

甘薯近缘野生种三浅裂野牵牛 EST 资源的 SSR 信息分析

贾赵东 马佩勇 郭小丁 谢一芝

(江苏省农业科学院 粮食作物研究所 江苏 南京 210014)

摘要: 利用生物信息学方法对 NCBI 公共数据库的 1 458 条三浅裂野牵牛 ESTs 序列进行 EST-SSRs 信息特征分析。结果显示,剔除低质量的和冗余的序列后,得到总长度为 334.368 kb 无冗余序列 868 条。这些无冗余序列中有 35 个微卫星简单重复序列(Simple sequence repeat, SSR),分布于 35 条 EST 中,出现频率为 4.03%,平均分布距离为 9.55 kb。在 2~6 核苷酸的重复类型中,二、三核苷酸 SSR 之和占总 EST-SSR 的 68.57%,其中三核苷酸 SSR 占总 SSR 数量的 45.71%。在所有重复单元中,所占比例最高的是 AT/AT(14.29%),其次是 TCT/AGA 为 11.43%,TAA/TTA 为 8.57% 列第 3 位。在所有的 SSR 中,EST-SSRs 重复次数 5 次的有 13 个,占全部 SSR 的 37.14%,其次是重复次数为 4 次和 7 次,各有 5 个,占全部 SSR 的 14.29%。结果说明,甘薯近缘野生种三浅裂野牵牛的 EST-SSR 出现频率较高、类型丰富、具有一定的研究利用价值。

关键词: 三浅裂野牵牛; 生物信息学; EST; SSR

中图分类号: S531.03 文献标识码: A 文章编号: 1000-7091(2012)增刊-0027-06

Analysis of SSR Information in ESTs Resource of *Ipomoea trifida*

JIA Zhao-dong, MA Pei-yong, GUO Xiao-ding, XIE Yi-zhi

(Institute of Food Crops, Jiangsu Academy of Agricultural Sciences, Nanjing 210014, China)

Abstract: In the study, 1 458 ESTs of *Ipomoea trifida* in the database of NCBI were downloaded and analyzed by bioinformatics tools and method. The results showed that we got 868 non-redundant ESTs with total length about 334.368 kb after the preprocessed, some redundants and low qualities sequences were removed. Totally 35 SSRs (Simple sequence repeat) were screened and distributed in 35 ESTs, accounting for 4.03% of the non-redundant ESTs, the average distribution distance of the EST-SSRs was about 9.55 kb. In the different motifs ranging from 2 to 6 nucleotides, except to mononucleotide, di-nucleotide and tri-nucleotide repeats were dominant repeat types and accounted for 68.57%, tri-nucleotide repeats types accounted for 45.71%. In addition, Among all repeat motifs of SSRs, AT/AT was most high-frequent repeat motif and accounted for 14.29%, followed by TCT/AGA accounted for 11.43%, thirdly TAA/TTA accounted for 8.57%. Among all SSRs, SSRs which repeated overstep 5 times accounted for 37.14%, that of repeated 5 and 7 times accounted for 14.29%. The result in this paper indicated that the EST-SSR of *Ipomoea trifida* had relevance utility value in the study of sweetpotato genetic and breeding.

Key words: *Ipomoea trifida*; Bioinformatics; Expressed sequence tags; Simple sequence repeat

甘薯是重要的粮食、饲料、工业原料,而且还是新型能源用块根作物。栽培种甘薯是六倍体作物,染色体数目多而且小(包括 90 条染色体),存在着自交不亲和以及同一不孕群内品种间杂交不亲和问题。其遗传背景比较复杂,所以甘薯分子遗传研究相对落后于水稻、玉米等其他作物。目前,在甘薯研

究中应用到的分子标记类型和数量都十分有限,仅有少数关于 RAPD、ISSR、AFLP、SRAP、SSR 等标记的报道,主要应用于甘薯遗传图谱构建和基因定位^[1-5]、遗传多样性分析^[6-10]、亲缘关系鉴定及系谱分析^[11-14]等研究中。然而,相对于较大的甘薯基因组,可有效利用的标记数量非常有限,为了从广度

