

文章编号:1673-1689(2008)02-0092-06

产甘油假丝酵母烯醇化酶基因的克隆及序列分析

姚雁, 沈微*, 饶志明, 王正祥, 诸葛健*

(食品科学与技术国家重点实验室, 江南大学, 江苏 无锡 214122)

摘要:从产甘油假丝酵母(*Candida glycerinogenes*)基因文库中分离了一个含烯醇化酶基因(CgENO1)的克隆子。插入片段全长2 456 bp,包含一段1 314 bp的没有内含子的蛋白质编码区,编码一个含438个氨基酸残基的多肽链,其氨基酸序列与来源于酿酒酵母的烯醇化酶相似性为74.3%。多重序列比对显示,产甘油假丝酵母烯醇化酶(CgENO1)含有酵母烯醇化酶全部保守区。对编码区上下游序列分析显示,克隆片段含有典型的酵母基因上下游调控区,是一个完整的酵母新基因。

关键词:产甘油假丝酵母;烯醇化酶;序列分析

中图分类号: Q 786

文献标识码: A

Cloning and Sequence Analysis of the *Candida glycerinogenes* Enolase Gene

YAO Yan, SHEN Wei*, RAO Zhi-min, WANG Zheng-xiang, ZHUGE Jian*

(National Key Laboratory of Food Science and Technology, Jiangnan University, Wuxi 214121, China)

Abstract: A DNA clone containing the putative *Candida glycerinogenes* enolase gene (CgENO1) was isolated from genomic library. The 2456 bp insert contained a intronless open reading frame (ORF) of 1314 bp encoding a peptide of 438 amino acid residues with 74.3% similarity to the *Saccharomyces cerevisiae* enolase 1. Multi-alignment comparison revealed the CgENO1 contained all the conserved regions of yeast enolase. Sequence analysis of 5'- and 3'-flanks of the ORF revealed the existence of several regulatory elements of the yeast genes, so the gene cloned was a complete new gene.

Key words: *Candida glycerinogenes*; enolase; sequence analysis

产甘油假丝酵母(*Candida glycerinogenes*)是中国发酵甘油工业的生产菌种,除过量合成甘油外,该酵母还具有耐受高渗透压的优良性能^[1]。我

国是世界上受荒漠化和土地盐碱化危害最严重的国家之一,采用转基因方法研究开发抗旱、抗盐植物对我国环境保护和农业发展具有重要的应用价

收稿日期:2007-03-06.

基金项目:国家自然科学基金项目(30570142).

作者简介:姚雁(1980-),女,河北怀来人,发酵工程硕士研究生.

* 通讯作者:沈微(1968-),男,江苏无锡人,工学博士,副教授,主要从事微生物学的研究.

Email: microshen@gmail. google. com

诸葛健(1939-),男,浙江金华人,教授,博士生导师,主要从事发酵工程和工业微生物领域的研究.

Email: jianzhuge@hotmail. com

值。烯醇化酶催化糖酵解途径中 2-磷酸甘油酸到磷酸烯醇式丙酮酸的反应,研究表明,植物烯醇化酶基因在干旱、高盐等环境下往往有较高的表达水平,被认为是抗逆相关基因^[2-3]。烯醇化酶基因的表达与酵母的渗透压适应有密切关系,Norbeck J 等对酿酒酵母的研究发现,当酿酒酵母从含 0.7 mol/L NaCl 的生长环境转移到 1.4 mol/L 的生长环境中时,有多个酶表达量有较大幅度上升,其中上升幅度最大的是烯醇化酶^[4]。产甘油假丝酵母具有耐高渗的优良性能,是提供抗逆相关基因特别是耐盐相关基因的优良物种。产甘油假丝酵母及其近缘种的烯醇化酶基因均未见报道,作者采用 PCR 检测的方法从产甘油假丝酵母基因库中克隆了包括完整的上下游调控区的产甘油假丝酵母烯醇化酶基因,为阐明烯醇化酶与产甘油假丝酵母耐高渗之间的关系以及烯醇化酶基因的应用研究奠定了基础。

