A peer-reviewed version of this preprint was published in PeerJ on 31 January 2020.

View the peer-reviewed version (peerj.com/articles/8498), which is the preferred citable publication unless you specifically need to cite this preprint.

Wang Y, Ma Y, Jia B, Wu Q, Zang D, Yu X. 2020. Analysis of the genetic diversity of the coastal and island endangered plant species *Elaeagnus macrophylla* via conserved DNA-derived polymorphism marker. PeerJ 8:e8498 https://doi.org/10.7717/peerj.8498

Analysis of genetic diversity in the coastal and island endangered plant *Elaeagnus macrophylla* by conserved DNA-derived polymorphism markers

Wang Yi [Equal first author, 1], **Yan Ma** [Equal first author, 1], **Bingyu Jia** [1], **Qichao Wu** [1], **Dekui Zang** [Corresp., 1], **Xiaoyan Yu** [1]

[1] Landscape architecture, College of Forestry, Shandong Agricultural University, Tai'an, Shandong province, China

Corresponding Author: Dekui Zang
Email address: zangdk@sdau.edu.cn

The genetic diversity and genetic structure of five natural populations of the island and coastal endangered plant *Elaeagnus macrophylla* were analyzed by conserved DNA-derived polymorphism (CDDP) molecular markers. A total of 289 discernible loci were obtained from 102 individuals using fifteen primers, and 100% of the loci were polymorphic. The observed number of alleles (Na) was 1.9654, the effective number of alleles (Ne) was 1.2604, the average of Nei's genetic diversity index (H) was 0.1724, and Shannon's information index was 0.2869, indicating that *E. macrophylla* had levels of genetic diversity lower than those reported for continental relatives and other coastal species. The average percentage of polymorphic loci (PPL) was 42.1%, and the maximum and minimum PPL values were 80.97% and 14.88%, belonging to the Nanji Island (NJD) and Liugong Island (LGD) populations, respectively. Populations of *E. macrophylla* were highly differentiated (Gst=0.3263). The unweighted pair group method with arithmetic mean (UPGMA) clustering results revealed 5 groups: one group comprising the Liugong Island (LGD) and Da Rushan (DRS) populations, one comprising the Lingshan Island (LSD), Laoshan (LS) and Daguan Island (DGD) populations, one comprising the Putuo Island (PTD), and two groups representing the Nanji Island (NJD) samples, which were differentiated into a northwestern group and a southeastern group. There was no cross-clustering among the samples, and the similarity of the relatives was strictly related to geographical location. R software analysis showed no correlation between genetic distance and geographic distance between populations ($r = 0.256579$, $p = 0.8309$).

Analysis of genetic diversity in the coastal and island endangered plant *Elaeagnus macrophylla* by conserved DNA-derived polymorphism markers

Yi Wang[1] Yan Ma[1] Bingyu Jia[1] Qichao Wu[1] Dekui Zang[1] Xiaoyan Yu[1]

[1]College of Forestry, Shandong Agricultural University, Key Laboratory of State Forestry Administration for Silviculture of the lower Yellow Rive , Tai'an, Shandong, China

Yi Wang and Yan Ma contributed equally to this work.

This study was funded by the project of Shandong Provincial Agricultural Elite Varieties Project (2016LZGC038) and Forestry Science & Technology Innovation Project of Shandong Province (LYCX01-2018-03) , China; Shandong Agricultural Seeds Engineering Project (Collection, Protection and Precision Identification of Forest Tree Germplasm Resources).

Corresponding Author:

Dekui Zang[1]

No. 61 Daizong Street, Shandong Agricultural University, Tai'an, Shandong, 271000, China

Email address: zangdk@sdau.edu.cn

Abstract

The genetic diversity and genetic structure of five natural populations of the island and coastal endangered plant *Elaeagnus macrophylla* were analyzed by conserved DNA-derived polymorphism (CDDP) molecular markers. A total of 289 discernible loci were obtained from 102 individuals using fifteen primers, and 100% of the loci were polymorphic. The observed number of alleles (Na) was 1.9654, the effective number of alleles (Ne) was 1.2604, the average of Nei's genetic diversity index (H) was 0.1724, and Shannon's information index was 0.2869, indicating that *E. macrophylla* had levels of genetic diversity lower than those reported for

continental relatives and other coastal species. The average percentage of polymorphic loci (PPL) was 42.1%, and the maximum and minimum PPL values were 80.97% and 14.88%, belonging to the Nanji Island (NJD) and Liugong Island (LGD) populations, respectively. Populations of *E. macrophylla* were highly differentiated (Gst=0.3263). The unweighted pair group method with arithmetic mean (UPGMA) clustering results revealed 5 groups: one group comprising the Liugong Island (LGD) and Da Rushan (DRS) populations, one comprising the Lingshan Island (LSD), Laoshan (LS) and Daguan Island (DGD) populations, one comprising the Putuo Island (PTD), and two groups representing the Nanji Island (NJD) samples, which were differentiated into a northwestern group and a southeastern group. There was no cross-clustering among the samples, and the similarity of the relatives was strictly related to geographical location. R software analysis showed no correlation between genetic distance and geographic distance between populations ($r = 0.256579$, $p = 0.8309$).

Key Words: *Elaeagnus macrophylla*, Genetic variation, CDDP markers, Conservation implications

Introduction

Elaeagnus macrophylla is an endangered evergreen shrub species of East Asian coastal areas and islands. It is distributed in China's Shandong, Zhejiang, and Jiangsu Provinces, mainly on offshore islands and in coastal lowlands (Chinese Flora, 1983). *E. macrophylla* has great value for studying coastal flora because of its unique geographical distribution pattern and can be widely used in coastal greening due to its tolerance of sea breeze, salinity, drought and thin soil. It also has potential economic value; for example, it can be used to produce fruit juice and wine. In recent years, with the rapid development of the economy and coastline, the intensification of human interference, and a continuous reduction in suitable environments, the number and size of natural populations have decreased sharply, causing the species to become endangered.

The genetic diversity of island species is generally lower than that of terrestrial species, and the risk of extinction is higher for island species (Raven, 1998). From the 17th century to the 20th century, 384 species of vascular plants went extinct worldwide, 139 of which were island plants. Forty percent of vulnerable or endangered vascular plant species are island species (Reid & Miller, 1989). Human disturbances, such as habitat destruction and invasion by alien species, are considered the main factors threatening island species (Wolf & Harrison, 2001). Studies of *Ilex integra* based on inter simple sequence repeat (ISSR) molecular markers (Leng et al., 2005) and *Neolitsea sericea* based on random amplified polymorphic DNA (RAPD) molecular markers (Wang et al., 2004) concluded that the geographical isolation of islands had a significant impact on the genetic differentiation of island populations and that the genetic diversity of island relatives was lower than that of close relatives on continents. However, studies of the genetic diversity of the island plant *E. macrophylla* have not been conducted.

