

甜瓜基因组 SSR 位点分布规律与多态性初探

高兴旺¹, 王贤磊², 杨雅婷², 张自强¹, 卢浩², 李冠²

(1. 新疆大学 资源与环境科学学院, 新疆 乌鲁木齐 830046; 2. 新疆大学 生命科学与技术学院, 新疆 乌鲁木齐 830046)

摘要:以 3 个甜瓜品种“BM7”、“伽师瓜”和“Y8”为试材, 选取甜瓜基因组的 3 个 DNA 区域进行 SSR 位点的分布规律分析, 进一步通过聚丙烯酰胺凝胶电泳检验甜瓜基因组 SSR 位点的多态性, 以期初步揭示甜瓜基因组中 SSR 位点的分布规律及其多态性, 进而为甜瓜及其它物种开展相关研究提供参考。试验表明: 甜瓜基因组中平均每 1.9 kb 就有一个 SSR 位点, 富含 A 和 T 的重复基序出现的频率远高于富含 C 和 G 的重复基序的频率, 没有发现全部由 C、G 构成的重复基序; 二、三、四碱基重复基序的 SSR 位点占绝大多数(83.3%), 五、六碱基重复基序的 SSR 位点分布较少(16.7%)。二、三、四、五、六核苷酸重复基序 SSR 位点平均长度依次为 21、15、14、16、20 bp。设计的 82 对 SSR 引物中 81 对扩增出产物, 多态比例最高为 48.1%。试验表明, 甜瓜基因组中 SSR 分布具有一定规律, 并且可以利用 SSR 分子标记开展相关研究; 试验同时对更好地开展 SSR 和 ISSR 分子标记的相关研究奠定了理论基础和依据。

关键词:甜瓜; 基因组; SSR; 分布; 多态性

中图分类号:S 652 **文献标识码:**A **文章编号:**1001-0009(2014)15-0093-05

SSR 全称为 Simple sequence repeat(s)^[1], 中文名称为简单重复序列, 又称串联重复序列 (Short Tandem Repeat, STR)^[2], 也称微卫星序列 (microsatellite)^[3], 是指以 1~6 个核苷酸为重复单元组成的多次串联重复的 DNA 序列, 序列全长一般在 100 个碱基以内。SSR 广泛存在于从酵母菌到动植物的几乎所有生物中^[4]。SSR 具有分布广, 共显性, 位点特异, 多态性高, 操作简单, 高度可重复等优点^[5-7]。SSR 分子标记技术在以下领域得到广泛应用, 如遗传图谱的构建、基因定位、分子标记辅助育种、种子纯度鉴定、亲子鉴定、遗传多样性研究等^[8]。尤其是随着测序技术的迅猛发展, 很多物种的全基因组序列可以下载获得, 采用相关软件就可以比较容易地在目标区域寻找并设计新的 SSR 分子标记。近年来, SSR 分子标记成为基因定位克隆、遗传图谱的构建、分子标记辅助育种等研究的强有力的常用技术, 获得越来越多的相关科研工作者的青睐。

在已有基因组序列可参考的物种中设计 SSR 分子

标记时, 科研人员往往会产生这样的迷惑, 在面对众多可选的 SSR 位点时, 究竟选择什么样的 SSR 位点设计引物才会得到较高的多态性, 从而高效定位克隆目标基因。这就需要对基因组的 SSR 位点的分布规律及其多态性进行一定分析。因此, 开展甜瓜基因组 SSR 位点分布规律及其多态性的相关研究具有一定现实意义。

甜瓜是一种重要的果蔬类经济作物, 在世界范围内广泛栽培, 属于新疆重要特色经济作物之一。随着甜瓜全基因组测序的完成^[9], 必将使得葫芦科这一模式植物的相关研究进程不断加快。分析甜瓜基因组中 SSR 位点的分布规律等相关信息, 可提高设计 SSR 标记时的选择效率, 提高可预见性, 同时也为其它物种 SSR 标记的开发利用提供一定参考。

