

• 药材与资源 •

布渣叶黄烷酮-2-羟化酶基因 cDNA 克隆及其生物信息学和表达分析

李坤平¹, 林爽², 张初梅¹, 潘天玲¹, 袁旭江³

1. 广东药科大学药学院, 广东 广州 510006

2. 广东药科大学中药学院, 广东 广州 510006

3. 广东药科大学中药开发研究所, 广东 广州 510006

摘要: 目的 从布渣叶中克隆黄酮碳苷合成途径关键酶黄烷酮-2-羟化酶 (flavanone 2-hydroxylase, F2H) 基因, 构建原核表达载体, 对其进行生物信息学和表达模式分析。方法 根据布渣叶转录组数据中注释为 F2H 的 Unigene 设计特异性引物, 利用 PCR 法扩增 MpF2H 基因的开放阅读框 (open reading frame, ORF), 基于在线工具对 cDNA 序列进行生物信息学分析。同时, 依据 F2H 序列, 设计特异引物, 采用 RT-qPCR 方法对其在芽、叶、枝、花、果等组织中的表达进行测定。结果 MpF2H 基因 ORF 全长为 1 557 bp (Genbank 登录号 KY652921), 编码 518 个氨基酸; 相对分子质量 54 500, 理论等电点 5.49, 含有信号肽, 不含跨膜域, 可能定位于叶绿体。同时, MpF2H 基因在不同组织中均有表达, 其中在叶中表达量最高, 在枝和花中表达量最低。结论 MpF2H 基因在不同组织中表达量差异较大。克隆了 MpF2H 基因 cDNA, 构建了其原核表达载体 pET30a-MpF2H, 为 MpF2H 基因的原核表达和功能验证奠定了基础。

关键词: 布渣叶; 黄烷酮-2-羟化酶; 黄酮碳苷; 类黄酮生物合成; 基因克隆; 原核表达载体

中图分类号: R282.12 **文献标志码:** A **文章编号:** 0253-2670(2018)01-0188-06

DOI: 10.7501/j.issn.0253-2670.2018.01.026

Cloning, bioinformatics, and expression analysis of MpF2H gene in *Microcos paniculata*LI Kun-ping¹, LIN Shuang², ZHANG Chu-mei¹, PAN Tian-ling¹, YUAN Xu-jiang³

1. School of Pharmacy, Guangdong Pharmaceutical University, Guangzhou 510006, China

2. School of Chinese Traditional Medicine, Guangdong Pharmaceutical University, Guangzhou 510006, China

3. Research & Development Institute of Chinese Materia Medica, Guangdong Pharmaceutical University, Guangzhou 510006, China

Abstract: **Objective** To obtain the key enzyme gene involved in flavone C-glycosides biosynthesis pathway, a flavanone 2-hydroxylase (F2H) gene was cloned from *Microcos paniculata*, and its bioinformatics analysis and gene expression pattern were also performed. **Methods** The specific primers were designed according to Unigene in F2H annotated in the transcriptome data of *M. paniculata*. The open reading frame (ORF) of MpF2H gene was amplified by PCR. Then the PCR product was purified and ligated to pET30a, and finally a prokaryotic expression vector pET30a-MpF2H was constructed. The bioinformation of F2H gene cDNA sequences was analyzed by some online tools. Using RT-qPCR with suitable primers, the quantitative expression analysis of MpF2H gene in different tissues, namely, buds, leaves, twigs, flowers and fruits was carried out. **Results** The length of MpF2H gene ORF was 1 557 bp (GenBank accession number KY652921), which encoded a protein with 518 amino acid residues, relative molecular weight of 54 500, theory isoelectric point of 5.49. In which was no transmembrane domain. It was hypothesized that this protein located in chloroplast. MpF2H gene was expressed in different tissues, with the highest expression in leaves and the lowest expression in twigs and flowers. **Conclusion** The expression of MpF2H gene varied widely in different tissues. The MpF2H gene was cloned from *M. paniculata* based on pET30a-MpF2H expression vector. This study will provide the fundamental information for the further preparation and functional research of MpF2H protein in flavone C-glycosides biosynthesis pathway.

Key words: *Microcos paniculata* L.; flavanone 2-hydroxylase; flavone C-glycosides; flavonoids biosynthesis; gene cloning; prokaryotic expression vector

收稿日期: 2017-07-11

基金项目: 国家自然科学基金资助项目 (31300273); 广东省科技计划项目 (2015A030302082)

作者简介: 李坤平 (1978—), 男, 博士, 副教授, 从事中药制药和药用植物资源方面的教学与研究。Tel: (020)39352118 E-mail: lkpcina@hotmail.com

布渣叶 *Microcos paniculata* L. 为椴树科破布叶属植物, 其叶药食两用, 最早见于清·何克谏著岭南医籍《生草药性备要》, 现收载于卫生部新资源食品目录(2010)和《中国药典》2015年版^[1]。前期研究发现布渣叶富含黄酮类成分, 其中黄酮碳-苷量亦较高^[2-3]。黄酮是一类具有广泛生理活性的小分子化合物, 在植物体内大部分以糖苷形式存在, 其中大多为氧-苷, 少数为碳-苷。黄酮碳苷的糖配基大多直接连接在其 A 环的 C-6/C-8 位, 因母核各异, 其种类也不尽相同。近年来, 随着研究的不断深入, 黄酮碳-苷及其生物活性引起了越来越多的关注^[4]。前期研究发现, 布渣叶中富含芹菜素碳苷类成分 (apigenin C-glycosides, ACGs), 如牡荆苷、异牡荆苷、佛来心苷、异佛来心苷等^[2-4], 而此类化合物通常具有较好的抗炎、护肝、抗肿瘤、降血糖等生物活性^[5-6]。

