

# 海岛棉 EST-SSR 引物的开发与应用研究

张艳欣\* 林忠旭\* 李武 涂礼莉 聂以春 张献龙†

(华中农业大学作物遗传改良国家重点实验室, 武汉 430070. \* 同等贡献. † 联系人, E-mail: [xlzhang@mail.hzau.edu.cn](mailto:xlzhang@mail.hzau.edu.cn))

**摘要** 现有的棉花 EST-SSR 引物大都开发于中棉和陆地棉, 尚没有海岛棉 EST-SSR 开发研究的报道. 本研究从实验室构建的海岛棉纤维发育的 cDNA 文库中先取 98 条 EST 序列设计了 119 对 EST-SSR 引物. 在这些 SSR 中, 三核苷酸重复基序 AAG 含量最为丰富, 达 11.76%. 用 36 份材料(13 份 A 基因组二倍体材料, 11 份 D 基因组二倍体材料和 12 份 AD 基因组异源四倍体材料)评价新 EST-SSR 引物, 119 对引物中有 76 对成功扩增, 共产生 313 条多态性条带, 平均每对引物产生 4.11 条. 这些引物的多态性信息含量(PIC)变化在 0.17~0.95 之间, 平均 0.53. 根据 Jaccard's 遗传相似系数对所用 36 份材料进行了聚类分析, 聚类结果为 3 大类, 分别为 A 基因组材料、D 基因组材料和 AD 基因组材料. 此外, 有 21 对引物在本实验室的种间回交群体((鄂棉 22 × Pima3-79) × 鄂棉 22) DNA 中表现多态性扩增, 产生 24 个多态性位点, 其中 22 个被锚定于该遗传连锁图的 12 条染色体. 本研究不仅详细地概括了这批新的 EST-SSR 引物的特征, 而且有力地证明了其在遗传多样性分析及遗传连锁图构建方面的应用价值.

**关键词** 海岛棉 表达序列标签(EST) 简单序列重复(SSR) 遗传多样性 遗传连锁图

棉花(*Gossypium* L)是世界上最重要的经济作物之一, 为人类提供了天然的纺织纤维和食用油. 棉纤维由于是植物中最长的单细胞结构而被作为研究细胞伸长和细胞壁生物合成的模型 [1,2]. 过去几年中, 纺织技术领域的深刻革新使得棉花纤维品质的改良变得越来越重要 [3]. 而随着基因组研究的飞速发展, 分子育种成为棉花纤维品质遗传改良的重要途径. 新的分子标记应不断被开发以利于棉花高度整合的遗传图谱的构建, 从而用于棉花纤维基因结构和功能的研究及棉纤维品质相关基因的分子标记辅助选择(marker-assisted selection, MAS).

DNA 标记是估计群体内或群体间遗传变异的有力可靠的工具 [4,5]. 限制性片段长度多态性(restriction fragment length polymorphism, RFLP)标记曾被用作最可靠的分子标记, 然而近年来, RFLP 和其他旧标记技术正渐渐被简单而有效的基于 PCR 的微卫星标记或称为简单序列重复(simple sequence repeat, SSR)所取代. 微卫星标记由 1~6 碱基的核苷酸排列组成, 是基因或基因组的普通组分 [6]. 由于它们丰富且遍布基因组, 具有超变异性 [5] 及位点专一性, 通常表现共显性和多等位基因性 [7], 在大多数真核生物基因组中以二、三或四核苷酸基序排列而易于识别 [8], 在多种作物中已被广泛用于遗传学研究 [9].

截至 2006 年 12 月 10 日, 棉花微卫星数据库(<http://www.mainlab.clemson.edu/cmd/Primer.shtml>)中已公布 5676 个微卫星标记, 包括 379 个 BNL, 392 个 CIR, 53 个 CM, 309 个 JESPR, 84 个 MGHES, 1316 个 MUSB, 1169 个 MUSS/MUCS, 1032 个 NAU, 192 个 STV 和 750 个 TMB 标记.

表达序列标签(expressed sequence tag, EST)工程提供了大量的植物物种的公开序列数据, 这些数据对于采集 SSR 很有用. EST-SSRs 可被用于植物遗传和进化的研究, 相对于基因组 SSR 标记它们也具有更高的到相关物种的转移性. 此外, EST-SSRs 的功能通常是可推测的.

许多植物中的 EST-SSR 标记已被开发出来, EST 工程也对具有重要经济价值的植物给予特别关注. 到 2006 年 12 月 10 日为止, 在 dbEST 数据库(<http://www.ncbi.nlm.nih.gov/dbEST>)中已可获得 280235 条棉属的 EST 序列, 包括陆地棉 177047 条、雷蒙德氏棉 63577 条、中棉 39230 条、草棉 247 条、海岛棉 134 条. 陆地棉的最多, 而海岛棉的最少. 大量的棉花 EST-SSR 标记已被开发并应用于科学研究 [10~16]. 这些源于 EST 的 SSR 标记对于性状表达中候选基因的位置及效应估计很有用. 然而, 现有的棉花 EST-SSR 引物大都开发于中棉和陆地棉, 几乎没有源于海岛棉

的。众所周知，海岛棉是最重要的棉种之一，可以从中海岛棉中发掘新的遗传材料用于陆地棉的纤维品质改良。

本实验室用 Pima3-79 (遗传标准系)构建了海岛棉的 cDNA 文库，基于此文库开发了一套新的海岛棉 EST-SSR 引物，本研究不仅详细地概括了这批新的 EST-SSR 引物的特征，而且评价了其在遗传多样性分析及遗传连锁图构建方面的应用价值。

### 1 材料和方法

( ) SSR 的鉴别和 EST-SSR 引物开发。从本实验室用 Pima3-79 构建的海岛棉 cDNA 文库中共采集 889 条 EST 序列，序列拼接去重复后获得 210 条单

一序列用于开发 EST-SSR 引物。用“简单序列重复识别工具”(SSRIT)(<http://www.gramene.org/db/searches/ssrtool>)识别 SSR，识别标准为二、三、四、五、六核苷酸重复基序最小长度分别达到 10, 12, 12, 10, 12 bp。用“Primer 3”软件(<http://redb.ncpgr.cn/modules/redbtools/primer3.php>)根据 SSR 的侧翼区域设计引物，命名为“HAU xxx”(HAU 是华中农业大学的英文缩写)。

