Genetic Diversity and Population Structure of Traditional Chinese Herb “Chai-Hu” Resources Using Genome-Wide SNPs Through Genotyping-By-Sequencing

Ming Jiang (jiangming314220@126.com)
Heilongjiang University of Chinese Medicine
Song Yan
Heilongjiang University of Chinese Medicine
Weichao Ren
Heilongjiang University of Chinese Medicine
Nannan Xing
Heilongjiang University of Chinese Medicine
Hongyuan Li
Heilongjiang University of Chinese Medicine
Meiqi Zhang
Northeast Forestry University
Meiqi Liu
Heilongjiang University of Chinese Medicine
Wei Ma
Heilongjiang University of Chinese Medicine

Research Article

Keywords: Bupleurum chinense, Genetic diversity, Population structure, Cluster analysis, Genotyping-by-sequencing (GBS)

Posted Date: January 12th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1194484/v1

Abstract

Bupleurum (named “Chai-hu”) is an important traditional Chinese medicine resource in China. It has been widely used since ancient times and has antipyretic, analgesic and cholagogic functions, but there is little research on its genetic diversity. In this study, genotyping-by-sequencing (GBS) was used to detect SNP loci in 39 *Bupleurum* germplasm resources from different regions in China and analyse their genetic diversity. A total of 25.1 Gb of data was obtained by sequencing, with an average of 0.64 Gb per sample. After screening, 83898 high-quality SNPs were obtained. The results of genetic research were obtained by phylogenetic tree, principal component analysis and population structure analysis, and the 39 experimental materials were divided into three groups. The average observed heterozygosity and expected heterozygosity of *Bupleurum* populations were 0.24 and 0.17, respectively, indicating that *Bupleurum* populations from five different provinces had a low level of genetic diversity. Population nucleotide diversity analysis and analysis of molecular variance showed that the percentage of intrapopulation variation was 120.88%, while the percentage of interpopulation variation was only 2.46%. There was relative aggregation of *Bupleurum* samples with the same geographical origin, but the division of population structure was not completely correlated with sample origin. The results showed that the genetic diversity of the materials was low and that the genetic variation was narrow. This provides a good basis for the genetic breeding and protection of species diversity of *Bupleurum*.

1 Introduction

Bupleurum chinense DC. is one of the most commonly used medicinal herbs belonging to the genus *Bupleurum* and is named “Chai-hu” (*Radix Bupleuri*). It has been broadly used for approximately 2000 years, with records in the historical Chinese book “The Herbal Classic of Shen Nong” [1]. *B. chinense* possesses diverse TCM value, including in the treatment of influenza, fever, hepatitis, pancreatitis, inflammation, malaria and depression [2, 3]. Extracts from the roots of *B. chinense* contain large amounts of bioactive flavonoids, saikosaponins (A, B, D, F), sesquiterpenes, alkaloid glycosides, saponins and volatile oils, which exhibit anti-inflammatory, antioxidant, antidepressant, antiapoptotic, and antiaddictive pharmacological functions [4–10].

As described in the Pharmacopoeia of the People's Republic of China, *B. chinense* is derived from the roots of *Bupleurum chinense* DC. (called ‘Bei Chaihu’) and *Bupleurum scorzonerifolium* Willd. (called ‘Nan Chaihu’) of the Umbelliferae [11]. Recently, more than 20 other counterfeit plants have been used to treat related diseases. Parts of other species have been used under same name of ‘Chaihu’, including *Dianthus superbus* L., *Bupleurum pyramidalis.*, *Bupleurum longiradiatum* Turez., *Stellaria dichotoma* L. var. *lanceolata* Bge., *Bupleurum smithii* Wolff var. *parvifolium* Shan et Y. Li and other species [12]. However, research shows that different species markedly influence the pharmacological effects and medicinal value of these herbal products. For this reason, the evaluation and utilization of the genetic resources of this important herb is of great significance for future breeding. The genetic background of *Bupleurum chinense* DC. has typically been determined based on inter simple sequence repeat (ISSR) and simple sequence repeat (SSR) molecular markers [13–14]. However, these two kinds of DNA markers have limitations, such as a low reproducibility and limited quantity [15].

