

EST-SSR markers identification based on RNA-sequencing and application in *Schisandra chinensis*

Guangli Shi

Jilin Agricultural University <https://orcid.org/0000-0001-6531-5546>

Zhenxing Wang

Jilin Agricultural University

Dan Sun

Jilin Agricultural University

Susu Zhang

Jilin Agricultural University

Jianhui Guo

Jilin Agricultural University

Jun Ai (✉ aijun1005@163.com)

<https://orcid.org/0000-0002-7977-1574>

Research article

Keywords: Expressed sequenced tag (EST), *Schisandra chinensis*, Simple sequence repeat (SSR), Genetic diversity analysis, Manual cultivar identification diagram

Posted Date: March 18th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-17573/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background: *Schisandra chinensis*, a climbing woody vine, is the best-known and representative genus of the Schisandraceae family which is an important plant in Chinese herbal medicine; however, the application of molecular breeding is restricted by the few genetic markers for this species.

Results: In this study, we performed transcriptome sequencing of *S. chinensis* using the Illumina HiSeq platform to establish a library of expressed sequence tag-simple sequence repeat (EST-SSR) markers. A total of 59,786 unigenes were obtained and 6254 putative SSR sites were detected with a frequency of 10.46%. The predominant type of repeat motif was dinucleotide (35.71%), followed by trinucleotide (13.22%), hexanucleotide (0.50%), tetranucleotide (0.06%), and pentanucleotide (0.22%). We randomly selected 50 EST-SSR primer pairs and used 14 of these for genetic diversity analysis in 42 *S. chinensis* genotypes. All 42 accessions were successfully identified and formed four major clusters, indicating that the SSR markers can be used for genetic diversity analysis and genetic linkage map construction. In addition, using the polymorphic bands associated with 10 markers as DNA fingerprints, we generated a manual cultivar identification diagram that can distinguish between the 42 accessions, with different individuals identifiable based on polymorphic band patterns.

Conclusion: *S. chinensis* transcriptome data is an effective resource for developing SSR markers. These results can provide a basis for the identification of *S. chinensis* accessions and construction of genetic linkage maps as part of future selective breeding and conservation efforts for this valuable plant.

Background

Schisandra chinensis (Turcz.) Baill. belonging to family Schisandraceae is an important plant in Chinese herbal medicine [1–2]. While only *S. sphenanthera* and *S. chinensis* are listed in the Chinese Pharmacopoeia, in fact, most *Schisandra* species have medicinal value [3–4]; specifically, the fruit has multiple beneficial therapeutic and physiologic properties including adaptogenic, hepatoprotective, anticancer, antioxidant, and anti-inflammatory properties. These are mostly attributable to dibenzocyclooctadiene lignans, which are widely referred to as *Schisandra* lignans [5–6]. About 150 lignan derivatives with a dibenzocyclooctadiene skeleton have been identified in fruit extracts [7]. Due to habitat loss and excessive economic exploitation, the abundance of *Schisandra* species has decreased markedly in recent years, such that they are now endangered [8].

Molecular marker-assisted selective breeding can be applied to the genetic diversification of *Schisandra* varieties. This requires the identification of DNA markers and generation of linkage maps for target gene localization. However, only few types of marker have been used to date in *Schisandra*—namely, inter simple sequence repeats (ISSRs) [9–10], simple sequence repeats (SSRs) [11–12], amplified fragment length polymorphisms (AFLPs) [13–14], and randomly amplified polymorphic DNA (RAPD) [15]. In order to increase the collection of available markers, in the present study, we screened EST-SSR markers for

Schisandra. Our results provide a foundation for research on germplasm resources and functional gene localization as well as for marker-assisted breeding of Schisandra.

Results

SSR site distribution in *S. chinensis*

We searched 59,786 unigene sequences in the transcriptome data and detected 6254 SSR sites in 4989 sequences. The SSR frequency was 10.46%. There were 897 unigene sequences containing two or more EST-SSR sites, and all sequences included a complex SSR site.

The types of EST-SSR detected in the transcriptome varied and their frequencies differed significantly (Table 2); mono-, di-, and trinucleotides were the most common, accounting for 60.06%, 31.61%, and 7.84%, respectively, of all SSRs. Tetra-, penta-, and hexanucleotides were rare, accounting for 0.49% of all SSR. There were 10 SSRs at most SSR sites (22.91%).