1 材料与方法

1.1 菌株及培养条件

产甘油假丝酵母 WL2002-5 由工业微生物研究室分离保藏,并用作染色体 DNA 来源。产甘油假丝酵母基因文库由作者所在实验室构建^[5-6],是在质粒 YEp51 基础上构建的,宿主菌为大肠杆菌 DH5 α 。宿主菌大肠杆菌 JM109 为复旦大学遗传学教研室惠赠。大肠杆菌细胞培养在含 100 μ g/mL 氨苄青霉素的 LB 培养基中(1 g/dL 蛋白胨,0.5 g/dL 酵母提取物,0.5 g/dL 氯化钠),培养温度为 37 $^{\circ}$ C。

1.2 酶和试剂

耐热 DNA 聚合酶(Taq 酶)、连接酶、pMD18-T 载体和 DNA 标准相对分子质量 Marker 等为宝生物工程(大连)有限公司产品。PCR 产物纯化试剂盒、从琼脂糖凝胶中回收 DNA 片断的试剂盒为上海华舜生物工程有限公司产品。引物由上海赛百盛基因技术有限公司合成。

1.3 DNA 重组及序列分析

DNA 重组、质粒提取、PCR 等操作均按文献[7]方法进行。用 DNAMAN 序列分析软件进行序列分析、序列比对及绘制进化树。在 NCBI (www.ncbi.nlm.nih.gov)中利用 BLAST 进行核苷酸及蛋白质同源性搜索。

1.4 产甘油假丝酵母烯醇化酶基因的克隆

从 NCBI 中查找已发表的酵母烯醇化酶基因序列,并进行多重序列比对,在序列高度保守区设计

简并引物:

Peno01:5'-GGTGGT(T,G)C(C,A)CA(C,T)GC(T,G)GG-3'

Peno02:5'-GAAACCAT(A,G)ACACCCCA(A,G)CC-3'

引物 Peno01 与酿酒酵母烯醇化酶结构基因编码链 469~485 位核苷酸序列对应,引物 Peno02 与酿酒酵母烯醇化酶基因 1099~2018 位核苷酸序列互补对应,以酿酒酵母烯醇化酶基因为依据,引物间 DNA 长为 650 bp。以产甘油假丝酵母染色体 DNA 为模板,利用简并引物 Peno01 和 Peno02 扩增产甘油假丝酵母烯醇化酶部分结构基因。PCR 按下述反应条件进行:1) 94 $^{\circ}$ C 4 min,2) 94 $^{\circ}$ C 50 s,45 $^{\circ}$ C 60 s,72 $^{\circ}$ C 90 s,循环 30 次,3) 72 $^{\circ}$ C 10 min。PCR 产物用 1 g/dL 凝胶电泳分离,小量胶回收试剂盒纯化回收。将扩增得到的约 650 bp 的 DNA 片断与 pMD18- $\uparrow$ 载体连接,重组质粒送上海基康生物技术公司测序。

分析测序结果,发现测序片断序列共 650 个碱基,所编码的氨基酸序列与其它酿酒酵母烯醇化酶相对应部分有 86% 的同源性。

将产甘油假丝酵母基因组文库适当稀释后涂布含氨苄青霉素的 LB 平板,控制每块平板的菌落数在 100 个左右,共涂布 150 块平板,37 $^{\circ}$ C 培养 16 h,将每块平板上形成的菌落影印到另一平板上,将新影印的平板放到 37 $^{\circ}$ C 培养而原先的平板则分别用生理盐水将菌落全部洗下,分别收集,形成 150 个分组文库(编号 001-150)。每组各取少量收集液,接种含氨苄青霉素的 LB 液体培养基,培养液提取质粒,分别以各组质粒为模板,用 Peno01、Peno02 为引物进行 PCR 检测,条件如下:1) 94 $^{\circ}$ C 4 min,2) 94 $^{\circ}$ C 50 s,55 $^{\circ}$ C 90 s,72 $^{\circ}$ C 1 min,循环 30 次;72 $^{\circ}$ C 10 min,电泳鉴定 PCR 扩增结果,如电泳显示 PCR 扩增产生 0.65 kb 左右的片段则为阳性,表明其中含有 CgENO 基因。将 PCR 检测为阳性的分组文库对应的影印平板上的菌落分别接种含氨苄青霉素的 LB 液体培养基,培养后提取质粒,以质粒为模板,用 Peno01、Peno02 为引物进行 PCR 鉴定,将结果为阳性的菌落划线分离后,送上海基康生物技术公司测序。

