

doi:10.3969/j.issn.1672-5565.2013.02.06

绿色木霉内切几丁质酶 cDNA 的克隆 及其编码蛋白的序列分析

于平*, 杨卫芳, 章慧慧

(浙江工商大学食品与生物工程学院, 浙江 杭州 310035)

摘要:根据已克隆的内切几丁质酶基因序列的同源性比较,设计引物,采用 PCR 技术从绿色木霉基因组中分离出一个大小为 1 467bp 的特异 DNA 片段,采用 RT-PCR 技术从绿色木霉总 RNA 中分离出大小约 1 276bp 的 cDNA 片段。序列对比后发现该内切几丁质酶 DNA 含有三个内含子,大小分别为 52bp, 69bp, 64bp。同源性分析表明其全长 cDNA 序列和已经报道的内切几丁质酶序列的同源性高达 95% 以上,预测其编码蛋白的氨基酸序列含 424 个氨基酸残基,分子量为 46 kDa,氨基酸序列分析表明该内切几丁质酶 164~172 位氨基酸是其活性中心,用同源建模法模拟其空间结构模型,为进一步研究其作用机制奠定了良好基础。

关键词:绿色木霉;内切几丁质酶;序列分析

中图分类号:Q518.2 文献标识码:A 文章编号:1672-5565(2013)-02-109-06

cDNA cloning of the endochitinase from *Trichoderma viride* and the sequence analysis of its encoding protein

YU Ping*, YANG Wei-fang, ZHANG Hui-hui

(College of Food Science and Biotechnology, Zhejiang Gongshang University, Hangzhou 310035, China)

Abstract: DNA of 1467bp and cDNA of 1 276 of the endochitinase from *T. viride* were amplified by PCR and RT-PCR using the genome and the total RNA as templates, respectively. The DNA sequence includes three introns with lengths of 52, 69 and 64 bp. The homology analysis shows that the cDNA sequence of the endochitinase from *T. viride* has a high homology of 95% with other ones. Deduced amino acids have 424 residues and the theoretical molecular weight is 46kDa. The amino acid sequence analysis shows that its active sites are 164~172 residues. The tertiary structure of the mature endochitinase protein was simulated with homology modeling method, which lays a good foundation for further studying its catalytic mechanism.

Key words: *T. viride*; Endochitinase; Sequence Analysis

几丁质酶(EC3.2.1.14) 是一种降解几丁质的糖苷酶,普遍认为它与植物体内防御系统有关^[1]。高等植物本身不含几丁质,但当其受到真菌、细菌或病毒感染时,植物几丁质酶活性迅速提高,通过抑制病原真菌的孢子萌发和菌丝生长、破坏细胞新物质的积累,使植物获得系统性抗性^[2]。因此几丁质酶在植物保护尤其在抗真菌病害中的重要作用,越来越受到研究者的关注,已经成为抗真菌病害的研究热点之一。目前,国内外研究人员已经从多种微生物中克隆得到了几丁质酶基因,包括放线菌、细菌、真菌、酵母及病毒等,其中以克隆 *Serratia marcescens* 和 *Trichoderma harzianum* 中的几丁质酶基因研究最多^[3-4]。绿色木霉是一种丝状真菌,在以胶态几丁质为唯一碳源诱导培养时能分泌几丁质酶,并产生该酶 mRNA。本研究以胶态几丁质为唯一碳源诱导培养绿色木霉,通过 PCR 和 RT-PCR 技术克隆了内切几丁质酶的 DNA 和 cDNA,对 cDNA 编码的蛋白从氨基酸组成、理化性质、亲/疏水性及三级结构等

物中克隆得到了几丁质酶基因,包括放线菌、细菌、真菌、酵母及病毒等,其中以克隆 *Serratia marcescens* 和 *Trichoderma harzianum* 中的几丁质酶基因研究最多^[3-4]。绿色木霉是一种丝状真菌,在以胶态几丁质为唯一碳源诱导培养时能分泌几丁质酶,并产生该酶 mRNA。本研究以胶态几丁质为唯一碳源诱导培养绿色木霉,通过 PCR 和 RT-PCR 技术克隆了内切几丁质酶的 DNA 和 cDNA,对 cDNA 编码的蛋白从氨基酸组成、理化性质、亲/疏水性及三级结构等

收稿日期:2012-02-26;修回日期:2012-03-07.

