

# 板栗颗粒型淀粉结合酶基因 cDNA 的克隆与生物信息学分析

王 猛, 梁 雪

(西南林业大学 林学院, 云南 昆明 650224)

**摘 要:**以板栗为试材,采用 RT-PCR 技术,克隆了颗粒性淀粉结合酶(GBSS)的基因,并进行了生物信息学分析。结果表明:从板栗果实中分离了 GBSS 基因 cDNA 全长,命名为 *Cgbss*,登录 GenBank 号为 KU162945。该基因全长 2 085 bp,编码区全长为 1 836 bp,编码 611 个氨基酸。生物信息学分析表明该基因编码的蛋白质等电点和分子量为 8.36 和 67.580 9 kDa,含有典型的 GT1 家族的糖基转移酶蛋白功能域。序列分析表明,该基因与所报道的植物 GBSS I 的基因有 85% 的同源性。

**关键词:**板栗;颗粒性淀粉结合酶;克隆;序列分析

**中图分类号:**S 664.203.6 **文献标识码:**A **文章编号:**1001-0009(2016)16-0099-05

板栗(*Castanea mollissima* Blume)属壳斗科(Fagaceae)栗属乔木经济植物,原产于我国,营养丰富,有“铁杆庄稼”之称,果实中富含蛋白质、淀粉,且口感有香糯等独特口味<sup>[1]</sup>。中国板栗品种约有 300 个,经过长期驯化,在不同地域的板栗果实性状和口感存在一定差异,甜香糯等口感等差异较大。其中糯性是主要差别之一,糯性是口腔感觉,是食物口感重要指标之一<sup>[1]</sup>。果实中直链淀粉和支链淀粉的含量,以及淀粉粒径的大小和分布,对果实糯性口感有重要的影响<sup>[2-5]</sup>。我国一些学者对板栗淀粉颗粒性质、流变性状、淀粉糊黏度及支链和直链淀粉含量进行过研究,并与马铃薯淀粉进行对比分析<sup>[6-8]</sup>。板栗果实富含淀粉,而控制淀粉生物合成的关键酶就是淀粉合成酶,它催化形成  $\alpha$ -1,4-糖苷键,将 ADP 葡萄糖连接在  $\alpha$ -1,4-葡聚糖的非还原端,延长  $\alpha$ -1,4-葡聚糖链<sup>[9]</sup>。淀粉合成酶由可溶性淀粉结合酶和颗粒性淀粉结合酶 2 种方式存在,在胚乳中合成直链淀粉的关键酶是颗粒性淀粉结合酶,即 GBSS。在对板栗花粉直感效应的研究上,赵志珩等<sup>[10]</sup>发现以“燕龙”和“燕山红栗”作为父本,云南板栗为母本授粉后,可以降低云南当地品种果实中直链淀粉含量。直链淀粉含量对板栗的口感和加工有重要影响,是影响糯性的重要因素,因此,现以板栗果实为原料,研究板栗果实中颗粒性淀粉合成酶的结构与功

能,以期改善板栗果实口感奠定基础。

## 1 材料与方法

### 1.1 试验材料

以“燕龙”板栗为父本、“云富”板栗为母本,授粉组合后采收板栗的果实为供试材料。

供试菌株与载体:感受态细胞 DH5 $\alpha$  购自 Tiangen 公司,克隆载体 pET-24a(+)购自 Promega 公司。

供试试剂:核酸分子质量标准、*Taq* DNA 聚合酶、T4 DNA 连接酶购自 TaKaRa 公司产品;RNAprep pure Plant Kit,RT-PCR 试剂盒购自 Tiangen 公司;DNA 凝胶纯化回收试剂盒购自北京鼎国生物技术公司。其它试剂均为国产分析纯。

