

青稞 *C4H* 基因的克隆及组织表达分析

罗小娇, 刘新春, 杨晓云, 袁金娥, 冯宗云

(四川农业大学农学院植物遗传育种学系大麦研究中心, 成都 611130)

摘要: 为揭示大麦中黄酮合成的分子调控机制, 利用反转录 PCR 结合同源克隆和 RACE 技术首次从青稞(裸大麦)叶片中克隆获得肉桂酸-4-羟化酶基因(*HvC4H*)的全长 cDNA 序列(Genbank 登录号: KF927086), 总长度 1951 bp, ORF 为 1518 bp, 编码 505 个氨基酸, 等电点 $PI=9.01$, 平均亲水指数(GRAVY)为 -0.170 , 属于亲水性碱性蛋白。高级结构分析表明其具有细胞色素 P450 家族保守域及 *C4H* 特异的功能性活性位点。利用实时荧光定量 PCR 分析胚乳发育 5 个时期不同组织(茎、叶及子粒)的表达情况, 结果显示 *HvC4H* 基因在青稞胚乳发育期的表达情况存在着明显的组织差异性, 在茎中的表达量最高。本研究为通过调控 *C4H* 基因的表达从而提高大麦黄酮的含量奠定了分子生物学基础, 对于改良大麦的品质、抗性、生长发育等性状具有重要意义。

关键词: 青稞(裸大麦); *C4H* 基因; 同源克隆; RACE; 实时荧光定量 PCR; 胚乳发育期

Cloning and Tissue Expression Analysis of *HvC4H* Gene in Hulless Barley

LUO Xiao-jiao, LIU Xin-chun, YANG Xiao-yun, YUAN Jin-e, FENG Zong-yun

(Barley Research Centre, Department of Plant Genetics and Breeding, College of Agronomy,
Sichuan Agricultural University, Chengdu 611130)

Abstract: In order to reveal the molecular regulatory mechanisms of flavonoids biosynthesis in barley (*Hordeum vulgare*), the full length cDNA (Genbank accession number: KF927086) encoding cinnamate-4-hydroxylase gene (*HvC4H*) was cloned from leaves of Qingke (hulless barley) through the methods of RT-PCR combining homologous clone and RACE technologies. The full length of *HvC4H* was 1951 bp and open reading frame was 1518 bp encoding 505 amino acids. Isoelectric point was 9.01 and grand average of hydropathicity (GRAVY) was -0.170 , which meant the *HvC4H* was hydrophilic and alkaline protein. Advanced structure analysis showed that *HvC4H* gene included family CYP450 conserved domains and specific functional active sites. The expression of *HvC4H* gene was analyzed in different tissues (stem, leaf and kernel) during 5 periods in endosperm development by real-time-fluorescence quantitative PCR. The results showed that obvious differences and the expression level in stem was pre-dominant. The study could lay molecular foundation to raise the content of flavonoid in barley by regulating the expression of *HvC4H* gene and provided useful informations to improve quality, resistibility and growth of barley.

Key words: Qingke (hulless barley); cinnamate-4-hydroxylase (*C4H*) gene; homologous cloning; rapid amplification of cDNA end (RACE); real-time-fluorescence quantitative PCR; endosperm development

青稞 (*Hordeum vulgare* L. var. *nudum* Hook. f.) 即裸大麦, 是我国青藏高原地区重要的种质资源^[1], 富含黄酮、 γ -氨基丁酸等功能成分, 具有较高的营养价

值和医疗保健作用。黄酮类物质在植物体内大部分以糖苷形式存在, 对植物的生长发育、抗病性、抗逆性及品质形成等有重要作用^[2-3], 对人体有抗氧化防

收稿日期: 2014-01-07 修回日期: 2014-03-03 网络出版日期: 2014-04-08

URL: <http://www.cnki.net/kcms/detail/11.4996.S.20140408.0845.013.html>

基金项目: 国家现代农业产业技术体系建设专项基金(CARS-05)

第一作者研究方向为大麦遗传育种与分子生物学。E-mail: ruyanruer@163.com

通信作者: 冯宗云, 研究方向为大麦遗传育种与分子生物学。E-mail: zyfeng49@126.com

衰老、抑制肿瘤及防治心血管疾病等保健功效^[4-5]。

黄酮类物质在植物体内的合成途径是:葡萄糖经过莽草酸途径和乙酸-丙二酸途径生成苯丙氨酸,经苯丙氨酸代谢途径生成黄酮、木质素、花青素等苯类化合物^[6]。苯丙氨酸代谢途径是植物最重要的次生代谢途径之一,多种含苯丙烷骨架的次生代谢产物都是由这一途径直接或者间接生成的^[7]。

肉桂酸-4-羟化酶(*C4H*, cinnamate-4-hydroxylase, EC1.14.13.11), 又称反式肉桂酸-4-单氧化酶^[8], 是苯丙氨酸代谢途径的第2个关键酶, 将苯丙氨酸途径的合成酶复合物固定在内质网上^[9], 在氧和 NADPH 存在下, 催化苯丙氨酸途径的第2步反应: 将反式肉桂酸苯环的4位羟基化生成对-香豆酸^[10]。*C4H* 不仅是植物细胞色素 P450 单氧化酶 CYP73 亚家族中第1个被鉴定出来的酶^[11], 也是第1个被克隆并确定了功能的植物 P450 酶, 在植物各组织中均具有很高的活性^[12-14], 研究表明 *C4H* 的蛋白质活性和转录丰度直接影响植物中黄酮类化合物和木质素的生物合成量^[15-17]。*C4H* 基因目前已从拟南芥、大豆、花生等50多种不同植物中克隆获得。目前大麦 *C4H* 基因克隆及表达分析的报道较少, GenBank 中还没有大麦 *C4H* 基因全长登录。本研究以前期实验室筛选的高黄酮青稞品种“94-19-1”为材料, 首次克隆获得了青稞 *HvC4H* 基因的 cDNA 全长, 同源比对后对序列进行了系统的生物信息学分析和进化树分析, 并通过实时荧光定量 PCR 技术检测了胚乳发育各时间段不同组织中 *HvC4H* 的表达水平。为深入研究 *C4H* 基因在大麦黄酮代谢中的分子调控机制提供依据, 为通过调控 *C4H* 基因的表达从而提高大麦黄酮的含量奠定基础, 对于改良大麦的品质、抗性、生长发育等性状具有重要意义, 并且有助于进一步阐明禾本科植物 *C4H* 基因的分子进化。

1 材料与方法

1.1 材料和主要试剂

1.1.1 材料 青稞“94-19-1”由甘肃省甘南藏族自治州农科所馈赠, 种植于四川省崇州市羊马镇国家大麦青稞产业体系成都综合试验站试验基地。

1.1.2 主要试剂 柱式植物 RNAout 试剂盒购自北京天恩泽基因工程有限公司; 反转录试剂盒 RevertAidTM First Strand cDNA Synthesis Kit 购自 Fermentas 公司; SMARTerTM RACE cDNA Amplification Kit 购自 Clontech 公司; SYBR[®] Premix Ex TaqTM II 试剂盒购自宝生物工程(大连)有限公司。