基金项目:浙江省自然科学基金(M303087),浙江省科学技术厅项目(2006C23073)和杭州市科技局项目(20061133B26)资助。

* 通讯作者:于平,男,博士,教授。Tel: 0571-88071024-8588; E-mail: yup9202@yahoo.com.cn.

方面进行预测和分析,为其后续高效表达该酶奠定了良好基础。

1 材料与方法

1.1 菌株与培养基

菌株绿色木霉(*T. viride* ACCC 30552)购自中国中国农业微生物菌种中心;马铃薯培养基(PDA):马铃薯 200g,蔗糖 20g,琼脂 20g,蒸馏水 1 000mL,pH 自然。把马铃薯去皮后,切成 1cm^3 的小块煮沸 30min,然后用纱布过滤,再加蔗糖和琼脂,并补足水至 1 000mL,加热使琼脂溶化后,分装,用于绿色木霉培养^[5]。液体种子培养基:葡萄糖 10g,蛋白胨 1g,营养盐 50mL,微量元素 1mL,加蒸馏水至 1 000mL,用 1M 柠檬酸调节 pH 至 4.8。营养盐溶液组成(g/L)(见 1): $(\text{NH}_4)_2\text{SO}_4$ 14, KH_2PO_4 20,尿素 3, $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ 3, $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ 3。微量元素盐溶液组成(g/L)(见 2): $\text{MnSO}_4 \cdot \text{H}_2\text{O}$ 1.6, $\text{ZnSO}_4 \cdot 2\text{H}_2\text{O}$ 1.4, $\text{CoCl}_2 \cdot 6\text{H}_2\text{O}$ 3.7。产酶培养基:蛋白胨 1g,吐温 80 1mL,0.75% 胶态几丁质,加蒸馏水至 1 000mL。用 1M 柠檬酸调节 pH 至 4.8。营养盐的量按所含氮源量以 C/N 比 8:1 加入。营养盐和微量元素盐组成同上,用于绿色木霉发酵培养^[6]。以上培养基于 $1.05\text{kg}/\text{cm}^2$, 121°C 条件下灭菌 20min。

1.2 实验方法

1.2.1 绿色木霉的培养方法

将购买的绿色木霉的冻干粉用少量无菌水溶解后,均匀涂布在 PDA 平板上, 28°C 培养 5~7 天。在 250 mL 三角瓶中,装入 25mL 液体种子培养基,从长满绿色木霉的 PDA 平板上接入一环绿色木霉的菌种, 28°C ,150rpm 振荡培养 2d 作为种子。在 250mL 三角瓶中,装入 50mL 产酶培养基,5% (v/v) 的接种量接入种子, 28°C ,150rpm 振荡培养 1~5 天后收集菌丝,用于基因组和总 RNA 的提取材料。绿色木霉基因组 DNA 的提取参照王源超等的真菌基因组提取方法^[7];绿色木霉总 RNA 的提取采用上海生工生物工程技术服务有限公司的 UNIQ-10 柱式 Trizol 总 RNA 抽提试剂盒。PCR 和 RT-PCR 的引物为:P1:5'-TCATG TTG GGCTTCCTCGGAAAATCC-3' 和 P2:5'-CTAGTTGAGACCGCTTCGGATGTT-3'。PCR 按常规实验方法进行。RT-PCR 按照上海生工生物工程有限公司提供的 RT-PCR 试剂盒进行。DNA 和 cDNA 序列测定由上海生工生物工程有限公司完成。

