

北柴胡糖基转移酶基因 BcUGT8 的克隆、序列分析及其原核表达载体构建

徐洁森^{1,2}, 魏建和^{1,2}, 陶韵文^{1,3}, 孙 晶^{1,4}, 隋 春^{1,2*}

1. 中国医学科学院 北京协和医学院药用植物研究所, 北京 100193

2. 中草药物质基础与资源利用教育部重点实验室, 北京 100193

3. 佳木斯大学药学院, 黑龙江 佳木斯 154007

4. 东北林业大学生命科学院, 黑龙江 哈尔滨 150040

摘要: 目的 克隆北柴胡 *Bupleurum chinense* 中可能参与柴胡皂苷生物合成的糖基转移酶基因 BcUGT8, 构建其原核表达载体, 为通过体外表达和纯化蛋白的催化活性, 分析验证其功能奠定基础。方法 在 Roche (454) GS FLX 系统高通量测序已获得部分 cDNA 序列基础上, 通过 cDNA 末端快速扩增技术 (RACE) 和长距离 PCR 扩增方法 (LD-PCR) 克隆全长 cDNA。设计含有酶切位点的 PCR 引物, PCR 扩增其开放阅读框 (ORF), 酶切后, 插入到 pET-28a (+) 载体中, 构建出原核表达载体。结果 扩增到北柴胡 1 个糖基转移酶 (UGT) 基因全长 cDNA, 构建了这一基因的原核表达载体。结论 通过 BcUGT8 基因的全长 cDNA 克隆、序列分析和原核表达载体的构建, 为后续开展体外酶蛋白表达、纯化和催化活性分析实验, 验证其生物功能奠定了基础。

关键词: 北柴胡; 糖基转移酶基因; 序列分析; 原核表达载体; PCR

中图分类号: R282.12 文献标志码: A 文章编号: 0253-2670(2013)17-2453-07

DOI: 10.7501/j.issn.0253-2670.2013.17.021

Cloning, sequence analysis, and prokaryotic expression vector construction of glycosyltransferase gene BcUGT8 from *Bupleurum chinense*

XU Jie-sen^{1,2}, WEI Jian-he^{1,2}, TAO Yun-wen^{1,3}, SUN Jing^{1,4}, SUI Chun^{1,2}

1. Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100193, China

2. Key Laboratory of Bioactive Substances and Resources Utilization of Chinese Herbal Medicine, Ministry of Education, Beijing 100193, China

3. College of Pharmacy, Jiamusi University, Jiamusi 154007, China

4. College of Life Science, Northeast Forestry University, Harbin 150040, China

Abstract: Objective To clone the full-length cDNA of the uridine diphosphate glycosyltransferase (UGT) gene in *Bupleurum chinense* (BcUGT8), which may be involved with the saikosaponin biosynthesis, and to construct the prokaryotic expression vector. The work will provide the foundation for its further function verification by *in vitro* expression and activity analysis of the purified protein. **Methods** RACE and LD-PCR were used to clone the full-length cDNA of BcUGT8, on the basis of its partial cDNA sequence obtained from our previous high-flux sequencing by Roche (454) GS FLX system. The open reading form (ORF) was PCR cloned using primers with corresponding restriction enzymes cutting sites. The PCR products were digested with corresponding restriction enzymes and then were inserted in expression vector pET-28a (+) to construct the recombinant expression vectors. **Results** The full-length cDNA of UGT gene was cloned from *B. chinense*, and the prokaryotic expression vector was obtained. **Conclusion** The full-length cDNA cloning, sequence analysis, and prokaryotic expression vector construction provide a substantial foundation for follow-up bio-function analysis of BcUGT8 through protein expression, purification, and activity analysis *in vitro*.

Key words: *Bupleurum chinense* DC.; uridine diphosphate glycosyltransferase gene; sequence analysis; prokaryotic expression vectors; PCR

北柴胡 *Bupleurum chinense* DC. 为伞形科柴胡属植物, 具有解表退热、疏肝解郁、升举阳气的功效, 其主要药效成分柴胡皂苷为齐墩果烷类型的三萜皂苷。三萜皂苷生物合成途径大致可分为 3 个阶

收稿日期: 2013-04-30

基金项目: 国家自然科学基金资助项目 (81072994); 北京市自然科学基金资助项目 (5102033); 中医药行业科研专项 (201107011)

