

## 条斑紫菜半胱氨酸合成酶基因的电子克隆与分析

易乐飞<sup>1</sup>, 王萍<sup>1</sup>, 周向红<sup>1</sup>, 刘楚吾<sup>2</sup>

(1. 淮海工学院 江苏省海洋生物技术重点实验室, 江苏 连云港 222005; 2. 广东海洋大学 水产学院, 广东 湛江 524088)

**摘要:** 从网络共享的条斑紫菜 (*Porphyra yezoensis* Ueda) 表达序列标签 (expressed sequence tag, EST) 数据库检索出与拟南芥 (*Arabidopsis thaliana*) 半胱氨酸合成酶 (cysteine synthase, CSase) 基因高度相似的EST序列, 构建EST叠连群 (contig), 并依据 contig 设计引物。提取条斑紫菜叶状体总 RNA, 采用 RT-PCR 方法, 扩增得到了条斑紫菜半胱氨酸合成酶基因 (*PyCSase-B*) 的 cDNA 序列 (GenBank accession: FJ232911)。该 cDNA 序列含有长 1 152 nt 的完整开放阅读框, 编码产物 (*PyCSase-B*) 的分子量 39.3 kD, 长 383 AA。PyCSase-B 的 N 端前 60 个氨基酸残基构成了叶绿体转运肽序列, 成熟的 PyCSase-B 定位在叶绿体中。PyCSase-B 与紫红紫菜 (*P. purpurea*)、莱茵衣藻 (*Chlamydomonas reinhardtii*)、拟南芥和水稻 (*Oryza sativa*) 的序列一致性分别为 94%、69%、64% 和 65%, 进化树显示 PyCSase-B 在进化上处在细菌和高等植物之间。PyCSase-B 含有与 5'-磷酸吡哆醛 (Pyridoxal 5'-phosphate) 和 O-乙酰丝氨酸 (O-acetylserine) 结合的保守氨基酸残基和序列, 含有将硫转移到半胱氨酸的保守残基; 成熟 PyCSase-B 单体的三维结构与拟南芥 CSase 相似, 由 2 个  $\alpha/\beta$  结构域组成。PyCSase-B 的成功克隆与生物信息学分析有助于进一步研究其在条斑紫菜抗逆调控中的作用。[中国水产科学, 2009, 16(4): 518-524]

**关键词:** 条斑紫菜; 半胱氨酸合成酶; 电子克隆; 生物信息学

**中图分类号:** Q94

**文献标识码:** A

**文章编号:** 1005-8737-(2009)04-0518-07

在植物和微生物细胞中催化丝氨酸合成半胱氨酸需要两步反应。首先在丝氨酸乙酰转移酶 (Serine acetyltransferase, EC 2.3.1.30) 的作用下, 丝氨酸与乙酰辅酶 A (acetyl-CoA) 反应生成 O-乙酰丝氨酸 (O-acetylserine, OAS); 然后半胱氨酸合成酶 (cysteine synthase, CSase, EC 4.2.99.8) 以 5'-磷酸吡哆醛 (Pyridoxal 5'-phosphate, PLP) 作为辅助因子, 催化 OAS 和硫化物反应生成半胱氨酸<sup>[1-2]</sup>。将无机硫掺入半胱氨酸是自然界硫循环的关键步骤。反应产生的半胱氨酸是唯一具有活性巯基 (-SH) 的氨基酸, 在许多酶的催化中心起关键作用, 也可进一步转化成谷胱甘肽 (Glutathione) 等重要物质<sup>[2-3]</sup>, 协助抵御有害物质、增加细胞活力。目前已在蜡梅 (*Chimonanthus praecox*)<sup>[2]</sup>、水稻 (*Oryza sativa*)<sup>[4]</sup>、拟南芥 (*Arabidopsis thaliana*)<sup>[5]</sup>、菠菜 (*Spinacia oleracea*)<sup>[6]</sup> 和金黄色葡萄球菌 (*Staphylococcus*

*aureus*)<sup>[7]</sup> 等多种植物和微生物中克隆了 CSase, 并已将 CSase 转入拟南芥和烟草 (*Nicotiana tabacum*) 中超表达, 发现转基因植株能够抗氧化应激、重金属等非生物胁迫<sup>[3,8-9]</sup>。然而目前尚无条斑紫菜 (*Porphyra yezoensis* Ueda) 中该基因的相关报道。本研究利用电子克隆 (in silico cloning) 技术<sup>[10]</sup>, 克隆了条斑紫菜半胱氨酸合成酶基因 (*PyCSase-B*) 的 cDNA 序列, 并对其序列进行了生物信息学分析。研究 *PyCSase-B* 有助于阐明条斑紫菜耐受重金属等非生物胁迫的作用机理, 对进一步利用超表达 *PyCSase-B* 转基因植株对重金属污染进行植物修复 (phytoremediation) 具有推动作用。

随着生物信息学资源 (包括数据库、计算机网络和应用软件等) 的发展, 电子克隆应运而生。电子克隆是利用计算机技术, 依托现有网络资源, 采用生物信息学方法, 对核酸序列进行组装和拼接, 最后利

收稿日期: 2008-12-17; 修订日期: 2009-03-03.

