

文章编号:1003-8701(2004)06-0011-05

简单重复序列的研究与应用

朴红梅,王玉民,刘宪虎,王景余,金成海

(延边大学农学院农学系,吉林 龙井 133400)

摘 要: SSR 是以几个碱基为基本单元的串联重复序列,广泛分布于真核生物基因组中。SSR 信息量高,覆盖整个基因组,共显性遗传,多态性水平高,易用 PCR 分析,是一种非常实用的分子标记。对 SSR 的种类、研究进展、以及 SSR 在品种鉴定、基因作图和标记辅助育种等方面的研究应用现状作了简要概述。

关键词: SSR; 分子标记; 研究进展; 品种鉴定; 基因作图

中图分类号: Q7

文献标识码: A

分子标记是指以生物大分子的多态性为基础的遗传标记,它能够反映植物在遗传物质 DNA 水平下的差异,是植物遗传育种领域内的一项新兴技术^[1]。目前开发的分子标记主要有:限制性片段长度多态性(RFLP)、随机扩增多态性 DNA(RAPD)、扩增片段长度多态性(AFLP)、简单重复序列(SSR)、序列特异扩增区域(SCAR)和数量可变串联重复(VN-TR)6 种^[2]。

SSR 标记是一种以 PCR 技术为基础的分子标记,它既有 RFLP 标记的稳定性、位置确定、共显性等优点,同时又具有 RAPD 标记的成本低、技术简单等特点,由于是共显性标记,在基因作图和基因定位等方面被广泛应用。简单序列重复 (Simple Sequence Repeats)简称 SSR。又称微卫星 DNA(Microsatellite),它是由几个核苷酸构成的一段 DNA 片段。它们普遍存在和随机分布于大多数真核生物的基因组中,重复数具有高频率的变异^[3]。对 SSR 的研究始于动物基因组,最早发现的是由 CA/GT 重复多次构成的一段 DNA。在人类基因组中二核苷酸 (CA)_n 多达 50 000 个,n 的大小在 10~60 之间,三核苷酸、四核苷酸也广泛存在于人类基因组中^[4]。

微卫星位点两侧区域的 DNA 序列较为保守和专一,且重复基因基序数变化不一;可用与两侧保守的 DNA 序列相互补的方式设计特定的寡聚核苷酸引物进行 PCR 扩增,扩增产物可用电泳进行分离。不同基因型的微卫星结构中的基序重复序列不同,重复单位的数量不同,产生了简单序列长度多态性(SSLP)和随机扩增微卫星多态性(RAMP),从而反映出高度的等位基因多样性,可作为一种有价值的遗传标记^[5]。同时,PCR 扩增结合聚丙烯酰胺凝胶电泳分离扩增片段,使得 SSLP 和 RAMP 易于分析使用。由于 SSR 高水平的多态性,容易检测和共显性特点,使得它们成为一个丰富的遗传标记来源。目前含有约 8 000 个位点的高饱和人类微卫星图谱已完成^[6]。一些哺乳动物的 SSR 图谱也被构建,一些主要栽培作物,如大豆、水稻和玉米等的 SSR 类型及频率也相继被评出。虽然微

收稿日期:2004-02-18

作者简介:朴红梅,女,吉林省延边人,延边大学农学院在读硕士,主要从事水稻遗传育种研究。

卫星分子标记研究起步相对较晚,但发展迅速。本文就 SSR 标记在植物尤其在农作物中的研究及应用进行简要概述。

1 SSR 标记的研究进展

随着生物技术在作物育种上广泛应用,微卫星分子标记作为一种比较理想的分子标记广泛应用于遗传图谱构建居群遗传学等各个领域的研究。Russell 等(1977)发现 11 个大麦 SSR 中利用 4 个 SSR 组成的 3 个不同组合均可区分 24 个大麦品种,错误概率仅为千分之一,来自相同亲本的不同基因型也可区分^[7]。关于微卫星的产生机制,普遍认为是 DNA 复制过程中滑动,或 DNA 复制和修复时滑动链与互补链碱基错配,导致一个或几个重复单位的插入或缺失。从统计学角度看,Valdas 等认为,微卫星的产生采用的是逐步突变模式 (Stepwise mutation model, SSM), 即每一次突变均增加一个或数个重复单位,从而导致一个新的等位基因的出现^[4]。微卫星的功能目前尚不甚明了,但已发现微卫星能参与遗传物质的结构改变,基因调控及细胞分化等过程,有自身特异性结合蛋白,还能直接编码蛋白质,是一种非常活跃的碱基序列。