作者简介: 徐洁森, 女, 硕士研究生, 研究方向为药用植物次生代谢途径及调控。

*通信作者 隋 春 E-mail: csui@implad.ac.cn

段,包括前体的形成(异戊烯焦磷酸 IPP 及其异构体二甲烯丙基焦磷酸 DMPP 等的形成)、骨架构建即三萜骨架的形成和后修饰过程(包括氧化及糖基化等)。其中糖基化是糖基转移酶(uridine diphosphate glycosyltransferase, UGT)催化糖基从活化的分子向特定的受体分子转移,形成糖苷键的过程。植物体的 UGT 为一大类,催化单糖苷、二糖苷、聚糖苷及其他非碳水化合物(蛋白、脂类、甾族化合物及其他一些小分子化合物)的糖苷形成。在皂苷合成途径下游,UGT 催化皂苷苷元糖基化形成单体皂苷。研究表明皂苷苷元的糖基化对苷元的生物活性、溶解性及在植物细胞内或整个植物体中的转运起调节作用。糖基侧链的结构对皂苷的多种生物活性也起至关重要的作用^[1-2]。虽然关于植物 UGT 的研究报道较多,如参与三萜皂苷生物合成的 UGT 基因可见报道的有蒺藜苜蓿 *Medicago truncatula* L. 中的 UGT71G1、UGT71K1 和 UGT73F3^[3-4],王不留行 *Saponaria vaccaria* L. 中的 UGT74M1^[5],大豆 *Glycine max* (L.) Merr. 中的 UGT73P2、UGT91H4、UGT73F4 和 UGT73F2^[6-7],且 UGT71G1 的晶体结构已解析^[8]。植物体内存在众多 UGT 基因和多种类单体皂苷,如拟南芥 *Arabidopsis thaliana* (L.) Heynh. 中有 112 个 UGT 基因和 8 个 UGT 假基因^[9],人参 *Panax ginseng* C. A. Mey. 和柴胡等可见报道的单体皂苷都已过百种^[10-11]。此外,UGT 参与多种植物次生代谢途径。因此,仍需大量的研究以最终探明 UGT 在众多单体皂苷的有机合成中是如何表达和调控的。本课题组以柴胡为研究材料,开展柴胡皂苷合成相关 UGT 基因的克隆和功能研究。利用转录组测序^[12-13],得到了 102 个 UGT 基因片段,对部分基因片段进行了茉莉酸甲酯(MeJA)诱导表达及组织表达特性分析。初步确定 3 个 UGT 可能参与柴胡皂苷生物合成^[13]。本实验利用 cDNA 末端快速扩增技术

(rapid-amplification of cDNA ends, RACE) 和长距离 PCR 扩增方法(long distance PCR, LD-PCR)克隆到了这一 UGT 全长 cDNA,命名为 BcUGT8,对 BcUGT8 进行了序列分析,构建了原核表达载体,为下一步开展该基因的功能研究奠定基础。

1 材料

植物材料为本实验室选育的北柴胡 *B. chinense* DC. 新品种“中柴 1 号”。经中国医学科学院药用植物研究所魏建和研究员鉴定。取盛花期“中柴 1 号”的新鲜根,用于总 RNA 提取。

2 方法

2.1 RNA 提取

利用 RNA 纯化试剂盒(Norgen, Canada)提取 RNA。1%琼脂糖凝胶电泳检测 RNA 完整性。

2.2 3'-RACE 和 5'-RACE PCR 扩增

根据 Roche (454) GS FLX 系统测序结果分别设计 3'-RACE 和 5'-RACE 的特异性引物(表 1),取约 1 μg RNA 利用 SMARTTM RACE PCR cDNA 合成试剂盒(Clontech, CA, USA)进行反转录,利用 Advantage[®] 2 PCR 试剂盒(Clontech, CA, USA)扩增。3'-RACE 反转录过程为 RNA 3.75 μL, 3'-RACE CDS primer A 1.00 μL, 72 °C, 3 min, 42 °C, 2 min, 短暂离心后,依次加入: 5×First-Strand 缓冲液 2.00 μL, DTT (20 mmol/L) 1.00 μL, dNTP Mix (10 mmol/L) 1.00 μL, RNase Inhibitor 0.25 μL, SmartScribe Reverse 1.00 μL, 42 °C, 1.5 h, 70 °C, 10 min, 再加入 20.00 μL Tricine-EDTA 缓冲液,离心混匀,-20 °C 保存备用。5'-RACE 反转录过程为 RNA 2.75 μL, 5'-RACE CDS primer A 1.00 μL, SMART IITM A Oligonucleotide 1.00 μL, 72 °C, 3 min, 42 °C, 2 min, 短暂离心后,依次加入: 5×First-Strand Buffer 2.00 μL, DTT (20 mmol/L) 1.00 μL, dNTP Mix (10 mmol/L) 1.00 μL, RNase Inhibitor 0.25 μL, SmartScribe Reverse 1.00 μL, 42 °C, 1.5 h, 70 °C,