Genotyping-by-sequencing (GBS) is a common simplified genome sequencing technology that is suitable for population research, germplasm identification, genetic improvement and trait mapping in a variety of sources. This method can generate high-density and low-cost genotyping data and is appropriate for assessing the genetic variation of organisms with very large genomes. The lack of a need for a reference genome makes it very useful in nonmodel plants [16–18]. For example, the genetic diversity and population structure of 103 popcorn germplasm samples were analysed by GBS, and the genetic diversity, population structure and heterosis patterns were revealed, which provided a basis for cross breeding with existing maize germplasm resources [19]. GBS analysis of 610 pumpkin germplasm resources from global sources revealed 2071 high-quality single nucleotide polymorphisms (SNPs) distributed on 20 chromosomes. The genetic structure of different germplasm resources worldwide was studied based on this valuable information, and new gene banks were established [20]. A total of 180 common wheat populations from Asia and Europe between 30° and 45° latitude were

analysed using GBS, and 24767 high-quality SNPs were used to analyse their genetic diversity and population structure. The results can not only help breeders understand the genetic diversity of wheat germplasm resources but also provide valuable information for wheat genetic improvement through the exploitation of new genetic variations in this region [21].

In this study, GBS was used for the first time to analyse the genetic diversity and population structure of 39 samples of *B. chinensis* collected from different provinces of China, providing a theoretical basis for the efficient utilization and genetic breeding of *B. chinensis* germplasm resources.

2 Materials And Methods

2.1 Plant materials and DNA extraction

Thirty-nine test materials of *B. chinensis* were collected from 14 various habitats of 5 different provinces, and these samples were checked by Professor Dezhi Ma (Table 1). Fresh leaves of these plants were collected for grinding, and DNA samples were extracted using a spin column reagent (Tiangen Biotech, Beijing, China) according to the isolation kit instructions and used for amplification. DNA purity and integrity were analysed by agarose gel electrophoresis, nanodrop detection (OD260/280 ratio) and Qubit accurate quantification methods. High-quality DNA after concentration and purity determination was used for GBS library construction and sequencing.

Table 1
Summary of the tested *Bupleurum* samples

Species	No.	Collection location (Longitude and Latitude)	Species	No.	Collection location (Longitude and Latitude)
B. chinense DC.	HB1	Hebei(117.298;40.908)	*B. chinense* DC.	HRB2	Heilongjiang(126.649;45.924)
	HB2	Hebei(117.275;40.869)		HRB3	Heilongjiang(126.906;45.902)
	HB3	Hebei(117.310;40.856)		DQ1	Heilongjiang(125.102;46.521)
	HB4	Hebei(117.737;41.355)		DQ2	Heilongjiang(125.150;46.443)
	HB5	Hebei(117.774;41.354)		DQ3	Heilongjiang(125.085;46.462)
	HB6	Hebei(117.774;41.297)		ZZ1	Heilongjiang(125.289;45.707)
	BJ2	Beijing(116.789;40.252)		ZZ2	Heilongjiang(125.307;45.699)
	BJ3	Beijing(116.858;40.427)		ZZ3	Heilongjiang(125.243;45.695)
	BJ4	Beijing(116.774;40.438)		DF4	Shaanxi(110.336;33.686)
	BJ5	Beijing(116.556;40.345)		DF5	Shaanxi(110.340;33.654)
	BJ7	Beijing(116.571;40.329)		DF6	Shaanxi(10.245;33.704)
	GS1	Gansu(105.036;33.387)	*B. scorzonerifolium* Willd.	MH1	Heilongjiang(129.699;44.619)
	GS2	Gansu(105.007;33.374)		MH3	Heilongjiang(129.620;44.653)
	GS3	Gansu(104.912;33.386)	*B. sibiricum* Vest.	Q1	Heilongjiang(129.699;44.619)
	GS4	Gansu(104.331;33.765)		Q2	Heilongjiang(129.620;44.655)
	GS6	Gansu(104.353;33.769)		Q3	Heilongjiang(129.571;44.641)
	BQ1	Heilongjiang(126.129;47.611)	*B. longiradiatum* Turcz.	DY1	Heilongjiang(129.692;44.622)
	BQ2	Heilongjiang(126.092;47.587)		DY2	Heilongjiang(129.595;44.631)
	BQ3	Heilongjiang(126.068;47.585)		DY3	Heilongjiang(129.494;44.625)
	HRB1 B2B1	Heilongjiang(126.476;45.936)			

2.2 GBS library construction and data analysis

The GBS library was constructed according to a procedure described earlier with slight revisions [22–24]. Briefly, total genomic DNA was diluted to 20 ng/μl with distilled water for the preparation for GBS analysis, and was digested with *Mse*I (New England Biolabs) at 80 ℃ for 30 min. The digested DNAs were ligated with barcoded adapters using T4 ligase, and all ligated fragments and PCR products were cleaned up using a QIAquick PCR purification kit (Qiagen). The cleaned PCR products were quantified, and the correct-sized fragments were selected with an E-gel system. The library concentration was estimated by a Qubit 2.0 fluorometer, and the library was sequenced on an Illumina HiSeq PE150 (Illumina). Then, clean reads were aligned to the simulated reference genome after analysis of clustering with stacks software. According to the suitable mapped reads from 39 *Bupleurum* samples, SNPs were detected and genotyped using SAMTOOLS software.

