

# 韭菜全长转录组 SSR 信息分析及分子标记开发

李延龙, 张华敏\*, 崔蕴刚, 陈建华, 吕爱芹, 李纪军, 李艺潇

(平顶山市农业科学院, 河南省韭菜工程技术研究中心, 河南平顶山 467001)

**摘要:** 利用 MISA 软件筛选韭菜 (*Allium tuberosum*) 全长转录组测序获得的 49 876 条 (大于 500 bp) 转录本序列 (89.44 Mb), 共检测出 13 111 个 SSR 位点, 分布在 10 332 条转录本中, SSR 出现频率为 26.29%, 平均分布距离为 6.82 kb。在搜索的 6 种 SSR 位点类型中, 1~3 个核苷酸重复类型占主要地位, 占总 SSR 位点数的 98.74%, 其中单核苷酸、二核苷酸和三核苷酸重复类型分别为 66.55%、12.52%、19.67%。发现重复序列的基序有 103 个, 其中二核苷酸和三核苷酸重复类型中出现频率高的基序有 AC/GT (2.54%)、CA/TG (2.29%)、GAA/TTC (1.85%) 和 AAG/CTT (1.60%)。重复序列的长度在 10~289 bp 之间, 其中小于 20 bp 的有 11 128 个 (84.88%), 大于 20 bp 的有 1 983 个 (15.12%)。根据这些 SSR 位点, 共设计出 3 311 对 SSR 引物, 在随机选取的 208 对引物中有 164 对 (78.85%) 能进行有效扩增, 其中 39 对引物在 24 份韭菜种质资源中表现出多态性。利用多态性引物 SSR41, 鉴定出 ‘马蔺韭’ × ‘791’ 后代中的 6 株有性生殖后代。以上结果表明, 基于韭菜全长转录组测序开发的 SSR 标记可以为韭菜的遗传多样性分析、种质资源鉴定和有性生殖后代筛选提供充足可靠的标记来源。

**关键词:** 韭菜; 全长转录组; SSR; 无融合生殖

**中图分类号:** S 633.3

**文献标志码:** A

**文章编号:** 0513-353X (2020) 04-0759-10

## Analysis on SSR Information in Full-length Transcriptome and Development of Molecular Markers in *Allium tuberosum*

LI Yanlong, ZHANG Huamin\*, CUI Yungang, CHEN Jianhua, LÜ Aiqin, LI Jijun, and LI Yixiao

(Pingdingshan Academy of Agricultural Sciences, Henan Chinese Chive Engineering Technology Research Center, Pingdingshan, Henan 467001, China)

**Abstract:** Based on 49 876 transcripts (> 500 bp, 89.44 Mb) obtained by full-length transcriptome sequencing of *Allium tuberosum*, 13 111 SSR loci distributed in 10 332 transcripts were detected by MISA software. The frequency of these SSR loci was 26.29% and the average distribution distance was 6.82 kb. Among the six types of SSR loci searched, 1 – 3 nucleotide repeats were dominant and accounted for 98.74% of the total SSR loci, of which mono-, di- and tri-nucleotide repeats were 66.55%, 12.52%, and 19.67%, respectively. There are 103 detected repeat motifs, of which AC/GT (2.54%), CA/TG (2.29%), GAA/TTC (1.85%) and AAG/CTT (1.60%) were high frequency repeat motifs in dinucleotide and trinucleotide repeats. The lengths of repeat nucleotide sequences for all SSR loci ranged from 10 bp to 289 bp, of which 11 128 (84.88%) were less than 20 bp, and 1 983 (15.12%) were larger than 20 bp.

**收稿日期:** 2019 – 11 – 07; **修回日期:** 2019 – 12 – 09

**基金项目:** 国家现代农业产业技术体系建设专项资金项目 (CARS-24-A-11)

\* 通信作者 Author for correspondence (E-mail: hmzhang111@126.com)

According to these SSR loci, a total of 3 311 pairs of SSR primers were designed. Two hundred and eight pairs of primers were randomly selected for amplification tests, 164 pairs of primers (78.85%) could amplify clear band effectively, and 39 pairs of primers showed polymorphism in 24 *Allium tuberosum* germplasm resources. Using the polymorphism primer SSR41, six sexual reproductive progenies from 'Malinjiu' × '791' were successfully identified. These results indicated that the SSR markers developed from the full-length transcriptome sequencing of *Allium tuberosum* could provide sufficient and reliable markers for analysis of genetic diversity, identification of germplasm resources and screening of sexual reproduction progeny of *Allium tuberosum*.

**Keywords:** *Allium tuberosum*; full-length transcriptome; SSR; apomixis

简单重复序列 (simple sequence repeat, SSR) 是由 1~6 个碱基组成的基元串联重复而成的 DNA 序列 (Tautz, 1989), 由于重复次数不同, 造成了每个位点的多态性。SSR 标记遍布整个基因组, 具有共显性、多态性高和重复性好的特点, 且 SSR 位点在属间和属内种间具有保守性 (Thomas & Scott, 1992)。SSR 标记根据来源可分为基因组 SSR 和表达序列标签 SSR (expressed sequence tag-SSR, EST-SSR)。对于大量未完成基因组测序的植物来说, EST-SSR 标记更容易获得, 虽然其多态性低于基因组 SSR 标记, 但序列保守性更高, 有较好的通用性, 且其多态性可能与基因功能直接相关 (Eujayl et al., 2002)。目前, EST-SSR 标记已在洋葱 (李满堂 等, 2015)、辣椒 (刘峰 等, 2012)、茄子 (魏明明 等, 2016)、白菜 (忻雅 等, 2006; Ding et al., 2015)、甘蓝 (陈琛 等, 2010)、萝卜 (崔娜 等, 2012; Zhai et al., 2014)、黄秋葵 (李永平 等, 2018) 等蔬菜作物上开发与应用。