The analysis of conserved DNA-derived polymorphism (CDDP) is based on a single prime amplification reaction, with primers designed to target conserved sequences of plant functional

genes mostly were transcription factor such as WRKY， MYB， MADs and so on. Because the strong conservation of some sequences of plant DNA, CDDP molecular marker technology can be used across different species. Studies of rice (*Oryza sativa*) have shown that CDDP molecular markers have many advantages, including convenience, low cost, and rich polymorphism, which can effectively produce markers related to the target traits (Collard & Mackill, 2009). Compared with traditional DNA molecular methods, CDDP is more practical because the primers used in CDDP are specific for conserved DNA sequences of genes. By amplifying these conserved sequences, which tend to be linked with phenotypic traits, CDDP can provide advantages in plant genetic diversity assessment (Andersen & Lübberstedt, 2003). Since Poczai first successfully used CDDP molecular markers and in-targeting markers to investigate the genetic diversity and group structure of *Solanum dulcamara* (Poczai et al., 2011), CDDP molecular markers have proven useful in the analysis of several other plant species, such as *Chrysanthemum* (Li et al., 2014), *Paeonia suffruticosa* (Li et al., 2013), *Vaccinium vitis-idaea* (Fang et al., 2016), and *Rosa rugosa* (Jiang et al., 2018). However, CDDP markers have not yet been used to study *E. macrophylla*.

In this study, the conserved DNA-derived polymorphism (CDDP) molecular markers were used to analyze the genetic diversity of and genetic relationships among major natural populations of *E. macrophylla* in China in order to reveal the level of genetic diversity and degree of genetic differentiation, analyze the relationships between populations and the influence of geographical isolation and human factors on the genetic structure, and provide a scientific basis for the protection and rational utilization of *E. macrophylla*.

Materials and Methods

Plant materials

A total of 102 individual leaf samples were collected from 7 islands and offshore sites (Fig. 1 and Table 1) from April to July 2018, sampling was conducted in the natural distribution area of *E. macrophylla*. Interval sampling was applied except for small populations (such as the Liugong Island population, whereas samples from all individual plants found were collected). Only one individual was found on each of Lingshan Island (LSD) and Putuo Island (PTD). After collection, silica gel was used to quickly dry the specimens, after which they were stored at -20°C.

DNA extraction and PCR amplification

Total DNA was extracted from *E. macrophylla* by the modified cetyl trimethyl ammonium bromide (CTAB) method (Doyle et al., 1987). The quality and purity of DNA were determined by 2% agarose gel electrophoresis (Fig. 2). All DNA samples were stored at -20℃ for later use.

The DNA from one sample per population was selected to screen 21 CDDP primers (Collard & Mackill, 2009) (synthesized by Sangon Biotech, China). According to the results, 15 primers with clear and reproducible amplification bands were screened out (Table 2).

PCR was conducted in a total reaction volume of 20 µl containing 10 µl 2× Ex Taq MasterMix (dye), 7.5 µl double-distilled H_2O (ddH2O), 1 µl 30 ng/µl DNA template, and 1.0 µl 10 pmol/µl primer (Sangon Biotech, China). A standard PCR cycle (RT-PCR 7500 ,Thermo Fisher Scientific Inc, USA) was used: an initial denaturation step at 94°C for 3 min; 35 annealing cycles of 94°C for 1 min, 50°C for 1 min and 72°C for 2 min; and a final extension of 5 min at 72°C. PCR products were stored at 4°C. The PCR products were electrophoresed on a 2% agarose gel at a voltage of 110 V and current of 110 mA for 1.5-2 h; a DL2000 marker was used as a size marker. The electrophoresis results were photographed and recorded by a gel imaging system.

Data statistics and analysis

The clearer bands (including weak but distinguishable bands) in the electropherogram were marked as "1", and lanes without a band or with one that was difficult to identify were marked as "0", forming a (0, 1) data matrix.

We used POPGEN v.1.32 (Yeh, Yang & Boyle, 1999) to compute the following: observed allele number (Na), effective allele number (Ne), Nei's genetic diversity index (H), Shannon's information index (I), polymorphic loci and polymorphic locus percentage (P, %), total genetic diversity (Ht), interspecific genetic diversity (Hs), intraspecific genetic diversity (Ht-Hs), the genetic differentiation coefficient (Gst), and gene flow (Nm=1-Gst/Gst).

Genetic distance (GD) and genetic identity (GS) were subjected to principal coordinate analysis (PCoA) with NTSYS-pc 2.10e software (RohlF, 1994), and a cluster analysis was performed according to the unweighted pair group method using arithmetic mean (UPGMA). The relationship between geographic distance and genetic distance was analyzed with R software.

Results and Analysis

Population- and species-level diversity of *E. macrophylla*

The DNA of 102 samples was amplified with 15 primers to obtain 289 bands, and the fragment length was between 500 and 2000 bp (Fig. 3). The number of amplified bands ranged from 11 to 30, and the average number of amplified bands was 19.3; the number of Pr2 and Pr6 amplification bands was the highest, at 30, and the number of Pr7 and Pr8 amplification bands was the lowest, at 11. The percentage of polymorphism reached 100% (Table 2), which indicated that the genomic DNA polymorphism of *E. macrophylla* was high.

There were 43 polymorphic loci in the LGD population, representing the lowest percentage of polymorphic loci (PPL, 14.88%), and 234 polymorphic loci in the NJD population, representing

the highest PPL (80.97%). The average values of Na, Ne, H, I and PPL were 1.9654, 1.2604, 0.1724, 0.2869 and 48.928%, respectively (Table 3). Na, Ne, H, I and PPL were consistent among populations, with the NJD population exhibiting the largest values and the LGD population exhibiting the lowest values, all of which were lower than the values observed at the species level. According to a survey taken at the time of sampling, the *E. macrophylla* on NJD is growing well, which may be related to the suitable climate of NJD allowing its growth: the island is located in the subtropical sea area and has a humid climate, warm temperatures in winter, cool temperatures in summer, and abundant rainfall. It has a typical mid-subtropical maritime monsoon climate (Xiao, 2007). Moreover, NJD is located far from the Chinese mainland and is relatively closed, and its exchange with the mainland is relatively limited, with less human interference, providing *E. macrophylla* with extremely favorable growth conditions. However, only two locations with individuals were found on LGD, and the individuals displayed poor growth and obvious insects on the leaves. In addition, because LGD is a famous scenic spot in China, it is extensively developed, which affects the habitat of *E. macrophylla*.

Genetic differentiation of the populations of *E. macrophylla*

The total genetic diversity (Ht), intraspecies genetic diversity (Hs) and interspecies genetic diversity (Dst= Ht-Hs) were 0.1706, 0.1149 and 0.0557, respectively, as calculated with POPGEN v.1.32 software. The interspecific genetic differentiation coefficient (Gst) was 0.3263, indicating that 67.37% of the variation was within the populations and 32.63% of the variation existed among the populations. A certain degree of genetic differentiation was observed among the populations. The gene flow (Nm) was 1.0325, indicating that the gene exchange between populations is limited and that genetic drift is the main cause of their genetic differentiation (Wright, 1931).

Genetic differentiation between populations can be further analyzed based on Nei's genetic distance (Gd) and genetic identity (Gs). For the five populations of *E. macrophylla* (Table 4), the genetic identity and the genetic distance were between 0.8656 and 0.9588 and between 0.0490 and 0.1443, respectively. The genetic similarity of the five populations was slightly higher, and the genetic distance was slightly smaller. The minimum genetic distance was observed between the NJD and DRS populations (0.0421, Table 4), and the genetic similarity between these two populations was the highest. The maximum genetic distance was observed between the LGD and NJD populations (0.1443, Table 4), with low population similarity and a large genetic difference.

