Assessment of Genetic Diversity and Identification of Core Germplasm of *Pueraria* in Guangxi Using SSR Markers

Xiao-Hong Shang
Guangxi Academy of Agricultural Science
Yun Zhou
Guangxi Medical University https://orcid.org/0000-0002-1957-9919
Liang Xiao
Guangxi Academy of Agricultural Science
Ying Wang
Guangxi Academy of Agricultural Science
Wen-Dan Zeng
Guangxi Academy of Agricultural Science
Sheng Cao
Guangxi Academy of Agricultural Science
Zheng-Dan Wu
Guangxi Academy of Agricultural Science
Hua-Bing Yan (h.b.yan@hotmail.com)
Guangxi Academy of Agricultural Science

Research Article

Keywords: Core germplasm, genetic diversity, SSR, Pueraria

Posted Date: November 8th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-1002949/v1

Abstract

Pueraria is not only one of the most important commercial crops, but a health supplement for human being. There are abundant *Pueraria* germplasm resources and a large planting scale in Guangxi. However, the genetic diversity and core germplasm of the *Pueraria* species in Guangxi are rarely understood. In this study, 272 individuals of *Pueraria* species in Guangxi combined with 23 pairs of simple sequence repeat primers were used to evaluate the genetic diversity and construct core germplasm. Ultimately, 118 alleles were identified and 112 alleles were polymorphic; the mean expected heterozygosity per locus is 0.1841, and the mean gene flow N_m is 1.7690. 272 individuals were divided into two main clusters, which is consistent with the results based on principal coordinate analysis and STRUCTURE cluster analysis. We proposed a core collection of 20 *Pueraria* accessions capturing 105 alleles. There was a non-significant relationship between genetic distance and geographical distance. The results could provide theoretical support for the scientific conversation of *Pueraria* genetic resources, which can serve as the basis for the breeding program of *Pueraria*.

Introduction

Pueraria lobata (2n=2x=22) is a semi-woody, perennial liana, which belongs to Leguminosae family, and widely distributed throughout Asia, including China, Japan, Korea and other regions in Southeast Asia, as well as in North and South America. As an economic crop, it contains puerarin and other functional components and is used in the production of both pharmaceuticals and health foods. The roots of both *Pueraria lobata* and *Pueraria lobata var. thomsonii* have been long used for treating fever, toxicosis, indigestion, and liver damage from alcohol abuse in traditional Chinese medicine[1], which was recorded in *The Divine Husbandman's Classic of Materia Medica* (*Shen Nong Ben Cao Jing*) compiled in the Eastern Han Dynasty (25-250 AD). China is probably the origin and distribution center of *Pueraria* species; however for a long time, the identification and the identification and breeding of germplasm resources has not received enough attention. With the development of urban society and excessive mining, many germplasm resources are facing the risk of loss or extinction. Genetic diversity provides a basis for the improvement of the crop for different desirable traits, evolutionary capability, species survival, management of germplasm collections, and breeding programs [2,3]. Therefore, it is necessary to fully understand the genetic diversity and genetic information of *Pueraria* core germplasm resources of the representative individuals, which can protect the key genetic resources and shorten the breeding process [4]. Most recently, RAPD (random amplified polymorphic DNA), ISSR (inter-simple sequence repeat) and SRAP (sequence-related amplified polymorphic), SSR (simple sequence repeats) markers have been used to analyze the genetic diversity in *Pueraria* [5-10]. Genic-SSRs have the advantages of the four markers because of the more comprehensive genetic information in the genome. Although the genic-SSR were used to evaluate the diversity of *Pueraria*, however, the population is just 44 lines [11].

Guangxi is a hotspot of *Pueraria* genetic resources in China [12]. However, the genetic diversity and core germplasm of the *Pueraria* species in Guangxi are rarely understood. In the present study, 272 individuals of *Pueraria* collected in Guangxi were used to estimate the extent of genetic diversity and construct the core germplasm. The findings of this study will be utilized for conservation and management of genetic resources in Guangxi, association mapping, and traits-based kudzu breeding.

Materials And Methods

Plant material and DNA extraction

A total of 272 individuals of *Pueraria* were collected in Guangxi from September 2017 to April 2019 (Figure 1, Table S1). Three to five fresh young leaves of each accession were collected and immediately frozen in liquid nitrogen and stored at -80℃ until DNA isolation.

Total genomic DNA was extracted from young leaf tissue of individual representative plants of each accession using a Plant DNA Isolation Reagent Kit (TaKaRa, Dalian, China). We measured the concentration and purity of the total DNA using both 1% agarose gel electrophoresis and a Nanodrop instrument (UV-2700). The total DNA extracts were stored at -20 ℃ until required for experiments.

SSR genotyping

The final concentration of DNA was adjusted to 50 ng/μl for PCR reaction. Based on the transcriptome of *P. lobata*, we designed and scored 28 SSR primers in six *Pueraria* collections from 229 SSRs [11]. Ultimately, 23 polymorphic markers were chosen for genetic diversity analysis. SSR amplification was carried out in a thermal cycler by Bio-Rad (MyCycler TM), in a final volume of 20 μl containing: 100 ng of genomic DNA, 10 μl of *Taq* DNA polymerase mix (TaKaRa, Dalian, China), and 10 μM each, forward and reverse non-fluorescent primers. The program used for PCR amplification was as follows: initial denaturation at 94 ℃ for 5 min; 30 cycles of denaturation at 94℃ for the 30 s, annealing at 50℃ for 30 s, extension at 72℃ for 30 s, and a final extension at 72℃ for 10 min. Amplified products were separated in 6% non-denaturing polyacrylamide gel electrophoresis (PAGE). The SSR markers amplified at sizes between 100 bp and 400 bp were converted into '0' and '1' codes denoting 'absence' and 'presence', respectively.

Data analysis

A sampling of accession for the core collection was carried out using corehunter in R. For each SSR locus, Popgene32 version 1.32 was used to analyze the gene frequency, number of allele (N_a), effective number of alleles (N_e), polymorphic loci, genetic distance (Nei's genetic distance), Shannon–Weaver diversity index (I^*), Homogeneity test index (H^*) and gene flow. Cluster analysis by unweighted pair group method with arithmetic mean (UPGMA) based on the jaccard method was also developed using the NTSYS-pc 2.10e [13] software. A principal component analysis (PCA) was performed using NTSYS 2.10.

The core germplasm set was developed using software Core hunter in R package [14] and validation of the core germplasm set is done based on maximum Shannon's diversity index. Genetic structure was inferred by STRUCTURE version 2.3.1 (Pritchard et al., 2000). The number of genetic clusters (K) was set from 1 to 20 with a burn-in period of 50000 steps followed by a run with 100000 iterations. Twenty independent runs were undertaken for each K value. The mean posterior probabilities [Lnp(D)] values of each K were calculated according to Pritchard et al. [15], along with ΔK [16] to explore the optimum number of clusters (K). We determined the most likely number of clusters using structure harvester [17].

Results

SSR polymorphisms

Encoding binary digit format for genotyping sequence format to exploit the utility of potential core SSRs to fingerprint *Pueraria* accessions. The utilization efficiency and 23 primers information are shown in Table 1 and Figure S1. A total of 118 alleles were detected among 272 *Pueraria* individuals, leading to a mean number of alleles per locus of 5.13 (ranging from two for PsSSR155 and PtSSR222, to 11 for PtSSR98 and PtSSR112). A total of 112 polymorphic alleles (94.92%) was identified with an average of 4.87 effective alleles per locus. Among the 118 alleles, 11 (9.4%) were rare alleles with frequency less than 1% and four of them were found to be only once in one individual. The average of the observed number of alleles (N_a) and the effective number of alleles (N_e) were 1.9492 and 1.2841, respectively.

Population genetic diversity

The population-level genetic diversity of the *Pueraria* accessions under study is present in Table 2. Nei's (1973) gene diversity ranged from 0 to 0.5 and Shannon's information index (I) ranged from 0 to 0.6931 across all 23 SSR loci with an average of 0.1778 and 0.2858, respectively. The average value of total expected heterozygosity (H_t) and N_m were recorded at 0.1841 and 1.7690, respectively. The degree of genetic differentiation among populations may decrease due to the existence of large gene flow ($N_m > 1$).

