

• 药材与资源 •

藏红花酸糖基转移酶 UGTCs4 基因的克隆、生物信息学及表达分析

赖成霞¹, 玛依拉·玉素音¹, 高燕¹, 李春平¹, 刘忠山¹, 石必显¹, 王博², 王玮²

1. 新疆农业科学院经济作物研究所, 新疆 乌鲁木齐 830091

2. 新疆农业科学院微生物应用研究所, 新疆 乌鲁木齐 830091

摘要: 目的 以藏红花悬浮细胞为研究对象, 克隆藏红花酸糖基转移酶(crocin glycosyltransferase)UGTCs4基因, 并对其进行生物信息学和表达分析。方法 采用同源克隆法和5'RACE(rapid amplification of cDNA ends)技术克隆UGTCs4基因的全长cDNA序列。运用生物信息学方法对该基因进行分析, 预测其编码蛋白的结构与功能, 并通过半定量RT-PCR(semi-quantitative reverse transcription and polymerase Chain reaction)方法检测不同诱导物条件下的基因表达情况。结果 UGTCs4基因的cDNA全长为1380 bp, 编码459个氨基酸。分子进化树分析, 该糖基转移酶基因与已知的藏红花酸糖基转移酶基因UGTCs2具有相同的糖基化藏红花酸的功能。RT-PCR的结果显示, UGTCs4基因能够响应吲哚乙酸(IAA)、脱落酸(ABA)、赤霉素(GA)、过氧化氢(H₂O₂)、茉莉酸甲酯(MeJA)多种处理, 这些处理均能促使其进行转录。结论 首次从藏红花悬浮细胞中分离并报道了UGTCs4的cDNA克隆, 并证实UGTCs4对不同诱导子的响应情况, 为进一步研究藏红花素的生物合成及表达调控奠定了基础。

关键词: 藏红花酸; 糖基转移酶; 克隆; 表达分析; 藏红花悬浮细胞

中图分类号: R282.12 文献标志码: A 文章编号: 0253-2670(2020)04-1044-08

DOI: 10.7501/j.issn.0253-2670.2020.04.031

Cloning, bioinformatics, expression mode analysis of crocetin glycosyltransferase UGTCs4

LAI Cheng-xia¹, MAYILA Yusuyin¹, GAO Yan¹, LI Chun-ping¹, LIU Zhong-shan¹, SHI Bi-xian¹, WANG Bo², WANG Wei²

1. Institute of Economic Crops, Xinjiang Academy of Agricultural Sciences, Urumqi 830091, China

2. Institute of Applied Microorganism, Xinjiang Academy of Agricultural Sciences, Urumqi 830091, China

Abstract: Objective With the aim to obtain crocetin glycosyltransferase UGTCs4 in cultured saffron suspension cell, and carry out its bioinformatics and expression mode analysis. **Methods** A homologous cloning strategy and 5' RACE methods were adopted, the full-length cDNA sequence of a crocetin glycosyltransferase, designated UGTCs4 (GeneBank number: KX398932), was obtained. The characteristics of physicochemical properties, structure and function of the deduced UGTCs4 protein were determined using a series of bioinformatics tools. Semi-quantitative PCR was used for gene expression analysis. **Results** The results showed that the full length cDNA of UGTCs4 was 1380 bp in length and encoding a 459 amino acid; UGTCs4 had high identities (83.2%) with UGTCs2 protein from saffron; UGTCs4 had the same evolutionary tree as UGTCs2. UGTCs4 transcripts were constitutively expressed in the leaves, stems, and roots. UGTCs4 gene could respond to multiple treatments of indoleacetic acid (IAA), abscisic acid (ABA), gibberellin (GA), hydrogen peroxide (H₂O₂), and methyl jasmonate (MJA), which promoted its transcription. **Conclusion** cDNA of crocetin glycosyltransferase was cloned from suspension cells for the first time and the response of UGTCs4 to different inducers was confirmed. Molecular characterization of UGTCs4 will be useful for further functional determination of the gene involving in the crocin biosynthesis and expression regulation.

Key words: crocetin; glycosyltransferase; cloning; expression analysis; saffron suspension cells

糖基化反应是植物体内的一种重要的次生代谢反应, 是通过糖基转移酶(glycosyltransferase, GT, EC 2.4.x.y)来完成这项生理功能的, 它能将糖基从激活的受体尿嘧啶核苷二磷酸-葡萄糖(UDP-

收稿日期: 2019-09-03

基金项目: 国家自然科学基金项目(C020409); 新疆自治区高技术项目(201211107)

作者简介: 赖成霞, 副研究员, 从事植物次生代谢调控研究。Tel: (0991)4534137 E-mail: lchxia2001@163.com

glucose) 上转移到不同的受体上, 是生物体内生物分子进行结构修饰^[1-2]、功能改变的一种重要方式, 通过这种方式能提高生物分子的溶解性、增加生理活性、加强膜的耐性, 使植物能适应生长发育的多个阶段, 如参与植物的防御反应^[3-4]、抗虫、脱毒反应^[5]、植物次生代谢^[6-7]以及调节植物激素平衡^[8]等多个方面。

藏红花是最昂贵的农作物之一, 主要原因是由于其药用成分集中在柱头上, 其水溶性良好, 可用作香辛料、食品添加剂, 具有抗氧化^[9]、镇痛^[10]、抗炎等作用。藏红花酸糖基转移酶是藏红花素生物合成代谢过程中最关键的酶, 能以添加 UDP-葡萄糖为供体, 催化形成一类由藏红花酸糖酯类组成的藏红花素类物质, 如藏红花素苷 1、藏红花素苷 2、藏红花素苷 3、藏红花素苷 4 及藏红花素苷 5^[11], 这些物质具有良好的水溶性, 被认为是天然食品添加剂之一^[12], 并被用于抗癌^[13-14]、抗氧化等药物研究^[15-16]。

