

稻瘟病菌 Subtilases 家族生物信息学分析 及其中一蛋白的分泌性检验

魏艺聪¹, 王宗华^{2,3}, 鲁国东²

(¹福建中医学院药理学系, 福建 福州 350108; ²福建农林大学植物保护学院, 福建 福州 350002;

³福建农林大学生命科学学院, 福建 福州 350002)

摘要:根据已有的研究结果推测稻瘟病菌 Subtilases 家族基因可能在其致病过程中具有重要的功能, 本研究以进化理论为基础, 结合一些生物信息学的软件对稻瘟病菌基因组中 Subtilases 家族基因进行了比较基因组、进化分析及功能域分析, 分析了 Subtilases 家族对稻瘟病菌致病性的重要性, 并推测家族中不同基因的结构、功能及参与的生理过程可能的异同之处, 以基因 MGG_07965.6 为例, 分析其可能具有的生物学功能, 并通过实验的方法证实了其基因表达产物确是分泌蛋白。

关键词: 稻瘟病菌; Subtilases; 分泌蛋白; 比较基因组; 进化分析

中图分类号: Q933; Q75

文献标识码: A

论文编号: 2009-2305

Bioinformatics Analysis of Subtilases in *M. grisea* and Secretion Detection of MGG_07965.6

Wei Yicong¹, Wang Zonghua^{2,3}, Lu Guodong²

(¹Department of Pharmacy, Fujian College of Traditional Chinese Medicine, Fuzhou 350108;

²College of Plant Protection, Fujian Agriculture and Forestry University, Fuzhou 350002;

³School of Life Sciences, Fujian Agriculture and Forestry University, Fuzhou 350002)

Abstract: Previous researches have suggested that the subtilases gene of *Magnaporthe grisea* may play important roles in pathogenesis. In this paper, bioinformatic tools and genome database were used to investigate the function of subtilases in *Magnaporthe grisea*. By oncogenomic and phylogenetic analyses of subtilases, we hypothesized that some of the subtilases may control *M. grisea* pathogenicity. The genes function and physiological roles were also speculated according to the genes structure. The MGG_07965.6 gene was as an example for further study. The protein product of MGG_07965.6 gene was confirmed to secrete outside of the fungal cells. The functional genomic analysis methods were tested in this study for *M. grisea* based on the evolutionary theory and bioinformatics softwares.

Key words: *Magnaporthe grisea*, subtilases, secreted proteins, oncogenomic analysis, phylogenetic analysis

0 引言

稻瘟病菌(*Magnaporthe grisea*) 是研究丝状真菌的重要模式生物, 也是引起水稻重要病害稻瘟病的病原菌。植物与病原菌相互作用的过程中, 涉及双方许多信号分子的识别与传导。在病原菌方面, 这些信号分子往往就是与植物受体蛋白起作用的激发子和其它致

病因子。分泌蛋白因其分泌到体外的特性, 最有可能是病原菌分泌到细胞外与寄主受体蛋白起作用的激发子和致病因子。

丝氨酸蛋白酶家族蛋白在烟曲霉等一些病原菌致病过程中的作用已有报道^[1-5], 迄今稻瘟病菌中丝氨酸蛋白酶家族蛋白的生物学功能尚不清楚。目前, 很多

基金项目: 国家自然科学基金资助项目(No. 30471132、30671348); 福建自然科学基金重点项目(B0520002)。

第一作者简介: 魏艺聪, 男, 1981 年出生, 福建漳浦人, 助教, 主要从事生物技术研究。通信地址: 350108 福建福州大学城福建中医学院药理学系, E-mail: yicongwei@126.com。

通讯作者: 鲁国东, 男, 1967 年出生, 福建上杭人, 教授, 博士, 主要从事分子植物病理学研究。通信地址: 350002 福建福州金山福建农林大学植物保护学院, Tel: 0591-83750663, E-mail: guodonglu@yahoo.com。