Genetic structure analysis

The clustering analyses using STRUCTURE under the admixture model suggested the optimum K was 2 by STRUCTURE HARVEST [17], which divided all sampled individuals into two groups. Correspondingly, the highest of adhoc measure (ΔK) analysis [16] gave a sharp peak at K=2 (Fig. 2A). Hence, the true number of groups were considered as two (Pop1 and Pop2). The accessions with the probability of more 80% were considered as pure and assigned to corresponding subgroups while less than 80% were categorized as admixture (Fig. 2B). Among 272 genotypes, 259 were pure and 13 *Pueraria* accessions were admixture. Subpopulation P1 showed 152 pure (97.5%) and 4 admixed (2.5%) landraces, P2 had 107 pure (92.2%) and 9 (7.8%) admixed landraces. In addition, all of the 272 individuals could be clustered into one of four groups when K=4 (Fig. 2C). However, within each of the four closely related species, a few individuals always contained an admixture of introgressed genetic material from another species.

Cluster analysis and Principal Component Analyses (PCA)

Although there was no clear demarcation in the clustering pattern in the present study, the UPGMA dendrogram (Fig. 3) showed that all the accessions were divided into two main clusters at 0.378 similarity coefficient, which showed similar results as structure analysis. Furthermore, 272 accessions were divided into four main clusters at 0.684 similarity coefficient. The minimum similarity is 0.587 for most other accessions (Fig. 3). There was no distinctive trend of accessions in these two clusters according to their place of origins. For instance, accessions from Longzhou county of Chongzuo (LZ-9 to LZ-13), were covered within all two clusters with no evident bias.

The PCA categorized all the accessions undertaken into two groups, which was in line with the results of UPGMA based phylogenetic tree and model-based STRUCTURE analysis. The first two axes of differentiation explained 89% of the total variation. The first coordinate explained 40% of the variation and the second coordinate explained 49% of the variation (Fig. 4). The result of PCA indicated that the genetic distance does not show a relationship with geographical distribution in this study.

Extraction of a core collection

105 SSR alleles found in this study could be represented by a core collection of 20 *Pueraria* accessions with 7.35% sampling proportion (Fig. S2, Table 3). When the core selection capacity reached 20, the allele number was 105, so it captured close to 93.75% (105/112) of the total polymorphic loci. The average of the value of N_a, N_e, h, I were 1.8898, 1.3716, 0.2359, and 0.3727, respectively. Based on the dendrogram, the

germplasm accessions could be divided into two main groups. The value of genetic similarity indices among 20 *Pueraria* germplasm clones varied between 0.31 and 0.60, which indicates that there was a relatedly narrow genetic variation within the different kudzu species belonging to the diverse geographic locations across the Guangxi (Fig. 5). In addition, our COREFINDER analysis highlighted that 10% of the entire core collection was represented by the *Pueraria* accession grouped in Cluster I, while Cluster II contribute to the core collection at 90%. Overall, the core collection identified in this study represented 7.3% of the *Pueraria* germplasm collection.

Discussion

Polymorphism of newly developed SSR markers

We detected a total of 118 alleles with 23 SSRs segregating in the 272 *Pueraria* accessions in Guangxi, with an average of 5.13 alleles per locus. This value is lower than the number of alleles per SSR locus reported in Zhou et al. [10]. The number of effective alleles per locus (4.87) obtained in the Guangxi *Pueraria* accessions appears to be higher than the number of effective alleles per SSR locus found in other crops, such as the value of 2.26 reported in rice [18], 3.17 in olive [19], but lower than the values of 7.2 in maize [20]. Moreover, our results showed that SSR allelic diversity of *Pueraria* germplasm was moderate (N_a =1.9492, N_e = 1.2841, h= 0.1778). Zhou et al. [10] reported an average of N_e = 1.4503 and h = 0.2865 in a collection of 184 *Pueraria* accessions from Jiangxi. The number of markers and individuals, the sexual propagules and type of plant material, the population size may be responsible for the level of polymorphism and discrimination power.

Genetic diversity of *Pueraria* species

Our results revealed that *Pueraria* species display moderate genetic variation throughout Guangxi. In our study, the STRUCTURE results revealed that the 272 accessions could be divided into two main clusters and admixed individuals, while the UPGMA dendrogram showed that 272 accessions were divided into two main clusters with 37.8% genetic similarity, four main clusters with 68.4% genetic similarity. With evidence for several admixtures within cluster I (code_collection number: 30_JCJ-30, 32_JCJ-32, 196_GL-32, 197_GL-33) or cluster II (code_collection number: 12_YZ-12, 26_LC-26, 27_LC-27, 28_HJ-28, 113_GP-21, 149_BS-13, 160_BS-24, 195_GL-31, 270_Y10), the genetic differentiation was not in congruence with their eco-geographical distribution in this study. Furthermore, the overall clustering patterns generated by the STRUCTURE and PCA did not clearly distinguish the sampling areas, which is consistent with the previous results [6,9,11,21]. The low genetic differentiation indicated that geographical isolation may not restrict gene exchange among *Pueraria* species populations in Guangxi. It is susceptible to external factors even though there was a certain correlation between genetic variation and geographical distribution based on RAPD in several studies [8,22,23]. As a result, it is thought that *Pueraria* species has been cultivated and utilized for a long period in Guangxi since native cultivars of *Pueraria* still exist in the major regions, which is similar to *Perilla* in Korea [24]. The selection by humans could be responsible for this clustering pattern and moderate genetic diversity.

However, previous studies revealed that *Pueraria* species possessed from moderate to the high level of genetic diversity with high clonal reproduction and perennial [10,11,20,21,25-29]. The inconsistencies observed, except for various taxon sampling and markers, could have originated by 1) the populations were found by sexual propagules could contribute to the maintenance of high genetic variation in clonal populations regardless of recruitment of sexual offspring [30]; 2) introductions from across its multiple native populations into novel habitats from seed stock [29]; 3) clonal populations with fewer genotypes still maintain higher genetic diversity at each locus [31].

Moreover, *Pueraria* species, as strictly self-pollinating and clonally persisting clumps plants, have considered heterozygosity (Table 2), like many clonal plants, e.g. *Castanea dentata* [32] and *Musa balbisiana* [33]. Our results showed that relatively low H_t (0.1841) and H_s (0.1435), which suggest that accessions were inbred due to little outcrossing during maintenance [34]. Moreover, we could not be ruled out a case that the existence of ancient clonality and the somatic mutation, which accumulates genetic variation within clonally persisting clumps may account for some of the heterozygosity, especially given rapid mutation of SSR fingerprints.

Implications for utilization of core germplasm resources in Guangxi

Core germplasm plays a key role in the conservation, management, and utilization of germplasm resources, which is critical for the development of plant breeding [35]. Individuals reflecting genetic information can be selected to build the core germplasm resources. China is the center of distribution of *Pueraria*, with a long history of growing *Pueraria* species. However, fewer excellent *Pueraria* germplasm has been established due to artificial over-mining, lack of conservation, and management of resources. *Pueraria* species are abundant in Guangxi, especially in Tengxian and Wuzhou [12]. *Pueraria* resources have a low level of genetic differentiation (N_m = 1.7690). Previous researches have shown that sampling proportion between 5 and 30% is enough to include at least 80% of the alleles representing the genetic diversity of the entire collection [36,37]. According to dynamic extracted results, our results revealed that when the samples collected reached 7.35% (20/272) of *Pueraria* accessions accounted for 105 alleles, accounting for approximately 93.75% of all alleles loci. Interestingly, the retention value of *Pueraria* core collection genetic diversity was lower than the allele retention values of 100%, 100%, and 97.5% in rosewood, licorice, eggplant, with sampling ratios reaching 12.4% [38], 16.84% [38] and 12.03% [40], respectively. The most likely reason was that the breeding of a majority of *Pueraria* accessions in Guangxi was still from layering breeding and self-crossing, and lacked extensive gene exchanges from cross-breeding, which led to a decrease in the ratio of

the core collection. Our findings will be useful in breeding programs for the introgression of noble alleles into modern cultivars by exploiting natural genetic variation existing in *Pueraria* genetic resources. Combined with the analysis of phenotypic diversity (e.g. puerarin, starches) of *Pueraria* species, we may detect the important polymorphic loci associated with the traits based on correlation analysis, which could provide a foundation for developing the molecular marker-assisted breeding or detection of target genes soon [4].

