

## 研究报告

## Research Report

# 大豆 KCS 基因的克隆及结构预测

赵艳✉, 邱增成, 翟莹, 徐红红, 国春晖

齐齐哈尔大学, 生命科学与农林学院, 齐齐哈尔, 161006

✉ 通讯作者, zhaoyan3053877@yahoo.com.cn; 作者

豆科基因组学与遗传学, 2013 年, 第 4 卷, 第 6 篇 doi: 10.5376/lgg.cn.2013.04.0006

收稿日期: 2016 年 11 月 18 日 接受日期: 2016 年 11 月 18 日 发表日期: 2016 年 11 月 18 日

© 2013 BioPublisher 生命科学中文期刊出版平台

本文首次发表在《分子植物育种》2013 年, 第 11 卷, 第 1 期上。现依据版权所有人授权的许可协议, 采用 [Creative Commons Attribution License](#), 协议对其进行授权, 再次发表与传播。只要对原作有恰当的引用, 版权所有人允许并同意第三方无条件的使用与传播。

**摘要** 以大豆叶片总 RNA 为模板, 利用 RT-PCR 方法, 克隆获得大豆 KCS 基因序列。采用生物信息学方法对其进行预测分析, 结果表明: 大豆 KCS 基因的 CDS 序列长度为 1 533 bp, 编码 510 个氨基酸; 大豆 KCS 基因编码的蛋白质理论分子量为 57.1 KD, 等电点为 9.03; 该基因编码的蛋白具有两个完全跨膜结构; 没有信号肽; 其蛋白质二级结构中  $\alpha$  螺旋占 47.65%, 无规则卷曲占 35.88%,  $\beta$  折叠占 16.47%; 在进化关系上, 与油茶、菟葵、苜蓿的亲缘关系相对较近, 该研究结果可为大豆 KCS 基因结构和功能的进一步研究提供理论参考。

**关键词** 大豆; KCS; 基因克隆; 结构预测

## Cloning of Soybean KCS Gene and Predicting of Its Structure

Zhao Yan✉, Zhai Ying, Qiu Zengcheng Xu, Honghong, Guo Chunhui

College of Life Science and Agroforestry, Qiqihaer University, Key Laboratory of Genetics, Qiqihaer, 161006

✉ Corresponding author, zhaoyan3053877@yahoo.com.cn; Authors

**Abstract** The sequence of soybean KCS gene was cloned by RT-PCR with the total RNA of soybean leaf as template and predicted by bioinformatics. The results showed that the coding sequence of soybean KCS gene consists of 1 533 bp nucleotides that encodes 510 amino acids. The molecular weight of the protein is 57.1 kD and the isoelectric point is 9.03. There are two full transmembrane domain and without signal peptide. The secondary structure comprises of 47.65% alpha helix, 35.88% random coil and 16.47% extended strand. The evolutionary relationship demonstrated that soybean has relatively close relationship with *Camellia oleifera*, *Eranthis hyemalis* and *Medicago truncatula*. This study would provide valuable information of supporting the further research in the field of structure and function in soybean.

**Keywords** Soybean, KCS, Gene cloning, Structure prediction

超长链单不饱和脂肪酸(VLCFAs)在生物体中起着综合的作用, 是三酰甘油、膜脂、鞘脂的必要生物成分, 也是蜡质层复合物的前体(Lee et al., 2009)。VLCFAs 的生物合成由脂肪酸酰基辅酶 A 延长酶(FAE)催化完成, FAE 是一种膜结合酶复合物, 包括 3-酮脂酰辅酶 A 合成酶(KCS), 3-酮脂酰辅酶 A 还原酶(KCR),  $\beta$ -羟脂酰-CoA 脱水酶(HCD)和反式烯脂酰(ECR)(倪郁和郭彦军, 2008; Joubès et al., 2008)。在所有植物及微生物的脂肪酸延长酶系统中都普遍存在着以上四种酶, 但 KCS 基因起着关键的作用(杨慧等, 2012)。

