

Identification of SSR markers using soybean (*Glycine max*) ESTs from globular stage embryos

Ai Qin Li^{1#} · Chuan Zhi Zhao^{1#} · Xing Jun Wang¹ ✉ · Zhan Ji Liu¹ · Li Feng Zhang² · Guo Qi Song¹ · Juan Yin¹ · Chang Sheng Li¹ · Han Xia¹ · Yu Ping Bi¹

1 High-Tech Research Center, Shandong Academy of Agricultural Sciences, Key Laboratory for Genetic Improvement of Crop, Animal and Poultry of Shandong Province, Key Laboratory of Crop Genetic Improvement and Biotechnology, Huanghuaihai, Ministry of Agriculture, Ji'nan 250100, P.R. China

2 Crop Research Institute, Shandong Academy of Agricultural Sciences, Ji'nan 250100, P.R. China

These authors contributed equally to this paper

✉ Corresponding author: xingjunw@hotmail.com
Received December 2, 2009 / Accepted May 26, 2010
Published online: September 15, 2010
© 2010 by Pontificia Universidad Católica de Valparaíso, Chile

Abstract Microsatellites, or simple sequence repeats (SSRs), in expressed sequence tags (ESTs) provide an opportunity for low cost SSR development. We looked for EST-SSRs in 403,511 ESTs (generated by 454 sequencing and representing 70,654 contigs and 52,082 singletons) from soybean globular stage embryos. Among 122,736 unique ESTs, 3,729 contained one or more SSRs. In total, 3,989 SSRs were identified including 304 mono, 1,374 di, 2,208 tri, 70 tetra, 13 penta and 20 hexanucleotide SSRs. Thirty three EST-SSRs were selected for primer design and polymorphism analysis using twenty soybean cultivars and one wild-type soybean. Successful amplification was obtained using 21 of these primer pairs, 11 of which detected polymorphisms in these soybean cultivars. These results demonstrated that 454 high throughput sequencing is a powerful tool for molecular marker development. From the 3,989 identified SSRs we expect to obtain a large number of makers with polymorphism among different soybean cultivars, which would be useful for analysis of genetic diversity and maker assisted selection in the soybean breeding programs.

Keywords: conserved 454 sequencing, EST, *Glycine max*, polymorphism, SSR

INTRODUCTION

Soybean (*Glycine max* L.) is an important crop plant worldwide and has been used in China for 5000 years as source of vegetable oil and food. The soybean genome has been proposed as a model for phaseoloid legumes due to its economic and biological importance, moderate genome size (1.12-1.81 x 10⁹ bp), and its extensive research history (Xia et al. 2007). There are more than 20,000 soybean varieties in China, most of which were derived from hybridization between limited fine strains and therefore have narrow genetic diversity (Qiu et al. 2000). To widen the genetic diversity of

cultivars in future breeding programs, breeders need to know the genetic background of the germplasms they are using. Characterization of current germplasms, for example, by labeling of agronomically valuable traits with molecular markers, is of fundamental importance in this aspect. Reliable molecular marker development and polymorphism analysis of germplasms are plausible first stages in this varietal characterization.

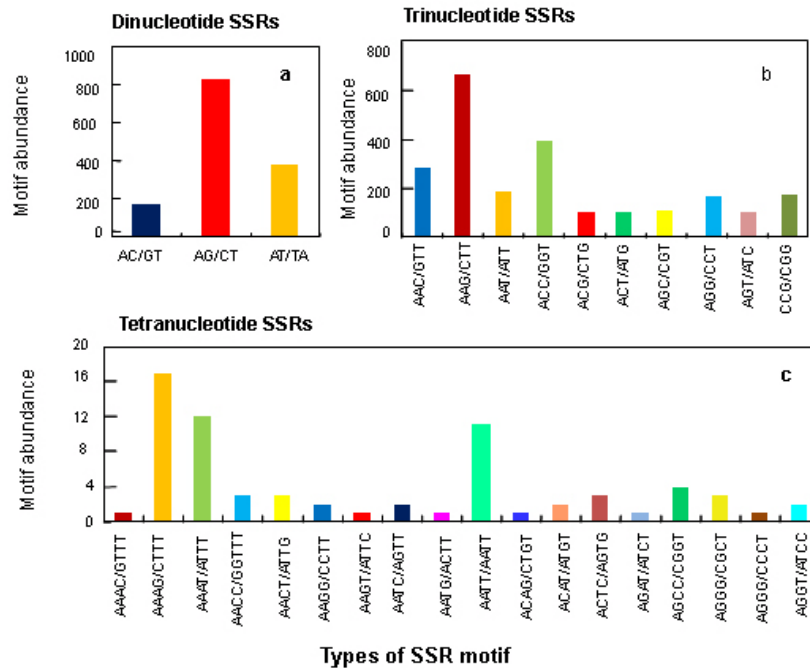


Fig. 1 Motif distribution of di, tri and tetranucleotide SSRs in ESTs from globular stage embryos: (a) Dinucleotide SSRs. (b) Trinucleotide SSRs. (c) Tetranucleotide SSRs.

