

## • 药材与资源 •

胡黄连 2-*C*-甲基-*D*-赤藓糖-4-磷酸胞苷转移酶基因的生物信息学分析郑玉娟<sup>1,2</sup>, 陈 容<sup>1,2</sup>, 周文化<sup>1,2\*</sup>

1. 稻谷及副产物深加工国家工程实验室, 湖南 长沙 410004

2. 中南林业科技大学食品学院, 湖南 长沙 410004

**摘 要:** 目的 研究胡黄连 *Picrorhiza kurrooa* 2-*C*-甲基-*D*-赤藓糖-4-磷酸胞苷转移酶(MCT)基因特征并预测 MCT 结构与功能位点。方法 以胡黄连 MCT 基因为研究对象, 利用 NCBI 网站及生物信息学软件对其碱基分布、氨基酸组成、亲疏水性、保守区及二级结构和三级结构进行预测和分析; 并对 9 个物种的 MCT 基因进行多重比对与进化分析。结果 胡黄连 MCT 基因的 mRNA 序列长为 1 216 bp (GenBank JQ991625.1), 编码由 399 个氨基酸组成的蛋白质, 该蛋白质相对分子质量为 44 448.5, 其中量最高的为 Ser (10.0%); 整个多肽中疏水性氨基酸占 59.5%, 平均疏水值为 0.050; 与丹参 *Salvia miltiorrhiza* MCT 氨基酸序列对比同源性最高, 达 99%。结论 胡黄连 MCT 基因处于稳定状态, 编码的蛋白为疏水性蛋白, 在进化过程中是相对保守的, 获得的保守区段序列信息为其他物种 MCT 基因的克隆奠定了基础。

**关键词:** 2-*C*-甲基-*D*-赤藓糖-4-磷酸胞苷转移酶; 胡黄连; 丹参; 生物信息学; 蛋白质结构

中图分类号: R282.12 文献标志码: A 文章编号: 0253-2670(2014)15-2218-06

DOI: 10.7501/j.issn.0253-2670.2014.15.019

Bio-informatics analysis of 2-*C*-methyl-*D*-erythritol 4-phosphate cytidylytransferase gene in *Picrorhiza kurrooa*ZHENG Yu-juan<sup>1,2</sup>, CHEN Rong<sup>1,2</sup>, ZHOU Wen-hua<sup>1,2</sup>

1. National Engineering Laboratory for Rice and Byproducts Processing, Changsha 410004, China

2. School of Food Science and Engineering, Center South University of Forestry and Technology, Changsha 410004, China

**Abstract: Objective** To study the characteristics of 2-*C*-methyl-*D*-erythritol 4-phosphate cytidylytransferase (MCT) gene and to predict the structure and function site of MCT. **Methods** Based on MCT gene of *Picrorhiza kurrooa*, NCBI website and bio-informatics software were used to predict and analyze the base distribution, amino acid composition, hydrophilicity or hydrophobicity, and secondary and three-level structures of hyper-conservative region. The results were compared with MCT gene sequences in other eight species and the related evolution analysis was followed. **Results** The mRNA sequence of MCT gene in *P. kurrooa* was 1 216 bp (GenBank: JQ991625.1), coding the protein containing 399 amino acids, and the relative molecular mass of protein was 44 448.5 with 10.0% Ser in volume; In polypeptide, hydrophobic amino acid was 59.5% in volume, and the average hydropathicity was 0.050; The homology was found as high as 99% compared with MCT gene in *Salvia miltiorrhiza*. **Conclusion** The MCT gene in *P. kurrooa* was quite stable and the coded protein was hydrophobic, which was quite conservative in evolution; The sequence information in conservative region gained in the test provides the basis for clone of novel gene in other species.

**Key words:** 2-*C*-methyl-*D*-erythritol 4-phosphate cytidylytransferase; *Picrorhiza kurrooa* Pennel; *Salvia miltiorrhiza* Alexander Bunge; bio-informatics; protein structure

胡黄连 *Picrorhiza kurrooa* Pennel 为玄参科胡黄连属的多年生草本植物, 分布区比较狭窄, 特产于横断山区和东喜马拉雅。其具有保肝利胆、抗菌消炎、抗哮喘、降血糖、神经保护、免疫调节等多

种药理作用, 临床上应用非常广泛, 环烯醚萜苷类成分是其作用成分之一<sup>[1-2]</sup>。近些年, 由于过度采挖, 野生胡黄连资源濒临枯竭, 其作为国家二级保护植物, 已被列入《中国珍稀濒危保护植物名

