Journal of Genetic Engineering and Biotechnology 22 (2024) 100379

Contents lists available at ScienceDirect

Journal of Genetic Engineering and Biotechnology

journal homepage: www.elsevier.com/locate/jgeb

Microsatellite marker-based analysis of the genetic diversity and population structure of three *Arnebiae Radix* in western China

Jinrong Zhao [a], Yanjiao Wang [b], Wenhuan Ding [c], Haiyan Xu [a,*]

[a] College of Traditional Chinese Medicine, Xinjiang Medical University, Xinjiang, China
[b] Department of Basic Medical Sciences, Xinjiang Medical University, Xinjiang, China
[c] Central Laboratory, Xinjiang Medical University, Xinjiang, China

ARTICLE INFO

Keywords:
Arnebiae Radix
Genetic diversity
Population genetic structure
ISSR molecular markers

ABSTRACT

Arnebiae Radix is an important medicinal and perennial herb found in Western China, particularly in the Xinjiang region. However, the assessment, utilization and conservation of *Arnebiae Radix* resources are still unexplored. In this study, we evaluated the genetic diversity of three *Arnebiae Radix* populations across 47 regions (Ae = 16, Ag = 16, Ad = 15) in Xinjiang, China, using inter-simple sequence repeat (ISSR) molecular markers. In total, 48 alleles were amplified by six pairs of primers screened with ISSR markers. The average number of effective alleles (Ne) was 1.5770. The percentage of interspecific genetic polymorphisms in *A. guttata* (Ag = 89.58 %) was greater than that in *A. euchroma*. and *A. decumbens* (Ae = Ad = 87.50 %). Intraspecific genetic polymorphisms, Bo Le (BL) population of *A. euchroma* exhibited the highest percentage of polymorphic bands (PPB% = 58.33 %, Na = 1.313, Ne = 1.467, I = 0.0.366, H = 0.255), which indicated high genetic diversity. In contrast, the Tuo Li (TL) population of *A. guttata* had the lowest values for these parameters (PPB% = 0.00 %, Na = 0.313, Ne = 1,000, I = 0.000, H = 0.000). The *Arnebiae Radix* germplasms were classified into two major groups (I and II) based on UPGMA cluster analysis (Fig. 8**a**) and principal coordinate analysis (PCOA). In addition, *A. decumbens* is placed in a separate category due to its high differentiation coefficient. The AMOVA and genetic differentiation coefficient results indicated that the genetic variation in *Arnebiae Radix* was predominantly due to intrapopulation differences (78 %). Additionally, the gene flow index (Nm) between populations was 2.4128, which further indicated that the genetic diversity of *Arnebiae Radix* was greater at the intrapopulation level. The destruction of the ecological environment leads to the continuous reduction and degradation of the genetic diversity of *Arnebiae Radix* germplasm resources. In this study, we used ISSR molecular markers to analyze the genetic diversity and relatedness of *Arnebiae Radix*, which revealed the genetic relationship of *Arnebiae Radix* germplasm resources at the molecular level and provided a scientific basis for future research on selecting and breeding good varieties, evaluating the quality of *Arnebiae Radix*, and conserving and utilizing its resources.

1. Introduction

Arnebiae Radix is a traditional Chinese medicine. It is the dried root bark of *A. euchroma* (Royle) Johnst. and *A. guttata* Bunge, according to the Chinese Pharmacopoeia (2020 version). The Xinjiang region is a well-known producer of *Arnebiae Radix*.[1–3] The Boraginaceae family is divided into three genera: *Arnebiae* Forsk., *Li thospermum* L., and *Oncsma* L. The roots of *Arnebiae* Forsk. are primarily valuable for medicinal purposes. *Arnebiae* Forsk. includes the following species: *A. euchroma* (Royle) Johnst., *A. guttata* Bunge., *A. tschimganica* (Fedtsch.) G. L. Chu, and *A. decumbens* (Vent.) Coss. and Kral[4,5] (Fig. 1). The active constituents of *Arnebiae Radix* are mainly fat soluble and water soluble, with the fat soluble constituents being mainly naphthoquinones and the water soluble constituents being mainly *Arnebiae Radix* polysaccharides. *Arnebiae Radix* has anti-inflammatory, antibacterial, antiviral, antitumor, antiallergic, hepatoprotective, enzyme-lowering, and hemostatic effects; promotes the healing of burns and wounds; and has other pharmacological effects. *Arnebiae Radix* is distributed mainly on the northern and southern slopes of Tianshan Mountain in Xinjiang, China, and west of Tibet. *Arnebiae Radix* is also distributed in Central Asia, such as Nepal, on the sunny slopes of mountains and grasslands at an altitude of

* Corresponding author at: College of Traditional Chinese Medicine, Xinjiang Medical University, Xinjiang, China.
E-mail address: xhy@xjmu.edu.cn (H. Xu).

https://doi.org/10.1016/j.jgeb.2024.100379
Received 23 November 2023; Received in revised form 18 March 2024; Accepted 18 April 2024

Fig. 1. (a). *Arnebia euchroma* (Royle) Johnst.; (b). *A. decumbens* (Vent.) Coss. And Kral; (c). *A. guttata* Bunge.

2000 – 4200 m. There are two ways to reproduce *Arnebiae Radix*, namely, sexual reproduction and asexual reproduction. Since the seed germination of *Arnebiae Radix* is strongly affected by internal and external factors, when relying on seeds for sexual reproduction, the germination rate is low when the external conditions are unsuitable and even lower when the external conditions are uncontrolled and unsuitable, which is one of the main reasons why wild populations of *Arnebiae Radix* cannot be recovered rapidly within a limited period of time. Therefore, the germplasm resources of *Arnebiae Radix* were further explored by reflecting the affinities among different populations through ISSR markers. We hope to promote the cultivation of *Arnebiae Radix* as soon as possible to alleviate its scarcity.

Genetic diversity indicates the extent of genetic variation present among individuals within a breed or population resulting from DNA recombination, mutation, or genetic drift during inheritance.[6] The diversity within plant populations occurs due to the variation in the genetic makeup of reproducing individuals. The genetic diversity of medicinal plants depends on the heritable variation present between and within populations.[7–8] It arises from genetic variation in DNA sequences, chromosomal mutations, and sexual reproduction, leading to alterations in DNA sequences, genetic maps, protein structures, and morphological traits.[9] Genetic diversity in plants generally occurs at three levels: diversity among species, diversity among populations within a species, and diversity among individuals within a population.[10,11] Genetic diversity is commonly referred to as the genetic variation among different populations within a species or among different individuals within a population. The interspecies diversity, interpopulation diversity within species, and interindividual diversity within populations represented interspecies variation, interpopulation variation, and interindividual variation within populations, respectively.

2. Materials and methods

2.1. Plant material

In this study, we used three *Arnebiae Radix* resources from *Arnebia* Forsk., namely, *A. euchroma* (Royle) Johnst., *A. decumbens* (Vent.) Coss. and Kral, and *A. guttata* Bunge, as our materials. The plants were dispersed across 47 districts in the Xinjiang Uyghur Autonomous Region and consisted of 159 samples (Fig. 2). Among them, 62 samples were of *A. euchroma* (Royle) Johnst. (Ae), 43 samples were from *A. guttata* Bunge (Ag), and 54 samples were from *A. decumbens* (Vent.) Coss. and Kral (Ad). *A. euchroma* is mainly distributed in the northern and southern Tianshan Mountains in Xinjiang, China, on the sunny slopes of mountains at an altitude of 2000–4200 m. In the southern region, it is mainly distributed in Bazhou, Kezhou, Aksu and Kashi. It is afraid of high temperature and cold tolerance and has high temperature and humidity requirements during the growing season. It is mainly distributed in Bazhou, Kezhou, Aksu and Kashgar in southern Xinjiang. In the north, *A. euchroma* was mainly distributed in Yili and Urumqi. *A. guttata* occurs in Xinjiang in the Front Range deserts, Gobi, etc. It grows in areas with a dry climate and annual precipitation not exceeding 200 mm. The population has a strong ability to regenerate and has a large distribution area but low abundance; only the outer skin of the roots shows a purple-red color without obvious embolization, and the yield of a single plant is not high, and the plant is not easy to dig up. The fungi were distributed mainly in FY, HJ, HM, and ATS. *A. decumbens* occurs in northern Xinjiang on mountain slopes, sandy places and wastelands at 200–1200 m above sea level. The main distribution areas are FK, JMSR, JH, and WS. While sampling, we collected at least one fresh leaf from each site and ensured that the linear distance between samples of the same germplasm was ≥1 m (Table 1).

