

• 药材与资源 •

黄花蒿 MCS 基因的克隆及其序列分析与原核表达

王 英¹, 贾伟章^{2,3*}, 谭明俊², 张 军², 王 丹², 刘顺会²

1. 华南农业大学园艺学院, 广东 广州 510640

2. 广东药学院生命科学与生物制药学院, 广东 广州 510006

3. 广东省生物技术候选药物研究重点实验室, 广东 广州 510006

摘 要: 目的 从黄花蒿中克隆 MEP 途径关键酶 2C-甲基-D-赤藓糖醇-2, 4-环焦磷酸合成酶 (MCS) 基因, 进行序列特征分析、原核表达以及组织表达的研究。方法 根据黄花蒿 MCS (AaMCS) 基因 EST 序列, 克隆 MCS cDNA 及基因组序列。将 MCS 编码区与表达载体 pET-21a (+) 重组, 构建 pET-21a (+)-MCS 的原核表达载体并转入大肠杆菌 BL21 (DE3), IPTG 诱导获得 MCS 融合蛋白。半定量 RT-PCR 检测 MCS 的组织表达情况。结果 AaMCS cDNA 全长 994 bp, 包含 681 bp 的开放阅读框, 编码 226 个氨基酸, 基因全长 2 540 bp, 包含 3 个外显子和 2 个内含子。构建 pET-21a (+)-MCS 重组质粒, 获得稳定的 pET-21a (+)-MCS 原核表达体系。AaMCS 在根、茎、叶、花中均有表达, 其中花中表达量较高, 根和茎中表达量低。结论 克隆了 AaMCS 基因, 建立 pET-21a (+)-MCS 稳定的原核表达体系, 为研究 AaMCS 蛋白的活性及其功能奠定了基础。关键词: 黄花蒿; 2C-甲基-D-赤藓糖醇-2, 4-环焦磷酸合成酶; 基因克隆; 序列分析; 原核表达

中图分类号: R282.12

文献标志码: A

文章编号: 0253-2670(2013)16-2288-06

DOI: 10.7501/j.issn.0253-2670.2013.16.019

Cloning and sequence analysis of MCS gene in *Artemisia annua* and its prokaryotic expressionWANG Ying¹, JIA Wei-zhang^{2,3}, TAN Ming-jun², ZHANG Jun², WANG Dan², LIU Shun-hui²

1. College of Horticulture, South China Agricultural University, Guangzhou 510640, China

2. College of Life Sciences and Biopharmaceutics, Guangdong Pharmaceutical University, Guangzhou 510006, China

3. Guangdong Provincial Key Laboratory of Biotechnology Candidate Drug Research, Guangzhou 510006, China

Abstract: Objective To obtain the indispensable key enzyme involved in the MEP pathway, the 2C-methyl-D-erythritol-2, 4-cyclodiphosphate synthase gene (MCS) was cloned from *Artemisia annua*, and bioinformatic analysis, prokaryotic expression, and tissue-specific expression were conducted. **Methods** According to AaMCS EST sequence, the full length of cDNA and genomic sequences were obtained by the design of specific primers, rapid-amplification of cDNA ends (RACE) and genome amplification. The coding region of MCS gene was cloned into the expression vector pET-21a (+) by gene recombination technology, the recombinant plasmid pET-21a (+)-MCS was transformed into *E. coli* BL21 (DE3), and the expression of recombinant protein was induced by IPTG. Semi-quantitative RT-PCR was used to detect the expression of AaMCS transcripts. **Results** The AaMCS cDNA was found to be 994 bp containing an open reading frame (ORF) of 681 bp that translated into a putative peptide of 226 amino acid. The full length of MCS was 2 540 bp consisting of three exons and two introns. A recombinant pET-21a (+)-MCS was constructed by genetically fusing the MCS to pET-21a (+) vector system, and was successfully expressed in *E. coli* BL21. The tissue expression patterns indicated that the expression level of AaMCS transcripts in flowers was higher than that in the roots and stems. **Conclusion** The MCS gene is cloned from *A. annua*, and the stable prokaryotic expression system of pET-21a (+)-MCS is constructed. This work is helpful for investigating the activities and other physiological functions of MCS protein.