2.3 Population genetic structure analysis

The pedigree relationship of *Bupleurum* resources was determined using the unsupervised clustering method with ADMIXTURE software [25]. The admixture model was run five independent times with K values ranging from 2 to 11. The individual samples were designated to each cluster after characterization of the number of K values according to the optimal grouping.

Evolution analysis was checked with TreeBest software (http://treesoft.sourceforge.net/treebest.shtml)using the NJ (neighbour-joining) method with 1,000 bootstrap values and shown with a phylogenetic tree. Principal component analysis (PCA) was performed with GCTA software (http://cnsgenomics.com/software/gcta/pca.html), and several components were analysed by a covariance matrix. We used the first three components with high eigenvalues for further analysis, and the results of PCA were visualized using R.

2.4 Genetic diversity analysis

Arlequin software [26] was used to calculate the observed heterozygosity (*Ho*) and expected heterozygosity (*He*). *Ho* is shown as the ratio of the number of heterozygous individuals to the total number of samples. *He* was determined according to Nei's method [27], which describes the average distance between samples within the same subpopulation. Nucleotide diversity (π) was also calculated to analyse the subpopulation-level polymorphism, and AMOVA revealed the genetic variation among the different subpopulations based on the filtered SNP loci.

3 Results

3.1 GBS analysis and SNP characterization

For the 39 *Bupleurum* accessions, GBS sequencing produced 1.3 billion clean reads, and more than 1.9 million reads per sample were retained for alignment. The number of clean reads ranged from 1,961,510 (HB2) to 6,702,434 (BJ2), with an average of 3,379207. The information for each of 39 accession reads was aligned to the simulated reference genome of BQ2, which has the highest tag number (5,192,250) (Table S1). In the 39 *Bupleurum* samples, an average of 2,657,398 (78.1%) reads were mapped to the reference genome. Among them, BJ3 showed the highest mapping rate (87.97%), and DY2 showed the lowest mapping rate (72.16%) (Table S2). After successful mapping, a total of 85,251 GBS SNPs were recognized, and 83,898 SNPs were obtained after filtering out multiple allele loci using Beagle software.

3.2 Population structure analysis

The K-value was calculated to evaluate the different clusters of these 39 *Bupleurum* samples based on 83,898 genome-wide SNPs with high-quality data. The cross-validation (CV) error was estimated to test the optimal K-value. From 2 to 11, the best K-value showed four groups with the highest probability of clustering according to the genetic background information instead of the sample collection location (Fig. 1). The four groups of 39 accessions accounted for 30.77%, 28.21%, 25.64% and 15.38%, respectively. Cluster □ contained 12 accessions of *B. chinense*, mainly from Hebei Province and Beijing, except for one sample from Heilongjiang Province (HB1, HB2, HB3, HB4, HB5, HB6, BJ2, BJ3, BJ4, BJ5, BJ7, MH1). Cluster □ included 11 accessions of *B. chinense* mainly from Shaanxi, Heilongjiang and Gansu Provinces (HRB1, HRB2, HRB3, DF4, DF5, DF6, GS1, GS2, GS3, GS4, GS6). Cluster □ included all 10 resources from Heilongjiang Province. However, these samples belong to different species. Three accessions were *B. chinense* (BQ1, BQ3, ZZ1), 3 samples were *B. sibiricum* Vest. (Q1, Q2, Q3), 3 accessions were *B. longiradiatum* Turcz. (DY1, DY2, DY3) and one sample was *B. scorzonerifolium* Willd. (MH3). Cluster □ contained the remaining 6 *B. chinense* resources from Heilongjiang Province (ZZ2, ZZ3, DQ1, DQ2, DQ3, BQ2).

Neighbour-joining (NJ) tree analysis was performed based on 83,898 SNPs and classified the 39 *Bupleurum* into three major groups: (1) *B. chinense* samples from Shaanxi Province; (2) *B. chinense* accessions from Gansu, Hebei Province and Beijing; and (3) *B. chinense*, *B. scorzonerifolium*, *B. sibiricum* and *B. longiradiatum* resources from Heilongjiang Province. Group □ was split into 2 subgroups on the basis of the population structure results. The first subgroup mainly corresponded

to samples from Gansu Province, and the second subgroup corresponded to samples from Hebei Province and the city of Beijing (Fig. 2).

The relationship between the 39 *Bupleurum* accessions was explored based on the broad geographical regions using PCA. The samples showed a relatively tight distribution in the PCA results (Fig. 3). The majority of the *B. chinense* accessions from Hebei Province and the city of Beijing were distributed in group Ⅱ, resources of *B. chinense* from Gansu and Shaanxi Provinces were clustered in group Ⅰ. Group Ⅲ contained the other three “Chai-hu” species from Heilongjiang Province, and group Ⅳ clustered accessions of *B. chinense* from Heilongjiang Province.