韭菜 (*Allium tuberosum* Rottl. ex Spreng.) 是同源四倍体作物 ( $2n = 4x = 32$ ), 且基因组巨大, 约 31.42 pg (程蛟文 等, 2013), 杂合度高, 标记开发困难, 目前所知基因组信息极其有限, 无法开发基因组 SSR 标记。另外, 韭菜性状遗传复杂, 种子繁殖方式特异, 具有兼性无融合生殖特性 (Gohil & Kaul, 1981; Kojima & Kawaguchi, 1989; Kojima et al., 1991; Kojima & Nagato, 1992a, 1992b, 1997; Yamashita et al., 2012), 杂交后代大多数是无融合生殖后代, 与母本性状完全一致, 是母本的复制, 只有极少数有性生殖后代是真正的杂交后代。通过表型鉴定有性繁殖后代不仅周期长且准确率低, 利用分子标记鉴定有性繁殖后代是一种快速高效的方法, 但是有关韭菜分子标记的研究很少, 只有少量应用 ISSR (李延龙 等, 2009; 张明 等, 2012) 和 RAPD (潘敏 等, 2005; 李万昌 等, 2012) 的报道。

利用韭菜全长转录组测序获得的数据进行 SSR 标记搜索, 分析其分布及组成特征, 并对其在韭菜中的可用性进行初步评价, 以期对韭菜的遗传多样性分析、种质资源鉴定及有性生殖后代筛选奠定基础。

## 1 材料与方法

### 1.1 全长转录组数据来源

韭菜全长转录组数据来源于平顶山市农业科学院韭菜课题组利用 PacBio RS II 三代测序平台对韭菜品种 '791' 的测序结果。2018 年 8 月, 韭菜开花后, 选择资源圃中生长正常的 2 年生植株进行取样, 分别取韭菜的叶、花序、薹、叶鞘和根, 提取 RNA 并等量混匀后委托北京百迈客生物科技有限公司进行单分子实时 (Single Molecular, Real-Time, SMRT) 测序, 获得了 21.22 Gb 的数据

量。测序后, 对高质量转录本以及校正后的转录本使用 CD-HIT 进行去冗余分析, 得到 51 719 条转录本序列, 以 500 bp 以上的 49 876 条转录本为背景数据进行 SSR 位点分析。

## 1.2 韭菜材料及其 DNA 提取

用于 SSR 引物筛选和可用性评价的材料为本课题组在全国各地收集和选育的 24 个不同类型的韭菜种质资源 (表 1)。用于有性生殖后代鉴定的材料是以韭菜品种 ‘马蔺韭’ ( $P_1$ ) 和 ‘791’ ( $P_2$ ) 为亲本, 配置的正、反交杂交组合的后代。所用试验材料在 2018 年 4 月上旬播种育苗, 6 月移栽定植, 9 月采集韭菜幼嫩叶片, 参照 Porebski 等 (1997) 的 CTAB 法进行基因组 DNA 的提取。

## 1.3 转录组 SSR 位点鉴别及引物设计

使用 MISA (MicroSatellite identification tool) 软件 (<http://pgrc.ipk-Gatersleben.de/misa>) 进行 SSR 位点查找, 查找标准为: 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的最少重复次数分别为 10、6、5、5、5 和 5 次。

采用 Primer Premier 5.0 设计引物。引物设计的主要参数为: 引物长度 18 ~ 28 bp, GC 含量 40% ~ 60% 之间, 退火温度 ( $T_m$ ) 55 ~ 65 °C 之间, 上、下游引物的退火温度值相差不大于 5 °C, 预计 PCR 产物长度在 100 ~ 300 bp 之间, 尽量避免引物二级结构如发卡结构、二聚体、错配的出现。

## 1.4 SSR 引物筛选

随机选取 208 对 SSR 引物交由生工生物工程(上海)股份有限公司合成。PCR 反应体系为 20  $\mu$ L, 其中包括 1  $\mu$ L 50 ng  $\cdot$   $\mu$ L<sup>-1</sup> 的 DNA 模板, 10  $\mu$ L 2 $\times$  Taq Master Mix for PAGE (Vazyme, 南京), 上、下游引物各 0.5  $\mu$ L (10  $\mu$ mol  $\cdot$  L<sup>-1</sup>), 8  $\mu$ L ddH<sub>2</sub>O。PCR 程序为: 94 °C 预变性 4 min, 94 °C 变性 30 s, 55 °C 退火 30 s (退火温度随引物不同而调整), 72 °C 延伸 30 s, 30 个循环, 最后 72 °C 延伸 8 min。PCR 产物用 6% 非变性聚丙烯酰胺凝胶电泳分离, 180 V 电压下电泳 90 min 左右, 电泳后的银染显色参照 Bassam 等 (1991) 的方法进行, 拍照后进行带型分析。

SSR 出现频率, 即  $fc(\%) = c/n \times 100$ ,  $c$  为搜索到的 SSR 数量,  $n$  为无冗余转录本数量; SSR 平均分布距离, 即  $fN = N/c$ ,  $N$  为无冗余转录本序列的总碱基数。

表 1 韭菜 SSR 多态性分析材料

Table 1 List of tested *Allium tuberosum* materials for SSR polymorphism analysis

序号 No.	名称 Name	来源 Source
1	航研 998 Hangyan 998	自育 Self breeding
2	勾头韭 Goutoujiu	河南洛阳 Louyang, Henan
3	马蔺韭 Malinjiu	内蒙古 Inner Mongolia
4	贵州大叶韭 Guizhou Dayejiu	贵州 Guizhou
5	紫根韭 Zigenjiu	河南周口 Zhoukou, Henan
6	大黄苗 Dahuangmiao	天津 Tianjin
7	长沙香韭 Changsha Xiangjiu	湖南长沙 Changsha, Hunan
8	平韭 4 号 Pingjiu 4	自育 Self breeding
9	汉中冬韭 Hanzhong Dongjiu	陕西汉中 Hanzhong, Shaanxi
10	环韭 Huanjiu	山西 Shanxi
11	平丰 8 号 Pingfeng 8	自育 Self breeding
12	桂林穿山韭 Guilin Chuanshanjiu	广西桂林 Guilin, Guangxi
13	四季青 Sijiqing	新疆库尔勒 Korla, Xinjiang
14	犀浦韭 Xipujiu	四川成都 Chengdu, Sichuan
15	春韭 Chunjiu	上海 Shanghai
16	津引 1 号 Jinyin 1	日本 Japan
17	白绵韭 Baimianjiu	陕西西安 Xi'an, Shaanxi
18	平韭 2 号 Pingjiu 2	自育 Self breeding
19	黄格子 Huanggezi	湖北武汉 Wuhan, Hubei
20	平丰 9 号 Pingfeng 9	自育 Self breeding
21	791	自育 Self breeding
22	嘉兴白根 Jiaxing Baigen	浙江杭州 Hangzhou, Zhejiang
23	青格子 Qinggezi	湖北武汉 Wuhan, Hubei
24	桂林大叶 Guilin Daye	广西桂林 Guilin, Guangxi