Key words: *Artemisia annua* L.; 2C-methyl-D-erythritol-2, 4-cyclodiphosphate synthase gene; gene cloning; sequence analysis; prokaryotic expression

收稿日期: 2013-03-27

基金项目: 广东药学院博士启动基金项目 (43555022)

作者简介: 王 英 (1979—) 女, 博士, 研究方向为药用植物生物技术研究。E-mail: wangying615@scau.edu.cn

*通信作者 贾伟章 Tel: (020)39352201 E-mail: jiawzh@gmail.com

网络出版时间: 2013-07-05 网络出版地址: <http://www.cnki.net/kcms/detail/12.1108.R.20130705.1531.006.html>

黄花蒿 *Artemisia annua* L. 属于菊科 (Asteraceae) 蒿属 *Artemisia* L. 一年生草本植物, 为传统中草药。青蒿素是 1972 年我国科研人员从黄花蒿中分离得到的抗疟疾有效单体^[1], 研究发现其具有免疫抑制、抗菌、抗肿瘤等较广泛的药理作用, 但黄花蒿中青蒿素量低, 青蒿素的生产不能满足市场需求^[2-3]。科研人员已利用表型及 SSR、SNP 等分子标记开展黄花蒿分子育种工作, 以期培育青蒿素量高的黄花蒿品系, 并且青蒿素的次生代谢途径及其合成机制的研究也取得显著进展^[4-5]。研究表明青蒿素的生物合成是以萜类化合物共同前体——异戊烯基焦磷酸 (IPP) 为基础, IPP 在法呢基焦磷酸合酶 (FPS) 的作用下生成法呢基焦磷酸 (FPP), FPP 经紫穗槐二烯合酶 (ADS) 环化生成紫穗槐-4, 11-二烯, 然后经细胞色素 P450 单加氧酶 (CYP71AV1)、青蒿醛 $\Delta 11(13)$ 双键还原酶 (DBR2)、青蒿醛脱氢酶 (ALDH1) 等一系列酶促反应合成青蒿素^[6]。植物中存在甲羟戊酸 (MVA) 和 2-甲基赤藓糖醇-4-磷酸 (MEP) 2 条 IPP 合成途径, 分别定位于不同的亚细胞空间以及具有不同的 IPP 生成机制, 但二者存在协同效应。Towler 等^[7]发现青蒿素的生物合成可以利用来源于 MVA 和 MEP 2 条途径的 IPP。要认识青蒿素的合成机制必须了解合成青蒿素的生源、中间体及其关键酶, 从而开展温度、光照等生理生化研究以探讨对青蒿素积累的影响^[8-10]。2C-甲基-D-赤藓糖醇-2, 4-环焦磷酸合成酶 (2C-methyl-D-erythritol-2, 4-cyclodiphosphate synthase, MCS) 是 MEP 途径的第 5 个关键酶, 催化 2-磷酸-4-(胞苷-5'-二磷酸)-2-C-甲基-D-赤藓糖醇转化为 2C-甲基-D-赤藓糖醇-2, 4-环焦磷酸。研究青蒿素合成的分子机制, 克隆青蒿素生物合成途径关键酶基因, 可以为青蒿素基因工程提供候选基因^[11]。本研究拟以黄花蒿为材料, 克隆 MEP 途径中的 MCS 基因并进行序列特征分析, 建立 pET-21a (+)-MCS 原核表达体系, 为研究青蒿素合成代谢以及利用基因工程手段提高次生代谢物的量奠定基础。

1 材料与试剂

1.1 材料

黄花蒿 *Artemisia annua* L. 叶片采自于广东省广东药学院中药园, 由广东药学院中药学院腾希峰讲师鉴定。植物材料用液氮速冻后 -80 °C 保存备用。大肠杆菌 (*Escherichia coli*) DH-5 α 和 BL21 菌

种由本实验室保存。

1.2 试剂

Trizol 试剂购自 Invitrogen 公司, RACE 试剂盒购自 Clontech 公司, Taq DNA 聚合酶、限制性内切酶 Nhe I 和 Xho I、T4-DNA 连接酶购自 Takara 公司。pMD18-T 和 pET-21a (+) 由本实验室保存。引物由上海生工合成, 其他化学试剂为进口或国产分析纯。

