

哈茨木霉 Thi4 基因的克隆与序列分析

黄彩红, 杨 谦, 刘志华

(哈尔滨工业大学 生命科学与工程系, 黑龙江 哈尔滨 150001)

摘 要: 哈茨木霉 (*Trichoderma harzianum*) 是一种丝状土壤真菌, 作为一种优秀的生防因子在植物病害防治过程中发挥着重要的作用。为研究其生防机制并获得生防相关基因, 构建了菌丝生长期 cDNA 文库, 通过对部分 ESTs 序列的测定与生物信息学分析, 成功获得了噻唑生物合成酶基因 Thi4 的全长 cDNA 序列。Thi4 基因全长 1311bp, 包含一个编码 322 个氨基酸残基长度为 969bp 的开放读码框, 编码蛋白理论分子量为 34.5kDa, 等电点为 5.18。基因的氨基酸序列含有多功能性位点且在进化上保守。将基因提交国际权威基因数据库 GenBank, 序列号为 DQ647956。

关键词: 哈茨木霉; 噻唑生物合成酶; 克隆; 序列分析

中图分类号: Q 949.32 **文献标识码:** A **文章编号:** 1001-0009(2009)01-0092-04

农药对提高现代农业生产贡献很大, 同时也给生态环境带来不利影响。单一农药的长期使用也会使病虫害、杂草产生抗性。因此, 对哺乳动物和有用生物低毒、对环境负荷小、对现有药剂抗性害虫具有高活性的新型化合物的研制就成为广泛研究的重点。噻唑是含一个硫原子和一个氮原子的五元杂环, 在植物体内是高效的治疗剂, 同时内吸性能好, 具有良好的治疗和保护的三重作用。将噻唑基因引入到各种不同的化合物结构中或在噻唑类化合物中引入不同取代基团, 通过结构修饰能产生一系列具有广谱生物活性的化合物, 从而使得噻唑类化合物在新型超高效农药创制中发挥出越来越重要的作用^[1-2]。此外, 噻唑结构作为很多重要化合物的内核而在医药方面占有重要位置, 而且不同取代基在不同位置的取代使得噻唑基团具有不同的味道, 可作为食品添加剂和香精应用于精细化工和食品工业^[3-4]。但是, 目前噻唑的生产主要依赖于化学合成, 而利用相关酶系进行的生物法合成却仍处于起始研究阶段。噻唑生物合成酶基因常被发现于酵母及古细菌中, 属于一个与 FAD/NAD(P) 结合相关的 Thi4 家族, 最早的编码基因 STI35 是在 1990 年作为与胁迫相关的基因被克隆和鉴定的, 其表达受维生素 B₁ 调控^[5-6]。迄今为止, 被研究

并报道的 Thi4 家族的功能主要体现在以下 3 个方面: 涉及噻唑和维生素 B₁ 的生物合成, 间接影响有机体的新陈代谢; 对线粒体 DNA 具有保护和修复作用^[7]; 与各种胁迫诱导有关。

1 材料与方法

1.1 菌种与质粒

大肠杆菌 (*E. coli*) DH5 α 和质粒 pBS-Thi4 均为该实验室保存。

1.2 培养基与试剂

LB 培养基用于大肠杆菌的培养。质粒提取试剂盒与胶回收试剂盒购自上海华舜生物工程公司; 各种限制性内切酶、T4 DNA 连接酶、EX-Taq DNA 聚合酶、IPTG 等均购自 TaKaRa 公司; 其他生化试剂均为国产分析纯。

1.3 试验方法

1.3.1 目的基因的筛选与测序 对 cDNA 文库中所得 EST 序列聚类分析并通过 NCBI 的 BLASTx 程序完成同源性分析, 将预计具有全长 cDNA 序列的质粒从保存的模板中挑取之后, 转化大肠杆菌 *E. coli* DH5 α 感受态细胞, 得到阳性克隆经酶切鉴定后, 用 T3 和 T7 通用引物进行双向测序。DNA 测序采用双脱氧链终止法, ABI 公司的 377 型全自动 DNA 测序仪和 BigDye Terminators Cycle Sequencing Kit, 由上海博亚生物工程公司进行 DNA 测序。