1.2 方法

1.2.1 RNA 提取和 cDNA 第一链合成 以柱式植物 RNAout 试剂盒分别提取材料3叶1心期的叶片总 RNA 以及胚乳发育5个时期(开花后第2、4、8、12、16天)茎、叶和子粒^[18]样品的 RNA; 用 RevertAidTM First Strand cDNA Synthesis Kit 反转录得到相应 cDNA; RACE 中采用 SMARTerTM RACE cDNA Amplification 试剂盒分别合成3'和5'末端 cDNA。

1.2.2 同源克隆 参照 GenBank 登录的稻属、高粱属等单子叶禾本科植物的17条 *C4H* 基因序列, 通过多序列比对在保守区域设计简并引物 JC4Hf、JC4Hr(表1)。以青稞叶片 cDNA 为模板进行 PCR 反应。反应体系 30 μ L: cDNA 1 μ L, JC4Hf (10 μ mol/L)、JC4Hr (10 μ mol/L) 各 1 μ L, 2 $\times$ Taq PCR Master Mix 15 μ L, ddH₂O 12 μ L。反应条件: 94 $^{\circ}$ C 50 s, 60 $^{\circ}$ C 50 s, 72 $^{\circ}$ C 1.25 min, 32 cycles。电泳分析后回收目的片段、亚克隆和测序, 将测序结果用 DNAMAN 6.0 比对分析后得到保守区序列。

1.2.3 RACE 克隆 3'-RACE 根据保守序列设计 3'-RACE 上游引物: 3'-GSP、3'-NGSP1、3'-NGSP2(表1)。以 3'-RACE-Ready cDNA 为模板进行第一轮 PCR 反应 (30 μ L): 3' cDNA 1.5 μ L, 3'-GSP (10 μ mol/L) 0.6 μ L, 10 $\times$ UPM (Universal Primer Mix) 3 μ L, Mix 15 μ L, ddH₂O 9.9 μ L; 反应条件: 94 $^{\circ}$ C 30 s, 70 $^{\circ}$ C 30 s, 72 $^{\circ}$ C 3 min, 5 cycles; 94 $^{\circ}$ C 30 s, 68 $^{\circ}$ C 30 s, 72 $^{\circ}$ C 3 min, 25 cycles。以稀释50倍的第1轮 PCR 产物为模板进行第2轮巢式 PCR (30 μ L): 50 $\times$ PCR1 1.5 μ L, 3'-NGSP1 (10 μ mol/L) 0.6 μ L, UPM 3 μ L, Mix 15 μ L, ddH₂O 9.9 μ L; 反应条件同第1轮。再以稀释50倍的第2轮 PCR 产物为模板进行第3轮巢式 PCR (30 μ L): 50 $\times$ PCR2 2 μ L, 3'-NGSP2 (10 μ mol/L) 0.6 μ L, UPM 3 μ L, Mix 15 μ L, ddH₂O 9.4 μ L; 反应条件同第1轮。电泳分析后回收目的片段、亚克隆和测序。

5'-RACE 根据保守序列设计 5'-RACE 下游引物: 5'-GSP、5'-NGSP1、5'-NGSP2(表1)。以 5'-RACE-Ready cDNA 为模板进行第1轮 PCR 反应 (30 μ L): 5' cDNA 1.5 μ L, 5'-GSP (10 μ mol/L) 0.6 μ L, 10 $\times$ UPM 3 μ L, Mix 15 μ L, ddH₂O 9.9 μ L; 反应条件: 94 $^{\circ}$ C 30 s, 72 $^{\circ}$ C 1.5 min, 5 cycles; 94 $^{\circ}$ C 30 s, 68 $^{\circ}$ C 30 s, 72 $^{\circ}$ C 1.5 min, 5 cycles; 94 $^{\circ}$ C 30 s, 65 $^{\circ}$ C 30 s, 72 $^{\circ}$ C 1.5 min, 25 cycles。以稀释500倍的第1轮 PCR 产物为模板进行第2轮巢式 PCR (30 μ L): 500 $\times$ PCR1 1 μ L, 5'-NGSP1 (10 μ mol/L) 0.6 μ L,

UPM 3 μL , Mix 15 μL , ddH₂O 10.4 μL ; 反应条件同第 1 轮。再以稀释 500 倍的第 2 轮 PCR 产物为模板进行第 3 轮巢式 PCR (30 μL): 500 $\times$ PCR2 1 μL , 5'-NGSP2 (10 $\mu\text{mol/L}$) 0.6 μL , UPM 3 μL , Mix 15 μL , ddH₂O 10.4 μL ; 反应条件同第 1 轮。电泳分析后回收目的片段、亚克隆和测序。

1.2.4 全长 将 3 批测序结果用 DNAMAN 拼接后经比对分析后得到 *HvC4H* 基因的 cDNA 全长序列, 提交 Genbank (登录号: KF927086)。

1.2.5 生物信息学分析 在 NCBI 数据库 Blast 进行多序列比对, 用 DNAMAN 6.0 推导氨基酸序列及进行同源比对, 分别利用瑞典生物信息研究所的 ExPASy 服务器的 Prot Param 在线软件和 SOPMA 软件对其酶蛋白的一、二级结构进行预测, 用 PHYRE2 Protein Fold Recognition Server (<http://www.sbg.bio.ic.ac.uk/phyre2>) 和 3D LigandSite server 预测三级结构模型, 用 MEGA 5.0 构建系统进化树。

1.2.6 胚乳发育期表达分析 以胚乳发育 5 个时期茎、叶、子粒的第一链 cDNA 为模板, 用 Bio-Rad CFX96 型荧光定量 PCR 仪进行 RTF-qPCR 反应。在靠近 5'-UTR 的 *HvC4H* 保守区域设计特异引物 qPCR-f、qPCR-r (表 1); 以大麦肌动蛋白基因 (β -actin) 和微管蛋白基因 (α -tubulin) 作为内参基因^[19] 根据其序列设计特异引物 actin-f、actin-r, tublin-f、tublin-r (表 1) 反应体系 (10 μL): cDNA 1 μL , SYBR Premix Ex Taq 5 μL , 上下游引物 (10 $\mu\text{mol/L}$) 各 0.3 μL , ddH₂O 3.4 μL 。反应条件: 95 $^{\circ}\text{C}$ 30 s; 95 $^{\circ}\text{C}$ 10 s 60 $^{\circ}\text{C}$ 30 s 40 cycles; 95 $^{\circ}\text{C}$ 15 s 60 $^{\circ}\text{C}$ 10 s 95 $^{\circ}\text{C}$ ∞ 。每个样品重复 3 次。

