

## 原壳小球藻番茄红素 $\epsilon$ 环化酶基因的克隆和分析

李婷, 施春雷, 淦志兵, 史贤明\*

(上海交通大学农业与生物学院食品科学系, 陆伯勋食品安全研究中心, 上海 200240)

**摘要** 【目的】番茄红素  $\epsilon$  环化酶(lycopene epsilon cyclase)是叶黄素代谢途径中处于分支位点的关键酶。它催化线性的番茄红素选择性环化,形成胡萝卜素,再进一步合成叶黄素(lutein)。本研究的目标是克隆原壳小球藻(*Chlorella protothecoides* CS-41) *LCYE* 基因的全长 cDNA,通过该基因的生物学信息分析预测其表达产物的空间结构与功能位点,并验证该表达产物的生物学活性。【方法】采用 cDNA 末端快速扩增(rapid-amplification of cDNA ends)和 RT-PCR 技术克隆原壳小球藻 *LCYE* 基因的全长 cDNA 序列。分别采用 PredictProtein、Pfam HMMs 和 Swiss-Model 等在线分析软件分析 *LCYE* 的基本生物学信息,预测蛋白质功能位点和空间结构。利用 pET28-a(+) 构建 *LCYE* 基因的原核表达质粒 pET-LCYE,并转入大肠杆菌 BL21(DE3),再利用 IPTG 诱导 *LCYE* 基因超量表达。此外,携带 pAC-LYC 质粒的大肠杆菌工程菌能够积累番茄红素,可用于验证 *LCYE* 基因编码蛋白的酶活。【结果】获得了原壳小球藻的 *LCYE* 基因的 cDNA 序列,长 2107 bp,GenBank 登录号为 FJ752528。序列分析表明:它含有 1731 bp 的完整开放阅读框(ORF),编码 576 个氨基酸,与高等植物和其他藻类的 *LCYE* 有很高的相似性,其中相似性最高的是莱茵衣藻(*Chlamydomonas reinhardtii* XM001696477.1),达到 67%。第 48~459 个氨基酸残基为一个典型的番茄红素环化酶蛋白(Lycopene cyclase protein)结构域(pfam05834),第 261~284 个氨基酸残基为一个典型的环化酶保守的模体。SDS-PAGE 检测表明,*LCYE* 基因在大肠杆菌 BL21(DE3)中得到了超量表达。携带 pAC-LYC 质粒的大肠杆菌工程菌在获得 *LCYE* 基因后,其颜色从粉红色变为黄色。【结论】原壳小球藻 *LCYE* 基因的全长 cDNA 为 2107 bp,预测出番茄红素  $\epsilon$  环化酶所特有的保守功能区域,构建了它的三维结构模型。同时,阐明了小球藻和衣藻的亲缘关系。最后,证实了本研究克隆到的 *LCYE* 基因编码的蛋白具有番茄红素  $\epsilon$  环化酶的功能与活性。

**关键词:** 原壳小球藻; 叶黄素; 番茄红素  $\epsilon$  环化酶

中图分类号: Q936 文献标识码: A 文章编号: 0001-6209(2009)09-1180-10

叶黄素(lutein)是含氧类胡萝卜素即类叶黄素(xanthophylls)中的一种<sup>[1]</sup>。医学实验证明,植物中所含的天然叶黄素是一种性能优异的抗氧化剂,可预防衰老和老年性眼视网膜黄斑退化<sup>[2]</sup>。另外叶黄素也可用作癌症预防剂、生命延长剂、溃疡抵制剂、心脏病与冠状动脉疾病的抵制剂<sup>[3]</sup>。富含叶黄

素的生物主要有万寿菊、金盏花、绿藻、羽衣甘蓝和苜蓿等,其中万寿菊为目前最主要的叶黄素源<sup>[3]</sup>。由于万寿菊中的叶黄素主要以各种酯化形式存在<sup>[4]</sup>,通常还需要进行皂化或酶水解以获得游离的叶黄素产品,这是天然叶黄素产业化的瓶颈问题。异养培养的小球藻细胞中叶黄素含量可达

基金项目: 国家 863 计划(2006AA02Z226)

\* 通信作者。Tel: +86-21-34206616, E-mail: xms@sjtu.edu.cn

作者简介: 李婷(1981-),女,云南昆明人,硕士研究生,主要研究方向为食品生物技术。E-mail: jiting8848@hotmail.com

收稿日期: 2009-02-19; 修回日期: 2009-06-04

2~4 mg/g 细胞干重<sup>[5]</sup>, 与万寿菊中的叶黄素含量接近。而且小球藻具有生长迅速, 可进行异养培养等优势。用流加培养产率可达 0.94 mg/L/h<sup>[6-7]</sup>, 是目前为止效率最高的叶黄素生产方式。因此, 异养小球藻可能取代万寿菊等植物而成为天然叶黄素的生产源。目前关于小球藻叶黄素的研究中, 主要采用培养条件的优化来提高叶黄素产量<sup>[8]</sup>, 而少见采用代谢途径改造的办法来实现这一目标。

叶黄素作为一种萜类化合物, 起源于类异戊二烯途径, 在高等植物、藻类、某些真菌和细菌中广泛存在<sup>[9]</sup>, 叶绿素、维生素、GA 及 ABA 等都属于该途径的分支产物。目前类胡萝卜素合成的主要代谢途径已经基本清楚<sup>[10]</sup>(图 1), 但对于类胡萝卜素的末端修饰、甲基化以及多烯链的修饰等复杂过程仍未完全了解<sup>[11]</sup>。番茄红素  $\epsilon$  环化酶(lycopene epsilon cyclase)催化线性的番茄红素选择性环化, 形成  $\delta$ -胡萝卜素, 是叶黄素代谢途径中处于分支位点的关键

酶<sup>[12]</sup>。Cunningham<sup>[13]</sup>等最早从蓝细菌 *Synechococcus* 中克隆了番茄红素环化酶基因。根据遗传互补分析<sup>[14]</sup>显示  $\epsilon$  环化酶只在番茄红素上引入一个  $\epsilon$  环, 形成  $\delta$ -胡萝卜素; 与之不同的是,  $\beta$  环化酶能够在番茄红素上连续引入两个  $\beta$  环, 形成  $\beta$ -胡萝卜素, 或将一个  $\beta$  环引入  $\delta$ -胡萝卜素, 生成一端  $\beta$  环另一端  $\epsilon$  环的  $\alpha$ -胡萝卜素<sup>[11]</sup>。这种  $\epsilon$ -环化酶在自然界中很少再引入第二个  $\epsilon$  环, 但是在莴苣(*Lactuca sp.*)等少数几种植物却合成带两个  $\epsilon$  环的  $\epsilon$ -胡萝卜素, 再进一步形成莴苣黄素(*Lactucaxanthin*)<sup>[15]</sup>。

