

# 长梗木霉内切葡聚糖酶 I 基因的克隆及其在毕赤酵母中的表达

刘海英 王 娟 舒正玉 吴松刚 黄建忠\*

(福建师范大学工业微生物教育部工程研究中心 生命科学学院 福建 福州 350108)

**摘 要:** 一株纤维素酶高产菌株经 ITS 序列鉴定并命名为长梗木霉 SSL (*Trichoderma longibrachiatum*, SSL)。利用 RT-PCR 的方法从该菌株中克隆出内切-1,4- $\beta$ -D-葡聚糖酶 I 的基因 (*egl1*)，该基因全长 1386 bp，编码 461 个氨基酸。序列分析表明：该基因序列与 *T. longibrachiatum egl1* 基因具有 90% 以上的同源性。将该基因的成熟肽编码序列插入到 *Pichia pastoris* 表达载体 pPIC9k 中，构建重组表达质粒 pPIC9k-*egl1*，转化 *P. pastoris* GS115。重组 *P. pastoris* 菌株，经甲醇诱导后，胞外重组内切葡聚糖酶 I 的活力达 73 U/mL。SDS-PAGE 中出现一条明显加强的蛋白质条带，其分子量大约为 58 kD。

**关键词:** 分子标记，内切葡聚糖酶 I，毕赤酵母，表达

## Cloning of Endoglucanase I Gene from *Trichoderma longibrachiatum* and Its Expression in *Pichia pastoris*

LIU Hai-Ying WANG Juan SHU Zheng-Yu WU Song-Gang  
HUANG Jian-Zhong\*

(Fujian Normal University, Engineering Research Center of Industrial Microbiology, Ministry of Education,  
The College of Life Science, Fuzhou, Fujian 350108, China)

**Abstract:** A cellulase high-yield strain was identified and named as *Trichoderma longibrachiatum* SSL by ITS sequence identification. The *endoglucanase1* gene (*egl1*) encoding endo-1,4- $\beta$ -D-glucanase I was amplified by RT-PCR method, which including 1386 bp and encoding 461 amino acid. Sequence analysis showed that: This gene has a more 90% homology with the *T. longibrachiatum egl1* gene. The *egl1* gene encoding the mature peptide was inserted into the *Pichia pastoris* expression vector pPIC9K, which resulted in construction of the recombinant expression plasmid, pPIC9k-*egl1*. The pPIC9k-*egl1* was then introduced into the host *Pichia pastoris* GS115. After the induction of methanol, extracellular recombinant endoglucanase I from the supernatant of the recombinant *Pichia pastoris* strain reached 73 U/mL. A clear strengthening of the protein bands, whose molecular weight is about 58 kD, appeared in the SDS-PAGE.

**Keywords:** Molecular marker, Cellulase endoglucanase I, *Pichia pastoris*, Expression

基金项目：福建省发改委重大项目[No. (2004) 477]；福建省科技厅重大项目(No. 2005Q007)

\* 通讯作者：Tel: 86-591-22868212; Fax: 86-591-22868195; ✉: hjz@fjnu.edu.cn

收稿日期：2008-09-25；接受日期：2008-12-09

纤维素是一种由多个葡萄糖单位通过  $\beta$ -1,4-糖苷键相连而形成的植物多糖, 占生物圈内生物量的一半, 是一种可再生资源。纤维素再循环利用的过程有希望用于解决能源危机、粮食危机等问题<sup>[1]</sup>。在纤维素再循环利用的自然过程中, 最重要的一步就是通过纤维素酶的催化活性将固态的纤维素降解为可溶性的葡萄糖。而纤维素酶是一个多组分酶系, 包括内切葡聚糖酶(Endoglucanase, EG)、纤维二糖水解酶(Cellobiohydrolase, CBH)和  $\beta$ -葡萄糖苷酶( $\beta$ -glucosidases)等<sup>[2]</sup>。

