Population genetic structure and variability in *Lindera glauca* (Lauraceae) indicates the low levels of genetic diversity and skewed sex ratio of natural populations in China mainland (#35773)

4th revision

Guidance from your Editor

Please submit by **5 Sep 2019** for the benefit of the authors (and your $200 publishing discount).

Structure and Criteria
Please read the 'Structure and Criteria' page for general guidance.

Raw data check
Review the raw data. Download from the materials page.

Image check
Check that figures and images have not been inappropriately manipulated.

Privacy reminder: If uploading an annotated PDF, remove identifiable information to remain anonymous.

Files
Download and review all files from the materials page.

1 Tracked changes manuscript(s)
1 Rebuttal letter(s)
4 Figure file(s)
4 Table file(s)
4 Other file(s)

Structure and Criteria

Structure your review

The review form is divided into 5 sections. Please consider these when composing your review:

1. **BASIC REPORTING**
2. **EXPERIMENTAL DESIGN**
3. **VALIDITY OF THE FINDINGS**
4. General comments
5. Confidential notes to the editor

You can also annotate this PDF and upload it as part of your review

When ready submit online.

Editorial Criteria

Use these criteria points to structure your review. The full detailed editorial criteria is on your guidance page.

BASIC REPORTING

- Clear, unambiguous, professional English language used throughout.
- Intro & background to show context. Literature well referenced & relevant.
- Structure conforms to PeerJ standards, discipline norm, or improved for clarity.
- Figures are relevant, high quality, well labelled & described.
- Raw data supplied (see PeerJ policy).

EXPERIMENTAL DESIGN

- Original primary research within Scope of the journal.
- Research question well defined, relevant & meaningful. It is stated how the research fills an identified knowledge gap.
- Rigorous investigation performed to a high technical & ethical standard.
- Methods described with sufficient detail & information to replicate.

VALIDITY OF THE FINDINGS

- Impact and novelty not assessed. Negative/inconclusive results accepted. *Meaningful* replication encouraged where rationale & benefit to literature is clearly stated.
- All underlying data have been provided; they are robust, statistically sound, & controlled.
- Speculation is welcome, but should be identified as such.
- Conclusions are well stated, linked to original research question & limited to supporting results.

Standout reviewing tips

The best reviewers use these techniques

Tip	*Example*
Support criticisms with evidence from the text or from other sources	*Smith et al (J of Methodology, 2005, V3, pp 123) have shown that the analysis you use in Lines 241-250 is not the most appropriate for this situation. Please explain why you used this method.*
Give specific suggestions on how to improve the manuscript	*Your introduction needs more detail. I suggest that you improve the description at lines 57- 86 to provide more justification for your study (specifically, you should expand upon the knowledge gap being filled).*
Comment on language and grammar issues	*The English language should be improved to ensure that an international audience can clearly understand your text. Some examples where the language could be improved include lines 23, 77, 121, 128 – the current phrasing makes comprehension difficult.*
Organize by importance of the issues, and number your points	*1. Your most important issue* *2. The next most important item* *3. …* *4. The least important points*
Please provide constructive criticism, and avoid personal opinions	*I thank you for providing the raw data, however your supplemental files need more descriptive metadata identifiers to be useful to future readers. Although your results are compelling, the data analysis should be improved in the following ways: AA, BB, CC*
Comment on strengths (as well as weaknesses) of the manuscript	*I commend the authors for their extensive data set, compiled over many years of detailed fieldwork. In addition, the manuscript is clearly written in professional, unambiguous language. If there is a weakness, it is in the statistical analysis (as I have noted above) which should be improved upon before Acceptance.*

Population genetic structure and variability in *Lindera glauca* (Lauraceae) indicates the low levels of genetic diversity and skewed sex ratio of natural populations in China mainland

Biao Xiong [Corresp., 1], **Limei Zhang** [2], **Shubin Dong** [2], **Zhixiang Zhang** [Corresp. 2]

[1] College of Tea Science, Guizhou University, Guiyang, China

[2] Beijing Forestry University, Beijing, China

Corresponding Authors: Biao Xiong, Zhixiang Zhang
Email address: bxiong@gzu.edu.cn, zxzhang@bjfu.edu.cn

Lindera glauca (Lauraceae) is a tree of economic and ecological significance reproducing sexually and asexually by apomictic seeds. It is widely distributed in low-altitudinal montane forests of East Asia. Despite the potential implications of a mixed reproductive system in terms of genetic diversity, few studies have focused on this aspect. In this study, the genetic structure of wild populations of *L. glauca* was investigated by a package of genetic analyses. Overall, 13 nuclear microsatellites (nSSRs) and 5 chloroplast microsatellites (cpSSRs) were used to genotype 300 individual plants, taken from 20 wild and 2 cultivated populations ranging across nearly the entire natural distribution in mainland China. The populations exhibited low levels of genetic diversity (nSSR: A_R = 1.75, *Ho* = 0.32, *He* = 0.36; cpSSR: *Nb* = 2.01, *Hrs* = 0.40) , and there was no significant effect of isolation by distance between populations, irrespective of marker type (nSSR: R^2 = 0.0401, *P* = 0.068; cpSSR: R^2 = 0.033, *P* = 0.091). Haplotype networks showed complex correlations among populations, and the H12 haplotype was predominant in most populations. Analyses of molecular variance observed with nuclear (*F*sc = 0.293, F_{ST} = 0.362) and chloroplast markers (*F*sc=0.299, F_{ST}=0.312) were similar. The migration ratio of pollen flow versus seed flow in this study was negative (*r*=−1.149). These results suggest that weak barriers to dispersal between populations and/or the similarity of founders shared between neighboring and distant populations, and gene flow between populations is more likely to involve seed. We infer that wild *L. glauca* in mainland China had highly skewed sex ratios with predominant females. In additional, some populations experienced a recent bottleneck effect, especially in Gujianshan, Chongqing, southwest China (population GJS). We suggest that few wild male individuals should be conserved to maintain overall genetic diversity in this species. Our findings provide important information for the sustainable utilization and preservation of the overall genetic diversity

of *L. glauca*.

Population genetic structure and variability in *Lindera glauca* (Lauraceae) indicates the low levels of genetic diversity and skewed sex ratio of natural populations in China mainland

Biao Xiong[1,2*], Limei Zhang[2], Shubin Dong[2], and Zhixiang Zhang[2]

[1] College of Tea Science, Guizhou University, Guiyang, China

[2] Beijing Forestry University, Beijing, China

Correspondence Author

Zhixiang Zhang, zxzhang@bjfu.edu.cn

No.35 Tsinghua East Road, Beijing 100083, China

Biao Xiong, bxiong@gzu.edu.cn

Jiaxiu South Road, Guiyang, Guizhou, 550025, China

Abstract

Lindera glauca (Lauraceae) is a tree of economic and ecological significance reproducing sexually and asexually by apomictic seeds. It is widely distributed in low-altitudinal montane forests of East Asia. Despite the potential implications of a mixed reproductive system in terms of genetic diversity, few studies have focused on this aspect. In this study, the genetic structure of wild populations of *L. glauca* was investigated by a package of genetic analyses. Overall, 13 nuclear microsatellites (nSSRs) and 5 chloroplast microsatellites (cpSSRs) were used to genotype 300 individual plants, taken from 20 wild and 2 cultivated populations ranging across nearly the entire natural distribution in mainland China. The populations exhibited low levels of genetic diversity (nSSR: A_R = 1.75, *Ho* = 0.32, *He* = 0.36; cpSSR: *Nb* = 2.01, *Hrs* = 0.40) , and there was no significant effect of isolation by distance between populations, irrespective of marker type (nSSR: R^2 = 0.0401, *P* = 0.068; cpSSR: R^2 = 0.033, *P* = 0.091). Haplotype networks showed complex correlations among populations, and the H12 haplotype was predominant in most populations. Analyses of molecular variance observed with nuclear (*F*sc = 0.293, F_{ST} = 0.362) and chloroplast markers (*F*sc=0.299, F_{ST}=0.312) were similar. The migration ratio of pollen flow versus seed flow in this study was negative (*r*=−1.149). These results suggest that weak barriers to dispersal between populations and/or the similarity of founders shared between neighboring and distant populations, and gene flow between populations is more likely to involve seed. We infer that wild *L. glauca* in mainland China had highly skewed sex ratios with predominant females. In additional, some populations experienced a recent bottleneck effect, especially in Gujianshan, Chongqing, southwest China (population GJS). We suggest that few wild male individuals should be conserved to maintain overall genetic diversity in this species. Our findings provide important information for the sustainable utilization and preservation of the overall genetic diversity of *L. glauca*.

Keywords: *Lindera glauca*, genetic structure, apomixis, SSR marker, gene flow

Introduction

Plant populations respond to changing environment and climate by phenotype shifts (Nicotra et al., 2010) and, mainly through sexual reproduction, by grouping high-fitness alleles that reside in different individuals (Whitton et al., 2008). In general, sexual reproduction is predominant in eukaryotes and a nearly universal characteristic of angiosperms. In a number of groups, however, some sexual plants can reproduce asexually by apomixis, which is the production of clonal seed in the absence of fertilization (Richards, 1986; Daniel et al., 2001), producing exact genetic replicas of maternal plants (Daniel et al., 2001). Apomixis occurs in fewer than 1% of flowering plant species, with an uneven distribution among lineages (Whitton et al., 2008). Among angiosperms, apomixis occurs sporadically (APG, 2003) (Whitton et al., 2008). In some genera (i.e., *Taraxacum*), apomictic clones could be widely distributed and are temporarily ecologically successful (van Dijk, 2003; Majesky et al., 2012). However, lack of diversity, the limited possibility of acquiring heritable variability (Richards, 1996), and an increased mutation load leading to the extinction of clones (van Dijk, 2003), give apomicts an adaptive disadvantage. In contrast, apomixis have lower reproductive costs over sexuals, high proportion of loci fixed in heterozygous conditions, and significant advantages over sexuals in colonizing new areas (Majesky et al., 2012). Because of these short-term advantages, natural populations of apomicts have become of interest for agricultural development.

