Electronic Journal of Biotechnology 62 (2023) 13–20

Contents lists available at ScienceDirect

Electronic Journal of Biotechnology

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

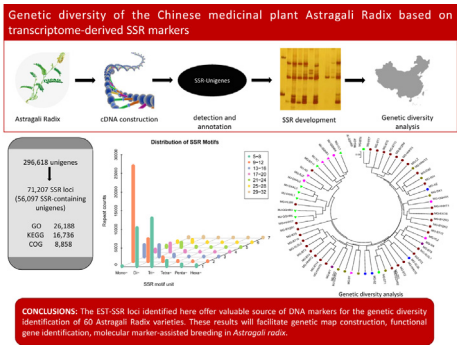
Research Article

Genetic diversity of the Chinese medicinal plant Astragali Radix based on transcriptome-derived SSR markers

Ming Jiang [a,b], Song Yan [a], Weichao Ren [a], Nannan Xing [a], Hongyuan Li [a], Meiqi Zhang [c], Meiqi Liu [a], Xiubo Liu [d], Wei Ma [a,*]

[a] School of Pharmacy, Heilongjiang University of Chinese Medicine, Harbin 150000, Heilongjiang, China
[b] Department of Science and Technology, Qiqihar Medical University, Qiqihar 161006, Heilongjiang, China
[c] School of Forestry, Northeast Forestry University, Harbin 150000, Heilongjiang, China
[d] School of Jiamusi, Heilongjiang University of Chinese Medicine, Harbin 150000, Heilongjiang, China

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 14 March 2022
Accepted 19 December 2022
Available online 26 December 2022

Keywords:
Astragali Radix
DNA markers
Fabaceae
Flavonoids
Genetic diversity
Medicinal plant
SSR loci
SSR markers
ST-SSRs
Transcriptome

ABSTRACT

Background: Astragali Radix is regarded as an important traditional medicinal plant and has been widely used as a Qi-Invigoratin medicine for more than 2000 years. However, the mechanism of its medicinal and nutritional components in different accessions is unexplored. Moreover, the lack of SSR markers for Astragali Radix has stalled genetic breeding and variety identification. To develop EST-SSRs, transcriptome sequencing was performed based on the Illumina platform.

Results: Approximately 99.6 million raw data points were collected, and 296,618 unigenes with an average length of 1,459 bp were obtained. A total of 71,207 SSR loci within 56,097 SSR-containing unigenes were identified. Of these SSR-containing unigenes, 26,188 (46.7%), 16,736 (29.8%), and 8,858 (15.8%) were related to items in the GO database, KEGG database, and COG database, respectively. The motifs A/T, AG/CT, AT/AT, and AAT/ATT were the most common types, accounting for 39.78%, 19.14%, 10.12%, and 4.60%, respectively, and the average length of the EST-SSRs was 17.16 bp. We acquired 8 functional markers linked to key genes correlated with the biosynthesis of flavonoids. The genetic similarity among the 60 AR varietal resources ranged from 0.2692 to 0.8077, with an average of 0.5742. The dendrogram results indicated that the 60 AR samples tested had high levels of genetic similarity.

Conclusions: The 71,207 EST-SSR loci identified here offer valuable source of DNA markers for the genetic diversity identification of 60 Astragali Radix varieties. These results will facilitate genetic map

Peer review under responsibility of Pontificia Universidad Católica de Valparaíso.
* Corresponding author.
E-mail address: maweihljucm@126.com (W. Ma).

https://doi.org/10.1016/j.ejbt.2022.12.001

construction, functional gene identification, variety identification, molecular marker-assisted breeding, and genetic diversity analysis in Astragali Radix.

How to cite: Jiang M, Yan S, Ren W, et al. Genetic diversity of the Chinese medicinal plant Astragali Radix based on transcriptome-derived SSR markers. Electron J Biotechnol 2023;62. https://doi.org/10.1016/j.ejbt.2022.12.001.

1. Introduction

Astragali Radix (AR) is a traditional Chinese medicinal herb belonging to the *Legume* family that has been widely used as a Qi-Invigoratin medicine for more than 2000 years; its first record is from the book "Shen Nong's Herbal Classic" [1]. AR shows high medicinal value and is widely used in the treatment of cardiovascular, cancer, immune, respiratory, blood sugar, and hepatic diseases [2]. Extracts from the stems/leaves, flowers, and roots are rich in active saponins, flavonoids, and polysaccharides, which have pharmacological functions such as immunity regulation, antiviral activities, liver and kidney protection, and antitumor activity [3,4,5].

AR contains two accessions in the Chinese Pharmacopoeia, *Astragalus membranaceus* (Fisch.) Bge. var. *mongholicus* (Bge.) Hsiao (*A. mongholicus*), and *Astragalus membranaceus* (Fisch.) Bge. (*A. membranaceus*) [6]. The commercial medicinal materials of AR were mainly collected from wild resources before the 1970s and were mainly distributed in Heilongjiang, Jilin, and Inner Mongolia Provinces. With the deterioration of the ecological environment and the increasing demand for artificial mining, the natural population and distribution of wild AR herbs have decreased sharply. Thus, AR has been included on the level 3 national protected endangered and rare plant list [2]. In recent years, the demand for cultivated AR has increased rapidly not only due to the development of medicinal decoctions and new drugs in China but also due to its inclusion in functional products in the United States [7]. Therefore, the artificial cultivation of AR has produced several different accessions that show unstable phenotypic traits, low levels of active ingredients, and poor quality. The utilization of AR as an important medicinal resource is restricted by the lack of genetic information.