( ) 材料及 DNA 抽提。用 A 基因组的 13 份二倍体材料、D 基因组的 11 份二倍体材料和 AD 基因组的 12 份异源四倍体材料评价新 EST-SSR 引物。本研究所用来自 17 个棉种的 36 份材料的情况见表 1。随后，来自种间 BC<sub>1</sub> 作图群体((鄂棉 22 × Pima3-79) ×

表 1 本研究所用 36 份材料的情况概括

编号	品种	缩写	倍性	物种	栽培/野生
1	罗甸官田小棉花	LDGTXMH	2x	<i>G. arboreum</i>	栽培
2	铁木赤	TMC	2x	<i>G. arboreum</i>	栽培
3	下寨小花 3	XZXH3	2x	<i>G. arboreum</i>	栽培
4	长绒中棉	CRZM	2x	<i>G. arboreum</i>	栽培
5	华中紫秆中棉	HZZGZM	2x	<i>G. arboreum</i>	栽培
6	东台小白花	DTXBH	2x	<i>G. arboreum</i>	栽培
7	束鹿白花白	SLBHB	2x	<i>G. arboreum</i>	栽培
8	宝山紫花	BSZH	2x	<i>G. arboreum</i>	栽培
9	长顺代化小花	CSDHXX	2x	<i>G. arboreum</i>	栽培
10	戊寨小花	WZXH	2x	<i>G. arboreum</i>	栽培
11	御系 3-3	YX3-3	2x	<i>G. arboreum</i>	栽培
12	完县小白花	WXXBH	2x	<i>G. arboreum</i>	栽培
13	阿非利加棉	G.HERB	2x	<i>G. herbaceum</i>	栽培
14	瑟伯氏棉	G.THUR	2x	<i>G. thurberi</i>	野生
15	三裂棉	G.TRIL	2x	<i>G. trilobum</i>	野生
16	拟似棉	G.GOSS	2x	<i>G. gossypoides</i>	野生
17	哈克尼西棉	G.HARK	2x	<i>G. harknessii</i>	野生
18	戴维逊氏棉	G.DAVI	2x	<i>G. davidsonii</i>	野生
19	克劳斯基棉	G.KLOT	2x	<i>G. klotzschianum</i>	野生
20	旱地棉	G.ARID	2x	<i>G. aridum</i>	野生
21	松散棉	G.LAXU	2x	<i>G. laxum</i>	野生
22	辣根棉	G.ARMO	2x	<i>G. armourianum</i>	野生
23	雷蒙德氏棉	G.RAIM	2x	<i>G. raimondii</i>	野生
24	裂片棉	G.LOBA	2x	<i>G. lobatum</i>	野生
25	黄褐棉	G.MUST	4x	<i>G. mustelinum</i>	野生
26	达尔文棉	G.DARW	4x	<i>G. darwinii</i>	野生
27	TM-1	TM-1	4x	<i>G. hirsutum</i>	栽培
28	Acala3080	Acala3080	4x	<i>G. hirsutum</i>	栽培
29	晋棉 6 号	Jimian6	4x	<i>G. hirsutum</i>	栽培
30	冀棉 5 号	Jimian5	4x	<i>G. hirsutum</i>	栽培
31	DH962	DH962	4x	<i>G. hirsutum</i>	栽培
32	邯郸 208	Handan208	4x	<i>G. hirsutum</i>	栽培
33	鄂棉 22	Emian22	4x	<i>G. hirsutum</i>	栽培
34	Pima3-79	Pima3-79	4x	<i>G. barbadense</i>	栽培
35	Pima90	Pima90	4x	<i>G. barbadense</i>	栽培
36	海 7124	Hai7124	4x	<i>G. barbadense</i>	栽培

鄂棉 22) 的 141 单株也被用来定位新的引物, 其亲本鄂棉 22 和 Pima3-79 包含于上述 36 份材料中. 每份材料基因组总 DNA 的提取参照 Paterson 等人 [17] 的方法.

( ) PCR 扩增、电泳和 DNA 条带记录. PCR 体系各组分的比例及 PCR 反应程序与本实验室以前所用方法 [18] 相同. 扩增产物通过 6% 变性聚丙烯酰胺凝胶电泳按分子量大小分离, 电泳参数和银染均依 Lin 等人 [18] 的方法. 每对引物所扩增 DNA 主带的大小参考 DNA Ladder 标准记载. 对于前 36 份材料, 条带的有或无记为 1 或 0. 对于 BC<sub>1</sub> 群体 141 单株的扩增片段, 与母本的基因型相同记为 A, 而与杂合的 F<sub>1</sub> 的基因型相同记为 H.

( ) PIC 值计算. 每对引物的多态性信息含量 (PIC) [15,19,20] 可用于估计其等位基因的变异, 它是根据所有材料中等位基因出现的频率计算的:

$$PIC = 1 - \sum_{i=1}^k P_i^2,$$

其中,  $P_i$  表示第  $i$  个等位基因出现的频率,  $k$  是此标记的所有等位基因的总数.

( ) 遗传相似系数分析和主坐标分析. 利用 NTSYS-pc 软件 (2.10 版) [21] 的 SIMQUAL 程序计算 36 份材料间的遗传相似系数. 然后, 基于遗传相似系数矩阵, 依据 UPGMA 算法并利用 NTSYS-pc 软件的 SHAN 聚类程序 [23] 生成聚类图 [22]. 同样, 基于遗传相似系数矩阵, 利用 NTSYS-pc 软件的 DCENTER 和 EIGEN 等主坐标分析程序进行主坐标 (PCO) 分析, 并生成前 3 个主坐标的三维图.

