

胡麻 EST-SSRs 标记分布特征及 PCR 反应体系的优化

赵 玮, 党占海, 张建平, 王利民, 党 照, 李闻娟, 赵 利

(甘肃省农业科学院作物研究所, 甘肃 兰州 730070)

摘要: 从美国 NCBI 数据库中下载了 286 868 条亚麻 EST, 利用 SSRIT 软件检索出符合要求的 SSR 重复序列 4 310 条, 占整个 EST 数据库的 1.50%。其中三核苷酸出现频率最高, 占总 SSR 的 40.09%; SSR 重复类型总共 235 种, 序列中占总 SSR 比例大于 1% 的重复类型共有 21 种, 占 SSR 序列总数的 68.63%。通过正交设计(L_{16})得到了胡麻 EST-SSR 最佳 PCR 反应体系: 即 20 μ L 反应体系里, 引物 0.2 μ mol/L、模板 DNA 50 ng、dNTP 0.10 mmol/L、 Mg^{2+} 1.5 mmol/L、Taq DNA 聚合酶 0.50 U/ μ L、2 μ L 10 \times buffer, 其余用 ddH₂O 补齐。并通过单因素优化试验, 对得到的胡麻 EST-SSR 最佳 PCR 反应体系进行了验证, 结果完全一致。

关键词: 胡麻; EST-SSR 标记; PCR 反应体系; 优化

中图分类号: S565.9 **文献标志码:** A **文章编号:** 1001-1463(2016)10-0001-05

doi: 10.3969/j.issn.1001-1463.2016.10.001

The Distribution Characteristics of EST-SSRs Marker and Its Optimization of PCR Reaction System in Oil Flax

ZHAO Wei, DANG Zhanhai, ZHANG Jianping, WANG Liming, DANG Zhao, LI Wenjuan, ZHAO Li
(Institute of Crops Research, Gansu Academy of Agricultural Sciences, Lanzhou Gansu 730070, China)

Abstract: In this study, 286 868 expressed sequence tags (ESTs) are downloaded from the USA NCBI database. 4 310 appropriate simple sequence repeats (SSRs) are picked out with software of SSRIT. These SSRs accounted for 1.50% of the EST database and 40.09% of them is trinucleotide which is the maximum type. A total of 235 SSR types are got, 21 of them accounted for 68.63% total SSRs picked out in this study, and each of the 21 SSRs is more than 1% of the total SSRs. The optimal PCR reaction system for EST-SSR of flax is got by orthogonal design as follow: primer 0.2 μ mol/L; template DNA, 50 ng, dNTP, 0.10 mmol/L, Mg^{2+} , 1.5 mmol/L, Taq DNA polymerase 0.50 U/ μ L, 2 μ L 10 \times buffer, and at last, filling 20 μ L with ddH₂O. The consistent result is got from single factor optimization experiment on the optimal PCR reaction system for EST-SSR of flax.

Key words: Oil flax; EST-SSR marker; PCR reaction system; Optimizing

随着国内外基因组测序技术的飞速发展, 胡麻的基因组研究逐渐成为热点, 但胡麻可利用基因和分子标记资源的匮乏, 严重制约其发展, 因此急需开发更多有效的分子标记。同时通过表达序列标签(*expressed sequence tags*, EST)进

行 SSR 标记有利于植物分子水平的研究, 对种质资源多样性的分析、连锁图谱构建、比较基因组学等研究也具有现实意义^[1-3]。简单序列重复(*simple sequence repeats*, SSRs)又称微卫星(*microsatellites*)标记, 因具有共显性、高多