纤维素酶的不同组分具有不同的酶学性质及催化活性, 而不同应用领域对纤维素酶各组分的需求也不相同。如纺织行业, 在布料的生物抛光、生物打磨及水洗过程中, 主要利用纤维素酶 EG 组分的催化活性<sup>[3,4]</sup>。因此, 为了强化纤维素酶复合酶系中某一组分的功能, 研究者通常利用其他表达体系对其组分进行单独表达, 然后对各组分进行不同比例的复配, 以适应不同应用领域的需求<sup>[5]</sup>。我们课题组从西双版纳原始森林中筛选获得了一株高产纤维素酶的长梗木霉 SSL, 羧甲基纤维素钠水解活力为 2114 U/mL。本研究探索了长梗木霉 SSL 纤维素酶组分中的 *egl* 基因在毕赤酵母中功能性表达的可行性。

## 1 材料与方法

### 1.1 菌株和质粒

纤维素酶高产菌株由本课题组从西双版纳原始森林土壤中分离并保存; *Escherichia coli* DH5 $\alpha$ 、*Pichia pastoris* GS115 及质粒 pPIC9k 由本实验室保存; pMD18-T Vector 购于 Sangon 有限公司(上海)。

### 1.2 试剂及培养基

T4 DNA 连接酶、各种限制性内切酶、DNA 聚合酶均购于 TaKaRa 公司(大连); DNA 纯化试剂盒购于 Sangon 有限公司(上海)。Mendels 培养基参照 Kubicek(1993)<sup>[6]</sup>, PDA 培养基、LB 培养基、YPD 培养基、MD 培养基、MM 培养基、BMGY 培养基、BMMY 培养基的配制均按照 Invitrogen 公司毕赤酵母操作手册进行。

### 1.3 菌种的分子学鉴定

CTAB 法提取菌株的基因组 DNA<sup>[7]</sup>。以基因组 DNA 为模板, ITS1(5'-TCCGTAGGTGAACCTGCG G-3')和 ITS4(5'-TCCTCCGCTTATTGATATGC-3')为

引物, 进行 PCR 扩增。PCR 扩增产物 T-Vector 克隆、测序。ITS 序列聚类分析, 构建进化树。

### 1.4 *egl* 基因的克隆

Mendels 培养基培养菌丝体, 收集 48h 的菌丝体提取总 RNA。RT-PCR 扩增 *egl* 基因的全长编码序列和成熟肽编码序列。扩增全长基因和成熟肽编码基因的引物分别为 F1/R1 和 F2/R2(下划线序列为引入的限制性内切酶识别位点):

F<sub>1</sub>: 5'-CGGAATTCATGGCGCCCTCAGTTACA CTGC-3'

R<sub>1</sub>: 5'-ATGCGGCCGCCTAAAGGCATTGCGAG TAGTCT-3'

F<sub>2</sub>: 5'-CGGAATTCCAGCAACCGGGCACCAG C-3'

R<sub>2</sub>: 5'-ATGCGGCCGCCTAAAGGCATTGCGAG TAGTCT-3'

PCR 扩增条件为: 94°C 5 min; 94°C 30 s, 53°C 30 s, 72°C 90 s, 35 个循环; 72°C 10 min。PCR 扩增产物 T-Vector 克隆、测序。

### 1.5 表达载体的构建、转化及转化子的诱导表达

将 *egl* 基因的成熟肽编码序列的 PCR 扩增产物用 *Eco*R I 和 *Not* I 进行双酶切, 与经过相同双酶切的 pPIC9k 载体连接, 构建重组表达质粒 pPIC9k-*egl* 并进行 5'端测序验证酶连正确。按照毕赤酵母表达系统操作手册, 转化 *P. pastoris* GS115。(pPIC9k 空载体转化 *P. pastoris* GS115 作为对照)使用 MD 培养基和 MM 培养基对转化子进行初筛, 通过 *egl* 基因的 PCR 扩增进行复筛。

复筛得到的转化子在 BMGY 培养基中培养至 OD<sub>600</sub>=2.0, 离心收集菌体将其转入 BMMY 培养基中, 30°C、250 r/min 甲醇诱导 *egl* 基因的表达<sup>[8]</sup>。

