

刘冬梅, 娄喜艳, 吴 狄, 等. 基于棉花转录组测序的 SSR 分子标记的开发[J]. 江苏农业科学, 2019, 47(7): 32–35.
doi:10.15889/j.issn.1002-1302.2019.07.008

基于棉花转录组测序的 SSR 分子标记的开发

刘冬梅¹, 娄喜艳², 吴 狄¹, 喻 拥¹, 裴冬丽¹

(1. 商丘师范学院生物与食品学院/植物与微生物互作省级重点实验室, 河南商丘 476000; 2. 商丘工学院土木工程学院, 河南商丘 476000)

摘要:通过对棉花转录组进行测序获得的 57 695 条 Unigenes, 利用 MISA 软件进行搜索, 得出简单重复序列 (simple sequence repeat, 简称 SSR) 的分布情况, 共挖掘检索得到 1 886 个 SSR 位点, SSR 的出现频率为 3.27%。所有 SSR 位点分布于 1 759 条 Unigenes 上, 发生频率为 3.05%, 平均每 20.8 kb 会出现 1 个 SSR 位点。有 117 条 Unigenes 含有 1 个以上 SSR 位点, 含有复合型 SSR 的 Unigenes 数目为 56 条, 其中三核苷酸基元类型分布最多, SSR 数量为 1 582 个, 占挖掘出总 SSR 数量的 83.88%, 而单、二、四、五、六核苷酸重复类型所占的比例较小。最后通过 Primer 3 程序进行 SSR 引物设计, 得到部分引物序列, 可用于棉花遗传多样性分析、分子标记辅助育种以及种质资源的保存等。

关键词:棉花; 转录组测序; SSR; 分子标记; 引物设计

中图分类号: S562.024 **文献标志码:** A **文章编号:** 1002-1302(2019)07-0032-03

微卫星 DNA, 即简单重复序列 (simple sequence repeat, 简称 SSR), 通常构成其重复基元的核苷酸个数为 1~6 个, 特点是长度长达几十个核苷酸, 为串联重复序列。单一序列的保守性较强, 在基因组中, 微卫星侧翼序列就有此特性。因此首先通过克隆相应的微卫星侧翼 DNA 片段, 对其数目进行扩增, 然后对扩增 DNA 进行全部测序, 再对微卫星的侧翼序列引物进行人工合成, 最终实现通过 PCR 扩增微卫星的目的。自分子领域发展以来, 人们最常用的分子标记技术主要有随机扩增多态性 DNA (RAPD)、扩增片段长度多态性 (AFLP)、序列标签位点 (STS)、限制性片段长度多态性 (RFLP)、单核苷酸多态性 (SNP)、SSR 等^[1-6]。与其他分子标记相比, SSR 标记具有其独特的优点, (1) SSR 标记的操作非常简便, 无需消耗大量试验器材即可进行, 整个过程消耗的时间较其他几种分子标记短, 且对于多态性而言, SSR 标记所发掘出的多态性较高; (2) SSR 标记所开发的位点具有多等位基因的特性, 且在分子水平上, 其提供的信息量较高; (3) SSR 标记识别出的基因位点呈共显性, 在整个基因组有着均匀分布, 且整个过程无放射、辐射危险。SSR 标记适用于 DNA 指纹图谱的构建^[7]、基因定位^[8]和遗传多样性分析^[9]等方面。高度变异是微卫星中重复基元的显著特性, 微卫星数目可呈现出整倍性变异, 并且重复基元序列中的序列有可能不完全相同, 因而造成多个位点的多态性。

目前关于棉花全基因组 SSR 分子标记开发的研究报道较少, 随着棉花品种数量的逐渐增多, 田间表型农艺性状鉴定方法已难以满足快速并准确鉴定品种、质量、亲缘关系等的需求。本研究以陆地棉洞 A 转录组的测序结果为试验材料, 对

全转录组 SSR 位点进行挖掘, 利用生物信息学技能和 MISA 软件, 发掘基于陆地棉洞 A 基因型的 SSR 分子标记数据, 在 RNA 水平上分析棉花 SSR 的特点及规律, 以期在分子水平上研究棉花种质资源、鉴定亲缘关系及进行分子辅助育种奠定基础。

1 材料与方法

1.1 棉花转录组数据来源

2014 年 8 月, 以陆地棉洞 A 不育和可育花药为材料, 委托深圳华大基因科技有限公司进行转录组测序, 获得的全基因序列共包含 57 695 条 Unigenes, 以此为材料进行 SSR 位点的挖掘。