收稿日期: 2009-11-04, 修回日期: 2009-11-08。

模式生物的基因组已经被测序,其中包括 25 多种真菌,既有动植物的病原菌类也有非病原菌类。这将方便于人们从比较基因组学及进化的角度去探索某些基因或基因家族的生物学功能。因而,进一步利用稻瘟病菌和其它真菌结构基因组学的成果^[6-8]及基因组公共数据库数据与生物信息学的软件,从基因组水平分析稻瘟病菌中丝氨酸蛋白酶家族成员的数目与分布,初步分析各基因的结构特点;分析不同真菌丝氨酸蛋白酶家族成员的演化特点。这些结果有利指导我们进行该家族基因功能的实验分析。

该研究中的稻瘟病菌等真菌中丝氨酸蛋白酶(serine protease)仅指其中的枯草杆菌蛋白酶(Subtilases)家族的成员,该家族是第二大的丝氨酸蛋白酶家族,该家族基因具有独立的比较一致的 Asp/Ser/His 催化残基,类似胰岛素丝氨酸蛋白酶,其结构是一个含有 7 链平行的 β 折叠的 $\alpha\beta$ 折叠。该研究以进化理论为基础,利用生物信息学的一些软件对稻瘟病菌中 Subtilases 家族的蛋白质功能进行分析,有益于指导研究者从具有重要功能的基因入手深入对该家族基因的功能进行研究,通过反向遗传学方法证实每个基因的功能。

1 材料与方法

1.1 试验材料

稻瘟病菌菌株 70-15 由美国普渡大学许金荣博士惠赠。用于超表达的质粒 pGLH55 由鲁国东博士提供,保存于本实验室。

1.2 试验方法

1.2.1 蛋白序列获得及系统演化关系分析 Subtilase 家族蛋白序列可从公共数据库(<http://www.broadinstitute.org/science-/data#>)获得,在以上数据库中,选所要分析物种的基因组数据库,在 Gene Search 栏中输入 Subtilase,即可得到该物种基因组中所有的丝氨酸蛋白酶(Subtilase 家族)蛋白序列。

多序列比对使用 Clustal X,参数使用默认值。序列系统关系树的分析则使用 PHYLIP-3.6b 程序,首先通过软件 seqboot 进行 bootstrap 分析(100 bootstrap replications)随机取样,使用软件 protdist 中的距离系统法(Jones-Talor-Thornton Matrix)进行距离测算,使用软件 Neighbor 邻接法(neighbor-joining method)进行建树,通过软件 consense 构建一致树,系统树通过 TreeView software (version 1.6.6)绘制。

1.2.2 蛋白结构域的分析 蛋白保守结构域分析通过 Pfam (<http://www.sanger.ac.uk/Software/Pfam/search.shtml>)以及 NCBI 数据库(<http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>)进行。蛋白质定位采用 CBS

提供的信号肽预测软件 SignalP3.0 (<http://www.cbs.dtu.dk/services/SignalP/>)以及 SoftBerry 提供的蛋白亚细胞分布预测软件 ProtComp6.0 (<http://sun1.softberry.com/berry.phtml?topic=protcompan&group=programs&subgroup=proloc>)

1.2.3 蛋白分泌性检测 MGG_07965.5 过量表达质粒 pGLH55 已经由鲁国东博士所构建,该质粒携带有 MGG_07965.5 基因,该基因克隆于稻瘟菌表达质粒 pDL1 上,置于 RP27 启动子下过量表达。由于 pDL1 载体加有 RGS-HIS tag 末端,表达产物可以用 Ni-NTA 进行纯化和用 RGS-HIS tag 抗体进行 Western blot 检测。把质粒 pGLH55 通过 PEG 介导法转化到稻瘟病菌 70-15 的原生质体中,进行 PCR 筛选阳性转化子,通过在细胞外提取特异目标蛋白的方法对阳性转化子进行鉴定。

