

研究报告



绿木霉 Gv29-8 丝氨酸蛋白酶 S8/S53 超家族基因特性及功能

王勇超 刘志华 安柳迪 张丛兴 王金杰*

东北林业大学 黑龙江 哈尔滨 150040

摘要:【背景】丝氨酸蛋白酶在木霉菌生物防治过程中发挥重要作用。【目的】研究绿木霉丝氨酸蛋白酶 S8/S53 超家族基因信息及其生物学功能,进而为该蛋白酶生防制剂的开发及基因改造提供理论支持。【方法】通过生物信息学分析方法,从绿木霉 Gv29-8 基因组中鉴定出 23 个丝氨酸蛋白酶基因,以少孢节丛孢菌 ATCC 24927 基因组中鉴定的 4 个丝氨酸蛋白酶基因作为对照,对这 27 个丝氨酸蛋白酶基因的特性、蛋白结构、进化地位、功能等进行预测分析。【结果】27 个基因结构差异较大,编码的蛋白具有典型的丝氨酸蛋白酶催化三联体结构,属于 S8/S53 超家族,分为 6 个亚家族,同一亚家族的蛋白酶保守区长度相近,相似性较高,催化残基附近序列比较保守。系统进化分析显示,同一亚家族丝氨酸蛋白酶聚为一类。【结论】绿木霉和少孢节丛孢菌的部分丝氨酸蛋白酶基因在结构和蛋白性质上相似性强,亲缘关系较近,均属于 S8_PCSK9_ProteinaseK_like 亚家族,推测绿木霉与少孢节丛孢菌该亚家族的丝氨酸蛋白酶具有相似的功能,可抑制植物病原真菌和降解线虫体壁。

关键词: 绿木霉, 丝氨酸蛋白酶, 生物信息学, 生物防治

Characterization of S8/S53 serine protease superfamily gene of *Trichoderma virens* Gv29-8

WANG Yong-Chao LIU Zhi-Hua AN Liu-Di ZHANG Cong-Xing WANG Jin-Jie*

Northeast Forestry University, Harbin, Heilongjiang 150040, China

Abstract: [Background] Serine protease plays an important role in the biological control of *Trichoderma* spp. [Objective] This study provided theoretical support for the development of the protease biocontrol preparation and genetic modification by studying the serine protease S8/S53 superfamily gene information and biological functions of *Trichoderma virens*. [Methods] Twenty-three serine protease genes were identified from *T. virens* Gv29-8 genome by bioinformatics analysis. Four serine protease genes identified from *Arthrobotrys oligospora* ATCC 24927 genome were used as controls. The characteristics, protein structure, evolutionary status and function of 27 serine protease genes were predicted and analyzed. [Results] The 27 genes were significantly different in structure, and the encoded proteins had typical serine protease catalytic triad structure, belonging to the S8/S53 superfamily, which were divided into 6 subfamilies. The conserved domain of proteases in the same subfamily was similar in length, with high similarity and conservative sequences near the catalytic residues. Phylogenetic analysis showed that the

Foundation items: The Youth Scientific Research Foundation (09039); Training Program of Innovation and Entrepreneurship for Undergraduate of Heilongjiang Province (201710225025)

*Corresponding author: E-mail: jinjieonly@126.com

Received: 28-04-2019; Accepted: 20-11-2019; Published online: 29-11-2019

基金项目: 青年科研启动基金(09039); 黑龙江省大学生创新创业训练计划(201710225025)

*通信作者: E-mail: jinjieonly@126.com

收稿日期: 2019-04-28; 接受日期: 2019-11-20; 网络首发日期: 2019-11-29

serine proteases of the same subfamily were clustered into one group. **[Conclusion]** Some serine protease genes of *T. virens* and *A. oligospora* have strong similarities in structure and protein properties and close genetic relationship, both belonging to the S8_PCSK9_ProteinaseK_like subfamily. It is speculated that *T. virens* and *A. oligospora* have similar functions of serine protease, which can inhibit plant pathogenic fungi and degrade the body wall of nematodes.

Keywords: *Trichoderma virens*, Serine protease, Bioinformatics, Biological control

木霉菌(*Trichoderma* spp.)广泛存在于土壤、植物根围、叶围、种子和球茎等生态环境中,是重要的植物病害生防菌^[1-2]。木霉菌产生的丝氨酸蛋白酶(serine protease)具有抗真菌和重寄生作用,以及诱导寄主植物防卫反应等作用^[3-4]。Geremia 等在诸多的研究后证明了哈茨木霉(*T. harzianum*)的丝氨酸蛋白酶 Prb1 与重寄生效应关系密切^[5]。Pozo 等发现丝氨酸蛋白酶基因 *tvsp1* 在绿木霉(*T. virens*)中的过表达显著提高了棉花幼苗抗立枯丝核菌(*Rhizoctonia solani*)的能力^[3]。

丝氨酸蛋白酶是许多真菌侵染线虫的一种关键酶,作为一种毒力因子,在降解线虫体壁的过程中发挥着非常重要的作用^[6]。近年来,人们已发现多种木霉对根结线虫具有一定的抑制作用^[7-8]。Szabó 等研究表明哈茨木霉的几丁质酶、丝氨酸蛋白酶和天冬氨酸蛋白酶等基因协同表达,在对线虫卵寄生过程中起重要作用^[9]。来自绿木霉(*T. virens*) G1-3 的培养滤液可以抑制南方根结线虫(*Meloidogyne incognita*)卵的孵化和二龄幼虫(J2s)的活动^[10]。土壤中最普遍的一种捕食性真菌少孢节丛孢菌(*Arthrobotrys oligospora*)也可以分泌丝氨酸蛋白酶,从而发挥杀线虫效应。Zhao 等扩增得到少孢节丛孢菌丝氨酸蛋白酶 P186 的基因序列,在毕赤酵母中成功表达,纯化后的蛋白酶 reP186 具有杀线虫作用^[11]。这些结果表明,丝氨酸蛋白酶在真菌生物防治过程中起到重要的作用。