表 1 引物序列

Table 1 Primer sequences

扩增目的	引物名称	引物序列 (5'→3')
3'-RACE	BcUGT8-3F	GCTTTCTCAGTCATTGTGGATGGGGTTC
5'-RACE	BcUGT8-5R	CAATCTCCTCCCACTTGAACCTGCCTTC
LD-PCR	BcUGT8-LD3	AAATGTATGCAATGACACGAGGT
	BcUGT8-LD5	TGGGATGGTGATTACCAAAATG
原核表达载体构建	BcUGT8-3	GCAAGCTTACATGTTTCTTTCTTCATCCTGC
	BcUGT8-5	GCGGATCCATGGAGAGTAAAAATGGAAGAATG

10 min, 再加入 20.00 μL Tricine-EDTA 缓冲液, 离心混匀, $-20\text{ }^{\circ}\text{C}$ 保存备用。3'-RACE 和 5'-RACE PCR 扩增体系均为超纯水 17.25 μL , 10 $\times$ Advantage 2 PCR 缓冲液 2.50 μL , dNTP Mix (50 mmol/L) 0.50 μL , 50 $\times$ Advantage 2 Polymerase Mix 0.50 μL , 混匀后加入反转录产物 1.25 μL , 10 $\times$ Universal Primer A Mix 2.50 μL , 特异性引物 (10 mmol/L) 0.50 μL 。PCR 扩增程序为 94 $^{\circ}\text{C}$ 预变性 5 min; 然后 94 $^{\circ}\text{C}$ 30 s, 72 $^{\circ}\text{C}$ 3 min, 5 个循环; 94 $^{\circ}\text{C}$ 30 s, 70 $^{\circ}\text{C}$ 30 s, 72 $^{\circ}\text{C}$ 3 min, 5 个循环; 94 $^{\circ}\text{C}$ 30 s, 68 $^{\circ}\text{C}$ 30 s, 72 $^{\circ}\text{C}$ 3 min, 25 个循环; 最后 72 $^{\circ}\text{C}$ 延伸 10 min。扩增产物用胶回收试剂盒[天根生化科技(北京)有限公司]回收, 连接到 pEASY-Blunt Zero 载体(北京全式金生物技术有限公司)上, 由北京金唯智生物科技有限公司完成测序。

2.3 LD-PCR 扩增

根据 3'-RACE 和 5'-RACE 结果设计 LD-PCR 两端引物(表 1), 利用保真性高的 2 $\times$ PCR Solution Prime STARTM HS Premix [宝生物工程(大连)有限公司] 进行 PCR 扩增。体系为 2 $\times$ Premix 25.00 μL , 正反向引物各 1.00 μL , 超纯水 20.50 μL , 模板为 5'-RACE 反转录产物 2.50 μL 。程序为 94 $^{\circ}\text{C}$ 预变性 5 min; 然后 98 $^{\circ}\text{C}$ 10 s, 58 $^{\circ}\text{C}$ 15 s, 72 $^{\circ}\text{C}$ 2 min, 35 个循环; 最后 72 $^{\circ}\text{C}$ 延伸 10 min。扩增产物电泳、回收、连接 T 载体 pEASY-Blunt Zero 上, 菌液 PCR 验证后选择阳性克隆送测序。