2 结果与分析

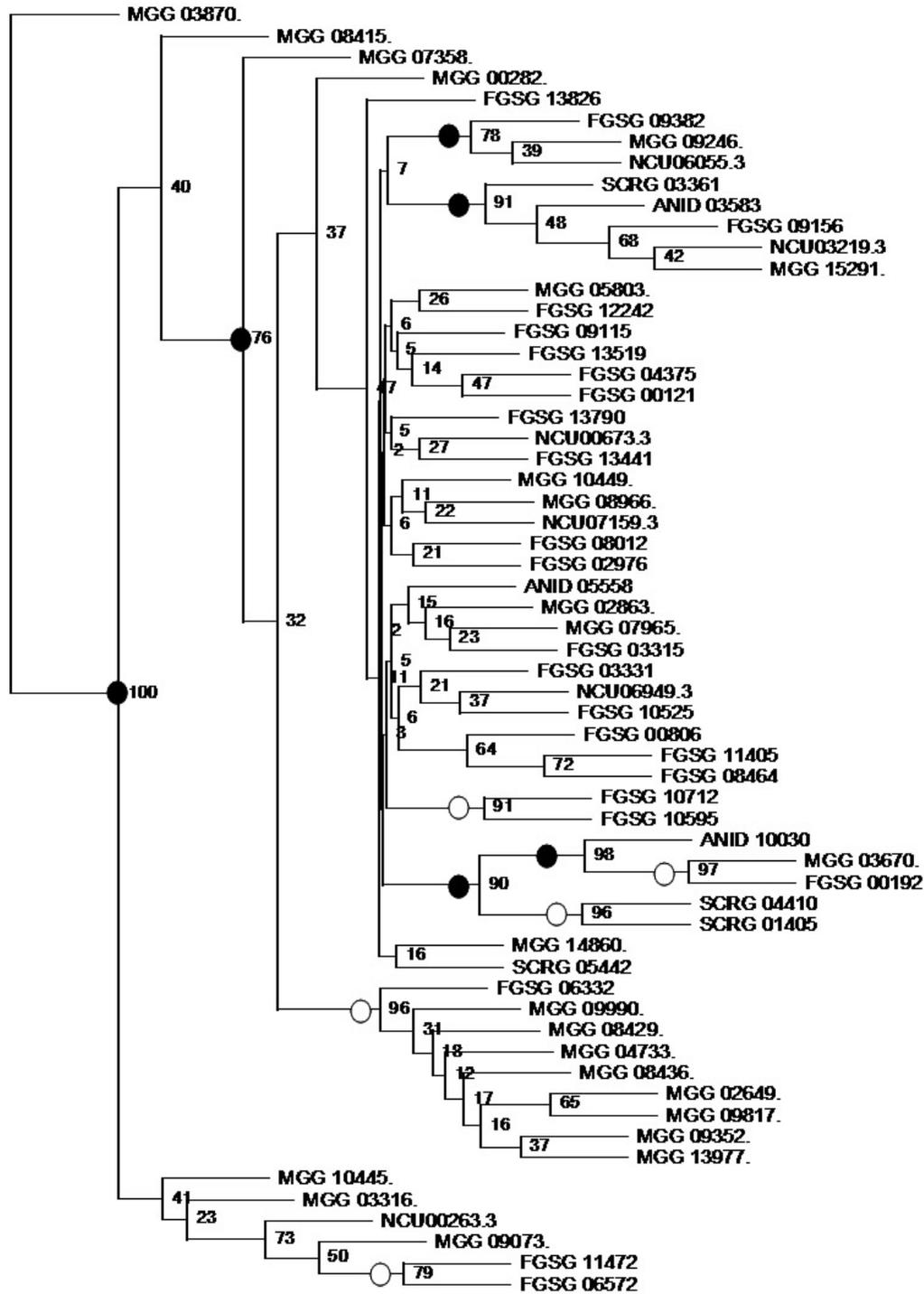
2.1 稻瘟病菌 Subtilase 家族比较基因组及进化分析

为了探索 Subtilase 家族成员的功能,在已经全基因组测序的真菌中选择了也是植物病原菌的禾谷镰刀菌,以及具有代表性的三种非植物病原菌类的构巢曲霉菌、粗糙脉孢霉菌及酿酒酵母菌的基因组,对这些真菌中的丝氨酸蛋白酶家族基因的关系进行了分析。首先分别对各个真菌的丝氨酸蛋白酶家族基因数目在全基因组水平进行了统计,然后对这些蛋白进行了进化方面的分析,构建了进化树。结果表明,稻瘟病菌被测序菌株 70-15 中具 Subtilases 结构域的基因有 24 个,同样是植物病原菌的禾谷镰刀菌被测序菌株中也有 24 个,而非植物病原菌类的构巢曲霉菌、粗糙脉孢霉菌及酿酒酵母菌的被测序菌株中分别只有 3 个、6 个及 4 个。很明显,相对于非病原菌类,丝氨酸蛋白酶家族基因数目在植物病原菌中有比较大的扩增,这或许跟这些真菌进化为植物病原菌相关。因此,丝氨酸蛋白酶家族基因很可能与稻瘟病菌致病性紧密相关。