Microsatellites, or simple sequence repeats (SSRs), have greater potential in polymorphism analysis and marker assisted selection than other types of molecular markers (Thiel et al. 2003; Jiang et al. 2006). However, the conventional development of SSR markers by SSR enriched genomic library construction and large-scale sequencing is costly and labor intensive (Varshney et al. 2002). The occurrence of SSRs in ESTs provides an opportunity to use currently available ESTs in databases for SSR development (Qiu et al. 2000). This strategy has been used successfully in plant species such as grape (*Vitis vinifera* L.) (Scott et al. 2000), wheat (*Triticum aestivum* L.) (Hackauf and Wehling, 2002; Peng and Lapitan, 2005), barley (*Hordeum vulgare* L.) (Thiel et al. 2003), strawberry (*Fragaria x ananassa*) (Folta et al. 2005), orange (*Citrus sinensis osbeck*) (Jiang et al. 2006), foxtail millet (*Setaria italica*) (Jia et al. 2007), ryegrass (*Lolium perenne*) (Asp et al. 2007), blackberry (*Rubus* L.) (Lewers et al. 2008).

Table 1. Summary of EST-SSRs identified.

	Number	Types of motifs	Proportion %	Frequency %	Average distance between two SSRs (kb)
Mononucleotide	304	2	7.62	0.24	90.54
Dinucleotide	1374	3	34.44	1.11	20.03
Trinucleotide	2208	10	55.35	1.79	12.46
Tetrnucleotide	70	18	1.75	0.06	393.23
Pentanucleotide	13	7	0.33	0.012	2117.41
Hexnucleotide	20	15	0.5	0.01	1376.32
Total	3989	55	100	3.25	6.9

SSRs are widely distributed in the soybean genome, with an average interval of about 15 Kb between SSRs (Wang et al. 1994). By the end of 2008, about 330,436 soybean ESTs from diverse tissues and organs were present in GmGI (http://www.tigr.org/tigr-scripts/tgi/T_index.cgi), in which 37,308 SSRs were identified (Jayashree et al. 2006). Recently, Dr. Goldberg's laboratory at the University of California, Los Angeles, submitted more than 400,000 ESTs from libraries of soybean globular stage embryos to the database (<http://estdb.biology.ucla.edu/seed/sequence/>). This huge number of ESTs was generated by 454 high throughput sequencing technology and provides a new opportunity for discovering molecular makers from genes that are active during early embryogenesis. The aim of this study is to develop EST-SSRs from globular stage embryos and examine its polymorphism among one wild type and 20 soybean cultivars. The polymorphic SSR makers would be important for genetic diversity analysis and marker assisted selection for soybean.

MATERIALS AND METHODS

We obtained EST sequences from the Goldberg laboratory, which used high throughput 454 sequencing to survey the transcriptome of soybean globular stage embryos (<http://estdb.biology.ucla.edu/seed/sequence>). Primer sequences were trimmed by SeqClean software (<http://compbio.dfci.harvard.edu/tgi/software/>). The 403,511 cleaned sequences were clustered and assembled by Tgicl and Cap3. Due to the large number of ESTs, Tgicl was unable to process all these sequences together. Instead, we grouped these sequences into three large groups and used Tgicl to process one group at a time. The resulted clusters were processed by Cap3 with $p = 90$ and $y = 30$. The contigs and singletons generated by Cap3 were subjected for microsatellite discovery by PERL5 script (MISA). The criteria used in this study were as follows: no less than sixteen repeats in single nucleotide SSRs; six or more repeats in di-nucleotide SSRs; and five or more repeats in tri-, tetra-, penta- and hexanucleotide SSRs. Randomly, 33 PCR primer pairs, which covered the di, tri, tetra, penta and hexanucleotide SSRs, were designed by primer premier 5.0 for amplification and polymorphism analysis.

