

## 药用植物丹参 EST-SSR 标记的鉴定

邓科君<sup>1</sup>, 张勇<sup>1\*</sup>, 熊丙全<sup>2</sup>, 彭金华<sup>1</sup>, 张韬<sup>1</sup>, 赵晓楠<sup>1</sup>, 任正隆<sup>1\*</sup>

(1. 电子科技大学生命科学与技术学院, 四川 成都 610054; 2. 成都农业科技职业学院农学院园艺分院, 四川 成都 611130)

**摘要:** 对重要大宗药用植物丹参尚没有高效特异分子标记开发应用研究, 为有效利用丹参现有 EST 资源开发功能性分子标记, 本文以生物信息学方法对 GenBank 的 dbEST 数据库中公开的 10 288 条丹参 (*Salvia miltiorrhiza*) EST 进行整理, 共获得非冗余性 EST 4 192 条。从中搜索到 159 个 SSR 位点, 出现频率为 3.79%, 平均相隔 12.74 kb 出现 1 个 SSR 位点。在 1 bp~6 bp 核苷酸重复基序中, 二核苷酸重复基序 SSR 出现频率最高 (77 个, 48.43%), 其次是三核苷酸 (41 个, 25.79%), 四核苷酸出现频率最低 (6 个, 3.77%)。在检出的 47 类重复基序中, 出现频率超过 5% 的 6 种重复基序依次为: GA/CT (16.35%)、AG/TC (15.09%)、TCA/AGT (10.69%)、AT/TA (6.29%)、GAAAAG/CAAAAC (6.29%)、TA/AT (5.03%)。针对这些序列, 设计开发了 83 对 EST-SSR 引物, 对 13 个不同居群丹参样品及 10 种鼠尾草属物种进行了扩增效率、多态性及通用性检测, 发现其中 72 对引物在供试材料中可进行特异性扩增, 共产生 279 个位点, 平均每对引物产生 3.88 个位点, 且所有引物均显示多态性扩增。可有效扩增引物在鼠尾 10 个异源物种中的可转移率为 60%~100%, 平均 85%。依据扩增结果进行遗传相似性分析, 结果表明 EST-SSR 分析可揭示供试样品不同水平的遗传多样性, 并可明确区分丹参及其他同属植物。上述工作为进行丹参及其同属物种的基因组差异性分析、遗传图谱构建及比较基因组等研究奠定了基础。

**关键词:** 丹参; 表达序列标签; 简单序列重复; 分子标记; 遗传多样性

中图分类号: R931.5

文献标识码: A

文章编号: 0513-4870 (2009) 10-1165-08

## Identification, characterization and utilization of simple sequence repeat markers derived from *Salvia miltiorrhiza* expressed sequence tags

DENG Ke-jun<sup>1</sup>, ZHANG Yong<sup>1\*</sup>, XIONG Bing-quan<sup>2</sup>, PENG Jin-hua<sup>1</sup>,  
ZHANG Tao<sup>1</sup>, ZHAO Xiao-nan<sup>1</sup>, REN Zheng-long<sup>1\*</sup>

(1. School of Life Sciences and Technology, University of Electronic Science and Technology of China, Chengdu 610054, China;

2. Department of Agriculture and Horticulture, Chengdu Vocational College of Agricultural Science and Technology, Chengdu 611130, China)

**Abstract:** Despite *Salvia miltiorrhiza* being one of the most important medicine plants in China, there is a limited availability of genomic resources, especially of the expressed sequence tag-based markers. In this study, we selected and characterized functional markers in *S. miltiorrhiza*, which consisted of 4 192 non-redundant expressed sequence tags (ESTs) from 10 288 identified *S. miltiorrhiza* ESTs in dbEST data bank. Among them, 159 simple sequence repeats (SSR) were detected, which amounted to 3.79% of the non-redundant starting sequence population. This incidence was equivalent to one EST-SSR in every 12.74 kb of *S. miltiorrhiza* ESTs. Among the different motifs ranging from 1 bp to 6 bp, di-nucleotide repeat motif was the most abundant (77, 48.43%), followed by tri-nucleotide (41, 25.79%), hexa-nucleotide (23, 14.47%), penta-nucleotide (12, 7.55%) and tetra-nucleotide (6, 3.77%). In 47 identified motif types, the detected frequency above 5% were GA/CT (16.35%), AG/TC (15.09%), TCA/AGT (10.69%), AT/TA (6.29%), GAAAAG/CAAAAC (6.29%) and TA/AT

收稿日期: 2009-04-16.

基金项目: 国家自然科学基金资助项目 (30670211); 国家博士后科学基金 (20070411158); 电子科技大学青年基金资助项目.

\*通讯作者 Tel: 86-28-83203928, Fax: 86-28-83208238, E-mail: zhangyong916@uestc.edu.cn; renzl@uestc.edu.cn

(5.03%)。Based on flank sequence of detected SSR, a total of 83 EST-SSR primer pairs were designed and tested for the amplification efficiency, polymorphism and transferability in thirteen *S. miltiorrhiza* samples and other ten species from the genus *Salvia*. The results showed that 72 primer pairs were successfully amplified in *S. miltiorrhiza* samples to yield and 279 loci with an average of 3.88 loci per primer pair. The cross-transferability of *S. miltiorrhiza* EST-SSR markers to other ten *Salvia* plants was very high, ranging from 60% to 100% with an average of 85%. Further analysis of the genetic similarity based on the polymorphic bands showed the EST-SSR could detect the genetic diversity on different levels among the whole test samples and distinguish the *S. miltiorrhiza* from other *Salvia* plants effectively. It is expected that the potential markers described here would add to the repertoire of DNA markers needed for genetic analysis, linkage mapping and comparative genomics studies in *S. miltiorrhiza* and related *Salvia* genus plants.