另外,从所构建的进化树中可以看出,同种菌中的 Subtilases 家族成员并没有完全的聚类,也就是有些丝氨酸蛋白酶家族基因在进化上物种间出现直向同源的现象,因此可推测在进化过程中,丝氨酸蛋白酶家族基因在这几种真菌分化之前就已有分化。理论上,基因的结构与功能紧密相关,生理过程与相关基因的功能紧密相关,在病原菌类与非病原菌类之间,当相应基因间的结构差异足够大可引起基因功能及相关生理过程的差异,则可以推论,在病原菌类与非病原菌类之间具有种间出现直向同源进化的分支(如图 1 标●处所示)中的基因可能与真菌某些共同或类似的生理过程有

关;反之在病原菌类共生同源的分支(如图1标○处所示)基因则可能在该物种具有相对比较重要的作用。而具有基因数目较大扩增的共生同源的分支中的基因可能对该菌具有重要的作用。本研究对 Bootstrapping 值大于75的分枝进行分析,结果显示:MGG_09246.6、MGG_1529-1.6与MGG_03670.6等3个基因与其它的病原菌的基因具有直向同源趋势,因此可推测他们

在稻瘟病菌中可能执行与其他真菌共同的一些生理功能。而MGG_08429.6、MGG_13977.6、MGG_09352.6、MGG_04733.6、MG-G_09990.6、MGG_08436.6、MGG_02649.6及MGG_09817.6等8个基因与同是病原菌的禾谷镰刀菌基因具有共生同源趋势,因此可推测这几个基因可能执行稻瘟病菌致病等一些独特的生理功能,而且他们的功能可能相互重叠或互补。



10

图1 部分真菌 Sublittase 家族蛋白之系统发育树

(MGG: 稻瘟病菌; FGSG: 禾谷镰刀菌; NCU: 粗糙脉孢霉菌; ANID: 构巢曲霉菌; SCRG: 酿酒酵母菌)

2.2 稻瘟病菌中Subtilase家族成员在基因组中分布、同源性与聚类分析

据统计,稻瘟病菌中24个丝氨酸蛋白酶基因不均匀的分布于7条染色体上,其中位于第5条染色体上的最多,有8个基因;在第1条染色体上的次之,有5个。但是都不是成簇存在的。其余11个基因分别散布于其他的5条染色体上(表1)。可在稻瘟数据库中通过Blastp分析他们之间的同源性,以基因MGG_07965.6为例,这些基因当中只有18个基因与MGG_07965.6具有同源性,同源性最高是基因MGG_02863.6,它们

的氨基酸序列的同源性(配对率)高达65%。可以推测这两个基因可能具有相同的功能。另外,对稻瘟病菌中24个丝氨酸蛋白酶基因进行了进化树的分析。结果显示,稻瘟病菌中24个丝氨酸蛋白酶基因聚为3类(图2箭头所示),以MGG_07965.6为例,聚类与进化距离的远近与基因间的序列同源性分值大小具有一定的正相关性,但并不是绝对的(图2下划线处所示);而与基因在染色体中的分布并没有明显的相关性。综合基因间的同源性、进化距离的分析,可以对家族中不同基因可能执行的功能的异同情况进行推测。

表1 稻瘟病菌中可能的Subtilase家族成员在基因组中的分布

染色体	丝氨酸蛋白酶基因	总数/个
I	MGG_08415.6,MGG_10449.6,MGG_10445.6,MGG_07358.6,MGG_08966.6	5
II	MGG_09246.6	1
III	MGG_09073.6,MGG_07965.6,MGG_05803.6	3
IV	MGG_03870.6,MGG_09352.6,MGG_04733.6	3
V	MGG_09990.6,MGG_03670.6,MGG_13977.6,MGG_08429.6,MGG_08436.6,MGG_03316.6,MGG_15291.6,MGG_14860.6	8
VI	MGG_00282.6,MGG_09817.6	2
VII	MGG_02863.6,MGG_02649.6	2