表 1 引物序列设计

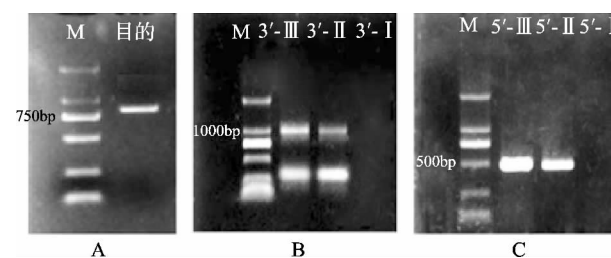
Table 1 Primer sequences designed

引物名称 Primer name	引物序列 (5'-3') Primer sequence	引物长度 (bp) Primer length
JC4Hf	CCGCCGGGCCCCATCSSLVRYBCCSATCKTC	30
JC4Hr	GTCTCGATCGCGGCGACRTTAAATGTTYTC	31
3'-GSP	ACCCGCCTCAAGCTCTTCAAGGATTAC	27
3'-NGSP1	CAGAGCAGAAGGGGAGATCAACGAGG	27
3'-NGSP2	GAGACAATCGTCAACGGCAGGCGT	25
5'-GSP	ACTGGTCGCGGTACACGGTGAAC	23
5'-NGSP1	ACCCGAACCTCCACGCCCTGC	20
5'-NGSP2	TGCGGTGGTTTCAGGTCGTCG	21
qPCR-f	TGCAAGGAGGTCAAGGAGAC	20
qPCR-r	CGACGTTGATGTTCTCGATG	20
actin-f	TCACGCTCAAGTACCCATCGA	22
actin-r	GGAGCTGTCTTGGCAGTCTCCA	23
tublin-f	CTCCATGATGGCCAAAGTGTG	20
tublin-r	TGGAGATCATGCACACAGCC	20

2 结果与分析

2.1 青稞 *HvC4H* 基因全长序列获得

使用简并引物 JC4Hf 和 JC4Hr 进行同源克隆获得约 800 bp 的特异片段 (图 1A), 测序分析后得到长为 733 bp 的保守片段, 用 DNAMAN 比对证实与其他植物 *C4H* 基因的保守区具有较高的同源性。使用巢式引物进行 3'-RACE 和 5'-RACE, 分别得到约 1000 bp (图 1B) 和 500 bp 的特异条带 (图 1C)。综合测序结果, 比对拼接后得到了长为 1951 bp 的青稞 *C4H* 基因 (*HvC4H*) 的 cDNA 序列, 其 ORF 为 1518 bp (125 ~ 1642), 5'-UTR 为 124 bp, 3'-UTR 为 309 bp 并包含 27 bp 的 Poly A 尾。



A: 同源克隆产物; B: 3'-RACE 三轮产物;

C: 5'-RACE 三轮产物; M: 标记 DL2000

A: Product of homologous clone; B: 3 PCR products of 3'-RACE;

C: 3 PCR products of 5'-RACE; M: DNA marker DL2000

图 1 青稞 *HvC4H* 基因扩增产物电泳图

Fig. 1 Amplified product of *HvC4H* gene in hulless barley

2.2 生物信息学分析

2.2.1 一、二级结构 将青稞 *C4H* 基因全长 cDNA 序列在 NCBI 中进行 Blastn 比对, 与二穗短柄草 (*B. distachyon*, XM003564447.1) 相似度达 94%, 与慈竹 (*N. affinis*, JN571418.1)、粟米 (*S. italica*, XM_004970328.1)、绿竹笋 (*B. oldhamii*, GU188741.1)、柳枝稷 (*P. virgatum*, JX845713.1) 等的相似度达 92%, 与玉米 (*Z. mays*, NM_001155686.1)、高粱 (*S. bicolor*, AY034143.1)、水稻 (*O. sativa*, AB207105.1) 等的相似度都达到 82%。利用 DNAMAN 推导出 *HvC4H* 基因的蛋白序列, 选取来自单子叶和双子叶植物的 7 个 *C4H* 蛋白序列进行同源比对, 同源性高达 88.53% (图 2)。

DNAMAN 的氨基酸序列分析结果表明, 该酶蛋白共编码 505 个氨基酸, 分子量为 57.984 kD, 等电点 (PI) 为 9.01, 是碱性蛋白。氨基酸组成中亮氨酸 (Leu) 最多 (11.5%), 半胱氨酸 (Cys) 最少 (0.6%), 该蛋白质中具有负电荷的氨基酸 (Asp +