目前还鲜见有关甜瓜基因组 SSR 位点分布规律方面的报道, 其它物种中的相关报道仅见于南方红豆杉^[10]、苹果^[11]及葡萄^[12]。该研究通过对甜瓜基因组 DNA 部分区域 SSR 位点的分布进行统计分析, 进一步通过试验检验甜瓜基因组 SSR 位点的多态性, 以期初步揭示甜瓜基因组中 SSR 位点的分布规律及其多态性, 为甜瓜及其它物种开展 SSR 的相关研究提供参考, 同时对开展 ISSR 的相关研究具有一定指导作用。

1 材料与方法

1.1 试验材料

用于 SSR 位点分布规律研究的甜瓜基因组数据下

第一作者简介:高兴旺(1983-), 男, 博士, 现主要从事植物分子生物学研究工作。E-mail: gwxw516@163.com

责任作者:李冠(1949-), 男, 硕士, 教授, 博士生导师, 现主要从事植物生理生化与分子生物学研究工作。E-mail: guanli@xju.edu.cn

基金项目:新疆高技术发展资助项目(201111120); 国家自然科学基金资助项目(31260258)。

收稿日期:2014-03-13

载自甜瓜基因组数据库 (<https://melonomics.net/>), 随机选取其中 3 条染色体(chr03、chr06、chr10)并分别从这 3 条染色体上随机截取 970.2 kb 的 DNA 序列用于数据分析, 选取的 3 个区域依次记为 A 区、B 区和 C 区。

用于 SSR 位点多态性分析的 3 个甜瓜品种为“BM7”、“伽师瓜”和“Y8”, 其幼嫩叶片采集自国家瓜类工程技术研究中心的试验地。

1.2 试验方法

1.2.1 SSR 位点分布规律分析 采用 SSRHunter 1.3^[13] 软件分析 A、B、C 区中 SSR 位点, 参数设置为构成重复基序的核苷酸数目最多为 6, 重复次数最少为 3。采用 Excel 和相关函数对生成的结果进行统计分析, 选取重复序列总长度不小于 12 的重复基序作为研究对象。

1.2.2 SSR 位点多态性分析 采用改良的 CTAB 法^[14] 从采集的甜瓜幼叶中提取基因组 DNA; 利用 Primer 5 设计 82 对 SSR 引物, 引物在北京华大基因研究中心合成; 以“BM7”、“伽师瓜”和“Y8”的基因组 DNA 作为模板, 采用常规 PCR 方法进行扩增, 聚丙烯酰胺凝胶电泳检测扩增产物多态性^[15]。

2 结果与分析

2.1 SSR 位点分布特征

由图 1 和表 1 可知, 3 个区域共发现 1 483 个 SSR 位点, A 区(970 192 bp)每 1.3 kb 就有一个 SSR 位点, B

区(970 230 bp)每 4 kb 有一个 SSR 位点, C 区(970 257 bp)每 1.8 kb 就有一个 SSR 位点, A 区 SSR 位点所占比例最高(48.1%), B 区 SSR 位点所占比例最低(16.4%), C 区 SSR 位点所占比例接近于 A 区(35.5%)。结果表明 SSR 位点的分布密度具有区域差异性。从 A 区和 C 区 SSR 位点分布可以看出, 二、三、四核苷酸重复基序的 SSR 位点占绝大多数(83.3%), 五、六核苷酸重复基序的 SSR 位点分布较少(16.7%)。整体来看, 平均每 1.9 kb 就有一个 SSR 位点(2 910 679 bp); 二、三核苷酸重复基序的 SSR 位点数最多(分别占 27.7%、28.7%), 六核苷酸重复基序的 SSR 位点数最少(5.5%)。

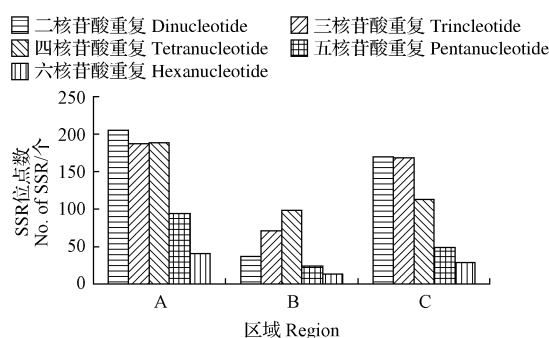


图 1 不同区域 SSR 位点数量分布特征

Fig. 1 Quantitative distribution characteristic of SSR loci in different region of melon genome