Characteristics of EST-SSRs

We analyzed di-, tri-, tetra-, penta-, and hexanucleotides; mononucleotides were excluded because of the potential for poor sequencing quality caused by homopolymerization [16]. In total, 82 types of motif were identified, with 8, 30, 25, 5, and 14 di-, tri-, tetra-, penta-, and hexanucleotide repeats, respectively (complementary sequences were considered one type of motif) (Table 3). The most highly represented EST-SSR type was dinucleotide (35.71%); TC/GA was the most common motif (48.46%) followed by CT/AG (39.52%), with other motif types constituting just 12.02% of all EST-SSR dinucleotides. There were 30 types of trinucleotide repeat motif among *S. chinensis* ESTs; the most frequent were GAA/TTC and AGA/TCT, which accounted for 16.59% (846) and 13.49% (688) of the total, respectively.

The frequency of different motifs in the EST datasets varied, with 20572 (79.39%) dinucleotides, 5100 (19.68%) trinucleotides, 206 (0.81%) tetranucleotides, 27 (0.1%) pentanucleotides, and six (0.02%) hexanucleotides (Fig. 1a). The number of motif repeats ranged from five to 12, with six being the most common number of repeats (Fig. 1b).

Development *S. chinensis* EST-SSR primer pairs and detection of polymorphisms

To obtain high quality SSR primers that could detect polymorphisms, we randomly selected 50 primer pairs to evaluate polymorphisms among four accessions of *S. chinensis* (Yanhong, Zaohong, Jinwuwei, and 12-(-2)-1). We identified 14 pairs of primer sets that were effective (Additional file 1), with a mean amplification rate of 28%.

Discrimination between different *S. chinensis* genotypes using EST-SSR primer pairs

In the genetic diversity analysis, the 14 EST-SSR primer pairs identified as described above could be used to differentiate between the 42 *S. chinensis* accessions. Using NTSYS-pc software to analyze genotype

data, the accessions were classified into four groups at a similarity index of 0.63 (Fig. 2). A dendrogram revealed clear distinctions between the accessions, reflecting a high genetic diversity that can be exploited for *S. chinensis* identification based on a DNA fingerprint. The relatedness of the 42 accessions was supported by similarity coefficients ranging between 0.61 and 0.97. Group I, which comprised 28 varieties mostly originating in Jilin, was the largest group with four subgroups and a similarity coefficient of 0.682. Group II included four varieties; most were from Heilongjiang, with one accession from Jilin. '18-10-3' and '162-1-4' did not cluster with any of the groups and were designated as group IV, and the remaining accessions constituted group III. All 42 accessions were distinguishable based on the 14 EST-SSR markers and their clustering pattern was concordant with their distribution, indicating that EST-SSR data obtained by transcriptome analysis can reveal the genetic relatedness of *S. chinensis* germplasm resources.

Accession identification

There is a need for a simple, practical, and reliable method for identifying *S. chinensis* accessions. Of the 14 primer pairs that were tested, ten were required to clearly distinguish between the 42 accessions (Fig. 3). All accessions were initially identified based on different combinations of the 220-, 270-, and 280-bp bands amplified by primer pair no. 30 (Fig. 3). The smallest group contained only two strains—18-10-3 and 17-N1-N1—that were further distinguished based on a 240-bp band amplified by primer set no. 11. Likewise, all five of the other groups could be differentiated using the primers shown in Figure 3. Thus, all of the accessions could be identified using 10 pairs of primers for the construction of a manual cultivar identification diagram (MCID).

Discussion

S. chinensis (Turcz.) Baill. is a plant used to treat asthma and cough in traditional Chinese herbal medicine that is mainly distributed in northeastern China as well as in Korea, far eastern Russia, and northern Japan [17–19]. However, appreciation of the medicinal value of *S. chinensis* has led to its over-exploitation as well as habitat destruction, which has severely depleted natural *S. chinensis* sources. There is growing interest among herbalists and across the general population to preserve natural sources of important herbal medicine plants, including *S. chinensis* [8, 20–21]. Sustainable use of *S. chinensis* requires an understanding of its population genetic structure and diversity.