2 方法

2.1 总 RNA 提取及 MCS 全长 cDNA 的克隆

按照 Trizol 试剂盒说明提取黄花蒿叶片总 RNA。根据 GenBank 中 AaMCS 基因 EST 序列设计一对特异性引物 MCS-F1 (5'-GTGCTTTGGGTCT-ACCTGAC-3') 和 MCS-R1 (5'-CCCAAAATTGCG-TCAACAACG-3'), 采用 BD Smart 5' RACE cDNA 扩增原理扩增 MCS 5' cDNA。3' 末端以接头 Olig-dT (5'-GGCCACGCGTCGACTAGTACT₁₇-3') 引导反转录合成一链 cDNA (PrimeScript II, Takara), 选择使用接头通用引物 3' Adapter (5'-GGCCACGCGTCGACTAGTAC-3') 和基因特异性引物 MCS-F1 扩增 MCS 3' cDNA。使用热启动和 Touch down PCR 程序: 94 °C 预变性 1 min 后, 10 循环: 94 °C、30 s, 66 °C、30 s (每个循环降低 1 °C), 72 °C、1 min; 25 个循环: 94 °C、30 s, 56 °C、30 s, 72 °C、1 min; 最后 72 °C、7 min。

2.2 MCS 基因的克隆及测序与分析

基因组 DNA 的提取采用酚氯仿法。根据已获得的 AaMCS 全长 cDNA 序列, 设计一对基因特异性引物 MCS-F2 (5'-CATGGCTACATCAACTGCTTC-3') 和 MCS-R2 (5'-CTTCTTCATCAAAAGTACAACG-3') 用于扩增 MCS 基因组全长。PCR 程序: 94 °C 预变性 3 min; 32 个循环: 94 °C、30 s, 56 °C、30 s, 72 °C、3 min; 最后 72 °C、10 min。MCS cDNA 和基因组扩增的 PCR 产物经 1.2% 的琼脂糖凝胶电泳, 用胶回收试剂盒 (Omega) 对目的片段进行纯化, 纯化产物与 pMD-18T 载体连接, 转化大肠杆菌 DH-5 α 感受态细胞, 对经 PCR 检测为阳性的克隆进行测序 (上海生工)。测序所得到的 cDNA 序列同源性比对和序列相似性搜索用 NCBI 网站 (<http://www.ncbi.nlm.nih.gov/blast>) BLAST 软件进行, 氨基酸序列的推断通过 ExPASy 网站 (<http://au.expasy.org>) 的程序翻译, 氨基酸序列的同源性比较由 ClustalW1.8 程序完成, 采用 MEGA 3.1 软件以邻位相联法构建进化树。

Fig. 1 Sequences of MCS gene and deduced amino acids

非编码区没有发现 mRNA 不稳定信号 (ATTTA), 仅有一个位于 PolyA 上游 28 bp 处的加尾信号 (AATAAA)。利用全长 MCS cDNA 序列在开放阅读框上游和下游分别设计特异性引物并成功扩增其基因组全长 (GenBank 登录号: KC142157)。AaMCS 基因全长为 2 540 bp, 包含 3 个外显子和 2 个内含子, 其中外显子 1 区域主要编码保守性较低的 N 末端。

3.2 MCS 氨基酸序列特征

AaMCS 氨基酸序列与其他已知同源蛋白质序列具有较高的同源性, 拟南芥 70.4%、葡萄 72.1%、

银杏 68.6%、长春花 70.8%。序列比对发现 AaMCS 蛋白质序列 N 末端 69 个氨基酸呈现高度变异, 而剩余氨基酸具有显著的同源性 (图 2)。进化树分析表明, AaMCS 和拟南芥、葡萄、长春花、大豆等植物的同源物集成一簇, 它们与玉米、高粱、水稻等植物的 MCS 源于共同祖先, 但分别处于进化上不同的分支。AaMCS 与银杏、三尖杉等裸子植物同源物的亲缘关系较远 (图 3)。因此, 研究结果基本反映 MCS 蛋白的分子进化过程, 同时也说明 AaMCS 是一种与植物 MEP 代谢途径相关的蛋白。