到目前为止, 已有较多学者对藏红花酸糖基转移酶进行了研究, 早在 2000 年, Cote 等^[17]就利用生物化学的手段, 从藏红花悬浮细胞中纯化到了 2 种糖基转移酶, 确定 2 种酶的相对分子质量均为 47 000, 等电点分别为 4.4 和 4.8, 并对 2 种蛋白进行体外实验, 发现它们均能以藏红花酸为受体底物发生糖基化反应, 并且二者还具有糖苷酶的活性, 能水解 4-硝基苯基-β-D-吡喃葡萄糖苷形成 4-硝基苯酚及 β-葡萄糖苷, 尽管其对糖基转移酶的生化性质进行了描述, 可是并没有鉴定出 2 种糖基转移酶的基因序列。2004 年, Moraga 等^[18]以藏红花柱头为材料, 采用同源片段克隆的方式, 首次克隆到 1 个藏红花酸糖基转移酶 UGTCs2 基因, 该基因的重组蛋白能在体外使藏红花酸发生糖基化反应形成一系列藏红花酸多糖酯。2012 年 Nagatoshi 等^[19]从梔子细胞中分离发现 UGT75L6 和 UGT94E5 2 个基因, 从而揭示了藏红花素合成的分子机制, 这 2 个糖基转移酶能参与藏红花素合成的 5 步糖基转移酶, 首先 UGT75L6 基因原核表达的蛋白粗提物具有顺序糖基化藏红花酸形成藏红花素苷 5、藏红花素苷 3 和藏红花素苷 4 的功能, UGT94E5 基因原核表达的蛋白粗提物具有糖基化藏红花素苷 5 形成藏红花素苷 2 和藏红花素苷 1 的功能。

尽管目前已经在梔子中诠释了藏红花素合成的分子机制, 而在藏红花中, 仅仅建立了以藏红花

酸为受体, 顺序叠加成藏红花酸多糖酯的机制, 而对于藏红花素的生物合成仍然还没有被揭示。

UGTCs4 基因是研究藏红花悬浮细胞中合成藏红花素代谢途径中, 利用同源克隆-RACE (rapid amplification of cDNA ends) 的方式获得的一个类藏红花酸糖基转移酶基因, 对其进行生物信息学分析及诱导表达分析, 以期为藏红花素的形成及分子调控机制的研究提供参考。

1 材料

藏红花细胞来自清华大学郭志刚教授实验室的馈赠; 细胞分裂素 (BA)、赤霉素 (GA)、烟酰胺腺嘌呤二核苷酸 (NAD)、吲哚乙酸 (IAA) 均购自上海生工有限公司; M-MLV、RNasin、pGEM-T easy vector (Promega 公司); oligod (T)、dNTPs、质粒小提试剂盒、胶回收试剂盒、PCR 产物回收试剂盒 (上海生工有限公司); 2×Taq PCR mix、大肠杆菌 DH5α 感受态细胞 [天根生化科技 (北京) 有限公司]; 3'RACE、5'RACE 试剂盒 (Takara 公司); 其他试剂均为国内分析纯, 引物为生工生物工程公司合成。

2 方法

2.1 细胞培养

本实验中植物材料为藏红花悬浮细胞, 该植物细胞培养温度为 (22±2) °C, 采用的培养方法是固液 2 步法, 即利用固体培养基繁殖细胞数量, 利用液体培养基培养细胞生成产物, 该细胞在 1/2 MS 液体培养基中, 添加 1 mg/L BA、2% 葡萄糖、100 mL/L 番茄汁、2 mg/L GA、0.1 mg/L NAD、40 μg/L H₂O₂、2 mg/L IAA, 在 23 °C 条件下, 诱导 2 d, 将细胞用 ddH₂O 多次洗涤后, 真空抽滤细胞表面的水分后, 置于 -80 °C 冰箱备存, 用于 RNA 提取。

2.2 RNA 的提取和 cDNA 的合成

本实验采用的试剂为 Trizol, 提取方法遵照实验生产商提供的技术规程进行, 凝胶电泳检测 RNA 的完整度, 28 S 与 18 S 比值为 1.5~2:1 可视为完整度较好, 紫外分光光度计法测定 RNA 含量及纯度, $A_{260}/A_{280}=1.9\sim2.0$, $A_{260}/A_{230}=1.9\sim2.0$, 这样纯度的 RNA 可用于 cDNA 第 1 链的合成。

2.3 UGTCs4 基因保守序列的克隆

参与植物次生代谢的糖基转移酶具有保守的序列 PSPG (plant secondary glycosyltransferase), 根据糖基转移酶这个特性, 利用文献报道^[20]的简并引物 PSPG: 5'-TTYGTIACICAYTGYGGITGG-

AA-3'及 Takara 公司的 3'RACE 的引物,通过 PCR 方法从藏红花愈伤组织的 RNA 反转录的产物为模板,扩增出保守序列 PSPG 第 1 轮产物,条件为预变性 94 ℃、5 min; 94 ℃、1 min, 58 ℃、1 min, 72 ℃、45 s, 40 个循环; 72 ℃、10 min。将第 1 轮产物稀释 100 倍后为模板进行第 2 轮扩增,条件为预变性 94 ℃、5 min; 94 ℃、1 min, 68 ℃、1 min, 72 ℃、45 s, 40 个循环; 72 ℃、10 min。PCR 产物经胶回收后,将第 1 轮 PCR 产物稀释 50×后的产物为模板继续进行第 2 轮 PCR,反应体系与上述相同,PCR 产物经 2% 琼脂糖凝胶电泳后,切胶纯化回收,将该保守序列克隆到 pGEM T-Easy Vector 上,命名为 pGEMT-PSPG。转化大肠杆菌 DH5 α 后提取重组质粒并测序,测序由上海生工公司完成。