Weising(1992)用(GATA)4 作探针在鹰嘴豆的 13 份材料中呈多态性,在 15 个番茄品种中也呈多态性。此外,该探针的 DNA 还可区分油菜和甜菜的特殊个体^[8]。Bryan 等(1997)对 200 余个 SSR 克隆进行测序,并对其中 153 个 SSR 设计了引物。选用 49 个引物对 10 个小麦品种进行 PCR 扩增,平均每个位点能检测到 3.5 个等位变异^[9]。梁凤山等利用与野生稻高产基因紧密连锁的 SSR 分子标记,对野生稻与超级杂交稻亲本 9311 的回交群体进行田间农艺性状选择和分子标记辅助选择,得到一些增产潜力很大的株系。并得出“同时具有多个高产基因标记的株系表现明显的增产优势”的结论^[10]。关荣霞等对 700 个大豆 SSR 位点进行筛选,得出 216 个多态性位点,选出 143 个带型清晰易辨的 SSR 位点,对 125 个 F₂ 个体的基因组 DNA 进行扩增,银染检测后分析。结果表明,117 个 SSR 位点构建了 28 个连锁群,覆盖大豆基因组 1 893.7 cm^[11]。谭君等应用 SSR 分子标记对 65 份玉米杂交种进行研究,从 96 对 SSR 引物中筛选出 22 对具有较高多态性的 SSR 引物,其中 7 对引物组合可区分所研究的 65 个材料,并将每个材料的 SSR-DNA 指纹转换为计算机指纹,由计算机辅助进行种子鉴定^[12]。杨振宇等应用 SSR 分子标记对中国研究所保存的中国东北栽培大豆 194 份(161 份地方品种/33 份栽培品种)、北海道及九州在日本大豆 59 份(28 份地方品种/31 份栽培品种),计 253 份供试材料,采用 7 个连锁群 12 个 SSR 标记对上述品种进行分子标记,进行中日大豆的遗传多样性分析。分析结果表明,中国大豆的等位基因的平均数为 5.4,而日本的为 3.6,说明中国大豆的遗传多样性较日本的丰富。在 SSR 分子标记的基因型的基础上进行了聚类分析,其结果为日本大豆几乎成为了中国大豆的一亚组(Fig.1)^[20]。大豆花叶病毒病(*Soybean mosaic virus*, SMV)也是一种世界性病害,在世界各大豆产区均有不同程度发生。Yu 等通过 PI96983 (R)× Lee68(S)的 107 株 F₂ 个体,利用 RFLP 和 SSR 标记将 G1 株系的抗病基因 Rsv 定位在 E 连锁群上。他们的结果还表明,RFLP 标记 pA186、pK644 及 SSR 标记 SM176 与 Rsv 紧密连锁,遗传距离分别为 0.5、1.5 和 2.1 CM^[21]。

2 植物中 SSR 的种类、数量、分布及多态性

SSR 在各种真核生物的基因组序列中普遍存在,而且它的分布是随机的,大约每隔 10~50 KB 就存在一个 SSR。SSR 既可以存在于内含子中,也可以存在于外显子中。在不