**Key words:** *Salvia miltiorrhiza*; expressed sequence tag; simple sequence repeat; molecular marker; genetic diversity

丹参是著名的活血化瘀药,《中华人民共和国药典》<sup>[1]</sup>收录的唇形科鼠尾草属植物 *Salvia miltiorrhiza* 是其唯一的基源植物,以干燥的根及根茎入药,具有祛瘀止痛、活血通经、清心除烦的功效,用于治疗心绞痛、冠心病等,是我国传统医药学中应用最早而且最广泛的药物之一<sup>[2]</sup>。近年来,对丹参的化学成分、药理作用等研究日趋深入,但对其群体遗传多样性、种质分类鉴定、个体基因组构成等方面的相关研究报道相对较少。特别是在丹参研究中能够应用的分子标记技术十分有限,仅有少数应用 RAPD<sup>[3]</sup>、ITS<sup>[4]</sup>、AFLP<sup>[5]</sup> 及 ISSR<sup>[6]</sup>通用引物进行多样性研究的报道,尚未开发出能够应用于丹参种质鉴定、遗传图谱构建、功能基因定位等研究的简便、高效、稳定且具有种属特异性的分子标记体系。

简单重复序列 (simple sequence repeat, SSR), 是指 DNA 分子中 1~6 个核苷酸的串联重复,广泛分布于真核生物基因组的编码区和非编码区<sup>[7]</sup>。传统 SSR 标记的开发通常需经过文库构建,包括 SSR 克隆筛选、测序、引物设计和 PCR 扩增与分析等步骤,要投入大量的人力物力,开发成本高。从表达序列标签 (expressed sequence tag, EST) 中开发的 SSR 标记 (EST-SSR) 是近年发展起来的新型分子标记,与基因组 SSR 相比,它反映了基因组的编码区域,可以直接获得基因表达的信息,为功能基因提供“绝对”的标记,这有可能对决定重要表型性状的等位基因进行直接鉴定,省去了 SSR 引物开发过程中的克隆和测序步骤,充分利用了现有测序数据,降低了开发成本。该技术在不同物种间通用性好,具有较高的转移性<sup>[8]</sup>。自 2000 年起,已相继在水稻<sup>[8, 9]</sup>、甘蔗<sup>[10]</sup>、葡萄<sup>[11]</sup>、大麦<sup>[8, 12]</sup>、小麦<sup>[13-16]</sup>、杏<sup>[11, 17, 18]</sup>、云杉<sup>[19]</sup>、苜蓿<sup>[20]</sup>、猕猴桃<sup>[21]</sup>、马铃薯<sup>[22]</sup>、棉花<sup>[23, 24]</sup>、柑橘<sup>[25, 26]</sup>、甜瓜<sup>[27]</sup>、莴苣<sup>[28]</sup>、鹰嘴豆<sup>[29]</sup>等多种植物中开展了

EST-SSR 标记开发,并广泛用于遗传图谱构建、遗传多样性分析、功能基因定位和比较作图等研究。

截止 2009 年 1 月 16 日,在 GenBank (<http://www.ncbi.nlm.nih.gov/dbEST/>) 中释放有鼠尾草属植物 EST 共 11 748 条,其中包括 10 288 条丹参 EST,但目前还没有利用这些 EST 开发 SSR 标记的报道。本研究针对丹参 EST 进行 SSR 标记挖掘,并对这些 EST-SSR 在丹参转录组中的分布特征进行全面分析,在此基础上开发丹参 EST-SSR 标记,并进行丹参居群间多样性分析以及鼠尾草属植物种间可转移性验证。本文结果将为丹参 EST-SSR 标记的后续开发以及应用于多样性分析、遗传图谱构建、功能基因定位、重要经济性状标记和比较基因组学研究等奠定基础,为更加有效利用丹参及其他鼠尾草属药用植物资源提供分子依据。

## 材料与方法

**材料** 选取 13 份丹参材料,包括 7 份野生居群和 6 份栽培居群,同时选取 10 份鼠尾草属近源物种用于检测丹参 EST-SSR 标记的通用性。其中丹参样品采自我国不同地区,由南开大学陈瑞阳教授鉴定,种植保存于本室丹参种质资源圃。鼠尾草属近源物种由美国 NPGS 提供,具体信息详见表 1。

**基因组 DNA 提取** 每个样本随机选取 3 个单株,采幼叶等量混合,取 0.5 g 用于 DNA 提取,提取按参考文献<sup>[30]</sup>报道的方法并做修改<sup>[31]</sup>。DNA 的纯度和浓度用 0.8%琼脂糖凝胶和 BioSpec-mini 分光光度仪 (Shimadzu, Japan) 检测。最后将 DNA 稀释至约 200 ng·μL<sup>-1</sup>, -20 °C 保存备用。