2.4 5'RACE 技术获取 UGTCs4 基因 5'末端序列

5'RACE 是步骤遵照 Takara 公司提供的技术规程进行 PCR 扩增,根据“2.3”项中获得的藏红花酸糖基转移酶 UGTCs4 保守片段的 cDNA 序列,应用 DNAMAN 设计 5'RACE 特异性引物,根据 M-MLV 逆转录酶的说明进行逆转录反应,利用特异性引物进行 PCR 扩增,第 1 轮 PCR 引物: 5'-CACTAACAGACTCCTAACCAACTTAC-3'; 第 2 轮 PCR 引物: 5'-CTCCCCACCATCGGCACACCA-AGCGC-3', 反应体系为 20 μL, 其中引物各 1 μL, 2×Taq mix 10 μL, cDNA 模板 1 μL。反应条件为 94 ℃、5 min; 94 ℃, 1 min, 63 ℃, 1 min, 72 ℃, 45 s, 40 个循环后 72 ℃延伸 10 min。将第 1 轮 PCR 产物稀释 50×后的产物为模板继续进行第 2 轮 PCR, 反应体系与上相同, 反应条件退火温度为 68 ℃相同, PCR 产物经 1% 琼脂糖凝胶电泳后, 参照“2.3”项中方法回收、克隆, 并送至生工公司测序。

2.5 UGTCs4 基因全长的拼接及 PCR 验证

利用 DNAMAN 软件将“2.2”中获得的含保守序列的 3'末端序列和“2.4”项中获得 5'末端序列进行拼接,获得了 UGTCs4 基因的全长 cDNA 序列。根据拼接的序列,设计扩增包含起始密码子的上游引物和包含终止密码子的下游引物。以逆转录反应的 cDNA 为模板, PCR 扩增 UGTCs4 的全长序列,反应体系为 20 μL,其中引物各 1 μL, 2×Taq mix 10 μL, cDNA 模板 1 μL。反应条件为 94 ℃, 5 min; 94 ℃, 1 min, 68 ℃, 1 min, 72 ℃, 45 s, 40 个循环后 72 ℃延伸 10 min。PCR

产物经 1% 琼脂糖凝胶电泳后,参照分子克隆中的方法回收、克隆,进行酶切鉴定后,送至上海生工公司测序验证。

2.6 UGTCs4 生物信息学分析

通过 DNAstar 生物学软件将 UGTCs4 基因的 cDNA 全长序列翻译成氨基酸序列, 使用 DNAstar 查找开放阅读框, 使用在线软件 <http://www.ncbi.nlm.nih.gov/Blast.cgi>, 将不同物种中功能各异的糖基转移酶的蛋白序列进行比对分析, 使用软件 MEGA 5.1 中 Neighbor-Joining 构建系统进化树; 利用 ProtParam (<http://au.expasy.org/tools/protparam.html>) 软件分析蛋白的生化特性。

2.7 UGTCs4 基因在藏红花细胞中的诱导表达

设置了如下实验对 UGTCs4 在藏红花细胞中的转录情况进行研究,对照处理组采用 1/2 MS 为基本培养基,按 1:10 的比例加入自制番茄汁。处理组采用 1/2 MS + 番茄汁为处理培养基,在 23 ℃条件下, 100 r/min 转速条件下, 将细胞处理 2 d 后, 在空白对照组中加入水作为处理,而在处理组中分别加入 50 μmol/L 的 ABA、GA、SA 等作为诱导子, 培养 24 h 后收取细胞, 用以今后提取 mRNA 的材料, 利用 DNAMAN 生物学软件, 设计半定量 PCR 扩增 UGTCs4 引物, 预计扩增 280 bp。上游引物: 5'-ACACCCGGTTCATCTCCAACTCCA-3', 下游引物: 5'-AGAGTTCCGGGTTAATGACTGGGTG-3', 扩增反应体系为 20 μL, 其中引物各 1 μL, 2×Taq mix 10 μL, cDNA 模板 1 μL。反应条件为 94 ℃、5 min; 94 ℃、1 min, 60 ℃、1 min, 72 ℃、15 s, 40 个循环, 72 ℃延伸 10 min。内标参比基因为 Tublin 基因, 上游引物: 5'-ATGATTCCAAC-TCGACCAGTGT-3'; 下游引物: 5'-ATACTCAT-CACCCTCGTCACCATC-3', 扩增反应体系为 20 μL, 其中引物各 1 μL, 2×Taq mix 10 μL, cDNA 模板 1 μL。反应条件为 94 ℃、5 min; 94 ℃, 1 min, 58 ℃, 1 min, 72 ℃, 15 s, 40 个循环, 72 ℃延伸 10 min。用半定量反转录 PCR (semi-quantitative reverse transcription PCR) 的方法检测在这几种诱导因子存在的条件下, UGTCs4 基因的转录情况。以不同处理的藏红花细胞的 RT 产物为模板进行扩增, PCR 产物经 1.5% 琼脂糖凝胶电泳后, 利用内参 Tublin 基因作为内标参比, 调节不同处理的 UGTCs4 的样品量, 最终解析不同处理条件下的 UGTCs4 的转录情况。

3 结果与分析

3.1 藏红花悬浮细胞 RNA 的提取及 UGTCs4 保守序列的扩增

利用 Trizol 试剂从藏红花悬浮细胞中提取 RNA，并将 RNA 中残余的 DNA 用 DNase I 去除，最终获得可供实验使用的 RNA（图 1-A）；根据 UGTCs4 保守序列 PSPG 设计简并引物，利用 3'RACE 技术从悬浮细胞中获得包含糖基转移酶基因保守序列 4 条不同长度的 DNA 片段的条带（图 1-B）。将其分别连接到 pGEM-T easy 载体转化大肠杆菌 DH5 α ，每个片段分别挑取 10 个菌落进行测序，共计 50 个菌落（编号 1~50），测序结果经 NCBI 的 Blast 比对。其中 4 号菌落所包含的 501 bp 长的核苷酸序列与黄酮类的糖基转移酶有高度的相似性，如藏红花糖基转移酶 UGT703B1 基因、芦荟糖基转移酶 GT5 基因等，因此推测 4 号菌落所包含的核苷酸序列与黄酮类糖基转移酶基因是具有同源性。21 号菌落中包含有 404 bp 长的核苷酸片段，基于核苷酸相似度的在线比对结果显示，21 号菌落中包含的核苷酸片段与 UGTCs4 基因具有高度相似性，相似度高达 77%，因此推测 21 号菌落包含的核苷酸序列与藏红花酸糖基转移酶 UGTCs2 基因有同源性，该片段可能为 UGTCs2 的异型体的一部分核酸序列。