UPGMA cluster analysis

The applied measure of genetic similarity was used to construct UPGMA dendrograms (Fig. 4). The clustering map shows that the five populations can be divided into three groups at the genetic coefficient of 0.94. One group represented the LGD population; the populations DRS, NJD and LS formed a second group, indicating that these three populations are closely related; the third group represented the DGD population. Populations with similar geographical distances were not clustered into the same group, indicating that the genetic distance between the populations of the *E. macrophylla* was not related to geographical distance. Consistent with these results, the Pearson correlation coefficient test revealed no significant correlation between geographic and genetic distance ($r = 0.256579$, $p = 0.8309$).

The UPGMA clustering map divides the 102 samples into five groups with a genetic coefficient of 0.786 (Fig. 5). The first group includes all the samples from the LGD and DRS populations, indicating that the two populations are closely related, and the genetic identity shows that compared with other populations, the LGD population is closely related to the DRS population. The second group includes all the samples from the LSD, DGD and LS populations: one sample from the LGD and the samples from the LS population cluster into one group, and those from the DGD population form another group. The third group includes 21 samples from the NJD population. The fourth group is a sample from the PTS population; the fifth group includes 12 samples from the NJD population. The affinity similarity between the tested samples is related to their geographical location. The samples with the same geographical origin tend to be clustered together, and there is no cross-clustering between the samples. Nos. 29-40 of the NJD population are grouped into one category, and Nos. 41-61 are grouped into another category. Based on the sampling location and latitude and longitude, samples 29-40 were collected in the northwestern part of NJD (121°3'24-121°3'8, 27°27'53-27°28'21), and samples 41-61 were collected in the southeastern part of NJD (121°5'52-121°6'11, 27°26'54-27°27'12). This pattern may have occurred because the waters of NJD are under the control of two streams in the Taiwan Warm Current and the Jiangsu and Zhejiang Coastal Currents. In the winter, NJD and its western waters are mainly controlled by the southeastern edge of the Donghai Coastal Current, and the northern part of the island borders the Yellow Sea Coastal Current. The sea area to the east of the NJD is mainly controlled by the northbound Taiwan Warm Current. The southern part of the NJD is connected to the South China Sea Warm Current and Kuroshio Current by the South China Sea Warm Current Continuum and a branch of the Kuroshio Current that passes northeastern Taiwan, respectively. Therefore, the winter temperature is higher in the southeastern part of NJD than in the northwestern part (Xiao, 2007).

In addition, according to the hydrological profile of the southeastern sea area of NJD, the upper 15 m in summer is controlled by mixed water from the Jiangsu and Zhejiang Coastal Currents and the Taiwan Warm Current, while the lower 15 m is controlled by the Taiwan Warm Current. In winter, the Jiangsu and Zhejiang Coastal Currents and the Taiwan Warm Current are in a left-right configuration. The Jiangsu and Zhejiang Coastal Currents control the waters to the west of the 25 m isobath, while the Taiwan Warm Current controls the deep waters to the east of the 45 m isobath. The hydrology of the sea to the northwest of Nanji Island is controlled by the Taiwan Warm Current in summer, the western part of which extends to the 15 m isobath area. During this period, the area is not affected by the Jiangsu and Zhejiang Coastal Currents. In winter, the area is completely controlled by the Jiangsu and Zhejiang Coastal Currents, but the bottom layer is affected by warm water from Taiwan (Xiao, 2007). Therefore, under the influence of ocean currents, the southeastern part of Nanji Island is generally warmer than the northwestern part, and as a result, the habitat may undergo subtle changes, leading to genetic differences that may result in the emergence of two populations on Nanji Island: the southeastern and northwestern populations.

Principal coordinate analysis

The principal coordinate (PCOA) analysis was conducted based on Nei's distance and used to calculate similarity coefficients. The positional relationships between individuals in the principal coordinate analysis map reflected their genetic similarities. The principal coordinate analysis was

performed on all the tested samples using NTSYS-pc 2.10e, and a two-dimensional clustering map of the main coordinates was obtained (Fig. 6). The classification results were largely consistent with the clustering results of the populations and samples. It can be seen from the figure that the LS, NJD and DRS populations tended to locate together, with the LGD and DGD populations being located far away. The LGD population was closest to the DRS population, and the DGD population was closest to the LS population. The 102 samples with the same geographical origin tended to cluster together, and there was no cross-clustering between samples. Consistent with the previous results, the NJD population was divided into two sub-populations. Thus, the principal coordinate analysis results confirmed the clustering results of the UPGMA partition.

Discussion

Genetic diversity and genetic differentiation of *E. macrophylla*

Frankham compared and analyzed the allelic diversity of 202 groups of land and island populations of various species, including mammals, birds, fish, reptiles, insects and plants; in 165 cases (81.7%), the genetic diversity of island populations was lower than that of terrestrial populations, with an average decrease of 29% (Frankham, 1997). The average PPL of the CDDP molecular markers detected in five populations of the island plant *E. macrophylla* was 48.928%, the average effective allele number (Ne) was 1.1801, the average of Nei's genetic diversity index (H) was 0.1149, and the average Shannon polymorphism information index (I) was 0.1848. All of these values are far below the genetic diversity values of populations of continental relatives and the endangered species *Elaeagnus mollis* (PPL=61.99%, Ne=1.6072, H=0.3166, and I=0.4603) estimated based on simple sequence repeat (SSR) molecular markers (Qin et al., 2010) and the genetic diversity values of four populations of the coastal plant *Eurya emarginata* (Ne = 1.223, H = 0.132, and I = 0.200) estimated based on ISSR molecular markers (Zhang et al., 2007). However, the present values are similar to those reported by Frankham. As an endangered coastal plant, *E. macrophylla* exhibits lower genetic diversity than its continental relatives and other coastal plants. The genetic diversity of island species is affected by breeding characteristics, dispersal capacity and effective population size (Frankham, 1997; Weller et al., 1996). The reasons for the low genetic diversity in the populations were as follows: first, *E. macrophylla* has a low seed-setting rate and produces fruit that is sweet and vulnerable to consumption by birds, limiting natural recruitment. Second, as an island species, *E. macrophylla* has small populations, occupies a fragile habitat and has a narrow distribution range, making it more vulnerable to extinction than terrestrial species (Francisco-Ortega et al., 2000). Frequent human activities, such as tourism development, felling and increased anthropogenic shoreline damage, have led to changes in the habitat of *E. macrophylla*, which has had a strong impact on its growth and reduced its genetic diversity. Finally, geographic isolation limits genetic communication between populations. The gene flow and genetic diversity of populations of *E. emarginata* are enhanced through the actions of ocean currents, which allow seed dispersal and

gene exchange between populations (Zhang et al., 2007). However, the cluster analysis of *E. macrophylla* grouped samples strictly according to geographical origin, with no exceptions, indicating that the limited genetic communication among populations might lead to increased inbreeding within the populations and the loss of rare alleles due to genetic drift, thereby reducing the genetic diversity of the populations (Emerson, 2002).