收稿日期: 2012-08-10

基金项目: 现代农业产业技术体系建设专项(CARS-41-C-03); 国家“863”计划(2012AA101204-4-7); 江苏省科技支撑计划(BE2011301-1); 江苏省农业科技自主创新基金(CX(11)1020)

作者简介: 贾赵东(1977-),男,山西晋城人,助理研究员,硕士,主要从事甘薯种质创新与遗传改良研究。

通讯作者: 谢一芝(1962-),男,江苏无锡人,研究员,主要从事甘薯遗传育种研究。

和深度上加快甘薯的基因组研究,需要开发更多更有效的分子标记。

众多分子标记中,SSR(Simple sequence repeat)分子标记从实用性和有效性方面均优越于其他标记类型,已经成为在植物上应用最广最为重要的分子标记。植物基因组的转录区和非转录区都存在 SSR 标记^[15],与基因组 SSR 标记相比,从基因表达序列标签 EST(Expressed sequence tags)中开发的 SSR(EST-SSR)更加经济、快速、简便,而且用 EST-SSR 构建遗传连锁图,就相当于定位功能已知的基因,这将对决定重要表型性状的等位基因进行直接鉴定,是快速揭示基因组容量的有效方法。此外,EST-SSR 来源于基因表达序列,通常其序列保守性程度较高,EST 标记在家族和种属间的通用性比来源于非表达序列的标记更高,也就是说它的通用性更高,这些都是基因组 SSR 所不具有的优点^[16]。近年来,许多作物大规模的 cDNA 单边测序并在网上公开释放,大大增加了基于 ESTs 的 SSR 标记开发能力。自 2000 年起,已经有许多作物开展了 EST-SSR 标记的开发及利用,并已广泛应用于基因组研究和分子育种^[17-23]。

截止 2009 年 8 月 27 日,在 GenBank (<http://www.ncbi.nlm.nih.gov/dbEST>) 中已经释放了甘薯(*Ipomoea batatas*) EST 序列 22 371 条,三浅裂野牵牛(*Ipomoea trifida*) EST 序列 1 458 条,基于上述甘薯 EST 序列信息,黄立飞等^[24]已对部分 EST 的 SSR 信息进行了初步研究。三浅裂野牵牛可能是甘薯的重要起源种之一,有研究表明,三浅裂野牵牛与栽培种甘薯有很紧密的亲缘性^[12-14],所以了解三浅裂野牵牛基因组信息有助于研究甘薯基因组和甘薯组的进化模式,揭示不同种基因之间的关系,从而研究物种的起源与进化。本研究利用 GenBank 中三浅裂野牵牛的 EST 序列信息,对其进行全面的 SSR 信息特征分析,探讨其 EST-SSR 开发的可能性。以明确三浅裂野牵牛 EST-SSR 发生频率和特点,为建立 EST-SSR 标记、构建遗传图谱和发掘有利基因,探索其在甘薯遗传育种中的应用奠定基础。

1 材料和方法

1.1 EST 序列的来源

2009 年 8 月 27 日从 dbEST/GenBank(<http://www.ncbi.nlm.nih.gov/entrez>) 中以 FASTA 格式下载了 1 458 条三浅裂野牵牛 EST 序列。

1.2 EST 序列预处理

序列预处理使用 phred/cross_match 软件包^[25]

完成。先后利用 phred 程序针对所有序列文件进行碱基判读与末端低质量区域去除,phd2fasta 程序进行数据格式转换,cross_match 程序用于屏蔽载体序列,脚本 EST_trimmer.pl(http://pgrc.ipk-gatersleben.de/misa/download/est_trimmer.pl)进行 ployA/T 及 X 屏蔽字符去除,进而获得高质量的 ESTs 序列。本研究中载体屏蔽的参数为: minmatch 13; minscore 20。去除 ployA/T 及 X 字符的参数为: 两端 100 bp 范围内,大于或等于 15 个 A 以及所有的 X 屏蔽字符均去除。