### 1.6 重组表达 EG I 定性与定量检测

**1.6.1 刚果红平板定性检测:** 刚果红定性检测平板的组成成分为: 0.5%(W/V)羧甲基纤维素钠(CMC-Na), 1.8%(W/V)琼脂糖, pH 4.8 的醋酸-醋酸钠缓冲液。取 100 $\times$ 浓缩的发酵液 100  $\mu$ L 加入到定性检测平板上的空穴中, 50°C 反应 1 h。0.1% 刚果红染液对反应后平板染色 1 h, 然后再用 1 mol/L NaCl 脱色 40 min<sup>[9]</sup>, 观察是否出现水解圈。

**1.6.2 3,5-二硝基水杨酸比色法定量检测:** 0.9 mL 1.5%(W/V)羧甲基纤维素钠作为底物, 50°C 温浴 3 min 后, 加入 0.1 mL 发酵液, 混匀, 50°C 反应 30 min。加入 1.5 mL 3, 5-二硝基水杨酸试剂, 沸水

水浴 10 min 后加水定容至 7.5 mL, 测定 540 nm 波长下的光吸收值。一个酶活单位(U)定义为: 每分钟水解 CMC-Na 产生 1  $\mu$ g 还原糖(葡萄糖)所需的酶量。

**1.6.3 SDS-PAGE 分析:** 取浓缩后的发酵液, 按照蛋白质技术手册进行 SDS-PAGE, 快速银染法染色。

## 2 结果与分析

### 2.1 纤维素酶高产菌株的分子鉴定

ITS 序列 PCR 扩增产物全长 637 bp(GenBank Accession No: EU935218)。ITS 序列的聚类分析结果如图 1。纤维素酶高产菌株 SSL 与 *T. longibrachiatum* 处于同一分枝, 亲缘关系最近。因此该菌株初步鉴定并命名为 *T. longibrachiatum* SSL。

### 2.2 *egl* 基因的克隆

*egl* 基因全长 1386 bp(GenBank Accession No: EU935217)。Blast 分析结果表明, 该基因序列与 *T. longibrachiatum egl* 基因序列(X60652)有 90%以上的同源性。

*egl* 基因编码的氨基酸序列与其他同源蛋白的

氨基酸序列进行比对, 结果如图 2。从比对结果可以看出: 该 *egl* 的氨基酸序列和其他同源蛋白 *egl* 的氨基酸序列具有高度的保守性, 特别是与 *T. longibrachiatum* 的 *egl* 有高达 97% 的同源性。其中 1~22 位氨基酸为信号肽序列; 23~393 位氨基酸构成酶的催化区域; 424~461 位氨基酸构成酶的底物结合区域; E<sup>218</sup>、D<sup>220</sup> 和 E<sup>223</sup> 三个氨基酸残基构成酶的活性中心。同时, 该蛋白序列中含有五个潜在的 N-糖基化位点, 分别为 N<sup>78</sup>TTL、N<sup>164</sup>GSL、N<sup>204</sup>GTL、N<sup>208</sup>TSH 和 N<sup>394</sup>STG<sup>[10-12]</sup>。

### 2.3 重组表达载体的构建

将 *egl* 基因的成熟肽编码序列的 PCR 扩增产物用 *Eco*R I 和 *Not* I 进行双酶切, 与经过相同双酶切的 pPIC9k 载体连接, 构建重组表达质粒 pPIC9k-*egl*。双酶切分析(图 3)和 5'端测序验证(序列未列出)表明重组表达质粒构建成功。

### 2.4 *egl* 基因的重组表达

转化子不同发酵时间的发酵液 CMC 定量检测结果如图 4。144 h 时, 发酵液中重组 *egl* 活力最高, 达到 73 U/mL。重组转化子所需诱导时间明显长于