There are two major types of apomixis, adventitious embryony and gametophytic apomixis that differ in the way embryos are formed (Whitton et al., 2008; Lo et al., 2009). The origin of former is somatic tissue surrounding ovule that must be fertilized and the latter one is an unreduced megagametophyte. Adventitious embryony is widely distributed in nature, and gametophytic apomixis is reported in a few families, e.g. Asteraceae, Poaceae, and Rosaceae. There are two ways for them to spread over space, direct dispersal via apomictic seed and indirect transmission via pollen (Whitton et al., 2008). For indirect transmission via pollen, the genes for maternal clonality

can be transmitted via male gametes, and this mode of transmission may well be important in the establishment and spread of apomixis (Brock, 2004; Preite et al., 2015). Therefore, the transmission of apomixis genes to sexuals via pollen may be of long-term importance for the spread of apomixis, especially for an agriculturally important tree such as *Lindera glauca*.

Lindera glauca (Sieb. et Zucc.) Blume (Lauraceae), a deciduous shrub or small tree with both apomixis (asexual reproduction by seeds) (Dupont, 2002) and sexual reproduction system (Tsui et al., 1982; Tsui et al., 2008), is extensively distributed in the low-altitudinal montane forests of central and southern China mainland, as well as in Japan, Korea, Vietnam and Taiwan island of China (Wang, 1972; Chang, 1976; Zheng,1983). One of the main trees making up the shrubbery and young forest ecosystems in the central and southern areas of China mainland, this species has both economic value and ecological importance. Its fruits are rich in fatty acids and aromatic oils, and contain terpenoids, flavonoids, and alkaloids, which are various applications in traditional medicine. The fruits are used as raw materials to produce medicines, lubricants, and biochemical products (Zheng, 1983; Wang et al., 1994; Kim et al., 2014; Suh et al., 2015; Qi et al., 2016). Some root extract components, like N-methyllaurotetanine, exhibit significant anti-tumor metastatic activity (Kim et al., 2014; Suh et al., 2015) and some volatile oils from the leaves are used in the industrial production pf spices (Qi et al., 2016). Additionally, *L. glauca* species has emerged as a novel potential source of biodiesel in China due to the high quality and quantity of its fruit oil (Lin et al., 2017; Xiong et al., 2018a). There has been increased scientific interest in the species, but we still know relatively little about its reproductive modes and their potential effects on genetic diversity in population dynamics and population differentiation.

L. glauca is native to the China mainland, diploid (2n=24) (Yang, 1999), with sexual reproduction and male plants have been known to exist in continental East Asia for several decades (Wang et al., 1972; Tsui et al., 1982;

Tsui et al., 2008).In a study conducted in Japan, Dupont (2002) found that female *L. glauca* could asexually reproduce by seed. Adult population sex ratios of other *Lindera* species observed in Japan ranged from equal to a strong male deviation (including *L. obtusiloba*, *L. umbellata*, and *L. erythrocarpa*) (Dupont, 2002). However, recent empirical studies revealed how males of *L. glauca* are very rare in China mainland, with females reproducing by apomixis. This indicates that natural populations have a mixed reproduction mode including apomixis and sexual propagation. Apomixis might play a major role in shaping the genetic structure of the species, by limiting gene flow within populations (Daniel et al., 2001). Interpopulation gene flow in plants is mediated by a combination of pollen and seed dispersal (Ennos, 1994). Some natural populations of apomicts retain residual sexual function as pollen donors and thus have the potential to spread apomixis via male gametes, thereby increasing the genetic diversity observed within apomictic populations (Whitton et al., 2008). Dioecious *L. glauca* has bisexual or functionally unisexual flowers(Tsui et al., 2008). and our results indicates that there was a very small amount of vital pollen from the staminode of female flowers (By 2, 3, 5-Triphenyltetrazolium chloride, or TTC method) (Hu, 1993), implying potential for natural pollination. Thus, the genetic diversity and structure of natural populations of *L. glauca* may well be more complex than previously thought. It is essential to study the gene flow and estimate the relative rates of pollen and seed migration among natural populations. Furthermore, population bottleneck effect is thought to be responsible for the very low levels of genetic variation found in a number of species that now have large population sizes (Pannell, 2013). Given that there are very few males of *L. glauca* in the last decade and many males existed on Taiwan island of China (Zhang, 2007), whether natural populations in China mainland experienced a bottleneck effect or not?

In the present study, we aimed to (1) investigate populations genetic diversity of *L. glauca* in the China mainland, (2) detect genetic variation within and differentiation among natural populations, (3) assess the relative

96 importance of pollen and seeds as agents of gene flow, (4) determine whether natural populations experienced a
97 decline in size (bottleneck effect). Molecular genetic analyses were performed and individuals in 20 wild
98 populations (and 2 cultivated populations) of *L. glauca* were genotyped using 13 nuclear and 5 chloroplast
99 microsatellite markers developed in our previous work (Xiong et al., 2016; Xiong et al., 2018b).

100 ## Materials and Methods

101 ### Sample Collection

102 During field expeditions carried out from 2013 to 2017, 300 individuals were sampled from 20 wild and 2
103 cultivated populations, representing nearly the entire natural distribution of *L. glauca* in the China mainland (**Table**
104 **1; Fig. 1a**). Most *L. glauca* individuals are able to form clones by vegetative reproduction via stolons (Tsui et al.,
105 1982), as found in a field survey. To avoid the collection of several ramets from the same genet, a single sample was
106 obtained from each cluster of shrubs in close proximity to or from a main tree, excluding surrounding young
107 branches growing on the ground. In some smaller populations, fewer than 10 plants of putatively nonclonal origin
108 were available. Overall sample sizes varied from 5 to 30 per wild population (**Table 1**). In the field, fresh leaves
109 were immediately dried in silica gel after collection, and preserved until DNA extraction.

110 ### DNA Extraction and Microsatellite Genotyping

111 Genomic DNA was extracted from 100–150 mg of dried leaves per sample using a modified
112 cetyltrimethylammonium bromide (CTAB) method (Doyle et al., 1987). Microsatellite loci of all 300 individuals of
113 *L. glauca* were screened, including 13 polymorphic nuclear microsatellite markers (EST-based microsatellites;
114 hereafter nSSRs; Table S1) and 5 polymorphic chloroplast microsatellite markers (hereafter cpSSRs; Table S2). All
115 nSSRs and cpSSRs were labeled with fluorescently labeled nucleotides (forward primer with M13F) and detected by

capillary gel electrophoresis. Subsequent\ steps and the PCR assay were according to Xiong et al. (2016). Genotyping was performed using an ABI 3730XL DNA Analyzer (Applied Biosystems, California, USA) with a GeneScan 500 LIZ Size Standard, and alleles for each locus were manually scored using GeneMarker version 2.2.0 software (SoftGenetics, State College, PA, USA).

Data Analyses

Raw data matrices containing the alleles and haplotypes information for 13 nSSR and 5 cpSSR loci were checked for scoring errors. All SSR analyses were conducted with 300 samples. Data editing and formatting were performed using GenAlEx v. 6.502 (Smouse et al., 2015).

The related indexes of genetic diversity were calculated as followed. For the nSSR data set, genetic diversity indices, including the number of alleles (N_A), observed heterozygosity (*Ho*), expected heterozygosity (*He*), percentage of polymorphic loci (PPB), Wright's inbreeding coefficient (F_{IS}), and Nei's (Nei, 1978) genetic distances, were estimated using GenAlEx v. 6.502 (Smouse et al., 2015) and POPGENE v. 1.32 (Yeh et al., 1999). The online package GENEPOP v. 4.1.4 (Rousset, 2008)was used to perform exact Hardy–Weinberg equilibrium (*HWE*) tests and to test for the presence of null alleles. The di□erentiation index F_{ST} was computed for pairs of populations using Arlequin v. 3.5.1.3 (Exco□er et al., 2005). Allelic richness (A_R) was calculated using the software FSTAT v. 1.2 (Goudet, 1995). For the cpSSR data set, number of haplotypes (*Nb*), genetic diversity (*Dv*), haplotype richness (*Hrs*), the number of private alleles (*Prv*), and the polymorphism information content (PIC) per locus were estimated using HAPLOTYPE v. 1.05 (Eliades and Eliades, 2009).

The analysis of population genetic structure was as follows. For the nSSR data set, the genetic structure of the 22 populations (20 wild and 2 cultivated populations) was analyzed using the Bayesian clustering approach implemented in STRUCTURE v. 2.3.4 (Pritchard et al., 2000), assuming an admixture model. To determine the

most appropriate number of genetic clusters or groups (K value), K was set from 1 to 20, and the analysis was run 20 times for each K with a burn-in of 1,000,000 generations followed by 50,000 generations for the Markov chain Monte Carlo (MCMC) simulation. The admixture level for each individual (Q) was also inferred. The program STRUCTURE HARVESTER v. 0.6.94 (Earl and von Holdt, 2012) was used to estimate the number of population clusters based on the ΔK parameter according to Evanno et al., (2005). Based on the most appropriate number of clusters suggested by Bayesian clustering, analysis of molecular variance (AMOVA) was performed using Arlequin, with 10,000 iterations for the permutation test. A neighbor-joining (NJ) tree was generated using POPTREE2 (Takezaki et al.,2010) based on pairwise Nei (1978) genetic distances between populations determined by GenAlEx. For the cpSSR data set, the Arlequin was used to determine pairwise F_{ST} values among all populations. A parsimony network illustrating genetic relationships among haplotypes of *L. glauca* populations was generated using TCS v.1.1 (Clement et al., 2000).

Considering that isolation by distance (IBD) can be a key factor keeping populations apart by limiting gene flow (Coyne and Orr, 2004), the IBD of wild *L. glauca* interpopulation in China mainland was teste. In view of the potential importance of pollen in the spread of apomixis (Mogie, 1992; Whitton et al., 2008), the pollen/seed migration ratio (*r*) was calculated. To examine IBD, the Mantel test was performed using GenAlEx, correlating the pairwise genetic distances [F_{ST} / (1− F_{ST})] with the pairwise geographic distances (in kilometers). To calculate *r*, we used the followed formula: $r = mp/ms = [(1/F_{ST}(n) - 1)(1+ F_{IS}) - 2(1/F_{ST}(cp) - 1)] / (1/F_{ST}(cp) - 1)$ (Ennos, 1994; Petit et al., 2005), where *mp* is the pollen migration rate, *ms* is the seed migration rate, F_{ST} values (overall F_{ST}) are population differentiation estimates derived from AMOVA, $F_{ST}(n)$ is the nuclear F_{ST} and $F_{ST}(cp)$ is the chloroplast F_{ST}.