本研究力图从原壳小球藻中克隆叶黄素代谢的关键分支点上的关键酶 LCYE 的编码基因, 并进行生物信息学分析, 以发掘特征性的结构域, 预测蛋白质空间结构, 并验证该基因表达产物的生物学活性。同时, 通过构建进化树, 探讨原壳小球藻 LCYE 基因的分子进化地位。

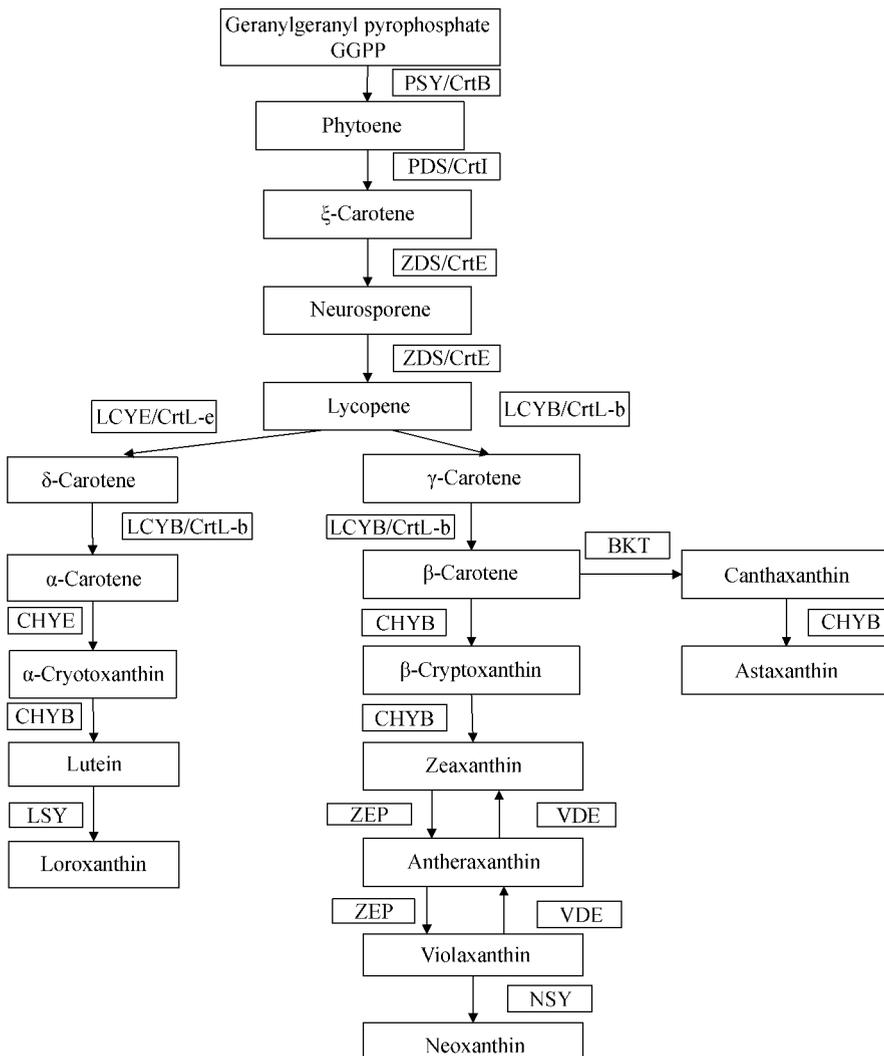


图 1 类胡萝卜素生物合成途径

Fig. 1 Metabolic pathways of carotenoid biosynthesis.

# 1 材料和方法

## 1.1 材料

**1.1.1 主要试剂和仪器:** SuperScript III reverse transcriptase 和 Trizol 购于美国 Invitrogen 公司; TaKaRa pMD18-T Vector Kit、TaKaRa One-Step RT-PCR Kit、3'-Full RACE Core Set Ver.2.0、5'-Full RACE Kit 和限制性内切酶购于日本 TaKaRa 公司; QIAprep Miniprep Kit 购于美国 QIAGEN 公司; AxyPrep™ DNA Gel Extraction Kit 和 AxyPrep™ Plasmid Extraction Kit 购于美国 Axygen 公司; GeneAmp® 9700 型 PCR 仪购于美国 ABI 公司。

**1.1.2 藻种和培养条件:** 原壳小球藻 (*Chlorella protothecoides* CS-41) 购自澳大利亚 CSIRO 微藻研究中心。原壳小球藻的培养采用改良式 Basal 培养基<sup>[5]</sup> 培养条件为: 28℃, 光照强度 30 μmol/(m<sup>2</sup>·s), 转速 180 r/min 培养时间 5 d。因为 LCYE 基因的表达丰度低, 为了提高其丰度, 需加用强光诱导。提取 RNA 前, 将葡萄糖浓度提高到 20 g/L, 将光照强度提高到 90 μmol/(m<sup>2</sup>·s), 诱导 5 h。

## 1.2 RNA 提取

取 15 mL 藻液, 离心收集沉淀后用液氮研磨, 采用美国 Invitrogen 公司的 Trizol 试剂盒抽提 RNA, 并溶于 30 μL 0.1% DEPC 的无菌去离子水中, -70℃ 保存备用。采用紫外分光光度计测定 OD 值来表示

RNA 的纯度,  $OD_{260/280}$  值达到 1.84 ~ 2.10 方可满足纯度要求。

## 1.3 cDNA 合成

采用美国 Invitrogen 公司的 Superscript III reverse transcriptase 试剂盒合成 cDNA, 其反应体积为 20 μL, 包括 2 μL RNA, 1 μL 10 mM dNTP Mix, 1 μL OligodT primers, 加 H<sub>2</sub>O (RNase free) 至 13 μL 后, 65℃ 条件下孵育 5 min 后冰浴 2 min, 后加入 4 μL 5 × First-Strand Buffer, 1 μL 0.1 mM DTT, 1 μL RNase Inhibitor, 0.5 μL Superscript III reverse transcriptase, 加 H<sub>2</sub>O (RNase free) 至 20 μL 体系, 混匀, 瞬时离心, 50℃ 孵育 60 min 后, 70℃ 热激 1 min, 后迅速冰浴 2 min, 于 -20℃ 保存。

