研究报告

长梗木霉纤维素酶基因的克隆及序列分析

石贤爱 1,2* 刘月 1 陈飞 1 杨锦 1

(1. 福州大学生物科学与工程学院 福建 福州 350108)

(2. 福建省医疗器械和医药技术重点实验室 福建 福州 350002)

摘 要: 从富含纤维素环境筛选获得一株纤维素降解菌株 FU05, 通过形态学特征及 ITS 序列分析确定其为长梗木霉(Trichoderma longibrachiatum)。PCR 扩增获得该菌株的 bgl2、cbh2 和 egl。序列分析表明,这3种纤维素酶基因与 GenBank 上其他木霉同种纤维素酶基因具有较高同源性: bgl2 基因与里氏木霉 bgl2 基因(AB003110)同源性达91%; cbh2 基因与康宁木霉 cbh2 基因 (DQ504304)同源性达99%; egl 基因与长梗木霉 egl 基因(X60652)同源性达95%。3种纤维素酶基因编码的相应氨基酸序列与其他木霉纤维素酶的氨基酸序列相似性也非常高。对上述纤维素酶基因编码的相应蛋白进行 PROSITE motif search, 对其 N 端糖基化位点、纤维素结合区、糖基水解酶家族特征结构区等进行了定位。

关键词: 长梗木霉, ITS 序列, 纤维素酶基因, 同源性

Cloning and Sequence Analysis of Cellulase Genes from *Trichoderma longibrachiatum*

SHI Xian-Ai^{1,2*} LIU Yue¹ CHEN Fei¹ YANG Jin¹

(1. College of Biological Science and Engineering, Fuzhou University, Fuzhou, Fujian 350108, China) (2. Fujian Key Laboratory of Medical Instrument and Pharmaceutical Technology, Fuzhou, Fujian 350002, China)

Abstract: Strain FU05 with cellulose-degrading activity was isolated from cellulose-rich environment and identified as *Trichoderma longibrachiatum* through morphology characteristics and ITS sequence analysis. The cellulase genes *bgl2*, *cbh2* and *eg1* were cloned by PCR. By conducting sequence alignment analysis with the reported data it was found that the homology of same cellulase genes between strain FU05 and other *Trichoderma* were: 91% to *bgl2* (AB003110) from *Trichoderma reesei*; 99% to *cbh2* (DQ504304) from *Trichoderma koninqii* and 95% to *eg1* (X60652) from *Trichoderma longibrachiatum*. Furthermore, the corresponding amino acid sequences were also quite similar. By means of PROSITE motif search, the locations of N-glycosylation site, cellulose-binding domain and conserved domains of glycosyl hydrolases family in the corresponding protein were confirmed.

Keywords: Trichoderma longibrachiatum, ITS sequence, Cellulase genes, Homology

基金项目: 福建省发改委资助项目(No. 闽发改投资[2008]256号)

^{*}通讯作者: Tel: 86-591-22866379; Fax: 86-591-22866379; ⊠: shixa@fzu.edu.c© 中国科学院微生物研究所期刊联合编辑部 http://journals.im.ac.cn 收稿日期: 2009-11-18; 接受日期: 2010-01-08

近年来,随着传统化石燃料储量的急剧减少, 开发新的可再生能源已迫在眉睫。木质纤维素是地球上来源最丰富的生物质资源^[1-3],若能将其有效降解成葡萄糖并用于燃料乙醇生产,对于改善目前资源紧张的状况将具有重要意义。纤维素酶是由内切葡聚糖酶(Endoglucanase, EC 3.2.1.4,简称 EG)、纤维二糖水解酶(Cellobiohydrolase, EC 3.2.1.91,简称 CBH)和β-葡萄糖苷酶(β-glucosidase, EC 3.2.1.21,简称 BGL)等组成的复合酶系^[4],在这 3 种酶的协调作用下纤维素被降解成葡萄糖。