( ) 遗传连锁作图. 经筛选, 在鄂棉 22 和 Pima3-79 之间有多态性的 EST-SSR 引物被用来扩增种间 BC<sub>1</sub> 作图群体 ((鄂棉 22 × Pima3-79) × 鄂棉 22) 的 141 单株的 DNA, 并利用软件 MapMaker3.0b [24] 将其整合到 BC<sub>1</sub> 遗传连锁图 (未发表). 计算中用到 “Group” (LOD = 5,  $r = 0.4$ ), “Three point”, “Order”, “Try”, “Map” 和 “Ripple” 等命令, 并用 Kosambi 作图函数 [25] 将重组频率转换为图距 (centi-Morgans). 根据其他已发表图谱上位置确定的 SSR 标记 (<http://www.mainlab.clemson.edu/cmd/Downloads.shtml>) 将 HAU 引物所在的连锁群与相应的染色体对应起来.

## 2 结果

### 2.1 EST-SSR 引物的特征

从 210 条 EST 序列中发现有 137 (65.24%) 条有

SSR 序列, 其中 33 条 EST 含 2 个 SSR, 4 条 EST 含 3 个 SSR, 最后从 98 条 EST 序列中设计得到 119 对引物. 这些引物的重复基序类型, 正向和反向引物序列, 及其 EST 序列在 GenBank 的登录号均已提交到棉花微卫星数据库 (<http://www.mainlab.clemson.edu/cmd/Primer.shtml>). 这些引物中出现频率最高的重复基序类型为三核苷酸 (46.22%), 其次是二核苷酸 (22.69%)、四核苷酸 (15.97%)、六核苷酸 (7.56%) 及五核苷酸 (7.56%) (图 1). 此结果与 EST 中三核苷酸比二核苷酸更为丰富的结论相符 [26-29].

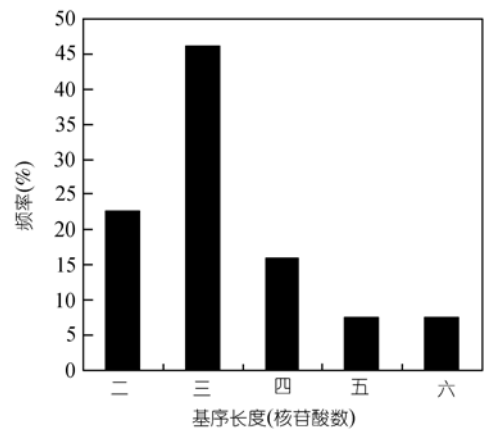


图 1 HAU 引物各种长度的基序出现的频率

在二核苷酸中, AT 出现频率最高, 达 9.24%, 其次是 AG (8.40%) 和 AC (5.04%). 三核苷酸中出现频率最高的是 AAG (11.76%), 而 ATC (8.40%), ACC (6.72%), GAC (4.20%) 及 GAT (4.20%) 次之. 四核苷酸中的 TAAT, AAAT 和 GTTT 出现频率相当, 都是 2.52% (图 2).

对于重复次数而言, 三核苷酸重复 4 次的达 30.25%, 四核苷酸重复 3 次的次之, 达 13.44%, 二核苷酸重复 5 次的达 10.92%, 三核苷酸重复 5 次的达 9.24% (图 3).

此外, EST-SSR 位点的总长度变化在 10~30 bp 之间, 其中长 12 bp 的最多, 占 44.54%; 长 15 和 18 bp 的分别占 15.97% 和 10.92%.

### 2.2 EST-SSR 引物的功能推测

用 TBLASTX 将设计引物的 98 条 EST 序列与 GenBank 数据库中其他序列进行同源性比对,  $E$  值设为大于或等于  $10^{-5}$ . 其中 83 条 (84.69%) 有搜索结果. 例如, 与拟南芥同源的有 36 条 (36.73%), 与烟草同源的有 3 条 (3.06%), 与棉花同源的有 18 条 (18.37%).

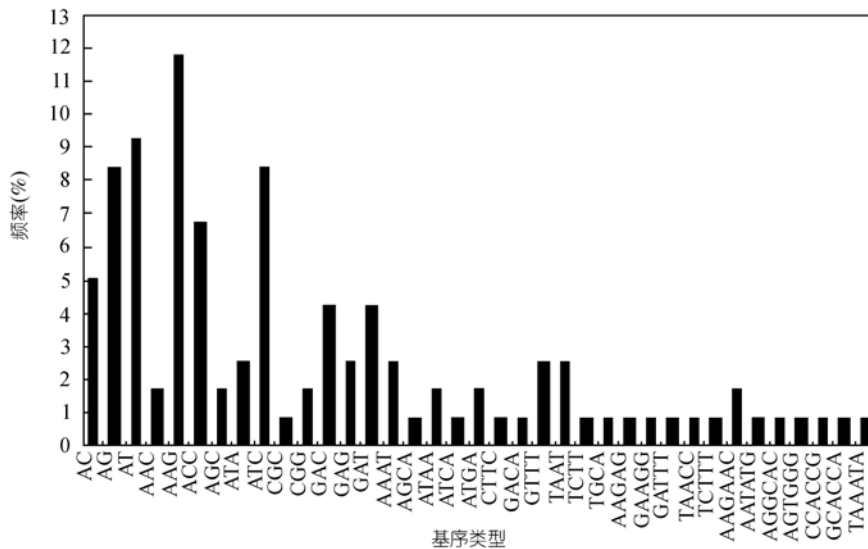


图2 HAU引物各类型基序出现的频率

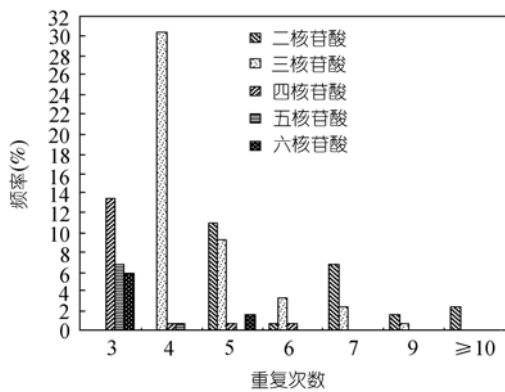


图3 HAU引物各种长度的基序出现不同重复次数的频率

### 2.3 EST-SSR 引物的扩增效率

用 119 对 EST-SSR 引物扩增 36 份材料, 76 对 (63.87%)引物有扩增产物, 共得到 313 个多态性片段, 平均每对引物产生 4.11 个多态性片段. 每对引物产生的多态性片段的数目 1~13 不等, 引物 HAU059 得到多态性片段最多, 达 13 个. 有 6 对 (0.05%)引物扩增得到非特异性产物, 37 对 (31.09%)引物在任何一份

材料中扩增都无产物. 图 4 所示为引物 HAU077 的扩增情况. 上述 76 对引物中, 在鄂棉 22 和 Pima3-79 之间表现多态性的有 33 对 (43.42%), 都被用来扩增 BC<sub>1</sub> 作图群体的 141 个单株的 DNA, 最后有 21 对引物产生了 24 个多态性位点.