## 2 结果与分析

### 2.1 韭菜全长转录组数据中 SSR 的位点的数量与分布

利用 MISA 软件对 49 876 条大于 500 bp 的转录本序列(总长为 89.44 Mb)进行搜索, 共在 10 332 条转录本序列中发现符合条件的 13 111 个 SSR 位点, 其中 2 063 条转录本含有两个或两个以上 SSR 位点。SSR 的发生频率(含 SSR 的转录本数/总转录本数)为 20.72%, SSR 出现频率为 26.29%, 平均 6.82 kb 出现 1 个 SSR, 其中复合型 SSR(包含至少两个 SSR, 且间距小于 100 bp)有 2 015 个, 占总 SSR 的 15.37%。

在搜索到的韭菜转录组 SSR 中, 单核苷酸、二核苷酸和三核苷酸重复类型出现的频率占优势, 分别占总 SSR 的 66.55%、12.52%、19.67%; 四核苷酸、五核苷酸和六核苷酸的重复类型出现的频率很低, 只占总 SSR 的 1.26%(表 2)。韭菜转录组 SSR 重复单元的重复次数分布在 5~289 次之间, 其中 5~10 次重复的 SSR 位点有 7 667 个, 占总数的 58.48%; 11~20 次重复的 SSR 位点有 4 220 个, 占总数的 32.19%; 20 次重复以上的 SSR 位点有 1 224 个, 占总数的 9.34%(表 2)。韭菜转录组 SSR 的长度主要集中在 10~289 bp, 其中长度在 10~15 bp 的 SSR 有 9 586 个, 占总数的 73.11%, 长度在 16~20 bp 的 SSR 有 1 542 个, 占总数的 11.76%, 长度在 21~25 bp 的 SSR 有 681 个, 占总数的 5.19%, 长度大于 25 bp 的 SSR 有 1 302 个, 占总数的 9.93%。

表 2 韭菜 SSR 的类型、数量及分布频率  
Table 2 Type, number and frequency of SSRs in *Allium tuberosum* Rottler ex Spr.

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	11~20	>20		
单核苷酸 Mononucleotide						3 600	3 972	1 154	8 726	66.55
二核苷酸 Dinucleotide		630	368	191	113	61	209	70	1 642	12.52
三核苷酸 Trinucleotide	1 571	548	210	121	55	35	39		2 579	19.67
四核苷酸 Tetranucleotide	73	21	7	3	6				110	0.84
五核苷酸 Pentanucleotide	9	3	1	1					14	0.11
六核苷酸 Hexanucleotide	33	4	1	1		1			40	0.31
总计 Total	1 686	1 206	587	317	174	3 697	4 220	1 224	13 111	100.00
百分比/% Percentage	12.86	9.20	4.48	2.42	1.33	28.20	32.19	9.34		

### 2.2 韭菜全长转录组 SSR 基序重复类型和频率特征

单核苷酸重复在测序时容易发生错配而导致测序质量不高, 因此在后续的试验分析中不予选择(李满堂等, 2015; 杜晓华等, 2019)。从韭菜全长转录组 SSR 核苷酸基序类型来看, 13 111 个 SSR 位点包含 103 种重复基序, 二核苷酸至六核苷酸重复分别有 8、30、26、11 和 28 种; 从 SSR 的重复次数来看, 二核苷酸主要分布在 6~9 次, 占二核苷酸总数的 79.29%, 三核苷酸主要分布在 5~8 次, 占三核苷酸总数的 95.00%, 四核苷酸、五核苷酸和六核苷酸主要分布在 5~6 次, 分别占其总数 85.45%、85.71%和 92.50%。

从表 3 可知, 二核苷酸重复基序中 AC/GT(333 个), CA/TG(300 个), TA/TA(296 个)的出现频率较高, 分别占二核苷酸重复总数的 20.28%、18.27%和 18.02%。三核苷酸中出现最多的重复基序是 GAA/TTC(242 个), 占三核苷酸总数的 9.38%, 其次是 AAG/CTT(210 个)、ATC/GAT(161 个)、AAT/ATT(155 个)、CAG/CTG(155 个)和 TCA/TGA(150 个), 分别占三核苷酸总数的 8.14%、6.24%、6.01%、6.01%和 5.82%, 最少的是 GTA/TAC(7 个), 只占三核苷酸总数的 0.27%。四核苷

酸中出现最多的重复基序是 TGTA/TACA (13 个) 和 AAAT/ATTT (13 个), 各占四核苷酸总数的 11.82%。五核苷酸中以 AAAAG/CTTTT (2 个)、TTTTC/GAAAA (2 个) 和 AGAAA/TTTCT (2 个) 出现频率最高, 各占五核苷酸总数的 14.29%。六核苷酸中出现频率最高的是 TTCAAA/TTTGAA (5 个), 占六核苷酸总数的 12.5%, 其次是 GATTCG/CGAATC (2 个) 和 GGAGCA/TGCTCC (2 个), 各占六核苷酸总数的 7.5%。

表 3 韭菜全长转录组中 SSR 基序的分布  
Table 3 Distribution of SSR motif in full-length transcriptome of *Allium tuberosum* Rottler ex Spr.

