***, p<0.005; ns, not significant.

DISCUSSION

In this study, a set of 24 polymorphic SSRs were developed for ancient Chinese arborvitae trees. We found that polymorphisms (*Na* and *He)* were moderately common. In a recent study focusing on ancient Chinese arborvitae at different historical places in Beijing, similar polymorphisms were also found for young trees, but polymorphisms were lower for the ancient trees (Zhu and Lou, 2013). This could be due to the way the SSRs were developed (Zhu and Lou, 2012) or the geographical locations and population sizes used (Zhu and Lou, 2013). The genetic diversity result (mean *He*=0.327) was also similar compared to another genetic diversity study (mean *He*=0.349) of a core breeding collection of Chinese arborvitae from central China (Jin et al., 2016). In the analysis of structure, we detected that ancient trees represented an original diversity that was not found in the local young trees. In contrast, those younger individuals shared an admixed structure. Similar results were also found in a previous study (Jin et al., 2016). These results indicated that there was a tendency of loss of diversity, especially between modern and oldest individuals. Until now, molecular markers for Chinese arborvitae, especially SSRs for ancient Chinese arborvitae, were quite limited and only a few were available (Zhu and Lou, 2013; Jin et al., 2016). These newly developed SSRs can be used for many applications in both ancient and young individuals, including genetic diversity evaluation, population structure analyses, marker-assisted breeding, and genome-wide association studies (Agarwal et al., 2008; Guichoux et al., 2011; Porth and El-Kassaby, 2014).

Two different techniques, including structure and DAPC, were used to analyze the population structure. Though structure (Pritchard et al., 2000; Falush et al., 2003) was one of the most widely applied programs for inferring population structuring, especially in assigning admixed individuals (Porras-Hurtado et al., 2013), the reliance on Bayesian clustering limits its efficiency and applicability (Jombart et al., 2010). In addition, estimating large population sizes can take considerable computational time to complete (Lee et al., 2009). In comparison, DAPC partitions the variance within and among groups without assumptions on linkage disequilibrium or Hardy-Weinberg equilibrium, and it does not rely on a particular population genetics model (Jombart et al., 2008; Jombart et al., 2010). Additionally, DAPC is a multivariate approach based on the distance among individuals. It minimizes the variance within groups, maximizes the variance between groups and can be used as a complementary approach to structure (Jombart et al., 2010). Although the optimal number of population structure was four, structure detected two clear populations (Cluster 1 and Cluster 2) and no clear detection for other individuals (Cluster 3). DAPC detected 4 groups partially related to the structure results (in particular for 2 clusters). It helped to clarify the structure within the third admixed cluster of the structure analysis. All young trees (HDL_50a) and most of the less ancient trees (HDL_1000a) were admixed in structure. In contrast, they were clearly separated into two distinct clusters in DAPC.

In this study, the 19 most ancient Chinese arborvitae located the yard of XuanYuan Temple are particularly interesting: 1) they are the most ancient group in the studied region; 2) their origins have long been a mystery; 2) these 19 ancient individuals have particular religious significance and aesthetic values; 4) their health status is currently deteriorating due to climate change and tourism. Population structure in structure and DAPC showed that most of them had significantly different genetic pools compared to individuals from other age groups. In addition, the genetic backgrounds of the 19 most ancient individuals were not always the same and could be subdivided into two groups. These results provided supporting genetic information for the long-term hypothesis that some individuals in the yard of XuanYuan Temple could have been brought from other places (representing different genetic backgrounds). This could have occurred during historical memorial ceremonies for the Yellow Emperor in ancient dynasties, which was a common and popular tradition in the history of China.

Based on legend, No. 1 (or HDL_XYM-1) was called as ‘Cypress Planted by Yellow Emperor’ and was said to have been planted by the Yellow Emperor himself. If this were true, the tree should be more than 5000 years old. Below the western steps of the hall of XuanYuan Temple, there was a comparatively smaller Chinese arborvitae called ‘Han Emperor WuDi’s Cypress for Hanging Armor’ (No. 2, or HDL_XYM-2). It is

said that Han Emperor WuDi hung his armor when he offered his sacrifice to the Yellow Emperor after returning from the north expedition. In this case, No. 2 could have been brought from the northern regions after returning from the north expedition in China. When combining the population assignment from structure and geographical location, these two legendary individuals represented the two main genetic pools of the 19 most ancient individuals (Figure 8). Notably, individuals sharing a close genetic background with No. 2 are highly significant compared to all others, indicating that they could also be from the northern regions of China. In contrast, the dominant genetic component of No. 1 also appeared in some individuals in HDL_2000a and HDL_1000a, possibly representing native genetic backgrounds (Figure 9).

