Research paper

Genetic diversity and population structure of *Indigofera szechuensis* complex (Fabaceae) based on EST-SSR markers

Li-na Goo\(^{a,b,c}\), Xin-fen Gao\(^{a}\)*

---

**A R T I C L E   I N F O**

**Keywords:**

*Indigofera*  
EST-SSRs  
Transcriptome  
Morphology  
Mantel test

**A B S T R A C T**

*Indigofera*, the third largest genus of Fabaceae, comprises approximately 750 species worldwide with a pantropical distribution. Eight *Indigofera* species, namely *I. calcicola*, *I. delavayi*, *I. franchetii*, *I. hancockii*, *I. lenticellata*, *I. pendula*, *I. rigioclada*, and *I. szechuensis*, are considered a species complex because of their morphological similarities and the phylogenetic analysis based on Internal Transcribed Spacer (ITS). Small populations of these species are allopatrically distributed in Hengduan Mountains in China. Although considerable EST-SSR markers have been developed from the transcriptome of *I. szechuensis*, no codominant markers have been applied to study population genetic structure of the complex. In this study, we selected 66 EST-SSR markers which were transferable in *Indigofera szechuensis* complex for estimating polymorphism, of which 44 EST-SSRs (66.67%) were polymorphic. Amplification with selected 23 polymorphic EST-SSRs revealed a moderately high genetic diversity in this complex. The mean value of Observed number of alleles, Expected heterozygosity, Polymorphism information content and Shannon's information index was 10.4783, 0.4335, 0.6228 and 1.4369 respectively. A total of 758 genotypes were detected, with an average of 32.9565 genotypes per locus. The Mantel test showed a significant correlation between genetic and geographic distance (\(r = 0.0748, P = 0.0100\)). High differentiation and limited gene flow were detected among all populations (\(F_{ST} = 0.3589, Nm = 0.5168\)). The PCA and structure analysis grouped 31 populations of *Indigofera szechuensis* complex into five main species. *I. delavayi* was obviously separated from other species, and the result was in accordance with that of morphology and phylogeny. *I. pendula* was a separate species, and had two distinct phenotypes. Four *I. szechuensis* populations (Pop23, Pop24, Pop27 and Pop30) were the same species. *I. calcicola*, *I. hancockii*, *I. rigioclada*, *I. franchetii*, *I. lenticellata*, and *I. szechuensis*, distributed in southwest of Jinshajiang, could be considered as the same species; the species of *I. franchetii*, *I. lenticellata*, and *I. szechuensis* distributed in northeast of Jinshajiang could be considered as another separate species. Given the above information, the morphological classification of *I. calcicola*, *I. franchetii*, *I. hancockii*, *I. lenticellata*, *I. rigioclada*, and *I. szechuensis* was incredible.

---

**1. Introduction**

*Indigofera szechuensis* complex belongs to the species-rich genus *Indigofera* in Fabaceae and includes eight species. Currently, this complex distributed in Hengduan Mountains is endemic to China (Gao et al., 2016). *Indigofera calcicola* Craib, *I. delavayi* Franchet, *I. franchetii* X.F. Gao & Schrire, *I. hancockii* Craib, *I. lenticellata* Craib, *I. pendula* Franchet, *I. rigioclada* Craib, and *I. szechuensis* Craib are considered a species complex, yet, the presence of intermediate morphotypes poses some taxonomic challenges (Yang et al., 2015).

While *I. delavayi* and *I. szechuensis* have been reported to have \(2n = 2x = 16\) chromosomes, the number of chromosomes in other six species remains unknown (Zheng et al., 2011). Pollen of this complex is a combination of 3-colporate pollen grains and microperforate tectum (Zhao et al., 2016). Phylogenetic studies have used Internal Transcribed Spacer (ITS) to evaluate low genetic diversity and differentiation within species (Yang et al., 2015). With the development of sequencing techniques, a large number of EST-SSRs have been developed from the transcriptome of *I. szechuensis* and the transferability to seven other closely related species in this complex have been examined.
(Guo et al., 2016). Using SSR markers, no genetic differentiation has been found in Indigofera szechuensis complex (Yang et al., 2015; Guo et al., 2016).