表 1 SSR 位点中不同核苷酸重复基序数量分布特征

Table 1 Quantitative distribution characteristic of SSR loci with different repeat motifs in different region of melon genome

区域及大小 Region and length	重复基序类型 Repeat type					合计 Total
	二核苷酸 Dinucleotide	三核苷酸 Trinucleotide	四核苷酸 Tetranucleotide	五核苷酸 Pentanucleotide	六核苷酸 Hexanucleotide	
A 区/970 192 bp	205	187	188	94	40	714
B 区/970 230 bp	37	71	98	24	13	243
C 区/970 257 bp	169	168	112	49	28	526
合计 Total	411	426	398	167	81	1 483

2.2 SSR 位点中不同核苷酸重复基序的组成特征

由图 1 可知, 二、三、四核苷酸重复基序的 SSR 位点占绝大多数。因此, 对 SSR 位点中二、三、四核苷酸重复基序的分布频率进行统计分析。考虑到序列互补特性和起始位点的不同, 二核苷酸重复基序的 SSR 位点可分为 3 类: AT/TA、AG/GA/CT/TC、CA/AC/TG/GT, 其中 AT/TA 重复所占比例最大, 并且在 A 区、B 区和 C 区中表现一致(图 2); 三、四、五、六核苷酸重复基序的类型更为多样, 但都表现出重复基序中 A、T 所占比例越高, 则该重复基序出现的几率越大的特征。所有重复基序中, 富含 A 和 T 的频率远高于富含 C 和 G 的频率, 没有发现全部由 C、G 构成的重复基序。

2.3 SSR 位点中不同核苷酸重复基序的长度分布特征

不同核苷酸重复基序 SSR 位点在长度上存在一定

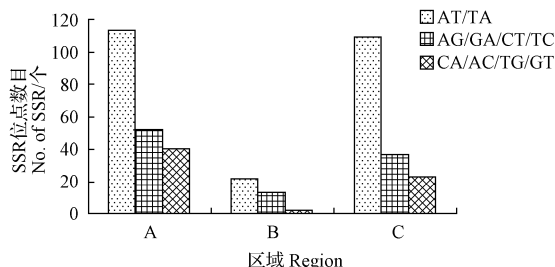


图 2 SSR 位点中不同二核苷酸重复基序类型的分布特征

Fig. 2 Distribution characteristics of different types of dinucleotide repeat motifs of SSR

变异, 由表 2 可知, 二、三、四、五、六核苷酸重复基序的 SSR 位点的长度变化范围依次为 12~74、12~78、12~36、15~30、18~36 bp, 二、三、四、五、六核苷酸重复基序

SSR 位点平均长度依次为 21、15、14、16、20 bp。结果表明,二、三核苷酸重复基序的 SSR 位点的长度变异较为明显,四、五、六核苷酸重复基序 SSR 位点的长度变异较小;不同核苷酸重复基序 SSR 位点的平均长度变异较小。

表 2 SSR 位点中不同核苷酸重复基序的长度分布特征

Table 2 Length distribution characteristics of SSR with different repeat motifs

区域 Region	二核苷酸 Dinucleotide			三核苷酸 Trinucleotide			四核苷酸 Tetranucleotide			五核苷酸 Pentanucleotide			六核苷酸 Hexanucleotide		
	百分比 Percentage /%	密度 Density /kb·个 ⁻¹	均长 Mean length /bp	百分比 Percentage /%	密度 Density /kb·个 ⁻¹	均长 Mean length /bp	百分比 Percentage /%	密度 Density /kb·个 ⁻¹	均长 Mean length /bp	百分比 Percentage /%	密度 Density /kb·个 ⁻¹	均长 Mean length /bp	百分比 Percentage /%	密度 Density /kb·个 ⁻¹	均长 Mean length /bp
A 区	28.7	4.7	20	26.2	5.2	15	26.3	5.2	14	13.2	10.3	17	5.6	24.3	22
B 区	15.2	26.0	24	29.2	13.7	14	40.3	9.9	13	9.9	40.4	15	5.4	74.6	19
C 区	32.1	5.7	20	31.9	5.8	16	21.4	8.7	14	9.3	19.8	17	5.3	34.7	19

