

## 黄花蒿 *HDS* 基因的克隆与功能分析

张祖荣<sup>1,2</sup>, 廖志华<sup>2</sup>, 彭梅芳<sup>2</sup>

1. 重庆文理学院 生命科学系, 重庆 402168

2. 西南大学生命科学学院 三峡库区生态环境教育部重点实验室, 重庆 400715

**摘要:** 目的 克隆获得黄花蒿 MEP 途径中必需关键酶——羟甲基丁烯基-4-磷酸合成酶基因 (*HDS*), 并进行生物信息学分析和功能互补分析研究。方法 对已知的其他种子植物 *HDS* 基因的核苷酸序列进行多重序列比对, 选取保守区域设计简并引物, 利用同源扩增和 cDNA 末端快速扩增技术从黄花蒿中获得目的基因; 利用 BLAST 进行序列比对, ORF Finder 寻找开放阅读框, 并用 MEGA3.0 中的邻位相联法构建进化树。结果 得到 1 条长 2 324 bp 的 *HDS* cDNA 序列, 其 ORF 框长 1 854 bp, 编码 617 个氨基酸残基的蛋白; 生物信息学分析显示, 黄花蒿 *HDS* 基因 *AaHDS* 与其他种子植物来源的 *HDS* 高度同源; 功能互补分析表明, *AaHDS* 能互补突变菌株 *Escherichia coli* MG1655 ara<>*HDS* 中缺失的 *HDS* 功能, 使突变菌株恢复生长, 证明 *AaHDS* 具有典型的 *HDS* 基因功能。结论 首次克隆获得黄花蒿 *HDS* 基因, 为青蒿素的代谢工程研究提供相应的基础。

**关键词:** 黄花蒿; 羟甲基丁烯基-4-磷酸合成酶基因; 基因克隆; 生物信息学分析; 青蒿素

**中图分类号:** R282.21 **文献标志码:** A **文章编号:** 0253-2670(2012)01-0148-07

## Cloning and function analyses of *HDS* gene from *Artemisia annua*

ZHANG Zu-rong<sup>1,2</sup>, LIAO Zhi-hua<sup>2</sup>, PENG Mei-fang<sup>2</sup>

1. Department of Life Science, Chongqing University of Arts and Sciences, Chongqing 402168, China

2. Key Laboratory of Eco-environments in Three Gorges Reservoir Region, Ministry of Education, School of Life Sciences, Southwest University, Chongqing 400715, China

**Abstract: Objective** To obtain the indispensable key enzyme—hydroxide methyl enylamino 4-cyclodiphosphate synthase (*HDS*) gene involved in the MEP pathway cloned from *Artemisia annua* and conduct bioinformatic and functional complementation analysis.

**Methods** To perform multiple sequence alignment for the nucleotide acid sequence of the other reported seed plants' *HDS* gene, to select conservative areas for designing degenerate primers, and to gain the aim gene from *A. annua* through homologous expanding and cDNA bottom speedily expanding technique. To perform sequence alignment using BLAST, to identify open reading frame (ORF) using ORF Finder, and to construct phylogenetic tree using neighbor joining (NJ) ways in MEGA3.0. **Results** The obtained *HDS* cDNA sequence was 2 324 bp containing a 1 854 bp ORF and encoding a 617-amino acid protein. Bioinformatic analysis showed that *AaHDS* was homologous with *HDS* derived from other seed plant species. Functional complementation analysis indicated that *AaHDS* could make up the short *HDS* function of mutant *Escherichia coli* MG1655 ara<>*HDS*. It could make the mutant get back to upgrowth, which showed *AaHDS* had typical *HDS* gene function. **Conclusion** The cloning *HDS* gene from *A. annua* for the first time provides a good basis for further study on the metabolism project of artemisinin.

**Key words:** *Artemisia annua* L.; *HDS*; gene cloning; bioinformatic analysis; artemisinin

青蒿素是从黄花蒿 *Artemisia annua* L. 中提取的一种治疗疟疾的特效药物<sup>[1-3]</sup>, 近年来的研究还发现青蒿素具有免疫抑制、抗血吸虫、抗病毒及抗肿瘤等多方面的药理作用<sup>[4]</sup>。但天然药源黄花蒿中青蒿素含量很低 (0.01%~0.6%)<sup>[5]</sup>, 大大限制了商业化的

生产, 青蒿素的生产远远不能满足国际市场的需求<sup>[6]</sup>。因此寻找和扩大青蒿素药源是目前亟待解决的问题。化学合成青蒿素, 成本高、难度大; 组织培养方法也不能有效提高青蒿素的产量, 目前青蒿素来源仍是依赖从栽培黄花蒿中提取<sup>[7-9]</sup>。

收稿日期: 2011-03-31

基金项目: 重庆市教委自然科学基金资助项目 (kj071204)

作者简介: 张祖荣 (1966—), 男, 重庆江津人, 重庆文理学院副教授, 长期从事药用植物的教学与科研工作, 现为西南大学访问学者。  
Tel: 13658304114 E-mail: 412587791@qq.com

近年来,随着青蒿素生物合成途径相关酶基因的克隆,基因工程成为提高青蒿素产量的有效途径之一<sup>[10]</sup>。为研究青蒿素生物合成的分子机制,以及为青蒿素代谢工程提供候选基因和作用靶点,本研究采用 RACE 技术<sup>[11]</sup>从黄花蒿中克隆获得了青蒿素代谢途径中的一个关键酶基因——羟甲基丁烯基-4-磷酸合成酶基因(*HDS*),并进行了相应的生物信息学分析和功能验证。