收稿日期: 2014-02-22

作者简介: 郑玉娟 (1990—), 女, 山东东营人, 硕士, 研究方向为特色农林产品深加工。E-mail: zhengyajuan9003@163.com

\*通信作者 周文化 (1969—), 男, 湖南澧县人, 教授, 博士, 研究方向为特色农林产品深加工, 传统食品工业化研究。

Tel: (0731)85623240 E-mail: zhouwenhua@126.com

录》<sup>[3]</sup>, 对其继续科学保护和合理利用已刻不容缓。萜类化合物 (terpenoids) 又称类异戊二烯, 是自然界中分布最广泛、结构最复杂的一类次生代谢化合物<sup>[4-8]</sup>。其在植物生理过程中发挥着重要作用, 部分萜类化合物还具有重要的药用价值和较高的经济价值。随着分子生物学和生物信息学等学科的发展, 萜类化合物的合成途径及合成过程中的关键酶成为植物次生代谢研究领域的热点之一。绝大多数萜类化合物的合成前体是异戊烯基焦磷酸 (IPP), 植物体内 IPP 的生物合成主要存在 2 条不同的代谢途径: 一是位于细胞质中的甲羟戊酸 (MVA) 途径, 另一条是位于质体中的 2-甲基-D-赤藓糖醇-4-磷酸 (MEP) 途径<sup>[9]</sup>。2-C-甲基-D-赤藓醇-4-磷酸脱氢酶 (MCT) 催化 MEP 途径中的第 3 个反应, 即催化 CDP 和 EP 在胞苷三磷酸 (CP) 的参与下合成 CDP-ME, 是 IPP 合成的关键酶之一<sup>[10-13]</sup>。

Rohdich 等<sup>[14]</sup>使用同位素标记法验证了 MCT 在大肠杆菌萜类前体物质生物合成中的重要作用。此外, MCT 缺失型 *E. coli* 菌株 NMW33 人为添加 MVA, 可以在 LB 平板上生长, 而缺失型不生长, 这说明了 MCT 是 IPP 生物合成所必需的<sup>[15]</sup>。近年来, 相继有文献报道多个物种的 MCT 基因得到解析, 但尚未有报道利用生物信息方法系统研究这些基因, 这制约其他物种中该基因的克隆与功能验证。生物信息学是生命科学和信息科学的交叉学科, 是后基因组时代的重要研究工具之一。本实验利用生物信息学的方法对胡黄连 MCT 基因以及 GenBank 上已发表的其他植物 MCT 基因进行序列分析和功能预测, 旨在为其他植物 MCT 基因的克隆、功能验证提供理论和实践参考依据。

## 1 材料

样品采自印度罗唐山隘和 Palampur 研究所保护地收集, 经笔者鉴定为胡黄连 *Picrorhiza kurrooa* Pennel。

## 2 方法

参照 NCBI 数据中报道的胡黄连 (JQ991625.1)、拟南芥 *Arabidopsis thaliana* Carl Linnaeus (NM\_126305.2)、丹参 *Salvia miltiorrhiza* Alexander Bunge (JN831096)、蓖麻 *Ricinus communis* James A. Duke (XM\_002540422.1)、罗汉果 *Siraitia grosvenorii* C. Jeffrey ex Lu et Z. Y. Zhang (HQ128560.1)、萝芙木 *Rauvolfia verticillata* Henri

Ernest Baillon (JX104650.1)、毛果杨 *Populus trichocarpa* John Torrey & Asa Gray (EU693021.1)、橡胶 *Hevea brasiliensis* Johannes Müller Argoviensis (AB294703.1) 和银杏 *Ginkgo biloba* Carl Linnaeus (DQ102360.1) 的 MCT 基因的 mRNA 序列以及其编码的蛋白质序列。

胡黄连 MCT 基因 mRNA 序列由叶组织中提取获得的 RNA 反转录合成 cDNA 特异性扩增获得<sup>[16-18]</sup>。利用 BioEdit 软件分析胡黄连 MCT 基因 mRNA 序列和氨基酸序列的相对分子质量、碱基; 利用 NCBI 网站的 ORF Finder 程序 (<http://www.ncbi.nlm.nih.gov/projects/gorf/orf.cgi>) 预测开放阅读框 (ORF) 利用 BioEdit 软件对胡黄连 MCT 基因的氨基酸序列进行初步疏水性分析, 再通过 ExPasy 的 ProtScale 程序 (<http://expasy.org/csibin/protScale.pl/>), 使用 Hphob/Kyte & Doolittle 方法对其进行详细的疏水性分析, 并获得氨基酸组成、碱基分布、氨基酸比例等; 将获得的氨基酸序列利用 Clustal X 软件进行氨基酸序列之间的多重比对分析; 利用 DNAMAN 软件对 9 种植物 MCT 基因的氨基酸序列构建进化树; 在 NCBI 中对胡黄连 MCT 蛋白的保守序列进行预测, 以及二级结构、三级结构的预测分析<sup>[19-22]</sup>。