## 1.4 引物设计和合成

采用 ClustalX 软件对 GenBank 中来源于莱茵衣藻 (*Chlamydomonas reinhardtii*, XM001696477.1) 拟南芥 (*Arabidopsis thaliana*, NM125085.3) 蜜橘 (*Citrus unshiu*, AB238230.1) 黄龙胆 (*Gentiana lutea*, AB017373.1) 胡萝卜 (*Daucus carota*, DQ192192.1) 等生物的 LCYE 基因序列进行比对, 找到氨基酸保守区 EEEWSYIPVGG 和 WXGFLG(A)SXLSS 为模板, 利用 Primer 5.0 和 Oligo 6.0 在保守区段设计简并引物扩增核心片段。然后根据核心序列设计 3'-RACE 和 5'-RACE 特异引物, 如表 1 所示。引物均由上海生工生物工程技术有限公司合成。

表 1 原壳小球藻的 LCYE 引物序列及 PCR 扩增得到目的片段大小

Table 1 PCR primers and amplicon size for *C. protothecoides* LCYE

Primers	Primer sequence (5'→3')	Product size /bp
Core fragment		
LF	CARACWGCWTATGGTGTGGA	368
LR	AGWSWGAIGGAAGRAAWCC	
LFS	GAAGAGGAATGGTCHTAYATHCC	
LRS	CGGAAGAAGWTRYGRAARAA	
5'-RACE		
FI	CGCGGATCCACAGCCTACTGATGATCAGTCGATG	
FO	GATGGCTACATGCTGACAGCCTA	
LFO1	GCGAGTGATGCTGTACCCCGTAGCA	
LFI1	CGCCCACAGGGATGTAGCTCCACTC	939
LF13	CGATTCCGTATGCTGTTTGTGCC	987
3'-RACE		
TI	CGCGGATCCTCCACTAGTGATTTCACTATAGG	861
TO	TACCGTCGTTCCACTAGTGATTT	
LTI	CTGACTTCTTACCACCTTCTTCCG	
LTO	GGCGCCAGACCTCTTCCAAGTGTT	
Prokaryotic expression		
CPL1	CCC <sup>gatcc</sup> ATGCTGCAGGGCGATGCGA ( BamH I )	1467
CPL2	CCC <sup>gaattc</sup> CTCCACCTCCCATTTGCCGC ( EcoR I )	

## 1.5 RACE 法克隆

**1.5.1 3'-RACE 法:**来自原壳小球藻的 RNA 被用于构建 3'-RACE cDNA 文库,该文库根据日本 TaKaRa 公司 3'-Full RACE Core Set Ver.2.0 试剂盒的操作程序构建。构建好的 3'-RACE cDNA 作为模板,以 LTO/TO(由 3'-Full RACE Core Set Ver.2.0 试剂盒提供)为引物进行第一轮 PCR。以第一轮 PCR 的产物为模板,以 LT1/TI(Nested Primer,由 3'-Full RACE Core Set Ver.2.0 试剂盒提供)为引物进行巢式 PCR。两次 PCR 的程序均为 94℃ 预变性 3 min;94℃ 变性 40 s,65℃ 退火 30 s,72℃ 延伸 1.5 min,32 个循环,72℃ 延伸 10 min。RACE 产物经琼脂糖凝胶电泳对 PCR 产物进行检测,以 200 bp DNA Ladder 为分子量参照物。

**1.5.2 5'-RACE 法:**来自原壳小球藻的 RNA 被用于构建 5'-RACE cDNA 文库,该文库根据日本 TaKaRa 公司 5'-Full RACE Kit 的操作程序构建。构建好的 5'-RACE cDNA 作为模板,以 LFO 和 FO(由 5'-Full RACE Kit 提供)为引物进行第一轮 PCR。以第一轮 PCR 的产物为模板,以 LF1/FK(Nested Primer,由 5'-Full RACE Kit 提供)为引物进行巢式 PCR。两次 PCR 的程序 94℃ 预变性 3 min,94℃ 变性 40 sec,65℃ 退火 30 sec,72℃ 延伸 1.5 min,32 个循环,72℃ 延伸 10 min。由于一次 5'-RACE 没有得到完整的 5'端序列,所以在已经得到的序列基础上再设计引物进行扩增,第一轮扩增同上,第二轮以第一轮产物为模板,LF2/FK(Nested Primer,由 5'-Full RACE Kit 试剂盒提供)为引物进行巢式 PCR。RACE 产物经琼脂糖凝胶电泳对 PCR 产物进行检测,以 200 bp DNA Ladder 为分子量参照物。

## 1.6 序列测定

扩增产物经琼脂糖凝胶电泳,切下目的片段,经美国 Axygen 公司的 AxyPrep™ DNA Gel Extraction Kit 纯化后,在 T4 连接酶的作用下,于 4℃ 过夜,连接于 pMD19-T 质粒载体上。将连接产物转化大肠杆菌感受态细胞,于 LB 琼脂平板(含 Amp 100  $\mu$ g/mL)过夜培养,以 M13 通用引物进行菌落 PCR 鉴定,选出阳性克隆转化子,于 LB 液体培养基(含 Amp 100  $\mu$ g/mL)中 37℃ 过夜培养,用 QIAprep Miniprep Kit 提取质粒 DNA,用于序列测定。所有的测序均由上海英骏生物技术有限公司完成,测序仪为 ABI3730xl 96-capillary DNA analyzer,测序引物为 M13 通用引物。

## 1.7 序列比对分析

用 GenBank 的 Blastn 和 Blastp 在线软件分别查

找与 LCYE 基因同源的核酸序列和对应的氨基酸序列,用 ClustalX 软件进行比对分析;用 PredictProtein (<http://www.predictprotein.org/>) ScanProsite (<http://www.expasy.ch/tools/scanprosite/>)和 Pfam (<http://pfam.sanger.ac.uk/search>)在线软件分析核酸序列的基本信息并进行结构域的预测;用 Swiss-Model (<http://swissmodel.expasy.org/workspace/>)在线软件预测蛋白质二级结构和三级结构;用 MEGA4.1 软件构建进化树。

