

三色堇转录组 SSR 分析及分子标记开发

杜晓华*, 杨雅萍, 朱小佩, 牛杨莉, 刘会超

(河南科技学院园艺园林学院, 河南省园艺植物资源利用与种质创新工程研究中心, 河南新乡 453003)

摘要: 为解决三色堇 (*Viola × wittrockiana*) 缺乏共显性 DNA 标记的问题, 利用 MISA 软件对其转录组测序获得的 167 576 条 unigene 进行筛选, 共检测出 23 791 个 SSR 位点, 出现频率为 14.20%, 平均分布距离为 6.75 kb。SSR 位点中主导类型为三核苷酸重复, 占总 SSR 的 28.31%; 其次是二核苷酸重复, 占总 SSR 的 12.86%。AG/CT 与 GAA/TTC 分别是二核苷酸与三核苷酸的优势重复单元, 分别占总 SSR 重复类型的 4.01% 和 1.85%。利用 Primer 3 共设计出 6 863 对 SSR 引物。随机选择 30 对引物进行 PCR 扩增, 其中 18 对引物可扩增出清晰、可重复条带, 并在 42 份三色堇资源中表现多态性。基于扩增的多态性 SSR 信息, 42 份三色堇种质资源的聚类结果与其遗传背景基本一致。本研究获得的 18 对 SSR 引物可为三色堇遗传图谱构建、功能基因定位等提供有力工具。

关键词: 三色堇; EST-SSR; 系统聚类

中图分类号: S 682

文献标志码: A

文章编号: 0513-353X (2019) 04-0797-10

Development of Genic-SSR Markers by Transcriptome Sequencing in *Viola × wittrockiana*

DU Xiaohua*, YANG Yaping, ZHU Xiaopei, NIU Yangli, and LIU Huichao

(School of Horticulture and Landscape Architecture, Henan Institute of Science and Technology, Henan Province Engineering Research Center of Horticultural Plant Resource Utilization and Germplasm Enhancement, Xinxiang, Henan 453003, China)

Abstract: To develop co-dominant DNA markers in pansy (*Viola × wittrockiana*), 167 576 unigenes from pansy transcriptomic database were screened by MISA software. A total of 23 791 simple sequence repeat (SSR) loci were detected in the unigenes with the frequency of 14.20%, and the average distribution distance of SSR was 6.75 kb. The predominant type of SSR was trinucleotide repetition amounting for 28.31% of total SSRs, followed by dinucleotide repetition accounting for 12.86%. AG/CT and GAA/TTC are the dominant elements in dinucleotide and trinucleotide repetitions, accounting for 4.01% and 1.85% of the total SSRs, respectively. Primer 3 was employed to design SSR primers and a total of 6 863 pairs of SSR primers were obtained. Thirty pairs of primers were randomly selected for PCR in pansies, in which 18 pairs of primers amplified clear and repeatable bands and showed polymorphism in 42 pansy germplasm accessions. Based on the amplified polymorphic SSR data, the dendrogram results were in accordance with the genetic backgrounds of 42 accessions. Therefore, a large amount of SSR markers can

收稿日期: 2018-11-08; **修回日期:** 2019-04-04

基金项目: 河南省科技攻关国际合作项目 (182102410029); 河南科技学院 2017 省部级成果奖培育项目 (2017CG02)

* E-mail: duxiaohua0124@sina.com

致谢: 邓衍明研究员对本文英文摘要进行了校正, 在此表示感谢!

be developed from transcriptome sequencing and employed in genetic linkage map construction and functional gene mapping in pansy.

Keywords: pansy; EST-SSR; hierarchical clustering

三色堇 (*Viola × wittrockiana*) 品种繁多, 花色丰富, 花期长, 耐寒, 是春、秋乃至温暖地区冬季的重要花坛花卉 (张其生 等, 2010)。当前开发和应用的 DNA 标记技术有 20 余种, 但在三色堇及堇菜属植物上应用的仅 4 种: RAPD (Ko et al., 1998; 王健和包满珠, 2007; Vemmos, 2015)、ISSR (Yockteng et al., 2003; Culley et al., 2007)、SRAP (王涛 等, 2012) 和 RSAP (李小梅 等, 2015)。这些均为显性标记, 不能区分杂合与纯合显性基因型, 揭示的遗传信息不够完整。简单重复序列 (simple sequence repeat, SSR) DNA 标记, 不但可区分杂合与纯合显性基因型, 而且基于 PCR 技术操作简便, 且多态性丰富, 已成为当前最为流行的 DNA 标记之一 (Li et al., 2014, 2018)。

由于 SSR 标记开发需要获知目标物种的部分基因组序列信息, 以往因测序费用偏高, 导致对基因组较大的非模式植物, 如三色堇 (Omka et al., 2011) 的 SSR 标记开发相对滞后。随着下一代测序技术 (Next Generation Sequencing, NGS) 不断发展, 测序成本持续下降。对基因组较大的植物, 可通过转录组测序, 高通量获取存在于转录基因区的 SSR 信息, 并进一步开发该物种的表达序列标签 SSR (Expressed Sequence Tag-SSR, EST-SSR) 标记, 且这些 SSR 多态性还可能与基因功能直接相关 (Eujayl et al., 2002), 具有很高的通用性及转化率。目前 EST-SSR 开发已在洋葱 (李满堂 等, 2015)、刺梨 (鲁敏 等, 2017)、万寿菊 (张华丽 等, 2018)、黄秋葵 (李永平等, 2018) 等多种植物上陆续开展。本研究中试图通过 Illumina 技术对热胁迫下三色堇转录组测序 (RNA-Seq), 开发三色堇 EST-SSR, 并在 42 份三色堇种质资源上对其有效性进行评价, 以期为三色堇种质资源研究、功能基因定位及标记辅助育种奠定基础。

