DOI:10.3969/j.issn.1000-3142.2013.04.005
张俊娥,李芬,孙蒙祥. 从烟草 EST 微卫星标记中筛选合子胚中优势表达的基因[J]. 广西植物,2013,33(4):456-459
Zhang JE, Li F, Sun MX. Polymorphic microsatellite markers from expressed sequences tag in *Nicotiana tabacum* for obtaining the genes of preponderant expression in zygotes[J]. Guihaia,2013,33(4):456-459

Polymorphic microsatellite markers from expressed sequences tag in *Nicotiana tabacum* for obtaining the genes of preponderant expression in zygotes

ZHANG Jun-E[1], LI Fen[2], SUN Meng-Xiang[3]
(1. *College of Life Sciences, Jiangxi Normal University*, Nanchang 330022, China; 2. *College of Life Sciences, Henan Normal University*, Xinxiang 453007, China; 3. *Key Laboratory of Ministry of Education for Plant Development Biology, College of Life Sciences, Wuhan University*, Wuhan 430072, China)

Abstract: In order to obtain the genes of preponderant expression in zygote cells of tobacco, CAP3 program was used to assemble ESTs from GenBank, and MISA program were used to screen microsatellites. Then polymorphic microsatellites were amplified in cDNA of tobacco zygotes library. The results showed that sixteen polymorphic SSR markers from tobacco ESTs deposited in a public sequence database were developed. Of these, nine genes could be successfully amplified in cDNA of tobacco zygotes library. The study would provide new channel to screen genes of preponderant expression in tobacco zygotes.

Key words: EST(expressed sequence tag); SSR(simple sequence repeats); genes of preponderant expression; zygotes; tobacco
中图分类号: Q948 **文献标识码:** A **文章编号:** 1000-3142(2013)04-0456-04

从烟草 EST 微卫星标记中筛选合子胚中优势表达的基因

张俊娥[1],李 芬[2],孙蒙祥[3]
(1. 江西师范大学 生命科学学院,南昌 330022; 2. 河南师范大学 生命科学学院,河南 新乡 453007; 3. 武汉大学 生命科学学院 植物发育生物学教育部重点实验室,武汉 430072)

摘 要: 为获得烟草合子胚中的优势表达基因,利用 CAP3 程序对来自烟草公共数据库的 EST 序列进行组装,利用 MISA 程序从组装后的 EST 中筛选 SSR 位点,将多态性的 SSR 位点在烟草合子文库中进行扩增并对等位基因进行分析。结果表明:具有多态性的 16 个 SSR 标记中,有 9 个基因能从烟草的合子库中成功扩增得到。该研究为筛选烟草合子胚中优势表达基因提供新的途径。

关键词: EST; SSR; 优势表达基因; 合子; 烟草

Tobacco(*Nicotiana tabacum*) is an important agricultural crop that plays a significant role in the economies of many countries(Davis & Nielsen,1999). Furthermore, tobacco is considered to be one of the most important model systems in biotechnology. Despite being an inbreeding, highly homozygous plant, as well as being of high-economic and scientific importance, it is surprising that the genetic analysis of tobacco is still in

收稿日期: 2012-11-06 修回日期: 2013-02-10
基金项目: 江西省自然科学基金(2010GZN0109); 江西省教育厅基金(GTT11386)
作者简介: 张俊娥,(1972-),女,山西榆社人,博士,副教授,研究方向为植物生物技术与遗传育种,(E-mail) zhangjune2001@aliyun.com。

its infancy. Although 684 microsatellite markers were screened from public domain databanks by Bindler *et al.* (2007) for a linkage map of tobacco, more microsatellite markers need to be developed to complement the studies. No markers can currently be used to study gene expression in early embryogenesis of tobacco.

Simple sequence repeats (SSRs) are multiallelic in nature, have abundant polymorphisms, high-information content, codominant inheritance, can be simply manipulated, and are easily detected by PCR. SSRs markers, therefore, have been widely used in genetic studies (Bindler *et al.*, 2007; Morgante & Oivieri, 1993; Powell *et al.*, 1996; Smulders *et al.*, 1997; He *et al.*, 2003). So far, a number of SSR markers have been identified from public EST databases and used for applications in economic plants, such as potato (Milbourne *et al.*, 1998), grape (Scott *et al.*, 2000), citrus (Jiang *et al.*, 2006), and tobacco (Bindler *et al.*, 2007). The 684 microsatellite markers of tobacco developed by Bindler *et al.* were predominantly derived from genomic sequences of the Tobacco Genome Initiative (TGI), through bioinformatics screening for microsatellite motifs. We searched GenBank to screen for another set of polymorphic microsatellite markers in order to obtain useful genes of preponderant expression in zygotes for studying gene expression in early embryogenesis of tobacco.