### 1.2 试验方法

1.2.1 总 RNA 提取与 cDNA 合成 取板栗幼果加入液氮碾磨后,转入 1.5 mL 的离心管,具体步骤按照 RNAprep pure Plant Kit 说明书操作,得到的总 RNA 样品储存-76 °C 超低温冰箱中。cDNA 第 1 链的合成按照天根的反转录试剂盒说明书进行操作,以 Oligo(dt)<sub>15</sub> 为引物,在 M-MLV 反转录酶作用下合成,得到的 DNA 第 1 链于-20 °C 保存备用。

1.2.2 引物设计与合成 根据 GenBank 中玉米 GBSS 基因和马铃薯 GBSS 基因中 CDS 序列以及前期研究中获得的板栗 GBSS 的基因片段,设计 2 对引物:P1:5'-ACTCTTCTCTGCTCGTCTCTC-3', P2:5'-GACCCTGTCCGATTCTATT-3'; P3:5'-GCTATGAGAAGC-CAGTGAAA-3', P4:GAAGTGAACCTTGAATGAGG。引物序列由上海生工生物工程技术服务公司合成。

1.2.3 基因克隆与测序 以 cDNA 第 1 链为模板,以 P1、P2 以及 P3、P4 引物分别分段进行克隆,PCR 反应体

**第一作者简介:**王猛(1980-),男,山东枣庄人,博士,讲师,现主要从事经济林和植物分子生物学等研究工作。E-mail:mengwang80@sina.com。

**基金项目:**西南林业大学林学学位授权学科经费资助项目;云南省教育厅科学研究基金资助项目(2013y125);西南林业大学科研启动基金资助项目。

**收稿日期:**2016-04-26

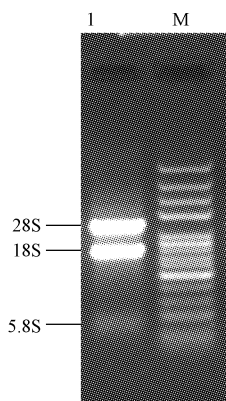
系为 25  $\mu$ L, 其中模板 cDNA 1  $\mu$ L, 10 $\times$  buffer 2.5  $\mu$ L, 引物各 1  $\mu$ L, *Taq* DNA 聚合酶 0.5  $\mu$ L, 用 ddH<sub>2</sub>O 补足到 25  $\mu$ L。具体扩增条件为 94  $^{\circ}$ C 预变性 2 min; 94  $^{\circ}$ C 40 s, 52  $^{\circ}$ C 退火 1 min, 72  $^{\circ}$ C 扩增 3 min, 30 个循环; 72  $^{\circ}$ C 延伸 10 min, 扩增完毕后 4  $^{\circ}$ C 终止反应。PCR 产物经 1% 琼脂糖凝胶电泳后回收、纯化, 并经分光光度计检验回收片段浓度。将纯化的 PCR 产物与 pET-24a(+) 载体连接, 重组子转化 DH5 $\alpha$  感受态大肠杆菌, 挑取阳性克隆于 37  $^{\circ}$ C 过夜培养, 并将菌液送上海生工公司测序。

1.2.4 生物信息学分析 用 DNAMAN 对获得的序列进行分析, 在 NCBI 数据库中进行核酸、氨基酸序列同源性比较。采用 NCBI (<http://www.ncbi.nlm.nih.gov/cdd/>) 保守功能区在线分析板栗 GBSSx 的保守序列; 用 Vetctor NTI 10.0 软件和 DNASTar 软件进行多重序列比对和进化树分析。推测蛋白的相关生物学信息在 <http://www.expasy.ch/tools/> 中进行预测分析; 使用 Signal P 3.0 Server 在线工具预测蛋白质信号肽。

## 2 结果与分析

### 2.1 总 RNA 的纯度鉴定

提取的板栗子叶总 RNA, 经电泳检测完整性, 测定 RNA 在 230、260、280 nm 的紫外吸收值, OD<sub>260</sub>/OD<sub>280</sub> 值为 1.96, OD<sub>260</sub>/OD<sub>230</sub> 值为 2.0。从图 1 可知, 总 RNA 电泳条带清晰, 没有降解, 完整性好, 纯度比较高, 可以用于后续研究。



注: 1. 总 RNA; M 100 bp DNA Ladder Marker。

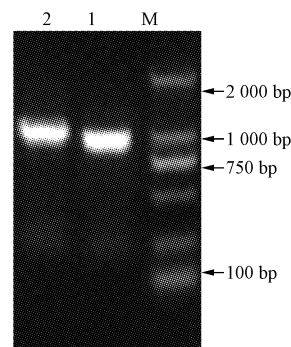
Note: 1. Total RNA; M 100 bp DNA Ladder Marker.

