

王凌云,郭明,赵艳,等. 谷子蔗糖合成酶基因家族鉴定及生物信息学分析[J]. 江苏农业科学,2017,45(15):30-34.
doi:10.15889/j.issn.1002-1302.2017.15.007

谷子蔗糖合成酶基因家族鉴定及生物信息学分析

王凌云¹, 郭明³, 赵艳¹, 瓮巧云¹, 马海莲¹, 宋晋辉¹, 袁进成¹, 董志平², 刘颖慧¹

(1. 河北北方学院,河北张家口 075000; 2. 河北省农林科学院谷子研究所,河北石家庄 050001;

3. 河北省沽源县农牧局,河北张家口 075000)

摘要:蔗糖合成酶(sucrose synthase, SuS)是蔗糖代谢的关键酶,在植物的代谢反应中起重要作用。以谷子基因组数据库为平台,借用生物信息学手段对谷子 *SiSuS* 基因家族进行挖掘和分析。结果表明,谷子的 *SiSuS* 基因家族包括 9 个基因,它们不均匀地分布在 5 条染色体上。该家族氨基酸序列长度为 809 ~ 1 088 aa,外显子数目为 10 ~ 16 个,大多数蛋白质为弱酸性。谷子 *SiSuS* 的氨基酸序列具有 9 个保守基序;进化树分析结果表明,谷子、水稻、高粱 *SiSuS* 蛋白聚在一起。

关键词:谷子;蔗糖合成酶;*SiSuS* 基因;生物信息学分析;蛋白结构

中图分类号: S515.01 **文献标志码:** A **文章编号:** 1002-1302(2017)15-0030-04

蔗糖合成酶(sucrose synthase, SuS)作为植物蔗糖代谢的重要酶之一,对植物的生长具有重要作用。蔗糖合成酶能够参与蔗糖的代谢和调控蔗糖的输入,影响细胞分化与纤维壁的形成,调节淀粉的合成,对作物产量和品质调控都有重要的意义^[1]。蔗糖合成酶基因广泛存在于高等植物中,目前已经从马铃薯、胡萝卜、玉米、柑橘等大量植物中获得 *SuS* 基因^[2]。大量研究显示, *SuS* 是一个小的基因家族,在大多数植物中至少含有 3 个 *SuS* 基因,拟南芥、苜蓿、橡树均含有 6 个 *SuS* 基因^[3-5],玉米中至少有 5 个 *SuS* 基因,水稻中已经发现 9 个 *SuS* 基因^[6],烟草中发现 14 个 *SuS* 基因^[7],但是对谷子 *SuS* 基因的研究尚未见报道,因此,本研究以谷子为对象进行分析。

谷子(*Setaria italica*)作为我国传统粮食作物,种植面积广,同时营养价值高,有“百谷之长”的美誉^[8]。谷子基因组小且为二倍体,与水稻、高粱、玉米共线性高,是禾本科基因组研究的模式植物之一,也是研究 C_4 植物的模式植物。目前谷子的全基因组测序已经完成,这为谷子分子生物学研究奠定了良好的基础。本研究利用生物信息学相关技术鉴定谷子 *SuS* 基因家族,并对该家族序列及蛋白系统进行比较分析,对了解 *SiSuS* 基因家族在谷子蔗糖代谢中的作用具有非常重要的意义。

1 材料与方法

1.1 谷子 *SuS* 基因的鉴定及序列分析

首先在 Pfam 数据库(<http://pfam.xfam.org/>)^[9]中下载 *SuS* 家族的隐马尔可夫模型文件(Pfam 号码: PF00862),从

收稿日期:2016-03-10

基金项目:国家科技重大专项(编号:2014ZX0800909B);河北省科技厅项目(编号:16226310D);河北北方学院重大项目(编号:ZD201305)。

作者简介:王凌云(1988—),女,山东莱州人,硕士研究生,主要从事谷子成分研究。E-mail:984112812@qq.com。

通信作者:刘颖慧,博士,教授,硕士生导师,主要从事植物基因功能研究。E-mail:leely519@126.com。

Gramene(<http://www.gramene.org/>)中输入 Pfam 号码进行相似性搜索,找到与谷子相关基因的 ID 及其蛋白序列,进行重复性比对,除去重复项和冗余,得到无重复的基因、转录本、蛋白 ID 以及蛋白序列和外显子数量,共得到 9 条蛋白并命名为 *SiSuS1* ~ *SiSuS9*。