青稞HORDEUM VULGARE	MEFVFEKLVGLLAAYVAVSVSKIRGRKRLPPGHLEVPFGNWLQVGGDLNHRNLAA	60
二穗短柄草BRACHYPODIUM DISTACHYON	MEFVFEKLVGLLAAYVAVSVSKIRGRKRLPPGHLEVPFGNWLQVGGDLNHRNLAA	60
慈竹NEOSINOCALAMUS AFFINIS	MEFVFEKLVGLLAAYVAVSVSKIRGRKRLPPGHLEVPFGNWLQVGGDLNHRNLAA	60
粟米SETARIA ITALICA	MEFVFEKLVGLLAAYVAVSVSKIRGRKRLPPGHLEVPFGNWLQVGGDLNHRNLAA	60
拟南芥ARABIDOPSIS THALIANA	MEFVFEKLVGLLAAYVAVSVSKIRGRKRLPPGHLEVPFGNWLQVGGDLNHRNLVD	60
油菜BRASSICA NAPUS	MEFVFEKLVGLLAAYVAVSVSKIRGRKRLPPGHLEVPFGNWLQVGGDLNHRNLVD	60
大豆GLYCINE MAX	MEFVFEKLVGLLAAYVAVSVSKIRGRKRLPPGHLEVPFGNWLQVGGDLNHRNLVD	60
水稻ORYZA SATIVA	MEFVFEKLVGLLAAYVAVSVSKIRGRKRLPPGHLEVPFGNWLQVGGDLNHRNLAA	60
青稞HORDEUM VULGARE	YARFEGEILLRMGCRNLVVSSFFIAREVLHTQGVFGSRTNRVVFDFITGGQCMVFT	120
二穗短柄草BRACHYPODIUM DISTACHYON	YARFEGEILLRMGCRNLVVSSFFIAREVLHTQGVFGSRTNRVVFDFITGGQCMVFT	120
慈竹NEOSINOCALAMUS AFFINIS	YARFEGEILLRMGCRNLVVSSFFIAREVLHTQGVFGSRTNRVVFDFITGGQCMVFT	120
粟米SETARIA ITALICA	YARFEGEILLRMGCRNLVVSSFFIAREVLHTQGVFGSRTNRVVFDFITGGQCMVFT	120
拟南芥ARABIDOPSIS THALIANA	YARFEGEILLRMGCRNLVVSSFFIAREVLHTQGVFGSRTNRVVFDFITGGQCMVFT	120
油菜BRASSICA NAPUS	YARFEGEILLRMGCRNLVVSSFFIAREVLHTQGVFGSRTNRVVFDFITGGQCMVFT	120
大豆GLYCINE MAX	YARFEGEILLRMGCRNLVVSSFFIAREVLHTQGVFGSRTNRVVFDFITGGQCMVFT	120
水稻ORYZA SATIVA	YARFEGEILLRMGCRNLVVSSFFIAREVLHTQGVFGSRTNRVVFDFITGGQCMVFT	120
青稞HORDEUM VULGARE	VYGHNRKMRIMTVFFFTNKVVQCNHGWDEAAFWVGNRADEPFAATGVRHRLQL	180
二穗短柄草BRACHYPODIUM DISTACHYON	VYGHNRKMRIMTVFFFTNKVVQCNHGWDEAAFWVGNRADEPFAATGVRHRLQL	180
慈竹NEOSINOCALAMUS AFFINIS	VYGHNRKMRIMTVFFFTNKVVQCNHGWDEAAFWVGNRADEPFAATGVRHRLQL	180
粟米SETARIA ITALICA	VYGHNRKMRIMTVFFFTNKVVQCNHGWDEAAFWVGNRADEPFAATGVRHRLQL	180
拟南芥ARABIDOPSIS THALIANA	VYGHNRKMRIMTVFFFTNKVVQCNHGWDEAAFWVGNRADEPFAATGVRHRLQL	180
油菜BRASSICA NAPUS	VYGHNRKMRIMTVFFFTNKVVQCNHGWDEAAFWVGNRADEPFAATGVRHRLQL	180
大豆GLYCINE MAX	VYGHNRKMRIMTVFFFTNKVVQCNHGWDEAAFWVGNRADEPFAATGVRHRLQL	180
水稻ORYZA SATIVA	VYGHNRKMRIMTVFFFTNKVVQCNHGWDEAAFWVGNRADEPFAATGVRHRLQL	180
青稞HORDEUM VULGARE	MMYNMGRIMFDRRFSEEDFHFQIRANNGERSRLQSFYNYGDFIFLRFFLRGYK	240
二穗短柄草BRACHYPODIUM DISTACHYON	MMYNMGRIMFDRRFSEEDFHFQIRANNGERSRLQSFYNYGDFIFLRFFLRGYK	240
慈竹NEOSINOCALAMUS AFFINIS	MMYNMGRIMFDRRFSEEDFHFQIRANNGERSRLQSFYNYGDFIFLRFFLRGYK	240
粟米SETARIA ITALICA	MMYNMGRIMFDRRFSEEDFHFQIRANNGERSRLQSFYNYGDFIFLRFFLRGYK	240
拟南芥ARABIDOPSIS THALIANA	MMYNMGRIMFDRRFSEEDFHFQIRANNGERSRLQSFYNYGDFIFLRFFLRGYK	240
油菜BRASSICA NAPUS	MMYNMGRIMFDRRFSEEDFHFQIRANNGERSRLQSFYNYGDFIFLRFFLRGYK	240
大豆GLYCINE MAX	MMYNMGRIMFDRRFSEEDFHFQIRANNGERSRLQSFYNYGDFIFLRFFLRGYK	240
水稻ORYZA SATIVA	MMYNMGRIMFDRRFSEEDFHFQIRANNGERSRLQSFYNYGDFIFLRFFLRGYK	240
青稞HORDEUM VULGARE	LCHEVRETRRLKDYVLDERRKLASTRAMNNG.LKCAQHILDAQCKGEINDNLVI	299
二穗短柄草BRACHYPODIUM DISTACHYON	LCHEVRETRRLKDYVLDERRKLASTRAMNNG.LKCAQHILDAQCKGEINDNLVI	299
慈竹NEOSINOCALAMUS AFFINIS	LCHEVRETRRLKDYVLDERRKLASTRAMNNG.LKCAQHILDAQCKGEINDNLVI	299
粟米SETARIA ITALICA	LCHEVRETRRLKDYVLDERRKLASTRAMNNG.LKCAQHILDAQCKGEINDNLVI	299
拟南芥ARABIDOPSIS THALIANA	LCHEVRETRRLKDYVLDERRKLASTRAMNNG.LKCAQHILDAQCKGEINDNLVI	299
油菜BRASSICA NAPUS	LCHEVRETRRLKDYVLDERRKLASTRAMNNG.LKCAQHILDAQCKGEINDNLVI	299
大豆GLYCINE MAX	LCHEVRETRRLKDYVLDERRKLASTRAMNNG.LKCAQHILDAQCKGEINDNLVI	300
水稻ORYZA SATIVA	LCHEVRETRRLKDYVLDERRKLASTRAMNNG.LKCAQHILDAQCKGEINDNLVI	295
青稞HORDEUM VULGARE	MEINVAAIETTLWSIEWGIRLVNHHIQCSHNDGMAELGCHQCTESGTHSLFYLQA	359
二穗短柄草BRACHYPODIUM DISTACHYON	MEINVAAIETTLWSIEWGIRLVNHHIQCSHNDGMAELGCHQCTESGTHSLFYLQA	359
慈竹NEOSINOCALAMUS AFFINIS	MEINVAAIETTLWSIEWGIRLVNHHIQCSHNDGMAELGCHQCTESGTHSLFYLQA	359
粟米SETARIA ITALICA	MEINVAAIETTLWSIEWGIRLVNHHIQCSHNDGMAELGCHQCTESGTHSLFYLQA	359
拟南芥ARABIDOPSIS THALIANA	MEINVAAIETTLWSIEWGIRLVNHHIQCSHNDGMAELGCHQCTESGTHSLFYLQA	359
油菜BRASSICA NAPUS	MEINVAAIETTLWSIEWGIRLVNHHIQCSHNDGMAELGCHQCTESGTHSLFYLQA	359
大豆GLYCINE MAX	MEINVAAIETTLWSIEWGIRLVNHHIQCSHNDGMAELGCHQCTESGTHSLFYLQA	360
水稻ORYZA SATIVA	MEINVAAIETTLWSIEWGIRLVNHHIQCSHNDGMAELGCHQCTESGTHSLFYLQA	354
青稞HORDEUM VULGARE	VKKTILRLMAIPILVPHMNLIDKLAGMDFAESKILVNANLANDEAKKHEERFE	419
二穗短柄草BRACHYPODIUM DISTACHYON	VKKTILRLMAIPILVPHMNLIDKLAGMDFAESKILVNANLANDEAKKHEERFE	419
慈竹NEOSINOCALAMUS AFFINIS	VKKTILRLMAIPILVPHMNLIDKLAGMDFAESKILVNANLANDEAKKHEERFE	419
粟米SETARIA ITALICA	VKKTILRLMAIPILVPHMNLIDKLAGMDFAESKILVNANLANDEAKKHEERFE	419
拟南芥ARABIDOPSIS THALIANA	VKKTILRLMAIPILVPHMNLIDKLAGMDFAESKILVNANLANDEAKKHEERFE	419
油菜BRASSICA NAPUS	VKKTILRLMAIPILVPHMNLIDKLAGMDFAESKILVNANLANDEAKKHEERFE	419
大豆GLYCINE MAX	VKKTILRLMAIPILVPHMNLIDKLAGMDFAESKILVNANLANDEAKKHEERFE	420
水稻ORYZA SATIVA	VKKTILRLMAIPILVPHMNLIDKLAGMDFAESKILVNANLANDEAKKHEERFE	414
青稞HORDEUM VULGARE	RSEEDFVEANGNDEPILPFGVGRSSCPGIIALPIGITHGRVQFELPPFGQDIL	479
二穗短柄草BRACHYPODIUM DISTACHYON	RSEEDFVEANGNDEPILPFGVGRSSCPGIIALPIGITHGRVQFELPPFGQDIL	479
慈竹NEOSINOCALAMUS AFFINIS	RSEEDFVEANGNDEPILPFGVGRSSCPGIIALPIGITHGRVQFELPPFGQDIL	479
粟米SETARIA ITALICA	RSEEDFVEANGNDEPILPFGVGRSSCPGIIALPIGITHGRVQFELPPFGQDIL	479
拟南芥ARABIDOPSIS THALIANA	RSEEDFVEANGNDEPILPFGVGRSSCPGIIALPIGITHGRVQFELPPFGQDIL	479
油菜BRASSICA NAPUS	RSEEDFVEANGNDEPILPFGVGRSSCPGIIALPIGITHGRVQFELPPFGQDIL	479
大豆GLYCINE MAX	RSEEDFVEANGNDEPILPFGVGRSSCPGIIALPIGITHGRVQFELPPFGQDIL	480
水稻ORYZA SATIVA	RSEEDFVEANGNDEPILPFGVGRSSCPGIIALPIGITHGRVQFELPPFGQDIL	474
青稞HORDEUM VULGARE	DTREKGGQFSMLIRKSTLVKKEVF.	505
二穗短柄草BRACHYPODIUM DISTACHYON	DTREKGGQFSMLIRKSTLVKKEVF.	505
慈竹NEOSINOCALAMUS AFFINIS	DTREKGGQFSMLIRKSTLVKKEVF.	506
粟米SETARIA ITALICA	DTREKGGQFSMLIRKSTLVKKEVF.	505
拟南芥ARABIDOPSIS THALIANA	DTREKGGQFSMLIRKSTLVKKEVF.	505
油菜BRASSICA NAPUS	DTREKGGQFSMLIRKSTLVKKEVF.	505
大豆GLYCINE MAX	DTREKGGQFSMLIRKSTLVKKEVF.	506
水稻ORYZA SATIVA	DTREKGGQFSMLIRKSTLVKKEVF.	500