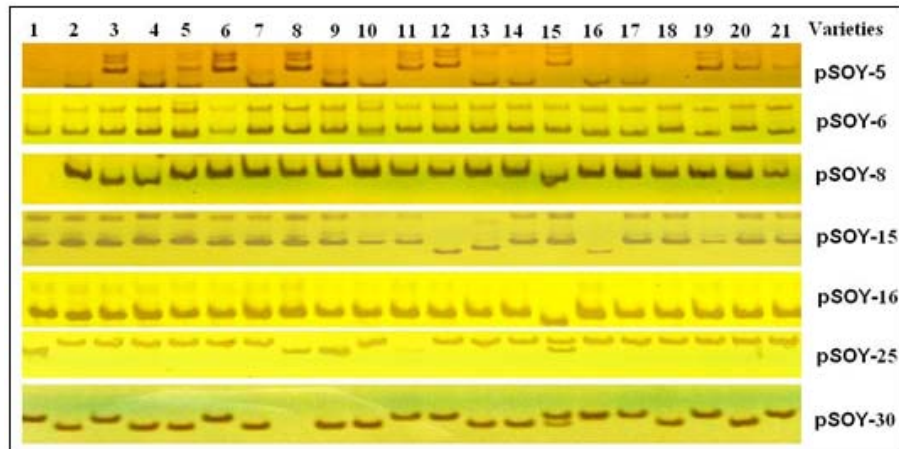


Fig. 2 Polymorphism analysis of EST-SSRs in twenty-one soybean varieties. The varieties are as follows: 1, Maoyandou; 2, Zhechuanjiwohuang; 3, Qiudou; 4, Miyunlaoyelian; 5, Xiataizimodadou; 6, Bendidahuangdou; 7, Yuandou; 8, Bailudou; 9, Qinyangxiaozihuan; 10, Tuerji; 11, Boaihongzaojiao; 12, Jlhuan52; 13, Chichenglvhuangdou; 14, Cansidou; 15, ZYD3263; 16, Luning1hao; 17, Zhongpin661; 18, Zhonghuang27; 19, Hedou13; 20, Fendou37 and 21, Youbian30.

Twenty Chinese cultivars and one wild-type soybean (ZYD3263) were used for polymorphism analysis. These cultivars included four varieties from Beijing (Miyunlaoyelian, Zhongpin661, Zhonghuang27, and Youbian30), seven varieties from Hebei (Maoyandou, Xiataizimodaodou, Bendidahuangdou, Yuandou, Tuerdou, and Jiguan52), three varieties from Henan (Zhechuanjiwohuang, Qinyangxiaozihuan, and Boaihongdouzaojiao), three varieties from Shandong (Qiudou, Luning1, and Hedou13) and three from Shanxi (Bailudou, Cansidou, and Fendou37). These soybeans were grown in pots containing soil mixture at 25°C for two weeks. Young leaves were collected for genomic DNA extraction using a Plant Genomic DNA Extraction Kit (Solarbio Inc., Beijing, China). PCR mixtures (20 µl volume) contained 1 x PCR buffer, 1.5 µM MgCl₂, 200 µM dNTPs, 0.2 µM of each primer, 0.1U Taq polymerase (TaKaRa Inc., Dalian, China) and 10 ng genomic DNA as the template. Amplifications were performed in a TaKaRa PCR thermal cycler Dice TP600 with the following conditions: 94°C for 4 min, 31 cycles of 94°C for 30 sec, 55°C for 30 sec and 72°C for 30 sec, and a final extension step of 72°C for 7 min. PCR products were separated on 6.5% w/v denaturing polyacrylamide gels (19:1 acrylamide/bisacrylamide, 7 M urea) and visualized by silver staining (Gao et al. 2009). The sizes of the amplified fragments were estimated based on Φ x 174-HaeIII digested DNA Markers (TaKaRa Inc., Dalian, China).

RESULTS AND DISCUSSION

Identification of SSRs in soybean

In total, 70,654 contigs and 52,082 singletons were generated from 403,511 cleaned ESTs. The average lengths of singletons and contigs were 177 bp and 268 bp, respectively. These non-redundant ESTs represented 28.1 Mb of cDNA. From these

122,736 non-redundant sequences 3,989 SSRs within 3,729 unique ESTs were identified. The SSR frequency in these soybean ESTs was approximately 3.3%, which was similar to frequencies obtained in several other plants (1.5% in maize, 4.7% in rice and 5.7% in lolium) (Kantety et al. 2002; Kumpatla and Mukhopadhyay 2005; Asp et al. 2007). However, there were fewer SSRs in these ESTs than that found in eucalyptus (29%) and citrus (21%) (Ceresini et al. 2005; Jiang et al. 2006). There were 235 (0.19%) unique ESTs that contained more than one SSR. This ratio was much lower than that in cattle (0.34%) (Yan et al. 2008) and citrus (3.4%) (Jiang et al. 2006). One important reason for this difference was that 454 sequencing-generated ESTs were shorter than ESTs generated by conventional methods. It is reasonable to assume that there is less chance of finding multiple SSRs in a short sequence than in a long one. Furthermore, the overall abundance of SSRs can vary depending on the criteria used to identify SSRs.