## 3 结果与分析

### 3.1 胡黄连 MCT 基因 mRNA 序列的分析

通过 BioEdit 软件和 NCBI 网站分析发现, 胡黄连 MCT 基因 mRNA 序列长为 1 216 bp, 其含有多个复制起始位点, ORF 在 148~1 083 bp, 共编码 311 个氨基酸, 相对分子质量为 367 599.00。从表 1 可以看出, A+T 量较高为 62.09%; G+C 量较少为 37.91%。

### 3.2 胡黄连 MCT 基因编码的蛋白质氨基酸组成分析

由表 2 可知, 胡黄连 MCT 基因编码的蛋白质中共有 20 种氨基酸, 共计 399 个氨基酸, 相对分子

表 1 胡黄连 MCT 基因 mRNA 序列的碱基组成

Table 1 Base composition of mRNA sequence of MCT gene in *P. kurrooa*

| 碱基类型 | 数目  | 含量 / % |
|------|-----|--------|
| A    | 357 | 29.36  |
| C    | 224 | 18.42  |
| G    | 237 | 19.49  |
| T    | 398 | 32.73  |

质量为 44 448.5, 其中量最高的为 Ser, 为 10.0%; Leu、Val、Lys、Ile 也较多, 分别为 9.8%、8.0%、7.8%、7.3%; 而量最低的是 Trp, 为 0.5%。N 端第一个氨基酸为 Gly。

表 2 胡黄连 MCT 编码的蛋白质的氨基酸组成  
Table 2 Composition of amino acid of MCT protein  
in *P. kurrooa*

| 氨基酸名称 | 数目 | 含量 / % |
|-------|----|--------|
| Ala   | 21 | 5.3    |
| Cys   | 13 | 3.3    |
| Asp   | 18 | 4.5    |
| Glu   | 21 | 5.3    |
| Phe   | 24 | 6.0    |
| Gly   | 20 | 5.0    |
| His   | 9  | 2.3    |
| Ile   | 29 | 7.3    |
| Lys   | 31 | 7.8    |
| Leu   | 39 | 9.8    |
| Met   | 6  | 1.5    |
| Asn   | 16 | 4.0    |
| Pro   | 22 | 6.0    |
| Gln   | 12 | 3.0    |
| Arg   | 11 | 2.8    |
| Ser   | 40 | 10.0   |
| Thr   | 19 | 4.8    |
| Val   | 32 | 8.0    |
| Trp   | 2  | 0.5    |
| Tyr   | 14 | 3.5    |

### 3.3 胡黄连 MCT 基因编码的氨基酸序列疏水性分析

组成蛋白质的氨基酸分子都具有一定的极性, 这些氨基酸分子的极性对蛋白质高级结构及功能性结构域的形成有重要的影响。分析可知该蛋白为稳定蛋白。从图 1 可知, 疏水值最大为 3.167, 375 bp 处的疏水性最强, 疏水值最小为 -2.244, 88 bp 处亲水性最强。整个多肽中含 59.5% 的疏水性氨基酸, 疏水性平均值为 0.050, 推断该蛋白为疏水性蛋白。

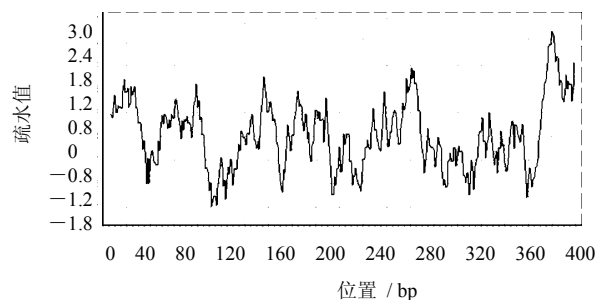


图 1 胡黄连 MCT 基因编码氨基酸的疏水性预测

Fig. 1 Hydrophobicity prediction of amino acid sequence of MCT gene in *P. kurrooa*

