

聂燕芳, 黄嘉瑶, 周玲苑, 等. 香蕉枯萎病菌热带 4 号小种基因组规模分泌蛋白的预测与分析[J]. 江苏农业学报, 2017, 33(2): 288-294.

doi: 10.3969/j.issn.1000-4440.2017.02.008

香蕉枯萎病菌热带 4 号小种基因组规模分泌蛋白的预测与分析

聂燕芳¹, 黄嘉瑶², 周玲苑², 涂晓欢², 陈慧妍², 王振中², 李云锋²

(1. 华南农业大学材料与能源学院, 广东 广州 510642; 2. 华南农业大学农学院/广东省微生物信号与作物病害重点实验室, 广东 广州 510642)

摘要: 为了明确香蕉枯萎病菌热带 4 号小种(*Fusarium oxysporum* f. sp. *cubense* tropical race 4, Foc TR4)与香蕉互作的分子机理, 本研究利用 SignalP、WoLF PSORT、TargetP、TMHMM 和 big-PI Predictor 等软件, 对 Foc TR4 全基因组 22 487 条蛋白质氨基酸序列进行了经典分泌蛋白的预测分析。结果表明, Foc TR4 全基因组编码蛋白质中有 1 054 个经典分泌蛋白, 占编码蛋白质总数的 4.7%。蛋白质特征分析结果表明, 其氨基酸序列长度集中在 100~500 个氨基酸, 信号肽长度集中在 17~22 个氨基酸, 信号肽切割位点以 SPase I 型为主。功能预测分析结果表明, 有 463 个经典分泌蛋白获得了注释, 主要涉及糖代谢、水解酶活性等。碳水化合物酶类(CAZymes)的分析结果表明, 有 281 个分泌蛋白为 CAZymes, 其中以 GH 家族最多。此外, 利用 SecretomeP 软件对 Foc TR4 非经典分泌蛋白进行了分析, 发现有 9 216 个蛋白质氨基酸序列, 占编码蛋白质总数的 41.0%。

关键词: 尖孢镰刀菌古巴专化型 4 号小种; 香蕉枯萎病; 分泌蛋白; 碳水化合物酶类

中图分类号: S435.111.41 **文献标识码:** A **文章编号:** 1000-4440(2017)02-0288-07

Genome-scale prediction and analysis of the secretory proteins of *Fusarium oxysporum* f. sp. *cubense* tropical race 4

NIE Yan-fang¹, HUANG Jia-yao², ZHOU Lin-wan², TU Xiao-huan², CHEN Hui-yan², WANG Zhen-zhong², LI Yun-feng²

(1. College of Materials and Energy, South China Agricultural University, Guangzhou 510642, China; 2. Guangdong Province Key Laboratory of Microbial Signals and Disease Control/College of Agriculture, South China Agricultural University, Guangzhou 510642, China)

Abstract: To better understand the molecular mechanisms of the *Fusarium oxysporum* f. sp. *cubense* tropical race 4 (Foc TR4)-banana interactions, a refined Foc TR4 secretome was predicted by combining bioinformatic approaches, including SignalP, WOLFPT, TargetP, TMHMM, and big-PI Predictor. Among the 22 487 amino acid sequences of Foc TR4, 1 054 were identified, accounting for 4.7% of the total proteins. The amino acids sequences contained 100 to 500 amino acids. The signal peptides included 17 to 22 amino acids, and the cleavage sites was dominated by SPaseI type. Among the 1 054 proteins, 463 were annotated for the involvement in carbohydrate metabolic process and hydrolase activity. 281 proteins were carbohydrate-active enzymes dominated by glycoside hydrolases superfamily.

In addition, 9 216 non-classically secretory proteins were identified by SecretomeP software, accounting for approximately 41.0% of the total proteins.

Key words: *Fusarium oxysporum* f. sp. *cubense* race 4; banana *Fusarium* wilt; secretory protein; carbohydrate-active enzyme

收稿日期: 2016-10-13

基金项目: 国家自然科学基金项目(31600663); 广东省科技计划项目(2016A020210098); 广东省普通高校青年创新人才项目[粤教科函(2015)3号]

作者简介: 聂燕芳(1979-), 女, 广东博罗人, 博士, 讲师, 研究方向为蛋白质组学。

通讯作者: 李云锋, (E-mail) yunfengli@scau.edu.cn

香蕉枯萎病是由尖孢镰刀菌古巴专化型(*Fusarium oxysporum* f. sp. *cubense*, Foc)引起的一种土传维管束病害,严重威胁着香蕉产业的可持续发展。在中国广东省香蕉主产区,香蕉枯萎病病株率一般为10%~40%,严重的超过90%,甚至造成全园毁灭、蕉园丢荒^[1-4]。Foc共有4个小种,其中以热带4号小种(Foc tropical race 4, Foc TR4)致病范围最广,几乎能侵染所有香蕉种类,尤其是能侵染抗Foc1的所有Cavendish(AAA)品系,使香蕉产业面临重大威胁^[5]。由于Foc TR4遗传背景复杂,其遗传机制虽一直得到研究者的重视,但其致病机理目前仍不清楚^[6]。