## 1 材料与方法

### 1.1 材料

黄花蒿全株采集于西南大学生命科学学院的重庆酉阳科研教学基地,用液氮速冻后在 $-70\text{ }^{\circ}\text{C}$ 保存备用。大肠杆菌 *Escherichia coli* T. Escherich 菌株 DH5 和 M15、质粒 pQE30、突变菌株 *E. coli* MG1655 *ara<>HDS* 以及其他实验材料均由该实验室廖志华教授鉴定,由西南大学生命科学学院天然产物与代谢工程实验室保存和提供。

### 1.2 方法

**1.2.1 黄花蒿总 RNA 的提取** 取 $-70\text{ }^{\circ}\text{C}$ 冰箱中保存的黄花蒿叶、茎混合材料 0.1 g,用 RNAlant 试剂盒(天根生物,北京)提取总 RNA,保存于 $-70\text{ }^{\circ}\text{C}$ 冰箱备用。

**1.2.2 黄花蒿 *HDS* 基因(*AaHDS*)核心片段的获得** 在美国国家生物技术信息中心(NCBI)网站上分别下载已报道的 *HDS* 基因的核苷酸序列[如甜菊 *Stevia rebaudiana* (Bert.) Hemsl.、拟南芥 *Arabidopsis thaliana* (L.) Heynh.、细长聚球藻 *Synechococcus elongates* Nag. 等],使用 Vector NTI 8.0 进行多重序列比对,选取最保守核苷酸序列设计简并引物<sup>[12]</sup>,正向引物为 *dfAaHDS*: 5'-TATGG(A/G)(A/C)G(A/G/T/C)GCAATGCG(C/A/T)ATTG-3',反向引物为 *drAaHDS*: 5'-CAATCTTCC(A/G)GG(A/T)G(C/A)-ACCACC-3'。cDNA 第一链合成使用 TaKaRa RNA PCR Kit(AMV),按照试剂盒操作手册合成 cDNA 链。以 cDNA 为模板进行 PCR 扩增,反应条件为:  $94\text{ }^{\circ}\text{C}$  预变性 4 min,  $94\text{ }^{\circ}\text{C}$  变性 45 s,  $50\sim 60\text{ }^{\circ}\text{C}$  温度梯度退火 45 s (每个循环降低  $0.5\text{ }^{\circ}\text{C}$ ),  $72\text{ }^{\circ}\text{C}$  延伸 1 min,共 30 个循环;最后  $72\text{ }^{\circ}\text{C}$  总延伸 10 min。PCR 产物纯化按照上海赛百盛基因技术有限公司的 PCR 产物纯化试剂盒说明操作,将回收产物与 pMD18-T 连接后转化 DH5 $\alpha$ ,进行蓝白斑筛选,挑取白斑进行菌落 PCR 检测,阳性克隆送至上海英骏生物技术有限公司测序,测序结果进行 BLAST 分

析以获得 *AaHDS* 核心片段。

**1.2.3 *AaHDS* 全长 cDNA 的获得** 根据 *AaHDS* 核心片段序列设计 RACE 引物。*AaHDS*3-1: 5'-TGCGGGTGCCAATACTGGTG-3'; *AaHDS*3-2: 5'-CTGCTACAAGGATGCAGGATG-3'; *AaHDS*5-1: 5'-GAGCCATCACGGTGGAGTACACC-3'; *AaHDS*5-2: 5'-CCTCTTCACCCTCCTTCTGCAC-3'。RACE-PCR 扩增按照 SMART RACE cDNA (Clontech, 美国) 扩增试剂盒的说明书进行。RACE 产物的纯化、克隆及测序与“1.2.2”相关操作相同。将 3'RACE、5'RACE 和核心片段进行电子拼接(Vector NTI Suite 8.0),获得全长 cDNA 序列,根据拼接序列设计全长特异性引物,正向引物 *fiAaHDS*: 5'-TTCTGAAGGGAGTCCTCTGTT-3';反向引物 *riAaHDS*: 5'-CCATCACCTTTAACTAGACGGT-3',克隆获得物理全长 cDNA 序列。

**1.2.4 *AaHDS* 序列分析** 序列比对均使用 Vector NTI Suite 8.0 软件和 BlastP2.2.3 (<http://www.ncbi.nlm.nih.gov>) 进行分析;开放性阅读框(open reading frame, ORF)的查找和核苷酸的翻译在 <http://www.ncbi.nlm.nih.gov> 网站的 ORF Finder 上完成;用 CLUSTALX<sup>[13]</sup>进行多重序列比对,用 MEGA3.0 中的临位相联法(neighbor-joining, NJ)<sup>[14]</sup>构建进化树,重复次数为 1 000 次;蛋白质基本性质及转运肽的分析使用 <http://www.expasy.org> 网站提供的相关生物信息学分析软件进行<sup>[15]</sup>。