2.4 SSR 位点多态性分析

在 A 区均匀选取 82 个 SSR 位点,在其两端设计 82 对引物,81 对引物可以成功扩增出产物。其中,21 对 SSR 引物在甜瓜 3 个品种(“BM7”、“伽师瓜”和“Y8”)中带型完全一致,没有表现出多态性(如图 3 中的 3G57);15 对 SSR 引物在这 3 个甜瓜品种中的带型完全不同(如图 3 中的 3G67);6 对 SSR 引物只在“BM7”中表现为多态,8 对 SSR 引物只在“伽师瓜”中表现为多态(如图 3 中

的 3G64),31 对 SSR 引物只在“Y8”中表现为多态(如图 3 中的 3G53)。结果表明,“BM7”和“伽师瓜”的多态性较为接近,“Y8”多态性最高。该结果与这 3 个甜瓜品种属性具有一定联系,“BM7”和“伽师瓜”都属于厚皮甜瓜,“Y8”属于薄皮甜瓜,分别属于 2 个不同的类群。该结果也进一步验证了 SSR 分子标记可有效用于甜瓜遗传多样性的分析。

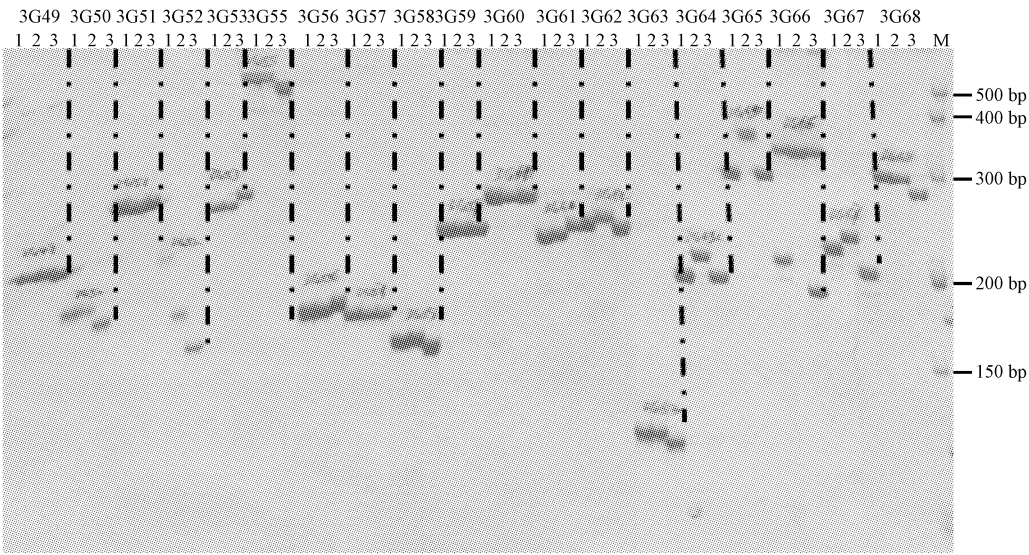


图 3 聚丙烯酰胺凝胶电泳检测 SSR 引物多态性

注:3G49~3G68:引物编号;1~3:模板 DNA 依次为“BM7”、“伽师瓜”和“Y8”。M 为 marker。

Fig. 3 Polyacrylamide gel electrophoresis results of SSR polymorphism

Note:3G49~3G68;Primer number;1~3;DNA template of melon line ‘BM7’, ‘Jiashi’ and ‘Y8’. M is marker.