1.3 去冗余-聚类

前处理后的 EST 通过软件 CAP3(<http://seq.cs.iastate.edu/cap3download.html>)进行片段重叠群分析和聚类^[26]。拼接时设定的初始装配参数为默认值。获得的非冗余 ESTs 序列用于 SSRs 标记挖掘与引物开发。

1.4 EST-SSR 筛选检测

利用简单重复序列识别工具 SSRIT(<http://www.gramene.org/db/searches/ssrtool>)和 SSRLocator 软件(<http://www.ufpel.tche.br/faem/fitotecnia/fitomelhoramento/faleconosco.html>)对聚类的非冗余 EST 进行 SSR 搜索,并结合人工查询。筛选标准为:二核苷酸重复次数 ≥ 7 ,三核苷酸重复次数 ≥ 5 ,四核苷酸重复次数 ≥ 4 ,五核苷酸重复次数 ≥ 4 ,六核苷酸重复次数 ≥ 3 ,复合型的 SSR 整体长度不小于 24 bp。

用 SSR 出现频率和 SSR 平均分布距离来描述 EST-SSR,同时定义出现频率前 2 位的重复基元作为优势重复基元。计算公式为: SSR 出现频率 $f_c = c/n \times 100\%$, c 为搜索到的 SSR 数量, n 为无冗余 EST 数量; SSR 平均分布距离 $f_N = N/c$, N 为无冗余 EST 数量的总碱基数。

2 结果与分析

2.1 源于三浅裂野牵牛 EST 的 SSRs 发掘

从 NCBI 公共数据库下载了 1 458 条三浅裂野牵牛 EST 序列,全部 EST 序列经过 EST-trimmer 软件处理后,用 CAP3 软件进行组装拼接,去除低质量的和冗余的序列后,得到全长为 334.368 kb 的无冗余 EST 序列(Uni-EST) 868 条,其中包括 255 个重叠群(Contigs)和 613 个单一序列(Singletons),用于 SSR 搜索和引物设计。

在 868 条无冗余 EST 序列(Uni-EST)中共搜索到 35 个 SSR 位点,分布在 35 条 EST 序列中,平均相隔 9.55 kb 出现 1 个 SSR,这些 SSR 出现的频率

4.03%。用 Weber^[27] 的分类标准可见 35 个 EST-SSRs 均为精密型(Perfect)。三浅裂野牵牛的 EST-SSR 种类十分丰富,二至六核苷酸重复类型都能检测到,但各种类型出现的频率相差很大(表 1)。在不同的核苷酸重复基元类型中,二、三核苷酸 SSR 之和占

总 EST-SSR 的 68.57%,其中三核苷酸 SSR 占总 SSR 数量的 45.71%,其次为二核苷酸 SSR22.86%,四核苷酸 SSR 和六核苷酸 SSR 占总 SSR 数量均为 11.43%,五核苷酸占的比较小,分别为 8.57%。

表 1 SSR-ESTs 序列中 SSR 的数量及分布频率
Tab.1 Number and frequency of SSRs in SSR-ESTs

重复单元类型 Repeat motif type		SSR 数目 Number of SSR	占全部 SSR 比例/% Proportion in all SSRs	出现频率/% Frequency	平均分布距离/kb Average distance
二核苷酸 Di-nucleotide	精密型	8	22.86	0.92	41.80
三核苷酸 Tri-nucleotide	精密型	16	45.71	1.84	20.90
四核苷酸 Tetra-nucleotide	精密型	4	11.43	0.46	83.59
五核苷酸 Penta-nucleotide	精密型	3	8.57	0.35	111.46
六核苷酸 Hexa-nucleotide	精密型	4	11.43	0.46	83.59
总计 Total		35	100	4.03	9.55

2.2 三浅裂野牵牛 EST-SSRs 的分布特征

在 35 个 EST-SSR 中,共观察到 24 种重复基元(Motif)。其中二核苷酸重复基元有 3 种,三、四、五、六核苷酸重复基元分别有 10、3、4 种(表 2)。从出现的频率来看,四、五、六核苷酸重复单元的各个类型均只出现了 1 次,二核苷酸重复基元中的优