鉴于此,本文利用生物信息学方法,对绿木霉 Gv29-8 菌株丝氨酸蛋白酶 S8/S53 超家族基因结构、蛋白酶理化性质、进化关系和结构等方面进行预测分析,旨在为研究该蛋白酶的结构功能、外源表达以及为绿木霉更有效地防治植物病害提供理论依据。

1 材料与方法

1.1 丝氨酸蛋白酶基因的鉴定

本实验室从中国农业微生物菌种保藏管理中心(ACCC)购买一株绿木霉 ACCC 32490 (*T. virens*),其生防效果较好,从中克隆了一个丝氨酸蛋白酶基因(GenBank 登录号为 MN399377),进一步对同种已知基因组信息的绿木霉 Gv29-8 丝氨酸蛋白酶进行研究,将该基因的 cDNA 序列与 JGI 数据库中绿木霉 Gv29-8 基因组信息(https://genome.jgi.doe.gov/TriviGv29_8_2/TriviGv29_8_2.home.html)进行 BLAST 比对,得到 23 个丝氨酸蛋白酶基因的 DNA 和 cDNA 序列,并在少孢节丛孢菌 ATCC 24927 基因组中(<https://genome.jgi.doe.gov/Artol1/Artol1.home.html>)搜索到 4 个丝氨酸蛋白酶基因,统计 27 个基因在染色体上的位置、外显子和内含子数量等。

1.2 丝氨酸蛋白酶保守结构域及家族预测

利用 NCBI 上的 BLASTp (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>)进行保守结构域及蛋白家族的预测,找出丝氨酸蛋白酶家族的保守结构域,并记录其氨基酸序列。利用 ExPASy PROSITE (<https://prosite.expasy.org/>)分析丝氨酸蛋白酶的功能区和活性位点。

1.3 丝氨酸蛋白酶基本性质预测

利用在线工具 ExPASy ProtParam (<https://web.expasy.org/protparam/>)分析蛋白的分子量、等电点、不稳定系数、总平均亲水性等理化性质。利用 SignalP4.1 (<http://www.cbs.dtu.dk/services/SignalP/>)预测蛋白的信号肽。利用 TMHMM 在线软件(<http://www.cbs.dtu.dk/services/TMHMM/>)预测蛋白的跨膜结构域。利用 CBS Prediction serves (<http://www.cbs.dtu.dk/services/TargetP/>)进行亚细胞定位预测。利用 NetNGlyc 1.0 Server (<http://www.cbs.dtu.dk>)

dk/services/NetNGlyc/)预测蛋白的 N-糖基化位点。

1.4 丝氨酸蛋白酶功能区多序列比对及进化分析

利用 BioEdit 软件对 27 个丝氨酸蛋白酶的氨基酸序列进行多序列比对, 对其功能区进行分析, 从而推测它们结构和功能上的联系。利用 MEGA 7.0 软件的 N-J 法(neighbor-joining method)构建系统进化树, Bootstrap 重复次数为 1 000 次, 对蛋白酶的进化关系进行分析。利用 DNAMAN 软件进行多序列相似性分析。

1.5 同源建模和三维结构分析

通过生物在线软件 SOPMA 和 SWISS-MODEL (<http://swissmodel.expasy.org/>)对丝氨酸蛋白酶的二级、三级结构进行预测。

2 结果与分析

2.1 丝氨酸蛋白酶保守结构域及家族预测

经过 BLASTp 比对分析, 27 个丝氨酸蛋白酶

氨基酸序列均具有丝氨酸蛋白酶家族的保守结构域, 属于丝氨酸蛋白酶家族 S8/S53 超家族。保守区预测显示, 27 个丝氨酸蛋白酶分为 S53、S8_5、S8_7、S8_Kp43_protease、S8_PCSK9_ProteinaseK_like、S8_Subtilisin_subset 这 6 个亚家族, 不同的亚家族保守区长度差别较大, 同一亚家族的保守区长度相近。其中少孢节丛孢菌的 4 个丝氨酸蛋白酶与绿木霉的 6 个蛋白酶同属于 S8_PCSK9_ProteinaseK_like 亚家族, 且保守区长度相近。

2.2 丝氨酸蛋白酶基因的结构和理化性质预测

27 个丝氨酸蛋白酶基因的内含子数量差异显著(图 1), 表明了生物体内基因的多多样性。这些基因编码蛋白的氨基酸数量差别较大, 由 265–1 023 个氨基酸组成, 蛋白分子量介于 29.10–111.69 kD (表 1)。它们的理论等电点大小不等, 除 SP41430、SP230947 外, 其余蛋白均为酸性蛋白。不稳定系数分析表

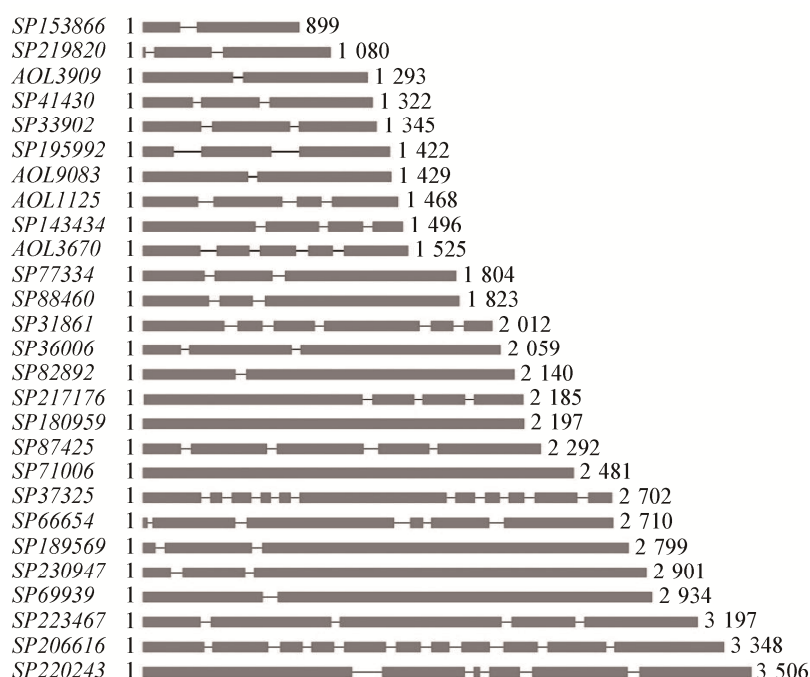


图 1 绿木霉和少孢节丛孢菌丝氨酸蛋白酶基因外显子及内含子分析

Figure 1 Exons and introns analysis of serine proteases genes in *T. virens* and *A. oligospora*

注: 方框和直线分别代表外显子和内含子, 数字表示基因的长度(bp). SP: 绿木霉丝氨酸蛋白酶基因; AOL: 少孢节丛孢菌丝氨酸蛋白酶基因。

Note: The boxes and lines represent exonic and intronic regions respectively. Numbers indicate the length of genes (bp). SP: Serine protease gene of *T. virens*; AOL: Serine protease gene of *A. oligospora*.