Population bottlenecks were evaluated using BOTTLENECK v. 1.2.02 (Piry et al., 1990) with the infinite

158 alleles model (IAM) that a single mutation is allocated at a time and the resulting number of alleles is computed,
159 stepwise mutation model (SMM) that is a Bayesian approach and generally more appropriate when testing
160 microsatellite loci, and two-phased model (TPM) that is a modified SMM. According to Piry et al.,, sign tests,
161 Wilcoxon tests, and mode-shift were applied, excluding standardized di□erences tests, which are useful when at
162 least 20 polymorphic loci are available.

163 ## Results

164 ### Genotypic Variation

165 A total of 74 alleles and 13 haplotypes were identified at 13 nSSRs and 5 cpSSRs across 300 individuals of *L.*
166 *glauca*, respectively. For each locus, the number of alleles for 13 nSSR loci ranged from 3 (P-298) to 8 (XBLG-
167 060), with a mean of 5.7 alleles (Table S1). In particular, A_R ranged from 1.807 to 2.774, with a mean of 2.329, and
168 PIC ranged from 0.363 to 0.711, with a mean of 0.556. *Ho* and *He* varied between 0.210 and 0.563, with a mean of
169 0.380, and between 0.380 and 0.754, with a mean of 0.602, respectively.

170 The total number of alleles across 13 nSSR loci varied from 16 (population SZY) to 44 (population ZJJ), with a
171 mean of 31.5, and allelic richness ranged from 1.231 to 2.011, with a mean of 1.746 (**Table 2**). Population *Ho* and
172 *He* ranged from 0.108 to 0.708 and from 0.106 to 0.477, respectively, with means of 0.318 and 0.355. The
173 percentage of polymorphic loci ranged from 23.08% to 100%, with a mean of 81.12%. Significant deviations from
174 the *HWE* indicating a heterozygote deficiency were detected in 9 of 22 populations. For some populations (KYS,
175 WYS, ZJS, WJS, GJS, NHS, FHS, and ZJJ), negative F_{IS} values within populations were observed, indicating more
176 heterozygotes than expected. However, none of the 13 nSSR loci with heterozygote excesses when calculating it on
177 all samples (Tables S1), and no evidence of null alleles within the data set. Among all wild samples (290), 277

individual plants had a unique multi-locus pattern after PCR amplification with 13 nSSRs primers, indicating that these samples were almost from different individual. Of the remaining 13 individual plants, 5 pairs exhibited the same multi-locus pattern in pairs, and 3 individuals exhibited the same multi-locus pattern. This implies that our sampling strategy was effective in avoiding the sampling of most clones (because clones contained nearly identical genotypes).

Genetic diversity parameters for cpSSR loci are summarized in Table S2. All 5 cpSSR loci exhibited 2–3 alleles per locus across all samples. *Dv* ranged from 0.115 to 0.249 per locus, and PIC varied from 0.155 to 0.218. Analyzing combinations of all alleles, there were 22 unique haplotypes (hereafter H; **Fig. 1**). All populations contained several haplotypes, except for populations ATM, YTH, NTB, FHS, TMS, and HZY (**Fig. 1a**). The network of plastid haplotypes was complex (**Fig. 1b**). Haplotype H12 exhibited the highest frequency and was detected in 18 of 20 wild populations (Table S3). Of the 22 haplotypes, 7 were identified just once within the dataset, and were present in only 1 population each. Excluding cultivated populations, *Hrs* per population ranged from 0 (populations ATM, NTB, YTH, TMS, and FHS) to 0.964 (population FJS), and the mean genetic distance between individuals (D^2_{sh}) varied from 0 to 39.638 (population JGS) (**Table 2**).

Genetic Clustering and Population Differentiation

A Bayesian analysis based on 13 nSSRs implemented in STRUCTURE showed the presence of 2 clusters (K = 2), with only slight admixture at the individual level in each population, except for population ATM (**Fig. 2a**). The ΔK statistic developed by Evanno et al., indicated that the overall differences were not substantial (**Fig. 2b**). Cluster orange included 13 wild populations (ATM, JGS, LDZ, SJG, NTB, YTH, DBS, HMF, TMS, SQS, LYS, KYS, and WYS), and the remaining 7 wild and 2 cultivated populations were assigned to cluster blue. The NJ tree (**Fig. 3a**)

and principal coordinates analysis (**Fig. 3b**) based on the nSSR dataset supported the results of STRUCTURE analysis, indicating that 22 populations could be grouped into 2 clusters. However, the network diagram of all 22 unique plastid haplotypes revealed by 5 cpSSRs was complex (**Fig. 1b**), and failed to support the 2 distinct clusters revealed using nuclear data.

The 2 clusters revealed by STRUCTURE analysis were set as groups for AMOVA based on either the nSSR or cpSSR dataset (**Table 3**). Using the nSSR dataset, the majority of genetic variation was detected within populations (63.82%), indicating a genetic differentiation mostly at individuals level. Nevertheless, a considerable proportion of the total variation (26.42%) was found among populations within groups, and a small amount of variation (9.76%) occurred among the 2 groups. In contrast to the nSSR results, the AMOVA based on the cpSSR dataset showed that a larger proportion of genetic variation could be attributed to variation within populations (68.84%) and among populations within groups (29.37%), and little variation among groups (1.79%). The overall F_{ST} values calculated by AMOVA were 0.362 ($P \leq 0.0001$) for the nSSR dataset and 0.312 ($P \leq 0.0001$) for the cpSSR dataset.

Isolation by Distance and Pollen/Seed Migration Ratio

The estimates of genetic differentiation (F_{ST} value) based on 13 nSSRs ranged from 0.023 (between WJS and GJS) to 0.427 (between LDZ and FHS) (Table S4), excluding 2 cultivated populations. Only 4 pairwise comparisons (ATM and JGS, WJS and NHS, NTB and YTH, and WJS and GJS) showed significant *F*st values ($P \leq 0.05$). Adopting a P-value of 0.001, no significant correlation between pairwise genetic distance [$F_{ST}/(1 - F_{ST})$] and geographic distance (in kilometers) was found using the Mantel test (**Fig. 4**) for the nSSR dataset ($R^2 = 0.0401$, $P = 0.068$) or the cpSSR dataset ($R^2 = 0.033$, $P = 0.091$), suggesting that *L. glauca* in China mainland does not exhibit significant IBD. The F_{ST} values for the nSSR and cpSSR were similar (**Table 3**), and the *r* was −1.149, indicating

that most gene flow among populations occurs via seed, rather than pollen.

Population Bottleneck Effect

Several populations had a significant excess of heterozygosity expected at mutation-drift equilibrium (i.e., *He>Heq*) (Piry et al., 1990) under the 3 models in the bottleneck analysis, indicating a deviation from mutation drift equilibrium in wild *L. glauca* populations (**Table 4**). In detail, population GJS exhibited a significant bottleneck event according to the sign and Wilcoxon tests in all 3 models, indicating a population size decline (bottleneck effect) in its history. The population WJS experienced a significant bottleneck event according to the sign test and Wilcoxon test for the IAM and TPM methods, and FHS exhibited a significant bottleneck event by Wilcoxon test in all 3 models. SJG, FJS, and ZJS only exhibited a significant bottleneck event based on the sign and Wilcoxon tests for the IAM method (**Table 4**).

Discussion

Genetic Variation within Populations

In this study, the sampling covered a large portion of the natural distribution, and overall genetic diversity across *L. glauca* populations exhibited a low level based on both nSSR (mean A_R = 1.75, *Ho* = 0.32, *He* = 0.36) and cpSSR (mean *Nb* = 2.01, *Hrs* = 0.40) loci. Our estimates of genetic diversity in *L. glauca* were almost half those of long-lived perennials (*Ho* = 0.63, *He* = 0.68), out-crossing species (*Ho* = 0.63, *He* = 0.65), and plants with wide distributions (*Ho* = 0.57, *He* = 0.62) (Nybom, 2004), and were lower than *Laurus nobilis* in Lauraceae (A_R = 3.22, *He* = 0.56) (Marzouki et al., 2009). There are several major factors influencing variation that can result, each by itself or in combination, in the low levels of genetic variation observed in wild *L. glauca* populations. Asexual

237 reproduction through apomictic seeds can decrease genetic variation in a population especially in apomictic
238 populations (Lo et al., 2009). Similarity, effective population size could seem limitation, due the population being
239 established from a limited number of individuals, and small sampling quantity. Furthermore, the estimated values of
240 genetic diversity herein are also lower than nSSR-based values found in literature (A_R = 2.61, *He* = 0.44) (Zhu et al.,
241 2016).Differences in genetic variation between this study and the previous one are likely to be explained by number
242 of sampling populations and individuals (previous study included 6 populations, a total of 96 individual plants, and
243 this study included 20 wild and 2 cultivated populations, a total of 300 individual plants).

244 On the other hand, a marked similarity in the molecular variance revealed by the 2 types of markers (overall
245 *F*sc = 0.293 and F_{ST} = 0.362 for nSSRs; 0.299 and 0.312 for cpSSRs, respectively) was observed, indicating general
246 consistency between chloroplast and nuclear DNA. In detail, predominant apomixes in wild *L. glauca* could explain
247 that F_{ST} observed for nuclear markers is a little higher than that for chloroplast markers, and vice versa. In addition,
248 apomictic (asexual) reproduction of *L. glauca* could also affect the results, because that means many individuals
249 have the identical genome. The use of clone-corrected data (removing data of clones from the same parent) is
250 necessary if clones are detected, because *L. glauca* tree could form clones by vegetative reproduction via stolons.