**EST 序列获得及预处理** 2009 年 1 月,从 GenBank (<http://www.ncbi.nlm.nih.gov/dbEST/>) 中以 FASTA 格式下载共计 10 288 条丹参 EST 序列。采用

**Table 1** Description of *S.miltiorrhiza* and other species from the genus *Salvia* used in EST-SSR test

Number	Code	Species	Source	Remark
1	Sm01	<i>S.miltiorrhiza</i>	Neixiang county, Henan province	Wild species
2	Sm02	<i>S.miltiorrhiza</i>	Jintang county, Sichuan province	Wild species
3	Sm03	<i>S.miltiorrhiza</i>	Emei county, Sichuan province	Wild species
4	Sm04	<i>S.miltiorrhiza</i>	Zhongjiang county, Sichuan province	Cultivar
5	Sm05	<i>S.miltiorrhiza</i>	Zhongjiang county, Sichuan province	Cultivar
6	Sm06	<i>S.miltiorrhiza</i>	Zhongjiang county, Sichuan province	Cultivar
7	Sm07	<i>S.miltiorrhiza</i>	Zhongjiang county, Sichuan province	Cultivar
8	Sm08	<i>S.miltiorrhiza</i>	Zhongjiang county, Sichuan province	Cultivar, little leaf
9	Sm09	<i>S.miltiorrhiza</i>	Zhongjiang county, Sichuan province	Cultivar
10	Sm10	<i>S.miltiorrhiza</i>	Zhongjiang county, Sichuan province	Wild species
11	Sm11	<i>S.miltiorrhiza</i>	Zhongjiang county, Sichuan province	Wild species
12	Sm12	<i>S.miltiorrhiza</i>	Zhongjiang county, Sichuan province	Wild species
13	Sm13	<i>S.miltiorrhiza</i>	Zhongjiang county, Sichuan province	Wild species
14	Ssca	<i>S.scapiformis</i>	Emei county, Sichuan province	Wild species
15	Sspl	<i>S.splendens</i>	Emei county, Sichuan province	Wild species
16	Scav	<i>S.cavaleriei</i>	Emei county, Sichuan province	Wild species
17	Ssub	<i>S.substolonifera</i>	Emei county, Sichuan province	Wild species
18	Soff	<i>S.officinalis</i>	USA National Plant Germplasm System	Wild species
19	Svir	<i>S.viridis</i>	USA National Plant Germplasm System	Wild species
20	Sver	<i>S.verbenacea</i>	USA National Plant Germplasm System	Wild species
21	Stes	<i>S.tesquicola</i>	USA National Plant Germplasm System	Wild species
22	Sscl	<i>S.sclarea</i>	USA National Plant Germplasm System	Wild species
23	Sazu	<i>S.azurea</i>	USA National Plant Germplasm System	Wild species

VecScreen (<http://www.ncbi.nlm.nih.gov/VecScreen>) 及 RepeatMasker (<http://www.repeatmasker.org>) 去除载体污染和重复序列, 以 Phrap (<http://www.phrap.org>) 进行 EST 重叠群分析和聚类去除冗余 EST 序列。

**EST-SSR 的筛选及引物设计** 采用 SSRIT (<http://www.gramene.org/db/markers/ssrtool>) 在拼接后的非冗余 Uni-EST 序列中搜索 SSR 位点, 搜索的限制条件为不同重复基元的 SSR, 其总重复序列长度不低于 20 bp, 二核苷酸、三核苷酸、四核苷酸、五核苷酸、六核苷酸至少重复次数分别为 10、7、5、4、4, 包括中间被少数间隔碱基 (>5 bp) 打断的不完全重复 SSR。依据 EST-SSR 侧翼区域序列, 应用 Primer3 (<http://frodo.wi.mit.edu/primer3>) 进行引物设计, 引物主要参数为: GC 含量 40%~60%, Tm 值 45 °C~65 °C, 引物长度 18 bp~23 bp, 预期扩增产物长度 120 bp~350 bp。引物序列经 Blastn 比对, 根据结果重点对 3' 端碱基及上下游引物间隔进行调整, 最大限度保证引物 3' 端碱基与扩增位点相匹配以及有效扩增长度, 由 Invitrogen 公司进行合成 (Invitrogen, America) (引物信息见, <http://www.mcgb.uestc.edu.cn/>)。

**EST-SSR 有效性、通用性检测及评价** 首先对 13 个不同丹参居群样品进行扩增, 以检测引物设计的可靠性、扩增模式及其可用性。然后进一步以不同

鼠尾草近源物种进行 PCR 扩增, 检测丹参 EST-SSR 标记在鼠尾草属种间的通用性。25 μL 扩增体系包括 1×PCR buffer、1.5 mmol·L<sup>-1</sup> MgCl<sub>2</sub>、200 μmol·L<sup>-1</sup> dNTPs (Sigma, America)、0.5 μmol·L<sup>-1</sup> 上下游引物、50 ng 基因组模板 DNA 以及 0.5 U Taq DNA 聚合酶 (Takara, Japan)。以 I-Cycler 热循环仪 (Bio-RAD, America) 进行扩增, PCR 扩增程序为: 94 °C 预变性 3 min; 94 °C 变性 30 s、65 °C 退火 30 s (每循环降 1 °C)、72 °C 延伸 60 s, 10 个循环; 94 °C 变性 30 s、55 °C 退火 30 s、72 °C 延伸 60 s, 23 个循环; 72 °C 延伸 7 min。反应完成后, 产物中加入 6×loading buffer 5 μL 混匀, 于 8% 非变性连续 PAGE (Acr : Bis=39 : 1)、150 V 恒压电泳 2 h 左右, 按 Bassam 等<sup>[32]</sup>的方法银染显带、拍照、记录结果。

**基于 EST-SSR 的遗传相似性分析** EST-SSR 为共显性标记, 按照扩增条带的有无分别记为 1 和 0, 只选用 150 bp~500 bp 间且在两次重复 PCR 反应中均出现的清晰条带记分。其结果为一个 0、1 二元矩阵, 将其输入 NTSYS-pc 处理分析, 以 SIMQUAL 子程序计算样本间的 Jaccard 相似系数。之后, 基于遗传相似系数矩阵, 依据 UPGMA 算法并以 SHAN 子程序进行聚类分析。同样, 基于遗传相似系数矩阵, 以 DCENTER 和 EIGEN 子程序进行主成分分析 (PCO), 生成前两