图 1 总 RNA 琼脂糖电泳凝胶

Fig. 1 Agarose gel electrophoresis of total RNA

### 2.2 GBSS 基因的扩增、克隆及鉴定

扩增产物经 1% 的琼脂糖凝胶电泳检测, 得到约 1 000 bp 的 DNA 片段(图 2)。测序表明, 插入片段分别为 1 104 bp 和 1 030 bp。用 DNAMAN 软件进行拼接后, 得到 2 085 bp 的片段, 其含有 1 个完整的开放阅读框, 其中起始密码子为 ATG, 终止密码子为 TGA, 长度为 1 836 bp。应用 Blast 在线软件, 把获得的基因序列与其它登录的基因序列进行同源性比较分析, 发现所获得片段与颗粒型淀粉结合酶 I 基因有较高的同源性, 达到 85%, 将其命名为 *Cgbss*, GenBank 序列号为 KU162945。



注: M. Marker DL 2 000; 1. P1、P2 引物对 PCR 扩增产物; 2. P3、P4 引物对 PCR 扩增产物。

Note: M. DL 2 000 Marker; 1. The PCR amplification of P1 and P2 primers; 2. The PCR amplification of P3 and P4 primers.

图 2 PCR 扩增琼脂糖凝胶电泳

Fig. 2 Agarose gel electrophoresis of PCR amplification

### 2.3 *Cgbss* 基因 cDNA 序列及蛋白序列分析

对 *Cgbss* 基因推导的氨基酸序列进行 ProtParam 程序预测分析, 该基因编码 611 个氨基酸, 相对分子量为 67.580 9 kDa, 理论等电点为 8.36, 亲水指数为 -0.161, 为亲水蛋白。负电荷残基 (Asp+Glu) 总数为 68 个, 正电荷残基 (Arg+Lys) 为 72 个。使用 MOTIFSCAN 对 *Cgbss* 基因序列的编码区分析, 结果显示, 序列的编码区含有 2 个酰胺化位点、8 个酪蛋白激酶 II 磷酸化位点、8 个 N 端豆蔻酰化、9 个蛋白激酶磷酸化位点、1 个酪氨酸激酶磷酸化位点、1 个糖原合成酶、1 个淀粉合成酶催化结构域、1 个糖基化转移位点。对获得的基因推导的氨基酸序列进行保守区分析发现其含有 GT1 家族的糖基转移酶结构域, 该家族含有糖原合成酶(或淀粉合成酶)(图 3)。用 SignalP 4.1server 程序推测 *Cgbss* 蛋白不含有信号肽。

