INTERNATIONAL JOURNAL OF AGRICULTURE & BIOLOGY
ISSN Print: 1560–8530; ISSN Online: 1814–9596
20F–076/202x/00–0–000–000
DOI: 10.17957/IJAB/15.1544
http://www.fspublishers.org

Full Length Article

Development of Pansies EST-SSR Markers and Analysis of Genetic Relationships among Pansy and Related Species

Xiaohua Du[1,2*], Hu Wang[1,2], Xiaopei Zhu[1,2], Jinyan Mu[1,2] and Huichao Liu[1,2]
[1]*School of Horticulture and Landscape Architecture, Henan Institute of Science and Technology, Xinxiang 453003, China*
[2]*Henan Province Engineering Research Center of Horticultural Plant Resource Utilization and Germplasm Enhancement, Xinxiang 453003, China*
[*]For correspondence: duxiaohua0124@sina.com
Received 15 April 2020; Accepted 04 June 2020; Published _____

Abstract

Although pansies (*Viola×wittrockiana*) are popular bedding flowers around the world, there is a limited availability of genomic resources, especially of the expressed sequence tag (EST)-SSR markers. In this study, 70 specific EST-SSR primers obtained from transcriptome sequencing of pansy leaves were selected to validate as available markers for pansies. Amplification across 35 pansy accessions revealed that 49 (70%) EST-SSR primers were successfully amplified DNA and generated a total of 309 amplicons and 283 polymorphic alleles. These markers exhibited high cross-species transferability by more than 80% from *V.×wittrockiana* to other species of *Viola* genus. The principal components analysis based on these EST-SSRs clearly separated two species of section *Viola* from the species of section *Melanium* and validated the proximity of *V.×wittrockiana* with *V. cornuta* and *V. tricolor*, confirming *V. cornuta* participating in the hybridization process of *V.×wittrockiana*, which was also supported by the results of analysis of molecular variance. The EST-SSR markers developed in this study can be used in molecular marker breeding and evolution analysis for *Viola*. © 2020 Friends Science Publishers

Keywords: *Viola×wittrockiana*; *Viola cornuta*; PcoA; EST; SSR

Introduction

Microsatellites or simple sequence repeats (SSRs) are short (1–6 bp) repeat motifs that can be found in both coding and non-coding DNA sequences of all higher organisms examined to date (Liu *et al*. 2020; Manee *et al*. 2020). They are usually associated with a high level of frequency of polymorphism, which provides a basis for the development of a marker system. Thanks to the characters of high level of polymorphism, co-dominant inheritance, adaptability to high-throughput genotyping, SSR marker technique, have been broadly used in genetic diversity analysis and linkage mapping (Röder *et al*. 1998; Liu *et al*. 2019).

Earlier experimental methods for developing SSRs involved isolating and sequencing clones containing putative SSR tracts, followed by designing and testing of flanking primers, which are laborious and costly (Schloss *et al*. 2002). With the development of next generation sequencing, obtaining high-throughput SSR information in the transcribed gene region and development of expressed sequence tag-SSR (EST-SSR) markers on large-scale is available. The EST-SSR markers provide the possibility of direct tagging of gene of interest (Xiao *et al*. 2014; Nie *et al*. 2017). They are likely to be more conserved across related species and therefore find higher levels of cross-species transferability than genomic SSRs (Cordeiro *et al*. 2001; Kantety *et al*. 2002; Decroocq *et al*. 2003), aiding in identification of conserved gene order across orthologous linkage groups for comparative analysis (Varshney *et al*. 2005). Development of EST-SSRs for different crops and ornamentals, such as oil palm (Xiao *et al*. 2014), tree peony (Wu *et al*. 2014), *Miscanthus* (Nie *et al*. 2017), *Tagetes erecta* (Zhang *et al*. 2018), *Hibiscus esculentus* (Li *et al*. 2018) and *glycyrrhiza* (Liu *et al*. 2019), has been carried out.

Pansies (*Viola ×wittrockiana*) are among the most popular garden flowers around the world. However, their DNA markers resources available are very limited. So far, only four DNA marker systems have been used in pansies, involving Random Amplified Polymorphic (RAPD) (Ko *et al*. 1998; Wang and Bao 2007; Vemmos 2015), Inter-Simple Sequence Repeat (ISSR) (Yockteng *et al*. 2003; Culley *et al*. 2007), Sequence-related Amplified Polymorphism (SRAP) (Wang *et al*. 2012; Du *et al*. 2019a) and Restriction Site Amplified Polymorphism (RSAP) (Li *et al*. 2015a). These DNA markers are usually dominant and unable to distinguish heterozygous from dominant homozygous resulting in insufficient genetic information. The co-dominant markers like EST-SSR for pansies are lacking.

Clausen (1926) reported that pansies were the hybrids of *Viol*a section *Melanium*, which originated from the crossing between a wild flower of Europe known as *V. tricolor* and a yellow Viola, *V. lutea*, and later further crossed with *V. cornuta*. But Zhang *et al*. (2010) believed that pansies were originally derived from the crossing between *V. tricolor* and *V. lutea*, and then the hybrid was crossed with a large and varied flower colored perennial *V. altaica.* Analysis of the genetic relationship among *V.×wittrockiana*, *V. tricolor* and *V. cornuta* in molecular level by utilizing DNA markers will be helpful to clarify this problem and the parent selection in pansies crossbreeding programs.

In this paper, bases on a *de novo* RNA-sequencing of pansies leaves at the transcriptome level (Du *et al.* 2019b), we designed the EST-SSR primers according to the flanking sequences of SSRs, then selected 70 primers to examine their efficiency of transferability and analysis ability on genetic diversity of pansies employing 42 pansies accessions and their related species. The objectives of this study were (i) to develop some EST-SSR markers for pansies, (ii) to examine the efficiency of marker transferability within *viola*, and (iii) to evaluate these EST-SSR markers in the genetic relationship analysis in pansies.

Materials and Methods

Plant material and DNA isolation

A total of 40 accessions of *Viol*a section *Melanium* including 35 breeding lines of *V. ×wittrockiana*, 3 breeding lines of *V. cornuta* and 2 lines of *V. tricolor*, and 2 wild species involving *V. hancockii* and *V. prionantha* of section *Viola* in Xinxiang, Henan province, China, were employed in this study (Table 1). All accessions were grown at the field site of Henan Institute of Science and Technology.

Genomic DNA was extracted from 0.2 g fresh leaves using SDS method. The quality of DNA was checked on a 0.6% (w/v) agarose gel and the concentration was determined by UV visible (Thermo Scientific NanoDrop2000, USA). All DNA samples were diluted to 20 ng L^{-1} and stored at -20°C prior to PCR amplification.