3.2 UGTCs4 基因的克隆及序列分析

以获得的藏红花悬浮细胞的 cDNA 为模板，根据获得的保守区序列设计引物，用 5'RACE 技术扩增到所需的片段，经琼脂糖凝胶电泳检测并测序后（图 2），将获得的序列与上述含 PSPG 保守序列的 3' 末端的进行拼接，获得含有开放阅读框（ORF）的 UGTCs4 基因完整序列，设计引物，对上述序列进

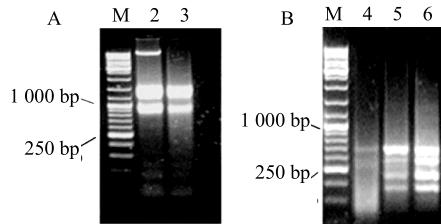
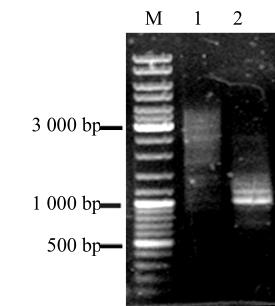


图 1 藏红花 RNA 的提取 (A) 及 UGTCs4 保守区的扩增 (B)
Fig. 1 RNA extraction (A) and amplification (B) of conserved region of saffron glycosyltransferase



M-Marker 1-UGTCs4 5'RACE 第 1 轮 PCR 2-UGTCs4 的 5'RACE 第 2 轮 PCR
M-Marker 1-round 1 PCR of 5'RACE for UGTCs4 2-round 2 PCR of 5'RACE for UGTCs4

图 2 UGTCs4 基因的 5'RACE-PCR 扩增

Fig. 2 5'RACE-PCR amplification for UGTCs4 gene

行 PCR 检测并测序验证，获得 UGTCs4 的 cDNA 序列（图 3），利用 NCBI 网上的在线软件 ORF finder，发现具有完整的 ORF，命名为 UGTCs4，genebank 登录号为 KX398932（图 3），该序列全长 1 377 bp，编码 459 个氨基酸，该基因 3' 非翻译区（UTR）有 85 bp，5' 非翻译区有 100，利用 NCBI 在线软件对 UGTCs4 的保守序列进行预测，在 UGTCs4 的 C 端存在受体底物结合的活性位点（图 4），N 端存在受体糖基结合的活性位点（图 3 中方框所示氨基酸），N 端的 22、23 及 25 位的 G（谷氨酸）^[21]、H（组氨酸）^[22]、N（天冬酰胺）或 R（精氨酸）残底物基^[23]，是所有参与植物次生代谢调控的保守序列，在 C 端存在供体糖基存在的活性位点，称为“PSPG 盒”，位于蛋白的 343~351 位，其序列为 HCGWNS，这个序列与二磷酸葡萄糖（UDPG）的结合位点^[24]。

3.3 UGTCs4 蛋白质理化性质及结构预测

利用 Protparam 在线程序 (<http://au.expasy.org/tools/protparam.html>) 分析其蛋白质的理化特性可知，该蛋白相对分子质量为 51 000，等电点为 5.41，包含 7 123 个原子，分子式为 $C_{2274}H_{3530}N_{612}O_{687}S_{20}$ ，其中该蛋白中丝氨酸含量最高，为 9.21%，组氨酸、半胱氨酸、色氨酸含量最低，为 2.2%，不稳定系数 43.12，为不稳定蛋白，脂肪系数 80.63，亲疏性分析该蛋白总平均亲水性为 -0.299。利用 TMpred 在线程序 (https://embnet.vital-it.ch/software/TMPRED_form.html) 对该蛋白的跨膜结构进行预测，得分超过 500 的是跨膜结构区，结果显示，该蛋白有 2 个跨膜结构区，