势重复基元为 AT/AT 和 TC/GA,出现频率分别为 14.29% 和 5.71%,三核苷酸重复基元中的优势重复基元为 TCT/AGA 和 TAA/TTA,出现频率分别均为 11.43% 和 8.57%(图 1)。在所有重复单元中,所占比例最高的是 AT/AT(14.29%),其次是 TCT/AGA 为 11.43%,TAA/TTA 为 8.57%,列第 3 位。

表 2 EST-SSRs 序列中不同重复基元的重复次数分布
Tab.2 Number of repeats of SSR motif in EST-SSRs

重复基元 Repeat motif	重复次数 Number of repeats									合计 Total
	3	4	5	6	7	8	9	10	> 10	
AT/AT	-	-	-	-	2	1	1	0	1	5
TC/GA	-	-	-	-	0	1	0	0	1	2
AG/CT	-	-	-	-	0	1	0	0	0	1
TCT/AGA	-	-	3	1	0	0	0	0	0	4
GTA/ATC	-	-	1	0	0	0	0	0	0	1
CGC/GCG	-	-	1	0	0	0	0	0	0	1
CCG/CGG	-	-	2	0	0	0	0	0	0	2
TAG/CTA	-	-	1	0	0	0	0	0	0	1
TAA/TTA	-	-	1	1	1	0	0	0	0	3
GCG/CGC	-	-	1	0	0	0	0	0	0	1
TAC/GTA	-	-	1	0	0	0	0	0	0	1
AAT/ATT	-	-	0	0	1	0	0	0	0	1
TTA/TAA	-	-	0	1	0	0	0	0	0	1
TTAT/ATAA	-	0	0	0	1	0	0	0	0	1
GCCT/AGGC	-	0	1	0	0	0	0	0	0	1
AAAC/GTTT	-	1	0	0	0	0	0	0	0	1
TGCA/TGCA	-	1	0	0	0	0	0	0	0	1
TTATT/AATAA	-	1	0	0	0	0	0	0	0	1
CAGAA/TTCTG	-	0	1	0	0	0	0	0	0	1
AAAAC/GTTTT	-	1	0	0	0	0	0	0	0	1
AAGCCT/AGGCTT	1	0	0	0	0	0	0	0	0	1
TTTTAT/ATAAAA	0	1	0	0	0	0	0	0	0	1
TGTTTT/AAAACA	1	0	0	0	0	0	0	0	0	1
TGATAA/TTATCA	1	0	0	0	0	0	0	0	0	1
合计 Total	3	5	13	3	5	3	1	0	2	35

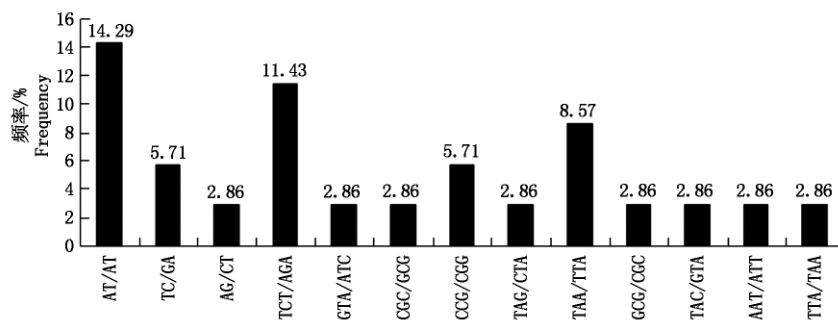


图1 EST-SSRs 序列中二、三核苷酸重复基元的分布频率

Fig. 1 Frequency of SSR repeat motif in Di-nucleotide and Tri-nucleotide

在所有二至六核苷酸重复类型中,重复次数分布见表2,二核苷酸 SSR 的重复次数为 7~13 次,三核苷酸 SSR 的重复次数为 5~7 次,四核苷酸 SSR 的重复次数为 4~7 次,五核苷酸 SSR 的重复次数为 4~5 次,六核苷酸 SSR 的重复次数为 3~4 次。在所有 24 种重复基元(Motif)中(表2),EST-SSRs 重复次数 5 次的有 13 个,占全部 SSR 的 37.14%,其次是重复次数为 4 次、7 次,各有 5 个,占全部 SSR 的 14.29%。