收稿日期: 2016-05-24

基金项目: 国家自然科学基金项目(31160291); 国家自然科学基金项目(31560347); 兰州市人才创新创业科技计划(2014-rc-44)部分内容。

作者简介: 赵 玮(1976—), 男, 甘肃景泰人, 副研究员, 研究方向为胡麻育种及抗逆生理。E-mail: 1048713911@qq.com。

通信作者: 赵 利(1974—), 女, 陕西武功人, 研究员, 硕士生导师, 主要从事胡麻遗传育种研究。E-mail: zhyj0801@126.com。

态性、多等位性、稳定性好以及操作简单等诸多优点而被广泛应用^[4-5]。胡麻即油用亚麻，属于亚麻科亚麻属，是一种重要的油料作物^[6-7]。胡麻油富含 α -亚麻酸和亚油酸这2种人体必需脂肪酸，具有促进人体智能、强身健脑、防止心血管疾病、抑制疾病基因等重要作用^[8-12]。随着对胡麻基因组学的深入研究，开发新的标记，进行分子标记辅助选择育种，对于提高胡麻育种效率具有重要意义。我们利用NCBI公布的亚麻EST序列查找EST-SSRs，通过分析SSR在亚麻中的分布特征设计SSR引物，并通过反应体系中引物和DNA浓度、 Mg^{2+} 、dNTP、Taq酶进行PCR反应体系优化，以期在亚麻高密度遗传图谱的构建、遗传多样性分析以及基因克隆和数量性状基因位点(QTL)定位等研究提供基础。

1 材料与方法

1.1 供试材料

供试材料为胡麻品种资源“STS”与“低亚麻酸”，由甘肃省农业科学院作物研究所提供。

1.2 试验方法

1.2.1 DNA的提取 采用改良CTAB法提取植物基因组DNA，用0.8%的琼脂糖凝胶和紫外分光光度计对其质量和浓度进行检测，提取的DNA溶解于100 μ l TE后放入4 $^{\circ}$ C冰箱备用。 Mg^{2+} 、dNTPs、Taq DNA聚合酶、Maker及引物合成及其他常规试剂均由自北京天根生物有限公司提供。

1.2.2 EST序列中SSR的检索 从NCBI(<http://www.ncbi.nlm.nih.gov/>)下载了286 868条亚麻EST，利用SSRIT (Simple Sequence Repeat Identification Tool)软件检索重复单元为二核苷酸、三核苷酸、四核苷酸、五核苷酸和六核苷酸的SSR，标准为重复序列长度 ≥ 20 bp，即二核苷酸、三核苷酸、四核苷酸、五核苷酸、六核苷酸重复次数分别大于或等于10、7、5、4、4。

1.2.3 L16正交反应体系的建立 对影响胡麻SSR-PCR反应体系的Taq酶浓度、 Mg^{2+} 含量、模板DNA用量、dNTP浓度及引物用量等5个因素采用正交设计在4个水平上进行试验，其中Taq酶浓度设定为0.5、1.0、1.5、2.0 U/ μ l； Mg^{2+} 含量

设为1.5、2.0、2.5、3.0 mmol/L；模板DNA用量设为50、100、150、200 ng；dNTP浓度设为0.1、0.2、0.3、0.4 mmol/L；引物用量设为0.2、0.4、0.6、0.8 μ mol/L。每管加入2.0 μ l 10 \times PCR loading buffer，反应体系20 μ l，其余用ddH₂O补足。试验设2次重复。体系优化使用的模板DNA为亲本STS，引物为ZS1 (PrimeA: CGACGACTTCAGCAT TG, primeB: TTGGTGGTCATCATCACTTATC)，ZS-2 (PrimeA: CTCCTTCAACTCGTCGTCT, prime B: TGCTGCTGCTACTTCACTC)。PCR运行程序设置为梯度PCR：94 $^{\circ}$ C预变性4 min，变性1 min；62 $^{\circ}$ C退火45 s；72 $^{\circ}$ C延伸1 min；94 $^{\circ}$ C变性1 min；55 $^{\circ}$ C退火45 s；72 $^{\circ}$ C延伸1 min；72 $^{\circ}$ C延伸10 min；16 $^{\circ}$ C延伸10 min；4 $^{\circ}$ C保存。

1.2.4 单因素试验优化反应体系 通过正交试验结果，将筛选出的反应体系作为PCR基本反应体系，再通过调节各因素对反应体系效果最好的浓度水平进行单因素优化试验，其他因素用量固定不变，进一步优化各反应成分的最终用量。单因素试验中各因素梯度设计为：引物(0.1、0.2、0.3、0.4、0.6、0.8、1.0 μ mol/L)；DNA模板(25、50、75、100、150、200、250 ng)；dNTP(0.05、0.10、0.15、0.20、0.30、0.40、0.50 mmol/L)； Mg^{2+} (0.25、0.50、0.75、1.00、1.50、2.00、3.00 mmol/L)；TaqDNA聚合酶(0.25、0.50、0.75、1.00、1.50、2.00、2.50 U/ μ l)。试验设3次重复。