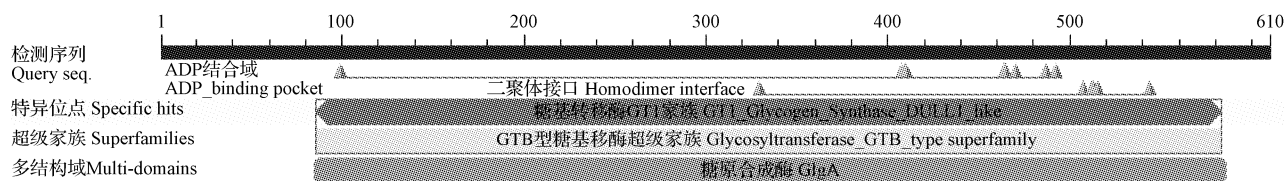


图 3 GBSS 蛋白保守区预测

Fig. 3 The prediction of GBSS protein conserved region



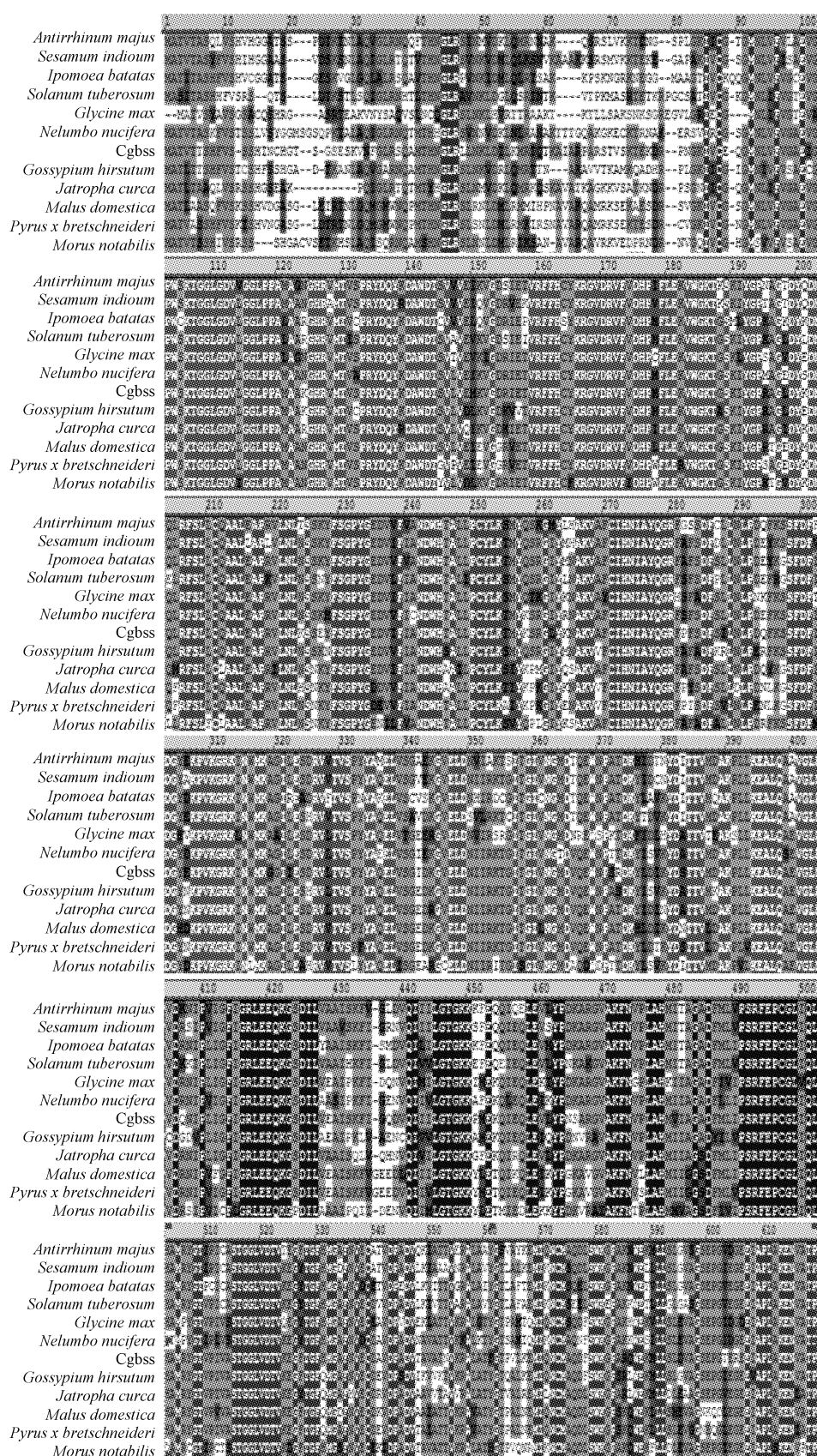


图4 不同植物GBSS蛋白氨基酸序列同源性比较

Fig. 4 Multiple alignment amino acid sequences of GBSS from different plant



## 2.4 不同物种 GBSS 蛋白序列分析同源性分析

在 GenBank 中对 GBSS 序列进行 Blast 比对,结果如图 4 所示。保守序列主要集中在中上游片段。根据 DNASTar MegAlign 软件,用 Cluster W 软件对芝麻 *Sesamum indicum*(XM\_011083059)、金鱼草 *Antirrhinum majus*(AJ006293)、马铃薯 *Solanum tuberosum*(XM\_006343701)、甘薯 *Ipomoea batatas*(AB524727)、大豆 *Glycine max*(XM\_003556383)、莲 *Nelumbo nucifera*(XM\_010253872)、棉花 *Gossypium hirsutum*(FJ713477)、麻疯树 *Jatropha curca*(XP\_012086631)、苹果 *Malus domestica*(NM\_001293995)、梨 *Pyrus bretschneideri*(XM\_009345784)、川桑 *Morus notabilis*(XM\_010101923) 相似性较高物种的 GBSS 蛋白序列进行同源性聚类分析,得到系统树。从图 5 可以看出,板栗 GBSS 蛋白与棉花亲缘关系最近,与川桑亲缘关系最远。