Genetic diversity is an important variable for wild plant conservation, and also a way to investigate the evolution of species in different habitats [8,9]. Plants have several variations of morphological features and quality in the actual production, such as stem color, flower color of AR. Thus, studying the genetic diversity of plants is necessary and will provide genetic background evidence for AR germplasm resources. In the past decade, many studies have employed molecular techniques to measure genetic diversity based on nucleotide information. Molecular markers are considered the most effective technique for evaluating and categorizing different germplasm resources [10]. Owing to their abundance, stability, cost-effectiveness, and ease of use, various DNA molecular markers have been used in the analysis of genetic diversity in traditional medicinal plants [11,12,13]. Different kinds of molecular markers were checked for their discrimination of AR types. A total of 25 ISSR (intersimple sequence repeat) primers were selected to differentiate 95 AR samples [14]. Dong et al. [15] discovered that *A. membranaceus* and *A. mongholicus* shared the most similar genetic information based on ITS (internal transcribed spacer region) sequences. Cheng et al. [16] analyzed *Astragalus* raw materials from stores in Taiwan on the basis of RAPD (random amplified polymorphic DNA) molecular markers. SCAR (sequence-characterized amplified region) markers were designed to differentiate Korean AR from Chinese AR based on cpDNA and ISSR sequences [17].

Compared to other types of DNA markers, SSR (simple sequence repeat) markers, also named microsatellite markers, are abundant, highly polymorphic, reproducible, codominant, reliable, and cost-effective [18]. SSRs are generally acquired from ESTs (expressed sequence tags); however, this process has previously required much time. In traditional medicinal plants, ESTs are regarded as an effective system for gene mining, genetic mapping, gene expression analysis, phylogenetic and diversity studies, variety identification, and molecular marker-assisted breeding. The use of ESTs for discovering DNA molecular markers and hence genetic diversity analysis among related accessions has been an effective pathway. With the steady decrease in time and cost for deep sequencing and high throughput methods, next-generation sequencing has been proven to be a powerful method for SSR identification. Thus, excavating genic SSRs based on a high-throughput transcriptome with its large scale and high efficiency has been considered to be the most appropriate technique. To date, many genetic diversity studies have been performed on traditional Chinese medicinal plants [19,20]. However, few studies of the genetic diversity of AR based on SSR markers have been published to date. The purpose of this study was to identify genic SSR molecular markers in a Chinese medicinal plant that has been widely applied for thousands of years in a number of countries and areas around the world and to estimate the genetic diversity of different AR accessions. Hence, in the present study, identification novel and highly polymorphic SSR markers from high through-put transcriptomic data and their successful application for genetic diversity evaluation of different AR natural populations will not only enrich the existing SSR molecular markers but also develop massive genotyping research to expand molecular marker-assisted breeding strategies of this medicinal plant.

2. Materials and methods

2.1. Plant materials and RNA/DNA extraction

Samples of *A. mongholicus* and *A. membranaceus*, a very famous Chinese traditional herbal plant, were used for this study. Three whole fresh plants of *A. membranaceus* were collected 120 days after sowing, frozen in liquid nitrogen immediately, and transferred to -80°C for RNA extraction. Then, total RNA was pooled using the TRIzol method based on the manufacturer's instructions and used for Illumina sequencing. Sixty different AR resources were collected representing eight natural populations, including Gansu, Shanxi, Heilongjiang, Shaanxi, Qinghai, Ningxia, Hebei, and Inner Mongolia Provinces of China (Table 1). Fresh leaves were used for DNA extraction based on the manufacturer's instructions for the DNA kit (TransGen Biotech, Beijing).

2.2. cDNA library construction, sequencing, SSR-containing unigene detection, and annotation

Illumina analysis was performed based on the HiSeq™ 2000 platform by Novogene Bioinformatics Technology (Beijing, China). Before assembly, the empty reads, low-quality reads, and

Table 1
Resources of the test samples.

Source	Species	Sample Number
China Gansu	*A. mongholicus*	6
China Shanxi	*A. mongholicus*	4
China Heilongjiang	*A. mongholicus*, *A. membranaceus*	12
China Shaanxi	*A. mongholicus*	2
China Qinghai	*A. mongholicus*	1
China Ningxia	*A. mongholicus*	1
China Hebei	*A. mongholicus*, *A. membranaceus*	2
China Inner Mongolia	*A. mongholicus*	32

adaptor-containing reads were deleted from the raw data. Then, *de novo* assembly was performed with these screened high-quality reads using Trinity [21]. The processed SSR-containing unigenes were screened and aligned to different databases, including the KEGG (Kyoto Encyclopedia of Genes and Genomes), COG (Cluster of Orthologous Groups), and GO (Gene Ontology) databases, to determine their annotation information.

2.3. Development of SSR primers

MISA software (http://pgrc.ipk-gatersleben.de/misa/) was utilized to identify the SSR loci in the SSR-containing unigenes. The parameters for motif repeat identification and default criteria for SSR locus calling were based on Li et al. [22]. After detecting microsatellites, primers were designed based on the sequence information with Primer 5.0. The standard conditions for the expected product lengths, melting temperatures, GC contents, and primer lengths can be found in Mishra et al. [23]. A total of 50 SSR primer pairs were selected randomly and used to study the genetic diversity.

2.4. Amplification of SSR markers

Sixty different AR accessions were collected and used to assess genetic diversity. The PCRs were performed in a volume of 20 μl containing 2 μl 10× PCR buffer, 0.8 μl (10 mmol/L) dNTPs, 1.0 μl (7.5 μmol/L) of each reverse and forward primer, 1 U DNA polymerase (TransGen Biotech, Beijing, China), 150 ng genomic DNA template, and 12 μl deionized distilled water.

The PCR amplification program was as follows: an initial step of 95°C for 5 min, followed by 35 cycles of 95°C for 30 s, 59°C for 60 s, and 72°C for 60 s, with a final extension step of 72°C for 10 min and a cooling step of 4°C for 5 min. The PCR products were examined and analyzed based on the method of Meng et al. [24] (Table S1).