1.2.5 电泳检测 凝胶电泳及银染扩增反应结束后，先在扩增产物中加入2 μ l 6 \times Loading buffer混匀，在8%的聚丙烯酰胺凝胶中，通过1 \times TBE的电极缓冲液中电泳检测。最后银染显色，照相保存。

2 结果与分析

2.1 EST序列中SSR的检索

从表1可知，从286 868条亚麻EST序列中共检测出符合要求的SSR重复序列4 310条，占整个EST数据库的1.50%。二、三、四、五和六核苷酸重复基元分别占总SSR的24.41%、40.09%、17.15%、9.07%和9.28%，其中三核苷酸出现频率最高，其次是二核苷酸和四核苷酸，五核苷酸和六核苷酸重复基元出现的频率较低且相近(表1)。

表1 EST-SSRs序列中不同重复核苷酸数的数量及百分比

重复核苷酸数	EST-SSRs条带数/条	占总EST-SSRs的比例/%
二核苷酸	1 052	24.41
三核苷酸	1 728	40.09
四核苷酸	739	17.15
五核苷酸	391	9.07
六核苷酸	400	9.28
合计	4 310	100.00

2.2 胡麻 EST-SSRs 标记分布特征

从表2可知, SSR重复类型总共235种。二核苷酸重复类型有8种, 其中, CT和AG出现频率较高, 分别占SSR序列总数的9.54%、和8.10%; 其次是TC, 占SSR序列总数的3.41%。三核苷酸重复类型共有54种, 以CTT出现的频

表2 EST-SSRs序列中不同重复类型及其数量和百分比

核苷酸名称	重复类型总数/种	重复类型	数量/条	占总SSR比例/%	占所在核苷酸种类的比例/%
二核苷酸	8	CT	411	9.54	39.07
		AG	349	8.10	33.17
		TC	147	3.41	13.97
		GA	59	1.37	5.61
		TA	45	1.04	4.28
三核苷酸	54	CTT	235	5.45	13.60
		TTC	133	3.09	7.70
		CCA	132	3.06	7.64
		GCT	122	2.83	7.06
		TCT	109	2.53	6.31
		GAA	96	2.23	5.56
		GAT	60	1.39	3.47
		TAA	49	1.14	2.84
		AAG	47	1.09	2.72
		AGA/CAG	43	1.00	2.49
四核苷酸	58	CTTT	341	7.91	46.14
		AGAA	87	2.02	11.77
		TACA	60	1.39	8.12
五核苷酸	60	TCTTC	206	4.78	52.69
		AGCTT	49	1.14	12.53
六核苷酸	55	CAGTAG	178	4.13	44.50

率最高, 占SSR序列总数的5.45%; 其次是TTC、CCA、GCT、TCT和GAA, 各占SSR序列总数的3.09%、3.06%、2.83%、2.53%和2.23%。四核苷酸重复类型出现58种, 其中CTTT出现的频率最高, 占SSR序列总数的7.91%, 其次是AGAA和TACA, 这2种核苷酸分别占SSR序列总数的2.02%和1.39%。五核苷酸重复类型60种, 其中以TCTTC出现的频率最高, 占SSR序列总数的4.78%。六核苷酸重复类型出现55种, 以CAGTAG出现的频率最高, 占SSR序列总数的4.13%。表2为EST-SSRs序列中占总SSR比例大于1%的重复类型, 这些重复类型总共有21种2 958条, 占SSR序列总数的68.63%, 其余214种核苷酸的重复类型分布比较分散, 出现频率均较低, 所占比例不足SSR序列总数的1%。

2.3 L_16 PCR反应体系的优化及建立

2.3.1 引物浓度优化 引物是PCR特异性扩增的关键, 引物浓度过低或过高都不能得出理想的扩增条带^[13-14]。PCR扩增结果显示(图1), 当引物浓度为0.1 $\mu\text{mol/L}$ 和0.2 $\mu\text{mol/L}$ 时扩增条带相对清晰; 浓度为0.50 $\mu\text{mol/L}$ 、0.75 $\mu\text{mol/L}$ 和1.00 $\mu\text{mol/L}$ 时, 随着引物浓度的增加条带逐渐变弱, 故选择扩增条带清晰、引物用量较少且背景较浅的0.20 $\mu\text{mol/L}$ 为最适反应浓度。