Generation of EST-SSRs and designing of primers

Using MISA software, a total of 23,791 potential SSRs were identified from 20, 679 unigene sequences after transcriptome sequencing of the pansies leaves. PRIMER3 (http://www-genome.wi.mit.edu/cgi-bin/primer/primer3_www.cgi) was employed to design EST-SSR primers with the following criteria: 18–23 bp primer length, 55–65°C melting temperature, 40–60% GC content, and 80–300 bp amplicon size. Finally, a total of 6,863 specific primer pairs were designed from 9,228 SSR-containing sequences. To test these primers availability, 70 primer pairs were selected for synthesis and screened in the experimental plant materials.

Amplification and detection of microsatellite alleles

PCR amplification was performed in a total volume of 10 μL containing 2 μL (40 ng) genomic DNA, 2 μL ddH_2O, 5 μL 1 × Taq PCR Master Mix (Beijing ComWin Biotech Co. Ltd., Beijing, China), and 0.5 μL (10 pmol) each reverse and forward primer. The following amplification protocol was performed: pre-denaturation at 95°C for 2 min, followed by 35 cycles of denaturation at 95°C for 30 s, annealing at 58–60°C for 30 s (different primer annealing temperatures are shown in Table 2) and extension at 72°C for 30 s, with a final extension at 72°C for 4 min. The PCR products were separated on 6% (w/v) denaturing polyacrylamide gels in 1×TBE buffer solution at 60 w of power for 2.5 h, and then stained using silver staining protocol. The separated DNA bands were visualized and estimated by comparing with 100 bp ladder molecular size standard (Solarbio).

Data analysis

The number of effective alleles (Ne), Shannon's information index (I), observed heterozygosity (H_O), expected heterozygosity (H_E), percentage of polymorphic alleles (PPA), and genetic differentiation coefficient (F_{ST}), gene flow (N_m), and Nei's gene diversity (H), genetic distances among different populations, were calculated using Popgene 32 (Quardokus 2000). A principal coordinate analysis (PCoA) based on simple matching similarity coefficients and unweighted pair group method arithmetic averages (UPGMA) were used to cluster all accessions using NTSYSpc 2.1 (Jensen 1989). Analysis of molecular variance between and within of section *Melanium* and section *Viola* was calculated using GeneAlEx v6.501 (Peakall and Smouse 2006; 2012).

Results

SSR marker development

Seventy EST-SSR primer pairs were tested on 42 pansies accessions involving 4 related species. Forty-nine primer pairs (70%) successfully amplified DNA for *V.×wittrockiana*, *V. tricolor* and *V. cornuta*. Of these, 40 primer pairs generated amplicons for two species of section *Viola*, and 36 primer pairs produced amplicons for all of the species tested. This suggested that most of EST-SSR markers developed from *V.×wittrockiana* can be transferable across species both in section *Melanium* and section *Viola*. The characterizations of these primer pairs and their amplicons sizes are presented in Table 2.

A total of 309 amplicons were produced by these primer pairs, with average of 6.3 amplicons per primer pair. The most amplicon-producing primer pair was P66, which produced 18 amplicons (Fig. 1). Nineteen EST-SSR primer pairs (39%) amplified a single amplicons and 30 primer pairs (61%) amplified two to five loci, resulting in 96 loci in

Table 1: The name, pedigree, species, flower type, and origin of the *Viola* accessions in this study

No.	Name	Pedigrees	Flower type	Species	Country/Company of origin
1	DFM-11-1-1	Frühblühende Mischung	large	*V.* ×*wittrockiana*	Germany/Gartenland Aschersleben
2	DFM-11-2-3	Frühblühende Mischung	large	*V.* ×*wittrockiana*	Germany/Gartenland Aschersleben
3	DFM-11-2-4-1	Frühblühende Mischung	large	*V.* ×*wittrockiana*	Germany/Gartenland Aschersleben
4	DFM-1-2-3-3	Frühblühende Mischung	large	*V.* ×*wittrockiana*	Germany/Gartenland Aschersleben
5	DFM-16-1-2-6	Frühblühende Mischung	large	*V.* ×*wittrockiana*	Germany/Gartenland Aschersleben
6	DFM-16-2-2	Frühblühende Mischung	large	*V.* ×*wittrockiana*	Germany/Gartenland Aschersleben
7	DFM-8-3-1-2	Frühblühende Mischung	large	*V.* ×*wittrockiana*	Germany/Gartenland Aschersleben
8	DSRAB-1-2-3	Schweizer Riesen Alpensee	large	*V.* ×*wittrockiana*	Germany/Dehner Seed
9	DSRAB-1-2-4	Schweizer Riesen Alpensee	large	*V.* ×*wittrockiana*	Germany/Dehner Seed
10	DSRAB-1-4-2	Schweizer Riesen Alpensee	large	*V.* ×*wittrockiana*	Germany/Dehner Seed
11	DSRFY-1-1-2	Schweizer Riesen Firnengold	large	*V.* ×*wittrockiana*	Germany/Dehner Seed
12	G10-1-1-1-3-3	229.10	medium	*V.* ×*wittrockiana*	China/JiuQuan Jinqiu Horticulture Seed
13	G10-1-3-1-2	229.10	medium	*V.* ×*wittrockiana*	China/JiuQuan Jinqiu Horticulture Seed
14	G10-1-3-1-4-2	229.10	medium	*V.* ×*wittrockiana*	China/JiuQuan Jinqiu Horticulture Seed
15	G1-1-1-1-1-4	229.01	medium	*V.* ×*wittrockiana*	China/JiuQuan Jinqiu Horticulture Seed
16	G10-1-1-1-3-2	229.10	medium	*V.* ×*wittrockiana*	China/JiuQuan Jinqiu Horticulture Seed
17	HAR2-1-14-1-1	Aalsmeerse Giants	large	*V.* ×*wittrockiana*	NL/Buzzy Seeds
18	JB-1-1-1	Penny Blue	small	*V. cornuta*	USA/Goldsmith seed
19	JB-1-1-6	Penny Blue	small	*V. cornuta*	USA/Goldsmith seed
20	JY-1-1-2	Penny Yellow	small	*V. cornuta*	USA/Goldsmith seed
21	MYB-1-2	Matrix™ Yellow Blotch	large	*V.* ×*wittrockiana*	USA/PanAmerican Seed
22	MYC-1-1-3-4	Matrix™ Yellow Clear	large	*V.* ×*wittrockiana*	USA/PanAmerican Seed
23	PXP-BT-4-1-1-1	Panola XP Blue True	medium	*V.* ×*wittrockiana*	USA/PanAmerican Seed
24	PXP-BT-4-1-1	Panola XP Blue True	medium	*V.* ×*wittrockiana*	USA/PanAmerican Seed
25	RCO-1-3-4	Clear orange of power mini	medium	*V.* ×*wittrockiana*	Japan/Takii Seed
26	RRB-1-3	Beacon blue of Dynamite	large	*V.* ×*wittrockiana*	Japan/Sakata Seed
27	RRB-2-7	Beacon blue of Dynamite	large	*V.* ×*wittrockiana*	Japan/Sakata Seed
28	RRB-3-1	Beacon blue of Dynamite	large	*V.* ×*wittrockiana*	Japan/Sakata Seed
29	XXL-G-1-1-2-3	XXL Golden e	extra large	*V.* ×*wittrockiana*	USA/PanAmerican Seed
30	XXL-G-1-1-3	XXL Golden	extra large	*V.* ×*wittrockiana*	USA/PanAmerican Seed
31	XXL-G-1-1-7-4	XXL Golden	extra large	*V.* ×*wittrockiana*	USA/PanAmerican Seed
32	EYO-1-2-1-4	Yellow large flower	large	*V.* ×*wittrockiana*	China/Shanghai Academy of Landscape Architecture Science and Planning
33	EYO-1-2-1-5	Yellow large flower	large	*V.* ×*wittrockiana*	China/Shanghai Academy of Landscape Architecture Science and Planning
34	EYO-1-1-4	Yellow large flower	large	*V.* ×*wittrockiana*	China/Shanghai Academy of Landscape Architecture Science and Planning
35	EWO-2-1-1	White large flower	medium	*V.* ×*wittrockiana*	China/Shanghai Academy of Landscape Architecture Science and Planning
36	EWO-1-1-3	White large flower	medium	*V.* ×*wittrockiana*	China/Shanghai Academy of Landscape Architecture Science and Planning
37	MW-1-1-1-1	Light blue flower	medium	*V.* ×*wittrockiana*	China/Henan Institute of Science and Technology
38	EWO-MW	Light blue flower	medium	*V.* ×*wittrockiana*	China/Henan Institute of Science and Technology
39	E01	Blue-purple small flower	small	*V. tricolor*	China/Shanghai Academy of Landscape Architecture Science and Planning
40	08H	Johnny Jump Up	small	*V. tricolor*	Germany/Dehner Seed
41	*V. hancockii*	Wild species	small	*V. hancockii*	China/Xinxiang
42	*V. prionantha*	Wild species	small	*V. prionantha*	China/Xinxiang