Diversity of SSRs in soybean

Investigation of the distribution of SSR motifs can provide insight into the composition of the genome (Yan et al. 2008). Our results showed that the motif types and abundance of soybean SSRs were not evenly distributed. Trinucleotide SSRs were the most abundant, accounting for 55% of the total SSRs identified in this study. The 1,374 (34.44%) dinucleotide SSRs represented the second most abundant type of motif. The mononucleotide, tetranucleotide, pentanucleotide and hexanucleotide SSRs were shown to be minor types of soybean SSRs. The length of these SSRs varied from 12 to 114 bp depending on number of repeats of the motif in each SSR (Table 1).

In total, 55 types of repeat motif were identified from these soybean ESTs. Among the mononucleotide repeats, the A/T motif (291) was much more prevalent than the C/G motif (13). AG/CT was the most abundant repeat, accounting for 60% of dinucleotide repeats, which agreed with reports in other crops (Cardle et al. 2000; Kantety et al. 2002; Asp et al. 2007). AT/TA and AC/GT repeat motifs were the second and third most abundant dinucleotide motifs, accounting for 27% and 13%, respectively (Figure 1A). AAG/CTT represented the most abundant (30%) trinucleotide motifs, followed by ACC/GGT (17%). The majority of these motifs were repeated 5-10 times in the ESTs examined (Figure 1B), which agreed with a report from citrus (Jiang et al. 2006). However, in foxtail millet (Jia et al. 2007) and eucalyptus (Ceresini et al. 2005), the most commonly identified trinucleotide motifs were CAG/TCT and CCG/CGG, respectively. Among the tetranucleotide motifs, AAAG/CTTT was the most abundant motif and AAAT/ATTT ranked as the second, which was in accordance with results from citrus (Figure 1C) (Jiang et al. 2006). In the case of pentanucleotide motifs, AAAAT/ATTTT and AACAT/ATTGT represented the most abundant identified SSRs. We found 15 different types of hexanucleotide repeat motifs, among which the A/T content (55.2%) was higher than C/G content (44.8%). The hexanucleotide repeat motifs occurred with roughly equal abundance.

Table 2. SSR primers designed from soybean SSR-containing EST sequences

Primer No.	Product size (bp)	Repeat unit	Forward primers(5'-3') Reverse primers(5'-3')	Chromatosome location*	Annotation*	Polymorphism
pSO Y5	200	(GA)15	F:AGCGTTATTGATTCGTT C R:TTTTGGTATCTCATTCT CACT	Gm15: 50201941- 50201714	Unknown (Glyma15g4 2780.1)	Yes
pSO Y6	190	(GA)9	F: GATGTGAGTGATAATGGT G R: ATCTTCTGCCTACTAAA	Gm06: 44897310- 44897074	No significant	Yes
pSO Y7	164	(AAC)8	F: GCTCTGCGAGATTCACTA R: TCTTATCAGGCTCTGGTT T	Gm08: 4635317- 4635110	Transcriptio n factor	No
pSO Y8	189	(CTT)10	F:TCACTCGCTCACTCCCG TTAA R: AGAGGTGCGACCCGATG A	Gm11: 16313110- 16313143	UDP- apiose/xylos e synthase	Yes
pSO Y11	202	(CAT)9	F:GAGATTTGGGTTGAGGT T R: GCGAGTTACGAGGCTATT	Gm18: 61901016- 61901449	Unknown (Glyma18g5 3630.1)	No
pSO Y12	186	(TGG)8	F:GCTTGGAGATGAGGTTA T R: CACTACTTCTCCCTAACT	Gm13: 5624222- 5623969	Unknown (Glyma13g0 5330.1)	No
pSO Y13	163	(TCT)7	F:GGGATGGTATTGGGTAT G R: GAGTGGATGGAGGGAGT G	Gm17: 4533609- 4533826	Unknown (Glyma17g0 6370.1)	Yes
pSO Y15	167	(AGG)5	F:TAACTCAGATTGGCATT C R: CTAAAGGTCAAATAGTG GT	Gm16: 418054- 418112	Unknown (Glyma16g0 0780.1)	Yes
pSO Y16	121	(GTG)8	F:GGGTGAGGGAATGAAG AA R: GAACCTCGCCACAACCTGC	Gm11: 15038048- 15038193	Unknown (Glyma11g1 8360.1)	Yes
pSO Y17	202	(TGG)10	F:GCCACCTAGAAGAACAA G R: TTCCATAATCCTCCAACA	Gm10: 50433819- 50433850	Unknown (Glyma10g4 3890.1)	No
pSO Y18	130	(TCT)12	F:GTAGTATCGGCGATCCA T R: GAAAGTCGAAACGAAAAC	Gm02: 3214600- 3214638	Unknown (Glyma02g0 4000.1)	Yes
pSO Y19	158	(AGG)8	F:CCTCGGCACCTTCAAAC G R: AACTCACCGCACCCATCA	Gm17: 258282- 258044	Transcriptio n factor	No
pSO Y20	137	(GAAA)6	F:TCTGTGACCTCCAATCC A R: AATCGAGGGCGAGAAAAG C	Gm10: 46938732- 46938612	Cleavage factor	No
pSO Y21	210	(TTTC)5	F:ATAATACGTGCATACAT AGC R: TCTTGTGGTGAGAACCC	Gm07: 17896053- 17895810	Ca2Yes/HY es exchanger	No
pSO Y22	166	(TTAA)6	F:AATAACTTCACATCTAA CCCAR:T	Gm08: 3210789-	Unknown (Glyma08g0)	No