由于木霉属能产生大量的胞外纤维素酶, 对纤 维素的降解能力较强, 因此对真菌产纤维素酶的研 究主要集中在木霉属[5-6]。目前对来源于里氏木霉、 绿色木霉和康宁木霉的纤维素酶基因研究较多,已 克隆 cbh1-3、eg1-5、bgl1-2 等基因, 且不少基因已 在酵母和大肠杆菌中成功表达[7-9]。但关于长梗木霉 纤维素酶的报道较少、且多见于纤维素酶的发酵条 件优化及提取[10-12]。目前国内外关于长梗木霉纤维 素酶基因表达仅见几篇报道: 王娟等[13]成功将长梗 木霉 XST1 cbh2 基因在毕赤酵母中表达,发酵上清 液水解 pNPC 的酶活为 18.1 U/L; 刘海英等[14]将长 梗木霉 SSL 的 egl 基因表达于毕赤酵母, 经甲醇诱 导后, 胞外重组内切葡聚糖酶 I 的活力达 73 U/mL; Gonzdlez等[15]克隆了长梗木霉的 egll 基因, 置于诱 导型启动子 cyc-gal 下游于酵母中表达, 获得的转化 子可分泌具有酶活性的蛋白。

本研究筛选获得一株纤维素降解菌长梗木霉FU05,测定了其液态发酵分泌表达纤维素酶的能力;通过PCR扩增获得3种纤维素酶基因 bgl2 、cbh2和 eg1 (其中首次从长梗木霉中克隆获得 bgl2 基因),并对其进行序列比对和同源性分析。本研究不仅有助于完善木霉纤维素酶的分子生物学理论基础,同时为长梗木酶纤维素酶基因的克隆、改造和高效表达创造了条件。

1 材料和方法

1.1 主要试剂

真菌基因组 DNA 提取试剂盒购自杭州博日科技有限公司; 质粒 DNA 提取试剂盒、胶回收试剂盒和 pUCm-T 克隆试剂盒购自 Bio Basic Inc.; 即用 PCR 扩增试剂盒(含 0.1 U/μL 普通 Taq DNA 聚合酶)购自上海生工生物工程技术服务有限公司; 其他试

剂为国产分析纯。

1.2 培养基

富集培养基 (g/L): NaNO₃ 0.5, KH₂PO₄ 1.0, MgSO₄ 0.5, KCl 0.5, Fe₂(SO₄)₃微量,自然 pH。使用时以滤纸为碳源,剪成长条状,放入试管内,略露出液面。

初筛培养基(g/L)^[16]: $(NH_4)_2SO_4$ 1.4, KH_2PO_4 2.0, $CaCl_2$ 0.3, $MgSO_4$ 0.3, 尿素 0.3, 羧甲基纤维素 钠 10.0, $FeSO_4$ ·7 H_2O 0.0016, $ZnSO_4$ ·7 H_2O 0.0014, $CoCl_2$ 0.002, 琼脂 20, 自然 pH_\circ

PDA 培养基: 土豆 200 g 洗净, 去皮, 切成小块, 煮沸 30 min, 6 层纱布过滤, 定容至 1 L, 加葡萄糖 20 g, 自然 pH。若制成平板, 加琼脂 20 g。

液体发酵培养基 (g/L): 蛋白胨 3, (NH₄)₂SO₄ 2, 酵母膏 0.5, KH₂PO₄ 4.0, CaCl₂ 0.3, MgSO₄ 0.3, Tween-80 0.2, 微晶纤维素 20, 麸皮 60。

LB 培养基: 胰蛋白胨 10 g, 酵母提取物 5 g, NaCl 10 g, 溶于 950 mL 水, 加 1 mol/L NaOH 调 pH 至 7.0 后补水至 1 L。若制成平板, 加琼脂 20 g。

1.3 菌株筛洗

从福州大学后山富含纤维素的环境采集样品,加入适量无菌水打散,制成悬液。吸取样品悬液于富集培养基,30°C、200 r/min 培养 7 d,将滤纸取出用无菌水打散制成悬液,适当稀释后涂布于初筛培养基平板,30°C培养 4 d。挑取生长较快的单菌落于PDA 平板上进行划线分离得到纯化的菌株。将该菌株点接于初筛培养基平板,30°C培养 1 d,再往培养皿内加适量 1 g/L 刚果红染液,染色 2 h。倾去染液,加入 1 mol/L NaCl 浸泡 1.5 h。选择透明圈较大的菌株于液体发酵培养基 30°C、200 r/min 培养,每天测定 1 次酶活。