基金项目: 国家科技支撑计划课题资助项目 (2007BAD29B03); 江苏省海洋生物技术重点实验室开放课题资助项目 (2008HS004).

作者简介: 易乐飞 (1975-), 男, 硕士, 讲师, 从事分子遗传学研究. E-mail: yileifei@yahoo.com.cn

通讯作者: 刘楚吾. Tel: 0759-2382044; E-mail: liucw@gdou.edu.cn

用RT-PCR 快速地获得部分乃至全长cDNA序列的方法。电子克隆能够快速获得同源新基因序列并能分析其功能特征,为同源新基因的实验研究提供了重要指导,对进一步研究基因蛋白家族组成、结构及其功能都具有重要的理论和现实意义。目前已有利用电子克隆结果指导实验,成功克隆新基因的大量报道,如人(*Homo sapiens*)的 $\alpha$ -葡萄糖苷酶基因<sup>[11]</sup>、小鼠(*Mus musculus*)黏蛋白基因<sup>[12]</sup>、蚯蚓(*Eisenia foetida*)纤溶酶基因<sup>[13]</sup>、大肠杆菌(*Escherichia coli*)NAD<sup>+</sup>/NADP<sup>+</sup>依赖性琥珀酸半醛脱氢酶基因<sup>[14]</sup>等。但电子克隆技术在条斑紫菜基因克隆中的应用相对较少,本研究的技术路线可为条斑紫菜功能基因的快速克隆以及进一步开发、利用条斑紫菜资源奠定基础。

## 1 材料与方法

### 1.1 实验材料和主要试剂

条斑紫菜叶状体(gametophytic blade)取自连云港近海养殖群体。TRIzol reagent购于Invitrogen公司,RevertAid<sup>TM</sup> H Minus First Strand cDNA Synthesis试剂盒和Lambda DNA/*EcoR* I + *Hind* III Marker购自Fermentas公司,*Taq* Plus I (*Taq*+*Pfu*)和dNTP购于BioBasic公司,引物由上海生工生物工程技术有限公司合成。

### 1.2 *PyCSase-B*的电子克隆

**1.2.1 数据库和软件** 核酸和蛋白数据库为美国国立生物技术信息中心(National Center for Biotechnology Information, NCBI)维护的GenBank数据库,相似性比对使用基本局部比对搜索工具(Basic Local Alignment Search Tool, BLAST)<sup>[15]</sup>,表达序列标签(expressed sequence tag, EST)的组装和拼接使用CAP3<sup>[16]</sup>。

**1.2.2 电子克隆技术路线** 以拟南芥CSase的氨基酸序列(GenBank accession: BAF02052)作为查询序列,使用tBlastn<sup>[15]</sup>检索GenBank中条斑紫菜EST数据库(dbEST)。将检索到的高相似且含CSase保守序列的EST序列用CAP3拼接,CAP3拼接参数中的P设置为94,O设置为45<sup>[17]</sup>。以拼接产物重复检索、拼接和延伸步骤直至不能延伸为止,最后得到*PyCSase-B* cDNA的叠连群(contig)。

### 1.3 含完整编码区cDNA的实验克隆

**1.3.1 总RNA提取纯化与引物设计** 条斑紫菜叶状体总RNA提取纯化参照TRIzol试剂说明书。用电泳检测RNA完整性,紫外分光光度计检测RNA含量和纯度。在contig序列基础上,针对开放阅读框(ORF)设计引物,预计扩增片段长1 392 bp,包含*PyCSase-B*的完整ORF。正向引物:5' TCGCAGTCCAACCTCGCCTCT 3',反向引物:5' GCTCGGCCACATCAATCGT 3'。

