



Research Article

Large-scale advances in SSR markers with high-throughput sequencing in *Euphorbia fischeriana* Steud

Hui Li, Yanshi Ma, Fangyi Pei, Haiyan Zhang, Jicheng Liu, Ming Jiang*

Department of Science and Technology, Qiqihar Medical University, Qiqihar 161006, Heilongjiang, China

ARTICLE INFO

Article history:

Received 22 April 2020

Accepted 24 November 2020

Available online 3 December 2020

Keywords:

Cross-species

EST-SSRs

Euphorbia fischeriana Steud

Genetic diversity analysis

High-throughput sequencing

Linkage maps

Medicinal herb

Molecular marker-assisted breeding

Molecular markers

SSR

ABSTRACT

Background: *Euphorbia fischeriana* Steud is a very important medicinal herb and has significant medical value for healing cancer, edema and tuberculosis in China. The lack of molecular markers for *Euphorbia fischeriana* Steud is a dominant barrier to genetic research. For the purpose of developing many simple sequence repeat (SSR) molecular markers, we completed transcriptome analysis with the Illumina HiSeq 2000 platform.

Results: Approximately 9.1 million clean reads were acquired and then assembled into approximately 186.3 thousand nonredundant unigenes, 53,146 of which were SSR-containing unigenes. A total of 76,193 SSR loci were identified. Of these SSR loci, 28,491 were detected at the terminal position of ESTs, which made it difficult to design SSR primers for these SSR-containing sequences, and the residual SSRs were thus used to design primer pairs. Analyzing the results of these markers revealed that the mononucleotide motif A/T (44,067, 57.83% of all SSRs) was the most abundant, followed by the dinucleotide type AG/CT (9430, 12.38%). Using 100 randomly selected primer pairs, 77 primers were successfully amplified in *Euphorbia fischeriana* Steud, and 79 were successfully amplified in three other related species. The markers developed displayed relatively high quality and cross-species transferability.

Conclusions: The large number of EST-SSRs exploited successfully in *Euphorbia fischeriana* Steud for the first time could provide genetic information for research on linkage maps, variety identification, genetic diversity analysis, and molecular marker-assisted breeding.

How to cite: Li H, Ma Y, Pei F, et al. Large-scale advances in SSR markers with high-throughput sequencing in *Euphorbia fischeriana* Steud. Electron J Biotechnol 2021;49. <https://doi.org/10.1016/j.ejbt.2020.11.004>

© 2020 Pontificia Universidad Católica de Valparaíso. Production and hosting by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Euphorbia fischeriana Steud is a traditional Chinese medicinal plant belonging to the *Euphorbiaceae* family (spurge family) and has a close relationship with other *Euphorbia* species, including *Euphorbia ebracteolata* Hayata and *Euphorbia esula* Linn, as well as *Stellera chamaejasme* Linn. The medicinal part of the plant is called “Lang-Du” and is mainly used in Northeast China as a traditional Chinese medicinal plant to treat cancer, edema and tuberculosis [1]. However, wild *Euphorbia fischeriana* Steud resources are extremely disturbed, and the germplasm population has decreased sharply because of overexploitation [2].

Recently, a number of counterfeit plants have been used as *Euphorbia fischeriana* Steud to treat associated diseases. Different species with the same name are found in Thymelaeaceae, Euphor-

biaceae, Araceae, Ranunculaceae and other families [3,4]. However, research indicates that species and other factors significantly influence the chemical components and medicinal effect of plants [5]. This provides a chance to characterize the genetic basis of key traits in *Euphorbia fischeriana* Steud and enhance the quality of this plant through molecular biotechnology [6]. The genetic diversity of *Euphorbia fischeriana* Steud and other medicinal plants has typically been assessed by random amplified polymorphic DNA (RAPD) [7] and amplified fragment length polymorphism (AFLP) [8] technologies.

The genetic diversity of *Euphorbia fischeriana* Steud populations in different habitats was previously researched with inter-simple sequence repeat (ISSR) markers [9]. However, these molecular markers have several limitations, such as poor reproducibility and incomplete dominance. Instead, simple sequence repeat (SSR) molecular markers display many benefits due to their rich content distribution, abundant polymorphism, reliable neutrality, high repeatability and codominance [10]. Therefore, SSR markers

* Corresponding author.

E-mail address: jiangming314220@126.com (M. Jiang).

have been widely adopted to research genetic diversity, geographical variation, molecular marker-assisted breeding, and variety identification in many vegetative plants [11,12,13,14].