1.4 纤维素酶活力测定

取发酵液于 6000 r/min 离心 15 min, 将上清作为粗酶液。滤纸酶活的测定:参照文献[17]略作改进,将 50 mg 滤纸浸入 1.5 mL 0.1 mol/L, pH 4.8 的柠檬酸缓冲液中,加入 0.5 mL 适当稀释的酶液,50°C 水浴反应 30 min,加入 1.5 mL DNS 试剂终止反应,沸水浴中煮沸 5 min,立即用流动水冷却,蒸馏水定容至 20 mL,于 540 nm 测定 OD 值。以 100°C 水浴灭活 5 min 的粗酶液为空白对照,通过葡萄糖标准曲线计算酶的活力。酶活定义: 1 min 内分解底物生成 1 μg 葡萄糖所需酶量定义为 1 个酶活单位(U)。

1.5 菌株鉴定

1.5.1 菌落形态观察:将盖玻片沿 45 度角斜插入 PDA 培养基平板,于盖玻片与培养基交接处接种菌 株孢子悬液 30°C培养,连续几天观察菌落边缘和表面的质地、颜色、纹饰等形态变化,并取下盖玻片,于光学显微镜下观察。

1.5.2 ITS序列分析: 按博日真菌基因组 DNA 提取 试剂盒法提取基因组 DNA, 合成引物^[18]进行 PCR 扩增, 1F: 5'-TCCGTAGGTGAACCTGCGG-3'; 1R: 5'-TCCTCCGCTTATTGATATGC-3'。PCR 反应条件: 94°C 7 min; 94°C 1 min, 55°C 30 s, 72°C 1 min, 30 个循环; 72°C 20 min。胶回收 PCR 扩增产物,经上海生工测序后,与 GenBank 上已知真菌 ITS 序列进行比对分析。

1.6 纤维素酶基因的扩增

根据 GenBank 上发表的纤维素酶基因序列设计 引物如下:根据 β-葡萄糖苷酶基因 *bgl2* 序列 (GenBank accession No. AY343988)设计引物: 2F: 5′-ATGTTGCCCAAGGACTTTCA-3′; 2R: 5′-TCAA GCTCTTTGCGCTCTTC-3′。根据外切葡聚糖酶基因 *cbh2* 序列 (DQ504304)设计引物: 3F: 5′-ATGAT TGTCGGCATTCTCACC-3′; 3R: 5′-TTACAGGAAC GATGGGTTTGC-3′。根据内切葡聚糖酶基因 *eg1* 序列 (M15665)设计引物: 4F: 5′-ATGGCGCCCTCAG TTAC-3′; 4R: 5′-CGCTCTAAAGGCATTGC-3′。反应 条件: 94°C 7 min; 94°C 1 min, 55°C 1.5 min, 72°C 2 min, 30 个循环; 72°C 10 min。

胶回收 PCR 扩增产物, 按 pUCm-T 克隆试剂盒步骤与 T 载体连接, 连接产物转化大肠杆菌感受态细胞, 挑取转化子, 用通用引物 M13/pUC Sequencing primer F 和 R 进行 PCR 验证, 挑选阳性转化子送上海生工进行测序。

1.7 序列比对分析

利用软件 MEGA 3 进行 ITS 序列聚类分析, DNAMAN、Primer Premier 5.0、GenBank 上的 BLAST、http://cubic.bioc.columbia.edu 网站上的 Predictprotein 等分析工具对扩增得到的长梗木霉 FU05 纤维素酶基因进行序列比对分析。

2 结果

2.1 产纤维素酶菌株的筛选

初筛得到的15株菌经刚果红染色法复筛后,均

可观察到透明圈,但透明圈的形态各不一样。有的菌株生长较快,但透明圈较小;有的菌株生长较慢,透明圈却很明显;有的菌株在平板上菌丝体呈疏松状蔓延生长,菌落大而蓬松,周围无明显透明圈,但通过平板背面观察可见菌丝覆盖处平板完全呈透明状态。综合平板上刚果红染色透明圈的情况以及菌株生长状态,确定7株菌用于液体发酵实验,并分别命名为:FU01、FU02、FU03、FU04、FU05、FU06、FU07。液体发酵产酶结果显示,酶活最大的菌株为FU05,培养6d时酶活达106.4 U/mL,如表1所示。