黄酮碳苷化学性质较氧-苷稳定, 其化学合成难度也较大, 迄今, 获取黄酮碳-苷的主要途径仍然是从植物原料中提取分离。随着现代生物技术的发展, 通过调控相应的生物合成途径以提高特定代谢产物累积的植物代谢工程的研究与应用已越来越成熟, 如青蒿素^[7]、人参皂苷^[8]等; 但对布渣叶而言, 其次生代谢研究鲜见报道, 亟待展开。从前期获得的大量布渣叶 RNA-Seq 数据分析中, 发现多个参与 ACGs 合成的 Unigene, 本实验对其中一个注释为黄酮酮-2-羟化酶 (flavanone 2-hydroxylase, F2H) 的序列进行了 cDNA 克隆, 并进行了相应的生物信息学分析和原核表达载体的构建, 以期对布渣叶 ACGs 生物合成研究提供参考。

1 材料与方法

1.1 材料与试剂

样品采自广东药科大学大学城校区药用植物园 (23°03'29.65N, 113°24'34.74"E), 经笔者鉴定为布渣叶 *M. paniculata* 新鲜组织, 所有样品采集后经液氮速冻, 于-80 °C超低温冰箱中保存。

植物总 RNA 提取试剂盒、PCR 产物纯化试剂盒购自天根生物科技公司; TransScript II Reverse Transcriptase、Ex Taq 酶、限制性核酸内切酶、T4 DNA 连接酶购自 Takara 公司; 引物由 BGI 合成。牡荆苷 (批号 111687-200602) 购自中国食品药品检定研究院; 甲醇 (色谱纯, Merck 公司), 磷酸 (分析纯, 广州试剂厂), 水 (蒸馏水, 屈臣氏)。

1.2 RNA 提取和 cDNA 合成

取约 100 mg 布渣叶的芽、叶、茎、花和果等

生物样本, 加液氮研磨, 根据植物总 RNA 提取试剂盒说明书提取总 RNA, 以 NanoDrop 2000 和 1% 琼脂糖凝胶电泳检测 RNA 的量和完整性, 用 TransScript II Reverse Transcriptase 试剂盒反转录合成 cDNA。

1.3 MpF2H 基因 cDNA 克隆和测序

根据已获得的布渣叶转录组数据, 针对 Unigene (MpF2H) 设计特异性引物, 正向引物: MpF2H-F (5'-CGGGATCCATGATGCCCATGATAC-TTGAAT-3', 下划线部分为 *Bam*HI 酶切位点); 反向引物: MpF2H-R (5'-CCCAAGCTTTTAGGTTG-CAAAGAGAGTGGA3', 下划线部分为 *Hind*III 酶切位点)。PCR 反应体系为 cDNA 模板 1 μL, 正反向引物各 1 μL, 10×ExTaq buffer 5 μL, 2.5 mmol/L dNTP 4 μL, ExTaq 酶 0.25 μL, 加水至 50 μL。PCR 扩增反应程序 (降落 PCR): 95 °C 预变性 5 min, 循环内 98 °C、10 s 变性, 从 65 °C 每个循环降 1 °C 退火, 72 °C 延伸 2 min, 10 个循环; 在 55 °C 退火, 进行 20 个循环; PCR 反应循环后 72 °C 延伸 7 min, 然后 16 °C 保存。PCR 产物经 1% 琼脂糖凝胶电泳检测和试剂盒纯化, 用限制性内切酶 *Bam*HI 和 *Hind*III 对 PCR 纯化产物和 pET-30a (+) 质粒进行双酶切, T4 DNA 连接酶 16 °C 连接过夜后转化大肠杆菌 *E. coli* BL21 (DE3) 感受态细胞, 在卡拉霉素抗性平板上进行筛选, 挑选阳性克隆经菌液 PCR 和电泳检测后, 送 BGI 测序。

1.4 MpF2H 基因的生物信息学分析^[9]

对 MpF2H 基因编码的蛋白质, 使用在线工具 ProtParam (<http://web.expasy.org/protparam/>) 预测其相对分子质量与理论等电点 (PI), PredictProtein (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_sopma.html) 在线软件进行二级结构预测, SWISS-MODEL (<http://swissmodel.expasy.org/>) 在线软件进行三维同源建模。SignalP 4.1 server (<http://www.cbs.dtu.dk/services/SignalP>) 分析蛋白信号肽, TargetP 1.1 server (<http://www.cbs.dtu.dk/services/TargetP>) 预测蛋白质亚细胞定位, TMHMM server (<http://www.cbs.dtu.dk/services/TMHMM/>) 进行跨膜区预测。对 MpF2H 蛋白的氨基酸序列在 NCBI 数据库进行在线搜索 (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>), 从中选取并下载与 MpF2H 基因编码的蛋白质相似性较高的其他植物 F2H 蛋白的氨基酸序列, 使用 MEGA6 软件的相邻连接法 (neighbor