三浅裂野牵牛 EST-SSR 基序长度分布情况见图2。由于搜索标准严格,部分 SSR 被过滤掉(特别是单核苷酸和基序长度较短的 EST-SSR),因此 EST-SSR 基序长度分布是不连续的,如长度为 17、19、22、23、27 bp 的 SSR 没有出现。大部分 EST-SSR 基序长度集中在 14~24 bp 范围内(32 个 SSR,占 91.43%),其次是 25~30 bp (3 个 SSR,占 8.57%),30 bp 以上的 SSR 数量没有筛选到,原因可能用于 EST-SSR 分析的 EST 序列相对偏少的缘故。

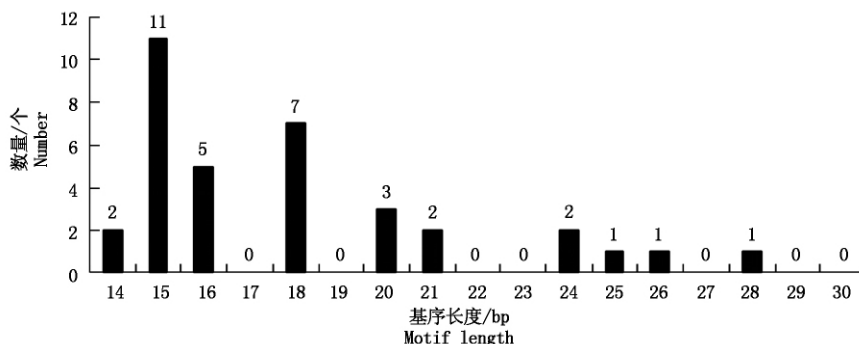


图2 EST-SSR 基序长度分布

Fig. 2 Distribution of motif length of EST-SSRs

3 结论与讨论

本试验对已公布的 1 458 条三浅裂野牵牛 EST 序列进行了 SSR 信息分析,在 868 条无冗余 EST 序列(Uni-EST)中共搜索到 35 个 SSR 位点,占被分析 EST 序列的 4.03%,平均每 9.55 kb 出现 1 个 SSR。对于 EST-SSR 的分布特征,不同研究者利用不同的材料以及不同的 EST-SSR 搜索标准导致了不同的结果。在拟南芥中,按照 Cardle 等^[28]的统计标准,平均 13.8 kb 出现 1 个 SSR;按照 Morgante 等^[29]的统计标准,平均 14.9 kb 出现 1 个 SSR。至于其他作物,按 Cardle 等^[28]的统计标准,水稻中出现频率为 3.4 kb,玉米为 8.1 kb,大豆为 7.4 kb,西红柿为 11.1 kb,棉花为 20.0 kb,杨树为 14.0 kb。若按照 Gao 等^[29]的统计标准,显示出水稻中平均 11.81 kb、玉米中平均 28.32 kb、大豆中平均 23.80 kb、小麦中平

均 17.42 kb 出现 1 个 SSR 的研究结果。在本研究中,与 Morgante 等^[29]的分类标准相比,将不同(单、二、三、四、五、六)重复基元的最小长度定为 15 bp,增加了对六核苷酸重复基元的筛选,结果显示在三浅裂野牵牛每 9.55 kb 的 EST 出现 1 个 SSR,35 条 ESTs(4.03%) 含有 SSR,也高于玉米(1.4%)、大麦(3.4%)、小麦(3.2%)和略低于水稻(4.7%)^[31]等其他作物。在薯类作物研究中,彭丁文等^[32]对木薯的 EST-SSR 信息分析显示,木薯中的平均出现距离为 6.02 kb,出现频率为 11.02%。黄立飞等^[24]对甘薯的 EST-SSR 信息分析显示,甘薯中的平均分布距离为 11.68 kb,其出现频率为 4.4%。就出现频率而言,三浅裂野牵牛与甘薯相当并低于木薯的出现频率,平均分布距离则均低于甘薯和高于木薯,不同研究结果的差异主要归因于 EST-SSR 搜索的 EST 序列数量多少、异或搜索 SSR 软件的不同算法及所