### 2.4 EST-SSR 引物的 PIC 值

依据 76 对引物在 36 份材料中扩增得到的等位基因的变异的信息来计算各引物的 PIC 值. 它们的 PIC 值变化在 0.17~0.95 之间, 平均为 0.53. 引物 HAU072 的 PIC 值最大, 达 0.95, 而引物 HAU100 的最小, 只有 0.17.

### 2.5 遗传相似系数分析和主坐标分析

依据 UPGMA 算法利用 NTSYS-pc 软件将本研究所用 36 份材料进行聚类分析, 结果分为 3 大类 (图 5): 第 1 大类包括来自 A 基因组的 13 个二倍体材料 (12 份中棉和 1 份草棉); 第 2 大类包括 12 份异源四倍体材料 (1 份黄褐棉、1 份达尔文棉、7 份陆地棉及 3 份海岛棉); 第 3 大类包括来自 D 基因组的 11 份二倍体

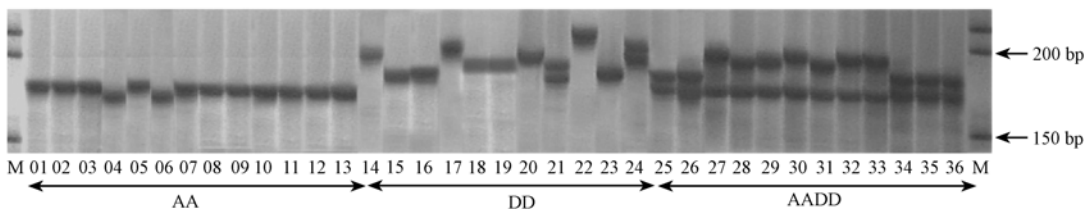


图4 引物 HAU077 对 36 份材料 PCR 扩增所得产物的 6% 变性聚丙烯酰胺凝胶电泳分离结果  
第一和最后一个泳道是 DNA Ladder, 中间 36 份材料的编号见表 1

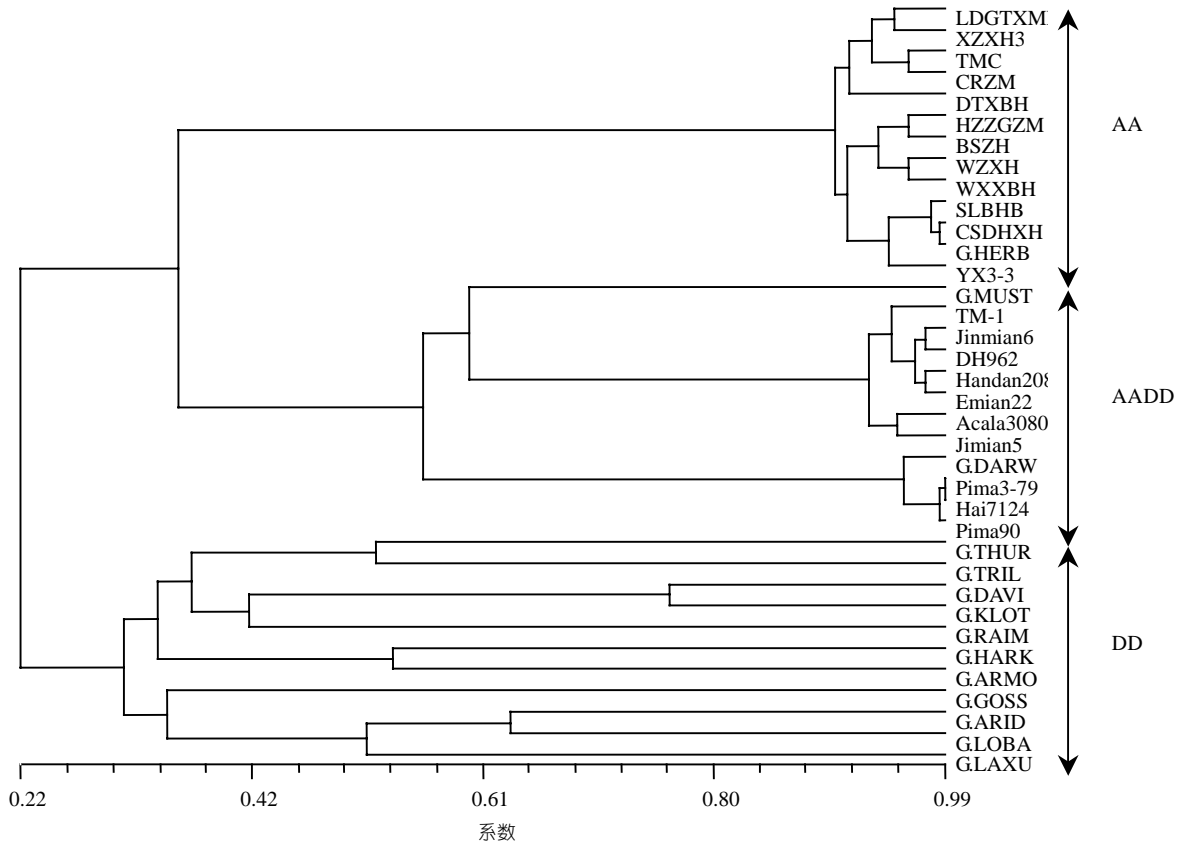


图 5 用 HAU 引物扩增 36 份材料的数据计算所得的 Jaccard 相似系数的 UPGMA 聚类图

材料, 均为野生棉. 聚类结果与 Fryxell<sup>[30]</sup> 的分类法一致. 基于遗传相似系数矩阵进行主坐标分析得到的三维图可以更直观地体现 36 份材料间的聚类关系(图 6).