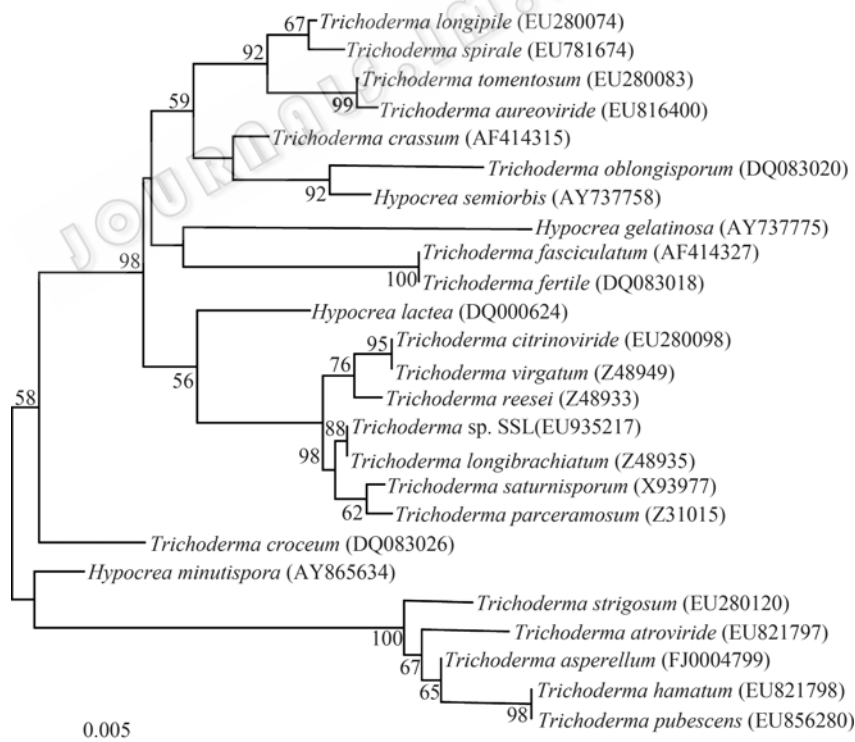


图 1 纤维素酶高产菌株 SSL ITS 序列的聚类分析结果

Fig. 1 Cluster analysis of ITS sequence of cellulase high-yield strain SSL

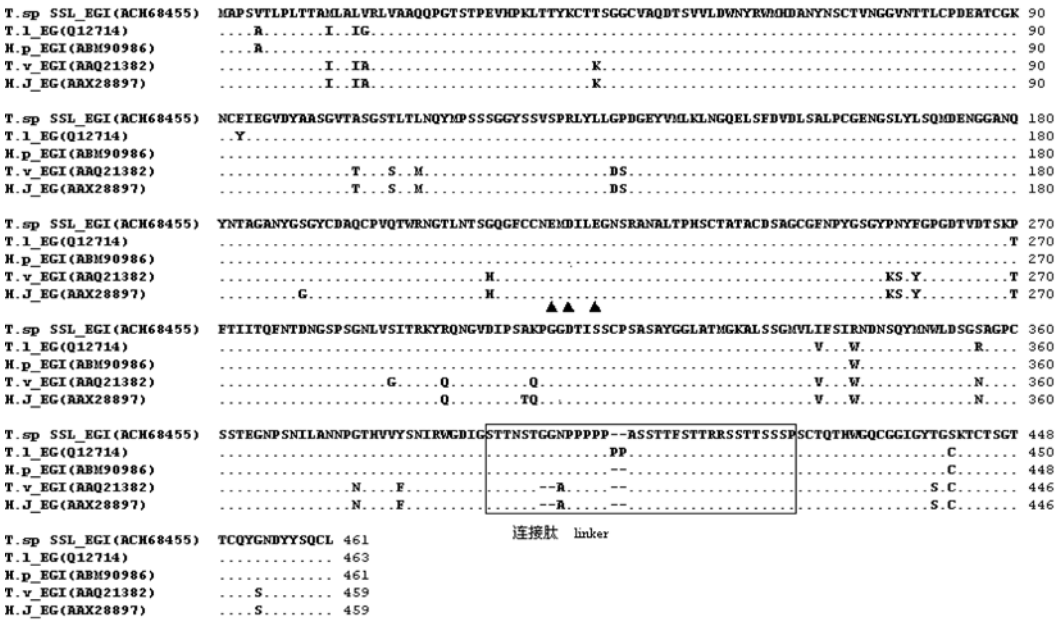


图 2 *T. longibrachiatum* SSL EGI 的氨基酸序列同源性比较

Fig. 2 Homologous alignment of the amino acid sequence of the *T. longibrachiatum* SSL EGI

注：▲：催化活性中心；□：连接肽。  
Note: ▲：Catalytic active site; □：Linking peptide.