利用 SMART 网站(<http://smart.embl-heidelberg.de/>)和 CDD(<http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>)网站^[10-11]对蛋白的结构域进行检测。在 ProtParam 网站(<http://web.expasy.org/protparam/>)对谷子 *SiSuS* 蛋白的相关信息(分子量、氨基酸数目、等电点)进行分析^[12]。

1.2 染色体定位

结合谷子基因组数据库中的信息,从 Ensembl Plants(http://plants.ensembl.org/Setaria_italica/Info/Index)中查询谷子染色体的长度,使用 Adobe illustrator CS5 软件绘制基因染色体上的相对位置。

1.3 *SiSuS* 基因家族蛋白的绘制

分析蛋白质结构域的位置,使用 PROSITE 在线网站(<http://prosite.expasy.org/>)^[13]分析 9 个 *SiSuS* 蛋白序列,得出每个蛋白相应的蔗糖合成酶结构域、糖基转移酶(glycosyltransferase)结构域、蔗糖磷酸合成酶(sucrose-phosphate synthase)结构域的位置。

1.4 *SiSuS* 基因家族进化树的绘制及内含子、外显子的结构分析

利用 Clustal X^[14]对比所得到的 *SiSuS* 蛋白序列并下载分析结果。使用 MEGA 6.0 软件采用邻接法(NJ 法)绘制进化树,Bootstrap 设置为 1 000。将 Fasta 格式的谷子 *SiSuS* 基因编码的核苷酸序列和相应的全核苷酸序列,通过 GSD 软件^[15]绘制谷子 *SiSuS* 基因的内含子和外显子的结构模式图。

1.5 谷子 *SiSuS* 基因编码的序列分析及模体分布

利用 MEME(<http://meme.nbcr.net/meme/cgibin/meme.cgi>)^[16]将 Fasta 格式的谷子 *SiSuS* 基因编码的氨基酸序列进行序列分析及保守性基序分布,设定序列长度范围为 20 ~ 300 aa。

1.6 谷子、高粱和水稻 *SuS* 基因家族的比较分析

将获得的谷子、高粱和水稻 *SuS* 基因的氨基酸序列以 Fasta 格式保存,采用“1.3”节的方法绘制这 3 种植物 *SuS* 基因的系统进化树,对不同物种 *SuS* 基因家族进行比较分析。

2 结果与分析

2.1 谷子 *SiSuS* 基因家族的鉴定及序列分析

由表 1 可知,通过对谷子基因组数据库的检索共鉴定出

9 个 *SuS* 基因,命名为 *SiSuS1* ~ *SiSuS9*。分析 9 个 *SiSuS* 蛋白序列发现不同蛋白差异很大,*SiSuS* 蛋白的氨基酸个数范围为 809 ~ 1 088 aa,分子量范围为 92.34 ~ 120.00 ku,等电点范围为 5.88 ~ 7.49,基因含有 10 ~ 16 个外显子。从蛋白的基本特点可知,*SiSuS* 蛋白无论从序列的长度还是蛋白的特性变化都很大,表示该基因家族蛋白执行不同的功能。值得注意的是,多数 *SiSuS* 蛋白的等电点小于 7.00,表示多数 *SiSuS* 基因可能编码弱酸性蛋白,在酸性的亚细胞环境中发挥作用。

