

李明瑾, 霍建飞, 韩长志, 等. 辣椒炭疽病病原菌中分泌蛋白的预测及特性分析[J]. 江苏农业学报, 2025, 41(1): 41-50.
doi: 10.3969/j.issn.1000-4440.2025.01.006

辣椒炭疽病病原菌中分泌蛋白的预测及特性分析

李明瑾^{1,2}, 霍建飞², 韩长志¹, 姚玉荣², 贵海燕², 郝永娟², 王万立²

(1. 西南林业大学林学院/云南省森林灾害预警与控制重点实验室, 云南 昆明 650224; 2. 天津市农业科学院植物保护研究所, 天津 300384)

摘要: 由辣椒炭疽病病原菌引起的辣椒炭疽病严重危害着中国辣椒产业的健康发展。分泌蛋白作为炭疽病病原菌中重要的致病因子, 能帮助病原菌侵入寄主组织、分解寄主细胞壁, 为病原菌生存和繁殖提供条件。然而, 学术界尚缺乏对辣椒炭疽病病原菌中分泌蛋白的深入解析和研究。本研究基于前期获得的辣椒炭疽病病原菌强致病菌株 TJNH1 全基因组序列, 充分利用 SignalP、ProtComp 等生物信息学分泌蛋白预测软件, 对该病菌中 13 419 条蛋白质氨基酸序列进行预测, 同时, 基于转录组测序数据, 对上述分泌蛋白的功能及特性展开研究。结果表明, 辣椒炭疽病病原菌株 TJNH1 中含有 388 个分泌蛋白, 占总蛋白质数量的 2.89%, 分泌蛋白氨基酸序列长度介于 55 aa 至 730 aa, 非极性氨基酸丙氨酸(Ala)在分泌蛋白里使用频率最高, 分泌蛋白信号肽长度集中于 17 aa 至 21 aa, 信号肽切割位点为 A-X-A 型。分泌蛋白均为亲水性蛋白, 高亲水性氨基酸和高疏水性氨基酸数量最多的分别是天冬酰胺(Asn)和丙氨酸(Ala), 分泌蛋白大多为稳定、酸性蛋白。同时运用 SMART、eggNOG、STRING V11.5 等软件对上述分泌蛋白的保守结构域、COG 功能分类、蛋白质互作关系等进行分析, 有 152 个分泌蛋白具有明显的保守结构域, 163 个分泌蛋白获得了功能注释, 32 个分泌蛋白存在互作关系。该研究结果为明确分泌蛋白的功能和辣椒炭疽病病原菌的致病机理提供了理论支撑, 也为筛选以分泌蛋白为靶标的新型农药和辣椒炭疽病绿色防控提供了理论依据。

关键词: 辣椒炭疽病病原菌; 分泌蛋白; 生物信息学分析

中图分类号: S436.418.1⁺1 **文献标识码:** A **文章编号:** 1000-4440(2025)01-0041-10

Prediction and characteristic analysis of secretory proteins in pepper anthracnose pathogens

LI Mingjin^{1,2}, HUO Jianfei², HAN Changzhi¹, YAO Yurong², BEN Haiyan², HAO Yongjuan², WANG Wanli²

(1. Forestry College of Southwest Forestry University/Yunnan Key Laboratory of Forest Disaster Warning and Control, Kunming 650224, China; 2. Institute of Plant Protection, Tianjin Academy of Agricultural Sciences, Tianjin 300384, China)

Abstract: Pepper anthracnose caused by the pathogens of pepper anthracnose seriously endangers the healthy development of pepper industry in China. As an important pathogenic factor in anthracnose pathogens, the secretory protein can help pathogens invade host tissues, decompose host cell walls, and provide survival and reproduction conditions for pathogens. However, there is still a lack of in-depth analysis and research on the secreted proteins in the pathogens of pepper anthracnose. In this study, based on the whole genome sequence of the strong pathogenic strain TJNH1 of pepper anthracnose pathogens obtained in the previous study, 13 419 protein amino acid sequences in