1.2.2 序列分析

测序后将 DNA 和 cDNA 测序结果利用 DAN-MAN 进行比对,得到内含子序列。通过 NCBI 网站(<http://www.ncbi.nlm.nih.gov>)上在线 BLAST 软件分析核苷酸的同源性。用软件 DANMAN 进行同源性分析、一级结构、二级结构、亲疏水性预测。用在线软件 SMRT (<http://smart.embl-heidelberg.de/>)预测蛋白质功能区,用在线软件 Prosite (<http://kr.expasy.org/prosite/>)对蛋白质功能位点进行分析。利用 ExPASy 服务器的 SWISS-MODEL(<http://swissmodel.expasy.org/>) Version 3.6 程序进行。建模过程分两步:首先搜索 SWISS-MODEL 的晶体图像数据库 EXPDB,寻找合适的蛋白晶体结构数据做模板;然后,将序列送交 SWISS-MODEL 程序,服务器将自动返回同源建模和修正、优化的结果。

2 结果与分析

2.1 绿色木霉基因组 DNA 和总 RNA 的提取

一般真核细胞基因组 DNA 常是采用在 EDTA 以及 SDS 等试剂存在下用蛋白酶 K 消化细胞,随后用酚抽提而实现的。对于提取真菌的基因组 DNA 的几个常用方法,如 CTAB 法等,进行了比较后,发现王源超等的真菌基因组提取方法,效果最好,得到的 DNA 质量较高。如图 1(A)所示,DNA 条带清晰完整,没有降解。而且 RNA 去除的也很干净,可以用作进行 PCR 反应的模板。木霉的内切几丁质酶为诱导型酶。Carolina Carsolio 等报道木霉在以几丁质为唯一碳源的培养基中生长时,其内切几丁质酶 mRNA 量在 24h 达到最高而之后逐渐消失^[8],所以设计试验在绿色木霉在产酶培养基中生长 6h、12h、18h、24h、36h 及 48h 时分别取样提取总 RNA。绿色木霉的总 RNA 提取采用上海生工生物工程技术服务有限公司的 UNIQ-10 柱式 Trizol 总 RNA 抽提试剂盒。该试剂盒适用于动物细胞及组织、细菌、酵母、植物组织以及丝状真菌,因此提取的总 RNA 效果较好。结果如图 1(B),分别为诱导培养 6h、12h、18h、24h、36h 及 48h 时的绿色木霉总 RNA,可以看到没有基因组 DNA 污染,而且 28S 和 18S 条带整齐,亮度 $>1:1$,说明总 RNA 完整性很好,而且纯度较高,可用来作为 RT-PCR 反应的模板。

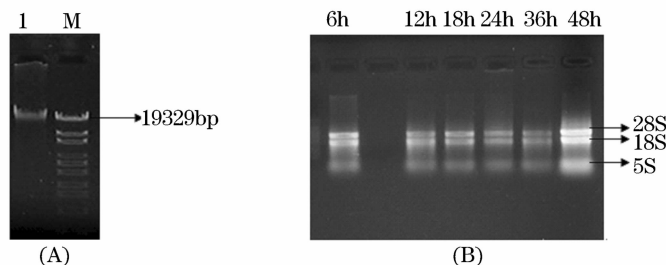


图 1 绿色木霉基因组 DNA(A) 和总 RNA(B) 示意图

Fig. 1 The schematic maps of the genomic DNA(A) and the total RNA(B) from *T. viride* ACCC 30552

2.2 PCR 和 RT-PCR

PCR 扩增的内切几丁质酶的 DNA 片段如图 2 (A) 所示,其中, M 代表 DNA Marker DL2000; 1 代表内切几丁质酶基因;大小约 1.5kb,与预期大小相符。RT-PCR 采用 TaKaRa 公司的 TaKaRa RNA PCR Kit(AMV),分为反转录和 PCR 两步。反转录反应体系为 10 μ L,模板为各个时间段的总 RNA,分别以 random 9 mers 和特异性下游引物为引物,反应条件参照相应试剂盒说明书。PCR 反应部分以 RT 产物为模板再按说明书加入相应的反应液,退火温度为 58 $^{\circ}$ C,得到了该菌编码的内切几丁质酶的 cDNA 片