**1.3.2 RT-PCR反应与产物的测序** 取2  $\mu$ g总RNA,以Oligo(dT)为引物,用RevertAid<sup>TM</sup> H Minus First Strand cDNA Synthesis试剂盒,按说明书的反应条件合成第一链cDNA。PCR反应体系为50  $\mu$ L,其中含5  $\mu$ L 10 $\times$  PCR反应缓冲液,1.75 mmol/L MgCl<sub>2</sub>,200  $\mu$ mol/L dNTP,上下游引物各0.5  $\mu$ mol/L,3 U *Taq* Plus I,1  $\mu$ L反转录产物为模板。反应条件为:95  $^{\circ}$ C预变性3 min,然后30个循环,每个循环95  $^{\circ}$ C变性60 s,57  $^{\circ}$ C退火30 s,72  $^{\circ}$ C延伸90 s,最后72  $^{\circ}$ C充分延伸5 min,4  $^{\circ}$ C保存。反应后取1  $\mu$ L电泳。将PCR产物送上海生工生物工程技术有限公司进行正反双向测序。

### 1.4 生物信息学分析

利用多种生物信息学软件分析*PyCSase-B*的cDNA序列和编码蛋白(PyCSase-B)。用ProtParam<sup>[18]</sup>分析PyCSase-B的各种基本理化特性,用PSORT<sup>[19]</sup>、ChloroP<sup>[20]</sup>和TargetP<sup>[20]</sup>分析PyCSase-B的亚细胞定位,以PyCSase-B为探针对GenBank的非冗余的蛋白质数据库进行Blastp<sup>[15]</sup>分析,用ClustalW<sup>[21]</sup>进行多序列比对,用ScanProsite<sup>[22]</sup>和CDD<sup>[23]</sup>进行功能结构域分析,用Swiss-model<sup>[24]</sup>、DALI<sup>[25]</sup>和UCSF Chimera<sup>[26]</sup>进行三维结构预测和三维比对,进而判断PyCSase-B是否为新序列以及推导其功能。

## 2 结果与分析

### 2.1 *PyCSase-B*的电子克隆结果

筛选条斑紫菜dbEST,共获得了10条符合聚类条件的条斑紫菜EST序列(GenBank accession: AU196735、AV432275、AV431559、AU188315、AU189347、AU189297、AV434527、AV435783、AV435569、

AV438271),最后拼接产生1条长1 481 bp的contig。

## 2.2 含完整编码区cDNA的实验克隆结果

PCR扩增后得到了长度为1 400 bp左右的单一特异性片段,与预期的1 392 bp十分相近(图1)。双向测序获得了*PyCSase-B*的cDNA序列(GenBank accession: FJ232911)。*PyCSase-B* cDNA序列与电子克隆所得contig序列比对结果表明两者的差异碱基2个,一致性为99.92%。*PyCSase-B*和contig的编码蛋白比对结果表明两者的差异残基1个,一致性为99.74%。如果将个体间多态性和EST固有的错误率等因素考虑进去,测序结果与电子克隆结果一致。

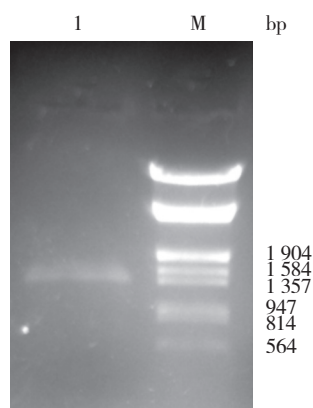


图1 PCR扩增产物电泳

1: 扩增结果; M:  $\lambda$  DNA/*EcoR* I + *Hind* III Marker.

Fig. 1 Gel electrophoresis of PCR product

1: PCR product; M:  $\lambda$  DNA/*EcoR* I + *Hind* III Marker.

## 2.3 *PyCSase-B* cDNA序列的生物信息学分析

在*PyCSase-B* cDNA序列的第94~1 245 nt处含有一连续ORF。第94~96 nt处是该cDNA序列的第一个ATG密码子,其下游的+4位是G,上游-3位是T,不符合典型的kozak规则;但在上游3个读码框存在终止密码子,再结合推导的氨基酸序列N端一致性比较结果,可以认定第一个ATG是该基因的起始密码子<sup>[27]</sup>。表明克隆的cDNA含完整ORF,编码1个含有383 AA的蛋白。*PyCSase-B*分子量为39.2 kD,其碱性氨基酸、酸性氨基酸、疏水性氨基酸和极性氨基酸数量分别为34、37、152和82个,等电点为pH5.9。