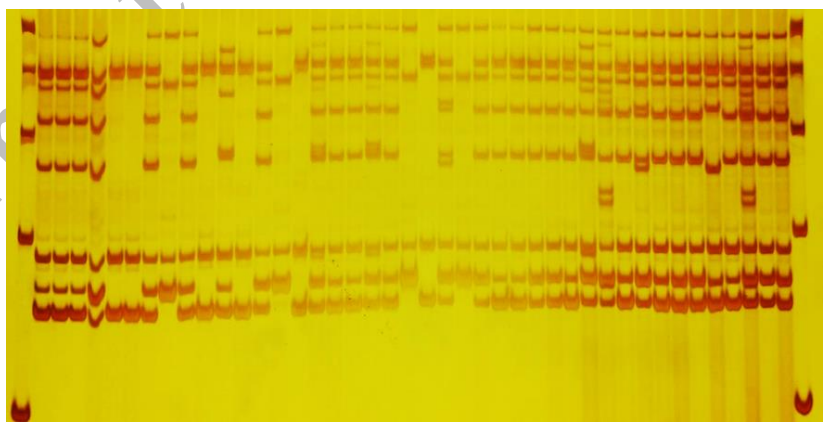

Fig. 1: The profile of amplification by EST-SSR primer pair P66

total (Table 3). The number of alleles per locus ranged from 1 to 13, with an average of 3.22 alleles per locus. Approximately 61% of the primer pairs amplified at least one PCR fragment size larger than expected. For example, the expected product size for primer P66 was 151 bp, but one of PCR amplicons was more than 400 bp.