2 结果与讨论

2.1 产甘油假丝酵母烯醇化酶基因的克隆

以产甘油假丝酵母染色体 DNA 为模板,用设计好的简并引物 Peno01 和 Peno02 扩增产甘油假

丝酵母烯醇化酶基因部分核苷酸序列,PCR 产物电泳显示获得 0.65 kb 的扩增片段,序列比对显示该片段编码的氨基酸序列与酿酒酵母烯醇化酶氨基酸序列有 86% 的同源性。充分证明所扩增的片断确实是来源于产甘油假丝酵母烯醇化酶基因。

以产甘油假丝酵母基因组文库为模板,以 Peno01 和 Peno02 为引物进行 PCR 检测,结果显示共有 3 个分组文库检测结果为阳性,即含有烯醇化酶基因。将这 3 个分组文库对应的影印平板上的菌落分别分组编号后进行 PCR 检测,最后获得 3 个阳性菌落(对应菌落编号分别为 008-021、031-089、128-047)。将这 3 个阳性菌落所含重组质粒分别进行测序。测序结果显示,3 个菌落所含质粒都是在 YEp51 多克隆位点中插入了长短不同的外源片段形成的重组质粒,其中 031-089 菌落所含片段最长,为 2 567 bp,序列分析显示该片段含有完整的烯醇

化酶结构基因、上游 699 个碱基和下游 554 个碱基。另两个菌落所含外源片段分别为 2 456 bp 和 1 527 bp,序列比对显示,031-089 号菌落所含插入片段包含了另两个质粒所含片段,而 3 个片段的对应部分序列完全相同。本研究所使用的基因文库是用核酸内切酶 *Sau3A I* 将产甘油酵母基因组 DNA 部分酶切后与经 *BamH I* 酶切的载体 YEp51 连接获得的,上述 3 个质粒所含插入片段很可能是同一段 DNA 片段被 *Sau3A I* 在不同部位酶切后获得的不同长度的片段。将菌落 031-089 所含插入片段核苷酸序列在 NCBI 登录,GenBank 序列号为 EF432726,序列名称为产甘油假丝酵母烯醇化酶基因(*C. glycerinogenes* enolase gene, CgENO1)。