2.5. Genetic diversity analysis

A genetic similarity matrix for 60 AR varietal resources was calculated according to the theory of "proportion of shared alleles" using NTSYS-2.1 software [25]. We further used the genetic similarity matrix to construct the unweighted pair group method with arithmetic mean (UPGMA) algorithm tree based on the illustration of genetic distances using MEGA 4.0 software [26].

3. Results

3.1. High-throughput sequencing and assembly

The transcriptome reads of Astragali Radix were sequenced using the Illumina HiSeq-2000 platform, and 99.6 million raw data were collected and used for analysis. The obtained data contained 296,618 unigenes with an average length of 1,459 bp. Among these sequences, there were 90,893 unigenes (30.6%) with a length of less than 500 bp, 92,235 (31.1%) with a length between 501 and 1,000 bp, 77,906 (26.3%) with a length between 1,001 and 2,000 bp, and 35,584 (12.0%) with a length greater than 2,000 bp.

3.2. Annotation and classification of SSR-containing unigenes

A total of 56,097 SSR-containing unigenes were detected across three public databases, 26,188 (46.7%) were related to items in the GO database, 16,736 (29.8%) to those in the KEGG database, and 8,858 (15.8%) to those in the COG database. A total of 26,188 unigenes were identified with GO functional classifications and related to known proteins based on BLAST results. Most of these annotated unigenes were assigned to the biological process group, followed by the molecular function and cellular component groups, and were divided into 43 classes at level 2 (Fig. 1a). Among the biological process categories, cellular process (28.6%) and metabolic process (25.3%) were the most abundant categories. In the molecular function group, binding (44.4%) and catalytic activity (37.0%) were the two major represented classes. The cellular anatomical entity (46.8%), intracellular (26.4%) and protein-containing complex (20.1%) classes within the cellular function category accounted for the largest proportion, followed by virion (2.9%) and virion part (2.9%).

To further understand the biological functions of these SSR-containing unigenes, the KEGG database was searched. A total of 16,736 unigenes were divided into 129 relevant metabolic pathways on the basis of the enrichment results. Among these unigenes, carbon metabolism (ko01200) and plant hormone signal transduction (ko04075) were the most abundant categories (Fig. 1b).

A total of 8,858 unigenes were matched to entries in the COG database, and 26 subgroups were identified (Fig. 2). The largest category was posttranslational modification, protein turnover, chaperones (1,231), followed by general function prediction only (1,136), signal transduction mechanisms (792) and intracellular trafficking, secretion, and vesicular transport (684).

3.3. Frequency and distribution of SSRs

SSRs were found to be highly abundant in all 296,618 clustered unigenes, and 56,097 unigenes containing 71,207 SSR loci were validated in this study. The frequency of occurrence in AR was one SSR per 4.34 kb of unigene sequence. Among all these unigene sequences, 5,641 SSRs showed a compound pattern, and 11,668 sequences contained more than one SSR locus. Among all repeat types, the mononucleotide motif was assessed to be the most abundant (30,381, 42.67%), followed by dinucleotide (23,689, 33.27%), trinucleotide (14,860, 20.87%), hexanucleotide (1,554, 2.18%), tetranucleotide (443, 0.62%), and pentanucleotide (280, 0.39%) motifs (Table 2).

Across all SSR loci, 156 different motif types were checked. The most common mononucleotide type was A/T (28,323, 39.78% of all SSRs), and AG/CT was the most abundant dinucleotide type (13,593, 19.14%), followed by AT/AT (7,203, 10.12%) and AC/GT (2,866, 4.02%). Among the trinucleotides, AAT/ATT was the most frequent type (3,264, 4.60%), followed by AAG/CTT (3,174, 4.47%) and ATC/ATG (2,446, 3.44%). Of the tetranucleotides, the AAAT/ATTT (407, 0.57%) motif was the most common type, followed by AAAG/CTTT (208, 0.29%) and AGAT/ATCT (179, 0.25%). The most abundant penta- and hexanucleotide types were AATAT/ATATT (35, 0.05%) and ACCTCC/AGGTGG (18, 0.03%) (Fig. 3). Overall, the number of repeats for the SSRs ranged from 5 to 20. Of the dinucleotides, 6 to 15 repeats were found for the most common type