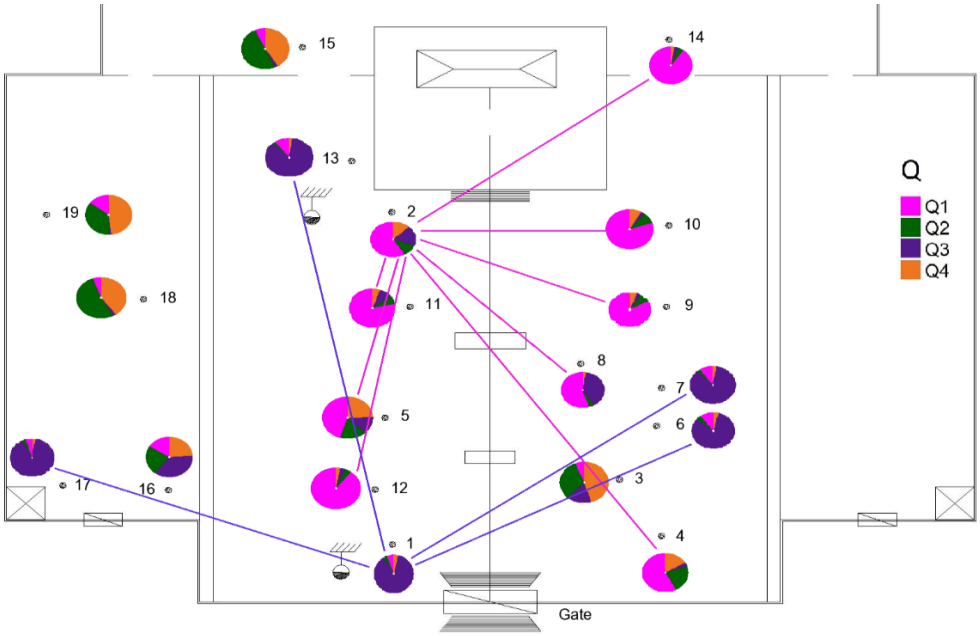

Figure 9. Geographical distributions and corresponding structure assigning from structure. Each number represents one individual and follows the same order. Pie plots represented the individual assignments based on the mean values of 30 independent runs at the optimal number of structure in structure.

It can, therefore, be viewed as the ancestor of the other ancient individuals in the region of the Mausoleum of the Yellow Emperor. Combining geographical location and DAPC, consistencies were found for several ancient trees (Figure 10), including HDL_XYM-6 and HDL_XYM-7; XYM-1 and HDL_XYM-17. However, more historical and genetic studies are needed to further validate their relationships and possible origins.

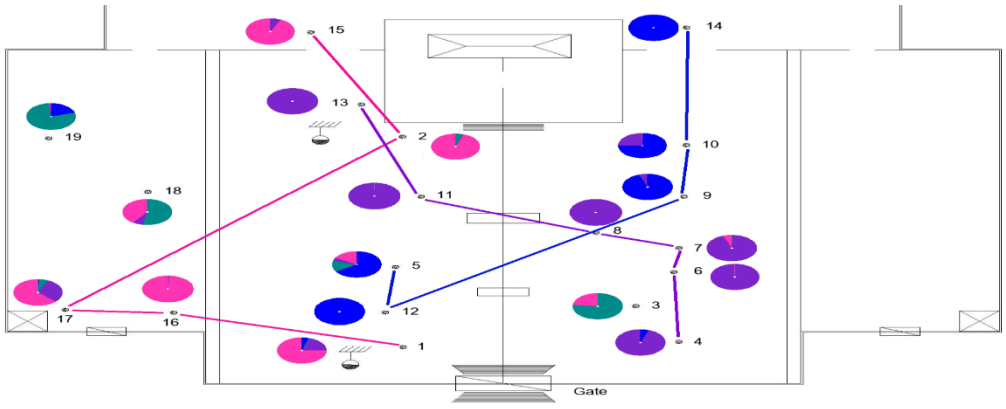

Figure 10. Geographical distributions and corresponding structure assigning from DAPC

CONCLUSION

Our results show that ancient trees have historical, aesthetic significance and that they also represent interesting genetic resources. Why do they have this genetic originality? Because they may come from distant non-local origins (the geographic originality aspect) or because they are very old and represent a gene pool that may have disappeared (the temporal originality aspect). The genetic originality is another reason to pay attention to these ancient trees and conserve them. The conservation of these ancient trees contributes to the conservation of genetic resources and biodiversity of the species of interest (Thompson et al., 2009; FAO, 2014). Better management of their ecosystems and possible vegetative propagation could play an important role in the genetic conservation management of these ancient trees (Read, 2013; Zhu and Lou, 2013).