### 3.4 氨基酸序列之间的多重比对分析

对 MCT 的氨基酸序列进行同源性分析(图 2), 结果表明胡黄连与丹参、拟南芥、罗汉果、萝芙木、橡胶、银杏具有较高同源性, 保守区主要集中在中部和尾部, 而 N 端变异性比较大。由于基因进化

|      |                                                                                                             |     |
|------|-------------------------------------------------------------------------------------------------------------|-----|
| 胡黄连  | ..... NPTI QLGLSLPTSSFLFLETNSI KFP I KLCSLFSLNPNYSKNYSFNAKYTH. KI SFKNPRFFH..... ARN TCSAD. AAAQGSVEV       | 82  |
| 丹参   | ..... NSTLQI TNLPSSAASPSVSPSS..... FPLFLGATLGLPTKFRPSFALTTVYNNRKI DSLKNPRFS..... RI SCSAG. GTD. TQVV        | 75  |
| 毛果杨  | ..... MALACNTVPSSLSLPHKSKLCI SSQSRHNL. YEFVGAKTSFQALCI LSPQRRHNLNAPPSARRSFQTVARFRGYGVPKRRPLRAFADKGADSGTDDV  | 11  |
| 银杏   | ..... NSI LQLSPPLSLSSPI LLGSN. YQI SERSLPLLSHNLSTLPHFKKSSLSSTKSNFRSPELS..... LLRVITCSAVPGESKSPDVV           | 99  |
| 萝芙木  | ..... NAMLQTLGF. I TSPITFLCPK. LKVKLNSYLWFSYR..... SQVQKLDKSKRVNRNRYKRDAL..... LLSTI KCSSTGTGDN. NVV        | 81  |
| 拟南芥  | ..... NASVSHPRMNCSPSS..... SPS..... KFP I KORM AAKNGQSI SHFSLQLTGHQHKFRNEKI LKS..... VI KCSAAEVGESS..... EF | 74  |
| 罗汉果  | ..... NGHLLHLNLTAVSSSSASSSSSS..... SSSI I SSSNSLFP..... CHKLTIW QKLPRI ATT..... TI KCSAKTENS..... AT        | 65  |
| 萝芙木  | ..... NGHLLHLNLTI I SPSASFSSNSLFPNCNHPQI PTNSFHLPSLSKSSVHKI SWI QKLPRI ATA..... TI KCSAKI ENS..... AA       | 77  |
| 橡胶   | ..... m s p n s s                                                                                           |     |
| 相同部分 | .....                                                                                                       |     |
| 胡黄连  | VREKSVSVLLAGCGKRRGASMPKQY PILLGQPIALYSFYTFSSKNPEVKEI VVCEPSYQI FEEAKENI CVDLKFALPGKERQDSVYSLGLAIDSNSE       | 182 |
| 丹参   | VREKSVSVLLAGCGKRRGASMPKQY PILLGQPIALYSFYTFSSKNPEVKEI VVCEPSYQI FEEAKENI III DLKFALPGKERQDSVYSLGLEEDLSE      | 175 |
| 毛果杨  | VREKSVSVLLAGCGKRRGASMPKQY PILLGQPIALYSFYTFSSKNPEVKEI VVCEPSYQI FEEPSQEQI NVDLKFALPGKERQDSVYSLGLEVDLSE       | 111 |
| 银杏   | VPEKSVSVLLAGCGKRRGASMPKQY PILLGQPIALYSFYTFSSKNPEVKEI VVCEPSYQI FEEARNSVKI NDLKFALPGKERQDSVYSLGLEVDLSE       | 199 |
| 萝芙木  | VREKSVSVLLAGCGKRRGASMPKQY PILLGQPIALYSFYTFSSKNPEVKEI VVCEPSYQI FEEAKENI EVDLKFALPGKERQDSVYSLGVANDLSA        | 181 |
| 拟南芥  | VREKSVSVLLAGCGKRRGASMPKQY PILLGQPIALYSFYTFSSKNPEVKEI VVCEPSYQI FEEAKENI DVLDFALPGKERQDSVYSLGLEVDLSE         | 173 |
| 罗汉果  | VREKSVSVLLAGCGKRRGASMPKQY PILLGQPIALYSFYTFSSKNPEVKEI VVCEPSYQI FEEAKENI DI DLKFALPGKERQDSVYSLGLEVDLSE       | 174 |
| 萝芙木  | VREKSVSVLLAGCGKRRGASMPKQY PILLGQPIALYSFYTFSSKNPEVKEI VVCEPSYQI FEEAKENI NVDLKFALPGKERQDSVYSLGLEVDLSE        | 165 |
| 橡胶   | VREKSVSVLLAGCGKRRGASMPKQY PILLGQPIALYSFYTFSSKNPEVKEI VVCEPSYQI FEEAKENI NVDLKFALPGKERQDSVYSLGLEVDLSE        | 177 |