KCS 基因最早在拟南芥中鉴定(Rossak et al., 2001), 之后在芸苔属芥菜、油菜, 油桐属的霍霍巴(Blacklock and

Jaworski, 2002), 旱金莲属旱金莲(Mietkiewska et al., 2004)等植物中都获得 KCS 基因。研究表明, 有些植物的 KCS 基因表达增加了长链脂肪酸的含量。2009 年, Guo 等将银扇草 KCS 基因在转基因拟南芥中表达, 种子油脂肪酸成分分析表明神经酸的含量增长 30 倍, 相似的转基因实验使埃塞俄比亚芥种子油中神经酸的含量增长 7~10 倍。拟南芥 KCS 基因在强半乳糖诱导型启动子调控下, 在酵母转化株中积累了 20:1、22:1 以及 24:1 的脂肪酸, 这些在非转化酵母培养物中不存在(Millar and Kunst, 1999)。本研究中根据已获得的植物中的 KCS 基因的序列, 在大豆基因组(<http://www.phytozome.net/soybean>)上同源搜索大豆 KCS 基

因, 克隆该基因, 并进行生物信息学预测分析, 旨在为该基因结构和功能的进一步研究提供理论依据。

## 1 结果与分析

### 1.1 大豆 KCS 基因序列的克隆

在公布的大豆基因组序列中, 根据其他植物 KCS 基因的氨基酸序列, 搜索得到的大豆 KCS 基因的序列位于大豆基因组第 14 条染色体的 6112931 至 6114767 序列之间。根据该序列设计引物 D1 和 D2, 以大豆叶片总 RNA 为模板, RT-PCR 扩增获得约 1 500 bp 的条带(图 1)。将该片段连接到 pMD18-T 载体上, 获得重组质粒 pMD18-T-KCS。将其转入大肠杆菌中, 提取质粒进行鉴定, 用 *EcoR* I 和 *Sal* I 双酶切鉴定, 得到与预期相符的片段(图 1)。利用 DNAMAN 软件, 将测序结果与大豆基因组上该段序列进行同源性比对, 同源性达到 100%, 说明该段序列(图 2)扩增正确, 且在不同大豆品种间的该基因无差别。由图 2 可知, 大豆 KCS 基因的 CDS 全长为 1 533 bp, 编码的蛋白质具有 510 个氨基酸, 序列数据已经登录(GenBank accession No. JX855159)。

### 1.2 大豆 KCS 基因的生物信息学预测

用 ProtParam 得知, 该蛋白的分子量 57.1 KD, 等电点(pI)为 9.03, 带正电残基(Arg+Lys)总数为 56, 负电残基(Asp+Glu)为 45, 总的亲水性平均系数为 0.035, 预测该蛋白属于疏水性蛋白。其氨基酸组成见表 1。

通过 TMHMM 对大豆 KCS 编码的蛋白进行跨膜区预测, 结果如图 3 所示。该序列包含 6 个跨膜区, 只有 TM1 和 TM2 完全跨膜, 分别位于第 43~65, 77~99 氨基酸残基之间; 利用 SignalP4.0 预测得知, 大豆 KCS 编码的蛋白不存在信号肽; 用 ProductProtein 预测蛋白的二级结构表明, 此蛋白的  $\alpha$  螺旋占 47.65%, 无规则卷曲占 35.88%,  $\beta$  折叠占 16.47%。

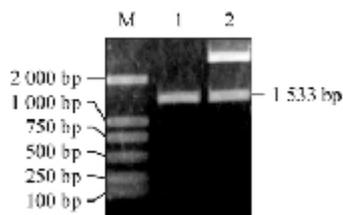


图1 PCR 扩增产物及pMD18-T-KCS 酶切鉴定

注: M: 2 000 bp DNA marker; 1: PCR 产物 2: *EcoR* I 和 *Sal* I 双酶切片片段

Figure 1 The product of PCR amplification and identification of pMD18-T-KCS by restriction enzyme digestion

Note: M: 2 000 bp DNA marker; 1: PCR amplification; 2: *EcoR* I and *Sal* I digestion of pMD18-T-KCS