2.2 产甘油假丝酵母烯醇化酶基因序列分析

菌落 031-089 所含插入片段核苷酸序列以及推测的氨基酸序列见图 1。

```

-699 GATCCTTCATAGTAGGTACCATATGCGACCACTAGCACAACAAACACCGACCCAGACATACAAATCA
-630 ACAACACTTCGCCAAAGACACCCCTTTCCAGGGAGGATCCCAACGCTCTCCATATGTCTGTGGCCCATGTCTGTGACACCGTAACCAACAC
-525 CAACCAACCCGTCATTGTACTGGGATGGTCGTCATAGACACCTCTCCACGGGGAACACCTCAITCGTAAACCGCCATGGTTACCGTTCCTCTGACTCGCCC
-420 CGTTGTTGATGCTCGCAGCTGTGGTGGCCCAACATGGTGTATATCGTGAACCAACACATGTGCAGCACATGTGTTAAAGAGTGTCTGGAGGTGG
-315 ATCATGATGGAAGTGGACTTTACCACTTGGGAAGTGTCTCCACTCCCGGGAAGAAAGACCCGGGTATCACGGGTTCCTCAATGGGGCAATTGGAAGGAGA
-210 AATATAGGGAATAACGCTGCTCTCGGACGGGAAGAGTTCAGACTATGAGGGGGGGGGGTGTAATATAAGACAGGAGATGTCCACCCCGAGAGAAGAA
-105 GAAAGTTGGAACCTTTAGAGAGAGAGATAAATTTTCCCAAGTGTCCATCAATACACAACCAACAGCAAACTCTATATTTACACATATAACCCCTCGAACCAACA
1  ATGTCCACTGTGTACCAAGTCTTTGCAAGATGTGTCTATGATTCCAGAGGTAAACCAACTGTGAAGTTGAATTAACCACTGCTGAAGGTGTCTTTAGAACCAATT
1  M S T V T K V F A R S V Y D S R G N P T V E V E L T T A E G V F R T I
106 GTTCCATCTGGTGCACTCTACCGGTGTCCAGAGGCTGTAGAACTAAGAGATGGCGACAAGTCCCAAGTGGTTAGGCAAGGCTGTGCAATGTGTC
38  V P S G A S T G V H E A L E L R D G D K S K W L G K Q V T K A V D N V
211 AACAAGATTATTCGCCAGCTTTAGTTGAGGCAAGGTTCCAGACTTGAACAATTGAAGATTGACCAATTCTTGATCAACCTTGACGGTACTCGCAACCAAGGGT
71  N K I A P A L V E A K V P V T E Q L K I D Q F L I N L D G T P N K Q
316 AAGCTGGGTGCAATGCCATTCTTGGTGTCTCCCTGGCTGTGTCTGAAGGCTGCTGCAGCTTATGAGAAGGTTCCGATTATACCAATCAATCAAGATCTCCGGT
108 K L G A N A I L G V S L A V A K A A A A Y E K V P L Y Q Y I H K I S G
421 TCCAGGATGCTTATGTCTTCCGACTTCCATTATGATGTGCTCAATGGTGGTCCGACGGGGGGTTCCTTCCCTTCCCAAGAAATTCATGATTGCTCCCAACC
141 S K D A Y V L P V P F M D V V N Q S H A G G S L F E F M I A P T
526 GGTGCTGAATCCTCTCTGAAAGCTTGAAGATGGGTTCTGAAGTTTACCAACCACTTGAAGTCTCTGACCAAGAAGACCTACGGTCAATCCGCTGGTAACGTTGGT
176 G A E S F S E A L R M G S E V Y H L K S L T K K T Y G Q S A Q N V G
631 GACGAAGGTGGTGTTCAGATATCCACACCCCAAGAGAGCTCTTGACCTGATTATGTCTGCTATTAAGGCTGCAGGTACGAGGGTAAGATTGACATTGCA
211 D E G G V A P D I H T P Q E A L D L I M S A I K A G Y E G K I D I A
736 ATGGAATCTGCTCCTCCGAAATCTACAGAAGATGTAAGTACGACTTGGCTTACAAGGATCCAACTCCGACGGTTCACAGGCCATCACTGGTTCGCAATTAGGT
246 M D S A S S E F Y K N G K Y D L A Y K D P N S D G S K A I T G S Q L G
841 GACTTGTACGCTGAATTTGTGAAGAACTACCAATTTGTCTCTGGAAGATCCATTTGCAAGAAGTACTGGGAACCTGGACTTAACCACTTTGAAAGCTCCCT
281 D L Y A E L C K N Y P I V S L E D P F A E D D W E T W T N N F E K L P
946 GTTCAAAATTTGTCAGATGACTTAACCGTCACCAACCCCAAGAAATCAAGACTGCTATTGAAAGAAAGGCTGCTTCTGCTTTGTTATGAAGGTCAACCAAAATT
316 V Q I V A D D L T V T N P Q R I K T A I E K K A A L L L K V N Q I
1051 GGTACCTTGTCTGAATCCATCCAGCTGCTAACGACTCCTTTGACAGAGGTTGGGTTGTCATGGTTTCCACAGATCTGGTGAACCGAAGATACCTTCATTGCA
351 G T L S E S I Q A A N D S F A A G W B V M V S H R S G E T E D T F I A
1156 GATTGACTGTGGTTTGAAGAACCGGTGAATGAAGTGTGGTCTGCTAGCTAGATCCGAAAGAAAGTGGTAAAGTACAACCAATTTAAGAAATGAAGAAGAAATTA
386 D L T V G L R T G E L K S G A P A R S E R N A K Y A L L L R I E E L
1281 GGTGACAAGGCTGTCTACGCTGGTAAGAGGTTGAGAACTGCCACGAAATATAATCTAACTAATGCTTCACTAAATCTAGTACAAATTTTACAGTCCCTAC
421 G D K A V Y A G K K F R T A H E L *
1386 GTTTATAAATGAAATTAATGAAAAAATATTTTGAACGATGTGTTTATTAAGTTGCGCTCTTCCGATAATCCCGACTTGGTTAATTTCTCAATGGGTTTT
1471 TTTTTTCAAAACCATTGTTGTAGTGTAAACAGACTTTAACAAGGACATCACTCTACAGGGCAGCTTTAAATCCCTCAGTGAATTTGTTCTTATTGATAACGT
1576 GGGAGTCAAGGAAATTTGCAAGAGTGTGTTGCTCAATAAAGACGAGAGTAATGGAAGAGTTATATTTTCTATCCCATAGGCATAACCAATGTAGCAATT
1681 TTTATATGTGCTCTTTGCTCTGAAAGGATATCACGCTGTGGGAATTTGATTGGCGGTTGATACGGCGAGATCTATTGTTCTCTTTATAGCAAACTGCATA
1788 TTATTCGTGTAAGAAATGATCAAAATTCAAAGAGGTTAAGACAAATTTACTGTAAATTTCTCTCTTCAAGACAGAAGATG