段。如图 2 (B) 所示,其中, M 代表 DNA Marker DL2000; 1~6 为 RT 部分以 random 9 mers 为引物,模板分别为 6h、12h、18h、24h、36h 及 48h 的总 RNA;7~12 为 RT 部分以特异性下游引物为引物,模板分别为 6h、12h、18h、24h、36h 及 48h 的总 RNA。样品 5,6,11,12 有与预期大小相符的条带,约 1.2kb,说明内切几丁质酶 mRNA 在诱导 36h 和 48h 时存在。由图亦可知以特异下游引物作为 RT 部分时的引物时扩增效果较好。回收 12 号管对应的与预期大小相符的条带进行测序。

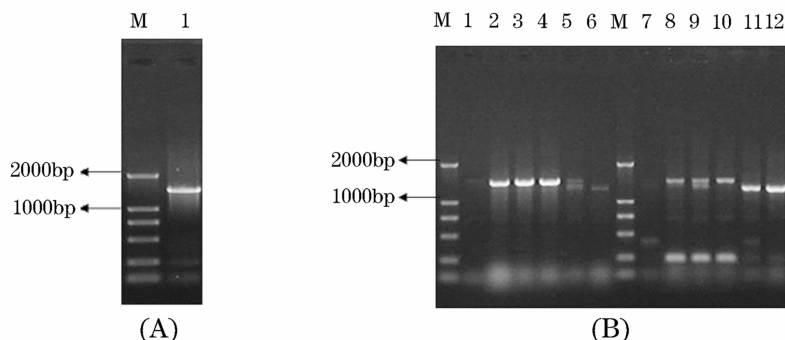


图 2 绿色木霉内切几丁质酶基因 PCR 和 RT-PCR 电泳结果

Fig. 2 PCR and RT-PCR results of endochitinase from *T. viride* ACCC30552

2.3 DNA 和 cDNA 序列分析

用 DNAMAN 软件对绿色木霉内切几丁质酶 DNA 和 cDNA 测序结果进行分析,发现 DNA 大小为 1 467bp,cDNA 大小为 1 276bp,经序列比对发现,绿色木霉内切几丁质酶 DNA 含有三个内含子,大小分别为 52bp、69bp、64bp。内含子的功能目前尚不清楚,只能推测出在形成 mRNA 过程中,RNA 前体除掉了全部内含子拼接成成熟的 mRNA。通过 NCBI 网站(<http://www.ncbi.nlm.nih.gov>)上在线 BLAST 软件分析,发现绿色木霉内切几丁质酶基因的 cDNA 序列与其他木霉的内切几丁质酶基因有很高的相似性。其中与 *T. harzianum* (IMI 206040) (编号为 X79381.1)、*T. viride* (编号为 AF188924) 的同源

性高达 100%,与 *T. sp.* w512 (编号 DQ462415.1)、*T. viride* (编号为 EF635427.1, AF208842.1)、*T. aureoviride* (编号为 AY850032.1)、*T. hamatum* (编号为 U88560.1, U49455.1) 同源性高达 95% 以上。

2.4 蛋白质一级结构分析

利用软件 DNAMAN 翻译开放阅读框的密码子,得出该基因编码一个含有 424 个氨基酸的多肽,氨基酸序列见图 3。用在线软件 SMRT(<http://smart.embl-heidelberg.de/>) 进行氨基酸序列分析,推导出前体蛋白质一级结构功能区域包括三个区域:1~22 为信号肽(signal peptide)区域,23~35 是前导(pro region)区域,36~423 是成熟的氨基酸。

signal peptide
pro region

```

[MLGFLGKSVALLAALQATFTSA]SPVTANDVSVSEKRASGYANAVYFTNWGIYGRNFQPQNLVASDITHVIYSFMNF
QADGTVWVGDAYADYQKHYYDDSWNDVGNNAAYGCVKQLFKLKKANRNLKVMLSIGGWTWSTNFPSAASDAN
RKNFAKTAITFMKDWGFDGIDVDWEYPADDTQATNMVLLLKEIRSQLDAYAAQYAPGYHFLLSIAAPAGPEHYSFL
HMSDLGQQLDYVNLMAVDYAGSWSSYSGHDANLFANPSNSNSSPYNTDQAIKDYIKGGVPASKIVLGMPIYGRSF
ESTGGIGQTYSGIGSGSWENGIWDYKVLPKAGATVQYDSTAQAYSYDPSKELISFDTPAMINTKVSYLKNLGLGG
SMFWEASADKTGSDSLIGTSHRALGSLDSTQNLGYPNSQYDNIRSLN
    
```