pSO Y 23	197	(TTAA) ₆	TTGACTATGCGGAACCC F:CTTATGTTATCTGTGGG ACG R: GTGACGGTGGTGGTTATG	3211056 Gm05: 39282498- 39282043	4540.1) Unknown (Glyma05g3 5170.1)	No
pSO Y 25	150	(GCCCA) ₆	F:GAGGCGAAGAAGAGCA GG R: GGAAATCCAAAGAGGGAA A	Gm03:37602 730- 37603206	Phosphate acyltransferase	Yes
pSO Y 26	140	(CACAA) ₅	F:GCTTCAATTCTGCCTTT A R: CTCTGCCAGTTCATTTT	Gm12: 32242085- 32241831	Unknown (Glyma12g2 8870.1)	No
pSO Y 29	172	(CAACA G) ₅	F:GGGATTCTTGC GG TCTT C R: ACTCGGTCCAGTTTGTCTG	Gm07: 4711040- 4710731	Unknown (Glyma07g0 6010.1)	Yes
pSO Y 30	117	(GTTAG G) ₅	F:CCCTCTTCCACGCTCTT G R: CCGCCATTGTTGTTTCTC	Gm13: 27016897- 27017149	Plastidic phosphate translocator	Yes
pSO Y 31	263	(GAGCA T) ₅	F:AAGTCATTGTTTCTTC GTT R: GCCAAGCATGTAGTTTCC	Gm07: 1150407- 1150774	Acid phosphatase	Yes

Chromosome location* and Annotation * based on The Soybean Genome Database (<http://www.phytozome.net/soybean>)

PCR amplification and polymorphism analysis

Twenty-one out of 33 designed primers gave successful amplification within soybean cultivars tested. The details of primer sequences, amplified fragment size, annotation (if available) of the corresponding ESTs, SSR sequences and polymorphisms among the tested soybeans are listed in Table 2. Importantly, 11 pairs of primers were able to detect polymorphisms in the tested soybean varieties (Figure 2). Interestingly, four ESTs corresponding to four of these SSRs, shared significant similarity to known proteins, namely UDP-apiose/xylose synthase, phosphate acyltransferase, plastidic phosphate translocator and acid phosphatase. It would be interesting to correlate these SSRs with a particular phenotype controlled by these genes. In this study, the functions of 3,097 (83%) sequences were not known. Only 632 (17%) of SSR-containing sequences were annotated. The function of these SSR-containing ESTs covered a wide range of biological processes, such as protein synthesis, signal transduction, cell structure, transport, metabolism, and disease and defense responses (Figure 3).

We demonstrated that 454 high throughput sequencing is a powerful approach for discovering SSRs in plants. 454 sequencing could generate more than 200 thousand reads each run in a short time and with an affordable cost. Sequences generated by this technology have an average length of 200 bp, which is enough for SNP and SSR discovery. This study led to the discovery of 3,989 EST-SSRs from 403,511 soybean ESTs obtained by 454 sequencing. This number of EST-SSR accounts for 10% of total soybean EST-SSR found from the ESTs in the database so far. Of the 3,792 SSR containing uni-ESTs, 632 (17%) represented protein encoding genes, which making these EST-SSRs valuable in marker assisted selection. While 3,097 (83%) have no homolog sequences found in GenBank. One reason of the poor annotation is due to the short length of EST sequence (the average lengths of singletons and contigs were 177 bp and 268 bp). Another possible reason is that the ESTs used in

this study derived from globular stage embryos and sequences available in the current database are mostly from mature tissues. However, ESTs contain less SSRs than that of the genomic sequences. Efficiency could be significantly improved, and more SSR markers could be identified, by applying 454 sequencing technology to SSR enriched genomic libraries.