2.4 BcUGT8 基因的序列分析

BcUGT8 基因编码蛋白的理化性质预测采用 ExPASy Proteomics Server 提供的在线工具 ProtParam (http://www.expasy.ch/tools/prot_param.html) 进行分析。采用 SWISS-MODEL (<http://swissmodel.expasy.org/>) 进行蛋白质二级结构分析和同源建模, 获得三级结构。利用在线工具 (<http://www.ebi.ac.uk/Tools/pfa/ipscan/>) 对 BcUGT8 蛋白质结构域进行分析。使用 SignalP 3.0 Server (<http://www.cbs.dtu.dk/services/SignalP/>) 和 TargetP 1.1 Server (<http://www.cbs.dtu.dk/services/TargetP/>) 进行分泌蛋白预测。利用 WOLF PSORT (<http://wolfsort.org/>) 和 CELLO (<http://cello.life.nctu.edu.tw/>) 进行蛋白的定位信号预测。利用在线软件 TMHMMServer v.2.0 (<http://www.cbs.dtu.dk/services/TMHMM/>) 预测 PnUGT1 的跨膜区。氨基酸序列比对利用美国国立生物技术信息中心

(NCBI) 的蛋白质序列数据库进行。

2.5 原核表达载体构建

根据 LD-PCR 所得 BcUGT39 的序列在翻译起始密码子和终止密码子处分别设计含有 *Bam*H I 和 *Hind* III 位点的 PCR 引物(表 1), 利用保真性高的 2 $\times$ PCR Solution PrimeSTARTM HS Premix[宝生物工程(大连)有限公司]进行 PCR 扩增。体系为 2 $\times$ Premix 25.00 μL , 正反向引物各 1.00 μL , 超纯水 20.50 μL , 模板为 5'-RACE 反转录产物 2.50 μL 。程序为 94 $^{\circ}\text{C}$ 预变性 5 min; 然后 98 $^{\circ}\text{C}$ 10 s, 56 $^{\circ}\text{C}$ 15 s, 72 $^{\circ}\text{C}$ 2 min, 5 个循环; 98 $^{\circ}\text{C}$ 10 s, 60 $^{\circ}\text{C}$ 15 s, 72 $^{\circ}\text{C}$ 2 min, 25 个循环; 最后 72 $^{\circ}\text{C}$ 延伸 10 min。PCR 产物电泳、回收、连接 T 载体 pEASY-Blunt Zero。菌液 PCR 验证后选择阳性克隆, *Bam*H I 和 *Hind* III 酶切, 酶切产物与同样内切酶酶切的载体 pET-28a (+) 连接, 转化 DH5 α 感受态细胞。菌液 PCR 鉴定的阳性克隆送测序。

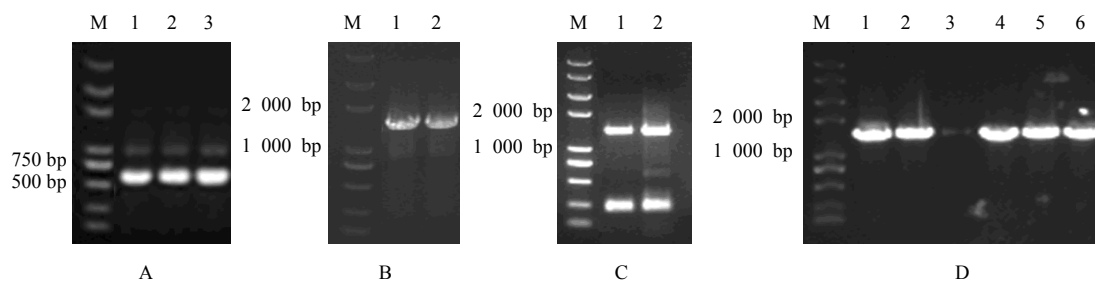
3 结果与分析

3.1 BcUGT8 全长 cDNA 的克隆

3'-RACE PCR 扩增产物琼脂糖凝胶电泳结果如图 1-A 所示, 得到明亮单一条带, 大小约为 600 bp。5'-RACE 扩增产物琼脂糖凝胶电泳结果如图 1-B 所示, 可得明亮单一条带, 大小约为 1500 bp。根据 3'-RACE 和 5'-RACE 所得序列设计 LD-PCR 引物, PCR 扩增可得明亮条带, 如图 1-C 所示, 大小约为 1500 bp。扩增片段连接 T-载体后菌液 PCR 验证结果如图 1-D 所示, 在 1500 bp 左右可见明亮单一条带。测序结果表明克隆得到 BcUGT8 全长 cDNA, 序列长度为 1552 bp, 提交到 GenBank, 登录号为 KC464461。