Complex genetic relationships among *Pueraria* species

Our results include new clues in genetic relationships among *Pueraria* species based on SSR markers. 13 *Pueraria* individuals of 272 accession were admixture. Meanwhile, the genetic clusters were not consistent with species delimitation and geographic distribution, which imply the complex evolutionary history with the human process blur the phylogenetic relationship among these species. *Pueraria* DC. (1825; Fabaceae, Phaseoleae) comprises ca. 20 species, occurring in tropical and E. Asia. Eight species and two varieties have been recorded in China [41], with four groups or three sections as infrageneric classification based on morphological traits [42,43]. However, molecular studies have revealed that *Pueraria* is not a monophyletic group [44,45]. At the species level, the problem is more severe, where morphological intermediacy, artificial selection, biochemical characteristics, and insufficient informative characters contribute to the lack of agreement between morphological-based classification schemes and recent molecular phylogenies. For example, taxonomically kudzu is placed under the genus *Pueraria*. *Pueraria lobata* var. *thomsonii* and *P. lobata* var. *montana* were treated as varieties for *P. lobata* in Flora of China. However, the phylogenetic relationship and classification among these three species are still confused based on various molecular markers and sampling taxon. There were two main taxonomic treatments: firstly, *Pueraria lobata* var. *thomsonii* and *P. lobata* var. *montana* should be treated as a single species respectively, other than varieties [46]; secondly, *Pueraria lobata* var. *thomsonii* should be treated as the variety to *Pueraria lobata*, while *P. lobata* var. *montana* may be a single species [47]. A wider taxon sampling with higher resolution genetic markers would increase confidence for the phylogenetic relationship among *Pueraria* species, efforts that are currently underway.

Declarations

Compliance with ethical standards

Conflict of Interest The authors declare that there is no conflict of interest arising from this work.

Ethical approval No human participants or animals were involved in this research.

Acknowledgments

This study was supported by the National Natural Science Foundation of China (31960420), Guangxi Key R&D Program Project (Guike AB1850028), Guangxi Natural Science Foundation Project (2018GXNSFBA294001 and 2019GXNSFBA245093), Science and Technology Development Fund of Guangxi Academy of Agricultural Sciences (Guinongke 2021JM11), and Special project for basic scientific research of Guangxi Academy of Agricultural Sciences (Guinongke 2021YT057).

Authors contributions

Yan H.B. planned and supervised the present research. Shang X.H. executed the whole experiments, statistical analysis and interpretation of data. Field experiments were sown and maintained by Wang Y., Zeng W.D., Cao S., Wu Z.D. Data analysis and writing of manuscript was done by Shang X.H. and Zhou Y. Zhou Y, Xiao L and Yan H.B. gave needful suggestions in manuscript.

References

1. Wang S, Zhang S, Wang S, Gao P, Dai L (2020) A comprehensive review on *Pueraria*: Insights on its chemistry and medicinal value. Biomed Pharmacother 131: 110734. https://doi.org/10.1016/j.biopha.2020.110734
2. Chen C, Zheng L, Ma Q, Zhou WB, Liu Y, Zhao YP, Fu CX (2018) Impacts of domestication on population genetics of a traditional Chinese medicinal herb, *Atractylodes macrocephala* (Asteraceae). J Syst Evol 57(3): 222–233. https://doi.org/10.1111/jse.12446
3. Ambati D, Phuke RM, Vani V, Prasad SVS, Singh JB, Patidar CP, Malviya P, Gautam A, Dubey VG (2020) Assessment of genetic diversity and development of core germplasm in durum wheat using agronomic and grain quality traits. Cereal Res Commun 48: 375–382. https://doi.org/10.1007/s42976-020-00050-z
4. Chen C, Chu Y, Ding C, Su X, Huang Q (2020) Genetic diversity and population structure of black cottonwood (*Populus deltoides*) revealed using simple sequence repeat markers. BMC Genet 21(1): 1–12. https://doi.org/10.1186/s12863-019-0805-1
5. Jing X, Xu L, Chen JY, Zeng XQ, Huang ZC, Zhu ZJ, Li H, Yuan C, Wang ZH, Li SM, Zhang JW, Chen DQ (2010) Genetic diversity of arrowroot (*Pueraria* L.) varieties revealed by RAPD analysis in Chongqing area. Chin Agric Sci Bull 26(24): 80–82.
6. Chen DX, Peng R, Li LY, Zhang X, Wang Y (2011) Analysis of genetic relationships of *Pueraria thomsonii* based on SRAP markers. China J Chin Mater Med 36(5): 538–541.

7. Guo YY, Chen CY, Huang JL, Yang XD, Lu JS, Wang AQ, He LF (2013) ISSR analysis of genetic relationships in Radix *Puerariae* from different original place. Popular Sci Technol 15(4): 134–136.
8. Zhou JH, Jie YC, Du XH, Xing HC, Xiong LF (2013) RAPD analysis on genetic relationship of Kudzu germplasm resources. Crop Res 27(04): 3447–350.
9. Yuan C, Zhong WJ, Gong YY, Pu DQ, Ji PC, Huang HY, Yang ZH, Zhang C (2017) Genetic diversity and trait association analysis of *Pueraria lobata* resources. J Plant Genet Resour 18(2): 233–241.
10. Zhou RR, Zhou JH, Nan TG, Jiang C, Duan HY, Zhao YP, Yuan Y (2019) Analysis of genomic SSRs in *Pueraria lobata* and *P. thomsonii* and establishment of DNA identity card for different germplasms of *P. thomsonii* of Jiangxi province. China J Chin Mater Med 44(17): 3615–3621.
11. Xiao L, Shang XH, Cao S, Xie XY, Zeng WD, Yan H.B (2019) Utilization of simple sequence repeat(SSR)markers developed from a de novo transcriptome assembly in *Pueraria thomsonii* benth. Acta Bot Boreal -Occid Sin 39(1): 0059–0067.
12. Shang XH, Cao S, Xiao L, Yan HB, Wang Y, Zeng WD, Xie XY, Lu LY, Lai DX (2020) Investigation and collection of *Pueraria* germplasm resources in Guangxi. J Plant Genet Resour 21(5): 1301–1307.
13. Rohlf FJ (2000) NTSYS pc2.1: Numerical taxonomy and multivariate analysis system version 2.1. New York: Applied Biostatistics Inc.
14. Thachuk C, Crossa J, Franco J, Dreisigacker S,Warburton M, Davenport GF (2009) Core hunter: an algorithm for sampling genetic resources based on multiple genetic measures. BMC Bioinformatics 10(1): 243. https://doi.org/10.1186/1471-2105-10-243
15. Pritchard JK, Stephens MJ, Donnelly PJ (2000) Inference of population atructure using multilocus genotype data. Genetics 155(2): 945–959. https://doi.org/10.1093/genetics/155.2.945.
16. Evanno GS, Regnaut SJ, Goudet J (2005) Detecting the number of clusters of individuals using the software structure: a simulation study. Mol Ecol 14(8): 2611–2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x
17. Earl DA, vonHoldt BM (2012) Structure harvester: a website and program for visualizing structure output and implementing the evanno method. Conserv Genet Resour 4(2): 359–361. https://doi.org/10.1007/s12686-011-9548-7
18. Verma H, Borah JL, Sarma RN (2019) Variability assessment for root and drought tolerance traits and genetic diversity analysis of rice germplasm using SSR markers. Sci Rep 9(1): 1–19. https://doi.org/10.1038/s41598-019-52884-1
19. Gómez-Rodríguez MV, Beuzon C, González-Plaza JJ, Fernández-Ocaña AM (2021) Identification of an olive (*Olea europaea* L.) core collection with a new set of SSR markers. Genet Resour Crop Evol 68(1): 117–133. https://doi.org/10.1007/s10722-020-00971-y
20. Park JY, Ramekar RV, Sa KJ, Lee JK (2015) Genetic diversity, population structure, and association mapping of biomass traits in maize with simple sequence repeat markers. Genes Genom 37: 725–735. https://doi.org/10.1007/s13258-015-0309-y
21. Pappert RA, Hamrick JL, Donovan LA (2000) Genetic variation in *Pueraria lobata* (Fabaceae), an introduced, clonal, invasive plant of the southereastern United States. Am J Bot 87: 1240–1245. https://doi.org/10.2307/2656716
22. Heider B, Fischer E, Berndl T, Schultze-Kraft R (2007) Analysis of genetic variation among accessions of *Pueraria montana* (Lour.) Merr. var. *lobata* and *Pueraria phaseoloides* (Roxb.) Benth. based on RAPD markers. Genet Resour Crop Evol 54: 529–542. https://doi.org/10.1007/s10722-006-0009-1
23. Ji BY, Pei LX, Chen SQ, Dong CM, Feng WS (2014) RAPD analysis of germplasm resource in *Pueraria lobata*. Chin J Exp Tradit Med Formulae 20(6): 56–59.
24. Oh JS, Sa KJ, Hyun DY, Cho GT, Ju KL (2020) Assessment of genetic diversity and population structure among a collection of Korean *Perilla* germplasms based on SSR markers. Genes Genom 42(12): 1419–1430. https://doi.org/10.1007/s13258-020-01013-8
25. Jewett DK, Jiang CJ, Britton KO, Sun JH, Tang J (2003) Characterizing specimens of kudzu and related taxa with RAPDs. Castanea 68: 254–60.
26. Hoffberg SL, Bentley KE, Lee JB, Myhre KE, Iwao K, Glenn TC, Mauricio R (2015) Characterization of 15 microsatellite loci in kudzu (*Pueraria montana* var. *lobata*) from the native and introduced ranges. Conserv Genet Resour 7: 403–405. https://doi.org/10.1007/s12686-014-0381-7
27. Sun JH, Li ZC, Jewett DK, Britton KO, Ye WH, Ge XJ (2005) Genetic diversity of *Pueraria lobata* (kudzu) and closely related taxa as revealed by inter-simple sequence repeat analysis. Weed Res 45: 255–60. https://doi.org/10.1111/j.1365-3180.2005.00462.x
28. Bentley K, Mauricio R (2016) High degree of clonal reproduction and lack of large-scale geographic patterning mark the introduced range of the invasive vine, kudzu (*Pueraria montana* var. *lobata*) in North America. Am J Bot 103: 1499–1507. https://doi.org/10.3732/ajb.1500434
29. Haynsen MS (2018) Population Genetics of *Pueraria montana* var. *lobata* (Doctoral dissertation, The George Washington University).
30. Ellstrand NC, Roose ML (1987) Patterns of genotypic diversity in clonal plant species. Am J Bot 74: 123–131.
31. Halkett FJ, Simon JC, Balloux FO (2005) Tackling the population genetics of clonal and partially clonal organisms. Trends Ecol Evol 20: 194–201. https://doi.org/10.1016/j.tree.2005.01.001
32. Stilwell KL, Wilbur HM, Werth CR, Taylor DR (2003) Heterozygote advantage in the American chestnut, *Catanea dentata* (Fagaceae). Am J Bot 90: 207–213. https://doi.org/10.3732/ajb.90.2.207
33. Ge XJ, Liu MH, Wang WK, Schaal BA, Chiang TY (2005) Population structure of wild bananas, *Musa balbisiana*, in China determined by SSR fingerprinting and cpDNA PCR-RFLP. Mol Ecol 14(4): 933–944. https://doi.org/10.1111/j.1365-294X.2005.02467.x