同的生物体间 SSR 的数量、重复次数、重复单位、拷贝数变异及染色体分布等有很大的差别^[13]。与哺乳动物相比,植物中 SSR 数量约为前者的 1/6。在人类中,AC/GT 重复数量最大,其次为 AG/TC,AC/TG 数量极小。在植物核基因中,各种 SSR 数量从大到小依次为 (AT)_n, (A)_n, (T)_n, (AG)_n, (CT)_n, (AAT)_n, (ATT)_n, (AAC)_n, (GTT)_n, (AGC)_n, (GCT)_n, (AAG)_n, (CTT)_n, (AATT)_n, (TTAA)_n, (AAAT)_n, (ATTT)_n, (AC)_n, (GT)_n。就分布频率而言,平均每 23.3 千碱基中有一个 SSR,而每 62 千碱基中有一个 AT/TA 类型 SSR。所有的单核苷酸及双核苷酸重复类型 SSR 均位于非编码区,而 57%左右的含 GC 碱基对的三核苷酸重复位于编码区。SSR 不仅存在于核 DNA 中,细胞质 DNA 中也有分布,但数量极少,其中 A/T 重复占绝对优势。AT/TA 重复仅发现于少数几种植物的叶绿体 DNA 中^[14]。上述数据均来自对现有基因数据库的搜索。鉴于绝大多数 SSR 存在于非编码区或许不能较好反映 SSR 存在的真实情况。所以,在研究 SSR 种类、数量及分布时,可利用 SSR 作探针与大片段基因文库进行杂交。但因 AT/TA,GC/CG 易在 DNA 分子中形成二级结构,降低了杂交效率,不能进行可靠的估计。

植物中 SSR 的多态性与动物相差不多。研究表明,不同作物中均存在 SSR 多态性。除个别作物外(番茄多态性水平较低),绝大多数作物如大豆、水稻、小麦和玉米等 SSR 位点上的等位基因数目可达十几个甚至几十个,表现出很高的多态性。对人类 SSR 的研究表明,核心序列的重复次数与 SSR 的长度多态性有密切关系,重复次数增加,SSR 的长度多态性增高,这在部分作物中也得到证实。不同类型 SSR 的多态性水平差异很大,二核苷酸重复构成的 SSR 的多态性水平要高于其他类型。同一类型 SSR 的不同位点之间多态性水平也不一致。鉴于(AT)_n 是植物中数量最大的一类 SSR,这一趋势加在其他作物中得以证实,可大大提高整个作物的 SSR 多态性水平。但由于(AT)_n 类型探针易形成二级结构,降低了筛选效率,影响了这一类型 SSR 的分离,但研究表明,(AT)_n 类型 SSR 还是可以分离的。

3 SSR 在研究中的应用

3.1 建立 DNA 指纹、品种鉴定及种子纯度检测

作物的性状大多属数量性状,由多基因控制,易受环境影响。利用传统形态学鉴定方法区分新品种时间长、费用高、难度较大和准确性差。现有的鉴定方法中,同工酶和蛋白质电泳可检测的位点少,蛋白质类型不多,多态性水平低,对同一作物不同品种不能有效区分。RFLP 技术繁琐,对人员和设备要求高,且多态性低于 SSR^[15],RAPD 可靠性较差,难以在品种鉴定中广泛应用。SSR 与其他技术相比,SSR 分布于整个基因组中,多态性水平高,易于分析,不受环境影响,在品种鉴定中显示出巨大的潜力。

1994 年 Yanagisawa 利用 SSR 研究了 43 份栽培大豆亲缘关系,获得了各个品种中的明显不同的指纹,品种中各个体间的 DNA 指纹基本相同。彭锁堂等用分布于 12 条染色体上的 26 对 SSR 引物对 9 个杂交水稻组合及亲本进行了 SSR 分析,能够有效区分所有恢系和大部分不育系^[16]。苏顺宇等筛选一批常用的高多态性的 SSR 引物对 12 个杂交种进行纯度鉴定。结果表明,SSR 鉴定的平均纯度与田间鉴定结果无显著差异,可用于杂交水稻种子纯度鉴定^[17]。SSR 遗传共显性的特点使其在杂交种鉴定中有极大应用价值,它可将双亲甚至一些异源花粉造成的杂交种区分开。这均表明 SSR 在农作物纯度鉴定上可望获得广泛应用。