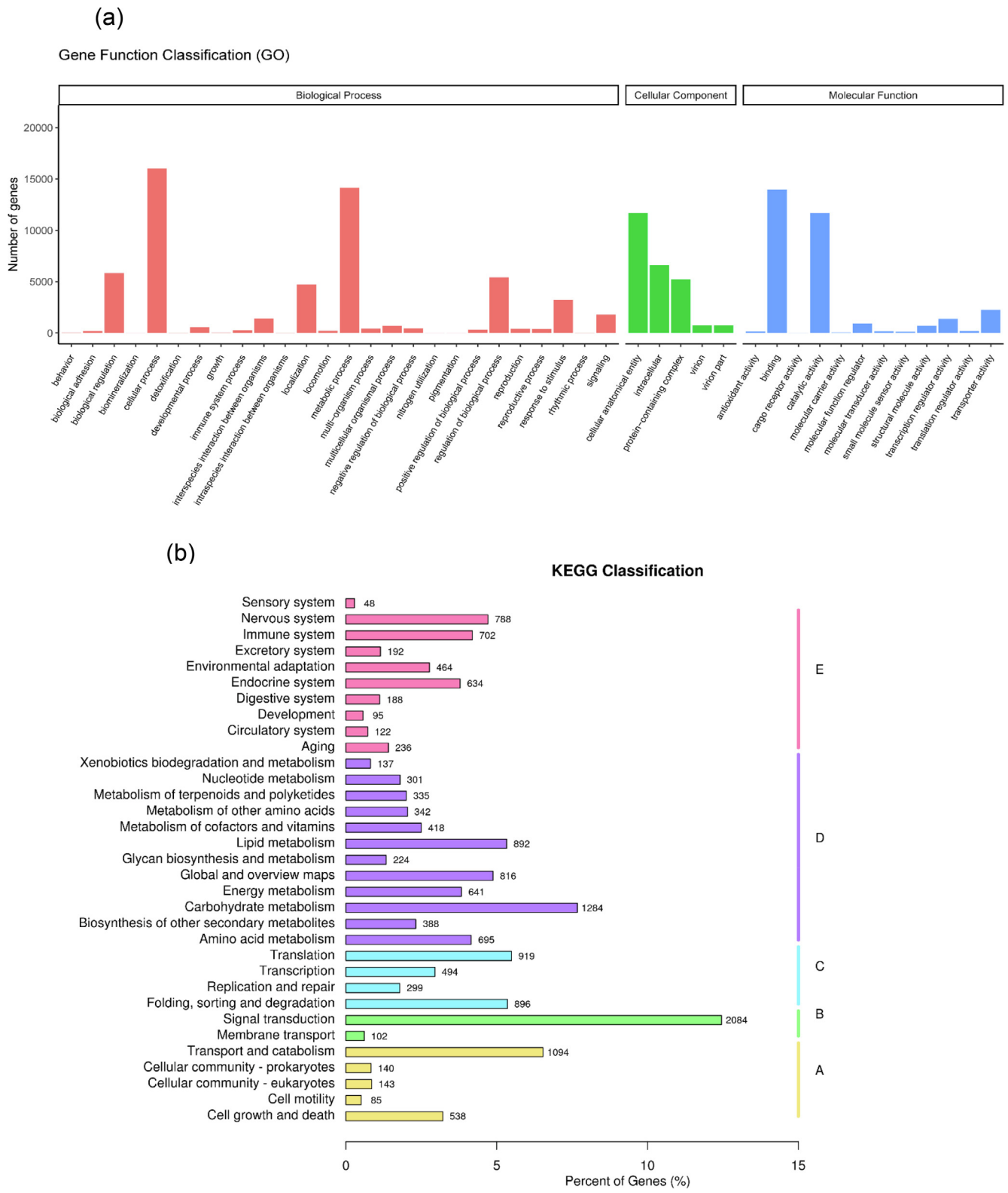


Fig. 1. **a** Gene Ontology identification of assembled unigenes in Astragali Radix; **b** Analysis of 129 relevant metabolic KEGG (Kyoto Encyclopedia of Gene and Genomes) pathways in Astragali Radix.

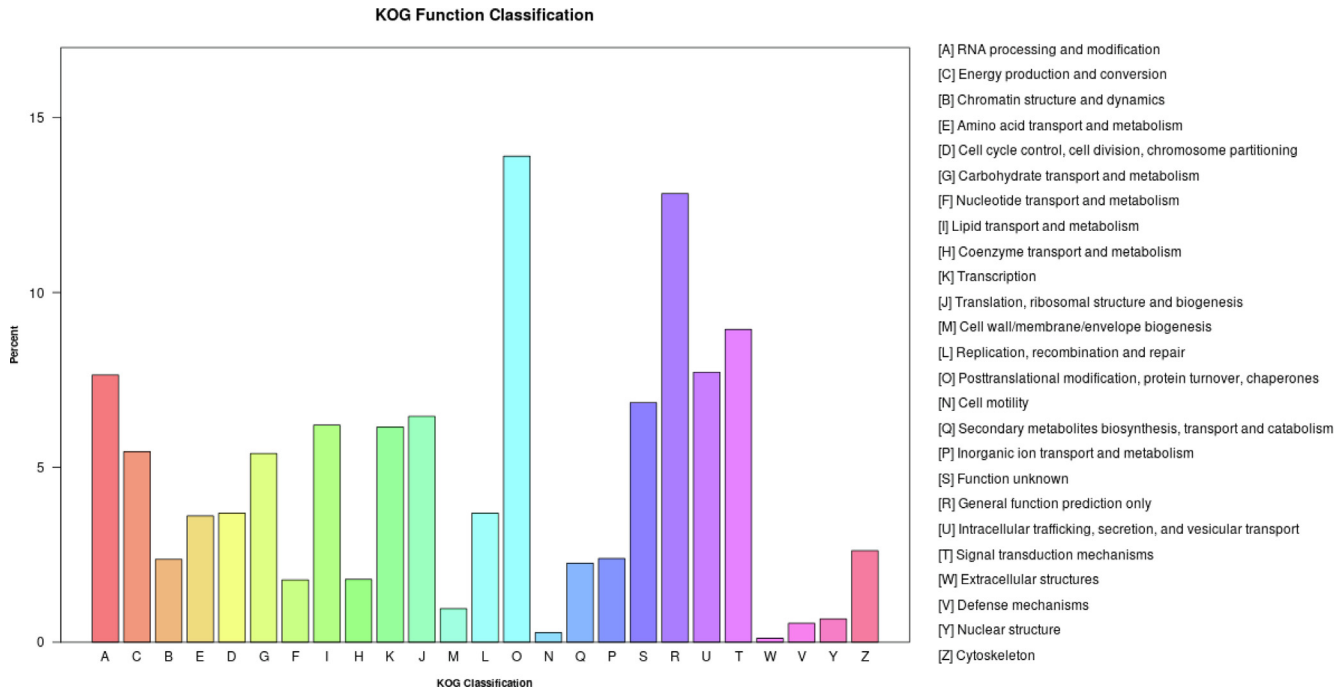


Fig. 2. KOG classification of Astragali Radix SSR-contained unigenes, and 26 subgroups were identified.