We extracted DNA samples using the TSINGKE DNA Plant Extraction Kit. Briefly, 20 mg of each of the silica gel-dried samples presented in Table 1 was extracted according to the manufacturer's instructions. The extracted genomic DNA was stored at –20 ˚C until further amplification.

2.2. PCR amplification

Eight highly variable DNAs (X2-3, X7-10, X8-3, H6-2, H7-1, H12-2, Y6-3, and Y7-16) were selected after DNA extraction. The eight highly variable genomic DNAs were amplified with 100 primer pairs based on the 100 universal ISSR primer sequences published by Columbia University.[12] The amplification program was as follows: pre-denaturation at 98 ˚C for 2 min; denaturation at 98 ˚C for 10 s; re-denaturation at 7 ˚C for 10 s; denaturation at 72 ˚C for 10 s for 35 cycles; extension at 72 ˚C for 5 min; and storage at 4 ˚C. Finally, six pairs of polymorphic primers were screened (UBC814, UBC821, UBC830, UBC860, UBC888, and UBC889) (Table 2).

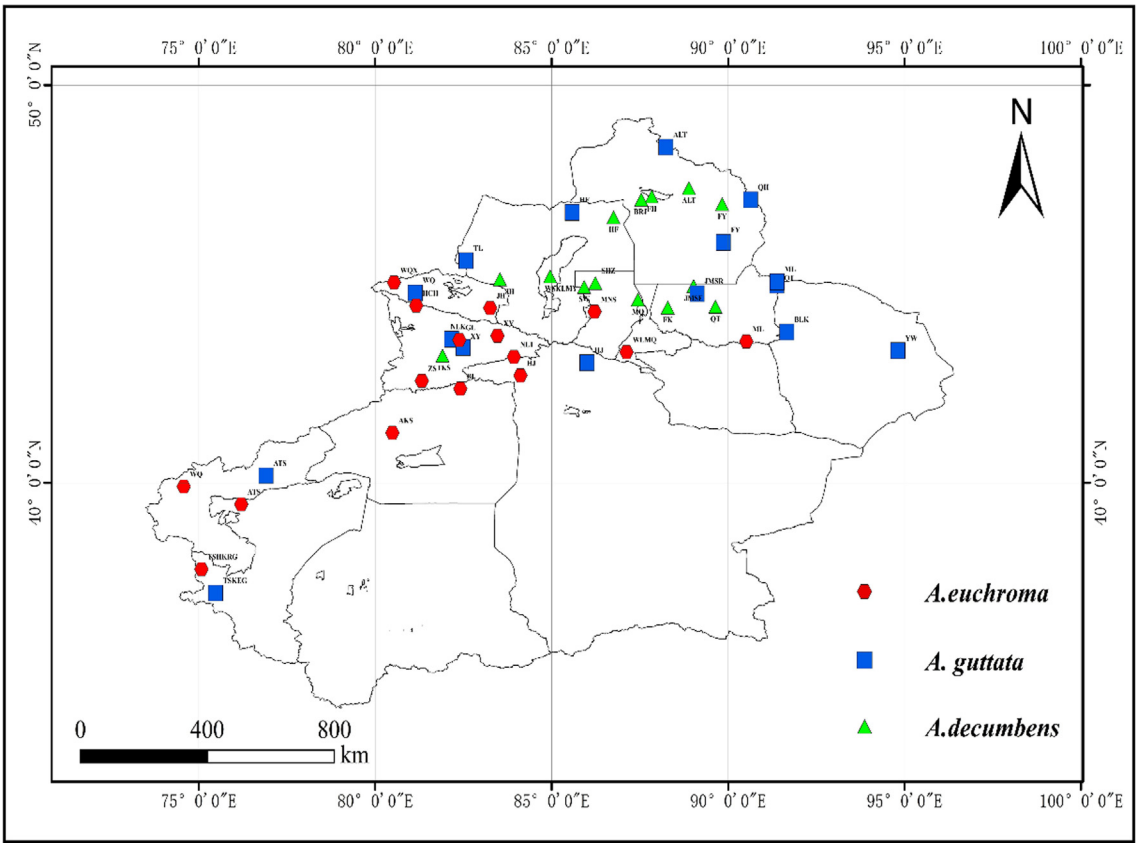


Fig. 2. The distributions of the three species of *Arnebiae Radix* in the study area are shown on a map.

2.3. Statistical analysis

The ISSR is a dominant marker, and each band obtained through electrophoresis is considered to be a molecular marker representing a distinct binding site of the primer. The molecular markers were converted into a binary matrix, with labeled bands represented as "1″ and no bands as "0".[13]

2.3.1. Genetic diversity

The percentage of polymorphic marker sites was calculated using the formula P (%) = (K/N) × 100, where K indicates the number of polymorphic sites and N indicates the total number of sites. The number of polymorphic sites was assessed after 2–3 PCR amplifications of the same ISSR primers, and after the electrophoresis, the gel plates showed clear and stable bands. The percentage of polymorphic sites (PPB), Shannon's information index (I), Nei's gene diversity, number of observed alleles (Na), and number of effective alleles (Ne) were calculated using GenAlEx V6.5 and POPGENE1.32.[14]

2.3.2. Structure analysis

The genetic structure of the population was analyzed using the population clustering method based on Bayesian modeling in STRUCTURE (version 2.3.3). Using the Markov chain Monte Carlo (MCMC) method, predetermined population groupings (K) were used to calculate, sample, and group individuals based on allele frequencies.[15,16] The optimal interval for the K value was determined following the guidelines of Evanno.[17] The K values ranged from 2 to 10, with 20 independent runs conducted for each value, and 100,000 replicate samples were obtained for each cycle. The appropriate population clustering number was determined by calculating the delta K value using STRUCTURE HARVESTER.[18] The samples were grouped into three *Arnebiae Radix* Latin abbreviations (Ae, Ag, Ad). The nomenclature used was Latin abbreviation + code (Table 1) + sample collection number. (e.g., Ae1-8 are collections of *A. euchroma* with MNS number 8).

2.3.3. Genetic differentiation

Genetic differentiation coefficients within and between populations were evaluated via analysis of molecular variance (AMOVA).[19] We also calculated Nei's diversity analysis indices for Hs, Ht, and Gst, along with the Nm (Nm = (1-Fst)/4Fst), to assess interpopulational diversity.[20]

2.3.4. Cluster analysis

Genetic distances were calculated using GenAlEx version 6.501.[21,22] Principal component analysis was conducted based on these calculations. Using the NTSYS calculations, a similarity matrix or distance matrix file was generated. We used an unweighted group averaging method (UPGMA) to build a clustering tree of individuals based on the similarity or distance matrix. This clustering tree showed kinship relationships between individuals. The sample grouping procedure was performed with STRUCTURE.

3. Results

3.1. Genetic diversity

In total, 48 alleles were amplified using six sets of primers. The number of amplified bands ranged from 5 to 10. UBC814 and UBC899 generated most of the bands (n = 10), while UBC830 generated the fewest bands (n = 5). Eight bands were amplified on average. The percentage of polymorphic sites (PPB) in *Arnebiae Radix* was found to reach 100 %. The total number of Ne was 10.9225, with a maximum value of 1.6403 (UBC814) and a minimum value of 1.3813 (UBC860). The average number of effective alleles per locus

Table 1
Source of the three species of *Arnebiae Radix* used in this study.