3.2 BcUGT8 序列分析

3.2.1 氨基酸序列同源性分析 DNAMAN 分析表明 BcUGT8 的开放阅读框(ORF)长 1335 bp, 编码 444 个氨基酸, 相对分子质量 5.01×10^4 (图 2)。NCBI blastp 在线搜索显示, 其氨基酸序列与本实验室之前克隆到的北柴胡 UGT5 (AFK79037.1) 一致性最高, 为 63%。与三七中推测参与皂苷形成的 2 个 UGT (AED99884 和 AF063526) 一致性分别为 46% 和 31%。与催化大豆皂苷合成的 UGT91F4 (NP_001240857) 一致性为 31%。另外与其他植物中催化黄酮类化合物生物合成的 UGT 氨基酸序列一致性也在 40% 左右。



A-3'-RACE PCR 扩增; B-5'-RACE PCR 扩增; C-LD-PCR 扩增; D-LD-PCR 扩增片段连接 T-载体菌液 PCR 验证; M-Marker

1~6-扩增产物 (除 D3 为阴性克隆外, 其他均为阳性克隆)

A-3'-RACE PCR amplification; B-5'-RACE PCR amplification; C-LD-PCR amplification; D-PCR validation for ligation of LD-PCR product to T-vector; M-Marker; 1—6-PCR amplification of bacterium (D3-negative clone, others-positive clone)

图1 BcUGT8 全长 cDNA 的克隆

Fig. 1 Full-length cDNA cloning of BcUGT8

```

1 ATGGAGAGTAAATTGGAAGAAAGTACTATAGTAATGGTACCATATTTAGCTCAGGGTCACATCTCTCCCTAC
1 M E S K I G R M S T I V M V P Y L A Q G H I S P Y
76 TTAGAAGTACTGCTAACTCACTCAAAAAAGAGTTTCAATATATATATTTGTTTACACCTGTCAAACCTTGCTTCC
26 L E L A K Q L T K R S F N I Y I C S T P V N L A S
151 ATCAAGAACAGAGTGCCTTCAAAATGATAATATACAACITGTAGAGCTTCATCTTCAATCTTCTCCAGATTGCGCA
51 I K N R V L Q N D N I Q L V E L H L Q S S P D L P
226 CCTCAATATCATTCACAAATGGCCTTCCATCTCATCTCATGCCAGTCCCTCGAGATGCGCTTGGAGAAAGCAGCT
76 P Q Y H S T N G L P S H L M P V L R D A L E K A A
301 CCAAACCTTTGTTAGCATTTCAAAAGACATCAACCCCTAATTTGGTTCATCTATGATTTTCATGCCCTCATGGCCTGCA
101 P N F V S I L K D I N P N L V I Y D F M P S W P A
376 GAGGTGCTATGTCCCTTAACATTCGGCTGTTTATTTTACGGTCAACGCAGCAGCAACCTCTTGCAAGGCTTA
126 E V A M S L N I P A V Y F T V N A A A T S C I G L
451 CATCCTTACAAAAGGGCAGGTGAAAAATTCCTATTTCCCGAAAATTTTGTCTTCTGTGACCAACCTCCAGTT
151 H P Y K R A G E K F P F P E I F V P S V D Q P P V
526 TCTGCAGATGTACTCAGGATTCCTGCAAACTTTTGTATGCTTGAACGATCATGCAATTTTGTGTTTGTGAAAA
176 S A D V L R I L R N F L L C F E R S C N F V L V K
601 AGTTGTAGAGAGTTGAAAGGAAATATATTGATCACCTTTCCGATCTGGCAGAGAAAGATATGATTCCTACCGGT
201 S C R E V E G K Y I D H L S D L A E K N M I P T G
676 CCACCTTTCATGATCCTACCGAAAATGAAGATGACAACATGAAGGATATCATGAAATGGCTCGACAAAGAGAAA
226 P L V H D P T E N E D D N M K D I M K W L D K K K
751 AAGTATTCGGTGGTGTGTTGTTGCTTGTAGTGAAGATTATTTATCTGCAGAGAGTATAGAGATGGCAAAAT
251 K Y S V V F V C F G S E N Y L S A E E V I E M A N
826 GCACTCGAAACAAAGTGTAATTTTCATATGGCTTTGAGATCAATCAAGGGGAGGAGAGGATAGTGCATTG
276 A L E T T K C N F I W A L R S I Q G E E E G S A L
901 CTACCAGATGTTTGTGCAAGAGTTGGAGATTGGGATTGATTTTGTCTGGGCGCCACAAACATGATTTTG
301 L P D G F V E R V G D L G L I L S W A P Q T M I L
976 CGACATCCAAGCACCGGTGCTTCTCAGTCATTGTGGATGGGGTTCTATGATCGAAGCATGAAATACGGGGTG
326 R H P S T G G F L S H C G W G S M I E S M K Y G V
1051 CCAATAATTGCCATGCCAATGAAGTGATCAGCCTATGAATGCTAGACTTGCAGAGAGATTGGGGTTAGCATG
351 P I I A M P M K V D Q P M N A R L A T E I G V S M
1126 GAGATAGTGAGAGACAATGAAGCAAGTTCAAGTGGGAGGAGATTGTCGCTAATAAGAATGGTGTGGTGGAA
376 E I V R D N E G K F K W E E I V R V I R M V L V E
1201 GAAAGTGGAGAGGTGTGAGGAGGAAAGCTAGAGAAATTGTTTGAATATGAAGGAGAGAGGAGGAGGAAGATAT
401 E S G E G V R R K A R E L C L N M K E R G E E E Y
1276 TTGACAAGGCAGCGCAGAGTTGGAGAAAATTTGCAGGATGAAGAAAGAAACATGTTAA
426 L D K A A A E L E K I C R M K K E T C *