Table 2: Characterization of 49 EST-SSR markers

Primer ID.	Forward primer sequence (5′→3′)	Reverse primer sequence (5′→3′)	T_m (°C)	Expected product size (bp)	Amplified product size (bp)	Motif	Corresponding unigene function	No. of Loci
P1	ACCTGAGCCTGATTCCAAGC	CCATCTCCGGTCACTGTTCC	60	203	260-480	(CTG)5	Uncharacterized protein	2
P2	AGGTCTGCGAGGAGGAAGAT	TGTATCCCATTGACCGCCAG	60	168	160-200	(GCG)5	hypothetical protein POPTR	2
P3	GCCTTGTCCTCAGCAAAACG	TGCAAGAGCTTTTCGTCAGC	60	219	210-500	(TCG)5	conserved hypothetical protein	3
P5	CCCAAACCTTAACCCGAGCT	GATACGGTTGGAGTGGACGG	60	224	165-300	$(CAC)_5$	uncharacterized protein	3
P9	CCCCCGCAATTTTGGTGAAG	CTGGGCATGGTTGATCAGGT	60	108	100-200	(TGA)6	formin homology 2 domain-containing family protein	2
P11	TCCTCAACCTCCTGCTCAGA	CCACTACCCAACAAACCCCA	60	238	160-170	$(TC)_6$	hypothetical protein POPTR	1
P12	GAGGGCTCGTTTCAAATGGC	GCAAATGGGTCGTCGTCAAC	60	185	180-410	$(CAG)_5$	transcription factor bHLH63 isoform X1	5
P16	CGCAGTCTCCGTCGATTACA	TGTCTCCGGCTAAAACCACC	60	170	160-340	$(CCG)_5$	catalytic	4
P17	TCTCTCCCTCACTTCTCCGT	GCTTGGCTCTGACGTAAGGT	60	175	165-280	(GCA)5	Tetratricopeptide repeat-like superfamily protein	1
P18	TTTCCACCTCCAAACCTCGG	TGTTTGATGCTGCAGGGGTA	60	289	250-360	$(CCA)_5$	pumilio homolog 1-like	2
P20	GAGCTGGAGATCCCGTTAGC	CCTCTGCTTCTGCTAACCCC	59	278	290-340	$(GCT)_5$	VQ motif-containing family protein	2
P21	AAGGTGGCTCAGTGCATCTC	GCAGTGAAGGAAACACACGC	60	229	190-300	(CTC)5	RNA-binding protein	3
P23	TGCCACCTGATTCCATTGCA	TGTGGCTGTTTGTTGTGCTG	60	203	200-300	$(AGG)_5$	transcription factor bHLH91-like	3
P24	GGTAGGAGACGCTGGGAAAC	GCCGCGTTACCATAGCTAGT	60	288	220-420	$(AGC)_5$	B3 domain-containing transcription factor NGA1-like isoform X1	3
P25	GGGAAGAGTGAACGAGGTGG	GGCATCTTGTTGCTGCTTCC	60	271	150-185	(TAC)6	transcription factor GTE6	1
P26	CCGCCTACTCCACTGAACTC	ACATGGAAGAGGAGCAAGCA	59	265	100-150	$(TCA)_5$	small RNA 2'-O-methyltransferase-like	2
P27	GCTTATGTGCAGTGTATGGCG	ACCTCTTTCTGCACACCACC	60	137	110-150	$(GCT)_8$	aha1 domain-containing family protein	2
P30	ACCGCAAACCAAGCAAACAA	TGAGGATGAAGGGGATGGGA	60	169	110-220	$(CAT)_6$	hypothetical protein POPTR	2
P32	GAAACTATCCACCACCGCCA	TCGGGAATACGGTGGTTGTG	60	167	167-210	$(CCA)_5$	carboxypeptidase Y	2
P33	ACCTCCCCCTCTTCCTCATC	TTTCAGCCGATCGACGTAGG	60	253	200-270	(CCG)5	hypothetical protein POPTR	1
P34	GGACCTGCTGCCTCATCAAG	CCAGGTCACAATTCCAACTGC	60	111	300-340	$(AAG)_5$	mitochondrial import receptor TOM20-2 family protein	2
P35	CCATTCGCTACAGCTTTGGC	CGGAGGAGGTTGTTTTGGGT	60	223	170-190	(CCA)5	protein OSB3	1
P36	CTCACTGAGTGGCTCATCCC	GAGGGGACATTGAGGCTGAC	60	128	128	(TCT)5	PWWP domain-containing family protein	1
P38	CGAAGAGCTTGAAGGCCCAA	TGATGCTGCCGAAACTAACG	59	239	170-240	$(CAA)_5$	7-deoxyloganetic acid glucosyltransferase-like	2
P39	CCCCTCCCACCTTTCCTTTC	CAGGCTGTTTGGTTGCTGAC	60	141	150-230	$(GGC)_5$	uncharacterized protein	2
P40	AGGCTCCTAGGGTCAAACCT	CGTCGCAAACAGTGAACACA	60	250	350-570	(GTG)5	Small nuclear ribonucleoprotein	1
P41	AGAACAGCAGCCCCTTTTGG	GGCCAGCCCCATTTTCATTG	60	196	190-210	(TGA)5	aluminum-activated malate transporter 9-like	1
P42	TGGCACTCTTCCTCGTTGTC	TGTCGTAGAGGCTGCCTACT	60	138	120-190	(CTC)5	cytochrome P450 98A2	1
P43	TTCAAAGCCATCCACCTCCC	AGCAGTGGAGAGGGGATCAT	60	255	200-240	(CT)6	nuclear acid binding protein	1
P44	AGCCAAGCCTCTCTCTCGTA	AGCAGTGGAGAGGGGATCAT	60	194	200-210	(AGC)5	nuclear acid binding protein	1
P45	CCTGGTGCGGAATTGTTGTG	GGGAGCTGGGTTTGTTGAGT	60	265	200-350	(CAC)5	uncharacterized protein LOC105644223 isoform X1	2
P46	AGGGTTGAGCCTCAGTCTCT	ACGCAATGAAACATGCCCTG	60	224	200-520	(AGG)5	uncharacterized PKHD-type hydroxylase At1g22950-like isoform X1	3
P47	GGCGATCGAGAAATGAGGCT	CGCTACCCATCATCTGTCTCC	60	286	260-370	$(TGC)_5$	lipoxygenase	3
P48	ACGGTGGTGGTTTATGGTGG	CTCTGGTGGTTCGAGTGGTC	60	273	200-500	(TTC)6	hypothetical protein POPT	2
P49	GTGGCAAAGCTGGGAACAAG	TGCTACTACCCGTTTTGCTCT	59	149	180-240	(CAG)5	hypothetical protein	1
P50	TGTCAACGGAGCAAAA TGGTC	GCCTGTGGAAAAAGCAAGCA	59	196	190-255	(ACT)6	transcriptional corepressor LEUNIG-like isoform X	1
P51	GATCCCACAGCGTTTACCCA	GCCGCGTTACCATAGCTAGT	60	224	200-360	$(AGC)_5$	B3 domain-containing transcription factor NGA1-like isoform X1	4
P52	ATTGCTACAGTCGCCATCCC	GAGCGGACCGGATGTGTTTA	60	196	180-190	(TC)6	amino acid transporter	1
P53	AGGCTTCCTCTTCGGTCTCT	GTCTGGATCCCGACGAATCC	60	171	170-230	(CTC)5	probable beta-1,3-galactosyl transferase 14	1
P57	TGTGACGACTGAAAAGGCCA	GCACAAACAACATAAGGGCGA	60	267	420-460	$(GAA)_5$	phenylalanyl-tRNA synthetase beta chain	1
P58	TTAGGACGAGCATGCACAGG	CGCAGTTCGTTTCACCGATG	60	279	220-450	(ATC)5	NADH dehydrogenase	2
P61	TCAGCTCAGCGAGAAACACA	AGGAAAGACACCACCACCAC	60	234	235-340	(CTG)5	Jatropha curcas protein tesmin/TSO1-like CXC 5	1
P62	TCACCGACCAGCAAACATCA	GGGGTTTTGTGGAAAGGTGC	60	198	190-200	(CTT)5	protein FD-like isoform X2	1
P63	ATGGGGAAATGGGCTCACAA	TCCCAAATGGCATCGGAACT	60	247	250-305	(AC)9	bidirectional sugar transporter SWEET2	2
P65	GGCCGTATGTCTTCCACACA	CAGGGGTGGGCAAAGATCAT	60	244	230-310	(ATC)5	casein kinase I-like	2
P66	CCTTCCGCTTACTCACTCCG	TGTACGGATGCGAATCGAGG	60	151	150-460	(AAG)5	Uncharacterized protein isoform 1	3
P67	TACCCAGAAAACTCCACCGC	ATCCGCCCAGTTTGTAGTGG	60	280	280	(AGA)5	probable AMP deaminase	1
P68	AAACCCCAAAAACCGCATGG	AAATCCCCTCCCTCTCCTCC	60	144	144-280	(GT)6	hypothetical protein CISIN	4
P70	TTTGTCGACGCCATCATCCA	GGGCGTATGCAGGACATGAT	60	276	276-610	(TGA)5	mitotic spindle checkpoint family protein	2
In total:								96

Genetic diversity

When the amplicons amplified were screened for length polymorphisms, 283 polymorphic alleles generated by 46 primer pairs were detected among 42 genotypes, with an average of 6.15 polymorphic alleles per primer pair. A total of 269 polymorphic alleles were produced for section *Melanium* by 46 primer pairs. Of these, 266 polymorphic alleles were for *V.×wittrockiana*, 84 polymorphic alleles for *V. cornuta*, and 50 polymorphic alleles for *V. tricolor*. The number of polymorphic alleles for *Viola* section was 44. The most polymorphic alleles were generated by primer P66, yielding 17 polymorphic alleles. However, three primer pairs including P36, P52 and P67 produced no polymorphic alleles.