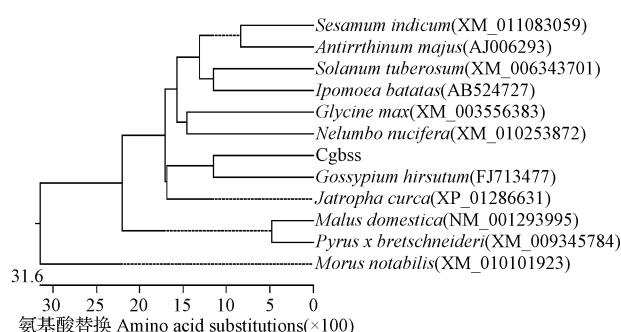


图 5 不同植物 GBSS 同源基因系统进化树分析

Fig. 5 Phylogenetic analysis of GBSS homologous genes

## 3 讨论

该研究在前期花粉直感效应的分子机理研究中,采用当地品种授粉组合为对照,以北方板栗(父本)授粉云南当地品种(母本),通过 cDNA-AFLP 技术对不同时序子叶的转录信息进行分析,筛选出差异常带进行测序与序列分析,然后根据获得序列设计引物,继续克隆研究,以期获得参与调控花粉直感效应相关基因的信息。在差异条带的回收以及克隆分析后发现颗粒性淀粉合成酶片段,由此可见该基因在花粉直感效应中起到一定作用。前期的研究发现,北方板栗花粉为父本,可以明显改变云南板栗外在和内在品质,例如增加果实的甜度及糯性,使得口感更佳<sup>[10]</sup>。在直链淀粉合成上,颗粒性淀粉合成酶有重要的作用。李良俊等<sup>[11]</sup>在研究莲藕直链

淀粉对莲藕品质的影响中发现,口感脆嫩爽口的莲藕直链淀粉含量高,莲藕品质较糯的品种其直链淀粉含量低。同时也有研究表明,GBSS 基因对淀粉含量多少及直链和支链的比例有密切关系<sup>[12]</sup>。直链淀粉和支链淀粉是相互影响的,GBSS 基因是影响果实中直链淀粉的重要因素,因此在板栗果实的糯性改良中,可以控制 GBSS 基因的表达,以此来改良板栗果实的糯性。多项研究表明,反义 GBSS 基因或者 RNAi 沉默 GBSSI 酶的基因,可以明显使得直链淀粉含量降低,从而改良培育出口感更好的新品种<sup>[13-15]</sup>。因此,可以根据获得的基因构建保守序列的 RNAi 的植物表达载体,可以提高板栗支链淀粉的含量,以培育出淀粉含量更高、口感较佳的板栗新品种。