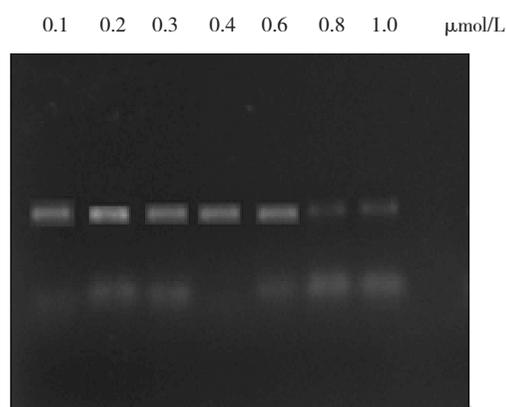


图1 引物浓度优化图谱

2.3.2 模板DNA浓度优化 DNA作为PCR反应的模板, 如果其含量过低则扩增产物减少, 含量过高则会引致非特异扩增^[15-16]。6个含量梯度的PCR扩增结果显示(图2), 当DNA含量达到50 ng后扩增产物的增加变得不明显, 且50-90 ng扩增

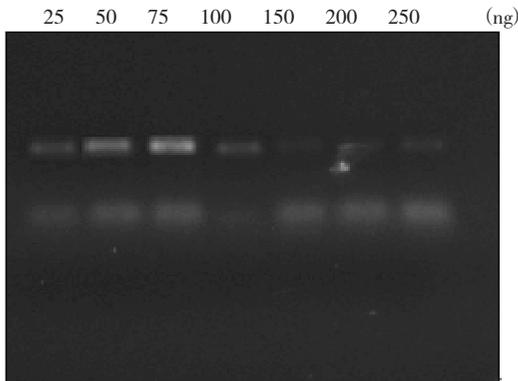


图2 模板 DNA 含量优化图谱

条带差异不大。因此,综合考虑,选择 50 ngDNA 含量为最适反应含量。

2.3.3 dNTP 浓度优化 dNTP 是 PCR 反应的原材料,其浓度过低或过高同样会影响 PCR 扩增效果。从实验结果分析(图3),当 dNTP 浓度为 0.05、0.10 mmol/L 时,扩增条带逐渐清晰,而浓度为 0.15、0.25、0.35、0.45 mmol/L 时,也同样能得到较清晰的条带,但随着浓度的增大扩增条带有逐渐变弱的趋势,结合成本等因素考虑,选择 0.10 mmol/L 为最适反应浓度。

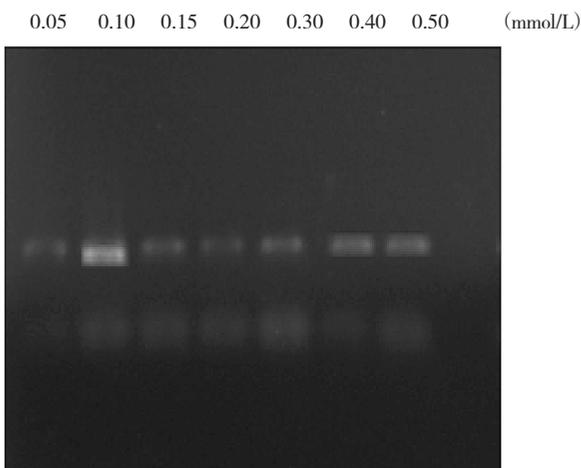
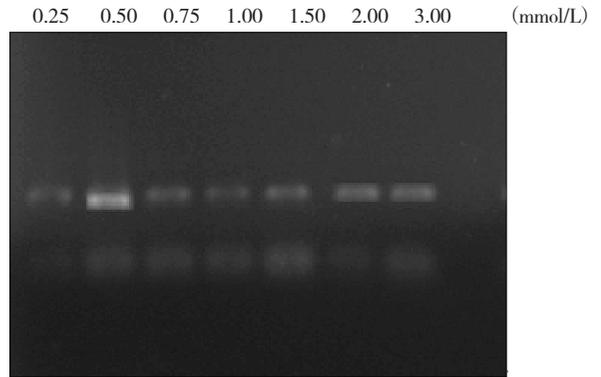