251 There were 9 populations that more he[ter]zygotes than expected. STRUCTURE analysis showed that the 9
252 populations were divided into one group and the rest into another, suggesting negative F_{IS} is an important factor
253 affecting group of population difference. Positive F_{IS} values within populations (JGS, LYS, and FJS) suggested
254 outcrossing and sexual reproduction did exist in these populations. Global multilocus test indicated that, for some
255 wild populations (ATM, LDZ, SJG, NTB, YTH, DBS, TMS, and SQS), F_{IS} values deviated significantly from
256 HWE, indicating inbreeding. Considering that clonality probably generates significant negative F_{IS} in some wild
257 plant populations with asexual reproduction (Stoeckel et al., 2006), the observed negative F_{IS} of wild populations

(KYS, WYS, ZJS, WJS, GJS, NHS, FHS, and ZJJ), coupled with the result grouped by NJ tree (blue cluster) (**Fig. 3a**), suggested that apomixis (asexual reproduction) may have been common in these populations. Asexual reproduction among individuals and generations could result in heterozygote excesses, thus explaining the F_{IS} pattern (Balloux, 2004; Stoeckel et al., 2006).

Genetic Differentiation among Populations

Populations often cluster according to habitat type or geographic distance. However, for species with predominantly asexual populations, like *Daktulosphaira vitifoliae* (Vorwerk and Forneck, 2006), *Crataegus douglasii* (Lo et al., 2009), and *Taraxacum officinale* (Majesky et al., 2012), there is no significant correlation between genetic distances and geographic distances. In this study, a correlation between genetic distances (as measured by $[F_{ST} / (1- F_{ST})]$ values) and geographic distances (in kilometers) was not detected in *L. glauca* populations, irrespective of marker type, suggesting that weak barriers to dispersal between populations and/or the similarity of founders shared between neighboring and distant populations, and apomixis within populations did not completely limit gene flow. Our results indicate that sexual production and apomixis co-occurred in the identical natural population. However, further research is needed to investigate the extent to which apomixis limits gene flow, and the exact rate of sexual production and apomixis occurred among and within populations.

According to Ennos' formula (1994), the migration ratio of pollen flow versus seed flow (r) in this study was negative (−1.149), suggesting that gene flow between populations is more likely to involve seed. The distorted r value could be explain by reproductive mode in different populations almost is apomict seeds of females, coupled with the result that no male was found at field observation sites for 5 consecutive years. Considering that there is likely to have a situation where a few individuals can self-pollinate, vital pollen from staminode of female flowers,

278 even a very few, may associated with the negative *r* value of pollen flow versus seed flow. However, because the
279 formula is derived for a hermaphrodite species, it needs to be modified to account for the disproportionate maternal
280 contribution from the females to the next generation if the exact *r* value of *L. glauca* would expect to be observed.

281 **Source of Evolutionary Potential of Apomicts**

282 Although 22 haplotypes were observed across 20 wild and 2 cultivated populations of *L. glauca*, the H12
283 accounted for 62.03% of overall haplotypes and was detected in all populations (Table S3), except for FHS, LYS,
284 SZY, and HZY. This haplotype existing in many populations separated by considerable geographical distances (e.g.,
285 greater than 1,720 km between KYS and GJS), coupled with the complex network (**Fig. 1b**), suggested three
286 inferences to account for the observed result. First, these individuals from different populations had a relatively
287 recent same female founder. Second, there was a very robust maternal (apomictic) lineage for this species within the
288 population. Third, the genome of this species has genetic traits controlling apomixis, which might be induced by
289 some factors (e.g., biological stimulation, environment influence, or climate changing, etc.) and coexist with sexual
290 reproduction in identical individual plant.

291 The first inference could be explained by a hypothesis that the H12 haplotype may be associated with the
292 migratory patterns of some birds responsible for the dispersion of apomictic seeds over a long distance. Despite
293 conducting field observation for 5 consecutive years, we found few birds eating grown fruits of *L. glauca*, and few
294 small mammals (e.g., *Paguma larvata*). The second inference means having an early maternal ancestor through
295 apomictic reproduction for many individuals. However, all remaining angiosperm apomicts are at the tips of the tree
296 of life (APG, 2003, 2016; Horandl, 2006; Thompson and Whitton, 2006; Lo et al., 2009), with no other higher-level
297 asexual taxa (families, genera). Consequently, it is almost impossible to have ancient asexual angiosperms, because

asexuals fail to maintain sex and recombination in populations that are limited in size, thereby inability to bring together high-fitness alleles that reside in different individuals. Third inference means that apomicts of *L. glauca* may be of very recent origin and have the ability to apomictic reproduce through mutations or losses of some sexual genes. A similar situation exists in some species, such as *Taraxacum officinale* that exhibits alternation of asexual and sexual histories of apomicts (Majesky et al., 2012), some hawthorns (*Crataegus*; Rosaceae) that have a population genetic structure of mixed diploid sexual and polyploid apomicts (Lo et al., 2009), and a marbled crayfish (*Procambarus virginalis*) that reproduce through parthenogenesis (Gerhard et al., 2003; Ewen, 2018). Thus, this haplotype (H12) is probably a predominant genotype existing in individuals that have the ability to apomixis.

In conclusion, we reject the former two inferences that most populations across such long geographical distances have a relatively recent founder or an ancient apomictic ancestor, and believe the last inference that apomixis caused by mutations or losses of related genes makes this maternal haplotype (H12) present in many individuals from different populations more likely.

Conjecture about Histories of Natural Populations

According to the relationship $Ne = 4(NmNf)/ (Nm + Nf)$, where *Nm* represents the number of males, and *Nf* represents the number of females (Beerli and Palczewski, 2010), we failed to calculate the accurate values of *Ne* because there are no males in the samples collected. Even so, based on the rarity of male individuals in the China mainland, dioecious reproduction reported in the past several decades (Tsui et al., 1982; Wang, 1972), some specimens of branches male individuals stored in the China National Herbarium (PE), and many males existed on Taiwan island of China (Zhang, 2007), we inferred that some natural populations of *L. glauca* recently experienced a severe bottleneck and male individuals experienced a decline, most likely resulting from human events. However,

318 to test the above hypotheses and obtain more accurate results, further field investigations are necessary, including
319 samples from male and female individuals in a population (especially males found on Taiwan island), correcting for
320 clone and apomictic reproduction, across the full range of *L. glauca*, including Japan, South Korea, and Taiwan.

321 **Implications for Conservation**

322 Genetic diversity is recognized as an important population attribute for both conservation and evolutionary
323 purposes (Cena et al., 2006). The purpose of conservation of endangered and threatened species is to maintain their
324 contribution to overall genetic diversity. People usually focus only on endangered species and provide protection,
325 whereas some species that are reduced in genetic diversity also require protection. This study detected a lower level
326 of genetic diversity in *L. glauca* than that of some other species in Lauraceae. We conclude that wild *L. glauca*
327 populations have female-skewed sex ratios, consistent with our field survey and sampling for 5 consecutive years.
328 The destruction of habitats to plant other commercial or medicinal crops and felling by local farmers may explain
329 the low frequency of male *L. glauca* in China mainland, although the species is common and widely distributed in
330 other regions. We proposed to find male individuals and promote sexual reproduction to maintain overall genetic
331 diversity in this species.

332 ## Conclusion

333 Our study has shown low levels of genetic diversity in *L. glauca* across nearly the entire natural distribution in
334 the China mainland. A complex correlation among populations was showed by haplotype networks. Genetic
335 structure within and among populations was similar at the nuclear and chloroplast levels. Furthermore, some
336 populations experienced a recent bottleneck, and gene flow between populations is more likely to involve seed. This
337 implies that wild *L. glauca* in China mainland have highly skewed sex ratios with predominant females.

338

Acknowledgements

The authors thank Richard A. Ennos for his explanation of the formula of pollen/seed migration ratio. The authors thank reviewers Anonymous and Alejandra Lobo for their constructive suggestions about original manuscript. The authors thank Rodrigo C. Gonçalves-Oliveira for his assistance with AMOVA analysis, and Jigongshan National Nature Reserve for a part of the material collection.

References

APG (The Angiosperm Phylogeny Group). 2003. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG II. *Botanical Journal of the Linnean Society*, **141**:399–436. DOI 10.1111/j.1095-8339.2009.00996.x

APG (The Angiosperm Phylogeny Group). 2016. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. *Botanical Journal of the Linnean Society*, **181**:1–20. DOI 10.1111/boj.12385

Balloux F. 2004. Heterozygote excess in small populations and the heterozygote-excess effective population size. *Evolution* **58**: 1891–1900. DOI 10.1554/03-692

Beerli P., Felsenstein J. 1999. Maximum likelihood estimation of migration rates and population numbers of two populations using a coalescent approach. Genetics 152, 763–773. DOI 10.1017/S0016672399003766

Beerli P, Palczewski M. 2010. Unified framework to evaluate panmixia and migration direction among multiple sampling locations. *Genetics* **185**: 313–326. DOI 10.1534/genetics.109.112532

Brock M. 2004. The potential for genetic assimilation of a native dandelion species, *Taraxacum ceratophorum* (Asteraceae), by the exotic congener *T. officinale*. *American Journal of Botany*, **91**: 656–663. DOI 10.3732/ajb.91.5.656

Cena CJ, Morgan GE, Malette MD, Heath DD. 2006. Inbreeding, outbreeding and environmental effects on genetic diversity

in 46 walleye (*Sander vitreus*) populations. *Molecular Ecology* 15: 303–320. DOI 10.1111/j.1365-294X.2005.02637.x

Chang CE. 1976. Volume 2: Lauraceae. In: Liu TS, eds. *Flora of Taiwan*. Taipei: Epoch press, 406–468 pp.

Clement M, Posada D, Crandall KA. 2000. TCS: A computer program to estimate gene genealogies. *Molecular Ecology* **9**: 1657–1659.

Cornuet JM, Luikart G. 1996. Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics* **144**: 2001–2014.

Coyne JA, Orr HA. 2004. Speciation. Sinauer Associates, Sunderland, Mass.

Daniel G, Olivier L, Enrico P, Ueli G. 2001. Developmental genetics of gametophytic apomixis. *Trends in Genetics* **17**: 597–604. DOI 10.1016/S0168-9525(01)02454-4

Doyle JJ, Doyle JL. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin* **19**: 11–15.