2.2 菌株鉴定

2.2.1 菌株形态学特征: 菌株 FU05 在 PDA 平板上 先以白色圆形绒状生长,随后迅速向四周拓展,然 后从菌落中央产生绿色孢子,使中央变成绿色,最 后整个平皿布满黄绿色菌落。显微镜下观察菌丝形 态如图 1 所示, FU05 菌丝体上可见明显的瓶颈状分 生孢子梗,且分枝与分生孢子梗呈直角,分生孢子 为圆形,聚集成圆团状于孢子梗顶端。参考真菌鉴 定手册[19]初步判断 FU05 为木霉属。

表 1 菌株产纤维素酶活力 Table 1 Activities of cellulases produced by strains		
菌株编号 Strain No.	达到最大酶活时间 Time of maximum activity (d)	最大酶活 Maximum enzyme activity (U/mL)
FU01	5	70.5
FU02	5	65.1
FU03	5	62.7
FU04	4	61.3
FU05	6	106.4
FU06	6	88.7
FU07	4	62.1

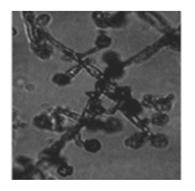


图 1 FU05 菌丝形态 (× 640)

Fig. 1 Morphology of strain FU05 (× 640)

2.2.2 ITS 序列分析: 将菌株 FU05 ITS 测序结果 (GenBank accession No. GU144295)在 GenBank 里进行搜索比对,与长梗木霉 (Trichoderma longibrachiatum) ITS 序列同源性达 99%,即确定所筛选的菌株 FU05 为长梗木霉。ITS 聚类分析结果如图 2 所示,可知 FU05 与长梗木霉处于同一分支,亲缘关系最近,且与里氏木霉、拟康氏木霉等也有较近的亲缘关系。

2.3 纤维素酶基因的扩增

以 FU05 基因组 DNA 为模板, 2F 和 2R、3F 和 3R、4F 和 4R 为引物, PCR 扩增纤维素酶基因, PCR 产物进行 1.2%琼脂糖凝胶电泳, 结果如图 3 所示。可以看出, PCR 扩增出 1000-2000 bp 间单一的 DNA 条带, 大小与预期一致(bgl2、cbh2、eg1 基因大小均

在 1500 bp 左右)。

2.4 纤维素酶基因序列分析

2.4.1 bgl2 基因: 长梗木霉 FU05 bgl2 基因全长 1431 bp (GenBank accession No. GU144296), 进行 BLAST 搜索比对,结果表明与里氏木霉 (AB003110)、绿色木霉(AY343988) bgl2 基因同源性较高,其中与里氏木霉 bgl2 基因(AB003110)同源性达91%。结合以上两种木霉 bgl2 基因内含子位置可知 FU05 bgl2 基因内含子位置可知 FU05 bgl2 基因编码的相应氨基酸序列进行 Protein blast搜索比对,结果显示 FU05 bgl2 基因编码 453 个氨基酸,属于糖基水解酶家族 1,与里氏木霉 bgl2 基因(AB003110)编码的氨基酸序列具有较高相似性,达 93%。

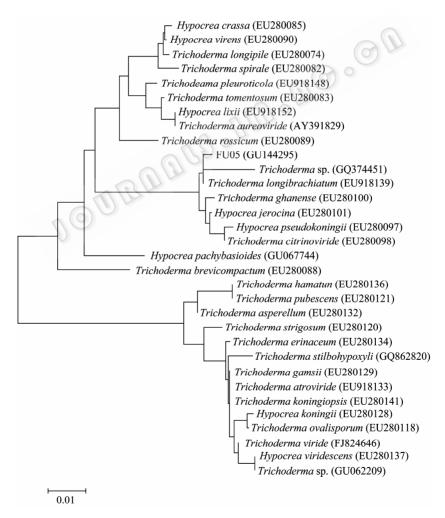


图 2 FU05 ITS 序列聚类分析结果(N-J 算法)

Fig. 2 Cluster analysis of ITS sequence (N-J algorithm)

Note: The number in parentheses represents GenBank accession number. The ruler represents evolutionary distance.

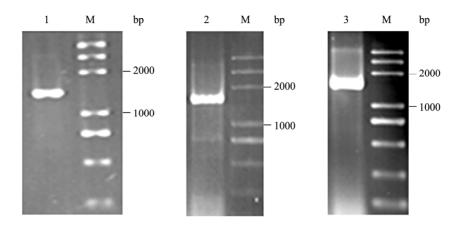


图 3 PCR 扩增纤维素酶基因片段琼脂糖凝胶电泳图 Fig. 3 Gel electrophoresis of cellulase genes obtained by PCR

Note: M: DNA marker; 1: bgl2; 2: cbh2; 3: eg1.