黄花蒿 (KC142157)	1	-----MATSTASIIYTTSPHITTTKAFSTRFPTRRSSILTRPTQHRAAVAVVAAVETENIPAVEKDIPAKVLP
拟南芥 (NP_850971)	1	-----MATSTQILLSSSLFHSQITEKPFILPATKIGVWRPEESLSLSCRPASVSAASSAVDVNESVISEKPTETLP
长春花 (Q9M4W3)	1	-----MAAAPIDHASSITKISSLEQIYP-----FTSHYRPSLVLSEHTLSLRSTLRFPVAAAATTAVETPRNSTAPSESIP
葡萄 (XP_002278406)	1	-----MAATSFYCSTAIPSEKIQDREMLFCSPWGGSETTPSYIRLSTRQSRILSLVWSAAASCAAVEAEPEFAAVTPSKIDS
银杏 (AAR95701)	1	-----MAAFSALVLPKWTIDGSLKSAISWVEMHQIRPINFSTHICVSRAPTLRRGQFQKRPVCSAMPVETIEQSAVAESAPVLP
黄花蒿 (KC142157)	70	FRVGHGFDLHRLFFGYPFIIGGDIIEHDRCCLASDGDVLIHCVDAILGALGLPDIGQIFPDSDFETEGAAASSVFLREAVRLDEAGYF
拟南芥 (NP_850971)	75	FRVGHGFDLHRLFFGYPFIIGGDIIEHDRCCLASDGDVLIHCVDAILGALGLPDIGQIFPDSDFETEGAAASSVFLREAVRLDEAGYF
长春花 (Q9M4W3)	75	FRVGHGFDLHRLFFGYPFIIGGDIIEHDRCCLASDGDVLIHCVDAILGALGLPDIGQIFPDSDFETEGAAASSVFLREAVRLDEAGYF
葡萄 (XP_002278406)	80	FRVGHGFDLHRLFFGYPFIIGGDIIEHDRCCLASDGDVLIHCVDAILGALGLPDIGQIFPDSDFETEGAAASSVFLREAVRLDEAGYF
银杏 (AAR95701)	83	FRVGHGFDLHRLFFGYPFIIGGDIIEHDRCCLASDGDVLIHCVDAILGALGLPDIGQIFPDSDFETEGAAASSVFLREAVRLDEAGYF
黄花蒿 (KC142157)	160	ICGLDATLILQRPKISSEKKTURATLSALIGADPAWVTLKAKTHEKVDISLGEYRSIAAETVULLIKK
拟南芥 (NP_850971)	165	ICGLDATLILQRPKISSEKKTURATLSALIGADPAWVTLKAKTHEKVDISLGEYRSIAAETVULLIKK
长春花 (Q9M4W3)	165	ICGLDATLILQRPKISSEKKTURATLSALIGADPAWVTLKAKTHEKVDISLGEYRSIAAETVULLIKK
葡萄 (XP_002278406)	170	ICGLDATLILQRPKISSEKKTURATLSALIGADPAWVTLKAKTHEKVDISLGEYRSIAAETVULLIKK
银杏 (AAR95701)	173	ICGLDATLILQRPKISSEKKTURATLSALIGADPAWVTLKAKTHEKVDISLGEYRSIAAETVULLIKK

一致氨基酸残基采用黑色背景表示, 其他保守的氨基酸残基采用灰色背景表示

The identical amino acids are highlighted in black, other highly conserved residues are shaded in grey

图 2 AaMCS 与其他物种 MCS 氨基酸序列同源性比较

Fig. 2 Homology comparison on MCS amino acids sequences between *A. annua* and other species