34. Kiran BU, Mukta N, Kadirvel P, Alivelu K, Senthilvel S, Kishore P, Varaprasa, KS (2015) Genetic diversity of safflower (*Carthamus tinctorius* L.) germplasm as revealed by SSR markers. Plant Genet Resour 15(1): 1–11. https://doi.org/10.1017/S1479262115000295
35. Singh N, Wu S, Raupp WJ, Sehgal S, Arora S, Tiwari V, Poland J (2019) Efficient curation of genebanks using next generation sequencing reveals substantial duplication of germplasm accessions. Sci Rep 9(1): 1–10. https://doi.org/10.1038/s41598-018-37269-0
36. Singh M, Bisht IS, Kumar S, Dutta M, Bansal KC, Karale M, Sarker A, Amri A, Kumar S, Datta SK (2014) Global wild annual *Lens* collection: a potential resource for lentil genetic base broadening and yield enhancement. PLoS ONE 9: e107781. https://doi.org/10.1371/journal.pone.0107781
37. Nie XH, Wang ZH, Liu NW, Li S, Yan BQ, Xing Y, Zhang Q, Fang KF, Zhao YL, Chen X, Wang GP, Qin L, Cao QQ (2021) Fingerprinting 146 Chinese chestnut (*Castanea mollissima* Blume) accessions and selecting a core collection using SSR markers. J In Agr 20(5): 1277–1286. https://doi.org/10.1016/S2095-3119(20)63400-1
38. Liu FM, Zhang NN, Liu XJ, Yang ZJ, Jia HY, Xu DP (2019) Genetic diversity and population structure analysis of *Dalbergia Odorifera* Germplasm and development of a core collection using microsatellite markers. Genes (Basel) 6: 10(4): 281.
39. Liu Y, Geng Y, Xie X, Zhang P, Hou J, Wang W (2020) Core collection construction and evaluation of the genetic structure of *Glycyrrhiza* in China using markers for genomic simple sequence repeats. Genet Resour Crop Evol 67(7): 1839–1852. https://doi.org/10.1007/s10722-020-00944-1
40. Miyatake K, Shinmura Y, Matsunaga H, Fukuoka H, Saito T (2019) Construction of a core collection of eggplant (*Solanum melongena* L.) based on genome-wide SNP and SSR genotypes. Breeding Sci 69(3): 498–502. https://doi.org/10.1270/jsbbs.18202
41. Wu DL, Thulin M (2010) *Pueraria*. Flora of China 10: 244–248.
42. Lacker JA (1977) A synopsis of the Phaseoleae (Leguminosae, Papilionoideae). Ph.D. dissertation, Iowa State University, Ames, Iowa, pp. 293.
43. Van der Maesen LJG (1985) Revision of the genus *Pueraria* DC. with some notes on Teyleria Backer (Leguminiosae). Agricultural University, Wageningen, pp. 5–13.
44. Stefanovic S, Pfeil BE, Palmer JD, Doyle JJ (2009) Relationships among phaseoloid legumes based on sequences from eight chloroplast regions. Syst Bot 34: 115–128. https://doi.org/10.1600/036364409787602221
45. Egan AN, Vatanparast M, Cagle W (2016) Parsing polyphyletic *Pueraria*: Delimiting distinct evolutionary lineages through phylogeny. Mol Phylogenet Evol 104: 44–59. https://doi.org/10.1016/j.ympev.2016.08.001
46. Zeng M, Yan JZ, Zhang HM, Zheng SQ, Su ZW (2000) Classification and authentication of plant *Pueraria* DC. in China using RAPD. Chin Tradit Herbal Drugs 31(8): 620–622.
47. Zeng M, Ma YJ, Zheng SQ, Xu JF, Di XH (2003) Studies on ribosomal DNA sequence analyses of Radix puerariae and its sibling species. Chin Pharmacol J 38(3): 173–175.