joint) 构建系统进化树。

1.5 MpF2H 基因的表达分析

按照“1.2”项的方法提取布渣叶生物样本中 RNA, 并逆转录合成 cDNA。选择 Tubulin 基因作为内参, 以 cDNA 为模板, 根据 MpF2H 基因设计 RT-qPCR 引物 F 5'-AACAAGCAAGGTCCTAATG-3' 和 R 5'-CAGAGCCAAGACGAAGATA-3', 内参基因引物 F 5'-TACCAGCAACCACGACAG-3' 和 R 5'-CACAAGAATCGTTGAGCC-3'。以 10 μ L 反应体系为基准, 包括 5 μ L SYBR Premix Ex Taq, 正反向引物各 0.2 μ L (10 μ mol/L), cDNA 模板 4.6 μ L。反应程序为 95 $^{\circ}$ C 预变性 30 s; 95 $^{\circ}$ C、5 s, 60 $^{\circ}$ C、20 s, 40 个循环; 数据采集; 65 $^{\circ}$ C、15 s; 溶解曲线 60~95 $^{\circ}$ C; 数据采集; 每个反应做 3 个重复。

1.6 布渣叶不同组织中 ACGs 的量分析

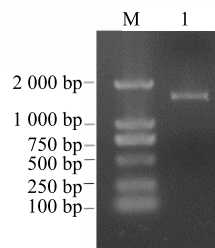
取适量布渣叶不同组织样品, 液氮研磨后微波干燥, 准确称取 100 mg 各干燥样品, 参考文献方法^[10]制备供试品, 并采用 HPLC 法测定^[10]布渣叶中 8 种黄酮苷的量, 以维采宁 2、牡荆苷、异牡荆苷、异夏胡塔苷、佛来心苷、异佛来心苷、紫云英苷 7 个黄酮碳苷类化合物的峰面积计算 ACGs 量; 每个组织部位取 3 份样品, 每份样品平行测定 3 次。

2 结果与分析

2.1 MpF2H 基因 cDNA 克隆

通过本实验室前期获得的布渣叶转录组数据, 发现多个参与 ACGs 合成途径的基因, 其中有一个功能注释为 F2H 的 Unigene, 长度为 1 557 bp, 利用美国国家生物科技信息中心 NCBI 上的 ORF Finder 程序 (<https://www.ncbi.nlm.nih.gov/orffinder/>) 分析其含有一个完整的 ORF, 根据该 ORF 序列设计特异性引物 MpF2H-F 和 MpF2H-R, 以布渣叶叶片总 RNA 反转录得到的 cDNA 为模板, PCR 扩增得到 1 条 1 500 bp 左右的片段, 电泳结果如图 1 所示, 与预期大小相符。

用限制性内切酶 *Bam*HI 和 *Hind*III 分别对 PCR 纯化产物和 pET-30a (+) 质粒进行双酶切, T4 DNA 连接酶连接后转化 *E. coli* BL21 (DE3) 感受态细胞, 卡拉霉素抗性筛选后挑取阳性单克隆, 进行 PCR 菌检和双酶切鉴定, 结果显示有 1 500 bp 左右的目的条带 (图 2)。将酶切正确的单克隆送 BGI 测序, 测序结果显示与目的基因 MpF2H 序列一致, 表明已成功获取 pET-30a (+)-MpF2H 重组体。MpF2H 基因 ORF 序列长为 1 557 bp, 编码 518 个氨基酸 (图 2), 序列信息已经提交到 GenBank, 登录号为 KY652921。

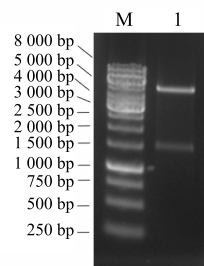


M-Marker 1-MpF2H 基因 PCR 扩增产物

M-Marker 1-PCR product of MpF2H gene

图 1 PCR 扩增 MpF2H 基因

Fig. 1 PCR amplification of MpF2H gene



M-Marker 1-*Bam*HI and *Hind*III 双酶切产物

M-Marker 1-double digestion results by *Bam*HI and *Hind*III

图 2 原核表达载体 pET30a-MpF2H 的双酶切鉴定

Fig. 2 Identification of expression vector pET30a-MpF2H

2.2 MpF2H 基因编码蛋白的生物信息学分析

MpF2H 基因编码 518 个氨基酸 (图 3), 利用 ExPASy Proteomics Server 的在线工具 ProtParam 对其编码蛋白的理化性质进行预测分析。推测分子式为 $C_{2411}H_{3840}N_{650}O_{742}S_{20}$, 相对分子质量 54 500, pI: 5.49; 正电残基 (Arg+Lys): 44, 负电残基 (Asp+Glu): 56。不稳定指数为 33.13, 说明 MpF2H 蛋白较稳定, 亲水性总平均系数为 -0.073, 说明 MpF2H 蛋白为亲水性蛋白。利用 <https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi> 进行 MpF2H 蛋白的保守结构域预测, 结果显示 MpF2H 蛋白含有细胞色素 P450 保守结构域。