1.3.2 引物设计与目的基因扩增 依据文库序列测定结果设计克隆噻唑生物合成酶基因 Thi4 的特异性双向引物, 同时添加 BamHI 和 XhoI 酶切位点以用于后续载体构建, 具体序列如下: ThiL: 5'—GACAT GGATC—CATGAGTCCTCCCGCCGCCGTCTCG—3'; ThiR: 5'—

第一作者简介: 黄彩红(1976-), 女, 博士, 主要从事微生物基因工程方向研究。E-mail: caihonghuang@163.com。

通讯作者: 杨谦。

基金项目: 国家“863”基金资助项目(2003AA241140); 黑龙江省自然科学基金重点资助项目(JZN03-04); “十一五”国家科技支撑计划重点资助项目(2006BAD07A00)。

收稿日期: 2008-09-26

AATGCTCGAGTGATGGCGTTTTCCTTGCGACGAGT—3’。以质粒 pBS-Thi4 为模板扩增基因 Thi4 的编码框序列,采用 TaKaRa 公司的 Ex-Taq 构建 25 μL 反应体系,经过反复摸索确定最佳退火温度(*T*_m)为 56℃,反应条件为:94℃预变性 5 min;94℃变性 30 s,56℃退火 40 s,72℃延伸 1 min,35 个循环;72℃延伸 10 min,4℃停止反应。

1.3.3 序列分析方法 将所获得的基因序列经

BLASTx 与 GenBank 中非冗余蛋白数据库比对后在核苷酸水平和氨基酸水平进行同源性比较;采用 DNA-MAN 软件分析序列基本信息;nnpredict 预测蛋白的二级结构;利用 Expzsy 的分子生物学服务器(Expzsy molecular biology server)上的 Scanprosite facility 软件对蛋白氨基酸序列进行结构域预测;利用 ClustalX (1.83) 软件进行同源序列的多重排定;应用 MEGA3.1 构建 NJ 系统进化树,节点支持率使用 1000 次 bootstrap 检验。