**1.2.5 *AaHDS* 基因的功能验证** 大肠杆菌突变菌株 MG1655 *ara<>HDS* 中,内源 *HDS* 基因被一个卡那霉素抗性基因所替代,染色体中仅含有一个单拷贝的受到 P<sub>BAD</sub> 启动子控制的 *HDS* 基因<sup>[16]</sup>。MEP 途径存在于大肠杆菌中,*HDS* 基因是大肠杆菌生存所必须的<sup>[17]</sup>,MG1655 突变菌株只能在含有阿拉伯糖(Ara)的培养基上生长,而不能在含有葡萄糖(Glc)的培养基上生长<sup>[16]</sup>。为了验证 *AaHDS* 的催化活性,在大肠杆菌突变体 MG1655 中进行互补试验<sup>[18]</sup>: *AaHDS* 的编码区被构建到原核表达载体 pQE30 上,并转化至大肠杆菌突变体 MG1655 中,获得 *AaHDS* 功能表达的工程菌,同时以空质粒 pQE30 转化大肠杆菌突变菌株 MG1655 *ara<>HDS* 感受态作为阴性对照,采用划线培养的方式将 MG1655 *ara<>HDS* 分别接种于 LB+Kan+0.2% Ara 与 LB+Kan+0.2%Glc 平板上,将携带 pQE30-*AaHDS* 质粒和空质粒的 MG1655 *ara<>HDS* 分别接种于 LB+Kan+

Amp+0.5 mmol/L IPTG+0.2% Glc 平板上, 过夜培养后, 观察生长情况以验证 *AaHDS* 的功能。

### 2 结果与分析

#### 2.1 总 RNA 的提取

以黄花蒿茎、叶混合取样为材料, 按照 RNAplant buffer 试剂盒操作说明提取黄花蒿总 RNA, 将抽提的 RNA 取 5 μL 进行电泳检测, RNA 明显呈 3 个条带, 无降解现象, 无蛋白杂质, 28S 条带明亮 (图 1), 说明获得了高质量的 RNA; 再用紫外法进行定量测定,  $A_{260}/A_{280}$  值均在 1.8~2.0, 表明所得 RNA 的纯度达到了相应实验的要求。

#### 2.2 基因 *AaHDS* 的克隆

黄花蒿总 RNA 反转录为 cDNA 后, 以 *dfAaHDS* 和 *drAaHDS* 为引物, cDNA 为模板进行 PCR 扩增。将获得的扩增条带进行目的片段回收, 亚克隆后的测序结果表明该片段长 1 420 bp。经 BLAST 分析初步鉴定为黄花蒿 *HDS* 基因核心片

段, 并命名为 *AaHDS*。根据所得片段设计基因特异性引物用于扩增 *AaHDS* 的 cDNA 末端, 通过 3'-RACE 和 5'-RACE 分别获得 602 bp 的 3'-末端和 1 045 bp 5'-末端。将 *AaHDS* 基因 3 个片段拼接后获得其全长 cDNA, 并最终扩增获得其物理全长, 测序结果表明 *AaHDS* 基因全长为 2 324 bp (图 2)。