个主成分的二维图。

结果

1 丹参 EST-SSR 发掘及特征分析

采用 VecScreen、RepeatMasker 及 Phrap 对 GenBank 的 dbEST 数据库中公布的丹参 EST 序列进行整理, 共获得总长约 2 026 kb 的 4 192 条 Uni-EST, 其中 contigs 1 276 条、singletons 2 916 条。从 4 192 条 Uni-EST 序列中共搜索到 159 个 SSR 位点, 占整个丹参 EST 序列的 3.79%, 平均每 100 kb 出现 7.85 个 SSR (表 2)。丹参 EST-SSR 种类也较为丰富, 二至六核苷酸重复类型均存在, 但出现的频率差异较大 (表 2)。二核苷酸重复单元检出频率最高, 共 77 个, 占所检出 SSR 总数的 48.43%, 其次是三核苷酸重复单元, 共 41 个, 占所检出 SSR 总数的 25.79%, 这两种类型核苷酸重复单元占多数, 合计占 SSR 总数的 74.22%。其余四核苷酸 (6 个, 3.77%)、五核苷酸 (12 个, 7.55%) 和六核苷酸 (23 个, 14.47%) 重复单元所占比例较小, 三者共计占 SSR 总数的 25.78%。

在检出的丹参 EST-SSR 中, 共观察到 47 种重复基序类型 (motif type), 其中二、三、四、五和六核苷酸重复基序类型分别为 6、13、6、9 和 13 种。出现频率超过 5% 的 6 种重复基序依次为: GA/CT (16.35%)、AG/TC (15.09%)、TCA/AGT (10.69%)、AT/TA (6.29%)、GAAAAG/CTTTTC (6.29%) 和 TA/AT (5.03%)。其余 41 种重复基序类型分布较为分散, 出现频率较低 (平均 0.98%、共计 40.25%) (图 1)。所有类型 SSR 重复单元长度在 20 bp~48 bp, 平均 SSR 单元长度为

27.2 bp。

2 丹参 EST-SSR 引物设计及扩增效率分析

针对丹参 EST 检出的 SSR 位点侧翼序列设计 PCR 扩增引物, 共得到 83 对 EST-SSR 引物, 涵盖了检出的所有 5 种 EST-SSR 重复类型, 其中二、三、四、五和六核苷酸重复类型分别为 36 (43.37%)、24 (28.92%)、5 (6.02%)、7 (8.43%) 和 11 (13.25%), 引物分布比例与相应 SSR 类型 (表 2) 基本相符。

Table 2 Occurrence of SSR in a set of non-redundant *S. miltiorrhiza* EST

Type	Number	Proportion in SSR/%	Frequency in test	Average distance /100 kb
Di-nucleotide	77	48.43	1.84	3.80
Tri-nucleotide	41	25.79	0.98	2.02
Tetra-nucleotide	6	3.77	0.14	0.30
Penta-nucleotide	12	7.55	0.29	0.59
Hexa-nucleotide	23	14.47	0.55	1.14
Total	159	100.00	3.79	7.85

选取 13 个不同居群丹参样品及 10 种鼠尾草属物种对设计的 EST-SSR 引物进行扩增效率及多态性分析。在 13 个丹参居群中, 72 对引物有特异性扩增条带 (86.75%), 5 对引物有非特异性扩增条带 (6.02%), 6 对引物无扩增产物 (7.23%)。其中 72 对引物共扩增 279 个位点, 平均每对引物产生 3.88 个位点, 所有引物在供试丹参样品中均显示多态性扩增。绝大多数在不同丹参居群中可进行有效扩增的 EST-SSR 引物在供试的其他 10 个异源物种中同样能产生有效扩增, 在鼠尾草属中的可转移率为 50%~100%, 平均为 85% (图 2)。