```

编码区上游序列中,单下划线部分为启动子区域,其中含框框部分(TATAAA)为酵母 TATA 框,另一个含框框部分(TCAA)是转录起始区,其上游双下划线部分为 CT 丰富区。编码区下游单下划线部分为转录终止信号,两个含框框部分为 poly A 信号。

图 1 产甘油假丝酵母烯醇化酶基因(CgENO1)序列分析

Fig. 1 Sequence analysis of *Candida glycerinogenes* Enolase gene

根据与酿酒酵母烯醇化酶基因序列比对的结果,这段序列包括一段1 314 bp的开放阅读框,所编码的氨基酸序列与酿酒酵母烯醇化酶氨基酸序列同源性达74.3%,显然是产甘油假丝酵母烯醇化酶基因(CgENO1)的结构基因。CgENO1结构基因起始翻译密码子上游第三个碱基(-3位碱基)为A,这符合酵母Kozak结构最主要的特征。CgENO结构基因+4位碱基(与ATG紧邻)为T,这与酵母Kozak结构不符。作者将来源于酵母的所有烯醇化酶基因进行多重序列比对,结果显示,所有基因的一3位碱基均为A而+4位碱基则没有明显的倾向性。对结构基因的分析显示,CgENO1基因编码含438个氨基酸的蛋白质,比酿酒酵母烯醇化酶基因少1个氨基酸,结构基因没有内含子。密码子偏好性分析显示,CgENO1结构基因密码子使用有明显的偏好性,见表1。在61个密码子中,CgENO1基因只使用了42个,而同样来源于产甘油假丝酵母的另两个基因URA3和TRP1虽然结构基因比CgENO短,但使用的密码子却明显比烯醇化酶基因多。Mason等研究表明,酵母中高表达的基因一般都有明显的密码子使用偏好性^[8]。URA3和TRP1基因分别编码酵母尿嘧啶和色氨酸合成途径的酶,表达量一般不高,而烯醇化酶是糖酵解途径的关键酶,表达量高,因此烯醇化酶有更明显的密码子偏好性。

表1 产甘油假丝酵母 ENO1、URA3、TRP1 基因密码子使用情况比较

Tab. 1 Codon usage of ENO1、URA3、TRP1 gene from *C. glycerinogenes*

基因	结构基因 全长/bp	使用密码 子数/个	文献
ENO	1 314	42	本文
URA3	789	52	[5]
TRP1	672	57	[9]

用真核生物启动子分析程序(http://www.fruitfly.org/seq_tools/promoter.html)对CgENO1结构基因上游的699个碱基分析显示,其中-165~-114间最有可能是启动子区域。该区域中-142~-137为TATA元件,与酵母TATA元件的共有序列TATAAA完全一致。酵母TATA元件和转录起始位点之间的距离往往不固定,转录起始位点一般位于TATA元件下游60~110碱基范围内。在这一范围内,-59~-56之间碱基