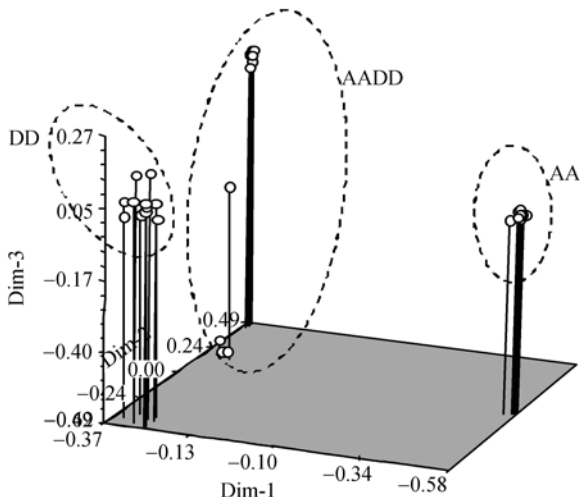


图 6 用 HAU 引物扩增 36 份材料的数据进行主坐标分析得到的第 1, 2 和 3 主坐标的三维图

36 份材料间的 Jaccard 两两相似系数在 0.149~0.991 之间. 中棉和草棉间高达 0.921 的相似系数揭示了二者极高的遗传相似度. 来自 D 基因组的 11 份二倍体野生棉间的相似系数变化在 0.245~0.763 之间, 这一广泛的变异反映了它们遗传关系较远. 通过比较平均每对引物在三大类材料中各自扩增的等位基因数, 同样可以得到上述结论. 扩增 D 基因组材料时, 平均每对引物得到 3.76 个等位基因, 但是在 A 基因组材料中仅得到 1.46 个, 在 AD 基因组材料中得到 2.23 个. 所以, D 基因组材料间存在高度的遗传多样性.

### 2.6 遗传作图

本研究所用 EST-SSR 引物扩增 BC<sub>1</sub> 作图群体得到的 24 个多态性位点中有 22 个已被整合到海陆种间遗传连锁图谱(结果未发表)上, 分别位于 12 条染色体(图 7). 这 22 个位点在整个基因组的分布较均衡. 有 11 个位点位于 A 基因组的 7 条染色体(Chr1, 4, 5, 6, 7, 9 和 12), 另 11 个位点位于 D 基因组的 5 条染色体(Chr15, 18, 19, 22 和 26).