表 1 丝氨酸蛋白酶的理化性质

Table 1 The physicochemical properties of serine proteases

基因名称		染色体上位置		外显子		内含子		氨基酸数		等电点		分子量		保守区		信号肽切割位点		不稳定系数		总平均亲水性	
Gene name	Scaffold location	Exons	Introns	Amino acids	Isoelectric point	Molecular weight (kD)	Conservative domain	Signal peptide cleavage sites	Instability index	GRAVY											
SP153866	scaffold_78: 228 150–229 048 (+)	2	1	265	5.86	29.10	17–199	No	30.13	–0.072											
SP219820	scaffold_4: 520 483–521 562 (–)	3	2	317	5.29	35.86	35–226	No	39.92	–0.452											
AOL3909	scaffold_00076: 18 422–19 714 (+)	2	1	409	6.06	41.91	131–389	21 and 22: AFA-GP	14.90	–0.055											
SP41430	scaffold_90: 1 384 959–1 386 280 (–)	3	2	401	7.17	41.86	126–378	20 and 21: VLA-AP	24.36	–0.037											
SP33902	scaffold_1: 1 090 650–1 091 994 (+)	3	2	407	6.27	43.78	134–387	20 and 21: VTA-YP	38.28	–0.170											
SP195992	scaffold_91: 1 209 708–1 211 129 (–)	3	2	363	6.67	40.61	141–294	No	42.38	–0.537											
AOL9083	scaffold_00170: 308 483–309 911 (+)	2	1	456	6.61	48.78	160–421	20 and 21: AFA-VP	25.91	–0.336											
AOL11255	scaffold_00215: 2 387 177–2 388 644 (–)	4	3	407	9.09	43.64	126–360	21 and 22: VAA-DL	30.99	–0.292											
SP143434	scaffold_3: 1 010 561–1 012 056 (–)	4	3	437	6.24	45.21	161–415	21 and 22: AAA-VP	18.48	0.113											
AOL3670	scaffold_00075: 38 460–39 984 (+)	5	4	406	9.92	43.73	130–384	21 and 22: ASA-DL	26.68	–0.281											
SP77334	scaffold_81: 665 581–667 384 (+)	3	2	409	6.59	42.34	130–388	20 and 21: VLA-AP	27.91	–0.016											
SP88460	scaffold_89: 1 031 951–1 033 773 (+)	3	2	391	5.98	41.42	117–367	15 and 16: VLA-AP	24.37	–0.009											
SP31861	scaffold_3: 1 924 815–1 926 826 (–)	6	5	556	5.29	60.74	171–552	16 and 17: AMA-TP	39.60	–0.271											
SP36006	scaffold_2: 97 156–99 214 (–)	3	2	650	4.76	71.48	223–641	24 and 25: SLA-KH	29.21	–0.334											
SP82892	scaffold_89: 1 169 367–1 171 506 (+)	2	1	539	5.81	57.31	160–437	No	36.80	–0.253											
SP217176	scaffold_91: 1 667 563–1 669 747 (–)	4	3	612	5.01	65.77	218–606	22 and 23: ISA-SV	38.08	–0.201											
SP180959	scaffold_78: 453 609–455 805 (+)	1	0	685	5.75	74.88	246–678	20 and 21: TDS-FV	36.49	–0.389											
SP87425	scaffold_91: 1 660 370–1 662 661 (–)	5	4	637	4.81	67.22	222–630	17 and 18: AIA-VP	37.21	–0.206											
SP71006	scaffold_42: 2 406–4 886 (–)	1	0	826	5.38	86.98	303–647	No	35.67	–0.229											
SP37325	scaffold_5: 292 363–295 064 (–)	11	10	707	4.97	77.67	82–417	No	24.73	–0.285											
SP66654	scaffold_88: 1 165 289–1 167 998 (–)	6	5	790	5.67	89.25	505–712	No	38.91	–0.342											
SP189569	scaffold_3: 2 495 345–2 498 143 (–)	3	2	892	5.15	95.37	116–606	18 and 19: VAG-ST	35.60	0.000											
SP230947	scaffold_74: 162 704–165 604 (–)	3	2	878	7.14	92.73	138–610	19 and 20: VIA-RG	31.18	–0.096											
SP69939	scaffold_86: 890 256–893 189 (–)	2	1	947	5.03	101.61	153–626	18 and 19: ASS-SL	38.00	–0.162											
SP223467	scaffold_76: 225 309–228 505 (+)	5	4	987	6.10	111.69	676–881	No	48.50	–0.573											
SP206616	scaffold_91: 513 223–516 570 (+)	11	10	901	6.00	101.21	614–783	No	46.45	–0.356											
SP220243	scaffold_4: 1 888 436–1 891 941 (–)	6	5	1 023	5.66	115.83	727–943	No	39.79	–0.481											

注: +: 正链编码; -: 负链编码。
Note: +: Plus strand; -: Minus strand.