For this species, the total genetic diversity (Ht) and intraspecies genetic diversity (Hs) were 0.1706 and 0.1149 respectively. Compared to the endangered plant *E. mollis* in previous RAPD- (Ht=0.359, Hs=0.302) and SSR-based (Ht=0.3173, Hs=0.2575) studies (Qin et al., 2006; Qin et al., 2010), *E. macrophylla* in the present study showed lower diversity. Similarly, the diversity values in the present study are lower than those estimated using ISSR markers for the coastal and island plants *Ilex integra* (Ht=0.223, Hs=0.153; Leng et al., 2005), *Machilus thunbergii* (Ht=0.269, Hs=0.186; Leng et al. (2006), and *Sonneratia caseolari* (Ht=0.2103, Hs=0.1468; Li et al. (2004)) but higher than those reported for *Neolitsea sericea* (Ht=0.1248, Hs=0.0793; (Wang et al., 2004). Genetic differentiation (Gst) is calculated as the ratio of between-population genetic variance to the total variance among populations(Wright, 1965). Although the results of different molecular markers are not fully comparable, they still provide some useful information. The interspecific differentiation coefficient (Gst) of the five populations of *E. macrophylla* was 0.3263, indicating that 32.63% of the variation existed among the populations. The genetic differentiation of the *E. macrophylla* populations was significant based on Nei's Gst classification criteria for genetic differentiation (low: Gst < 0.05, medium: Gst = 0.05~0.15, and high: Gst > 0.15) (Nei, 1978). Its genetic differentiation value was higher than the average value of 23 species (28.06%) of the Carinla Islands (Francisco-Ortega et al., 2000), it is also proved that the genetic differentiation between *E. macrophylla* populations is higher. The prevention of gene flow was the main cause of genetic differentiation among populations (Ouborg, 1999).

The term gene flow refers to the process by which a biological individual disperses from its place of origin, followed by the exchange of genes between populations. Such exchange may occur between biological populations of the same species or different species, and it is essential to the evolution of many plant populations (Grant, 1991; Gerber et al., 2014). The populations of *E. macrophylla* displayed little gene flow (Nm=1.0325), there was no hybridization among individuals, and the genetic communication between each pair of populations was low. These results were mainly due to the geographical isolation of islands (mainly barriers posed by sea water), which limited the range of dispersal by pollen- and seed-dispersing birds: the maximum range of bird-mediated propagation is approximately 480-680 m (Chung et al., 2002). This distance is shorter than the shortest distance between islands (8100 m, between the LS and DGD); as a result, water currents are the main medium of genetic communication between islands (Kwon et al., 2002). The fruits of *E. macrophylla* fall into the sea due to sea breezes and disperse with the currents over short distances; however, but the germination rates of fruits or seeds of most higher plants are significantly reduced after soaking in sea water for long periods (Angelique, 2000). Thus, the probability of seeds being dispersed by currents is small, resulting in limited gene flow between populations.

Conservation of *E. macrophylla* diversity

The CDDP molecular markers showed that the genetic diversity of the natural population of *E. macrophylla* was low (PPL=48.928%, Ne=1.1801, H=0.1149, and I=0.1848); only the population of Nanji Island showed high genetic diversity (PPL=80.28%, Ne=1.2410, H=0.1580,

and I=0.2613), which was due to NJD being located far away from the Chinese mainland and relatively closed. Compared with the limitations of pollen and seed transmission, human destruction is the main cause of the shrinking populations and reduced genetic diversity. To protect *E. macrophylla*, we should strengthen local protection, increase the awareness of local residents and tourism management personnel, and provide suitable habitats. Additionally, the introduction and exchange of populations should be increased, breaking the barrier to gene flow. Finally, expanding the populations by artificial cultivation and tending would improve the genetic diversity of the populations.

Conclusions

The present study is the first genetic investigation of *E. macrophylla* using CDDP markers to investigate this species' distribution and genetic variation. The results showed that CDDP molecular markers can be effectively used to study the genetic diversity of *E. macrophylla* populations and revealed that *E. macrophylla* populations have low genetic diversity and high genetic differentiation. The low levels of gene flow among populations are the main cause of the high levels of genetic differentiation. Based on these findings, some conservation measures for *E. macrophylla* are proposed.

Acknowledgements

We thank Qing Zhang of the Shandong Agricultural University for assistance with the experimental methods.

References

Andersen JR, Lübberstedt T. 2003. Functional markers in plants. *Trends Plant Sci* **8:** 554-560.

Appliqueing Q, Debusschea M. 2000. Seed dispersal and germination patterns in a rare Mediterranean island endemic (*Anchusacrispa Viv, Boraginaceae*). *Acta Oecologica* **21 (6):** 303-313.

Chinese Flora Editorial Board of the Chinese Academy of Sciences. 1983. Chinese Flora. Beijing: Science Press, **Volume 52(2)** : 015.

Chung GC, Chung MY, Oh GS, et al. 2002. Spatial genetic structure in a *Neolitsea sericea* population (Lauraceae). *Heredity* **85:** 490.

Collard BCY, Mackill DJ. 2009. Conserved DNA-derived polymorphism (CDDP): a simple and novel method for generating DNA markers in plants. *Plant Molecular Bioloty Reporter*, **27:** 558－562.

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

Emerson BC. 2002. Evolution on oceanic islands: molecular phylogenetic approaches to understanding pattern and process. *Molecular Ecology* **11:** 951-966.

Fang WX, Xia XY, An LJ, Peng Q. 2016. Genetic Diversity and Clustering Analysis of Blueberry Resources by CDDP Markers. *Chinese Agricultural Science Bulletin* **32(28):**136-143.

Francisco-Ortega J, Santos-Guerra A, Kim SC, et al. 2000. Plant genetic diversity in the Canary Islands: a conservation respective. *American Journal of Botany* **87(7):** 909-919.

Frankham R. 1997. Do island populations have less genetic variation than mainland populations? *Heredity*, **78:** 311–327.

Gerber S, Chadoeuf J, Gugerli F, Lascoux M, Buiteveld J, Cottrell J, Dounavi A, Fineschi S, Forrest LL, Fogelqvist J, Goicoechea PG, Jensen JS, Salvini D, Ven-dramin GG, Kremer-A. 2014. High rates of gene flow by pollen and seed in Oak populations across Europe. *PLOS ONE* **9:** 1-16. DOI 10.1371/journal.pone.0085130.

Grant V. 1991. The evolutionary process: a critical study of evolutionary theory. New York: Columbia University Press.

Kwon JA, Morden CW. 2002. Population genetic structure of two rare tree species (*Colubrina oppositifolia* and *Alphitonia ponderosa*, *Rhamnaceae*) from Hawaiian dry and mesic forests using random amplified polymorphic DNA markers. *Molecular Ecology* **11:** 991-1001.

Leng X, Wang ZS, An SQ, Feng J, Liu YH, Wang GM. 2005. ISSR analysis of genetic diversity of *Ilex integra*, an insular endemic plant. *Biodiversity*, **13 (06)**: 546-557.