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We thank Li. M for suggestions on data analysis and manuscript. We appreciate the helps of Li. Y, and other colleges in collecting the samples.

AUTHOR CONTRIBUTIONS

Y.X., J.T., Z. and Z.Z. conceived and designed the experiments. Y.X., J.-T., Z. and S.Z. conducted the experiments. Y.X. and J.-T., Z. analyzed the data and wrote the manuscript. All authors reviewed the manuscript and approved the submission. We really appreciate insightful suggestions of the two reviewers.

REFERENCES

Agarwal M, Shrivastava N and Padh H (2008) Advances in molecular marker techniques and their applications in plant sciences. *Plant Cell Reports*. 27: 617–631. https://doi.org/10.1007/s00299-008-0507-z

ATF (2008a) Ancient Tree Guide no.4: What are ancient, veteran and other trees of special interest? The Woodland Trust, Grantham.

ATF (2008b) Ancient Tree Guide no.2: Trees in historic parks and landscape gardens. The Woodland Trust, Grantham.

Doyle JLJJ (1973) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem Bull.* 28: 330.

Earl DA and Von Holdt BM (2012) structure HARVESTER: A website and program for visualizing structure output and implementing the Evanno method. *Conserv Genet Resour*. 4: 359–361. https://doi.org/10.1007/s12686-011-9548-7

Evanno G, Regnaut S and 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

Falush D, Stephens M and Pritchard JK (2003) Inference of population structure using multilocus genotype data linked loci and correlated allele frequencies. *Genetics* 164: 1567–1587. https://doi.org/10.1111/j.1471-8286.2007.01758.x

FAO (2014) The State of the World's Forest Genetic Resources.

Fay N, Mynors C and Davis C (2000) Veteran trees: A guide to risk and responsibility. English Nature, Peterborough.

Guichoux E, Lagache L, Wagner S, Chaumeil P, et al. (2011). Current trends in microsatellite genotyping. *Mol Ecol Resour*. 11: 591–611. https://doi.org/10.1111/j.1755-0998.2011.03014.x

Hu XG, Jin Y, Wang XR, Mao JF and Li Y (2015) Predicting impacts of future climate change on the distribution of the widespread conifer *Platycladus orientalis*. *PLoS One*. 10: e0132326. https://doi.org/10.1371/journal.pone.0132326

Jakobsson M and Rosenberg NA (2007) CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* 23: 1801–1806. https://doi.org/10.1093/bioinformatics/btm233

Jin Y, Ma Y, Wang S, Hu XG, et al. (2016) Genetic evaluation of the breeding population of a valuable reforestation conifer *Platycladus orientalis* (Cupressaceae). Sci Rep. 6: 34821. https://doi.org/10.1038/srep34821

Jombart T (2008). adegenet: A R-package for the multivariate analysis of genetic markers. *Bioinformatics* 24: 1403–1405. https://doi.org/10.1093/bioinformatics/btn129

Jombart T, Devillard S and Balloux F (2010) Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genet*. 11: 94. https://doi.org/10.1186/1471-2156-11-94

Li G, Du S and Wen Z (2016) Mapping the climatic suitable habitat of oriental arborvitae (*Platycladus orientalis*) for introduction and cultivation at a global scale. *Sci Rep.* 6: 30009. https://doi.org/10.1038/srep30009

Hu XG, Liu H, Jin YQ, Sun YQ, et al. (2016) *De novo* transcriptome assembly and characterization for the widespread and stress-tolerant conifer *Platycladus orientalis*. *PLoS One*. 11: e148985. https://doi.org/10.1371/journal.pone.0148985

Li M, Chen X, Huang M, Wu P, et al. (2017) Genetic diversity and relationships of ancient Chinese fir (*Cunninghamia lanceolata*) genotypes revealed by sequence-related amplified polymorphism markers. *Genet Resour Crop Evol.* 64: 1087–1099. https://doi.org/10.1007/s10722-016-0428-6

Lee C, Abdoo l A and Huang CH (2009) PCA-based population structure inference with generic clustering algorithms. *BMC Bioinformatics*. 10: S73. https://doi.org/10.1186/1471-2105-10-s1-s73

Lonsdale D (2013) Ancient and other veteran trees: Further guidance on management. *Arboricultural Journal* 35: 110-111 https://doi.org/10.1080/03071375.2013.823318

Mosseler A, Thompson I and Pendrel BA (2003) Overview of old-growth forests in Canada from a science perspective. *Environ Rev.* 11: S1–S7. https://doi.org/10.1139/a03-018