1 材料与方法

1.1 转录组数据来源

转录组数据来源于本课题组 2017 年三色堇 DFM-16 材料叶片进行 Illumina 高通量深度测序的结果。测序时选取 6~8 片真叶期的三色堇苗, 2 个重复, 用 TaKaRa 公司 RNA 提取试剂盒提取 RNA, 反转录成 cDNA 后, 送北京诺禾致源生物信息科技有限公司进行转录组测序, 并通过 Trinity 方法 (Grabherr et al., 2011) 组装得到 167 576 条 unigene, 作为分析背景数据。

1.2 三色堇材料及其 DNA 提取

用于引物筛选和可用性评价的 42 份材料为本课题组选育的 34 份三色堇自交系、1 份三色堇杂交种、5 份角堇 (*V. cornuta*) 自交系和 2 份野生堇菜 (表 1)。试材于 2016 年 9 月 20 日播种, 2017 年 3 月移栽大田, 5 月采集成株期幼叶, 采用 SDS 法提取基因组 DNA (王关林和方宏筠, 2009)。

1.3 转录组 SSR 位点鉴别及 SSR 引物设计

使用软件 MISA 对 unigene 进行 SSR 位点分析, 单核苷酸、二核苷酸、三核苷酸、四核苷酸、五核苷酸和六核苷酸最少重复次数分别为 10、6、5、5、5 和 5 次。使用 Primer 3 设计 SSR 引物, 引物长度在 18~23 bp; 退火温度 (T_m) 55~65 °C, 上、下游引物 T_m 相差 < 2 °C; PCR 产物大小

80 ~ 300 bp; GC 含量 40% ~ 60%; 尽量避免出现发卡结构、引物二聚体和错配等。从中随机挑选 30 对 SSR 引物进行评价。引物由生工生物工程 (上海) 股份有限公司合成。

表 1 三色堇 SSR 多态性分析材料
Table 1 List of accessions used in assessing the level of SSR in *Viola × wittrockiana*

种属 Species	材料 Material	系谱来源 Pedigrees	编号 No.	名称 Name		
<i>Viola × wittrockiana</i>	自交系 Inbred line	早开花组合 Frühblühende Mischung	1	DFM-11-1-1		
			2	DFM-11-2-3		
			3	DFM-11-2-4-1		
			4	DFM-1-2-3-3		
			5	DFM-16-1-2-6		
			6	DFM-16-2-2		
			7	DFM-8-3-1-2		
			瑞士巨人阿尔卑斯湖 Schweizer Riesen Alpensee	8	DSRAB-1-2	
				9	DSRAB-1-2-4	
				10	DSRAB-1-4-2	
				瑞士巨人粒雪金 Schweizer Riesen Firmengold 229.10	11	DSRFY-1-1-2
					12	G10-1-1-1-3-3
					13	G10-1-3-1
					14	G10-1-3-1-4-2
				229.01	15	G10-1-1-1-3-2
					16	G1-1-1-1-1-4
					阿斯米尔巨人 Aalsmeerse Giants	17
		超级宾格黄色带斑 Matrix™ Yellow Blotch	21			MMYB-1-2
			22		MYC-1-1-3-4	
		超级宾格黄色 Matrix™ Yellow Clear	23		PXP-BT-4-1-1-1	
			24		PXP-BT-4-1-1	
		潘诺拉蓝色 Panola XP Blue True	25		RCO-1-3-4	
			26		RRB-1-3	
		超凡橙色 Clear Orange of Power Mini	27		RRB-2-7	
			28	RRB-3-1		
		革命者蓝色信号灯 Beacon Blue of Dynamite	29	XXL-G-1-1-2-3		
			30	XXL-G-1-1-3		
		超大花型 Extra flower size	31	XXL-G-1-1-7-4		
			黄色花 Yellow flower	32	EYO-1-2-1-4	
				33	EYO-1-2-1-5	
			白色花 White flower	34	EYO-1-1-4	
				35	EWO-2-1-1	
			浅蓝色花 Light blue	36	EWO-1-1-3	
				37	MW-1-1-1-1	
			<i>Viola cornuta</i>	杂交种 Hybrid	38	EWO × MW
					自交系 Inbred line	18
				19		JB-1-1-6
		20		JY-1-1-2		
		39		E01		
		40		08H		
		西山堇菜 <i>V. hancockii</i>		41		西山堇菜变种 <i>V. hancockii</i>
				早开堇菜 <i>V. prionantha</i>		42

1.4 PCR 扩增与数据统计

10 μL PCR 扩增反应体系包括: 2 × Es Taq Master Mix (康为世纪生物科技有限公司) 5 μL; 上、下游引物 (10 μmol · L⁻¹) 各 0.5 μL, 模板 DNA (40 ng · μL⁻¹) 2 μL, 超纯水 2 μL。

PCR 反应程序为: 95 °C 预变性 5 min; 然后 35 个循环, 每循环 95 °C 变性 40 s, 退火 (退火温度因不同引物而异) 30 s, 72 °C 延伸 30 s; 最后 72 °C 延伸 10 min, 4 °C 保存。扩增产物用 6% 非变

性聚丙烯酰胺凝胶检测，90 V 电压下电泳 4 h，银染显色。

电泳图上清晰条带记为“1”，同位置无带或不易分辨的弱带记为“0”，建立原始数据矩阵。采用 DPS 软件（7.55 版）按非加权组平均法（unweighted pair-group method with arithmetic means, UPGMA）法进行聚类分析。