| 相同部分 | vkeksvsvllaggkgrngasmpkqylpllgqpi alysfytfss m evkei vvvdpsyqdi fedake i vdlkf alpgkerqdsyvslglevdl nse     |     |
| 胡黄连  | LVCVHDSARPLVLTETVAKVLKDGWLL GAAVGVPAKATI KEANSEFVVVTL DRKTLVEMQTPQVI KPDLLKKGFEHVNRGLEVTEDVSI VEHLKHPV      | 282 |
| 丹参   | LVCVHDSARPLVLTETVAKVLKDGWLL GAAVGVPAKATI KEANSEFVVVTL DRKTLVEMQTPQVI EPDILLKKGFEHVNRGLEVTEDVSI VEHLKHPV     | 275 |
| 毛果杨  | LVCVHDSARPLVLTETVAKVLKDGWLL GAAVGVPAKATI KEADGESFVVVTL DRKTLVEMQTPQVKNPELLKKGFEHVNRGLEVTEDVSI VEHLKHPV      | 211 |
| 银杏   | LVCVHDSARPLVLTETVAKVLKDGWLL GAAVGVPAKATI KEADGESFVVVTL DRKTLVEMQTPQVI KPNLLKKGFEHVNRGLEVTEDVSI VEHLKHPV     | 299 |
| 萝芙木  | LVCVHDSARPLVLTETVAKVLKDGWLL GAAVGVPAKATI KEADGESFVVVTL DRKTLVEMQTPQVI KPDLLKKGFEHVNRGLEVTEDVSI VEHLKHPV     | 281 |
| 拟南芥  | LVCVHDSARPLVLTETVAKVLKDGWLL GAAVGVPAKATI KEANSEFVVVTL DRKTLVEMQTPQVI KPDLLKKGFEHVNRGLEVTEDVSI VEHLKHPV      | 273 |
| 罗汉果  | LVCVHDSARPLVLTETVAKVLKDGWLL GAAVGVPAKATI KEANSEFVVVTL DRKTLVEMQTPQVI KPDLLKKGFEHVNRGLEVTEDVSI VEHLKHPV      | 274 |
| 萝芙木  | LVCVHDSARPLVLTETVAKVLKDGWLL GAAVGVPAKATI KEANSEFVVVTL DRKTLVEMQTPQVI KPDLLKKGFEHVNRGLEVTEDVSI VEHLKHPV      | 256 |
| 橡胶   | LVCVHDSARPLVLTETVAKVLKDGWLL GAAVGVPAKATI KEANSEFVVVTL DRKTLVEMQTPQVI KPDLLKKGFEHVNRGLEVTEDVSI VEHLKHPV      | 277 |
| 相同部分 | lvci hdsarplvlt edvekvldkgw gaavlgvpakatikeans sfvv tldrktl vemqtpqvi kp lllkkgfelvnrglevtdvsi vchlkhpv     |     |
| 胡黄连  | YI TEGSYTNI RVITPDDLLAERI LNPD.....                                                                         | 311 |
| 丹参   | YI TEGSYTNI RVITPDDLLAERI LNPD.....                                                                         | 304 |
| 毛果杨  | YI TEGSYTNI RVITPDDLLAERI LNSTS.....                                                                        | 240 |
| 银杏   | YI TEGSYTNI RVITPDDLLAERI LKQS.....                                                                         | 327 |
| 萝芙木  | YI TEGSYTNI RVITPDDLLAERI LNPTSAKSS.....                                                                    | 314 |
| 拟南芥  | YVSGSYTNI RVITPDDLLAERI SEDS.....                                                                           | 302 |
| 罗汉果  | FI TEGSYTNI RVITPDDLLAERI NLDS.....                                                                         | 303 |
| 萝芙木  | FI TEGSYTNI RVITPDDLLAERI NMDS.....                                                                         | 285 |
| 橡胶   | YI TEGSYTNI RVITPDDLLAERI NLNSES.....                                                                       | 310 |
| 相同部分 | yi tegsytni kvittpddlllaeri ln s                                                                            |     |