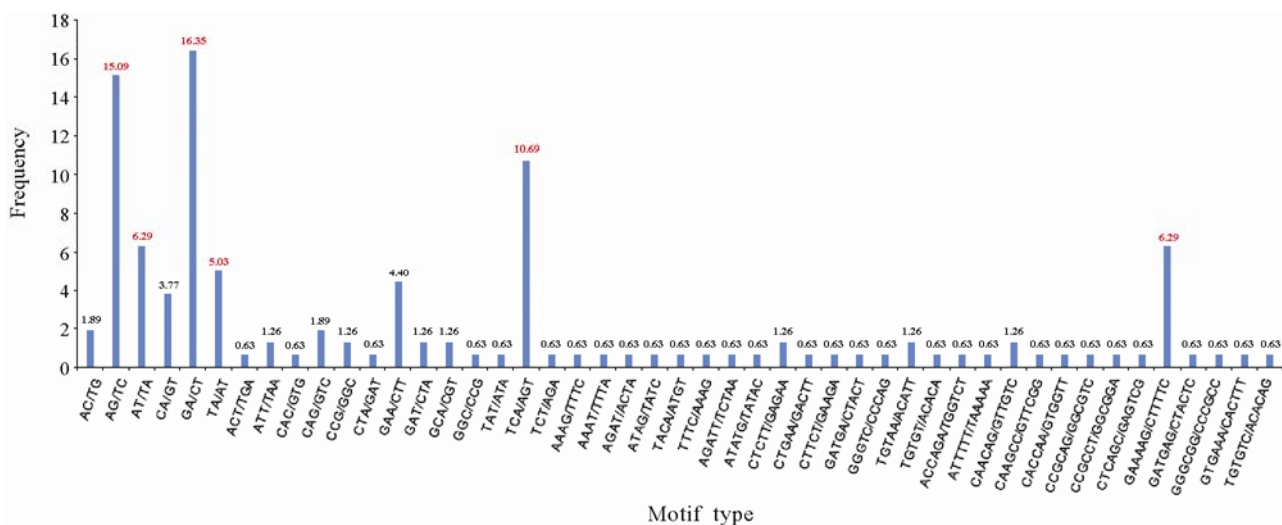
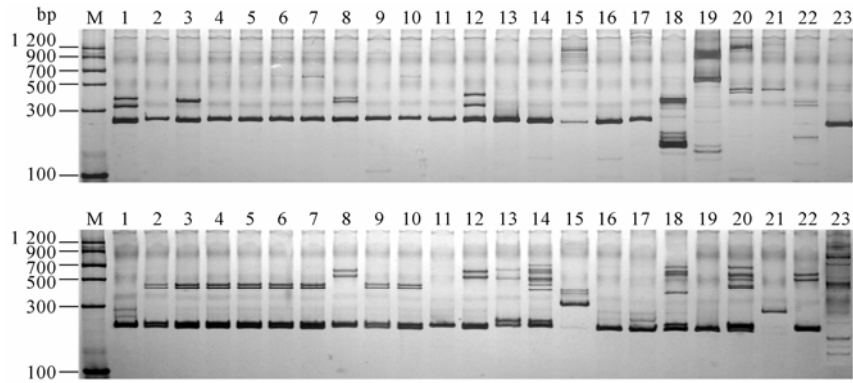
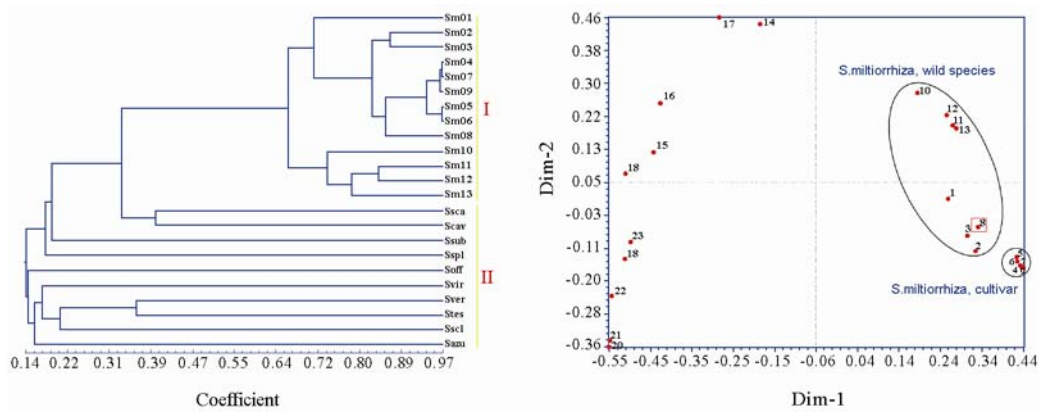


Figure 1 Frequency distribution of *S. miltiorrhiza* EST-SSR based on the motif sequence type



**Figure 2** Amplification of *S.miltiorrhiza* and other species from the genus *Salvia* with EST-SSR primer. (M: DNA molecular marker; 1-23: the numbers for each lane correspond with the plant given in Table 1)



**Figure 3** UPGMA (left) and PCA (right) analysis of genetic relationships of *S.miltiorrhiza* and other species from the genus *Salvia* based on the EST-SSR analysis

### 3 基于 EST-SSR 的遗传多样性分析

依据 EST-SSR 在丹参及其同属植物的扩增结果, 统计扩增位点并转化为 0、1 二元矩阵, 以 NTSYS pc21 计算供试材料间的 Jaccard 相似性系数, 得到的相似性系数介于 0.105 7~0.968 7, 均值为 0.357 5。丹参 13 个居群间相似性系数介于 0.588 2~0.968 7 (均值为 0.753 0), 其中采自中江的栽培丹参居群间相似性系数最高, 介于 0.831 9~0.968 7 (均值为 0.918 0)。与此相对, 在其他 10 种鼠尾草属植物间检测到较高的多样性, 相似性系数介于 0.105 7~0.398 1, 均值为 0.159 1。

依据 Jaccard 相似性系数, 以 UPGMA 算法对所有供试材料进行聚类分析, 结果表明 EST-SSR 分析可明确区分丹参及其他同属植物, 所有分析材料可聚为两大类: 丹参和其他鼠尾草属植物 (图 3 左)。将 Jaccard 相似性系数做 DCENTER 转换, 进行主成分分析, 结果显示供试丹参居群间存在丰富的多样性, 野生居群与栽培居群间分化明显 (图 3 右)。

### 讨论

本研究从丹参 Uni-EST 中鉴定出 159 个 SSR 位点, 平均每 12.74 kb EST 序列含 1 个 SSR 位点。基于不同的 EST-SSR 搜寻标准, 不同研究者已经对许多植物基因组中的 EST-SSR 分布特征进行了研究。在拟南芥中, 依据 Cardle 等<sup>[33]</sup>的统计标准, 平均 13.8 kb 检出一个 SSR 位点; 依据 Morgante 等<sup>[7]</sup>的统计标准, 平均 2.1 kb 可检出一个 SSR 位点。同样, 在水稻中按照 Cardle<sup>[33]</sup>的统计标准, 平均 3.4 kb 可检出一个 SSR 位点; 而按照 Gao 等<sup>[13]</sup>的统计标准, 则平均 11.8 kb 才可检出一个 SSR 位点。不同研究结果的差异主要归因于发掘 EST-SSR 时所采用的标准不一样 (同时也和 EST 不同的物种、组织、时期来源不同有关), 在水稻中当搜索标准由 12 bp 增加到 30 bp 时, EST-SSR 检出率从 50% 减少到 1%<sup>[34]</sup>。在本研究中, 使用了更为严格的标准 (二核苷酸、三核苷酸、四核苷酸、五核苷酸、六核苷酸至少重复次数分别为 10、7、5、4、4), 但 EST-SSR 检出率依然高于小麦 (15.6 kb、