2 结果与分析

2.1 三色堇转录组中的 SSR 位点的数量与分布

通过对三色堇转录组 167 576 条 unigene 序列搜索，共在 20 679 条 unigene 序列中发现符合条件的 23 791 个 SSR 位点，SSR 发生频率（含 SSR 的 unigene 数与总 unigene 数之比）为 12.34%，出现频率（检出的 SSR 个数与总 unigene 数之比）为 14.20%。其中 2 594 条 unigene 含有两个或两个以上 EST-SSR 位点，1 008 条 unigene 序列含有复合型 SSR 位点，平均 6.75 kb 出现 1 个 SSR。

三色堇的转录组 SSR 各种重复类型均有，但其出现频率存在较大差异（表 2），其中单核苷酸、二核苷酸和三核苷酸重复出现频率占优势，分别占总 SSR 的 57.36%、12.86%和 28.31%；四核苷酸、五核苷酸和六核苷酸重复类型数量很少，总计只占到总 SSR 的 1.47%。所有 SSR 中，10 次重复的 SSR 最多，占 28.92%；其次为 5 次重复的，占 20.62%；6 次重复的占 12.92%。

三色堇转录组 SSR 重复单元的重复次数分布在 5~39 次之间，其中 5~10 次重复的 SSR 位点有 16 713 个，占总数的 70.25%；11~20 次重复的有 6 624 个，占 27.84%；20 次重复以上的有 454 个，只占 1.91%。三色堇转录组 SSR 的长度从 10~60 bp 不等，分布在 10~15 bp 的约有 18 502 个，占整个 EST-SSR 的 77.77%；16~20 bp 的有 3 857 个，占 16.21%；21~25 bp 的有 1 408 个，占 5.92%；长度大于 25 bp 的有 24 个，占 1.01%。

表 2 三色堇 EST-SSR 的类型、数量及分布频率

Table 2 Type, number and frequency of EST-SSRs in *V. × wittrockiana*

重复类型 Repeat type	重复次数 Repeat number							总计 Total	百分比/% Percentage
	5	6	7	8	9	10	> 10		
单核苷酸 Mononucleotide						6 733	6 913	13 646	57.36
二核苷酸 Dinucleotide		1 598	621	326	207	146	161	3 059	12.86
三核苷酸 Trinucleotide	4 614	1 429	654	35	1		2	6 735	28.31
四核苷酸 Tetranucleotide	226	43					1	270	1.13
五核苷酸 Pentanucleotide	62		1		2		1	66	0.28
六核苷酸 Hexanucleotide	6	4		4		1		15	0.06
总计 Total	4 908	3 074	1 276	365	210	6 880	7 078	23 791	100.00
百分比% Percentage	20.62	12.92	5.36	1.53	0.88	28.92	29.75		

2.2 转录组 SSR 基序重复类型和频率特征

单核苷酸重复会因均聚物的存在导致测序质量不高（Gilles et al., 2011），因此本试验中不予选用。从三色堇转录组 SSR 核苷酸基序的类型来看，其 23 791 个 SSR 位点包含 149 种重复基序，二、三、四、五、六核苷酸重复各有 8、30、69、31 和 11 种；从 SSR 重复频率来看，二核苷酸主要分布在 6~11 次，三核苷酸主要分布在 5~7 次，四核苷酸主要分布在 5~6 次，五核苷酸和六核苷酸主要分布在 5 次。

从三色堇 SSR 基序的分布频率来看，三核苷酸重复基序最多，占总 SSR 的 28.32%；在三核苷酸重复基序中又以 GAA/TTC、ATC/GAT、TCA/TGA 和 CTC/GAG 最多，分别占三核苷酸 SSR 的

6.55%、6.34%、6.31%和 6.09%，占总 SSR 的 1.85%、1.80%、1.79%和 1.72%。其次是二核苷酸重复基序，占总 SSR 的 12.86%，其中以 AG/CT 和 GA/TC 为主，占二核苷酸 SSR 的 31.15%和 21.22%，占总 SSR 的 4.01%和 2.73%。四、五、六核苷酸重复基序出现频率均较低，分别占总 SSR 的 1.13%、0.28%和 0.06%（表 3）。

四核苷酸中以 AAAG/CTTT、AAAT/ATTT 最多，分别占四核苷酸 SSR 的 6.67%和 5.19%。五核苷酸中以 ATCTA、CCTCT 最多，各占五核苷酸 SSR 的 9.09%。六核苷酸中以 GGTATG 最多，占六核苷酸 SSR 的 20%。

表 3 三色堇转录组中 SSR 基序的分布

Table 3 Distribution of SSR motif in transcriptome of *V. × wittrockiana*

重复类型 Repeat type	基序 Motif	重复次数 Repeat number	重复类型 Repeat type	基序 Motif	重复次数 Repeat number
单核苷酸 Mononucleotide		13 646		ATG/CAT	254
二核苷酸 Dinucleotide		3 059		CAA/TTG	287
	AC/GT	256		CAC/GTG	178
	AG/CT	953		CAG/CTG	325
	CA/TG	296		CCA/TGG	261
	GA/TC	649		CCG/CGG	78
	AT	443		CCT/AGG	292
	TA	449		CGA/TCG	20
	CG	8		CTA/TAG	57
	GC	5		CTC/GAG	410
三核苷酸 Trinucleotide		6 737		GAA/TTC	441
	AAC/GTT	166		GAC/GTC	48
	AAG/CTT	270		GCA/TGC	329
	AAT/ATT	202		GCC/GGC	60
	ACA/TGT	238		GCG/CGC	79
	ACC/GGT	222		GGA/TCC	334
	ACG/CGT	14		GTA/TAC	116
	ACT/AGT	128		TAA/TTA	169
	AGA/TCT	396		TCA/TGA	425
	AGC/GCT	402	四核苷酸 Tetranucleotide		270
	ATA/TAT	107	五核苷酸 Pentanucleotide		66
	ATC/GAT	427	六核苷酸 Hexnucleotide		15