分泌蛋白(Secreted protein)作为一类重要的致病因子,已被证实在枯萎病菌侵染香蕉的过程中起着重要作用。例如,Foc分泌的多种角质酶和细胞壁降解酶(Cell wall degrading enzymes, CWDEs)对香蕉组织具有降解破坏作用^[7-9]。贾慧升等^[9]应用生物信息学方法对Foc1和Foc TR4的细胞壁降解酶(CWDEs)基因进行了分析,结果表明Foc基因组中存在编码各种CWDEs的基因,其中56%的预测基因可以分泌到胞外而发挥作用。董章勇等^[10]对Foc1和Foc TR4的CWDEs进行了分析,认为2个生理小种间CWDEs的差异可能是导致2个小种致病性差异的原因之一。但迄今为止,我们对Foc TR4分泌蛋白质的组成等仍缺乏系统的了解。

近年来,应用生物信息学技术对重要植物病原菌的全基因组数据进行挖掘,预测分析其分泌蛋白质的组成和功能等已成为解析病原菌致病分子机理的重要方法之一。目前,已有许多植物病原真菌的分泌蛋白质得到了预测分析^[11-14]。随着Foc TR4基因组测序及基因组注释的完成^[8],为开展Foc TR4分泌蛋白的研究带来了新的契机。本研究结合SignalP、WoLF PSORT、TMHMM、big-PI Predictor、TargetP和SecretomeP等软件,对Foc TR4分泌蛋白进行了预测分析,拟为进一步了解Foc TR4与香蕉互作的分子机理等提供参考。

1 材料与方法

1.1 Foc TR4 基因组编码蛋白质的数据来源

用于Foc TR4分泌蛋白分析的22 487个氨基酸序列来源于数据库(http://www.broadinstitute.org/annotation/genome/fusarium_group/MultiDownloads.html)。

1.2 Foc TR4 经典分泌蛋白的生物信息学分析

1.2.1 经典分泌蛋白的预测标准 参考聂燕芳等^[15]的方法,经典分泌蛋白需同时满足以下4个特征:(1)具有N-端信号肽;(2)亚细胞定位为胞外分泌类型,不含有细胞质膜、细胞质及其他亚细胞预测定位信号;(3)无跨膜结构域;(4)无GPI锚定位点。

1.2.2 Foc TR4 经典分泌蛋白的软件预测分析 利用SignalP 4.1 Server (<http://www.cbs.dtu.dk/services/SignalP/>)^[16],对Foc TR4基因组编码的22 487个蛋白质氨基酸序列进行N-端信号肽的预测分析。利用WoLF PSORT (http://www.genscript.com/psort/wolf_psort.html)^[17]和TargetP 1.1 Server (<http://www.cbs.dtu.dk/services/TargetP/>)^[18],对具有N-端信号肽的蛋白质进行亚细胞定位分析。利用TMHMM 2.0 Server (<http://www.cbs.dtu.dk/services/TMHMM/>),对具有N-端信号肽且定位于胞外的蛋白质进行跨膜结构域分析^[19]。利用big-PI Predictor (http://mendel.imp.ac.at/sat/gpi/gpi_server.html)^[20],对具有N-端信号肽、定位于胞外、且不含有跨膜结构域的蛋白质进行GPI锚定位点分析。利用LipoP 1.0 Server (www.cbs.dtu.dk/services/LipoP/)对经典分泌蛋白信号肽切割位点的氨基酸组成进行分析^[21]。

1.3 经典分泌蛋白的功能预测及碳水化合物酶类分析

利用UniProt数据库(<http://www.uniprot.org/>),对Foc TR4经典分泌蛋白进行功能预测分析。利用CAZy数据库(<http://www.cazy.org/>),对Foc TR4经典分泌蛋白与碳水化合物酶类(Carbohydrate-active enzymes, CAZymes)序列进行BlastP同源性比对,分析Foc TR4经典分泌蛋白中的CAZymes(具体参数为 E 值 $\leq 10^{-5}$ 、且序列一致性 $\geq 30\%$),并对CAZymes进行家族功能的分类分析^[22-23]。

1.4 Foc TR4 非经典分泌蛋白的预测分析

利用SecretomeP 2.0 Server (<http://www.cbs.dtu.dk/services/SecretomeP/>)对不含有N端信号肽的蛋白质氨基酸序列进行非经典分泌蛋白预测分析,将氨基酸序列 $NN\text{-score} > 0.5$ (人工神经网络算法, Neural network)的蛋白质定义为非经典分泌蛋白^[24]。