1	ATTCCTCGATAAAATAATCGAATAAAATATCCAATTGAAAATTCAAGTATTCAGTGATT
61	TAATTCACAGTACCAACTTACATCATTCCGTACACCTCGTCCAGTCGCCATGGCGTCTTT
121	CGGTGGTAGTCCGTCCAGACGGGATCATGTCTCTCAGGTCCATCATGAGTCTCCCGCC
41	M S P P A
181	GCCGTCCTCGCCACCCAGCACACCGCTGAGCTTGTGAAGCCCGTCTCCAAGCTGGCTGTT
61	A V S P T Q H T A E L V K P V S K L A V
241	GCTCAGAACCAAGGCCCTGGCCAAGACTGTCTCCGAGATGCACGGCCAATGGGACTCG
81	A Q N Q G P L A K T V S E M H G Q W D S
301	TTCACCTTTGCCCCATCCGCGAGTCTCAGGTCTCGCGGCCATGACCCGTGCTACTTC
101	F T F A P I R E S Q V S R A M T R R Y F
361	GAAGACCTCGACAACACGCGGAGTCCGACATCGTCATCATCGCGCGCGGCTCTGCGGT
121	E D L D N Y A E S D I V I I G A G S C G
421	CTTAGCACCGGTACGTCCTCGGCACTCAGCGTCCGGACCTCAAGATGCCATCATCGAG
141	L S T A Y V L G T Q R P D L K I A I I E
481	GCCTCTGTCTCGCCTGGTGGAGGTGCCTGGCTGGGAGGCCAGCTCTTCTCCGCCATGGTG
161	A S V S P G G G A W L G G Q L F S A M V
541	ATGCGCAAGCCGGCGATGCTTCTGCGCGAGATTGGCGTTCCTACGAGGACGAGGGC
181	M R K P A D A F L R E I G V P Y E D E G
601	AACTACGTCGTCGTAAGCACGCGGCGCTCTTACCTCGACCATCATGGCCAAGGTGCTG
201	N Y V V V K H A A L F T S T I M A K V L
661	CAGTGCCCAACGTCAAGCTCTTCAACGCCACTTGCGTCGAGGACCTCATCACTCGTCCT
221	Q L P N V K L F N A T C V E D L I T R P
721	TCTGCTGAGGGAGTCCGCATTGCTGGTGTGTGACCAACTGGACTCTTGCTCGATGCAT
241	S A E G V R I A G V V T N W T L V S M H
781	CACGATGACCAGTCTTGATGGACCCCAACCATCAATGCTCCCTTGTCATCTCCACT
261	H D D Q S C M D P N T I N A P L V I S T
841	ACCGGCCACGATGGCCCAATGGGTGCCTTTTGGCTCAAGCGTCTGTGACATGGGCCGT
281	T G H D G P M G A F C V K R L V S M G R
901	ATTGAGAAGCTGGGCGCATGCGCGGCTCGATATGAACAGGGCTGAAGATGCTATTGTC
301	I E K L G G M R G L D M N R A E D A I V
961	AAGAACTCGTGAGGTTGTTCGGGCTCATTGTCGGAGGCATGGAGTCTCTGAGATT
321	K N T R E V V P G L I V G G M E L S E I
1021	GATGGAGCCAACCGTATGGGTCTTACCTTTGGTGCTATGGCTCTCAGCGGTGTCAAGGCT
341	D G A N R M G P T F G A M A L S G V K A
1081	GCTGAAGAGGCTCTCAAGGTCTTTGAGACTCGTCGAAGGAAAACGCCATCTAGATGTGA
361	A E E A L K V F E T R R K E N A I *
1141	TTTTTGGGATATACATACCAACCACATTCATTTGATACTAGTTCTTTGTGTGACTGGTC
1201	GCTCATTTGATGGAAGAAGTAAATCTGAAAATGAAAAAATAGTTATTCCGCTGCCAA
1261	CCAGATTGAATACAAATTTGCAGCTATAAAAAAAAAAAAAAAAAAAAAA

图 1 噻唑生物合成酶基因 Thi4 的核苷酸及推定的氨基酸序列

2 结果与分析

2.1 基因 Thi4 全长 cDNA 序列的获得

经 BLASTx 与 GenBank 中非冗余蛋白数据库比对分析后,发现哈茨木霉 cDNA 文库中一条 EST 序列与镰刀菌 (*F.solani*) 的噻唑生物合成酶基因具有较高的同源性。通过以 T3 和 T7 通用引物双向测通后,获得了哈茨木霉噻唑生物合成酶的全长 cDNA 序列(图 1)。该序列长度为 1311bp,GC 含量为 53.32%,含有 969 个碱基组成的完整开放阅读框,编码 322 个氨基酸组成的蛋白,该蛋白理论分子量为 34.5kDa,理论等电点(pI)为 5.92。

2.2 基因 Thi4 的 PCR 扩增结果

以质粒 pBS-Thi4 为扩增模板,采用 TaKaRa 公司的 Ex-Taq 构建 25 μ L 反应体系,经过反复摸索确定最佳退火温度(T_m)为 58 $^{\circ}$ C,其它反应条件参照相应试剂盒说