重复基序类型 Repeat type	数量 Number	重复基序类型 Repeat type	数量 Number	重复基序类型 Repeat type	数量 Number
A/T	8 645	GCG/CGC	19	ATATA/TATAT	1
G/C	81	ACT/AGT	19	CCAGC/GCTGG	1
AC/GT	333	ACG/CGT	13	TCTCC/GGAGA	1
CA/TG	300	CGA/TCG	12	CCTCT/AGAGG	1
TA/TA	296	GTA/TAC	7	ATTTT/AAAAT	1
AT/AT	269	TGTA/TACA	13	ACAGC/GCTGT	1
AG/CT	235	AAAT/ATTT	13	GGGCT/AGCCC	1
GA/TC	191	ATGT/ACAT	10	TTCAAA/TTTGAA	5
CG/CG	11	TATG/CATA	9	GATTCG/CGAATC	3
GC/GC	7	TTAT/ATAA	9	GGAGCA/TGCTCC	3
GAA/TTC	242	ATAC/GTAT	7	CAGCA/TTGCTG	2
AAG/CTT	210	TTCT/AGAA	6	ATTGA/TCAAAT	2
ATC/GAT	161	TTTA/TAAA	6	CGCCTC/GAGGCG	2
AAT/ATT	155	TTTC/GAAA	5	AACCCC/GGGGTT	2
CAG/CTG	155	AAAG/CTTT	4	TATATG/CATATA	1
TCA/TGA	150	GACT/AGTC	3	TGGAAC/GTTCCA	1
TAA/TTA	138	GTTT/AAAC	3	ACTGGA/TCCAGT	1
AGA/TCT	131	TTGT/ACAA	3	CAATTT/AAATTG	1
GCA/TGC	126	TGTT/AACA	3	CGAGTT/AACTCG	1
ATA/TAT	106	AAGA/TCTT	2	ACCCGA/TCGGGT	1
CTC/GAG	99	GTGC/GCAC	1	GAAGAT/ATCTTC	1
ATG/CAT	94	GAAT/ATTC	1	AGAGGC/GCCTCT	1
GGA/TCC	83	CCTT/AAGG	1	ATCAGA/TCTGAT	1
CAC/GTG	80	ATGC/GCAT	1	GCAACA/TGTTGC	1
CAA/TTG	78	TCCT/AGGA	1	CCAAAC/GTTTGG	1
CCT/AGG	75	CCTC/GAGG	1	CTCCGC/GCGGAG	1
AAC/GTT	43	CTCC/GGAG	1	CCTCAA/TTGAGG	1
GAC/GTC	31	CTCA/TGAG	1	AACCCA/TGGGTT	1
AGC/GCT	75	TGCA/TGCA	1	ATGGGC/GCCCAT	1
CCA/TGG	73	TGCA/TGCA	1	CTTCTG/CAGAAAG	1
ACC/GGT	65	AATT/AATT	1	TTTGAG/CTCAAA	1
GCC/GGC	44	AAAAG/CTTTT	2	AGGAGA/TCTCCT	1
ACA/TGT	39	TTTTC/GAAAA	2	GGAAGA/TCTTCC	1
CCG/CGG	36	AGAAA/TTTCT	2	CAAATG/CATTTG	1
CTA/TAG	20	ACACA/TGTGT	1	AAAGAA/TTCTTT	1

2.3 韭菜全长转录组 SSR 引物的设计、筛选与多态性分析

除去复合型 SSR 和单核苷酸重复 SSR, 对 3 648 条 2~6 核苷酸重复的 SSR 位点进行引物设计, 共设计出 3 311 对 SSR 位点特异引物。为验证其引物的有效性, 随机挑选了 208 对引物对韭菜品种 ‘791’ 的基因组 DNA 进行 PCR 扩增, 其中 164 对引物能扩增到预期产物, 其余 44 对引物为无效引物, 引物扩增有效率为 78.85%。用 164 对有效引物对 24 份韭菜种质资源进行扩增, 分析评价这些有效引物的多态性 (图 1), 其中 39 对引物呈现多态性 (表 4), 占有效引物的 23.78%。每对引物产生的条带数在 2~4 之间, 共得到 100 个条带, 其中多态性条带有 51 个, 每对引物平均产生 1.31 个多态性条带。

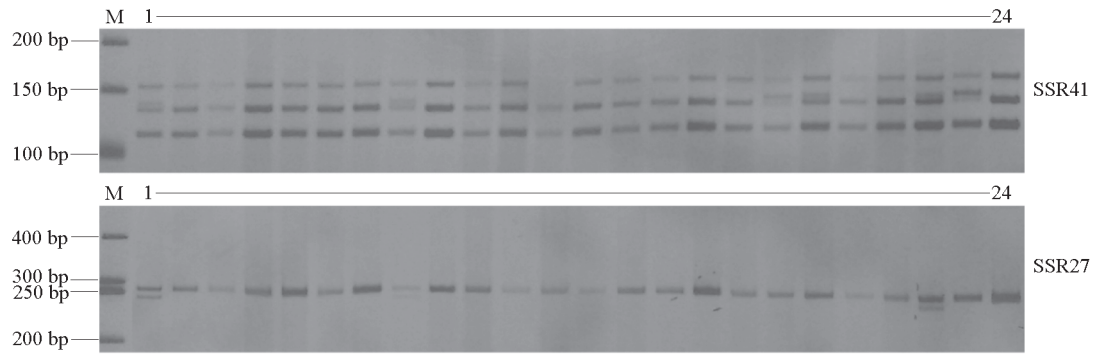


图1 引物 SSR41 和 SSR27 在 24 份韭菜种质资源 (编号名称见表 1) 中的多态性

Fig. 1 Polymorphisms showed 24 *Allium tuberosum* germplasm (Code names are shown in Table 1) by primer SSR41 and SSR27

表 4 韭菜 39 对多态性 SSR 引物信息

Table 4 Information of 39 pairs of polymorphic primers developed from *Allium tuberosum*