表 1 大豆 KCS 编码蛋白的氨基酸组成

Table 1 The amino acid composition of KCS encoding protein of soybean

氨基酸	比例	氨基酸	比例
Amino acid	Proportion	Amino acid	Proportion
Ala(A)	7.1	Phe(F)	4.7
Arg(R)	5.3	Pro(P)	4.7
Asn(N)	4.5	Ser(S)	7.3
Asp(D)	4.5	Thr(T)	4.9
Cys(C)	2.2	Trp(W)	1.2
Gln(Q)	3.7	Tyr(Y)	3.1
Glu(E)	4.3	Val(V)	7.8
Gly(G)	5.9	Pyl(O)	0
His(H)	2.5	Sec(U)	0
Ile(I)	5.7	Asx(B)	0
Leu(L)	12.4	Glx(Z)	0
Lys(K)	5.7	Xaa(X)	0
Met(M)	2.5		

利用 MEGA5.05 构建 KCS 家族的蛋白质序列系统进化树, 家族中各成员的进化关系如图 4 所示, 结果表明大豆 KCS 和油茶的 KCS 的亲缘关系最近, 来自于同一个分枝, 其次是菟葵、苜蓿, 来自同一个次分枝。从该图中可以看出, 根据 KCS 构建的系统进化树中, 比较符合传统的系统分类, 玉米和小麦在同一个次分枝, 同属禾本科植物; 花生和棉花来自同一分枝, 都属于油料作物; 银扇草、水芹、海甘蓝、拟南芥、芥菜、独行菜、小果亚麻芥和亚麻芥来自同一分枝, 都属于十字花科植物。

## 2 讨论

KCS 基因在调控长链脂肪酸合成中起着关键的作用, 在多种植物中发现并克隆, 但大豆 KCS 基因的获得鲜有报道。本文利用同源搜索克隆了大豆 KCS 基因的序列, 并通过生物信息学相关软件对该基因编码蛋白的理化性质、信号肽、跨膜结构域、蛋白质二级结构及系统进化等方面进行预测, 这些初步预测的结果有助于有目的性的进行该基因所编码蛋白的功能研究。

研究表明某些植物的 KCS 基因表达主要集中在种子中 (Mietkiewska et al., 2004)。本文中根据 KCS 构建系统进化树的植物来自于豆科植物、主要的粮食作物、油料作物及含有较高浓度的芥子油的十字花科植物, 因此 KCS 基因极可能在这些植物种子长链脂肪酸的积累中具有重要作用。由图 4 结果可知, 多数已知植物的分类符合传统的系统分类, 说明该进化树的构建合理。