3.3 Genetic diversity analysis

The tested *Bupleurum* materials were divided into 11 populations according to geographical origin and species (Table 1). The observed heterozygosity (Ho) ranged from 0.21 to 0.27 with an average of 0.24. The expected heterozygosity (He) value ranged from 0.15 to 0.19 with an average of 0.17. The polymorphic loci number (PLN) ranged from 30,724 to 45,778, and the average was 37,358. The percentage of polymorphic loci (PPL) ranged from 38.97 to 58.07, and the average PPL was 47.38. The mean heterozygosity, Ho, was greater than the expected heterozygosity, He, indicating that the genetic diversity of *Bupleurum* was narrow (Table 2).

Table 2
Genetic diversity of *Bupleurum* from different geographical resources and species

Origin	Species	Number	Ho	He	PLN	PPL/%
Hebei	*B. chinense* DC.	6	0.2427	0.1718	41196	52.2533
Beijing		5	0.2082	0.1731	43285	54.903
Gansu		5	0.2419	0.1862	45778	58.0652
Harbin		3	0.2485	0.1774	38116	48.3466
Daqing		3	0.2452	0.1766	38409	48.7183
Zhaozhou		3	0.2157	0.1538	32949	41.7928
Qiqihar		3	0.2701	0.1919	41160	52.2077
Shaanxi		3	0.2262	0.1566	33302	42.2405
Heilongjiang	*B. scorzonerifolium* Willd.	2	0.2359	0.1599	30724	38.9706
Heilongjiang	*B. sibiricum* Vest.	3	0.2299	0.157	32587	41.3336
Heilongjiang	*B. longiradiatum* Turcz.	3	0.2719	0.1711	33439	42.4143
Average			0.24	0.17	37358.63	47.38

The genetic distance (GD) value between 11 populations from different geographical resources and species was -0.3916 to 0.2845, indicating that the genetic distance between populations was small and that the genetic basis was relatively narrow. Among these different species, *B. scorzonerifolium* Willd. showed the maximum genetic distance from *B. chinense* DC. from Heilongjiang Province (0.068), indicating that these resources possess different genetic background information but have the same effect as the Chinese traditional medicine “Chai-hu” (Table 3). The nucleotide diversity analysis of each population found that the mean nucleotide diversity (π) of the total loci was 0.2019, indicating that polymorphisms within populations were relatively rare. The AMOVA results revealed that the genetic variation within populations (23.36%) was greater than that among populations (2.46%), indicating that relatively high gene information exchange occurred within populations and low differentiation occurred among populations (Table 4).

Table 3
Genetic distance matrix of *Bupleurum* samples from different geographical and species resources

	HB	BJ	GS	BQ	HRB	DQ	MH	ZZ	Q	DY	DF
HB	0										
BJ	0.0821	0									
GS	0.2138	0.1201	0								
BQ	0.1817	0.0212	0.1724	0							
HRB	0.2854	0.1422	0.2380	0.1409	0						
DQ	0.1899	0.1932	0.1028	0.2958	0.1436	0					
MH	0.4505	0.2377	0.1085	0.6804	0.3124	0.3964	0				
ZZ	0.1506	0.2650	0.0173	0.2029	0.2369	0.4547	0.4427	0			
Q	0.4181	0.2036	0.0861	0.3977	0.0866	0.3710	0.5262	0.4284	0		
DY	0.3040	0.0640	0.1581	0.2257	0.0289	0.1311	0.3916	0.2587	0.4742	0	
DF	0.2559	0.1151	0.2566	0.2600	0.0259	0.4367	0.3370	0.2132	0.5405	0.4465	0

Table 4
Analysis of molecular variance (AMOVA) for 39 *Bupleurum* accessions

Source of variation	df	Sum of squares	Variance components	Percentage of variation
Among populations	10	71131.032	191.35821 Va	2.46
Among individuals within populations	28	161578.25	-1815.00744 Vb	-23.34
Within individuals	39	366626	9400.66667 Vc	120.88
Total	77	599335.282	7777.01743	

ZZ2	2603018	1741851	66.92%	27.39	28.65%	17.28%
ZZ3	4030106	3343969	82.97%	40.52	36.55%	22.55%

4 Discussion

The traditional herb “Chai-hu” is a medicinal plant-based remedy used worldwide that contains the roots of several species from the genus *Bupleurum* L. according to the WHO in 1999 [28] and shows uneven quality among these different resources in China [29]. It is essential to estimate the diversity and relationships among “Chai-hu” resources to identify excellent breeding materials. Previous research on the population structure and genetic diversity of “Chai-hu” has been carried out based on molecular markers including RAPD, AFLP, ISSR and ITS sequences [30–32]. SNP markers are commonly applied because of their low cost, high throughput and high abundance. Although there is a lack of studies on the thorough collection of Chai-hu accessions in China, the population structure and genetic diversity of 39 *Bupleurum* accessions from 11 different areas are reported with 83,898 SNPs detected by GBS analysis in this work.