SSRs are normally obtained by genomic sequencing of SSR-containing sections or library construction of expressed sequence tags (ESTs). However, these methods are considered laborious and time-consuming, and genomic SSR markers mainly originate from intergenic DNA fragments, which have unclear locations in the genome. In contrast, transcriptome analysis with high-throughput DNA sequencing technologies on a large scale has been used to rapidly identify genic SSRs with high efficiency. Thus, many SSR molecular markers have been developed and analyzed with this technology in many herbal plants [15,16,17,18,19]. However, few SSR markers for *Euphorbia* have been published to date. Although studies have characterized the transcriptome of *Euphorbia fischeriana* Steud, SSR loci in this species have not been reported [20].

The purpose of the present research was to exploit the EST-SSR molecular markers in a Chinese herbal plant. A *Euphorbia fischeriana* Steud cultivar called “Lang-Du-Da-Ji” was chosen for SSR locus identification and marker development from unigenes assembled on the basis of transcriptome analysis.

Moreover, the cross-species transferability of *Euphorbia fischeriana* Steud SSR markers was examined.

2. Materials and methods

2.1. Plant materials and RNA and DNA extraction

A cultivar of *Euphorbia fischeriana* Steud, a very important Chinese medicinal herb, was used for transcriptome analysis. Three fresh roots of this plant were collected 150 d after planting, frozen in liquid nitrogen and preserved at -80°C immediately for further study. Total RNA was extracted with TRIzol reagent (TransGen Biotech, Beijing, China) on the basis of the product description and used for high-throughput transcriptome sequencing. Fresh leaves of *Euphorbia fischeriana* Steud and three other species (*Stellera chamaejasme* Linn, *Euphorbia ebracteolata* Hayata and *Euphorbia esula* Linn) were gathered, and DNA samples were isolated using a Plant Genomic DNA Kit (Tiangen, China) based on the user instructions.

2.2. cDNA library construction, sequencing, and assembly

Illumina analysis was performed by Novogene Bioinformatics Technology Co., Ltd., Beijing, China (www.novogene.cn). The cDNA library of *E. fischeriana* was constructed with suitable RNA fragments (approximately 180 bp). After that, the paired-end strategy was carried out with the Illumina HiSeq™ 2000 platform. To obtain high-quality reads, we removed the empty reads, N-containing reads and common adaptor sequences from the raw sequences. After that, *de novo* assembly was performed with these filtered clean reads using Trinity [21].

2.3. Gene annotation of SSR-containing unigenes

For the purpose of annotating the assembled SSR-containing unigenes, which were first probed from different databases, including the public NCBI nonredundant protein (Nr) database, the Swiss-Prot protein sequence database, the NCBI nonredundant nucleotide sequence (Nt) database, the Gene Ontology (GO) database, the Kyoto Encyclopedia of Genes and Genomes (KEGG) knowledge database, the Pfam database and the Cluster of Orthologous Groups (COG) database, GO functional classification was executed with the software WEGO [22,23]. GO terms were sepa-

rated into three components: biological processes (BPs), cellular components (CCs) and molecular functions (MFs).

2.4. Discovery of SSR loci and development of markers

The SSR loci were identified from the managed unigenes with Microsatellite Identification Tool (MISA) software (<http://pgrc.ipk-gatersleben.de/misa/>). The default criteria were utilized to select a minimum of ten repeats for mononucleotide motifs, six for dinucleotide motifs and five for the other four motif types (trinucleotide, tetranucleotide, pentanucleotide, and hexanucleotide). SSR primers flanking the putative SSR loci were designed based on the EST sequences with Primer 3.0. The standard parameters related to target amplified fragment lengths, annealing temperatures, GC contents and primer sizes were set according to Mishra et al. [24].

2.5. Amplification of SSR markers

Euphorbia fischeriana Steud and 3 other species (*Euphorbia ebracteolata* Hayata, *Stellera chamaejasme* Linn and *Euphorbia esula* Linn) were selected to assess one hundred chosen SSR primers (LDDJ0001-LDDJ0100) and their cross-species universality. The PCR amplification system included a total volume of 20 μl consisting of 2 μl $10 \times$ PCR buffer, 0.8 μl (10 mmol/L) dNTPs, 0.2 μl (5 U/ μl) DNA polymerase (TransGen Biotech, Beijing, China), 1.0 μl (7.5 $\mu\text{mol/L}$) each SSR primer, 3 μl genomic DNA (50 ng/ μl), and 12 μl deionized distilled water. The PCR program was as follows: 5 min at 95°C , followed by 35 cycles of 30 s at 95°C , 1 min at 59°C , 1 min at 72°C , and a sufficient extension step at 72°C for 10 min and a final cooling step of 5 min at 4°C . The PCR analysis and silver staining method were as described by Meng et al. [25].