Table 3: Genetic diversity of locus level estimated from 42 accessions of *Viola*

Locus	*N*	*Ne*	*I*	*He*	*Ho*	F_{ST}	N_m	*H*
V1185	1	1.707	0.605	0.414	0.419	0.100	4.482	0.418
V1200	1	1.049	0.114	0.047	0.048	0.022	22.328	0.047
V2300	2	1.445	0.483	0.306	0.310	0.321	1.683	0.282
V2330	3	1.505	0.466	0.302	0.305	0.566	0.417	0.316
V3180	2	1.849	0.650	0.458	0.464	0.767	0.152	0.437
V3220	4	1.655	0.573	0.387	0.392	0.417	1.125	0.389
V5170	4	1.320	0.318	0.195	0.198	0.433	6.338	0.201
V5230	4	1.485	0.473	0.307	0.311	0.238	3.999	0.312
V5270	3	1.194	0.289	0.157	0.159	0.083	6.877	0.159
V9190	2	1.062	0.135	0.059	0.059	0.114	12.133	0.060
V9220	2	1.337	0.384	0.234	0.237	0.554	1.304	0.256
V1123	2	1.986	0.690	0.497	0.503	0.192	5.600	0.493
V1218	2	1.354	0.410	0.250	0.253	0.541	2.824	0.126
V1222	4	1.359	0.426	0.259	0.263	0.259	2.640	0.269
V1226	3	1.611	0.534	0.358	0.362	0.377	1.245	0.368
V1231	4	1.707	0.598	0.409	0.414	0.289	5.252	0.416
V1241	2	1.725	0.609	0.418	0.424	0.550	0.410	0.443
V1617	4	1.230	0.271	0.162	0.164	0.530	0.874	0.053
V1621	4	1.548	0.464	0.310	0.314	0.499	5.802	0.321
V1627	4	1.492	0.448	0.293	0.296	0.400	1.865	0.283
V1632	4	1.449	0.468	0.297	0.301	0.238	3.229	0.304
V1727	2	1.698	0.581	0.396	0.401	0.612	0.420	0.403
V1732	4	1.349	0.364	0.227	0.230	0.123	19.811	0.231
V1736	3	1.392	0.453	0.280	0.284	0.310	2.929	0.305
V1827	3	1.451	0.394	0.259	0.263	0.153	16.675	0.254
V1835	4	1.226	0.297	0.170	0.172	0.663	2.148	0.200
V2032	3	1.439	0.409	0.269	0.273	0.497	1.161	0.285
V2120	4	1.250	0.327	0.188	0.190	0.492	2.924	0.137
V2128	3	1.628	0.523	0.353	0.357	0.546	0.571	0.362
V2133	3	1.552	0.513	0.336	0.340	0.488	1.272	0.350
V2135	4	1.192	0.245	0.142	0.143	0.505	0.885	0.158
V2324	3	1.551	0.504	0.330	0.334	0.436	1.166	0.343
V2330	2	1.801	0.628	0.438	0.443	0.792	0.133	0.456
V2426	2	1.239	0.335	0.190	0.192	0.081	5.798	0.191
V2431	4	1.726	0.596	0.409	0.414	0.468	1.262	0.419
V2438	5	1.482	0.483	0.310	0.314	0.374	1.855	0.311
V2545	3	1.331	0.414	0.248	0.251	0.402	2.195	0.269
V2621	3	1.928	0.673	0.480	0.486	0.464	0.627	0.484
V2624	4	1.478	0.391	0.261	0.264	0.298	11.395	0.256
V2712	3	1.662	0.550	0.373	0.377	0.499	1.313	0.381
V2714	3	1.373	0.322	0.206	0.209	0.199	12.432	0.204
V3012	2	1.724	0.595	0.408	0.413	0.482	0.928	0.404
V3018	6	1.492	0.492	0.319	0.323	0.509	1.963	0.290
V3227	3	1.568	0.538	0.355	0.359	0.238	2.020	0.355
V3240	5	1.379	0.377	0.238	0.240	0.237	8.068	0.247
V3321	6	1.496	0.443	0.290	0.294	0.372	666.560	0.293
V3417	2	1.655	0.581	0.393	0.398	0.645	0.610	0.255
V3419	3	1.900	0.666	0.473	0.479	0.663	0.278	0.483
V3518	3	1.580	0.546	0.362	0.367	2.477	-0.250	0.360
V3822	2	1.600	0.509	0.339	0.343	0.256	2.676	0.334
V3830	4	1.726	0.603	0.414	0.419	0.343	2.381	0.408
V3845	3	1.819	0.632	0.442	0.447	0.222	2.668	0.444
V3911	3	1.706	0.594	0.406	0.411	0.666	0.384	0.327
V3919	3	1.389	0.375	0.236	0.239	0.223	6.521	0.228
V4050	3	1.339	0.372	0.227	0.230	0.571	2.175	0.231
V4120	3	1.761	0.596	0.412	0.419	-0.018	1333.196	0.405
V4219	2	1.995	0.692	0.499	0.506	0.732	0.183	0.490
V4324	2	1.494	0.377	0.258	0.261	0.381	1.031	0.253
V4421	1	1.084	0.169	0.077	0.078	1.244	-0.098	0.078
V4520	2	1.925	0.673	0.480	0.486	0.570	0.457	0.489
V4531	3	1.431	0.403	0.257	0.260	0.360	2.374	0.248
V4624	4	1.843	0.635	0.446	0.451	0.338	1.229	0.449
V4645	4	1.665	0.529	0.363	0.367	0.424	3.958	0.371
V4722	8	1.412	0.438	0.275	0.278	0.289	3.999	0.242
V4822	2	1.940	0.677	0.484	0.491	0.507	0.731	0.493
V4850	1	2.000	0.693	0.500	0.507	0.206	1.932	0.500
V4918	1	1.888	0.663	0.470	0.477	2.732	-0.317	0.462
V5025	2	1.466	0.498	0.318	0.322	-1.552	2000.000	0.323
V5120	2	1.626	0.561	0.376	0.381	0.394	0.955	0.385
V5124	2	1.490	0.510	0.328	0.332	0.342	1.308	0.305
V5128	4	1.447	0.439	0.283	0.286	0.264	2.317	0.274
V5322	5	1.748	0.594	0.409	0.414	0.327	2.228	0.407
V5727	5	1.350	0.355	0.222	0.225	0.431	5.670	0.216
V5825	4	1.523	0.506	0.329	0.333	-0.199	999.869	0.330
V5835	4	1.696	0.568	0.388	0.393	2.984	1499.882	0.392
V6129	5	1.421	0.357	0.232	0.235	0.318	9.420	0.238
V6219	1	1.159	0.264	0.137	0.139	0.111	4.019	0.141
V6326	3	1.732	0.611	0.421	0.426	0.293	1.694	0.429
V6329	2	1.409	0.466	0.290	0.294	0.481	1.526	0.306
V6530	3	1.316	0.336	0.204	0.207	0.278	1.604	0.198
V6616	3	1.568	0.453	0.311	0.315	0.222	15.931	0.299
V6618	1	1.049	0.114	0.047	0.048	0.022	22.328	0.047
V6634	13	1.400	0.353	0.227	0.230	0.252	3.595	0.224
V6816	2	1.478	0.438	0.284	0.287	0.312	1.441	0.275
V6821	2	1.284	0.319	0.194	0.196	0.106	12.225	0.183
V6827	6	1.485	0.404	0.268	0.271	0.264	8.098	0.263
V7035	3	1.145	0.229	0.121	0.123	0.226	16.793	0.134
V7060	4	1.584	0.511	0.343	0.347	0.314	3.341	0.337
Total	283	1.523	0.468	0.308	0.312	0.440	0.637	0.304