Pop	District	Code	Longitude (E)	Latitude (N)	Altitude (m)	Sample
***A. euchroma*(Ae)**	MNS	Ae-1	86°13′11.03″	44°18′13.51″	2010	4
	ML	Ae-2	90°31′09.79″	43°33′21.42″	2568	3
	BL	Ae-3	81°84′55.37″	41°82′11.06″	2606	4
	NLT	Ae-4	83°56′10.92″	43°10′11.24″	2500	4
	TSKEG	Ae-5	75°04′52.20″	37°49′58.71″	4234	4
	WQ	Ae-6	74°34′44.80″	39°54′36.53″	3323	4
	WQX	Ae-7	80°32′18.03″	45°02′18.16″	2299	4
	HJ	Ae-8	84°07′13.31″	42°42′00.50″	2456	3
	HCH	Ae-9	81°09′53.83″	44°27′32.39″	2502	4
	ZS	Ae-10	81°19′39.02″	42°33′56.4″	2194	4
	GL	Ae-11	82°23′19.10″	43°35′42.30″	2530	4
	XY	Ae-13	83°27′66.19″	43°41′55.40″	2201	3
	WLMQ	Ae-14	87°07′28.67″	43°17′19.09″	2507	4
	JH	Ae-15	83°15′44.71″	44°23′42.60″	2144	4
	AKS	Ae-16	80°29′31.22″	41°15′33.28″	2030	4
	ATS	Ae-17	76°12′33.08″	39°27′43.50″	2300	4
***A. guttata*(Ag)**	ATS	Ag-1	76°54′38.11″	40°10′42.03″	1313	3
	FY	Ag-2	98°52′02.03″	46°02′46.01″	1065	2
	BLK	Ag-3	91°39′48.60″	43°47′25.60″	1632	2
	NLK	Ag-4	82°10′03.02″	43°36′41.03″	765	2
	TSKEG	Ag-5	75°28′58.30″	37°13′44.70″	3780	3
	QT	Ag-6	91°22′52.90″	44°58′00.51″	1167	3
	HJ	Ag-7	86°00′23.56″	43°01′01.89″	2253	2
	HF	Ag-8	85°35′11.01″	46°47′59.07″	1292	4
	TL	Ag-10	82°34′40.01″	45°35′00.01″	1532	3
	ML	Ag-11	91°23′22.51″	45°03′20.05″	1312	3
	QH	Ag-12	90°38′25.03″	46°67′07.03″	1193	3
	WQ	Ag-13	81°8′18.04″	44°46′32.07″	1801	3
	XY	Ag-14	82°29′45.02″	43°24′17.02″	894	2
	YW	Ag-17	94°48′51.30″	43°19′25.31″	1414	2
	JMSR	Ag-19	89°07′19.46″	44°45′17.46″	1067	4
	ALT	Ag-20	88°14′19.46″	47°86′17.46″	1067	2
***A. decumbens* (Ad)**	JMSE	Ad-1	89°01′23.07″	44°57′07.02″	778	3
	MQ	Ad-3	87°26′47.11″	44°36′42.03″	707	4
	JH	Ad-4	82°92′11.33″	44°67′13.12″	444	4
	WS	Ad-5	84°57′40.12″	45°11′48.03″	379	3
	SHZ	Ad-6	86°14′28.06″	45°01′42.05″	472	3
	BRJ	Ad-7	86°92′34.51″	47°07′33.08″	497	2
	SW	Ad-8	85°55′17.15″	44°55′52.11″	564	4
	KLMY	Ad-9	84°57′40.13″	45°11′48.01″	436	4
	FK	Ad-10	88°17′34.11″	44°24′39.10″	487	4
	TKS	Ad-11	81°54′41.78″	43°12′10.64″	1375	4
	FY	Ad-12	89°50′02.03″	47°00′46.01″	565	4
	HF	Ad-13	86°45′20.03″	46°40′50.01	1285	4
	FH	Ad-14	87°50′24.01″	47°12′07.01″	543	4
	QT	Ad-15	89°38′36.05″	44°25′60.04″	530	4
	ALT	Ad-17	88°53′36.03″	47°24′53.08″	650	3

Table 2
ISSR primers and their optimal annealing temperatures.

Number	Primer	Annealing temperature (°C)
814	CTCTCTCTCTCTCTCTA	51
821	GTGTGTGTGTGTGTGTT	51
830	TGTGTGTGTGTGTGTGG	53
860	TGTGTGTGTGTGTGTGRA	57
888	BDBCACACACACACACA	50
889	DBDACACACACACACAC	50

was 1.5770. The maximum value of Shannon's index (I) was 0.5486 (UBC888) and the minimum value was 0.3433 (UBC860). The average Shannon index value was 0.5060. The mean Nei's gene diversity was 0.3383 **(**Table 3**).** Agarose gel electrophoresis images of the PCR products are shown **(**Fig. 3**).**

We evaluated the interspecific genetic diversity of three populations of *A. euchroma* (Ae), *A. guttata* (Ag), and *A. decumbens* (Ad). The observed allele count (Na) and the effective allele count (Ne) of the three populations ranged between 1.750 and 1.792 (mean

Table 3
Polymorphism of ISSR Primer Amplification Products.

Primer	NSB	NPB	PPB	Na	Ne	I	H	PIC
UBC814	10	10	100 %	2	1.6403	0.5389	0.3652	0.8245
UBC821	8	8	100 %	2	1.6031	0.5203	0.3507	0.6735
UBC830	5	5	100 %	2	1.4853	0.4704	0.3050	0.4603
UBC860	6	6	100 %	2	1.3813	0.3433	0.2220	0.5247
UBC888	9	9	100 %	2	1.5992	0.5486	0.3653	0.7538
UBC889	10	10	100 %	2	1.6363	0.5389	0.3638	0.8061
Mean	8	8	100 %	2	1.5770	0.5060	0.3383	0.6342

NSB = number of scorable bands, NPB = number of polymorphic bands, PPB = percent polymorphism bands, Na = no. of different alleles; Ne = no. of effective alleles; I = Shannon's information index; H = Nei's gene diversity; PIC = polymorphic information content.

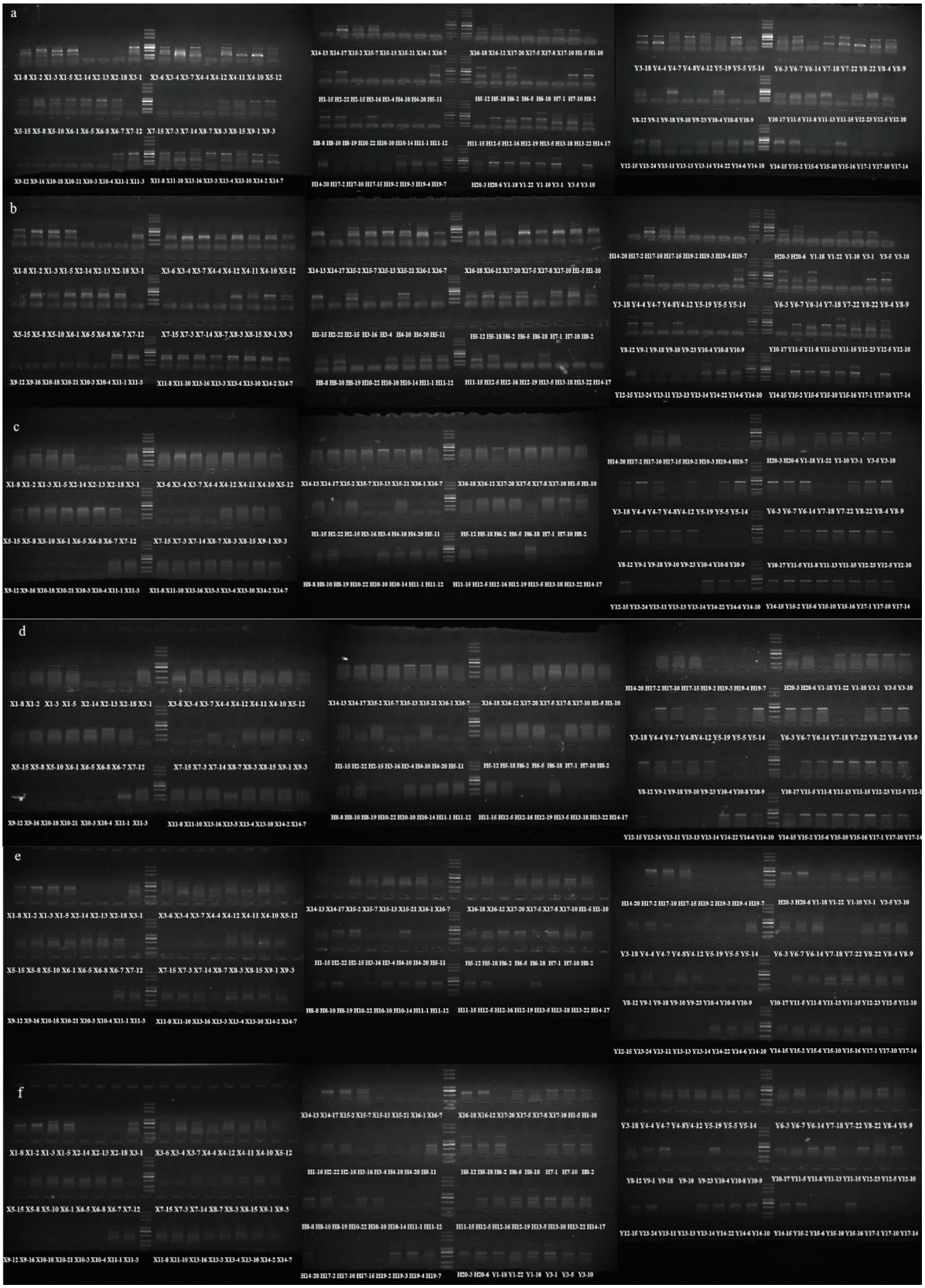


Fig. 3. ISSR PCR profile results of three *Arnebiae Radix* strains produced with six primers (a. UBC814, b. UBC821, c. UBC830, d. UBC860, e. UBC888, f. UBC889).