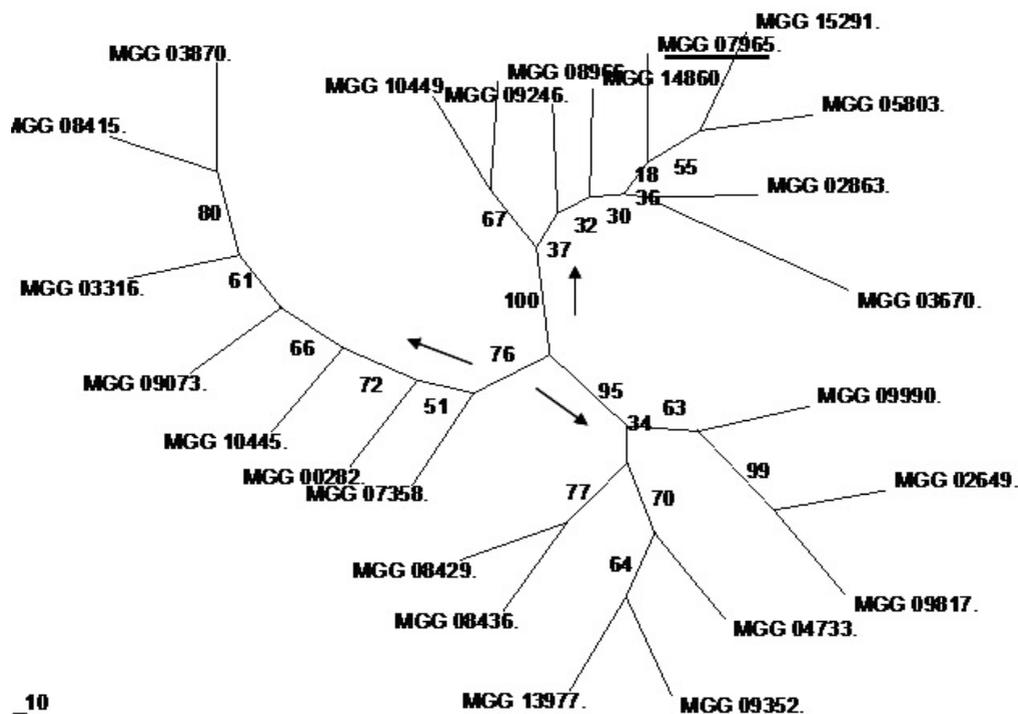


图2 稻瘟病菌中Subtilase家族蛋白之系统发育树

2.3 稻瘟病菌中丝氨酸蛋白酶家族基因的功能域分析

经SignalP3.0分析,预测稻瘟病菌中24个丝氨酸蛋白酶蛋白均含有信号肽,进一步经ProtComp6.0分析,这24个蛋白在细胞内外均有不同程度的分布。其

中分泌到体外的分值3.0以上的有8个,从大到小分别为MGG_0896.6、MGG_07965.6、MGG_10449.6、MGG_02863.6、MGG_15291.6、MGG_09246.6、MGG_14860.6和MGG_03670.6等。病菌分泌到体外的蛋白

水解酶可能在病菌与水稻互作中起到关键作用,因此这8个基因的功能可能与稻瘟病菌的致病性相关。同时,预测结果也显示这些蛋白在细胞内的各个亚细胞结构中都有不同程度的分布,这或许预示着这些蛋白酶也与稻瘟病菌的生长发育相关。另外,据分析,这24个基因中,MGG_09817.6、MGG_08436.6、MGG_-13977.6、MGG_09990.6、MGG_02469.6、MGG_07358.6与MGG_09073.6等7个基因含有Protease-associated (PA) domain,基因MGG_015291.6含有Pro-protein convertase P-domain,这两个功能域都可能跟蛋白与蛋白间的互作有关。基因MGG_10445.6还存在未知功能的结构域。