图 3 绿色木霉内切几丁质酶的氨基酸序列

Fig. 3 The amino acid sequence of *T. viride* endochitinase

经分析,含信号肽的绿色木霉内切几丁质酶的分子量约为 46kDa,等电点 $pI = 4.95$,属于酸性蛋白。氨基酸组成成分如表 1 所示,绿色木霉内切几丁质酶肽链中富含 Ala10.14% 和 Ser10.85%,疏水性氨基酸 33.73%,酸性氨基酸 9.2%,碱性氨基酸 6.84%。除去信号肽而形成的成熟的蛋白质分子量约为 42kDa,等电点 $pI = 4.83$ 。

表 1 推测氨基酸序列中的氨基酸组成

Table1 Amino acid composition of the deduced amino acid sequence

Amino acid	Num	% mol/mol	% w/w	
A	Ala	43	10.14	7.11
C	Cys	1	0.24	0.22
D	Asp	31	7.31	7.66
E	Glu	8	1.89	2.19
F	Phe	17	4.01	5.21
G	Gly	38	8.96	5.30
H	His	7	1.65	2.02
I	Ile	19	4.48	4.62
K	Lys	21	4.95	5.70
L	Leu	33	7.78	8.03
M	MET	10	2.36	2.77
N	Asn	28	6.60	6.87
P	Pro	15	3.54	3.21
Q	Gln	16	3.77	4.34
R	Arg	8	1.89	2.59
S	Ser	46	10.85	8.97
T	Thr	23	5.42	5.09
V	Val	21	4.95	4.57
W	Trp	10	2.36	3.79
Y	Tyr	29	6.84	9.75

2.5 蛋白质功能区的预测

用在线软件 Prosite (<http://kr.expasy.org/prosite/>)对蛋白质功能位点进行分析可知,绿色木霉内切几丁质酶属于 18 家族几丁质酶,其 164 ~ 172 位氨基酸是活性位点,序列为 FDGID VDWE。

2.6 蛋白质亲/疏水性分析及二级结构的预测

利用 DNAMAN 软件分析成熟内切几丁质酶亲/疏水性以及二级结构的结果见图 4 所示。蛋白质分子中亲水的极性部分在分子表面,疏水的非极性部分在分子内部,所以可以根据氨基酸序列及残基的极性来了解蛋白一级结构中不同肽段的极性及其在分子中的定位。由图 4 可知,该蛋白表现疏水性。通过 Chou-Fasman 分析,该蛋白二级结构由许多 α 螺旋、转角和少量 β 折叠构成。

2.7 蛋白质三级结构预测

利用 SWISS-MODEL (<http://swissmodel.expasy.org/>)同源建模方法,得到成熟内切几丁质酶蛋白三维构象,平面截图如图 5 所示,可以看到该蛋白质三级结构具有 18 家族几丁质酶的相似结构,即具有 α 螺旋和 β 折叠构成的圆桶形结构,中心是有平行的 β -折叠结构组成的内桶,依次为 $\beta 1 \sim \beta 8$,由 α 螺旋结构将它们逐个连接起来,在 α -螺旋和 β -折叠之间有一段无规卷曲连接。