Tables

Table1 Amplification results and polymorphism information of 23 SSR primers

No.	Primer name	Sequences(5'-3')	Total number of lands	Number of polymorphic bands	Polymorphism rate/%
1	PtSSR36	Fw: CTGAGTCTCTGCAAAGCCCA	10	10	100
		Rv: TGTCACTGTGCTCCAACTCC			
2	PtSSR98	Fw: CATTCGGACCTCCATACCCG	11	10	90.9
		Rv: CCGCATCCAACCCTGATCAA			
3	PtSSR99	Fw: GCTTTCCGCTGCTACCATTC	7	7	100
		Rv: GCAACCCCAATGCTTCACAG			
4	PtSSR104	Fw: CACCCTCCCACCACTACAAC	3	3	100
		Rv: GCAATGTCCTCCTCAGCTGT			
5	PtSSR108	Fw: AGCGTGCCCAACTCAGTTAA	3	3	100
		Rv: CGACGGAGAAGGAGGGAATG			
6	PtSSR109	Fw: CAACCTGGCTTCTGTTGTGC	5	4	80
		Rv: CTCTGAAACGCTGGGCAATG			
7	PtSSR121	Fw: ACACTCAACACTCCACCACC	3	2	66.67
		Rv: AGGGTTTCCACCTTGAACCG			
8	PtSSR122	Fw: GGGGTTTCTTCTCGGCTGAA	11	11	100
		Rv: CACCCCCTTCACGCTTCATA			
9	PtSSR130	Fw: ATCAGTGTCTACGTGGGGGA	5	4	80
		Rv: CACTGCAGCCACAACAACAT			
10	PtSSR135	Fw: GATCCGCACCCTATCTGTGG	8	8	100
		Rv: CTGCGACAGCTCCGATCTTA			
11	PtSSR144	Fw: TGTTGCTTTGAACACTAACATGCT	3	2	100
		Rv: TGCCCTTGTCAGACACAACA			
12	PtSSR155	Fw: TTCAACATTCCCCCAACCCC	2	2	100
		Rv:AAGAAGAGGAACACCAGGCC			
13	PtSSR168	Fw: GATCCCACCCACCACTTCTG	5	5	100
		Rv: GGCTCTAGTTCTGGTGCTGG			
14	PtSSR172	Fw:TCTCCAAAACAAGAAGGAAACTCC	4	3	75
		Rv: TCTTTCCTCTTCTGGTATCCCA			
15	PtSSR174	Fw: CAAAGAAGAAGCAGCCGCAG	6	6	100
		Rv: GTCAATCCCGAAGCACTTGC			
16	PtSSR175	Fw: CTGAGTCTCTGCAAAGCCCA	7	7	100
		Rv: TGTCACTGTGCTCCAACTCC			
17	PtSSR186	Fw: TGTTGCTTTGAACACTAACATGCT	4	4	100
		Rv: TGCCCTTGTCAGACACAACA			
18	PtSSR187	Fw: TGTTGCTTTGAACACTAACATGCT	4	4	100
		Rv: TGCCCTTGTCAGACACAACA			
19	PtSSR190	Fw: AACTGCAGGAGGAGCATGAC	5	5	100
		Rv: GAGCCTCCAGGTTCTTGTCC			

20	PtSSR191	Fw: GGAAGCATTGCGGTTTGGTT	3	3	100
		Rv:TCACATCACATGCTGCCACT			
21	PtSSR196	Fw: GCAAGAACCTGTGCTCCTCT	3	2	66.67
		Rv: TGCCAATGCCATTGTGGTTG			
22	PtSSR201	Fw:GCCTCTTCCAGCGAGAACTT	4	4	100
		Rv: TGATCCTCCCCAACAAGCTG			
23	PtSSR222	Fw: TGTGCAAGAAGGATGGGTGA	2	2	100
		Rv: GGTTGCATTCGGAAGCAACA			
Total			118	112	
Avarage			5.08	4.87	95.72

Table2 Genetic characteristics for 112 polymorphic microsatellite loci in 272 individuals of *Pueraria* species in the present study.

Locus	Sample Size	N_a*	N_e*	h*	I*	H_t	H_s	G_{st}	N_m*
36-1	272	2	1.0112	0.011	0.0344	0.0129	0.0128	0.0066	75.8301
36-2	272	2	1.0112	0.011	0.0344	0.0129	0.0128	0.0066	75.8301
36-3	272	2	1.0463	0.0443	0.1082	0.0517	0.0503	0.0273	17.8197
36-4	272	2	1.0074	0.0074	0.0244	0.0086	0.0086	0.0043	114.4978
36-5	272	2	1.1017	0.0923	0.1941	0.1063	0.1007	0.0531	8.9206
36-6	272	2	1.0843	0.0778	0.1696	0.0905	0.086	0.0499	9.5227
36-7	272	2	1.2654	0.2098	0.3650	0.2394	0.2025	0.1541	2.7446
36-8	272	2	1.9322	0.4824	0.6755	0.4915	0.4157	0.1543	2.7395
36-9	272	2	1.034	0.0329	0.0849	0.0309	0.0308	0.0032	156.1603
36-10	272	2	1.0074	0.0074	0.0244	0.0086	0.0086	0.0043	114.4978
98-1	272	2	1.0188	0.0184	0.0527	0.0216	0.0213	0.011	44.8945
98-2	272	2	1.1442	0.126	0.2473	0.145	0.1336	0.0786	5.8651
98-3	272	2	1.1499	0.1304	0.2539	0.1509	0.1374	0.0895	5.0838
98-4	272	2	1.0584	0.0552	0.1291	0.0634	0.0616	0.0275	17.6531
98-5	272	2	1.0383	0.0369	0.0933	0.0431	0.0421	0.0225	21.6887
98-6	272	2	1.0383	0.0369	0.0933	0.0431	0.0421	0.0225	21.6887
98-7	272	2	1.0671	0.0629	0.1433	0.0733	0.0704	0.0396	12.1273
98-8	272	2	1.0671	0.0629	0.1433	0.0733	0.0704	0.0396	12.1273
98-9	272	2	1.16	0.1379	0.2652	0.1595	0.1442	0.0957	4.7224
98-10	272	1	1.0000	0.0000	0.0000	0	0	****	****
98-11	272	2	1.0149	0.0147	0.0436	0.0161	0.016	0.003	166.8934
99-1	272	2	1.0188	0.0184	0.0527	0.0216	0.0213	0.011	44.8945
99-2	272	2	1.0615	0.058	0.1343	0.0592	0.0592	0.0007	711.3620
99-3	272	2	1.288	0.2236	0.3830	0.2029	0.1857	0.0849	5.3899
99-4	272	2	1.2065	0.1712	0.3129	0.1855	0.1781	0.04	12.0065
99-5	272	2	1.14	0.1228	0.2424	0.1422	0.1304	0.0835	5.4886
99-6	272	2	1.3643	0.267	0.4375	0.2622	0.261	0.0043	114.6222
99-7	272	2	1.5399	0.3506	0.5353	0.3448	0.3421	0.0076	65.5416
104-1	272	2	1.034	0.0329	0.0850	0.0352	0.0351	0.0039	128.2046
104-2	272	2	1.9558	0.4887	0.6818	0.4964	0.3986	0.1969	2.0392
104-3	272	2	1.9992	0.4998	0.6929	0.4957	0.3738	0.2459	1.5335
108-1	272	2	1.6292	0.3862	0.5746	0.393	0.3882	0.0122	40.3512
108-2	272	2	1.4728	0.321	0.5016	0.3331	0.3234	0.0292	16.6414
108-3	272	2	1.6831	0.4058	0.5958	0.3972	0.3883	0.0223	21.9303
109-1	272	2	1.2934	0.2269	0.3872	0.2568	0.2167	0.1562	2.7010
109-2	272	2	1.0149	0.0147	0.0436	0.0161	0.016	0.003	166.8934
109-3	272	2	1.2934	0.2269	0.3872	0.2568	0.2167	0.1562	2.7010
109-4	272	2	1.0149	0.0147	0.0436	0.0161	0.016	0.003	166.8934
109-5	272	1	1	0	0.0000	0	0	****	****
121-1	272	2	1.0753	0.07	0.1561	0.0805	0.0775	0.0374	12.8581