Due to the conserved genomic regions between different genome, SSR marker could be transferable among congeneric species (Boutin et al. 1995; Isobe et al. 2003; Kuleung et al. 2004). High transferability of EST-SSRs (34%) and genomic SSR (25%) between soybean and cultivated peanut were reported (He et al. 2006; Hong et al. 2010). As an important oil crop, marker development of peanut was far behind soybean and many other crops. These soybean EST-SSRs provided a useful molecular marker source for peanut.

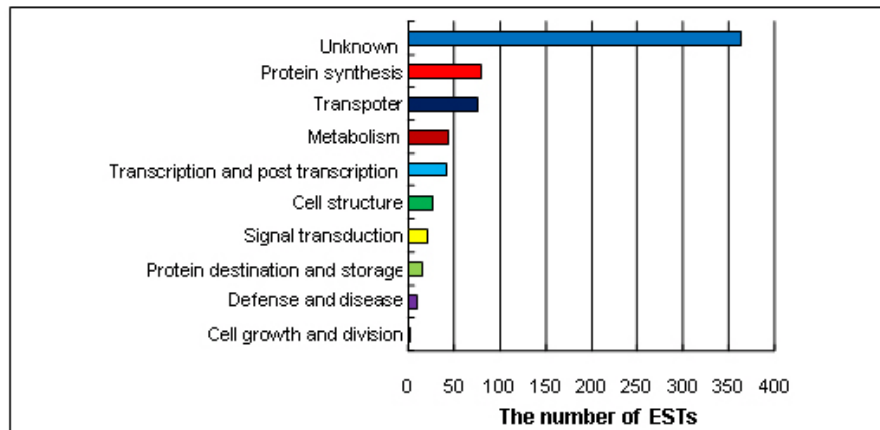


Fig. 3 Functional categories of SSR-containing ESTs.

In this study we randomly selected 33 SSR containing ESTs and designed primers to amplify the SSRs. Twenty-one achieved successful amplification. The twelve instances where amplification was not achieved were mainly due to poor primer quality. The length of these ESTs ranged from 100 bp to 250 bp. If the SSR were located close to one end of the sequence, it would be very difficult to use the flanking region for primer designing. Integrating these soybean 454 sequences with soybean sequences from other sources available in the database could partially resolve this problem. Out of the 21 amplified SSRs, 11 showed polymorphism among the tested soybean cultivars and wild-type soybean. Based on this ratio we could expect to identify more than 1,300 SSRs with polymorphisms in different soybean cultivars. The polymorphism screening experiment was undertaken using all EST-SSRs identified in this study. The results of this study will greatly facilitate the study of soybean genetic diversity analysis and marker assisted selection.

ACKNOWLEDGMENTS

We would like to thank Professor Robert B. Goldberg of University of California, Los Angeles, for allowing us to use the soybean EST sequences prior to their submission to the public database.

Financial support: This work was supported by Shandong Provincial Natural Sciences Foundation, China (Y2008D44), Shandong province "Taishan Scholar" foundation, Doctoral Foundation of Shandong Province (2006BS06008) and grants from Shandong Academy of Agricultural Sciences (2007YC001, 2006YBS001).