DNA markers are useful tools for genetic diversity analysis owing to their abundance, codominance, reproducibility, and high degree of polymorphism. A variety of markers including *rbcL*, internal transcribed spacers, AFLPs, ISSRs, and SSRs have been described in *S. chinensis* [11–13, 20, 22]. However, although DNA markers have been used to analyze genetic diversity in *Schisandrae*, there are no reports of DNA fingerprints or linkage maps for different genotypes. In the present work, transcriptome data for 42 *Schisandra* accessions were screened for EST-SSR markers. We found 6254 SSR sites in 59,786 unigenes, corresponding to a frequency of 10.46%. This value was influenced by the SSR search threshold and large number of single nucleotide repeats. In order to facilitate comparisons with other

species, we analyzed the results in two groups: one including and the other excluding mononucleotide repeats. In the former context, the SSR frequency in *S. chinensis* was 10.46%, which is higher than onion (5.57%) [23] and *Phoebe nanmu* (9.90%) [24] but lower than oil palm (11.26%) [25], *Miscanthus* (14.44%) [26], and eggplant (18.32%) [27]. Without considering mononucleotide repeats, the SSR distribution frequency was 5.79%, which is higher than corn (1.5%), barley (3.4%), rice (4.7%) [28], cucumber (4.03%) [29], *Taxus chinensis* (2.24%) [30], and Korean pine (4.24%) [31]; similar to peanut (6.63%) [32], pepper (7.83%) [33], and dal (7.6%) [34]; and lower than *Rosa roxburghii* (20.37%) [35], oak (27.61%) [36], and precocious trifoliolate orange (26.88%) [37]. These differences may reflect actual variations in SSR abundance across species if we ignore the influence of the amount and source of EST data [35]. Thus, transcriptome SSRs are abundant in *Schisandra*, making these an ideal markers for genetic mapping.

Dinucleotides (37.83%) and trinucleotides (14.00%) were the main types of SSR in *Schisandra*; the two times higher frequency of the former compared to the latter is consistent with the trend observed in peanut [32], precocious trifoliolate orange [37], and litchi [20], although there is also evidence that trinucleotides are the most common SSR type [20, 23, 26, 28, 31]. This difference may be related to the characteristics and quantity of EST-SSR and EST resources in plants. TC/GA was the most common dinucleotide repeat unit in *S. chinensis*, which is the same as pigeonpea [34] but in contrast to barley, wheat, corn, sorghum, rice [28], eggplant [27], peanut [32], *Miscanthus* [26], thorn pear [35], and *T. chinensis* [30]. GAA/TTC was the most abundant trinucleotide repeat motif, as reported for pansy [38].

Of the 50 EST-SSR primer pairs that we tested, 14 produced stable and polymorphic bands of the expected size, with an amplification efficiency of 28%; this is comparable to *Morinda officinalis* How (24%) [39] but higher than pigeon pea [34] and lower than onion (60%) [23], thorn pear (54.76%) [35], and *T. chinensis* (53.23%) [30]. Using these primer sets, the accessions could be divided into four distinct groups. To date, only nine EST-SSR markers have been identified to date for *S. chinensis*. In this study, we used 10 EST-SSR markers to generate an MCID identification map for different *S. chinensis* resources. The large number of SSR loci and polymorphic markers identified here based on transcriptome data can facilitate the identification of accessions and construction of genetic linkage maps in *S. chinensis* as part of broader efforts for the selective breeding and conservation of this valuable plant.

Conclusion

In this study, specific SSR markers were developed based on the transcriptome sequencing data of *S. chinensis*, and the frequency and distribution of SSR markers were analyzed. Using these markers, we successfully analyzed the genetic diversity and identified different accessions of *S. chinensis*. The results demonstrated that transcriptome sequencing is an effective method to identify molecular markers. Our work may lay a foundation for genetic diversity, genetic mapping, and marker-assisted selection in *S. chinensis*. It may facilitate *S. chinensis* breeding, as well as studies with other *Schisandra* plants with economic and medicinal value.

Materials And Methods

Sample collection and total RNA extraction

The plant materials were obtained from germplasm resource garden of Jilin Agriculture University (125°24'15"E, 43°48'5"N; Changchun Jilin Province, China) in July, 2019. These accessions originated from the wild and the information on 42 accessions is shown in Table1, and there are no permissions were necessary to obtain branches of *Schisandra chinensis*. 'Yanhong' and 'Jinwuwei' were examined and approved by Jilin Provincial Variety Examination and Approval Committee in 2012 and 2016, respectively. Young leaves (0.5–1.0 g) were collected in July 2019 and ground into powder after freeze drying in liquid nitrogen. Total RNA was extracted using a modified version of the cetyltrimethylammonium bromide (CTAB) method [41]. RNA was visualized by 1.5% agarose gel electrophoresis.