SignalP 4.1 server 预测 MpF2H 蛋白含有信号肽 (图 4), 为分泌性蛋白, 前导肽切割位点在序列的第 15、16 位之间; TargetP 1.1 server 预测 MpF2H 蛋白很可能定位于叶绿体内; TMHMM server 预测该蛋白不含跨膜域 (图 5)。用 PredictProtein 对 MpF2H 蛋白的二级结构进行预测 (图 6), 结果显示在该蛋白中 α -螺旋结构占 48.65%, 延伸链占 15.06%, β -转角占 7.72%, 无规卷曲占 28.57%。

```

1 ATGATGCCCATGATACTTGAATTCCTATCATATCTAGCTCCCTTAGTTGTTTCATTCTCTCTCGTTAAGAGAATACTC
1 M M P M I L E F L S Y L A P L V V S F L L V K R I L
79 ATCTCTACCAACAAGGTCCTAATGCCAAGCAGCTTCTTCCCCAGGCCCAATCGCCTTACCGATCATTGGCCAC
27 I S T N K Q G P N A K Q L L P P G P I A L P I I G H
157 CTCCATCTCTGGGCCCTTCTTCATCAAACTTCCGCAAGCTCTCTCGGCTACGGTCCCTTAATGTATCTTCGT
53 L H L L G P F L H Q T F R K L S S R Y G P L M Y L R
235 CTGGCTCTGTGGCTGCGTGGTGGCTTCCAATCCAGAGTTTGCCAAAGAGCTTCTCAAACTTACGAGCTCGGCTTC
79 L G S V G C V V A S N P E F A K E L L K T Y E L A F
313 GCTGCCCGCATGCACCCGCTGCCATCACCCACCTGACCTACAACCTGCTCTTGGCTTCGCACCGTACGGAACGTAC
105 A A R M H T A A I T H L T Y N S S F A F A P Y G T Y
391 TGGAAATTCATCAGAAAGTTCAGCAGTACGAGCTCTAGGCAACCGTACTCTCGCCAGTTTCTTCCCGTTTCAAGC
131 W K F I R K F S T Y E L L G N R T L A Q F L P V R T
469 AAGGAATTGCACCATCTCCTTTCAGTTTGTCTTGGCAAGGCTAAAGCAGGGGAAAGCGTGAATTTAAACCAAGAGCTG
157 K E L H H L L Q F V L G K A K A G E S V N L T Q E L
547 TTGAACTAAACAACAACCATATCGCAGATGATGCTGAGCATGCGCTTTCGGGGACAGGAAACCCGCGCAGCTCG
183 L K L T N N T I S Q M M L S M R C S G T G N P A D S
625 GTTAGAATCTTGTGCGAGAGGTGACAGATCTTCGAGAGTTCAACATCTCAGATAGCATATGGTTTTTAAAAAC
209 V R T L V R E V T E I F G E F N I S D S I W F L K N
703 TGGGACTTGCAGGGATTGAGAAAGAGATTCGAGACCTGCATAGGAGTTTGTATGCAATTGTGGAGATGATTATGAGA
235 W D L Q G F R K R F E D L H R R F D A L L E M I M R
781 GAACGTGACGAGTAAGATCAGAAAGCAAGCAAAAGAGGCGACAATGTTAACAAGGTAAGAGATTTCTGGACATC
261 E R E R V R S E S K Q K K G D N V N K V K D F L D I
859 ATGCTTGACGTTTGGAAAAGGATACTCGGAGTTCGAGGTAGATTTTACCAGAAATCACATCAAGGCGCTTAATTTTG
287 M L D V L E K D N S E S E V D F T R N H I K A L I L
937 GATTTCTTACAGCCGCTACAGATACAACGGCAATTGTACTTGAATGGCAATGGCAGAGTTGATCAACCAACCGAAG
313 D F F T A A T D T T A I V L E W A M A E L I N H P K
1015 CTACTAAAAATAGCTCAGCAAGAAATTGATCAAGTTGTGGGAAAAGCAGGTTGGTGGGAAGATCTGACAGCCACAT
339 L L K I A Q Q E I D Q V V G K G R L V E E S D S P H
1093 CTCCATTACATCAAGCCATTATTAAGAAACATTTCCGCTTCAACCACAGTCCCAATGATCAATAGGAAATCAATC
365 L H Y I Q A I I K E T F R L H P P V P M I N R K S I
1171 CAAACATGCCAGGTTAAGGGATACAAATCCCTGCTGAATGTATGGTGTGTGTGAACGTTTGGGCTATCGGAAGAGAT
391 Q T C Q V K G Y T I P A E C M V F V N V W A I G R D
1249 CCCAAGGTTTGGACAGATCCATTGAAGTTTCAGCCTGAGAGATTCTGAAATCCGATGATTCTATAGATGTTAGAGGG
417 P K V W T D P L K F Q P E R F L K S D D S I D V R G
1327 CAACATTACGAGCTGTTCCTTTCGGGTCAGGAGGAGGGGCTGCCCTGGCGCTTCTTTGGCGTTGCAGGAGCTGCCC
443 Q H Y E L L P F G S G R R G C P G A S L A L Q E L P
1405 ACCACTTTGGCTGCCATGATTCACTGCTTGTACTGGAAGGTGGCGAATGGCGTTGTGACATGGTTGAAAGGCGCTGGA
469 T T L A A M I Q C F D W K V A N G V V D M V E R P G
1483 CTTACGGCTCCAGGGCTAAGGATCTCGAGTGTGTCTGTGTACGCTTCACTCCACTCTCTTTGCAACCTAA
495 L T A P R A K D L E C V P V V R F T P T L F A T *