## 参考文献

- [1] 郑金贵. 农产品品质学:1 卷[M]. 厦门:厦门大学出版社,2004:48,498.
- [2] 曹颖妮,胡卫国,王根平,等. 糯性和非糯性小麦灌浆期胚乳直/支链淀粉积累及其相关酶活性研究[J]. 西北植物学报,2010,30(10):1995-2001.
- [3] DONG-SOON S, AO Z H, JAY-LIN J, et al. 淀粉结构及应用[J]. 广西轻工业,2005(4):27-29.
- [4] 王广鹏,孔德军,刘庆香. 板栗支链淀粉含量与坚果糯性的定量关系研究[J]. 河北果树,2008(4):6-7.
- [5] 侯蕾,韩小贤,郑学玲,等. 不同直链淀粉含量玉米淀粉研究进展[J]. 粮食与油脂,2013,26(6):11-14.
- [6] 梁丽松,徐娟,王贵禧,等. 板栗淀粉糊化特性与淀粉粒粒径及直链淀粉含量的关系[J]. 中国农业科学,2009,42(1):251-260.
- [7] 梁建兰,刘秀凤,李志博,等. 板栗淀粉与板栗变性淀粉性质的比较[J]. 中国粮油学报,2011,26(1):65-68.
- [8] 张笃芹,木泰华,孙红男. 马铃薯蛋白及马铃薯蛋白-马铃薯淀粉复合物乳化、热及流变学特性的研究[J]. 食品科技,2015(4):223-231.
- [9] 沈宝云,刘玉汇,张俊莲,等. 马铃薯块茎 GBSSI 基因的 cDNA 克隆及其序列特征分析[J]. 草业学报,2010,19(5):1-8.
- [10] 赵志珩,杨柳,石卓功,等. 板栗花粉直感效应在坚果内在品质上的表现[J]. 经济林研究,2012,30(4):9-12.
- [11] 李良俊,林惠鸣,曹曙生. 江苏省莲藕无公害生成中存在的几个问题[J]. 中国农学通讯,2003,19(4):156-158.
- [12] 唐秀丽,周玉琼,郑丽莉,等. 复合调控 Wx 与 SBE3 基因对水稻籽粒直链淀粉含量的影响[J]. 激光生物学报,2012,21(1):16-21.
- [13] 李洲. 以 GBSS、SSIII、SSII 基因为靶标的 RNAi 载体对马铃薯的转化及鉴定[D]. 兰州:甘肃农业大学,2009.
- [14] 李加瑞,赵伟,李全梓. Waxy 基因的 RNA 沉默使转基因小麦种子中直链淀粉含量下降[J]. 遗传学报,2005(8):846-854.
- [15] 康国章,肖向红,王永华,等. 小麦淀粉 GBSSII 基因的克隆、反义和 RNAi 载体的构建[J]. 华北农学报,2008,23(1):7-11

## Cloning and Bioinformatics Analysis of the GBSS Gene From Chinese Chestnut

WANG Meng, LIANG Xue

(College of Forestry, Southwest Forestry University, Kunming, Yunnan 650224)

# 基于叶绿体 *ndhA* 基因内含子序列的 DNA 条形码在秋海棠属物种鉴定中的应用

赵 博<sup>1</sup>, 李景剑<sup>2</sup>, 毛世忠<sup>1</sup>, 唐文秀<sup>1</sup>

(1. 中国科学院 广西植物研究所, 广西 桂林 541006; 2. 华南农业大学 林学与风景园林学院, 广东 广州 510642)

**摘 要:**以 18 种秋海棠属植物为试材, 基于叶绿体 *ndhA* 基因内含子片段, 利用 Codon Code Aligner 软件拼接序列后, 通过 MEGA 软件进行比对并计算 19 个秋海棠属植物种间及种内遗传距离, 最后利用邻接法构建分子系统树, 研究了该片段作为 DNA 条形码对秋海棠属植物进行物种鉴定的可行性。结果表明: 秋海棠属 19 个种类共 59 个个体的 *ndhA* 基因内含子序列长度为 1 182 bp, 在所考察的候选秋海棠属植物中具有最大的种间变异和较小的种内变异, 且二者存在极显著差异。在系统树中, 秋海棠属植物每一物种能够分别形成各自独立的分支, 表现出良好单系性。基于 *ndhA* 基因内含子的 DNA 条形码在识别秋海棠属植物物种方面和传统形态学基本一致。该研究表明, 以 *ndhA* 基因内含子作为秋海棠属植物 DNA 条形码进行物种鉴定具有一定的可行性。