Transcriptome analysis

Transcriptome data were obtained in 2018 by Illumina (San Diego, CA, USA) high-throughput deep sequencing. RNA was extracted with the CTAB method from six to eight *S. chinensis* seedlings at the true leaf stage and then reverse-transcribed into cDNA, which was sent to Biomarker Technologies (Beijing, China) for transcriptome sequencing. A total of 59,786 unigenes were assembled with the Trinity method [42] and used as background data for analysis.

SSR site identification and primer design based on transcriptome data

SSR loci of unigenes were analyzed using MISA software (<https://omictools.com/misa-tool>); the minimum number of mono-, di-, tri-, tetra-, penta-, and hexanucleotide repeats was 10, 6, 5, 5, 5, and 5, respectively. Primer 3 (<http://bioinfo.ut.ee/primer3-0.4.0/>) was used to design primers with the following features: (1) length between 18 and 23 bp; (2) annealing temperature of 55°C–65°C, with a difference of less than 2°C between forward and reverse primers; (3) amplification product size between 80 and 300 bp; and (4) GC content of 40%–60%. In total, 50 SSR primers were randomly selected and synthesized by Shengong Biological Engineering (Shanghai, China).

PCR amplification and data analysis

The 50 EST-SSR primers were tested to identify those that were stable and could detect polymorphisms. The PCR reaction, which had a total volume of 16 µl, contained 8 µl of 2× Ex Taq Master Mix, 0.8 µl of each of primer, 5.4 µl ultrapure water, and about 20 ng DNA template. PCR amplification was carried out as previously described [10]. The products were separated on a 5% polyacrylamide gel; bands were visualized by silver staining. A clear band was assigned a value of 1 and a weak or no band at the same position was given a value of 0.

Declarations

Abbreviations

EST-SSR: Expressed Sequence Tag-Simple Sequence Repeat; ISSRs: Inter Simple Sequence Repeats; AFLPs: Amplified Fragment Length Polymorphisms; RAPD: Randomly Amplified Polymorphic DNA; CTAB: Cetyltrimethylammonium bromide; MCID: Manual Cultivar Identification Diagram.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

The raw data of RNA-seq are deposited in Sequence Read Archives Database

(<http://www.ncbi.nlm.nih.gov/bioproject/609148>) under accession number PRJNA609148. Other dataset supporting the conclusions of this article are included within the article (and its additional file).

Competing interests

The authors declare that they have no conflict of interest.

Funding

This work was supported by National Key R&D Program of China (2019 YFC1710701); The Research project in Jilin Province Science and Technology Department (grant number 20180201008YY); National Natural Science Fund of China (grant number 31870673).

Authors' contributions

SGL and AJ conceived and designed the research. SGL performed the research and wrote this manuscript. WZX, SD, ZSS and GJH analyzed the data. All authors read and approved the manuscript.

Acknowledgments

Not Applicable

References

1. Szopa A, Ekiert H. In vitro cultures of *Schisandra chinensis* (Turcz.) Baill. (Chinese Magnolia Vine)—a potential biotechnological rich source of therapeutically important phenolic acids. *Appl Biochem Biotechnol*. 2014; 166: 1941–1948. <https://doi.org/10.1007/s12010-012-9622-y>.
2. Yan BQ, Wang J, Chen GP, Wang T. Isolation and characterization of polymorphic microsatellite loci in a traditional Chinese medicinal plant, *Schisandra sphenanthera*. *Conserv Genet*. 2009; 10: 615–