明书,得到了哈茨木霉编码 Thi4 噻唑生物合成酶的特异性 DNA 片断,约 969 bp 符合预期大小(图 2)。

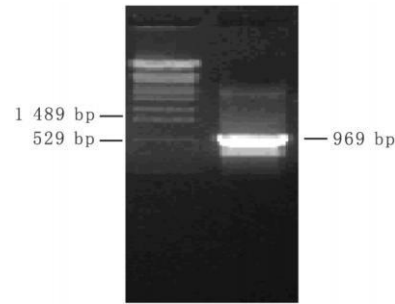


图 2 哈茨木霉 Thi4 基因的 PCR 电泳结果
M: Marker λ -Eco T14 I digest; 1: Thi4 约 969bp

THI4_FUSSH	MSPPAAVSPPARSAELASAPAVKLPVGLSKNSAAATTVEEMEGKUDDFKFAPIRESQVSR	60
THI4_FUSOX	MAPPAAVSPPARSAELATS--TKLPV-MSKN-INTKTVEEMLGQDDFKFAPIRESQVSR	56
THI4_HAZUM	MSPPAAVSPPTQHTAELVKP-VSKLAVAQNG-PLAKTVSEMHGQWDSFTFAPIRESQVSR	58
THI4_ASPOR	MSPPAAIYEPTVAATGLKGKVVVSETVPVVEGASQTKLLDHFGGKWDEFKFAPIRESQVSR	60
Q76B84_EMENI	MSPPAAIFEPTVAPTGIKGVVVPETATIPGDSQTKLLDHFGGKWDFNFKFAPIRESQVSR	60
THI4_FUSSH	AMTRRYFQDLNVAESDIVIGAGSCGLSTRYILGKKRPDLKIAIEASVSPGGGAWLGG	120
THI4_FUSOX	AMTRRYFQDLNVAESDIVIGAGSCGLSAAVILGKKRPDLKIAIEASVSPGGGAWLGG	116
THI4_HAZUM	AMTRRYFEDLDNVAESDIVIGAGSCGLSTAYVLGTORPDLKIAIEASVSPGGGAWLGG	118
THI4_ASPOR	AMTRRYFEDLDKYAESDVVIVGAGSCGLSTAYVLAKARPDKIAIEASVSPGGGAWLGG	120
Q76B84_EMENI	AMTRRYFQDLDRVAESDVVIVGAGSCGLSTAYVLAKARPDKIAIEASVSPGGGAWLGG	120
THI4_FUSSH	QLFSAMVMRKPADAFLREVGVPEDEGN---YVVVKHAAFLTSTIMSKVLQLPNCKLFNA	177
THI4_FUSOX	QLFSAMIMRKPADAFLREVGVPEDEGN---YVVVKHAAFLTSTIMSKVLQMPNICKLFNA	173
THI4_HAZUM	QLFSAMVMRKPADAFLREIGVPVEDEGN---YVVVKHAAFLTSTIMAKVLQLPNCKLFNA	175
THI4_ASPOR	QLFSAMVMRPFAEVFLNELGVPEEDAN-PNVYVVVKHASFSTIMSKVLSFPNVKLFNA	179
Q76B84_EMENI	QLFSAMVMRPFALFLNELGVPEEDPDMPNVYVVVKHASFSTLLSKVLSFPNVKLFNA	180
THI4_FUSSH	TCVEDLITRPSKEGVR----ISGVVTNWTLVSMHDDQSCMDPNTINAPLVIISTTGHDGP	233
THI4_FUSOX	TCVEDLITRPSKEGVR----IAGVVTNWTLVSMHDDQSCMDPNTINAPLIISTTGHDGP	229
THI4_HAZUM	TCVEDLITRPSAEGVR----IAGVVTNWTLVSMHDDQSCMDPNTINAPLVIISTTGHDGP	231
THI4_ASPOR	TAVEDLITRPTENGVP----QIAGVVTNWTLVTLHDDHSCMDPNTINAPVIISTTGHDGP	236
Q76B84_EMENI	TCVEDLITRPGFNGNAQEVQIAGVVTNWTLVTLHDDHSCMDPNTINAPVIISTTGHDGP	240
THI4_FUSSH	MGAFCKRLVSMGRIEKLGGMRGLDMNVAEDAIVKGTREIVPGLIVGGMELSEVDGANRM	293
THI4_FUSOX	MGAFCKRLVSMQRIEKLGGMRGLDMNLAEDAIVKGTREIVPGLIVGGMELSEVDGANRM	289
THI4_HAZUM	MGAFCKRLVSMGRIEKLGGMRGLDMNRAEDAIVKNTREVVPGLIVGGMELSEIDGANRM	291
THI4_ASPOR	FGAFCAKRLVSMGSDVKLGGMRGLDMNSAEDAIVKNTREVTKGLIIGGMELSEIDGFNRM	296
Q76B84_EMENI	FGAFSAKRLVSMSTTIDKLGMRGLDMNSAEDAIVKNTREVAKGLIIGGMELSEIDGFNRM	300
THI4_FUSSH	GPTFGAMVLSGLKAAEEALKVIDIRQKQNSF	324
THI4_FUSOX	GPTFGAMALSLGLKAAEEALKIFDTRKKQNDL	320
THI4_HAZUM	GPTFGAMALSGVKAEEALKVFETRKKENAI	322
THI4_ASPOR	GPTFGAMVLSGVKAAEEALKVDFERQRECAE	327
Q76B84_EMENI	GPTFGAMVLSGVKAAEEALRVFDDRKRECAE	331