3. Results

3.1. High-throughput sequencing and read assembly

The transcriptome data of *Euphorbia fischeriana* Steud were analyzed by the Illumina HiSeq 2000 sequencing platform, and 9.1 million raw reads were acquired from the library and used for assembly. The dataset included 186,384 nonredundant unigenes with a total length of 270.6 Mb and an average length of 1452 bp (Table 1). Among these unigenes, there were 8861 (4.8%) with a length of not more than 300 bp, 26,065 (14.0%) with a length between 301 and 500 bp, 52,210 (28.0%) with a length between 501 and 1000 bp, 54,429 (29.2%) with a length between 1001 and 2000, and 44,819 (24.0%) with a length greater than 2000 bp (Fig. 1).

3.2. Gene annotation and classification of SSR-containing unigenes

The SSR-containing unigene sequences were validated and annotated with seven different public databases. Among these 53,146 unigenes, 42,395 (79.77%) were related to known protein information in the NR database; 37,898 (71.31%), in the NT data-

Table 1
Summary of assembled result with deep sequencing.

Clear reads	91,413,326
Clean bases (Gb)	13.72
Unigene number	186,384
Total length of unigene (Mb)	270.6
Average length of unigene (bp)	1452
Unigene number annotated	144,356 (77.45%)

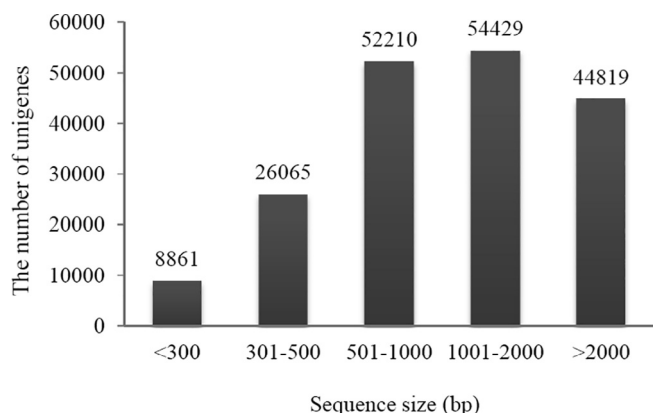


Fig. 1. Overview of the length distribution of unigenes about the *Euphorbia fischeriana* Steud transcriptome assembly.

base; 29,709 (55.9%), in the GO database; and 33,402 (62.85%), in the Swiss-Prot database (Fig. 2a).

GO functional classification of the clustered unigenes identified 29,713 unigenes with BLAST results corresponding to known proteins. Of these, assignments to the biological process category were the most abundant (22,891; 77.04%), followed by cellular component (15,986; 53.80%) and molecular function (26,013; 87.55%) ontologies (Fig. 2b).

The KEGG pathway database was used for the next step to understand the biological functions of SSR-containing unigenes. A total of 16,120 unigenes were matched to 130 relevant metabolic pathways according to the enrichment results. Most of the pathways were associated with carbon metabolism (ko00500) and amino acid biosynthesis (ko00270) (Fig. 2c).

3.3. Frequency and distribution of SSRs

All 186,384 clustered unigenes were used to check possible SSR loci in this study, and 53,146 unigenes including 76,193 SSR loci were checked. The frequency of detection for *Euphorbia fischeriana* Steud was 1 SSR per 3.55 kb of EST sequence. Among the unigenes, 6224 SSRs presented a compound formation, and 16,101 included more than one SSR. The mononucleotide motif was the most abundant type (44,810, 58.81%) among the illustrated SSRs, followed by the di- (14,245, 18.69%), tri- (14,226, 18.67%), hexa- (1651, 2.17%), tetra- (696, 0.91%), and pentanucleotide (565, 0.74%) types (Table 2).

In all SSR loci, 185 different motif types were recognized. The most abundant mononucleotide type was A/T (44,067, 57.83% of all SSRs), and there were 743 C/G nucleotides in total. AG/CT was the most abundant dinucleotide type (9430, 12.38%), followed by AT/AT (3373, 4.43%) and AC/GT (1417, 1.86%). Of the trinucleotides, AAG/CTT was the most common type (6404, 8.41%), followed by ATC/ATG (1897, 2.49%) and ACC/GGT (1604, 2.11%). The AAAG/CTTT (384, 0.50%) motif was the most frequent tetranucleotide type, and the most common pentanucleotides and hexanucleotides were AAAAG/CTTTT (96, 0.13%) and AAAAGG/CCTTTT (29, 0.04%), respectively (Fig. 3). Most of the SSR repeat numbers ranged from 5–20. Of the dinucleotide motif types, 6 repeats was the most common type (8116 SSRs), followed by 7 repeats. Of the trinucleotides, the SSR numbers decreased with decreasing repeat numbers. Most of the SSR lengths (47.2%) ranged from 12 to 18 bp, followed by 19 to 25 bp (13.9%), and 2455 (3.2%) were longer than 30 bp.