引物名称 Primer name	SSR 类型 SSR motif	正反向引物序列 (5' - 3') Forward, reverse primer sequence	产物预期 大小/bp Expected size	扩增带数 Number of bands	多态性带数 Number of polymorphic bands
SSR01	(ATT) <sub>5</sub>	F: TTTTGACGCTAGCAACATCG; R: GTGCACAGCATGCACAC	228	3	1
SSR08	(GATTCG) <sub>5</sub>	F: AAATCGATTCTTTTCGCCCTT; R: CAATTTTCGGTCTGCTCCTC	180	3	1
SSR10	(CCAGC) <sub>5</sub>	F: ACGTCTGGCTTTGGGTTGTA; R: GTTTGGCCATTGCAAGAAT	263	2	1
SSR13	(AAAG) <sub>5</sub>	F: AGGAGGCATAGCATCGGATA; R: GCATCTCAACTCTCCTTGAC	190	3	1
SSR15	(TTAT) <sub>5</sub>	F: TGGCTGCTTTTGTGAAACAC; R: GAAGACAACCCAAAACGAA	254	2	1
SSR16	(GTTT) <sub>5</sub>	F: TTGAGAGCTGACATGGATGC; R: TGCTTCATTGGAGCAGAAA	239	2	2
SSR21	(TTAT) <sub>5</sub>	F: TGCTACAATGACTCATCCCTTG; R: AACTGCAACAATGCCATGAA	141	2	1
SSR25	(ATGT) <sub>5</sub>	F: GTCTCAATAGCCGCTCCCTA; R: CGACCATAGGACCTCATTCTG	208	3	1
SSR27	(ATGC) <sub>5</sub>	F: CTTCAATGGAGCAGCAACCT; R: TGAGTGCATCTATCGCCTCTT	260	2	1
SSR30	(TTTT) <sub>6</sub>	F: CCATTTGATCAAGGAATCGG; R: AATTCAGCAACAAGCCCATC	279	2	1
SSR32	(TATG) <sub>6</sub>	F: ACCGGTTTTCCAAATTTCTCC; R: AACACACATGCATGCACAAA	222	2	1
SSR33	(TTTA) <sub>5</sub>	F: CCAAATTGACTTTTGGCTTTT; R: CGGACTGCTTAGTAGGGCAC	120	2	1
SSR41	(GAAGAT) <sub>5</sub>	F: GGTGGAGATGATTTCCTGTA; R: CACATTCATCCATCATGCC	155	4	3
SSR45	(AATT) <sub>5</sub>	F: CAGCCCTCATCCAGAGTTA; R: AACTGGACTGAAGCATCCAT	157	3	1
SSR47	(TCTCC) <sub>5</sub>	F: AAATATTTGTGCCGCTTTGC; R: GATGGGGTTTTTGAGGGATT	230	2	2
SSR48	(CCTCT) <sub>5</sub>	F: TTCGTCAACCTCCTCTCCTC; R: GATGGGGTTTTTGAGGGATT	186	2	1
SSR54	(TCCT) <sub>5</sub>	F: TGGACCATGATTCAAGCAA; R: AATCCTCTCCACATCAGCAAA	150	3	2
SSR64	(TTTT) <sub>5</sub>	F: GGATCAAAGCACATATCCGA; R: TTATACATGGAAGGAGTGAA	213	2	1
SSR65	(ATT) <sub>5</sub>	F: TATCAAAAAGGCAAGCTGGG; R: TGCTGAGTGCCATTTTAAAGC	217	2	1
SSR74	(TTAT) <sub>5</sub>	F: GACGTGGTGATTGGAGGACT; R: GCTCTTTCTCCAACCATTGAA	219	4	2
SSR75	(AACCCA) <sub>6</sub>	F: CCAAGACCACCTGTCCTAA; R: GTTGAGTTGGAGTGCCTTGA	189	3	1
SSR92	(AAAT) <sub>5</sub>	F: GGAGATGCAGCTTCACATCTT; R: TGCAATCACAAGTTGTACCGA	222	3	1
SSR94	(TATG) <sub>6</sub>	F: TTGAATGTTCCGTTGTGCTG; R: GCATGGTAGCAGCTTGTA	261	4	2
SSR98	(TGTA) <sub>9</sub>	F: GATGTGCGTGAATCATGGAG; R: TTTCTAGCAACAAGCATGCA	142	2	1
SSR99	(ATAC) <sub>6</sub>	F: GAGCGAGTGAGGATTGAAC; R: TGCCATGGAATTGAGAACAA	231	4	2
SSR102	(TGTA) <sub>6</sub>	F: GATTTACGGTGATGTGCGTG; R: CAAGCCACCATTATATTAGCACA	130	2	1
SSR105	(TTCAAA) <sub>5</sub>	F: CTCGTTCCCATCTCTCCAAA; R: CTGAAGAGTCCATGTCCCGT	213	2	1
SSR106	(TGTA) <sub>9</sub>	F: CTTGTGGTGGGATTTGAAGAA; R: AAGCCACCATTATATTAGCACA	249	3	2
SSR108	(TGTA) <sub>6</sub>	F: GATGTGCGTGAATCATGGAG; R: CTTACGTGCCTTCCAACCTC	175	3	1
SSR109	(TAT) <sub>6</sub>	F: CAATGCGCTATAGTTCAACACG; R: ATTGAAAGCCTAGATCGCCA	231	2	1
SSR111	(TGA) <sub>7</sub>	F: AACCTTGCTGATGCTGTTTG; R: AACCAGGTATCGTCGTCGTC	236	3	2
SSR117	(GGA) <sub>6</sub>	F: ACGTGGACAAGGAGATGAGG; R: TAAACGGGTGGGGATTGATA	177	2	1
SSR126	(CAG) <sub>6</sub>	F: CTTCAGAACAGCAATGGCAA; R: GCCAGCATTTCTTCTCCAT	219	4	2
SSR130	(ATT) <sub>5</sub>	F: CAGAGGGGAGTGAGCTTCAG; R: TCACAATTCCATATCATGCAGAA	171	2	1
SSR131	(GAA) <sub>5</sub>	F: TGGATTCTCTCTTGGCAAACT; R: CGGAGAATAGCGATGATGGT	151	2	1
SSR132	(TTC) <sub>5</sub>	F: ATCCACGAAACCAGAAACCA; R: GCGTAGGATGAAGATGGAA	210	2	2
SSR135	(CTT) <sub>5</sub>	F: TTCGTCAAGCAATTAACCTTCA; R: AGAATTCTTTGGGGGTGAGG	131	2	1
SSR140	(AAG) <sub>6</sub>	F: CAAAGCAGAGAAGAAACCCG; R: TTCCCATAGCTTTGCTCGAT	242	2	1
SSR172	(ACA) <sub>8</sub>	F: TGGGCTATATGCAGCACAAT; R: AATCGTGTGTTCCGCTCTCC	195	3	1

## 2.4 SSR 标记在韭菜有性生殖后代鉴定中的应用

韭菜具有兼性无融合生殖特性, 所产生的种子后代 90%以上是母本的复制, 只有极少量的有性生殖后代 (Kojima et al., 1991)。根据 SSR 引物多态性分析结果, 引物 SSR41 的扩增产物在韭菜种质资源 ‘791’ 中有一条特异性条带, 而 ‘马蔺韭’ 中没有。利用引物 SSR41 对 56 株 ‘马蔺韭’ × ‘791’ 杂交后代进行检测, 在 6 株杂交后代中检测到 ‘791’ 的特异条带, 可确认为有性生殖后代。由于引物 SSR41 在 ‘马蔺韭’ 中没有显性特异条带, 因此不适用于检测以 ‘791’ 为母本的反交组合后代, 所有后代检测结果都与母本一致 (图 2)。上述结果表明利用 SSR 标记从大量杂交后代中快速筛选出有性生殖后代是一种可行的方法, 但是需要筛选出更多有多态性的 SSR 引物。