设定的参数不同造成的,同时也和不同作物 EST 序列来源的组织、器官不同有关,有待于进一步研究验证。

从目前的文献报道来看,大多数植物的 EST-SSR 都是以二、三核苷酸重复基元为主(不计单核苷酸重复),但主要的重复基序类型有所差异。Cardle 等^[28]在对水稻、玉米、大豆、番茄、棉花、杨树和拟南芥的 EST 序列分析后发现,在 EST-SSR 中,最多的是三核苷酸重复类型,其次是二核苷酸。Kantety 等^[31]在大麦、小麦、玉米、燕麦、黑麦、高粱和水稻等主要的禾谷类作物的 EST 序列分析中发现,三核苷酸重复出现次数最多(54%~78%),其次为二核苷酸重复(17%~40%)。另外棉花^[18]、小麦^[33]、黄瓜^[34]、香蕉^[35]、花生^[36]的 EST 中三核苷酸重复在检出的所有 SSR 中所占比例最大(不计单核苷酸重复),分别为 54.8%、52.4%、78%、46.7%、62.8%。在本研究中,二、三核苷酸 SSR 之和占总 EST-SSR 的 68.57%,其中三核苷酸 SSR 为 45.71%,二核苷酸 SSR 占总 SSR 数量的 22.86%,四核苷酸 SSR 和六核苷酸 SSR 占总 SSR 数量均为 11.43%,五核苷酸占的较小,分别为 8.57%。普遍认为,重复基元为三核苷酸的 EST-SSRs 出现频率最高,主要归因于 ESTs 是 cDNA 单边测序的一段序列,除 3'或 5'的非翻译区外,其余属于编码区。三核苷酸重复出现频率高是与其编码区三联体密码子相对应的,而非三核苷酸重复则面临移码突变的选择压^[37]。Hu 等^[14]对甘薯 EST 分析表明,二、三、六核苷酸重复的 SSR 最多,其中三核苷酸重复比例最大,其次为二核苷酸和六核苷酸。黄立飞等^[24]研究显示甘薯 EST-SSR 主要集中在二、三、六核苷酸,其中六核苷酸比例最大,占总 EST-SSR 的 30.96%,其次为三核苷酸重复,再次为二核苷酸重复。不同学者研究结果出现这种差异的原因,可能是各作物供分析的 EST 序列的组织来源不同所致。在对棉花 EST-SSR 分布特征的研究中,王长彪等^[38]和李华盛等^[18]用不同组织来源的 EST 进行分析,得出的结果相差较大,其中三核苷酸重复出现的频率分别为 38.31% 和 54.80%。许多研究发现,大多数植物的 EST-SSRs 主要是三核苷酸重复类型含量最高,其次为二核苷酸重复类型与四核苷酸重复类型。但在三核苷酸重复类型中重复基元出现最高频率依不同作物有所差异^[30]。Kantety 等^[31]统计表明,(CCG)_n 重复基元在小麦中占 32%,高粱中则为 49%。Gupta 等^[39]则认为(AAG)_n 是植物三核苷酸重复的主要类型。本研究结果中,二、三核苷酸重复基元中的优势重复基

元为 AT/AT(14.29%) 和 TCT/AGA(11.43%),二、三核苷酸重复基元优势序列也与 Hu 等^[14] 和黄立飞等^[24] 的研究结果有差异,这可能是由于供分析 EST-SSR 数量偏少导致重复基元表现出明显的偏倚性。

利用甘薯野生近缘种进行 EST-SSRs 筛选和分析,结果表明 SSR 在甘薯基因组中广泛存在,出现频率较高且类型丰富,甘薯野生种 EST-SSR 的标记的开发对丰富甘薯属 EST-SSR 的来源具有积极意义。目前,NCBI 数据库中甘薯及其野生种 EST 序列数量相比较其他作物而言还十分有限。随着数据库的不断更新,我们正在对这些新增的甘薯及其相关野生种 EST 序列作进一步的分析,以期从中开发出更多 EST-SSR 标记。依据本研究结果并结合甘薯 EST-SSRs 的分析,甘薯组 EST-SSR 引物的开发和应用研究,已在本实验室全面展开。相信随着研究的不断深入,更多的甘薯及其野生种 SSR 标记的开发将有助于六倍体甘薯起源进化、基因组结构和功能的研究。