1.2 棉花转录组 SSR 位点检索

利用 SSR 位点挖掘软件 MISA 对陆地棉洞 A 花药的 57 695 条 Unigenes 进行 SSR 位点挖掘, 共挖掘出 6 种不同类别的 SSR, 分别是单、二、三、四、五、六核苷酸 SSR。

1.3 SSR 引物挖掘

利用 MISA 软件进行 SSR 位点挖掘之前首先将单、二、三、四、五、六核苷酸重复次数的操作参数分别设置为 ≥ 12 bp、 ≥ 6 bp、 ≥ 5 bp、 ≥ 5 bp、 ≥ 4 bp、 ≥ 4 bp; 其次对长度条件参数进行设置, 当 2 个微卫星可以组合为 1 个复合微卫星时, 这 2 个微卫星之间的距离必须小于 100 bp。

1.4 引物设计

通过 Primer 3 程序进行 SSR 引物设计。

2 结果与分析

2.1 棉花转录组序列中 SSR 位点的数量

通过 MISA 软件对棉花转录组获得的 Unigenes 进行挖掘, 从表 1 可以看出, 共挖掘出 1 886 个 SSR 位点, SSR 出现频率 (搜索出的 SSR 数量与搜索序列总数的比值) 为 3.27%。所有 SSR 位点分布于 1 759 条 Unigenes 上, 发生频率 (搜索出的含有 SSR 序列的数量与搜索序列总数的比值) 为 3.05%, 平均每 20.8 kb 会出现 1 个 SSR 位点 (搜索出的序列总长度

收稿日期: 2018-05-31

基金项目: 国家自然科学基金 (编号: 31571997); 河南省高等学校重点科研项目 (编号: 16A210036)。

作者简介: 刘冬梅 (1978—), 女, 黑龙江哈尔滨人, 博士, 讲师, 主要从事棉花种植资源利用与创新研究。E-mail: liudongmei78@126.com。

通信作者: 裴冬丽, 博士, 教授, 主要从事植物与微生物互作研究。

E-mail: peidongli@126.com。

与 SSR 总数量的比值)。其中含有 1 个以上 SSR 位点的 Unigenes 数量较少,有 117 条;含复合型 SSR 的序列数更少,为 56 条。

表 1 棉花转录组中 SSR 分析结果

项目	数量
搜索序列总数(条)	57 695
搜索序列总长度(bp)	39 312 432
SSR 位点总数(个)	1 886
含 SSR 位点的序列数(条)	1 759
含 1 个以上 SSR 位点的序列数(条)	117
含复合型 SSR 的序列数(条)	56
SSR 出现频率(%)	3.27
SSR 发生频率(%)	3.05
平均距离(kb)	20.8

2.2 棉花转录组 SSR 基元频率特征及引物设计

由表 2 可知,棉花转录组的 SSR 种类较为丰富,本研究共挖掘到 86 种重复的 SSR 类型,且各重复类型之间的数量差别较大,分布极其不平衡。棉花转录组中三核苷酸 SSR 重复基元类型的分布最多,SSR 重复基元数量为 1 582 个,占挖掘出总 SSR 数量的 83.88%;SSR 位点数范围为 1~393 个,最多的为 AAG/CTT,其次是 ATC/ATG,数量为 325 个,AGC/CTG 的数量为 216 个,ACC/GGT 数量为 209 个。单、二、四、五、六核苷酸重复类型所占的比例较小,SSR 重复基元数量分别为 66、92、15、11、120 个,总比例之和为 16.12%。在棉花转录组 Unigenes 中,重复基元类型最多的是六核苷酸,为 61 种;其次是三核苷酸重复类型,有 10 种。从图 1 可以看出,在三核苷酸重复基元类型中,AAG/CTT 的 SSR 位点数量最多,为 393 个,占总 SSR 位点数量的 20.84%,发生频率为 0.68%;其次是 ATC/ATG,占总 SSR 位点数量的 17.23%;再次是 AGC/CTG,占 11.45%;ACC/GGT 较少,占 11.08%;其他三核苷酸重复基元类型占总 SSR 位点的 23.28%。单、二、四、五、六核苷酸重复类型数量分别为 2、4、3、6 种,共占重复类型总数的 17.44%。通过对棉花转录组 57 695 条序列进行 SSR 位点搜索,共搜索得到 1 886 个 SSR 位点,利用 Primer 3 引物设计程序所得出的部分 SSR 位点特异引物见表 3。