617. <https://doi.org/10.615-617>. 10.1007/s10592-008-9590-5.
3. Liu YH. Schisandreae. In: Flora Reipublicae Popularis Sinicae, Science Press, Beijing, China, 1996.
 4. Tian K, Qi S, Cheng S, Chen XG, Hu ZD. Separation and determination of lignans from seeds of *Schisandra* species by micellar electrokinetic capillary chromatography using ionic liquid as modifier. *J Chromatogr A*. 2005; 1078: 181–187. <https://doi.org/10.1016/j.chroma.2005.05.018>
 5. Hancke JL., Burgos RA., Ahumada F. *Schisandra chinensis* (Turcz.) Baill. *Fitoterapia*, 1999; 70: 451–471. [https://doi.org/10.1016/s0367-326x\(99\)00102-1](https://doi.org/10.1016/s0367-326x(99)00102-1).
 6. World Health Organization (WHO). WHO Monographs on Selected Medicinal Plants. Fructus *Schisandrae*, 30: 296. Geneva: World Health Organization, 2007.
 7. Szopa, A, Barnaś, M, Ekiert H. Phytochemical studies and biological activity of three Chinese *Schisandra* species (*Schisandra sphenanthera*, *Schisandra henryi* and *Schisandra rubriflora*): current findings and future applications. *Phytochem Rev*. 2019; 18: 109–128, <https://doi.org/10.1007/s11101-018-9582-0>.
 8. Yu LH, Sun QL, Dou SS, Wang XJ. A studies on the investigation and evaluation germplasm of *Schisandra chinensis*. *Asia Pac Tradit Med*. 2005; 3: 84–87.
 9. Luo C. Population genetic structure and genetic diversity of *Schisandra sphenanthera* in Qinling of China. Dissertation, Shanxi Normal University, 2012.
 10. Sun Y, Wen XY, Huang HW. Population genetic differentiation of *Schisandra chinensis* and *Schisandra sphenanthera* as revealed by ISSR analysis. *Biochem Syst Ecol*. 2010 38: 257–263. <https://doi.org/10.1016/j.bse.2010.01.005>.
 11. Xu M, Wu S, Liu XX. Genetic diversity analysis of *Schisandra sphenanthera* in Qinling mountains based on SSR markers. *J Chin Med Mater*. 2013; 36: 1215–1218. <https://doi.org/10.13863/j.issn1001-4454.2013.08.046>.
 12. Yan B, Wang T, Hu L. Population genetic diversity and structure of *Schisandra sphenanthera*, a medicinal plant in China. *Chin J Ecol*. 2009; 28: 811–819. <https://doi.org/10.13292/j.1000-4890.2009.0152>.
 13. Gu W. Studies on genetic diversity of germplasm in *Schisandra sphenanthera* et Wils. PhD Dissertation, Shanxi Normal University, 2010.
 14. Valente M. Clonality and patterns of genetic diversity in *Schisandra glabra* (Brickell) Rehder (Schisandraceae), a threatened woody vine. MSc Dissertation, University of Tennessee, Knoxville, TN, USA, 2007.
 15. Gao JP. Comparative studies on *Schisandra chinensis* and *Schisandra sphenanthera*. PhD Dissertation, Fudan University, 2003.
 16. Gilles A, Megléczy E, Pech N, Ferreira S, Malausa T, Martin JF. Accuracy and quality assessment of 454 GS-FLX Titanium pyrosequencing. *BMC Genomics*. 2011; 12: 245. <https://doi.org/10.1186/1471-2164-12-245>