1.771) and between 1.321 and 1.443 (mean 1.380), respectively. Nei's gene diversity (H) ranged from 0.202 to 0.264 (mean 0.234), whereas the Shannon diversity index (I) ranged from 0.322 to 0.402 (mean 0.366). The percentage of polymorphic loci (PPB) ranged from 87.50 % to 89.58 % (Table 4).

We also examined the genetic variation within individuals of the species using 159 samples from this collection. Among them, 13 populations of *A. euchroma* (Royle) Johnst. (Ae), 10 populations of *A. guttata* Bunge (Ag), and 12 populations of *A. decumbens* (Vent.) Coss. and Kral (Ad) 120 samples were evaluated. The mean diversity values of the three *Arnebiae Radix* strains were Na = 0.739, Ne = 1.216, I = 0.173, and H = 0.120 (Table 5).

For *A. euchroma* (Royle) Johnst. (Ae), the observed allele count (Na) and effective allele count (Ne) ranged from 0.583 to 1.313 and from 1.142 to 1.467, respectively. Nei's genetic diversity index (H) and Shannon diversity index (I) ranged from 0.083 to 0.255 and from 0.116 to 0.366, respectively. The highest genetic diversity was found in the Bo Le (BL) population, for which the diversity parameters were Na = 1.313, Ne = 1.467, I = 0.366, and H = 0.255. The number of observed alleles (Na) and the number of effective alleles (Ne) in *A. guttata* Bunge (Ag) ranged from 0.271 to 0.979 and from 1.000 to 1.313, respectively. Nei's genetic diversity index (H) and Shannon's diversity index (I) ranged from 0.000 to 0.157 and from 0.000 to 0.225, respectively. The number of observed alleles (Na) and effective alleles (Ne) of *A. decumbens* (Vent.) Coss. and Kral (Ad) ranged from 0.375 to 1.125 and from 1.083 to 1.521, respectively, and Nei's genetic diversity index (H) and Shannon's diversity index (I) ranged from 0.042 to 0.260 and from 0.058 to 0.361, respectively. The highest genetic diversity was found in the Wu Su (WS) population, for which the diversity parameters were Na = 1.125, Ne = 1.521, I = 0.361, and H = 0.260.

The effective allele count (Ne) in the *Arnebiae Radix* population might indicate the number of alleles influencing the observed alleles and thus can indicate allelic richness.[23] The populations with a higher

Table 4
Interspecific genetic diversity of *Arnebiae Radix*.

Pop	Na	Ne	I	H	PPB (%)
Ae	1.750	1.374	0.374	0.238	87.50 %
Ag	1.792	1.321	0.322	0.202	89.58 %
Ad	1.771	1.443	0.402	0.264	87.50 %
Mean	1.771	1.380	0.366	0.234	88.19 %
SE	0.053	0.028	0.020	0.014	0.69 %

Na = number of different alleles; Ne = no. of effective alleles; I = Shannon's information index; H = Nei's gene diversity; PPB = percent polymorphism bands.

Table 5
Genetic diversity of individuals within *Arnebiae Radix* species based on ISSR.

Pop	District	Na	Ne	I	H	PPB (%)
***A. euchroma* (Ae)**	MNS	0.750	1.225	0.180	0.125	29.17 %
	BL	1.313	1.467	0.366	0.255	58.33 %
	NLT	0.979	1.296	0.257	0.174	43.75 %
	TSHKEG	0.917	1.308	0.253	0.174	41.67 %
	WQ	0.667	1.142	0.123	0.083	20.83 %
	HJ	0.625	1.167	0.116	0.083	16.67 %
	HCH	0.583	1.233	0.186	0.130	29.17 %
	GL	0.813	1.233	0.198	0.135	33.33 %
	XY	0.875	1.275	0.227	0.156	37.50 %
	WLMQ	0.833	1.292	0.233	0.161	37.50 %
	JH	0.813	1.196	0.163	0.112	27.08 %
	AKS	0.896	1.233	0.214	0.143	37.50 %
	ATS	0.813	1.233	0.198	0.135	33.33 %
***A. guttata* (Ag)**	ATS	0.979	1.283	0.225	0.157	35.42 %
	FY	0.479	1.188	0.130	0.094	18.75 %
	TSKEG	0.271	1.042	0.029	0.021	4.17 %
	QT	0.604	1.233	0.186	0.130	29.17 %
	HJ	0.771	1.313	0.217	0.156	31.25 %
	HF	0.646	1.183	0.146	0.102	22.92 %
	TL	0.313	1.000	0.000	0.000	0.00 %
	ML	0.375	1.067	0.053	0.037	8.33 %
	YW	0.646	1.183	0.146	0.102	22.92 %
	ALT	0.500	1.125	0.087	0.063	12.50 %
***A. decumbens* (Ad)**	JMSE	0.375	1.083	0.058	0.042	8.33 %
	MQ	0.667	1.154	0.134	0.091	22.92 %
	JH	0.542	1.183	0.146	0.102	22.92 %
	WS	1.125	1.521	0.361	0.260	52.08 %
	SHZ	0.688	1.083	0.066	0.046	10.42 %
	SW	0.875	1.204	0.181	0.122	31.25 %
	KLMY	0.875	1.333	0.262	0.182	41.67 %
	TKS	1.083	1.375	0.306	0.211	50.00 %
	FY	0.729	1.163	0.152	0.102	27.08 %
	FH	0.896	1.242	0.201	0.138	33.33 %
	QT	0.667	1.100	0.094	0.063	16.67 %
	ALT	0.875	1.217	0.172	0.120	27.08 %
	Mean	0.739	1.216	0.173	0.120	27.86 %
	SE	0.021	0.009	0.007	0.005	2.29 %

Na = number of different alleles; Ne = no. of effective alleles; I = Shannon's information index; H = Nei's gene diversity.

Ne were ranked as follows: Ad Wu Su (WS) > Ae Bo Le (BL) > Ag He Jing (HJ). Nei's index of genetic diversity (H) and Shannon's diversity index (I) indicate the level of genetic diversity among populations. The populations with high Shannon's indices were Ae Bo Le (BL) > Ad Wu Su (WS) > Ag A Tu Shi (ATS), and the Nei populations were Ad Wu Su (WS) > Ae Bo Le (BL) > Ag A Tu Shi (ATS). In contrast, the Ag Tuo Li (TL) population (0.000, 0.000) had the lowest genetic diversity.

3.2. Structure analysis

Allelic diversity and population differentiation can be effectively used for plant species conservation and breeding genetics[16]. We determined appropriate K values from the structure analysis following the method described by Evanno et al.[17] When the K value was approximately 3–8, the populations were indistinct and more heterogeneous.[24,25] The analysis showed that when the K value was 2, the delta K reached its maximum value (Fig. 4), and the 35 populations representing *Arnebiae Radix* could be divided into two groups. When the K value was 2, the differences between Ae and Ag were not prominent. All the Ae, Ag, and a few Ad (MQ, JH, and WS) strains were very similar and formed one group (the green part). Most of the Ad (SHZ, SW, KLMY, TKS, FY, FH, and QT) were similar and formed another group (the red part) (Fig. 5).

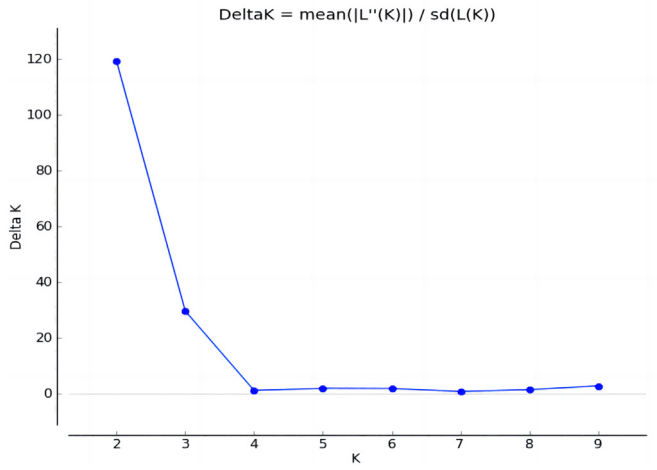


Fig. 4. Deltak change plot of *Arnebiae Radix* with K = 2–9.