3.2 种质资源保存、评价和利用

在种质资源保存研究中,可利用 SSR 的高分辨率鉴定品种真实性,去除重复种子,防止混杂,评价不同个体、群体之间的近缘程度,分析其遗传构成,区分个体、品种及群体内的变异,还可分析现代作物品种相对于原始资源的遗传构成。赵勇等^[18]利用 16 对水稻功能基因的 SSR 的引物研究了 23 份不同来源的水稻种质资源的遗传多样性,共检测出 78 个等位基因变异。说明水稻功能基因在不同亚种、不同产地来源和不同生态型的水稻之间存在差异。类似研究有利于丰富作物基因库,评价育种家对种质资源中变异的利用程度。

3.3 基因标记与作图

SSR 的高度变异性,在整个基因组均匀分布,共显性及易于分析等特点使其成为构建基因组图谱的最佳选择。利用 SSR 作图极易获得双亲间等位基因的差异。由于植物中 SSR 数量仅为哺乳动物的 1/6,构建完全由 SSR 组成的遗传图谱尚不现实。但可用 SSR 补充利用 RFLP 构建的遗传图谱,填补其空白区域。

在水稻中已把 120 个 SSLP 整合到现成的 RFLP 遗传图谱上^[19]。这些 SSR 分布到水稻全部 12 条染色体中。在大豆、大麦和番茄等作物中也已将 SSR 定位到连锁图中。寻找与目标性状紧密连锁的分子标记是分子辅助选择育种的基础。迄今已有越来越多关于 SSR 多态性与抗病基因、品质性状、QTL 位点之间紧密连锁的报道。如在水稻中发现一个高多态性 SSR 位于水稻蜡质基因的 5 端,可用来判断水稻的淀粉性状。因此,可利用 SSR 与控制性状基因连锁从种质资源中筛选抗病虫特殊性状等基因。

3.4 系谱分析及分子标记辅助育种

在系统育种中,大多数品种都有详细的系谱记录。SSR 标记可分析系谱育种过程中等位基因的转移。在水稻中,已经证明 SSR 标记能十分稳定和可靠地在系谱中追溯目的单基因或数量性状基因座流^[4]。目前已开发出可把分子标记数据、系谱关系及表现型综合在一起的软件,大大方便了在育种中的应用。SSR 标记是辅助选择育种的理想分子标记。应用 SSR 标记进行大豆抗胞囊线虫的分子辅助选择育种便是成功的典范。另外一些重要性状的有效 SSR 标记也可用于分子标记辅助选择。

除上述应用外,SSR 还可用于估算遗传距离、亲缘关系研究、预测杂种优势和利用细胞质单亲遗传特性进行群体遗传进化研究等。

4 结论及展望

SSR 标记是较为理想的遗传标记。与其他标记相比,以微卫星序列为基础的 SSR 标记显示了独特的优越性:SSR 标记数量丰富,覆盖整个基因组,揭示的多态性高;具有复等位基因的特性,提供的信息量高;以孟德尔方式遗传,呈共显性;每个位点由设计的引物序列决定,便于不同实验室相互交流合作开发引物,获得的资料能够在不同的实验室重复并共享,虽然 SSR 引物的开发费用相当高,但由于其巨大的应用潜力,在一些主要农作物中 SSR 引物的开发正在进行中,并每隔一段时间就有新的引物增加,为该技术在指纹图谱中的应用提供了良好的条件^[2]。尽管需要设计合成引物,且花费较大,但由于它具有异常丰富、分布广泛、多态性强和共显性遗传易于分析使用等特点,使得 SSR 在农作物研究中具有良好的应用前景。

目前,SSR 在玉米、水稻、大豆、番茄和棉花等许多重要作物上刚刚开始研究和应用。各类作物中发现的 SSR 位点较少,大多为二核苷酸重复类型。由于不同类型及不同位点的 SSR 多态性水平不同,研究者应针对研究目标,确定基因组的研究区域,开发更

多的信息量更高的 SSR 标记。随着更多 SSR 位点的发现及研究应用,在水稻、玉米、大豆和小麦等主要农作物上,可构建完全由 SSR 标记组成的基因组图谱,进一步推动 SSR 标记在种子纯度检验、DNA 指纹建立、遗传标记、基因作图及其他研究领域中更广泛的应用。