3.4. SSR marker development

Of the 76,193 SSR loci, 28,491 were positioned at the ends of an EST, which made it difficult to design primer pairs related to

regions surrounding these loci. We designed 1000 primer pairs for analyzing markers randomly selected from these sequences. These SSR markers were assigned individual numbers (HQ0001–HQ1000).

3.5. Quality estimation of the SSR markers

We amplified 100 SSRs to estimate the applicability of these developed SSR markers with PCR technology. The result was considered successful if a single fragment was amplified with PCR and/or one of the fragments displayed an accurate length when there were several fragments amplified by a primer pair.

In the end, 77 primer pairs achieved exact amplification. Among these primer pairs, 33 produced multiple fragments, and 44 resulted in a single fragment from PCR amplification. Among the sole-amplicon primers, 21 fragments showed accurate sizes, 21 showed longer lengths than expected, and the other 2 showed smaller fragments than expected.

3.6. Transferability evaluation of SSR markers

The species universality of these 100 SSR markers was estimated in three different species. For these 100 markers, 45, 72 and 53 primer pairs were used to obtain PCR fragments from the three species (*Stellera chamaejasme* Linn, *Euphorbia ebracteolata* Hayata and *Euphorbia esula* Linn), respectively (Fig. S1). PCR amplicons obtained by 12 primer pairs from *Euphorbia ebracteolata* Hayata, 28 primer pairs from *Stellera chamaejasme* Linn and 15 primer pairs from *Euphorbia esula* Linn contained multiamplicons. For these primer pairs, it was complicated to distinguish specific fragments from nonspecific fragments because the accurate lengths were unknown. Thus, the SSRs were considered successful and transferable because of special sites combined with the genomic resources of the other species. Consequently, 44 markers showed high universality in all three other species, and 79 could amplify DNA products in at least one of those species.

4. Discussion

In medicinal plants, expressed sequence tags (ESTs) are regarded as an effective strategy for aspects of gene discovery, genomic sequence analysis and important genomic functional annotation, including genetic map construction, molecular marker-assisted breeding, variety identification, microarrays and gene expression profiling. Hence, many researchers have constructed and sequenced cDNA libraries for model plants, such as *Prosopis cineraria* [26], manihot [27] and maize [28], to obtain an EST database. Traditional sequencing strategies are not only time-consuming but also costly, which has greatly limited the number of ESTs developed in nonmodel plants, especially Chinese medicinal herbs. With the application of less time-consuming and less expensive deep-sequencing methods, many ESTs for nonmodel medicinal herbs have become available. To date, the sequencing of several nonmodel plants has been finished, and hundreds of thousands of ESTs have been obtained [29,30,31].

The lack of SSR molecular markers is an important hurdle to genomic analysis of *Euphorbia* species, with only a few SSR markers having been reported, making genetic map construction and species identification difficult. In this research, 1000 SSR primer pairs were used to obtain ESTs based on transcriptome analysis. One hundred SSRs were amplified to evaluate the quality of these primers, and 77 primer pairs finally amplified the target fragments (77%). Because these SSR loci originated from EST sequences, 23% of the SSRs were unsuccessful, mainly because the primers were positioned in introns and/or at splice sites within target fragments.