#### 2.3 基因 *AaHDS* 的生物信息学分析

*AaHDS* 的 cDNA 全长 2 324 bp, 包含一段长

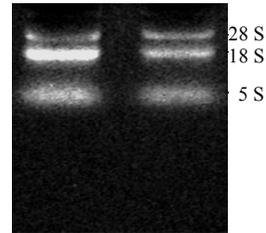


图 1 黄花蒿总 RNA 电泳图

Fig. 1 Electrophoresis of total RNA from *A. annua*

```

01                                     acgc gggatcgggaaact tggccttgcct tc
31 tgaaggagatcctctctgttagttcccgctacagaagtatttgaatccacatataaaaaaccatcaggggaaaaccgcaacagtaattggttgg
21 agatgtaccctcttggtagcgcacccataagaatcacaacaatgacacaaatgactaactaaggaactgctgcaaacgtgagaacaggtt
11 ATGCAAACTCCTGACAGGAGCTGACAT TGTT CGAATAACTGTCAGGAAAAGAGAGGCAATGCATGTTT TGA AATCAAGAACC
M Q I A D R G A D I V R I T V Q G K R E A D A C F E I K N T
01 CTTGTT CAGAAGAATTATAACA TAACC TCTGGTGCCAGACATTCATT TTCC TCCATCAGTTGCCATGCGAGTTGC TGAATGCTTTGACAAA
L V Q K N Y N I P L V A D I H F A P S V A L R V A E C F D K
91 ATTCGTGTCAACCCCGGAGATT TTGC TGA TCOC CGAGCTCAGT TTGAGCAGTT GGAATACAGAAGATGATTATCAACACGGAACCTTGAC
I R V N P G D F A D R A Q F E Q L E Y T E D D Y Q Q E L E
81 CATATTGAAAAGGTTT TTA CTC CATTGGT TGA A A A A T G C A A G A A A T A T G G A A G G C C A A T G C A A T A G G A A C A A A C C A T G G G A G T C T T T C A
H I E K V F T P L V E K C K K Y G R A M R I G T N H G S L S
71 GATCGTATCATGAGTTACTATGGTGA TTC TCCT AGG G G A A T G C T T G A G T C T G C A T T T G A A T T C G A A G A A T C T G T A G G A A G T T A G A C T T C
D R I M S Y Y G D S P R G M V E S A F E F A R I C R K L D F
61 CA TAAC TTTGTA TTTTCAA TGA AAGCAAGCAAT CCA GTGATCA TGG TTCAAGCATA TCGT CTT CTAG TGG CCGA AAT GAA TGT CAGGGA
H N F V F S M K A S N P V I M V Q A Y R L L V A E M N V Q G
51 TGGAC TAT C C T T T G C A C T T G G G G T T A C T G A A G C T G C G A G G T G A G G A T G G A C G T A T G A A A T C T G C A A T T G C A T T G G C A C A C T T C T T
W D Y P L H L G V T E A G E G E D G R M K S A I G I G T L L
41 CAGGATGGTCTTGGTGATAC TA TCAGGGT TTCC CTCACA GAAC CAC CAGAGGAAGAATA GAC CCGT GTAGGAAATTGGC CAA C C T T G G T
Q D G L G D T I R V L S T E E E I D P C R K L A N L
31 ATGAAAGCA TCCCAAC TTC AAC AAGGAGT GGCA CCG TTT GAAGAAAAGCA CAGGCG TTAT TTT GATT TTCAACGGAGAACT TGGTAT TTG
M K A S Q L Q Q G V A P F E E K H R R Y F D F Q R R T T G D L
21 CC TGT GCAGAAGGAGCGGTGAAGAGGTAGA TTA CAGA GGTGAT CCA CCG TGA TGG CTC TGTTATCA TGT CTG TTAC C C T G A T C A A C T C
P V Q K E G E E V D Y R G V L H R D G S V I M S V T L D Q L
11 AAGAC C C C T G A A C T C T T C A C A G A T C A T T A G C A G C A A A A C T T G T T C T T G C T A T G C C A T T A A G G A T C T T G C A A C A G T T G S T T A T C T T A
K T P E L F Y R S L A A K L V L G M P F K D L A T V D S I L
01 TTAAGAGAGCTTCCA C C A G C A G A T A A A G A T G C T C G T T T A G C T C T T A A G A G G T T G A T C G A T G T A A G T A T T A C C C C T T T G
L R E L P P A D D K D A R L A L K R L I D V S M G I I T P L
91 TCAGACAG TTGACA AAGC C T T T G C C A A G C C T A T T G T C T T G G T A A A C T T G A A T G A A C T C C A C A G G A G C T C A T A A C C T T T T G C C A G A A
S E Q L T K P L P N A I V L V N L N E L S T G A H K L L P E
81 GGCACAGT TTGCTGGTGT CAG T A C G T G G T G A T G A G C C T T A T G A G G A C T T G A T A T T C T T A A A A C C A C T G A T G C T A C A A T G A T T C T T C A T
G T R L V V S V R G D E P Y E E L D I L K T T D A T M I L H
71 GAAC T G C T T A T G C T G A A G A A A A A C T G C A G A G T A C A T G C T G C T A G G A G T T G T T G A G T A C C T T T C T G A A A A T T C A C T G A A T T T C C C T
E L P Y A E E K T G R V H A A R R L F E Y L S E N S L N F P
61 GTTCT TCAC C A C A T A A A G T T C C A A A G G A A T T C C C A G G A T G A T T T G G T T A T C A G T G C G G G T G C C A A T G C T G T G C T T T A T T A G T T G A C
V L H H I K F P K G I P R D D L V I S A G A N A G A L L V D
51 GGACT TGGAGATGGTATCTTAT TAGAAGC TCCGGATCAAGACT TCGAATT TATCAGGAA TACA TCTT TCAATC TGCTACAAGGATG CAGG
G L G D G I L L E A P D Q D F E F I R N T S F N L L Q G C R
41 ATGGC CAACACTAAGACAGAAATATGCTCTCATGCCCACTCTGTGGTGAAC TCTATT TGA TCTTCAAG TGA TTAGTGGC GAAAT TAGA GAA
M R N T K T E Y V S C P S C G R T L F D L Q V I S A E I R E
31 AAAACATCA CAT TTG C C T G G T C T C G A T T G C G A T T A T G G G T T G C A T A G T C A A T G G A C C A G G G A G A T G G C T G A O G C A G A T T T G G A T A T
K T S H L P G V S I A I M G C I V N G P G E M A D A D F G Y
21 CTGGTGGTGGCTCTCGAAGA TTGATCTTTA TGTTGAAAGCA C G G T G G T A C A C A G G A A A T T G C A A T G G A A G T G C A A C G A T G C A T T G
V G G A P G K I D L Y V G K T V V Q R G I A M E G A T D A L
11 ATTCAACTAATCAAAAGACATGGCCGCTGGGTGATCTCTGTTGAAGAGTACagggatgcatactccatgtatatacaaaataaaaagt
I Q L I K D H G R W V D P P V E E *
01 ggaggagctgcaaat ttgtat ttaaattttatattat tttttac ttgaaac ctagg gcaac aagaacat taaaggtaac atgtttat tat
91 ttgagatgataaggaaaaacgtctatgtaaaaggtatggtttttaaattgctt gga ttcaattttgc cat tggc tat ttaaaaaa
2281 gatat gaaag gagg caagaat tcaaat caaaa aaaa aaaaa

```

编码区和翻译的氨基酸用大写字母表示, 非编码区 (UTR) 用小写字母表示, 终止密码子用\*表示

Coding and its deduced amino acid sequences are shown by capital letters, UTR is shown by lowercase letters; \* indicates stop codon