REFERENCES

- ASP, Torben; FREI, Ursula K.; DIDION, Thomas; NIELSEN, Klaus K. and LÜBBERSTEDT, Thomas. Frequency, type, and distribution of EST-SSRs from three genotypes of *Lolium perenne*, and their conservation across orthologous sequences of *Festuca arundinacea*, *Brachypodium distachyon*, and *Oryza sativa*. *BMC Plant Biology*, July 2007, vol. 7, no. 36. [\[CrossRef\]](#)
- BOUTIN, S.R.; YOUNG, N.D.; OLSON, T.C.; YU, Z.H.; VALLEJOS, C.E. and SHOEMAKER, R.C. Genome conservation among three legume genera detected with DNA markers. *Genome*, October 1995, vol. 38, no. 5, p. 928-937. [\[CrossRef\]](#)
- CARDLE, Linda; RAMSAY, Luke; MILBOURNE, Dan; MACAULAY, Malcolm; MARSHAL, David and WAUGH, Robbie. Computational and experimental characterization of physically clustered simple sequence repeats in plants. *Genetics*, October 2000, vol. 156, no. 2, p. 847-854.
- CERESINI, Paulo Cesar; PETRAROLHA SILVA, Cristina Lacerda Soraes; MISSIO, Robson Fernando; SOUZA, Elaine Costa; FISCHER, Carlos Norberto; GUILLHERME, Ivan Rizzo; GREGORIO, Ivo; DA SILVA, Eloiza Elena Tajara; BARRETO CICARELLI, Regina Maria; ALVES DA SILVA, Marco Tulio; GARCIA, Jose Fernando; AVELAR, Gustavo Arbex; PORTO NETO, Laercio Ribeiro; MARÇON, Andre Ricardo; BACCI JUNIOR, Maurício and MARINI, Danyelle Cristine. Satellyptus: Analysis and database of microsatellites from ESTs of Eucalyptus. *Genetics and Molecular Biology*, 2005, vol. 28, no. 3, p. 589-600. [\[CrossRef\]](#)
- FOLTA, Kevin M.; STATON, Margaret; STEWART, Philip J.; SOOK, Jung; DAWN, H. Bies; JESDURAI, Christopher and MAIN, Dorrie. Expressed sequence tags (ESTs) and simple sequence repeat (SSR) markers from octoploid strawberry (*Fragaria xananassa*). *BMC Plant Biology*, June 2005, vol. 5, no. 1. [\[CrossRef\]](#)
- GAO, Dong; DU, Fei and ZHU, You-Yong. Low-background and high-resolution contracted silver-stained method in polyacrylamide gels electrophoresis. *Hereditas (Beijing)*, July 2009, vol. 31, no. 6, p. 668-673. [\[CrossRef\]](#)
- HACKAUF, B. and WEHLING, P. Identification of microsatellite polymorphisms in an expressed portion of the rye genome. *Plant Breeding*, February 2002, vol. 121, no. 1, p. 17-25. [\[CrossRef\]](#)
- HE, G.; WOULLARD, F.E.; MARONG, I. and GUO, B.Z. Transferability of soybean SSR markers in peanut (*Arachis hypogaea* L.). *Peanut Science*, January 2006, vol. 33, no. 3, p. 22-28. [\[CrossRef\]](#)
- HONG, Yan-Bin; CHEN, Xiao-Ping; LIU, Hai-Yan; ZHOU, Gui-Yuan; LI, Shiao-Xiong; WEN, Shi-Jie and LIANG, Xuan-Qiang. Development and utilization of orthologous SSR markers in *Arachis* through Soybean (*Glycine max*) EST. *Acta Agronomica Sinica*, 2010, vol. 36, no. 3, p. 410-421. [\[CrossRef\]](#)
- ISOBE, S.; KLIMENKO, I.; IVASHUTA, S.; GAU, M. and KOZLOV, N.N. First RFLP linkage map of red clover (*Trifolium pratense* L.) based on cDNA probes and its transferability to other red clover germplasm. *TAG Theoretical and Applied Genetics*, December 2003, vol. 108, no. 1, p. 105-112. [\[CrossRef\]](#)
- JAYASHREE, B.; PUNNA, R.; PRASAD, P.; BANTTE, K.; HASH, C.T.; CHANDRA, S.; HOISINGTON, D.A. and VARSHNEY, R.K. A database of simple sequence repeats from cereal and legume expressed sequence tags mined in silico: survey and evaluation. *In Silico Biology*, 2006, vol. 6, no. 6, p. 607-620.