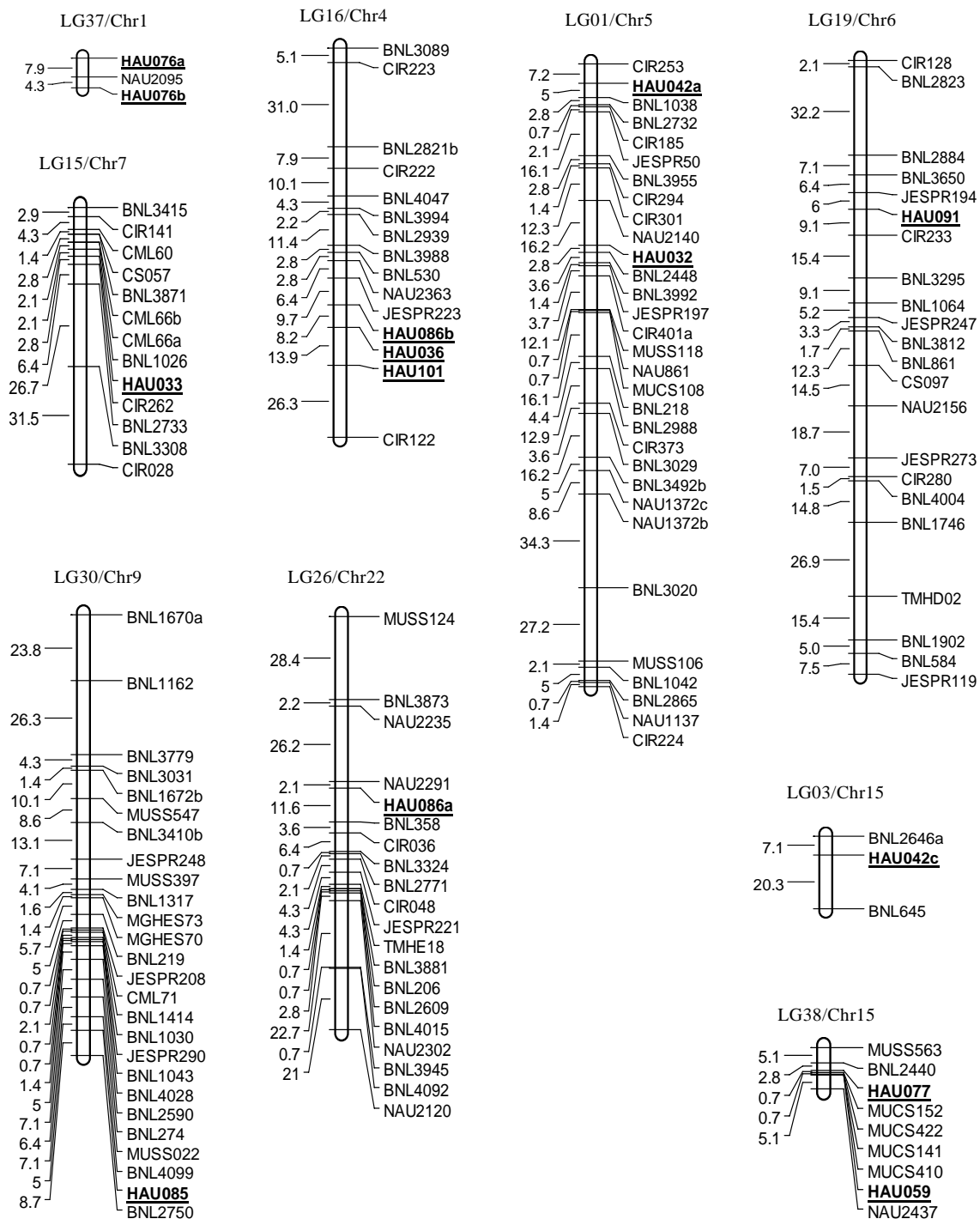


图7 22个HAU多态性位点在种间BC<sub>1</sub>遗传连锁图的13个连锁群的分布情况  
HAU引物均用下划线并加黑突出显示

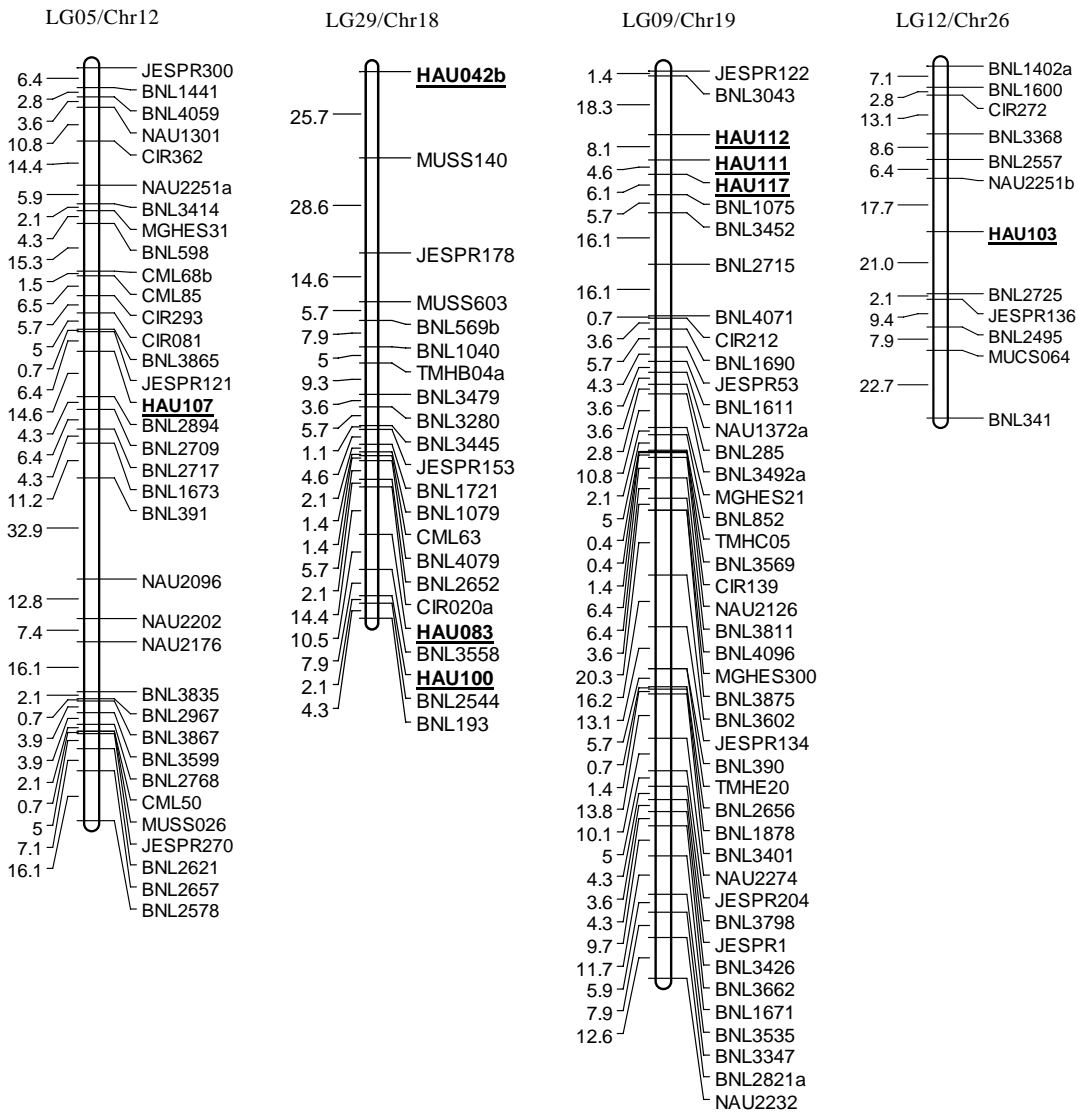


图 7(续) 22 个 HAU 多态性位点在种间 BC<sub>1</sub> 遗传连锁图的 13 个连锁群的分布情况

### 3 讨论

由于传统基因组 SSR 标记的开发费时费力且效率低,限制了其在更多作物上的应用.随着功能基因组学的发展,EST 被大规模测序并且序列信息被公开在公共数据库中,这为我们开发 SSR 标记提供了丰富而宝贵的资源.

截至 2006 年 12 月 10 日,在 dbEST 数据库(<http://www.ncbi.nlm.nih.gov/dbEST>)中棉花的 EST 序列已达到 280235 条,包括陆地棉、中棉、草棉、海岛棉和雷蒙德氏棉 5 个棉种.与 2004 年 4 月 30 日的 54141 条序列 [13] 相比,增加了 226094 条,其中增长最快的

是陆地棉,其次是雷蒙德氏棉和中棉,而草棉和海岛棉无增长.由于陆地棉和中棉的 EST 序列最为丰富,已经从中开发出多套 EST-SSR 引物,陆地棉的有 573 对,中棉的有 1865 对;由于海岛棉的 EST 序列太少,此前还没有任何 SSR 引物从中开发.本研究的 119 对 EST-SSR 引物均开发自海岛棉,丰富了棉属 EST-SSR 的来源.

前人对其他植物的研究表明,三核苷酸是存在最广泛的重复基序类型 [26-29],在本研究中也得到了同样的结论.此外,本研究结果也与王长彪等人 [31] 开发自雷蒙德氏棉的 EST-SSR 引物的研究结果一致,即二核苷酸和三核苷酸中出现最多的重复基序分别

为AT和AAG; 但在水稻、小麦和大麦等其他植物中<sup>[32,33]</sup>却与棉花不同, 分别为AG和CCG最多, 或许这是单子叶植物区别于双子叶植物的特征之一<sup>[34]</sup>.