2.3 三色堇转录组 SSR 引物设计与筛选

除去复合 SSR，对 9 228 条二核苷酸 ~ 六核苷酸重复的 SSR 位点进行引物设计，共获得 SSR 位点特异引物 6 863 对，占总 SSR 位点的 74.37%。为验证其引物的有效性，随机挑选了 30 对 SSR 引物，用 4 份三色堇自交系 G10-1-3-1-4B、MYC-1-1-3-1、RRB-1C 和 PXP-BT-D-1-1-6E 的基因组 DNA 为模板进行 PCR 扩增和引物筛选。结果（表 4）表明，30 对引物中的 18 对可有效扩增，扩增率 60%。

2.4 多态性分析

选取 42 份三色堇资源对 18 对 EST-SSR 引物的扩增多态性进行评价。结果表明，18 对引物均呈现出多态性。18 对引物在 42 份三色堇资源上扩增共得到 251 条多态性片段，各引物产生的多态性片段数在 5 ~ 17 条之间，平均每对引物产生 13.94 个多态性片段。图 1 为引物 24 号的扩增情况，其他引物的扩增及多态性情况见表 4。基于扩增的多态性 EST-SSR 进行的聚类分析（图 2），将 42 份材料分为 2 大类。第一类为三色堇和角堇，均属于美丽堇菜亚属；第二类为西山堇菜变种和早开堇菜，均属于堇菜亚属。在第一类中，大多源于相同亲本的自交选系聚在一起。

以上结果显示，基于转录组开发的 EST-SSR 数据能较好地反映三色堇种质资源间的遗传关系。

表 4 18 对 SSR 引物及其扩增情况

Table 4 18 pairs of primers developed from *V. × wittrockiana* and their amplification

编号 No.	正反向引物序列 (5'→3') Forward, reverse primer sequence (5'→3')	重复单元 Repeat motif	预期产物/bp Expected size	扩增带数 Number of bands	多态性带数 Number of polymorphic bands
1	GAAACTATCCACCACCGCCA, TCGGGAATACGGTGGTTGTG	(CCA) ₅	167	13	13
2	GGACTGCTGCCTCATCAAG, CCAGGTCACAATTCCAAGTGC	(AAG) ₅	111	17	17
3	CGAAGAGCTTGAAGGCCAA, TGATGCTCCGAAACTAACG	(CAA) ₅	239	14	14
5	CCCAAACCTTAACCCGAGCT, GATACGGTTGGAGTGGACGG	(CAC) ₅	224	18	18
9	CAGGCTGTTGGTTGTGAC, CCCCTCCCACCTTTCCTTTC	(GGC) ₅	141	16	16
11	CCACTACCAACAAACCCCA, TCCTCAACCTCTGTCTAGA	(TC) ₆	238	14	14
12	GAGGGTCGTTTCAAATGGC, GCAAATGGGTCGTCGTCAAC	(CAG) ₅	185	16	16
16	TGTCTCCGGCTAAAACCACC, CGCAGTCTCCGTCGATTACA	(CCG) ₅	170	10	10
17	GCGATCGAGAAATGAGGCT, CGCTACCCATCATCTGTCTCC	(TGC) ₅	286	11	11
18	TTCCACCTCAAACCTCGG, TGTTGATGCTGCAGGGGTA	(CCA) ₅	289	12	12
19	AAGTGCCATGCTTGTCCCTT, GCTGTCGCTGCTAAACTTGC	(GCT) ₅	261	12	12
21	GATCCCACAGCGTTTACCCA, GCCGCGTTACCATAGCTAGT	(AGC) ₅	224	19	19
23	TGCCACCTGATCCATTGCA, TGTGGCTGTTGTTGTGCTG	(AGG) ₅	203	5	5
24	GGTAGGAGACGCTGGGAAAC, GCCGCGTTACCATAGCTAGT	(AGC) ₅	288	19	19
25	TGTGACGACTGAAAAGGCCA, GCACAAACAACATAAGGGCGA	(GAA) ₅	267	11	11
26	CCGCTACTCCACTGAACTC, ACATGGAAGAGGAGCAAGCA	(TCA) ₅	265	13	13
27	GCTTATGTGAGTGTATGGCG, ACCTCTTCTGCACACCACC	(GCT) ₈	137	14	14
30	ACCGCAAACCAAGCAAACAA, TGAGGATGAAGGGGATGGGA	(CAT) ₆	169	17	17

编号 No.	对应 unigene 功能预测 Responded unigene function prediction	NR E-value
1	羧肽酶 Y Carboxypeptidase Y	5.2E-06
2	线粒体输入受体 TOM20-2 家族蛋白 Mitochondrial import receptor TOM20-2 family protein	3.2E-70
3	7-脱氧糖苷转移酶类 7-deoxyloganetic acid glucosyltransferase-like	5.3E-150
5	无特征蛋白 Uncharacterized protein	2.1E-50
9	无特征蛋白 Uncharacterized protein	1.8E-23
11	假设蛋白 POPTR Hypothetical protein POPTR	8.4E-70
12	转录因子 bHLH63 异构体 X1 Transcription factor bHLH63 isoform X1	1E-129
16	催化剂 Catalytic	8.3E-99
17	脂氧合酶 Lipoxigenase	3.5E-54
18	Pumilio 同源物 1 类 Pumilio homolog 1-like	0
19	核 mRNA 前体调节类蛋白 2 内域 Regulation of nuclear pre-mRNA domain-containing protein 2-like	4.4E-168
21	B3 域内转录因子 NGA1 类异构体 X1 B3 domain-containing transcription factor NGA1-like isoform X1	3.2E-51
23	转录因子 bHLH91 类 Transcription factor bHLH91-like	1.1E-79
24	B3 域内转录因子 NGA1 类异型 X1 B3 domain-containing transcription factor NGA1-like isoform X1	1.4E-52
25	苯丙氨酰 tRNA 合成酶 β 链 Phenylalanyl-tRNA synthetase beta chain	6.9E-91
26	小 RNA 2'-O-甲基转移酶类 Small RNA 2'-O-methyltransferase-like	0
27	含 aha1 域的家庭蛋白 aha1 domain-containing family protein	1E-152
30	假设蛋白 POPTR Hypothetical protein POPTR	1.1E-117