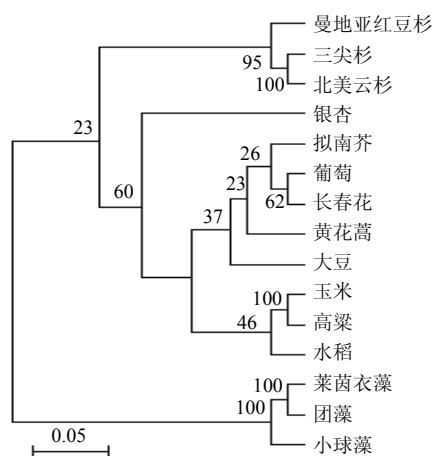


图 3 MCS 家族系统进化树

Fig. 3 Phylogenetic tree of MCS family

3.3 MCS 原核表达

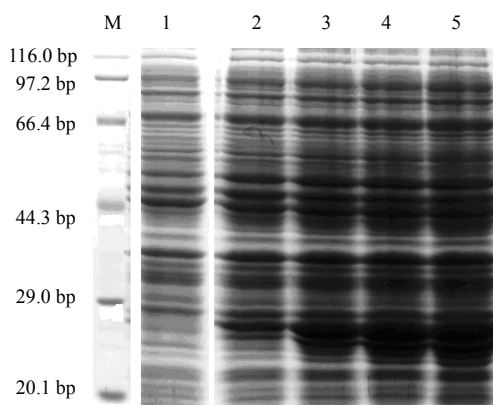
重组质粒经 *Nhe* I 和 *Xho* I 双酶切及测序鉴定。AaMCS 基因开放阅读框大小为 681 bp, 其编码的目的蛋白相对分子质量约为 2.438×10^4 。取 pET-21a (+)-MCS 阳性工程菌进行 SDS-PAGE 分析, 在

$2.90 \times 10^4 \sim 2.01 \times 10^4$ 可见特异性目的蛋白条带, 而对照组在相应位置未见明显条带, 其特异条带大小与通过生物信息学软件预测的目的蛋白大小一致; 通过 Clustal W 进行序列比对 16 种 MCS 蛋白质序列, MEGA3.1 构建 N-J 树 (图 4)。MCS 蛋白质序列选自 GenBank 和 SwissPort。黄花蒿 (KC142157), 拟南芥 (NP_850971), 葡萄 (XP_002278406), 长春花 (Q9M4W3), 大豆 (NP_001237894), 水稻 (NP_001047741), 玉米 (NP_001150687), 高粱 (XP_002452756), 银杏 (AAR95701), 曼地亚红豆杉 (ABB88956), 三尖杉 (ABD73009), 北美云杉 (ACN40993), 莱茵衣藻 (XP_001690985), 团藻 (XP_002947986), 小球藻 (EFN54075)。研究结果表明重组质粒已成功转入 BL21 中并能正常表达 AaMCS 的蛋白, 随着 IPTG 诱导表达时间的延长, 目的基因的蛋白表达产物明显增加。

3.4 MCS 组织特异性表达

检测 AaMCS 基因在不同组织的表达情况。分别提取根、叶、茎、花等组织总 RNA, 以 β -Actin

作为内参,进行 RT-PCR 半定量分析。结果表明 MCS 基因在黄花蒿不同组织中的表达量存在差异,花中表达量较高,根、茎中表达量低(图 5)。



M-Marker 1-对照 2~5-IPTG 诱导 1.5、3、4.5、6 h 后 MCS 的表达

M-Marker 1-control 2~5-MCS expression induced by IPTG for 1.5, 3, 4.5, and 6 h

图 4 重组 MCS SDS-PAGE 电泳
Fig. 4 SDS-PAGE electrophoretograms of recombinant MCS

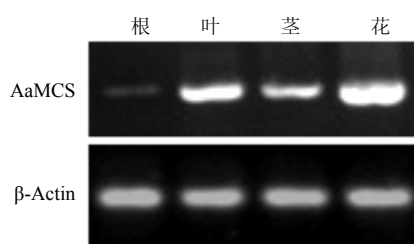


图 5 MCS 组织表达

Fig. 5 Expression of MCS in different tissues

4 讨论

本研究报道了在黄花蒿中克隆了 MCS 基因,该基因编码的氨基酸序列与其他植物 MEP 途径 MCS 蛋白的序列同源性高,表明 AaMCS 为 MEP 途径的第 5 个关键酶 MCS。黄花蒿青蒿素生物合成途径的研究已取得了显著进展,其代谢途径中多个编码关键酶基因已经被克隆^[12-13]。本研究在 GenBank EST 序列的基础上,设计特异引物并通过 RACE 和基因组扩增的方法克隆获得了编码 AaMCS 的基因。对 AaMCS 编码的氨基酸序列进行生物信息学分析,结果显示 AaMCS 与来源于其他植物的 MCS 同源性高达 68%~72%,并且发育树也反映 AaMCS 与来源于被子植物的 MCS 亲缘关系最近。植物可以通过在胞质 MVA 途径和质体 MEP