SSRs are now widely employed for assessing genetic diversity, phylogeographic, and population genetic relationships (Li et al., 2013; Jia et al., 2015; Li et al., 2016). SSRs have also been used to construct linkage maps, and identify molecular markers for marker-assisted selection (Fraser et al., 2009; Guo et al., 2011; Rajaram et al., 2013). EST-SSRs differ from, and have several advantages over traditional genomic SSR markers due to its codominant, abundant, highly reproducible, lower development cost and higher transferability across related species (Mian et al., 2005). Therefore, EST-SSR markers developed in the transcriptome of I. szechuensis provide an additional and distinct solution for investigating population genetic structure of the Indigofera szechuensis complex.

Although some morphological variation exists within the Indigofera szechuensis complex, assessment of their genetic diversity using just ITS has been inadequate. The resulting data are limited, however, and can’t fully reveal genetic relationships among the Indigofera szechuensis complex. In this study, EST-SSR primers developed in our laboratory by transcriptome were used to identify polymorphisms and to analyze the genetic diversity and structure of 31 populations comprising 8 main species. This information may be useful to solve some intermediate morphotypes taxonomic.

2. Materials and methods

2.1. Plant materials and DNA extraction

A total of 558 accessions collected from Hengduan Mountains were provided by Chengdu Institute of Biology, Chinese Academy of Sciences (CAS). These accessions were divided into 31 populations according to sources (Table S1). Each population was represented by 16 to 20 individuals except the population 16 includes only 12 individuals, and each of the sampled individuals was kept > 100 m apart to minimize the genetic relationships among the sample trees. Populations 1, 2, 7, 19 are I. calicula (20 accessions), I. delavayi (20 accessions), I. hancockii (16 accessions), I. rigioclada (20 accessions) respectively. Populations 3–6 are four I. franchetii populations (72 accessions), 8–16 are nine I. lenticellata populations (156 accessions), 17–18 are two I. pendula populations (40 accessions) and 20–31 are twelve I. szechuensis populations (214 accessions). Another 18 individuals of I. mairei were used as an outgroup. Samples for DNA extraction were dried in silica gel. Total genomic DNA was extracted using a Plant Genomic DNA Kit (Tiangen, China) according to the manufacturer’s protocol.

2.2. Primer screening and microsatellite amplification

A total of 66 primers pairs (Table S2) were selected according to their transferability in previous studies (Guo et al., 2016). Then these EST-SSRs were used to identify polymorphisms in 124 non-repeated accessions randomly selected from the 31 populations. The polymorphisms were checked by 8% polyacrylamide gel electrophoresis, with DNA bands visualized by silver nitrate staining. All primers were synthesized by Sangon biotech co., Ltd. (Sichuan, China). A total of 44 primer pairs produced clear and reproducible polymorphic fragments among the 124 accessions. In the end, we selected 23 EST-SSRs to assess the genetic diversity of all 558 accessions.

PCR amplification was conducted in a total volume of 25 μl containing 1.0 μl genomic DNA (10 ng/μl), 1.0 μl of the specific primer (10 pmol/μl), 2.5 μl 10 × EasyTaq buffer (Tiangen, China), 0.5 μl dNTP mix (10 mM), 0.3 μl EasyTaq DNA polymerase (5 unit/μl, Tiangen) and 18.7 μl ddH2O. PCR amplification was conducted as follows: PCR mixtures were held at 94 °C for 4 min, followed by 35 cycles of 94 °C for 45 s, 58–60 °C for 1 min and 72 °C for 45 s. The final extension was performed at 72 °C for 10 min. Forward primers were labeled with different fluorescence: FAM, HEX and TAMRA. PCR products were then separated by capillary electrophoresis in an ABI 3730XL Genetic Analyzer (Applied Biosystems, Foster City, CA, USA), and the fragment size was assessed using genemarker Version 1.90 software.

2.3. Data analysis

SSR data were scored as two alleles per locus distinguished by their size. We did not found multi-locus cases. Observed number of alleles (Na), effective number of alleles (Ne), the number of genotypes where each locus amplified alleles (Ng), and Nei’s gene diversity (H) were computed using POPGENE32 software (Yeh et al., 1999). The observed heterozygosity (Ho), expected heterozygosity (He), Shannon’s information index (I), gene flow (Nm = [(1/Fst) – 1] / 4), wright’s F statistics parameters (Fis, Fst and Fis) and Wright’s fixation index (F = 1-Ho/He) were calculated using GenAlEx 6.5 (Peakall and Smouse, 2012). GenAlEx 6.5 was also used to conduct Mantel test, PCoA and AMOVA. The Simpson diversity index for each SSR loci, also known as the polymorphism information content (PIC), was calculated using the program PIC-CALC 0.6 (Liu et al., 2015). In addition, HWE test was conducted through Genepop on the web (http://genepop.curtin.edu.au/) (Raymond and Rousset, 1995).