3 讨论

通过对甜瓜基因组 3 个区域的 SSR 分布特征的分析可知,富含 A 和 T 的重复基序出现的频率远远高于富含 C 和 G 的重复基序的频率,这与其他学者在南方红豆杉^[10]、苹果^[11]及葡萄^[12]中的研究结果基本一致。该研究没有发现全部由 C、G 构成的重复基序,前人研究发现双子叶植物中 C/G 重复单元 SSR 位点含量远远少于单子叶植物^[16],该结果与前人研究结果相吻合。甜瓜基因组中二、三、四核苷酸重复基序的 SSR 位点占绝大多

数(83.3%),五、六核苷酸重复基序的 SSR 位点分布较少(16.7%),易官美等^[10]发现南方红豆杉中仅二核苷酸重复基序的 SSR 位点就占 88.6%,关玲等^[11]在苹果中发现一、二核苷酸重复基序的 SSR 位点占 91.67%,由此可以看出,不同物种中不同重复基序类型的 SSR 位点分布具有一定差异性。该研究发现甜瓜中平均每 1.9 kb 就有一个 SSR 位点,而苹果平均每隔 3.22 kb 存在 1 个^[11],二者存在一定差异。在 SSR 重复基序中,A 和 T 的含量远高于 C 和 G,短重复基序的含量也明显高于长重复基序,这些特征在甜瓜、南方红豆杉、苹果及葡萄中得到

了进一步验证;至于分布密度的差异可能是由于不同物种本身的特性导致的。

该研究仅对甜瓜基因组的部分区域进行了 SSR 位点分布特征的分析,为了提高结果的可信度,该研究随机选取了甜瓜基因组的 3 个区域进行相关分析。从 SSR 位点分布特征的分析结果来看,A 区和 C 区的分布特征较为相似,B 区的差异较大,这表明 SSR 位点在甜瓜基因组上的分布密度具有差异:多数区域的分布特征可能与 A 区和 C 区相似,少数区域与 B 区相似。从 SSR 位点多态性分析结果来看,具有 2~6 核苷酸重复基序的 SSR 引物均表现出多态性。由于不同类型的 SSR 引物数量有限,达不到统计分析的要求,但笔者建议在可选 SSR 标记充足的前提下,可优先选用二、三核苷酸重复基序的 SSR 引物,且优先考虑 A/T 含量相对较高的重复基序,最后考虑 G/C 含量相对较高的重复基序。在可选用的 SSR 标记有限的情况下也可以选用多核苷酸重复基序的 SSR 引物。

该研究为在其它不具有基因组序列的物种中进行 SSR 分子标记的开发工作具有一定指导意义。例如,对于双子叶植物,由于基因组中二、三核苷酸重复基序的 SSR 位点数占绝大多数,且重复基序中富含 A、T,而没有发现全部由 G、C 构成的重复基序,因此,在选用探针时就可以少用或者不用全部由 G/C 组成重复基序的引物,多选用以 AT/TA、AG/GA/CT/TC、CA/AC/TG/GT 组成重复基序的引物。由于 ISSR 分子标记技术是在 SSR 的基础上发展起来的一种分子标记技术,因此该研究也可以为 ISSR 引物的选择提供一定参考:同样,在选用 ISSR 引物时少用或者不用全部由 G/C 组成重复基序的引物,多选用以 AT/TA、AG/GA/CT/TC、CA/AC/TG/GT 组成重复基序的引物。

参考文献

[1] Howard J J, Klaus L, Stephen E L, et al. Genetic mapping of a gene causing hypertension in the stroke-prone spontaneously hypertensive rat[J]. Cell, 1991, 67(1): 213-224.

