

• 药材与资源 •

怀地黄 3-酮酯酰 CoA-硫解酶基因的克隆、序列特征和时空表达分析

周延清^{1,2}, 张永华^{1,2}, 张 喻^{1,2}, 陈艳梅^{1,2}, 白妍妍^{1,2}, 魏海方^{1,2}, 段红英^{1,2}, 周春娥^{1,2}

1. 河南师范大学生命科学学院, 河南 新乡 453007

2. 河南省高校道地药材保育与利用研究工程技术中心, 河南 新乡 453007

摘要: 目的 对怀地黄 3-酮酯酰 CoA-硫解酶 (*Rehmannia glutinosa* f. *hueichingensis* 3-ketoacyl CoA-thiolase, RghKAT) cDNA 全长基因进行克隆及分析, 为怀地黄分子育种提供候选基因和理论依据。方法 根据其他植物 ARGOS 基因序列的保守结构域设计简并引物, 采用 RT-PCR 和 RACE 技术, 获得 RghKAT cDNA 全长序列; 通过生物信息学技术对其核苷酸序列和氨基酸序列进行比对; 利用实时荧光定量 PCR 技术检测了其在 2 个时期、10 个组织的表达。结果 RghKAT 基因全长 1 713 bp, 包含了 1 395 bp 的开放阅读框 (ORF), 编码 464 个氨基酸; 同源比对和系统进化分析表明, RghKAT 的核苷酸序列与葡萄、番茄、毛果杨、拟南芥和小麦的 KAT 核苷酸序列同源性分别达 84%、82%、82%、79%、73%; RghKAT 编码的氨基酸序列与矮牵牛、葡萄、黄瓜、拟南芥和小麦的 KAT 氨基酸同源性分别为 88%、88%、86%、87%、78%; 各物种 KAT 酶进化树符合物种进化规律; 理化性质表明该蛋白为略成碱性的稳定蛋白质, 蛋白质二级结构主要由 α -螺旋、不规则卷曲、 β -折叠和 β -转角构成; 在 N 端存在一个由 70 个氨基酸残基组成的信号肽; RghKAT 蛋白三维结构具有硫解酶典型的特征序列; 表达谱分析表明, RghKAT mRNA 在各时期、各组织中均有表达, 盛花期花瓣中表达最强, 而在幼苗期叶中表达量最低。结论 成功克隆了 RghKAT cDNA 全长序列, 具有 KAT 基因的结构特性及其产物硫解酶典型的特征序列, 其在盛花期花瓣中表达量最高。

关键词: 怀地黄; 3-酮酯酰 CoA-硫解酶; 生物信息学分析; 时空表达分析; 蛋白质二级结构; 蛋白三维结构

中图分类号: R282.12

文献标志码: A

文章编号: 0253-2670(2013)01-0076-09

DOI: 10.7501/j.issn.0253-2670.2013.01.015

Gene cloning, features of sequence, and analysis on temporal and spatial expression of *Rehmannia glutinosa* f. *hueichingensis* 3-ketoacyl CoA-thiolase

ZHOU Yan-qing^{1,2}, ZHANG Yong-hua^{1,2}, ZHANG Yu^{1,2}, CHEN Yan-mei^{1,2}, BAI Yan-yan^{1,2},
WEI Hai-fang^{1,2}, DUAN Hong-ying^{1,2}, ZHOU Chun-e^{1,2}

1. College of Life Sciences, Henan Normal University, Xinxian 453007, China

2. Engineering Technology Research Center of Nursing and Utilization of Genuine Chinese Medicinal Materials, University and Colleges of Henan Province, Xinxian 453007, China

Abstract: Objective To clone and analyze the full-length cDNA of *Rehmannia glutinosa* f. *hueichingensis* 3-ketoacyl CoA-thiolase (RghKAT) gene and to provide the candidate genes and theoretical basis for the molecular breeding of *R. glutinosa* f. *hueichingensis*. **Methods** The full-length cDNA sequence of RghKAT gene was amplified by quantitative RT-PCR and RACE techniques with degenerated primers being designed based on the conserved domain of ARGOS base sequences from other plants. The nucleotide and amino acid sequences were compared by bioinformatics technology. The temporal and spatial expression levels in 10 tissues of regenerated plantlets at two stages were detected by quantitative RT-PCR. **Results** The full-length of RghKAT gene was 1 713 bp, including an open reading frame (1 395 bp) encoding a 464-amino acid protein. Homology comparison and phylogenetic analysis revealed that RghKAT shared the high nucleotide sequence identity to those of *Vitis vinifera* (84%), *Solanum lycopersicum* (82%), *Populus trichocarpa* (82%), *Arabidopsis thaliana* (79%), and *Triticum aestivum* (73%), respectively. Meanwhile, the amino acid sequence coded by RghKAT gene shared the high identity to those of *Petunia x hybrida* (88%), *V. vinifera* (88%), *Cucumis sativus* (86%), *A. thaliana*

收稿日期: 2012-09-28

基金项目: 河南省基础与前沿技术研究计划项目 (092300410009); 河南省教育厅自然科学研究计划项目 (2011A180022); 河南省高校道地药材保育与利用研究工程技术中心开放课题基金项目 (20100911); 河南师范大学第三届学生创新创业论坛暨大学生创新性实验计划项目

*通信作者 周延清 Tel: (0373)3326340 E-mail: yqzhou@htu.cn

网络出版时间: 2012-12-25 网络出版地址: <http://www.cnki.net/kcms/detail/12.1108.R.20121225.0911.002.html>

(87%), and *T. aestivum* (78%). The phylogenetic tree of RghKAT was consistent with the evolutionary relationship among the species. The physicochemical properties of RghKAT indicated that it was a slightly alkaline and stable protein, and the secondary structure was made of α -helixes, random coil, β -sheets, and β -turns. There was a putative signal peptide composed of 70 amino acid residues in the N-terminus, and the three-dimensional structure of RghKAT protein showed a typical characteristic sequence of thiolases. The quantitative RT-PCR assay demonstrated that the RghKAT mRNA could be detected in all 10 tissues examined at seedling and blooming stages with the highest expression level in the petals at blooming stage and the lowest one in the young leaves at seedling stage. **Conclusion** The full-length cDNA of RghKAT is cloned successfully, with the structural properties of KAT gene and the typical characteristic sequence of product thiolases, which has the highest expression level in the petals of blooming stage.