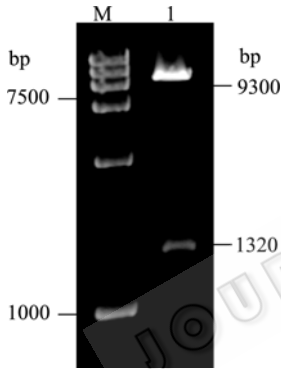


图 3 重组表达质粒 pPIC9k-egI 的双酶切验证  
Fig. 3 Identification of recombinant expression plamid pPIC9k-egI by double digestion  
M: DL15000 bp marker; 1: pPIC9k-egI digested with *EcoR* I and *Not* I.

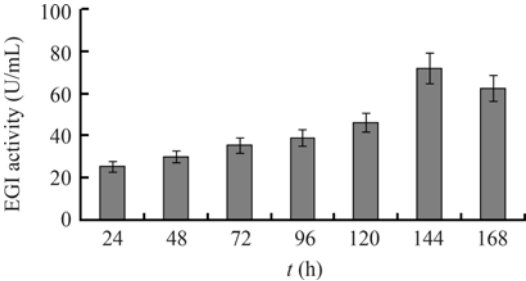


图 4 转化子重组菌株诱导产酶情况  
Fig. 4 Status of EGI production of recombinants

乔宇等构建的表达 EGII 的转化子 Gp2025<sup>[13]</sup>，可能与得到的重组菌株是甲醇利用缓慢型有关<sup>[14]</sup>。浓缩发酵液的刚果红平板定性检测结果如图 5。刚果红平板上可以看到清晰的水解圈。重组 EGI 的定量和定性分析结果均表明 *egI* 基因在 *P. pastoris* 中实现了分泌型的功能表达。

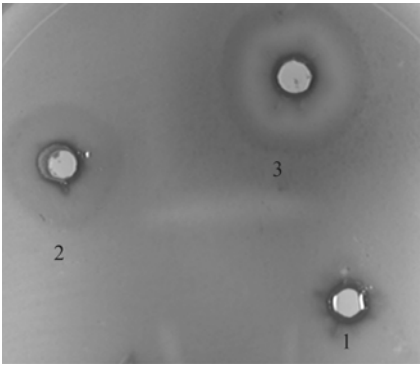


图 5 重组内切葡聚糖酶 I 的刚果红平板定性检测  
Fig. 5 Qualitative detection of recombinant EGI on agar medium  
1: Control; 2: Transformants; 3: *T. longibrachiatum* SSL.

浓缩发酵液的 SDS-PAGE 结果如图 6。与对照相比，重组转化子的发酵液中含有一条明显加强的蛋白质条带，其分子量大约为 58 kD，较根据氨基酸序列预测的分子量(48 kD)稍大，这可能与重组蛋白的糖基化有关<sup>[15]</sup>。

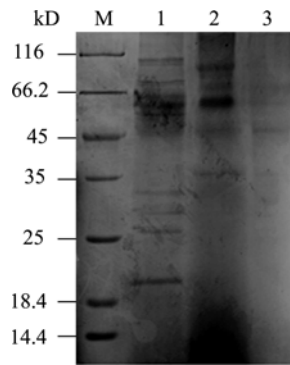


图 6 重组 EGI 的 SDS-PAGE 电泳分析

Fig. 6 Electrophoretic analysis of the recombinant EGI by SDS-PAGE

M: Protein molecular weight marker; 1: Supernatant of *T. longibrachiatum* SSL; 2: Supernatant of recombinants; 3: Supernatant of control.

### 3 讨论

国外长梗木霉内切葡聚糖酶 I 的研究主要集中在该基因的克隆与表达及酶学性质等方面, 1992 年 Gonzalez R 克隆了一个长梗木霉的 *eglI* 基因, 并在在酿酒酵母中实现了表达, 但没有活性的报道。而国内对长梗木霉的研究主要集中于产纤维素酶条件的优化、菌株的诱变选育, 而有关该菌种的纤维素酶基因克隆与表达方面的研究尚未见任何报道。本文首次从长梗木霉中克隆到 *eglI* 基因, 并在 *P. pastoris* 中实现了功能性表达。