1　Materials and Methods

1.1 Plant materials

Twelve tobacco cultivars, "Burley 21", "Taiyan No. 7", "KY14", "Samsun", "Yunyan No. 87", "Souther Rhodesia", "SR1", "Smoke HAMA", "Nephrite", "Virginia331", "W38" and "Gexin No. 1" were used in present work. The plants were grown in a greenhouse at 25 ℃ with a 16 h light period.

1.2 Search for SSR loci

A total of 133 609 ESTs derived from tobacco were obtained by searching GenBank (http://www.ncbi.nlm.nih.gov/). These ESTs were assembled with CAP3 (Huang & Madan, 1999). The nonredundant dataset was used for screening microsatellites using the MISA program (http://pgrc.ipk-gatersleben.de/misa/). We adopted a criteria of searching for SSR loci with a minimum of 12-bp repeat length.

1.3 The primer pairs designed, polymorphism analysis and genes amplification from tobacco zygotes library

Several primer pairs flanking the SSR loci were screened for functionality. The primer pairs were designed using the software Primer3 (http://redb.ncpgr.cn/modules/redbtools/primer3.php), and were synthesized for detecting polymorphisms among tobacco cultivars. ESTs containing polymorphic microsatellites were searched against the GenBank nonredundant database using BLASTX algorithms (http://www.ncbi.nlm.nih.gov/BLAST) with expected value $< 10^{-7}$ for annotation of putative functions.

Twelve tobacco cultivars were employed for polymorphism evaluation. PCR conditions were as follows, 20-30 ng temple DNA, 0.2 μmol · L^{-1} of each primer, 0.2 mmol · L^{-1} dNTP, 10 mmol · L^{-1} Tris-Cl (pH 8.3), 50 mmol · L^{-1} KCl, 3 mmol · L^{-1} $MgCl_2$, 1 unit of Taq DNA polymerase in a volume of 20 μL. PCR was performed on PTC-100 Peltier thermal cycler with a program of 94 ℃ for 3 min, 32 cycles of 30 s at 94 ℃, 30 s at appropriate annealing temperature (Table 1), 30 s at 72 ℃, and a final extension of 4 min at 72 ℃. The PCR products were detected on ethidium bromide-stained 1.5% agarose gels, and then on denatured PAGE by silver staining.

Using the twenty primer pairs synthesized to amplify genes from tobacco zygotes library (The zygotes library from Key Laboratory of Ministry of Education for Plant Development Biology, College of Life Sciences, Wuhan University), PCR conditions of the amplification genes were as above. The PCR products were detected on ethidium bromide-stained 1.5% agarose gels.

2　Results and Analysis

A total of 133 609 ESTs derived from tobacco GenBank were assembled with CAP3, resulting in 64 732 unigenes (including 18 039 contigs and 46 693 singlets). The nonredundant dataset was analyzed for screening microsatellites using the MISA program. A

total of 40 894 SSR loci were identified in the unigenes. Twenty primer pairs flanking the SSR loci were screened and designed using the software Primer3, and were synthesized for detecting polymorphisms among tobacco cultivars.

Tobacco was an inbreeding, highly homozygous plant, so there were few polymorphisms among the tobacco cultivars. Twelve tobacco cultivars were employed for polymorphism evaluation for further screening out cross-parents with polymorphisms.

SSR markers were defined as either functional or non-functional (Bindler *et al.*, 2007). Of the 20 screened potential SSR markers, 17 primer pairs were functional. For the functional primers, 16 had polymorphisms among the tested tobacco cultivars, which were BP750276, EB445309, EB683490, EB683247, EB683081, EB683181, EB681632, EB682048, EB681345, EB679090, EB680649, EB683715, EB698114, EB690260, DW003872, DW004801. Of them, polymorphisms of DW004801 among the tested tobacco cultivars were shown in Fig. 1. Functions for the polymorphic SSR associated ESTs were determined by BLASTX in GenBank, and 11 SSR loci showed significant similarities to known genes (Table 1).