图 3 噻唑生物合成酶基因多序列比较

注: THI4_FUSSH 镰刀根腐菌(*F. solani f. sp. phaseoli*), THI4_FUSOX 尖镰孢菌(*F. oxysporum*), THI4_HAZUM 哈茨木霉(*T. harzianum*), THI4_ASPOR 米曲霉(*A. oryzae*), Q76B84_EMENI 构巢曲霉(*A. nidulans*)。

2.3 噻唑合成酶基因 Thi4 的氨基酸水平多序列比对

从 GenBank/ EMBL/ DDBJ 数据库中下载其它噻唑生物合成酶基因,将 THI4 的氨基酸序列与数据库中其他序列进行同源比对,利用 ClustaX 软件将其中同源性较高的 4 种做多序列比对分析,结果表明噻唑生物合成酶的氨基酸长度从 320~331 不等,但其中有 213 个氨基酸完全保守,说明 Thi4 基因在进化上非常保守。

2.4 噻唑合成酶 Thi4 的氨基酸结构域预测

推导的氨基酸序列通过 Prosite 数据库查询,发现该序列包括 1 个酪氨酸激酶磷酸化位点,5 个酪蛋白激酶 II 磷酸化位点,12 个 N-豆蔻酰化位点,2 个 N-糖基化位点,3 个蛋白激酶 C 磷酸化位点。二级结构预测表明该蛋白含有 28.26% helix 和 12.11% strand。

2.5 噻唑合成酶 Thi4 的系统进化树的构建

将同源性较高的 Thi4 基因的 14 种真菌构建了系统进化树,包括米曲霉(*A. oryzae*)、裸孢壳菌(*E.*

nidulans)、镰刀根腐菌(*F. phaseoli*)、尖镰孢菌(*F. oxysporum*)、粗糙脉孢菌(*N. crassa*)、裂殖酵母(*S. pombe*)、光滑假丝酵母(*C. glabrata*CBS138)、白色念珠菌(*C. albicans*)、假囊酵母(*E. gossypii*)、乳酸克鲁维酵母(*K. lactis*NRRL Y-1140)、汉逊氏德巴利酵母菌(*D. hansenii*CBS767)2个、酿酒酵母(*S. cerevisiae*)和蚕豆锈病菌(*U. viciae-fabae*)。由进化树可以看出哈茨木霉(*T. harzianum*)与镰刀根腐菌(*F. solanif.sp. phaseoli*)、尖镰孢菌(*F. oxysporum*)、米曲霉(*A. oryzae*)和构巢曲霉(*A. nidulans*)处于同一分支上,说明这5种真菌在噻唑生物合成酶在进化上具有较近的亲缘关系,可以为哈茨木霉噻唑生物合成酶的后续研究提供有价值的参考。