图 2 *AaHDS* cDNA 全长序列和由此推测的氨基酸序列

Fig. 2 *AaHDS* cDNA full-length sequence and its deduced amino acid sequence

度为 1 854 bp 的 ORF, 还有长度为 210 bp 的 5'-端非翻译区 (untranslated region, UTR) 和 260 bp 的 3'-UTR, 终止密码子为 TAG (图 2)。AaHDS 编码含 617 个氨基酸残基的蛋白, 其相对分子质量预测为  $6.87 \times 10^4$ , 等电点预测为 5.3。在 ExPasy 中进一步预测 AaHDS 的二级结构<sup>[19]</sup>, 显示其中的随机卷曲占 45.38%,  $\alpha$ -螺旋占 40.19%, 延伸链占 14.42% (图 3)。将 AaHDS 与来源于其他植物以及藻类、细菌的 HDS 进行多重序列比对。

结果图 4 表明, 来源于植物与来源于藻类和细菌

的 HDS 差异较大, 尤其是在 N 端, 植物中 HDS 的 N 端比细菌和藻类多出一段大约含 80 个氨基酸残基的区域, 这段序列的功能已经在拟南芥 *A. thaliana* 中得到了证实<sup>[20]</sup>, 是引导 HDS 定位于质体的植物特有的转运肽, 这与在植物中 MEP 途径定位于质体中的观点一致<sup>[21]</sup>; 同是来源于植物的 HDS 在 N 端的氨基酸序列差异也较大, 主要因为这部分序列是非催化区域, 而催化区域的 HDS 氨基酸序列非常保守, 其保守一致性高达 80%~95%。