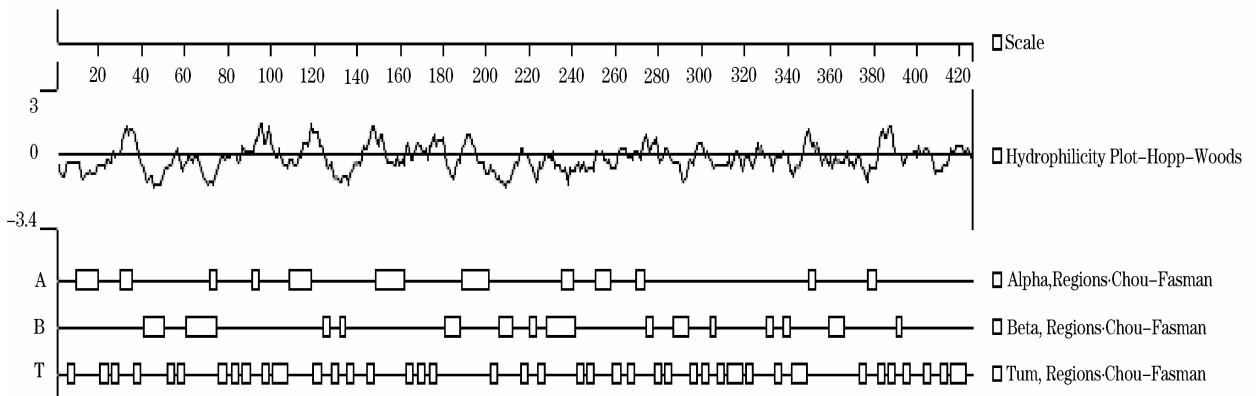


图 4 内切几丁质酶亲/疏水性分析及二级结构预测

Fig. 4 Hydrophilicity/hydrophobicity Plot and secondary structure prediction



图 5 通过 SWISS - MODEL 预测的内切几丁质酶的三维模型

Fig. 5 Three - dimensional Structure of Endochitinase
Predicted by SWISS - MODEL

2.8 氨基酸序列同源性分析

通过软件 DNAMAN 把绿色木霉内切几丁质酶氨基酸序列和其它微生物的内切几丁质酶基因的氨基酸序列进行同源性分析,构建了同源树,见图 6。

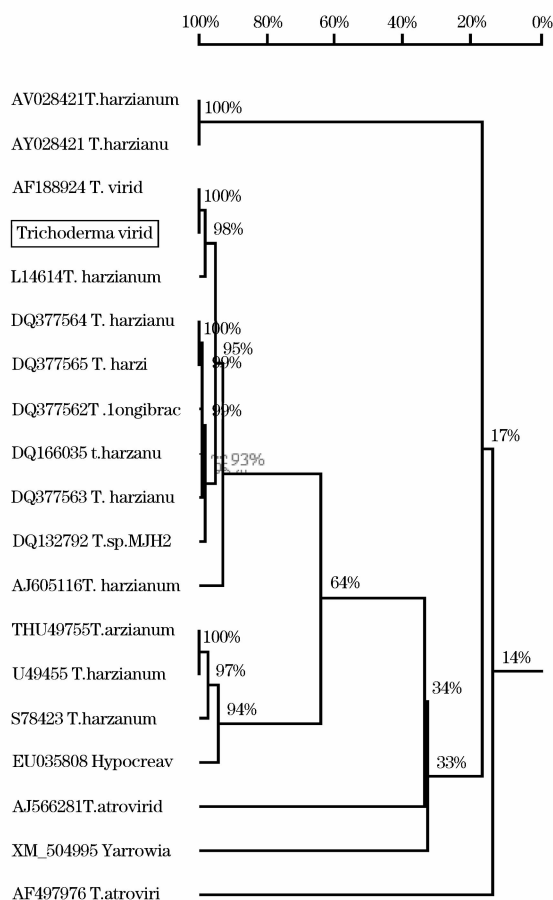


图 6 18 种内切几丁质酶氨基酸序列同源树

Fig. 6 Homologue tree of eighteen endochitinase

大部分木霉的内切几丁质酶同源性都在 64% 以上。实验菌株绿色木霉 (*T. viride*) 用黑框标记,

找到与实验菌株氨基酸同源性为 100% 的绿色木霉内切几丁质酶。但是,同为哈茨木霉产生的内切几丁质酶的同源性也有很大差异。

3 讨论

绿色木霉的 DNA 采用了两种方法来提取。因为很多文献报道丝状真菌的 DNA 提取多采用 CTAB 法^[9-10],所以实验最初采用了 CTAB 法来提取 DNA,发现效果不好,条带弥散, DNA 质量不高。后来采用了王源超等的真菌基因组提取方法,发现效果很好,得到的 DNA 质量较高,可以用作进行 PCR 反应的模板。绿色木霉内切几丁质酶基因为诱导酶,必须在以几丁质为唯一碳源的产酶培养基中进行培养,才能提取到含有内切几丁质酶基因的总 RNA。总 RNA 提取的关键问题是要在其提取过程中严格防止 RNA 酶的污染,并设法抑制其活性。由于所有的组织中均存在 RNA 酶,人的皮肤、手指、试剂、容器等均可能被污染,因此全部实验过程中均需戴手套操作并经常更换(使用一次性手套)。所用的塑料耗材经 0.1% 的 DEPC 溶液浸泡处理过夜,灭活 RNA 酶,后经 1.05kg/cm²、于 121℃ 条件下灭菌 30min,以去除残余的 DEPC。玻璃器皿经 180℃ 干烤 3h 以上灭活 RNA 酶。在实验开始,采用的是单步法 RT-PCR 试剂盒,过程中出现无扩增信号的现象,这与 PCR 扩增不同,除了要考虑引物特异性和退火温度外,还要考虑是否是因为 RNA 模板降解或者是诱导时间不合适,导致目的基因的 mRNA 被降解或目的基因还未转录从而使得反转录得到的 cDNA 模板中没有目的基因片断或者目的基因片断不完整,使得扩增不到目的片断。由于因素众多,提高了实验难度和成本。考虑到上述种种,改用了两步法 RT-PCR,通过在绿色木霉诱导培养的不同阶段提取总 RNA 作为模板进行 RT,在随后的 PCR 反应中再通过常规手段,比如温度梯度来确定最佳退火温度,不同 cDNA 模板量来确定最佳模板量。这样可以有效解决实验中遇到的无扩增条带的问题。利用各种软件对绿色木霉内切几丁质酶测序结果经过分析后,预测了蛋白质一级结构、二级结构、亲/疏水性,模拟了三级结构的三维模型,为进一步研究该蛋白的高级结构、活性中心位点以及 PCR 定点突变奠定了基础。