收稿日期: 2024-06-03

基金项目: 天津市农业科学院青年科研人员创新研究与实验项目(2022011); 云南省“兴滇英才支持计划”青年人才专项(YNWR-QNBJ-2020-188); 天津市科技计划项目(23ZYCSN00220)

作者简介: 李明瑾(1992-), 女, 云南昆明人, 硕士研究生, 主要从事农业资源利用与植物保护研究。(E-mail) 1614950297@qq.com

通讯作者: 霍建飞, (E-mail) hjf2203@163.com; 韩长志, (E-mail) swfuhcz@163.com

the pathogen were predicted by using bioinformatics secreted protein prediction software such as SignalP and ProtComp. Meanwhile, based on transcriptome sequencing data, the functions and characteristics of the above-mentioned secreted proteins were studied. The results showed that TJNH1 contained 388 secretory proteins, accounting for 2.89% of the total proteins. The length of amino acid sequences of secreted proteins ranged from 55 aa to 730 aa. The non-polar amino acid alanine (Ala) was the most frequently used in secreted proteins. The length of the secreted protein signal peptide was concentrated from 17 aa to 21 aa, and the signal peptide cleavage site was A-X-A type. The secreted proteins were all hydrophilic proteins. Asparagine (Asn) and alanine (Ala) had the largest number of highly hydrophilic amino acids and highly hydrophobic amino acids, respectively. Most of the secreted proteins were stable and acidic proteins. At the same time, SMART, eggNOG, STRING V11.5 and other software were used to analyze the conserved domains, COG functional classification and protein interaction of the above secreted proteins. There were 152 secreted proteins with obvious conserved domains, 163 secreted proteins were functionally annotated, and 32 secreted proteins had interactions. The results of this study provide a theoretical support for clarifying the functions of secreted proteins and the pathogenic mechanisms of pepper anthracnose pathogens, and also provide a theoretical basis for screening new pesticides targeting secreted proteins and green prevention and control of pepper anthracnose.

Key words: pepper anthracnose pathogens; secreted proteins; bioinformatics analysis

辣椒是中国常见的经济作物,也是重要的蔬菜品种,风味多样、营养丰富,深受消费者喜爱。2020 年,中国辣椒种植面积为 $7.375 \times 10^5 \text{ hm}^2$,产量为 $1.668 \times 10^{10} \text{ kg}$,约占全球辣椒产量的 46.16%,是全球最大的辣椒生产国^[1]。近年来,中国辣椒产业发展迅速,已助力多个地区实现脱贫致富,然而随着种植面积的不断扩大,炭疽病、疫病等病害频发,造成其产量和品质下降,严重制约了辣椒产业的高质量发展^[2]。

辣椒炭疽病是由炭疽菌引起的最具破坏性的真菌性病害之一,主要危害辣椒叶片和果实,引起叶片脱落和果实腐烂,严重影响辣椒的品质和产量^[3]。目前已报道的能引发辣椒炭疽病的炭疽菌有 20 多种,例如斯高维尔炭疽菌 (*Colletotrichum scovillei*)^[4]、平头炭疽菌 (*C. truncatum*)^[5]、尖孢炭疽菌 (*C. acutatum*)^[6]、松针炭疽菌 (*C. fiorinae*)^[7] 等。应该指出, *C. scovillei* 是引发辣椒炭疽病的重要病原菌,在中国天津、四川、甘肃等地均有报道,对辣椒产量和品质均造成严重影响。前人研究发现,炭疽菌的致病机理是病原菌通过分泌系统将分泌蛋白传输至宿主植物组织中,分泌蛋白可以帮助病原菌在宿主植物组织内定殖和侵染,进而引起发病,可见分泌蛋白是宿主植物与病原菌之间相互作用的关键因子,在病原菌入侵和危害寄主植物的过程中起重要作用^[8-9]。Prasanth 等^[10]对甘蔗红腐病病原菌 (*C. falcatum*) 的候选分泌蛋白进行比较转录组分析发现,病原菌入侵寄主植物是通过分泌蛋白辅助的,因此该研究把分泌蛋白确定为甘蔗红腐病病原菌关键