明, 除 SP195992、SP206616、SP223467 为不稳定蛋白, 其余蛋白均为稳定蛋白。总平均亲水性预测显示, 除 SP143434 和 SP189569 外, 其余蛋白均为亲水性蛋白, 总平均亲水性为-0.573。N-糖基化位点预测发现, 27 个蛋白均有 N-糖基化位点。亚细胞定位显示有 17 个蛋白为分泌蛋白, 它们均有 N-端信号肽, 主要分布在 S8_PCSK9_ProteinaseK_like、S53 和 S8_5 亚家族。跨膜区预测显示, 只有 SP41430 和 SP143434 蛋白具有跨膜区, 其余蛋白均位于膜外。

2.3 丝氨酸蛋白酶功能区多序列比对

多序列比对结果显示(图 2), 这些丝氨酸蛋白酶的保守性低, 但是在催化残基(Glu/Asp/His/Ser)及周围的序列都很保守, 说明其在丝氨酸蛋白酶功能行使上的重要性, 对酶的催化活性是必需的。其中, SP31861、SP217176、SP180959、SP36006 和 SP87425 这 5 个蛋白含有 S53 家族的 Glu-Asp-Ser 催化三联体结构, 其余蛋白除 SP153866 外, 均含有 S8 家族的 Asp-His-Ser 催化三

联体结构^[12], 这与之前的亚家族预测一致。

2.4 丝氨酸蛋白酶进化分析和相似性比较

使用 MEGA 7.0 软件对 27 个丝氨酸蛋白酶的氨基酸序列进行聚类分析, 并建立系统进化树(图 3)。结果表明, 进化树主要分为 6 支, 同一亚家族的蛋白酶聚在一起。来自少孢节丛孢菌的 4 个蛋白酶与绿木霉的 6 个蛋白酶较有规律地排列在进化树上并形成第一个分支, 表明其亲缘关系非常近。S8_PCSK9_ProteinaseK_like 和 S53 亚家族的蛋白酶位于系统进化树的后端, 处于较为进化的地位, 其中绿木霉 S8_PCSK9_ProteinaseK_like 亚家族的蛋白酶可能具有与少孢节丛孢菌的 4 个蛋白酶功能相似的杀线虫作用, 推测杀线虫作用可能是后来随蛋白酶进化产生的, 而 S53 亚家族的蛋白酶组氨酸活性位点缺失, 出现谷氨酸活性位点, 可能导致其发生进化。相似性比较发现这些蛋白酶的相似性相差较大, 但同一亚家族的蛋白酶相似性较高(表 2)。以上结果表明丝氨酸蛋白酶存在种内差

SP31861	EANVDIQYAVSM :257	-----	GTSVSSPVFAA :470
SP217176	EANLDVQNILTI :302	-----	GTSAAASPVVAS :520
SP180959	ESDIDIDIIQTL :333	-----	GTSLAAPLWAS :610
SP36006	EAMLDIIVAYPI :311	-----	GTSSAAPIVAA :573
SP87425	ESDLDFEIAIPI :308	-----	GTSMATPLFAG :551
SP41430	AYVLDSGINIAH :164	HGTHVAGTIGG :198	GTSMAHAHVSG :352
SP77334	AYVVDGINTSH :168	HGTHVSGTIGG :202	GTSMATPHVVG :361
SP33902	AYVIDTGINTKH :172	HGTHVSGIIGG :206	GTSMATPHITG :360
SP88460	AYIIDTGVDATH :155	HGTHVAGTIGS :189	GTSMATPHIVG :342
SP82892	AYVIDTGTNIDH :202	HGTHCSGTVAG :237	GTSMA SPHICG :401
SP143434	AYVVDTGIRVTH :198	HGSHVAGTIGG :232	GTSMATPHVAG :388
AOL3909	VYVIDTGIRISH :171	HGTHCAGTIAG :210	GTSMACPHVAG :226
AOL9083	VYVIDTGKREH :200	HGTHCAGTIGG :240	GTSMATPHVAG :391
AOL11255	IYIVDSGVRISH :167	HGTHVAGTAAG :202	GTSMA SPHVAG :353
AOL3670	AYVVDGVRITH :170	HGTHVAGTIAG :205	GTSMAAPHVAG :356
SP153866	-----	HGTQMANLICA :47	GSSVATAIAAG :187
SP195992	VAIIDNGVLSVS :73	HGTQMANLICA :165	GSSVATAITAG :281
SP223467	IAIIDDGIDTTL :691	HGSLVAWLVLCL :741	GESAATAIAAG :894
SP220243	IALIDNGVYSAA :740	HGTQMAQCIIYK :782	GSSVATAFAAG :936
SP206616	VCVIDTGFNPVAV :628	HGTISASIILS :664	GTSTATAAAAAG :820
SP219820	IAVLDTGVDPSN :48	HGTTSVNLVFS :84	GTSFSTAIAAG :226
SP66654	IAVLDTGVDPTD :518	HGTTLVNLIFD :554	GTSFSTAIAATG :698
SP71006	ISISDSGFDRCG :322	HGTHVCGSALG :384	GTSMATPLVAG :612
SP37325	IAIIDTGVDYTH :109	HGTHVAGIIAA :165	GTSMAAPYIAG :339
SP230947	LAVIDTGIDYTH :165	HGSHVAGIIAA :217	GTSMACPLTAA :538
SP189569	IGFIDTGIYYKH :143	HATHVAGILGM :194	GTSMATPYIAG :537
SP69939	VAIIDSGVDYRH :179	HGTHTAGIVGM :234	GTSMATPFAAG :558

图 2 27 个丝氨酸蛋白酶催化基序多序列比对

Figure 2 Multiple sequence alignment of catalytic motif of 27 serine proteases

注: 催化残基用灰色底纹显示。SP: 绿木霉丝氨酸蛋白酶; AOL: 少孢节丛孢菌丝氨酸蛋白酶。

Note: The catalytic residues are shown in gray shade. SP: Serine protease of *T. vires*; AOL: Serine protease of *A. oligospora*.