Leng X, Wang ZS, An SQ, Wang GM, Zheng JW, Feng Y. 2006. The Influence of Insular Geographical Isolation on Population Genetic Structure of *Machilus thunbergii Journal of Nanjing Forestry University(Natural Sciences Edition)* **(02):** 20-24.

Li HS, Chen GZ. 2004. Genetic diversity of mangrove plant *Sonneratia caseolaris* in Hainan Island based on ISSR analysis.*Acta Ecologica Sinica* **(08):** 1657-1663.

Li T, Guo JE, Zheng CS, Sun X, Sun XZ. 2014. Genetic diversity and construction of fingerprinting of *chrysanthemum* cultivars by CDDP marker. *Journal of Beijing Forestry University* **36(4):** 94-101.

Liyuan Jiang, Dekui Zang. 2018. Analysis of genetic relationships in *Rosa rugosa* using conserved DNA-derived polymorphism markers. *Biotechnology & Biotechnological Equipment* **32:1** 88-94, DOI: 10.1080/13102818.2017.1407255.

Li YY. 2013. Screening of CDDP Molecular Marker Primers for Tree Peony and Analysis of Its Polymorphism. *Journal of Nuclear Agricultural Sciences* **27(08)**: 1099-1105.

Nei M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* **89(3):** 583-590.

Ouborg NJ. 1999. Population genetics, molecular markers and the study of dispersal in plants. *Journal of Ecology* **87:** 551-568.

Poczai P, Varga I, Bell NE, Hyvönen J. 2011. Genetic diversity assessment of bittersweet (*Solanum dulcamara, Solanaceae*) germplasm using conserved DNA-derived polymorphism and intron-targeting markers. *Ann Appl Biol* **159:** 141-153.

Qin YY, Wang YL, Zhang QD, Bi RC, Yan GQ. 2010. Analysis on the Population Genetic Diversity of an Endangered Plant (*Elaeagnus mollis*) by SSR Markers. *Wuhan botanical research*, **28 (04):** 466-472.

Qin YY, Zhang QD, Yan GQ. 2006. RAPD Analysis of Genetic Diversity in Populations of the Endangered Plant of Elaeagnus mollis Diels. *Molecular Plant Breeding*, **(S2):** 31-36.

Raven PH. 1998. Plant conservation in a changing world. *Aliso*, **16 :** 121-126.

Reid WV, Miller KR. 1989. Keeping Options Alive : The Scientific Basis for Conserving Biodiversity. Washingt on DC : World Resources Institute.

Rohl FJ. 1994. NTSYS-PC. Version 2. 10, New York : Appllied Biostatistics Inc.

Wang ZS, An SQ, Leng X, Ding FM, Zheng JW, Chen L. 2004. Population genetic diversity of the insular plant *Neolitsea sericea* based on random amplified polymorphic DNA (RAPD). *Acta ecologica sinica*, **(03):** 414-422.

Weller SG, Sakai AK, Straub C. 1996. Allo zyme Diversity and genetic identity in *Schiedea* and *Alsinidendron* (*Caryophyllaceae: Alsinoideae*) in the Hawaiian Islands. *Evolution* **50:** 23-34.

Wolf A , Harrison SP. 2001. Effects of habit at size and pat chisolation on reproductive success of the Serpentine Morning Glory. *Conservation Biology*, **15:** 111-121.

Wright S. 1931. Evolution in Mendelian population. *Genetics*, **6:** 91–159.

Wright S. 1965. The interpretation of population structure by F-statistics with special regard to systems of mating. *Evolution* **19:** 395-420.

Xiao JM. 2007. Ecosystem Assessment of Nanji island Based on PSR Model. Xiamen university.

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

Zhang F, Wang ZS, An SQ, Leng X, Feng J, Wei N, Chen SN. 2007. Effect of excursion of plant seeds floating on genetic differentiation of seaside species *Eurya emarginated* Makine around the Zhoushan Archipelago. *Ocean science* **(05):** 25-31.

Figure 1

Fig.1 Geographical location of the 7 sampling points (including Putuo Island and Lingshan Island where only one sample was collected) of *E. macrophylla* in China.

The yellow areas represent sampling provinces in China, the red dots represent sampling sites, and the black circles represent zooming in on the area.

Figure 2

Fig.2 Sample DNA extraction results of the Daguan Island *E. macrophylla* population.

The samples from left to right are DGD1-12, and Marker is DL5000.

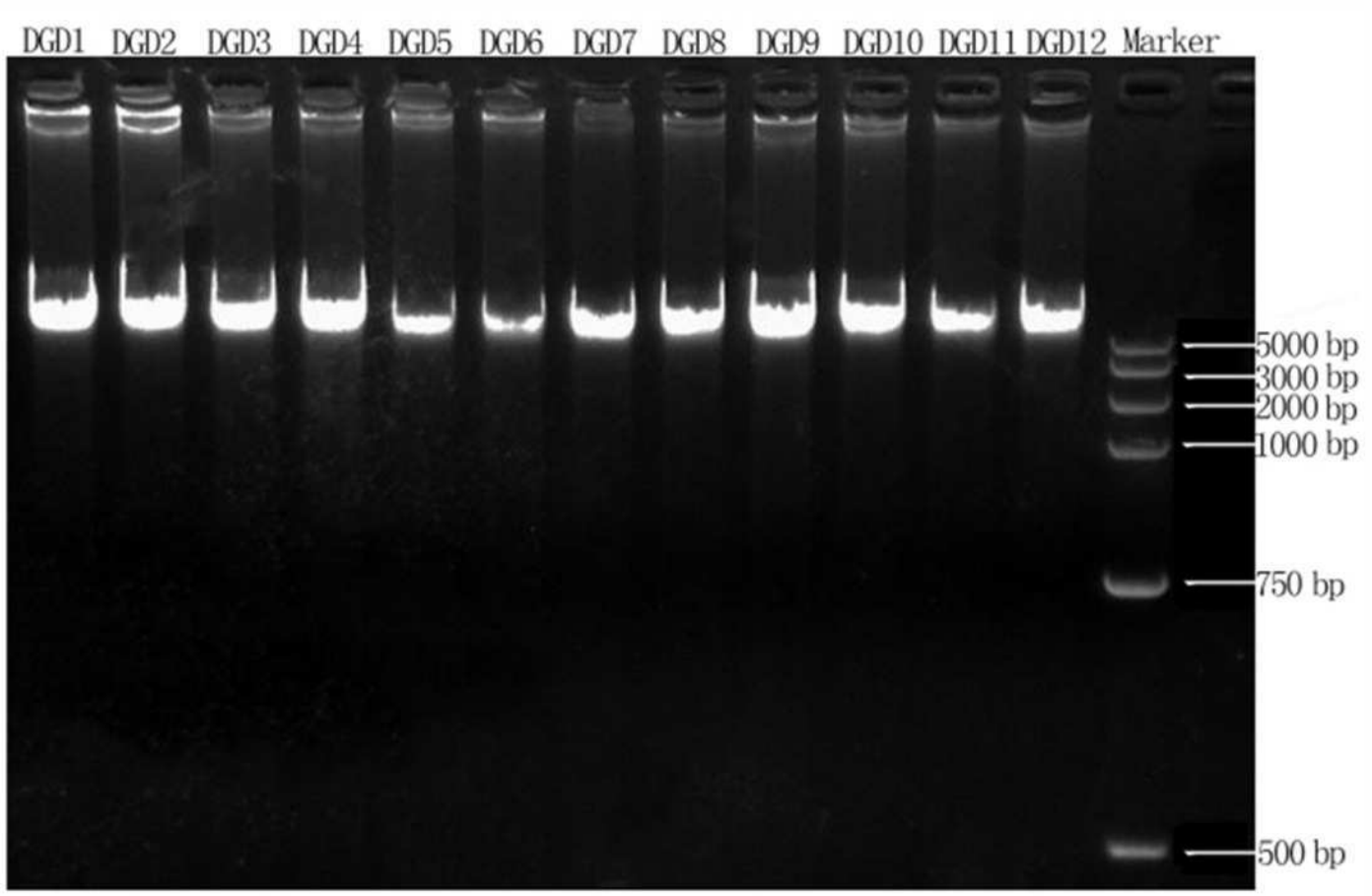

Figure 3

Figure 3 Amplification results for MYB1 in the Liugong Island population and the Da Rushan population.