3.3. AMOVA

Nei's diversity analysis of the populations of the three *Arnebiae* Discussion *Radix* species using POPGNE 1.32 showed that Nei's total genetic diversity index (Ht) was 0.3344, Nei's intrapopulation genetic diversity index (Hs) was 0.2770, and the interpopulation genetic differentiation coefficient (Gst) was 0.1717[26] (Table 6). The AMOVA results revealed that the coefficients of genetic differentiation (Φst) within and among populations were 78 % and 22 %, respectively (Table 7). This indicated that 78 % of the genetic differentiation was recorded within populations, whereas 22 % of the genetic differentiation was recorded among populations. The value of Gst ranges from 0 to 1. When Gst tends to 0, Ht ≈ Hs, and there is little differentiation between populations. Ht exists mostly within populations. The main genetic variation in *Arnebiae Radix* was found within the populations (Fig. 6). This result was similar to that of Nei's genetic diversity. Additionally, the gene flow index (Nm) between populations was 2.4128 (i.e., >1), which further indicated that the genetic diversity of *Arnebiae Radix* was greater at the intrapopulation level.[27]

4. Discussion

Arnebiae Radix is a traditional Chinese herbal medicine that is widely used in various fields. Several researchers around the world have investigated the active ingredients of *Arnebiae Radix*, including their pharmacological effects and clinical application.[28–30] In China, *Arnebiae Radix* is found mainly in the high-altitude mountainous areas of Xinjiang or the western region of Tibet. Its special growing environment indicates that wild *Arnebiae Radix* has specific requirements for

Table 6
Analysis of the genetic structure of the populations.

Index		Method
Ht	0.3344	Nei's analysis of genes in subdivided populations
Hs	0.2770	
Gst	0.1717	
Nm	2.4128	
$Φst^{v}$(among Pops)	22 %	The Analysis of Molecular Variance (AMOVA)
Φst(within Pops)	78 %	

Ht = Total genotype diversity; Hs = Within-population diversity; Gst = Mean coefficient of gene differentiation value; Nm = estimate of gene flow from Gst. Nm = 0.5 (1-Gst)/Gst; Φst = Genetic differentiation coefficient.

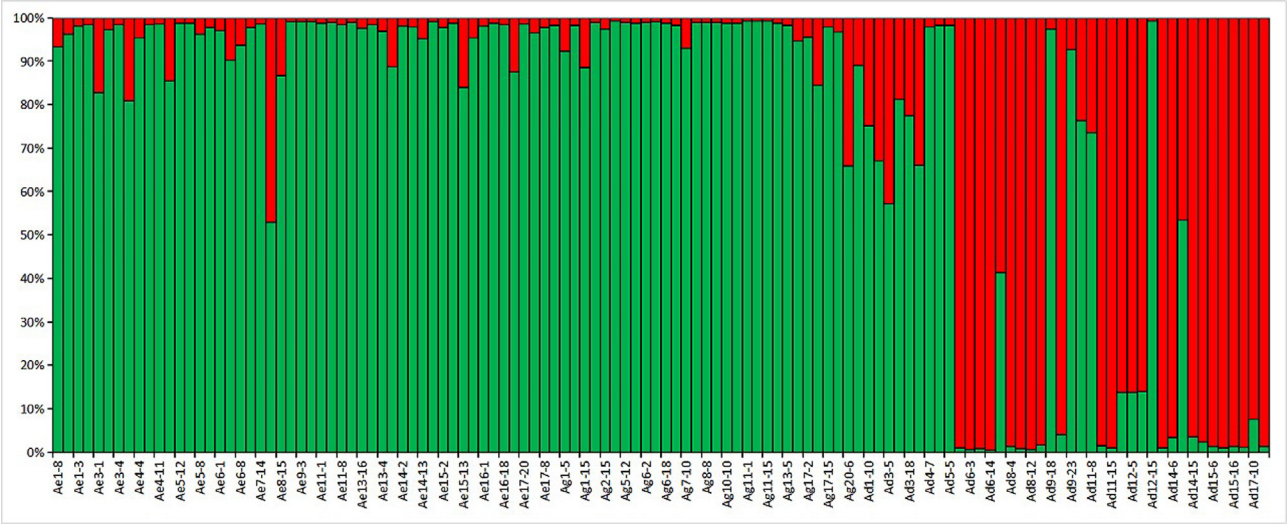


Fig. 5. Structure of the three species of *Arnebiae Radix* with K = 2.

Table 7
Analysis of molecular variance using ISSR markers.

Source	df	SS	MS	Est. Var.	%
Among Pops	2	162.760	81.380	1.905	22 %
Within Pops	117	811.682	6.937	6.937	78 %
Total	119	974.442		8.843	100 %

df, degree of freedom; SS, sum of squares; MS, mean square; Est. Var., estimated variation.

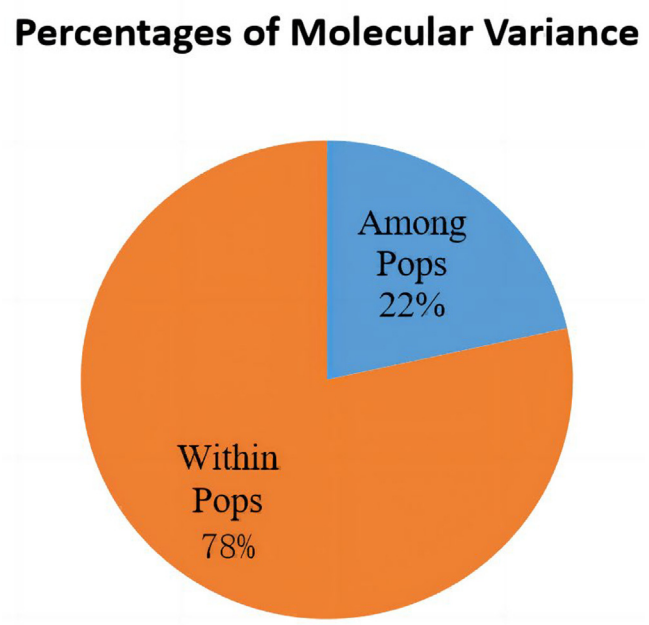


Fig. 6. A graphical representation of the percentage of genetic differentiation.

Table 8
Nei's genetic identity and genetic distance between populations.

Pops	Ae	Ag	Ad
Ae	–	0.9392	0.8548
Ag	0.0627	–	0.8488
Ad	0.1569	0.1640	–

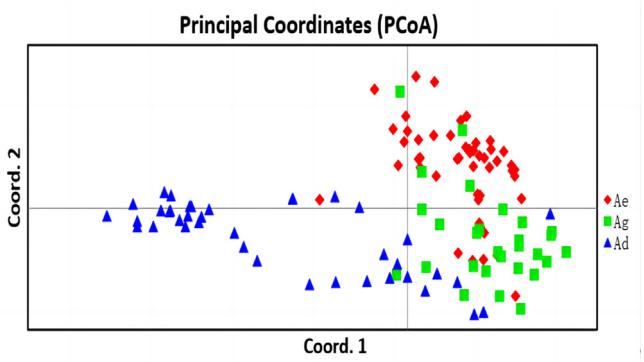
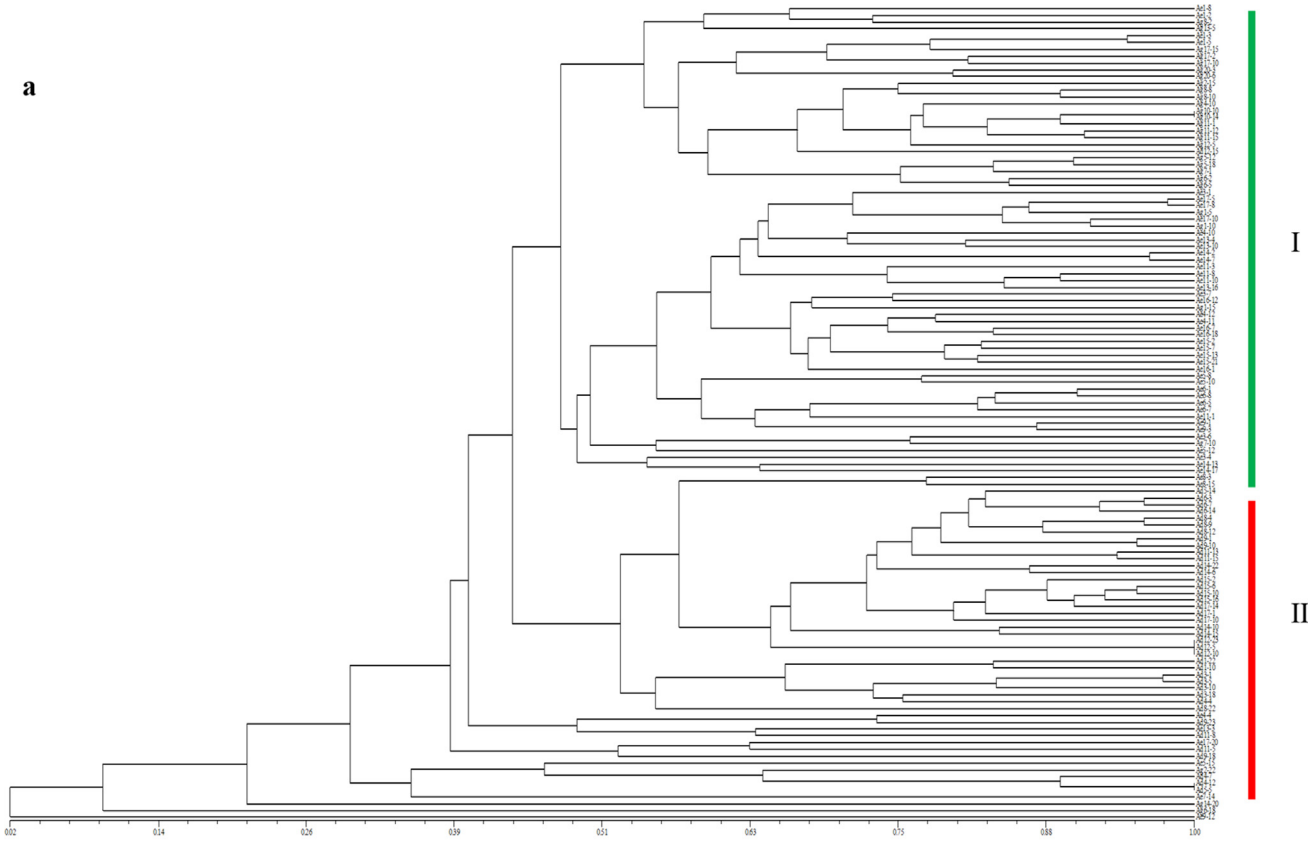