17.42 kb)<sup>[8, 13]</sup>、玉米 (28.32 kb) 和大豆 (23.80 kb)<sup>[13]</sup>、棉花 (14.8 kb)<sup>[23]</sup>、拟南芥 (13.83 kb) 和杨树 (14.0 kb)<sup>[33]</sup>，这充分说明丹参 EST 中分布着非常丰富 SSR 位点。

从现有的报道分析，大多数植物 EST-SSR 主要属二、三核苷酸重复基序类型，但主导重复基序的类型则有所差异。例如，在拟南芥、小麦<sup>[8]</sup>、甘蔗<sup>[10]</sup>、大麦<sup>[12]</sup>、棉花<sup>[24]</sup>、柑橘<sup>[25]</sup>、水稻、玉米和大豆<sup>[33]</sup>中以三核苷酸重复为主；而在杏、桃、蔷薇<sup>[18]</sup>和猕猴桃<sup>[21]</sup>中则以二核苷酸重复为主。本文研究表明，丹参 EST-SSR 种类丰富，二至六核苷酸重复类型均存在，以二核苷酸重复基序检出频率最高 (77 个、48.43%)，三核苷酸重复基序次之 (41 个、25.79%)。然而需要指出的是，EST-SSR 基序检出类型也与检索标准密切相关，在水稻中当搜索标准由 12 bp 增加到 30 bp 时，二核苷酸重复基序数量从三核苷酸重复基序的 1/10 增加至二者基本接近，而当搜索标准增加到 40 bp 时，二核苷酸重复基序数量则超过了三核苷酸重复基序<sup>[34]</sup>。本文中，当将丹参 EST-SSR 搜索标准进行不同严谨度设定时，检出的 EST-SSR 主导重复基序类型在二、三核苷酸重复中有所不同。

在小麦、水稻、玉米、大豆中二核苷酸重复基序出现频率最高的都是 AG/TC<sup>[8]</sup>，而三核苷酸重复基序主要类型在不同植物中表现各异<sup>[8, 12, 23-25, 33]</sup>。本文中，在检出的丹参 EST-SSR 中，共观察到 47 种重复基序类型，其中两种二核苷酸重复基序以 GA/CT (16.35%) 和 AG/TC (15.09%) 为主，而三核苷酸重复基序则以 TCA/AGT (10.69%) 为主，这种表现与前人的研究结果基本相符。

相关研究中，EST-SSR 标记扩增效率在 60%~90%，而当设计引物跨越 mRNA 剪切位点、扩增产物包含的内含子过长或者扩增区域存在比较复杂的高级结构时，均会造成比例不等的 EST-SSR 无效扩增<sup>[18, 24, 25]</sup>。针对丹参 EST 序列检出的 SSR 位点侧翼序列设计 PCR 扩增引物，共筛选到 72 对可产生清晰可辨扩增位点的 EST-SSR 引物，占全部设计引物的 86.75%。获得相对较高的扩增效率，可能原因是在引物设计时重点注意了对 3' 端碱基及上下游引物间隔进行调整，使得最大限度的保证了引物 3' 端碱基与扩增位点的匹配及有效扩增长度，从而提高了扩增有效性。

多态性是评价分子标记的一个重要依据，一般来说 EST-SSR 标记在揭示多态性方面要低于基因组

SSR 标记<sup>[15]</sup>。但在不同的研究中，EST-SSR 标记依然有较高的多态性检出率，例如：Saha 等<sup>[35]</sup>用高秆酥油草开发的 EST-SSR 对黑麦草和高秆酥油草两亲本进行筛选，多态检出率为 66%；Wang 等<sup>[23]</sup>用雷蒙德氏棉 EST-SSR 标记检测海、陆四倍体栽培棉种间多态性，多态检出率为 43%。依据 Temnykh 等<sup>[9]</sup>在水稻不同品种间的分析发现，当 SSR 长度大于或等于 20 bp 时，多态性检出率高于长度在 12~20 bp 的 SSR，而长度小于 12 bp 以下的 SSR 多态性检出率十分有限，并且二核苷酸重复基序的 EST-SSR 揭示多态性水平高于其他类型的重复基序<sup>[36]</sup>。本研究中，将 EST-SSR 检索条件设定为 20 bp，最大限度的过滤了多态检出率不高的 EST-SSR 位点。同时，在丹参 EST 中，二核苷酸重复基序检出频率最高 (48.43%)，在所设计的引物中也体现了这种二核苷酸 EST-SSR 位点的高比例分布 (43.37%)。因此，本实验中所有可产生有效扩增的 EST-SSR 引物在供试的 13 个丹参居群材料中均可检测到多态性分布，这与之前的研究结果相一致<sup>[9, 36]</sup>。另外需要指出的是，目前的 EST-SSR 分析多集中于小麦、黑麦、玉米、水稻、棉花等经过长期人工选择的栽培物种中，多态性检测多以不同的栽培品种为材料，而丹参作为药用植物，大多以野生资源形式加以利用，较少受到人工选择的干预，不同地理及栽培居群间遗传异质性高于大多数栽培作物，这种检测材料的特殊性应该也是本文中 EST-SSR 多态性检出率较高的一个原因。