Medicinal *Bupleurum* is widely distributed in all regions of China except Hainan Island; clearly, *Bupleurum* has displayed a wide genetic diversity over a long period of domestication to adapt to diverse environments. *Bupleurum* resources can be divided into different groups according to their traditional uses. Using amplified fragment length polymorphisms, 11

Bupleurum strains were classified into 3 different subsets, among which 2 varieties of *Bupleurum chinense* DC. From Henan and Hebei Provinces showed the closest relationship [31]. ISSR markers were used to analyse the genetic distance of 11 *Bupleurum chinense* samples, indicating that the genetic variation has a certain relation with geographic distribution [33]. In recent decades, seeds of *B. chinense* DC. resources grown in several provinces or areas in China have mainly originated from Gansu and Shaanxi Provinces; this phenomenon is strongly supported by genetic diversity results with genetic distances ranging from 0.0173–0.4547, which suggested low genetic diversity within *B. chinense* DC. samples. The nearest accessions were from Heilongjiang and Gansu Provinces, indicating that cultivated samples in Heilongjiang Province were introduced from Gansu Province, which is consistent with the present situation regarding the origin of seeds on the medicinal market in China.

Species grouping of *Bupleurum* strains was also demonstrated with PCA. Accessions of *B. sibiricum* Vest. (Q1, Q2, Q3), 3 accessions of *B. longiradiatum* Turcz. (DY1, DY2, DY3) and 2 samples of *B. scorzonerifolium* Willd. (MH1, MH3) were clearly distinguished from accessions of *B. chinense* DC., emphasizing the trouble with traditional discrimination methods based on phenotypes because of the substantial divergence among these different species. However, the relationship between these different species was also relatively close, ranging from 0.0289–0.6804, suggesting that they may have evolved from a common genetic origin with a lack of diversity at the DNA level. The data collected in this research will be helpful for the establishment of quality criteria for “Chai-hu” and for the formulation of reasonable clinical use strategies. At the same time, the utilization of SNP markers could not only be applied in resource discrimination but also offer genetic information for screening and molecular breeding of medicinal plant varieties.

Declarations

Acknowledgements This work was supported by the Key project at central government level: The ability establishment of sustainable use for valuable Chinese medicine resources (2060302) and the State Key Laboratory of Pharmaceutical New-tech for Chinese Medicine (SKL2020M0302).

Author contributions M J wrote the manuscript, and checked and analyzed data. S Y, W-C R, and N-N X collected data and revised the manuscript. H-Y L, M-Q Z and M-Q L helped revised the manuscript. W M critically designed the experiment.

Conflict of interest The authors declare that they have no conflicts of interest.

Ethical approval This aitical does not contain any experiment with human participants or animals.

References

1. Feng YJ, Wu ZW, Luo YY et al (2019) A new triterpene diglycoside from the roots of *Bupleurum chinense* DC. and its inhibitory effect on adipogensis in 3T3-L1 cells[J]. Med Chem Res 28:239–245
2. Pan SL (2006) Bupleurum species: scientific evaluation and clinical applications. CRC Press, Boca Raton
3. Law YK, Mo JF, Wong KW (2014) Autophagic effects of Chaihu (dried roots of *Bupleurum Chinense* DC or Bupleurum scorzoneraefolium WILD) [J]. Chin Med 9:21
4. Yu J, Deng A, Wu L et al (2013) Osteoclast-inhibiting saikosaponin derivatives from *Bupleurum Chinense*[J]. Fitoterapia 85:101–108
5. Li DQ, Wu J, Liu LY et al (2015) Cytotoxic triterpenoid glycosides (saikosaponins) from the roots of *Bupleurum chinense*[J], vol 25. Bioorganic & Medicinal Chemistry Letters, pp 3887–3892
6. Li HY, Zhao YH, Zeng MJ et al (2017) Saikosaponin D relieves unpredictable chronic mild stress induced depressive-like behavior in rats: involvement of HPA axis and hippocampal neurogenesis[J]. Psychopharmacology 234:3385–3394
7. Wang Y, Qiang G, Cheng Z et al (2017) New saikosaponins from the roots of *Bupleurum chinense*[J]. Phytochem Lett 21:183–189