Key words: *Rehmannia glutinosa* Libosch. f. *hueichingensis* (Chao et Schih) Hsiao; 3-ketoacyl CoA-thiolase; bioinformatics analysis; temporal and spatial expression analysis; secondary structure of protein; three-dimensional structure of protein

地黄 *Rehmannia glutinosa* (Gaertn.) Libosch. ex Fisch. et Mey. 属玄参科 (Scrophulariaceae) 地黄属, 是多年生的草本植物, 广泛分布于中国、韩国和日本等地^[1]。地黄在中国的河南、山东、陕西和山西等地广泛分布, 但河南“怀庆府”的“怀地黄”最为著名, 质最优^[2]。地黄块根入药, 富含梓醇、糖类、维生素、多种氨基酸、多种微量元素, 具有防癌、抗癌、增强免疫力、延缓衰老、增强造血能力、降低血糖、防治糖尿病并发症的功能^[1,3]和重要经济价值^[4]。因此, 克隆地黄功能基因, 研究其功能和表达模式, 对进一步了解地黄药理分子机制和改良地黄品种具有重要意义。迄今, 国内外有少量地黄基因克隆、功能分析和表达模式研究的报道。用 RT-PCR 方法克隆地黄肌动蛋白基因片段, 可用作内参基因^[5]; 用 RT-qPCR 技术分析地黄中 7 个传统内参基因 18S、EF-1、ACT11、UBQ10、UBQ5、TUB5、GAPDH 和 4 个新报道的内参基因 PP2A、RPII、HSP90、TIP41 的 mRNA 表达差异情况^[6]; 用 SSH 方法克隆了地黄块根中 RgPR-10 基因, 研究其在地黄根和茎以及原核表达情况^[3,7]; 利用 Solexa 测序技术、生物信息学和荧光定量 PCR 分析从正茬地黄和重茬地黄中鉴定了 89 个 miRNAs, 预测了其中 7 个新型 miRNAs 的 15 个靶基因, 在正茬地黄和重茬地黄之间构建了 miRNAs 差异表达谱^[4,8]; 用 PCR 技术克隆了怀地黄基因片段及其中转座酶基因^[9]; 进行了地黄 ABA、GA 和部分寡糖合成代谢关键酶基因以及扩展蛋白基因 RgExpA1 的克隆与表达分析^[10-11]。由此可见, 地黄基因研究多限于内参基因, 功能基因及其表达特性研究很少。 β -酮酰辅酶 A 硫解酶 (3-ketoacyl CoA thiolase, KAT) 催化 β 氧化循环中硫解反应, 广泛存在于真、原核生物细胞中。KAT 在真核细胞中有两类, 分别位于线粒体和过氧化物体中^[16]。迄今, 已经从南瓜^[17]、

拟南芥^[18-19]、葡萄 (XM002285619.1)、番茄 (AK327019.1)、毛果杨 (XM002299248.1)、小麦 (AB539589.1) 和黄瓜 (CAA47926.1) 等植物中克隆了 KAT 基因。虽然 KAT 基因已经从上述植物中克隆, 但是, 地黄 KAT 基因及其表达模式研究尚未见报道。本研究用 RT-PCR 和 RACE 技术克隆怀地黄 KAT 基因, 用实时荧光定量 PCR 分析其时空表达规律, 为进一步研究其分子作用机制和在地黄分子育种中的应用奠定基础。

1 材料

怀地黄 *Rehmannia glutinosa* Libosch. f. *hueichingensis* (Chao et Schih) Hsiao 85-5 品种 30 日龄幼苗和开花 10 d 的植株由河南师范大学李敬敬、徐守真提供并鉴定。分别取前者的叶、根、茎和后者的根、茎、叶、雄蕊、雌蕊、花托、花瓣。

大肠杆菌 DH5 α 菌株为本实验室保存; RNA 提取主要试剂、DNA 凝胶回收试剂盒、DNA Marker 购自上海莱枫生物科技有限公司, 反转录试剂盒购自 New Industrytech 公司, LA Taq DNA 聚合酶、克隆载体 pMD19-T Vector、3'-Full RACE kit、5'-Full RACE kit 购自 TaKaRa 公司, 引物合成和基因测序由北京华大基因有限公司完成。

2 方法

2.1 引物设计与合成

根据 GenBank 中玉米 (HQ875338.1)、水稻 (DQ641272.1)、拟南芥 (NM_115853.4)、大白菜 (FJ171724.1) 和葡萄 (FJ171725.1) 的器官膨大基因 ARGOS 的序列设计引物 P1 用于扩增 KAT cDNA 中间片段; 根据其中间序列, 设计 P2 引物与 P3 引物进行 3'-RACE 和 5'-RACE; 根据三者拼接成的 cDNA 全长序列, 设计引物 P4, 扩增其开放阅读框 (ORF); 根据 ORF 序列, 设计引物 P5 及地黄内参基因 TIP41 引物 P6^[6]用于该基因的时空表达分析。所有引物 (表 1)

表1 所用引物来源、名称及其序列

Table 1 Origions, names, and their sequences of primers used

基因	引物名称	引物序列
KAT	P1	AATBGGRACYAAYARCARCATA
	P2	OuterF:GCCAAGTGAGTGATGGTGCTGGAGCT InnerF:CATGAAGAGAAGTGTTGCCATGC
	P3	OuterR: GTAATACCCATTGGAAGGAGAC InnerR: CACAGTTCTAACAGGTACAGTTTCAGGG
3'RACE kit	P2'	OuterR:TACCGTCGTTCCACTAGTGATT InnerR: CGCGGATCCTCCACTAGTGATTTCATATAGG
5'RACE kit	P3'	OuterF: CATGGCTACATGCTGACAGCCTA InnerF: CGCGGATCCACAGCCTACTGATGATCAGTCGATG
	P4	F: CGCGGATCCATGGAGAAAGCAACTG R: CCGCTCGAGCTTGCAAAGTGACAT
	P5	F: CATGAAGAGAAGTGTTGCCATGC R: TCGATGTCCCAAGTTCAAGACC
	P6	F: TGGCTCAGAGTTGATGGAGTGCT R: CTCTCCAGCAGCTTTCTCGGAGA