## 1.8 LCYE 基因的原核表达及表达产物的鉴定

以 LCYE 基因为模板,设计其原核表达引物 CPL1 和 CPL2(表 1),下划线分别为 BamHI 和 EcoRI 的识别序列。以原壳小球藻的第一链 cDNA 为模板进行扩增,将该产物命名为 D-LCYE。将 D-LCYE 经纯化回收后,导入克隆载体 pMD19-T 中,提取阳性克隆质粒 pMD-D-LCYE,同时用 BamHI 和 EcoRI 分别双酶切 pMD-D-LCYE 及表达载体 pET28- $\alpha$ (+)并将其完全酶切的产物回收纯化后在 T<sub>4</sub> DNA 连接酶的作用下进行酶连,酶连产物转化大肠杆菌 BL21(DE3)菌株。涂布 LB 琼脂平板(含 Kan 50  $\mu$ g/mL)37℃ 倒置过夜培养后挑取单菌落,少量制备质粒,经 BamHI 和 EcoRI 双酶切鉴定后,将获得的阳性表达质粒命名为 pET-LCYE。将携带表达质粒的 BL21(DE3)菌株接种于含 Kan 50  $\mu$ g/mL 的 LB 液体培养基中 37℃ 振荡培养,当 OD<sub>600</sub> 为 0.4~0.6 时加入 IPTG 至终浓度为 1 mM 进行诱导表达,诱导后的菌液离心,加上样缓冲液 100℃ 煮沸 10 min,离心取上清液进行 SDS-PAGE 电泳分析以检测目的蛋白的表达。

## 1.9 原壳小球藻 LCYE 酶的功能验证

将 pET-LCYE 转化到带有质粒 pAC-LYC<sup>[13]</sup>(由美国马里兰大学的 Elisabeth Gantt 教授馈赠)的 BL21(DE3)感受态细胞中。pAC-LYC 含有细菌来源的类胡萝卜素代谢途径中的 3 个基因  *crtE/crtI/crtB* (图 2),在大肠杆菌中能积累 LCYE 的催化底物番茄红素,在没有 LCYE 引入时菌体呈粉红色,一旦引入 LCYE 后菌体会由粉红色变为黄色,可以利用颜色改变效应检测和鉴定 LCYE 的环化酶活性。质粒 pET-LCYE 转化到 BL21(DE3)后挑单菌落于含 50  $\mu$ g/mL Kan 和 50  $\mu$ g/mL Cm 的 LB 培养基中 37℃ 振荡培养,当 OD<sub>600</sub> 值为 0.4~0.6 时加入终浓度为 1 mmol/L 的 IPTG 诱导,诱导后的菌液离心,观察颜色变化,以空载体 pET-28 $\alpha$ (+)转化 BL21(DE3)的菌体作为阴性对照。

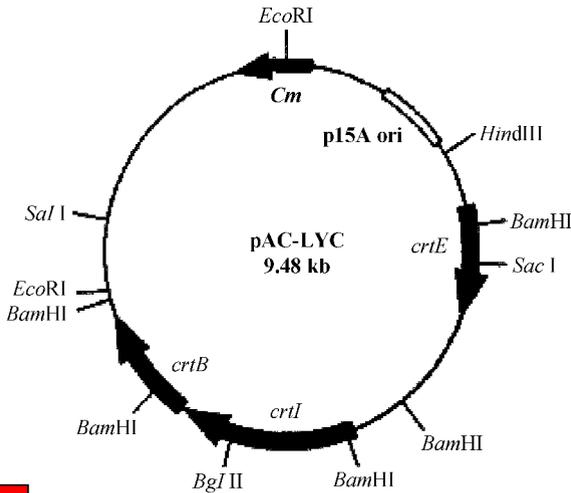


图2 质粒 pAC-LYC 的结构<sup>[13]</sup>

Fig.2 Structure of the plasmid pAC-LYC.

## 2 结果和分析

### 2.1 小球藻总 RNA 的制备

采用 Trizol 试剂盒,获得的小球藻总 RNA,其  $OD_{260}/OD_{280}$  大于 1.9;经过电泳检测,发现其纯度高、完整性较好。

### 2.2 LCYE 核心片段的克隆

根据包括 *C. reinhardtii* (XM001696477.1) 及 *A. thaliana* (NM125085.3) 及 *C. unshiu* (AB238230.1) 及 *G. lutea* (AB017373.1) 及 *D. carota* (DQ192192.1) 及 *Lelyocyc* (Y14387.1) 等生物 LCYE 基因核酸序列的比对结果,找出保守区段,设计 2 对简并引物(LF/LR, LFS/LFR)进行巢式 PCR(表 1),经反转录 PCR(reverse transcriptase PCR, RT-PCR)扩增,得到 368 bp 的产物(图 2)。将其在 GenBank 上进行 Blastn 比对后,发现所得序列和莱茵衣藻(XM\_001696477.1) LCYE 基因的 cDNA 序列中的一段有 76% 的同源性,推测该 368 bp 的核酸序列是 LCYE 基因的核心片段。

### 2.3 LCYE 高保真 PCR 扩增产物

根据核心片段设计特异引物进行 3'-RACE 和 5'-RACE。3'-RACE 的扩增巢式引物为 TO/LTO 和 TI/LTI,产物经测序后,片段大小为 861 bp;第一次 5'-RACE 的扩增巢式引物为 FI/FI1 和 FO/FO1,产物经测序后,片段大小为 939 bp;第二次 5'-RACE 的扩增巢式引物为 FI/FI3 和 FO/FO1,产物片段大小为 987 bp(图 3)。

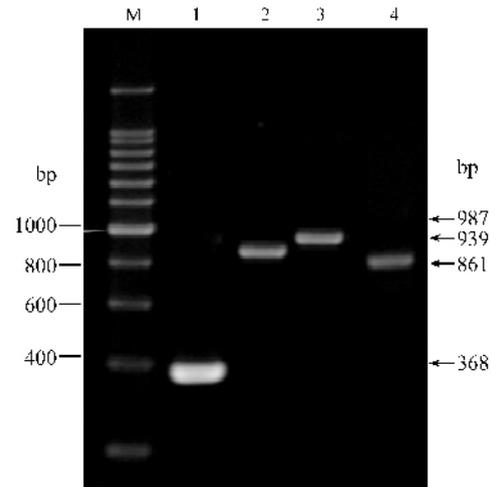


图3 LCYE 基因高保真 PCR 扩增产物的电泳图

Fig.3 The electrophoresis of high fidelity PCR products of the LCYE gene. 1. Core fragment; 2.5'-RACE product with primers FI/FI1 + FO/FO1; 3.5'-RACE product with primers FI/FI3 + FO/FO1; 4.3'-RACE product with primers TO/LTO + TI/LTI; M, 200 bp DNA ladder.