N = Number of polymorphic alleles per locus; N_e = Effective number of alleles; *I* = Shannon's Information index; H_O = Observed heterozygosity; H_E = Expected heterozygosity; F_{ST} = Genetic differentiation coefficient; N_m = Gene flow; *H* = Gene diversity

Table 4: Genetic diversity parameters of *Viola* sections and species

Section	Species	N_L	*N*	*PPA* (%)	*Na*	*Ne*	*I*	*H*
Melanium		40	269	94.70	1.922	1.495	0.444	0.296
	Viola ×wittrockiana	35	266	93.99	1.940	1.496	0.444	0.293
	V. cornuta	3	97	34.28	1.343	1.232	0.197	0.133
	V. tricolor	2	50	17.67	1.177	1.125	0.107	0.073
Viola		2	48	16.96	1.186	1.132	0.1125	0.077
Total		42	283	100.00	2.000	1.506	0.456	0.300

Note: N_L = Number of breeding lines; *N* = Number of polymorphic alleles; *PPA* = Percentage of polymorphic alleles; *Na* = Observed number of alleles; *Ne* = Effective number of alleles, *I*= Shannon's Information index; *H* = Nei's gene diversity

At the locus level, a total of 283 polymorphic alleles were present in 88 loci. The polymorphism level of the loci (*I*) ranged from 0.114 (at the locus V6618) to 0.693 (V4850), with an average of 0.468. The mean observed homozygosity (*Ho*) was 0.312, ranging from 0.048 (at the locus V6618) to 0.507 (V4850), and the expected heterozygosity (*He*) ranged from 0.047 (at the locus V6618) to 0.500 (V4850), with an average of 0.308 (Table 3). With respect to the population level, the genetic diversity (*H*) ranged from 0.073 for *V. tricolor* to 0.415 for *V.×wittrockiana* (Table 4).

Genetic relationship

Based on 283 polymorphic alleles detected by 46 EST-SSR markers, the genetic distances between section *Viola* and section *Melanium* were greater than those among species of section *Melanium* (Table 5). PCoA partitioned 8.84 and 7.08% of the total variance to the first two axes, cumulating in 15.91% of the total variation. PCoA clearly separated two accessions of the section *Viola* from those of section *Melanium* (Fig. 2), while there was no obvious distinction

Table 5: Genetic distances among *Viola* section or species tested

Population ID	*V.* ×*wittrockiana*	*V. cornuta*	*V. tricolor*	*Viola* section
Viola ×*wittrockiana*		0.9172	0.8357	0.7755
V. cornuta	0.0865		0.8023	0.7381
V. tricolor	0.1795	0.2202		0.6622
Viola section	0.2542	0.3037	0.4122	

Nei's genetic identity (above diagonal) and genetic distance (below diagonal)

Table 6: Analyses of molecular variance (AMOVAs) for two *Viola* sections and three species of section *Melanium*

Source	df	Sum of squares	Variance components	Percentage of variation	*P*-value
1. Total	41	2029.143	72.171	100%	
Among sections	1	142.418	25.003	35%	0.005**
Within sections	40	1886.725	47.168	65%	
2. *Melanium* section	39				
Among species	2	140.001	5.258	10%	0.002**
Within species	37	1709.724	46.209	90%	

Note: d.f. = degree of freedom; $^{**}P < 0.01$

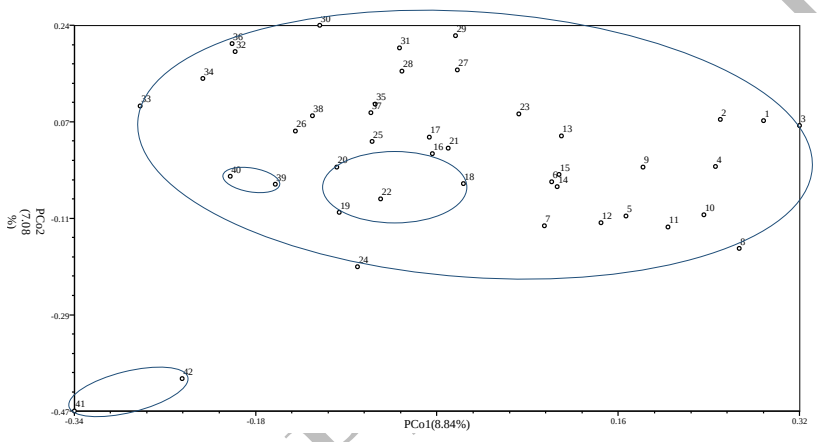

Fig. 2: Principal coordinates analysis (PCoA) based on the matrix of Nei's unbiased genetic distance among 42 accessions of *Viola*

between the accessions of *V.*×*wittrockiana* and those of the other two species (*V. tricolor* and *V. cornuta*) of section *Melanium*. AMOVAs revealed that 35% of the genetic diversity was presented between sections *Melanium* and *Viola*, whereas only 10% of the genetic variation occurred among species of section *Melanium* (Table 6).