Mosseler A, Major JE and Rajora OP (2003) Old-growth red spruce forests as reservoirs of genetic diversity and reproductive fitness. *Theor Appl Genet*. 106: 931–937. https://doi.org/10.1007/s00122-002-1156-1

Pauls SU, Nowak C, Bálint M and Pfenninger M (2013) The impact of global climate change on genetic diversity within populations and species. *Mol. Ecol.* 22: 925–946. https://doi.org/10.1111/mec.12152

Peakall R and Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics* 28: 2537–2539. https://doi.org/10.1093/bioinformatics/bts460

Pritchard JK, Stephens M and Donnelly P (2010) Inference of population structure using multilocus genotype data. *Genetics*155: 945–959. https://doi.org/10.1111/j.1471-8286.2007.01758.x

Porras-Hurtado L, Ruiz Y, Santos C, Phillips C, et al. (2013) An overview of structure: Applications, parameter settings, and supporting software. *Front Genet*. 4: 1–13. https://doi.org/10.3389/fgene.2013.00098

Porth I and El-Kassaby YA (2014) Assessment of the genetic diversity in forest tree populations using molecular markers. *Diversity*. 6: 283–295. https://doi.org/10.3390/d6020283

Read H (2013) Veteran Trees: A guide to good management. Veteran trees: A guide to good management. English Nature, Peterborough.

Reisch C, Mayer F, Rüther C and Nelle O (2007) Forest history affects genetic diversity - Molecular variation of *Dryopteris dilatata* (Dryopteridaceae) in ancient and recent forests. *Nord J Bot*. 25: 366–371. https://doi.org/10.1111/j.0107-055x.2008.00188.x

Ruzza DAC, Rossi AAB, Bispo RB, Tiago AV, et al. (2017) The genetic diversity and population structure of *Genipa americana* L. in Northern Mato Grosso, Brazil. *Genet. Mol. Res.* 17: 1-9.

Short FT, Polidoro B, Livingstone SR, Carpenter KE, et al. (2011) Extinction risk assessment of the world's seagrass species. *Biol Conserv.* 144: 1961–1971.

Tabbasam N, Zafar Y, Mehboob-ur-Rahman (2014) Pros and cons of using genomic SSRs and EST-SSRs for resolving phylogeny of the genus Gossypium. *Plant Syst Evol*. 300: 559–575. https://doi.org/10.1007/s00606-013-0891-x

Thompson I, Mackey B, McNulty S and Mosseler A (2009) Forest resilience, biodiversity, and climate change. Secretariat of the Convention on Biological Diversity, Montreal.

Van Oosterhout C, Hutchinson WF, Wills DPM and Shipley P (2004) MICRO-CHECKER: Software for identifying and correcting genotyping errors in microsatellite data. *Mol Ecol Notes* 4: 535–538. https://doi.org/10.1111/j.1471-8286.2004.00684.x

Wickham H (2009) ggplot2 elegant graphics for data analysis. 2edn. Media. Springer Publishing Company, New York. https://doi.org/10.1111/j.1541-0420.2011.01616.x

Zhang S, Zhang L, Zhao Z, Li Y, et al. (2016) Root transcriptome sequencing and differentially expressed drought-responsive genes in the *Platycladus orientalis* (L.). *Tree Genet Genomes.* 12: 1-15. https://doi.org/10.1007/s11295-016-1042-7

Zhao J, Xu Y, Ding Q, Huang X, et al. (2016) Association mapping of main tomato fruit sugars and organic acids. *Front Plant Sci.* 7: 1–11. https://doi.org/10.3389/fpls.2016.01286

Zhao Y, Ruan C J, Ding G J and Mopper S (2017) Genetic relationships in a germplasm collection of *Camellia japonica* and *Camellia oleifera* using SSR analysis. *Genet. Mol. Res.* 16: r16019526. https://doi.org/10.4238/gmr16019526

Zhu L and Lou A (2012) Development and characterization of nine highly polymorphic microsatellite primers for *Platycladus orientalis* (Cupressaceae). *American Journal of Botany* :99 https://doi.org/10.3732/ajb.1100583

Zhu L and Lou A (2013) Old-growth *Platycladus orientalis* as a resource for reproductive capacity and genetic diversity. *PLoS One.* 8: e56489. https://doi.org/10.1371/journal.pone.0056489

Zhu P, Wang ZF, Ye WH and Cao HL (2013) Maintenance of genetic diversity in a small, isolated population of ancient tree *Erythrophleum fordii*. *J Syst Evol.* 51: 722–730. https://doi.org/10.1111/jse.12047