A pairwise comparison of the genetic differentiation (Fst) outlier test was performed using LOSITAN software (Beaumont, 2005; Antao et al., 2008) in order to identify candidate non-neutral SSR loci, which might have biased the genetic diversity analysis. The identified neutral SSR loci were used for further genetic diversity and population genetic analyses. Genetic relationships were determined by cluster analysis based on the unweighted pair group method of mathematical averages (UPGMA) as implemented in MEGA version 7.0 (Kumar et al., 2016). We also used Structure version 2.3.4 to identify genetic groups within the 31 Indigofera populations (Pritchard et al., 2000). Structure applies a Bayesian method for inferring K values without using a priori classification. K values ranging from 1 to 10 were adopted to infer the number of clusters for twenty replicate runs, with a 200,000 iterations burn-in period followed by 1,000,000 iterations Markov chain Monte Carlo (Hubisz et al., 2009). The most favorable grouping number (K) was determined by calculating ΔK using Structure Harvester Web version 0.6.94 (Evanno et al., 2005; Earl and Vonholdt, 2012).

3. Results

3.1. Genetic diversity of the Indigofera szechuensis complex

Using the 23 EST-SSRs that produced clear polymorphic fragments during preliminary screening, we detected 241 alleles and 758 genotypes in the 558 studied accessions. Observed number of alleles (Na) is one of the most important indexes of genetic differentiation is associated with populations, types, and geographical sites (Wang et al., 2012). Among the 558 accessions, Na per locus varied from 4 to 22, with a mean value of 10.4783, and the number of amplified genotypes varied from 8 to 108, with an average of 32.9565. The effective number of alleles (Ne) for each locus varied between 1.1560 and 7.7856, with an average of 3.5960 per locus. The frequency of more than half of the alleles (145, 60.17%) was < 5% and 87 were unique (frequency < 1%). Approximately 26.09% of polymorphic SSR loci were associated with nine alleles (Table 2). Values of Shannon’s information index (I) varied from 0.3079 to 2.3798 per locus, with an average of 1.4369, while expected heterozygosity (He) and observed heterozygosity (Ho) ranged from 0.0616 to 0.6938 (mean = 0.4335) and 0.0486 to 0.6728 (mean = 0.4017), respectively. The value of genetic diversity which calculated according to Nei's 1973 (H) ranged from 0.1349 (Gln57) to 0.8716 (Gln99), with an average of 0.6591. Polymorphism information content (PIC) values for each SSR ranged from 0.1293 (Gln57) to 0.8601 (Gln99), with an average of 0.6288, indicating a high level of genetic diversity in Indigofera szechuensis.
heterozygosity; He, Expected heterozygosity; HF is Fit Fst − 0.5651 (mean = 0.4654), respectively. The lowest genetic differentiation (Fst) ranged from 0.2029 (Gln99) to 0.6446 (Gln6), with a mean of 0.3589, which indicated the limited gene flow and high differentiation among populations (Table 1). Based on AMOVA analysis, most variation (59.74%) was detected within individual accessions, with only 32.79% attributed to variation among the 31 populations. The inbreeding coefficient at total populations; Nm, Gene flow (Nm = [(1/Fis) − 1] / 4); PIC, Polymorphism information content, and HWE, Hardy-Weinberg equilibrium and ** represent the high significant (P < 0.0000).

The HWE test showed significant deviations for 3 EST-SSRs (Gln29, Gln57 and Gln101), and high significant deviations for 7 EST-SSRs (Gln21, Gln36, Gln71, Gln81, Gln84, Gln88 and Gln99) which may result from a deficiency of heterozygosity among Indigofera populations (Table 1). There were only 4 populations (Pop3, Pop7, Pop21 and Pop22) showed no significant deviations while the remaining 27 populations demonstrated high significant deviations which indicated almost all populations maybe affected with evolution of the factors of interference such as hybrid introgression, mutation and selection of migration (Table 4).