为TCAA,是典型的酵母转录起始位点,该位点上游17个碱基(-76~-60)中T和C占13个,符合酵母转录起始位点上游C/T丰富区的特征,因此TCAA是CgENO1基因的转录起始区。

对CgENO1结构基因终止密码子TAA下游591个碱基序列分析显示,其中1 679~1 686碱基之间序列为TTTTTATA,与典型的酵母转录终止信号完全一致,在其上游有两个区域:1385~1394、1610~1615,核苷酸序列分别为AAAAAAAAA和AATAAA,与酵母RNA腺嘌呤末端转移酶识别序列(PolyA信号)完全一致。

由上述分析可见,作者获得的CgENO1基因除具有完整的蛋白质编码区外,还包含了典型的酵母基因上下游调控区,是一条完整的基因。

2.3 产甘油假丝酵母烯醇化酶氨基酸序列与其他酵母烯醇化酶氨基酸序列比对

将由基因序列推测的产甘油假丝酵母烯醇化酶(CgENO1)氨基酸序列与在NCBI中搜索到的其它酵母烯醇化酶氨基酸序列进行多重序列比对。结果显示,不同酵母来源的烯醇化酶氨基酸序列有多个高度保守区,而作者克隆获得的CgENO1拥有所有的保守区,这充分证明克隆得到的基因确实是烯醇化酶基因。(为节约篇幅,图2只列出了CgENO1与酿酒酵母烯醇化酶、白假丝酵母烯醇化酶以及亲缘关系较远的粟酒裂殖酵母烯醇化酶2氨基酸序列多重序列比对的结果)。产甘油假丝酵母烯醇化酶基因与其它酵母烯醇化酶氨基酸序列相似性列于图2。

由表2可见,产甘油假丝酵母烯醇化酶与*Eremothecium gossypii*的烯醇化酶相似性最高,为75.9%,与其它几个酵母烯醇化酶基因的序列相似性也均在60%以上,可见烯醇化酶基因有较高的保守性,这与糖酵解途径中其它酶的情况相似。上述13种酵母烯醇化酶的亲缘关系见图3。

由图3可知,尽管产甘油假丝酵母烯醇化酶序列与*E. gossypii*和*K. lactis*的烯醇化酶基因相似性最高,但用DNAMAN绘制进化树时,产甘油假丝酵母与这两个酵母并不在同一进化簇内,这是因为*E. gossypii*和*K. lactis*的氨基酸序列与*C. glabrata*以及酿酒酵母烯醇化酶1、2的同源性很高,均在85%左右或更高,而产甘油假丝酵母与这几种酵母烯醇化酶氨基酸序列的相似性远没有这么高,因此没有与这几种酵母的分成一类。