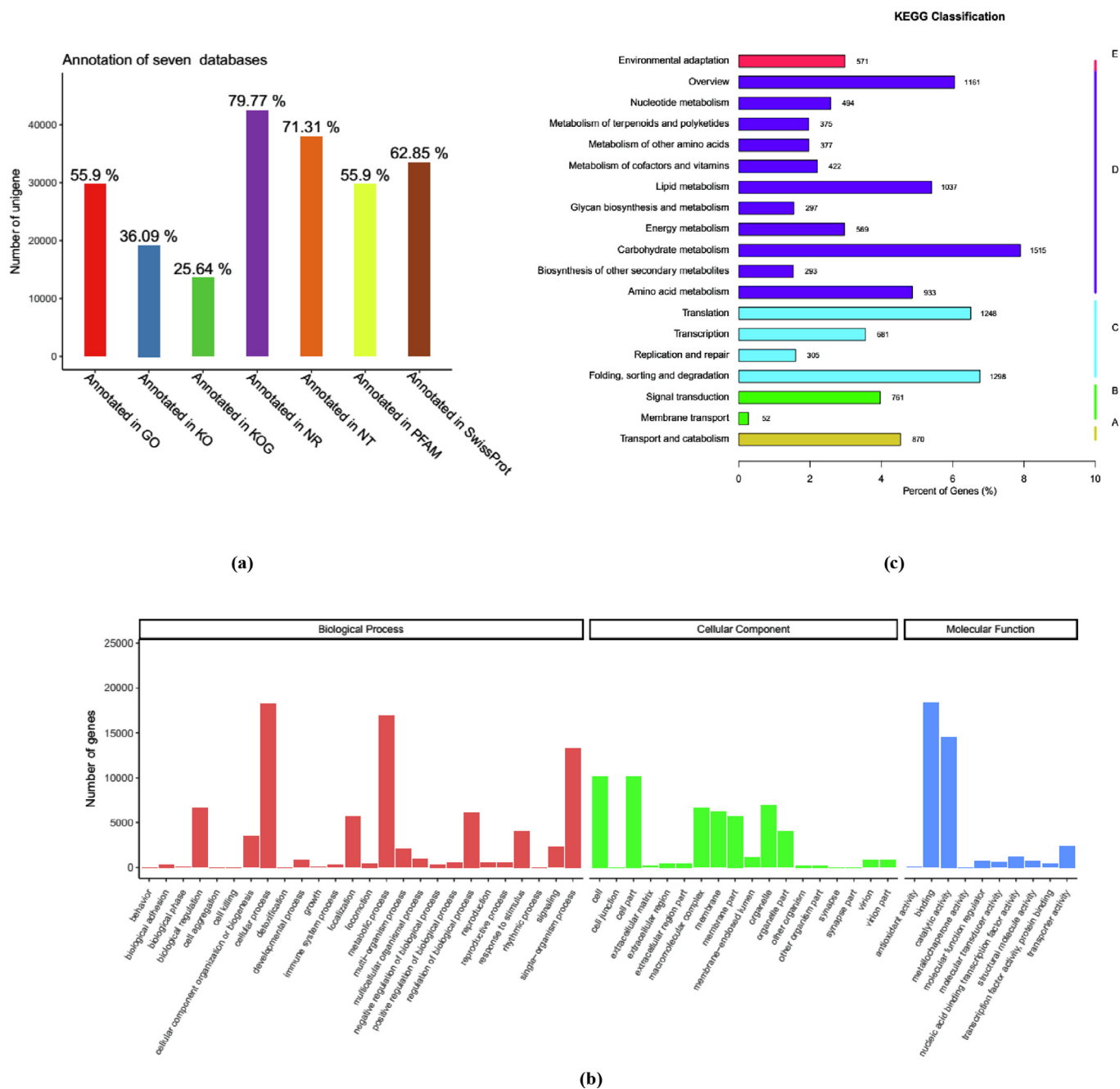


Fig. 2. (a) SSR-contained unigenes frequency associated to the different databases. (b) Gene Ontology identification of assembled unigenes; the results contained three major types: biological process, cellular component, and molecular function. (c) Distribution of Kyoto Encyclopedia of Gene and Genomes (KEGG) pathways in *Euphorbia fischeriana* Steud SSR-contained genes.

For the 100 SSRs, 44 primer pairs amplified only one amplicon, indicating that 44% of the SSRs obtained in the research showed high usefulness. Thirty-three primer pairs resulted in nonspecific fragments, which may be a result of the highly repetitive nature of herbal genome sequences. Even so, the number of nonspecific fragments resulting from the SSR primers was also far smaller than the numbers resulting from AFLP and RAPD techniques. According to the location information, we obtained 10 SSR markers linked to key genes related to biosynthetic pathway of prostratin (3 linked to *3-hydroxy-3-methylglutaryl-CoA reductase*, 2 to *1-deoxy-D-xylulose 5-phosphate reductoisomeras*, 1 to *4-diphosphocytidyl-2-C-methyl-D-erythritol kinase*, 1 to *1-deoxy-D-xylulose 5-phosphate synthase*, 1 to *4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase*, 1 to *far-*

nesyl pyrophosphate synthase, 1 to *Geranylgeranyl pyrophosphate synthase*), which is one particular terpenoid in this plant [20]. Thus, the markers acquired in this study might provide an important resource in future studies for molecular marker-assisted selection.

The 3.55-kb interval discovered in *Euphorbia* EST-SSRs in this study showed that the frequency of EST-SSRs was equivalent to the frequency found in tea (3.5 kb) [32] and higher than that in most other plants [33,34]. Among the repeat motifs, mononucleotide repeats were the most abundant (58.81%), followed by dinucleotides (18.69%) and trinucleotides (18.67%). The number of dinucleotides was the same as that of trinucleotides but not that of mononucleotide motifs, which is contradictory to previous

Table 2
Description of SSR in transcriptome of *Euphorbia fischeriana* Steud.

