

Novel polymorphic EST-based microsatellite marker isolation and characterization from *Poncirus trifoliata* (Rutaceae)

Manosh Kumar BISWAS¹, Peng CHEN¹, Mohamed Hamdy AMAR^{1,2}, Xiuxin DENG (✉)¹

¹ Key Laboratory of Horticultural Plant Biology (MOE), National Center of Citrus Breeding, Huazhong Agricultural University, Wuhan 430070, China

² Desert Research Center, Egyptian Deserts Gene Bank, North Sinai, Egypt

Abstract Novel *Poncirus trifoliata* simple sequence repeat (SSR) markers were developed to evaluate their utility for genetic diversity and breeding studies of *P. trifoliata* and related species. A total of 108 primer pairs were characterized by PCR amplification experiments. Among these, 61 were polymorphic and transferable to other citrus species. The number of alleles per locus ranged from 2 to 6, with an average of 2.37 alleles per locus. The expected heterozygosity and observed heterozygosity ranged from 0 to 0.83 and 0 to 1.00, respectively. These novel polymorphic SSR markers will be useful for citrus cultivar identification and evaluation as well as breeding studies.

Keywords microsatellite, polymorphism, *Poncirus trifoliata*

graft compatibility with other citrus species make it a perfect rootstock for commercial propagation of cultivated citrus species.

SSR markers have been widely used in genetic diversity, population genetic studies, linkage map construction as well as several breeding applications due to the superior characteristics of SSR markers over other molecular markers^[2]. Previous studies suggested that Expressed Sequence Tag (EST) databases are valuable sources from which to develop SSR markers^[3–7]. The development of SSR markers specific to *P. trifoliata* is important and critical for improving its varieties and assisting breeding programs. Till to date, no EST-SSR markers have been reported for *P. trifoliata*. Hence, the aim of this study was to develop *P. trifoliata* specific EST-SSR markers, investigate their cross-taxa transferability and utility for germplasm characterization.

1 Introduction

Poncirus trifoliata (L.) Raf. (Trifoliate orange) is a member of the family Rutaceae and closely related to the genus *Citrus*. It is also known as Chinese bitter orange and its center of origin is northern China. Sometimes it is included as a species in the genus *Citrus*. Swingle^[1] recorded and compared the characters of *Poncirus* and *Citrus* and demonstrated that some pronounced characters, namely, being a deciduous tree with compound leaves, pubescent fruit, shrub-like tree structure and very thorny branches, make it difficult to place *Poncirus* as an ordinary member in the genus *Citrus*. The resistance of *Poncirus* against the *Phytophthora* root rot, the citrus tristeza virus (CTV) and the citrus nematode (CN), its cold tolerance and

2 Materials and methods

2.1 Plant materials

A total of 42 accessions (Appendix A, Table S1) of citrus and its relatives were obtained from the Citrus Germplasm and Conservation Center of Huazhong Agricultural University, China. DNA of those accessions was extracted from the young leaves using the CTAB (Hexadecyltrimethylammonium bromide) methods described by Cheng et al.^[8].

2.2 EST processing and marker development

EST sequences of *P. trifoliata* were downloaded from the NCBI (National Center for Biotechnology Information; <http://www.ncbi.nlm.nih.gov>) on the 18th April, 2012. A perl script `est_trimmer.pl` was used to remove unusual EST sequences, vector contamination, poly A and poly T bases

Received February 11, 2015; accepted March 18, 2015

Correspondence: xxdeng@mail.hzau.edu.cn

from the downloaded *P. trifoliata* EST sequences. After that, high quality EST sequences were assembled, redundancy of the ESTs were removed using Cap3 (<http://mobyle.pasteur.fr/cgi-bin/portal.py#forms::cap3>) software. Assembled ESTs were screened for SSR marker development by using SSRLocator software^[9] using a minimum SSR repeat length of 16 bp, an expected product size of 100–300 bp, an annealing temperature of 50–55°C and a GC content of 45%–50%. To identify putative functions of the developed *P. trifoliata* EST-SSR primers, flanking regions of the primer pairs were blasted against the NCBI protein database using Blast2Go tool (www.blast2go.com). Virtual PCR strategy was applied to select *in silico* polymorphic SSR markers. Selected SSR primers were synthesized by the Sangon Company, Shanghai, China.