3.2. Comparative diversity of Indigofera szechuensis complex from different populations

We detected 2105 alleles at 17 neutral SSR loci (Fig. 1) in the 558 accessions, with the total number of alleles in each population ranging from 40 to 88 (Table 4). Na per population ranged from 2.3529 to 5.1765 and averaged 3.9943, while Ne ranged from 1.2666 to 3.1243, with a mean of 2.3625. Using the 17 EST-SSRs, we detected 36 private alleles in the eight studied species. The Shannon's information index (I) parameter values (Ne, Ho, He, and PIC) were found in accessions (mean = 0.1010), and the Shannon's information index (I) parameter values (Ne, Ho, He, and PIC) were found in accessions (mean = 0.1010), and the Shannon's information index (I) parameter values (Ne, Ho, He, and PIC) were found in accessions (mean = 0.1010), and the Shannon's information index (I) parameter values (Ne, Ho, He, and PIC) were found in accessions (mean = 0.1010), and the Shannon's information index (I) parameter values (Ne, Ho, He, and PIC) were found in accessions (mean = 0.1010), and the Shannon's information index (I) parameter values (Ne, Ho, He, and PIC) were found in accessions (mean = 0.1010), and the Shannon's information index (I) param
for all 23 polymorphic markers. Six outlier loci were identified (Fig. 1): Gln6, Gln31, Gln41, Gln53, Gln69 and Gln91. Finally, we selected 17 polymorphic and neutral EST-SSR markers, with which to assess population genetic structure among the 31 Indigofera populations.

To illustrate the genetic relationships among the 31 populations, an unweighted pair-group method with arithmetic (UPGMA) dendrogram (Fig. 2) was constructed based on the Nei's genetic distances (Table S3). UPGMA cluster analysis grouped 31 populations of Indigofera szechuensis complex into three main clusters at a genetic similarity value of 0.4342. As seen in the dendrogram, cluster A was further subdivided into three subgroups, with the populations in A1 were all collected from Sichuan and A2 mainly comprising populations from I. lentellicata. Cluster B were consisting of I. pendula and half of I. szechuensis populations, and cluster C only involved population 2 (I. delavayi) which was the most divergent population. Pop3, Pop6, Pop9, Pop17, Pop19, Pop20 and Pop30 were collected from Yunnan, and the other populations were collected from Sichuan (Table S1). These results indicated that the populations in Sichuan had no distinct geographic structure difference with the populations in Yunnan.

The Mantel test revealed that a significant positive correlation was found between genetic distance and geographic distance in Indigofera szechuensis complex (r = 0.0748, P = 0.0100). These results provided further evidence that geographic isolation was one possible factor leading to high genetic differentiation of the complex.

The genetic diversity of 558 accessions was also confirmed by PCoA. The PCoA analysis revealed that large genetic diversity exists in Indigofera szechuensis complex. The first three axes explained 24.55% of cumulative variation (Table 5). The first and second coordinates accounted for 13.48% and 19.58% of the molecular variation, respectively. In PCoA, eight Indigofera species labeled with 31 different colours which represent the 31 populations were grouped into three clusters (Fig. 3). I. delavayi (Pop2) showed distinct grouping was the different.
The individuals of *I. pendula* (Pop17 and Pop18) and *I. szechuensis* (Pop23, Pop24, Pop26, Pop27, Pop29 and Pop30) were grouped together. The others from *I. calcicola*, *I. franchetii*, *I. hancockii*, *I. lenticellata*, *I. rigioclada*, and *I. szechuensis* were the third group.