```

图2 BcUGT8 的 cDNA 序列及其预测的编码蛋白质的氨基酸序列

Fig. 2 cDNA sequences of BcUGT8 and predicted amino acid sequences coding protein

3.2.2 理化性质分析 利用 ExPASy ProteomicsServer 在线软件 ProtParam 对 BcUGT8 基因编码蛋白质的理化性质进行预测, 结果显示 BcUGT8 蛋白的分子式为 $C_{2235}H_{3547}N_{591}O_{650}S_{31}$; 等电点 (pI) 为 5.64, 带负电残基 (Asp+Glu) 为 58, 带正电残基 (Arg+Lys)

为 49。该蛋白的不稳定系数 (instability index) 为 53.10, 属于不稳定蛋白; 脂肪系数 (aliphatic index) 为 92.64, 亲水性系数 (grand average of hydropathicity, GRAVY) 为 -0.124。

3.2.3 蛋白二级结构及三维建模 在线预测结果表

在线工具 WOLF PSORT 预测该蛋白的亚细胞定位情况: 叶绿体的定位系数为 9.0 (chlo: 9.0)、细胞质的定位系数为 3.0 (cyto: 3.0)、线粒体的定位系数为 1.0 (mito: 1.0)。同样预测了目前已证实参与三萜皂苷生物合成的 UGT 定位情况: UGT71G1 为 chlo: 4.0, nucl: 3.0, extr: 3.0, cyto: 2.0, mito: 1.0; UGT73K1 为 cyto: 8.0, chlo: 4.0, nucl: 1.0; UGT74M1 为 chlo: 8.0, cyto: 5.0; UGT73F3 为 nucl: 6.0, cyto: 6.0, chlo: 1.0; UGT73P2 为 cyto: 9.5, cyto_E.R.: 5.5, chlo: 2.0, extr: 1.0; UGT91H4 为 cyto: 9.5, cyto_E.R.: 5.5, chlo: 2.0, nucl: 1.0; UGT73F4 为 chlo: 11.0, mito: 2.0; UGT73F2 为 chlo: 10.0, mito: 3.0。由此, BcUGT8 蛋白预测的亚细胞定位与 UGT71G1、UGT74M1、UGT73F4 和 UGT73F2 类似。利用 CELLO 预测, BcUGT8 蛋白位于细胞质中, 预测的

可靠性指数 (reliability index, RI) 为 2.379, 预测精度 (expected accuracy) 为 85.1%。因此, BcUGT8 蛋白最可能定位于细胞质及叶绿体中。

3.3 BcUGT8 原核表达载体构建

本实验用 pET 系列载体 pET-28a (+) 构建了 BcUGT39 的原核表达载体。如图 6 所示, 以反转录产物为模板, 采用高保真 DNA 聚合酶 PCR 得到了目标条带 (图 6-A)。目标条带经切胶回收后, 首先与 T-载体连接, 经菌液 PCR 验证以选择阳性克隆 (图 6-B)。测序后选择其中的阳性克隆进行酶切 (图 6-C)。将酶切下来的目标条带, 连接同样酶切的 pET-28a (+)。菌液 PCR 验证如图 6-D, D2、D4、D6 为阴性, 其他均为阳性。选择阳性克隆送测序。测序结果表明, 获得了 BcUGT8 的原核表达载体。