Type	Number
Total unigene number	186,384
SSR loci	76,193
SSR-contained unigenes	53,146
Unigenes contained more than one SSR	16,101
Number of SSRs presented a compound formation	6224
SSR motif of mononucleotide	44,810
SSR motif of dinucleotide	14,245
SSR motif of trinucleotide	14,226
SSR motif of tetranucleotide	1651
SSR motif of pentanucleotide	696
SSR motif of hexanucleotide	565

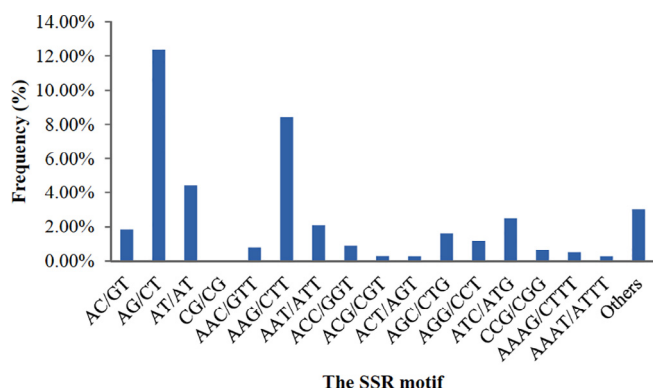


Fig. 3. Frequency characterize of EST-SSRs on the basis of motif types.

results in other plants [35,36]. The different results related to motif frequency, SSR number and average distance could be caused by differences in search criteria and the abundance of ESTs in the databases obtained in previous studies [37]. The A/T, AG/CT and AAG/CTT types were the most abundant, accounting for 57.83%, 12.38% and 8.41% of all the repeat types, respectively. Within GC-rich motifs, CG/CG and CCG/CGG were detected, accounting for 0.03% and 0.65%, respectively. The same result, i.e., a low abundance of GC repeats, was obtained in dicot plants [14].

Among these 100 SSRs from *Euphorbia fischeriana* Steud, 79 displayed universality in at least one of three other species, 44 of which showed universality in all three species. High transferability has also been found in many other plants, including *Chrysanthemum nankingense* [38], *Dysosma versipellis* [39], and *Pinus densiflora* [40]. The high transferability of *Euphorbia fischeriana* Steud SSRs to other species suggests that these species might have a relatively close genetic relationship, and the newly developed markers could have new applications in other *Euphorbia* species, along with other members of the *Euphorbiaceae* family, thus serving as a useful tool for acquiring reliable genomic data in future studies.

Financial support

This work was supported by the Fundamental Research Funds for Education Department of Heilongjiang Province (no. 2017-KYYWF-0107).

Informed consent

This article does not involve any informed consent.

Ethical approval

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

Conflict of interest

The authors declare that they have no conflict of interest.

Acknowledgments

This work was supported by the Fundamental Research Funds for the Education Department of Heilongjiang Province (no. 2017-KYYWF-0107).

Supplementary material

<https://doi.org/10.1016/j.ejbt.2020.11.004>.