图 2 MCT 序列氨基酸的多重比对分析

Fig. 2 Multiple comparison analysis of amino acid of MCT sequence

等原因,也存在较大差异的区段,这些区段可能为基因进化理论研究提供一些参考依据。BlastP 的结果表明 MCT 属于糖基转移酶 (GT-A) 家族。

### 3.5 进化树的构建与分析

由图 3 可知,胡黄连与丹参亲缘关系最近,属于同一目即唇形目,与萝芙木分支较近,与蓖麻、

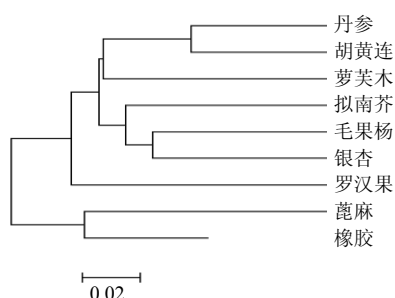


图 3 MCT 基因编码氨基酸序列进化树

Fig. 3 Evolutionary tree of amino acid sequence of MCT gene

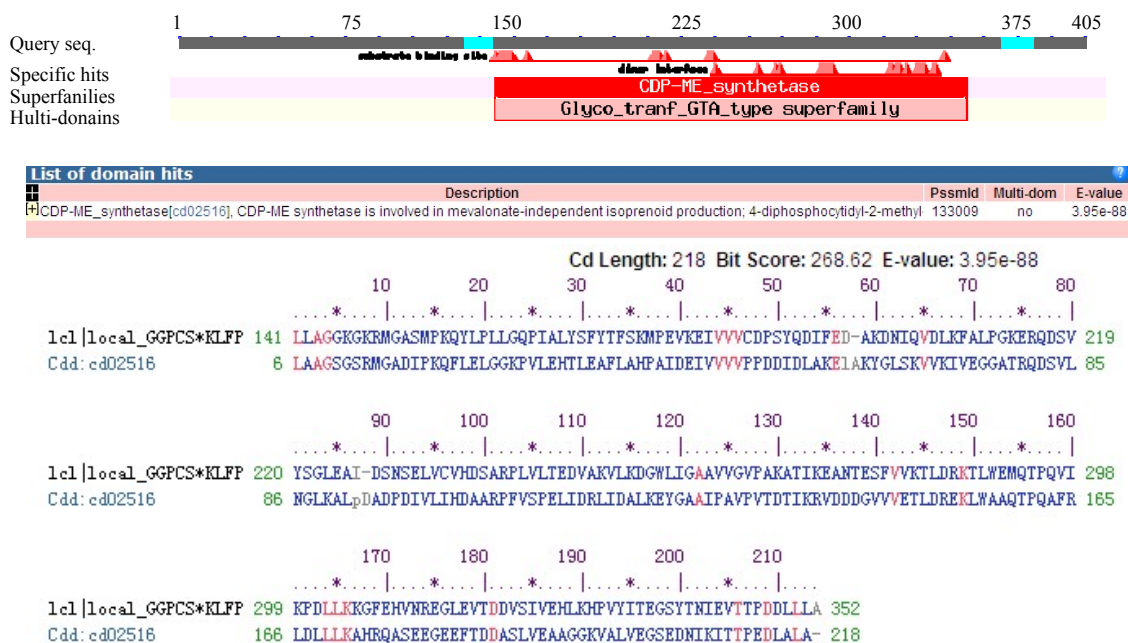


图 4 胡黄连 MCT 保守结构区域的预测

Fig. 4 Prediction of conserved domain of MCT protein in *P. kurrooa*

表 3 SOPMA 对胡黄连 MCT 二级结构分析  
Table 3 Analysis on secondary structure of MCT in *P. kurrooa* by SOPMA

| 折叠方式              | 数目  | 比例 / % |
|-------------------|-----|--------|
| $\alpha$ -螺旋 (Hh) | 64  | 20.58  |
| 延长带 (Ee)          | 91  | 29.26  |
| $\beta$ -转角 (Tt)  | 30  | 9.65   |
| 无规则卷曲 (Cc)        | 126 | 30.51  |

橡胶、罗汉果、银杏较远。

### 3.6 MCT 基因的保守区预测

保守序列 (conserved sequence) 对推测进化结果有特殊意义。在 NCBI 的蛋白保守结构域数据库 (conserved domain database, CDD) 中对 MCT 进行蛋白保守区预测,结果表明 (图 4) 与该基因匹配的蛋白为 MCT,保守序列在中间和 C 端,这也印证了不同物种的氨基酸序列之间的多重比对分析结果。

### 3.7 MCT 编码基因的二级结构分析与三级结构预测

用 SOPMA 对胡黄连 MCT 二级结构分析表明 (表 3 和图 5),该蛋白由 20.58% 的  $\alpha$ -螺旋、29.26% 的延伸链、9.65% 的  $\beta$ -转角及 30.51% 的无规则卷曲组成。无规则卷曲是胡黄连 MCT 中最多的结构元件。