### 3.4. Population genetic structure of the Indigofera szechuensis complex

We evaluated population structure and differentiation of the 558 accessions from 31 populations using a Bayesian Markov chain Monte Carlo approach as implemented in Structure version 2.3.4 as a first step. Here, ΔK values computed for all classes indicated a strong signal for K = 2 (ΔK = 3458.71). The results indicated that the 31 populations under study were grouped into two clusters (Fig. 4a). Group 1 consisted of 22 populations, all from *I. calcicola*, *I. franchetii*, *I. hancockii*, *I. lenticellata*, *I. rigioclada*, and *I. szechuensis* (Pop20, Pop21, Pop22, Pop25, Pop28, Pop31). Group 2 consisted of 9 populations, all from *I. delavayi* (Pop2), *I. pendula* (Pop17, Pop18), and *I. szechuensis* (Pop23, Pop24, Pop26, Pop27, Pop29, Pop30). Several individuals displayed an intermixed composition, which could be attributed to a historical gene flow between them. Further, based on the membership fractions, individuals under different populations were categorized as pure or admixture. The individuals with the probability more than ≥ 0.80 score was considered as pure and < 0.80 as an admixture. Pop2, Pop17 and Pop 18 showed 19 pure (95.00%) and 1 admixed (5.00%) individuals; Pop4, Pop11 and Pop29 showed 17 pure (94.44%) and 1 admixed (5.56%) individuals; Pop14 and Pop20 showed 16 pure (88.89%) and 2 (11.11%) admixed individuals; Pop3 showed 15 pure (83.33%) and 3 (16.67%) admixed individuals; Pop7 showed 14 pure (85.71%) and 2 (14.29%) admixed individuals; Pop8 showed 13 pure (86.67%) and 2 (13.33%) admixed individuals; Pop9 showed 12 pure (83.33%), 13 (72.22%), and 9 (75.00%) admixed individuals respectively. The other populations had no admixed individuals.

Then the three populations in which the admixed individuals were > 70% and the admixed individuals in other populations were removed to identify genetic subgroups within the group 1 and group 2 respectively. Group 1 was further subdivided into two subgroups (Fig. 4b). Nest structure analysis separated Pop1, Pop3, Pop6, Pop8, Pop11, Pop13, Pop19 and Pop 20 from the other ten populations. The group 2 was further structured into three subgroups, revealing that at least three distinct groups existed among the 9 populations (Fig. 4c). Pop2 (*I. delavayi*) showed distinct grouping was the first group. Pop17 (*I. pendula*), Pop18 (*I. pendula*), Pop26 (*I. szechuensis*) and Pop29 (*I. szechuensis*) were grouped together. The others were the third group.
These results closely mirrored the pattern of diversity revealed by the UPGMA dendrogram and PCoA analysis.

The AMOVA and PCoA of group 1 and group 2, based on the structure analysis, were analyzed. In group 1, AMOVA analysis detected a variation of 69.91% within individuals, while 21.83% of the variation was attributed in 19 subgroups. The overall $F_{st}$ in the 19 subgroups was 0.2183 (Table S4). AMOVA analysis was also performed on the 9 populations in group 2. Among the 9 populations 35.45% variance was recorded, whereas, among individuals, 7.62% variance and within individuals, 56.93% variance was found. The overall $F_{st}$ in the 9 subgroups was 0.3545 (Table S5).

In PCoA, group 1 was further divided into two subgroups, which was essentially identical to those determined by structure analysis (Fig. 5a). The first three axes explained 22.61% of cumulative variation. The first and second principal coordinates account for 9.89% and 16.88% of the total variation, respectively. PCoA revealed that large genetic diversity exists in group 2. A PCoA plot showed that the first and second coordinates accounted for 19.64% and 30.62% of the molecular variation, respectively. The first coordinate clearly separated Pop2 (I. delavayi) from the other 8 populations in group 2. The second coordinate separated Pop17 (I. pendula), Pop18 (I. pendula), Pop26 (I. szechuensis) and Pop29 (I. szechuensis) from the other 4 I. szechuensis populations (Fig. 5b). The results were in agreement with structure analysis.

4. Discussion

4.1. SSR polymorphism and genetic diversity

Reliable and polymorphic SSR markers are essential to study the genetic diversity and structure of such species. As indicated, 66 EST-SSRs which are transferability in Indigofera szechuensis complex have already been published (Guo et al., 2016). Nevertheless, the number of polymorphic SSRs didn’t developed in that study. We, therefore, screened them for polymorphisms using the 124 representatives randomly selected from 31 populations. Of the 66 primer pairs, 44 (66.67%) produced clear, reproducible and polymorphic fragments. In the present study, 23 primer pairs were applied to characterize 558 Indigofera accessions. The different colours represent the 31 populations.
4.2. Genetic diversity in Indigofera szechuensis complex