参考文献:

- [1] 郭小平,等. SSR 技术及其在植物遗传育种中的应用[J]. 华北农学报, 1998, 13(3): 73-76.
- [2] 危文亮,等. 分子标记在作物育种中的应用[J]. 生物技术通报, 2000, (2): 12-16.
- [3] 朱振东,等. 小麦 SSR 标记的发展及应用[J]. 遗传 HEREDITAS(Beijing), 2003, 25(3): 355-360.
- [4] 李汝玉,等. 简单序列重复(SSR)及其在品种鉴定及种子检测中的应用[J]. 种子, 1999, (5): 33-35.
- [5] 万平,等. SSRs 标记与植物遗传育种研究[J]. 安徽农业大学学报, 1998, 25(1): 92-95.
- [6] Weissenbach J. The human genome project: from mapping to sequencing[J]. Clin Chem Lab Med, 1998, 36: 511-514.
- [7] 赵婷,等. 种子纯度检验技术及进展[J]. 世界农业, 2000, (11): 26-28.
- [8] 张文兰,等. 作物品种纯度快速鉴定技术及其应用[J]. 山东农业科学, 2001, (2): 45-48.
- [9] 马渐新,等. 小麦抗条锈病基因定位及分子标记研究进展[J]. 生物技术通报, 1999, (1): 2-6.
- [10] 梁凤山,等. 应用 SSR 方法对野生稻高产基因的分子标记辅助选择[A]. 2003 年全国作物遗传育种学术研讨会论文集[C]. 中国作物学会, 2003.
- [11] 关荣霞,等. 大豆 SSR 标记遗传连锁图谱的构建研究[A]. 2003 年全国作物遗传育种学术研讨会论文集[C]. 中国作物学会, 2003.
- [12] 谭君,等. 利用 SSR-DNA 指纹的计算机分析鉴定玉米种子[A]. 2003 年全国作物遗传育种学术研讨会论文集[C]. 中国作物学会, 2003.
- [13] 朱宏波,等. 利用水稻基因组序列数据开发 SSR 标记的方法[J]. 分子植物育种, 2003, 3(2): 273-276.
- [14] 李汝玉. 简单序列重复(SSR)及其在农作物研究中的应用[J]. 山东农业科学, 1999, (4): 45-49.
- [15] 朱作峰,等. 水稻品种 SSR 与 RFLP 及其与杂种优势的关系比较研究[J]. 遗传学报, 2001, 28(8): 738-745.
- [16] 彭锁堂,等. 我国主要杂交水稻组合及其亲本 SSR 标记各纯度鉴定[J]. 中国水稻科学, 2003, 17(1): 1-5.
- [17] 苏顺宗,等. 利用 SSR 鉴定水稻杂交种子纯度的研究[J]. 种子, 2003, (1): 23-25.
- [18] 赵勇,等. 利用水稻功能基因 SSR 标记鉴定水稻种质资源[J]. 中国农业科学, 2002, 35(4): 349-353.
- [19] McCouch S R, et al. Microsatellite marker development, mapping and applications in rice genetics and breeding. Plant Molecular Biology, 1997, 35: 89-97.
- [20] 邹继军,等. 大豆抗病基因定位的分子标记研究进展[J]. 中国油料作物学报, 2000, (4): 75-78.

A Review on Studies and Application of Simple Sequence Repeats

PIAO Hong-mei, WANG Yu-min, LIU Xian-hu, et al.

(College of Agronomy Yanbian University, Longjing 133400, China)

Abstract: SSR(simple sequence repeats) is a simple, tandem repeated several nucleotide sequence motifs. SSR is widely distributed over the genome of eukaryote. SSR is very usually as genetic marker because it contains plenty of information, covers the whole genome, co-dominant heredity, and is easily detected by PCR. Types of SSR, progress in studies, and its application in identification of varieties, genetic mapping and marker-assisted selection were briefly reviewed in the paper.

Key words: SSR; Molecular marker; Review of progress; Identification of varieties; Genetic mapping