氨基酸的排列顺序即蛋白质的一级结构决定了蛋白质的高级结构,而该蛋白的生物学功能由高级

结构所决定。采用 Swiss-Model 得到 MCT 蛋白的三级结构 (图 6),结果表明 MCT 包含 13 个  $\alpha$ -螺旋和 20 个  $\beta$ -片层,有助于理解蛋白质结构与功能之间的相关性。所有已知植物 MCT 中都具有 2 个保守的活性位点: (1) 底物 MEP 结合位点; (2) 依赖于  $Mg^{2+}$  的底物 CTP 结合位点,不同植物来源的 MCT 在 N 端非催化区域的氨基酸序列差异很大,在催化区域序列非常保守,萝芙木 MCT 中,CTP 结

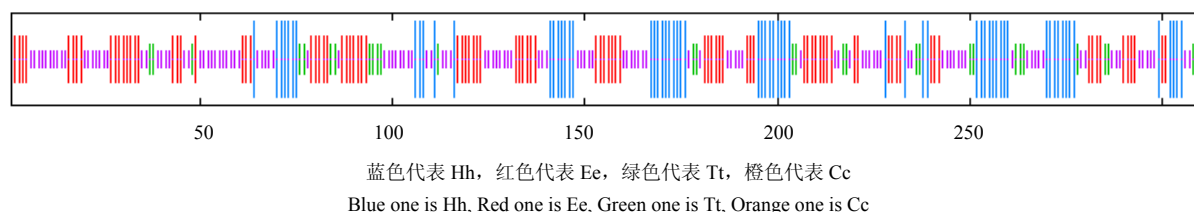


图5 胡黄连 MCT 二级结构预测

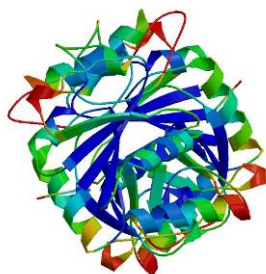
Fig. 5 Secondary structure prediction of MCT in *P. kurrooa*

图6 胡黄连 MCT 三级结构预测

Fig. 6 Prediction of three-level structure of MCT in *P. kurrooa*

合位点由 Arg99-Lys106 组成, MEP 结合位点是 Arg236-Lys292, 由此可以推测胡黄连 MCT CTP 结合位点和 MEP 结合位点分别由 Arg99-Lys106 和 Arg236-Lys292 组成, CTP 和 MEP 结合位点在一级结构上距离较远, 由于 MCT 全酶结构是一个二聚体, 所以 2 个酶活性中心在一级结构是可能存在距离的<sup>[20-26]</sup>。

#### 4 讨论

通过 NCBI 网站和生物信息学软件研究胡黄连 MCT 基因的碱基分布、氨基酸组成, 结果发现, 该基因中 G+C 的碱基量约为 37.91%, 低于 50%, 表明发生错配的概率较低, 核苷酸处于稳定状态。通过 Clustal X 等生物信息学软件将胡黄连 MCT 与其他物种基因序列进行多重性比对和进化分析, 序列同源性较高, 部分区段同源性达 100%, 说明其在进化中是十分保守的, 也有助于其他物种 MCT 基因的克隆与分离。这些保守区基因信息为其他物种中该基因的克隆提供了十分有价值的序列信息, 为加快基因克隆及 MCT 基因的分子调控研究奠定基础。较大差异的基因区段可能与基因的进化与变异相关, 这些仍待确定。通过 NCBI 对该基因的蛋白质进行二级和三级结构预测, 也有助于未来该基因在生物体内的调控研究。