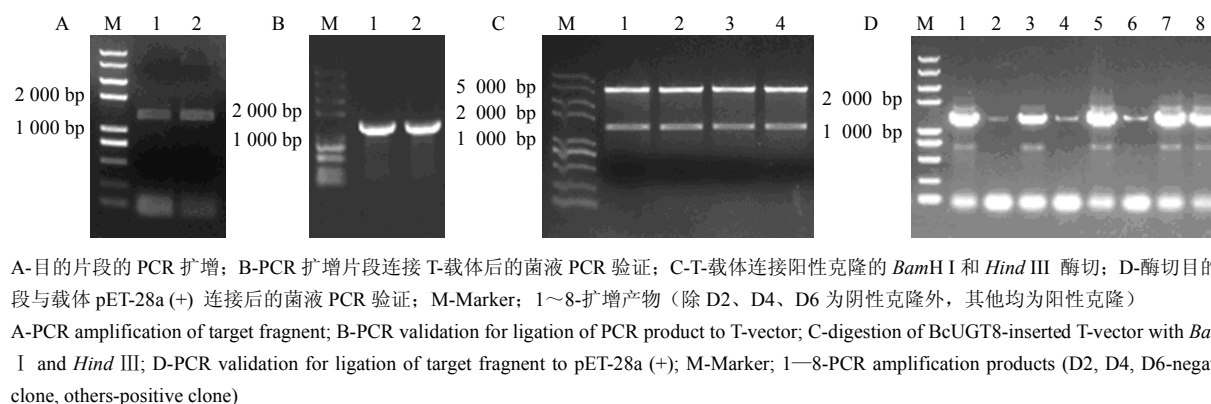


图 6 BcUGT8 原核表达载体构建

Fig. 6 Construction of prokaryotic expression vectors for BcUGT8

4 讨论

糖基转移酶普遍存在于生物体中, 催化多种植物次生代谢产物如黄酮类、皂苷、甾体生物碱等的生物合成。这些植物次生代谢产物常作为食品添加剂、药物和化工原料等, 广泛用于人们生活和生产^[14]。因此解析 UGT 基因功能和调控机制有重要理论价值和应用前景。与植物体内单体皂苷种类和存在的众多 UGT 基因相比, 目前已得到功能验证的 UGT 基因还是少数, 还有很大的空间有待探索, 尤其是植物体在不同环境条件下单体皂苷在种类和含量上的调节, 以及调控机制。

基因克隆是研究基因功能的基础, 已有报道利用转录组测序和表达特性分析筛选参与重要次生代谢产物生物合成相关基因, 如人参中 6 个 UGT 基因^[15], 罗汉果 *Siraitia grosvenorii* (Swingle) C.

Jeffrey ex A. M. Lu & Zhi Y. Zhang 中 5 个 UGT 基因^[16], 以及三七 *Panax notoginseng* F. H. Chen 中 PnUGT1 基因^[17], 本课题组前期也克隆了另外一个 UGT 基因^[18]。本研究在前期实验结果基础上, 对筛选出的可能参与柴胡皂苷生物合成的 BcUGT8 基因片段, 进行了全长 cDNA 克隆和序列分析, 以期为后续该基因研究提供参考。

植物体内众多 UGT 基因的存在阻碍了通过同源性鉴定单一基因可能参与的特定代谢途径的研究。UGT 蛋白立体结构是决定催化反应底物的关键。松叶菊中的 2 个 UGT 基因编码的氨基酸序列只有 19% 的一致性, 但催化相同的底物^[19]。因此, 目前仍不能仅仅通过序列分析确定基因编码酶的确切生物功能^[3]。体外表达蛋白催化活性分析是基因功能研究的重要手段, 本研究构建了 BcUGT8 的原

核表达载体, 为开展这方面的研究奠定了基础。除本课题组外, 以往国内外未见柴胡 UGT 基因克隆和功能研究报道, 柴胡 UGT 基因的研究不仅将促进柴胡次生代谢的分子机制研究, 基因工程在柴胡品种改良上的应用, 且对于植物 UGT 基因和皂苷合成研究也将具有一定贡献。