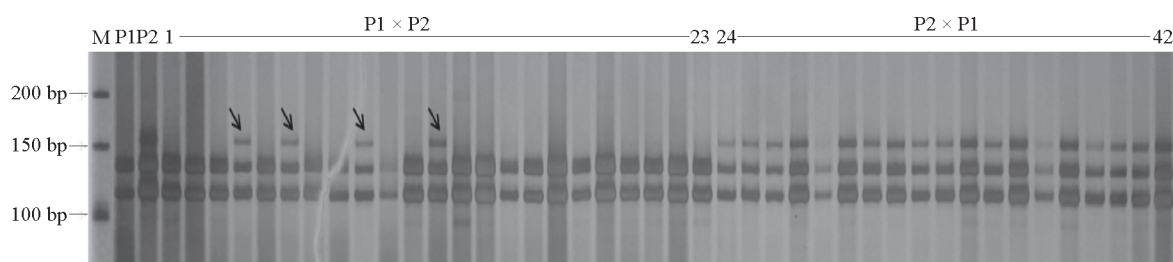


图 2 SSR41 检测杂交后代中有性生殖后代的结果

M: DNA marker; P1: ‘马蔺韭’; P2: ‘791’。

Fig. 2 Detection of sexual reproductive progenies in hybrid progenies by SSR41

M: DNA marker; P1: ‘Malinjiu’; P2: ‘791’。

## 3 讨论

目前常用的转录组测序技术有基于 Illumina 平台的二代转录组测序技术和基于 PacBio 平台的三代全长转录组测序技术。全长转录组测序技术无需拼接即可直接获得完整的全长转录本, 克服了二代测序技术读长较短, 得到的测序片段需要拼接, 可能会产生拼接错误和较多的嵌合体, 从而不能得到完整转录本的缺点。前人对于 EST-SSR 标记的开发都是基于 Illumina 平台的二代转录组测序结果, 未见有利用三代全长转录组测序结果开发 SSR 标记的报道。本研究利用基于 PacBio 平台的三代测序技术进行的韭菜全长转录组测序, 从 49 876 条大于 500 bp 的转录本序列 (总长为 89.44 Mb) 中搜索出了 13 111 个 SSR 位点, SSR 出现频率为 26.29%, 平均 6.82 kb 出现 1 个 SSR。由于三代全长转录组测序获得的转录本平均长度显著高于二代转录组测序拼接的 unigene 的平均长度, 用于搜索 SSR 位点的转录本数量比其他二代转录组测序拼接的 unigene 的数量要少得多。因此, 本研究中 SSR 的出现频率高于利用基于 Illumina 平台进行二代转录组测序的同属植物韭菜和洋葱, 其中在韭菜的 60 031 条 unigene 中仅搜索到 2 279 个 SSR 位点, SSR 出频率仅为 3.80%, 在洋葱的 106 932 条 unigene 中仅搜索到 5 959 个 SSR 位点, SSR 出频率仅为 5.57% (Zhou et al., 2015; 李满堂 等, 2015)。另外, 本研究中 SSR 的出现频率也高于茄子 (18.32%)、三色堇 (14.19%) 和黄秋葵 (20.59%) 等其他科植物 (魏明明 等, 2016; 李永平等, 2018; 杜晓华 等, 2019)。除测序方法外, 物种间 SSR 位点出现频率不同也可能与物种中 SSR 位点进化速度不同有关, 还可能受样品来源、测序深度、拼接质量以及 SSR 位点检索的参数设置的影响。

本研究中发现在 6 种韭菜 SSR 位点重复类型中, 单核苷酸重复 (8 726 个) 最多, 占总 SSR 的 66.55%, 其次是三核苷酸重复 (2 579 个) 占总 SSR 的 19.67%, 二核苷酸重复 (1 642 个) 占总 SSR



的 12.39%，与茄子（魏明明 等，2016）和三色堇（杜晓华 等，2019）的研究结果相同。本研究在韭菜中发现了 103 个 SSR 位点的重复基序类型，比万寿菊（45 个）（张华丽 等，2018）和茄子（54 个）（魏明明 等，2016）的类型丰富，但是比三色堇（149 个）（杜晓华 等，2019）和黄秋葵（181 个）（李永平 等，2018）中发现的类型少。在不统计单核苷酸重复时，大多数植物的 SSR 位点都是三核苷酸重复出现频率最高，只有樱桃（宗宇 等，2016）和越橘（方茜 等，2018）等少数物种的二核苷酸重复出现频率最高。本研究中出现频率最高的三核苷酸重复基序是 GAA/TTC（占 SSR 总数的 1.85%），与三色堇（杜晓华 等，2019）相同，与黄秋葵（李永平 等，2018）等植物不同。在葱属植物中，前人研究都发现三核苷酸重复基序 AAG/CTT 出现频率最高（李满堂 等，2015；Zhou et al., 2015），而本研究中 AAG/CTT 的出现频率仅次于 GAA/TTC，占 SSR 总数的 1.60%。由于识别重复基序时移动 1 个核苷酸即可将 GAA/TTC 识别为 AAG/CTT，这很可能是本研究与 Zhou 等（2015）和李满堂等（2015）研究结果不同的原因。本研究中出现频率最高的二核苷酸重复是 AC/GT，与 Zhou 等（2015）和李满堂等（2015）在韭菜和洋葱等葱属作物中的发现相同，与茄子（魏明明 等，2016）和莲雾（魏秀清 等，2018）等其他科植物不同。不同物种间 SSR 位点的重复类型是否有特定的规律还需要进一步的研究。

本研究中随机选了 208 对 SSR 引物在韭菜品种‘791’中进行检测，有 164 对引物能有效扩增出预期产物，有效率为 78.85%，高于洋葱（60%）（李满堂 等，2015）和三色堇（60%）（杜晓华 等，2019）、低于黄秋葵（87%）（李永平 等，2018）和万寿菊（80.56%）（张华丽 等，2018）。部分引物扩增失败可能是由于引物位置跨越内含子剪接位点引起的。在 24 份韭菜种质资源中对 164 对有效引物的多态性进行检测，发现 39 对引物有多态性，占有效引物的 23.78%，占总选择引物的 18.75%，表明韭菜转录组序列中的 SSR 位点有效扩增率高，但是多态性较低，要想获得足够的有多态性的 SSR 引物用于种质资源遗传多样性分析和有性生殖后代的鉴定，还需要大量筛选候选 SSR 位点。