1 AAGTCGCATGCTTCCGGCCCATGGCGGCCGGGAAATTGATTGCGGATCCACAGCCTACTGATGATCAGTCGAT
 79 GGAAAAGCCAATATCACCCCATGGAGCAGAAAGATGAGAACGGCAACAACGCCACATCCCTCCCTCCCTATC
 1 CAGCACAAGGCCACATCAACCCCATGCTCCAAATTGCGCAAGCGACTAGCCCTCTCGCAACCTCCGACCCACTCGTCA
 15 L L P Y P A Q **G H I N** P M L Q F G K R L A S R N L R
 235 ACACCCGGTTCATCTCCAACCCACCCAAACCCGACCCCGGTCCGTCACACCCGAGTGTGATATCCGACGGATTGATT
 31 T T L V N T R F I S N S T K P D P G P V N T E C I S
 313 CGCGTGGGCTGAATGCCGACCTAGTCGCGAGGCCATTITGCCGACTCGAGTCGGTCCGGTCAGAAACACTGTTTG
 57 D G F D S G G L N A A P S R E A Y F D R L E S V G S
 391 GCCTCATCGAGTCACTTCGATCCAGAGGCCGACCCGGTCATGTTTGGTCTACGACCCGTTCTTGCCTTGGCAACGAA
 83 E T L F G L I E S L R S R G R P A H V L V Y D P F L
 469 ACGTAGCCGAGCGGTCCGGTTGGCTGGTAGCATTCTTAACAGCGGTGTGCAAGTGGACACAAATCTACCGCC TCG
 109 P W A T N V A E R S G L R S V A F F T Q P C A V D T
 547 TGTTGGGAGGGACAATTAAGGGTCCGTAACCGAGCCGGTCCGGTCCGCTGGTTCCCCCAACTCGAGCCTA GTGATC
 135 I Y R L V W E G Q I K V P V T E P V R L P G F P P L
 625 TGCCGTCGTTCTGAACGGATTGCAACCCAGTCATTAACCCGGAACCTCTGCCACTCTAGTGAACCAGCACA AAAAATT
 161 E P S D L P S F V T D S H P V I N P E L L P L L V N
 703 TGGAGAAAGCTGATATGATGCTGATTAACCTGGTACAGAGTTAGAGCATGAGGAATTGATTGGATGAAGTCCCATT
 187 Q H K N L E K A D M M L I N S V Y E L E H E E F D W
 781 TACCA TTACCGGTGAAGTCGATGGGCCGACCGTGGCTACTTATCTGACAACCGAATCCGCGGATCACACT
 213 M K S R L P L P V K S I G P T V P S T Y L D N N R I P
 859 AATGGATTCAATCTATACACCCCTGACACTACACCCCTACTTGGACTGGCTGACTCCAAGGACCCGAACCTGTGTCATC
 239 S D S H Y G F N L Y T P D T T P Y L D W L D S K D P
 937 TTGTCCTCATTTGGTAGTATGCCAGTCTGAGTCGGATCAAACCAATGAGATTGCGTCGGGCTGATGCCACCAACA
 265 N S V I Y V S F G S M A S L I S P D Q T N E I A S G L I
 1015 AAAGCTTCATTTGGGGTAGCAGCACTTCGGAGTTGCCAAACTCCCTACCAACTTCAACCCAAAGAGAAATGCGAGCCGTG
 291 A T N K S F I W V V V R T S E L A K L P T N F T Q E N
 1093 GTTGGTTGTGACCTGGTCAGCCAACACTAGATGCTTGCACAGCTGGCTGACTCCAAGGACCCGAACCTGTGTCATC
 317 A S R G L V V T W C S Q L D V L A H V A T G C F V T
 1171 GGAACCTCGATATGGAAGGTATCGCGCTTGGTGTGCCGATGGTGGAGTGCCTGGGACCCAGCCTATGAATG
 343 H C G W N S T M E G I A L G V P M V G P Q W S D Q
 1249 CCAAGTATGIAAGAGGAATGTTGGAAAGTCGGCTGAGGGCGAAGACTTATGGGAAGGATTCTGTGAGCGGGAGGAGT
 369 P M N A K Y V E D V W K V V G V R A K T Y G K D S V S
 1327 CAAAAGATGTGGAGGGGGTAGGTGATGGAGAGAGAGGGATGGAAAGGATTAGAGAGAATGCTGCAAGGTGGAGTAAGTT
 295 G E E F K R C V E E V M D G E R S G K I R E N A A R
 1405 GTTAAGGAGTCTGTTAGTGGAGGGGGAGTTGCTGACAAAGTGTATTAGGAGTCTGATCTGATCATTGTGAAATG
 421 W S K L V K E S V S E G G S S D K C I K E F I D H C C K
 1483 CTCCAAAGTTCTTAGTGTGATTCATCTGTGATAAAAGTTGTTCATGTCAACCTGATCAATCTACTAAGAGAAAA
 1561 TTAG

起始、终止密码子用黑体标出，受体底物活性位点 G 谷氨酸、H 组氨酸、N 天冬酰胺用方框标出，供体底物活性位点用下划线标出。The start codon and stop codon marked in boldface, the receptor substrate active sites G (glutamic acid), H (histidine), N (asparagine) marked with a square box, and the donor substrate active sites marked with an underscore

图 3 UGTCs4 基因的全长序列及推测的氨基酸序列

Fig. 3 Full length sequence of UGTCs4 and its inferred amino acid sequence

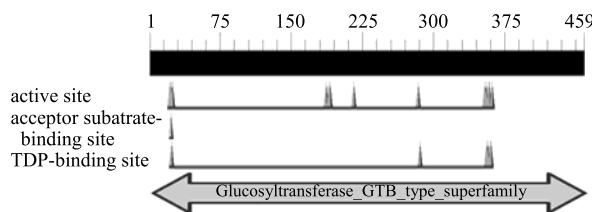


图 4 UGTCs4 蛋白功能性位点分析

Fig. 4 Functional site analysis of UGTCs4 protein

均为 α 螺旋形式跨膜，方向均为从内到外，第1个跨膜区从第332个氨基酸残基到第357个氨基酸残基，核心氨基酸残基在341位，第2个跨膜区从第346个氨基酸残基到第372个氨基酸残基，核心氨基酸残基在366位。利用SignalP在线软件(<http://www.cbs.dtu.dk/services/SignalP/>)预测，该蛋白，该蛋白没有信号肽，不属于分泌蛋白，利用NetPhos3.1在线程序(<http://www.cbs.dtu.dk/services/NetPhos/>)对该蛋白的磷酸化位点进行预测，结果显示，有27个丝氨酸，13个苏氨酸及7个酪氨酸残基上可能发生磷酸化，利用NCBI的在线软件对UGTCs4的保守序列进行预测(图4)，结果表明该蛋白属于糖基转移酶超家族，具有多个保