参考文献

- [1] Bowles D, Isayenkova J, Lim E K, *et al.* Glycosyltransferases: Managers of small molecules [J]. *Curr Opin Plant Biol*, 2005(8): 254-263.
- [2] Wang X Q. Structure, mechanism and engineering of plant natural product glycosyltransferases [J]. *FEBS Lett*, 2009, 583: 3303-3309.
- [3] Achnine L, Huhman D V, Farag M A, *et al.* Genomics-based selection and functional characterization of triterpene glycosyltransferases from the model legume *Medicago truncatula* [J]. *Plant J*, 2005, 41: 875-887.
- [4] Naoumkina M A, Modolo L V, Huhman D V, *et al.* Genomic and coexpression analyses predict multiple genes involved in triterpene saponin biosynthesis in *Medicago truncatula* [J]. *Plant Cell*, 2010, 22: 850-866.
- [5] Meesapyodsuk D, Balsevich J, Reed D W, *et al.* Saponin biosynthesis in *Saponaria vaccaria* cDNAs encoding β -amyrin synthase and a triterpene carboxylic acid glucosyltransferase [J]. *Plant Physiol*, 2007, 143: 959-969.
- [6] Sayama T, Ono E, Takagi K, *et al.* The *Sg-1* glycosyltransferase locus regulates structural diversity of triterpenoid saponins of soybean [J]. *Plant Cell*, 2012, 24: 2123-2138.
- [7] Shibuya M, Nishimura K, Yasuyama N, *et al.* Identification and characterization of glycosyltransferases involved in the biosynthesis of soyasaponin I in *Glycine max* [J]. *FEBS Lett*, 2010, 584: 2258-2264.
- [8] Shao H, He X Z, Achnine L, *et al.* Crystal structures of a multifunctional triterpene/flavonoid glycosyltransferase from *Medicago truncatula* [J]. *Plant Cell*, 2005, 17: 3141-3154.
- [9] Paquette S, Moller B L, Bak S. On the origin of family 1 plant glycosyltransferases [J]. *Phytochemistry*, 2003, 62(3): 399-413.
- [10] Qi L W, Wang C Z, Yuan C S. Isolation and analysis of ginseng: advances and challenges [J]. *Nat Prod Rep*, 2011, 28(3): 467-495.
- [11] Ashour M L, Wink M. Genus *Bupleurum*: A review of its phytochemistry, pharmacology and modes of action [J]. *J Pharm Pharmacol*, 2011, 63: 305-321.
- [12] 隋 春, 魏建和, 战晴晴, 等. 北柴胡鲨烯合酶基因及其编码区 cDNA 克隆与序列分析 [J]. *园艺学报*, 2010, 37(2): 283-290.
- [13] Sui C, Zhang J, Wei J H, *et al.* Transcriptome analysis of *Bupleurum chinense* focusing on genes involved in the biosynthesis of saikosaponins [J]. *BMC Genomics*, 2011, 12(1): 539-543.
- [14] Gachon C M M, Langlois-Meurinne M, Saindrenan P. Plant secondary metabolism glycosyltransferases: the emerging functional analysis [J]. *Trends Plant Sci*, 2005, 10(11): 542-549.
- [15] Chen S, Luo H, Li Y, *et al.* 454 EST analysis detects genes putatively involved in ginsenoside biosynthesis in *Panax ginseng* [J]. *Plant Cell Rep*, 2011, 30: 1593-1601.
- [16] Tang Q, Ma X, Mo C, *et al.* An efficient approach to finding *Siraitia grosvenorii* triterpene biosynthetic genes by RNA-seq and digital gene expression analysis [J]. *BMC Genomics*, 2011, 12: 343-346.
- [17] 向 丽, 郭 淑, 牛云云, 等. 三七 *PnUGT1* 基因的全长 cDNA 克隆和生物信息学分析 [J]. *药学学报*, 2012, 47(8): 1085-1091.
- [18] 隋 春, 徐洁森, 赵立子, 等. 北柴胡 UGT 基因的克隆及其过量表达和 RNAi 转基因载体的构建 [J]. *中国中药杂志*, 2012, 37(5): 558-563.
- [19] Vogt T. Substrate specificity and sequence analysis define a polyphyletic origin of betanidin 5- and 6-O-glucosyltransferase from *Dortheanthus bellidiformis* [J]. *Planta*, 2002, 214: 492-495.