- JIA, Xiao-Ping; SHI, Yun-Su; SONG, Yan-Cun; WANG, Guo-Yin; WANG, Tian-Yu and YU, Li. Development of EST-SSR in foxtail millet (*Setaria italica*). *Genetic Resources and Crop Evolution*, March 2007, vol. 54, no. 2, p. 233-236. [\[CrossRef\]](#)
- JIANG, Dong; ZHONG, Guang-Yan and HONG, Qi-Bin. Analysis of microsatellites in citrus unigenes. *Acta Genetica Sinica*, April 2006, vol. 33, no. 4, p. 345-353. [\[CrossRef\]](#)
- KANTETY, Ramesh V.; LA ROTA, Mauricio; MATTHEWS, David E. and SORRELLS, Mark E. Data mining for simple sequence repeats in expressed sequence tags from barley, maize, rice, sorghum and wheat. *Plant Molecular Biology*, March 2002, vol. 48, no. 5-6, p. 501-510. [\[CrossRef\]](#)
- KULEUNG, C.; BAENZIGER, P.S. and DWEIKAT, I. Transferability of SSR markers among wheat, rye, and triticale. *TAG Theoretical and Applied Genetics*, 2004, vol. 108, no. 6, p. 1147-1150. [\[CrossRef\]](#)
- KUMPATLA, Siva P. and MUKHOPADHYAY, Snehasis. Mining and survey of simple sequence repeats in expressed sequence tags of dicotyledonous species. *Genome*, December 2005, vol. 48, no. 6, p. 985-998. [\[CrossRef\]](#)
- LEWERS, Kim S.; SASKI, Chris A.; CUTHBERTSON, Brandon J.; HENRY, David C.; STATON, Meg E.; MAIN, Dorrie S.; DHANARAJ, Anik L.; ROWLAND, Lisa J. and TOMKINS, Jeff P. A blackberry (*Rubus* L.) expressed sequence tag library for the development of simple sequence repeat markers. *BMC Plant Biology*, June 2008, vol. 8, no. 1, p. 69. [\[CrossRef\]](#)
- PENG, J.H. and LAPITAN, N.L. Characterization of EST-derived microsatellites in the wheat genome and development of eSSR markers. *Functional & Integrative Genomics*, April 2005, vol. 5, no. 2, p. 80-96. [\[CrossRef\]](#)
- QIU, Lijuan; CHANG, Ruzhen; SUN, Jangyin; LI, Xianghua; XU, Zhanyou and LIU, Lihong. Perspectives of evaluation and utilization of soybean Germplasm in China. *Review of China Agricultural Science and Technology*, 2000, no. 5, p. 58-61.
- SCOTT, K.D.; EGGLE, P.; SEATON, G.; ROSSETTO, M.; ABLETT, E.M.; LEE, L.S. and HENRY, R.J. Analysis of SSRs derived from grape ESTs. *TAG Theoretical and Applied Genetics*, March 2000, vol. 100, no. 5, p. 723-726. [\[CrossRef\]](#)
- THIEL, T.; MICHALEK, W.; VARSHNEY, R.K. and GRANER, A. Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.). *TAG Theoretical and Applied Genetics*, February 2003, vol. 106, no. 3, p. 411-422. [\[CrossRef\]](#)
- VARSHNEY, Rajeev K.; THIEL, Thomas; STEIN, Nils; LANGRIDGE, Peter and GRANER, Andreas. *In silico* analysis on frequency and distribution of microsatellites in ESTs of some cereal species. *Cellular & Molecular Biology Letters*, June 2002, vol. 7, no. 2A, p. 537-546.
- WANG, Z.; WEBER, J.L.; ZHONG, G. and TANKSLEY, S.D. Survey of plant short tandem DNA repeats. *TAG Theoretical and Applied Genetics*, April 1994, vol. 88, no. 1, p. 1-6. [\[CrossRef\]](#)
- XIA, Zhengjun; TSUBOKURA, Yasutaka; HOSHI, Masako; HANAWA, Masayoshi; YANO, Chizuru; OKAMURA, Kayo; TALAAT, A. Ahmed; ANAI, Toyooki; WATANABE, Satoshi; HAYASHI, Masaki; KAWAI, Takashi; HOSSAIN, Khwaja G.; MASAKI, Hirokazu; ASAI, Kazumi; YAMANAKA, Naoki; KUBO, Nakao; KADOWAKI, Koh-ichi; NAGAMURA, Yoshiaki; YANO, Mashairo; SASAKI, Takuji and HARADA, Kyuya. An integrated high-density linkage map of soybean with RFLP, SSR, STS, and AFLP markers using a single F₂ population. *DNA Research*, 2007, vol. 14, no. 6, p. 257-269. [\[CrossRef\]](#)
- YAN, Qiulang; ZHANG, Yinghan; LI, Hongbin; WEI, Caihong; NIU, Lili; GUAN, Shan; LI, Shangang and DU, Lixin. Identification of microsatellites in cattle unigenes. *Journal of Genetic and Genomics*, May 2008, vol. 35, no. 5, p. 261-266. [\[CrossRef\]](#)

How to cite this article:

LI, A.Q.; ZHAO, C.Z.; WANG, X.J.; LIU, Z.J.; ZHANG, L.F.; SONG, G.Q.; YIN, J.; LI, C.S.; XIA, H. and BI, Y.P. Identification of SSR markers using soybean (*Glycine max*) ESTs from globular stage embryos. *Electronic Journal of Biotechnology*, September 2010, vol. 13, no. 5. <http://dx.doi.org/10.2225/vol13-issue5-fulltext-5>