表 1 谷子 *SiSuS* 基因的鉴定及特性

| 基因名称 | 转录名称 | 染色体 (条) | 基因位置 | 外显子数 (个) | 开放阅读框 长度(bp) | 蛋白质预测 | | |
|---------------|-----------|------------|-------------------------|-------------|-----------------|-----------|---------|------|
| | | | | | | 氨基酸个数(aa) | 分子量(ku) | 等电点 |
| <i>SiSuS1</i> | Si000130m | 5 | 44 828 497 ~ 44 833 956 | 12 | 5 460 | 1 088 | 120.00 | 6.26 |
| <i>SiSuS2</i> | Si005783m | 4 | 32 007 425 ~ 32 016 125 | 13 | 8 701 | 978 | 109.22 | 7.40 |
| <i>SiSuS3</i> | Si005845m | 4 | 40 061 875 ~ 40 065 988 | 16 | 4 114 | 856 | 97.60 | 7.49 |
| <i>SiSuS4</i> | Si005859m | 4 | 2 940 746 ~ 2 949 224 | 15 | 8 479 | 830 | 95.01 | 5.88 |
| <i>SiSuS5</i> | Si013170m | 6 | 16 670 386 ~ 16 680 956 | 14 | 10 571 | 1 061 | 117.93 | 5.96 |
| <i>SiSuS6</i> | Si016229m | 1 | 5 697 446 ~ 5 705 143 | 13 | 7 698 | 964 | 108.36 | 6.52 |
| <i>SiSuS7</i> | Si020148m | 1 | 41 988 015 ~ 41 991 404 | 10 | 3 390 | 849 | 96.89 | 6.95 |
| <i>SiSuS8</i> | Si034282m | 9 | 43 201 562 ~ 43 208 087 | 15 | 6 526 | 816 | 92.95 | 5.99 |
| <i>SiSuS9</i> | Si034293m | 9 | 46 886 615 ~ 46 892 456 | 14 | 5 842 | 809 | 92.34 | 6.10 |

2.2 谷子 *SiSuS* 基因家族内含子、外显子的分析

为进一步研究 *SiSuS* 基因的特性,使用 GSDS 2.0 软件绘制 *SiSuS* 基因的结构模式图,得到谷子 *SiSuS* 基因的内含子-外显子结构。由图 1 可见,所有的 *SiSuS* 基因都含有内含子,

除 *SiSuS7* 基因含有 9 个内含子外,其他 8 个基因的内含子数均大于 10,*SiSuS3* 基因的内含子最多,含有 15 个内含子。由此可知,*SiSuS* 基因结构较为复杂。

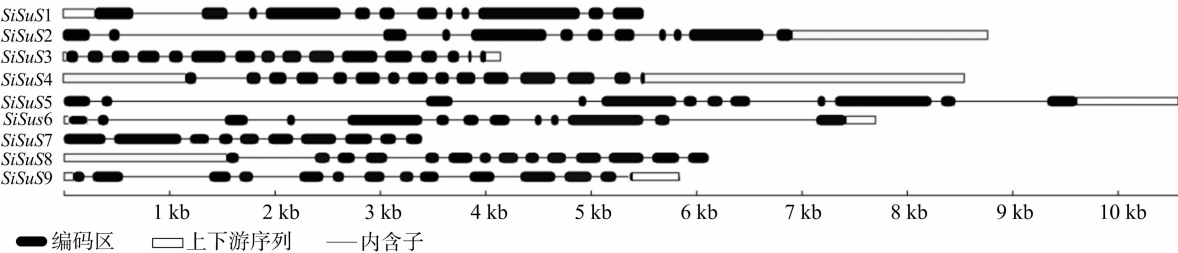


图1 谷子 *SiSuS* 基因的内含子-外显子结构

2.3 谷子 *SiSuS* 基因染色体定位

由图 2 可知,*SiSuS* 基因在染色体上分布不均匀,在谷子的 9 条染色体中除 2、3、7、8 号染色体外,其余染色体上均含有 *SiSuS* 基因。5 号、6 号染色体上各只有 1 个 *SiSuS* 基因;1 号、9 号染色体上各含有 2 个 *SiSuS* 基因;4 号染色体上含有 3 个 *SiSuS* 基因,数量最多。其中,除基因 *SiSuS6*、*SiSuS4*、*SiSuS5* 位于染色体的中上端外,其余基因均位于染色体的下端。