与其他植物的胞质型基因相比,本研究中克隆

的*PyCSase-B*的N端多出60~70个氨基酸残基,这段序列与定位到叶绿体和线粒体的转运肽(Transit peptide)具有共通特征<sup>[6]</sup>:富含碱性氨基酸和疏水性氨基酸。TargetP(cTP Score: 0.735)、ChloroP(cTP: Y, cTP-length: 60)以及PSORT(chlo:8.5)分析结果表明,*PyCSase-B*的N端前60个氨基酸构成了叶绿体转运肽序列,*PyCSase-B*定位在叶绿体中。生化分析表明高等植物拥有多个不同亚细胞定位的CSase异形体(isoform)<sup>[6]</sup>,例如在菠菜叶片中,2种主要的CSase活性分别位于胞质和叶绿体中,另一种次要的CSase活性位于线粒体中。通常将定位到胞质、叶绿体和线粒体的异形体分别用A、B和C表示,所以将本实验克隆的基因命名为*PyCSase-B*。植物细胞的CSase分布在所有有蛋白质合成的亚细胞区域,可能是半胱氨酸在细胞内的不同区域间不能自由运输,蛋白质合成所需的半胱氨酸必须在不同的区域内独立合成<sup>[28]</sup>。

在GenBank的非冗余蛋白质数据库中发现紫红紫菜(*P. purpurea*)、莱茵衣藻(*Chlamydomonas reinhardtii*)、拟南芥、水稻等多种藻类和高等植物的各类型CSase与*PyCSase-B*高度相似(similarity),但在条斑紫菜的蛋白质数据库没有找到与之高度相似的序列。将已知亚细胞定位的不同生物的CSase<sup>[29]</sup>、部分藻类和细菌CSase以及*PyCSase-B*进行多序列比对并构建系统发生进化树(图2),结果表明,在进化上条斑紫菜与同属红藻的紫红紫菜遗传距离最近。除去转运肽的成熟*PyCSase-B*与紫红紫菜、莱茵衣藻、拟南芥、水稻的对应序列一致性(identity)分别为94%、69%、64%和65%;若考虑功能相似性(similarity)氨基酸之间的保守替换,与它们的序列相似性分别为97%、80%、79%和79%,说明CSase基因家族在植物中具有较高的保守性。从进化树上可以看出,胞质型CSase(CSase-A)和细菌的CSase各自聚为一簇,叶绿体型CSase(CSase-B)和线粒体型CSase(CSase-C)聚为一大簇,暗示着细胞器型CSase与胞质型CSase在进化上分歧较早,但叶绿体型和线粒体型CSase分歧不明显。进化树上藻类的CSase在进化上处在细菌和高等植物之间,符合其经典的分类学地位。

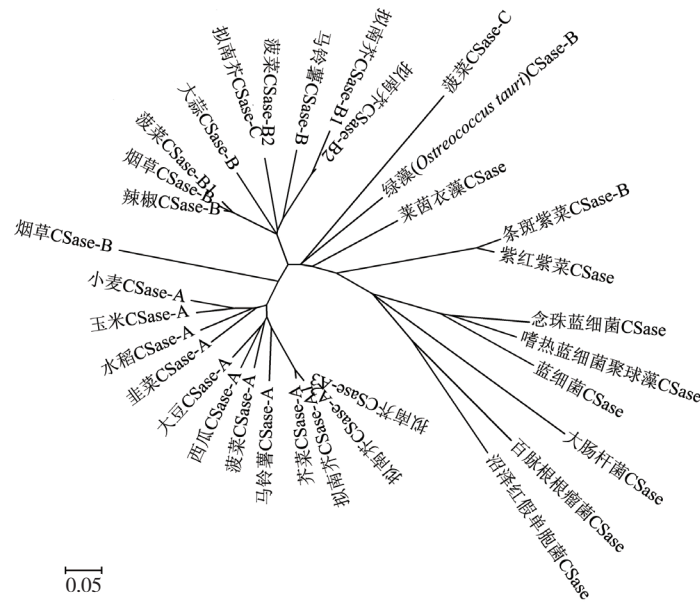


图2 不同生物CSase的系统发生进化树

用neighbour-joining (NJ) 法构建了进化树. CSase序列分别来自下列生物: 条斑紫菜、莱茵衣藻、紫红紫菜、拟南芥、芥菜、菠菜、西瓜、水稻、玉米、韭菜、小麦、辣椒、大肠杆菌、蓝细菌、大豆、马铃薯、念珠蓝细菌、嗜热蓝细菌聚球藻、百脉根根瘤菌、烟草、绿藻、大蒜、沼泽红假单胞菌、菠菜.

CSase-A、CSase-B和CSase-C分别表示胞质型、叶绿体型和线粒体型CSase.