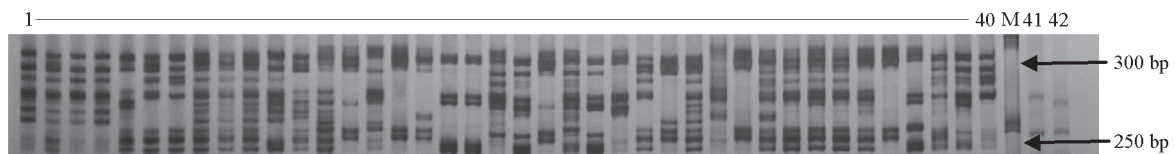


图 1 引物 24 号在 42 份三色堇资源中的扩增图谱

M: Marker; 材料编号同表 1。

Fig. 1 Profile of amplification in 42 *V. × wittrockiana* germplasms by primer No. 24

M stands for molecular weight marker; Plant material codes correspond to those in Table 1.

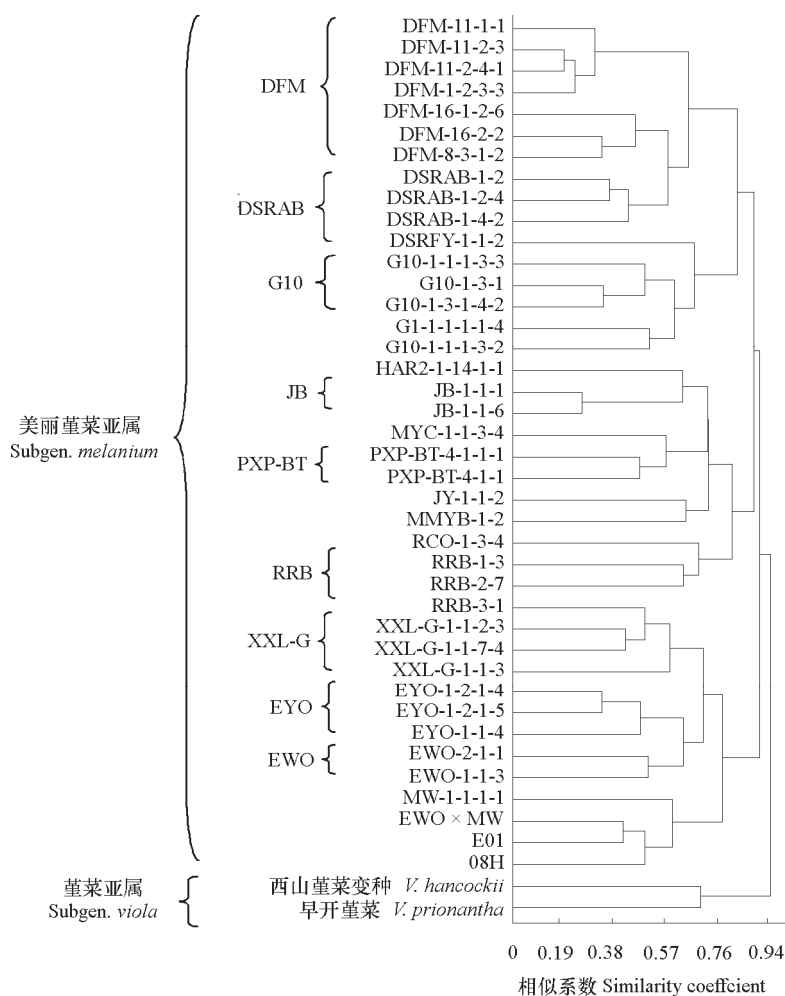


图2 供试三色堇种质资源的 UPGMA 聚类图
Fig. 2 Dendrogram for *V. × wittrockiana* germplasms by UPGMA

3 讨论

三色堇经种间杂交和多倍化培育,属于基因组较大的非模式植物(张其生 等,2010; Omka et al., 2011),其共显性标记开发较为滞后。本研究中以热胁迫下的三色堇 DFM-16 叶片转录组数据为基础,利用 Trinity 软件拼接组装成 167 576 条 unigene,从中检测出 23 791 个 SSR 位点,出现频率为 14.20%。因 SSR 出现频率受到 SSR 搜索阈值影响,尤其是单核苷酸重复数量多,是否统计会明显影响结果。为便于与其他物种相比较,这里将以往研究分为两组,一组为统计单核苷酸重复,另一组不统计。统计单核苷酸重复时,三色堇 SSR 出现频率 14.20%,高于洋葱(5.57%,李满堂 等,2015)、楠木(9.90%,时小东 等,2016)、油棕(11.26%,Xiao et al.,2014),接近芒草(14.44%,Nie et al.,2017),但低于茄子(18.32%,魏明明,2016)。不统计单核苷酸重复时,三色堇的 SSR 分布频率为 6.05%,高于玉米(1.5%)、大麦(3.4%)、水稻(4.7%,Kantety et al.,2002)、黄瓜(4.03%,Guo et al.,2010)、南方红豆杉(2.24%,李炎林 等,2014)、红松(4.24%,张振 等,2015),与花生(6.63%,Zhang et al.,2012)、辣椒(7.83%,刘峰 等,2012)、木豆(7.6%,Dutta et al.,2011)