图3 dNTP 浓度优化图谱

2.3.4 Mg^{2+} 浓度优化 Mg^{2+} 浓度变化能够影响 DNA 聚合酶活性,并间接影响 PCR 扩增反应,浓度不合适往往会引起扩增失败。实验结果显示(图4),浓度 0.5、1.5 mmol/L 扩增条带清晰,而浓度为 2.0、3.0、4.0 mmol/L 时,且随着 Mg^{2+} 浓度的增大扩增条带明显逐渐变弱。我们综合各种因素,选择 1.5 mmol/L 为最适反应浓度。

图4 Mg^{2+} 浓度优化图谱

2.3.5 Taq DNA 聚合酶浓度优化 Taq DNA 聚合酶用量是影响的关键因素之一,其用量也会对 PCR 反应体系产生影响。实验结果表明(图5),Taq 聚合酶用量为 0.25、0.50 U/ μ l 时,扩增条带由弱变清晰,浓度为 1.0、1.5、2.0、2.5 U/ μ l 时,扩增条带依然较清晰,之后随着浓度的增大条带变弱,同时背景也在明显加深。选择 0.50 U/ μ l 为最适反应用量。

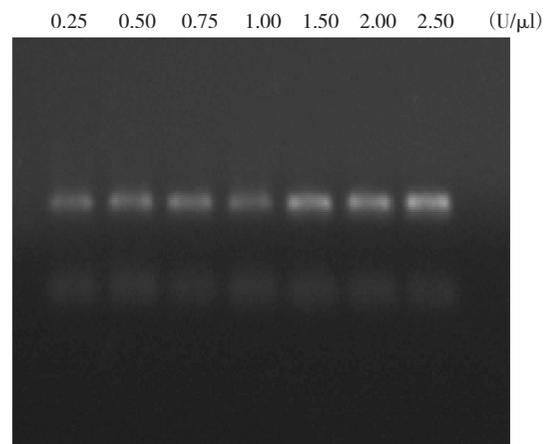


图5 Taq DNA 聚合酶浓度优化图谱

通过以上单因素试验的结果,得到的最佳反应体系浓度为:引物 0.2 μ mol/L、模板 DNA 50 ng、dNTP 0.10 mmol/L、 Mg^{2+} 1.5 mmol/L、Taq DNA 聚合酶 0.50 U/ μ l、2 μ L 10 \times buffer、ddH₂O 补齐至 20 μ L。

2.4 引物 ZS16 对优化反应体系的验证

以引物 ZS16 对筛选出的最佳反应体系进行稳定性及可靠性验证(图6)。在 10 个来源不同的胡麻品种中均扩增出了多态性条带,表明优化后的反应体系可以应用于胡麻 EST-SSR 标记的开发及

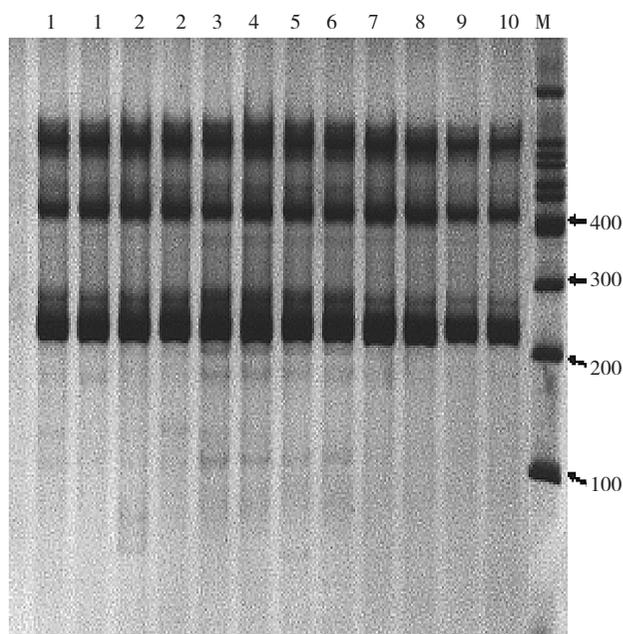


图6 ZS16在不同胡麻品种中的扩增条带

遗传多样性的分析等研究中。

3 小结与讨论

利用甘肃省农业科学院作物研究所胡麻课题开发的 EST-SSR 引物, 通过 L16 正交设计试验结合单因素设计, 首次对胡麻 EST-SSR 反应体系进行了优化, 得到胡麻 EST-SSR 最佳 PCR 反应体系为 0.20 $\mu\text{mol/L}$ 引物、50 ng 模板 DNA、0.10 mmol/L dNTP、1.5 mmol/L Mg^{2+} 、0.50 U/ μl Taq DNA 聚合酶、2 μL 10 \times buffer、ddH₂O 补齐至 20 μL 。经验证该体系能够扩增出了清晰而稳定的多态性条带。