```

*代表终止密码子

*stands for the termination codon

图 3 布渣叶 F2H 基因 (MpF2H)cDNA 编码区序列及氨基酸序列

Fig. 3 Coding sequence and amino acid sequence of MpF2H gene cDNA

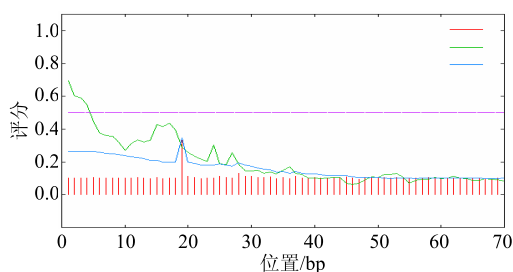


图 4 MpF2H 蛋白信号肽预测分析

Fig. 4 Prediction of signal peptide of MpF2H protein

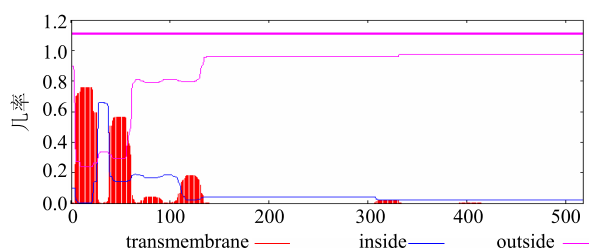
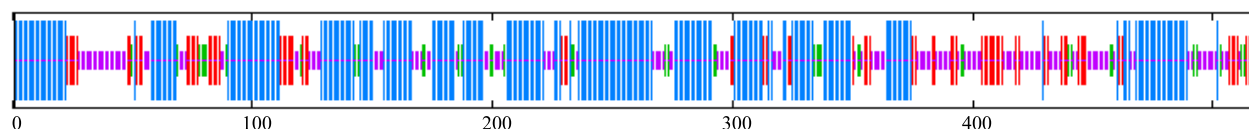


图 5 MpF2H 蛋白的跨膜区预测分析

Fig. 5 Prediction of transmembrane domains of MpF2H protein



蓝色代表 α -螺旋, 红色代表延伸链, 绿色代表 β -转角, 紫色代表无规卷曲

Blue, red, green and purple stand for alpha helix, extended strand, beta turn and random coil, respectively

图 6 MpF2H 基因编码蛋白质的二级结构预测

Fig. 6 Prediction of secondary structure of MpF2H protein encoding by MpF2H gene

用 SWISS-MODEL 以人 P450 1A2 蛋白(PDB ID: 2HI4) 为模板^[11], 预测 MpF2H 蛋白的三维结构, 序列相似度为 35%, 在第 56~512 位氨基酸建模, 模型覆盖率 85%, MpF2H 蛋白预测的三维模型见图 7。

2.3 MpF2H 氨基酸序列多重比对与系统进化分析

将 MpF2H 编码的蛋白在 NCBI 数据库进行 Blast 分析, 发现其与可可 *Theobroma cacao* Linn. Le Cacotier.(EOY17412.1)、长蒴黄麻 *Corchorus olitorius*

Linn. (OMO95638.1)、陆地棉 *Gossypium hirsutum* Linn. (XP_016704458.1) 和毛果杨 *Populus trichocarpa* Torr. & A. Gray ex Hook (XP_002319566.2) 等木本植物的细胞色素 P450 蛋白或黄酮合酶 II 的氨基酸序列相似度较高, 分别达到了 78%、71%、68%和 63%, 相应的多序列比对结果如图 8 所示。同时, 从 NCBI 数据库中下载与 MpF2H 相似度较高其他物种的蛋白质序列, 用 MEGA 6.0 构建系统进化树, 如图 9 所示。进化树上可以看到布渣叶 MpF2H 基因和同科植物黄麻、长蒴黄麻, 以及可可的亲缘性最近。

2.4 MpF2H 基因的表达特性

RT-qPCR 结果表明, MpF2H 基因在芽、叶、茎、花和果中均有表达, 在叶中的表达量最高, 芽、果次之, 在茎中表达量最低(图 10); 这与布渣叶不同组织部位中总黄酮素碳苷的含量(图 11)呈现正相关性, 提示 MpF2H 基因为布渣叶 ACGs 合成关键酶基因, 调控 MpF2H 基因的表达能影响到其 ACGs 的生物合成。