Dupont YL. 2002. Evolution of apomixis as a strategy of colonization in the dioecious species *Lindera glauca* (Lauraceae). *Population Ecology* **4**: 293–297. DOI 10.1007/s101440200033

Earl DA, von Holdt BM. 2012. Structure Harvester: a website and program for visualizing structure output and implementing the evanno method. *Conservation Genetics Resources* **4**: 359–361. DOI 10.1007/s12686-011-9548-7

Eliades NG, Eliades DG. 2009. HAPLOTYPE ANALYSIS: software for analysis of haplotype data. Distributed by the authors, Forest Genetics and Forest Tree Breeding, Georg-August University Goettingen, Germany.

Ennos RA. 1994. Estimating the relative rates of pollen and seed migration among plant populations. *Heredity* **72**: 250–259. DOI 10.1038/hdy.1994.35

Evanno G, Regnaut S, Goudet J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* **14**: 2611–2620. DOI 10.1111/j.1365-294X.2005.02553.x

Ewen C. 2018. Geneticists unravel secrets of super-invasive crayfish. *Nature*, **554**: 157–158. DOI 10.1038/d41586-018-01624-y

Exco□er L, Laval G, Schneider S. 2005. Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics* **1**: 47–50.

Gerhard S, Anke B, Laura T, Reimann A, Mittmann B, Lukhaup C, Steuerwald F, Vogt G. 2003. Ecology: Parthenogenesis in an outsider crayfish. *Nature*, 4216925:806. DOI 10.1038/421806a

Gonçalves-Oliveira RC, Wöhrmann T, Benko-Iseppon AM, Krapp F, Alves M, Wanderley MGL, Weising K. 2017. Population genetic structure of the rock outcrop species *Encholirium spectabile* (Bromeliaceae): The role of pollination vs. seed dispersal and evolutionary implications. *American Journal of Botany* **104**: 868. DOI 10.3732/ajb.1600410

Goudet J. 1995. FSTAT (version 1.2): A computer program to calculate F-statistics. *Journal of Heredity* **86**: 485–486. DOI 10.1093/oxfordjournals.jhered.a111627

Horandl E. 2006. The complex causality of geographical parthenogenesis. *New Phytologist* **171**:525–538. DOI 10.1111/j.1469-8137.2006.01769.x

Hu SY. 1993. Experimental methods in plant embryology (Ⅰ) determination of pollen viability. *Chinese Bulletin of Botany* **10**: 60–62.

Kim KH, Moon E, Ha K, Suh WS, Kim HH, Kim SY, Choi SU, Lee KR. 2014. Bioactive lignan constituents from the twigs of *Lindera glauca*. *Chemical & Pharmaceutical Bulletin* **62**: 1136–1140. DOI 10.1248/cpb.c14-00381

Lin ZX, An JY, Wang J, Niu J, Ma C, Wang LB, Yuan GS, Shi LL, Liu LL, Zhang JS, Zhang ZX, and Lin SZ. 2017. Integrated analysis of 454 and Illumina transcriptomic sequencing characterizes carbon flux and energy source for fatty acid synthesis in developing *Lindera glauca* fruits for woody biodiesel. *Biotechnology for Biofuels* **10**: 134. DOI 10.1186/s13068-017-0820-2

Lo YYE, Stefanovic S, Dickinson TA. 2009. Population genetic structure of diploid sexual and polyploid apomictic hawthorns

(*Crataegus*; Rosaceae) in the Pacific Northwest. *Molecular Ecology* **18**: 1145–1160 DOI 10.1111/j.1365-294X.2009.04091.x

Majesky L., Vasut RJ., Kitner M., Travnicek B. 2012. The pattern of genetic variability in apomictic clones of *Taraxacum officinale* indicates the alternation of asexual and sexual histories of apomicts. *PLoS ONE* **7**: e41868. DOI 10.1371/journal.pone.0041868

Marzouki H, Nasri N, Jouaud B, Bonnet C, Khaldi A, Bouzid S, Fady B. 2009. Population genetic structure of *Laurus nobilis* L. inferred from transferred nuclear microsatellites. *Silvae Genetica* **58**: 5–6.

Mogie M. 1992. The evolution of asexual reproduction. London: Chapman &Hall, 276 pp.

Nei M. 1978. Estimation of average heterozygosity and genetic distance for small number of individuals. *Genetics* **89**: 583–590.

Nicotra AB, Atkin OK, Bonser SP, Davidson AM, Finnegan EJ, Mathesius U., Poot P., Purugganan MD., Richards CL., Valladares F., van Kleunen M. 2010. Plant phenotypic plasticity in a changing climate. *Trends in Plant Science* **15**: 684–692. DOI 10.1016/j.tplants.2010.09.008

Nybom H. 2004. Comparison of di□erent nuclear DNA markers for estimating intraspecific genetic diversity in plants. *Molecular Ecology* **13**: 1143–1155. DOI 10.1111/ j.1365-294X.2004.02141.x

Pannell JR. 2013. Bottleneck Effect. *Brenner's Encyclopedia of Genetics* 362–365. DOI 10.1016/b978-0-12-374984-0.00164-9

Petit RJ, Duminil J, Fineschi S, Hampe A, Salvini D, Vendramin GG. 2005. Invited review: comparative organization of chloroplast, mitochondrial and nuclear diversity in plant populations. *Molecular Ecology* **14**: 689–701. DOI 10.1111/j.1365-294X.2004.02410.x

Piry S, Luikart G, Cornuet JM. 1990. BOTTLENECK: a computer program for detecting recent reductions in the e□ective size using allele frequency data. *Journal of Heredity* **90**: 502–503. DOI 10.1093/jhered/90.4.502

Preite V., Snoek LB., Oplaat C., Biere A., van der Putten WH., Verhoeven KJ. 2015. The epigenetic footprint of poleward

range-expanding plants in apomictic dandelions. *Molecular Ecology* **24**: 4406–4418. DOI 10.1111/mec.13329

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

Qi J, Xiong B, Zhang ZX, Zhen AG, Pu FG. 2016. Diversity analysis on fatty acid composition of *Lindera glauca* fruit from different provenances. *Journal of the Chinese Cereals and Oils Association* **31**: 72–78. ([In Chinese])

Richards AJ. 1986. Plant breeding systems. London: Unwin Hyman, 529 pp.

Richards AJ. 1996. Genetic variability in obligate apomicts of the genus Taraxacum. *Folia Geobot Phytotax* **31**: 405–414.

Rousset F. 2008. GENEPOP'007: a complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources* **8**: 103–106. DOI 10.1111/j.1471-8286.2007.01931.x

Smouse PE, Whitehead MR, Peakall R. 2015. An informational diversity framework, illustrated with sexually deceptive orchids in early stages of speciation. *Molecular Ecology Resources* **15**: 1375–1384. DOI 10.1111/1755-0998.12422.

Stoeckel S, Grange J, Fernandez-Manharres F, Bilger I, Frascaria-Lacoste N, Mariette S. 2006. Heterozygote excess in a self-incompatible and partially clonal forest tree species—*Prunus avium* L. Molecular Ecology **15:** 2109–2118. DOI 10.1111/j.1365-294X.2006.02926.x

Suh WS, Kim KH, Kim HK, Choi SU, Lee KR. 2015. Three new lignan derivatives from *Lindera glauca* (Siebold et Zucc.) Blume. *Helvetica Chimica Acta* **98**: 1087–1094. DOI 10.1002/hlca.201500002

Takezaki N, Nei M, Tamura K. 2010. POPTREE2: software for constructing population trees from allele frequency data and computing other population statistics with Windows-interface. *Molecular Biology and Evolution* **27**: 747–752.

Thompson SL, Ritland K. 2006. A novel mating system analysis for modes of self-oriented mating applied to diploid and polyploid arctic Easter daisies (*Townsendia hookeri*). *Heredity* **97**:119–126. DOI 10.1038/sj.hdy.6800844

Tsui HP, Werff H. 2008. Volume 7: Lauraceae and Hernandiaceae. In: Li HW, eds. *Flora of China*. Beijing &St. Louis: Science

Press & Missouri Botanical Garden Press, 146–147 pp.

Tsui HP, Xia ZD, Li JL. 1982. Volume 31: Lauraceae and Hernandiaceae. In: Li HW, eds. *Flora of China Vol*. Beijing: Science Press, 393–394 pp.

van Dijk PJ. 2003. Ecological and evolutionary opportunities of apomixis: insights from *Taraxacum* and *Chondrilla*. *Philosophical Transactions of the Royal Society Biological Sciences* **358**: 1113–1121.

Vorwerk S, Forneck F. 2006. Reproductive mode of grape phylloxera (*Daktulosphaira vitifoliae*, Homoptera: Phylloxeridae) in Europe: molecular evidence for predominantly asexual populations and a lack of gene flow between them. *Genome* **49**:678–687. DOI 10.1139/g06-028

Wang WC. 1972. Higher plants iconography of China. Beijing: Science Press, 853 pp.

Wang Y, Gao X, Yu X, Cheng S, Kong L. 1994. Study on the resource and its utilizations of *Lindera glauca* in China. *Henan Science* **12**: 331–334. ([In Chinese])

Whitton J., Sears CJ., Baack EJ., Otto SP. 2008. The dynamic nature of apomixis in the angioperms. *International Journal of Plant Sciences*, 169: 169–182. DOI 10.1086/523369

Xiong B, Dong SB, Qi J, Zhang LM, Ha DL, Ju YX, Zhang ZX. 2016. Development and characterization of microsatellite loci for *Lindera glauca* (Lauraceae). *Applications in Plant Science* **4**: 1600088. DOI 10.3732/apps.1600088

Xiong B, Dong SB, Zhang ZX. 2018a. Biodiesel from *Lindera glauca* oil, a potential non-food feedstock in Southern China. *Industrial Crops & Products* **122**: 107–113. DOI 10.1016/j.indcrop.2018.05.026

Xiong B, Zhang LM, Xie L, Dong SB, Zhang ZX. 2018b. Complete chloroplast genome of a valuable economic tree, *Lindera glauca* (Lauraceae) and comparison with its congeners. *Pakistan Journal of Botany* **50**: 2189–2196.