Fig. 2 Cladogram for the phylogenetic relationship of CSase proteins

The phylogenetic tree was constructed using neighbour-joining methods. The source organisms for CSase sequences were as follows: *Porphyra yezoensis*, *Chlamydomonas reinhardtii*, *Porphyra purpurea*, *Arabidopsis thaliana*, *Brassica juncea*, *Spinacea oleracea*, *Citrullus lanatus*, *Oryza sativa*, *Zea mays*, *Allium tuberosum*, *Triticum aestivum*, *Capsicum annum*, *Escherichia coli*, *Synechocystis sp.*, *Glycine max*, *Solanum tuberosum*, *Nostoc sp.*, *Thermosynechococcus elongates*, *Mesorhizobium loti*, *Nicotiana tabacum*, *Ostreococcus tauri*, *Allium sativum*, *Rhodospseudomonas palustri* and *Spinacia oleracea*. CSase-A, CSase-B and CSase-C is refer to cytoplasmic, chloroplastic and mitochondrial CSase respectively.

结构域分析结果(图3)显示PyCSase-B第76~366位氨基酸序列属于类CBS亚群(CBS\_like subgroup, 该亚群包含胱硫醚 $\beta$ 合酶和半胱氨酸合成酶), 第78~358位属于色氨酸合成酶 $\beta$ 超家族(trypthophan synthase beta superfamily (fold type II)). 第99~117位氨基酸为半胱氨酸合成酶/胱硫醚 $\beta$ 合酶的PLP连接位点并具有下列保守序列K-x-E-x(3,4)-[PAF]-[STAGC]-x-S-[IVAPM]-K-x-R-x-[STAG]-x(2)-[LIVM], 其中活性位点中的第二个赖

氨酸残基(Lys<sup>110</sup>)负责连接PLP。酶活性位点的19个氨基酸残基和Lys<sup>110</sup>与其他物种CSase氨基酸序列完全匹配。图3显示PyCSase-B含有许多与酶活性密切相关的氨基酸残基。例如与PLP共价结合的Lys<sup>110</sup>, 与PLP形成氢键并维持其正确排列的Asn<sup>141</sup>、Ser<sup>331</sup>、Thr<sup>247</sup>、His<sup>221</sup>、Gly<sup>243</sup>、Gly<sup>245</sup>、Thr<sup>244</sup>, 与OAS结合的Asn<sup>141</sup>、Gln<sup>211</sup>, 参与将硫转移到半胱氨酸的Thr<sup>138</sup>、Ser<sup>139</sup>, 与SAT-O

三维结构(图4)显示,成熟PyCSase-B的三维结构类似于拟南芥CSase(PDB ID:1Z7W)<sup>[1]</sup>。成熟PyCSase-B单体由2个 $\alpha\beta$ 结构域构成;N端结构域较小,其中心由4链反平行 $\beta$ 片层( $\beta 3B-\beta 6B$ )构成,外周是4条 $\alpha$ 螺旋( $\alpha 1-\alpha 4$ );C端结构域较大,含有6链 $\beta$ 片层( $\beta 1A-\beta 2A$ 和 $\beta 7A-\beta 10A$ ),其旁侧为4条 $\alpha$ 螺旋( $\alpha 5-\alpha 8$ ),来自N端的残基构成了C端 $\beta$ 片层的头2条链( $\beta 1A-\beta 2A$ );另有一条 $\alpha$ 螺旋( $\alpha 9$ )参与了二聚体的形成。但与拟南芥相比,PyCSase-B在 $\beta 1A-\beta 2A$ 环多2氨基酸残基,在 $\alpha 5-\beta 7A$ 环

少2氨基酸残基。使用DALI对PDB(Protein Data Bank)进行搜索,结果显示PyCSase-B在三维结构上不仅与多种植物的CSase高度相似,还与胱硫醚 $\beta$ 合酶、苏氨酸合酶、色氨酸合酶、1-氨基环丙烷-1-羧酸脱氨酶相似。虽然PyCSase-B与他们的氨基酸序列一致性较低,但它们都需要PLP作为辅助因子,因此相似的拓扑结构便于PLP的结合,活性位点上的序列差异又可以识别不同的底物,执行不同的功能<sup>[1]</sup>。