的致病因素。Inoue 等^[11]对黄瓜炭疽病病原菌 (*C. orbiculare*) 的分泌蛋白基因进行敲除试验发现,大多数高表达的无宿主特异性分泌蛋白是病原菌侵染寄主植物的重要毒力因子。臧睿等^[12]基于全基因组数据对葡萄座腔菌 (*Botryosphaeria dothidea*) 的分泌蛋白进行预测,得到 578 个分泌蛋白的功能注释,并筛选到 11 个与致病性紧密联系的潜在分泌蛋白。邢启凯等^[13]基于可可毛色二孢菌 (*L. theobromae*) 全基因组序列,通过生物信息学预测得到 238 个候选非经典分泌蛋白编码基因,经反转录荧光定量 PCR 分析结果显示,分泌蛋白参与了病原菌的侵染过程。

分泌蛋白作为植物病原真菌侵染植物的重要致病因子,在危害植物方面发挥着重要作用,不仅可以作为效应物参与侵染寄主过程以及触发寄主防卫反应,还可以改变寄主细胞的结构和功能^[14]。近年来,学术界对于分泌蛋白的预测主要采用生物信息学和计算生物学等方法,大多是结合多种生物信息学方法以提高预测的准确性和可靠性。为了明确辣椒炭疽病病原菌中分泌蛋白的特征情况,本研究基于辣椒炭疽病病原菌 *C. scovillei* 全基因组序列数据,根据分泌蛋白所具有的特征^[15],利用生物信息学软件对辣椒炭疽病病原菌的分泌蛋白进行预测,并对分泌蛋白进行理化性质分析和蛋白质功能分析,以期揭示该病原菌侵染寄主植物过程中分泌蛋白的致病作用奠定理论基础。

1 材料与方法

1.1 辣椒炭疽病病原菌蛋白质氨基酸序列来源

辣椒炭疽病病原菌菌株 TJNH1 是从天津市宁

河区感染炭疽病的辣椒果实中分离鉴定出来的^[16],送至北京百迈客生物科技有限公司基于 Nanopore 技术进行了全基因组测序,然后通过 NovaSeq 6000 获得其转录组数据,并使用 EVIDENCEModeler (EVM) 对基因组进行注释,得到 13 419 个蛋白质编码基因,其 Assembly 登录号为 GCA_011075155.1。

1.2 辣椒炭疽病原菌分泌蛋白预测

植物病原菌分泌蛋白具有在 N 端含有信号肽、无跨膜结构域、无糖基化磷脂酰肌醇 (GPI) 锚定位点,且不定位于胞内细胞器等典型特征。根据分泌蛋白的这些特征,利用 SignalP v6.0^[17]、ProtComp v9.0、TMHMM v2.0 分别预测蛋白质信号肽、亚细胞定位以及跨膜结构域,筛选出含有信号肽、定位于胞外且不具有跨膜结构的蛋白质。再用 BIG-PI Fungal predictor 和 TargetP v2.0 网站预测 GPI 锚定位点和细胞器蛋白质分析,去除具有 GPI 锚定位点的蛋白质和胞内细胞器蛋白质,最终获得分泌蛋白。

1.3 辣椒炭疽病原菌分泌蛋白基本理化性质分析

利用 TBtools-II^[18] 软件的 Protein Paramter 功能进行蛋白质氨基酸数量、理论等电点及不稳定系数分析;通过 ExPasy-ProtScale 在线软件进行蛋白质氨基酸亲疏水性和最强亲疏水性分析。