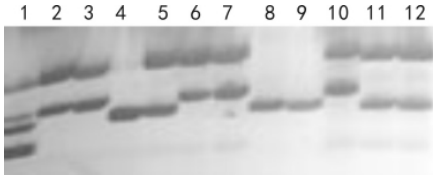

Fig. 1 PCR products of DW004801 on denatured PAGE by silver staining **1**. Burley 21; **2**. Taiyan No. 7; **3**. KY14; **4**. Samsun; **5**. Yunyan No. 87; **6**. Souther Rhodesia; **7**. SR1; **8**. Smoke HAMA; **9**. Nephrite; **10**. Virginia331; **11**. W38; **12**. Gexin No. 1.

After screening out the genes with polymorphisms, cross parents were confirmed for further studying gene expression in early embryogenesis of tobacco. The sixteen genes with polymorphisms were amplified from tobacco zygotes library. Nine genes (EB445309, EB683490, EB683181, EB681632, EB682048, EB683715, EB698114, EB690260, DW003872) could be successfully amplified in cDNA of tobacco zygotes library (Fig. 2). Their length was about 200 bp. Of the nine genes, six genes (EB683490, EB681632, EB682048, EB683715, EB698114, DW003872) showed significant similarities to known genes.

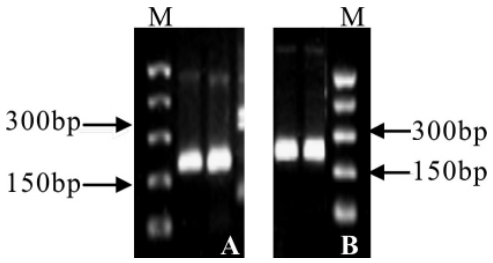

Fig. 2 Genes amplified from cDNA of tobacco zygotes library (two repeats) **A**. DW003872; **B**. EB683490; **M**. 50 bp DNA ladder.

3 Conclusion and Discussion

In order to study the expression pattern of genes during early embryogenesis in plants, enough viable and healthy cells of early embryogenesis must be isolated from them. In addition, appropriate markers must be obtained. We selected tobacco as the experimental material to study the expression pattern of genes during early embryogenesis in plants, as at present we have developed reliable technique to isolate healthy cells of early embryogenesis (including zygote, two-cell, four-cell, eight-cell, 16-cell, 32-cell, and so on) from tobacco. The technique resolves one main difficulty of above-mentioned, and it is also crucial to reliability of the experimental results. However, obtaining appropriate markers are also difficult, because polymorphism is few due to self-pollinated characteristic of tobacco. Twelve tobacco cultivars and twenty primer pairs flanking the SSR loci were choosed to select polymorphic parents for further studying the expression pattern of genes during early embryogenesis. The results showed that sixteen polymorphic SSR markers from tobacco ESTs deposited in a public sequence database were developed. After selecting out polymorphic SSR markers, they must be amplified in cDNA of tobacco zygotes library for further confirming whether or not expression in zygote cells. The results showed that nine genes can be successfully amplified in cDNA of tobacco zygotes library. By above