2 结果与分析

2.1 Foc TR4 经典分泌蛋白的预测分析

2.1.1 基于 N-端信号肽的分泌蛋白预测分析 具有 N-端信号肽是判断经典分泌蛋白的必要条件,其中 SignalP 是一种常用的预测真核生物信号肽的在线预测软件。本研究利用 SignalP 4.1 Server 对 Foc TR4 基因组编码的 22 487 条蛋白质氨基酸序列进行了信号肽预测分析,结果表明有 1 793 个氨基酸序列具有分泌型信号肽,占总数的 8.0%。

2.1.2 基于亚细胞定位的分泌蛋白预测分析 利用 WoLF PSORT 软件,对具有信号肽的 1 793 个分泌型蛋白质进行亚细胞定位预测分析。结果表明有 1 401 个蛋白质分泌到胞外,属于胞外分泌蛋白类型。有 392 个蛋白质不分泌到胞外,而是转运到各亚细胞器中,其中转运到细胞质有 83 个、线粒体有 92 个、细胞核有 39 个、细胞质膜有 169 个、其他细胞器有 9 个,所占比例分别为 21.2%、23.5%、9.9%、43.1% 和 2.3%,这些转运到各亚细胞器的蛋白质不属于经典的分泌蛋白。

2.1.3 基于亚细胞定位的分泌蛋白验证分析 为了增加经典分泌蛋白预测的精确性,本研究利用 TargetP 1.1 Server 进一步对 WoLF PSORT 预测的 1 401 个分泌到胞外的蛋白质进行了验证分析。结果表明,含有分泌型信号肽的蛋白质有 1 370 个,所占比例为 97.8%,含有线粒体靶向肽和叶绿体转运肽的蛋白质则分别有 23 个和 8 个。

2.1.4 基于跨膜结构域的分泌蛋白预测分析 含有分泌型信号肽的蛋白质若同时含有跨膜区,则表明其可能为膜受体、膜锚定蛋白或离子通道蛋白,而不是经典分泌蛋白^[22]。本研究利用 TMHMM Server v2.0 对上述 1 370 个含有分泌型信号肽的蛋白质进行了跨膜结构域分析。结果表明有 1 184 个(所占比例为 86.4%)蛋白质没有跨膜螺旋,具有典型分泌蛋白的特征。此外,有 148 个蛋白质氨基酸序列含有 1 个跨膜螺旋区(所占比例分别为 10.8%),38 个蛋白质氨基酸序列有 2 个以上的跨膜螺旋区(所占比例分别为 2.8%),不属于经典的分泌蛋白。

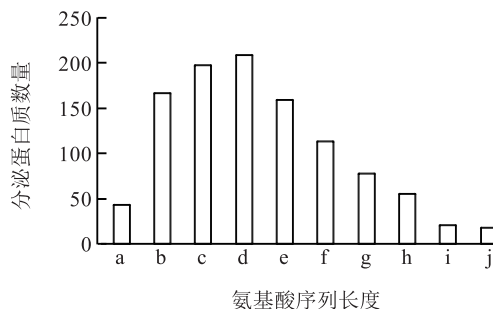
2.1.5 基于 GPI 锚定位点的分泌蛋白预测分析 利用 big-PI Predictor 软件对上述 1 184 个没有跨膜螺旋的蛋白质进行分析。结果表明有 1 054 个蛋白质氨基酸序列不具有 GPI 锚定位点,具有经典分泌

蛋白的特征,另有 130 个蛋白质具有 GPI 锚定位点,不符合经典分泌蛋白的特征。

综合上述不同生物信息学软件的分析结果,发现 Foc TR4 全基因组 22 487 条蛋白质氨基酸序列中,有 1 054 个具有经典分泌蛋白质的特征,所占比例为 4.7%。

2.2 Foc TR4 经典分泌蛋白特征分析

2.2.1 Foc TR4 经典分泌蛋白的氨基酸序列长度分析 对 Foc TR4 经典分泌蛋白的氨基酸序列长度分析结果(图 1)表明,其大多集中在 100~500 个氨基酸。其中,以 300~399 个氨基酸长度为最多,占经典分泌蛋白总数的 19.7%;以 200~299 个氨基酸长度次之,占经典分泌蛋白总数的 18.7%;以 100~199 个氨基酸长度为第 3,占经典分泌蛋白总数的 15.7%。



a: <100 个氨基酸; b: 100~199 个氨基酸; c: 200~299 个氨基酸; d: 300~399 个氨基酸; e: 400~499 个氨基酸; f: 500~599 个氨基酸; g: 600~699 个氨基酸; h: 700~899 个氨基酸; i: 900~1199 个氨基酸; j: >1200 个氨基酸。

图 1 Foc TR4 经典分泌蛋白的氨基酸序列长度分布

Fig.1 The amino acid length of classically secretory proteins in Foc TR4

2.2.2 Foc TR4 经典分泌蛋白信号肽长度分析 对 Foc TR4 经典分泌蛋白信号肽长度分析结果(图 2)表明,其信号肽长度介于 15~36 个氨基酸,信号肽长度为 17~22 个氨基酸的经典分泌蛋白分布最为集中,占经典分泌蛋白总数的 82.0%。其中,信号肽长度为 19 个氨基酸的最多,占经典分泌蛋白总数的 18.1%;信号肽长度为 20 个氨基酸的数量次之,占经典分泌蛋白总数的 15.8%;信号肽长度为 21 个氨基酸的数量为第 3,占经典分泌蛋白总数的 14.2%。