2.3 PCR amplification and PAGE analysis

PCR amplifications were performed in a reaction volume of 15 µL, which consisted of 1U Taq polymerase, 1× reaction buffer (750 mmol·L⁻¹ Tris HCl (pH 9), 50 mmol·L⁻¹ KCl, 200 mmol·L⁻¹ (NH₄)₂SO₄, 0.001% BSA), 0.1 mmol·L⁻¹ of each dNTP, 1.5 mmol·L⁻¹ MgCl₂, 3 µmol·L⁻¹ of each primer, and 25 ng DNA sample. The PCR reactions were conducted in a MJ-PTC-200 thermal controller (MJ Research, Waltham MA) under the following conditions: initial denaturing at 94°C for 3 min followed by 35 cycles at 94°C for 30 s, 50–55°C for 30 s, and 72°C for 45 s, with a final extension of 7 min at 72°C. The annealing temperatures ranged from 50 to 55°C for different primer pairs (Appendix A, Table S2). PCR product was separated by 6% denaturant PAGE (Poly-Acrylamide Gel Electrophoresis) (60 cm × 30 cm × 0.4 cm), and bands visualized by silver-staining according to the protocol of Ruiz et al.^[10].

2.4 Data analysis

Allele numbers, expected heterozygosity, observed heterozygosity and polymorphic information content (PIC) values were estimated with PowerMarker V3.25.

3 Results

In this study, 25388 non-redundant *P. trifoliata* EST sequences (obtained after clustering ESTs from NCBI) were screened for microsatellites using the SSRLocator software^[9]. The search was restricted to repeats with a minimum number of 8, 7, 5, 4 and 4 complete repeat units for di-, tri-, tetra-, penta- and hexa-nucleotide motifs, respectively. In total, 976 SSRs were identified from 887 ESTs (Appendix A, Table S3). Thus 3.49% of the *P. trifoliata* EST contained one or more microsatellites with one SSR per 4 kb EST (976 SSR in 22.7 Mb). Of the 887

SSR containing EST sequences, 695 (78%) ESTs are suitable for *P. trifoliata* EST SSR (PteSR) marker development. These sequences were further compared by blast against the NCBI non-redundant protein database with an *E*-value cut-off of 1.e⁻³, in order to identify functional EST sequences. As a result, 699 sequences containing di (333), tri (170), tetra (47), penta (60) and hexa (59) microsatellites units were identified as functional *P. trifoliata* specific EST SSR motifs (Fig. 1a). Microsatellites with higher repeat numbers and having a significant similarity to existing proteins in the NCBI database were favored because they are generally considered to be more polymorphic and functional. With these criteria, 542 primer pairs were selected for the *in silico* PCR (VPCR) amplification analysis, which was performed using SSRLocator. Whole genomic sequences of *Citrus sinensis* and *C. clementina* were used as templates for the virtual PCR amplification analysis and results reveal that 226 (41%) primer pairs had positive VPCR amplification. To determine the genomic distribution of developed PteSR marker, we mapped 542 marker on nine chromosome of *C. sinensis* and results revealed that 198 (36%) markers distributed among nine chromosomes of *C. sinensis* with an average marker density 0.83 Mb (Fig. 1b, Fig. 2). Chromosome mapping of the markers showed highest frequency of markers on chromosome 5 (38 markers, 7%) and lowest on chromosome 1 (8, 1.4%) (Fig. 1b). *In silico* cross genera transferability was performed using an *in silico* PCR strategy to select a sub-set of PteSR marker for subsequent estimation of their utility for genetic diversity of citrus and relative species. The results showed the highest transferability (42%) in *C. sinensis* and the lowest (1%) in *Arabidopsis* with an average of 12% of PteSR markers transferable in citrus and non citrus species (Fig. 1c). Based on the *in silico* cross-genera transferability results the best 108 primer pairs were selected for subsequent PCR amplification and analysis for transferability to citrus species and for determining their polymorphism in eight *Citrus* spp. (*C. sinensis*, *C. reticulata*, *C. grandis*, *C. aurantifolia*, *C. limon*, *C. medica*, *C. ichangensis*, *C. paradisi*, *P. trifoliata* and *Fortunella hindsii*). Among all 108 primer pairs tested, 18 (16%) failed to amplify the expected products. The remaining 90 microsatellite loci produced the expected products, of which 61 (56%) were found to be polymorphic and transferable to citrus related species.