守位点，在N端具有受体底物结合的位点，在C端具有TDP结合位点。

3.4 UGTCs4 的分子进化树构建

为了确定克隆到UGTCs4是否具有藏红花酸糖基转移酶同样的功能，将其氨基酸序列与已知具有藏红花素糖基化功能的糖基转移酶UGTCs2、UGT75L6、UGT94E的氨基酸进行了序列比对(图5)，结果表明，UGTCs4的氨基酸序列与UGTCs2具有69.9%相似性，与UGT75L6的氨基酸序列具有47.7%选择的相似性。而与UGT94E5的氨基酸序列仅为5.8%，同时本课题组对来自藏红花、栀子、雏菊、棉花、长春花、拟南芥等植物中已经报道的23个糖基转移酶进行聚类分析，利用Clustaw软件对这些基因的氨基酸序列进行多重比对，结合MEGA 5.1软件，建立了糖基转移酶的分子进化树，由图6可知，不同植物来源糖基转移酶被聚类分为2个分支，藏红花的UGTCs4蛋白(KX398932)与藏红花的UGTCs2蛋白(AY262037)以及栀子的UGT75L6蛋白(A B733668)处于相同分支中，说明这UGTCs4可能具有与UGTCs2、UGT75L6相同的生物学功能，能催化藏红花酸的糖基化。

UGT94E5	1	-MW-	-AYGHS-	-YAKETD-	R	
UGT75L6	1	-MVRHVT-	-YAGHN-	-AABRMG-	-VLTATSVYASRMTKESGSGSTK	
UGTCs2	1	-MNGNRCH-	-CAGHN-	-GKHAS-	-HNTTVNTR-SNSTRS-	
UGTCs4	1	MEQKDENGNNCHILLLPYP	ACGHINPMLQFGKPLASRNLR	EVNTREISNSTRFDPGPV		
UGT94E5	25	GAYGSTNGKKRTGKYSVTKVHDTHWHTTN		GHMATRANGAKSNKTRDVYDAT		
UGT75L6	37	GGAT-SDGY-DGRKGVDHTYMSAKGSNT		GRVNTSADG-CVTCVYV-		
UGTCs2	32	GVNC-SDG--DGC-TAASRRA-YDRSR-		GKHWGSRSGRG-RGACGRV-		
UGTCs4	62	NTECISDGFD-SCGILNAPSREAYEDRLESVGSETLEGLIESLRSRG-RPAHVLYYDPFL				
UGT94E5	67	TWTATRVAHHNAVKTSS-VSMAY-		CHMKG-	AYS	
UGT75L6	61	WAAALVA-ECHSAAVVA-	VMDYYYYRGYDD-	VKNNS-	MPT	
UGTCs2	73	WAAHVA-RSG-RSVA-	TCAVDT-YRIVWGRKV-	MAVRG-	SAC	
UGTCs4	119	PWATKVAERSGLRSVAFFTQPCAVDTIYRIVWEGQIKVPTEPVRLP	GFPPLFSDLPSF			
UGT94E5	96	DAKARTAAADARA-	DANDAARNR-	DCDSVRSR		
UGT75L6	115	WSGSMMKAQDSS-	DNYSATKNTDRKVWNT	DAAKASYNA		
UGTCs2	109	VRNGGRVVNDREV-	RHKN-DKADMGRNSYAD-	DGSRVRSGT		
UGTCs4	179	VTDSPHPVINPELLPLLVNQHKNL-EKADMMLINNSVYELEHEEFDWMSKSLPLPVNSIGPTV				
UGT94E5	126	AGKY-DYD-MKKMVG-	-MVVKDDGDNNSN-	GTRSRSTVVSGT	-YTKMAH	
UGT75L6	152	GTSAA-DGK-DSTSSG-	-DLSKDYKVN-NSRD-	-GSVYVVS	-GSTMAR	
UGTCs2	147	VSTY-DNR-SDSHYG-	-NYTDAT-	-YDW-DSR-	-ANSVYVS	-GSSSDTHAS
UGTCs4	239	PSTYLDNRIPSDSHYC	FNLTYTPD-TPYLDWLDSDKIDPN	SVIYVSFGMASLSPDQTNIEIAS		
UGT94E5	169	G--SVN--WWVRAMGIRD-	-AGRVGDRG-PVCLASV-	AESTGGC-HCGWN		
UGT75L6	194	G--KSGR-WV-RAK-	-NGKDRCMGVWCSV-	-EH-SGC-VTHCGWN		
UGTCs2	189	G--ATNKS-WWVRTSAKAN-	-TNASRG-VVTTWCD-	-AHVATGC-VTHCGWN		
UGTCs4	299	GLIATNKSFIWWVRTSELAKLPTNFTQENASRGLVVWTCQLDVL	LAHVATGC	FVTHCGWN		
UGT94E5	210	SVVSGVVV-AMMH-	-DNAR-VVVGAGVVVRDTGKD-			
UGT75L6	227	ST-SVC-GV-WVVAH-	-WDG-TNAK-DVW-	-VSD-KRCT-VMD		
UGTCs2	232	STM-GVA-GV-MVGV-	-WSD-MNAKVV-	-DVWVKGVRKATYGD-VRG-KRCV-VMD		
UGTCs4	359	STMEGHALGVPMVCVPQWSDQPMNAKYVEDDVWVKGVRKATYGD	SVGEFFKRCVEEVMD			
UGT94E5	239	--RKARA-KDAMVKGNTRAKMDVKGRVKKD-VATVTTTSSN-				
UGT75L6	269	DGKGVRL-RNAAR-KARADGSS-	-DKNKARADAGKGYAS-			
UGTCs2	278	G-RSGK-RNAARAPCK-AKDSVS-	GGSSDKC-	-RHC-CDSDSKSV		
UGTCs4	419	GERSGKIRENAARPSK-LVKEESVS-	GGSSDKCIKEFIDHCK-			

图 5 不同藏红花酸糖基转移酶的氨基酸序列比对分析

Fig. 5 Comparison and analysis of amino acid sequences of different glycosyltransferases