8. Wang HW, Liu M, Zhong TD et al (2015) Saikosaponin-d attenuates ventilator-induced lung injury in rats[J]. Int J Clin Exp Med 8:15137–15145
9. Chen XQ, Chen SJ, Liang WN et al (2018) Saikosaponin A attenuates perimenopausal depression-like symptoms by chronic unpredictable mild stress[J]. Neurosci Lett 662:283–289
10. Lorrai I, Maccioni P, Carai M et al (2017) Suppressing effect of saikosaponin A, an active ingredient of Bupleurum falcatum, on chocolate self-administration and reinstatement of chocolate seeking in rats[J]. Neurosci Lett 638:211–217
11. National Pharmacopoeia Committee. Pharmacopoeia of Peoples Republic of China [M]. Part 1. Beijing: Chemical Industry Press (2020) : Appendix 2: 293
12. Sui C, He WJ, Lin CS et al (2009) Development of genomic SSR and potential EST-SSR markers in *Bupleurum chinense* DC.[J]. Afr J Biotechnol 8:6233–6240
13. Zhao X, Liu C, Xue W et al (2015) ISSR Research on Germplasm of *Bupleurum chinense* DC. in Beijing[J]. Modern Chinese Medicine 10:1008–1013
14. Yang W, Bai Y, Ji-Ying HU (2013) ISSR Research on Germplasm of *Bupleurum chinense* DC. in Baokang[J]. Research and Practice on Chinese Medicines, 2:25-27
15. Lee KJ, Lee JR, Sebasti NR et al (2019) Genetic Diversity Assessed by Genotyping by Sequencing (GBS) in Watermelon Germplasm[J]. Genes 10:822–834
16. Fu YB, Peterson GW (2011) Genetic Diversity Analysis with 454 Pyrosequencing and Genomic Reduction Confirmed the Eastern and Western Division in the Cultivated Barley Gene Pool[J]. Plant Genome 4:226–237
17. Peterson BK, Weber JN, Kay EH et al (2012) Double Digest RADseq: An Inexpensive Method for De Novo SNP Discovery and Genotyping in Model and Non-Model Species[J]. PLoS ONE 7:e37135
18. Peterson G, Dong Y, Horbach C et al (2014) Genotyping-By-Sequencing for Plant Genetic Diversity Analysis: A Lab Guide for SNP Genotyping[J]. Diversity 6:665–680
19. Yu D, Wang H, Gu W et al (2021) Genetic diversity and population structure of popcorn germplasm resources using genome-wide SNPs through genotyping-by-sequencing[J]. Genetic Resources and Crop Evolution, pp 1–11
20. Lee HY, Jang S, Yu CR et al (2020) Population Structure and Genetic Diversity of Cucurbita moschata Based on Genome-Wide High-Quality SNPs[J]. Plants 10:56–65
21. Yang X, Tan B, Liu H et al (2020) Genetic Diversity and Population Structure of Asian and European Common Wheat Accessions Based on Genotyping-By-Sequencing[J].Frontiers in Genetics,1–11
22. Li H, Durbin R Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics,2009a, 25:1754-1760
23. Li H, Handsaker B, Wysoker A et al The Sequence Alignment/Map format and SAMtools. Bioinformatics,2009b, 25:2078-2079
24. Wang K, Li M, Hakonarson H (2010) ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res 38:e164
25. Alexander DH, Novembre J, Lange K (2009) Fast model-based estimation of ancestry in nrelated individuals. Genome, Res 19:1655–1664
26. Excoffier L, Laval G, Schneider S (2005) Arlequin (version 3.0): an integrated software package for population genetics data analysis[J]. Evol Bioinform Online 1:47–50
27. Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees.[J]. Molecular Biology & Evolution, (4):406
28. World Health Organization (1999) *WHO Monographs on Selected Medicinal Plants*[M], vol 1. World Health Organization, Geneva

29. Huang HQ, Zhang X, Xu ZX et al (2009) Fast determination of saikosaponins in *Bupleurum* by rapid resolution liquid chromatography with evaporative light scattering detection[J]. J Pharm Biomed Anal 49(4):1048–1055
30. Huang W, Sun P, Zhang WS et al (2008) Genetic Diversity of *Bupleurum chinense* DC. populations from different altitudes in Dongling mountain district in Beijing[J]. Plant Genetic Resources 9(4):453–457
31. Ke SY, Shi LL, Ma YZ et al (2015) Evaluation of the genetic diversity of *Bupleurum* using amplified fragment length polymorphism analysis[J], vol 14. GENET MOL RES, pp 2590–25991
32. Du SM, Wang G, Liu YM et al (2013) Study on Biological Materials with Genetic Diversity of *Bupleurum marginatum* in Northwest of Hubei Province of China Base on ISSR[J]. Adv Mater Res 830:463–468
33. Li YH, Yu XL, Ou XJ et al (2018) Genetic Diversity of Different Origin Bupleurum chinese Detected by ISSR Analysis[J]. Lishizhen Medicine and Materia Medica Research 29:1728–1731