### 2.4 原壳小球藻的 LCYE 基因全长 cDNA 的序列分析

核心片段和 RACE 产物拼接后得到一条长度为 2107 bp 的 cDNA 序列。利用 GenBank 的 ORF Finder 软件进行开放阅读框(ORF)判定,从第 140 位碱基到 1870 位碱基为完整的阅读框,阅读框全长为 1731 bp, G + C 含量为 62.62%, A + T 含量为 37.38%, 编码 576 个氨基酸,等电点为 5.84, 分子量为 62.9986 kDa, 编码区的起始密码子位于第 140 bp, 终止密码子位于第 1870 bp。5' 端的非翻译区为 1~140 bp, 3' 端的非翻译区为 1871~2107 bp, TGTA<sup>[16]</sup>这段序列被认为是藻类中多聚腺苷酸(PolyA)信号,出现在 2043~2048 bp 之间。将 2107 bp 的核苷酸测序结果用 Blastn 软件在 GenBank 中检索,发现这一 cDNA 序列的 274 bp~1551 bp 这段和莱茵衣藻(XM\_001696477.1) LCYE 基因的 cDNA 序列的 599~1876 bp 之间高达 67% 的相似性,推断此序列为原壳小球藻 LCYE 基因的 cDNA 序列,将这一序列登录至 GenBank,获得登录号为 FJ752528。

### 2.5 LCYE 酶的功能位点预测和进化树分析

原壳小球藻 LCYE 的氨基酸序列与 *C. reinhardtii*、*A. thaliana*、*C. unshiu*、*G. lutea*、*D. carota* 等的 LCYE 分别有 60%、43%、44%、43% 和 44% 的相似性。根据 PredictProtein 数据库推测 LCYE 氨基酸序列的功能位点,发现了两个 cAMP-和 cGMP-磷酸化激酶蛋白激酶位点(cAMP- and cGMP-

dependent protein kinase phosphorylation site)RRQT 和 RRAS;四个蛋白激酶磷酸化位点(Protein kinase C phosphate site)SSK、SLR、TLK、SLR;九个 CK2 磷酸化位点(Protein kinase CK2 site)SKQD、STAD、SYLE、TEGE、SLRE、SSVE、TLDE、SWAE、TSFE;一个酪氨酸激酶磷酸化位点(Tyrosine kinase phosphorylation site)KVHEEWSY;十个 N 端酰基化位点(N-myristoylation site)GAGPAG、GLALAA、GLEHTL、GQTAGK、GLYDGT、GLWGTE、GITVTK、GAAASL、GGAMAG、GAMAGD。

采用 ClustalX 软件比对已经报道的 *LCYE* 基因的核酸序列 21 个,然后以来自 *Prochlorococcus*

*marinus* 的 *LCYE* 基因作为外类群,用 MEGA4.1 构建 *LCYE* 基因的 UPGMA 进化树(图 4),采用 1000 重复的自展检验来完成。从图 4 中可以看出,*LCYE* 基因截然分为了 3 个类群:高等植物、藻类以及作为外类群的 *P. marinus*。原壳小球藻的 *LCYE* 基因最先与莱茵衣藻聚类合并,最后才与高等植物聚类合并,这个结果与 *LCYE* 的核苷酸与氨基酸序列同源比对所得结果基本一致,表明不同物种的 *LCYE* 基因进化速率存在差异。在藻类中小球藻和衣藻的亲缘关系比较近,它们的 *LCYE* 基因有可能起源于一个共同的基因。