2.2.3 Foc TR4 经典分泌蛋白信号肽中氨基酸残基的出现频率分析 对 1 054 个 Foc TR4 经典分泌蛋白信号肽中氨基酸残基的出现频率进行了分析,结果表明:(1)在 20 种氨基酸残基中,丙氨酸出现频

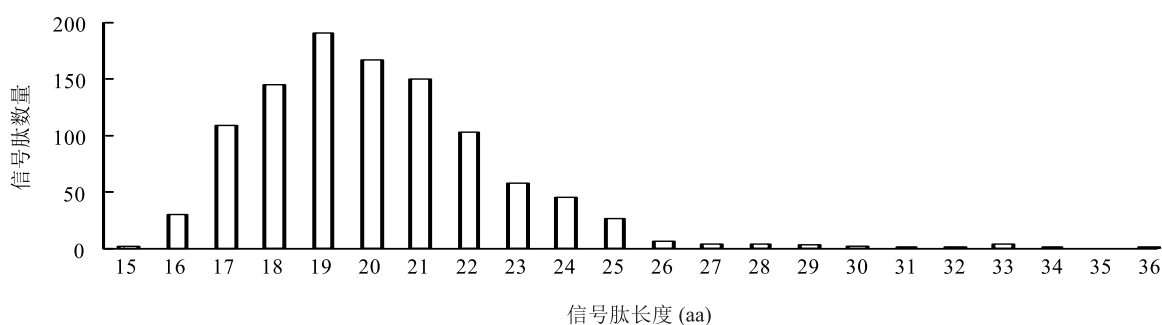


图2 Foc TR4 经典分泌蛋白的信号肽长度分布

Fig.2 The signal peptides length of classically secretory proteins in Foc TR4

率最高,为 21.0%;其次为丝氨酸,所占比例为 11.7%;谷氨酰胺、亮氨酸、甘氨酸、苏氨酸和缬氨酸所占比例则依次为 10.3%、8.8%、6.7%、5.5% 和 5.2%。(2) 非极性氨基酸出现频率最高,占 42.8%;有极性不带电荷的氨基酸次之,占 39.1%;此外,带正电荷的碱性氨基酸占 11.9%;带负电荷的酸性氨基酸占 6.2%。(3) 在信号肽组成中,使用频率大于 5% 的氨基酸残基都属于脂肪族氨基酸,主要为中性和含有羟基或硫氨基的氨基酸,可能与分泌蛋白的属性相关,使信号肽更容易穿过细胞质膜,从而行使信号指导功能^[25]。

2.2.4 Foc TR4 经典分泌蛋白信号肽切割位点的氨基酸组成分析 对 Foc TR4 经典分泌蛋白信号肽的切割位点-3 位到+3 位的氨基酸组成进行了分析,结果表明:(1) 在-3、-2、-1、+1、+2、+3 位置上,使用频率最高的氨基酸分别为丙氨酸、亮氨酸、丙氨酸、丙氨酸、脯氨酸、亮氨酸,所占比例分别为 44.7%、19.0%、68.9%、19.6%、29.9%、7.9%。(2) 信号肽切割位点-3 位、-2 位、-1 位的氨基酸组成为丙氨酸-X-丙氨酸(其中 X 代表任意氨基酸)的分泌蛋白有 330 个(占经典分泌蛋白总数的 31.3%),属于 SPase I 型信号肽识别位点,即这些分泌蛋白具有 Sec-信号肽,通过 Sec 途径进行分泌。该结果与禾谷炭疽菌(*Colletotrichum graminicola*)、马铃薯晚疫病病菌(*Phytophthora infestans*)等重要植物病原真菌分泌蛋白的信号肽酶切位点类型相一致,具有相对的保守特点^[11-12]。

2.3 Foc TR4 与重要植物病原真菌经典分泌蛋白的比较分析

对一些重要植物病原真菌所预测的分泌蛋白和 Foc TR4 进行了比较分析。结果(表 1)表明,不同植

物病原真菌分泌蛋白占总蛋白的比例存在差异,其范围大多为 3%~11%。例如,Heard 等^[26]对灰霉菌(*Botrytis cinerea*)和核盘菌(*Sclerotinia sclerotiorum*)的分泌蛋白质进行了预测分析,结果表明 2 个病原菌中分别有 499 个和 432 个分泌蛋白,均约占蛋白质总数的 3.0%。对尖孢镰刀菌不同专化型预测的分泌蛋白进行比较发现,其分泌蛋白也存在较大差异,如尖孢镰刀菌甜瓜专化型(*F. oxysporum* f. sp. *melonis*)预测的经典分泌蛋白占总蛋白质总数的 4.3%^[15],而尖孢镰刀菌番茄专化型(*F. oxysporum* f. sp. *lycopersici*)预测的经典分泌蛋白占蛋白质总数的 10.2%^[27]。