本文虽然实现了 *eglI* 基因在 *P. pastoris* 中的分泌性功能表达, 但表达水平比较低, 发酵液酶活仅 73 U/mL, 可能与毕赤酵母自身产生的蛋白水解酶有关, 或者表达系统对外源蛋白过度糖基化, N-多糖阻遏了酶与底物的结合, 进而造成重组蛋白的活性降低。在后续研究中, 我们将通过构建蛋白水解酶基因缺失的表达系统或对 N-糖基化位点突变进一步提高 *eglI* 基因在 *P. pastoris* 中的表达量。

### 参 考 文 献

- [1] Wood BE, Ingram LO. Ethanol production from cellobiose, amorphous cellulose, and crystalline cellulose by recombinant *Klebsiella oxytoca* containing chromosomally integrated *zymomonas mobilis* genes for ethanol production and plasmids expressing thermostable cellulase genes from *Clostridium thermocellum*. *Applied Environmental Microbiology*, 1992, **58**: 2103–2110.
- [2] Penttila ME, Andre L, Saloheimo M, *et al.* Expression of two *Trichoderma reesei* endoglucanase in the yeast *saccharomyces cereviae*. *Yeast*, 1987, **3**: 75–85.
- [3] Heikinheimo L, Buchert J, Miettinen-Oinonen A, *et al.* Treating denim fabrics with *Trichoderma reesei* cellulase. *Textile Research Journal*, 2000, **70**: 969–973.
- [4] 张海生, 张同亮, 陈德兆. 纤维素酶对纤维素的作用机理及其在纺织上的应用. *四川纺织科技*, 2003, **6**: 3–5.
- [5] 汤 新, 刘 刚, 田生礼, 等. 里氏木霉内切葡聚糖苷酶 IV 在毕赤酵母中的表达. *微生物学通报*, 2005, **32**(6): 47–51.
- [6] Kubicek CP, Messner R, Gruber F, *et al.* Triggering of cellulase biosynthesis by cellulose in *Trichoderma reesei*. Involvement of a constitutive, sophorose-inducible, glucose-inhibited beta-diglucoside permease. *Journal of Biological Chemistry*, 1993, **268**(26): 19364–19368.
- [7] 熊 涛, 肖 满, 曾哲灵, 等. 松乳菇组织分离菌株的 rDNA ITS 序列分子鉴定. *微生物学通报*, 2006, **33**(4): 1–4.
- [8] 乔 宇. 里氏木霉纤维素酶基因的克隆和表达的研究. 北京师范大学硕士学位论文, 2004.
- [9] Teacher RM, Wood PJ. Use of congored-polysaccharide interactions in enumeration and characterization of cellulolytic bacteria from the bovine rumen. *Applied Environmental Microbiology*, 1982, **43**: 777–780.
- [10] Il Kwon, KeisukeE, MasatoshiG, *et al.* Heterologous expression and characterization of endoglucanase I(EGI) from *Trichoderma viride* HK-75. *Bioscience Biotechnology Biochemistry*, 1999, **63**(10): 1714–1720.
- [11] Kleywegt GJ, Zou JY, Divne C, *et al.* The crystal structure of the catalytic core domain of endoglucanase I from *Trichoderma reesei* at 3.6Å resolution, and a comparison with related enzymes. *Journal Molecular Biology*, 1997, **272**: 383–397.
- [12] Gonzalez R, Ramon D, Perez-Gonzalez JA. Cloning, sequence analysis and yeast expression of the *eglI* gene from *Trichoderma longibrachiatum*. *Applied Environmental Microbiology*, 1992, **38**: 370–375.
- [13] 乔 宇, 毛爱军, 何永志, 等. 里氏木霉内切-β-葡聚糖酶 II 基因在毕赤酵母中的表达及酶学性质研究. *菌物学报*, 2004, **23**(3): 388–396.
- [14] Sue MP, Marianal LF, Brian MN, *et al.* Heterologous protein production using the *Pichia pastoris* expression system. *Yeast*, 2005, **22**: 249–270.
- [15] 顾 圆, 诸欣平, 王少华. 毕赤酵母表达蛋白质的糖基化. *生命的化学*, 2004, **24**: 353–355.