相近,但低于刺梨(20.37%,鄢秀芹等,2015)、橡树(27.61%,An et al.,2016)和早实枳(26.88%,龙荡,2014)。如果忽略 EST 数据量及来源等对以上差异的影响(鄢秀芹等,2015),这种差异可能反映了物种间真实的 SSR 信息差异,表明三色堇转录组 SSR 较为丰富,是遗传作图的理想标记之一。

本研究中发现三色堇 SSR 类型以三核苷酸(28.31%)和二核苷酸(12.86%)为主,三核苷酸出现频率是二核苷酸(12.86%)2倍以上,这与前人对其他植物的研究(Kota et al.,2001;Kantety et al.,2002;孙清明等,2011;李满堂等,2015;张振等,2015;Nie et al.,2017)结果相符,这可能是因为密码子是以三核苷酸为1个功能单位的,其位移对表达基因的阅读框造成的影响较小,因此在 EST 序列中出现最多的类型是三核苷酸重复 SSR(张振等,2015)。而花生(Zhang et al.,2012)、早实枳(龙荡,2014)、荔枝(孙清明等,2011)等出现最多的是二核苷酸重复 SSR,这差异可能与植物自身 EST-SSR 特点、数量以及 EST 数据来源等密切相关。三色堇中三核苷酸重复基序中以 GAA/TTC、ATC/GAT、TCA/TGA 和 CTC/GAG 最多,这与辣椒(AAC/GTT)(刘峰等,2012)、水稻(AGG/TCC)(Kantety et al.,2002)、洋葱(AAG/CTT)(李满堂等,2015)和花生(AAG/CTT)(Zhang et al.,2012)等不同。在二核苷酸重复基元中频率最高的是 AG/CT,这与大麦、小麦、玉米、高粱、水稻(Kantety et al.,2002)、茄子(魏明明,2016)、花生(Zhang et al.,2012)、芒草(Nie et al.,2017)、刺梨(鄢秀芹等,2015)、南方红豆杉(李炎林等,2014)、早实枳(龙荡,2014)等多种植物中的研究结果一致。

本试验中选用 30 对 SSR 引物在三色堇种质材料上有 18 对实现了有效扩增,可扩增率为 60%,这与洋葱(60%,李满堂等,2015)、刺梨(54.76%,鄢秀芹等,2015)、南方红豆杉(53.23%,李炎林等,2014)相近,低于芒草(75.3%,Nie et al.,2017)和茄子(81.2%,魏明明,2016)。根据 EST 设计的引物,部分引物不能成功扩增的原因可能是侧翼引物延伸穿过具有大内含子或嵌合体的剪接位点。本研究中同一份材料上出现多个 SSR 位点,这可能与三色堇为多倍体植物有关(张其生等,2010)。采用 18 对引物对 42 份三色堇资源分析的系统聚类结果清晰地将其分为美丽堇菜亚属和堇菜亚属 2 大类,且属于美丽堇菜亚属的源于自交选系基本按亲本来源都聚在一起,说明聚类结果与材料的遗传背景基本一致,表明利用三色堇转录组数据开发的 SSR 标记具有很好的实用性。聚类结果并没有将三色堇与角堇单独分开,或与角堇作为亲本之一曾参与三色堇杂交过程有关(张其生,2009)。本研究中批量开发的三色堇 SSR 标记为首次报道的三色堇共显性 DNA 标记,将为三色堇的目标性状基因定位及分子标记辅助育种奠定基础。

References

- An M, Deng M, Zheng S, Song Y. 2016. *De novo* transcriptome assembly and development of SSR markers of oaks *Quercus austrocochinchinensis*, and *Q. kerrii*, (Fagaceae). *Tree Genetics & Genomes*, 12 (6): 103.
- Culley T M, Sbita S J, Wick A. 2007. Population genetic effects of urban habitat fragmentation in the perennial herb *Viola pubescens* (Violaceae) using ISSR markers. *Annals of Botany*, 100 (1): 91 - 100.
- Dutta S, Kumawat G, Singh B P, Gupta D K, Singh S, Dogra V, Gaikwad K, Sharma T R, Raje R S, Bandhopadhyaya T K, Datta S, Singh M N, Bashasab F, Kulwal P, Wanjari K B, Varshney K R, Cook D R, Singh N K. 2011. Development of genic-SSR markers by deep transcriptome sequencing in pigeonpea [*Cajanus cajan* (L.) Millspaugh]. *BMC Plant Biology*, 11: 17.
- Eujayl I, Sorrells M, Banm M, Wolters P, Powell W. 2002. Isolation of EST-derived microsatellite markers for genotyping the A and B genomes of wheat. *Theor Appl Genet*, 104 (2): 399 - 407.
- Guo S, Zheng Y, Joung J G, Liu S Q, Zhang Z H, Crasta O R, Sobral B W, Xu Y, Huang S W, Fei Z G. 2010. Transcriptome sequencing and comparative analysis of cucumber flowers with different sex types. *BMC Genomics*, 11 (1): 384.