17. Ai J, Wang ZX, Qin HY, Xu PL, Liu YX, Fan ST, et al. Status and prospect of *Schisandra chinensis* (Turcz.) Baill. J Northeast Agric Univ. 2012; 43: 14–20. <https://doi.org/10.19720/j.issn.1005-9369.2012.10.003>.
18. Chinese Pharmacopoeia Commission. Pharmacopoeia of the People's Republic of China, Vol. 1. People's Medical Publishing House, 2005, Beijing.
19. Saunders RMK. Monograph of *Schisandra* (Schisandraceae). Syst Bot Monogr. 2000; 58: 1–146. <https://doi.org/10.2307/25027879>
20. Sun QM, Ma WC, Ma SP, Zhao JS, Bai LJ. Characteristics of SSRs derived from ESTs and development of EST-SSR markers in litchi (*Litchi chinensis* Sonn). Sci Agric Sin, 2011; 44: 4037–4049. <https://doi.org/10.3864/j.issn.0578-1752.2011.19.014>.
21. Zhang H, Wang WQ, Yang Q, Zhang Y, Hou FQ, Xu YX. Report on investigation of the resources of *Schisandra chinensis* (Turcz) Baill in the Northeast of China. LiShiZhen Med Mater Med Res. 2007; 18: 1340–1341.
22. Sun Y, Liu YF, Huang HW. Isolation and characterization of polymorphic microsatellite markers in *Schisandra chinensis* (Turcz.) Baill. (Schisandraceae). Conserv Genet Resour. 2009; 1: 119–121. <https://doi.org/10.1007/s12686-009-9029-4>
23. Li MT, Zhang SL, Deng P, Hou XL, Wang JJ. Analysis on SSR information in transcriptome of onion and the polymorphism. Acta Hort Sinica. 2015; 42: 1103–1111. <https://doi.org/10.16420/j.issn.0513-353x.2015-0046>.
24. Shi XD, Zhu XH, Sheng YZ, Zhuang QG, Chen F. Development of SSR markers based on transcriptome sequence of *Phoebe zhennan*. Sci Silvae Sin. 2016; 52: 71–78. <https://doi.org/10.11707/j.1001-7488.20161109>.
25. Xiao Y, Zhou L, Xia W, Mason AS, Yang YD, Ma ZL et al. Exploiting transcriptome data for the development and characterization of gene-based SSR markers related to cold tolerance in oil palm (*Elaeis guineensis*). BMC Plant Biol. 2014; 14: 384. <https://doi.org/10.1186/s12870-014-0384-2>.
26. Nie G, Tang L, Zhang Y, Huang LK, Ma X, Cao X. et al. Development of SSR markers based on transcriptome sequencing and association analysis with drought tolerance in perennial grass *Miscanthus* from China. Front Plant Sci. 2017; 8: 801. <https://doi.org/10.3389/fpls.2017.00801>.
27. Wei MM. Development and application of SSR markers in eggplant based on transcriptome sequencing. MSc Dissertation, Chinese Academy of Agricultural Sciences, 2016.
28. Kantety RV, Rota ML, Matthews DE, Sorrells ME. Data mining for simple sequence repeats in expressed sequence tags from barley, maize, rice, sorghum and wheat. Plant Mol Biol. 2002; 48: 501–510. <https://doi.org/10.1023/a:1014875206165>.
29. Guo S, Zheng Y, Joung JG, Liu SQ, Zhang ZH, Crasta OR et al Transcriptome sequencing and comparative analysis of cucumber flowers with different sex types. BMC Genomics. 2010; 11: 384. <https://doi.org/10.1186/1471-2164-11-384>.
30. Li YL, Yang XX, Zhang JY, Huang SW, Xiong XY. Studies on SSR molecular markers based on transcriptome of *Taxus chinensis* mairei. Acta Hort Sinica. 2014; 41: 735–745. <https://doi.org/>

10.16420/j.issn.0513-353x.2014.04.028.

31. Zhang Z, Zhang HG, Mo C, Zhang L. Transcriptome sequencing analysis and development of EST-SSR markers for *Pinus koraiensis*. *Sci Silvae Sin.* 2015; 51: 114–120.
<https://doi.org/10.11707/j.1001-7488.20150815>.
32. Zhang J, Liang S, Duan J, Wang J, Chen SL, Cheng ZS, et al. De novo assembly and characterization of the transcriptome during seed development, and generation of genic-SSR markers in peanut (*Arachis hypogaea*). *BMC Genomics.* 2012; 13: 90. <https://doi.org/10.1186/1471-2164-13-90>
33. Liu F, Wang YS, Tian XL, Mao ZC, Zou XX, Xie BY. SSR mining in pepper (*Capsicum annuum*) transcriptome and the polymorphism analysis. *Acta Hort Sinica,* 2012; 39: 168–174.
<https://doi.org/10.16420/j.issn.0513-353x.2012.01.006>.
34. Dutta S, Kumawat G, Singh BP, Gupta DK, Singh S, Dogra V, et al. Development of genic-SSR markers by deep transcriptome sequencing in pigeonpea [*Cajanus cajan* (L.) Millspaugh]. *BMC Plant Biol.* 2011; 11: 17. <https://doi.org/10.1186/1471-2229-11-17>.
35. Yan XQ, Lu M, An HM. Analysis on SSR information in transcriptome and development of molecular markers in *Rosa roxburghii*. *Acta Hort Sinica.* 2015; 42: 341–349.
<https://doi.org/10.16420/j.issn.0513-353x.2014-0882>.
36. An M, Deng M, Zheng SS, Song YG. De novo transcriptome assembly and development of SSR markers of oaks *Quercus austrocochinchinensis*, and *kerrii* (Fagaceae). *Tree Genet Genomes.* 2016; 12: 103. <https://doi.org/10.1007/s11295-016-1060-5>.
37. Long D. Large scale development of SSR markers based on transcriptome sequencing of precocious trifoliolate orange. MSc Dissertation, Huazhong Agriculture University, 2014.
38. Du XH, Yang YP, Zhu XP, Niu YL, Liu HC. Development of genic-SSR markers by transcriptome sequencing in *Viola × wittrockiana*. *Acta Hort Sinica.* 2018; 46: 797–806.
<https://doi.org/10.16420/j.issn.0513-353x.2018-041>.
39. Liao BY, Lee SY, Meng KK, Yin QY, Huang CY, Fan Q, et al. Characterization and novel EST-SSR marker development of an important Chinese medicinal plant, *Morinda officinalis* How (Rubiaceae). *Biotechnol. Biotechnol. Equip.* 2019; 33: 1311–1318.
<https://doi.org/10.1080/13102818.2019.1664322>.
40. Doyle J, Doyle J. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem Bull.* 1987; 19: 11–15.
41. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA; Amit I, et al. Full length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology Italic.* 2011; 29: 644-652. <https://doi.org/10.1038/nbt.18>.