KCS 基因的表达主要集中在种子中, 而且有些植物的 KCS 基因的启动子是种子特异性启动子 (Rossak et al., 2001; Guo et al., 2009), 大豆 KCS 基因的表达方式的研究, 及其启动子的调控特性(即是否具有种子特异表达的特性), 也将是进一步研究的方面。

```
ATGACAGTGACGATGAGTGGAGAAGAA3AAGCTGCGGTGGG3AGTTCCAGATCCAAACAGAAGAGCAGAATGGTGCCTCCAGATTTCCTCCAGA
GCCTGAACTCAAGTACGCTGAAATTAGGTTACCCTACTTGTATCTCCAACTCGTTTACCTTATTCCTGGTCCCTTTGATCCTCGTCAAGCT
AATCCAGGTTTTCAGACCACCGACCTCCGCGACCTTCGGTCCACCTCCACATACAACTCCCTCACCATCTCCACCTCCGCTCCGCGTCCCTC
GTTTTCGGCTTACCTTCTACGCGCTTACCTTCCCCACGCGCCGCTTACCTTCCCTCGACTCCGCTTCCGCTTCCGCGCCCGCCGACCACTCAAGG
CCCCCTTCCGGAGCTTCATGGATCACTCCCGCCTCCACGGGAGACTTCGAGGAGTCTCCCTGGAGTTCCAGCGGAGATCCTCGAAGCGCTC
TGGCTCCGGGAGGAAACCTACCTCCCGACCGCCATGCACTCTATTCCCGCCCGAGCCCTCCATGCGCCCTCCAGCGCCCGAGGCGGAAACAG
GTATGTTTGGTTCCTTGGACAACCTCTTCCAAAGCACCAACATAAAACCAAAGGATATTGATATCTTATTTGTGAATTCAGCTTGTITIA
ACCTACCCCTCACTGTCTTCCATGATTTGTTAAACAAGTACAAATTAAGGG3TAACATTAGGAGCTTTAATTTGGGAGGGATGGGGTGCAG
TGCTGGAGTCACTTGTGTGATCTTGTCTAAGGACCTTTTGCAGGTCATAG3AACACATATGCTGTGGTGTGTAGCACTGAGAACATTACT
CA3AATTGGTACTTTTGGAAACAAAGAAATCCATGCTCATTCCAAATTCCTTTTTCGCGTTTCGCTGCTCCGCTGCTGCTCTCCAAACAAGC
CGCGGATCGAAGGAGGGCAAGTACC3GCTTGTCCATTTGTGAGGACTCATCGCGGTGCGGACGACAAAGGCAITTCGCGTGTGTTTACCA
GGAGCAGGATGATGCAAGGAGACTGTTGTTTCTTGTCTAAGGATTTGATGGCAATTTGCTGGTGGAGCATTGAAGACCAACATCACAACA
CTTGGCCCTTTTGGTGTGTTCCAAATTAGTGGAGCAGCTTGTGTTTTTGGTGGACTTGTGCTGATGAAGAAGTGTGTTAAGGCTGATGTGAAGCCIT
ACATACCGGATTTCAAGCTTGTCTTTTGTATCATTTCTGCATCCATGCGCGGG3CAGGGCCGTGATTGATGAGTTGGAGAAGAACCTGCAGCT
GCTTCTGAGCATGTGAGGCTTCTAGAATGACCCCTCATAGATTTGGGACACTTCTCAGCTCCATTTGGTATGAGTTGGCTTACAT
GAAGCCAAAGG3AGGATCAAGAAGGTTAACAGGATTTGCCAAATTTGTTTTCGCAATGCTTTTCAAATGTAACACTCCGCTTTGGCAGGCTC
TGAGAATGTGAGGCTTCTCTAATGGACCATGGGAAGATTGATGATAAGTATCCTGTGGAAATAGTCTCATAG
```

编码的蛋白质:  
Coding protein

```
MIVTMSDEEEAAVQVIQKSRMVLPDFLQSVNLRVVKLOYHYLISNLTFLVPLLIVTLIQVSTTDLRHLWLHLQYNLLTILTCSAVL
VFLTLIYAVTQPRVYLLDSACFRPADHLKAPFRSEFMHSRLTGDFFEESELEFQRKILERSGLGEETVVPDAMHSIPFQPSMAAARAEAEQ
VMFGALDNLFSQINIKPDIIGLIVNCSLNFPTFSLSEMIVNYKLRGNIRSFNLGGMGCSAGVIAVDLAKDLIQVHENTYAVVSTENIT
QWYFQNKZSMLIFNCLFRVQCSVLLSNKPADRRRAKYRLVHVVRTHRGADDKAFRCVYQEQQDDAGKTGVSLSKDLMAIAGGALKTNIT
LGPLVLPISQQLLFFVTLMLKMLFKADVRFYIPDFKLAFDHFCIHAGGRAVIDELEKMLQLLPEHVEASRMTLHRFGTSSSSIWYELAYI
EAKGEIRKENRIWQIAGSGFKNSAVWQALRNVRPSPNGPWEDCIDKYPVEIVS
```

图 2 大豆 KCS 基因的 CDS 及蛋白质的序列

Figure 2 The sequences of CDS and protein of soybean KCS gene

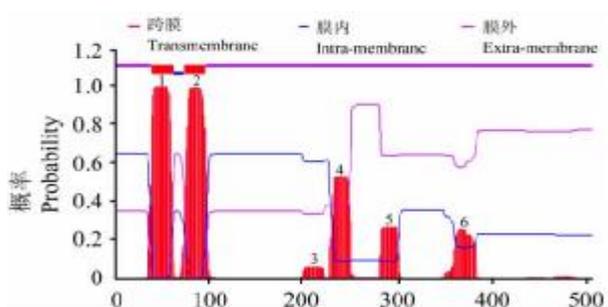


图 3 大豆 KCS 编码蛋白的跨膜结构分析

Figure 3 Transmembrane analysis of the KCS encoding protein of soybean

### 3 材料与方法

#### 3.1 供试材料

大豆品种吉豆 2 号、大肠杆菌 DH 5 $\alpha$  为本实验室保存; 限制性内切酶 *EcoR* I、*Sal* I、pMD18-T 克隆载体、ExTaq、*T*<sub>4</sub> 连接酶、RNAiso Reagent 试剂盒、逆转录试剂盒均购自 Takara 公司, DNA 凝胶回收试剂盒购自维特洁公司, PCR 引物由上海生工生物工程公司合成, 其他试剂均为进口或国产分析纯。