✱ C4H 的功能性氨基酸残基

✱ functional amino acids of C4H

图2 青稞 HvC4H 蛋白序列的同源比对

Fig. 2 Homologous alignment of HvC4H protein sequence

Glu) 为 61 个,正电荷氨基酸(Arg + Lys) 为 67 个。该酶蛋白的不稳定系数大于 40(43. 23) ,属于不稳定蛋白,其平均亲水指数(GRAVY ,grand average of hydropathicity) 为 -0. 170 ,属于亲水性蛋白,预测半衰期为 30 h。

利用 PHYRE2 Server 并结合 SOPMA 软件预测二级结构,预测结果(图 3) 表明,该酶蛋白的二级结构主要由 α -螺旋(50. 1%)、 β 股(约 8%)、无规则卷曲(33. 86%)、延伸链(12. 4%) 及少量 β -转角(3. 37%) 组成。

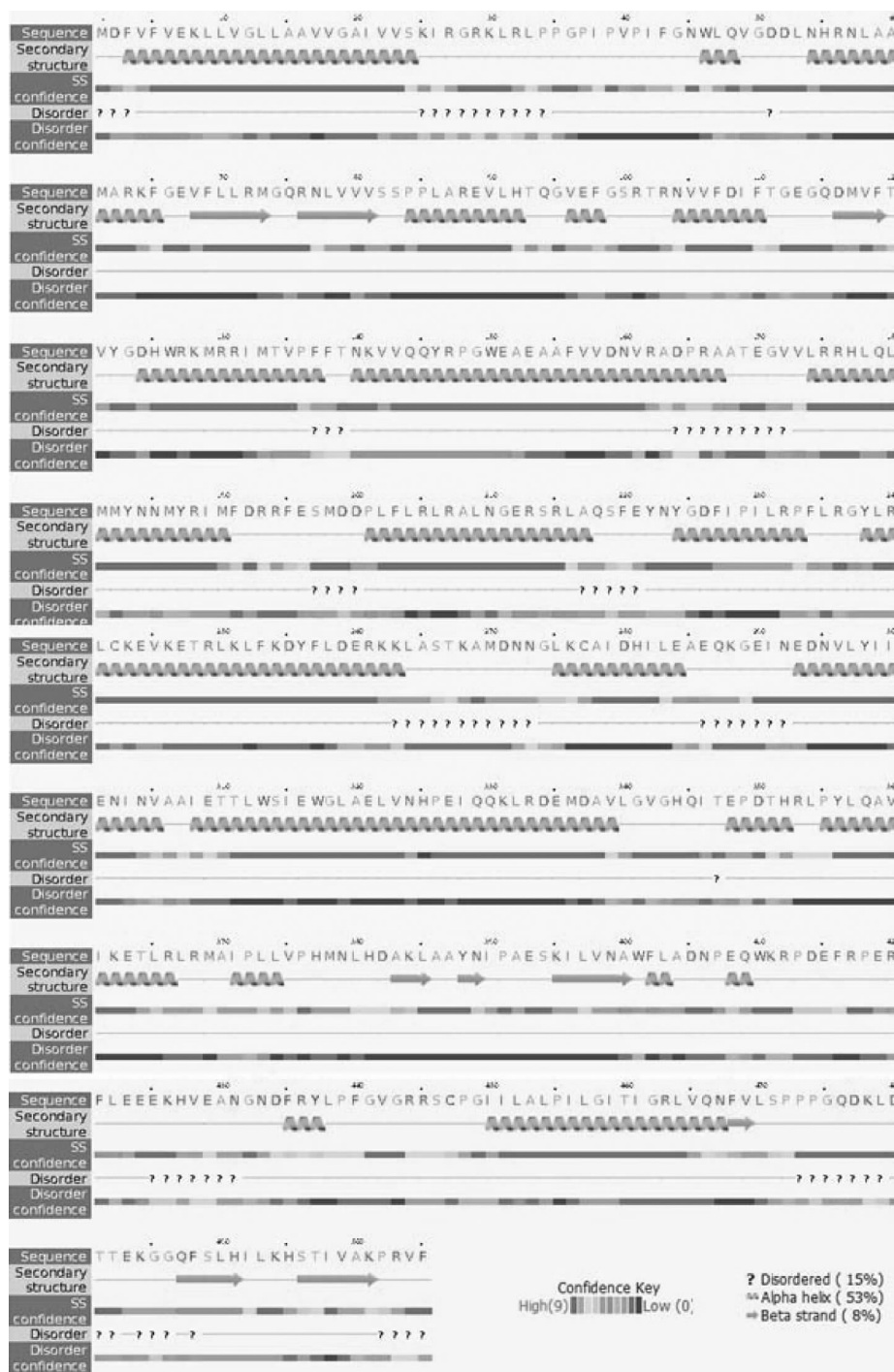


图3 PHYRE2 二级结构预测

Fig. 3 Secondary structure prediction by PHYRE 2

为了进一步对高级结构功能进行研究,利用丹麦 CBS 研究所中 NetNGly 和 NetPho 2.0 进行蛋白质的 N-糖基化和 N-磷酸化预测,没有发现糖基化位点,却存在着丰富的磷酸化位点,共包括 7 个 Ser、6 个 Thr 和 3 个 Tyr 存在磷酸化活性位点(图 4)。使用 SignalP-4.1 进行信号肽预测,信号肽存在机率较低(0.313),锚定机率为 0.265,将其和 TargetP-1.1