利用 Predictprotein 进行 PROSITE motif search, 结果表明 bgl2 基因编码的蛋白含 $3 \land N$ 端糖基化位点($N^{55}RTA \lor N^{135}RTE \lor N^{367}GTS$)、4 个蛋白激酶 C 磷酸化位点、10 个酪蛋白激酶 II 磷酸化位点、3 个 N 端酰基化位点和 1 个糖基水解酶家族 1 的 N 端特征序列($F^{6}OWGFATASYOIE GA$)。

2.4.2 cbh2 基因: 长梗木霉 FU05 cbh2 基因全长 1584 bp (GenBank accession No. GU144297), 与康宁木霉 3.2774 cbh2 基因(DQ504304)的同源性非常高,达 99%,仅 4个碱基不同。结合康宁木霉 3.2774 cbh2基因序列比对分析,FU05 cbh2基因内含子位于92-142 bp、528-583 bp、831-894 bp处。氨基酸序列比对结果显示 FU05 cbh2 基因编码 470 个氨基酸,属于糖基水解酶家族 6,与康宁木霉 3.2774 cbh2 基因编码的氨基酸序列具有非常高的相似性,达 99%,仅第 30 位的氨基酸不同(长梗木霉 FU05 为 Val,康宁木霉 3.2774 为 Ala)。

PROSITE motif search 分析表明: *cbh2* 基因编码的蛋白含 3 个 N 端糖基化位点(N³⁸WSG、N³¹²ASS、N³³³ITS)、8 个蛋白激酶 C 磷酸化位点、4 个酪蛋白激酶 II 磷酸化位点、11 个 N 端酰基化位点、2 个糖基水解酶家族 6 的特征序列(V¹⁹⁰VYDLPDRD CAALASNG, L²³⁸LVIEPDSLA)及纤维素结合域(C³⁴GGQN WSGPTCCASG STCVYSNDYYSQC)。

2.4.3 *eg1* 基因: 长梗木霉 FU05 *eg1* 基因全长 1565 bp (GenBank accession No. GU144298), 与长梗木霉 *eg1* 基因(X60652)同源性较高, 达95%; 另外与拟康氏木霉(EF185865)、木霉属(EU935217) *eg1*

基因同源性也较高。结合长梗木霉和拟康氏木霉 egl 基因序列比对分析, FU05 egl 基因内含子位于771-892 bp、1489-1545 bp 处。FU05 egl 基因编码461 个氨基酸,属于糖基水解酶家族 7,与拟康氏木霉 egl 基因(EF185865)编码的氨基酸序列具有非常高的相似性,达99%,仅第5、152位的氨基酸不同(长梗木霉 FU05 第5、152位分别为 Val、Leu,而康宁木霉 3.2774 相应位置为 Ala 和 Phe)。

PROSITE motif search 分析表明: egI 基因编码的蛋白含 5 个 N 端糖基化位点(N⁷⁸TTL、N¹⁶⁴GSL、N²⁰⁴GTL、N²⁰⁸TSG、N³⁹⁴STG)、1 个 cAMP-cGMP依赖性蛋白激酶磷酸化位点(R⁴¹⁴RSS)、7 个蛋白激酶 C 磷酸化位点、4 个酪蛋白激酶 II 磷酸化位点、1 个酪氨酸激酶磷酸化位点、16 个 N 端酰基化位点、及纤维素结合域(C⁴³³GGIGYTGCKTCTSGTTCQY GNDYYSQC)。

3 讨论

从富含纤维素环境中分离获得一株纤维素降解菌长梗木霉 FU05, 通过 PCR 扩增获得 3 种纤维素酶基因 bgl2、cbh2 和 egl。经序列比对分析发现,上述 3 种纤维素酶基因与同属内其他木霉同类基因的同源性很高,达 90%以上;并对这 3 种纤维素酶基因编码的相应蛋白进行 PROSITE motif search,为其进一步的分子生物学研究提供参考依据。