References

- Jiang M, Li H. High-throughput sequencing analysis of *Euphorbia fischeriana* Steud provides insights into the molecular mechanism of pharmaceutical ingredient biosynthesis. *3 Biotech* 2018;8(10). <https://doi.org/10.1007/s13205-018-1475-9>.
- Wang R. Advance of research on the exploitation and utilization of poisonous plants resources in Heilongjiang province. *J Anhui Agric Sci* 2009;37:4221–4. <https://doi.org/10.3969/j.issn.0517-6611.2009.09.142>.
- Su X-L, Lin R-C, Wong S-K, et al. Identification and characterisation of the Chinese herb Langdu by LC-MS/MS analysis. *Phytochem Anal* 2003;14(1):40–7. <https://doi.org/10.1002/pca.685>.
- Qiu LB, Sun ZR, Yu JN. Research on the original plants of *Stellera Chamaejasme*. *J Sichuan Traditional Chin Med* 2012;9:52–4 (In Chinese).
- Kamenetsky R, Shafir IL, Zemah H, et al. Environmental control of garlic growth and florogenesis. *J Am Soc Hortic Sci* 2004;129(2):144–51. <https://doi.org/10.21273/JASHS.129.2.0144>.
- Kamenetsky R. Garlic: botany and horticulture. *HorticRev* 2007;33:123–72. <https://doi.org/10.1002/9780470168011.ch2>.
- Das A, Kesari V, Satyanarayana VM, et al. Genetic diversity in ecotypes of the scarce wild medicinal crop *Zingiber moran* revealed by ISSR and AFLP marker analysis and chromosome number assessment. *Giornale Botanico Italiano* 2015;149(1):111–20. <https://doi.org/10.1080/11263504.2013.795197>.
- Muriira NG, Muchugi A, Yu A, et al. genetic diversity analysis reveals genetic differentiation and strong population structure in Calotropis plants. *Sci Rep* 2018;8(1). <https://doi.org/10.1038/s41598-018-26275-x>.
- Li W, Li L, Chen ZJ. Genetic diversity of germplasm resources of *Euphorbia fischeriana* Steud based on ISSR markers. *J Liaoning Univ TCM* 2014;16:50–2. <https://doi.org/10.13194/j.issn.1673-842x.2014.03.018>.
- Liu T, Zeng L, Zhu S, et al. Large-scale development of expressed sequence tag-derived simple sequence repeat markers by deep transcriptome sequencing in garlic (*Allium sativum* L.). *Mol Breeding* 2015;35(11). <https://doi.org/10.1007/s11032-015-0399-x>.
- Chen CL, Xu ML, Wang CP, et al. Characterization of the *Lycium barbarum* fruit transcriptome and development of EST-SSR markers. *PLoS one* 2017;12:11. <https://doi.org/10.1371/journal.pone.0187738> PMID: 29125846.
- Su T, Zhao X, Sui G, et al. Validation of a set of informative simple sequence repeats markers for variety identification in Pak-choi (*Brassica rapa* L. ssp. chinensis var. communis). *Plant Breeding* 2017;136:410–9. <https://doi.org/10.1111/pbr.12483>.
- Fu XQ, Yang FX, Lu XK, et al. Molecular identification of variety purity in a cotton hybrid with unknown parentage using DNA-SSR markers. *Genetics Mol Res* 2017;16:1–8. <https://doi.org/10.4238/gmr16039799>.
- Liu Y, Geng Y, Song M, et al. Genetic structure and diversity of glycyrrhiza populations based on transcriptome SSR markers. *Plant Mol Biol Rep* 2019;37(5-6):401–12. <https://doi.org/10.1007/s11105-019-01165-2>.
- Wu J, Cai C, Cheng F, et al. Characterisation and development of EST-SSR markers in tree peony using transcriptome sequences. *Mol Breeding* 2014;34(4):1853–66. <https://doi.org/10.1007/s11032-014-0144-x>.
- Du F, Wu Y, Zhang L, et al. De Novo assembled transcriptome analysis and SSR marker development of a mixture of six tissues from *Lilium* oriental hybrid 'sorbonne'. *Plant Mol Biol Rep* 2015;33(2):281–93. <https://doi.org/10.1007/s11105-014-0746-9>.
- Wang L, Yang Y, Zhao Y, et al. De Novo characterization of the root transcriptome and development of EST-SSR markers in *Paris polyphylla* Smith var. *yunnanensis*, an endanger medicinal plant. *J Agric Sci Technol* 2016;18:437–52.
- Neeraja C, Mayur D, Mittapelli SR, et al. De novo assembly of leaf transcriptome in the medicinal plant *Andrographis paniculata*. *Front Plant Sci* 2016;7:1–13. <https://doi.org/10.3389/fpls.2016.01203> PMID: 27582746.
- Liao B, Lee SY, Meng K, et al. Characterization and novel Est-SSR marker development of an important Chinese medicinal plant, *Morinda officinalis* How (Rubiaceae). *Biotechnol Biotechnol Equip* 2019;33(1):1311–8. <https://doi.org/10.1080/13102818.2019.1664322>.
- Barrero RA, Chapman B, Yang Y, et al. De novo assembly of *Euphorbia fischeriana* root transcriptome identifies prostratin pathway related genes. *BMC Genomics* 2011;12(1). <https://doi.org/10.1186/1471-2164-12-600>.