Yeh FC, Yang RC, Boyle T. 1999. POPGENE Version 1.31. Microsoft window-based freeware for population genetic analysis. University of Alberta and Centre for International Forestry Research.

Zhang YM. 2007. The search for the observation of *Lindera glauca*, Lauraceae in Dadu mountain. *Taiwan Academy of Ecology*, **15**.

Zheng WQ. 1983. Dendrology of China. Beijing: Chinese Forestry Publishing House Press, 701–702 pp.

Zhu SS, Ding YQ, Yap ZY, Qiu YX. 2016. De novo assembly and characterization of the floral transcriptome of an economically important tree species, *Lindera glauca* (Lauraceae), including the development of EST-SSR markers for population genetics. *Molecular Biology Reports* **43**: 1243–1250.

Fig. 1. (A) Chloroplast haplotypes and sampling location present in *Lindera glauca* populations analyzed in the present study (see Table 1 for details). Each population is represented by a triangle, and pie charts are shown when a population was present in more than one haplotype. The green background shows the provincial-level distribution of the species in China. (B) Haplotype network was generated by program TCS. Each haplotype is represented by a single color, and circle sizes correspond to the relative frequency of a particular haplotype in the total sample.

Fig. 2. (A) Bayesian inference using STRUCTURE (K = 2) based on 13 nSSR markers from 22 populations of *L. galuca*, (B) K = 2 appeared to be the optimal number of clusters by showing the ΔK at its peak.

Fig. 3. (A) Neighbor-joining (NJ) dendrogram based on Nei's (1978) genetic distances among populations. (B) Principal coordinates analysis (PCoA) of genetic variation across 22 populations of *L. glauca* based on 13 nSSR markers.

Fig. 4. (A) Figure plot of geographical distance against genetic distance for 22 populations of *L. glauca* based on 13 nSSR markers (B) and 5 cpSSR markers.

Table 1(on next page)

Table 1 Plant material of *Lindera glauca* analyzed in the current study. Sample locations and abbreviations of 20 wild and two cultivated population s used in the main text are listed below.

Table 1 Plant material of *Lindera glauca* analyzed in the current study. Sample locations and abbreviations of 20 wild and two cultivated populations used in the main text are listed below.

Population Code	Number	Samples accession no.	Longitude (E°)	Latitude (N°)	Elevation (m)	Location
ATM	10	A14-10	115.8602778	31.2243972	646–834	Tianma, Jinzhai, Anhui
JGS	30	J13-09	114.0883389	31.8658833	203–317	Jigongshan, Xinyang, Henan
LDZ	20	L14-04	114.2575139	31.9452917	154–261	Dongzhai, Luoshan, Henan
SJG	15	S14-10	115.5421083	31.7488389	243–476	Jingangtai, Shangcheng, Henan
NTB	10	N14-04	113.4231917	32.3289139	241–256	Tongbaishan, Nanyang, Henan
YTH	10	Y14-04	115.8647444	31.0572889	647–734	Taohuachong, Yingshan, Hubei
DBS	10	D14-09	115.8391528	31.0087028	834–1003	Dabieshan, Yingshan, Hubei
HMF	10	H14-09	113.0076417	28.4467750	224–257	Heimifeng, Wangcheng, Hunan
TMS	10	T14-09	119.4495389	30.3255389	359–432	Tianmushan, Lin'an, Zhejiang
SQS	5	S15-08	118.0738639	28.9580278	567–572	Sanqingshan, Yushan, Jiangxi
LYS	5	LY15-07	118.2822889	32.2834417	134–138	Langyashan, chuzhou, Anhui
KYS	5	K15-09	121.7357528	37.2661972	136–141	Kunyushan, Muping, Shandong
FJS	8	F15-05	108.7698111	27.8495556	586–597	Fanjingshan, Tongren, Guizhou
WYS	7	W16-04	117.9581530	27.6423810	206–217	Wuyishan, Wuyishan, Fujian
ZJS	30	Z17-04	118.8264001	32.0671503	238–256	Zijinshan, Nanjing, Jiangsu
WJS	15	WJ17-04	107.4861333	31.2340670	763–791	Wangjiangshan, Dazhou, Sichuan
GJS	22	G17-04	106.6017170	28.9666002	1104–1186	Gujianshan, Qijiang, Chongqing
NHS	30	HS17-05	112.7194830	27.2638000	233–255	Nanyuehengshan, Hengyang, Hunan
FHS	8	FH17-04	108.4688137	32.8432889	813–897	Fenghuangshan, Hanyin, Shanxi
ZJJ	30	ZJ17-05	110.5002670	29.1373330	291–368	Huilongshan, Zhangjiajie, Hunan
SZY*	5	SZ16-10	121.1781194	31.0778861	45	Shanghai Botanical Garden, Shanghai
HZY*	5	HZ15-09	113.3594389	23.1793639	11	South China Botanical Garden, Guangzhou, Guangdong

Note: * cultivated population. SZY cultivars were introduced from Jiangsu between 1973-1977. HZY cultivars were introduced from Guangdong in before 1985; All Samples accession numbers refer to voucher specimens deposited in the Beijing Forestry University (BJFU); Geographic coordinates and elevation were obtained with portable GPS receiver.

Table 2(on next page)

Table 2 Genetic diversity within populations of *L. glauca* revealed by 13 nSSR and 5 cpSSR markers.

Table 2 Genetic diversity within populations of *L. glauca* revealed by 13 nSSR and 5 cpSSR markers.

Population	nSSRs						cpSSRs				
	$N_{A\text{nSSR}}$	A_R	*Ho*	*He*	F_{IS}	PPB(%)	$N_{A\text{cpSSR}}$	*Nb*	*Prv*	*Hrs*	D^2sh
ATM (10)	31	1.823	0.262	0.400	0.3917*	84.62	1	1.000	0	0.000	0.000
JGS (30)	40	1.837	0.390	0.400	0.043	100	8	4.018	1	0.777	39.638
LDZ (20)	34	1.689	0.192	0.344	0.4608*	92.31	4	1.527	0	0.363	1.091
SJG (15)	38	2.011	0.241	0.471	0.5139*	92.31	2	1.142	0	0.133	0.427
NTB (10)	35	1.916	0.162	0.431	0.6561*	92.31	1	1.000	0	0.000	0.000
YTH (10)	41	1.992	0.239	0.456	0.5169*	84.62	1	1.000	0	0.000	0.000
DBS (10)	29	1.588	0.108	0.288	0.6571*	84.62	2	1.220	0	0.200	4.000
HMF (10)	30	1.636	0.115	0.313	0.6617*	92.31	4	2.381	1	0.644	13.227
TMS (10)	38	1.987	0.239	0.447	0.5071*	100	1	1.000	0	0.000	0.000
SQS (5)	26	1.711	0.231	0.317	0.3717*	69.23	3	2.273	0	0.700	3.200
LYS (5)	29	1.747	0.277	0.319	0.238	61.54	3	2.778	0	0.800	16.480
KYS (5)	30	1.734	0.415	0.319	-0.200	76.92	2	1.471	0	0.400	0.320
FJS (8)	28	1.771	0.327	0.380	0.205	92.31	7	6.400	1	0.964	24.343
WYS (7)	22	1.231	0.264	0.188	-0.333	53.85	2	1.324	1	0.286	8.229
ZJS (30)	37	1.431	0.708	0.477	-0.470	100	4	1.230	1	0.193	3.915
WJS (15)	31	1.392	0.518	0.396	-0.275	76.92	3	1.923	0	0.514	2.011
GJS (22)	28	1.929	0.479	0.382	-0.233	76.92	3	2.142	0	0.558	22.940
NHS (30)	40	1.823	0.469	0.404	-0.145	100	5	1.531	0	0.359	2.231
FHS (8)	26	1.748	0.462	0.338	-0.307	69.23	1	1.000	0	0.000	0.000
ZJJ (30)	44	1.806	0.472	0.453	-0.025	100	9	4.787	1	0.818	7.342
SZY[a] (5)	16	1.692	0.192	0.108	-0.667	23.07	2	2.000	1	1.000	0.800
HZY[a] (5)	21	1.921	0.231	0.182	-0.059	61.54	1	1.000	0	0.000	0.000
Mean	31.5	1.746	0.318	0.355	-0.171	81.12	3.1	2.007	0.318	0.396	6.827

Notes: [a] cultivated population; N_{anSSR} = number of alleles across 13 nuclear SSR loci; A_R = allelic richness; *Ho* = observed heterozygosity; *He* = expected heterozygosity; F_{IS} = inbreeding coefficient; PPB = percentage of polymorphic loci; N_{acpSSR} = number of alleles across five chloroplast SSR loci; *Nb* = number of haplotypes; *Hrs* = haplotype richness; *Prv* = private haplotypes; D^2sh = the mean genetic distance between individuals; * significant deviations from *HWE* determined by a global multilocus test implemented in GENEPOP ($P < 0.005$).

Table 3(on next page)

Table 3 Analysis of molecular variance (AMOVA) and degrees of freedom (df) based on 13 nuclear SSR and five chloroplast SSR markers for populations of *L. glauca*. The groups revealed by a Bayesian STRUCTURE analysis ($K = 2$) were considered for

Table 3 Analysis of molecular variance (AMOVA) and degrees of freedom (df) based on 13 nuclear SSR and five chloroplast SSR markers for populations of *L. glauca*. The groups revealed by a Bayesian STRUCTURE analysis were set as 2 for AMOVA.

Source of variation	nSSRs			cpSSRs		
	df	% of Variation	*F* -statistics	df	% of Variation	*F* -statistics
Among groups	1	9.76	F_{CT} = 0.09756*	1	1.79	F_{CT} = 0.01793
Among populations within groups	20	26.42	F_{SC} = 0.29281*	20	29.37	F_{SC} = 0.29903*
Within populations	273	63.82	F_{ST} = 0.36180*	273	68.84	F_{ST} = 0.31160*

Notes: F_{CT} = differentiation among groups; F_{SC} = differentiation among populations within groups; F_{ST} = differentiation among populations; * significant values with P ≤ 0.0001.