- [2] Edwards A L, Civitello A, Hammond H A, et al. DNA typing and genetic mapping with trimeric and tetrameric tandem repeats[J]. American Journal of Human Genetics, 1991, 49(4): 746.
- [3] Litt M, Luty J A. A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene[J]. American Journal of Human Genetics, 1989, 44(3): 397.
- [4] Tautz D, Renz M. Simple sequences are ubiquitous repetitive components of eukaryotic genomes[J]. Nucleic Acids Research, 1984, 12(10): 4127-4138.
- [5] Wayne P, Gordon C M, Jim P. Polymorphism revealed by simple sequence repeats[J]. Trends in Plant Science, 1996, 1(7): 215-222.
- [6] Schlötterer C. The evolution of molecular markers-just a matter of fashion? [J]. Nature Reviews Genetics, 2004, 5(1): 63-69.
- [7] Dimitry A C, Bart H, Filip A M, Volckaert. Microsatellites and their genomic distribution, evolution, function and applications: A review with special reference to fish genetics[J]. Aquaculture, 2006, 255(1-4): 1-29.
- [8] Swarup K P, Sanjay K K, Sunita K, et al. Informative genomic microsatellite markers for efficient genotyping applications in sugarcane[J]. Theoretical and Applied Genetics, 2009, 118(2): 327-338.
- [9] Jordi G M, Andrej B, Walter S, et al. The genome of melon (*Cucumis melo* L.) [J]. Proceedings of the National Academy of Sciences, 2012, 109(29): 11872-11877.
- [10] 易官美, 黎建辉, 王冬梅, 等. 南方红豆杉 SSR 分布特征分析及分子标记的开发[J]. 园艺学报, 2013, 40(3): 571-578.
- [11] 关玲, 章镇, 王新卫, 等. 苹果基因组 SSR 位点分析与应用[J]. 中国农业科学, 2011, 44(21): 4415-4428.
- [12] 蔡斌, 李成慧, 姚泉洪, 等. 葡萄全基因组 SSR 分析和数据库构建[J]. 南京农业大学学报, 2009, 32(4): 28-32.
- [13] Li Q, Wan J M. SSR Hunter: Development of a local searching software for SSR sites[J]. Yichuan, 2005, 27(5): 808-810.
- [14] Sue P L, Grant B, Bernard R B. Modification of a CTAB DNA extraction protocol for plants containing high polysaccharide and polyphenol components[J]. Plant Molecular Biology Reporter, 1997, 15(1): 8-15.
- [15] Ning X F, Wang X L, Gao X W, et al. Inheritances and location of powdery mildew resistance gene in melon Edisto47[J]. Euphytica, 2014, 195(3): 345-353.
- [16] Michele M, Michael H, Wayne P. Microsatellites are preferentially associated with nonrepetitive DNA in plant genomes[J]. Nature Genetics, 2002, 30(2): 194-200.

Distribution and Polymorphism Characteristics of SSR Loci in Melon Genome

GAO Xing-wang¹, WANG Xian-lei², YANG Ya-ting², ZHANG Zi-qiang¹, LU Hao², LI Guan²

(1. College of Resources and Environmental Science, Xinjiang University, Urumqi, Xinjiang 830046; 2. College of Life Science and Technology, Xinjiang University, Urumqi, Xinjiang 830046)

Abstract: Taking three kinds of melon including 'BM7', 'Jiashi' and 'Y8' as materials, three fragments of the melon genome sequence were randomly selected for investigation of distribution characteristics of SSR in melon genome; polyacrylamide gel electrophoresis was used to test the polymorphism of self designed SSR primers. The results showed that the mean distribution density of SSR was one SSR loci per 1.9 kb. Motif rich in A and T appears much higher than motif rich in C and G, motif comprised all by C and/or G was not found. The largest proportion of SSR loci was the

芫荽遗传多样性 RAPD 分析

李小梅, 张立微, 张景涛

(哈尔滨市农业科学院 蔬菜花卉分院, 黑龙江 哈尔滨 150029)

摘要:以 96 份芫荽为试材, 采用 RAPD 分析方法, 研究了芫荽种质资源遗传多样性。结果表明: 从 300 条随机引物中筛出 23 条扩增性强、清晰度高和稳定性好的引物, 扩增出 53 条带, 其中 34 条为多态性带, 多态性比例 64.2%。通过聚类分析, 生成 96 份芫荽亲缘关系树状图, 在遗传距离 0.50 处, 试材划分为 6 类。RAPD 分析揭示了芫荽的遗传背景并为其分类提供了依据。