Amplification results for MYB1 in the Liugong Island population (1-8) and the Da Rushan population (9-20), Marker=DL2000.

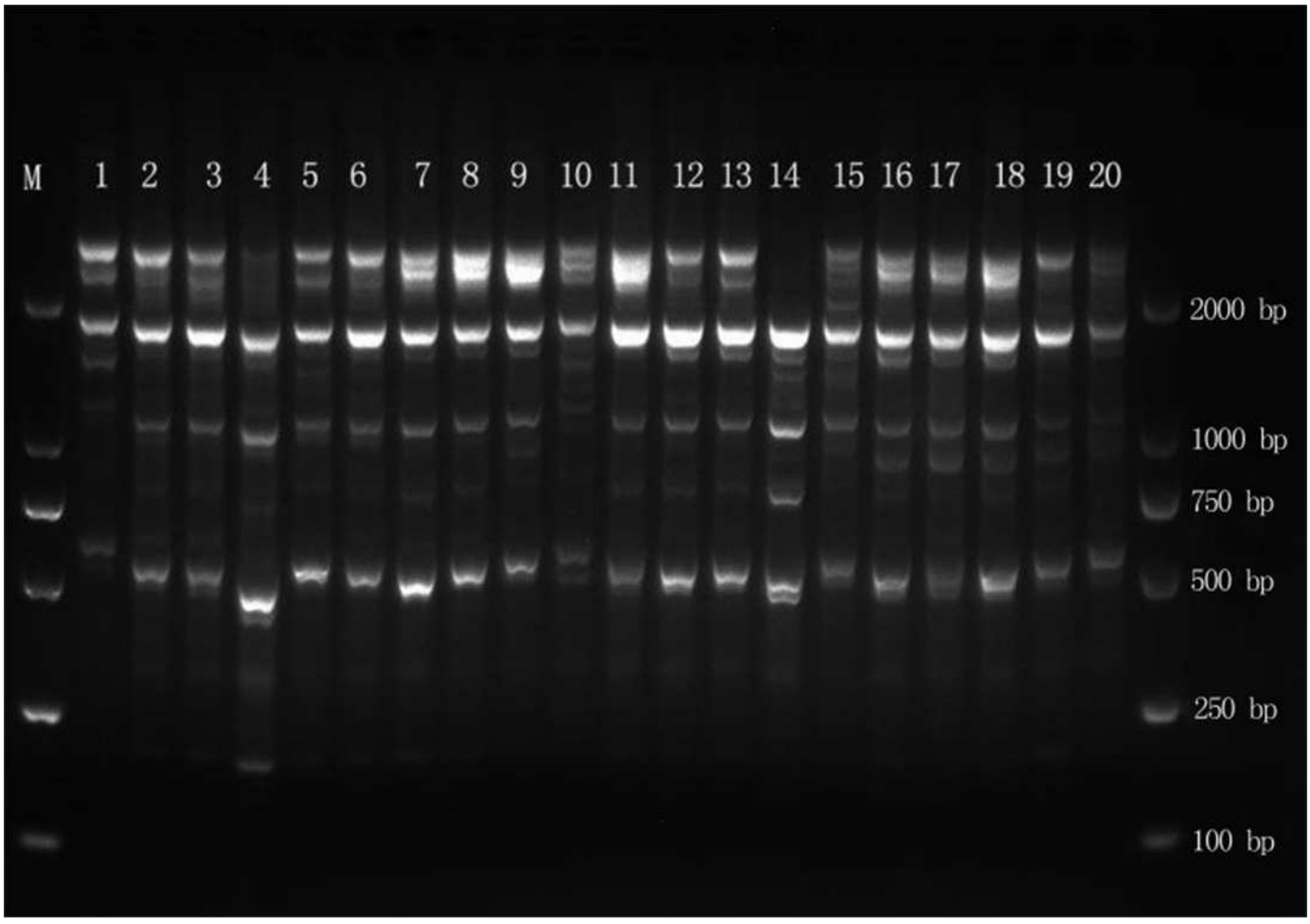

Figure 4

Fig.4 UPGMA cluster analysis of genetic similarity of 5 populations.

There are 7 sampling points, but since LSD and PUD only collect one sample, they cannot be counted as a population, so there are only cluster results of 5 populations.