#### 参考文献

- [1] 何希瑞, 李倩, 张春玲, 等. 胡黄连化学成分及单体化合物药理活性研究新进展 [J]. 环球中医药, 2012, 5(9): 708-713.
- [2] Lauris E K, Charles S B, William N H. Structure of a tetragonal crystal form of *Escherichia coli* 2-C-methyl-D-erythritol 4-phosphate cytidyltransferase [J]. *Biol Crystallogr*, 2003, 59: 607-610.
- [3] 国家环境保护局. 中国珍稀濒危保护植物名录 (第 1 册) [M]. 北京: 科学出版社, 1987.
- [4] Ma Y M, Yuan L C, Wu B, et al. Genome-wide identification and characterization of novel genes involved in terpenoid biosynthesis in *Salvia miltiorrhiza* [J]. *J Exp Bot*, 2012, 63(7): 2809-2823.
- [5] Liu L H, Li C H, Yang J L, et al. Bioinformatics analysis of 4-hydroxy-3-methyl-2-butenyl diphosphate synthase (HDS) in *Populus trichocarpa* [J]. *Agric Sci Technol*, 2012, 13(1): 24-28.
- [6] 黄瑛, 曾庆平. 萜类生物合成的基因操作 [J]. 中国生物工程杂志, 2006, 26(1): 60-64.
- [7] 刘娟, 刘斌, 折改梅, 等. 单萜萜类化合物及其药理活性研究进展 [J]. 现代药物与临床, 2010, 25(2): 81-93.
- [8] 郭建华, 田成旺, 刘晓, 等. 中药环烯醚萜类化合物研究进展 [J]. 药物评价研究, 2011, 34(4): 293-297.
- [9] 马靓, 丁鹏, 杨广笑, 等. 植物类萜生物合成途径及关键酶的研究进展 [J]. 生物技术通报, 2006(1): 22-30.
- [10] 李莉, 高凌云, 董越, 等. 植物类异戊二烯生物合成相关酶基因研究进展 [J]. 浙江师范大学学报: 自然科学版, 2008, 31(4): 461-466.
- [11] 郑月. 萝芙木 MCT、HDS、SCG 基因的克隆与遗传转化体系的建立 [D]. 重庆: 西南大学, 2011.
- [12] 张长波, 孙红霞, 巩中军, 等. 植物萜类化合物的天然合成途径及其相关合酶 [J]. 植物生理学通报, 2007, 43(4): 779-786.
- [13] 汪婉宜, 申业, 黄璐琦, 等. 萜类合成甲羟戊酸途径的关键基因的克隆及表达分析 [J]. 中国食物与营养, 2011, 17(9): 24-28.

- [14] Rohdich F, Wungsintaweekul J, Fellermeier M, *et al.* Cytidines 5'-triphosphate-dependent biosynthesis of isoprenoids: YgbP Protein of *Escherichia coli* catalyzes the formation of 4-diphosphocytidyl-2-C-methylerythritol [J]. *Proc Natl Acad Sci USA*, 1999, 96(21): 11758-1163.
- [15] Kim S M, Kuzuyama T, Chang Y J, *et al.* Cloning and functional characterization of 2-C-methyl-D-erythritol 4-phosphate cytidyltransferase (GbMECT) gene from *Ginkgo biloba* [J]. *Phytochemistry*, 2006, 67(14): 1435-1441.
- [16] Gahlan P, Singh H R, Shankar R, *et al.* De novo sequencing and characterization of *Picrorhiza kurrooa* transcriptome at two temperatures showed major transcriptome adjustments [J]. *BMC Genomics*, 2012, 13(1): 1-21.
- [17] Singh H, Gahlan P, Kumar S. Cloning and expression analysis of ten genes associated with picrosides biosynthesis in *Picrorhiza kurrooa* [J]. *Gene*, 2013, 515(2): 320-328.
- [18] Kawoosa T, Singh H, Kumar A, *et al.* Light and temperature regulated terpene biosynthesis: hepatoprotective monoterpene picroside accumulation in *Picrorhiza kurrooa* [J]. *Justice Peace*, 2010, 10(3): 393-404.
- [19] 张慧敏, 李剑芳, 郭敏辰. 圆弧青霉脂肪酶基因序列的生物信息学分析 [J]. 食品与生物技术学报, 2010, 29(4): 602-608.
- [20] 张文红, 刘新平, 王 琰, 等. NdrG2 基因及 NDRG2 蛋白的生物信息学分析 [J]. 第四军医大学学报, 2003, 24(17): 1537-1539.
- [21] 石鸥燕, 李艳云, 蔡春泉, 等. FZ6 基因及其蛋白的生物信息学分析 [J]. 现代生物医学进展, 2010, 10(6): 1052-1054.
- [22] 侯 琳, 钱敏平, 朱云平, 等. 转录因子结合位点生物信息学研究进展 [J]. 遗传, 2009, 31: 365-373.
- [23] Martinelli F, Tonutti P. Flavonoid metabolism and gene expression in developing olive (*Olea europaea* L.) fruit [J]. *Plant Biosystems*, 2012, 146: 164-170.
- [24] Serrano I, Olmedilla A. Histochemical location of key enzyme activities involved in receptivity and self-incompatibility in the olive tree (*Olea europaea* L.) [J]. *Plant Sci*, 2012, 197: 40-49.
- [25] 刘亮伟, 秦天苍, 王 宝, 等. 木聚糖酶的分子进化 [J]. 食品与生物技术学报, 2007, 26(6): 110-116.
- [26] Gabrielsen M, Kaiser J, Rohdich F, *et al.* The crystal structure of a plant 2C-methyl-D-erythritol 4-phosphate cytidyltransferase exhibits a distinct quaternary structure compared to bacterial homologues and a possible role in feedback regulation for cytidine monophosphate [J]. *FEBS J*, 2006, 273(5): 1065-1073.