均使用 DNAMAN 4.0 和引物 Premier 5.0 软件设计。

2.2 总 RNA 的提取

地黄不同发育时期、不同组织的总 RNA 的提取参照上海莱枫生物科技有限公司的 RNA 提取试剂盒说明书操作。琼脂糖凝胶电泳测定 RNA 的质量, A_{260}/A_{280} 变化值计算 RNA 的浓度和纯度。

2.3 cDNA 第一链合成和 RT-PCR 中间片段的克隆、3'-RACE 和 5'-RACE

cDNA 反转录按照 New Ndustrytech 的第一链合成说明书进行。总 RNA 5 μ g, DEPC 水 6 μ L, Olig(dT)₁₈ 2 μ L, 70 $^{\circ}$ C、5 min, 冰上放置 2 min; RNase inhibitor 0.5 μ L, 5 $\times$ 第一链缓冲液 4 μ L, 0.1 mol/L DTT 1 μ L, 200 U M-MLV, 42 $^{\circ}$ C 温浴 50 min, 95 $^{\circ}$ C 变性 5 min, 加 DEPC 水稀释到 50 μ L 反应体系。

RT-PCR 中间片段的克隆 PCR 反应条件: 94 $^{\circ}$ C、5 min; 94 $^{\circ}$ C、30 s, 46 $^{\circ}$ C、30 s, 72 $^{\circ}$ C、1 min, 30 个循环; 72 $^{\circ}$ C、10 min; 4 $^{\circ}$ C 备用。25 μ L 反应体系为: 1 μ L cDNA、1.25 U La Taq 聚合酶、 1.25×10^{-8} mmol P1、200 μ mol/L dNTP。

3'-RACE: 按 TaKaRa 的 3'-Full RACE kit 进行。3'-Full RACE Outer PCR 程序为 94 $^{\circ}$ C、3 min; 94 $^{\circ}$ C、30 s, 55 $^{\circ}$ C、30 s, 72 $^{\circ}$ C、1 min, 20 个循环; 72 $^{\circ}$ C、10 min。3'-Full RACE Inner PCR 程序为 94

$^{\circ}$ C、3 min; 94 $^{\circ}$ C、30 s, 52 $^{\circ}$ C、30 s, 72 $^{\circ}$ C、1 min, 30 个循环; 72 $^{\circ}$ C、10 min。

5'-RACE: 按 TaKaRa 的 5'-Full RACE kit 进行。5'-Full RACE Outer PCR 程序为 94 $^{\circ}$ C、3 min; 94 $^{\circ}$ C、30 s, 50 $^{\circ}$ C、30 s, 72 $^{\circ}$ C、1 min, 20 个循环; 72 $^{\circ}$ C、10 min。5'-Full RACE Inner PCR 程序为 94 $^{\circ}$ C、3 min; 94 $^{\circ}$ C、30 s, 57.6 $^{\circ}$ C、30 s, 72 $^{\circ}$ C、1 min, 25 个循环; 72 $^{\circ}$ C、10 min。

PCR 产物经 1% 琼脂糖凝胶电泳检测, 回收特异条带, 连接、转化, 阳性克隆在北京华大基因有限公司测序。

2.4 RghKAT 全长 cDNA 编码区的克隆

应用 DNAMAN 4.0 软件对基因的 3'端、中间序列和 5'端序列进行拼接, 并在 NCBI 网站上预测其 ORF。根据 ORF 序列设计特异引物 P4 进行 ORF 扩增验证。PCR 反应条件: 94 $^{\circ}$ C、5 min; 94 $^{\circ}$ C、30 s, 58 $^{\circ}$ C、30 s, 72 $^{\circ}$ C、2 min, 35 个循环; 72 $^{\circ}$ C、10 min。25 μ L 反应体系为: 1 μ L cDNA、1.25 U La Taq 聚合酶、 1.25×10^{-8} mmol P4、200 μ mol/L dNTP。PCR 产物经 0.8% 琼脂糖凝胶电泳检测, 切下目的条带, 经回收后与 pMD19-T Vector 连接, 转化大肠杆菌 DH5 α , 鉴定并测序。

2.5 基因序列的生物信息学分析

基于克隆 cDNA 序列, 用 NCBI 网站 (<http://>

blast.ncbi.nlm.nih.gov/Blast.cgi)上 BLAST、BLASTP 软件进行同源基因的 cDNA、氨基酸序列搜索;用 NCBI 的 ORF Finder (<http://www.ncbi.nlm.nih.gov/gorf/gorf.html>) 分析基因的 ORF;利用生物软件 MEGA5.1 将同源序列根据它们的序列一致性高低构建系统发育树^[14];通过 ExPASy Proteomics Server 网站 (<http://www.expasy.ch/tools/protparam.html>) 分析蛋白的理化性质包括相对分子质量、等电点、消光系数、氨基酸组成等;SignalP 4.0 预测氨基酸残基组成的信号肽;SOPMA 在线预测氨基酸的二级结构以及功能基序;通过 NCBI 网站 (<http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml>) 预测氨基酸的保守域和预测氨基酸的三维结构 (<http://www.ncbi.nlm.nih.gov/Structure/MMDB/mmdb.shtml>)。

2.6 实时荧光定量 PCR

选用 2×UltraSYBR 混合物(含 ROX)试剂盒在 ABI7500 分析仪对 2 个时期、10 个组织 RghKAT 转录的 mRNA 丰度进行定量监测。不同组织 RghKAT 转录的 mRNA 拷贝数使用地黄内参基因 TIP41 进行校正,目的基因荧光定量 PCR 引物用 P5,内参基因的引物为 P6。采用 20 μL 反应体系,其中包括 1 μL cDNA 模板、10 μL 2×UltraSYBR 混合物(含 ROX)、5 μmol/L 引物,其余用水补齐。扩增条件为 95 ℃、10 min;95 ℃、15 s,60 ℃、1 min 共 40 个循环。溶解曲线分析 95 ℃、15 s,60 ℃、1 min,95 ℃、15 s,60 ℃、15 s。每个样本设 3