**关键词:**秋海棠属; *ndhA* 基因内含子; DNA 条形码; 物种鉴定

**中图分类号:**Q 948 **文献标识码:**A **文章编号:**1001-0009(2016)16-0103-05

秋海棠属(*Begonia* L.)为秋海棠科(Begoniaceae)一年或多年生草本、灌木或小树枝状植物, 全世界约 1 500~1 600 种<sup>[1]</sup>。在中国约有 173 种, 分布于长江以南的地区, 常见于云南、广西和贵州<sup>[2-3]</sup>。由于秋海棠属植物在习性、生境、性别、形态等方面均表现出很高的多样性, 品种繁多, 其中大部分具有极高的观赏价值, 在国内外花卉市场颇受人们的喜爱<sup>[4]</sup>。另外, 秋海棠属植物也是我国少数民族中常见的民族药, 具有清热解毒、散瘀消肿的功效<sup>[5-6]</sup>。由此可见, 秋海棠属植物中蕴藏着丰富的观赏、药用资源有待进一步的开发利用。然而,

秋海棠属植物性状受环境影响较大, 即使同一种类的不同居群, 在某些性状上的变异也比较显著, 很多近缘种类难以鉴定<sup>[7]</sup>。同时该属大多是多汁草本植物, 具有肉质的茎、娇嫩的花和果实, 传统方法压制成标本后, 在叶片颜色、花纹、毛被颜色和花被片颜色等方面都有很大程度的失真, 果实和子房的形态解剖学特征等是属下分组不可缺少的重要鉴定特征, 在多汁的果实被压干以后容易变形, 给传统的分类鉴定带来较大的困难<sup>[8-9]</sup>。因此, 为秋海棠属植物寻找快速、高效的分类鉴定方法一直都是研究的热点。

DNA 条形码(DNA barcoding)是通过使用短的标准 DNA 片段, 对物种进行快速、准确的识别和鉴定<sup>[10-13]</sup>。DNA 条形码不仅是传统植物分类与鉴定的强有力补充, 而且能突破对经验的过度依赖, 可以在较短时间内建成易于利用的应用系统, 在物种分类和鉴定方面展示出了强大的生命力<sup>[14]</sup>。植物 DNA 条形码片段主要在叶

**第一作者简介:**赵博(1981-), 女, 河南南阳人, 博士, 副研究员, 研究方向为植物细胞学与分子系统发育。E-mail: calljone@163.com.

**责任作者:**唐文秀(1973-), 女, 本科, 副研究员, 现主要从事珍稀濒危植物保护等研究工作。E-mail: tangwx99@163.com.

**基金项目:**广西自然科学基金资助项目(2012GXNSFBA053075)。

**收稿日期:**2016-04-26

**Abstract:** Taking chestnut as test material, GBSS gene was cloned from the first strand of *Castanea mollissima* seed's cDNA through RT-PCR methods, and the bioinformatics characters were analysed. The results showed that a full length cDNA of *Cgbss* (GenBank accession number: KU162945) was obtained. The length of *Cgbss* gene was 2 085 bp, which contained an open reading frame encoding 611 amino acids. The bioinformatics analysis indicated that the pI and MW of amino acids encoded by *Cgbss* were predicted to be 8.36 and 67.580 9 kDa, respectively, and the protein had one GT1\_glycogen\_synthase domain. And the fragment had 85% homology identical to the already reported sequence of GBSS I.

**Keywords:** *Castanea mollissima*; granule-bound starch synthase; clone; sequence analysis