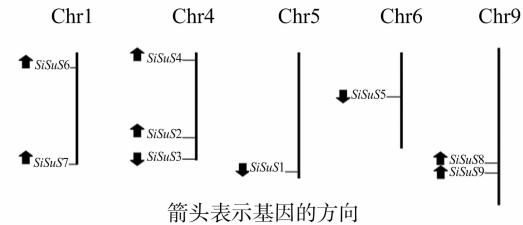


图2 谷子 *SiSuS* 基因的染色体分布及扩增模式

2.4 谷子 *SiSuS* 蛋白的进化树与结构域分析

运用 MAGA 5.1 软件分析谷子 9 个 *SiSuS* 蛋白的进化,

2.5 谷子 *SiSuS* 蛋白的序列分析及模体分布

模体(motif)是序列中局部的保守区域。进一步分析谷

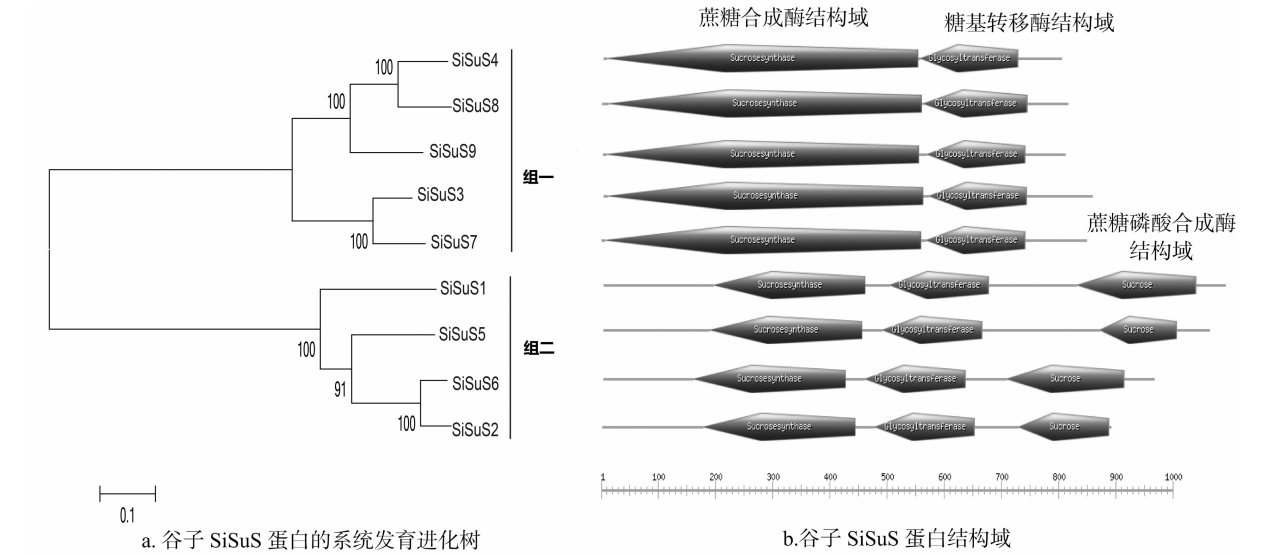


图3 谷子 SiSuS 蛋白的进化树和结构域

子的 9 个 SiSuS 蛋白结构域,利用 MEME 对 SiSuS 蛋白进行基序分析,设定序列长度为 20 ~ 300 aa,设定大范围的长度序列可多样性搜索蛋白的模式序列。由表 2 可知,SiSuS 蛋白 3 个结构域又可以细分为 9 个基序,9 条序列氨基酸个数范围在 55 ~ 200 aa 之间。其中,motif1、motif3、motif4、motif7、motif9

5 个保守基序包含的氨基酸数量均大于 100 aa,基序 motif 2、motif 5 包含的氨基酸数量较少,均为 55 aa。通过软件分析得出 motif 3、motif4、motif7、motif8 共同组成了谷子蔗糖合成酶结构域,motif1、motif2、motif5、motif6 共同组成了糖基转移酶结构域,motif 9 构成了蔗糖磷酸合成酶结构域。