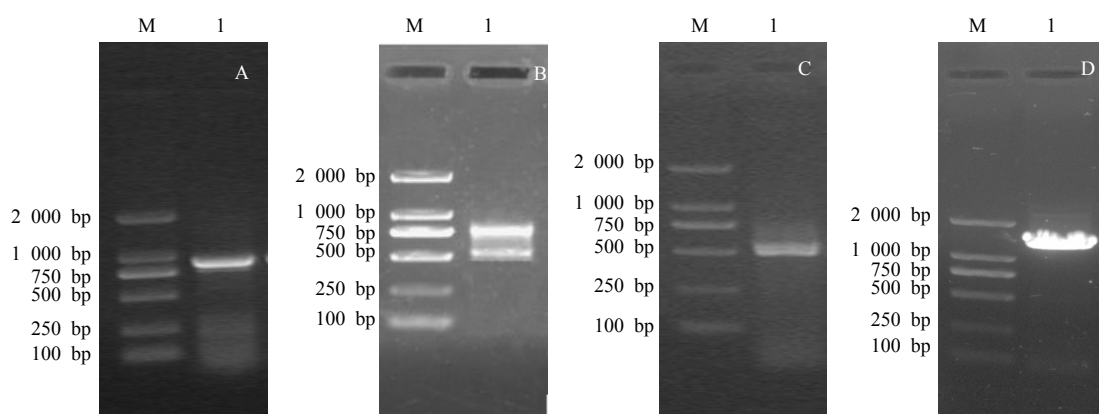
个重复,分别使用 P5 和 P6 进行扩增。使用溶解曲线判定产物特异性,根据定量扩增曲线达到平台期时的 Ct 值,以 TIP41 为参比计算各时期、各组织 RghKAT 转录的 mRNA 相对表达量。

3 结果与分析

3.1 RghKAT 全长 cDNA 和 ORF 的克隆

3.1.1 RghKAT 全长 cDNA 的克隆 以怀地黄 85-5 幼苗幼叶 RNA 逆转录的 cDNA 为模板,以简并引物 P1 为引物,进行中间片段的 PCR 扩增,经测序其核苷酸序列为 885 bp(图 1-A)。在 NCBI 上进行 BLAST 分析表明,该片段与多种植物的 KAT 具有较高的同源性。根据该中间片段,设计特异引物 P2 和 P3,结合 TaKaRa 3'-Full RACE kit、5'-Full RACE kit 自带的特异引物 P2'和 P3',采用 3'RACE 和 5'RACE 巢式扩增法分别扩增中间片段的 3'与 5'端(图 1-B、1-C),均进行克隆测序,结果显示 3'与 5'端分别为 250/177 bp 和 68 bp。经与其他物种的 KAT 比对,确认这两条序列为地黄 KAT 基因的 3'和 5'端序列。

3.1.2 RghKAT 基因全长 cDNA 编码区的克隆 根据 RghKAT 的中间片段、3'端和 5'端序列,利用 DNASTar 软件进行基因全长序列的拼接,然后在 cDNA 序列的 5'和 3'非编码区设计引物对 P4,扩增其 ORF。结果得到与预期片段大小一致的产物(图 1-D)。测序确认该片段为地黄 RghKAT 的 ORF 序列。将克隆的 1 713 bp 全长 cDNA 序列提交 NCBI



A-RghKAT 中间片段的电泳 B-3'-RACE 扩增片段的电泳 C-5'-RACE 扩增片段的电泳 D-RghKAT ORF 的电泳 M-DNA marker 1-扩增片段
A-electrophoresis of RghKAT middle fragments B-electrophoresis of 3'-RACE amplification fragments C-electrophoresis of 5'-RACE amplification fragments D-electrophoresis of RghKAT ORF M-DNA marker 1-amplification fragments

图1 RghKAT 基因中间片段(A)、3'端片段(B)、5'端片段(C)和 ORF(D) 的扩增片段电泳
Fig. 1 Electrophoreses of RghKAT middle fragments (A), 3'-end fragments (B), 5'-end fragments (C), and ORF (D) amplification fragments

公布, GenBank 登录号为 JX290369。

3.2 RghKAT 基因的生物信息学分析

3.2.1 RghKAT 基因的核苷酸序列及氨基酸序列分析

RghKAT 基因的 cDNA 全长序列为 1 640/1 713 bp, 包含 1 个 68 bp 的 5'UTR, 1 个 1 395 bp 的 ORF 和 1 个 250/177 bp 的 3'UTR。利用 ORFfinder 推导其编码 464

个氨基酸, 包含 20 种常见的氨基酸, 其氨基酸中色氨酸量最低, 为 0.2%; 丙氨酸量最高 11.9%, 其他氨基酸的量在 1.3%~10.6%。碱性氨基酸量略高于酸性氨基酸, 分别为 12.9%和 10.9% (图 2)。