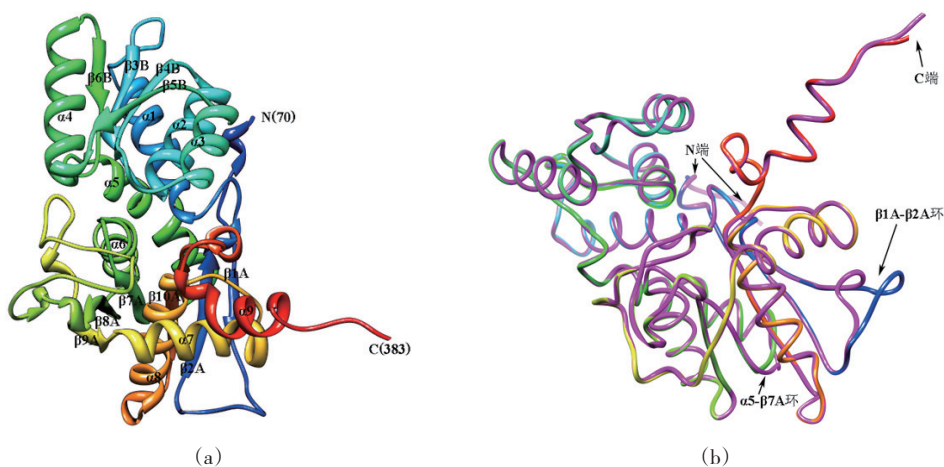


图4 成熟PyCSase-B单体以及与拟南芥CSase单体的空间结构比对  
(a)成熟PyCSase-B单体的三维结构,N端和C端残基数以及二级结构进行了标记。  
(b)成熟PyCSase-B单体与拟南芥CSase空间结构的比较,箭头指示差异处。

Fig.4 Tertiary structure of mature PyCSase-B monomer and structural comparison between mature PyCSase-B monomer and *A. thaliana* CSase

(a) Ribbon diagram of the mature PyCSase-B monomer. The N-terminal and C-terminal residues are indicated with residue number. Secondary structure features are labeled. (b) Overlay of the C $\alpha$  traces of *A. thaliana* CSase and mature PyCSase-B monomers. The arrows indicate regions of structural variation.

### 3 讨论

虽然CSase在细胞内是持续表达,但环境应激可以改变CSase的表达和酶活性,例如硫、氮、碳缺乏以及盐、硫、重金属等非生物胁迫能增强CSase的表达。而且CSase在转基因植物中的超表达能够提高转基因植物对重金属的耐受力,提高对硫胁迫的耐受,提供对氧化应激的保护<sup>[1,3,6,8-9]</sup>。CSase参与催化的反应是半胱氨酸生物合成的最后一步,那么它是如何参与植物的耐胁迫过程呢?在植物中,CSase催化产生的半胱氨酸对环境无机硫的同化起中心作用,并且为

甲硫氨酸、谷胱甘肽、植物络合素(phytochelatins)、铁-硫簇(iron-sulfur clusters)、维生素辅助因子和多种次级代谢物的生物合成提供唯一的硫化物供体<sup>[1]</sup>。半胱氨酸是合成谷胱甘肽的前体物之一,强烈调节谷胱甘肽的合成,过量表达CSase能够增加半胱氨酸和谷胱甘肽的含量,由于谷胱甘肽是胞内代谢过程和植物遭受氧化胁迫产生的过氧化物有效清除剂之一,因此过量表达CSase能够提高植物对过氧化物的清除能力从而提高植物本身抗氧化胁迫能力<sup>[30]</sup>。植物络合素是植物和微生物中一类小分子的能与重金属结合

的富含半胱氨酸的多肽 $[(\gamma\text{-Glu-Cys})_n\text{Gly}(n=2\sim 11)]$ ,在重金属及重金属复合物的胁迫下,植物络合素在胞质中合成,然后转移到液泡中,参与镉(Cd)等重金属的结合<sup>[31]</sup>,因此过量表达CSase能够提高植物抗镉等重金属污染的能力。

传统的条斑紫菜基因克隆方法相对繁杂和耗时,例如抗坏血酸过氧化物酶基因<sup>[32]</sup>和液泡型的H<sup>+</sup>-ATPase的C亚基基因<sup>[33]</sup>的克隆都需要构建cDNA文库。相比而言,电子克隆技术更为快速、简便且目的性强,例如钙调素基因<sup>[34]</sup>的克隆仅需进行生物信息学预测、RT-PCR和测序,就能获得其cDNA序列。本研究也采用电子克隆技术快速获得了PyCSase-B的cDNA序列,有助于进一步研究PyCSase-B在条斑紫菜抗逆调控中的作用。随着更多海洋生物测序的开展以及各类数据库的不断完善,电子克隆技术必将成为海洋生物新基因快速克隆的重要手段之一。