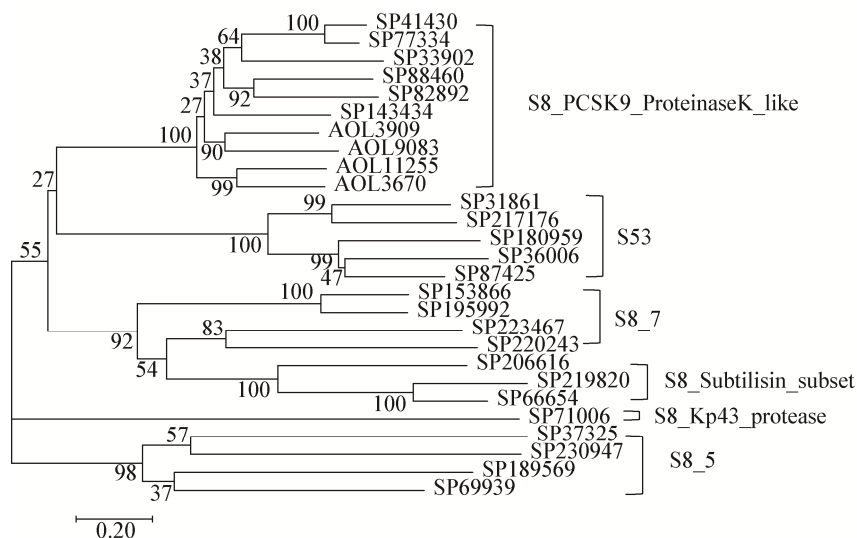


图3 MEGA 7.0 构建 27 个丝氨酸蛋白酶的系统进化树

Figure 3 Phylogenetic tree of 27 serine proteases generated by MEGA 7.0

表2 27 个丝氨酸蛋白酶的相似性比较

Table 2 Similarity comparison of 27 serine proteases

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
1	100																										
2	78.2	100																									
3	42.4	41.6	100																								
4	38.7	40.4	35.5	100																							
5	34.3	35.6	31.5	39.9	100																						
6	39.0	38.8	36.5	41.1	32.0	100																					
7	37.0	38.9	34.6	42.0	39.9	42.9	100																				
8	36.4	36.8	33.7	41.1	35.6	40.7	55.9	100																			
9	38.4	36.8	33.8	36.5	33.3	38.4	46.2	43.8	100																		
10	39.0	39.6	32.8	38.6	34.9	39.8	45.6	44.4	61.8	100																	
11	10.7	12.2	11.5	8.9	10.9	9.2	10.0	8.3	10.9	9.3	100																
12	10.6	10.9	9.2	9.1	8.8	10.0	11.0	9.5	11.0	9.8	21.2	100															
13	9.2	9.0	9.1	10.5	8.5	9.4	10.0	8.1	9.3	11.6	9.5	9.9	100														
14	7.5	8.7	9.6	10.0	9.4	9.6	11.7	11.5	11.5	12.1	9.0	8.9	25.1	100													
15	10.9	10.4	9.3	11.0	9.3	10.1	10.3	11.4	8.8	11.1	8.9	15.3	13.4	14.3	100												
16	15.2	16.6	15.0	12.3	13.0	14.9	16.1	17.5	15.3	14.7	7.7	8.0	10.1	10.0	7.5	100											
17	14.4	15.4	11.1	14.4	10.2	13.5	12.5	14.1	14.2	13.0	8.6	6.6	9.6	8.7	7.5	57.8	100										
18	14.7	14.3	11.2	13.7	12.5	12.1	12.6	11.7	11.7	11.7	8.2	7.5	8.2	6.0	8.1	19.7	15.6	100									
19	14.0	13.6	10.0	12.7	12.2	10.5	12.7	13.0	11.8	12.8	7.0	7.8	7.8	7.8	8.9	21.3	15.3	11.9	100								
20	12.8	12.2	13.4	14.2	10.8	11.6	15.0	13.7	13.7	12.5	8.1	7.2	8.1	6.1	6.6	20.5	13.4	9.7	11.6	100							
21	14.3	14.3	15.2	15.6	12.7	12.9	13.4	11.8	13.5	12.2	8.4	6.2	9.7	7.9	8.8	15.7	13.0	15.4	13.6	31.5	100						
22	11.2	10.6	13.7	14.8	12.0	11.1	12.8	12.5	12.7	11.7	7.2	5.8	7.1	6.9	8.4	17.6	11.4	9.3	11.5	18.0	53.3	100					
23	9.1	9.4	7.1	8.7	7.8	9.8	10.8	8.3	9.1	9.6	6.3	6.7	7.3	5.6	6.3	8.7	8.8	7.0	7.2	6.4	5.8	6.3	100				
24	9.6	9.2	8.0	10.6	10.4	10.0	7.9	8.2	8.9	8.4	6.5	5.5	5.2	7.3	6.0	7.5	6.6	7.4	6.1	6.4	5.4	5.6	8.2	100			
25	6.8	7.8	9.3	9.0	6.6	6.9	9.1	7.8	8.8	7.9	5.1	8.1	7.2	7.3	9.3	4.2	5.8	5.9	5.2	4.9	7.6	6.0	6.6	6.0	100		
26	7.5	7.6	7.4	10.0	8.1	7.1	6.8	8.0	7.6	8.1	6.3	8.2	7.5	4.8	8.3	6.4	8.8	7.1	5.7	5.3	4.7	7.2	7.5	6.3	8.2	100	
27	11.3	10.5	7.1	6.4	7.5	6.6	5.9	7.1	6.9	7.1	8.6	6.7	6.5	6.5	8.4	5.7	6.3	7.5	6.2	6.1	7.6	5.9	6.7	5.9	11.7	11.8	100

Note: 1: SP41430; 2: SP77334; 3: SP33902; 4: SP88460; 5: SP82892; 6: SP143434; 7: AOL3909; 8: AOL9083; 9: AOL11255; 10: AOL3670; 11: SP31861; 12: SP217176; 13: SP180959; 14: SP36006; 15: SP87425; 16: SP153866; 17: SP195992; 18: SP223467; 19: SP220243; 20: SP206616; 21: SP219820; 22: SP66654; 23: SP71006; 24: SP37325; 25: SP230947; 26: SP189569; 27: SP69939.