2.4 稻瘟病菌中Subtilase家族蛋白分泌性的检测

随机挑选基因MGG_07965.6(MGSP3)进行其分泌性的分析。将MGG_-07965.6过量表达质粒pGLH55通过PEG介导法转化到稻瘟病菌70-15的原生质体,获得的转化子中,通过PCR筛选,获得阳性的单孢菌落OEsp3-6-1与OEsp3-17-1等。然后在OEsp3-6-1与OEsp3-17-1的培养液中提取总蛋白质进行分析,结果显示与野生型70-15的总蛋白对比,转化子OEsp3-17-1与OEsp3-6-1的菌液总蛋白有一大小41kDa的特异条带(如图3的上图箭头所示),进一步通过Ni-NTA小量制备法提纯目标蛋白质,从这个两个转化子培养液中获得了目标蛋白质(如图3的下图箭头所示),表明外源基因表达产物已被分泌到细胞外,证实MGG_07965.6基因表达产物是分泌蛋白质。

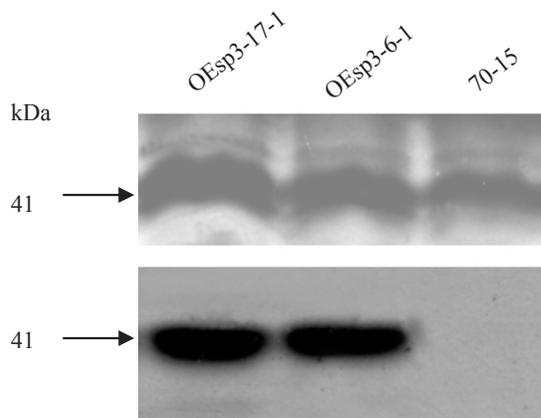


图3 稻瘟病菌MGSP3过量表达转化子目标分泌蛋白的验证
(上泳道:提取胞外总蛋白;下泳道:特异提取胞外mgsp3蛋白)

3 小结与讨论

基因组测序的完成,极大方便了人们进行功能基因组的研究。当今,常用的功能基因组分析方法有:同源比较、表达序列标签、基因表达系列分析、微阵列分析、反向遗传学方法等。在研究过程当中,这几种方法

经常综合使用,有机结合以初步确认某一个或某些基因的功能。其中第一种方法以进化理论为基础进行预测,后三种方法则是通过差异表达方法进行大规模的分析,但是只有最后的反向遗传学方法才能比较直接确认某个基因的功能。当然,前四种方法有助于反向遗传学方法进行基因功能研究的开展,指导研究者从哪些功能基因入手进行深入的研究。以进化理论为基础进行预测及利用生物信息学的一些软件进行基因功能分析是功能基因组研究的重要环节,本研究对这一环节进行探索,从比较基因组及进化分析的角度,对稻瘟病菌中Subtilases家族的蛋白质功能进行分析,得出一些推论。一是从整体上,由该基因家族在致病菌有比较大的扩增现象可以推测丝氨酸蛋白酶家族基因与病菌的致病性紧密相关;二是系统进化分析,可推测MGG_09246.6、MGG_15291.6与MGG_03670.6等三个基因在稻瘟病菌中可能执行与其他真菌共同的一些生理活动。而MGG_08429.6、MGG_13977.6、MGG_09352.6、MGG_04733.6、MGG_09990.6、MGG_08436.6、MGG_02649.6及MGG_09817.6等八个基因可能执行稻瘟病菌一些独特的生理功能,而且他们的功能可能相互重叠或互补。三是从家族中的基因之间的同源性分析,可以探索家族中不同基因可能执行的功能的异同情况,以基因MGG_07965.6为例,基因MGG_02863.6与它的氨基酸序列的同源性(配对率)高达65%,二者的生物学功能可能一致。此外,还可以通过生物信息学其他一些软件对基因的功能域进行分析,本研究发现稻瘟病菌该家族24个基因表达产物均含有信号肽,在细胞内外均有不同程度的分布。以基因MGG_07965.6为例,其分泌到体外的分值3.0以上,在细胞内也有分布,说明该基因可能不仅与稻瘟病菌的营养吸收及与水稻的互作有关,还与稻瘟病菌本身生长发育的生理过程有关。本研究还证实了基因MGG_07965.6表达产物确实分泌蛋白,而且已有的研究结果也表明稻瘟病菌致病相关的过程有丝氨酸蛋白酶基因的表达^[9-12],特别是2004年,Dean等的研究表明稻瘟病菌Guy11中的丝氨酸蛋白酶基因MGG_07965.6,在产孢培养基中的孢子内及氮胁迫下培养的菌丝中有表达^[6],说明该基因可能与产孢和致病相关。这些结果印证了对基因MGG_07965.6可能具有的生物学功能分析的结果。