SSR-PCR 反应体系的优化一般通过正交试验、单因素试验或完全试验进行。但是, 完全随机设计试验由于组合数较多导致工作量较大, 较费时费力, 因此目前大多选用单因素试验和正交试验。我们采用 L16 正交试验设计结合单因素试验, 对胡麻 SSR-PCR 反应体系中的主要成分进行了优化, 建立了一套适用于胡麻的 SSR-PCR 反应体系, 得到了理想的结果, 为今后胡麻 EST-SSR 分子标记的开发以及应用于指纹图谱的构建、遗传多样性分析、基因定位、比较基因组学等方面研究打下了基础。

参考文献:

[1] 杨金娥, 黄庆德, 周琦, 等. 冷榨和热榨亚麻籽油

挥发性成分比较[J]. 中国油料作物学报, 2013, 35(3): 321-325.

[2] 兰孟焦, 吴问胜, 黄水金, 等. 番薯属甘薯与牵牛 EST-SSR 标记开发及通用性分析[J]. 分子植物育种, 2015, 13(12): 2761-2766.

[3] 毕影东, 郭东林, 杨冬鹤, 等. 表达序列标签(EST)及其在抗孢囊线虫大豆研究中的应用[J]. 黑龙江农业科学, 2006(3): 90-93.

[4] 倪州献, 白天道, 蔡恒, 等. 马尾松基因组 SSR 标记在松属其他树种中的通用性分析[J]. 分子植物育种, 2015, 13(12): 2811-2817.

[5] 房冬梅, 吕品, 侯建华. 油葵 SSR-PCR 反应体系的优化及引物筛选[J]. 中国农学通报, 2015, 31(12): 205-209.

[6] 郝荣楷, 严兴初, 党占海, 等. 我国胡麻育成品种的遗传多样性分析[J]. 中国油料作物学报, 2014, 36(3): 334-342.

[7] 赵玮, 党占海, 张建平, 等. 胡麻胚状体诱导及植株再生的研究[J]. 北方园艺, 2010(21): 174-176.

[8] 陈海华. 亚麻籽的营养成分及开发利用[J]. 中国油脂, 2004, 29(6): 72-75.

[9] 王利民, 张建平, 米君, 等. 国外引进油用亚麻品种资源农艺性状分析与评价[J]. 中国油料作物学报, 2011(4): 356-361.

[10] 党照, 党占海, 杨崇庆, 等. 两个胡麻杂交组合在云南元谋制种播期试验[J]. 甘肃农业科技, 2013(9): 6-8.

[11] 张运晖, 赵瑛, 罗俊杰. 甘肃胡麻产业发展浅议[J]. 甘肃农业科技, 2013(7): 54-55.

[12] 王利民. 我国胡麻生产现状及发展建议[J]. 甘肃农业科技, 2014(4): 60-61.

[13] 魏利斌, 张海洋, 郑永战, 等. 芝麻 EST-SSR 标记的开发和初步研究[J]. 作物学报, 2008, 34(12): 2077-2084.

[14] 朱海生, 卢丽芳, 温庆放, 等. 南瓜 ISSR-PCR 反应体系的建立与优化[J]. 分子植物育种, 2014, 12(4): 802-809.

[15] 何仁锋, 冯尚国, 陈喆, 等. 药用菊花 SSR-PCR 反应体系优化及引物筛选[J]. 分子植物育种, 2015, 13(2): 367-378.

[16] 田艳伶, 李志辉, 杨模华, 等. 钩粟 ISSR-PCR 反应体系的建立与优化[J]. 中南林业科技大学学报, 2015, 35(2): 32-37.

(本文责编: 陈伟)