关键词:芫荽; RAPD; 遗传多样性

中图分类号:S 573⁺.2 **文献标识码:**A **文章编号:**1001-0009(2014)15-0097-04

芫荽(*Coriandrum sativum*) 属伞形科一二年生草本绿叶蔬菜, 耐寒, 喜冷凉忌炎热, 芫荽可四季栽培, 春秋两季栽培品质好, 是重要的蔬菜和调味料^[1]。芫荽叶、根、茎、籽均可入药, 可治麻疹, 食物积滞, 感冒风寒等症^[2-3]。

种质资源是遗传育种和品种改良的基础^[4]。分子标记是种质资源亲缘关系及检测其多样性的有效工具^[5]。遗传多样性是物种长期进化的结果, 也是培育优良品种的基础^[6]。芫荽研究在我国集中在种植、成分分离和功能验证方面, 而在分子水平上研究芫荽遗传多样性在国内外鲜见报道。Lo'pez 等^[7]利用 AFLP 分子标记技术对芫荽进行了遗传多样性及聚类分析。DNA 分子标记的发展及应用, 为种质资源遗传多样性的研究提供了更多手段^[8]。RAPD 方法简便、灵敏、费用低, 可以检测到大量多态位点, 已被广泛地用于遗传多样性分析^[9]。该研究利用 RAPD 分子标记对芫荽遗传多样性进行分析, 确定其亲缘关系, 以期品种选育、生产加工、资源研究及保护提供理论依据^[10]。

第一作者简介:李小梅(1982-), 女, 硕士研究生, 农艺师, 现主要从事辣椒及芫荽栽培与育种等研究工作。E-mail: eileen828@126.com.

责任作者:张景涛(1963-), 男, 硕士, 研究员, 现主要从事蔬菜育种等研究工作。E-mail: chiillii@126.com.

收稿日期:2014-02-10

1 材料与方法

1.1 试验材料

以 2009、2010 年在哈尔滨市农业科学院试验基地种植的 96 份芫荽为试验材料, 材料名称及来源详见表 1。

供试仪器: PCR 仪(Eppendorf 公司), 离心机(Eppendorf 公司 Centrifuge 5804R), 紫外分光光度计(日本岛津公司 UV-2550), 电泳仪(北京六一厂 DYY-11 型), 凝胶成像分析系统(北京赛智创业科技有限公司 ChampGel 6000), 水浴锅, 超低温冰箱, Eppendorf 微量移液器。

供试试剂: dNTP 购自天泽基因工程有限公司, Buffer(10×PCR Buffer)、Taq DNA 聚合酶购自北京鼎盛生物技术有限公司; RNaseA、DNA Marker DL 2000 购自宝生物公司; 随机引物购自上海生工; CTAB、EDTA、Tris、HCl、NaCl、PVP、β-巯基乙醇、无水乙醇、氯仿、异戊醇等均为国产分析纯试剂。

1.2 试验方法

1.2.1 基因组 DNA 的提取 采用 CTAB 法提取芫荽总 DNA。

1.2.2 引物的筛选 从 300 条随机引物中筛选出多态性高的引物做 3 次重复。

1.2.3 基因组 DNA 的扩增和电泳检测 扩增体系为: 总体积 20 μL, 模板 DNA 浓度 1.0 ng/μL, dNTP 浓度

dinucleotide, trinucleotide and tetranucleotide repeat motifs, and the total ratio was 83.3%. The proportion of pentanucleotide and hexanucleotide was low, with the ratio of 16.7%. The average lengths of dinucleotide to hexanucleotide repeat motifs were 21, 15, 14, 16, 20 bp respectively. The PCR products can be amplified by 81 of total 82 designed self designed SSR primers successfully, and the maximal polymorphism was 48.1%. The results showed that the distribution of SSR in melon genome was relative stable, and it is helpful for future work in SSR and ISSR.

Key words: melon; genome; SSR; distribution; polymorphism