121-2	272	2	1.0753	0.07	0.1561	0.0805	0.0775	0.0374	12.8581
121-3	272	1	1	0	0.0000	0	0	****	****
122-1	272	2	1.0304	0.0295	0.0778	0.0345	0.0339	0.0179	27.4911
122-2	272	2	1.1351	0.119	0.2367	0.1379	0.1268	0.0805	5.7097
122-3	272	2	1.0304	0.0295	0.0778	0.0345	0.0339	0.0179	27.4911
122-4	272	2	1.0887	0.0815	0.1760	0.0948	0.0898	0.0525	9.0192
122-5	272	2	1.0887	0.0815	0.1760	0.0948	0.0898	0.0525	9.0192
122-6	272	2	1.0671	0.0629	0.1433	0.0733	0.0704	0.0396	12.1273
122-7	272	2	1.0343	0.0332	0.0856	0.0388	0.038	0.0202	24.2677
122-8	272	2	1.4253	0.2984	0.4752	0.2929	0.2912	0.0058	85.4766
122-9	272	2	1.0037	0.0037	0.0134	0.0043	0.0043	0.0022	230.4989
122-10	272	2	1.5187	0.3415	0.5250	0.3673	0.3146	0.1435	2.9847
122-11	272	2	1.2973	0.2291	0.3902	0.2608	0.2152	0.1748	2.3604
130-1	272	2	1.2025	0.1684	0.3090	0.194	0.1703	0.1221	3.5945
130-2	272	2	1.2025	0.1684	0.3090	0.194	0.1703	0.1221	3.5945
130-3	272	2	1.0383	0.0369	0.0933	0.0431	0.0421	0.0225	21.6887
130-4	272	2	1.0383	0.0369	0.0933	0.0431	0.0421	0.0225	21.6887
130-5	272	1	1	0	0.0000	0	0	****	****
135-1	272	2	1.0149	0.0147	0.0438	0.0172	0.0171	0.0088	56.4956
135-2	272	2	1.0587	0.0554	0.1296	0.0647	0.0624	0.0346	13.9494
135-3	272	2	1.3477	0.258	0.4264	0.2931	0.2295	0.2171	1.8032
135-4	272	2	1.0111	0.011	0.0343	0.0118	0.0118	0.0013	397.0141
135-5	272	2	1.3127	0.2382	0.4017	0.2716	0.219	0.1934	2.0859
135-6	272	2	1.0149	0.0147	0.0436	0.0161	0.016	0.003	166.8934
135-7	272	2	1.7785	0.4377	0.6295	0.4706	0.1885	0.5994	0.3341
135-8	272	2	1.9656	0.4913	0.6844	0.4999	0.1558	0.6884	0.2263
144-1	272	2	1.0343	0.0332	0.0856	0.0388	0.038	0.0202	24.2677
144-2	272	2	1.9251	0.4805	0.6736	0.479	0.4776	0.0029	174.8282
144-3	272	2	1.6998	0.4117	0.6020	0.3877	0.3213	0.1713	2.4193
155-1	272	2	1.0698	0.0652	0.1475	0.0687	0.0684	0.0049	101.9746
155-2	272	2	1.2483	0.1989	0.3507	0.2281	0.1936	0.1512	2.8079
168-1	272	2	1.501	0.3338	0.5162	0.3124	0.2825	0.0956	4.7296
168-2	272	2	1.9254	0.4806	0.6736	0.4892	0.4321	0.1166	3.7876
168-3	272	2	1.5113	0.3383	0.5214	0.3181	0.2904	0.0869	5.2557
168-4	272	2	1.9569	0.489	0.6821	0.4954	0.4328	0.1262	3.4613
168-5	272	2	1.0074	0.0073	0.0243	0.0075	0.0075	0.0001	2000.0000
172-1	272	2	1.2368	0.1915	0.3407	0.2198	0.1882	0.1438	2.9771
172-2	272	2	1.0112	0.011	0.0344	0.0129	0.0128	0.0066	75.8301
172-3	272	1	1	0	0.0000	0	0	****	****
172-4	272	2	1.0037	0.0037	0.0134	0.0032	0.0032	0.0016	310.4992
174-1	272	2	1.0074	0.0074	0.0244	0.0086	0.0086	0.0043	114.4978

174-2	272	2	1.074	0.0689	0.1542	0.0741	0.0734	0.0099	50.0991
174-3	272	2	1.1249	0.111	0.2241	0.1278	0.1192	0.0669	6.9725
174-4	272	2	1.0037	0.0037	0.0134	0.0043	0.0043	0.0022	230.4989
174-5	272	2	1.7565	0.4307	0.6221	0.4621	0.2446	0.4706	0.5625
174-6	272	2	1.9722	0.4929	0.6861	0.5	0.2054	0.5892	0.3486
175-1	272	2	1.0074	0.0074	0.0244	0.0086	0.0086	0.0043	114.4978
175-2	272	2	1.0932	0.0852	0.1823	0.0991	0.0937	0.0552	8.5594
175-3	272	2	1.0074	0.0074	0.0244	0.0086	0.0086	0.0043	114.4978
175-4	272	2	1.2069	0.1714	0.3132	0.1965	0.1735	0.117	3.7725
175-5	272	2	1.8504	0.4596	0.6522	0.4769	0.3654	0.2337	1.6394
175-6	272	2	1.7047	0.4134	0.6038	0.39	0.3252	0.1659	2.5132
175-7	272	2	1.0301	0.0292	0.0771	0.0298	0.0298	0.0003	1480.2742
186-1	272	2	1.0421	0.0404	0.1005	0.0462	0.0453	0.0181	27.1466
186-2	272	2	1.7726	0.4359	0.6276	0.4059	0.2739	0.3252	1.0373
186-3	272	2	1.7848	0.4397	0.6316	0.4698	0.2325	0.5051	0.4899
186-4	272	2	1.6977	0.411	0.6013	0.3808	0.2793	0.2664	1.3765
187-1	272	2	1.042	0.0403	0.1002	0.0449	0.0444	0.0122	40.4142
187-2	272	2	1.7641	0.4331	0.6247	0.4021	0.2655	0.3397	0.9718
187-3	272	2	1.875	0.4667	0.6594	0.4905	0.1571	0.6796	0.2357
187-4	272	2	1.6636	0.3989	0.5883	0.3687	0.2779	0.2465	1.5286
190-1	272	2	1.0993	0.0903	0.1908	0.0833	0.082	0.0164	29.9125
190-2	272	2	1.046	0.044	0.1076	0.0385	0.0377	0.02	24.4900
190-3	272	2	1.1535	0.1331	0.2580	0.1211	0.1166	0.037	13.0124
190-4	272	2	1.6024	0.3759	0.5634	0.4076	0.2996	0.2651	1.3862
190-5	272	2	1.9999	0.5	0.6931	0.4956	0.323	0.3483	0.9354
191-1	272	2	1.9284	0.4814	0.6745	0.498	0.1126	0.7738	0.1462
191-2	272	2	1.6789	0.4044	0.5942	0.3708	0.2541	0.3146	1.0893
191-3	272	2	1.1641	0.141	0.2696	0.1261	0.1192	0.0552	8.5656
196-1	272	2	1.2295	0.1867	0.3342	0.2136	0.1856	0.1312	3.3100
196-2	272	2	1.0074	0.0074	0.0244	0.0086	0.0086	0.0043	114.4978
196-3	272	1	1	0	0.0000	0	0	****	****
201-1	272	2	1.0037	0.0037	0.0134	0.0043	0.0043	0.0022	230.4989
201-2	272	2	1.9978	0.4995	0.6926	0.4989	0.3531	0.2923	1.2103
201-3	272	2	1.404	0.2877	0.4625	0.3224	0.2509	0.2219	1.7536
201-4	272	2	1.4033	0.2874	0.4621	0.2798	0.2768	0.011	45.0326
222-1	272	2	1.3039	0.233	0.3951	0.2651	0.2176	0.1791	2.2910
222-2	272	2	1.7711	0.4354	0.6271	0.4617	0.3031	0.3435	0.9556
Mean	272	1.9492	1.2841	0.1778	0.2858	0.1841	0.1435	0.2204	1.7690
St. Dev		0.2206	0.3277	0.1741	0.2397	0.0305	0.0166		

* N_a = Observed number of alleles

* N_e = Effective number of alleles [Kimura and Crow (1964)]

* h = Nei's (1973) gene diversity

* I = Shannon's Information index [Lewontin (1972)]

* G_{st} = coefficient of gene differentiation

* N_m = estimate of gene flow from Gst or Gcs. E.g., Nm = 0.5(1 - Gst)/Gst;

* H_t = Total expected hetrogygosity,

* H_s =the average expected heterozygosity within subpopulations

Table 3 Summary of the extraction of a core collection.