#### 3.2 大豆 KCS 基因序列的克隆

利用 RNA 提取试剂盒提取大豆叶中的总 RNA, 以提取的总 RNA 为模板, 按照逆转录试剂盒的说明合成大豆 cDNA 的第一条链。根据已获得的植物中的 *KCS* 基因的氨基酸序列(拟南芥 AL023094.2, 苜蓿 XP\_003606927.1, 花生 AC251240.1, 棉花 ABA01490.1, 银扇草 EU871787.1), 在大

豆基因组序列(<http://www.phytozome.net/soybean/>)中, 同源搜索大豆的 *KCS* 基因序列。根据该基因的序列, 设计引物 (D<sub>1</sub>: 5'-ATGACAGTGACGATGAGTGGGA-3'; D<sub>2</sub>: 5'-CTATGAGAC-TATTTCCACAGGATAC-3'), RT-PCR 扩增大豆 *KCS* 基因序列。反应条件为预变性 95 $^{\circ}$ C 5 min; 94 $^{\circ}$ C 30 s, 51 $^{\circ}$ C 40 s, 72 $^{\circ}$ C 1 min, 共 30 个循环; 72 $^{\circ}$ C 后延伸 7 min。将 PCR 扩增片段连接到 pMD18-T 克隆载体上, 获得重组质粒并进行测序。

#### 3.3 大豆 KCS 基因的生物信息学分析

利用 ExPASy Proteomics Server (<http://expasy.org/tools/protparam.html/>)在线软件预测蛋白基本理化性质; 利用 TMHMM Server v.2.0、SignalP4.0、Prodictpro-tein 分别预测可能的跨膜结构域、信号肽和蛋白质的二级结构; 应用多序列比对工具 ClustalX, 以氨基酸全序列联配的结果为基础, 采用 MEGA5.05 软件包中的邻接法构建系统发育树。

#### 作者贡献

赵艳是本研究的实验设计和实验研究的执行人; 邱增成和翟莹参与实验设计和实验结果的分析; 徐红红和国春晖参与基因的克隆工作。全体作者都阅读并同意最终的本文。

#### 致谢

本研究由黑龙江省教育厅科学技术研究项目 (12521611)、作物生物学国家重点实验室开放基金资助

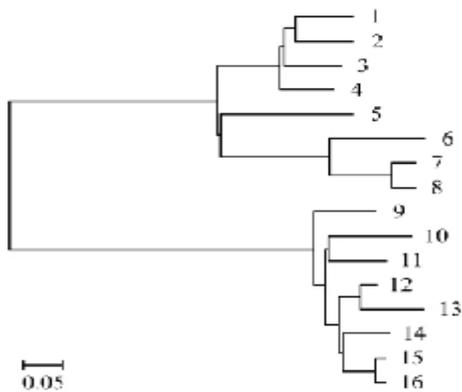


图 4 大豆 KCS 与其它 KCS 的进化树分析

注: 1: 油茶(ACQ41892.1); 2: 大豆(JX855159); 3: 菟葵(ABS-18382.1); 4: 苜蓿(XP-003606927.1); 5: 玉米(ACG36525.1); 6: 小麦(ABY82171.2); 7: 花生(ACZ51240.1); 8: 棉花(ABA01490.1); 9: 银扇草(EU871787.1); 10: 水芹(Ef186003.1); 11: 海甘蓝(AY793549.1); 12: 拟南芥(AL023094.2); 13: 芹菜(AF491877.1); 14: 独行菜(FJ907545.1); 15: 小果亚麻芥(GU929435.1); 16: 亚麻芥(GU929421.1)