3.2.2 多序列比对与系统进化树分析 利用基因 RghKAT (GenBank 登录号: JX290369) cDNA 全

```

1   gaaaatagagtgttaatttttcggtataatatacaaatgttaatttagagagaaagaatagagg
69   ATGGAGAAAGCAACTGAGAGGCAGAGAGT TTTGCTGCAGCATCTCCGTCCTTCACCTCGTCGCTCTCTA
1   M E K A T E R Q R V L L Q H L R P S F T S S S L
141  GAAGATATTGAATCTTCCGTTTCCGCATCTATATGTTTCAGCAGGGGACAGTGCAGCATACCATCGTTCTTCT
25   E D I E S S V S A S I C S A G D S A A Y H R S S
213  GTCTTTGGTGACGATGTGGTTATAGTGGCTGCCTATCGAACTGCACTGTGCAAGTCTAAGAGAGGTGGCTTC
49   V F G D D V V I V A A Y R T A L C K S K R G G F
285  AAAGATACCTATCCTGATGATTACTAGCACCTGTTTGGAGGCAGTGGTAGAAAAAACAAATGTAAATCCA
73   K D T Y P D D L L A P V L R A V V E K T N V N P
357  AATGAAGTTGGGGACATTGTTGTGGGCACGGTGTGGCACCGGGCTCCCAAAGAGCAAGTGAATGCAGGATG
97   N E V G D I V V G T V L A P G S Q R A S E C R M
429  GCTGCATTTTATGCTGGTTTCCCTGAAACTGTACCTGTTAGAACTGTGAACAGGCAATGTTTCATCTGGCCTT
121  A A F Y A G F P E T V P V R T V N R Q C S S G L
501  CAAGCTGTAGCTGATGTAGCTGCAGCTATCAAAGCTGGATTTTATGACATTGGGATTGGTGTCTGGGTTGGAG
145  Q A V A D V A A A I K A G F Y D I G I G A G L E
573  TCGATGACCGTTAATCCTATGGCTTGGGAAGGATCAGTCAATCCAAGAGTAAAATCGATGGCACAAGCACAA
169  S M T V N P M A W E G S V N P R V K S M A Q A Q
645  GATTGTCTCCTTCCAATGGGTATTACTTCAGAAAATGTCGCCCATCGTTTGGAGTGACAAGGCAGGAGCAA
193  D C L L P M G I T S E N V A H R F G V T R Q E Q
717  GATCAGGCTGCAGTTGATTCGCACAGAAAGGCTGCTGCTGCCTGCAATCAGGCAAATTTAAAGATGAGATA
217  D Q A A V D S H R K A A A A T A S G K F K D E I
789  ATACCAGTGAACCAAGATTGTGGACCCTAAATCTGGGGATGAGAAACCGGTTACGATATCTGTTGATGAT
241  I P V K T K I V D P K S G D E K P V T I S V D D
861  GGCATACGACCAATACCTCAGTGGCAGACCTGGGCAAGCTGAAACCTGTTTCAAGAAAGATGGCTCAACC
265  G I R P N T S V A D L G K L K P V F K K D G S T
933  ACTGCTGGGAATTCTAGCCAAGTGAGTGATGGTGTCTGGAGCTGTGCTCCTCATGAAGAGAAGTGTGCCATG
289  T A G N S S Q V S D G A G A V L L M K R S V A M
1005 CAGAAAGGGCTTCCCATCCTTGGTGTATTACAGGAGCTTTGCTGCAGTTGGTGTAGATCCTGCGATCATGGGT
313  Q K G L P I L G V F R S F A A V G V D P A I M G
1077 GTTGGTCCAGCC GTTGCAATTCCTGCAGCAGTCAAATCTGCTGGTCTTGAACCTGGGGACATCGATCTTTTC
337  V G P A V A I P A A V K S A G L E L G D I D L F
1149 GAGATAAATGAGGCATTTGCTTCGCAATTCGTGACTGCCGTAAGAAGCTTGAACCTGATCCTGAGAAGATC
361  E I N E A F A S Q F V Y C R K K L E L D P E K I
1221 AATGTGAATGGAGGTGCCATGGCAATTGGGCACTCTTGGGTGTTACAGGAGCTCGCTGTGTTGTACCTTA
385  N V N G G A M A I G H P L G V T G A R C V V T L
1293 TTGCATGAAATGAAGCGTCGTGGCAAAGATTGCCGCTTGGCGTGGTGTCAATGTGTATAGGCACTGGGATG
409  L H E M K R R G K D C R F G V V S M C I G T G M
1365 GGCGTGCTGCTGTTTTTGAAGGGGTGACTCCTGTGACGAGCTTTCCAATACACGAACAGTCGGAACACAC
433  G A A A V F E R G D S C D E L S N T R T V G T H
1437 AATTTTCTGTCCAAGGATGCTCGA TAAactccaatgtccacttgcaggtccaataaattttaaatctgca
457  N F L S K D A R *
1509 ttcagaagtcaataataactgcaataaattacttactgattatcacacagctgtggtgcagtttaatga
1581 tctctgtatttttctgactgatgtatttctgtacttggatcattttgccccttttttctctata
1653 atctataatgataaagcatgtgagtgatgaatactaccaattttatcaaaaaaaaaaaaa

```

小写字母为 5'端非编码区和 3'端非编码区, 大写字母为编码区; 上面为核苷酸序列, 下面为氨基酸序列; *表示终止子; 起始密码子和终止密码子用方框标出; 多聚腺苷酸尾用粗斜体表示; 下划线标出 1 640/1 713 bp 全长 cDNA 不同部分

lowercase letters show 3', 5'-untranslated regions, capital letters show coding region sequence; upper sequences indicate nucleotide and lower show amino acid; * refers to stop codon; start codon and the stop codon are in box; Poly adenine nucleotides tail are in bold italic; different regions of 1 640/1 713 bp full-cDNA are underlined

图 2 RghKAT 基因的核苷酸及其编码蛋白质的氨基酸序列

Fig. 2 Nucleotide sequences of RghKAT gene and amino acid sequences of their encoding proteins

长序列与来自葡萄 (XM002285619.1)、番茄 (AK327019.1)、毛果杨 (XM002299248.1)、拟南芥 (NM 100351.4) 和小麦 (AB539589.1) 等植物的 KAT 的碱基序列进行比对分析, 序列的相似性分别高达 84%、82%、82%、79% 和 73%。利用基因 RghKAT 预测的氨基酸序列与来自矮牵牛 (ACV70033.1)、葡萄 (XP002285653.1)、黄瓜 (CAA47926.1)、拟南芥 (NP180873.1)、小麦 (BAI66423.1) 的 KAT 的氨基酸序列进行比对分析, 同源性分别为 88%、88%、86%、87% 和 78% (表 2)。

可见, 基因 RghKAT 碱基序列及其编码蛋白质的氨基酸序列与双子叶植物矮牵牛、葡萄的同源性最高。利用 MEGA5.1 软件 Neighbor-Joining 法构建 RghKAT 和其他物种 KAT 的氨基酸同源序列的系统进化树 (图 3), 进行系统分析。其氨基酸序列与 BLAST 检索到的真菌、动植物的 KAT 的氨基酸序列的同源性大于 53%, 说明其在生物进化上的保守性很大。但是, 它与双子叶植物的亲缘关系最近, 与单子叶植物的亲缘关系次之, 与绿藻、真菌和动物由近到远的亲缘关系依次为绿藻、真菌和动物。

表 2 地黄与其他物种 KAT 的同源性比较

Table 2 Homologous comparison of KAT between *R. glutinosa* f. *hueichingensis* and other species