参考文献 (References)

- [1] Kolaskar A. S., and Reddy B. V. B. Contextual constraints on codon pair usage: structural and biological implications [J]. J Biol Struc Dyn, 1986, 3: 725-728.

- [2] Folley L. S. , and Yams M. Codon contexts from weakly expressed genes reduce expression in vivo [J]. *J Mol Biol*, 1989, 209: 359, 378.
- [3] David J. A. Fungal cell wall chitinases and glucanase [J]. *Microbiology*, 2004, 150:2029-2035.
- [4] Wang H. L. , Wu D. , Deng F. , Peng H. Y. Characterization and phylogenetic analysis of the chitinase gene from the *Helicoverpa armigera* single nucleocapsid nucleopolyhedrovirus [J]. *Virus Research*, 2004, 100: 179-189.
- [5] 沈萍, 范秀容, 李广武. 微生物学实验 [M], 北京:高等教育出版社, 1999:57.
- [6] 吴绵斌, 夏黎明, 岑沛霖. 绿色木霉合成几丁质酶条件的研究 [J]. *化学反应工程与工艺*, 1999, 15(2): 179-185.
- [7] 王源超, 张正光, 郑小波. 核糖体基因 ITS 作为苎麻疫霉、恶疫霉分类辅助性状的研究 [J]. *菌物系统*, 2000, 19(4): 485-491.
- [8] Carolina Carsolio, Ana Gutiurrez, Beatriz Jimenez, Marc Van Montagu. Characterization of ech42, a *Trichoderma harzianum* endochitinase gene expressed during mycoparasitism [J]. *Proc. Natl. Acad. Sci. USA*, 1994, 91: 10903-10907.
- [9] Moller E. M. , Bahnwee, H. G. , Sandermann A. A simple and efficient protocol for isolation of high molecular weight DNA from filamentous fungi, fruit bodies and infected plant tissues [J]. *Nucleic Acids Research*, 1992, 20(22): 61154-61161.
- [10] 朱衡, 瞿峰, 朱立煌. 利用氯化苄提取适于分子生物学分析的真菌 DNA[J]. *真菌学报*, 1994, 13(1): 34-40.