2.4 Foc TR4 经典分泌蛋白的功能预测分析

为了进一步分析 Foc TR4 经典分泌蛋白的作用,本研究利用 UniProt 数据库进行了功能预测分析(表 2)。结果表明,在 1 054 个经典分泌蛋白中,有 463 个获得了注释,占分泌蛋白总数的 43.9%。在有功能注释的分泌蛋白中,有 201 个具有生物学过程注释,其中与糖代谢过程相关分泌蛋白最多(154 个);有 449 个蛋白质具有分子功能注释,其中以水解酶活性的分泌蛋白最多(273 个)。

2.5 Foc TR4 经典分泌蛋白的 CAZymes 分析

CAZymes 是指参与碳水化合物合成和降解的重要酶类,主要包括 GHs(糖基水解酶)、GTs(糖基转移酶)、PLs(多糖裂解酶)、CEs(碳水化合物酯酶)、CBMs(碳水化合物结合模块)及 AAs(辅助模块酶类)等 6 类。本研究对 Foc TR4 经典分泌蛋白进行了分析,结果表明有 281 个为 CAZymes,分属于 6 个 CAZymes 家族,占经典分泌蛋白总数的 26.7%。其中,以 GH 家族最多(10.5%),AA 家族次之(6.1%),CE 家族第三(4.0%),而 CBM、PL 和 GT 家族较少(分别为 3.1%、1.5%和 1.4%)。

表 1 部分重要植物病原真菌预测的分泌蛋白比较分析

Table 1 Comparison of predicted secretory proteins in some phytopathogenic fungi

病原真菌	总蛋白数量	经典分泌蛋白数量	分析软件	文献
<i>Phytophthora infestans</i>	22 658	671	SignalP、PSORT、TMHMM、TargetP、THUMBUP、big-PI Predictor	周晓罡等 ^[12]
<i>Colletotrichum graminicola</i>	12 006	630	SignalP、ProtComp、TMHMM、big-PI Fungal Predictor、TargetP	韩长志 ^[13]
<i>F. oxysporum</i> f. sp. <i>melonis</i>	26 811	1 145	SignalP、WoLF PSORT、TargetP、TMHMM、big-PI Predictor	聂燕芳等 ^[15]
<i>F. graminearum</i>	11 640	606	SignalP、TargetP、Big-PI Predictor、TMHMM	于钦亮等 ^[27]
<i>F. oxysporum</i> f. sp. <i>lycopersici</i>	17 735	1 803	SignalP、TargetP、TMHMM、Phobius	Ma 等 ^[28]
<i>F. verticillioideis</i>	14 179	1 549	SignalP、TargetP、TMHMM、Phobius	Ma 等 ^[28]
<i>F. fujikuroi</i>	14 813	1 336	SignalP、WoLF PSORT、TargetP、TMHMM	Wilemann 等 ^[29]
<i>F. graminearum</i>	13 937	574	SignalP、TargetP、TMHMM、big-PI Predictor、ProtComp、WoLF PSORT	Brown 等 ^[30]
<i>Ustilago maydis</i>	6 902	426	SignalP3.0、TargetP、ProtComp 6.0	Kamper 等 ^[31]
<i>Mycosphaerella graminicola</i>	11 182	492	SignalP、TargetP、ProtComp 6.0、ProtComp	Morais do Amaral 等 ^[32]
<i>Botrytis cinerea</i>	16 488	499	SignalP、TargetP、TMHMM、big-PI Predictor、WoLF PSORT	Heard 等 ^[26]
<i>Sclerotinia sclerotiorum</i>	14 522	432	SignalP、TargetP、TMHMM、big-PI Predictor、WoLF PSORT	Heard 等 ^[26]
<i>Melampsora lini</i>	26 443	1 085	SignalP、TargetP、TMHMM	Nemri 等 ^[33]

表 2 Foc TR4 经典分泌蛋白的功能注释分析

Table 2 Functional annotation and categories of classically secretory proteins in Foc TR4

生物学功能	分泌蛋白数量	分子功能	分泌蛋白数量
糖代谢	154	水解酶活性	273
过氧化氢代谢	6	氧化还原酶活性	62
生物合成	3	激酶活性	1
转运过程	3	ATP 结合 ATP	3
细胞壁大分子代谢	7	核酸结合	24
其他	28	裂解酶活性	32
-	-	催化活性	26
-	-	其他	28