物 种	GenBank 登录号	氨基酸同源性 / %
矮牵牛 <i>Petunia x hybrida</i>	ACV70033.1	88
葡萄 <i>Vitis vinifera</i>	XP002285653.1	88
芒果 <i>Mangifera indica</i>	CAA53078.1	88
拟南芥 <i>Arabidopsis thaliana</i>	NP180873.1	87
黄瓜 <i>Cucumis sativus</i>	CAA47926.1	86
南瓜 <i>Precursor-cucurbit</i>	S72532	86
大白菜 <i>Brassica napus</i>	CAA63598.1	86
向日葵 <i>Helianthus annuus</i>	AAQ77242.1	86
蓖麻 <i>Ricinus communis</i>	XP002518136.1	85
大豆 <i>Glycine max</i>	XP003555712.1	83
油棕 <i>Elaeis guineensis</i>	AEQ94132.1	81
短柄草 <i>Brachypodium distachyon</i>	XP003570403.1	79
玉米 <i>Zea mays</i>	NP001241698.1	79
小麦 <i>Triticum aestivum</i>	BAI66423.1	78
蒺藜苜蓿 <i>Medicago truncatula</i>	XP003604976.1	78
衣藻 <i>Chlamydomonas reinhardtii</i>	XP001697225.1	66
微单胞菌属 <i>Micromonas</i> sp. RCC299	XP002500295.1	64
胶球藻 <i>Coccomyxa subellipsoidea</i> C-169	EIE21619.1	63
马 <i>Equus caballus</i>	XP001488609.1	60
天竺鼠 <i>Cavia porcellus</i>	XP003464147.1	60
原鸡 <i>Gallus gallus</i>	NP001184217.1	59
网柄菌 <i>Dictyostelium discoideum</i> AX4	XP643924.1	59
变色龙 <i>Anolis carolinensis</i>	XP003221941.1	57
小鼠 <i>Mus musculus</i>	NP666342.1	57
大鼠 <i>Rattus norvegicus</i>	BAA14107.1	56
罗非鱼 <i>Oreochromis niloticus</i>	XP003456538.1	56
猕猴 <i>Macaca mulatta</i>	XP001088834.2	56
人 <i>Homo sapiens</i>	NP001598.1	56
猩猩 <i>Pongo abelii</i>	XP002813979.1	56
熊猫 <i>Ailuropoda melanoleuca</i>	XP002914680.1	56
野猪 <i>Sus scrofa</i>	XP003132151.1	56
普通狨 <i>Callithrix jacchus</i>	XP002755029.2	55
紫海胆 <i>Strongylocentrotus purpuratus</i>	XP003724378.1	53

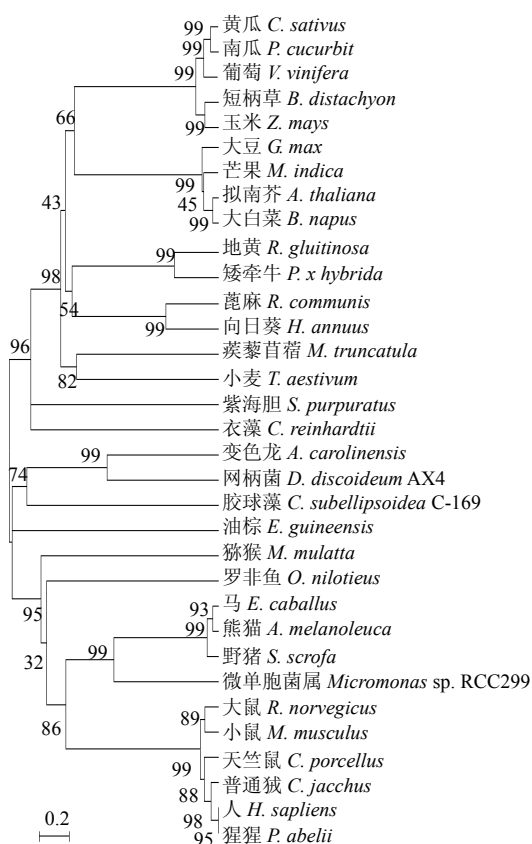


图3 基于 Neighbor-Joining 法构建地黄和其他物种全长 KAT 氨基酸序列的系统进化树

Fig. 3 Phylogenetic tree of full-length KAT amino acid sequences between *R. glutinosus* f. *hueichingensis* and other species using Neighbor-Joining method

3.2.3 RghKAT 编码蛋白质结构预测 RghKAT 编码蛋白质特征 ProtParam 分析显示, 相对分子质量为 4.9×10^4 , 等电点为 7.98, 摩尔消光系数为 15 065, 不稳定系数为 31.58, 为略碱性的稳定蛋白质。SignalP4.0 显示地黄 KAT 的 N 端存在一个由 70 个氨基酸残基组成的信号肽 (图 4-A), 信号肽序列为 “MEKATERQVRLLQHLRPSFTSSSLEDIESSVSASIC” 和 “SAGDSAAYHRSSVFGDDVVIVAAYRTALCKSK-RG”, 分界位点在 70 和 71 号氨基酸残基之间。用 SOPMA 预测地黄 KAT 的二级结构 (图 4-D), 结果为 α -螺旋为 38.58%, 不规则卷曲 35.78%, β -折叠为 15.30%, β -转角为 10.34%, 这 4 个二级结构构成地黄 KAT 的基本结构。氨基酸的保守域见图 4-C; 氨基酸的三维结构见图 4-B, RghKAT 蛋白具有硫解酶典型的特征序列。