1.4 辣椒炭疽病原菌分泌蛋白保守结构域分析

通过 SMART 网站在线分析辣椒炭疽病原菌分泌蛋白保守结构域。

1.5 辣椒炭疽病原菌分泌蛋白功能注释

利用 eggNOG 数据库预测辣椒炭疽病原菌分泌蛋白的功能,并用 COG 直系同源蛋白质数据库比对其功能进行注释及分类。

1.6 辣椒炭疽病原菌分泌蛋白互作分析

利用 STRING v11.5 软件对辣椒炭疽病原菌分泌蛋白进行蛋白质互作分析。

2 结果与分析

2.1 辣椒炭疽病原菌分泌蛋白特征预测

2.1.1 信号肽预测 信号肽是蛋白质氨基酸序列末端的一段疏水性氨基酸序列,用于引导新合成蛋白质向通路转移。分泌蛋白所具有的典型特征是在 N 端具有信号肽,根据信号肽酶的不同,分泌蛋白信号肽类型也不同,SignalP 6.0 可以区分 Sec/SP I、Sec/SP II、Sec/SP III、Tat/SP I 和 Tat/SP II 5 种信号肽类型。在线预测网站对 *C. scovillei* 蛋白质氨基酸

序列进行预测分析,结果表明 *C. scovillei* 的 13 419 条蛋白质氨基酸序列中,有 1 717 条蛋白质氨基酸序列具有典型的信号肽,占总蛋白质氨基酸序列数的 12.8%,信号肽类型均为 Sec/SP I 型(即由 Sec 易位子转运并被信号肽酶 I 切割的“标准”分泌信号肽)。进一步分析可知,信号肽长度多集中在 17 aa 至 22 aa,占比 65.4%,其中以 19 aa 最多,为 226 个,占比 13.2%;组成信号肽的氨基酸主要为丙氨酸、亮氨酸和丝氨酸,占比依次为 21.2%、18.2%、10.2%。

2.1.2 亚细胞定位预测 对含有信号肽的 1 717 个蛋白质进行亚细胞定位预测,结果发现定位在细胞外的蛋白质数量为 415 个,膜结合细胞外受体蛋白为 676 个,定位到质膜的蛋白质为 230 个,定位到线粒体的蛋白质为 179 个,定位到溶酶体的蛋白质为 55 个,定位到内质网的蛋白质为 53 个,剩余蛋白质定位至细胞质、高尔基体、过氧化物酶体、细胞核、液泡中,分别为 49 个、29 个、15 个、9 个和 7 个(表 1)。

表 1 分泌蛋白的亚细胞定位结果

Table 1 Subcellular localization of secreted proteins

亚细胞定位	蛋白质数量 (个)	亚细胞定位	蛋白质数量 (个)
膜结合细胞外受体	676	细胞质	49
细胞外	415	高尔基体	29
质膜	230	过氧化物酶体	15
线粒体	179	细胞核	9
溶酶体	55	液泡	7
内质网	53		

2.1.3 跨膜结构域预测 使用 TMHMM v2.0 网站对 415 个胞外蛋白质进行跨膜结构域分析,结果发现,有 380 个蛋白质没有跨膜结构域,35 个蛋白质有 1 个跨膜结构域。本研究沿用前期课题小组筛选标准,选择不含跨膜结构域和含有 1 个跨膜结构域的蛋白质,因此 415 个蛋白质都具有分泌蛋白特征。

2.1.4 GPI 锚点预测 使用 BIG-PI Fungal predictor 网站对 415 条蛋白质氨基酸序列进行 GPI 锚点定位预测,结果显示,有 23 条蛋白质氨基酸序列具有 GPI 锚点位点,3 条蛋白质氨基酸序列由于氨基酸数量较少未能预测,剩余 389 条蛋白质氨基酸序列没有 GPI 锚点。