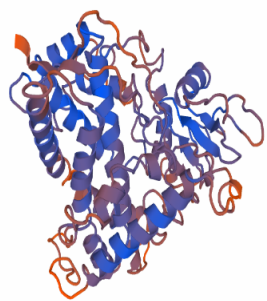


图 7 MpF2H 蛋白分子三维结构预测

Fig. 7 Prediction of 3D structure of MpF2H protein

布渣叶 <i>M. paniculata</i>	... MPMI LEFLSYLAPLVVSFLVKRI L STNKGPNKCLPPGP ALPI GHLLHLGFLHCTFRKLSRYGPNLYL	77
长蒴黄麻 <i>C. olitorius</i>	... NVAVGPLYATLFI LPLVI VKAI LI SI KK. RPKLPPGPVSLPIVGHLLHGLSFLHCTFCTLSSTYGPLNLYL	70
可可 <i>T. cacao</i>	... NVNVGPLYATLFI LPLVI VKAI LI SI KK. RPKLPPGPVSLPIVGHLLHGLSFLHCTFCTLSSTYGPLNLYL	70
毛果杨 <i>P. trichocarpa</i>	... M FELII AGFAVLLLT VFI LTN. KG. HGSLLPPGPVSLPIVGHLLHGLSFLHCTFCTLSSTYGPLNLYL	65
陆地棉 <i>G. hirsutum</i>	NVLCLANLPLFSYAALFI VSFLI VR. TVI SI RN. RSKLPPGPATLPIVGHLLHGLSFLHCTFCTLSSTYGPLNLYL	74
Consensus	l p p p p l p i g h l h l l p h f l s g y l	
布渣叶 <i>M. paniculata</i>	RLGSVGVASINPEFAKELLKTYELAFARNITAI THLYNSSFAPAFYGYVVFIRKSTYELLGRLACFLPVRTK	157
长蒴黄麻 <i>C. olitorius</i>	RLGSVGVASINPDALAKELLKTHETFSARKHTAI DHLYNSSFAPAFYGYVVFIRKSTYELLGRLACFLPVRTK	150
可可 <i>T. cacao</i>	RLGSVGVASINPDALAKELLKTHETFSARKHTAI DHLYNSSFAPAFYGYVVFIRKSTYELLGRLACFLPVRTK	150
毛果杨 <i>P. trichocarpa</i>	RLGSVGVASINPEFAKELLKTHETFSARKHTAI DHLYNSSFAPAFYGYVVFIRKSTYELLGRLACFLPVRTK	145
陆地棉 <i>G. hirsutum</i>	RLGSVGVASINPEFAKELLKTHETFSARKHTAI DHLYNSSFAPAFYGYVVFIRKSTYELLGRLACFLPVRTK	154
Consensus	r g s v c v s p a e l i k l f f r h a i h t y s s f a p a p y g y v f i k s t e l g r l q f l p v r k	
布渣叶 <i>M. paniculata</i>	ELHHL GFVLCAKAGESVNTGELLKLTNNI SCMLSMRCSGTGNFADSVRTLVRVTEIFGEFNSDSIVFLKNVLD	237
长蒴黄麻 <i>C. olitorius</i>	ELHHL GFVLCAKAGESVNTGELLKLTNNI SCMLSMRCSGTGTGADGIRTLVRVTEIFGEFNSDSIVFLKNVLD	230
可可 <i>T. cacao</i>	ELHHL GFVLCAKAGESVNTGELLKLTNNI SCMLSMRCSGTGTGADGIRTLVRVTEIFGEFNSDSIVFLKNVLD	230
毛果杨 <i>P. trichocarpa</i>	ELRCFI GVLHKSSEVCSVNTGELLKLTNNI SCMLSMRCSGTGTGADGIRTLVRVTEIFGEFNSDSIVFLKNVLD	225
陆地棉 <i>G. hirsutum</i>	ELHHL GFVLCAKAGESVNTGELLKLTNNI SCMLSMRCSGTGTGADGIRTLVRVTEIFGEFNSDSIVFLKNVLD	234
Consensus	e l h h l s i l l e k s s e i v n t g e l l k l t n n i s c m l s m r c s g t g n f a d s v r t l v r v t e i f g e f n s d s i v f l k n v l d	
布渣叶 <i>M. paniculata</i>	GGRRKRFEDLHRRYCALLLENI NRRERVRVRS ESCKKGDVNVKVKDLIDMLDVLEKDSSEVDFTRNH KALI LDFLTA	317
长蒴黄麻 <i>C. olitorius</i>	GGRRKRFEDLHRRYCALLLENI NRRERVRVRS ESCKKGDVNVKVKDLIDMLDVLEKDSSEVDFTRNH KALI LDFLTA	304
可可 <i>T. cacao</i>	GGRRKRFEDLHRRYCALLLENI NRRERVRVRS ESCKKGDVNVKVKDLIDMLDVLEKDSSEVDFTRNH KALI LDFLTA	304
毛果杨 <i>P. trichocarpa</i>	GGRRKRFEDLHRRYCALLLENI NRRERVRVRS ESCKKGDVNVKVKDLIDMLDVLEKDSSEVDFTRNH KALI LDFLTA	298
陆地棉 <i>G. hirsutum</i>	GGRRKRFEDLHRRYCALLLENI NRRERVRVRS ESCKKGDVNVKVKDLIDMLDVLEKDSSEVDFTRNH KALI LDFLTA	308
Consensus	g r k r f d h r d l l e i r e k k d l d n l d e s e i c l t r h k a l i c f t a	
布渣叶 <i>M. paniculata</i>	ATDTTA VLEVAALNI NHPKLLKI ACCEI DGVVGKRLVGSSESPHLYI CAI KETFRLLHPPVPMI NRKSI GTCCVK	397
长蒴黄麻 <i>C. olitorius</i>	ATDTTA VVEVSLAELNHTVLRKAGKEI DGVVGKRLVGSSESPHLYI CAI KETFRLLHPPVPMI NRKSI GTCCVK	384
可可 <i>T. cacao</i>	ATDTTA VVEVSLAELNHTVLRKAGKEI DGVVGKRLVGSSESPHLYI CAI KETFRLLHPPVPMI NRKSI GTCCVK	384
毛果杨 <i>P. trichocarpa</i>	ATDTTA ATEVALAELNI NHPKLLKI ACCEI DGVVGKRLVGSSESPHLYI CAI KETFRLLHPPVPMI NRKSI GTCCVK	378
陆地棉 <i>G. hirsutum</i>	ATDTTA AVEVALAELNI NHPKLLKI ACCEI DGVVGKRLVGSSESPHLYI CAI KETFRLLHPPVPMI NRKSI GTCCVK	388
Consensus	a t d t t a e w a e l n l a e d v v g r l v e d l y q a i k e f r l h p p p n i r k g	
布渣叶 <i>M. paniculata</i>	YTI PAECNVFVNVAI GRDPKVVDPLKFGPERFLKSD. DSIVRGCHYELLPFSGRRGCPGSLAACEVPT	469
长蒴黄麻 <i>C. olitorius</i>	YTI PACSLLFVNVAI GRDPKVVDPLKFGPERFLKSD. DSIVRGCHYELLPFSGRRGCPGSLAACEVPT	456
可可 <i>T. cacao</i>	YTI PACSLLFVNVAI GRDPKVVDPLKFGPERFLKSD. DSIVRGCHYELLPFSGRRGCPGSLAACEVPT	456
毛果杨 <i>P. trichocarpa</i>	YTI PKATMLFVNVAI GRDPKVVDPLKFGPERFLKSD. DSIVRGCHYELLPFSGRRGCPGSLAACEVPT	457
陆地棉 <i>G. hirsutum</i>	YTI PACSLLFVNVAI GRDPKVVDPLKFGPERFLKSD. DSIVRGCHYELLPFSGRRGCPGSLAACEVPT	468
Consensus	y t i p f v n w g r d w p l f p e r f k d g h y l l c f g g r r c p g l a c e v p	
布渣叶 <i>M. paniculata</i>	TLAANI CCFCVKVAN. GVVDMAERGLTAPRAKDLGVVVFRTPTLTA	517
长蒴黄麻 <i>C. olitorius</i>	TLAANI CCFCVKLPGEA. AVVDMAERGLTAPRAKDLGVVVFRTPTLTA	504
可可 <i>T. cacao</i>	TLAANI CCFCVKLPGEA. AVVDMAERGLTAPRAKDLGVVVFRTPTLTA	506
毛果杨 <i>P. trichocarpa</i>	SLAANI CCFCVKVADPHGVKI KGNALVDNTERGLTAPRALDGVVVFRTPTLTA	514
陆地棉 <i>G. hirsutum</i>	TLAANI CCFCVKLPVNG. GVVDMAERGLTAPRAKDLGVVVFRTPTLTA	518
Consensus	l a a i q c f w k v d m e r g l t a p r a d l g v v v f r t p t l t a	