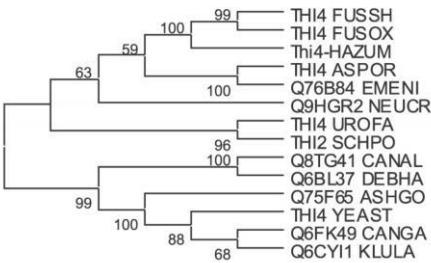


图 4 用 NJ 法构建的真菌噻唑生物合成酶氨基酸序列的分子进化树

3 结论

丝状真菌哈茨木霉(*Trichoderma harzianum*)是一

种重要的植物病害生物防治菌,通过重寄生作用、溶菌作用、产生抗性物质以及对空间和营养的竞争等多种机制的协同作用达到防治植物病原菌的目的。因此,基于分子水平对其生防机制的研究是十分必要的,该实验室构建哈茨木霉菌丝体时期 cDNA 文库而获得了大量的 ESTs 序列,通过对随机克隆的测序与分析获得基因 Thi4 的全长序列。Thi4 基因的克隆及序列分析为后续的生物学功能验证提供了参考,对于木霉菌中噻唑生物合成的机制研究具有一定的推动作用,为今后噻唑类化合物的工业化生产及新型农药的生物合成提供的理论依据。

参考文献

[1] 熊兴平. 浅谈新杀菌剂“噻菌铜”的市场前景和发展潜力[J] . 湖北植保, 2005(4): 53.
[2] 陈才俊, 宋宝安. 噻唑类杀虫剂的研究进展[J] . 农药, 2005, 44 (2): 53-55, 65.
[3] Park J H, Dorrestein P C, Zhai H, et al. Biosynthesis of the thiazole moiety of thiamin pyrophosphate (vitamin B1) [J] . Biochemistry, 2003, 42: 12430-12438.
[4] 孙宝国, 田红玉, 郑福平, 等. 肉香味含硫化合物的核心分子结构及应用[J] . 精细化工, 2005, 22(1): 29-31.
[5] Choi G H, Marek E T, Schardl L C, et al. sti35, a stress-responsive gene in Fusarium spp[J] . Journal of Bacteriology, 1990, 172: 4522-4528.
[6] Kazuto N, Mari O, Hiroyuki K, et al. Genetic regulation mediated by thiamin pyrophosphate-binding motif in Saccharomyces cerevisiae [J] . Molecular Microbiology, 2005, 58(2): 467-479.
[7] Machado C R, Praekelt U M, Meacock P A, et al. Dual role for yeast THI4 gene in thiamine biosynthesis and DNA damage tolerance [J] . Journal of Molecular Biology, 1997, 273: 114-121.

Cloning and Sequence Analysis of Gene Thi4 From *Trichoderma harzianum*

HUANG Cai-hong, YANG Qian, LIU Zhi-hua

(Department of Life Science and Engineering Harbin Institute of Technology, Harbin, Heilongjiang 150001, China)

Abstract: *Trichoderma harzianum* is a filamentous soil fungus known as an effective biocontrol agent for a wide range of economically important airborne and soilborne plant pathogens. To explore the integrated bio-control mechanism of *T. harzianum* and acquire some bio-control associated novel genes, a *T.harzianum* cDNA library has been constructed and thereby randomly selected clones were sequenced and analyzed by bioinformatics analysis. The full-length cDNA, encoding thiazole biosynthetic enzyme (THI4) with 1318bp nucleotides included the ORF of 969 bp which encoding 397 amino acids with deduced molecular weight 34. 5kDa and deduced theoretical isoelectric point 5.18. The amino acids sequences of gene Thi4 included multi-functional sites and were conservative in evolutionary development. The sequences of gene Thi4 were submitted to the GenBank and were published as accession numbers of DQ647956.

Key words: *Trichoderma harzianum*; Thiazole biosynthetic enzyme; Cloning; Sequence analysis