以丹参 EST 开发的 SSR 位点扩增引物在鼠尾草属 10 个异源物种中同样能进行有效扩增，可转移率为 50%~100%，平均为 85%。依据 EST-SSR 的扩增模式进行聚类分析，在鼠尾草属植物种间、丹参不同地理居群间、相同区域内野生及栽培丹参居群间的水平上均可进行有效区分，体现了不同程度的遗传分化及多样性分布 (图 3)。基于 EST-SSR 的分析结果与之前应用 RAPD、ITS 和 AFLP 得出的相关结果总体相近，这充分证实了 EST-SSR 在遗传多样性分析方面的应用能力。然而，由于本文所用材料相关物种测试数目有限，所检测到的种间及种内居群间的遗传变异还不够全面，有待于增加测试材料和使用引物的数目，以取得更有说服力的结论。

## 结论

本文在系统整理丹参 EST 序列的基础上，首次对蕴藏于丹参 EST 中的 SSR 位点进行了分析，成功

开发出一套适用于丹参及鼠尾草属植物的高效、高特异性 EST-SSR 扩增引物, 进而对供试材料进行了遗传多样性分析。本研究的结果证明, 以丹参 EST 开发 SSR 标记是可行的, 丹参 EST-SSR 标记的开发丰富了丹参及鼠尾草药用属植物的遗传标记。总之, 丹参 EST-SSR 标记的建立对于加速丹参 EST 资源的进一步发掘和利用、种质多样性研究、遗传作图、辅助选择、基因挖掘以及相关分子生物学研究奠定了基础, 同时也为其他药用植物进行分子基础研究拓展了空间。

## References

- [1] Chinese Pharmacopoeia Committee. Pharmacopoeia of the People's Republic of China (中华人民共和国药典) [S]. Beijing: Chemical Industry Press, 2005: 52.
- [2] Du GH, Zhang JT. The general situation and progress of the modern research of red sage root (radix *Salviae miltiorrhizae*) [J]. Her Med (医药导报), 2004, 23: 355-360.
- [3] Guo BL, Lin S, Feng YX, et al. Primary research on genetic relationship among main populations of *Salvia miltiorrhiza* and genuineness of herb [J]. Chin Tradit Herb Drugs (中草药), 2002, 33: 1113-1116.
- [4] Wang Y, Li DH, Zhang YT. Analysis of ITS sequences of some medicinal plants and their related species in *Salvia* [J]. Acta Pharm Sin (药学报), 2007, 42: 1309-1313.
- [5] Wang B, Zhang Y, Chen CB, et al. Analysis on genetic diversity of different *Salvia miltiorrhiza* geographical populations in China [J]. China J Chin Mater Med (中国中药杂志), 2007, 32: 1988-1991.
- [6] Xu H, Wang YY, Wei DH, et al. Genetic analysis and identification of radix et rhizoma *Salvia miltiorrhizae* from different habitats by ISSR [J]. Tradit Chin Drug Res Clin Pharmacol (中药新药与临床药理), 2007, 18: 454-457.
- [7] Morgante M, Hanafey M, Powell W. Microsatellites are preferentially associated with nonrepetitive DNA in plant genomes [J]. Nat Genet, 2002, 30: 194-200.
- [8] Kantety RV, La Rota M, Matthews DE, et al. Data mining for simple sequence repeats in expressed sequence tags from barley, maize, rice, sorghum and wheat [J]. Plant Mol Biol, 2002, 48: 501-510.
- [9] Temnykh S, DeClerck G, Lukashova A, et al. Computational and experimental analysis of microsatellites in rice (*Oryza sativa* L.): frequency, length variation, transposon associations, and genetic marker potential [J]. Genome Res, 2001, 11: 1441-1452.
- [10] Cordeiro GM, Casu R, McIntyre CL, et al. Microsatellite markers from sugarcane (*Saccharum spp.*) ESTs cross transferable to *erianthus* and *sorghum* [J]. Plant Sci, 2001, 160: 1115-1123.
- [11] Decroocq V, Favé MG, Hagen L, et al. Development and transferability of apricot and grape EST microsatellite markers across taxa [J]. Theor Appl Genet, 2003, 106: 912-922.
- [12] Thiel T, Michalek W, Varshney RK, et al. Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.) [J]. Theor Appl Genet, 2003, 1106: 411-422.
- [13] Gao LF, Jing RL, Huo NX, et al. One hundred and one new microsatellite loci derived from ESTs (EST-SSRs) in bread wheat [J]. Theor Appl Genet, 2004, 108: 1392-1400.
- [14] Gao LF, Tang JF, Li HW, et al. Analysis of microsatellites in major crops assessed by computational and experimental approaches [J]. Mol Breed, 2003, 12: 245-261.
- [15] Nicot N, Chiquet V, Gandon B, et al. Study of simple sequence repeat (SSR) markers from wheat expressed sequence tags (ESTs) [J]. Theor Appl Genet, 2004, 109: 800-805.
- [16] Peng JH, Lapitan NL. Characterization of EST-derived microsatellites in the wheat genome and development of eSSR markers [J]. Funct Integr Genomics, 2005, 5: 80-96.
- [17] Xu Y, Ma RC, Xie H, et al. Development of SSR markers for the phylogenetic analysis of almond trees from China and the Mediterranean region [J]. Genome, 2004, 47: 1091-1104.
- [18] Jung S, Abbott A, Jesudurai C, et al. Frequency, type, distribution and annotation of simple sequence repeats in Rosaceae ESTs [J]. Funct Integr Genomics, 2005, 5: 136-143.
- [19] Rungis D, Bérubé Y, Zhang J, et al. Robust simple sequence repeat markers for spruce (*Picea spp.*) from expressed sequence tags [J]. Theor Appl Genet, 2004, 109: 1283-1294.
- [20] Barrett B, Griffiths A, Schreiber M, et al. A microsatellite map of white clover [J]. Theor Appl Genet, 2004, 109: 596-608.
- [21] Fraser LG, Harvey CF, Crowhurst RN, et al. EST-derived microsatellites from Actinidia species and their potential for mapping [J]. Theor Appl Genet, 2004, 108: 1010-1016.
- [22] Feingold S, Lloyd J, Norero N, et al. Mapping and characterization of new EST-derived microsatellites for potato (*Solanum tuberosum* L.) [J]. Theor Appl Genet, 2005, 111: 456-466.
- [23] Wang CB, Guo WZ, Cai CP, et al. Characterization, development and exploitation of EST-derived microsatellites in *Gossypium raimondii* Ulbrich [J]. Chin Sci Bull (科学通报), 2006, 51: 557-561.
- [24] Zhang YX, Lin ZX, Li W, et al. Studies of new EST-SSRs derived from *Gossypium barbadense* [J]. Chin Sci Bull (科学通报), 2007, 52: 2522-2531.