经典分泌蛋白的功能注释与分类均来自 UniProt 数据库。

每个 CAZymes 家族包含了较多的亚家族,其中部分亚家族可以直接参与植物细胞壁组分的降解,被认为是植物病原真菌的一类重要致病因子^[34]。因此,本研究进一步对具有降解寄主植物细胞壁组分的各 CAZymes 亚家族进行了分析。结果表明,参与植物细胞壁组分果胶降解的 GH 和 PL 亚家族共有 51 个(含 1 个 GH2、10 个 GH28、2 个 GH35、11 个 GH43、1 个 GH51、2 个 GH53、1 个 GH54、1 个 GH78、10 个 PL1、6 个 PL3、1 个 PL4、1 个 CE1、4 个 CE8)、参与纤维素降解的 GH 亚家族共有 13 个(含 5 个 GH3、4 个 GH5、2 个 GH7 和 2 个 GH45)、参与木聚糖降解的 GH 亚家族共有 27 个(含 5 个 GH3、4 个

GH10、3 个 GH11、12 个 GH43、1 个 GH62、1 个 GH67 和 1 个 CE1)。

2.6 Foc TR4 非经典分泌蛋白的预测分析

利用 SecretomeP 2.0 Server 软件,本研究对 Foc TR4 中不具有信号肽的 20 694 条氨基酸序列进行了分析。结果表明,有 9 216 个蛋白质氨基酸序列符合非典型分泌蛋白的特征,占基因组编码蛋白质总数的 41.0%。与尖孢镰刀菌甜瓜专化型(*F. oxysporum* f. sp. *melonis*) (有 8 471 个非经典分泌蛋白,占编码蛋白总数的 31.6%)相比^[15],2 个病原菌的非经典分泌蛋白有较大的差异。

3 讨论

分泌蛋白作为植物病原真菌的重要致病因子,在其侵染植物的过程中起着重要作用。随着许多植物病原真菌基因组测序的完成,应用生物信息学技术来分析其分泌蛋白的组成已成为解析病原真菌致病机理的重要内容之一。目前,已有一些重要植物病原真菌的分泌蛋白得到了预测分析^[35-36]。本研究利用 SignalP、WoLF PSORT、TMHMM、big-PI Predictor 和 TargetP 等软件对 Foc TR4 经典分泌蛋白进行了精细的预测分析,发现有 1 054 个为经典分泌蛋白,占编码蛋白质总数的 4.7%。该结果与很多植物病原真菌所预测的分泌蛋白数占总蛋白质数的比例基本相符,如油菜茎基溃疡病菌(*Leptosphaeria*

maculans) (分泌蛋白占总蛋白质的比例为4.4%)、玉米黑粉菌(*Ustilago maydis*) (5.5%)、小麦秆锈菌(*Puccinia graminis* f. sp. *tritici*) (5.5%)、番茄叶霉菌(*Cladosporium fulvum*) (5.0%)、小麦壳针孢叶枯菌(*Mycosphaerella graminicola*) (4.7%)、甘蓝黑斑病菌(*Alternaria brassicicola*) (4.8%) 和小麦黄斑叶枯病菌(*Pyrenophora tritici-repentis*) (5.5%) 等^[36]。与尖孢镰刀菌其他专化型相比,Foc TR4 与尖孢镰刀菌甜瓜专化型(*F. oxysporum* f. sp. *melonis*, FOM) 分泌蛋白占总蛋白质的比例(4.3%)^[15] 相似,而尖孢镰刀菌番茄专化型(*F. oxysporum* f. sp. *lycopersici*) 预测的经典分泌蛋白则占蛋白质总数的10.2%^[28]。

病原菌可以产生大量的 CAZymes 来降解植物的多糖类物质,以利于其对植物的侵染并获取其自身所需的营养^[37]。由于 CAZymes 在病原菌侵入及成功侵染寄主的过程中起着重要作用,因此其也被认为是病原菌的一类重要致病因素。研究结果表明,不同病原菌分泌的 CAZymes 存在较大的差异。例如,疫霉属4个种(*Phytophthora parasitica*、*P. infestans*、*P. sojae* 和 *P. ramorum*) 分别含有431个、435个、379个和310个 CAZymes^[38-39],而松杨栅锈菌(*Melampsora larici-populina*) 仅含有126个 CAZymes^[40]。聂燕芳等^[15] 对 FOM 的 CAZymes 进行了分析,发现 FOM 有277个 CAZymes,占经典分泌蛋白总数的24.2%。本研究结果表明,Foc TR4 有281个分泌蛋白为 CAZymes,占经典分泌蛋白总数的26.7%。与不同属的病原菌相比,其 CAZymes 数量差异较大,这可能与病原菌不同属间(甚至不同种)的侵染方式存在差异有关。与尖孢镰刀菌不同专化型相比,其 CAZymes 数量相似,可能与不同专化型间具有相似的侵染方式有关。

在真菌中,CAZymes 中的 GH、PL 和 CE 亚家族具有细胞壁降解酶的活性,在致病过程中起着重要作用^[38]。本研究对具有降解寄主植物细胞壁组分的 Foc TR4 各 CAZymes 亚家族进行了分析,发现共有91个 GH、CE 和 PL 亚家族酶类参与了植物细胞壁组分的降解,这有助于我们通过进一步分析来获得感兴趣的重要功能蛋白。但这些酶类在 Foc TR4 致病过程中的具体作用及其生物学功能,仍有待进一步的研究证实。此外,我们对于 Foc TR4 与香蕉互作条件下 CAZymes 的真实分泌情况等也不清楚。因此,进一步开展 Foc TR4 侵染香蕉过程中的定量