的分泌情况预测相结合,发现该酶蛋白为分泌性蛋白。利用 Clavious 大学的蛋白质二硫键预测软件(<http://clavious.bc.edu/~clotelab/DiANNA/>)分析表明,在 Cys²⁷⁷ 和 Cys⁴⁷⁷ 间可能存在一个二硫键。访问日本 PSORT II Prediction 网站和在线分析软件 TMHMM,对 HvC4H 蛋白进行亚细胞定位,确定该酶蛋白定位于内质网外膜上。

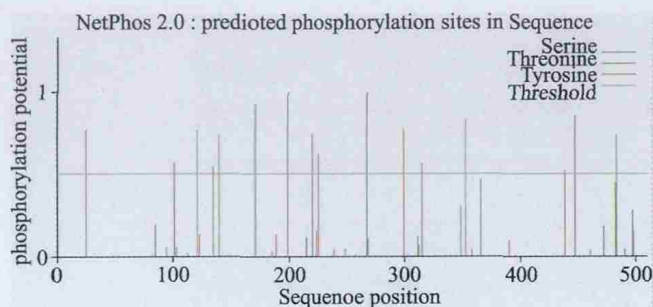


图4 NetPhos 2.0 磷酸化位点预测

Fig. 4 Predicted phosphorylation sites prediction by NetPhos 2.0

2.2.2 三级结构 用 PHYRE 2 及伦敦帝国理工学院 (Imperial College London) 的 3D LigandSite server 进行 HvC4H 酶蛋白的三级结构及活性中心的预测, 预测结果表明, 91% 的 HvC4H 酶蛋白序列 (460 个氨基酸残基) 与 Phyre2 数据库中 CYP450 家族各蛋白序列最高达到了 100% 的相似度。以 CYP450 家族 d1nr6a 编码蛋白的三级结构为模板构建出 HvC4H 酶蛋白三级结构 (图 5), 该三级结构是典型的 P450 家族球状蛋白, 主要包含 α -domain 和 β -domain 两个结构域, 预测其中主要位于 α -domain 的血红素结合域 (heme binding domain) 及周围的底物结合域 (substrate recognition site) 构成了该酶蛋白的活性中心。

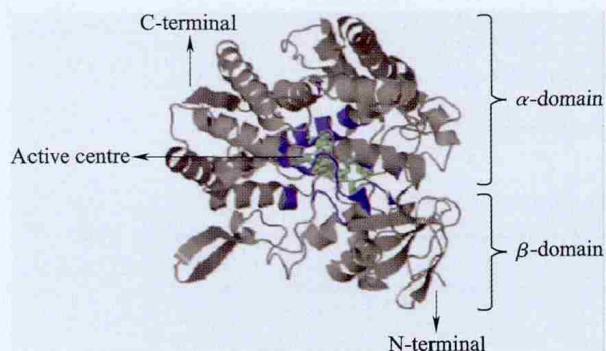


图5 HvC4H 的蛋白质三级结构及活性中心预测

Fig. 5 Protein tertiary structure and active centre predictions of HvC4H protein

利用 PHYRE 2 的在线分析, 采用拓扑学对跨膜螺旋 (transmembrane helix) 进行了预测 (图 6), 将预测结果与 PHYRE 2 的二级结构预测比较分析后可知, HvC4H 酶蛋白很可能在 V⁶-S²⁴、H¹⁷⁷-D¹⁹²、A²¹⁷-R²³² 处主要以 α 螺旋多次跨膜构成跨膜螺旋, 疏水区在脂双层内与脂类分子的疏水尾部相互作用, 亲水区则露在膜的两侧。

2.2.3 系统进化树分析 在同源比对的序列基础上, 再加入部分被子植物 C4H 的蛋白序列, 利用

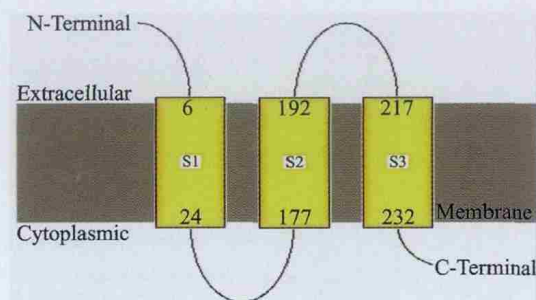


图6 跨膜螺旋预测

Fig. 6 Transmembrane helix prediction

MEGA 5.0 软件使用 Neighbor-joining 法重复 1000 次抽样构建系统进化树 (图 7), 该进化树基本可以分为 2 个亚簇: 双子叶植物 (cluster I) 和单子叶植物 (cluster II), 结果表明, 青稞与单子叶植物的山羊草、二穗短柄草以及禾本科单子叶的粟米、毛绿竹和慈竹等的亲缘关系最近, 相对双子叶植物如大豆等的亲缘关系较远。

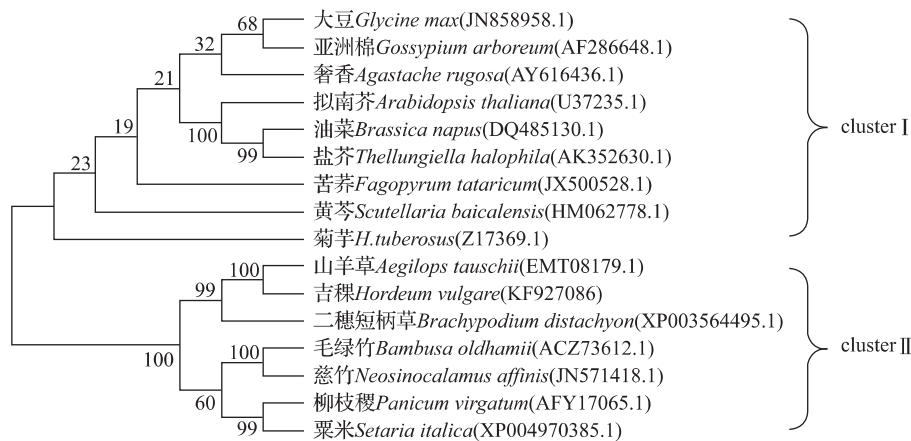
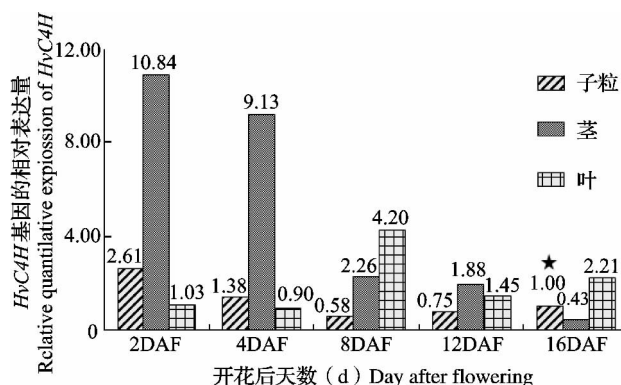
2.3 胚乳发育期的定量表达