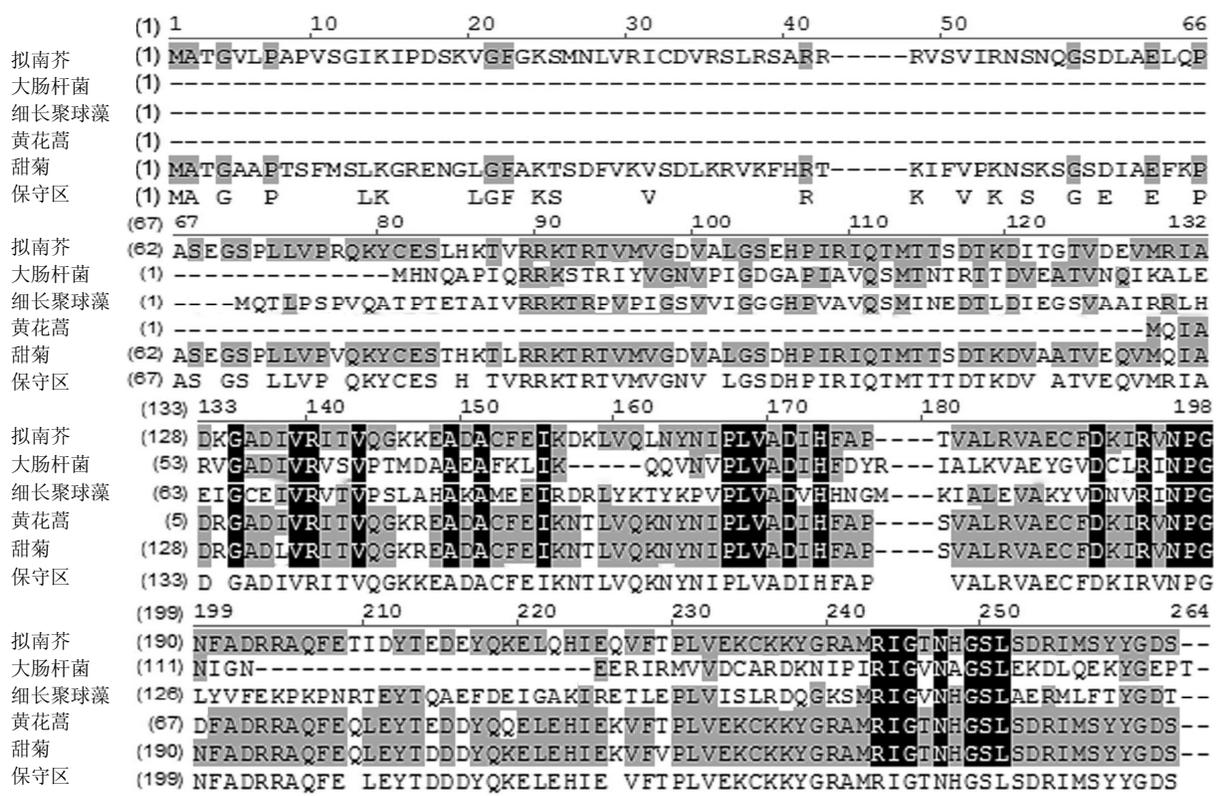
用 Clustalx 和 MEGA3.0 软件对 AaHDS 进行分

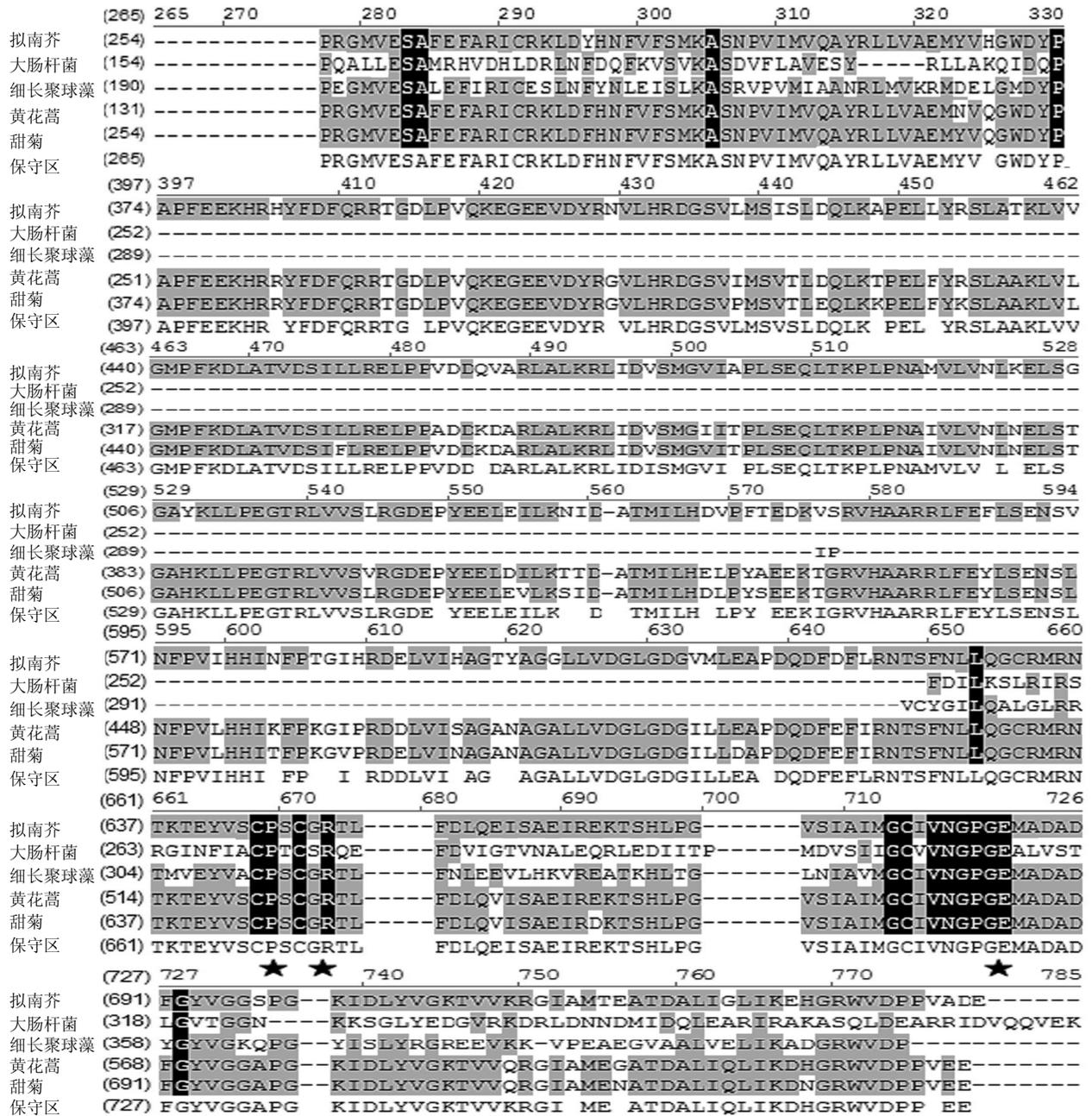
```

MQIADRGADIVRITVQGKREADACFEIKNTLVQKYNINPLVADIHAFAPSVALRVAECFDKIRVNPFGDFADRRRAQFEQLEYTEDDYQQELE
ccccccccchhhhhhhccccchhhhhhhccccccccccccccccccccccccchhhhhhhccccccccccccchhhhhhhhhhhccccchhhhhhh
HIEKVFTPLVEKCKKYGRAMRIGTNHGSLSDRIMSYYGDSPRGMVESAFEFARICRKLDFHNFVFSMKASNPVIMVQAYRLLVAEMNVQGG
hhhhhhchhhhhccccccccccccccccccccccccccccchhhhhhhhhhhhhccccccccccccccccchhhhhhhhhhhhhcccc
WDYPLHLGVTAGEGEDGRMKSAGIGITLLQDGLGDTIRVSLTEPPEEIEIDPCRKLANLGMKASQLQGGVAPFEEKHRRYDFDQRRTGDL
ccccccccccccccccchhhhhhhhhccccccccccccccccccccccccchhhhhhhccccchhhhhhhccccchhhhhhhhhhhhhcccc
PVQKEGEEVDYRGLVLRDGSVIMSVTLDLQKTPELFYRSLAAKLVGMFPKDLATVDSILLRELPPADDDARLALKRLIDVSMGIITPL
ccccccccccccccccccccccccccccccccccccccccchhhhhhhhhhhhhhhccccccccchhhhhhhhhccccchhhhhhhhhhhhhcccc
SEQLTRKPLPNAIVLVNLSLSTGAHKLLPEGTRLVVSVRGDEPYEELDILKTTDATMILHELPYAEKTRVHAARRLFYELSENSLNF
ccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccchhhhhhhhhhhhhhhcccccccccccccccc
VLHIIKFKPGIPRDDLVISAGANAGALLVDGLGDGILLEAPDQDFEFIRNTSFNLLQGCRRMNTKTEYVSCPSCGRITLFDLQVISAETRE
eeeeccccccccccccchhhhhhhhhhhhhhhcccccccccccccccccccccccccccccccccccccccccccccccccccccccccccc
KTSHLPGVSIAMGCIVNGPGEMADADFGYVGGAPGKIDLYVGKTVVQRGIAMEGATDALIQLIKDHGRWVDFPVEE+
ccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccchhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhh
    
```