表 2 谷子 SiSuS 蛋白的基序组成

| 基序 | 氨基酸个数 (aa) | 序列 |
|---------|------------|---|
| motif 1 | 152 | QKGPLIIGNYS DGNL VASLLSHKLGVTQCTIAHALEKTKYPDSDIYWKKFDDKYHFSCQFTADLIAMNHADFIITSTYQEIAGSKDTVGQYESHYAFTMPGLYRVRVHGIDVDFDPKFNIVSPGADMSIYFPFTEKQKRLTDLHPEIEELIYS |
| motif 2 | 55 | QHNVRVRNGELYRLIADTKGAFVQPAFYEAFFGLTVIEAMTCGLPTEATKNGGPAEI |
| motif 3 | 200 | SMKPLLDFFLLAHNYKGMTMLNDRIDSLSKLQSAALLKAEFLSGLPADTPYSKFHRFQEWGLEKGWGD TAERCLETIHLILLDLLQAPDPSNLEKFLGRIPMIFNVVIFSPHGYFGQANVLGLPDTGGQVYVILDQVRALEDEMILLRIKQGLDITPKILIVTRLIPDAKGTKCNQRLEKVEGTEHCDILRVFPRTENGI |
| motif 4 | 134 | PVWPAVIHGHYADAGDAAALLSGALNVPMVFTGHFLGCKDKLEGLLKQCRQRTREEINMTYKIMRRIEAEELSLDASEI VIASTRQEIEEQWGLYDGFVDVILARKLRARVKRGANCYGRFMPRMVHIPPGMEFGHI |
| motif 5 | 55 | NDEHMRFLTDPNKPIIFSMARLDPVKNTIGLVEAFGECRRRLRELANLVLVMGNRD |
| motif 6 | 76 | FEKCKEDPSYWNKISQAGLQRIYEKYTWKIYSERLMTLGGVYGFWKYVSKLERRRTRYLEMFYALKFRELAKTVP |
| motif 7 | 105 | RERGHFSPARYFVEEVITGYDETDLTKTWLRANAMRSPQERNTRLENMTWRIWNLARKKKFEFEKEEACRLSKRRLETEKARADATADLSEDLFEGEKGEDAGDPS |
| motif 8 | 77 | DKLYIVLISLHGLIRGENMELGRDSDTGGQVKYVVELAKALSSSPGVYRVDLLTRQILAPDFDRSYGEPTMLASTS |
| motif 9 | 110 | QVIFEDEEHSSTYCLAFRVVPNHLPLKELRKLMIQSLRCHALYNHDATRLSVIPIHASRSQALRYLCVRWGIELPNMAVVGESGSDSYEELLGGLHKTHILKGEFN |

进一步分析谷子 SiSuS 蛋白 9 个基序的保守程度(图 4 - a),在相同部位蔗糖合成酶结构域中 SiSuS3、SiSuS4、SiSuS7、SiSuS8、SiSuS9 蛋白含有 motif 3, SiSuS1、SiSuS2、SiSuS5、SiSuS6 则含有 motif 7、motif 4、motif 8,这表明该蛋白序列可能存在缺失。进一步分析发现,9 个 motif 中 motif 8 的保守性较高,保守比例为 71.4%;而含 200 个氨基酸的 motif 3 保守性最低,其中许多位置的氨基酸多样性程度较高,其保守比例仅为 42%(图 4 - b)。

2.6 谷子、高粱、水稻 SuS 基因家族进化树比较

为了更深入研究谷子 SuS 基因家族的进化关系,选取高粱、水稻 SuS 与谷子 SuS 进行比较和分析。其中,高粱和水稻的基因均为 9 个,共 27 个 SuS 基因。如图 5 所示,27 个基因

分成了 4 个分支,命名为 Class A ~ Class D。除 Class C 亚支含有 2 种植物的 SuS 基因外,其他 3 个亚支都含有 3 种植物(高粱、水稻、谷子)的 SuS 基因;Class A 亚支中 SuS 基因最多,有 10 个基因,Class C 亚支中 SuS 基因最少,仅有 2 个基因。通过对谷子、水稻、高粱 SuS 基因的进化关系分析,发现 9 对物种间的直系同源基因,1 对水稻的旁系同源基因,约占所有基因的 74.1%,表明这些基因在谷子、水稻和高粱基因组中,按照各自物种的特异方式进行了扩展,这种现象在其他植物基因家族中也普遍存在。