Tables

Due to technical limitations, Tables 1 - 3 are only available for download from the Supplementary Files section.

Figures

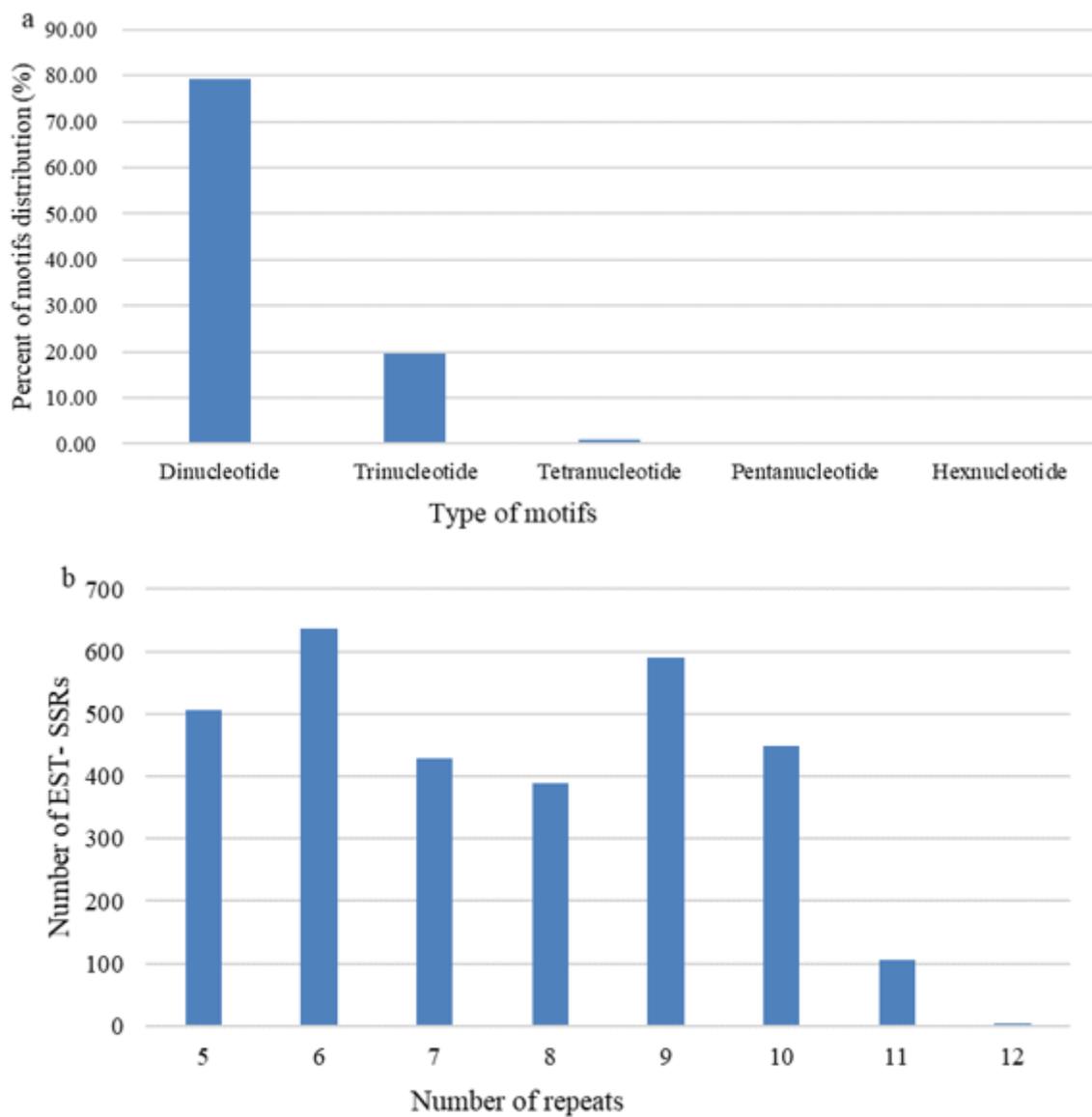


Figure 1

Frequency distribution of EST-SSRs identified from *S. chinensis* ESTs, according to (a) motif type and (b) number of repeats.