Sampling proportion	Sample number	N_a	N_e	h	I	Number of polymorphic loci	Percentage of polymorphic loci	Percentage of polymorphic loci
5%	14	1.8644±0.3438	1.3839±0.3116	0.2413±0.1639	0.3778±0.2243	102	91.07%	86.44
7.00%	19	1.8644±0.3438	1.3779±0.3123	0.2381±0.1635	0.3736±0.2242	102	91.07%	86.44
7.35%	**20**	**1.8898±0.3144**	**1.3716±0.3084**	**0.2359±0.1598**	**0.3727±0.2170**	**105**	**93.75%**	**88.98**
7.70%	21	1.8983±0.3035	1.3658±0.3063	0.2333±0.1580	0.3702±0.2139	106	94.64%	89.83
8%	22	1.8983±0.3035	1.3655±0.3049	0.2333±0.1581	0.3699±0.2145	106	94.64%	89.83
10%	27	1.8983±0.3035	1.3577±0.2991	0.2297±0.1575	0.3648±0.2154	106	94.64%	89.83
15%	41	1.9068±0.2920	1.3451±0.3055	0.2208±0.1613	0.3518±0.2203	107	95.54%	90.68
20%	54	1.9237±0.2666	1.3384±0.3106	0.2161±0.1621	0.3459±0.2195	109	97.32%	92.37
30%	82	1.9322±0.2525	1.3204±0.3052	0.2060±0.1625	0.3314±0.2217	110	98.21%	93.22
40%	109	1.9407±0.2372	1.3180±0.3158	0.2024±0.1664	0.3249±0.2272	111	99.11%	94.07
50%	136	1.9322±0.2525	1.3146±0.3273	0.1980±0.1703	0.3171±0.2327	110	98.21%	93.22
100%	272	1.9492±0.2206	1.2841±0.3277	0.1778±0.1741	0.2858±0.2397	112		94.92

* N_a = Observed number of alleles

* N_e = Effective number of alleles [Kimura and Crow (1964)]

* h = Nei's (1973) gene diversity

* I = Shannon's Information index [Lewontin (1972)]

Figures

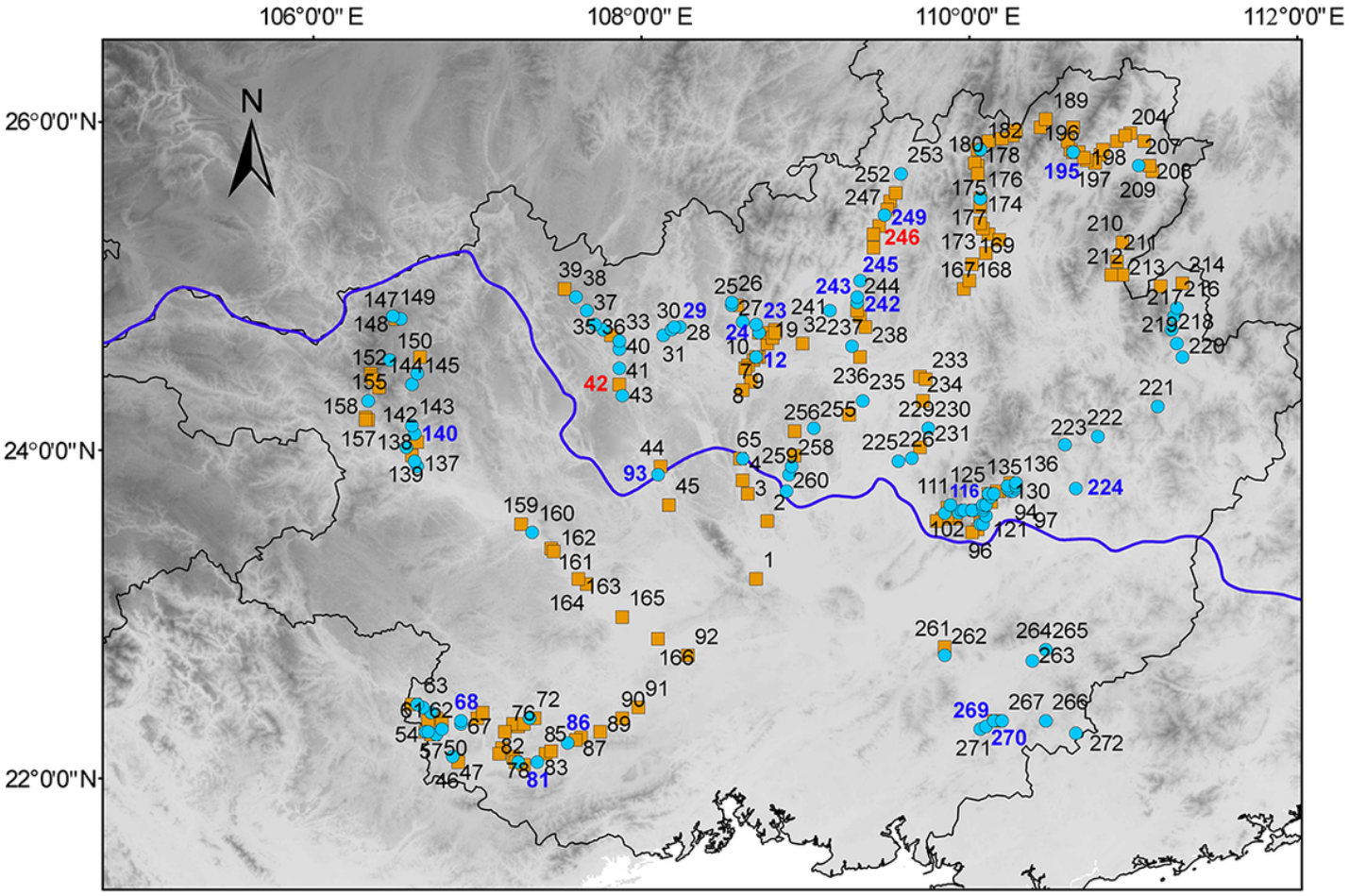

Figure 1

Geographical distribution of the accession collected in Guangxi. The number represents the code in Table S1.

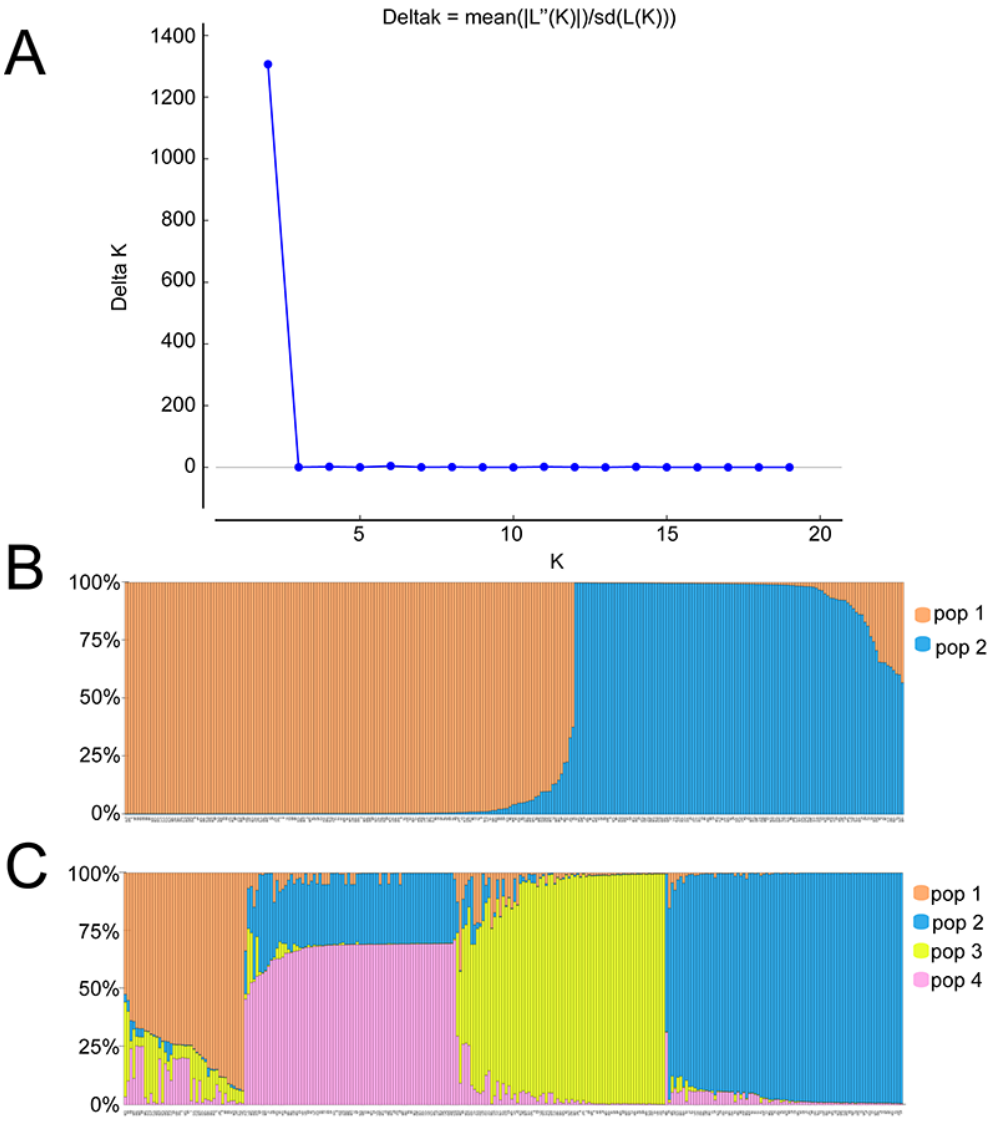

Figure 2

Bar plots of all 272 individuals from Pueraria germplasm grouped into two or four genetic clusters with assignment probabilities obtained from STRUCTURE analyses of polymorphisms at 23 simple sequence repeat loci. A, Distribution of delta K=1-20. B, C, Histogram of the STRUCTURE assignment test when K=2 or K=4, respectively. The number represents the code in Table S1.