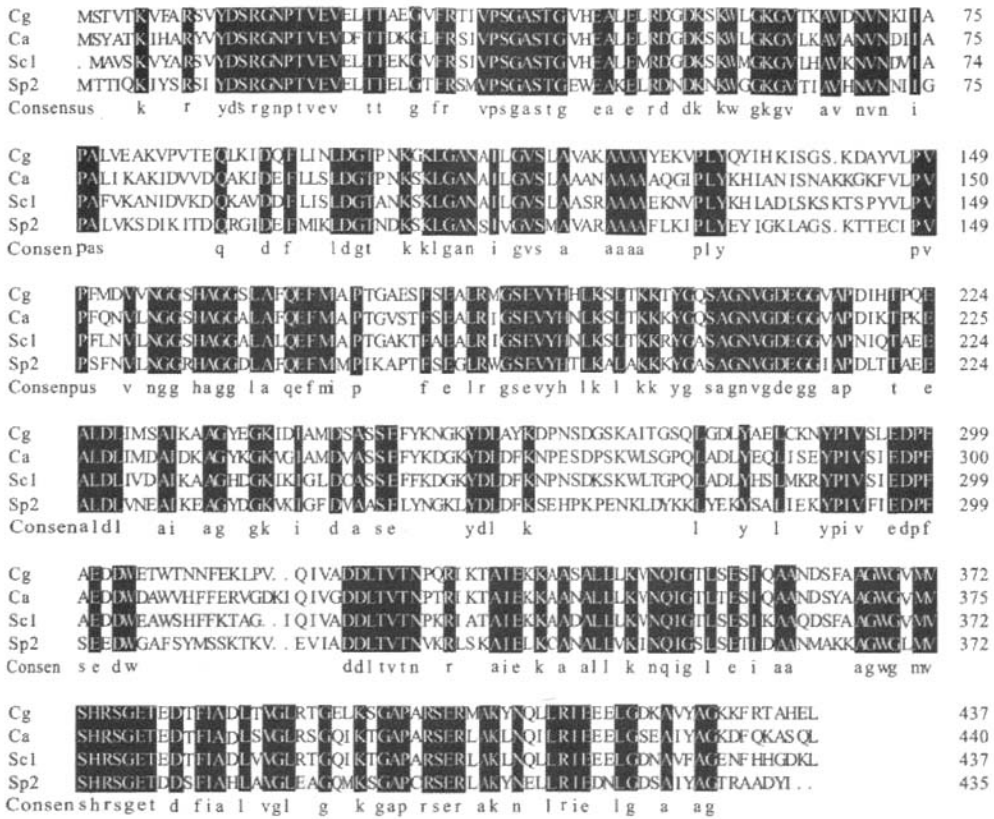


图 2 产甘油假丝酵母烯醇化酶与其它酵母烯醇化酶氨基酸序列多重比对
Fig. 2 Multiple Sequence alignment of enolases from *C. glycerinogenes* and other yeasts

表 2 产甘油假丝酵母烯醇化酶与其他酵母烯醇化酶氨基酸序列相似性
Tab. 2 Similarity of amino acid sequence between enolases from *C. glycerinogenes* and other yeasts

基因来源	Accession N.	相似性
<i>Candida glycerinogenes</i>	EF432726	100%
<i>Eremothecium gossypii</i>	Q756H2	75.9%
<i>Kluveromyces lactis</i>	Q70CP7	75.7%
<i>Candida albicans</i>	P30575	75.7%
<i>Debaryomyces hansenii</i> 2	Q6BI20	75.5%
<i>Candida glabrata</i> 1	Q6FQY4	74.5%
<i>Saccharomyces cerevisiae</i> 1	NP_011770	74.3%
<i>Saccharomyces cerevisiae</i> 2	NP_012044	72.9%
<i>Debaryomyces hansenii</i> 1	Q6BTB1	71.6%
<i>Yarrowia lipolytica</i>	Q6C1F3	68.7%
<i>Candida glabrata</i> 2	XP_446328	68.3%
<i>Shizosaccharomyces pombe</i> 1	AAA70080	63.7%
<i>Shizosaccharomyces pombe</i> 2	CAD31742	60.4%

所有氨基酸序列均来源于美国国家生物技术信息中心 (www.ncbi.nlm.nih.gov)

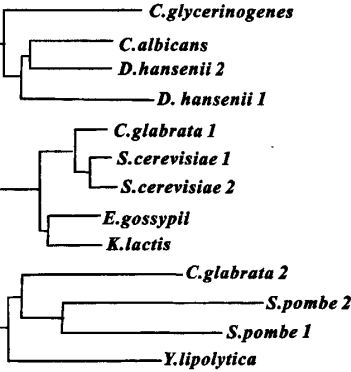


图 3 不同酵母烯醇化酶进化树
Fig. 3 Phylogenetic tree of enolase from different yeasts

2.4 产甘油假丝酵母中密码子 CTG 编码亮氨酸

研究酵母基因时,密码子 CTG 的特殊性是一个需要关注的问题,在绝大多数的生物体中密码子 CTG 编码亮氨酸。Kawaguchi Y 等在研究 *Candida cylindracea* 时首次发现密码子 CTG 在该酵母中编码丝氨酸,这一差别导致许多基因,特别是遗