Discussion

EST-SSR marker is one of most popular DNA makers nowadays due to its codominant, highly informative, locus-specific and adaptable to high-throughput genotyping, as well as gene tagging of interest traits and higher levels of cross-species transferability. With the development of next-generation sequencing, obtaining high-throughput information and development of EST-SSR markers on large-scale through RNA-sequencing has become an efficient means. Using transcriptome sequencing, we obtained 6,863 specific EST-SSR primers for pansies. Preliminary screening of seventy primers of them showed that 70% of these EST-SSR primers successfully amplified DNA and 66% generated polymorphic alleles for pansies (Table 2). The success of amplified primers in pansies was higher than that in *Rosa roxburghii* (Yan *et al.* 2015) and onion (Li *et al.* 2015b), but lower than that in eggplant (Wei 2016) and *Tagetes erecta* (Zhang *et al.* 2018). A possible reason for some primers failing to produce amplicons is either an intron occurred within the primer sequences interrupting amplification, or a large intron disrupted PCR extension (Yu *et al.* 2004).

Because EST-SSR markers are developed in relatively conserved gene sequences, this allowed to develop EST-SSR primers that could amplify orthologous loci in multiple species. This study showed EST-SSRs were not only highly conserved among the relative species in section *Melanium*, but also among more distantly related species in section *Viola* with 81.6% of transferability (Table 4). It is reported that SSRs were highly conserved in barley and wheat (Holton *et al.* 2002; Kantety *et al.* 2002; Yu *et al.* 2004).

The occurrence of approximately 61% of primers amplified at least one PCR fragment size larger than expected in this study was also found in the study on hexaploid wheat (Yu *et al.* 2004). The cause for this phenomenon is not likely due to polymorphism of repeat length within the SSRs, rather the result of insertion-deletion

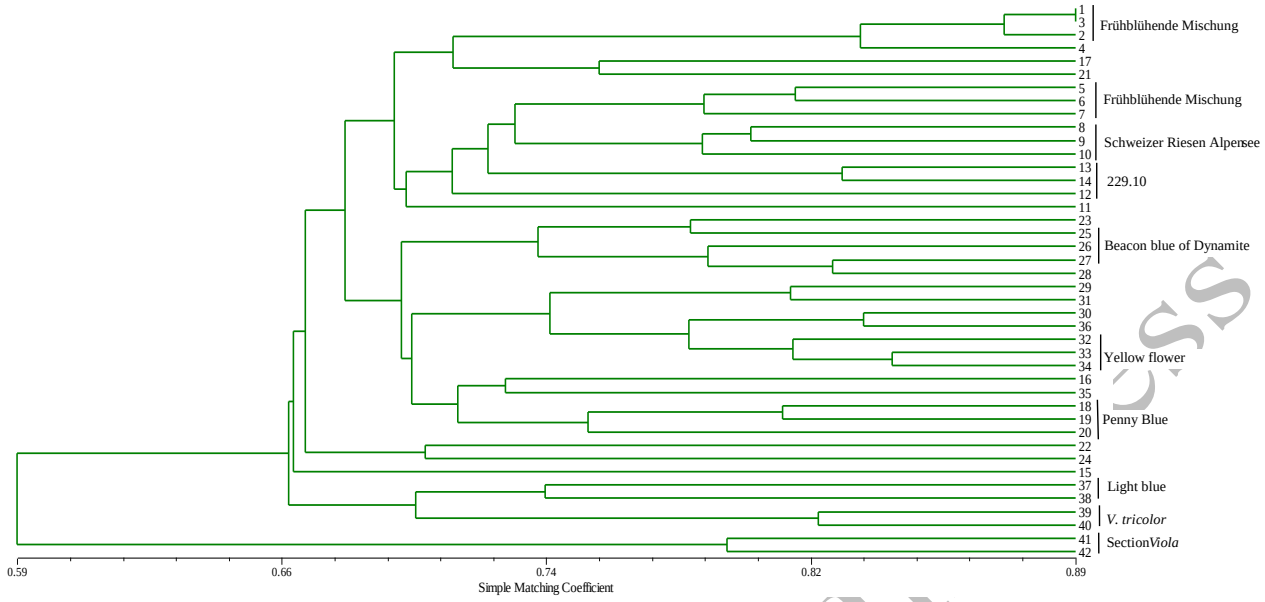

Fig. 3: UPGMA Dendrogram of 42 pansies accessions and their related species based on EST-SSR markers (Note: the labels at the right side indicate from the same parent or belonging to the same specie s or section)

variability within the amplicon. Some of EST-SSR primer pairs amplified more than one locus in pansies, which also happened in hexaploid wheat (Yu *et al*. 2004). These multi-loci detecting markers appeared possibly owing to sequence conservation in coding regions (Röder *et al*. 1998), polyploidy, and gene duplication (Anderson *et al*. 1992).

The UPGMA of all accessions showed the most breeding lines derived from the same parents were firstly clustered together (Fig. 3), indicating the genetic relationships among the accessions revealed based on the EST-SSRs was generally consistent with their pedigrees. The PCoA (Fig. 2) and the UPGMA (Fig. 3) clearly separated two accessions of the section *Viola* from those of section *Melanium*, and the result was further verified by the results of AMOVA (Table 6). This observation was in concurrence with the botanical classification. All of the above revealed the genetic relationships based on the EST-SSR markers are reliable.

The PCoA based on the EST-SSR markers developed in this study also revealed no obvious distinction among the accessions of *V.×wittrockiana* and those of *V. tricolor* and *V. cornuta* (Fig. 2). This confirmed that *V. tricolor* and *V. cornuta* both participated in the hybridization process of *V.×wittrockiana* (Clausen 1926).

Conclusion

Preliminary screening of 70 EST-SSR primers obtained from transcriptome sequencing of *V.×wittrockiana* developed 49 EST-SSR markers for pansies and showed high level of transferability by more than 80% from *V.×wittrockiana* to other species of *Viola* genus. These markers generated a total of 309 amplicons and 283 polymorphic alleles across 42 accessions of pansies and their related species. Based on the polymorphic alleles detected, the genetic relationships revealed that there was no obvious distinction between the accessions of *V.×wittrockiana* and those of *V. tricolor* and *V. cornuta*, confirming *V. tricolor* and *V. cornuta* both participating in the hybridization process of *V.×wittrockiana*.

Acknowledgements

This work was financially supported by the International Science and Technology Cooperation Research Programme of Henan Province (Grant No.182102410029) and Henan Institute of Science and Technology Provincial and Ministerial Achievement Award Cultivation Project (Grant No. 2017CG02).

Author Contributions

XD planned and wrote the paper, HW and JM performed the experiments, XZ statistically analyzed the data and made illustrations, and HL reviewed the paper.