分别以青稞胚乳发育 5 个时期的茎、叶、子粒的 first strand cDNA 为模板, 以 β -actin 和 α -tubulin 为内参, 采用 TaKaRa SYBR[®] Premix Ex Taq[™] II 试剂盒进行 RTF-qPCR。熔解曲线走势正常, 无杂峰及引物二聚体, 解链温度均一, 说明扩增产物具有极高的特异性; 通过标准曲线计算得到的扩增效率均 > 95%, 符合实时荧光定量 PCR 对引物扩增效率的要求。

用 $2^{-\Delta\Delta C_t}$ 法得到各个时期基因的表达量, 以开花后第 16 天子粒的表达量作为对照 (设定为 1), 对其他样本的表达量进行校正, 最后的数值代表所有样本相对于对照样本基因表达量的倍数差异。对青稞材料 94-19-1 在胚乳发育期 HvC4H 基因的表达量进行分析, 结果表明, HvC4H 基因在青稞胚乳发育期的表达情况存在着组织差异性 (图 8)。从茎部来看, 4 d 前 HvC4H 的表达量较高, 在 4 ~ 8 d 之间, HvC4H 的表达量有一个急剧下降的过程 (8 d 仅为 2 d 表达量的 1/5); 而在叶中, HvC4H 存在着一个在 4 d 前表达量较低、在 8 d 达到最高 (4.20) 然后又下降的过程; 在子粒部位, HvC4H 基因表达量处于较低水平, 2 d 子粒表达量最高, 但也不到同时期茎部的 1/4。总的看来, 在整个胚乳发育期, 前期茎中 HvC4H 基因的表达量较高, 而子粒中的表达则逐渐趋于稳定。

3 讨论

HvC4H 酶蛋白属于细胞色素 P450 亚族 CYP73 家族成员, CYP450 家族蛋白的保守域和特征基序

图 7 *C4H* 蛋白质序列的系统进化树Fig. 7 Phylogenetic tree of *C4H* protein sequences图 8 *HvC4H* 基因定量表达分析Fig. 8 Quantitative analysis of *HvC4H* by RTF-qPCR

也在 *HvC4H* 蛋白序列中得到印证(图 2),如血红素结合域“P⁴³⁹FGVGRRSCPG⁴⁴⁹”^[20];特征性底物识别位点(SRSs, substrate recognition sites)^[21-22];脯氨酸铰链结构(proline hinge)“P³⁴PGPIPVP⁴¹”^[23];T 结合槽基序 A³⁰⁶A IET T³¹¹以及 E³⁶³-R³⁶⁶-R⁴²⁰三联体^[9];另外, N-端的内质网锚定信号基序(endoplasmic reticulum signal-anchor sequence)“M¹-S²⁴”使 *C4H* 定位于内质网外膜上^[24],与 PHYRE 2 的跨膜螺旋预测结果相同,参与苯丙氨酸代谢中多酶复合体的亚细胞定位和内质网上电子链的化学传递^[25]。从同源比对结果可以看出,这些区域在单子叶、双子叶植物 *C4H* 蛋白的一级结构中都是高度保守的,对 *HvC4H* 形成准确的高级结构和与内质网外膜正确结合具有重要的意义。同时也发现参与了 *HvC4H* 与肉桂酸特异性识别和结合的氨基酸残基 Glu³⁰¹、Asn³⁰²、Val³⁰⁵、Ala³⁰⁶及 Thr³¹⁰^[26](图 2),这些功能性氨基酸残基参与高级结构所形成的活性中心^[27],使 *C4H* 在结构和功能上与其他 P450 基因区别开来^[28]。从三级结构预测可以看出, α -domain 含有丰

富的 α 螺旋结构, β -domain 由 β 折叠、 α 螺旋和一些无规则卷曲构成。同时,预测到的跨膜螺旋(图 6)对 *HvC4H* 酶蛋白在内质网上的定位有重要的指导意义。结合高级结构的预测结果可以看出, *HvC4H* 酶蛋白的活性中心处于内质网外膜内侧的 C-末端,结合血红素上的铁原子,在 NADPH 和 O₂ 的作用下进行苯丙氨酸代谢途径的第一个氧化反应:将反式肉桂酸羟基化生成对-香豆酸。此外,对 *HvC4H* 酶蛋白预测到丰富的磷酸化位点,磷酸化可能在 *HvC4H* 的活性调节上发挥作用,这还有待进一步探究。*C4H* 的进化树分析表明,其聚类方式总体上与传统的植物分类学相符,双、单子叶植物主要分为 cluster I 和 cluster II 两簇,可以看出 *C4H* 在禾本科内比较保守,但在不同科之间的保守性不高,较好地说明了 *C4H* 基因的分化路径。

已有研究发现,黄酮积累量与 *C4H* 基因表达量在黑莓果实发育的不同时期表现出同步增减的变化趋势^[29]。本研究定量分析表明, *HvC4H* 基因在青稞的茎、叶和子粒中都有表达,并且 *HvC4H* 基因在苯丙烷类物质富集的茎中^[30]的表达占明显优势。从青稞胚乳发育各个时期不同组织的表达量来看,除了 12 d 外, *HvC4H* 基因在同一时期不同部位表达差异较大(最高表达量约为最低表达量的 10 倍),基因相对表达量存在着明显的组织差异性。从总的变化趋势来看, *HvC4H* 在胚乳发育期呈逐渐下降的趋势,这可能与 *C4H* 处于苯丙氨酸代谢途径的上游^[31],受到其上游结构基因或转录因子和下游的代谢产物(对-香豆酸)的反馈抑制等方面的调控有关。

对代谢途径相关基因的研究可以增强对青稞种子性状形成分子机制的了解,克隆出青稞 *HvC4H* 基因的 cDNA 序列,对其进行系统的生物信息学分析

和比较全面的功能预测,有助于进一步研究青稞黄酮合成的分子调控机制和作用机理。胚乳发育期不同组织的差异性表达有利于进一步了解 *HvC4H* 在青稞生殖生长过程中调控黄酮合成的作用,为青稞功能品质分子育种奠定基础。