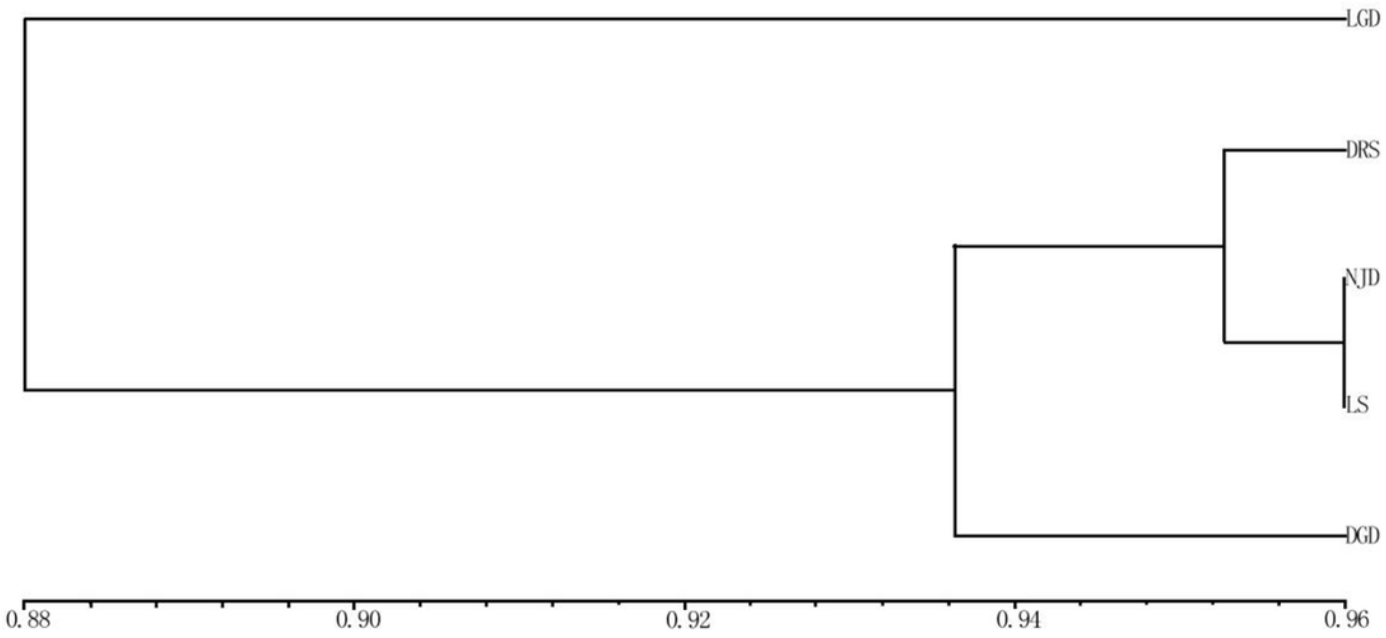

Figure 5

Figure 5 UPGMA cluster analysis of 102 samples of *E. macrophylla.*

Liugong Island (1-8), Da Rushan (9-28), Nanji Island (29-61), Putuo Island (62), Lingshan Island (63), Laoshan (64-76), and Daguan Island (77-102). Sample clustering contains one sample of Lingshan Island and one sample of Putuo Island.

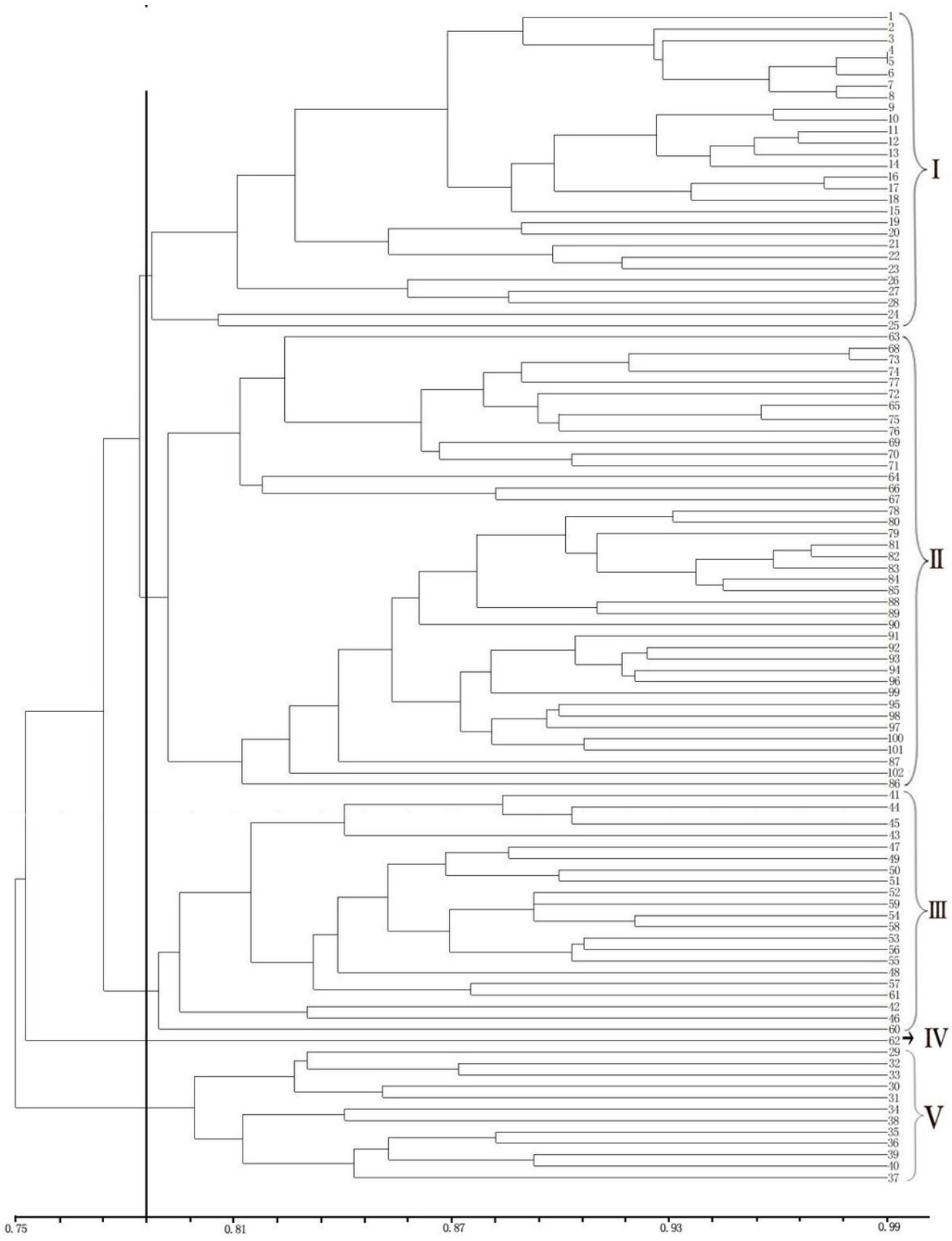

Figure 6

Fig.6 Two-dimensional PCoA analysis of 102 samples.

The numbers in the figure are consistent with those in Fig.5.