- [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(7):644–52. <https://doi.org/10.1038/nbt.1883>.
- [22] Conesa A, Gotz S, Garcia-Gomez JM, et al. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 2005;21(18):3674–6. <https://doi.org/10.1093/bioinformatics/bti610>.
- [23] Ye J, Fang L, Zheng H, et al. WEGO: a web tool for plotting GO annotations. *Nucl Acids Res* 2006;34:W293–7. <https://doi.org/10.1093/nar/gki031>. PMID: 16845012.
- [24] 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 Breeding* 2012;131:118–24. <https://doi.org/10.1111/j.1439-0523.2011.01926.x>.
- [25] Meng YS, Zhao N, Li H, et al. SSR fingerprinting of 203 sweetpotato (*Ipomoea batatas* (L.) Lam.) varieties. *J Integrative Agric* 2018;17(1):86–93. [https://doi.org/10.1016/S2095-3119\(17\)61687-3](https://doi.org/10.1016/S2095-3119(17)61687-3).
- [26] Padaria JC, Thuy NT, Tarafdar A, et al. Development of a heat-responsive cDNA library from *Prosopis cineraria* and molecular characterisation of the *Pchsp17.9* gene. *J Hortic Sci Biotechnol* 2015;90(3):318–24. <https://doi.org/10.1080/14620316.2015.11513189>.
- [27] Zhang Y, Chen X, Wang H, et al. Annotation and validation of genes involved in photosynthesis and starch synthesis from a *Manihot* full-length cDNA library. *Front Agr Sci Eng* 2016;3(4):308. <https://doi.org/10.15302/J-FASE-2016113>.
- [28] Zong N, Li XJ, Wang L, et al. Maize *ABP2* enhances tolerance to drought and salt stress in transgenic *Arabidopsis*. *J Integr Agric* 2018;17(11):2379–93. [https://doi.org/10.1016/S2095-3119\(18\)61947-1](https://doi.org/10.1016/S2095-3119(18)61947-1).
- [29] Miao N, Zhang L, Li M, et al. Development of EST-SSR markers for *Taxillus nigrans* (Loranthaceae) in Southwestern China using next-generation sequencing. *Appl Plant Sci* 2017;5(8):1700010. <https://doi.org/10.3732/apps.1700010>. PMID: 28924510.
- [30] Hu W. Development of 31 EST-SNP markers in *Glycyrrhiza uralensis* Fisch (Leguminosae) based on transcriptomics. *Conservation Genet Resour* 2020;12(2):219–23. <https://doi.org/10.1007/s12686-019-01101-2>.
- [31] Zhang Z, Xie W, Zhao Y, et al. EST-SSR marker development based on RNA-sequencing of *E. sibiricus* and its application for phylogenetic relationships analysis of seventeen *Elymus* species. *BMC Plant Biol* 2019;19(1). <https://doi.org/10.1186/s12870-019-1825-8>.
- [32] Sharma R, Bhardwaj P, Negi R, et al. Identification, characterization and utilization of unigene derived microsatellite markers in tea (*Camellia sinensis* L.). *BMC Plant Biol* 2009;9(1):53. <https://doi.org/10.1186/1471-2229-9-53>.
- [33] Pan L, Xia Q, Quan Z, et al. Development of novel EST-SSRs from sacred lotus (*Nelumbo nucifera* Gaertn) and their utilization for the genetic diversity analysis of *N. nucifera*. *J Hered* 2010;101(1):71–82. <https://doi.org/10.1093/jhered/esp070>.
- [34] Kumar Yadav H, Ranjan A, Asif MH, et al. EST-derived SSR markers in *Jatropha curcas* L.: development, characterization, polymorphism, and transferability across the species/genera. *Tree Genet Genomes* 2011;7(1):207–19. <https://doi.org/10.1007/s11295-010-0326-6>.
- [35] Raji AAJ, Anderson JV, Kolade OA, et al. Gene-based microsatellites for cassava (*Manihot esculenta* Crantz): prevalence, polymorphisms, and cross-taxa utility. *BMC Plant Biol* 2009;9(1):118. <https://doi.org/10.1186/1471-2229-9-118>. PMID: 19747391.
- [36] Wu T, Luo S, Wang R, et al. The first Illumina-based *de novo* transcriptome sequencing and analysis of pumpkin (*Cucurbita moschata* Duch.) and SSR marker development. *Mol Breeding* 2014;34(3):1437–47. <https://doi.org/10.1007/s11032-014-0128-x>.
- [37] Aggarwal RK, Hendre PS, Varshney RK, et al. Identification, characterization and utilization of EST-derived genic microsatellite markers for genome analyses of coffee and related species. *Theor Appl Genet* 2007;114(2):359–72. <https://doi.org/10.1007/s00122-006-0440-x>. PMID: 17115127.
- [38] Wang H, Jiang J, Chen S, et al. Next-generation sequencing of the *Chrysanthemum nankingense* (Asteraceae) Transcriptome permits large-scale unigene assembly and SSR marker discovery. *Plos One* 2013;8(4):e62293. <https://doi.org/10.1371/journal.pone.0062293>. PMID: 23626799.
- [39] Guo R, Mao Y-R, Cai J-R, et al. Characterization and cross-species transferability of EST-SSR markers developed from the transcriptome of *Dyosma versipellis* (Berberidaceae) and their application to population genetic studies. *Mol Breeding* 2014;34(4):1733–46. <https://doi.org/10.1007/s11032-014-0134-z>.
- [40] Liu L, Zhang S, Lian C. De novo transcriptome sequencing analysis of cDNA library and large-scale unigene assembly in Japanese Red Pine (*Pinus densiflora*). *Int J Mol Sci* 2015;16:29047–59. <https://doi.org/10.3390/ijms161226139>. PMID: 26690126.