2.3 棉花转录组序列中 SSR 基元重复次数分布

SSR 位点的多态性是由基元重复次数的变化所决定的,通过棉花 SSR 位点重复次数分布统计结果(图 2)可以得出,棉花转录组基元随着重复次数的增加,其数量和比例逐渐减少。棉花转录组 SSR 中基元重复次数主要集中在 4~7 次,占总数的 94.91%,其中重复 8~10 次的 SSR 位点数量为 28 个,占总数的 1.48%;重复 11~14 次数的 SSR 位点为 61 个(3.23%),重复 15~17 次的 SSR 位点有 7 个(0.37%),其中重复 16 次的 SSR 位点有 0 个。重复次数多于 20 次为较高重复次数,棉花转录组中没有重复次数高于 17 次的 SSR 位点,即无高重复次数的 SSR 位点,低重复次数(重复 1~10 次)与一般重复次数(重复 11~20 次)的 SSR 位点较多。

3 讨论

本研究利用棉花花药的转录组序列数据,从 57 695 条序列中成功挖掘检索出 1 886 个 SSR 位点,棉花转录组序列中

表 2 SSR 重复基元类型及数量

重复类型	类型数量(种)	重复基元类型	重复基元数量(个)	发生频率(%)	比例(%)
单核苷酸	2	A/T	61	0.01	3.234
		C/G	5	0.00	0.265
二核苷酸	4	AC/GT	55	0.09	2.916
		AG/CT	33	0.03	1.750
		AT/AT	3	0.00	0.159
		CG/CG	1	0.00	0.053
三核苷酸	10	AAG/CTT	393	0.68	20.838
		ATC/ATG	325	0.56	17.232
		AGC/CTG	216	0.37	11.453
		ACC/GGT	209	0.36	11.082
		其他类型	439	0.75	23.277
四核苷酸	3	AAAG/CTTT	11	0.01	0.583
		AAAT/ATTT	3	0.00	0.159
		ACAT/ATGT	1	0.00	0.053
五核苷酸	6	AAAAT/ATTTT	4	0.00	0.212
		AAAAG/CTTTT	2	0.00	0.106
		AAGAG/CTCTT	2	0.00	0.106
		AAAAC/GTTTT	1	0.00	0.053
		其他类型	2	0.00	0.106
六核苷酸	61	所有类型	120	0.20	6.363
合计	86		1 886	3.05	100

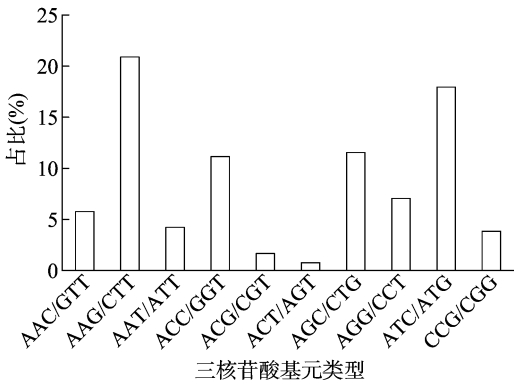


图 1 SSR 中三核苷酸基元类型分布

SSR 位点的总发生频率为 3.05%,通过对比杜仲(2.90%)^[10]、南方红豆杉(2.07%)^[11]可以发现,棉花转录组序列中 SSR 位点的发生频率略高,但略低于红松(4.24%)^[12],且明显低于樱桃(15.62%)^[13]。棉花转录组 SSR 标记重复次数最多的是三核苷酸重复基元类型,高级基元重复次数较为微小,且跨度大,可为基元的重复次数与 SSR 标记的多态性呈正相关提供依据^[14]。三核苷酸中的 AAG/CTT 基元的丰度最高,这与海甘蓝三核苷酸为丰度最高的基元^[15]一致,即棉花转录组 SSR 中三核苷酸 SSR 在理论上具备更高的多态性,可作为潜在的 SSR 重复基序进行有关的引物挖掘及开发。由此推测,通过利用棉花转录组得到的 SSR 引物,同样可以用来进行棉花的遗传多样性分析、分子标记辅助育种以及种质资源的保存等。

参考文献:

[1] 涂明月,李杰,何亚丽,等. 利用 RAPD 标记鉴定草地早熟禾种质资源的遗传多样性[J]. 草业学报,2017,26(7):71-81.
[2] 石悦,于肖夏,于卓,等. 高丹草 AFLP 分子遗传连锁图谱的

表 3 部分 SSR 重复类型及引物序列

编号	SSR 类型	引物序列 (5'→3')
MH01	(CCT) ₅	GGTGCTTACATTCCGAGCCTGAAGAGGAGGAGGAGGAGGG
MH05	(TCT) ₅	ACGCACTTGTAGCAGCTGATCCGAGTCATGCCAGTCACAA
MH10	(AGA) ₅	GAGTTGGAGGAATCCAGGAACAGTGGCCTGGGACTCCAATC
MH13	(TGATGG) ₄	GCTTAATTAAGACGTGGCAAGCTCCCATCACCATCACCATCAC
MH14	(CCATCA) ₄	GCTCCCATCACCATCACCATCAGCCTTAATTAAGACGTGGCA
MH16	(TTC) ₅	TCCTCCTTTGAACCTGCTGACCACTCATGTGATTCTGCCGA
MH22	(ATGGAA) ₄	TGCTGGAGGTGCGAGGTGAATCCTTCTTTGCCGCTCCTCC
MH24	(ATGGAA) ₄	TGCTGGAGGTGCGAGGTGAATCCTTCTTTGCCGCTCCTCC
MH32	(CAC) ₅	TGTTGTTTGGCATGGCTGTGTGGAATGGCTGCCAAATTCGA
MH34	(TGG) ₅	TGTTGTTTGGCATGGCTGTGTGGAATGGCTGCCAAATTCGA
MH46	(GGC) ₅	GTTCAGCCAAGCCAAACGCTAACCTGGAATCCGTTTGC
MH53	(TTTGCT) ₅	TTGCCCTTCACCGCACTTTAGCTCCGCTTATCTTTTGGCG
MH55	(GTT) ₅	AATCTCCTCGCCACCGATTGCGGCATGAGTTTTCGTTGCA
MH56	(TTTGCT) ₅	TGCTCCTCAACTCTGCGTTCCGGCATGAGTTTTCGTTGCA
MH62	(GGT) ₅	ACCTGCAGTTCCATGGCTACGTGCCCGATGTTGCAAATGT
MH65	(TTC) ₅	TGACGTTGCGCTTCTCTTCTGAACTGCAATTTCCCCCT
MH72	(TAA) ₅	ATCACTTCCCTTTGCTCCGGCCCATGTGCCAAACTGATG
MH77	(AGA) ₆	CTGCAGCTATGGAGGTAGCCCTTGGGAAGCTCGGCTACAA
MH80	(GTG) ₆	CTGCAGCTATGGAGGTAGCCCTTGGGAAGCTCGGCTACAA
MH97	(CAGCCT) ₆	TCGAGGTCAACCAAAGAGGCCGTTGCTTTGGAAACCTGG
MH98	(CAGCCT) ₇	TCGAGGTCAACCAAAGAGGCCGTTGCTTTGGAAACCTGG
MH106	(GCTGGA) ₄	TCACATCCTCCGCAAACTCCGCGGTCTTAGGGAATGAGG
MH110	(TCT) ₅	TCACATCCTCCGCAAACTCCGCGGTCTTAGGGAATGAGG
MH122	(GGAGAA) ₄	GAAGACCTTCCACAGCCTGCGTTAATCGGGGATGCTAGTTGC
MH129	(TCCATC) ₄	ACCCCTTTCTTCCATCTCCACTGAAATCTTCGCAAGCTGC
MH135	(AATGAG) ₆	GCAAGAGATCCCGTACTCCCTCAGCAGGCAATGGACGATT
MH151	(GCG) ₆	GATCTTCCATCCCACCTCCACTTGGTAGCCCTTGTGATGCA
MH155	(CAA) ₅	GGGCAATTGCGTTTTCAAACCGCACGGAGCTTTTTCGTTGTT
MH159	(AGAAA) ₄	AAAAGCCGAATGGAAAACGCTTTCTGCCCTCATCTCTCCC
MH161	(CTT) ₇	TGACTCCATCTTCTTAACAGCTGGAAGAAGCCGACATGAACG
MH166	(AGGCGG) ₄	GGTTCACCCAGCGGAGTATCGGATGCCAGTGACCAAAGGT
MH168	(GA) ₇	ACTCCAGTCCCAAAGTGACGACGCTCTTGCTTGTCTTCCA
MH171	(TC) ₆	GTTTACTCGCTCTCTCGCCGACGGGAAAGGAAGCACTCA
MH173	(TC) ₈	CCATCTGCTGGCTGTTTCAAACCTAACAGAACAGCAGGGT
MH177	(CTG) ₇	GTGCCATTCCCATCCTTCCAAACTCTGTGCAGATCCCGTC
MH188	(AG) ₆	GCGAGCGGAATTATCGGAGACTTGCAGCAGCTTTGCTTGA
MH97	(CAGCCT) ₆	TCGAGGTCAACCAAAGAGGCCGTTGCTTTGGAAACCTGG
MH98	(CAGCCT) ₇	TCGAGGTCAACCAAAGAGGCCGTTGCTTTGGAAACCTGG