3 讨论与结论

SuS 基因家族是参与蔗糖代谢的重要因子,在植物的生

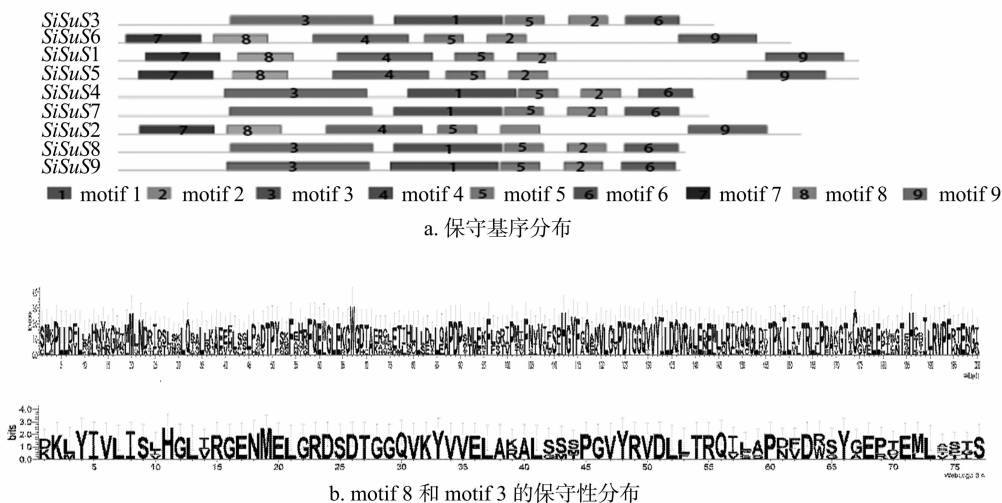
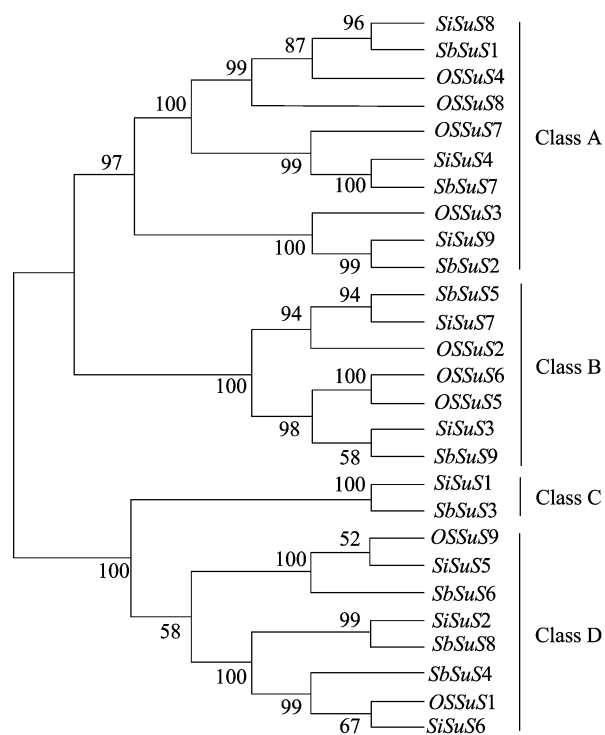


图4 谷子 SuS 蛋白保守基序分布及部分基序保守性分析



SiSuS—谷子 SuS 基因家族; OSSuS—水稻 SuS 基因家族;
SbSuS—高粱 SuS 基因家族

图5 高粱、水稻、谷子 SuS 基因家族的系统发育树

长发育过程中发挥着至关重要的作用。1955 年 Cardini 等首次在小麦胚芽中发现了蔗糖合成酶^[17], 自此以后蔗糖合成酶基因从多种植物中分离得到, 其中多数为玉米、水稻等淀粉储存植物, 在拟南芥、烟草等蔗糖储存植物中也相继获得。谷子是我国传统的淀粉植物, 也是 C₄ 植物的模式植物, 但是目前还没有对其 SuS 基因家族进行系统研究和报道, 基于此, 本研究开展了谷子全基因扫描, 深入分析 SuS 基因家族的特征、特性。

在许多植物中蔗糖合成酶以不同亚型形式存在, 这些亚型至少由 2 个基因编码, 甚至更多, 玉米中至少有 5 个 SuS 基因, 水稻中已经发现 9 个 SuS 基因^[18]。这些基因的结构、功

能和染色体定位各不相同, 白杨中蔗糖合成酶基因家族的 15 个成员分别在 2、4、6、15、17、18 号染色体上^[19], 谷子中蔗糖合成酶基因分别在 1、4、5、6、9 号染色体上。蔗糖合成酶在植物中催化蔗糖 + 二磷酸尿(嘧啶)苷 ⇌ 果糖 + 尿嘧啶核苷 - 5'-二磷酸葡萄糖可逆反应, 已经公认蔗糖合成酶基因含有 2 个结构域: 蔗糖合成酶结构域、糖基转移酶结构域, 在拟南芥中所有的 AtSUS 家族都有这 2 个结构域, 本研究的结果与之相符。