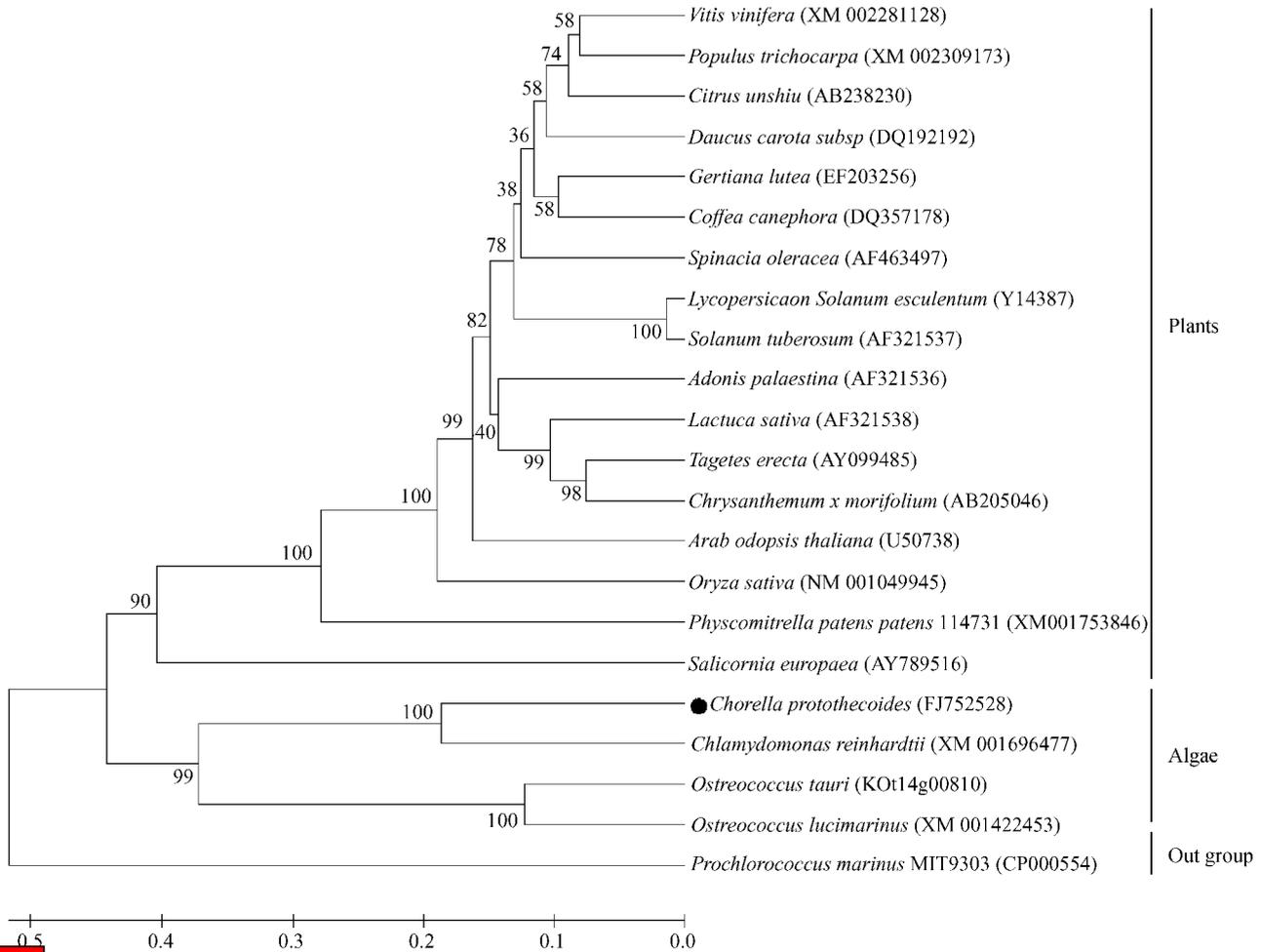


图 4 基于 UPGMA 构建的 *LCYE* 基因系统发育树

Fig.4 Phylogenetic tree of *LCYE* gene based on its nucleotide sequence using UPGMA method. Numbers in parentheses represent the sequences' accession number in GenBank. The number at each branch point is the percentage supported by bootstrap. Bar ,0.10% sequence divergence.

## 2.6 蛋白质序列多重比对和功能结构域确定

利用原壳小球藻 *LCYE* 氨基酸序列作参照,在 GenBank 中查找到了 21 个同源序列,用 ClustalX1.83 软件进行序列多重比对分析。这些同源序列分别来源于 *A. thaliana* (NM125085.3)、*C. reinhardtii* (XM001696477.1)、*C. unshiu* (AB238230.1)、*D.*

*carota* (DQ192192.1)、*V. vinifera* (gbAAS48096.1)、*S. oleracea* (AAL69394)等。根据 Pfam 的 HMMs 数据库分析可能的结构域,并对比已报道的番茄红素环化酶的保守区域<sup>[13,17]</sup>,显示在原壳小球藻 *LCYE* 酶中,存在以下结构域(图 5):靠近 N 端,有一段高度保守的模体,其序列是 DX4GXGXAX4A,这是双核苷酸结



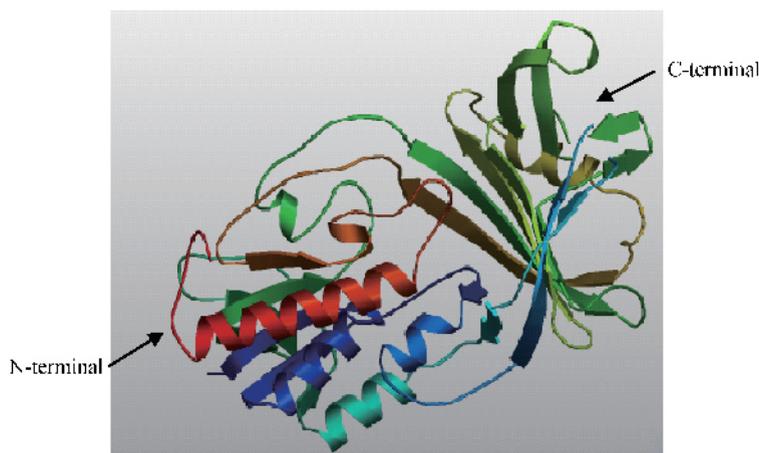


图6 原壳小球藻 LCYE 蛋白的三级结构预测图

Fig.6 The protein predicted 3D structure of the LCYE of *C. protothecoides* CS-41.

## 2.7 LCYE 基因的原核表达及其产物的鉴定

将 *LCYE* 基因的结构域编码序列 D-LCYE 导入原核表达载体 pET-28a(+), 构建原核表达质粒 pET-LCYE, 经测序验证, 表达质粒中 D-LCYE 序列正确, 表明原核表达载体 pET28-LCYE 构建成功。表达质粒 pET-LCYE 转化到宿主菌 BL21(DE3) 后, 经 37℃ 振荡培养至  $OD_{600}$  值为 0.5 左右, 加入 IPTG 分别诱导 1h、3h、6h, 收集菌体加上样缓冲液煮沸, 经 SDS-PAGE 电泳, 如图 7 所示, 诱导后的胞内蛋白在 53 kDa 左右出现一条超量表达的蛋白条带, 与预期的 LCYE 蛋白大小(约 53 kDa)一致, 而空白载体和未诱导的菌体未发现这一大小的条带, 以上结果表明, 原壳小球藻的 *LCYE* 基因在大肠杆菌中得到了超量表达。

## 2.8 原壳小球藻 LCYE 酶的功能验证

将 pET-LCYE 转化到带有质粒 pAC-LYC 的 BL21(DE3) 中, 如图 8 所示, 1 号菌落为携带 pET-28a(+ ) 空载体的阴性对照, 菌体颜色为粉红色; 3 号为带有 pET-LCYE 的质粒未加 IPTG 诱导的结果, 颜色已经有些许偏黄; 2 号为带有 pET-LCYE 的质粒加 IPTG 诱导的结果, 颜色为黄色。通过这个酶活试验, 表明 *LCYE* 基因的表达产物能够将工程菌中积

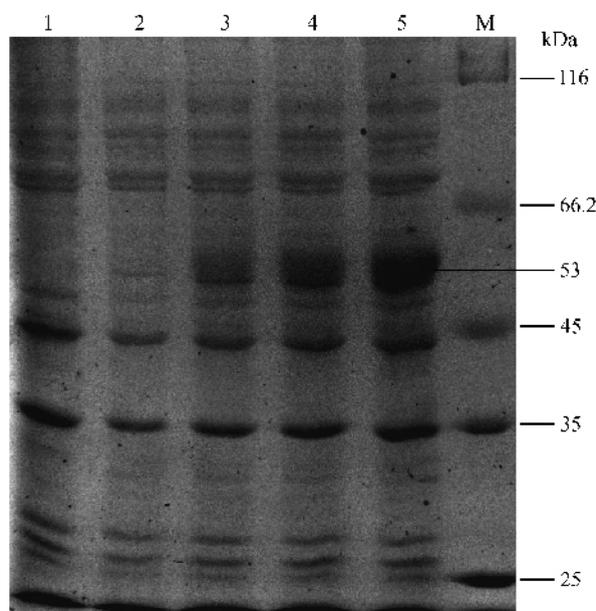
图7 SDS-PAGE 电泳检测在 *E. coli* 中表达的重组蛋白

Fig.7 SDS-PAGE profile of recombinant proteins in *E. coli*. Lane 1, pET-28a(+); lane 2, pET-LCYE (-IPTG); lane 3-5, pET-LCYE (+IPTG) for 1 h, 3 h and 6 h respectively; M, Mid-range protein Marker.