Table 1　Characteristics of 16 polymorphic EST-derived SSRs for tobacco

Locus (Accession No.)	Size range (bp)	Repeat unit	Primers sequence (5′-3′)	T_a (℃)	No. of allele	Protein homologue (Accession No.)	E-value
BP750276	189-239	$(TTA)_{16}$	F: TCTCCACTTCCTTGCTCCA R: CGCATGAAAAGAGACGTCAA	53	5	Pepsin A [Arabidopsis thaliana]. (NP_188636)	1e-28
EB445309	198-235	$(GCCAAT)_3$	F: AAGAAGGCAGAAGCACAACC R: TTCTCTTCCATGCTGTGGTG	53	3	No significant similarity	—
EB683490	203-241	$(CAA)_6$	F: AATCATTCGAACGCCAACTC R: GCGGACTATTGTCCTGCTTT	60	4	Avr9/Cf-9 rapidly elicited protein 194 [Nicotiana tabacum]. (AAG43553)	7e-55
EB683247	195-249	$(GA)_{25}$	F: CCTTCTCGGAACCAAACAGA R: GGCCAGCCACTTCAGTAATTT	53	5	No significant similarity	—
EB683081	193-241	$(TGG)_6$	F: GCACCTGAAAAAGGAAGTCG R: TATCATCTCTCCGCCACCA	60	3	hypothetical protein [Nicotiana tabacum]. (BAC53934)	2e-22
EB683181	204-233	$(TCA)_5$	F: GGGGCTAGAATCGGAACAA R: CCACTGTCCCTGATTTCCAT	53	3	Os05g0270400 [Oryza sativa (japonica cultivar-group)]. (NP_001055058)	6e-09
EB681632	193-229	$(CTT)_5$	F: ATCTCCGGTTCGTTGTTCAC R: ACGACGTCGGAGAACTCATC	53	2	Thioredoxin x [Arabidopsis thaliana] (AAF15952)	8e-46
EB682048	183-231	$(AAG)_6$	F: GAGATAATGGCGAAGGTGGA R: AGGGAGAGGCATAGGGTTGT	54	3	forever young oxidoreductase [Solanum bulbocastanum]. (AAL60069)	2e-113
EB681345	194-237	$(TCT)_6$	F: CGACCAAACTTGGGTTCAGT R: GTGGCCTTAGGGGAGAGAAC	55	3	unknown protein [Arabidopsis thaliana]. (NP_192933)	4e-91
EB679090	189-241	$(ACA)_6$	F: GATGCAGCCAGGAACACATA R: TGTGAAGTCCACTGCTTCCA	53	2	AN3(ANGUSITFOLIA3) [Arabidopsis thaliana]. (NP_198216)	2e-23
EB680649	195-236	$(CTG)_6$	F: TCACCTTCAGCTTTGTCGTG R: TTGTTTGCGGAAGCATACTG	52	3	No significant similarity	—
EB683715	172-203	$(TTG)_6$	F: CACCAGGTGCACAGAGACAT R: CAATCCGTTCCACATGTTTG	53	4	Os09g0493800 [Oryza sativa (japonica cultivar-group)]. (NP_001063548)	2e-07
EB698114	201-258	$(AAC)_6$	F: CCAGCAGATAAGGCAAATGA R: TCGCAATTGGTAAATCAGCA	50	6	PnFL-2 [Ipomoea nil] (AAG49896)	1e-30
EB690260	206-232	$(AG)_9$	F: GGGGAGGAAGATCTGATGAA R: CCCTTCTGGAAGGAGAAGAGA	54	2	No significant similarity	—
DW003872	190-212	$(GAA)_6$	F: AATCAATTCCAGGCTCATCG R: TTTTTCCCCTCTTATGCCACT	50	2	No significant similarity	—
DW004801	189-248	$(CGC)_6$	F: GACGCAGGAGGTGATGAAAT R: TATTGAGAAGCAACGCCACA	52	4	unknown protein [Arabidopsis thaliana]. (NP_565494)	9e-19

Notes: T_a. Annealing temperatures, putative homology acquired by searching the NCBI nonredundant database with BLASTX with an expected value $< 10^{-7}$.

work, a foundation of further studying the expression pattern of genes during early embryogenesis was laid.

It is obvious that the genes of preponderant expression in zygote cells can be found from tobacco public database, which will establish the foundation for further studying the genes expression of early embryogenesis of tobacco. These informative EST-SSR markers will be useful in studying gene expression in early embryogenesis of tobacco.

However, the technique of EST-SSR exists its limitation of few polymorphism while comparing it with the traditional SSR technique. For instance, low density is in constructing genetic map. It is thus clear that tobacco EST database for gene mapping and marker selection is relatively limited. The study used only the data for embryogenesis development, so polymorphism was not high. Therefore, the selected markers for the genes of preponderant expression in zygotes were not comprehensive. But, the method was still a very effective to obtain a part of genes of preponderant expression in zygotes.

(下转第 555 页 Continue on page 555)

现明显正相关,高浓度反而会削弱镇痛活性,由此可见,当油樟叶挥发油作为中枢镇痛药物时,其对浓度依赖性小。醋酸扭体试验是化学刺激剂致痛,引起的是内脏痛,主要是用于筛选外周镇痛药物。给予不同剂量的油樟叶挥发油的小鼠扭体次数均低于对照组,尤其是高剂量组和中剂量与溶剂对照组和油樟叶低剂量组有显著差异,提示油樟叶挥发油具有良好的外周镇痛活性,该生物学活性可能与 1,8-桉叶油素相关(Le *et al.*,2001)。

参考文献:

A dzu B, Amos S, Kapu SD, *et al.* 2003. Anti-inflammatory and anti-nociceptive effects of *Sphaeranthus senegalensis* [J]. *J Ethnopharmacol*, **84**(2-3): 169-173

Elisabetsky E, Amador TA, Albuquerque RR, *et al.* 1995. Analgesic activity of *Psychotria colorata* (Willd. ex R. & S.) Muell. Arg. alkaloids [J]. *J Ethnopharmacol*, **48**(2): 77-83