异性,可能产生不同的生物学特性,并且在同一亚家族内具有一定的保守性。

2.5 同源建模和三级结构分析

利用 SOPMA 在线程序预测结果表明,27 个丝氨酸蛋白酶的二级结构以 α -螺旋(alpha helix)、无规则卷曲(random coil)、延伸链(extended strand)为主,超过 85%,而 β -转角(beta turn)少于 15%。从每个亚家族选择一种蛋白酶,利用 SWISS-MODEL 获得蛋白质的三级结构(图 4),A、B、C、D、E、F 分别是对 SP69939、SP71006、SP66654、SP217176、SP77334、AOL3909 的结构预测。由于 S8_7 亚家族丝氨酸蛋白酶未具有较为适合的同源建模模板或模板相似性较低,因此未能构建模拟出三级结构的模型。预测结果显示,这些蛋白质三级结构的相似度较低,但它们每个结构域均含有至少一个折叠片,构成一个折叠桶,活性部位位于两个折叠桶之间,为 Asp-His-Ser 或 Glu-Asp-Ser 催化三联体。

SP77334 和 AOL3909 的三维结构高度相似,两个 Ca^{2+} 结合位点和催化三联体的位置基本相同,说明它们很可能具有相似的功能,这与之前的分析

结果一致,推测绿木霉 S8_PCSK9_ProteinaseK_like 亚家族的蛋白酶和少孢节丛孢菌的丝氨酸蛋白酶功能相似,均对线虫有抑制作用。

3 讨论与结论

丝氨酸蛋白酶家族成员具有多种多样的生理功能,在生物机体中发挥着重要的生理作用,具有广泛的研究和应用价值^[13]。本研究从绿木霉 Gv29-8 菌株的基因组信息中得到了 27 个丝氨酸蛋白酶基因的信息,它们的长度相差较大,内含子数量差异显著。这些基因编码的丝氨酸蛋白酶分为 6 个亚家族,同一亚家族的蛋白酶保守区长度相近,理化性质相似,说明同一亚家族的蛋白酶具有一定的保守性。

丝氨酸蛋白酶家族氨基酸序列差异较大,某些蛋白质序列的相似度甚至低于 20%,但该蛋白质的活性结构域却极其保守,都有一个类似于口袋状的核心催化结构域 Asp-His-Ser^[14]。通过进一步对 27 个丝氨酸蛋白酶的结构研究发现,虽然它们的相似度较低,但在催化残基(Asp/His/Ser 或

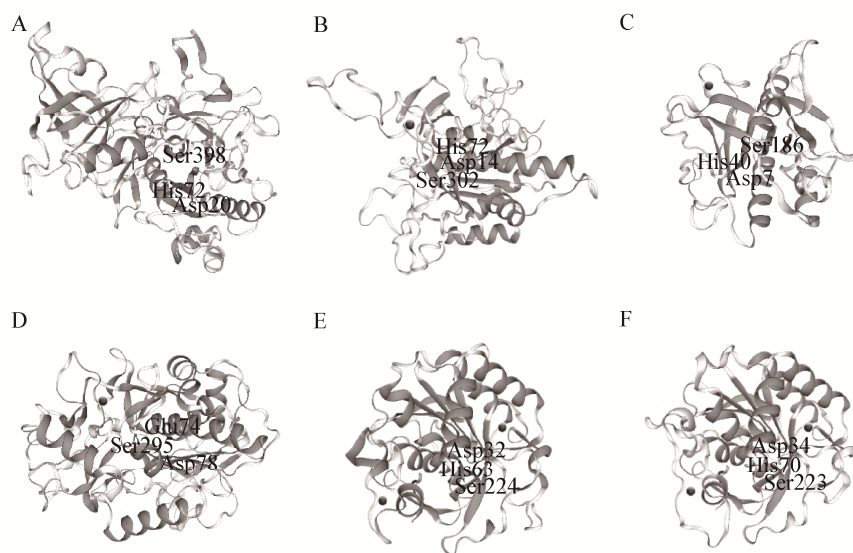


图 4 丝氨酸蛋白酶三级结构模型

Figure 4 The model on tertiary structure of serine proteases

注:圆球代表 Ca^{2+} 结合位点。A: SP69939; B: SP71006; C: SP66654; D: SP217176; E: SP77334; F: AOL3909。

Note: The Ca^{2+} binding sites are shown as spheres. A: SP69939; B: SP71006; C: SP66654; D: SP217176; E: SP77334; F: AOL3909.

Glu/Asp/Ser)及周围的序列都很保守,同一亚家族的蛋白酶保守性更高。正是由于丝氨酸蛋白酶在结构上的微小变化,导致其在功能上的进化^[13]。在少孢节丛孢菌的丝氨酸蛋白酶中也有相似的高度保守片段,与绿木霉 S8_PCSK9_ProteinaseK_like 亚家族的蛋白酶结构相似,这说明真菌丝氨酸蛋白酶基因在物种进化上具有很强的保守性,并且暗示着它们可能在某类重要途径中具有相似的功能。少孢节丛孢菌的这 4 个丝氨酸蛋白酶均具有杀线虫作用,而且已报道的体壁降解酶主要是 S8_PCSK9_ProteinaseK_like 和 S8_Subtilisin_subset 亚家族中的成员,是真菌重要的毒力因子^[11,15-19]。系统进化树将同一亚家族的蛋白酶聚在一起,亲缘关系较近,并且相似性分析显示,同一亚家族的蛋白酶相似性更高,表明丝氨酸蛋白酶可能由于催化基序的氨基酸残基发生变化,导致不同亚家族的蛋白酶具有独特的功能。已有研究表明,寄生真菌能够通过表达其他蛋白酶而补偿由于寄主产生丝氨酸蛋白酶抑制剂而损失的蛋白酶活力^[20-21],一些蛋白酶的相似性不高,但都具有丝氨酸蛋白酶结构域和保守的催化基序,表明它们的存在可能是为了补偿侵染性蛋白酶的酶活力。同源建模分析显示,不同亚家族的丝氨酸蛋白酶在三维结构上差别较大,但都具有催化三联体结构和 β 折叠桶,它们结构的差异可能是这些酶具有独特水解作用的原因。丝氨酸蛋白酶的活性及稳定性主要和 Ca^{2+} 结合位点有关,SP69939 没有预测到 Ca^{2+} 结合位点,SP77334 和 AOL3909 具有 2 个 Ca^{2+} 结合位点,SP71006、SP66654 和 SP217176 具有一个 Ca^{2+} 结合位点,距酶活性中心有一定距离,有助于稳定构象^[22]。