随着生物信息学的发展和完善, 植物蔗糖合成酶基因家族的进化与分类研究备受关注。进化分析表明, SuS 基因家族可分为 4 族: 单子叶族、双子叶 SuS1 族、双子叶 SuS2 族、NG 族。本研究初步分析了谷子、水稻、高粱 SuS 基因家族的进化关系, 结果表明, 这些基因的同源性高, 可为深入研究 SuS 家族对谷子生长调控的机制提供参考, 对深入揭示蔗糖合成酶在谷子的生物学功能以及了解整个糖代谢过程具有重要意义。

参考文献:

- [1] Koch K E, Wu Y, Xu J. Sugar and metabolic regulation of genes for sucrose metabolism: potential influence of maize sucrose synthase and soluble invertase responses on carbon partitioning and sugar sensing [J]. Journal of Experimental Botany, 1996, 47(301): 1179 - 1185.
- [2] 卢合全, 沈法富, 刘凌霄, 等. 植物蔗糖合成酶功能与分子生物学研究进展[J]. 农业生物技术学报, 2005, 21(7): 34 - 37.
- [3] Baud S, Vaultier M N, Rochat C. Structure and expression profile of the sucrose synthase multigene family in Arabidopsis [J]. Journal of Experimental Botany, 2004, 55(396): 397 - 409.
- [4] Horst I, Welham T, Kelly S, et al. TILLING mutants of Lotus japonicus reveal that nitrogen assimilation and fixation can occur in the absence of nodule - enhanced sucrose synthase [J]. Plant Physiology, 2007, 144(2): 806 - 820.
- [5] Xiao X, Tang C, Fang Y, et al. Structure and expression profile of the sucrose synthase gene family in the rubber tree; indicative of roles in stress response and sucrose utilization in the laticifers [J]. FEBS Journal, 2014, 281(1): 291 - 305.
- [6] Cho J I, Kim H B, Kim C Y, et al. Identification and characterization of the duplicate rice sucrose synthase genes OsSUS5 and OsSUS7 which are associated with the plasma membrane [J]. Molecules and

徐 雯,瞿印权,陈剑成,等. 绿竹 ISSR-PCR 反应体系的建立与优化[J]. 江苏农业科学,2017,45(15):34-38.
doi:10.15889/j.issn.1002-1302.2017.15.008

绿竹 ISSR-PCR 反应体系的建立与优化

徐 雯¹,瞿印权¹,陈剑成¹,何天友²,荣俊冬¹,陈礼光¹,郑郁善¹

(1. 福建农林大学林学院,福建福州 350002; 2. 福建农林大学艺术学院园林学院,福建福州 350002)

摘要:以绿竹基因组 DNA 为 ISSR-PCR 扩增模板,采用单因素试验方法,对 dNTPs、Mg²⁺、Taq DNA 聚合酶、引物、模板 DNA 设置 7 个不同浓度梯度进行研究,并对退火温度和循环次数进行筛选,建立绿竹 ISSR-PCR 最佳反应体系和扩增程序,并利用优化后的体系对 100 条 ISSR 引物进行筛选。最终确定的最佳反应体系为:20 μL 的扩增体系中,dNTPs 浓度为 0.2 mol/L,Mg²⁺ 浓度为 2.0 mmol/L,Taq DNA 聚合酶用量为 1.0 U,引物浓度为 0.4 μmol/L,DNA 用量为 50 ng,10 × PCR buffer 体积为 2 μL、剩余体积用灭菌 ddH₂O 补全。扩增程序为:94 ℃ 预变性 5 min;94 ℃ 变性 45 s,(根据引物的退火温度)复性 30 s,72 ℃ 延伸 90 s,循环 38 次;72 ℃ 延伸 10 min,4 ℃ 保存。以此体系为基础进行引物筛选,在 100 条 ISSR 引物中筛选出 14 条扩增条带清晰、多态性较高、重复性好的引物。本研究建立了绿竹 ISSR-PCR 最佳反应体系并筛选出高多态性引物,为绿竹的指纹图谱、遗传多样性分析、分子育种和品种鉴定等研究提供了基础。