Figures

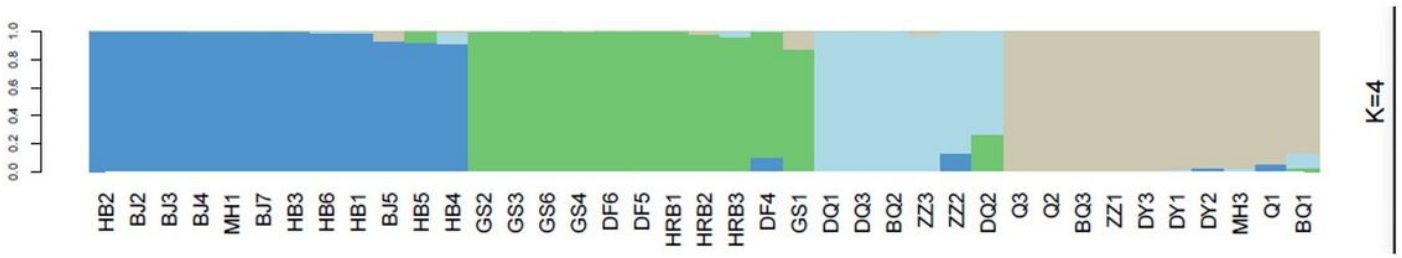

Figure 1

Population structure of 39 genotypes with K=4. Each genotype is represented by a single vertical line