注:括号内表示 SSR 基元;括号外数字表示重复次数。

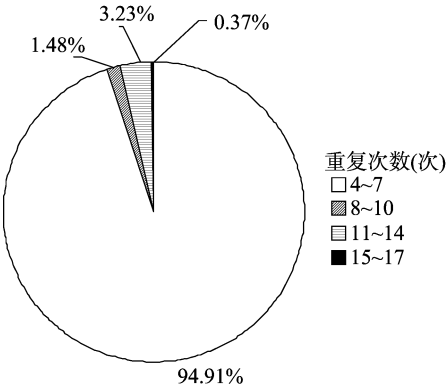


图2 SSR 位点数重复次数分布

构建[J]. 西北农林科技大学学报(自然科学版),2017,45(10): 7-14.

[3] 曹雯梅,刘述忠,杨青华,等. STS 标记鉴定河南历史主推小麦品种春化光周期基因及与品种冬春性的相关性[J]. 分子植物育种,2016,14(1):117-124.

[4] 王 婵,赵 泓,张丽英,等. 大葱细胞质雄性不育系及保持系线粒体基因组 RFLP 分析[J]. 农业生物技术学报,2015,23(7): 888-893.

[5] 匡 猛,王延琴,周大云,等. 基于单拷贝 SNP 标记的棉花杂交种纯度高通量检测技术[J]. 棉花学报,2016,28(3):227-233.

[6] Fan C J, Liu Q Y, Zeng B S, et al. Development of simple sequence repeat (SSR) markers and genetic diversity analysis in blackwood (*Acacia melanoxylon*) clones in China[J]. Silvae Genetica,2016,65(1):49-54.

司振书,殷国政,路建彪,等. H9N2 亚型 AIV 双重 RT-PCR 检测方法的建立[J]. 江苏农业科学,2019,47(7):35-37.

doi:10.15889/j.issn.1002-1302.2019.07.009

H9N2 亚型 AIV 双重 RT-PCR 检测方法的建立

司振书,殷国政,路建彪,李玉保

(聊城大学农学院,山东聊城 252000)

摘要:为 H9N2 亚型 AIV 进行快速、准确鉴定与诊断,根据 H9N2 亚型禽流感病毒的 HA、NA 基因各设计 1 对引物,建立了一种对 2 个表面基因同时进行扩增的双重 RT-PCR 检测方法。HA、NA 基因预期扩增的目的片段长度分别为 700、423 bp。通过对 H9N2 亚型禽流感病毒不同稀释度的尿囊液进行检测,证实病毒尿囊液的最低检出量为 $1 \times 10^{4.25}$ EID₅₀/100 μ L。采用该方法对 H1~H15 和 N1~N9 等亚型流感病毒及新城疫病毒等进行检测,结果只有 H9N2 亚型 AIV 出现 2 个特异性目的条带,与其他常见禽病原无交叉反应;用建立的双重 RT-PCR 和病毒分离 2 种方法同时对送检样品进行检测,结果 2 种方法的符合率达 99.77%,说明该检测方法特异性强、敏感性较高,在临床诊断和流行病学调查方面具有较好的应用前景。