Furthermore, a subset of 12 primer pairs was used for genotyping 42 accessions from six species (*P. trifoliata*, *Fortunella* sp. *C. grandis*, *C. sinensis*, *C. reticulata*, *C. limon*). Our results reveal that the number of alleles per locus varied from 1 to 6 with an average of 2.38 (Table 1). Values of Ho, He and PIC ranged from 0 to 1.00, from 0 to 0.83 and 0 to 0.81, respectively. The highest number of alleles per locus (2.75) and diversity (He = 0.453) was recorded for *C. reticulata*, followed by *P. trifoliata* (number of allele = 2.667, He = 0.501). The single taxon

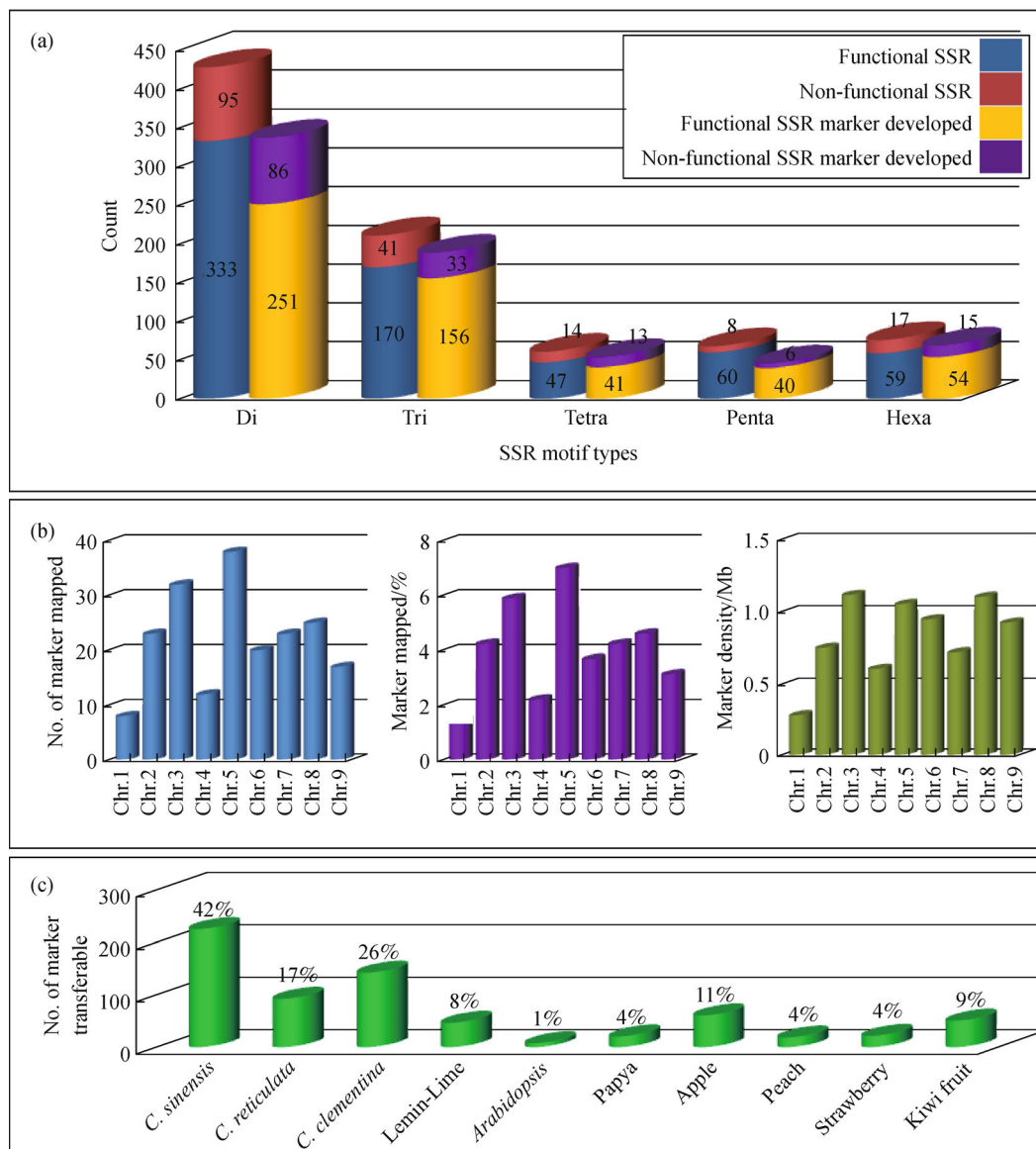


Fig. 1 (a) Distribution of functional SSR, non-functional SSR, functional SSR marker and non-functional SSR markers among different classes of repeat motifs; (b) chromosomal distribution of marker, percent of markers and density; (c) percent transferability of *P. trifoliata* EST-SSR marker in different citrus and non citrus species.

from *Fortunella* appeared to be the least polymorphic for the markers investigated with only 1.66 alleles per locus, a lower expected heterozygosity and lower level of the PIC value.

4 Discussion

A large amount of EST data in the public domain provides a valuable resource for genetic research, including molecular marker identification, and SSRs have been discovered from EST data for many plant species^[7,11–16]. The SSR frequency in the *P. trifoliata* ESTs was 1/4.0 kb and a similar SSR frequency has been reported in poplar

(1/4.0 kb), citrus (1/5.2 kb), *Arabidopsis* (1/6.0 kb) and sweet potato (1/7.1 kb)^[17,18], although much higher and lower SSR frequencies have also been reported in plants, viz., rice (1/40.0 kb) and cucumber (1/1.8 kb)^[19,20]. Direct comparison of the frequency of occurrence of SSRs in different reports is difficult given that estimates are dependent on the SSR mining tools, search criteria, size of the data set and the redundancy of EST sequences. Dinucleotide repeats were the most frequent motif types in the transcribed region of the *P. trifoliata* genome, as has often been observed in other plant species^[21]. In addition, the frequency of motif types varies greatly according to the SSR search criteria and search tools. The percentage of polymorphic *P. trifoliata* markers was slightly less than