在119对引物中有76对成功扩增了36份材料的DNA, 扩增效率仅63.87%. 可能的原因是正向引物或反向引物或二者都恰好跨过了mRNA的剪切位点, 也可能是由于在相应的基因组DNA中存在较大的内含子, 这两个原因都会导致引物无法与模板DNA结合而不能扩增. 此外, 76对可扩增的引物中只有33对(43.42%)在双亲鄂棉22与Pima3-79之间有多态性, 其中21对在作图群体中有多态性, 这一较低的多态性率与识别SSR时对重复基序的参数设置不够严格有关, 导致所开发EST-SSR引物之间变异较小. 此外, 亲本间有多态性的但表现为父本显性的引物在BC<sub>1</sub>群体中不能扩增出差异的带型, 因为这些引物在BC<sub>1</sub>后代中杂合子与纯合子表现相同的带型.

PIC值可用来估计每对扩增引物的等位基因的变异, 而PIC值的大小是基于所检测到的等位基因的数目及分布频率<sup>[19]</sup>, 所以在一定意义上, PIC也是引物检测多态性的一种度量. 引物的PIC值越高, 它揭示等位基因变异的能力就越强, 便于分子标记作图时双亲的选择及对一个群体内遗传变异的估计<sup>[20]</sup>. 前人得到的不同作物EST-SSR引物的PIC值各不相同, 取决于他们所用引物的数目和被扩增的基因型的数目. 不论是从等位基因的数目还是PIC值的大小看, 基于基因的EST-SSR标记的多态性确实低于基因组SSR<sup>[35]</sup>. Park等人<sup>[15]</sup>开发的1232对EST-SSR(MUSS)及EST-CSR(MUCS)引物的PIC值在0.12~0.73之间, 平均0.42, 稍稍低于HAU引物的PIC值(0.17~0.95, 平均0.53). 可能的原因是他们的引物数目较多且研究材料是183个种间重组自交系, 而本研究所用引物较少, 测试的棉种较多, 变异丰富.

本研究所用的材料中有两个野生的异源四倍体棉种, 其中黄褐棉与陆地棉聚为一类, 而达尔文棉与海岛棉聚为一类, 遗传相似系数分别为0.595和0.958, 揭示了黄褐棉与陆地棉有较近的亲缘关系, 达尔文棉与海岛棉也有极高的遗传相似性. 以上聚类结果充分证实了EST-SSR在遗传多样性分析方面的应用能力. 然而, 由于所用材料中个别棉种的测试材料的数目有限, 所检测到的它们的种内及种间的遗传变异还不够全面, 有待于增加测试材料和引物的数目取得更有说服力的结论.

Han等人<sup>[13]</sup>于2004年将111个EST-SSR位点(NAU引物)整合到其种间BC<sub>1</sub>作图群体构建的遗传图谱上, 随后于2006年又增加了123个EST-SSR位点(NAU引物)<sup>[14]</sup>. Park等人<sup>[15]</sup>用183个种间重组自交系为作图群体构建的遗传图谱上有121个EST-SSR位点(73个MUSS, 48个MUCS)和72个基因组SSR位点. 他们所用的EST-SSR引物均开发自中棉和陆地棉的纤维EST序列. 本研究所用EST-SSR引物扩增BC<sub>1</sub>作图群体得到的24个多态性位点中有22个已被整合到海陆种间遗传连锁图谱, 位于包括第12和第26号染色体在内的12条染色体上, 而第12和第26号染色体正是以往报道定位纤维发育相关基因最多的. 此张遗传连锁图谱上也包括MGHES, NAU和MUSS/MUCS等其他系列的EST-SSR引物, 均来自棉花微卫星数据库(<http://www.mainlab.clemson.edu/cmd/Primer.shtml>). 本研究所整合到图谱上的新的EST-SSR位点不仅增加了图谱的长度, 使之覆盖了棉花基因组的更大比例, 更丰富了功能标记的数目. 以上结果表明, 在遗传图谱构建方面EST-SSR引物不失为新的更宝贵的资源.

在GenBank数据库中进行同源性搜索发现棉花比其他大多数作物与拟南芥具有更近的亲缘关系, 这更有利于开展基因组学研究和比较作图<sup>[36]</sup>. 而EST-SSR标记在比较作图方面更具应用潜力<sup>[37-39]</sup>. 此外, EST-SSR标记对于分子标记辅助选择育种也很有帮助, 一旦发现控制目标性状(如纤维品质或抗性)的基因中有EST-SSR存在, 便可以直接而有效地对其等位基因进行选择.

总之, EST-SSR标记易于开发、成本较低, 其功能可以通过序列同源性比对获得. 而且由于来自于转录谱, 能够反映出转录区的差异, 使得EST-SSR标记在遗传多样性分析、比较作图和分子标记辅助选择育种的研究中具有更高的应用价值.

## 参 考 文 献

- 1 Hof J V, Saha S. Cotton fibers can undergo cell division. *Am J Bot*, 1997, 84: 1231—1235 [\[DOI\]](#)
- 2 Kim H J, Triplett B A. Cotton fiber growth in planta and *in vitro*. Models for plant cell elongation and cell wall biogenesis. *Plant Physiol*, 2001, 127: 1361—1366 [\[DOI\]](#)
- 3 Shen X L, Guo W Z, Zhu X F. Molecular mapping of QTLs for fiber qualities in three diverse lines in Upland cotton using SSR markers. *Mol Breed*, 2005, 15: 169—181 [\[DOI\]](#)
- 4 Qamaruz Z F, Michael F F, Parker J S, et al. Molecular techniques employed in the assessment of genetic diversity: A review focusing