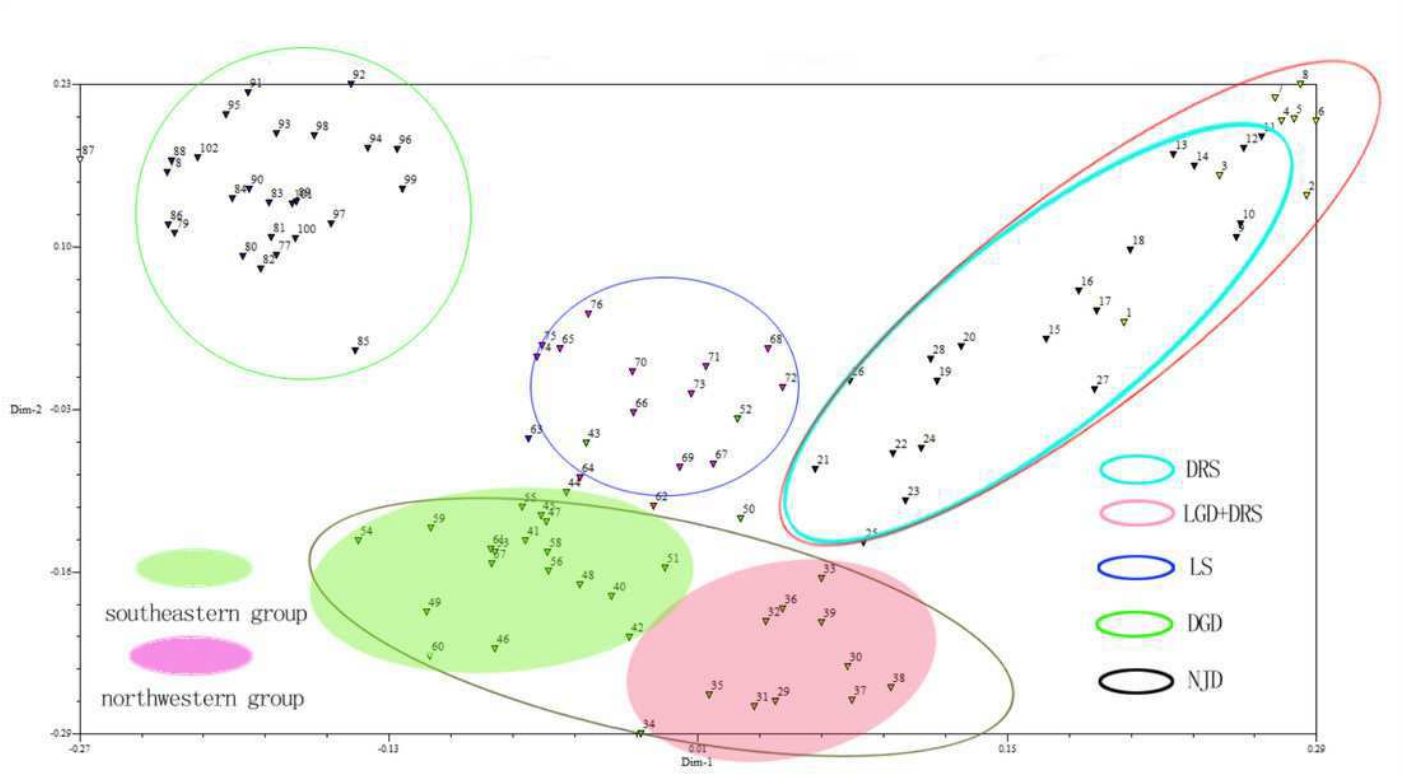

Table 1(on next page)

Table 1 Sampling information for *E. macrophylla*

The table shows information such as population name, abbreviation, geographic location, latitude and longitude, and altitude.

1 Table 1 Sampling information for *E. macrophylla*

Population	Number	Locality	Geographical location	Altitude (meters)
Liugong Island (LGD)	8	Weihai Bay, Weihai city, Shandong Province	37°30'N, 122°10'E	22
Da Rushan (DRS)	20	Rushan city, Weihai city, Shandong Province	36°45'N, 121°30'E	5.2
Lingshan Island (LSD)	1	Huangdao District, Qingdao city, Shandong Province	36°27'N, 121°58'E	20
Daguan Island (DGD)	26	Laoshan District, Qingdao city, Shandong Province	36°13'N, 120°46'E	11
Laoshan (LS)	13	Laoshan District, Qingdao city, Shandong Province	36°7'N, 120°39'E	20
Putuo Island (PTD)	1	Zhoushan Islands, Zhoushan city, Zhejiang Province	30°0'N, 122°24'E	96
Nanji Island (NJD)	33	Pingyang County, Wenzhou city, Zhejiang Province	27°28'N, 121°3'E	42

2

Table 2(on next page)

Table 2 Site information for 15 CDDP markers and genetic diversity parameters at each locus of *E. macrophylla.*

Table contains inforations such as p rimer coding, primer names, sequence, annealingTemperature, number of bands recorded, number of polymorphic bands and PPL.

Table 2 Site information for 15 CDDP markers and genetic diversity parameters at each locus of *E. macrophylla*

Primer coding	Primer name	Sequence (5'-3')	Annealing Temperature	Number of bands recorded	Number of polymorphic bands	PPL/%
Pr1	WRKY-F1	TGGCGSAAGTACGGCCAG	50	21	21	100
Pr2	WRKY-R1	GTGGTTGTGCTTGCC	52	30	30	100
Pr3	WRKY-R3	CCGCTCGTGTGSACG	50	21	25	100
Pr4	MYB1	GGCAAGGGCTGCCGC	50	19	19	100
Pr5	MYB2	GGCAAGGGCTGCCGG	50	13	13	100
Pr6	ERF1	CACTACCGCGGSCTSCG	50	30	30	100
Pr7	ERF2	GCSGAGATCCGSGACCC	50	11	11	100
Pr8	ERF3	TGGCTSGGCACSTTCGA	50	11	11	100
Pr9	KNOX-1	AAGGGSAAGCTSCCSAAG	50	21	21	100
Pr10	KNOX-2	CACTGGTGGGAGCTSCAC	50	19	19	100
Pr11	KNOX-3	AAGCGSCACTGGAAGCC	50	15	15	100
Pr12	MADS-1	ATGGGCCGSGGCAAGGTGC	50	14	14	100
Pr13	MADS-4	CTSTGCGACCGSGAGGTG	50	28	28	100
Pr14	ABP1-1	ACSCCSATCCACCGC	50	14	14	100
P15	ABP1-3	CACGAGGACCTSCAGG	50	18	18	100

Table 3(on next page)

Table 3 Genetic diversity in five populations of *E. macrophylla.*

The table contains information such as population name, number of samples, Na, Ne,H, I, PPL, and standard deviation in parentheses.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27850v1 | CC BY 4.0 Open Access | rec: 11 Jul 2019, publ: 11 Jul 2019

Table 3 Genetic diversity in five populations of *E. macrophylla* (standard deviation in parentheses)

Population name	Number of samples	*Na*	*Ne*	*H*	*I*	*PPL (%)*
Liugong Island (LGD)	8	1.1488 (0.3565)	1.0739 (0.2178)	0.0446 (0.1210)	0.0690 (0.1785)	14.88
Da Rushan (DRS)	20	1.5398 (0.4993)	1.2002 (0.2850)	0.1290 (0.1622)	0.2070 (0.2381)	53.98
Nanji Island (NJD)	33	1.8028 (0.3986)	1.2410 (0.2931)	0.1580 (0.1566)	0.2613 (0.2198)	80.28
Laoshan (LS)	13	1.4998 (0.4983)	1.1941 (0.3020)	0.1208 (0.1680)	0.1898 (0.2458)	50.52
Daguan Island (DGD)	26	1.5502 (0.4983)	1.1912 (0.2878)	0.1222 (0.1615)	0.1968 (0.2355)	44.98
Mean	20	1.5083	1.1801	0.1149	0.1848	48.928
Species level	100	1.9654 (0.1831)	1.2601 (0.2845)	0.1724 (0.1532)	0.2869 (0.2098)	96.54

Note: The Putuo Island (PTD) and Lingshan Island (LSD) populations are not include because there was only one sample for each.

Table 4(on next page)

Table 4 Nei’s genetic identity (above diagonal) and genetic distance (below diagonal) for five populations.

Liugong Island (LGD), Da Rushan (DRS), Nanji Island (NJD), Laoshan (LS), Daguan Island (DGD).

Table 4 Nei's genetic identity (above diagonal) and genetic distance (below diagonal) for five populations

Population	LGD	DRS	NJD	LS	DGD
LGD	****	0.9253	0.8656	0.8730	0.8697
DRS	0.0776	****	0.9588	0.9569	0.9431
NJD	0.1443	0.0421	****	0.9522	0.9391
LS	0.1358	0.0440	0.0490	****	0.9427
DGD	0.1396	0.0585	0.0628	0.0590	****

Note: Liugong Island (LGD), Da Rushan (DRS), Nanji Island (NJD), Laoshan (LS), Daguan Island (DGD).