参考文献:

- [1] Kriegner A, Cervantes J C, Burg K, et al. A genetic linkage map of sweetpotato (*Ipomoea batatas* (L.) Lam.) based on AFLP markers [J]. Mol Breed, 2003, 11(3): 169-185.
- [2] 蒲志刚,唐静,谭文芳,等.甘薯抗黑斑病材料 AFLP 标记分子鉴定初步研究[J].西南农业学报,2005,21(1):93-95.
- [3] 吴洁,谭文芳,何俊蓉,等.甘薯 SRAP 连锁图构建淀粉含量 QTL 检测[J].分子植物育种,2005,3(6):841-845.
- [4] Cervantes-Flores J C, Yench G C, Kriegner A, et al. Development of a genetic linkage map and identification of homologous linkage groups in sweetpotato using multiple-dose AFLP markers [J]. Mol Breed, 2008, 21(4): 511-532.
- [5] Kuang-Yi CHANG, Hsiao-Feng L O, Yung-Chang LAI, et al. Identification of quantitative trait loci associated with yield-related traits in sweetpotato (*Ipomoea batatas*) [J]. Botanical studies, 2009, 50: 43-55.
- [6] Zhang D P, Ghislain M, Huaman Z, et al. RAPD variation in sweetpotato (*Ipomoea batatas* (L.) Lam) cultivars from South America and Papua New Guinea [J]. Genet Resour Crop Evol, 1998, 45(3): 271-277.
- [7] Wang J X, He G H, Prakash C S, et al. Analysis of genetic diversity in Chinese sweetpotato (*Ipomoea batatas* (L.) Lam.) germplasm using DNA amplification fingerprinting [J]. Plant Genet Res Newsl, 1998, 113: 13-16.
- [8] Zhang D P, Cervantes J, Huaman Z, et al. Assessing genetic diversity of sweetpotato (*Ipomoea batatas* (L.) Lam.) cultivars from tropical America using AFLP [J]. Genet Resour Crop Evol, 2000, 47(6): 659-665.