Table 2
Description of SSR in transcriptome of Astragali Radix.

Type	Number
Total unigene number	296,618
SSR loci	71,207
SSR-contained unigenes	56,097
Unigenes contained more than one SSR	11,668
Number of SSRs presented a compound formation	5,641
SSR motif of mononucleotide	30,381
SSR motif of dinucleotide	23,689
SSR motif of trinucleotide	14,860
SSR motif of tetranucleotide	1,554
SSR motif of pentanucleotide	443
SSR motif of hexanucleotide	280

(18,555, 78.32%), while 5 to 8 repeats were the most abundant type for the trinucleotides (12,790, 86.07%). The majority of the SSR lengths ranged from 12 to 18 bp (28,153, 39.64%), followed by 19 to 25 bp (8,962, 12.62%), and 7,277 (10.25%) were longer than 30 bp.

3.4. SSR primer design and quality assessment

Of the 71,027 SSR loci, 5,461 were found at the end position of a sequence and could not be used to design appropriate primers for amplification. Thus, we designed 200 primers for the analysis of randomly selected unigenes. We selected 50 (HQ001-HQ050) primer pairs for PCR usability verification. Successful amplification was confirmed when a single fragment was found and/or a fragment with an accurate length was found when several different fragments were amplified for one primer. After validation, 23 pairs were amplified successfully. Among these primers, 10 amplified single fragments, and 13 produced multiple fragments.

3.5. Genetic diversity analysis

The genetic similarity among the 60 AR varietal resources ranged from 0.2692 to 0.8077, with an average of 0.5742, based on the SSR data from the 23 different primer pairs. Based on UPGMA analysis, the genetic distance matrix was used to construct a dendrogram for evaluating the genetic relationships of accessions (Fig. 4). The dendrogram results showed that AR resources from the same area did not form the same subgroup, yet accessions from different regions were clustered in the same group. For instance, the seven resources from the city of Hohhot were dispersed and not clustered together in line with the sample collection regions. The samples with the highest similarity coefficient of 0.8077 were the *A. mongholicus* accessions MG-BT1 and MG-BT6, which are from the city of Baotou, Inner Mongolian region of China, as well as MG-BT6 and MG-GY (from the city of Guyang, Inner Mongolian region). The resources with the lowest coefficient of 0.2692 were *A. mongholicus* variety MG-BT1 and *A. membranaceus* variety MJ-DXAL3 from the Daxing'an Mountains area, Heilongjiang Province. These results indicated that the 60 AR samples tested had high levels of genetic similarity.

4. Discussion

SSRs are widely developed and well-used in research on the genetic diversity of non-model species [10,27]. The lack of SSR information for AR was regarded as a challenge for genetic mapping and variety identification. In this study, 71,207 SSR loci were identified from 56,097 SSR-containing unigenes based on transcriptome information. The type of mononucleotide was calculated in total loci in order to research the most abundant SNPs in future study. A total of 200 primer pairs were designed, 50 SSR primer pairs were amplified to assess quality, and 23 pairs ultimately produced amplicons (46%). The other 27 primer pairs failed mainly because these SSR loci were positioned at splice sites or in introns of the target fragments. Of the 23 EST-SSRs, 10 primer pairs produced a single fragment, showing that 43.5% of the effective SSRs

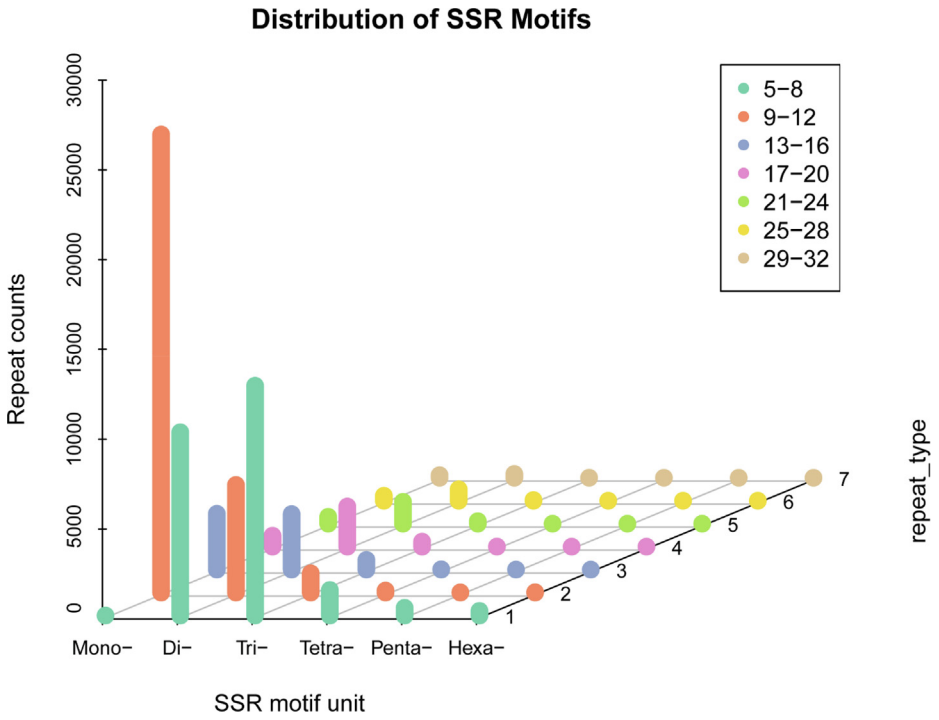


Fig. 3. Distribution of different EST-SSR repeat motif types in Astragali Radix. The most common type was A/T, followed by AG/CT, AT/AT, AC/GT type. Among the trinucleotides and tetranucleotides, AAT/ATT and AAAT/ATTT were the most frequent type, respectively.