传工程上常用的一些标记基因在该酵母中不能表现出功能^[9],之后其它研究者在白假丝酵母等共计近百种酵母中发现了同样的情况^[10]。产甘油假丝酵母与白假丝酵母等有较近的亲缘关系,其CTG密码子是否也编码丝氨酸是一个值得关注的问题。对CgENO结构基因分析显示,其中共有4个CTG,分别编码第107、117、197、228位氨基酸(亮氨酸或丝氨酸)。对照图2,其中第107位氨基酸KLGAN是一个在4种酵母中保守的区域,实际上将13种酵母的氨基酸序列进行多重比对时,LGAN是一个在所有酵母中完全保守的区域,可见在产甘油假丝酵母中密码子CTG在这个位置上编码亮氨酸而不是丝氨酸。进一步分析可知,第228位也是一个在13种酵母中完全保守的位点,全部都是亮氨酸。197位是一个相对保守的区域,除产甘油假丝酵母之外的12种酵母烯醇化酶中,只有粟酒裂

酵母的烯醇化酶1在这一位置是异亮氨酸,其余11种均为亮氨酸。由上述分析可见在产甘油假丝酵母中密码子CTG编码亮氨酸,与通用密码子一致。

3 结 语

从产甘油假丝酵母(*Candida glycerinogenes*)基因文库中分离了一个含烯醇化酶基因(CgENO1)的克隆子。插入片段全长2 456 bp,包含一段1 314 bp的没有内含子的蛋白质编码区,编码一个含438个氨基酸残基的多肽链,其氨基酸序列与来源于酿酒酵母的烯醇化酶相似性为74.3%。多重序列比对显示,产甘油假丝酵母烯醇化酶(CgENO1)含有酵母烯醇化酶全部保守区。对编码区上下游序列分析显示,克隆片段含有典型的酵母基因上下游调控区,是一个完整的酵母新基因。

参考文献(References):

- [1] Zhuge J, Fang H Y, Wang Z X, et al. Glycerol production by a novel osmotolerant yeast, *Candida glycerinogenes*[J]. *Appl Microbiol Biotechnol*, 2001, 55: 686-692.
- [2] Forsthoefel N R, Cushman M A F, Cushman J C. Posttranscriptional and Posttranslational control of enolase expression in the facultative crassulacean acid metabolism plant *Messembryanthemum crystallinum*[J]. *Plant Physiol*, 1995, 108: 1185-1195.
- [3] Shailesh K L, Chwenfang L, Martin M. Differential regulation of enolase during anaerobiosis in maize[J]. *Plant Physiol*, 1998, 118: 1285-1293.
- [4] Norbeck J, Blomberg. Metabolic and regulation changes associated with growth of *Saccharomyces cerevisiae* in 1.4 mol/L NaCl[J]. *J Biol Chem*, 1997, 272(9): 5544-5554.
- [5] Li Y L, Shen W, Wang Z X, et al. Isolation and sequence analysis of the gene URA3 encoding the orotidine-5-phosphate decarboxylase from *Candida glycerinogenes* WL2002-5, an industrial glycerol producer[J]. *Yeast*, 2005, 22: 423-430.
- [6] 王正祥, 诸葛健. 运用穿梭载体建立产甘油假丝酵母质粒基因文库[J]. *生物技术*, 1998, 8(2): 9-11.
WANG Zheng-xiang, ZHUGE Jian. Construction of *Candida glycerinogenes*'s genetic plasmid bank based on a shuttle vector[J]. *Biotechnology*, 1998, 8(2): 9-11. (in Chinese)
- [7] Sambrook 著. 分子克隆实验指南[M]. 金冬雁译. 北京: 科学出版社, 2002.
- [8] Mason A B, Buckley H R, Gorman J A. Molecular cloning and characterization of the *Candida albicans* enolase gene[J]. *J Bacteriol*, 1993, 175(9): 2632-2639.
- [9] 李艳丽. 产甘油假丝酵母筛选标记基因和整合位点基因的克隆[D]. 无锡: 江南大学, 2006.
- [10] Kawaguchi Y, Honda H, Taniguchi-Morimura J, et al. The codon CUG is read as serine in an asporogenic yeast *Candida cylindracea*[J]. *Nature*, 1989, 341: 164-166.
- [11] Ohama M, Suzuki T, Mori M, et al. Non-universal decoding of the leucine codon CUG in several *Candida species*[J]. *Nucl Acids Res*, 1993, 21: 4039-4045.

(责任编辑:李春丽)