韭菜的种子繁殖具有兼性无融合生殖特性，无论是自交还是杂交，超过 90% 的种子后代都是母本的复制（Kojima et al., 1991），有性生殖后代极少，通过表型性状筛选有性生殖后代非常困难，不仅耗时而且准确性低，非常不利于韭菜种质资源创新和新品种选育。本研究中利用筛选出的多态性标记 SSR41，在‘马蔺韭’×‘791’的杂交后代中检测出了 6 株有性生殖后代，与前人利用酯酶同工酶、RAPD 和 SCAR 标记检测有性生殖后代的研究相比（Kojima et al., 1991；Nakazawa et al., 2006；Yamashita et al., 2012），SSR 标记具有来源丰富，重复性好，通用性强，多态性条带清晰易读的优点。但是，由于韭菜基因组的多倍性和种质资源的高度杂合性，SSR 引物在有性生殖后代的检测中还有一定的局限性，一方面存在遗漏的可能，用 1 个标记不能将所有的有性生殖后代都检测出来，另一方面 1 对 SSR 引物只能在少数杂交组合中应用。下一步将继续筛选在韭菜种质资源中有多态性的 SSR 引物，为韭菜种质资源鉴定，有性生殖后代筛选，分子标记辅助育种和功能基因挖掘等研究奠定基础。

## References

- Bassam B J, Caetano-Anollés G, Gresshoff P M. 1991. Fast and sensitive silver staining of DNA in polyacrylamide gels. *Analytical Biochemistry*, 196 (1): 80 - 83.
- Chen Chen, Zhuang Mu, Li Kangning, Liu Yumei, Yang Limei, Zhang Yangyong, Cheng Fei, Sun Peitian, Fang Zhiyuan. 2010. Development and utility of EST-SSR marker in cabbage. *Acta Horticulturae Sinica*, 37 (2): 221 - 228. (in Chinese)
- 陈琛, 庄木, 李康宁, 刘玉梅, 杨丽梅, 张扬勇, 程斐, 孙培田, 方智远. 2010. 甘蓝 EST-SSR 标记的开发与应用. *园艺学报*, 37 (2): 221 - 228.



- Cheng Jiaowen, Wu Zhiming, Cui Junjie, Li Weipeng, Tan Shu, Hu Kailin. 2013. Statistical and comparative analysis of vegetable genome size. *Acta Horticulturae Sinica*, 40 (1): 135 - 144. (in Chinese)
- 程蛟文, 吴智明, 崔竣杰, 李卫鹏, 谭 澍, 胡开林. 2013. 主要蔬菜作物基因组含量统计与比较分析. *园艺学报*, 40 (1): 135 - 144.
- Cui Na, Qiu Yang, Li Xixiang, Shen Di, Wang Haiping, Song Jiangping. 2012. Data mining for SSRs in EST resources and EST-SSR markers development in radish. *Acta Horticulturae Sinica*, 39 (7): 1303 - 1312. (in Chinese)
- 崔 娜, 邱 杨, 李锡香, 沈 镒, 王海平, 宋江萍. 2012. 萝卜 EST 资源 SSR 信息分析及 EST-SSRs 标记开发. *园艺学报*, 39 (7): 1303 - 1312.
- Ding Q, Li J, Wang F, Zhang Y, Li H, Zhang J, Gao J. 2015. Characterization and development of EST-SSRs by deep transcriptome sequencing in Chinese cabbage (*Brassica rapa* L. ssp. *pekinensis*). *International Journal of Genomics*, 2015: 473028.
- Du Xiaohua, Yang Yaping, Zhu Xiaopei, Niu Yangli, Liu Huichao. 2019. Development of genic-SSR marker by transcriptome sequencing in *Viola × wittrockiana*. *Acta Horticulturae Sinica*, 46 (4): 797 - 806. (in Chinese)
- 杜晓华, 杨雅萍, 朱小佩, 牛杨莉, 刘会超. 2019. 三色堇转录组 SSR 分析及分子标记开发. *园艺学报*, 46 (4): 797 - 806.
- Eujayl I, Sorrells M E, Baum M, Wolters P, Powell W. 2002. Isolation of EST-derived microsatellite markers for genotyping the A and B genomes of wheat. *Theoretical and Applied Genetics*, 104 (3): 399 - 407.
- Fang Qian, Zhang Yuanyuan, Yang Yuting, Huang Miaomiao, Fu Qiaoli, Zhou Huisha, Chen Wenrong, Zong Yu, Guo Weidong. 2018. SSR mining and polymorphism analysis in leaf transcriptome of blueberry. *Acta Horticulturae Sinica*, 45 (7): 1359 - 1370. (in Chinese)
- 方 茜, 张园园, 杨钰婷, 黄苗苗, 符巧丽, 周慧莎, 陈文荣, 宗 宇, 郭卫东. 2018. 越橘叶片转录组 SSR 发掘及其多态性研究. *园艺学报*, 45 (7): 1359 - 1370.
- Gohil R N, Kaul R. 1981. Studies on male and female meiosis in Indian *Allium* II. Autotetraploid *Allium tuberosum*. *Chromosoma*, 82: 735 - 739.
- Kojima A, Nagato Y. 1992a. Diplosporous embryo-sac formation and the degree of dipospority in *Allium tuberosum*. *Sexual Plant Reproduction*, 5: 72 - 78.
- Kojima A, Nagato Y. 1992b. Pseudogamous embryogenesis and the degree of parthenogenesis in *Allium tuberosum*. *Sexual Plant Reproduction*, 5: 79 - 85.
- Kojima A, Nagato Y. 1997. Discovery of highly apomictic and highly amphimictic dihaploids in *Allium tuberosum*. *Sexual Plant Reproduction*, 10: 8 - 12.
- Kojima A, Nagato Y, Hinata K. 1991. Degree of apomixes in Chinese chive (*Allium tuberosum*) estimated by esterase isozyme analysis. *Japanese Journal of Breeding*, 41: 73 - 83.
- Li Mantang, Zhang Shilin, Deng Peng, Hou Xilin, Wang Jianjun. 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 Wanchang, Qiao Baojian, Wang Junwei, Yu Jiaojiao. 2012. Analysis of genetic diversity in leek cultivars based on RAPD. *Journal of Henan Agricultural Sciences*, 41 (3): 116 - 119. (in Chinese)
- 李万昌, 乔保建, 王俊伟, 余娇娇. 2012. 基于 RAPD 技术的韭菜品种间遗传多样性分析. *河南农业科学*, 41 (3): 116 - 119.
- Li Yanlong, Zhang Ming, Cao Xiumin, Zhang Wei, Wang Liya. 2009. Optimization of ISSR-PCR reaction system in *Allium tuberosum* Rottler ex Spr. *Journal of Henan Agricultural Sciences*, (1): 91 - 93, 99. (in Chinese)
- 李延龙, 张 明, 曹秀敏, 张 伟, 王利亚. 2009. 韭菜 ISSR-PCR 反应体系的优化. *河南农业科学*, (1): 91 - 93, 99.
- Li Yongping, Liu Jianting, Chen Mindong, Zhang Qianrong, Zhu Haisheng, Wen Qingfang. 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, 陈敏氢, 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 Yunsheng, Tian Xueliang, Mao Zhenchuan, Zhou Xuexiao, Xie Bingyan. 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.