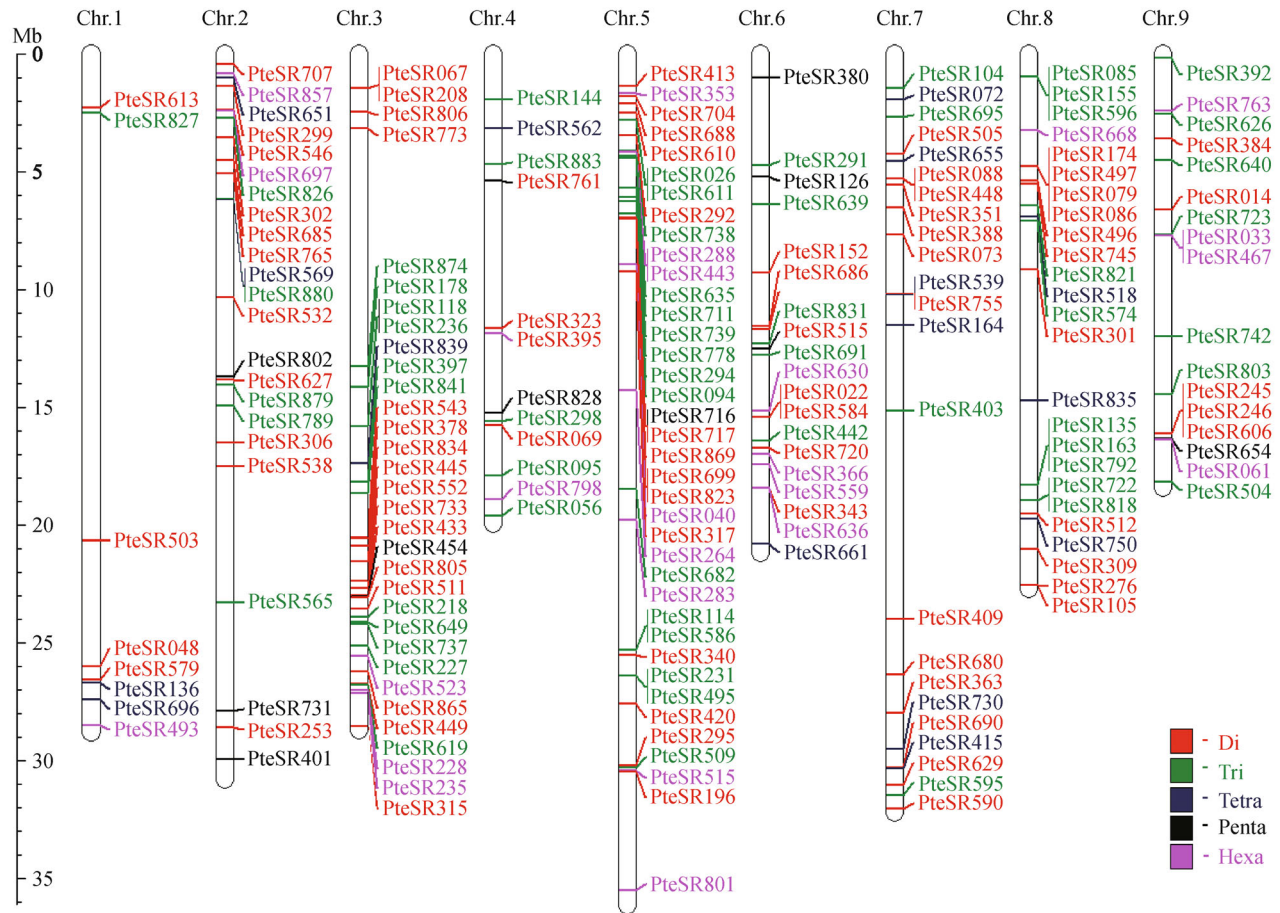


Fig. 2 Physical map of *P. trifoliata* EST-SSR marker based on 198 functionally relevant primers

Table 1 Results of the initial marker screening in citrus and related species

Locus ID	<i>P. trifoliata</i> (8)				<i>Fortunella</i> (8)				<i>C. grandis</i> (8)			
	AN	He	Ho	PIC	AN	He	Ho	PIC	AN	He	Ho	PIC
PteSR046	2	0.48	0.40	0.36	1	0.00	0.00	0.00	1	0.00	0.00	0.00
PteSR666	4	0.64	0.80	0.58	2	0.50	1.00	0.38	4	0.69	1.00	0.63
PteSR168	2	0.50	1.00	0.38	2	0.38	0.50	0.30	2	0.38	0.00	0.30
PteSR581	2	0.50	1.00	0.38	2	0.50	1.00	0.38	2	0.22	0.25	0.19
PteSR484	2	0.48	0.40	0.36	2	0.50	1.00	0.38	1	0.00	0.00	0.00
PteSR661	2	0.50	0.60	0.38	1	0.00	0.00	0.00	2	0.38	0.50	0.30
PteSR563	3	0.46	0.60	0.41	2	0.50	1.00	0.38	2	0.22	0.25	0.19
PteSR755	2	0.28	0.33	0.24	2	0.50	1.00	0.38	3	0.66	0.75	0.58
PteSR299	2	0.48	0.40	0.36	1	0.00	0.00	0.00	3	0.66	0.75	0.58
PteSR550	3	0.63	0.50	0.55	1	0.00	0.00	0.00	2	0.50	1.00	0.38
PteSR493	4	0.58	0.60	0.54	2	0.50	1.00	0.38	2	0.50	1.00	0.38
PteSR366	4	0.48	0.60	0.45	2	0.50	1.00	0.38	2	0.38	0.00	0.30
Mean	2.667	0.501	0.603	0.416	1.667	0.323	0.625	0.247	2.167	0.383	0.458	0.319

(Continued)