h- $\alpha$ -螺旋, e-延伸链, c-随意卷曲  
h-alpha helix e-extended strand c-random coil

图 3 AaHDS 的二级结构预测  
Fig. 3 Secondary structure prediction of AaHDS





一致氨基酸残基采用白色字体黑色背景表示，保守氨基酸残基采用黑色字体灰色背景表示，其他氨基酸残基均用黑色字体白色背景表示；★表示半胱氨酸残基保守位点

Identical amino acid residues were shown in white with black background and conserved amino acid residues were shown in black with gray background, other amino acid residues were shown in black with white background; ★ indicates position of conserved cysteine residues

图 4 来自不同物种的 HDS 氨基酸序列多重比对

Fig. 4 Multi-alignment of HDS amino acid sequences from various plant species

子系统发育树分析，结果显示，来源于植物、藻类、细菌的 HDS 在发育进化树上明显各聚为一支（图 5），并且在进化树上的距离也与其所属类别的亲缘关系相符。其中 AaHDS 与甜菊 *S. rebaudiana* 的 HDS 亲缘关系最接近，这与黄花蒿和甜菊同属菊科植物的事实相符。

#### 2.4 基因 AaHDS 的遗传功能互补分析

由于 MEP 途径对细菌的生长是必需的，所以

HDS 基因对大肠杆菌的生长是必需的。而 *E. coli* MG1655 ara<>HDS 突变菌株经过人工改造后，其的内源性 HDS 基因由卡拉霉素抗性表达盒取代，因此其只能在含有 Ara 的培养基上生长，而不能在含有 Glc 的培养基上生长。将携带 AaHDS 基因的重组表达质粒 pQE30-AaHDS 转化进突变菌株 *E. coli* MG1655 ara<>HDS 后，该菌株获得

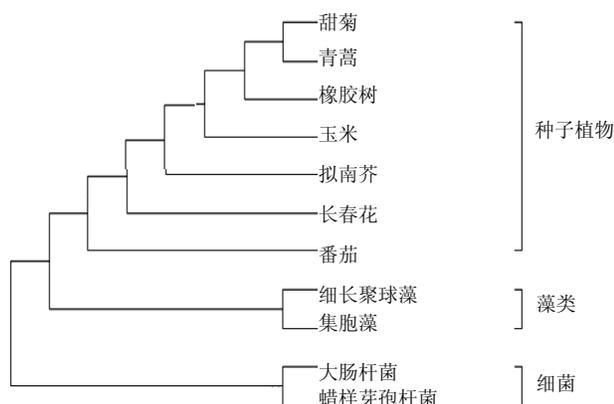


图 5 在 MEGA3.0 平台上用 NJ 方法构建的 *HDS* 分子进化树  
Fig. 5 *HDS* phylogenetic tree constructed by NJ method on MEGA 3.0

外源性的 *HDS* 基因，在含有葡萄糖的培养基上可以生长，而对照菌株则不能生长，从而证明 *AaHDS* 互补了大肠杆菌缺失的 *HDS* 的功能。

大肠杆菌突变菌株 MG1655 *ara<>HDS* 能在含有 0.2% Ara 的 LB 培养基上生长，却不能在含有 0.2% Glc 的培养基上生长（图 6-a, b）；当将携带 *AaHDS* 编码区的质粒 pQE30-*AaHDS* 转入 *E. coli* MG1655 *ara<>HDS* 之后，能使大肠杆菌 *HDS* 突变菌株恢复生长，在含有 0.2% Glc 和 0.5 mmol/L IPTG 的培养基上能生长（图 6-d），而作为对照的空载体 pQE30 则不能在此培养基上生长（图 6-c）。

### 3 讨论

青蒿素作为治疗疟疾的特效药，在国际市场上供不应求。自从青蒿素被发现以来，研究者们不断尝试各种方法以期提高青蒿素量：采用化学合成青蒿素，需要经过一系列繁杂的反应，且成本高、得率低、毒性大；组织培养也不能有效提高青蒿素的产量。近年来，随着基因转化技术的广泛运用，不少研究者致

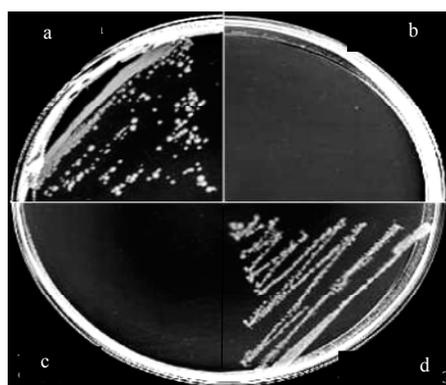


图 6 *AaHDS* 基因的功能验证  
Fig. 6 Functional demonstration of *AaHDS* gene

力于研究利用转基因技术和次生代谢工程手段来提高黄花蒿中青蒿素的量，但其前提是必需对青蒿素生物合成代谢途径的分子生物学和生物化学有深入的了解<sup>[22]</sup>。