#### 参考文献:

- [1] Bonner E R, Cahoon R E, Knapke S M, et al. Molecular basis of cysteine biosynthesis in plants: Structural and functional analysis of O-acetylserine sulfhydrylase from *Arabidopsis thaliana* [J]. J Biol Chem, 2005, 280 (46): 38 803–38 813.
- [2] 蒋安. 蜡梅半胱氨酸合成酶基因CpCysA的克隆分析与表达载体构建[D]. 重庆: 西南大学, 2006: 1–47.
- [3] Dominguez-solis J R, Gutierrez-alcala G, Romero L C, et al. The cytosolic O-Acetylserine (thiol) lyase gene is regulated by heavy metals and can function in cadmium tolerance [J]. J Biol Chem, 2001, 276 (12): 9 297–9 302.
- [4] Nakamura T, Yamaguchi Y, Sano H. Four rice genes encoding cysteine synthase: isolation and differential responses to sulfur, nitrogen and light [J]. Gene, 1999, 229 (1–2): 155–161.
- [5] Saito K, Miura N, Yamazaki M, et al. Molecular cloning and bacterial expression of cDNA encoding a plant cysteine synthase [J]. PNAS, 1992, 89 (17): 8 078–8 082.
- [6] Takahashi H, Saito K. Subcellular localization of spinach cysteine synthase isoforms and regulation of their gene expression by nitrogen and sulfur [J]. Plant Physiol, 1996, 112 (1): 273–280.
- [7] Lithgow J K, Hayhurst E J, Cohen G, et al. Role of a cysteine synthase in *Staphylococcus aureus* [J]. J Bacteriol, 2004, 186 (6): 1 579–1 590.
- [8] Youssefian S, Nakamura M, Orudjev E, et al. Increased cysteine biosynthesis capacity of transgenic tobacco overexpressing an O-acetylserine (thiol) lyase modifies plant responses to oxidative stress [J]. Plant Physiol, 2001, 126 (3): 1 001–1 011.
- [9] Noji M, Saito M, Nakamura M, et al. Cysteine synthase overexpression in tobacco confers tolerance to sulfur-containing environmental pollutants [J]. Plant Physiol, 2001, 126 (3): 973–980.
- [10] Gill R W, Sanseau P. Rapid in silico cloning of gene using expressed sequence tags (ESTs) [J]. Biotechnol Annu Rev, 2000, 5: 25–44.
- [11] Hirschhorn R, Huie M L, Kasper J S. Computer assisted cloning of human neutral alpha-glucosidase C (GANC): A new paralog in the glycosyl hydrolase gene family 31 [J]. PNAS, 2002, 99 (21): 13 642–13 646.
- [12] Chen Y I, Zhao Y U, Wu R E. In silico cloning of mouse *Muc5b* gene and upregulation of its expression in mouse asthma model [J]. Am J Respir Crit Care Med, 2001, 164 (6): 1 059–1 066.
- [13] Zhao X Y, Gao S, Cui D L, et al. In silico cloning of *Efp-O*, a novel earthworm fibrinolytic enzyme gene, and verification of its coding region by RT-PCR [J]. Chinese Journal of Biotechnology, 2006, 22 (6): 897–901.
- [14] Fuhrer T, Chen L, Sauer U, et al. Computational prediction and experimental verification of the gene encoding the NAD<sup>+</sup>/NADP<sup>+</sup>-dependent succinate semialdehyde dehydrogenase in *Escherichia coli* [J]. J Bacteriol, 2007, 189 (22): 8 073–8 078.
- [15] Altschul S F, Madden T L, Schaffer A A, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs [J]. Nucleic Acids Res, 1997, 25 (17): 3 389–3 402.
- [16] Huang X, Madan A. CAP3: A DNA sequence assembly program [J]. Genome Res, 1999, 9 (9): 868–877.
- [17] 张利达, 袁德军. 一种新的EST聚类方法[J]. 遗传学报, 2003, 30 (2): 147–153.
- [18] Wilkins M R, Gasteiger E, Bairoch A, et al. Protein identification and analysis tools in the ExPASy server [J]. Methods Mol Biol, 1999, 112: 531–552.
- [19] Nakai K, Horton P. PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization [J]. Trends Biochem Sci, 1999, 24 (1): 34–36.
- [20] Emanuelsson O, Nielsen H, Von H G. ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites [J]. Protein Sci, 1999, 8 (5): 978–984.
- [21] Fukami-kobayashi K, Saito N. How to make good use of CLUSTALW [J]. Tanpakushitsu Kakusan Koso, 2002, 47 (9): 1 237–1 239.
- [22] De C E, Sigrist C J, Gattiker A, et al. ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins [J]. Nucleic Acids Res, 2006, 34 (suppl-2): 362–365.
- [23] Marchler-bauer A, Anderson J B, Cherukuri P F, et al. CDD: a conserved domain database for protein classification [J]. Nucleic Acids Res, 2005, 33 (Database issue): 192–196.
- [24] Schwede T, Kopp J, Guex N, et al. SWISS-MODEL: an automated protein homology-modeling server [J]. Nucleic Acids Res, 2003, 31 (13): 3 381–3 385.