关键词:绿竹;ISSR;反应体系;扩增程序;引物筛选

中图分类号:S795.501 **文献标志码:**A **文章编号:**1002-1302(2017)15-0034-05

绿竹 [*Dendrocalamopsis oldhami* (Munro) Keng f.] 是禾本科竹亚科 (Bambusoideae) 绿竹属 (*Dendrocalamopsis*) 的模式种,是绿竹属中分布最广、栽培最多的竹种^[1]。绿竹分布在中亚热带南部、南亚热带的北部和中部,主要分布于东南亚等国家^[2]。绿竹原产我国台湾省淡水,我国主要分布在浙江南

部、福建、台湾、广东、广西和海南等省区,是我国南方优良速生的笋材两用丛生竹种之一,具有较高的观赏价值、经济价值和社会价值^[3]。

分子标记技术经过三代的发展,已经出现了十几种分子标记技术,目前使用比较广泛且成熟的技术有 SSR、RAPD、ISSR、SRAP、RFLP、AFLP、SNP 等^[4-5]。ISSR 技术是 1994 年 Zietkiewicz 等在 SSR 分子标记技术的基础上发展起来的一种新的分子标记方法,它在 SSR 序列的 3' 或 5' 末端加上 1~4 个核苷酸作为反应的引物,从而扩大 DNA 序列^[6]。此种分子标记技术具有简单快捷易操作、DNA 用量较少、稳定性较好、试验成本低等优点,且扩增出的条带多态性较好,重复性较高^[7-8]。目前 ISSR 技术已经广泛应用于种质资源的鉴定、基

收稿日期:2017-01-10

基金项目:福建省科技重大专项(编号:2013NZ0001);福建省科技重大项目(编号:2010N5002)。

作者简介:徐 雯(1993—),女,江苏扬州人,硕士研究生,研究方向为森林培育理论与技术研究。E-mail:1243051350@qq.com。

通信作者:郑郁善,教授,博士生导师,主要从事森林培育与园林植物栽培研究。E-mail:zys1960@163.com。

Cells,2011,31(6):553-561.

[7] Wang Z, Wei P, Wu M, et al. Analysis of the sucrose synthase gene family in tobacco: structure, phylogeny, and expression patterns[J]. *Planta*, 2015, 242(1): 153-166.

[8] 陈卫军,魏益民,张国权,等. 国内外谷子的研究现状[J]. *杂粮作物*, 2000, 20(3): 27-29.

[9] Punta M, Coghill P C, Eberhardt R Y, et al. The Pfam protein families database[J]. *Nucleic Acids Research*, 2012, 40(1): 290-301.

[10] Schultz J, Milpetz F, Bork P, et al. SMART, a simple modular architecture research tool: identification of signaling domains[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 1998, 95(11): 5857-5864.

[11] Marchler-Bauer A, Anderson J B, Chitsaz F, et al. CDD: specific functional annotation with the Conserved Domain Database[J]. *Nucleic Acids Research*, 2009, 37: 205-210.

[12] Wang Y, Deng D, Shi Y, et al. Diversification, phylogeny and evolution of auxin response factor (ARF) family: insights gained from analyzing maize *ARF* genes[J]. *Molecular Biology Reports*, 2012, 39(3): 2401-2415.

[13] Sigrist C J, Cerutti L, de Castro E, et al. PROSITE, a protein domain database for functional characterization and annotation[J]. *Nucleic Acids Research*, 2010, 38: 161-166.

[14] Larkin M A, Blackshields G, Brown N P, et al. Clustal W and clustal X version 2.0[J]. *Bioinformatics*, 2007, 23(21): 2947-2948.

[15] Kiefer F, Arnold K, Künzli M, et al. The SWISS-MODEL repository and associated resources[J]. *Nucleic Acids Research*, 2009, 37: 387-392.

[16] Bailey T L, Boden M, Buske F A, et al. MEME SUITE: tools for motif discovery and searching[J]. *Nucleic Acids Research*, 2009, 37: 202-208.

[17] Cardini C E, Leloir L F, Chiriboga J. The biosynthesis of sucrose[J]. *Journal of Biological Chemistry*, 1955, 214(1): 149-155.

[18] An X, Chen Z, Wang J, et al. Identification and characterization of the *Populus* sucrose synthase gene family[J]. *Gene*, 2014, 539(1): 58-67.

[19] Zhang D Q, Xu B H, Yang X H, et al. The sucrose synthase gene family in *Populus*: structure, expression, and evolution[J]. *Tree Genetics & Genomes*, 2011, 7(3): 443-456.