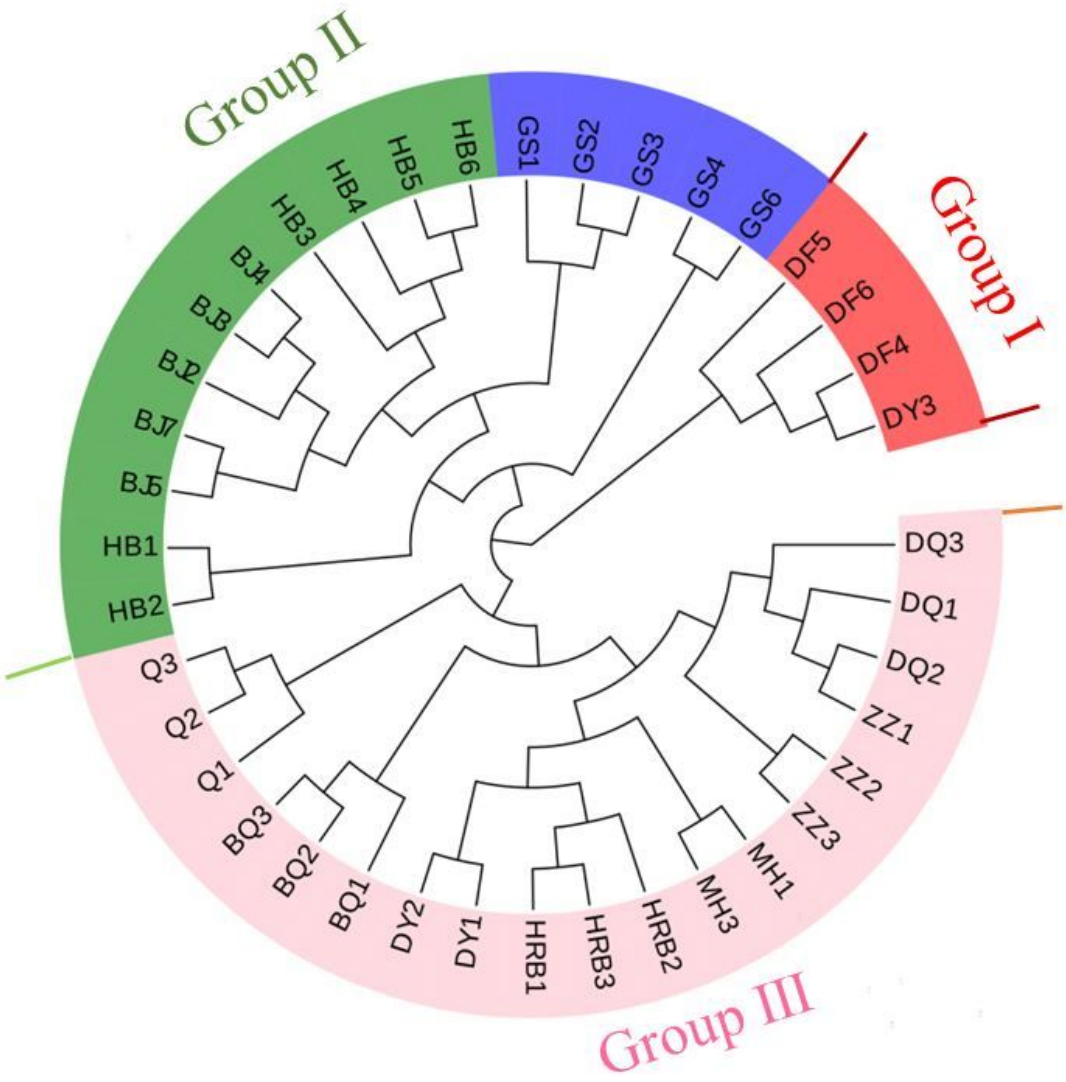

Figure 2

Neighbour-joining (NJ) tree analysis of 39 *Bupleurum* accessions based on population structure

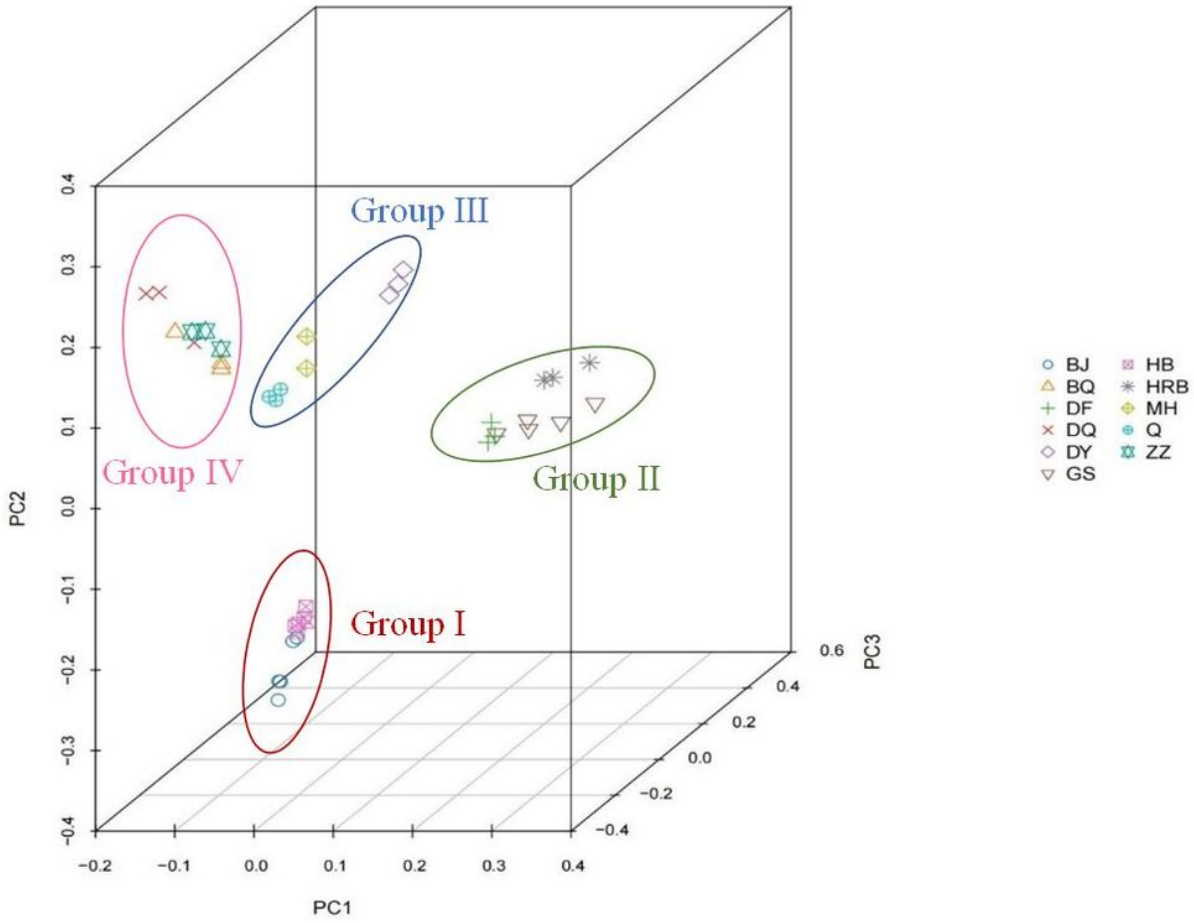

Figure 3

PCA (principal component analysis) based on SNP information with genetic distances among 39 *Bupleurum* accessions

Supplementary Files

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

- JiangmingMBRsupplementarytable.docx