Indigofera szechuensis complex is characterized by its trichomes typically medifixed (T-shaped), floral morphology, height, and leaflet size across its distributional range (Fang and Zheng, 1989, 1994; Gao et al., 2010). As recorded in descriptions and data standards for this complex, they show a low degree of variation in morphological features and are often indistinguishable between each other. Attempts to use morphological variation and ITS sequence to distinguish the interspecies have not been successful (Yang et al., 2015). Up to now, no molecular markers have been used to study this complex. In our study, the 23 SSR primers possessed an average of 10.4783 alleles per locus, and the mean value of PIC and He was 0.6228 and 0.4335, respectively. The average value of $F_{st}$ was 0.3589, which stands for a high level of population differentiation. This was also shown in the AMOVA, which indicated that 32.79% of variation was attributable to population diversity. Limited gene flow could be the cause of high genetic differentiation. We suggest that the geographic isolation, inbreeding and limited gene flow may be the foremost evolutionary force in Indigofera szechuensis complex. Furthermore, the Mantel test displayed a significant positive correlation between genetic and geographic distance in the sampled populations ($r = 0.0748$, $P = 0.0100$).

4.3. Population relationships and structure in Indigofera szechuensis complex

The structure analysis grouped the 31 populations into group 1 and group 2 at $K = 2$ (Fig. 4a). After removing the admixed samples, group 1 was further subdivided into two subgroups at $K = 2$. In group 1, Pop1 (I. calcicola), Pop3 (I. franchetii), Pop6 (I. franchetii), Pop7 (I. hancockii), Pop8 (I. lenticellata), Pop11 (I. lenticellata), Pop13 (I. lenticellata), Pop19 (I. rigiolada) and Pop 20 (I. szechuensis) were clustered together, which were mainly distributed in southwest of Jinshajiang. While Pop4 (I. franchetii), Pop5 (I. franchetii), Pop12 (I. lenticellata), Pop14 (I. lenticellata), Pop15 (I. lenticellata), Pop21 (I. szechuensis), Pop22 (I. szechuensis), Pop25 (I. szechuensis), Pop28 (I. szechuensis) and Pop 31 (I. szechuensis) were separated across the Qinghai Tibet Plateau (QTP) and the subsequent climate changes might have promoted species diversification in Indigofera szechuensis complex.

5. Conclusions

In conclusion, the Indigofera-specific SSR markers developed in this study can be served as effective molecular tools for the assessment of genetic diversity and the elucidation of population structure in Indigofera szechuensis complex. The Indigofera species exhibited moderately high genetic diversity, high differentiation and limited gene flow using the selected 23 polymorphic EST-SSRs. The PCoA and structure analysis grouped 31 populations of Indigofera szechuensis complex into five main species. Population genetics analysis this study and previous morphology and phylogeny research revealed that I. delavayi was the most divergent species; I. pendula was a separate species, and had two distinct phenotypes. Pop23, Pop24, Pop27 and Pop30 were the same species; I. calcicola, I. hancockii, I. rigiolada, I. franchetii, I. lenticellata, and I. szechuensis, distributed in southwest of Jinshajiang, could be considered as the same species; the species of I. franchetii, I. lenticellata, and I. szechuensis distributed in northeast of Jinshajiang could be considered as another separate species. Given the above information, we claimed that morphological classification of I. calcicola, I. franchetii, I. hancockii, I. lenticellata, I. rigiolada, and I. szechuensis was not credible. The results of cluster and genetic structure analysis revealed by our study have provided us basic genetic knowledge of these species, and this information will be very useful to policy makers on Indigofera.
breeding plan and species conservation.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.gene.2017.04.047.

Conflicts of interest

The authors declare that they have no conflicts of personal, communication or financial interests.

Acknowledgments

This research was supported by the National Natural Science Foundation of China (grant nos. 31370230, 31070173).

References

Antao, T., Lopes, A., Lopes, R., Beja-Pereira, A., Luikart, G., 2008. LOSITAN: a workbench to detect molecular adaptation based on a F-statistic method. BMC Bioinformatics 9, 323.

Beaumont, M.A., 2005. Adaptation and speciation: what can Fst tell us? Trends Ecol. Evol. 20, 435–440.

Earl, D.A., Vonholdt, B.M., 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv. Genet. Resour. 4, 359–361.

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

Fang, Y.Y., Zheng, C.Z., 1989. A study on the genus Indigofera Linnaeus from China. Acta Phytotaxon. Sin. 27, 161–177.

Fang, Y.Y., Zheng, C.Z., 1994. Indigofera Linnaeus. In: Wei, Z. (Ed.), Fl Reipubl Popularis Sin. Vol. 40. Science Press, Beijing, pp. 239–325.