- [9] Gichuki S T, Berenyi M, Zhang D P *et al.* Genetic diversity in sweetpotato (*Ipomoea batatas* (L.) Lam.) in relationship to geographic sources as assessed with RAPD markers [J]. *Genet Resour Crop Evol* 2003, 50: 429–437.
- [10] He G, Prakash C S, Jarret R L. Analysis of genetic diversity in a sweetpotato (*Ipomoea batatas*) germplasm collection using DNA amplification fingerprinting [J]. *Genome* 1995, 38(5): 938–945.
- [11] 贺学勤, 刘庆昌, 翟红, 等. 用 RAPD, ISSR 和 AFLP 分析系谱关系明确的甘薯品种的亲缘关系 [J]. *作物学报* 2005, 31(10): 1300–1304.
- [12] Huang J C, Sun M. Genetic diversity and relationships of sweetpotato and its wild relatives in *Ipomoea* series *Batatas* (*Convolvulaceae*) as revealed by inter-simple sequence repeat (ISSR) and restriction analysis of chloroplast DNA [J]. *Theor Appl Genet* 2000, 100(7): 1050–1060.
- [13] Hu J, Nakatani M, Lalusin A G. *et al.* Genetic Analysis of sweetpotato and wild relatives using inter-simple sequence repeats (ISSRs) [J]. *Breed Sci* 2003, 1(53): 297–304.
- [14] Hu J J, Nakatani M, Lalusin A G *et al.* New microsatellite markers developed from reported *Ipomoea trifida* sequences and their application to sweetpotato and its related wild species [J]. *Scientia Horticulturae* 2004, 102(4): 375–386.
- [15] Brown S M, Hopkins M S, Mitchell S E *et al.* Multiple-methods for the identification of polymorphic simple sequence repeats (SSRs) in sorghum (*Sorghum bicolor* L. Moench) [J]. *Theor Appl Genet* 1996, 93: 190–198.
- [16] Varshney R K, Sigmund R, Börner A *et al.* Interspecific transfer ability and comparative mapping of barley EST-SSR markers in wheat, rye and rice [J]. *Plant Sci* 2005, 168: 195–202.
- [17] Thiel T, Michalek W, Varshney R K *et al.* Exploiting EST database for the development and characterization of gen-derived SSR markers in barley (*Hordeum vulgare* L.) [J]. *Theor Appl Genet* 2003, 106: 411–422.
- [18] Yu J K, La R M, Kantety R V *et al.* EST derived SSR markers for comparative mapping in wheat and rice [J]. *Mol Gen Genet* 2004, 271: 742–751.
- [19] 李华盛, 范术丽, 沈法富. 从棉花 ESTs 数据库中筛选微卫星标记的初步研究 [J]. *棉花学报* 2005, 17(4): 211–216.
- [20] 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.
- [21] Chen C X, Zhou P, Choi Y A *et al.* Mining and characterizing microsatellites from citrus EST [J]. *Theor Appl Genet* 2006, 112: 1248–1257.
- [22] 李小白, 张明龙, 崔海瑞. 油菜 EST 资源的 SSR 信息分析 [J]. *中国油料作物学报* 2007, 29: 20–25.
- [23] Jia X P, Shi Y S, Song Y C. Development of EST-SSR in foxtail millet (*Setaria italica*) [J]. *Genet Resour Crop Evol* 2007, 54: 233–236.
- [24] 黄立飞, 房伯平, 陈景益, 等. 甘薯 EST 资源的 SSR 信息分析 [J]. *植物生理学通讯* 2009, 45(1): 23–27.
- [25] Ewing B, Green P. Base-calling of automated sequencer traces using phred II Error probabilities [J]. *Genome Res* 1998, 8: 186–194.
- [26] Huang X, Madan A. CAP3: A DNA sequence assembly program [J]. *Genome Res* 1999, 9: 868–877.
- [27] Weber J L. Informativeness of human (dC-dA) n (dG-dT) n polymorphisms [J]. *Genomics* 1990, 7: 524–530.
- [28] 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.
- [29] Morgante M, Hanafey M, Powell W. Microsatellites are preferentially associated with nonrepetitive DNA in plant genomes [J]. *Nat Genet* 2002, 30: 194–200.
- [30] Gao L F, Tang J F, Li H W *et al.* Analysis of microsatellites in major crops assessed by computational and experimental approaches [J]. *Mol Breed* 2003, 12: 245–261.
- [31] Kantety R V, La R M, Matthews D E *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.
- [32] 彭丁文, 郑柳城, 朱宏波. 木薯 EST 资源的 SSR 信息分析 [J]. *中国农学通报* 2008, 24(2): 433–436.
- [33] 陈军方, 任正隆, 高丽锋, 等. 从小麦 EST 序列中开发新的 SSR 引物 [J]. *作物学报* 2005, 31(2): 154–158.
- [34] 关媛, 李效尊, 潘俊松, 等. 黄瓜果实 EST-SSR 标记的开发 [J]. *分子植物育种* 2007, 5(5): 725–728.
- [35] 王静毅, 陈业渊, 刘伟良, 等. 香蕉 EST-SSRs 标记的开发与应用 [J]. *遗传* 2008, 30(7): 933–940.
- [36] 梁炫强, 洪彦彬, 陈小平, 等. 花生栽培种 EST-SSRs 分布特征及应用研究 [J]. *作物学报* 2009, 35(2): 246–254.
- [37] Metzgar D, Bytof J, Wills C. Selection against frameshift mutations limits microsatellite expansion in coding DNA [J]. *Genome Res* 2000, 10: 72–80.
- [38] 王长彪, 郭旺珍, 蔡彩平, 等. 雷蒙德氏棉 EST-SSRs 分别特征及开发与利用 [J]. *科学通报* 2006, 51(3): 316–320.
- [39] Gupta P K, Balyan H S, Sharma P C. Microsatellites in plants: a new class of molecular markers [J]. *Current Sci* 1996, 70: 45–53.