Fig. 7. PCoA of the three *Arnebiae Radix* species.

growth. However, with the overharvesting of wild resources and the imperfection of artificial cultivation technology, wild *Arnebiae Radix* resources are endangered.[31] The number of studies on *Arnebiae Radix* resources has decreased in recent years. Most of those studies were performed only by DNA barcoding and high-resolution melting curve (HRM) technology to establish a method for identifying *Arnebiae Radix* and applying PCR-RFLP to differentiate between Chinese Pharmacopoeia varieties and non-Chinese Pharmacopoeia varieties of marketed *Arnebiae Radix* herbs and tablets.[32,33] Given that traditional physical and chemical identification methods are unsatisfactory for detailed studies of wild *Arnebiae Radix* resources, in this study, we analyzed the population genetic structure of 159 *Arnebiae Radix* samples from 47 populations at the molecular level using ISSR molecular markers. ISSR molecular markers are widely used in the identification of germplasm resources, relationship analysis, genetic diversity determination, etc. Compared with AFLP, RFLP, and RAPD molecular marker technology, ISSR markers have several advantages; for example, information on the genome sequence of the samples is not needed, and the procedure is simple and efficient.[34,35] Therefore, elucidating the genetic diversity of *Arnebiae Radix* germplasm resources through molecular markers is conducive not only to identifying plant varieties but also to rational conservation, cultivation, breeding, sustainable development, and utilization.[36–39] In this study, the differences in genetic structure and differentiation among the populations of *A. euchroma* (Royle) Johnst. (Ae), *A. guttata* Bunge (Ag), and *A. decumbens* (Vent.) Coss. and Kral (Ad) were more pronounced. The genetic variation between populations was investigated by analysis of molecular variance (AMOVA) with measures of population differentiation (Fst). The results showed that 22 % of the total genetic variation occurred between populations, and the genetic variation in the three *Arnebiae Radix* species was mainly intraspecific (78 %). POPGENE 1.32 software revealed that Fst had significant genetic differentiation in the range of 0.0627 to 0.9392 (Table 8). The relationships between genetic distance and genetic identity showed that the Ag population had the smallest genetic distance (0.0627) from the Ae population, and the genetic distance between the Ag and Ad populations was the greatest (0.1640). In contrast, the Ag population had the highest genetic identity (0.9392) with the Ae population and the lowest identity (0.8488) with the Ad population. Therefore, the Ag population had the highest genetic identity with the Ae population, with almost no genetic differentiation. Conversely, the Ag population had the greatest genetic distance from the Ad population, with the highest genetic differentiation. In the POCA, three populations could be divided into two clusters; Ae and Ag represented a class of clusters together, whereas Ad represented a class of clusters alone.[40,41] The results of the principal component analysis also showed that the Ae and Ag populations were concentrated in the second and third quadrants, respectively, and the Ad population was mainly concentrated in the first and fourth quadrants (Fig. 7).

The genetic structure of a population is a fundamental genetic characteristic of a species and is manifested by genetic differentiation between and within species groups.[42] Based on the UPGMA analysis, we divided 120 samples into two major clustering groups (I and II). The Jaccard similarity coefficients ranged between 0.02 and 1.00 (Fig. 8). Cluster I mainly consisted of *A. euchroma* (Royle) Johnst. (Ae) and *A. guttata* Bunge (Ag). Cluster II mainly consisted of *A. decumbens* (Vent.) Coss. and Kral (Ad) (Fig. 8**a**). These results were consistent with the structural and PCA results. All three species of *Arnebiae Radix* were grouped into two main categories, A and B. For the species *A.*



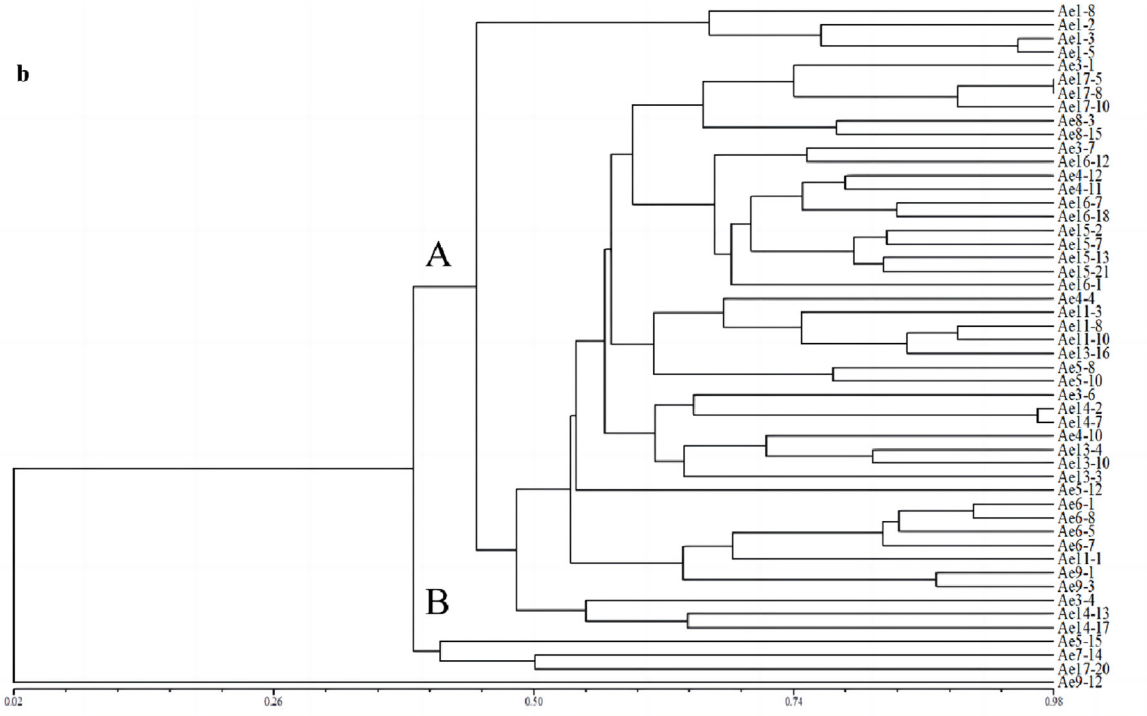


Fig. 8. (a). UPGMA cluster analysis of three *Arnebiae Radix* species. (b). *A. euchroma* (Ae) (c). *A. guttata* (Ag) (d). *A. decumbens* (Ad).