Table 4(on next page)

Table 4 Bottleneck analyses for 20 wild populations of *L. glauca.*

Table 4 Bottleneck analyses for 20 wild populations of *L. glauca*

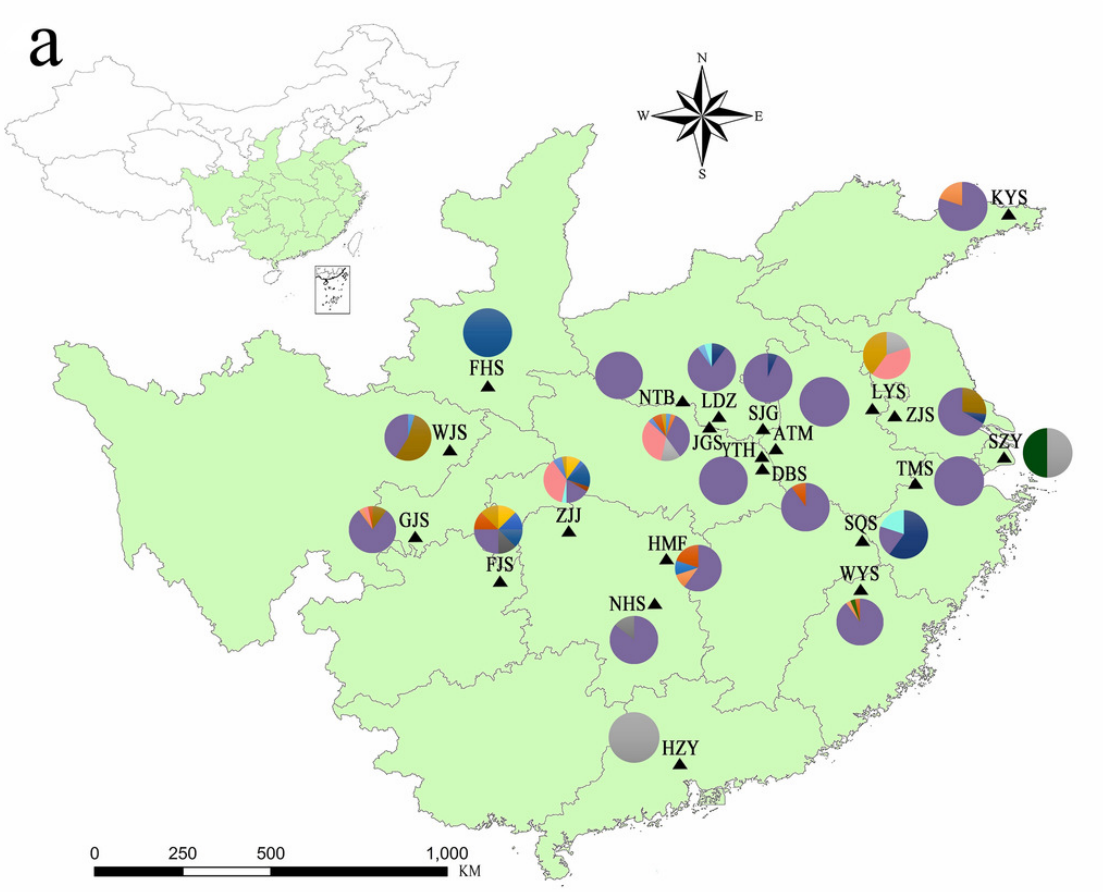
Population	IAM		TPM		SMM	
	Sign test	Wilcoxon test	Sign test	Wilcoxon test	Sign test	Wilcoxon test
ATM	0.0366*	0.0674	0.0596	0.1230	0.2577	0.2061
JGS	0.3046	0.2439	0.3968	1.0000	0.4540	0.7354
LDZ	0.3681	0.3804	0.5193	0.9097	0.4268	0.4697
SJG	0.0047**	0.0105*	0.0469*	0.0522	0.1750	0.1514
NTB	0.2277	0.0923	0.2999	0.1514	0.3827	0.6772
YTH	0.0782	0.2402	0.3010	0.8984	0.2911	0.4131
DBS	0.5698	0.9658	0.4498	0.7002	0.3485	0.3652
HMF	0.3258	0.9697	0.4538	0.9097	0.4342	0.5693
TMS	0.3221	0.2734	0.5841	0.5879	0.0534	0.8394
SQS	0.1909	0.0371*	0.3072	0.3594	0.2811	0.7344
LYS	0.5500	0.4609	0.3552	0.8438	0.2784	0.4609
KYS	0.6099	0.8457	0.1162	0.4316	0.1274	0.1934
FJS	0.0234*	0.0134*	0.1536	0.0923	0.4000	0.2661
WYS	0.4625	0.9375	0.3955	0.9375	0.3416	0.5781
ZJS	0.0117*	0.0067**	0.0247*	0.0803	0.1177	0.2439
WJS	0.0105*	0.0049**	0.0208*	0.0137*	0.1171	0.1309
GJS	0.0066**	0.002**	0.0106*	0.0098**	0.0189*	0.0098*
NHS	0.1557	0.6355	0.2295	1.0000	0.5266	0.7869
FHS	0.1297	0.0098**	0.1693	0.0098**	0.1244	0.0371*
ZJJ	0.3329	0.1909	0.3724	0.5417	0.3253	0.5417

Notes: * significant values with $P \leq 0.05$, ** significant values with $P \leq 0.01$.

Figure 1

Figure 1

Fig. 1. (A) Chloroplast haplotypes and sampling location present in *Lindera glauca* populations analyzed in the present study (see Table 1 for details). Each population is represented by a triangle, and pie charts are shown when a population was present in more than one haplotype. The green background shows the provincial-level distribution of the species in China. (B) Haplotype network was generated by program TCS. Each haplotype is represented by a single color, and circle sizes correspond to the relative frequency of a particular haplotype in the total sample.


a
N
W
E
S
KYS
FHS
NTB
LDZ
SJG
ATM
JGS
YTH
DBS
LYS
ZJS
SZY
TMS
WJS
GJS
ZJJ
FJS
HMF
SQS
WYS
NHS
HZY
0
250
500
1,000
KM

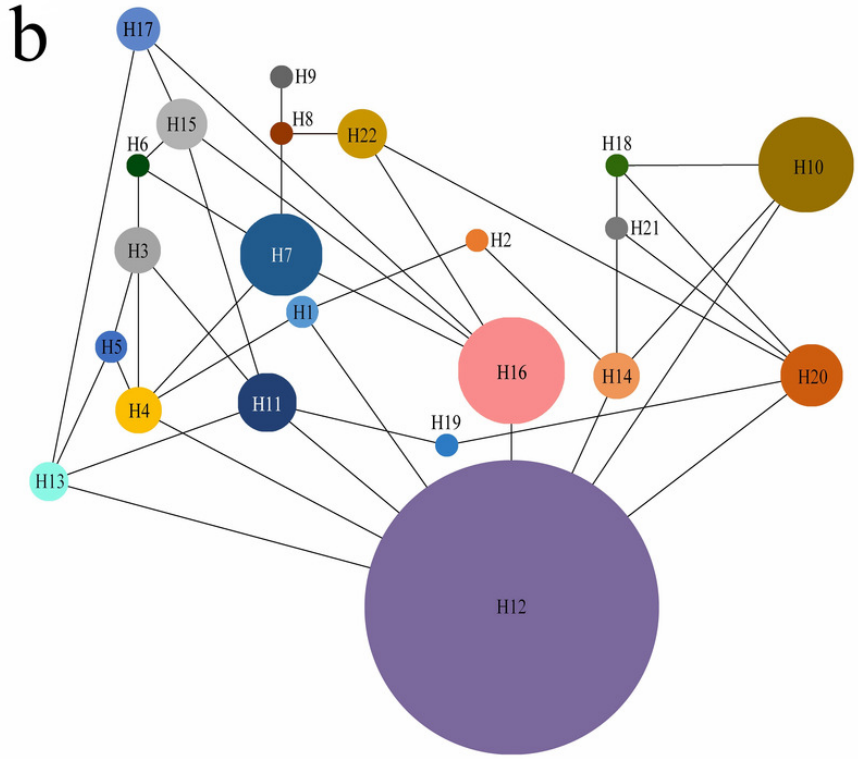

b
H17
H9
H8
H15
H22
H6
H18
H10
H21
H3
H7
H2
H1
H5
H16
H14
H20
H4
H11
H19
H13
H12

Figure 2

Figure 2 (A) Bayesian inference using STRUCTURE (K = 2) based on 13 nSSR markers from 22 populations of *L. galuca*, (B) K = 2 appeared to be the optimal number of clusters by showing the ΔK at its peak.