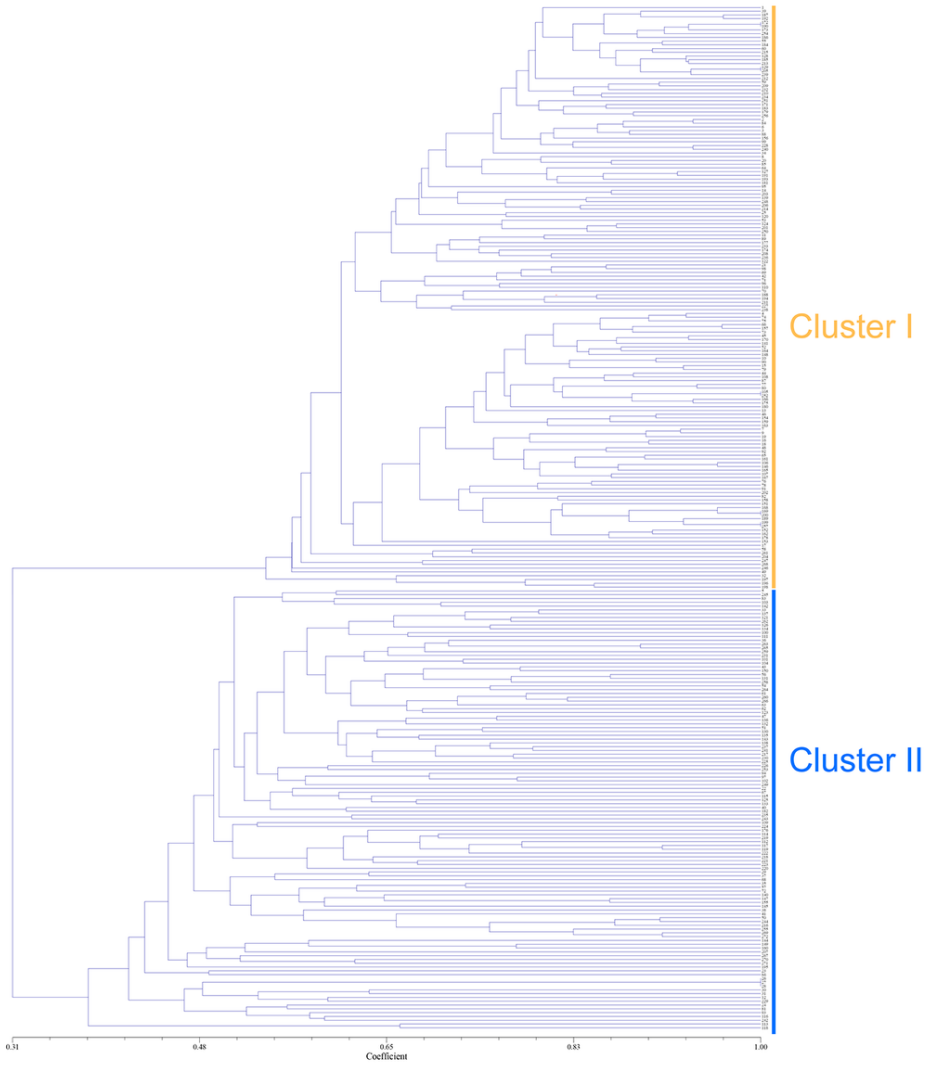

Figure 3

Cluster diagram based on jaccard by UPGMA analysis calculated from alleles derived from 272 Pueraria accessions. The number represents the code in Table S1.

Figure 4

PCA of Pueraria accessions based on dissimilarity matrix (Jaccard). The number represents the code in Table S1.

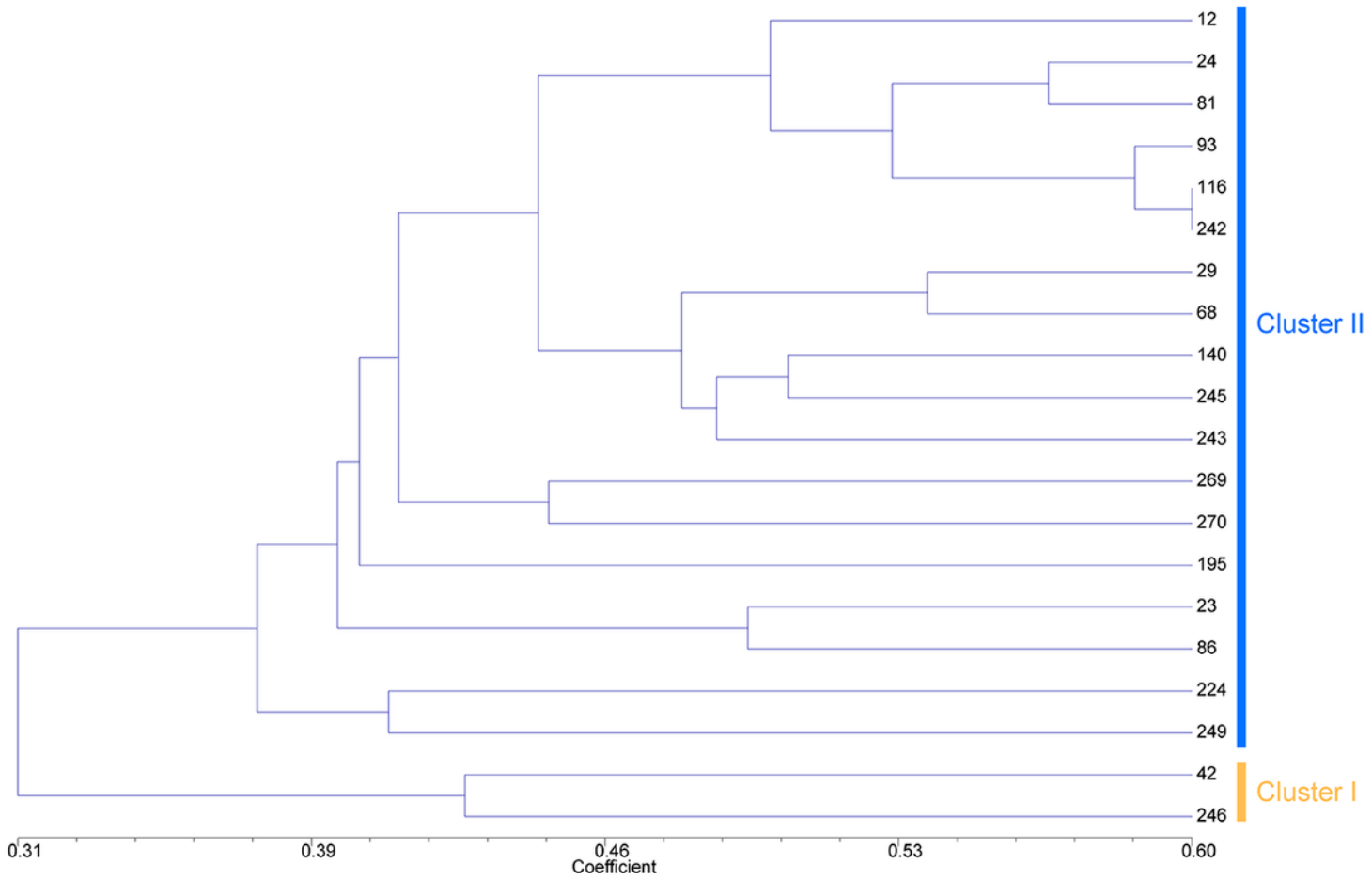

Figure 5

Cluster diagram based on jaccard by UPGMA analysis calculated from alleles derived from 20 Pueraria accessions of core germplasm. The number represents the code in Table S1.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- FigS1finger.tif
- FigS2corefinger.tif
- TableS1.docx