累的番茄红素(粉红色)转化成  $\delta$ -胡萝卜素(黄色), 证明了 this 蛋白具有番茄红素  $\epsilon$  环化酶的功能与

图8 原壳小球藻 *LCYE* 基因表达产物的酶活验证

Fig.8 The enzyme activity assay of the expression product of *LCYE* gene from *C. protothecoides*.

1 pAC-LYC + pET-28a(+ ) (+ IPTG) 2 pAC-LYC + pET-LCYE(+ IPTG) 3 pAC-LYC + pET-LCYE(-IPTG).

活性。

### 3 讨论

番茄红素环化酶基因是一个大家族,不同 *LCYE* 基因家族成员编码的氨基酸序列差异较大,但仍有一些高度保守的区域存在,这些高度保守的残基与催化功能有关。原壳小球藻的 *LCYE* 基因对应的氨基酸序列在中、后部与其它植物 *LCYE* 的氨基酸序列同源性很高,且包含 *LCYE* 所共有的“FLYA(V)MP”序列和双核苷酸结合位点(Dinucleotides-binding site)。此外,原壳小球藻的 *LCYE* 氨基酸序列与衣藻的同源性明显高于其他植物,与进化分析的结果相吻合。

*LCYE* 三级结构预测发现,靠近 N 端有一个  $\beta$  折叠- $\alpha$  螺旋- $\beta$  折叠构型,这与保守功能域分析相符合。但从整体构型来说,其  $\beta$  折叠和  $\alpha$  螺旋分别集中在两个区域中,这两个区域在空间上分别向相反的方向延伸,预示着它们在功能上有着明确的分工,但是各自在酶催化过程中行使哪种功能尚未可知,还需整合更多有关番茄红素环化酶的进一步研究结果才能定论。另外,该酶的底物结合位点也需要加以进一步的研究。

*LCYE* 和 *LCYB* 功能分析显示,*LCYB* 可在番茄红素分子的两端各形成一个  $\beta$  环,而 *LCYE* 只能在番茄红素分子的一端形成一个  $\epsilon$  环。此环形成后,将在某种程度上限制或抑制 *LCYE* 在另一端再形成一个  $\epsilon$  环。莴苣的 *LCYE* 氨基酸序列与拟南芥 *LCYE* 的氨基酸序列一致性高达 77%,这一酶在大肠杆菌杂合系统中可以将番茄红素有效地转化为双  $\epsilon$  环胡萝卜素<sup>[15]</sup>。进一步研究表明,莴苣 *LCYE* 蛋白第 457 位的 H 残基决定着 *LCYE* 催化番茄红素的两端形成  $\epsilon$  环,而拟南芥 *LCYE* 蛋白第 448 位的 L 残基决定了 *LCYE* 只能催化番茄红素的一端形成  $\epsilon$  环。对原壳小球藻 *LCYE* 序列分析表明,其对应的第 405 位有一个 L 残基,有可能也决定了在原壳小球藻中没有双  $\epsilon$  环胡萝卜素。

### 4 结论

采用 RACE 和 RT-PCR 方法,克隆得到的原壳小球藻 *LCYE* 基因的全长 cDNA 为 2107 bp,具有 1731 bp 的完整开放阅读框。预测氨基酸序列中第 48~459 个氨基酸残基为一个典型的番茄红素环化酶蛋白结构域(pfam05834),具有环化酶特有的保守功能区域,即双核苷酸结合位点、环化酶模体(Cyclase motif1)以及预测跨膜区。另外还发现一段高保守的区域,其模体结构是 E-X-E-X-P[LMV]-

GN ]X[ PST ]H ILP ]P-X-Q-X[ AG ]X-G[ AG ] [AST ]A[ AGNS ]H LM ]V-H-P[ AS ]H ST ]G[ FY ]X[ ILV ]H AGSTV ]X[ FIL ]X-A-P-X-A。进化树分析结果表明,在藻类中小球藻和衣藻的亲缘关系比较近,它们拥有的 *LCYE* 基因可能来自一个共同的基因。利用 pET28-a(+)构建的原壳小球藻番茄红素  $\epsilon$  环化酶基因的原核表达质粒 pET-*LCYE* 转入大肠杆菌 BL21(DE3),经 IPTG 诱导成功表达。在积累番茄红素的大肠杆菌工程菌中导入 *LCYE* 基因,酶活试验证实了本研究克隆到的 *LCYE* 基因编码的蛋白具有番茄红素  $\epsilon$  环化酶的功能。

致谢 真诚感谢美国马里兰大学 Elisabeth Gantt 教授馈赠质粒 pAC-LYC,为 *LCYE* 的酶活验证提供了很大的帮助。衷心感谢香港大学 Chen Feng 教授和 Huang Jun Chao 博士在实验方法上的指导和讨论。

### 参考文献

- [1] Pfander H. Carotenoids: an overview. *Methods in Enzymology* 1992, 21: 33-35.
- [2] Alves-Rodrigues A, Shao A. The science behind lutein. *Toxicology Letters* 2004, 150(1): 57-83.
- [3] Berendschot TT, Goldbohm RA, Klöpping WA, et al. Influence of lutein supplementation on macular pigment, assessed with two objective techniques. *Investigative Ophthalmology and Visual Science* 2000, 41(11): 3322-3326.
- [4] Piccaglia R, Marotti M, Grandi S. Lutein and lutein ester content in different types of *Tagetes patula* and *T. erecta*. *Industrial Crops and Products* 1998, 8(1): 45-51.
- [5] Shi XM, Chen F. High yield production of lutein by heterotrophic *Chlorella protothecoides* in fed-batch systems. *Algae and Their Biotechnological Potential* 2001, 107-109.
- [6] Shi X-M, Chen F, Yuan J-P, et al. Heterotrophic production of lutein by selected *Chlorella* strains. *Journal of Applied Phycology* 1997, 9(5): 445-450.
- [7] Zhang XW, Shi XM, Chen F. A kinetic model for lutein production by the green microalga *Chlorella protothecoides* in heterotrophic culture. *Journal of Industrial Microbiology & Biotechnology* 1999, 23(6): 503-507.
- [8] Shi XM, Wu ZY, Chen F. Kinetic modeling of lutein production by heterotrophic *Chlorella* at various pH and temperatures. *Molecular Nutrition & Food Research* 2006, 50(8): 763-768.
- [9] Cunningham FX, Gantt E. Genes and enzymes of carotenoid biosynthesis in plants. *Annual Reviews in Plant Physiology and Plant Molecular Biology* 1998, 49: 557-583.
- [10] Lohr M, Grossman AR. Genome-Based Examination of Chlorophyll and Carotenoid Biosynthesis in *Chlamydomonas reinhardtii*. *Plant Physiology* 2005, 138(1): 490-515.