- [25] Holm L, Kaariainen S, Rosenstrom P, et al. Searching protein structure databases with DaliLite v.3 [J]. *Bioinformatics*, 2008; 507.
- [26] Pettersen E F, Goddard T D, Huang C C, et al. UCSF Chimera-a visualization system for exploratory research and analysis [J]. *J Comput Chem*, 2004, 25 (13): 1 605–1 612.
- [27] 王冬冬, 朱延明, 李勇, 等. 电子克隆技术及其在植物基因工程中的应用 [J]. *东北农业大学学报*, 2006, 37 (3): 403–408.
- [28] Warrilow A, Hawkesford M. Separation, subcellular location and influence of sulphur nutrition on isoforms of cysteine synthase in spinach [J]. *J Exp Bot*, 1998, 49 (327): 1 625–1 636.
- [29] Chronis D, Krishnan H B. Sulfur assimilation in soybean: Molecular cloning and characterization of O-acetylserine (thiol) lyase (cysteine synthase) [J]. *Crop Science*. 2003, 43: 1 819–1 827.
- [30] 刘明坤, 刘关君, 魏志刚, 等. 西伯利亚蓼半胱氨酸合成酶基因的克隆与表达 [J]. *遗传*, 2008, 30 (10): 1 363–1 371.
- [31] Wojas S, Clemens S, Hennig J, et al. Overexpression of phytochelatin synthase in tobacco: distinctive effects of *AtPCS1* and *CePCS* genes on plant response to cadmium [J]. *J Exp Bot*, 2008, 59 (8): 2 205–2 219.
- [32] 马凌波, 张凤英, 马春艳, 等. 条斑紫菜抗坏血酸过氧化物酶基因的 cDNA 克隆和序列分析 [J]. *中国水产科学*. 2005, 12 (3): 281–287.
- [33] Kitade Y, Yamazaki S, Watanabe T, et al. Structural features of a gene encoding the vacuolar H<sup>+</sup>-ATPase c subunit from a marine red Alga, *Porphyra yezoensis* [J]. *DNA Res*, 1999, 6 (5): 307–312.
- [34] 王孟强. 条斑紫菜功能基因组特性与抗逆相关基因表达分析 [D]. 青岛: 中国海洋大学, 2007: 165–178.

## In silico cloning and analysis of cysteine synthase gene from *Porphyra yezoensis* Ueda

YI Le-fei<sup>1</sup>, WANG Ping<sup>1</sup>, ZHOU Xiang-hong<sup>1</sup>, LIU Chu-wu<sup>2</sup>

(1. Jiangsu Key Laboratory of Marine Biotechnology, Huaihai Institute of Technology, Lianyungang 222005, China; 2. Fisheries College, Guangdong Ocean University, Zhanjiang 524088, China)

**Abstract:** The cloning and characterization of cDNA of *P. yezoensis* cysteine synthase gene (*PyCSase-B*) by in silico cloning were reported. Cysteine synthase from *Arabidopsis thaliana* was used as a probe to search the *P. yezoensis* dbEST. ESTs showing great similarity to *CSase* were assembled into contig. According to the ORF of the contig, the primers were designed. Total RNA was isolated from gametophytic blade of *P. yezoensis*. And the cDNA (GenBank accession: FJ232911) of *PyCSase-B* was cloned successfully through RT-PCR. This cDNA contains an 1 152 nt continuous complete ORF encoding a 383 amino acids polypeptide (PyCSase-B) with a calculated molecular mass of 39.3 kDa. The N-terminal 60 amino acids of PyCSase-B formed a chloroplastic transit peptide, which led and located the mature *PyCSase-B* in chloroplast. Sequence comparison of the mature PyCSase-B revealed 94%, 69%, 64% and 65% amino acid sequence identity with CSase from *P. purpurea*, *Chlamydomonas reinhardtii*, *A. thaliana* and *Oryza sativa*. The phylogenetic tree indicated that PyCSase-B was situated between the bacteria and the higher plants in the evolution. PyCSase-B contained conserved amino acids residues and domains which interacted with pyridoxal 5' -phosphate and O-acetylserine, and transferred sulfur to cysteine. The mature PyCSase-B possesses similar tertiary structure to *A. thaliana* CSase, and is constructed with 2  $\alpha/\beta$  domains. The cloning and bioinformatical analysis of *PyCSase-B* gene laid a foundation for subsequent research of its expression regulation under abiotic stresses. [Journal of Fishery Sciences of China, 2009, 16 (4): 518–524]

**Key words:** *Porphyra yezoensis*; cysteine synthase; in silico cloning; bioinformatics

**Corresponding author:** LIU Chu-wu. Tel: 0759-2382044; E-mail: liucw@gdou.edu.cn