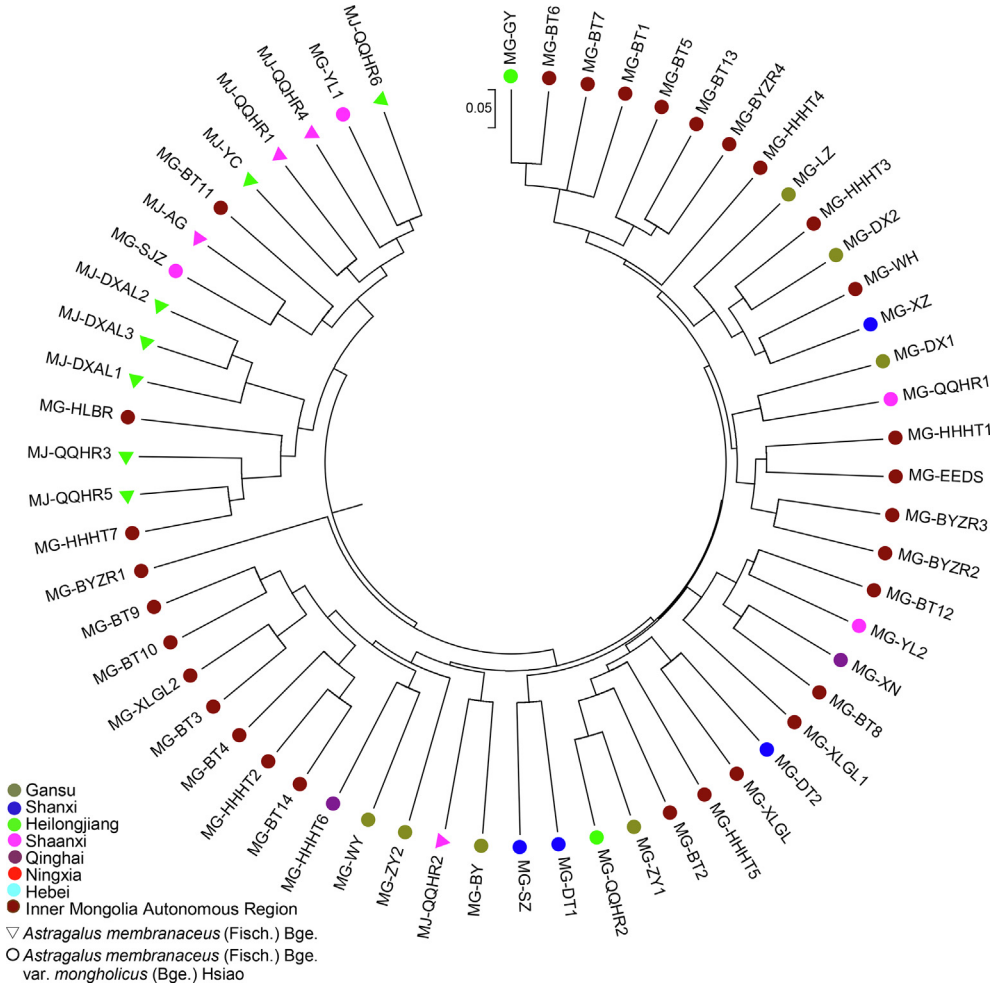


Fig. 4. Unweighted pair group method with UPGMA dendrogram of 60 Astragali Radix accessions based on EST-SSR molecular markers. The dendrogram results showed that AR resources from the same area did not form the same subgroup, yet varieties from different regions were clustered in the same group.

Table 3
Summary statistic for different types of SSR in Astragali Radix

Repeat type	EST-SSR Length (bp)											Total number	Percentage of total SSRs
	10–11	12–13	14–15	16–17	18–19	20–21	22–23	24–25	26–27	28–29	≥30		
Mono-nucleotide	22837	4136	1255	579	310	185	195	202	132	99	451	30381	42.7%
Di-nucleotide	0	4782	3082	2334	1816	1268	1380	1524	840	791	5872	23689	33.3%
Tri-nucleotide	0	0	6835	0	3175	1558	0	1246	364	0	1682	14860	20.9%
Tetra-nucleotide	0	0	0	0	0	871	0	430	0	123	130	1554	2.2%
Penta-nucleotide	0	0	0	0	0	0	0	329	0	0	114	443	0.6%
Hexa-nucleotide	0	0	0	0	0	0	0	0	0	0	280	280	0.4%
Total	22837	8918	11172	2913	5301	3882	1575	3731	1336	1013	8529	71207	100.0%
%	0.32	0.13	0.16	0.04	0.07	0.05	0.02	0.05	0.02	0.01	0.12		

were of high value. Thirteen primers amplified nonspecific amplicons, mainly due to the highly repetitive characteristics of the plant genome. Compared to genomic SSRs, although EST-SSRs are less abundant, they can form functional DNA markers with the advantages of lower cost and greater efficiency. Based on position information, we acquired 8 markers linked to key genes correlated with the biosynthesis of flavonoids (Cluster-8005.151416 linked to *anthocyanidin 3-O-glucoside 2'''-O-xylosyltransferase*, Cluster-8005.168708 linked to *flavonoid 3'-monooxygenase*, Cluster-8005.218030 linked to *flavonoid 3',5'-hydroxylase*, Cluster-8005.117575 linked to *shikimate O-hydroxycinnamoyltransferase*, Cluster-8005.149780 linked to *chalcone isomerase*, Cluster-8005.128384 linked to *trans-cinnamate 4-monooxygenase*, Cluster-8005.162143 linked to *isoflavone 2'-hydroxylase*, and Cluster-8005.114594 linked to *isoflavone-7-O-methyltransferase*), which is one of the most important active ingredients in Astragali Radix.