分泌蛋白质组学研究,有利于全面了解 Foc TR4 与香蕉互作的分子机理。

参考文献:

- [1] 曾莉莎,吕顺,杜彩娟,等. 香蕉枯萎病苗期抗性鉴定的水培接种法研究[J]. 南方农业学报,2015,46(5):817-822.
- [2] LI C, SHAO J, WANG Y, et al. Analysis of banana transcriptome and global gene expression profiles in banana roots in response to infection by race 1 and tropical race 4 of *Fusarium oxysporum* f. sp. *cubense* [J]. BMC Genomics, 2013, 14(1):851.
- [3] 王飞,周登博,廉法钦,等. 香蕉枯萎病拮抗放线菌 8N-10 的分离鉴定及抑菌效果评价[J]. 江苏农业科学,2016,44(4):179-183.
- [4] 彭埃天,宋晓兵,凌金锋,等. 香蕉枯萎病菌4号生理小种分子检测与枯萎病生物防治研究进展[J]. 果树学报,2009,26(1):77-81.
- [5] 左存武,李斌,李春雨,等. 香蕉对尖孢镰刀菌热带4号小种的抗性评价方法的建立[J]. 园艺学报,2016,43(5):876-884.
- [6] LI M H, XIE X L, LIN X F, et al. Functional characterization of the gene FoOCH1 encoding a putative α -1,6-mannosyltransferase in *Fusarium oxysporum* f. sp. *cubense* [J]. Fungal Genetics and Biology, 2014, 65:1-13.
- [7] 李梅婷,张绍升. 香蕉枯萎病菌细胞壁降解酶的诱导及其对香蕉组织的降解[J]. 中国农学通报,2010,26(5):228-231.
- [8] GUO L, HAN L, YANG L, et al. Genome and transcriptome analysis of the fungal pathogen *Fusarium oxysporum* f. sp. *cubense* causing banana vascular wilt disease [J]. PLoS One, 2014, 9(4):e95543.
- [9] 贾慧升,薛玉潇,方晓东,等. 香蕉枯萎病菌基因组编码的植物细胞壁降解酶系统生物信息学分析[J]. 广东农业科学,2012,39(9):129-132.
- [10] 董章勇,王琪,秦世雯,等. 香蕉枯萎病菌1号和4号生理小种细胞壁降解酶的比较[J]. 植物病理学报,2010,40(5):463-468.
- [11] CHOI J, PARK J, KIM D, et al. Fungal secretome database: integrated platform for annotation of fungal secretomes [J]. BMC Genomics, 2010, 11(1):105.
- [12] 周晓罡,侯思名,陈铎文,等. 马铃薯晚疫病病菌全基因组分泌蛋白的初步分析[J]. 遗传,2011,33(7):785-793.
- [13] 韩长志. 全基因组预测禾谷炭疽菌的分泌蛋白[J]. 生物技术,2014,24(2):36-41.
- [14] 苏源,李成云,赵之伟,等. 稻瘟菌基因组规模分泌蛋白的预测分析[J]. 云南农业大学学报,2006,21(3):271-275,292.
- [15] 聂燕芳,周淦,黄嘉瑶,等. 尖孢镰刀菌甜瓜专化型基因组规模分泌蛋白的预测与分析[J]. 华中农业大学学报,2016,35(3):24-29.
- [16] DYRL V, BENDTSEN J, NIELSEN H, et al. Improved prediction of signal peptides: SignalP 3.0 [J]. Journal of Molecular Biology, 2008, 373(2):919-924.