- Nakazawa Y, Namai K, Kojima A, Kobayashi S, Tasaki K, Amagai M. 2006. Estimating the inheritance for parthenogenesis in tetraploid Chinese leek (*Allium ramosum*, syn. *A. tuberosum*). *Breeding Research*, 8: 89 – 98.
- Pan Min, Yang Jianping, Cao Dehang, Li Xincheng, Liu Lianhang, Li Yongxiang. 2005. Genetic diversity of cultivated Chinese chive germplasm by ISSR and RAPD markers. *Chinese Agricultural Science Bulletin*, 21 (4): 44 – 47, 58. (in Chinese)
- 潘 敏, 杨建平, 曹德航, 李新成, 刘连航, 李永祥. 2005. 韭菜栽培品种遗传多样性的 ISSR 和 RAPD 研究. *中国农学通报*, 21 (4): 44 – 47, 58.
- Porebski S, Bailey L G, Baum B R. 1997. Modification of a CTAB DNA extraction protocol for plants containing high polysaccharide and polyphenol components. *Plant Molecular Biology Reporter*, 15: 8 – 15.
- Tautz D. 1989. Hypervariability of simple sequence as a general source of polymorphic DNA markers. *Nucleic Acids Research*, 17: 6463 – 6471.
- Thomas M R, Scott N S. 1992. Microsatellite repeats in grapevine reveal DNA polymorphisms when analysed as sequence-tagged sites (STSs). *Theoretical and Applied Genetics*, 86 (8): 985 – 900.
- Wei Mingming, Chen Yuhui, Liu Fuzhong, Zhang Ying, Lian Yong. 2016. Development of SSR markers for eggplant with transcriptome sequencing data. *Journal of Plant Genetic Resources*, 17 (6): 1082 – 1091. (in Chinese)
- 魏明明, 陈钰辉, 刘富中, 张 映, 连 勇. 2016. 基于转录组测序的茄子 SSR 标记开发. *植物遗传资源学报*, 17 (6): 1082 – 1091.
- Wei Xiuqing, Xu Ling, Zhang Xijuan, Yu Dong, Chen Zhifeng, Zhang Limei, Chen Changzhong, Xu Jiahui. 2018. Analysis on SSR information in transcriptome and development of molecular markers in wax apple. *Acta Horticulturae Sinica*, 45 (3): 541 – 551. (in Chinese)
- 魏秀清, 许 玲, 章希娟, 余 东, 陈志锋, 张丽梅, 陈长忠, 许家辉. 2018. 莲雾转录组 SSR 信息分析及其分子标记开发. *园艺学报*, 45 (3): 541 – 551.
- Xin Ya, Cui Hairui, Lu Meizhen, Yao Yanling, Jin Jiqiang, Lin Rongshao, Cui Shuilian. 2006. Data mining for SSRs in ESTs and EST-SSR marker development in Chinese cabbage. *Acta Horticulturae Sinica*, 33 (3): 549 – 554. (in Chinese)
- 忻 雅, 崔海瑞, 卢美贞, 姚艳玲, 金基强, 林容杓, 崔水莲. 2006. 白菜 EST-SSR 信息分析与标记的建立. *园艺学报*, 33 (3): 549 – 554.
- Yamashita K, Nakazawa Y, Namai K, Amagai M, Tsukazaki H, Wako T, Kojima A. 2012. Modes of inheritance of two apomixis components, diplospory and parthenogenesis, in Chinese chive (*Allium tuberosum*) revealed by analysis of the segregation population generated by back-crossing between amphimictic and apomictic diploids. *Breeding Science*, 62: 160 – 169.
- Zhai L, Xu L, Wang Y, Cheng H, Chen Y, Gong Y, Liu L. 2014. Novel and useful genic-SSR markers from de novo transcriptome sequencing of radish (*Raphanus sativus* L.). *Molecular Breeding*, 33 (3): 611 – 624.
- 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 Ming, Li Yanlong, Wang Zhen, Yin Shouheng, Liu Hongmin. 2012. The primary exploring of genetic diversity of leek varieties based on ISSR markers. *Acta Agriculturae Boreali-occidentalis Sinica*, 21 (1): 146 – 150. (in Chinese)
- 张 明, 李延龙, 王 贞, 尹守恒, 刘宏敏. 2012. 基于 ISSR 标记的韭菜种质资源遗传多样性初探. *西北农业学报*, 21 (1): 146 – 150.
- Zhou S, Chen L, Liu S, Wang X, Sun X. 2015. *De novo* assembly and annotation of the Chinese chive (*Allium tuberosum* Rottler ex Spr.) transcriptome using the Illumin platform. *PLoS ONE*, 10 (7): e0133312.
- Zong Yu, Wang Yue, Zhu Yonyin, Shao Xu, Li Yongqiang, Guo Weidong. 2016. Development and validation of SSR markers based on transcriptomic data of Chinese cherry (*Prunus pseudocerasus*). *Acta Horticulturae Sinica*, 43 (8): 1566 – 1576. (in Chinese)
- 宗 宇, 王 月, 朱友银, 邵 姁, 李永强, 郭卫东. 2016. 基于中国樱桃转录组的 SSR 分子标记开发与鉴定. *园艺学报*, 43 (8): 1566 – 1576.