Figure 4 Phylogenetic relationships between KCS from soybean and other KCS proteins

Note: 1: *Camellia oleifera* (ACQ41892.1); 2: *Glycine max* (JX855-159); 3: *Eranthis hyemalis* (ABS18382.1); 4: *Medicago truncatula* (XP-003606927.1); 5: *Zea mays* (ACG36525.1); 6: *Triticum aestivum* (ABY82171.2); 7: *Arachis hypogaea* (ACZ51240.1); 8: *Gossypium hirsutum* (ABA01490.1); 9: *Lunaria annua* (EU8717-87.1); 10: *Teesdalia nudicaulis* (Ef186003.1); 11: *Gramme abyssinica* (AY793549.1); 12: *Arabidopsis thaliana* (AL023094.2); 13: *Brassica juncea* (AF491877.1); 14: *Lepidium campestre* (FJ9075-45.1); 15: *Camelina microcarpa* (GU929435.1); 16: *Camelina sativa* (GU929421.1)

(2012KF13)和齐齐哈尔大学青年教师科研启动支持计划项目(2010K-M09)共同资助。

## 参考文献

Blacklock B.J., and Jaworski J.G., 2002, Studies into factors contributing to substrate specificity of membrane-bound 3-ketoacyl-CoA synthases, *Eur. J. Biochem.*, 269(19): 4789-4798

Guo Y., Mietkiewska E., Francis T., Katavic V., Brost J.M., Giblin M., Barton D.L., and Taylor D.C., 2009, Increase in nervonic acid content in transformed yeast and transgenic plants by introduction of a *Lunaria annua* L.

3-ketoacyl- CoA synthase (KCS) gene, *Plant Mol. Biol.*, 69(5): 565-575

Joubès J., Raffaele S., Bourdenx B., Garcia C., Laroche-Traineau J., Moreau P., Domergue F., and Lessire R., 2008, The VLCFA elongase gene family in *Arabidopsis thaliana*: phylogenetic analysis, 3D modelling and expression profiling, *Plant Mol. Biol.*, 67(5): 547-566

Lee S.B., Jung S.J., GoY.S., Kim H.U., Kim J.K., Cho H.J., Park O.K., and Suh M.C., 2009, Two *Arabidopsis* 3-ketoacyl CoA synthase genes, *KCS20* and *KCS2/DAISY*, are functionally redundant in cuticular wax and root suberin biosynthesis, but differentially controlled by osmotic stress, *Plant J.*, 60(3): 462-475

Mietkiewska E., Giblin E.M., Wang S., Barton D.L., Dirpaul J., Brost J.M., Katavic V., and Taylor D.C., 2004, Seed-specific heterologous expression of a nasturtium *FAE* gene in *Arabidopsis* results in a dramatic increase in the proportion of erucic acid, *Plant Physiol.*, 136(1): 2665-2675

Millar A.A., and Kunst L., 1999, The natural genetic variation of the fatty-acyl composition of seed oils in different ecotypes of *Arabidopsis thaliana*, *Phytochemistry*, 52(6): 1029-1033

Ni Y., and Guo Y.J., 2008, Progress in the study on genes encoding enzymes involved in biosynthesis of very long chain fatty acids and cuticular wax in plants, *Yichuan (Hereditas)*, 30(5): 561-567 (倪郁, 郭彦军, 2008, 植物超长链脂肪酸及角质层蜡质生物合成相关酶基因研究现状, *遗传*, 30(5): 561-567)

Rossak M., Smith M., and Kunst L., 2001, Expression of the *FAE1* gene and *FAE1* promoter activity in developing seeds of *Arabidopsis thaliana*, *Plant Mol. Biol.*, 46(6): 717-725

Yang H., Wei J.B., Ruan Y., and Liu C.L., 2012, Clone, sequence analysis and expression vector construction of *CarKCS* gene full-length CDS in *Cardamine hirsute*, *Zuowu Yanjiu (Crop Research)*, 26(3): 213-218 (杨慧, 魏解冰, 阮颖, 刘春林, 2012, 碎米荠 *CarKCS* 基因全长 CDS 克隆、序列分析及表达载体构建, *作物研究*, 26(3): 213-218)