3.3 RghKAT 时空表达差异分析

为了明确 RghKAT 在地黄不同组织中的表达模

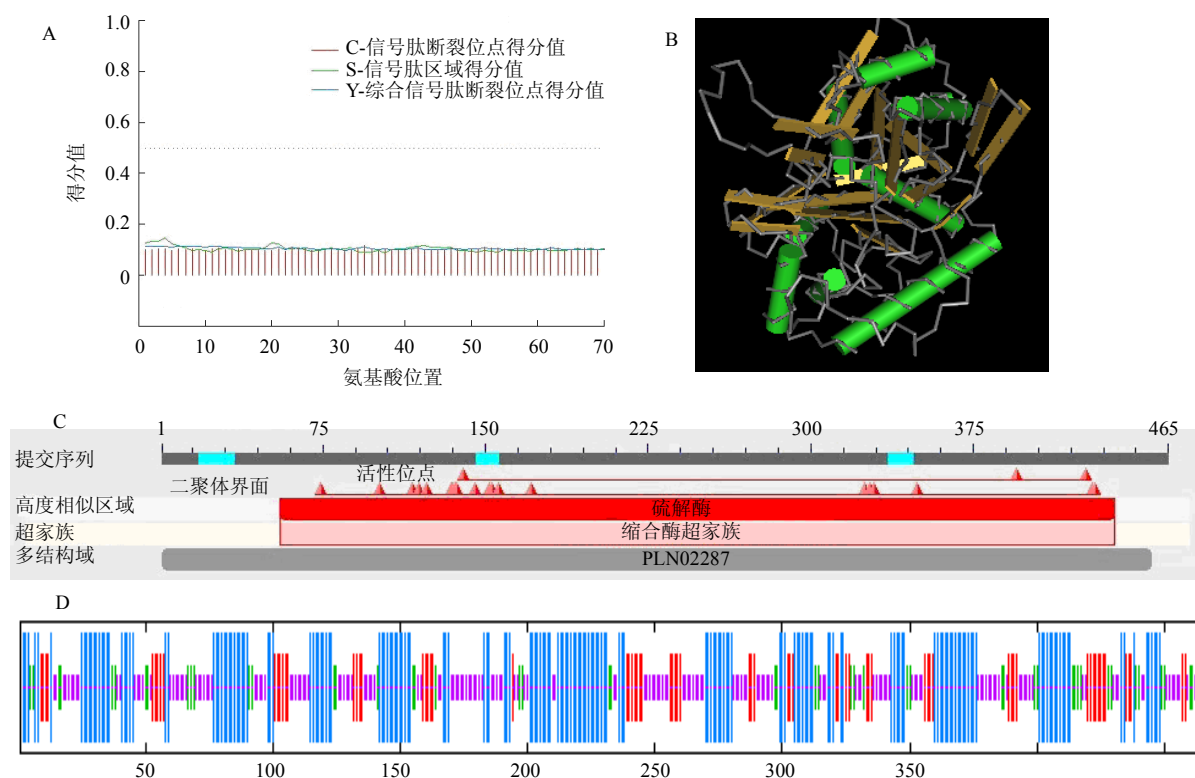
式, 分别提取怀地黄组培幼苗期的根、茎、叶和盛花期的花瓣、花托、雄蕊、雌蕊、根、茎与叶的总 RNA, 反转录合成 cDNA 后进行 qRT-PCR 分析 (图 5)。结果发现, RghKAT 在被检测的各个组织中均有表达 mRNA, 但是其不同组织中表达量存在显著差异。其中, 盛花期花瓣的表达量最高, 其次是花托、雄蕊、根、茎、叶, 雌蕊中的表达量较弱; 组培幼苗茎的表达量较高, 根次之, 叶的表达量最低; 比较其在幼苗期和盛花期的根、茎、叶中的表达, 发现其在幼苗期茎中的表达高于盛花期茎的表达, 而其在幼苗期根和叶中的表达低于盛花期根和叶的表达。

4 讨论

KAT (EC 2.3.1.16) 是在真菌、动植物细胞中广泛存在的、具有催化脂肪酸 β -氧化反应中硫解断裂作用的一些酶^[12,16]。在 CoASH 存在下, β -酮脂酰 CoA 在该酶催化下生成一分子乙酰 CoA 和少了两个碳原子的脂酰 CoA。迄今, 已经从南瓜^[13]、拟南芥^[14-15]、葡萄 (XM002285619.1)、番茄 (AK327019.1)、毛果杨 (XM002299248.1)、小麦 (AB539589.1) 和黄瓜 (CAA47926.1) 等植物中克隆了 KAT 基因。但是, 尚未见克隆地黄 KAT 基因的报道。

在植物中已经发现一些与植物器官膨大有关的功能基因 ARGOS (auxin-regulated gene involved in organ size)^[18-19]。本研究根据 NCBI GenBank 核酸数据库中 5 种已知植物 ARGOS 基因碱基序列的保守区, 设计几对简并引物, 采用 RT-PCR 和 RACE 技术克隆了几个新的地黄功能基因。其中之一是 β -酮脂酰辅酶 A 硫解酶基因 RghKAT。

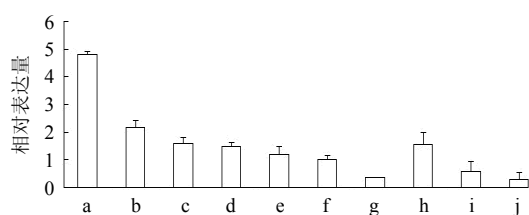
该基因编码的蛋白质具有硫解酶典型的特征序列, N 端含一个信号肽, 可以推测克隆的 RghKAT 是完整的 KAT 基因。序列同源性分析表明, RghKAT 与已知的葡萄、黄瓜、拟南芥、小麦等植物 KAT 具有高度同源性, 在核苷酸水平上序列同源性高达 73%~84%, 其编码的氨基酸同源性高达 78%~88%, 进一步说明 RghKAT 蛋白具有相当高的结构保守性, 可归于 KAT 酶类。另外, 葡萄、黄瓜和拟南芥依次分属于双子叶植物葡萄科、葫芦科和十字花科, 其间的亲缘关系远; 小麦属于单子叶植物, 其间的亲缘关系更远。然而, 它们的 KAT 同源性之高表明 KAT 可能是进化上很保守的蛋白。该结果与彭丹等^[19]的结果相似。



A-信号肽预测 B-蛋白二级结构预测 (蓝色表示 α -螺旋, 绿色表示 β -转角, 红色表示 β -折叠, 紫色表示不规则卷曲) C-保守区预测 D-三维结构预测
A-prediction of signal peptide B-prediction of secondary structure (blue bars show α -helix, green bars show β -extended strand, red bars show β -turn, purple bars show random coil) C-prediction of conserved domain D-prediction of three-dimensional structure