关键词:禽流感病毒;H9N2 亚型;双重 RT-PCR;检测方法

中图分类号: S852.65⁺7 **文献标志码:** A **文章编号:** 1002-1302(2019)07-0035-03

H9N2 亚型禽流感病毒为中低致病力毒株,分布广泛,司振书等对山东省鸡群进行了 H9 亚型 AIV 的感染情况进行了调查,疑似流感鸡病料 RT-PCR 阳性率为 8.41%^[1]。该病毒能造成鸡群的免疫抑制,如果与其他病原协同感染则可引起禽类严重发病,如与金黄色葡萄球菌混合感染可引起肉鸡发病,死亡率达 6%^[2]。2013 年,我国发生了 H7N9 感染人事件,6 个内部基因全部来自于 H9N2 AIV^[3-4],1999 年来发生了多次 H9N2 亚型 AIV 感染人的事件^[5],给人类健康带来威胁。禽流感病毒有多种检测方法,包括血凝(HA)、血凝抑制(HI)试验,神经氨酸酶抑制(NI)试验,琼脂扩散试验(AGP),酶联免疫吸附试验(ELISA),病毒的中和试验(NT),免疫荧光抗体技术(IFA)和 real-time RT-PCR 等。司振书等建立了针对 H9 亚型 AIV 的 RT-PCR 检测方法^[6],包红梅

等建立了一步法 RT-PCR 检测方法,具有较好地特异性和敏感性^[7],但以上方法没有对 NA 基因进行扩增,不能确定 NA 亚型。司振书等对 H9N2 亚型禽流感病毒的 HA、NA 基因分别进行扩增,操作较麻烦且多耗试验材料^[8]。传统的血清学方法则需要多种亚型的血清和抗原,费时费力。根据 H9N2 亚型禽流感病毒的 HA 基因和 NA 基因各设计 1 对引物,初步建立了一种针对 H9N2 亚型 AIV 的双重 RT-PCR 检测方法,该方法特异性强、敏感性较高。该试验于 2012 年在中国农业大学动物流感研究室完成。

1 材料与方法

1.1 毒株

H9N2 亚型 AIV 标准参考毒株、H1~H15 亚型 AIV、N1~N9 亚型 AIV、鸡新城疫病毒、传染性支气管炎病毒、传染性法氏囊炎病毒、传染性喉气管炎病毒、马立克氏病病毒及鸡痘病毒等毒株,由中国农业大学动物流感研究室保存并提供。

1.2 引物的设计、合成与筛选

针对 HA、NA 基因的保守区序列分别设计特异性引物,用 Oligo 4.0 软件分析上下游引物是否匹配,将分析合格的引物送至生工生物工程(上海)股份有限公司北京合成部合成。

收稿日期:2017-11-12

基金项目:聊城大学博士启动基金(编号:318051310);聊城大学实验技术研究基金(编号:LDSY2014046)。

作者简介:司振书(1971—),女,山东冠县人,博士,副教授,主要从事动物病毒学研究。E-mail:dckszs@163.com。

通信作者:李玉保,博士,教授,主要从事家禽疫病防控研究。E-mail:liyubao@lcu.edu.cn。

[7] 宋海斌,崔喜波,马鸿艳,等. 基于 SSR 标记的甜瓜品种(系) DNA 指纹图谱库的构建[J]. 中国农业科学,2012,45(13):2676-2689.

[8] 叶卫军,杨勇,周斌,等. 分子标记在绿豆遗传连锁图谱构建和基因定位研究中的应用[J]. 植物遗传资源学报,2017,18(6):1193-1203.

[9] 周丽霞,吴翼,肖勇. 基于 SSR 分子标记的油棕遗传多样性分析[J]. 南方农业学报,2017,48(2):216-221.

[10] 黄海燕,杜红岩,乌云塔娜,等. 基于杜仲转录组序列的 SSR 分子标记的开发[J]. 林业科学,2013,49(5):176-181.

[11] 李炎林,杨星星,张家银,等. 南方红豆杉转录组 SSR 挖掘及分

子标记的研究[J]. 园艺学报,2014,41(4):735-745.

[12] 张振,张含国,莫迟,等. 红松转录组 SSR 分析及 EST-SSR 标记开发[J]. 林业科学,2015,51(8):114-120.

[13] 宗宇,王月,朱友银,等. 基于中国樱桃转录组的 SSR 分子标记开发与鉴定[J]. 园艺学报,2016,43(8):1566-1576.

[14] Gao L F, Tang J F, Li H W, et al. Analysis of microsatellites in major crops assessed by computational and experimental approaches[J]. Molecular Breeding, 2003, 12(3):245-261.

[15] 戚维聪,程计华,黄邦全,等. 基于海甘蓝 RNA-Seq 序列开发 EST-SSR 分子标记[J]. 江苏农业学报,2014,30(5):997-1002.