到目前为止，青蒿素生物合成途径的研究已取得非常可喜的进展，其代谢途径上多个编码关键酶的基因已被克隆<sup>[7]</sup>。*HDS* 催化 ME-cPP 转化生成 HMBPP，是青蒿素生物合成上游途径 MEP 途径中必需的一个关键酶<sup>[23]</sup>。本研究在前人研究的基础上，采用 RACE 技术，以重庆酉阳的优质高产黄花蒿为研究材料，克隆获得了编码黄花蒿 *HDS* 的 *AaHDS* 基因。对其编码的氨基酸序列进行生物信息学分析，结果表明，*AaHDS* 与来源于其他植物的 *HDS* 同源性高达 80%~95%，且具有典型的 *HDS* 催化活性位点，在发育进化树上，*AaHDS* 与来源于植物甜菊的 *HDS* 亲缘关系最接近。所有这些证据都表明，本研究对 *AaHDS* 基因的克隆是非常成功的。该成果为更加深入地研究青蒿素生物合成代谢途径奠定了重要的分子生物学基础，也为研究青蒿素的代谢工程提供了更多可能的候选基因和调控靶点。

### 参考文献

- [1] Klayman D L. Qinghaosu (artemisinin): an antimalarial drug from China [J]. *Science*, 1985, 228: 1049-1055.
- [2] Lin C X, Xiao P G, Peng Y, et al. Challenges in research and development of traditional Chinese medicines [J]. *Chin Herb Med*, 2009, 1(1): 1-28.
- [3] 王鸿博, 肖 皖, 华会明, 等. 黄花蒿的化学成分研究进展 [J]. *现代药物与临床*, 2011, 26(6): 430-434.
- [4] 陆金健. 青蒿素类化合物抗肿瘤研究进展 [J]. *中国药理学通报*, 2010, 26(6): 818-820.
- [5] 赵 兵, 王玉春, 欧阳藩. 青蒿素生物合成机理研究现状 [J]. *广西植物*, 1999, 19(2): 154-158.
- [6] 刘春朝, 王玉春, 欧阳藩, 等. 青蒿素研究进展 [J]. *化学进展*, 1999, 11(1): 41-48.
- [7] 孔建强, 王 伟, 朱 平, 等. 青蒿素生物合成研究进展 [J]. *中国药学杂志*, 2008, 43(11): 804-808.
- [8] 唐 煜, 朱建华, 于荣敏. 黄花蒿悬浮培养细胞对二氢青蒿酸的生物转化研究 [J]. *中草药*, 2010, 41(8): 1358-1361.
- [9] 周 洁, 张 霁, 郭兰萍, 等. 稀土元素镧对黄花蒿光合作用及青蒿素积累的影响 [J]. *中草药*, 2010, 41(8): 1371-1374.
- [10] Brown G D. Cadinanes from *Artemisia annua* that may be intermediates in the biosynthesis of artemisinin [J].

- Phytochemistry*, 1994, 36(3): 637-641.
- [11] 魏群编译. 基因克隆和 DNA 分析 (第 5 版) [M]. 北京: 高等教育出版社, 2007.
- [12] 刘万宏. 紫杉醇前体合成途径两个关键酶基因克隆和分析 [D]. 重庆: 西南大学, 2008.
- [13] Thompson L D, Gibson T J, Plewniak F, *et al.* The clustal X windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools [J]. *Nucl Acids Res*, 1997, 24: 4876-4882.
- [14] Kumar S, Tamura K, Nei M. MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment [J]. *Brief Bioinform*, 2004, 5: 150-163.
- [15] Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees [J]. *Mol Biol Evol*, 1987, 4: 406-425.
- [16] McAteer S, Coulson A, McLennan N, *et al.* The *lytB* gene of *Escherichia coli* is essential and specifies a product needed for isoprenoid biosynthesis [J]. *J Bacteriol*, 2001, 183: 7403-7407.
- [17] Hsieh M H, Goodman H M. The *arabidopsis* IspH homolog is involved in the plastid nonmevalonate pathway of isoprenoid biosynthesis [J]. *Plant Physiol*, 2005, 138: 641-653.
- [18] 刘秋云, 李蕾, 蒋一帆. 粗糙脉孢霉瓜氨酸需求型突变株的分离 [J]. 中山大学学报: 自然科学版, 2000, 39(1): 118-120.
- [19] Garnier J, Gibrat J, Robson B. GOR secondary structure prediction method version IV [J]. *Methods Enzymol*, 1996, 266: 540-553.
- [20] Querol J, Campos N, Imperial S, *et al.* Functional analysis of the *Arabidopsis thaliana* GCPE protein involved in plastid isoprenoid biosynthesis [J]. *FEBS Lett*, 2002, 514: 343-346.
- [21] Lichtenthaler H K, Rohmer M, Schwender J. Two independent biochemical pathways for isopentenyl diphosphate and isoprenoid biosynthesis in higher plants [J]. *Plant Physiol*, 1997, 101: 643-52.
- [22] 刘春明译. 植物代谢途径的生物工程和分子生物学 [M]. 北京: 科学出版社, 2009.
- [23] Rohmer M, Knani M, Simonin P, *et al.* Isoprenoid biosynthesis in bacteria: a novel pathway for the early steps leading to isopentenyl diphosphate [J]. *Biochem J*, 1993, 295: 517-524.