The 4.34-kb interval per SSR found in this research indicated that the frequency of EST-SSRs in AR was higher than that in most other medicinal plants, such as *Gardenia jasminoides* and *Lycium barbarum,* which had intervals of 2.9 kb and 3.5 kb, respectively [28,29]. However, this value is lower than that for *Epimedium sagittatum,* which was 6.9 kb [25]. Among the repeat motifs, the mononucleotide repeat type was the most abundant (42.67%), followed by di- (33.27%) and trinucleotides (20.87%). Tetra- (2.18%), penta- (0.62%), and hexanucleotides (0.39%) only represent a small percentage of EST-SSR repeat types. Excluding mononucleotide repeats, the EST-SSRs with 6 repeats were the most abundant (20.81%), followed by those with 5 (20.02%), 7 (11.72%), 8 (8.91%), 9 (5.45%), and 10 (3.98%) repeats. The relationship between repeat number and abundance showed a significant negative correlation. The motifs A/T, AG/CT, AT/AT, and AAT/ATT were the most common types, accounting for 39.78%, 19.14%, 10.12%, and 4.60%, respectively. The analysis results indicated that the average length of AR EST-SSRs was 17.16 bp, while the shortest and longest lengths were 10 and 83 bp, respectively. The differentiation between diverse medicinal plants based on EST-SSR length has been performed [25], and the EST-SSRs found in this study mainly ranged from 10 bp to 17 bp, with only 11.98% being larger than 30 bp. The results showed that SSRs in this medicinal plant had shorter rather than longer repeat lengths (Table 3).

The 23 primer pairs developed here were utilized to detect genetic diversity and relationships among 60 AR varieties. The 60 accessions were divided into 2 subgroups containing both species, *A. mongholicus* and *A. membranaceus*, from different origins, which were named "Meng-Gu-Huang-QI" and "Mo-Jia-Huang-Qi" in Chinese, respectively. MG-BT (1, 5, 6, 7, 13), MG-BYZR4, and MG-HHHT4, which were collected from the same province, were clustered into one group, suggesting that AR had the high genetic similarity. The seeds of AR that have been planted every year were mainly from self-preserved collections or were bought from the same area; the lack of breeding for fine accessions breeding may have caused the low genetic distance between types [14]. And some accessions are introduced to different places, so the relationship between accessions shows no good geographical isolation. In general, the results suggested that the EST-SSRs obtained in this study showed high value and could be used for effectively evaluating genetic relationships. Therefore, the 71,207 identified EST-SSR loci offer a valuable basis for DNA markers that will facilitate genetic map construction, genetic diversity analysis, molecular marker-assisted breeding, and functional gene identification in AR.

Ethical approval

This article does not include any studies with human participants or animals performed by any of the authors.

Financial support

This work was supported by the National Natural Science Foundation of China (no. 82173927), the Talent Training Project funded by the Central Government to Support the Reform and Development of Local Colleges and Universities (no. ZYRCB2021008), the National Key Research and Development Projects of Common Key Technologies of Rural Industries (no. 2021YFD1600901), and the Qiqihar Academy of Medical Sciences project (no. QMSI2021M-13). We would like to thank Springer Nature for providing linguistic assistance for manuscript preparation.

Conflict of interest

The authors declare that they have no conflict of interest.

Supplementary material

https://doi.org/10.1016/j.ejbt.2022.12.001.

References

[1] Li SY, Wang D, Li XR, et al. Identification and activity evaluation of *Astragalus Radix* from different germplasm resources based on specific oligosaccharide fragments. Chinese Herbal Med 2020;13(1):33–42. https://doi.org/10.1016/j.chmed.2020.07.004. PMid: 36117754.

[2] Zhang CH, Yang X, Wei JR, et al. Ethnopharmacology, phytochemistry, pharmacology, toxicology and clinical applications of *Radix Astragali*. Chin J Integr Med 2021;27:229–40. https://doi.org/10.1007/s11655-019-3032-8. PMid: 31502185.

[3] Ping L, Zhao H, Luo Y. Anti-aging implications of Astragalus membranaceus (Huangqi): A well-known Chinese tonic. Aging Dis 2017;8(6):868–86. https://doi.org/10.14336/AD.2017.0816. PMid: 29344421.

[4] Niu Y, Wang H, Xie Z, et al. Structural analysis and bioactivity of a polysaccharide from the roots of *Astragalus membranaceus* (Fisch) Bge. var. mongolicus (Bge.) Hsiao. Food Chem 2011;128(3):620–66. https://doi.org/10.1016/j.foodchem.2011.03.055.

[5] Li K, Li SY, Wang D, et al. Extraction, characterization, antitumor and immunological activities of hemicellulose polysaccharide from *Astragalus radix* herb residue. Molecules 2019;24(20):3644. https://doi.org/10.3390/molecules24203644. PMid: 31601012.