途径合成 IPP,银杏 *Gingo biloba* L.、红豆杉 *Taxus chinedsis* (Pilger) Rehd. 及母菊 *Matrica riarecutita* L. 等植物合成多种萜类化合物的 IPP 单体来源于 MVA 和 MEP 2 条途径^[14],黄花蒿中青蒿素的生物合成也可以利用两条途径的 IPP 合成青蒿素^[7,15]。因此,对黄花蒿质体中 MEP 途径中相关酶基因的研究对进一步理解青蒿素生物合成的调控机制具有重要意义。与银杏 MCS 基因的检测结果相似^[16],AaMCS 在检测的 4 种组织中均表达,其中叶和花中表达量相对较高,而根和茎中表达量较低。将 MCS 编码区序列与表达载体 pET-21a(+) 重组,构建 pET-21a(+)-MCS 的原核表达载体并转入大肠杆菌 BL21,获得稳定的 pET-21a(+)-MCS 原核表达体系且目的蛋白表达量较高。原核生物具有生长周期短、繁殖快等特点,为下一步纯化 AaMCS 蛋白以及进行生物功能分析奠定了基础。以体外实验来模拟体内条件生产生物体内的化学成分正是目前科学界的研究热点,采用生物技术缓解资源短缺问题也是中药现代化的目的之一。该研究将有助于深入认识青蒿素生物合成代谢途径的分子基础,也为青蒿素的基因工程提供候选基因。

参考文献

- [1] 青蒿素结构研究协作组. 一种新型的倍半萜内酯——青蒿素 [J]. 科学通报, 1977, 22(3): 142.
- [2] 陆金健. 青蒿素类化合物抗肿瘤研究进展 [J]. 中国药理通报, 2010, 26(6): 818-820.
- [3] 张晓蓉, 彭光花, 陈功锡, 等. 黄花蒿残渣挥发油化学成分及其抑菌活性分析 [J]. 中草药, 2011, 42(12): 2418-2420.
- [4] 韦树根, 马小军, 冯世鑫, 等. 中国黄花蒿主产区种质资源评价 [J]. 中国中药杂志, 2008, 33(3): 241-244.
- [5] Graham I A, Besser K, Blumer S, et al. The genetic map of *Artemisia annua* L. identifies loci affecting yield of the antimalarial drug artemisinin [J]. *Science*, 2010, 327(5963): 328-331.
- [6] O'Neill P M, Barton V E, Ward S A. The molecular mechanism of action of artemisinin-the debate continues [J]. *Molecules*, 2010, 15(3): 1705-1721.
- [7] Towler M J, Weathers P J. Evidence of artemisinin production from IPP stemming from both the mevalonate and the nonmevalonate pathways [J]. *Plant Cell Rep*, 2007, 26: 2129-2136.
- [8] 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.

- [9] 杨瑞仪, 卢元媛, 杨雪芹, 等. 低温诱导黄花蒿中青蒿素的生物合成及其机制研究 [J]. 中草药, 2012, 43(2): 350-354.
- [10] 孙年喜, 李隆云, 钟国跃, 等. 不同生长期土壤水分处理对黄花蒿生理特性及产量的影响 [J]. 中国中药杂志, 2009, 34(4): 386-389.
- [11] Liu C, Zhao Y, Wang Y. Artemisinin: current state and perspectives for biotechnological production of an antimalarial drug [J]. *Appl Microbiol Biotechnol*, 2006, 72(1): 11-20.
- [12] Covello P S, Teoh K H, Polichuk D R, *et al.* Functional genomics and the biosynthesis of artemisinin [J]. *Phytochemistry*, 2007, 68(14): 1864-1871.
- [13] Wen W, Yu R M. Artemisinin biosynthesis and its regulatory enzymes: Progress and perspective [J]. *Pharmacogn Rev*, 2011, 5(10): 189-194.
- [14] Dubey V S, Bhalla R, Luthra R. An overview of the non-mevalonate pathway for terpenoid biosynthesis in plants [J]. *J Biosci*, 2003, 28(5): 637-646.
- [15] Schramek N, Wang H, Römisch-Margl W, *et al.* Artemisinin biosynthesis in growing plants of *Artemisia annua* A $^{13}\text{CO}_2$ study [J]. *Phytochemistry*, 2010, 71(2/3): 179-187.
- [16] Gao S, Lin J, Liu X F, *et al.* Molecular cloning, characterization and functional analysis of a 2C-methyl-D-erythritol 2, 4-cyclodiphosphate synthase gene from *ginkgo biloba* [J]. *J Biochem Mol Biol*, 2006, 391(5): 502-510.