丝氨酸蛋白酶广泛存在于捕食线虫真菌中,其杀线虫作用已被广泛研究,这些蛋白酶大多为生化特性相似、序列相似的底物酶,是对线虫体识别作用中重要的酶类之一,其中类枯草杆菌(subtilisin-like)丝氨酸蛋白酶在捕食线虫真菌对线虫的致病性进化过程中可能是至关重要的^[18,23]。

越来越多的研究证明,木霉丝氨酸蛋白酶 S8/S53 超家族的基因在木霉抗真菌和杀线虫过程中起着重要的作用^[3,10,24-28],而且它们大部分属于 S8_PCSK9_ProteinaseK_like 亚家族。在该研究中,绿木霉有 6 个丝氨酸蛋白酶和少孢节丛孢菌的 4 个丝氨酸蛋白酶同属于 S8_PCSK9_ProteinaseK_like 亚家族,它们的相似性较高,多序列比对显示它们的催化基序十分保守,系统进化树将它们归为一支,并且三级结构非常相似。少孢节丛孢菌的 4 个丝氨酸蛋白酶均为已报道的体壁降解丝氨酸蛋白酶,它们都参与到侵染、降解线虫体壁的过程,可能发挥类似的作用^[11,15,18,21,29],绿木霉的 SP77334 蛋白酶为已报道的 Tvsp1 丝氨酸蛋白酶^[3]。综合以上结果,我们推测绿木霉 S8_PCSK9_ProteinaseK_like 亚家族的蛋白酶可能具有降解细胞壁或体壁,杀菌或杀线虫的作用。

目前对另外几个亚家族的丝氨酸蛋白酶功能报道较少,也无法预测其具体功能,还需要通过基因克隆及表达分析来验证。Szabó 等发现哈茨木霉 S53 亚家族的 P5216 蛋白酶与其他蛋白酶共同表达,可能在线虫卵寄生过程中起关键作用^[9]。虽然这个家族和 S8 家族属于同一个超家族,但是两个家族之间的催化机制是非常不同的^[30]。来自芽孢杆菌(*Bacillus* sp.) KSM-KP43 菌株的丝氨酸蛋白酶 KP-43 属于 S8_Kp43_protease 亚家族,由于其具有水解活性,可用作洗涤剂酶^[31]。但经过与绿木霉丝氨酸蛋白酶基因比对,这两个蛋白与该亚家族蛋白酶相似性较低,无法判断绿木霉丝氨酸蛋白酶是否具有相似的功能。

来自绿木霉的 23 个基因编码的蛋白质均具有丝氨酸蛋白酶的催化残基,它们可能作为一种重要的毒力因子在抑制植物病原真菌和降解线虫体壁的过程中发挥着重要的作用,但其结构的不同导致其与底物结合部位的差异决定了各自对底物的专一性。了解绿木霉丝氨酸蛋白酶 S8/S53 超家族的生物学信息后,将为进一步研究丝氨酸蛋白酶的生物学功能和开发高效、稳定的生防制剂奠定理论基础。