- [25] Chen C, Zhou P, Choi YA, et al. Mining and characterizing microsatellites from citrus ESTs [J]. *Theor Appl Genet*, 2006, 112: 1248-1257.
- [26] Luro FL, Costantino G, Terol J, et al. Transferability of the EST-SSRs developed on *Nules clementine* (*Citrus clementina* Hort ex Tan) to other Citrus species and their effectiveness for genetic mapping [J]. *BMC Genomics*, 2008, 9: 287.
- [27] Fernandez-Silva I, Eduardo I, Blanca J, et al. Bin mapping of genomic and EST-derived SSRs in melon (*Cucumis melo* L.) [J]. *Theor Appl Genet*, 2008, 118: 139-150.
- [28] Simko I. Development of EST-SSR markers for the study of population structure in lettuce (*Lactuca sativa* L.) [J]. *J Hered*, 2009, 100: 256-262.
- [29] Choudhary S, Sethy NK, Shokeen B, et al. Development of chickpea EST-SSR markers and analysis of allelic variation across related species [J]. *Theor Appl Genet*, 2009, 118: 591-608.
- [30] Murray MG, Thompson WF. Rapid isolation of high molecular weight plant DNA [J]. *Nucleic Acids Res*, 1980, 8: 4321-4325.
- [31] Zhang Y, Liu ZH, Liu C, et al. Analysis of DNA methylation variation in wheat genetic background after alien chromatin introduction based on methylation-sensitive amplification polymorphism [J]. *Sci Chin C: Life Sci (中国科学 C: 生命科学)*, 2008, 53: 58-69.
- [32] Bassam BJ, Caetano-Anolles G, Gresshoff PM. Fast and sensitive silver staining of DNA in polyacrylamide gels [J]. *Anal Biochem*, 1991, 196: 80-83.
- [33] Cardle L, Ramsay L, Milbourne D, et al. Computational and experimental characterization of physically clustered simple sequence repeats in plants [J]. *Genetics*, 2000, 156: 847-854.
- [34] La Rota M, Kantety RV, Yu J, et al. Nonrandom distribution and frequencies of genomic and EST-derived microsatellite markers in rice, wheat, and barley [J]. *BMC Genomics*, 2005, 6: 23.
- [35] Saha MC, Mian MA, Eujayl I, et al. Tall fescue EST-SSR markers with transferability across several grass species [J]. *Theor Appl Genet*, 2004, 109: 783-791.
- [36] Dreisigacker S, Zhang P, Warburton ML, et al. SSR and pedigree analyses of genetic diversity among CIMMYT wheat lines targeted to different mega environments [J]. *Crop Sci*, 2004, 44: 381-388.

---

### 欢迎订阅 2010 年《中国药师》杂志

《中国药师》由国家食品药品监督管理局主管，1998 年 6 月创刊，为科技部中国科技论文统计源期刊、中国科技核心期刊，且为国内许多大型检索数据库和《国际药学文摘》收录。设有“研究论文”“继续教育”“药学进展”“研究报告”“药学与临床”“综述与专论”“药品监管”“科技交流”“医药信息”等深受广大药师、医师欢迎的栏目。杂志由《中国药师》编辑部出版，国内统一刊号 CN42-1626/R，月刊，大 16 开 152 页，每期定价 16.00 元，全年 192.00 元。欢迎到当地邮局订阅，邮发代号 38-325，漏订者可向编辑部补订。地址：湖北省武汉市兰陵路 2 号《中国药师》发行部。邮编：430014，电话：027-82778580，82835077。电子信箱：yh@cnjpe.org，zgys@periodicals.net.cn。网址：<http://www.zgys.org>。

我们向广大作者承诺：在您的帮助配合下，90 天内可获知稿件处理结果；省、部级基金或重要成果的首发论文 60 天内刊登；国家级基金首发论文 30 天内刊登。欢迎踊跃投稿。本刊已全面启用网上《中国药师》用户中心 (<http://tg.zgys.org>)，读者、作者可在网上实现在线投稿、查询，会议注册，参加继续教育项目等，作者与编者的互动交流更方便、快捷。