- 2004,340(4):783-795.
- [17] HORTON P, PARK K J, OBAYASHI T, et al. WoLF PSORT: protein localization predictor[J]. Nucleic Acids Research, 2007, 35 (Web Server): W585-W587.
 - [18] EMANUELSSON O, BRUNAK S, VON HEIJINE G, et al. Locating proteins in the cell using TargetP, SignalP, and related tools[J]. Nature Protocols, 2007, 2(4):953-971.
 - [19] KROGH A, LARSSON B, VON HEIJINE G, et al. Predicting transmembrane protein topology with a hidden markov model: application to complete genomes[J]. Journal of Molecular Biology, 2001, 305(3):567-580.
 - [20] EISENHABER B, SCHNEIDER G, WILDPANER M, et al. A Sensitive predictor for potential GPI lipid modification sites in fungal protein sequences and its application to genome-wide studies for *Aspergillus nidulans*, *Candida albicans*, *Neurospora crassa*, *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*[J]. Journal of Molecular Biology, 2004, 337(2):243-253.
 - [21] JUNCKER A S, WILLENBROCK H, VON HEIJINE G, et al. Prediction of lipoprotein signal peptides in Gram-negative bacteria[J]. Protein Science, 2003, 12(8):1652-1662.
 - [22] LOMBARD V, GOLACONDA R H, DRULA E, et al. The Carbohydrate-active enzymes database (CAZy) in 2013[J]. Nucleic Acids Research, 2014, 42:490-495.
 - [23] PARK B H, KARPINETS T V, SYED M H, et al. CAZymes Analysis Toolkit (CAT): web service for searching and analyzing carbohydrate-active enzymes in a newly sequenced organism using CAZy database[J]. Glycobiology, 2010, 20(12):1574-1584.
 - [24] BENDTSEN J D, JENSEN L J, BLOM N, et al. Feature based prediction of non-classical and leaderless protein secretion[J]. Protein Engineering Design & Selection, 2004, 17(4):349-356.
 - [25] 罗灯涛, 范继英, 范成明, 等. 根癌农杆菌 Ti 和 AT 质粒基因组中的分泌型信号肽分析[J]. 农业生物技术学报, 2006, 14(2):265-268.
 - [26] HEARD S, BROWN N A, HAMMOND-KOSACK K. An interspecies comparative analysis of the predicted secretomes of the necrotrophic plant pathogens *Sclerotinia sclerotiorum* and *Botrytis cinerea*[J]. PLoS One, 2015, 10(6):e0130534.
 - [27] 于钦亮, 马莉, 刘林, 等. 禾谷镰刀菌基因组中含寄主靶向模式分泌蛋白功能的初步分析[J]. 生物技术通报, 2008, (1):160-165.
 - [28] MA L J, VAN DER DOES H C, BORKOVICH K A, et al. Comparative genomics reveals mobile pathogenicity chromosomes in *Fusarium*[J]. Nature, 2010, 464(7287):367-373.
 - [29] WIEMANN P, SIEBER C M, VON BARGEN K W, et al. Deciphering the cryptic genome: genome-wide analyses of the rice pathogen *Fusarium fujikuroi* reveal complex regulation of secondary metabolism and novel metabolites[J]. PLoS Pathogen, 2013, 9(6):e1003475.
 - [30] BROWN N A, ANTONIW J, HAMMOND-KOSACK K E. The predicted secretome of the plant pathogenic fungus *Fusarium graminearum*: a refined comparative analysis[J]. PLoS One, 2012, 7(4):e33731.
 - [31] KAMPER J, KAHMANN R, BOLKER M, et al. Insights from the genome of the biotrophic fungal plant pathogen *Ustilago maydis*[J]. Nature, 2006, 444(7115):97-101.
 - [32] MORAIS DO AMARAL A, ANTONIW J, RUDD J J, et al. Defining the predicted protein secretome of the fungal wheat leaf pathogen *Mycosphaerella graminicola*[J]. PLoS One, 2012, 7(12):e49904.
 - [33] NEMRI A, SAUNDERS D G, ANDERSON C, et al. The genome sequence and effector complement of the flax rust pathogen *Melampsora lini*[J]. Frontiers in Plant Science, 2014, 5(98):98.
 - [34] BATTAGLIA E, BENOIT I, VAN DEN BRINK J, et al. Carbohydrate-active enzymes from the zygomycete fungus *Rhizopus oryzae*: a highly specialized approach to carbohydrate degradation depicted at genome[J]. BMC Genomics, 2011, 12(1):38.
 - [35] SPERSCHNEIDER J, WILLIAMS A H, HANE J K, et al. Evaluation of secretion prediction highlights differing approaches needed for oomycete and fungal effectors[J]. Frontiers in Plant Science, 2015, 6(98):1168.
 - [36] SONAH H, DESHMUKH R K, BÉLANGER R R. Computational prediction of effector proteins in fungi: opportunities and challenges[J]. Frontiers in Plant Science, 2016, 7(e3):126.
 - [37] ZHAO Z, LIU H, WANG C, et al. Comparative analysis of fungal genomes reveals different plant cell wall degrading capacity in fungi[J]. BMC Genomics, 2014, 15(1):1-15.
 - [38] OSPINA-GIRALDO M D, GRIFFITH J G, LAIRD E W, et al. The CAZome of *Phytophthora* spp.: a comprehensive analysis of the gene complement coding for carbohydrate-active enzymes in species of the genus *Phytophthora*[J]. BMC Genomics, 2010, 11(1):525.
 - [39] BLACKMAN L M, CULLERNE D P, HARDHAM A R. Bioinformatic characterisation of genes encoding cell wall degrading enzymes in the *Phytophthora parasitica* genome[J]. BMC Genomics, 2014, 15(1):785.
 - [40] LORRAIN C, HECKER A, DUPLESSIS S. Effector-mining in the poplar rust fungus *Melampsora larici-populina* secretome[J]. Frontiers in Plant Science, 2015, 6(1051):1-7.

(责任编辑:陈海霞)