Le Bars D, Gozariu M, Cadden SW. 2001. Animal models of nociception [J]. *Pharmacol Rev*, **53**(4): 597-652

Lee HJ, Hyun EA, Yoon WJ, *et al.* 2006. *In vitro* anti-inflammatory and anti-oxidative effects of *Cinnamomum camphora* extracts [J]. *J Ethnopharmacol*, **103**(2): 208-216

Ling J, Liu WY. 1996. Cytotoxicity of two new ribosome-inactivating proteins, cinnamomin and camphorin, to carcinoma cells [J]. *Cell Biochem Funct*, **14**(3): 157-161

Luo ZJ(罗中杰), Li WY(李维一), Wei Q(魏琴), *et al.* 2001. The status Quo and future of *Cinnamomum longepaniculatum* in Yibin(宜宾油樟的现状及未来 [J]. *J Sichuan Norm Univ* (四川师范大学学报). **24**(3): 317-319

Santos FA, Rao VS. 2000. Antiinflammatory and antinociceptive effects of 1,8-cineole a terpenoid oxide present in many plant essential oils [J]. *Phytother Res*, **14**(4): 240-244

Wei Q(魏琴), Li Q(李群), Luo Y(罗扬), *et al.* 2006. Anti-fungal activity of leaf essential oil from *Cinnamomum longepaniculatum*(油樟油对植物病原真菌活性的抑制作用) [J]. *Chin J Oil Crop Sci* (中国油料作物学报), **28**(1): 63-66

Wei Q(魏琴), Zhou YK(周宇科), Zhou LJ(周黎军), *et al.* 2009. Inhibitive activity of *Cinnamomum* oil against bacteria(油樟油抑制细菌生长的活性试验) [J]. *Chin J Trop Agric* (热带农业科学), **29**(1): 5-7

Xu H, Liu WY. 2004. Cinnamomin-a versatile type II ribosome-inactivating protein [J]. *Acta Biochim Biophys Sin*, **36**(3): 169-176

(上接第 459 页 Continue from page 459)

Nine genes had been screened by this method and had been used in studying gene expression in early embryogenesis of tobacco. They had high similarity to known database. Now, we are carrying out verification of the genes screened to determine their biological function. In addition, we will also compare the screened genes with tobacco genome which have been publicized to open function in March 2012 in *Nicotiana tomentosiformis* Goodspeed genome and *Nicotiana sylvestris* genome for further obtaining some valuable information.

The study would provide an effective and applied way to find molecular markers for preponderant expression in zygotes.

References:

Bindler G, Hoeven R, Irfan GI. 2007. A microsatellite marker based linkage map of tobacco [J]. *Theoret Appl Gen*, **114**: 341-349

Davis DL, Nielsen MT. 1999. Tobacco-Production, Chemistry and Technology [M]. Oxford: Blackwell

He C, Poysa V, Yu K. 2003. Development and characterization of simple sequence repeat (SSR) markers and their use in determining relationships among *Lycopersicon esculentum* cultivars [J]. *Theoret Appl Gen*, **106**: 363-373

Huang X, Madan A. 1999. CAP3: a DNA sequence assembly program [J]. *Gen Resh*, **9**: 868-877

Jiang D, Zhang GY, Hong QB. 2006. Analysis of microsatellites in *Citrus* unigenes [J]. *Acta Gen Sin*, **33**(4): 345-353

Milbourne D, Meyer RC, Collins AJ, *et al.* 1998. Isolation, characterization and mapping of simple sequences repeat loci in potato [J]. *Mol & Gen Gen*, **259**: 233-245

Morgante M, Oivieri AM. 1993. PCR-amplified microsatellites as markers in plant genetics [J]. *Plant J*, **3**: 175-182

Powell W, Machray GC, Provan J. 1996. Polymorphism revealed by simple sequence repeats [J]. *Trends Plant Sci*, **1**: 215-222

Smulders MJM, Bredemeijer G, Rus-Kortekaas W, *et al.* 1997. Use of short microsatellites from database sequences to generate polymorphisms among *Lycopersicon esculentum* cultivars and accessions of other *Lycopersicon* species [J]. *Theoret Appl Gen*, **94**: 264-272

Scott KD, Eggler P, Seaton G, *et al.* 2000. Analysis of SSRs derived from grape ESTs [J]. *Theoret Appl Gen*, **100**: 723-726