纤维素酶包括3种酶,每种酶又各有多种类型。 BGLII是 Takashima S 等 1999 年发现的一种新型纤维素酶,目前与该酶相关的研究报道并不多; CBH II 含量虽远低于 CBH I,但该酶特异性强且酶活明显高于 CBH I,其降解微晶纤维素的能力比 CBH I 高两倍。而 EG I 是内切葡聚糖酶中的中的重要组分,其表达分泌量约占木霉胞外蛋白总量的 10%。因此,本文对长梗木霉 FU05 选取上述 3 种纤维素酶的基因 bgl2、cbh2 和 eg I 作为代表进行克隆和序列分析。有关长梗木霉其他类型纤维素酶基因的克隆、分析,以及选择合适的载体及表达宿主实现长梗木霉纤维素酶基因的高效表达,将在后续工作中进行系统和深入的研究。

参考文献

- [1] Sukumaran RK, Singhania RR, Mathew GM, *et al.* Cellulase production using biomass feed stock and its application in lignocellulose saccharification for bio-ethanol production. *Renewable Energy*, 2009, **34**(2): 421–424.
- [2] Kadam KL, McMillan JD. Availability of corn stover as a sustainable feedstock for bioethanol production. *Bioresource Technology*, 2003, **88**(1): 17–25.
- [3] Wyk VJPH, Mohulatsi M. Biodegradation of wastepaper by cellulase from *Trichoderma viride*. *Bioresource Technology*, 2003, **86**(1): 21–23.
- [4] Gray KA, Zhao LS, Emptage M. Bioethanol. *Current Opinion in Chemical Biology*, 2006, **10**(2): 141–146.
- [5] Sun Y, Cheng JY. Hydrolysis of lignocellulosic materials for ethanol production: a review. *Bioresource Technology*, 2002, **83**(1): 1–11.
- [6] Weber J, Agblevor FA. Microbubble fermentation of *Trichoderma reesei* for cellulase production. *Process Biochemistry*, 2005, **40**(2): 669–676.
- [7] 张梁, 石贵阳, 王正祥, 等. 酿酒酵母 GPD1 中整合表达纤维二糖酶基因用于纤维素酒精发酵的研究. 西北农林科技大学学报:自然版, 2006, **34**(10): 164-177.
- [8] 刘北东,杨谦,周麒,等. 绿色木霉 AS313711 的葡聚糖 内切酶Ⅲ基因的克隆与表达. 环境科学, 2004, **25**(5):

127-132.

- [9] 黄艳,凌敏,覃拥灵,等.康氏木霉内切葡聚糖酶 (EGI)基因的克隆及表达.生物技术,2008,18(2):10-13.
- [10] Gama FM, Faro CJ, Teixeira JA, et al. New methodology for the characterization of endoglucanase activity and its application on the *Trichoderma longibrachiatum* cellulolytic complex. Enzyme and Microbial Technology, 1993, 15(7): 57-61.
- [11] Sandhu DK, Kalra MK. Production of cellulase, xylanase and pectinase by *Trichoderma longibrachiatum* on different substrates. *Transactions of the British Mycological Society*, 1982, 79(3): 409–413.
- [12] 谭宏, 刘淑欢, 李剑英, 等. 长梗木霉纤维素酶的产生及提取. 微生物学通报, 1993, **20**(2): 90-93.
- [13] 王娟, 刘海英, 朱亚然, 等. 长梗木霉外切纤维素酶 CBH II 基因的克隆及表达. 药物生物技术, 2009, **16**(2): 95-98.
- [14] 刘海英,王娟,舒正玉,等.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.微生物学通报,2009,36(3):355-359.
- [15] Gonzdlez R, Ramon D, Perez GJA. Cloning, sequence analysis and yeast expression of the *egl1* gene from *Trichoderma longibrachiatum*. *Applied Environmental Microbiology*, 1992(38): 370–375.
- [16] 刘清锋, 支晓鹏, 徐惠娟, 等. 纤维素降解菌青霉 T24-2 的分离及产酶特性. 工业微生物, 2007, **37**(3): 15-19.
- [17] 白洪志,杨谦,王希国,等.纤维素降解菌绿色木霉 C-08 的筛选及酶学特性研究.安徽农业科学,2007, **35**(17):5033-5034.
- [18] 陈应龙, 弓明钦, Dell B. 分子生物学技术在菌根研究中的应用及其进展. 土壤与环境, 1999, **8**(3): 230-234.
- [19] 魏景超. 真菌鉴定手册. 上海: 上海科学技术出版社, 1979: 493-494.