Locus ID	<i>C. sinensis</i> (5)				<i>C. reticulata</i> (8)				Lemone-Lime (5)			
	AN	He	Ho	PIC	AN	He	Ho	PIC	AN	He	Ho	PIC
PteSR046	2	0.18	0.20	0.16	3	0.56	0.00	0.50	1	0.00	0.00	0.00
PteSR666	2	0.50	1.00	0.38	4	0.74	1.00	0.69	6	0.83	1.00	0.81
PteSR168	3	0.58	0.20	0.49	4	0.58	0.60	0.54	3	0.67	0.67	0.59
PteSR581	2	0.42	0.60	0.33	2	0.50	0.20	0.38	2	0.50	1.00	0.38
PteSR484	3	0.56	0.40	0.50	1	0.00	0.00	0.00	1	0.00	0.00	0.00
PteSR661	2	0.42	0.60	0.33	2	0.42	0.60	0.33	2	0.50	1.00	0.38
PteSR563	2	0.50	0.20	0.38	2	0.18	0.20	0.16	3	0.61	1.00	0.54
PteSR755	1	0.00	0.00	0.00	4	0.58	0.80	0.54	4	0.75	1.00	0.70
PteSR299	3	0.56	0.00	0.50	2	0.32	0.00	0.27	1	0.00	0.00	0.00
PteSR550	4	0.70	0.80	0.65	4	0.72	1.00	0.67	3	0.61	0.33	0.54
PteSR493	2	0.50	1.00	0.38	3	0.66	0.80	0.59	3	0.61	1.00	0.54
PteSR366	3	0.34	0.20	0.31	2	0.18	0.20	0.16	2	0.50	1.00	0.38
Mean	2.417	0.438	0.433	0.368	2.750	0.453	0.450	0.403	2.583	0.465	0.667	0.405

Note: For each primer pair the number of alleles (AN), expected heterozygosity (He), observed heterozygosity (Ho) and the polymorphic information content (PIC) are given. The sample size for each population is shown in parentheses.

observed in a previous study with *C. grandis* EST SSR markers, in which 57% of the primers were polymorphic^[5]. Transferability to the related genera of SSR markers make them useful for genetic research, such as genotyping, linkage map construction, association mapping. *P. trifoliata* EST-SSR markers showed transferability to the *Fortunella* and *Citrus* genera. The rate of SSR transferability in different species is related to the genetic distance between the species from which the SSRs were developed and the other species^[22]. We also found a similar trend for *P. trifoliata* EST-SSR markers transferability. As expected 36% anchored in the *C. sinensis* genome, which might be due to genomic similarity of these two species. This suggests that *P. trifoliata* EST-SSR markers might be useful for *C. sinensis* breeding research, such as in mapping, genotyping and population structure estimation.

5 Conclusions

In this study we developed, characterized and utilized the first set of functionally relevant SSR markers for *P. trifoliata*. These novel highly transferable polymorphic microsatellite markers will be useful for germplasm diversity, genetic mapping, and population structure analysis of *Citrus* and its relatives, and could help reveal the inter-specific relationships, origin and evolution of cultivated *Citrus* species.

Acknowledgements This work was supported by the Ministry of Science and Technology of China (2011CB100601) and the National Natural Science Foundation of China (31330066).

Supplementary materials The online version of this article at <http://dx.doi.org/10.15302/J-FASE-2015048> contains supplementary material (Appendix A).

doi.org/10.15302/J-FASE-2015048 contains supplementary material (Appendix A).

Compliance with ethics guidelines Manosh Kumar Biswas, Peng Chen, Mohamed Hamdy Amar and Xiuxin Deng declare that they have no conflict of interest or financial conflicts to disclose.

This article does not contain any studies with human or animal subjects performed by any of the authors.

References

- Swingle W T. The botany of *Citrus* and its wild relatives of the orange subfamily: I. Oakland: *University of California Press*, 1943
- Biswas M K, Chai L, Amar M H, Zhang X, Deng X. Comparative analysis of genetic diversity in *Citrus* germplasm collection using AFLP, SSAP, SAMPL and SSR markers. *Scientia Horticulturae*, 2011, **129**(4): 798–803
- Setsuko S, Uchiyama K, Sugai K, Yoshimaru H. Rapid development of microsatellite markers for *Pandanus boninensis* (Pandanaceae) by pyrosequencing technology. *American Journal of Botany*, 2012, **99**(1): e33–e37
- Setsuko S, Uchiyama K, Sugai K, Yoshimaru H. Isolation and characterization of EST-SSR markers in *Schima mertensiana* (Theaceae) using pyrosequencing technology. *American Journal of Botany*, 2012, **99**(1): e38–e42
- Biswas M K, Chai L J, Qi X, Deng X X. Generation, functional analysis and utility of *C. grandis* EST from a flower-derived cDNA library. *Molecular Biology Reports*, 2012, **39**(7): 7221–7235
- Huang H, Lu J, Ren Z, Hunter W, Dowd S E, Dang P. Mining and validating grape (*Vitis* L.) ESTs to develop EST-SSR markers for genotyping and mapping. *Molecular Breeding*, 2011, **28**(2): 241–254