础,对病原物的生物防控也具有重要的理论和实践意义。

REFERENCES

- [1] Schuster A, Schmoll M. Biology and biotechnology of *Trichoderma*[J]. Applied Microbiology and Biotechnology, 2010, 87(3): 787-799
- [2] Qualhato TF, Lopes FAC, Steindorff AS, et al. Mycoparasitism studies of *Trichoderma* species against three phytopathogenic fungi: evaluation of antagonism and hydrolytic enzyme production[J]. Biotechnology Letters, 2013, 35(9): 1461-1468
- [3] Pozo MJ, Baek JM, García JM, et al. Functional analysis of *tvsp1*, a serine protease-encoding gene in the biocontrol agent *Trichoderma virens*[J]. Fungal Genetics and Biology, 2004, 41(3): 336-348
- [4] Ferre FS, Santamarina MP. Efficacy of *Trichoderma harzianum* in suppression of *Fusarium culmorum*[J]. Annals of Microbiology, 2010, 60(2): 335-340
- [5] Geremia RA, Goldman GH, Jacobs D, et al. Molecular characterization of the proteinase-encoding gene, *prb1*, related to mycoparasitism by *Trichoderma harzianum*[J]. Molecular Microbiology, 1993, 8(3): 603-613
- [6] Tang LY, Yang L, Xing XJ, et al. Advances in pathogenicity-related serine proteases in nematophagous fungi[J]. Microbiology China, 2017, 44(4): 911-919 (in Chinese)
汤丽燕, 杨乐, 邢新婧, 等. 食线虫真菌致病相关丝氨酸蛋白酶的研究进展[J]. 微生物学通报, 2017, 44(4): 911-919
- [7] Daragó A, Szabó M, Hrács K, et al. *In vitro* investigations on the biological control of *Xiphinema index* with *Trichoderma* species[J]. Helminthologia, 2013, 50(2): 132-137
- [8] Affokpon A, Coyne DL, Htay CC, et al. Biocontrol potential of native *Trichoderma* isolates against root-knot nematodes in West African vegetable production systems[J]. Soil Biology and Biochemistry, 2011, 43(3): 600-608
- [9] Szabó M, Urbán P, Virányi F, et al. Comparative gene expression profiles of *Trichoderma harzianum* proteases during *in vitro* nematode egg-parasitism[J]. Biological Control, 2013, 67(3): 337-343
- [10] Meyer SLF, Massoud SI, Chitwood DJ, et al. Evaluation of *Trichoderma virens* and *Burkholderia cepacia* for antagonistic activity against root-knot nematode, *Meloidogyne incognita*[J]. Nematology, 2000, 2(8): 871-879
- [11] Zhao HL, Qiao J, Meng QL, et al. Expression of serine proteinase P186 of *Arthrobotrys oligospora* and analysis of its nematode-degrading activity[J]. Antonie Van Leeuwenhoek, 2015, 108(6): 1485-1494
- [12] Yan HB. Genome-wide analysis of proteases and inhibitors sequences identified through bioinformatics data mining in four tapeworms[D]. Beijing: Doctoral Dissertation of Chinese Academy of Agricultural Sciences, 2013 (in Chinese)
- [13] 闫鸿斌. 四种绦虫蛋白酶及其抑制剂的系统挖掘与功能分析[D]. 北京: 中国农业科学院博士学位论文, 2013
- [13] Wang SH, Wang WY, Huang YZ, et al. On serine protease[J]. Fujian Journal of Agricultural Sciences, 2007, 22(4): 453-456 (in Chinese)
汪世华, 王文勇, 黄益洲, 等. 丝氨酸蛋白酶研究进展[J]. 福建农业学报, 2007, 22(4): 453-456
- [14] Ji XL, Sun ZR. Structure based evolution analysis of the serine protease superfamily[J]. Acta Electronica Sinica, 2001, 29(12A): 1756-1758 (in Chinese)
季星来, 孙之荣. 基于结构的丝氨酸蛋白酶超家族进化分析[J]. 电子学报, 2001, 29(12A): 1756-1758
- [15] Wang JW, Meng QL, Qiao J, et al. The recombinant serine protease *XAoz1* of *Arthrobotrys oligospora* exhibits potent nematocidal activity against *Caenorhabditis elegans* and *Haemonchus contortus*[J]. FEMS Microbiology Letters, 2013, 344(1): 53-59
- [16] Zhao HL, Meng QL, Qiao J, et al. Cloning and bioinformatics analysis of associated gene of serine protease family of *Arthrobotrys oligospora*[J]. Southwest China Journal of Agricultural Sciences, 2015, 28(1): 414-418 (in Chinese)
赵海龙, 孟庆玲, 乔军, 等. 少孢节丛孢菌丝氨酸蛋白酶家族基因的克隆及生物信息学分析[J]. 西南农业学报, 2015, 28(1): 414-418
- [17] Wang RB, Yang J, Lin C, et al. Purification and characterization of an extracellular serine protease from the nematode-trapping fungus *Dactylella shizishanna*[J]. Letters in Applied Microbiology, 2006, 42(6): 589-594
- [18] Åhman J, Ek B, Rask L, et al. Sequence analysis and regulation of a gene encoding a cuticle-degrading serine protease from the nematophagous fungus *Arthrobotrys oligospora*[J]. Microbiology, 1996, 142(7): 1605-1616
- [19] Li YS. The molecular mechanism underlying resistance against entomopathogenic fungi by protease inhibitors in *Bombyx mori*[D]. Chongqing: Doctoral Dissertation of Southwest University, 2013 (in Chinese)
李游山. 家蚕蛋白酶抑制剂抵御昆虫致病性真菌入侵的分子机理研究[D]. 重庆: 西南大学博士学位论文, 2013
- [20] Zhao ML. Study on gene cloning and expression of infectious extracellular proteases from nematophagous fungi[D]. Kunming: Doctoral Dissertation of Yunnan University, 2003 (in Chinese)
赵明莲. 食线虫真菌侵染性胞外蛋白酶的基因克隆与表达[D]. 昆明: 云南大学博士学位论文, 2003
- [21] Yang JK, Wang L, Ji XL, et al. Genomic and proteomic analyses of the fungus *Arthrobotrys oligospora* provide insights into nematode-trap formation[J]. PLoS Pathogens, 2011, 7(9): e1002179
- [22] Wang N. Heterogeneous expression of protease gene *SS42* from *Trichoderma atroviride* and characterization of

- recombinant protein in biocontrol progress[D]. Harbin: Master's Thesis of Northeast Forestry University, 2014 (in Chinese)
- 王娜. 深绿木霉蛋白酶基因 SS42 异源表达及重组蛋白生防特性分析[D]. 哈尔滨: 东北林业大学硕士学位论文, 2014
- [23] Han HB. Induction of traps and analysis of serine protease gene sequence of *Arthrobotrys oligospora*[D]. Hohhot: Master's Thesis of Inner Mongolia Agricultural University, 2008 (in Chinese)
- 韩海滨. 少孢节丛孢菌捕食器诱导及丝氨酸蛋白酶基因序列分析[D]. 呼和浩特: 内蒙古农业大学硕士学位论文, 2008
- [24] Chen LL, Liu LJ, Shi M, et al. Characterization and gene cloning of a novel serine protease with nematocidal activity from *Trichoderma pseudokoningii* SMF2[J]. FEMS Microbiology Letters, 2009, 299(2): 135-142
- [25] Yan L, Qian Y. Cloning and heterologous expression of SS10, a subtilisin-like protease displaying antifungal activity from *Trichoderma harzianum*[J]. FEMS Microbiology Letters, 2009, 290(1): 54-61
- [26] Liu ZH, Yang XX, Sun DM, et al. Expressed sequence tags-based identification of genes in a biocontrol strain *Trichoderma asperellum*[J]. Molecular Biology Reports, 2010, 37(8): 3673-3681
- [27] Vieira PM, Coelho AS, Steindorff AS, et al. Identification of differentially expressed genes from *Trichoderma harzianum* during growth on cell wall of *Fusarium solani* as a tool for biotechnological application[J]. BMC Genomics, 2013, 14: 177
- [28] Liu Y, Yang Q. Cloning and heterologous expression of serine protease SL41 related to biocontrol in *Trichoderma harzianum*[J]. Journal of Molecular Microbiology and Biotechnology, 2013, 23(6): 431-439
- [29] Yang JK, Liang LM, Li J, et al. Nematicidal enzymes from microorganisms and their applications[J]. Applied Microbiology and Biotechnology, 2013, 97(16): 7081-7095
- [30] Elad Y, Kapat A. The role of *Trichoderma harzianum* protease in the biocontrol of *Botrytis cinerea*[J]. European Journal of Plant Pathology, 1999, 105(2): 177-189
- [31] Saeki K, Ozaki K, Kobayashi T, et al. Detergent alkaline proteases: enzymatic properties, genes, and crystal structures[J]. Journal of Bioscience and Bioengineering, 2007, 103(6): 501-508