- Gilles A, Meglecz E, Pech N, Ferreira S, Malausa T, Martin J F. 2011. Accuracy and quality assessment of 454 GS-FLX Titanium pyrosequencing. *BMC Genomics*, 12: 245.
- Grabherr M G, Haas B J, Yassour M, Levin J Z, Thompson D A, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q D, Chen Z H, Mauceli E, Hacohen N, Gnirke A, Rhind N, Palma F D, Birren B W, Nusbaum C, Lindblad-Toh K, Friedman N, Regev A. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, 29: 644 - 652.
- Kantety R V, Rota M L, Matthews D E, Sorrells M E. 2002. Data mining for simple sequence repeats in expressed sequence tags from barley, maize, rice, sorghum and wheat. *Plant Molecular Biology*, 48: 501 - 510.
- Ko M K, Yang J, Jin Y H, Lee C H, Oh B J. 1998. Genetic relationships of *Viola* species evaluated by random amplified polymorphic DNA analysis. *Journal of Pomology & Horticultural Science*, 73 (5): 601 - 605.
- Kota R, Varshney R K, Thiel T, Dehmer K J, Graner A. 2001. Generation and comparison of EST-derived SSRs and SNPs in barley (*Hordeum vulgare* L.). *Hereditas*, 135 (2 - 3): 145 - 151.
- Li Man-tang, Zhang Shi-lin, Deng Peng, Hou Xi-lin, Wang Jian-jun. 2015. Analysis on SSR information in transcriptome of onion and the polymorphism. *Acta Horticulturae Sinica*, 42 (6): 1103 - 1111. (in Chinese)
李满堂, 张仕林, 邓鹏, 侯喜林, 王建军. 2015. 洋葱转录组 SSR 信息分析及其多态性研究. *园艺学报*, 42 (6): 1103 - 1111.
- Li X, Gao W, Guo H, Zhang X, Fang D D, Lin Z. 2014. Development of EST-based SNP and InDel markers and their utilization in tetraploid cotton genetic mapping. *BMC Genomics*, 15: 1046 - 1047.
- Li X, Yu H L, Li Z Y, Liu X P, Fang Z Y, Liu Y M, Yang L M, Zhuang M, Lv H H, Zhang Y Y. 2018. Heterotic group classification of 63 inbred lines and hybrid purity identification by using SSR markers in winter cabbage (*Brassica oleracea* L. var. *capitata*). *Horticultural Plant Journal*, 4 (4): 158 - 164.
- Li Xiao-mei, Du Xiao-hua, Mu Jin-yan, Liu Hui-chao. 2015. Genetic diversity analysis of 41 pansy germplasms based on RSAP markers. *Acta Botanica Boreali-Occidentalia Sinica*, 35 (10): 1989 - 1997. (in Chinese)
李小明, 杜晓华, 穆金燕, 刘会超. 2015. 基于 RSAP 标记的大花三色堇遗传多样性分析. *西北植物学报*, 35 (10): 1989 - 1997.
- Li Yan-lin, Yang Xing-xing, Zhang Jia-yin, Huang San-wen, Xiong Xing-yao. 2014. Studies on SSR molecular markers based on transcriptome of *Taxus chinensis* var. *mairei*. *Acta Horticulturae Sinica*, 41 (4): 735 - 745. (in Chinese)
李炎林, 杨星星, 张家银, 黄三文, 熊兴耀. 2014. 南方红豆杉转录组 SSR 挖掘及分子标记的研究. *园艺学报*, 41 (4): 735 - 745.
- Li Yong-ping, Liu Jian-ting, Chen Min-dong, Zhang Qian-rong, Zhu Hai-sheng, Wen Qing-fang. 2018. SSR markers excavation and germplasm analysis using the transcriptome information of *Hibiscus esculentus*. *Acta Horticulturae Sinica*, 45 (3): 579 - 590. (in Chinese)
李永平, 刘建汀, 陈敏东, 张前荣, 朱海生, 温庆放. 2018. 利用黄秋葵转录组信息挖掘 SSR 标记及用于种质分析. *园艺学报*, 45 (3): 579 - 590.
- Liu Feng, Wang Yun-sheng, Tian Xue-liang, Mao Zhen-chuang, Zhou Xue-xiao, Xie Bing-yan. 2012. SSR mining in pepper (*Capsicum annuum* L.) transcriptome and the polymorphism analysis. *Acta Horticulturae Sinica*, 39 (1): 168 - 174. (in Chinese)
刘峰, 王运生, 田雪亮, 茆振川, 邹学校, 谢丙炎. 2012. 辣椒转录组 SSR 挖掘及其多态性分析. *园艺学报*, 39 (1): 168 - 174.
- Long Dang. 2014. Large scale development of SSR markers based on transcriptome sequencing of precocious trifoliate orange[M. D. Dissertation]. Wuhan: Huazhong Agricultural University. (in Chinese)
龙荡. 2014. 基于转录组测序大规模开发早实枳标记[硕士论文]. 武汉: 华中农业大学.
- Lu Min, Yan Xiuqin, Bai Jing, Zhang Huaishan, Wang Daoping, An Huaming. 2017. Construction of core collection in wild *Rosa roxburghii* from Guizhou Province using EST-SSR markers and fruits quality traits. *Acta Horticulturae Sinica*, 44 (8): 1486 - 1495. (in Chinese)
鲁敏, 鄢秀芹, 白静, 张怀山, 王道平, 安华明. 2017. 基于 EST-SSR 标记与果实品质性状的贵州野生刺梨核心种质构建. *园艺学报*, 44 (8): 1486 - 1495.
- Nie G, Tang L, Zhang Y, Huang L, Ma X, Cao X, Pan L, Zhang X, Zhang X. 2017. Development of SSR markers based on transcriptome sequencing and association analysis with drought tolerance in perennial grass *Miscanthus* from China. *Frontiers in Plant Science*, 8: 801.
- Omka A S, Siwińska D, Wolny E B, Kellner K, Kuta E B. 2011. Influence of a heavy-metal-polluted environment on *Viola tricolor* genome size and chromosome number. *Acta Biologica Cracoviensia*, 53 (1): 7 - 15.
- Shi Xiao-dong, Zhu Xue-hui, Sheng Yu-zhen, Zhuang Guo-qing, Chen Fang. 2016. Development of SSR markers based on transcriptome sequence