7. Uchiyama K, Fujii S, Ishizuka W, Goto S, Tsumura Y. Development of 32 EST-SSR markers for *Abies firma* (Pinaceae) and their transferability to related species. *Applications in Plant Sciences*, 2013, **1**(2): 1–5
8. Cheng Y J, Guo W W, Yi H L, Pang X M, Deng X X. An efficient protocol for genomic DNA extraction from *Citrus* species. *Plant Molecular Biology Reporter*, 2003, **21**(2): 177–178
9. da Maia L C, Palmieri D A, de Souza V Q, Kopp M M, de Carvalho F I, Costa de Oliveira A. SSR Locator: tool for simple sequence repeat discovery integrated with primer design and PCR simulation. *International Journal of Plant Genomics*, 2008: 1–9
10. Ruiz C, Paz Breto M, Asíns M J. A quick methodology to identify sexual seedlings in citrus breeding programs using SSR markers. *Euphytica*, 2000, **112**(1): 89–94
11. Yan Y, Huang Y, Fang X, Lu L, Zhou R, Ge X, Shi S. Development and characterization of EST-SSR markers in the invasive weed *Mikania micrantha* (Asteraceae). *American Journal of Botany*, 2011, **98**(1): e1–e3
12. Terzoli S, Beritognolo I, Sabatti M, Kuzminsky E. Development of a novel set of EST-SSR markers and cross-species amplification in *Tamarix africana* (Tamaricaceae). *American Journal of Botany*, 2010, **97**(6): e45–e47
13. Ouyang J, Yang G, Zhou R, Chen S, Guo M, Wu W, Huang Y, Shi S. Development and characterization of 18 EST-SSR markers in *Sonneratia caseolaris*. *American Journal of Botany*, 2011, **98**(4): e78–e80
14. Ma J Q, Zhou Y H, Ma C L, Yao M Z, Jin J Q, Wang X C, Chen L. Identification and characterization of 74 novel polymorphic EST-SSR markers in the tea plant, *Camellia sinensis* (Theaceae). *American Journal of Botany*, 2010, **97**(12): e153–e156
15. Jewell M C, Frere C H, Prentis P J, Lambrides C J, Godwin I D. Characterization and multiplexing of EST-SSR primers in *Cynodon* (Poaceae) species1. *American Journal of Botany*, 2010, **97**(10): e99–e101
16. Hou X G, Guo D L, Wang J. Development and characterization of EST-SSR markers in *Paeonia suffruticosa* (Paeoniaceae). *American Journal of Botany*, 2011, **98**(11): e303–e305
17. Wang Z, Li J, Luo Z, Huang L, Chen X, Fang B, Li Y, Chen J, Zhang X. Characterization and development of EST-derived SSR markers in cultivated sweetpotato (*Ipomoea batatas*). *BMC Plant Biology*, 2011, **11**(1): 139
18. Chen C, Zhou P, Choi Y A, Huang S, Gmitter Jr F G. Mining and characterizing microsatellites from citrus ESTs. *TAG Theoretical and Applied Genetics*, 2006, **112**(7): 1248–1257
19. Tangphatsornruang S, Somta P, Uthapaisanwong P, Chanprasert J, Sangsrakru D, Seehalak W, Sommanas W, Tragoonrungs S, Srinives P. Characterization of microsatellites and gene contents from genome shotgun sequences of mungbean (*Vigna radiata* (L.) Wilczek). *BMC Plant Biology*, 2009, **9**(1): 137
20. Cavagnaro P F, Senalik D A, Yang L, Simon P W, Harkins T T, Kodira C D, Huang S, Weng Y. Genome-wide characterization of simple sequence repeats in cucumber (*Cucumis sativus* L.). *BMC Genomics*, 2010, **11**(1): 569
21. Biswas M K, Xu Q, Mayer C, Deng X. Genome wide characterization of short tandem repeat markers in sweet orange (*Citrus sinensis*). *PLoS ONE*, 2014, **9**(8): e104182
22. Wang H, Walla J A, Zhong S, Huang D, Dai W. Development and cross-species/genera transferability of microsatellite markers discovered using 454 genome sequencing in chokecherry (*Prunus virginiana* L.). *Plant Cell Reports*, 2012, **31**(11): 2047–2055