euchroma (Royle), Johnst. (Ae) Individuals from the TSKEG, WQX, ATS, and BL areas formed one category. Among them, the BL population with PPB% = 58.33 %, Na = 1.313, Ne = 1.467, I = 0.0.366, H = 0.255 had the highest index of genetic diversity and bifurcated into a single cluster; the rest of the population belonged to class A **(**Fig. 8**b)**. For the species *A. guttata* Bunge (Ag), individuals from the

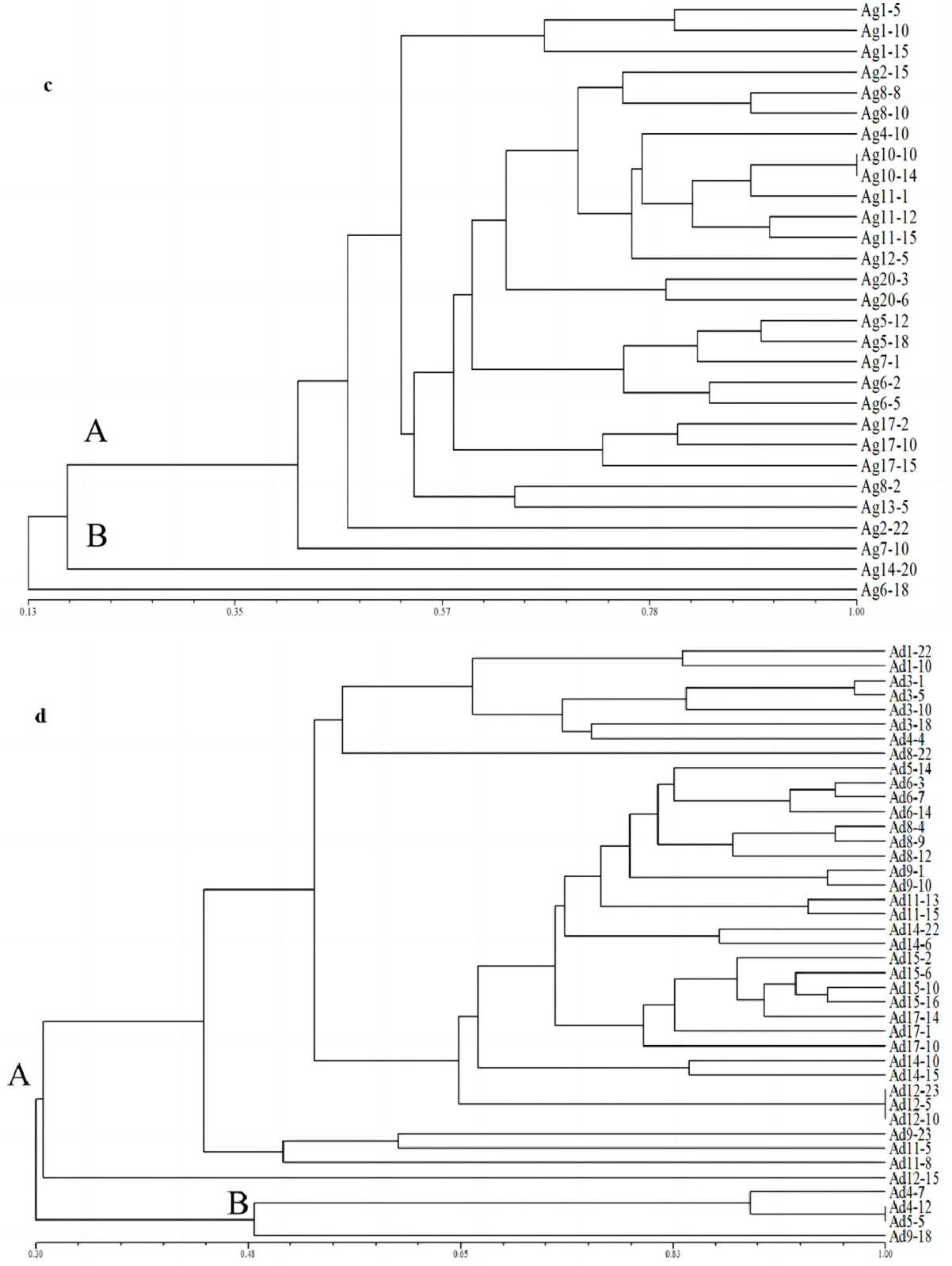


Fig. 8 (*continued*)

XY region were placed in one category. The individuals from the QT region with PPB% = 29.17 %, Na = 0.604, Ne = 1.233, I = 0.186, and H = 0.130 showed high genetic diversity and were placed in one category. The remaining individuals were placed in one category (Fig. 8**c**). For the species *A. decumbens* (Vent.) Coss. and Kral (Ad), the individuals from the JH and WS areas were in one category, whereas those from the KLMY areas had genetic diversity indices of PPB% = 41.67 %, Na = 0.875, Ne = 1.333, I = 0.262, H = 0.182 and formed another category; the remaining individuals were placed in one category (Fig. 8**d**).[43–45]

5. Conclusion

In this study, 159 samples of three species of *Arnebiae Radix* were molecularly characterized by amplifying 48 alleles with six primer pairs. We discussed the interspecific and intraspecific genetic differences among *A. euchroma* (Royle) Johnst., *A. guttata* Bunge, and *A. decumbens* (Vent.) Coss, and Kral. We found that the three *Arnebiae Radix* species could be divided into two major groups: the first two species formed a group that exhibited a close relationship, whereas the third species formed a separate group. Genetic diversity is a very important indicator of the degree of variation in a population; the greater the degree of variation, the greater the genetic diversity of the population, and the richer the genetic resources that can be used to improve the adaptability of species to climate change and historical events. Genetic resources can also facilitate the conservation of species resources and the selection of excellent populations.

Funding

1. The National Natural Science Foundation of China (No. 82060696).
2. Xinjiang Tianshan talent projects (TSYC2021001).
3. Xinjiang Medical University Doctoral Research Initiation Fund Project (B202205).
4. Xinjiang Medical University Graduate Student Innovation and Entrepreneurship Program (CXCY2023013).