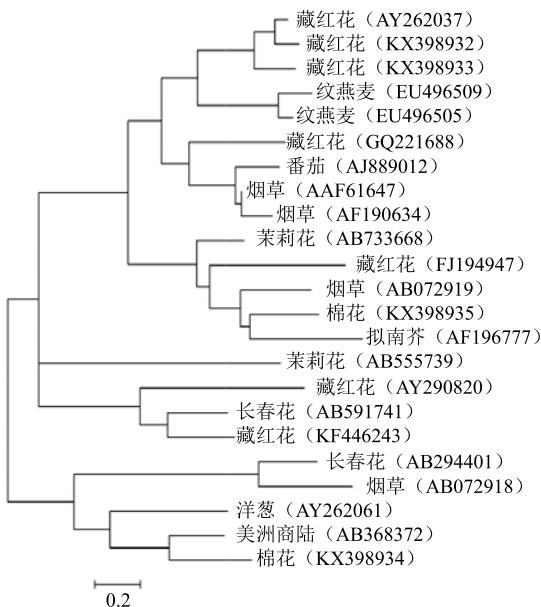


图 6 不同植物 UGTs 分子进化树

Fig. 6 Phylogenetic tree of different plants UGTs

3.5 UGTCs4 基因对不同处理的表达分析

一般情况下，基因的转录调控可以通过启动子、内含子、3'非翻译区等这些单元内的功能性位点与相应的结合蛋白的相互作用进行调控行使功能。为了进一步了解这些因素是否参与了对UGTCs4 及UGTCs5 的转录调控，而UGTCs4 基

因则对所有的诱导处理都有所响应。结果如图 7 所示，UGTCs4 基因在空白对照组中有微量的表达，而加入 IAA、ABA、GA、H₂O₂、MeJA 等诱导因子后，基因的表达增强，说明 UGTCs4 基因能够响应 IAA、ABA、GA、H₂O₂、MJA，能促使其进行转录。利用半定量 RT-PCR 技术检测了不同处理条件下 UGTCs4 基因的表达情况，结果显示（图 7），UGTCs4 基因在空白对照组中有微量的表达，而加入 IAA、ABA、GA、H₂O₂、MeJA 等诱导因子后，基因的表达增强，说明 UGTCs4 基因能够响应 IAA、ABA、GA、H₂O₂、MeJA，能促使其进行转录。对半定量 PCR 的结果进行灰度扫描分析，将目的基因的灰度值与参比内标基因的灰度值进行比较作柱状图，结果如

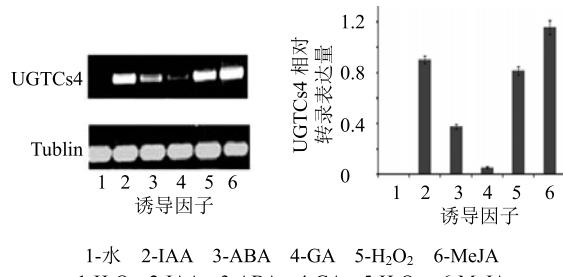


图 7 UGTCs4 基因在不同诱导因子处理下转录情况

Fig. 7 Transcription analysis of UGTCs4 gene under different inducible factors

图 7 所示, 与半定量 PCR 结果一致, 其对不同诱导因子的响应顺序依次为 MeJA>IAA>H₂O₂>ABA>GA>H₂O。

4 讨论

在本实验中根据糖基转移酶具有保守序列 PSPG 盒这个特性, 建立了包含保守序列及 polyA 信号的 cDNA 文库, 这是一个新的尝试, 利用这个方法 Nagatoshi 等^[19]从桉树悬浮细胞中成功地克隆到了一个糖基转移酶基因。在糖基转移酶的保守区中仅有 HCGWNS 这几个氨基酸, 因为其可能为 UDP-葡萄糖中尿苷酸的结合位点, 因此相对比较保守。而 PSPG 其他的部分的氨基酸保守性相对只是针对具有同样底物类型的糖基转移酶中, 因此, 仅利用 PSPG 这个保守序列克隆基因, 势必会提高了简并引物的简并性, 降低了引物与模板的有效结合, 造成了直接利用 PCR 的方法获得糖基转移酶保守序列的几率就大大降低了。利用 Moraga 等^[18]于 2004 年发表的简并引物, 结果发现, 因为其简并性比较高的原因, 没有得到任何片段。在本实验中扩增糖基转移酶的保守区的引物设计就是基于这个原理, 降低简并引物的简并性, 在其中加入次黄嘌呤 I, 结合真核基因转录后产生 1 个 polyA 尾巴, 在其 polyA 尾巴加上适当的引物, 这样就比较容易地获得包含部分保守序列的 cDNA 文库, 为糖基转移酶的克隆提供了一个思路。

本研究克隆的 UGTCs4 与已报道的其他植物来源的糖基转移酶相比, 都具有保守的 PSPG 盒, 其生化功能分析显示, 该蛋白在水溶液中稳定性差, 并且其与来源于藏红花酸的 UGTCs2 的氨基酸相似性高达 69.6%, 其分子进化树分析可知, 两者共同属于一个分支, 说明其可能是藏红花酸糖基转移酶 UGTCs2 的一个异型体, 并具有相同的功能——糖基化藏红花酸。

目前藏红花素的生物合成机制在栀子中已有揭示, 而在藏红花细胞中尚未被阐明, 尽管 2004 年从藏红花中已经分离获得的 UGTCs2 基因, 其原核表达的重组蛋白的粗提物在体外能糖基化藏红花酸, 但是它能在藏红花酸上添加多个糖基, 从这个角度说, 该基因并未真正参与藏红花素的代谢调控过程。本实验是首次从藏红花悬浮细胞中克隆了藏红花酸糖基转移酶基因 UGTCs4, 生物信息学分析结果表明, 该基因具有参与次生代谢的保守序列 PSPG, 并且为藏红花酸糖基转移酶基因 UGTCs2 的

异型体, 并证实了该基因对 IAA、GA、ABA、MeJA、H₂O₂ 的处理等都具有响应, 其中对 IAA、H₂O₂、MeJA 的响应最为明显, 这为藏红花细胞中藏红花酸糖基转移酶 UGTCs4 的表达调控奠定了基础。