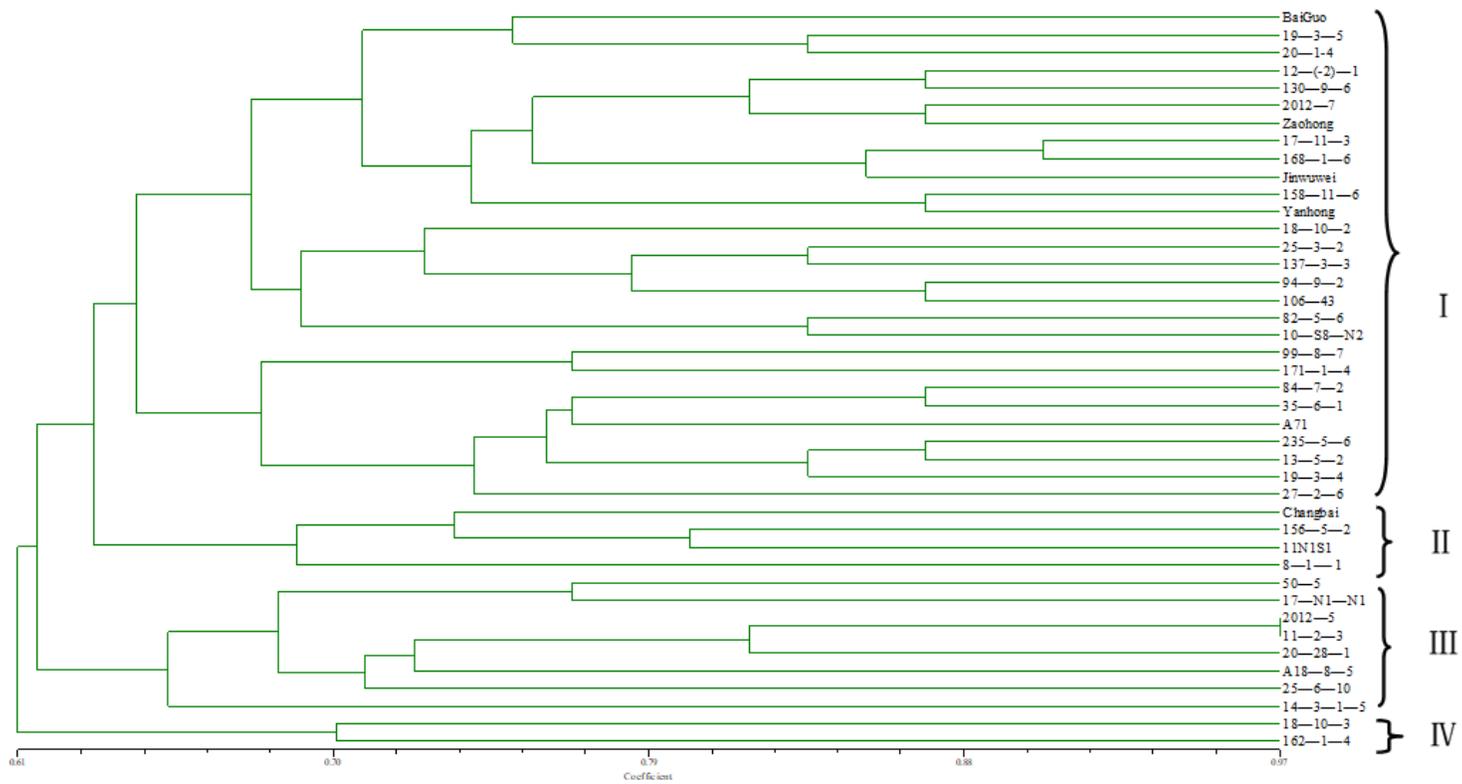


Figure 2

Dendrogram of *S. chinensis* germplasms generated using the unweighted pair group method with arithmetic mean.

Supplementary Files

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

- [Tables.docx](#)