图 8 MpF2H 与其他物种同源蛋白的氨基酸序列比对结果

Fig. 8 Alignments of amino acid sequences of MpF2H and similar proteins from other species

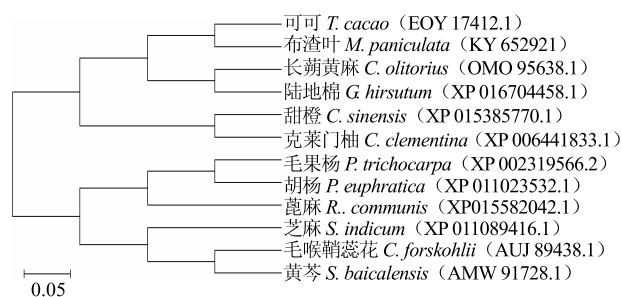


图 9 MpF2H 蛋白与其他物种 F2H 蛋白的系统进化树分析

Fig. 9 Phylogenetic tree of MpF2H protein and F2H proteins from other species

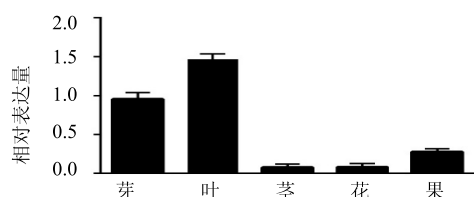


图 10 布渣叶不同组织中 MpF2H 相对表达量

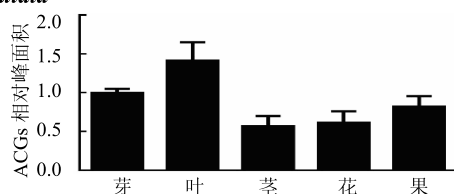
Fig. 10 Relative expression of MpF2H in different organs of *M. paniculata*