- of *Phoebe zhennan*. *Scientia Silvae Sinicae*, 52 (11): 71 - 78. (in Chinese)
- 时小东, 朱学慧, 盛玉珍, 庄国庆, 陈放. 2016. 基于转录组序列的楠木 SSR 分子标记开发. *林业科学*, 52 (11): 71 - 78.
- Sun Qing-ming, Ma Wen-chao, Ma Shuai-peng, Zhao Jun-sheng, Bai Li-jun, Chen Jie-zhen, Cai Chang-he, Xiang Xu, Ou Liang-xi. 2011. Characteristics of SSRs derived from ESTs and development of EST-SSR markers in litch (*Litchi chinensis* Sonn). *Scientia Agricultura Sinica*, 44 (19): 4037 - 4049. (in Chinese)
- 孙清明, 马文朝, 马帅鹏, 赵俊生, 白丽军, 陈洁珍, 蔡长河, 向旭, 欧良喜. 2011. 荔枝 EST 资源的 SSR 信息分析及 EST-SSR 标记开发. *中国农业科学*, 44 (19): 4037 - 4049.
- Vemmos S N. 2015. Characterisation of genetic relationships in pansy (*Viola wittrockiana*) inbred lines using morphological traits and RAPD markers. *Journal of Pomology & Horticultural Science*, 80 (5): 529 - 536.
- Wang Guan-lin, Fang Hong-jun. 2009. *Plant genetic engineering*. Beijing: Science Press. (in Chinese)
- 王关林, 方宏筠. 2009. 植物基因工程. 北京: 科学出版社.
- Wang Jian, Bao Man-zhu. 2007. Application of RAPD on analysis of genetic relationships between inbred lines and prediction of heterosis in pansy (*Viola wittrockiana*). *Journal of Wuhan Botanical Research*, 25 (1): 19 - 23. (in Chinese)
- 王健, 包满珠. 2007. RAPD 在三色堇 (*Viola wittrockiana*) 自交系遗传关系研究及杂种优势预测中的应用. *武汉植物学研究*, 25 (1): 19 - 23.
- Wang Tao, Xu Jin, Zhang Xi-xi, Zhao Liang-jun. 2012. Genetic relationship of 43 cultivars of *Viola tricolor* and *Viola cornuta* using SRAP marker. *Scientia Agricultura Sinica*, 45 (3): 496 - 502. (in Chinese)
- 王涛, 徐进, 张西西, 赵梁军. 2012. 43 份三色堇、角堇材料亲缘关系的 SRAP 分析. *中国农业科学*, 45 (3): 496 - 502.
- Wei Ming-ming. 2016. Development and application of SSR markers in eggplant based on transcriptome sequencing [M. D. Dissertation]. Beijing: Chinese Academy of Agricultural Sciences. (in Chinese)
- 魏明明. 2016. 基于转录组测序的茄子 SSR 标记开发及应用 [硕士论文]. 北京: 中国农业科学院.
- Xiao Y, Zhou L, Xia W, Mason A S, Yang Y, Ma Z, Peng M. 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 Biology*, 14: 384.
- Yan Xiu-qin, Lu Min, An Hua-ming. 2015. Analysis on SSR information in transcriptome and development of molecular markers in *Rosa roxburghii*. *Acta Horticulturae Sinica*, 42 (2): 341 - 349. (in Chinese)
- 鄢秀芹, 鲁敏, 安华明. 2015. 刺梨转录组 SSR 信息分析及其分子标记开发. *园艺学报*, 42 (2): 341 - 349.
- Yockteng R, Ballard H E J, Mansion G, Dajoz I, Nadot S. 2003. Relationships among pansies (*Viola*, section *Melanium*) investigated using ITS and ISSR markers. *Plant Systematics & Evolution*, 241 (3 - 4): 153 - 170.
- Zhang Huali, Cong Richen, Wang Maoliang, Dong Aixiang, Xin Haibo, Yi Mingfang, Guo Hua. 2018. Development of SSR molecular markers based on transcriptome sequencing of *Tagetes erecta*. *Acta Horticulturae Sinica*, 45 (1): 159 - 167. (in Chinese)
- 张华丽, 丛日晨, 王茂良, 董爱香, 辛海波, 义鸣放, 郭华. 2018. 基于万寿菊转录组测序的 SSR 标记开发. *园艺学报*, 45 (1): 159 - 167.
- Zhang J, Liang S, Duan J, Wang J, Chen S, Cheng Z, Zhang Q, Liang X, Li Y. 2012. *De novo* assembly and characterisation of the transcriptome during seed development, and generation of genic-SSR markers in peanut (*Arachis hypogaea* L.). *BMC Genomics*, 13 (1): 90.
- Zhang Qi-sheng. 2009. Studies on inheritance of flower color and flower color patten and tissue culture of pansy (*Viola × wittrockiana*) and *Viola (V. cornuta)* [M. D. Dissertation]. Wuhan: Huazhong Agricultural University. (in Chinese)
- 张其生. 2009. 三色堇与角堇花色、花斑遗传规律及组织培养的研究 [硕士论文]. 武汉: 华中农业大学.
- Zhang Qi-Sheng, Bao Man-Zhu, Lu Xing-Xia, Hu Hui-Rong. 2010. Research advances in breeding of pansy (*Viola × wittrockiana*). *Chinese Bulletin of Botany*, 45 (1): 128 - 133. (in Chinese)
- 张其生, 包满珠, 卢兴霞, 胡惠蓉. 2010. 大花三色堇育种研究进展. *植物学报*, 45 (1): 128 - 133.
- Zhang Zhen, Zhang Han-guo, Mo Chi, Zhang Lei. 2015. Transcriptome sequencing analysis and development of EST-SSR markers for *Pinus koraiensis*. *Scientia Silvae Sinicae*, 51 (8): 114 - 120. (in Chinese)
- 张振, 张含国, 莫迟, 张磊. 2015. 红松转录组 SSR 分析及 EST-SSR 标记开发. *林业科学*, 51 (8): 114 - 120.