参考文献

- [1] 王勇, 赵晓珍, 薛伟, 等. 尿核苷二磷酸糖基转移酶及其抑制子在植物激活剂中研究与应用 [J]. 中国农学通报, 2018, 34(3): 58-69.
- [2] 张桂芝, 林继山, 李燕洁, 等. 植物激素糖基化修饰研究进展 [J]. 植物学报, 2014, 49(5): 515-523.
- [3] Kosaka A, Manickavelu A, Kajihara D, et al. Altered gene expression profiles of wheat genotypes against *Fusarium* head blight [J]. *Toxins (Basel)*, 2015, 7(2): 604-620.
- [4] Langlois-Meurinne M, Gachon C M, Saindrenan P. Pathogen-responsive expression of glycosyltransferase genes UGT73B3 and UGT73B5 is necessary for resistance to *Pseudomonas syringae* pv tomato in *Arabidopsis* [J]. *Plant Physiol*, 2005, 139(4): 1890-1901.
- [5] Poppenberger B, Berthiller F, Lucyshyn D, et al. Detoxification of the *Fusarium* mycotoxin deoxynivalenol by a UDP-glucosyltransferase from *Arabidopsis thaliana* [J]. *J Biol Chem*, 2003, 278(48): 47905-47914.
- [6] 郭淑, 罗红梅, 宋经元, 等. 糖基转移酶在植物次生代谢途径中的研究进展 [J]. 世界科学技术—中医药现代化, 2012, 14(6): 2126-2130.
- [7] Choi D W, Jung J, Ha Y I, et al. Analysis of transcripts in methyl jasmonate-treated ginseng hairy roots to identify genes involved in the biosynthesis of ginsenosides and other secondary metabolites [J]. *Plant Cell Rep*, 2005, 23(8): 557-566.
- [8] Wandall H H, Rumjantseva V, Sorensen A L, et al. The origin and function of platelet glycosyltransferases [J]. *Blood*, 2012, 120(3): 626-635.
- [9] 陈可, 曲绮雯, 陈霞连, 等. 藏红花花瓣粗多糖的体外抗氧化活性研究 [J]. 四川大学学报: 自然科学版, 2016, 53(2): 448-452.
- [10] 孟庆山, 尹恒, 胡建恩, 等. 植物糖基转移酶在植物抗病过程中的研究进展 [J]. 生物技术通讯, 2013, 2(10): 290-293.
- [11] Li N, Lin G, Kwan Y W, et al. Simultaneous quantification of five major biologically active ingredients of saffron by high-performance liquid chromatography [J]. *J Chromatogr A*, 1999, 849(2): 349-355.
- [12] Jose B M, Alonso S G, Jimenez-Monreal A M, et al.

- Saffron: An old medicinal plant and a potential novel functional food [J]. *Molecules*, 2017, 23(1): 30-50
- [13] Kim S H, Lee J M, Kim S C, et al. Proposed cytotoxic mechanisms of the saffron carotenoids crocin and crocetin on cancer cell lines [J]. *Biochem Cell biol*, 2014, 92(2): 105-111.
- [14] Hoshyar R, Mollaei H. A comprehensive review on anticancer mechanisms of the main carotenoid of saffron, crocin [J]. *J Pharm Pharmacol*, 2017, 69(11): 1419-1427.
- [15] Chatterjee S, Poduval T B, Tilak J C, et al. A modified, economic, sensitive method for measuring total antioxidant capacities of human plasma and natural compounds using Indian saffron (*Crocus sativus*) [J]. *Clin Chim Acta*, 2005, 352(1/2): 155-163.
- [16] Asdaq S M, Inamdar M N. Potential of *Crocus sativus* (saffron) and its constituent, crocin, as hypolipidemic and antioxidant in rats [J]. *Appl Biochem Biotechnol*, 2010, 162(2): 358-372.
- [17] Cote F F, Dufresne C C, Willermot C. Properties of a glucosyltransferase involved in crocin synthesis [J]. *Plant Sci*, 2000, 153: 55-56.
- [18] Moraga A R, Nohales P F, Perez J A, et al. Glucosylation of the saffron apocarotenoid crocetin by a glucosyltransferase isolated from *Crocus sativus* stigmas [J]. *Planta*, 2004, 219(6): 955-966.
- [19] Nagatoshi M, Terasaka K, Owaki M, et al. UGT75L6 and UGT94E5 mediate sequential glucosylation of crocetin to crocin in *Gardenia jasminoides* [J]. *Febs Lett*, 2012, 586(7): 1055-1061.
- [20] Shigeyuki N, Satoko T, Yutaka O, et al. Cloning and characterization of glucosyltransferase cDNA from *Eucalyptus perriniana* cultured cells [J]. *Plant Biotech*, 2004, 21(5): 343-348.
- [21] Hans J, Brandt W, Vogt T, et al. Site-directed mutagenesis and protein 3D-homology modelling suggest a catalytic mechanism for UDP-glucose-dependent betanidin 5-O-glucosyltransferase from *Dorotheanthus bellidiformis* [J]. *Plant J*, 2004, 39(3): 319-333.
- [22] Ghose K, Selvaraj K, McCallum J, et al. Identification and functional characterization of a flax UDP-glycosyltransferase glucosylating secoisolariciresinol (SECO) into secoisolariciresinol monoglucoside (SMG) and diglucoside (SDG) [J]. *BMC Plant Biol*, 2014, 14(3): 82-98.
- [23] Osmani S A, Bak S, Imbert A, et al. Catalytic key amino acids and UDP-sugar donor specificity of a plant glucuronosyltransferase, UGT94B1: Molecular modeling substantiated by site-specific mutagenesis and biochemical analyses [J]. *Plant Physiol*, 2008, 148(3): 1295-1308.
- [24] Mageroy M H, Jancsik S, Man S Y M, et al. A conifer UDP-sugar dependent glycosyltransferase contributes to acetophenone metabolism and defense against insects [J]. *Plant Physiol*, 2017, 175(2): 641-651.