[6] National Pharmacopoeia Committee. Pharmacopoeia of Peoples Republic of China. Part 1. Beijing: Chemical Industry Press, 2020;Appendix 1:315.
[7] Xiao WL, Motley TJ, Unachukwu UJ, et al. Chemical and genetic assessment of variability in commercial Radix Astragali (*Astragalus* spp.) by ion trap LC-MS and nuclear ribosomal DNA barcoding sequence analyses. J Agric Food Chem 2011;59(5):1548–56. https://doi.org/10.1021/jf1028174. PMid: 21322556.
[8] Peng JY, Shi C, Wang DW, et al. Genetic diversity and population structure of the medicinal plant *Docynia delavayi* (Franch.) Schneid revealed by transcriptome-based SSR markers. J Appl Res Med Aromat Plants 2021;21:100294. https://doi.org/10.1016/j.jarmap.2021.100294.
[9] Li XL, Wang J, Fan ZQ, et al. Genetic diversity in the endangered *Camellia nitidissima* assessed using transcriptome-based SSR markers. Trees 2020;34:543–52. https://doi.org/10.1007/s00468-019-01935-1.
[10] Zavinon F, Adoukonou-Sagbadja H, Keilwagen J, et al. Genetic diversity and population structure in Beninese pigeon pea [*Cajanus cajan* (L.) Huth] landraces collection revealed by SSR and genome wide SNP markers. Genet Resour Crop Evol 2020;67:191–208. https://doi.org/10.1007/s10722-019-00864-9.
[11] Zhao X, Liu C, Xue W, et al. ISSR Research on Germplasm of *Bupleurum chinense* DC. in Beijing. Modern Chinese Med 2015;17(10):1008–13.
[12] Yao XH, Deng JY, Huang HW, et al. Genetic diversity in *Eucommia ulmoides* (Eucommiaceae), an endangered traditional Chinese medicinal plant. Conserv Genet 2012;13:1499–507. https://doi.org/10.1007/s10592-012-0393-3.
[13] Sarwat M, Das S, Srivastava PS. Analysis of genetic diversity through AFLP, SAMPL, ISSR and RAPD markers in *Tribulus terrestris*, a medicinal herb. Plant Cell Rep 2008;27:519–28. https://doi.org/10.1007/s00299-007-0478-5. PMid: 18074139.
[14] Liu Y, Zhang P, Zhang R, et al. Analysis on genetic diversity of Radix *Astragali* by ISSR Markers. Adv Biosci Biotechnol 2016;7(101):381–91. https://doi.org/10.4236/abb.2016.710037.
[15] Dong T, Xiao QM, Clarke C, et al. Phylogeny of *Astragalus* in China: molecular evidence from the DNA sequences of 5S rRNA spacer, ITS, and 18S rRNA. J Agric Food Chem 2003;51(23):6709–14. https://doi.org/10.1021/jf034278x. PMid: 14582964.
[16] Cheng C. The First New Botanical Drug Approved in Taiwan–PG2 Injection. Proceedings for Annual Meeting of the Japanese Pharmacological Society, 18th World Congress of Basic and Clinical Pharmacology July 01-06, 2018: WCP2018:SY41-4. https://doi.org/10.1254/jpssuppl.WCP2018.0_SY41-4.
[17] Yang JY, Kim HK, Park SJ. Development of genetic marker specific for Korean *hwanggi* medicine (Radix Astragali). Food Sci Biotechnol 2011;20:1561–7. https://doi.org/10.1007/s10068-011-0216-1.
[18] Grover A, Sharma PC. Development and use of molecular markers: past and present. Crit Rev Biotechnol 2016;36(2):290–302. https://doi.org/10.3109/07388551.2014.959891. PMid: 25430893.
[19] Zheng JY, Wang H, Chen XX, et al. Microsatellite markers for assessing genetic diversity of the medicinal plant *Paris polyphylla* var. *chinensis* (Trilliaceae). Genet Mol Res 2012;11(3):1975–80. https://doi.org/10.4238/2012.August.6.1. PMid: 22911581.
[20] Yang HS, Li XP, Liu DJ, et al. Genetic diversity and population structure of the endangered medicinal plant *Phellodendron amurense* in China revealed by SSR markers. Biochem Syst Ecol 2016;66:286–92. https://doi.org/10.1016/j.bse.2016.04.018.
[21] Grabherr MG, Haas BJ, Yassour M, et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol 2011;29:644–52. https://doi.org/10.1038/nbt.1883. PMid: 21572440.
[22] Li H, Ma YS, Pei FY, et al. Large-scale advances in SSR markers with high-throughput sequencing in *Euphorbia fischeriana* Steud. Electron J Biotechnol 2021;49:50–5. https://doi.org/10.1016/j.ejbt.2020.11.004.
[23] Mishra RK, Gangadhar BH, Nookaraju A, et al. Development of EST-derived SSR markers in pea (*Pisum sativum*) and their potential utility for genetic mapping and transferability. Plant Breed 2012;131(1):118–24. https://doi.org/10.1111/j.1439-0523.2011.01926.x.
[24] Meng YS, Zhao N, Li H, et al. SSR fingerprinting of 203 sweetpotato (*Ipomoea batatas* (L.) Lam.) varieties. Journal of Integrative. Agriculture 2018;17(1):86–93. https://doi.org/10.1016/S2095-3119(17)61687-3.
[25] Rohlf FJ. NTSYS-pc: Numerical taxonomy and multivariate analysis system. Setauket, New York: Applied Biostatistics Inc., External Software; 1992.
[26] Tamura K, Dudley J, Nei M, Kumar S. MEGA4: Molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol Biol Evol 2007;24(8):1596–9. https://doi.org/10.1093/molbev/msm092. PMid: 17488738.
[27] Shahabzadeh Z, Mohammadi R, Darvishzadeh R, et al. Genetic structure and diversity analysis of tall fescue populations by EST-SSR and ISSR markers. Mol. Biol. Rep. 2020;47:655–69. https://doi.org/10.1007/s11033-019-05173-z. PMid: 31707600.
[28] Chen C, Xu M, Wang C, et al. Characterization of the *Lycium barbarum* fruit transcriptome and development of EST-SSR markers. PLoS One 2017;12(11):e0187738. https://doi.org/10.1371/journal.pone.0187738. PMid:29125846.
[29] Deng SY, Wang XR, Zhu PL, et al. Development of polymorphic microsatellite markers in the medicinal plant *Gardenia jasminoides* (Rubiaceae). Biochem Syst Ecol 2015;58:149–55. https://doi.org/10.1016/j.bse.2014.11.009.