图 11 布渣叶不同组织中 ACGs 的相对量

Fig. 11 Relative contents of ACGs in different organs of *M. paniculata*

3 讨论

在植物类黄酮的生物合成研究上, 黄酮氧-苷的代谢途径、关键酶基因筛选及功能鉴定等报道较多, 但黄酮碳苷的生物合成研究相对较少。F2H 基因是控制整个黄酮碳苷合成的关键酶之一^[12-13], 具有重要的研究价值。迄今, 对于 F2H 的研究, 主要集中在小麦、大豆、玉米和水稻等^[13-16]作物上, 而鲜有药用植物的相关研究报道。本研究首次从布渣叶叶片中克隆得到了 MpF2H 基因, 并发现其表达量与布渣叶 ACGs 含量呈明显相关性, 这对布渣叶黄酮碳苷生物合成研究具有重要意义。

黄酮碳苷是植物类黄酮中较少见的一部分, 植物体内的量也很低, 如何提高黄酮碳苷量是一个热门的研究方向。近年来, 随着黄酮类化合物生物合成途径及其关键酶基因研究的深入, 以分子生物学技术为基础的合成生物学和植物代谢工程, 成为提高植物类黄酮产量最有力的途径之一^[17]。本实验首次从布渣叶中

克隆了 F2H 基因, 并利用 pET-30a (+) 质粒构建了 pET-30a (+)-MpF2H 重组体, 为后续 MpF2H 重组蛋白的诱导表达、蛋白结构和生物学功能研究奠定了基础。

参考文献

- [1] 中国药典 [S]. 一部. 2015.
- [2] 李坤平, 陈艳芬, 岳春华, 等. 布渣叶保护急性心肌缺血损伤有效部位化学成分研究 [J]. 中草药, 2014, 45(23): 3373-3376.
- [3] 冯世秀, 刘梅芳, 魏孝义, 等. 布渣叶中三萜和黄酮类成分的研究 [J]. 热带亚热带植物学报, 2008, 16(1): 51-56.
- [4] 邵莹, 吴启南, 乐巍, 等. 碳苷黄酮保护心肌缺血损伤作用的研究进展 [J]. 中草药, 2015, 46(1): 128-139.
- [5] Xiao J, Capanoglu E, Jassbi A R, *et al.* Advance on the flavonoid C-glycosides and health benefits [J]. *Crit Rev Food Sci Nutr*, 2016, 56(Suppl 1): 29-45.
- [6] Choi J S, Islam M N, Ali M Y, *et al.* Effects of C-glycosylation on anti-diabetic, anti-Alzheimer's disease and anti-inflammatory potential of apigenin [J]. *Food Chem Toxicol*, 2014, 64: 27-33.
- [7] Paddon C J, Westfall P J, Pitera D J, *et al.* High-level semi-synthetic production of the potent antimalarial artemisinin [J]. *Nature*, 2013, 496(7446): 528-532.
- [8] Yan X, Fan Y, Wei W, *et al.* Production of bioactive ginsenoside compound K in metabolically engineered yeast [J]. *Cell Res*, 2014, 24(6): 770-773.
- [9] 赵乐, 马利刚, 杨方方, 等. 独行菜磷酸甲羟戊酸激酶 LaPMK 基因克隆、生物信息学分析及原核表达 [J]. 中草药, 2016, 47(17): 3087-3093.
- [10] 刘昌乐, 林爽, 何卓儒, 等. 一测多评法测定布渣叶总黄酮部位中 8 种黄酮苷成分 [J]. 中草药, 2017, 48(9): 1872-1877.
- [11] Sansen S, Yano J K, Reynald R L, *et al.* Adaptations for the oxidation of polycyclic aromatic hydrocarbons exhibited by the structure of human P450 1A2 [J]. *J Biol Chem*, 2007, 282: 14348-14355.
- [12] Brazier-Hicks M, Evans K M, Gershater M C, *et al.* The C-glycosylation of flavonoids in cereals [J]. *J Biol Chem*, 2009, 284(27): 17926-17934.
- [13] Du Y G, Chu H, Chu I K, *et al.* CYP93G2 is a flavanone 2-hydroxylase required for C-glycosylflavone biosynthesis in Rice [J]. *Plant Physiol*, 2010, 154: 324-333.
- [14] Maria F F, Eduardo R, Maria I C, *et al.* Identification of a bifunctional maize C- and O-glucosyltransferase [J]. *J Biol Chem*, 2013, 288(44): 31678-31688.
- [15] Nagatomo Y, Usui S, Ito T, *et al.* Purification, molecular cloning and functional characterization of flavonoid C-glucosyltransferases from *Fagopyrum esculentum* (buckwheat) cotyledon [J]. *Plant J*, 2014, 80(3): 437-448.
- [16] Hirade Y, Kotoku N, Terasaka K, *et al.* Identification and functional analysis of 2-hydroxyflavone C-glucosyltransferase in soybean (*Glycine max*) FEBS [J]. *FEBS Lett*, 2015, 589(15): 1778-1786.
- [17] Melissa B H, Robert E. Metabolic engineering of the flavone-C-glycoside pathway using polypeptide technology [J]. *Metabolic Eng*, 2013, 16: 11-20.