参考文献

- [1] 刘仙俊. 大麦特异种质资源的分子生物学研究[D]. 成都: 四川农业大学, 2011
- [2] Winkel-Shirley B. Flavonoid biosynthesis a colorful model for genetics, biochemistry, cell biology, and biotechnology [J]. Plant Physiol 2001, 126(2): 485-493
- [3] Harakava R. Genes encoding enzymes of the lignin in biosynthesis pathway in *Eucalyptus* [J]. Genet Mol Biol, 2005, 28(3S): 601-607
- [4] 邹浩军, 明亮. 黄酮类化合物抗炎及抗衰老药理研究进展[J]. 安徽卫生职业技术学院学报, 2003, 2(6): 48-50, 57
- [5] 农朝赞, 黄华艺. 黄酮类化合物抗肿瘤作用的研究应用[J]. 中国药房, 2004, 15(9): 568-569
- [6] Boerjan W, Ralph J, Baucher M. Lignin biosynthesis [J]. Annu Rev Plant Biol 2003, 54: 519-546
- [7] Li X H, Park Nam Il, Xu H, et al. Differential expression of flavonoid biosynthesis genes and accumulation of phenolic compounds in common buckwheat (*Fagopyrum esculentum*) [J]. J Agric Food Chem 2010, 58(23): 12176-12181
- [8] 王安娜, 王婵婵, 吴蕾, 等. 大豆 *C4H* 基因克隆及生物信息学分析[J]. 东北农业大学学报, 2010, 41(4): 12-15
- [9] 陈安和, 李加纳, 柴友荣, 等. 羽衣甘蓝中一个突变型肉桂酸-4-羟化酶基因的克隆及分析[J]. 园艺学报, 2007, 34(4): 915-922
- [10] 李莉, 赵越, 马君兰. 苯丙氨酸代谢途径关键酶: PAL、C4H、4CL 研究新进展[J]. 生物信息学, 2007, 5(4): 187-189
- [11] Russell D W, Conn E E. The cinnamic acid 4-hydroxylase of pea seedlings [J]. Arch Biochem Biophys, 1967, 122(1): 256-258
- [12] Fahrendorf T, Dixon R A. Stress responses in alfalfa (*Medicago sativa* L.). XVIII: Molecular cloning and expression of the elicitor-inducible cinnamic acid 4-hydroxylase cytochrome 450 [J]. Arch Biochem Biophys, 1993, 305(2): 509-515
- [13] Mizutani M, Ward E, DiMaio J, et al. Molecular cloning and sequencing of a cDNA encoding mung bean cytochrome P450 possessing cinnamate 4-hydroxylase activity [J]. Biochem Biophys Res Commun, 1993, 190: 875-880
- [14] Teutsch H G, Hasenfratz M P, Lesot A, et al. Isolation and sequence of a cDNA encoding the Jerusalem artichoke cinnamate 4-hydroxylase, a major plant cytochrome P450 involved in the general phenylpropanoid pathway [J]. Proc Natl Acad Sci USA, 1993, 90: 119-126
- [15] Millar D J, Long M, Donovan G, et al. Introduction of sense constructs of cinnamate 4-hydroxylase (CYP73A24) in transgenic tomato plants shows opposite effects on flux into stem lignin and fruit flavonoids [J]. Phytochemistry, 2007, 68(11): 1497-1509
- [16] Ryan K G, Swinney E E, Markham K R, et al. Flavonoid gene expression and UV photoprotection in transgenic and mutant *Petunia* leaves [J]. Phytochemistry, 2002, 59(1): 23-32
- [17] Sewalt V J H, Ni W, Blount J W, et al. Reduced lignin content and altered lignin composition in transgenic tobacco down-regulated in expression of L-phenylalanine ammonia-lyase or cinnamate 4-hydroxylase [J]. Plant Physiol, 1997, 115(1): 41-50
- [18] Han Z X, Wu F, Deng G B, et al. Structural and expression analysis of the B-hordein genes in Tibetan hull-less barley [J]. Genetics, 2010, 138: 227-239
- [19] Kaczmarczyk A, Bowra S, Elek Z, et al. Quantitative RT-PCR based platform for rapid quantification of the transcripts of highly homologous multigene families and their members during grain development [J]. BMC Plant Biol, 2012, 12: 184
- [20] Werck-Reichhart D, Feyereisen R. Cytochromes P450: A success story [J]. Genome Biol, 2000, 1(6): 3003.1-3003.9
- [21] Chapple C. Molecular-genetic analysis of plant cytochrome P450-dependent monooxygenases [J]. Annu Rev Plant Physiol Plant Mol Biol, 1998, 49(49): 311-343
- [22] Schoch G A, Attias R, Ret M L, et al. Key substrate recognition residues in the active site of a plant cytochrome P450, CYP73A1 [J]. Eur J Biochem, 2003, 270(18): 3684-3695
- [23] Yamazaki S, Sato K, Suhara K, et al. Importance of the proline-rich region following signal-anchor sequence in the formation of correct conformation of microsomal cytochrome P450s [J]. J Biochem, 1993, 114(5): 652-657
- [24] Williams P A, Cosme J, Sridhar V, et al. Mammalian microsomal cytochrome P450 monooxygenase: Structural adaptations for membrane binding and functional diversity [J]. Mol Cell, 2000, 5(1): 121-131
- [25] Zhao J, Dixon R A. The ins and outs of flavonoid transport [J]. Trends Plant Sci, 2009, 15(2): 72-80
- [26] 陈鸿翰, 袁梦求, 李双江, 等. 苦荞肉桂酸羟化酶基因 (*FtC4H*) 的克隆及其 UV-B 胁迫下的组织表达 [J]. 农业生物技术学报, 2013, 21(2): 137-147
- [27] Rupasinghe S, Baudry J, Schuler M A. Common active site architecture and binding strategy of four phenylpropanoid P450s from *Arabidopsis thaliana* as revealed by molecular modeling [J]. Protein Eng, 2003, 16(10): 721-731
- [28] 刘宛, 李培军, 周启星, 等. 植物细胞色素 P450 酶系的研究进展及其与外来物质的关系 [J]. 环境污染治理技术与设备, 2001, 2(5): 1-9
- [29] Baek M H, Chung B Y, Kim J H, et al. cDNA cloning and expression pattern of cinnamate 4-hydroxylase in the Korean black raspberry [J]. Biochem Mol Biol Rep, 2008, 41(7): 529-536
- [30] Koopmann E, Logemann E, Hahlbrock K. Regulation and functional expression of cinnamate 4-hydroxylase from parsley [J]. Plant Physiol, 1999, 119(1): 49-56
- [31] Gravot A, Larbat R, Hehn A, et al. Cinnamic acid 4-hydroxylase mechanism-based inactivation by psoralen derivatives: cloning and characterization of a C4H from a psoralen producing plant-*Ruta graveolens*-exhibiting low sensitivity to psoralen inactivation [J]. Arch Biochem Biophys, 2004, 422(1): 71-80