另外,2001年日本北海道大学的Satoru FUKIYA等鉴定了第一个稻瘟病菌的丝氨酸蛋白酶基因SPM1(MGG_03670.6),该基因是从对附着胞形成早期的cDNA文库的序列分析中得到,并进一步在附着胞形成

阶段的EST中找到相应的基因,预测该基因编码产物属泡囊蛋白^[9]。2006年,美国N.M.Donofrio等在稻瘟菌氮胁迫相关基因研究中发现,基因MGG_03670.6在氮胁迫条件下其转录水平上调,进一步通过基因敲除方法发现该基因缺失时将导致稻瘟病菌产孢量的下降、附着胞发育减慢和致病性的大幅度减弱^[12]。同时发现基因MGG_02863.6在氮胁迫条件下其转录水平也上调,而且该基因与*M. poae*中致病相关基因同源^[12]。这些结果都预示着稻瘟病菌中Subtilases家族可能在其致病过程起着很重要的作用。

参考文献

- [1] Shen H D ,Lin W L , Tam M F , et al . Identification of vacuolar serine proteinase as a major allergen of *Aspergillus fumigatus* by immunoblotting and N-terminal amino acid sequence analysis[J]. Clin Exp Allergy, 2001,31(2) :295-302.
- [2] Iraneta S G,Duschak V G, Rodriguez S M, et al . Serine proteinases with gelatinolytic activity in an *Aspergillus fumigatus* allergenic extract[J]. J Investig Allerg Clin Immunol ,2002 ,12 (4) :257-262.
- [3] Iadarola P , Lungarella G, Martorana P A , et al . Lung injury and degradation of extracellular matrix components by *Aspergillus fumigatus* serine proteinase[J]. Exp Lung Res ,1998 ,24 (3) :233-251.
- [4] Reichard U , Cole G T , Hill T W, et al . Molecular characterization and influence on fungal development of ALP2 ,a novel serine proteinase from *Aspergillus fumigatus*[J]. Int J Med Microbiol ,2000 , 290 (6) :549-558.
- [5] Shen H D ,TamM F ,Chou H, et al . The importance of serine proteinases as aeroallergens associated with asthma[J]. Int Arch Allergy Immunol, 1999,119(4) ,259-264.
- [6] Dean R A, Talbot N J, Ebbole D J, et al. The genome sequence of the rice blast fungus *Magnaporthe grisea*[J]. Nature, 2005,434(7036): 980-986.
- [7] Mannhaupt G, Montrone C, Haase D, et al. What's in the genome of a filamentous fungus? Analysis of the *Neurospora* genome sequence[J]. Nucleic Acids Research, 2003 ,31(7):1944-1954.
- [8] Nierman W C, Pain A, Anderson M J, et al. Genomic sequence of the pathogenic and allergenic filamentous fungus *Aspergillus fumigatus*[J]. Nature, 2005,438(7071):1151-1156.
- [9] Satoru F,Takao,KuGE, et al .Identification of a Putive Vacuolar Serine Protease Gene in the Rice Blast Fungus ,*Magnaporthe grisea* [J]. Biosci,Biotechnol,Biochem.2002.66(3): 663-666.
- [10] Ebbole D J . Jin. Y et al. Gene discovery and gene expression in rice blast fungus *Magnaporthe grisea*:analysis of the expressed sequence tags[J]. PMPI, 2004,17(12):1337-1347.
- [11] Matsumura H ,Irie T,Saitoh . H. Serial analysis of gene expression (SAGE) of *Magnaporthe grisea* : genes involved in appressorium formation.[J]. Mol Gen Genomics,2003,270: 181 - 189.
- [12] Donofrio N.M., Oh Y. , Lundy R., et al. Global gene expression during nitrogen starvation in the rice blast fungus, *Magnaporthe grisea*[J]. Fungal Genetics and Biology, 2006,43(9): 605-617.