References

Anderson JA, Y Ogihara, ME Sorrells, SD Tanksley (1992). Development of a chromosomal arm map for wheat based on RFLP markers. *Theor Appl Genet* 83:1035–1043

Clausen J (1926). Genetical and cytological investigations on *Viola tricolor* L. and *V. arvensis* MURR. *Hereditas* 8:1–156

Cordeiro GM, R Casu, CL McIntyre, JM Manners, RJ Henry (2001). Microsatellite markers from sugarcane (*Saccharum spp*.) ESTs cross transferable to erianthus and sorghum. *Plant Sci* 160:1115–1123

Culley TM, SJ Sbita, A Wick (2007). Population genetic effects of urban habitat fragmentation in the perennial herb *Viola pubescens* (Violaceae) using ISSR markers. *Ann Bot* 100:91–100

Decroocq V, MG Fave, L Hagen, L Bordenave, S Decroocq (2003). Development and transferability of apricot and grape EST microsatellite markers across taxa. *Theor Appl Genet* 106:912–922

Du X, X Zhu, X Li, J Mu, H Liu (2019a). Genetic diversity and structure in a germplasm collection of pansies using SRAP markers. *Intl J Agric Biol* 22:1244–1250

Du X, X Zhu, Y Yang, Y Wang, P Arens, H Liu (2019b). *De novo* transcriptome analysis of *Viola×wittrockiana* exposed to high temperature stress. *PloS One* 14; Article e0222344

Holton TA, JT Christopher, L McClure, N Harker, RJ Henry (2002). Identification and mapping of polymorphic SSR markers from expressed gene sequences of barley and wheat. *Mol Breed* 9:63–71

Jensen RJ (1989). NTSYS-PC numerical taxonomy and multivariate analysis system version 1.40. *Quat Rev Biol* 64:250–252

Kantety RV, ML Rota, DE Matthews, ME Sorrells (2002). Data mining for simple sequence repeats in expressed sequence tags from barley, maize, rice, sorghum and wheat. *Plant Mol Biol* 48:501–510

Ko MK, J Yang, YH Jin, CH Lee, BJ Oh (1998). Genetic relationships of Viola species evaluated by random amplified polymorphic DNA analysis. *J Pomol Hort Sci* 73:601–605

Li X, X Du, J Mu, H Liu (2015a). Genetic diversity analysis of 41 pansy germplasms based on RSAP markers. *Acta Bot Bor-Occidentalia Sin* 35:1989–1997

Li M, S Zhang, P Deng, X Hou, J Wang (2015b). Analysis on SSR information in transcriptome of onion and the polymorphism. *Acta Hortic Sin* 42:1103–1111

Liu Y, S Li, Y Wang, P Liu, W Han (2020). De novo assembly of the seed transcriptome and search for potential EST-SSR markers for an endangered, economically important tree species: *Elaeagnus mollis* Diels. *J For Res* 31:759–767

Li Y, J Liu, M Chen, Q Zhang, H Zhu, Q Wen (2018). SSR markers excavation and germplasm analysis using the transcriptome information of *Hibiscus esculentus*. *Acta Hortic Sin* 45:579–590

Liu Y, Y Geng, M Song, P Zhang, J Hou, W Wang (2019). Genetic structure and diversity of *glycyrrhiza* populations based on transcriptome SSR markers. *Plant Mol Biol Rep* 37:401–412

Manee MM, BM Al-Shomran, MB Al-Fageeh (2020). Genome-wide characterization of simple sequence repeats in Palmae genomes. *Genes Genome* 42:597–608

Nie G, L Tang, Y Zhang, L Huang, X Ma, X Cao, L Pan, X Zhang, X Zhang (2017). Development of SSR markers based on transcriptome sequencing and association analysis with drought tolerance in perennial grass *Miscanthus* from China. *Front Plant Sci* 8; Article 801

Peakall R, PE Smouse (2006). GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Mol Ecol Notes* 6:288–295

Peakall R, PE Smouse (2012). GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics* 28:2537–2539

Quardokus E (2000). PopGene. *Science* 288:458–459

Röder MS, V Korzun, K Wendehake, J Plaschke, NH Tixier, P Leroy, MW Ganal (1998). A microsatellite map of wheat. *Genetics* 149:2007–2023

Schloss S, S Mitchell, G White, R Kukatla, J Bowers, A Paterson, S Kresovich (2002). Characterization of RFLP probe sequences for gene discovery and SSR development in *Sorghum bicolor* (L.) Moench. *Theor Appl Genet* 105:912–920

Varshney RK, R Sigmund, A Bo¨rner, V Korzun, N Stein, ME Sorrells, P Langridge, A Graner (2005). Interspecific transferability and comparative mapping of barley EST-SSR markers in wheat, rye and rice. *Plant Sci* 168:195–202

Vemmos SN (2015). Characterisation of genetic relationships in pansy (*Viola wittrockiana*) inbred lines using morphological traits and RAPD markers. *J Pomol Hort Sci* 80:529–536

Wang J, M Bao (2007). Application of RAPD on analysis of genetic relationships between inbred lines and prediction of heterosis in pansy (*Viola wittrockiana*). *J Wuhan Bot Res* 25:19–23

Wang T, J Xu, X Zhang, L Zhao (2012). Genetic relationship of 43 cultivars of *Viola tricolor* and *Viola cornuta* using SRAP marker. *Sci Agric Sin* 45:496–502

Wei M (2016). *Development and Application of SSR Markers in Eggplant based on Transcriptome Sequencing*. Chinese Academy of Agricultural Sciences, Beijing, China

Wu J, C Cai, F Cheng (2014). Characterization and development of EST-SSR markers in tree peony using transcriptome sequences. *Mol Breed* 34:1853–1866

Xiao Y, L Zhou, W Xia, AS Mason, Y Yang, Z Ma, M Peng (2014). 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* 14; Article 384

Yan X, M Lu, H An (2015). Analysis on SSR information in transcriptome and development of molecular markers in *Rosa roxburghii*. *Acta Hort Sin* 42:341–349

Yockteng R, HEJ Ballard, G Mansion, I Dajoz, S Nadot (2003). Relationships among pansies (*Viola* section *Melanium*) investigated using ITS and ISSR markers. *Plant Syst Evol* 241:153–170

Yu J, TM Dake, S Singh, D Benscher, W Li, B Gill, ME Sorrells (2004). Development and mapping of EST-derived simple sequence repeat markers for hexaploid wheat. *Genome* 47:805–818

Zhang H, R Cong, M Wang, A Dong, H Xin, M Yi, H Guo (2018). Development of SSR molecular markers based on transcriptome sequencing of *Tagetes erecta*. *Acta Hort Sin* 45:159–167

Zhang Q, M Bao, X Lu, H Hu (2010). Research advances in breeding of pansy (*Viola×wittrockiana*). *Chin Bull Bot* 45:128–133