图 4 RghKAT 蛋白二级结构、保守区和三维结构预测

Fig. 4 Prediction of secondary structure, conserved domain, and three-dimensional structure of RghKAT proteins



a-花瓣 b-花托 c-雄蕊 d~f-盛花期根、茎、叶 g-雌蕊
h~j-幼苗期茎、根、叶
a-petals b-torus c-stamens d~f-roots, stems, and leaves of blooming stage g-pistils h~j-stems, roots, and leaves of seedling stage

图 5 地黄组培幼苗期、盛花期各组织 KAT 表达的 qRT-PCR 监测

Fig. 5 Expression of KAT transcript in various tissues of *R. glutinosah f. hueichingensis* at seedling and blooming stages by qRT-PCR monitoring

实时荧光定量 PCR 技术分析 RghKAT 的时空表达模式, 发现其在怀地黄幼苗期和盛花期的 10

个不同组织中均差异表达, 但在花瓣中表达量最高, 在幼苗叶中表达量最低。这与小麦 Ta14S 基因的时空表达模式^[20]相似。比较其在幼苗期和盛花期的根、茎、叶中的表达, 发现其在幼苗期茎中的表达高于盛花期茎的表达, 而其在幼苗期根和叶中的表达低于盛花期根和叶的表达。这是首次报道 KAT 基因在植物中时空表达模式。

本研究怀地黄 RghKAT 的克隆和时空表达分析结果为进一步探讨其功能以及分子作用机制提供了重要基础。

参考文献

- [1] Young S K, Jin A R, Byoung S K. Discrimination of *Korean Rehmannia glutinosa* from Chinese *Rehmannia glutinosa* using sequence-characterized amplified region marker [J]. *Korean Soc Appl Biol Chem*, 2012, 55: 1-6.
- [2] Zhou Y Q, Gu F P, Zhou C N, *et al*. Genetic diversity of *Rehmannia glutinosa* cultivars based on sequence-related amplified polymorphism markers [J]. *Sci Hort*, 2010,

- 125: 789-794.
- [3] Peng S, Guo Y H, Qi J J, *et al.* Isolation and expression analysis of tuberous root development related genes in *Rehmannia glutinosa* [J]. *Mol Biol Rep*, 2010, 37(2): 1069-1079.
- [4] Yang Y H, Chen X J, Chen J Y, *et al.* Differential miRNA expression in *Rehmannia glutinosa* plants subjected to continuous cropping [J]. *BMC Plant Biol*, 2011, 11: 53-56.
- [5] 孙 鹏, 郭玉海, 祁建军, 等. 地黄肌动蛋白基因片段的克隆与序列分析 [J]. 安徽农业科学, 2008, 36(20): 8470-8471.
- [6] 侯维海, 孙 鹏, 陈全家, 等. 地黄实时定量 PCR 内参基因的筛选 [J]. 中国农学通报, 2011, 27(17): 76-82.
- [7] 孙 鹏, 郭玉海, 祁建军, 等. 地黄 *RgPR-10* 基因的克隆与表达 [J]. 西北农业学报, 2009, 18(1): 300-304.
- [8] Yang Y H, Chen X J, Chen J Y, *et al.* Identification of Novel and Conserved microRNAs in *Rehmannia glutinosa* L. by Solexa sequencing [J]. *Plant Mol Biol Rep*, 2011, 29: 986-996.
- [9] 周延清, 姚焕灵, 段红英, 等. 怀地黄基因片段及其中转座酶基因的克隆与序列分析 [J]. 河南农业科学, 2012, 41(1): 106-109.
- [10] 侯维海. 地黄 ABA、GA 和部分寡糖合成代谢关键酶基因克隆与表达分析 [D]. 乌鲁木齐: 新疆农业大学, 2011.
- [11] 臧亚超, 孙 鹏, 杨太新, 等. 地黄扩展蛋白基因 *RgExpA1* 的克隆与表达分析 [J]. 生物技术通报, 2012, 4: 420-25.
- [12] Ramasubramanian S, Elena M, Magnus S, *et al.* The crystal structure of a plant 3-ketoacyl-CoA thiolase reveals the potential for redox control of peroxisomal fatty acid β -oxidation [J]. *Mol Biol*, 2006, 359: 347-357.
- [13] Akira K, Makoto H, Yuka T, *et al.* cDNA cloning and expression of a gene for 3-ketoacyl-CoA thiolase in pumpkin cotyledons [J]. *Plant Mol Biol*, 1996, 31: 843-852.
- [14] Germain V, Rylott E L, Larson T R, *et al.* Requirement for 3-ketoacyl-CoA thiolase-2 in peroxisome development, fatty acid β -oxidation and breakdown of triacylglycerol in lipid bodies of *Arabidopsis* seedlings [J]. *Plant J*, 2001, 28(1): 1-12.
- [15] Chris C, Murcha M W, Millar A H, *et al.* Nine 3-ketoacyl-CoA thiolases (KATs) and acetoacetyl-CoA thiolases (ACATs) encoded by five genes in *Arabidopsis thaliana* are targeted either to peroxisomes or cytosol but not to mitochondria [J]. *Plant Mol Biol*, 2007, 63: 97-108.
- [16] Marco F, Ségolène A, Marie C C, *et al.* A role for the peroxisomal 3-ketoacyl-CoA thiolase B enzyme in the control of PPAR α -mediated upregulation of SREBP-2 target genes in the liver [J]. *Biochimie*, 2011, 93: 876-891.
- [17] Gonzalez N, Gerrit T S B, Dirk Inze, *et al.* What can the tiny weed *Arabidopsis* teach us to improve biomass production in crops [J]. *Curr Opin Plant Biol*, 2009, 12: 157-164.
- [18] Krizek B A. Making bigger plants: Key regulators of final organ size [J]. *Curr Opin Plant Biol*, 2009, 12: 17-22.
- [19] 彭 丹, 周 波, 赵小英, 等. 甘蓝型油菜 *BnCOPI* 基因编码区全长 cDNA 的克隆与功能研究 [J]. 中国生物化学与分子生物学报, 2011, 27(1): 69-77.
- [20] 任江萍, 刘海伦, 王新国, 等. 小麦逆境胁迫相关基因 *Ta14S* 的克隆及表达分析 [J]. 中国农业科学, 2012, 45(6): 1226-1234.