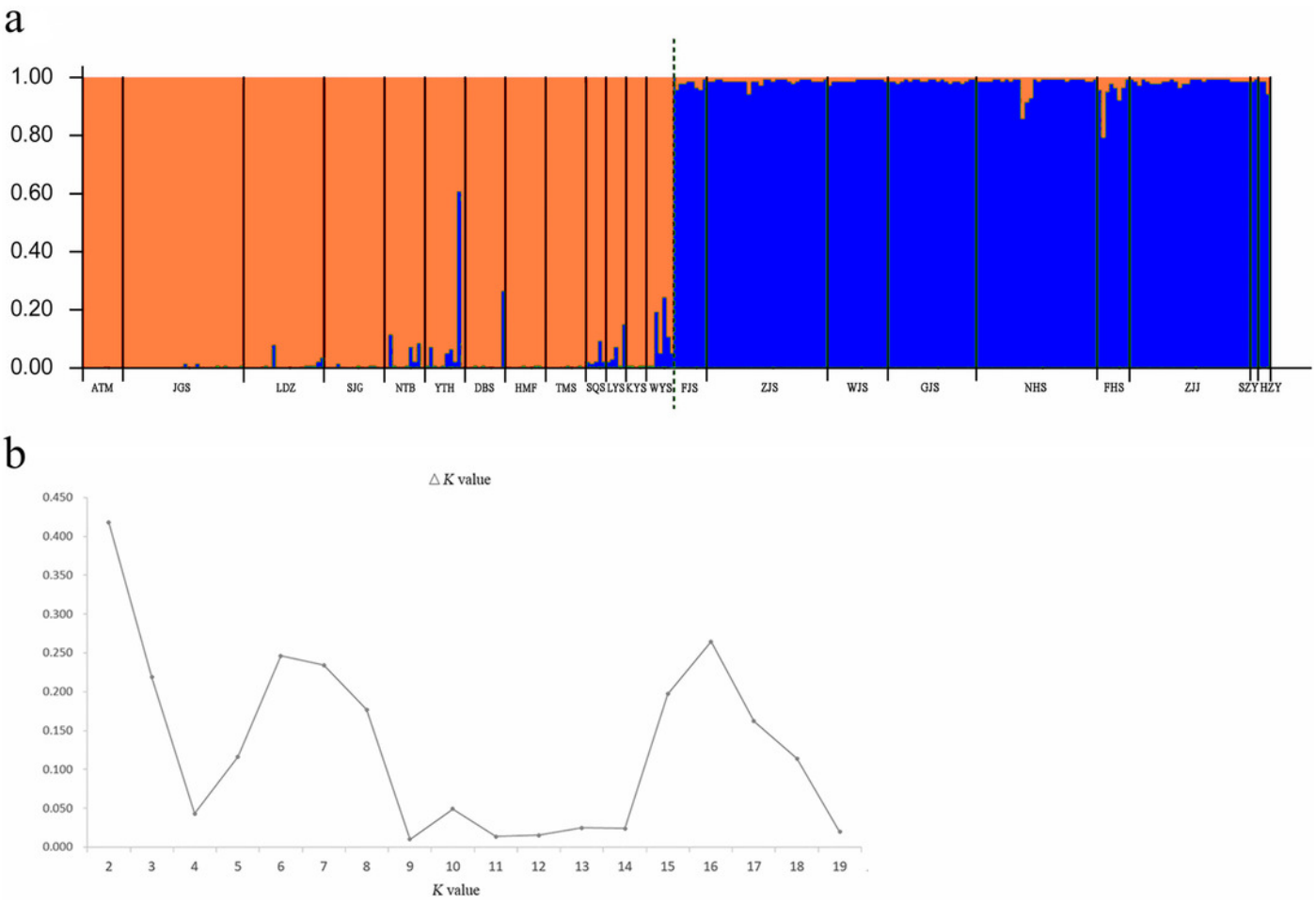
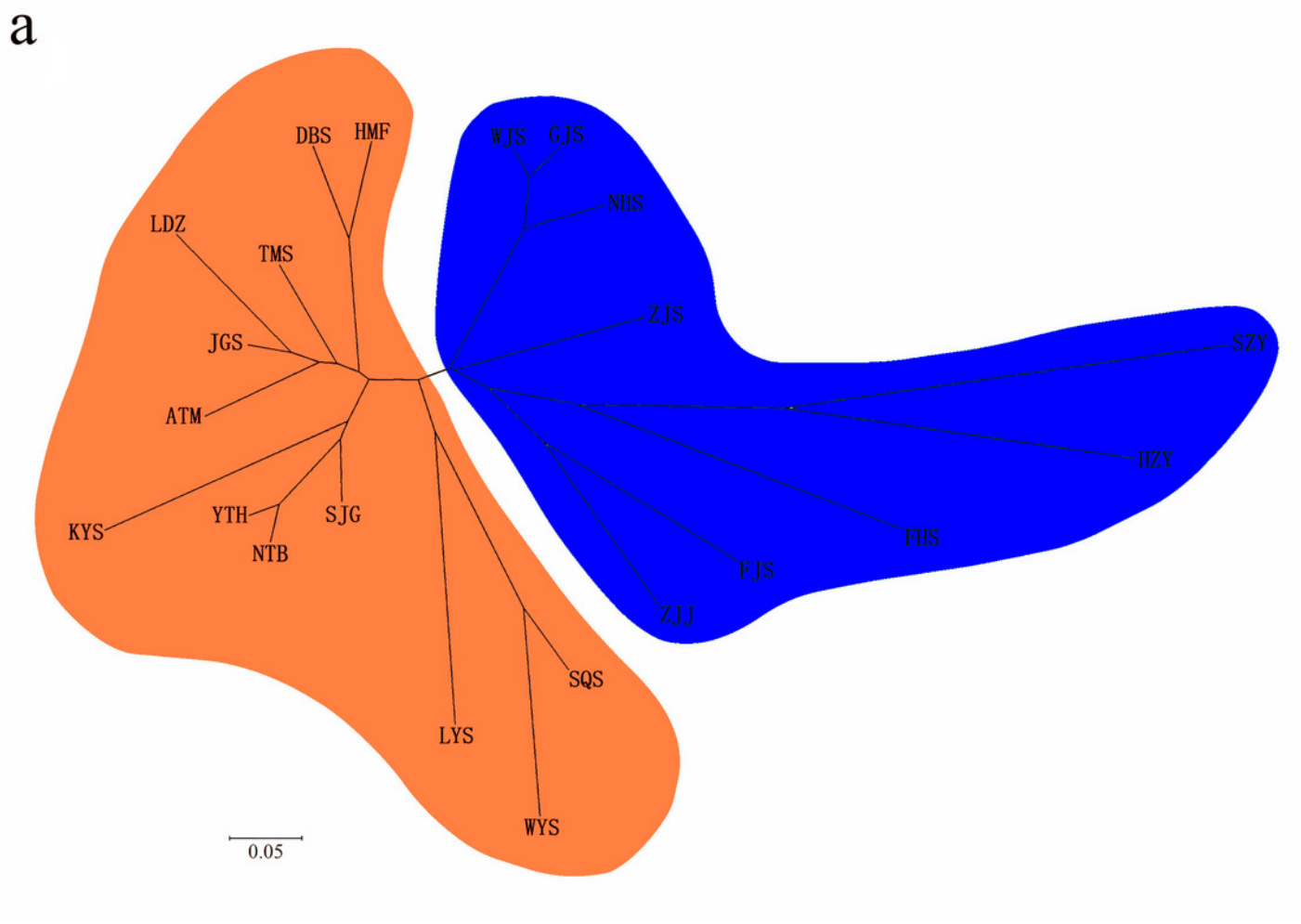


Figure 3

Figure 3 (A) Neighbor-joining (NJ) dendrogram based on Nei's (1978) genetic distances among populations. (B) Principal coordinates analysis (PCoA) of genetic variation across 22 populations of *L. glauca* based on 13 nSSR markers.


a
DBS
HMF
WJS
GJS
NHS
LDZ
TMS
ZJS
SZY
JGS
ATM
HZY
KYS
YTH
SJG
FHS
NTB
FJS
ZJJ
SQS
LYS
WYS
0.05

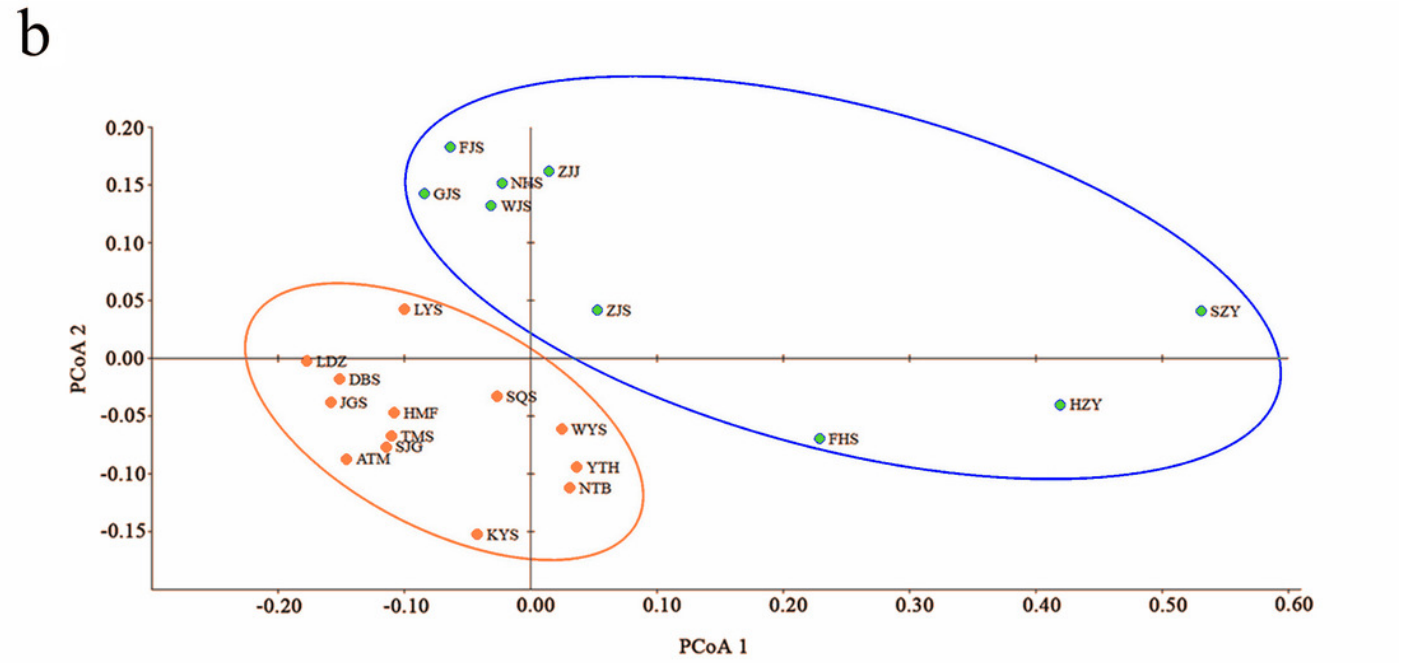

b
0.20
0.15
0.10
0.05
0.00
-0.05
-0.10
-0.15
PCoA 2
FJS
ZJJ
NHS
GJS
WJS
LYS
ZJS
SZY
LDZ
DBS
JGS
SQS
HZY
HMF
WYS
TMS
FHS
SJG
ATM
YTH
NTB
KYS
-0.20
-0.10
0.00
0.10
0.20
0.30
0.40
0.50
0.60
PCoA 1

Figure 4

Figure 4 (A) Figure plot of geographical distance against genetic distance for 22 populations of *L. glauca* based on 13 nSSR markers (B) and 5 cpSSR markers.