- [ 11 ] Cunningham FX ,Pogson B ,Sun Z ,et al. Functional Analysis of the beta and epsilon Lycopene Cyclase Enzymes of *Arabidopsis* Reveals a Mechanism for Control of Cyclic Carotenoid Formation. *Plant Cell* ,1996 ,8( 9 ) :1613 - 1626.
- [ 12 ] Dharmapuri S , Rosati C , Pallara P , et al. Metabolic engineering of xanthophyll content in tomato fruits. *FEBS Letters* ,2002 ,519( 3 ) :30 - 40.
- [ 13 ] Cunningham FX ,Sun Z ,Chamovitz D , et al. Molecular structure and enzymatic function of lycopene cyclase from the cyanobacterium *Synechococcus* sp strain PCC7942. *Plant Cell* ,1994 ,6( 8 ) :1107 - 1121.
- [ 14 ] Li T ,Magallanes-Lundback M ,Musetti V ,et al. Functional analysis of beta and epsilon ring carotenoid hydroxylases in *Arabidopsis*. *Plant Cell* ,2003 ,15( 6 ) :1320 - 1332.
- [ 15 ] Cunningham FX Jr ,Gantt E. One ring or two ?Determination of ring number in carotenoids by lycopene epsilon-cyclases. *Proc Natl Acad Sci U S A* ,2001 ,98( 5 ) :2905 - 2910.
- [ 16 ] Huang JC ,Wang Y ,Sandmann G , et al. Isolation and characterization of a carotenoid oxygenase gene from *Chlorella zofingiensis*. *Applied Microbiology and Biotechnology* ,2006 ,71 :473 - 479
- [ 17 ] Huguene P ,Badillo A ,Chen HC , et al. Metabolism of cyclic carotenoids : a model for the alteration of this biosynthetic pathway in *Capsicum annuum* chromoplasts. *The Plant Journal* ,1995 ,8( 3 ) :417 - 424.
- [ 18 ] Lee PC ,Schmidt-Dannert C. Metabolic engineering towards biotechnological production of carotenoids in microorganisms. *Applied Microbiology and Biotechnology* ,2002 ,60 :1 - 11.
- [ 19 ] Diretto G , Tavazza R , Welsch R , et al. Metabolic engineering of potato tuber carotenoids through tuber-specific silencing of lycopene epsilon cyclase. *BMC Plant Biology* ,2006 ,6 :13.
- [ 20 ] Phillip DM ,Young AJ. Preferential inhibition of the lycopene epsilon-cyclase by the substituted triethylamine compound MPTA in higher plants. *The Journal of Plant Physiology* ,2006 ,163( 4 ) :383 - 391.

## Cloning and analysis of the gene encoding lycopene epsilon cyclase in *Chlorella protothecoides* CS-41

Ting Li ,Chunlei Shi ,Zhibing Gan ,Xianming Shi \*

( Department of Food Science & Technology and Bor Luh Food Safety Center ,School of Agriculture and Biology ,Shanghai Jiaotong University ,Shanghai 200240 ,China )

**Abstract [ Objective ]** Lycopene epsilon cyclase ( LCYE ) is the key enzyme in the lutein synthesis pathway and catalyses linear lycopene to form cyclic  $\epsilon$ -carotene ,a precursor of lutein. We aimed to clone the full-length cDNA of *LCYE* gene from *Chlorella protothecoides* CS-41 ,to predict the functional sites and the three-dimensional structure of LCYE through bioinformatics analysis and to confirm its activities and functions. **[ Method ]** We used RACE( rapid-amplification of cDNA ends ) essay and RT-PCR for the cloning of the full-length cDNA of *LCYE* from *C. protothecoides* CS-41. The online software such as PredictProtein ,Pfam HMMs and Swiss-Model were used in bioinformatics analysis of the amino acid sequence of LCYE protein. We constructed the expression vector for *LCYE* gene with pET-28a( + ) and transformed into *Escherichia coli* BL21( DE3 ). Furthermore ,the *E. coli* strain containing the pAC-LYC plasmid which could accumulate lycopene was used for the functional confirmation of LCYE from *C. protothecoides* CS-41. **[ Results ]** A 2107 bp cDNA( GenBank Accession No. FJ752528 )sequence was cloned with 1731 bp open reading frame ,encoding a putative LCYE from *C. protothecoides* CS-41. Homology studies showed that the deduced amino acid sequence of *LCYE* gene had a significant similarity with the corresponding sequences of other green algae and higher plants. It shared the highest sequence identity up to 67% ,with the *LCYE* gene from *Chlamydomonas reinhardtii*. One typical lycopene cyclase protein domain ( Pfam05834 ) was predicted between the 48<sup>th</sup>-459<sup>th</sup> amino acid. In addition ,the sequence between 261<sup>th</sup>-284<sup>th</sup> was one typical conserved lycopene cyclase protein motif. The SDS-PAGE result showed that the *LCYE* gene was over-expressed in *Escherichia coli* BL21( DE3 ) after the addition of IPTG. The prokaryotically expressed LCYE protein was able to transfer the color of the *E. coli* strain containing the pAC-LYC plasmid from pink to yellow. **[ Conclusion ]** The full-length cDNA sequence of *LCYE* gene was successfully cloned with the size of 2107 bp. Several typical motifs were found and the three-dimensional structure of LCYE was constructed from Bioinformatics analysis. The generated phylogenetic tree showed the closest relationship between *C. protothecoides* CS-41 and *C. reinhardtii* among the listed organisms. Finally ,the expression product of the *LCYE* gene cloned in the study was confirmed to hold the function and activity of lycopene epsilon cyclase.

**Keywords :** *Chlorella protothecoides* CS-41 ; Lutein ; LCYE

( 本文责编 :王晋芳 )