Fraser, L.G., Tsang, G.K., Datson, P.M., De Silva, H.N., Harvey, C.F., Gill, G.P., Crowhurst, R.N., McNeilage, M.A., 2009. A gene-rich linkage map in the dioecious species Actinidia chinensis (kiwifruit) reveals putative X/Y sex-determining chromosomes. BMC Genomics 10, 102.

Guo, L.N., Zhao, X.L., Gao, X.F., 2016. De novo assembly and characterization of leaf transcriptome for the development of EST-SSR markers of the non-model species Indigofera szechuanensis. Biochem. Syst. Ecol. 68, 36–43.

Hubisz, M.J., Falush, D., Stephens, M., Pritchard, J.K., 2009. Inferring weak population structure with the assistance of sample group information. Mol. Ecol. Resour. 9, 1322–1332.

Jia, H.M., Jiao, Y., Wang, G.Y., Li, Y.H., Jia, H.L., Wu, H.X., Chai, C.Y., Dong, X., Guo, Y.P., Zhang, L.P., et al., 2015. Genetic diversity of male and female Chinese bayberry (Myrica rubra) populations and identification of sex-associated markers. BMC Genomics 16, 394.

Kumar, S., Stecher, G., Tamura, K., 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol. Biol. Evol. 33, 1870–1874.

Li, X.W., Meng, X.Q., Jia, H.J., Yu, M.L., Ma, R.J., Wang, L.R., Cao, K., Shen, Z.J., Niu, L., Tian, J.B., et al., 2013. Peach genetic resources: diversity, population structure and linkage disequilibrium. BMC Genomics 14, 84.

Li, C.Y., Chiang, Y.C., Chung, C.C., Hsu, H.H., Ge, X.J., Huang, C.C., Chen, C.T., Hung, K.H., 2016. Cross-species, amplifiable EST-SSR markers for Armentosus species obtained by next-generation sequencing. Molecules 21, 67.

Liu, H.Y., Li, C.Y., Xiong, F., 2015. Isolation and characterization of 19 polymorphic microsatellite loci from Neosalix rubens, a rapidly invasive and adaptive species. Biochem. Syst. Ecol. 61, 121–123.

Mian, M.A.R., Saha, M.C., Hopkins, A.A., Wang, Z.Y., 2005. Use of tall fescue EST-SSR markers in phylogenetic analysis of cool-season forage grasses. Genome 48, 637–647.

Nei, M., 1973. Analysis of gene diversity in subdivided populations. Proc. Natl. Acad. Sci. USA 70, 3321–3323.

Peakall, R., Smouse, P.E., 2012. GenAlEx 6.5: genetic analysis in excel. Population genetic software for teaching and research—an update. Bioinformatics 28, 2537–2539.

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

Rajaram, V., Nepolean, T., Senthilvel, S., Varshney, R.K., Vadez, V., Srivastava, R.K., Shah, T.M., Supriya, A., Kumar, S., Kumari, B.R., et al., 2013. Pearl millet [Pennisetum glaucum (L.) R. Br.] consensus linkage map constructed using four RIL mapping populations and newly developed EST-SSRs. BMC Genomics 14, 159.

Raymond, M., Rousset, F., 1995. Genepop (version 1.2)—population-genetics software for exact tests and ecumenicism. J. Hered. 86, 248–249.

Wang, C.F., Jia, G.Q., Zhi, H., Niu, Z.G., Chai, Y., Li, W., Wang, Y.F., Li, H.Q., Lu, P., Zhao, B.H., et al., 2012. Genetic diversity and population structure of Chinese foxtail millet [Setaria italica (L.) Beauv.] landraces. G3 Genes Genom. Genet. 2, 769–777.

Yang, X.H., Zhao, X.L., Gao, X.F., 2015. Morphological variation and ITS sequence analysis of the Indigofera szechuanensis complex. Plant. Sci. J. 33, 727–735.

Yeh, F.C., Yang, R.C., Boyle, T., 1999. POPGEPHE Software Package Version 1.31 for Population Genetic Analysis. University of Alberta, Edmonton, AB, Canada.

Zhao, X.L., Gao, X.F., Xu, B., 2016. Pollen morphology of Indigofera (Fabaceae) in China and its taxonomic implications. Plant Syst. Evol. 302, 469–479.

Zheng, X., Gao, X.F., Xu, B., Tu, W.G., 2011. Chromosomal studies on seven species of Indigofera L. Plant Sci. J. 29, 417–422.