- on orchid conservation. *Lindleyana*, 1998, 13: 259—283
- 5 Powell W, Machray G C, Provan J. Polymorphism revealed by simple sequence repeats. *Trends Plant Sci*, 1996, 1: 215—222
- 6 Tautz D, Schotterer C. Simple sequences. *Curr Opin Genet*, 1994, 4: 832—837[DOI]
- 7 Gupta P K, Balyan H S, Sharma P C, et al. Microsatellites in plants: a new class of molecular markers. *Curr Sci*, 1996, 70: 45—54
- 8 Weber J L, May P E. Abundance class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *Am J Hum Genet*, 1989, 44: 388—396
- 9 Combes M C, Andrzejewski S, Anthony F, et al. Characterization of microsatellite loci in *Coffea arabica* and related coffee species. *Mol Ecol*, 2000, 8: 1171—1193
- 10 Saha S, Karaca M, Jenkins J N, et al. Simple sequence repeats as useful resources to study transcribed genes of cotton. *Euphytica*, 2003, 130: 355—364[DOI]
- 11 Qureshi S N, Saha S, Kantety R V, et al. EST-SSR: A new class of genetic markers in cotton. *J Cotton Sci*, 2004, 8: 112—123
- 12 Chee P W, Rong J, Williams C D, et al. EST derived PCR-based markers for functional gene homologues in cotton. *Genome*, 2004, 47: 449—462[DOI]
- 13 Han Z G, Guo W Z, Song X, et al. Genetic mapping of EST derived microsatellites from the diploid *Gossypium arboreum* in allotetraploid cotton. *Mol Genet Genom*, 2004, 272: 308—327[DOI]
- 14 Han Z G, Wang C B, Song X L, et al. Characteristics, development and mapping of *Gossypium hirsutum* derived EST-SSRs in allotetraploid cotton. *Theor Appl Genet*, 2006, 112: 430—439[DOI]
- 15 Park Y H, Alabady M S, Ulloa M, et al. Genetic mapping of new cotton fiber loci using EST-derived microsatellites in an interspecific recombinant inbred (RIL) cotton population. *Mol Genet Genom*, 2005, 274: 428—441[DOI]
- 16 Taliercio E, Allen R D, Essenberg M, et al. Analysis of ESTs from multiple *Gossypium hirsutum* tissues and identification of SSRs. *Genome*, 2006, 49: 306—319[DOI]
- 17 Paterson A H, Brubaker C, Wendel J F. A rapid method for extraction of cotton (*Gossypium* spp.) genomic DNA suitable for RFLP or PCR analysis. *Plant Mol Biol Rep*, 1993, 11: 122—127
- 18 Lin Z X, He D H, Zhang X L, et al. Linkage map construction and mapping QTLs for cotton fiber quality using SRAP, SSR and RAPD. *Plant Breed*, 2005, 124: 180—187[DOI]
- 19 Botstein D, White R L, Skolnick M, et al. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am J Human Genet*, 1980, 32: 314—31
- 20 Anderson J A, Churchill G A, Autrique J E, et al. Optimizing parental selection for genetic linkage maps. *Genome*, 1993, 36: 181—186
- 21 Rohlf F J. NTSYS-pc: Numerical Taxonomy and Multivariate Analysis System, Version 2.1, User Guide. Exeter Software, New York, 2000
- 22 Sokal R R, Michener C D. A statistical method for evaluating systematic relationships. *Univ Kansas Sci Bull*, 1958, 28: 1409—1438
- 23 Sneath P H, Sokal R R. Numerical Taxonomy: The Principal and Practice of Numerical Classification. San Francisco: W. H. Freeman and Company, 1973
- 24 Lander E S, Green P, Abrahamson J, et al. MAPMAKER: An interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. *Genomics*, 1987, 1: 174—181[DOI]
- 25 Kosambi D D. The estimation of map distances from recombination values. *Ann Eugen*, 1994, 12: 172—175
- 26 Chin E C L. Maize simple repetitive DNA sequences: Abundance and allele variation. *Genome*, 1996, 156: 847—854
- 27 Temnykh S, Park W D, Ayres N, et al. Mapping and genome organization of microsatellite sequences in rice (*Oryza sativa* L). *Theor Appl Genet*, 2000, 100: 697—712[DOI]
- 28 Morgante M, Hanafey M, Powell W. Microsatellites are preferentially associated with nonrepetitive DNA in plant genomes. *Nat Genet*, 2002, 30: 194—200[DOI]
- 29 Thiel T, Michalek W, Varshney R K, et al. Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L). *Theor Appl Genet*, 2003, 106: 411—422
- 30 Fryxell P A. A revised taxonomic interpretation of *Gossypium* L. (Makvaceae). *Rheede*, 1992, 2: 108—165
- 31 王长彪, 郭旺珍, 蔡彩平, 等. 雷蒙德氏棉EST-SSRs分布特征及开发与利用. *科学通报*, 2006, 51: 316—320
- 32 Gao L F, Jing R L, Huo N X, et al. One hundred and one new microsatellite loci derived from ESTs (EST-SSRs) in bread wheat. *Theor Appl Genet*, 2004, 108: 1392—1400[DOI]
- 33 Varshney R K, Graner A, Sorrells M E. Genic microsatellite markers in plants: Features and applications. *Trends Biotechnol*, 2005, 23: 48—55[DOI]
- 34 Frelichowski J E, Palmer M B, Main D, et al. Cotton genome mapping with new microsatellites from Acala 'Maxxa' BAC-ends. *Mol Gen Genom*, 2006, 275: 479—491[DOI]
- 35 Blair M W, Giraldo M C, Buendía H F, et al. Microsatellite marker diversity in common bean (*Phaseolus vulgaris* L). *Theor Appl Genet*, 2006, 113: 100—109[DOI]
- 36 Rong J, Bowers J E, Schulze S R, et al. Comparative genomics of *Gossypium* and *Arabidopsis*: Unraveling the consequences of both ancient and recent polyploidy. *Genome Res*, 2005, 15: 1198—1210[DOI]
- 37 Scott K D, Eggler P, Seaton G, et al. Analysis of SSRs derived from grapes EST. *Theor Appl Genet*, 2000, 100: 723—726[DOI]
- 38 Eujayl I, Sorrells M E, Wolters P, et al. Isolation of EST-derived microsatellite markers for genotyping the A and B genomes of wheat. *Theor Appl Genet*, 2002, 104: 399—407[DOI]
- 39 Kantety R V, Rota M L, Matthews D E, et al. Data mining for simple sequence repeats in expressed sequence tags from barely, maize, rice, sorghum and wheat. *Plant Mol Biol*, 2002, 48: 501—510[DOI]