CRediT authorship contribution statement

Jinrong Zhao: Writing – review & editing, Software, Formal analysis. **Yanjiao Wang:** Project administration, Methodology, Investigation. **Wenhuan Ding:** Methodology, Formal analysis, Data curation. **Haiyan Xu:** Validation, Supervision, Project administration, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. SHEN Jie, SHEN Wei, CAI Xue, et al. Determination of active components of *Arnebiae Radix* oil by high performance liquid chromatography and its application in the preparation of *Arnebiae Radix* oil by supercritical fluid extraction. Chromatography, 2021; 39: 708-714.
2. WEN E, LIU WENHUI, SONG Hailong, et al. Resource investigation of wild *A. guttata* Bunge in Hami, Xinjiang. Chinese Modern Traditional Medicine, 2016; 18: 1479-1483. 10.13313/j.issn.1673-4890.2016.11.019.
3. J. J., J. Y. Li, G. Q. Li, et al. Materia medica of the medicinal plant *Arnebiae Radix*. Anhui Agricultural Science, 2019; 47: 199-202.
4. Guo H, Sun J, Li D, et al. Shikonin attenuates acetaminophen-induced acute liver injury via inhibition of oxidative stress and inflammation. *Biomed Pharmacother*. 2019;112. https://doi.org/10.1016/j.biopha.2019.108704 108704.
5. ZHU Guoqiang, LIAO Jing, LI Xiaojin, et al. Overview of the resources and distribution of medicinal plants of the family *Arnebiae Radix* in Xinjiang. Chinese Modern Traditional Medicine, 2013; 15: 291-294.10.13313/j.issn.167 3-4890.2013.04.016.
6. Wang H, Kim MK, Kwon WS, et al. Molecular authentication of panax ginseng and ginseng products using robust snp markers in ribosomal external transcribed spacer region. *J Pharm Biomed Anal*. 2011;55:972–976. https://doi.org/10.1016/j.jpba.2011.03.037.
7. Manimekalai R, Suresh G, Govinda Kurup H, et al. Role of NGS and SNP genotyping methods in sugarcane improvement programs. *Crit Rev Biotechnol*. 2020;40:865–880. https://doi.org/10.1080/07388551.2020.1765730.
8. LIU Jie, FANG Wenliang, TANG Zhe, et al. Establishment of an RFLP-HRM Identification Method for *Arnebiae Radix* Herbs Based on DNA Barcoding and HRM Technology Journal of Pharmaceutical Analysis.2022; 42(08): 1354-1362. 10.16155/j.0254-1793.2022.08.08.
9. Tehseen MM, Istipliler D, Kehel Z, et al. Genetic diversity and population structure analysis of Triticum aestivum L. Landrace Panel from Afghanistan. *Genes (Basel)*. 2021;12. https://doi.org/10.3390/genes12030340.
10. Sthapit SR, Ruff TM, Hooker MA, et al. Population structure and genetic diversity of U.S. Wheat varieties. *Plant Genome*. 2022;15:e20196. https://doi.org/10.1002/tpg2.20196.
11. Soliman ERS, El-Shazly H, Börner A, et al. Genetic diversity of a global collection of maize genetic resources in relation to their subspecies assignment s, geographic origin, and drought tolerance. *Breed Sci*. 2021;71:313–325. https://doi.org/10.1270/jsbbs.20142.
12. LI Xiaohong, ZHANG Hui, WANG De yuan, et al. ISSR analysis of the genetic structure of the endemic plant Cycadophyllum in China. Journal of Ecology. Journal of Ecology, 2013; 33: 4892-4901.
13. XU Rong, CHEN Jun. Application of DNA molecular markers in the study of medicinal plant germplasm resources. World Science and Technology, 2006; 58-62 +66.
14. Gupta S, Bharalee R. Genetic diversity and population structure of a medicinal herb Houttuynia cordata Thunb. of North-East India. *Plant Mol Biol Report*. 2021;39:434–442. https://doi.org/10.1007/s11105-020-01260-9.
15. Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. *Genetics*. 2000;155:945–959. https://doi.org/10.1093/genetics/155.2.945.
16. Laidò G, Mangini G, Taranto F, et al. Genetic diversity and population structure of tetraploid wheats (triticum turgidum L.) estimated by ssr, dart and pedigree data. *PLoS One*. 2013;8:e67280.
17. Evanno G, Regnaut S, Goudet J. Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Mol Ecol*. 2005;14:2611–2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x.
18. Tamura K, Stecher G, Kumar S. MEGA11: Molecular evolutionary genetics analysis version 11. *Mol Biol Evol*, 2021; 3810.1093/molbev/msab120:3022-3027.
19. de Paula CF, Paneto GG. Recent patents on high-throughput single nucleotide polymorphism (SNP) genotyping methods. *Recent Pat DNA Gene Seq*. 2012;6:122–126. https://doi.org/10.2174/187221512801327370.
20. Wu W, Chen F, Yeh K, et al. ISSR analysis of genetic diversity and structure of plum varieties cultivated in Southern China. *Biology (Basel)*. 2018;8. https://doi.org/10.3390/biology8010002.
21. Peakall R, Smouse PE. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research–An update. *Bioinformatics*. 2012;28 (19):2537–2539. https://doi.org/10.1093/bioinformatics/bts460.
22. Sheikh ZN, Sharma V, Shah RA, et al. Genetic diversity analysis and population structure in apricot (Prunus armeniaca L.) grown under north-western himalayas using ISSR markers. *Saudi J Biol Sci*. 2021;28:5986–5992. https://doi.org/10.1016/j.sjbs.2021.06.051.
23. Salgotra RK, Chauhan BS. Genetic diversity, conservation, and utilization of plant genetic resources. *Genes (Basel)*. 2023;14. https://doi.org/10.3390/genes14010174.
24. Sabreena, Nazir M, Mahajan R, et al. Deciphering allelic variability and population structure in buckwheat: An analogy between the efficiency of ISSR and SSR markers. *Saudi J Biol Sci*, 2021; 28: 6050–6056.10.1016/j.sjbs.2021.0 7.061.
25. Naugžemys D, Patamsytė J, Žilinskaitė S, et al. Genetic structure of native blue honeysuckle populations in the Western and Eastern Eurasian Ranges. *Plants (Basel)*. 2022;11. https://doi.org/10.3390/plants11111480.
26. Muli JK, Neondo JO, Kamau PK, et al. Genetic diversity and population structure of wild and cultivated Crotalaria species based on genotyping-by-sequencing. *PLoS One*. 2022;17:e0272955. https://doi.org/10.1371/journal.pone.0272955.
27. Hossain MI, Ahmad K, Vadamalai G, et al. Phylogenetic analysis and genetic diversity of Colletotrichum falcatum isolates causing sugarcane red rot disease in Bangladesh. *Biology (Basel)*. 2021;10. https://doi.org/10.3390/biology10090862.
28. MA Shengjun, GENG Yang, MA Lu, et al. Progress of medicinal *Arnebiae Radix* research. *China Modern Chinese Medicine*, 2021; 23: 177-184.10.13313/ j.issn.1673-4890.20190719004.
29. Yingying KANG, Fang XU, Yiping PU, et al. Determination of eight naphthoquinone compounds in *A. euchroma* and optimization of their extraction process. *Chin J Tradit Chin Med*. 2022;37:4720–4724.
30. Wenhuan DING, Xuehua YAN, Liang GE, et al. Determination of naphthoquinone compounds in the roots of four species of Xinjiang *Arnebiae Radix* by HPLC. *Chin Patent Med*. 2019;41:936–939.
31. Guoying LI, Mahmuti M. Distribution of *Arnebiae Radix* in Xinjiang and research progress. *J Xinjiang Med Univ*. 2009;32:386–388.
32. REN Yangyang, JIANG Chunli, ZHANG Peng, et al. Application of DNA barcode identification and determination of efficacy components in quality evaluation of *Arnebiae Radix*. *Chinese Journal of Experimental Formulas*, 2017; 23: 38-43. 10.13422/j.cnki.syfjx.2017120038.
33. LI Qian, LIU Jie, GAO Linong, et al. A PCR-RFLP identification study of *arnebiae radix* based on ITS2 sequence journal of pharmaceutical analysis. *Journal of Pharmaceutical Analysis*, 2016; 36: 1611-1617. 10.16155/j.0254-1793.2016.09.14.
34. Al-Khayri JM, Mahdy EMB, Taha HSA, et al. Genetic and morphological diversity assessment of five kalanchoe genotypes by SCOT, ISSR and RAPD-PCR markers. *Plants (Basel)*. 2022;11. https://doi.org/10.3390/plants11131722.
35. Aghaei M, Hassani A, Nazemiyeh H, et al. Analysis of population structure and genetic diversity of Iranian Wild Salicornia (Salicornia iranica Akhani) population. *J Genet Eng Biotechnol*. 2022;20:61. https://doi.org/10.1186/s43141-022-00337-0.
36. Zhang Rui Xian. A “Codex” of chinese herbal medicines–introduction of “plant and animal breeding and processing technology for medicinal use”. *Northern Pastoralism*, 2003; 23.

37. Yao Yuan. M.S. Studies on the Chemical Composition of *A. guttata* Bunge. M.D. Thesis, Central South Nationalities University, 2018.
38. WANG Sheng, LI Xuan, ZHOU Liangyun, et al. Breeding biology and artificial cultivation of *A. euchroma*. *Modern Chinese Traditional Medicine*, 2011; 13: 18-22 +56. 10.13313/j.issn.1673-4890.2011.11.002.
39. Liwei DING. Current situation of production and marketing of *Arnebiae Radix* and analysis of future market. *Chin Mod Trad Med*. 2011;13:59–61.
40. Tong YY, Liu Y, Wang JW, et al. application of molecular biotechnology in pharmacognosy. *Zhongguo Zhong Yao Za Zhi*. 2016;41:654–658. https://doi.org/10.4268/cjc mm20160418.
41. Niazian M. Application of genetics and biotechnology for improving medicinal plants. *Planta*. 2019;249:953–973. https://doi.org/10.1007/s00425-019-03099-1.
42. Qi H, Sun X, Yan W, et al. Genetic relationships and low diversity among the tea-oil Camellia species in Sect. Oleifera, a bulk woody oil crop in China. *Front Plant Sci*. 2022;13. https://doi.org/10.3389/fpls.2022.996731 996731.
43. Rani R, Raza G, Tung M H, et al. Genetic diversity and population structure analysis in cultivated soybean (glycine max [l.] merr.) using SSR and EST-SSR markers. *PLOS One*, 2023; 18: e0286099. 10.1371/journal.pone.0286099.
44. Hong Z, Peng D, Tembrock LR, et al. Chromosome-level genome assemblies from two sandalwood species provide insights into the evolution of the Santalal- les. *Commun Biol*. 2023;6:587. https://doi.org/10.1038/s42003-023-04980-2.
45. Yin X, Xiang Y, Huang FQ, et al. Comparative genomics of the medicinal plants Lonicera macranthoides and L. japonica provides insight into genus genome evolution and hederagenin-based saponin biosynthesis. *Plant Biotechnol J*. 2023. https://doi.org/10.1111/pbi.14123.