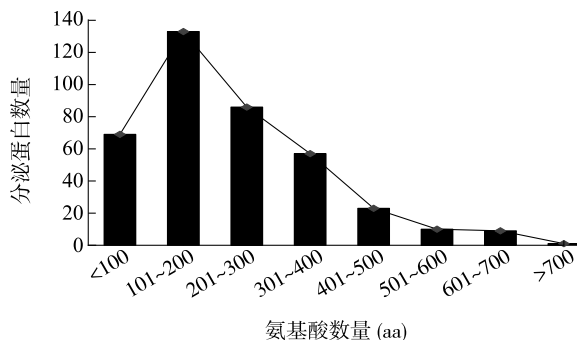
2.1.5 TargetP v2.0 预测 使用 TargetP v2.0 对 389 个

没有 GPI 锚定位点的蛋白质信号肽进行验证分析,结果显示,388 个蛋白质具有信号肽特征,1 个蛋白质不含信号肽,不含信号肽的蛋白质氨基酸序列编号为 KAF4784779.1。通过上述分析,最终明确了辣椒炭疽病原菌中具有典型特征的分泌蛋白数量为 388 个,占总蛋白质数量的 2.89%(图 1)。薄淑文等^[19]对茶树炭疽病原菌、希金斯炭疽病原菌中分泌蛋白进行研究,得到分泌蛋白数量占总蛋白质数量的 2.14%和 3.76%,本研究得到辣椒炭疽病原菌分泌蛋白数量占总蛋白质数量的 2.89%,与薄淑文等的研究结果相符。

2.2 辣椒炭疽病原菌分泌蛋白的理化性质

2.2.1 辣椒炭疽病原菌分泌蛋白的氨基酸组成和数量

运用 TBtools-II 软件的 Protein Paramter 对辣椒炭疽病原菌 388 个分泌蛋白氨基酸序列进行分析,结果发现辣椒炭疽病原菌分泌蛋白的氨基酸数量在 55 aa 至 730 aa 之间,平均值为 225.34 aa,中位数为 190.5 aa,分泌蛋白的氨基酸数量主要分布于 55~350 aa,占比 82.73%。其中位于 101~200 aa 之间的有 133 条,占比 34.27%;201~300 aa 之间有 86 条,占比 22.16%。同时,对上述分泌蛋白



A: 丙氨酸 (Ala); G: 甘氨酸 (Gly); S: 丝氨酸 (Ser); T: 苏氨酸 (Thr); L: 亮氨酸 (Leu); V: 缬氨酸 (Val); D: 天冬氨酸 (Asp); N: 天冬酰胺 (Asn); P: 脯氨酸 (Pro); K: 赖氨酸 (Lys); I: 异亮氨酸 (Ile); E: 谷氨酸 (Glu); R: 精氨酸 (Arg); Q: 谷氨酰胺 (Gln); F: 苯丙氨酸 (Phe); Y: 酪氨酸 (Tyr); C: 半胱氨酸 (Cys); H: 组氨酸 (His); M: 甲硫氨酸 (Met); W: 色氨酸 (Trp)。

图 2 分泌蛋白氨基酸数量和组成

Fig.2 Amino acid quantity and composition of secreted proteins

2.2.2 辣椒炭疽病原菌分泌蛋白的稳定性和理论等电点

通过对辣椒炭疽病原菌中 388 个分泌蛋白的不稳定性系数进行预测分析,结果发现,稳定蛋白有 265 个,不稳定蛋白有 123 个,稳定蛋白占比 68.30%。对辣椒炭疽病原菌分泌蛋白的理论等电点进行分析发现,理论等电点位于 5.01~6.00 的蛋白质数量最多,有 113 个,所占比例为 29.12%;位于 4.01~5.00 的蛋白质数量位居第 2,有 95 个,占

比 24.48%。理论等电点最低值为 3.87,最高值为 12.28,小于 7.0 的蛋白质为酸性蛋白质,数量为 272 个,占比 70.10%。

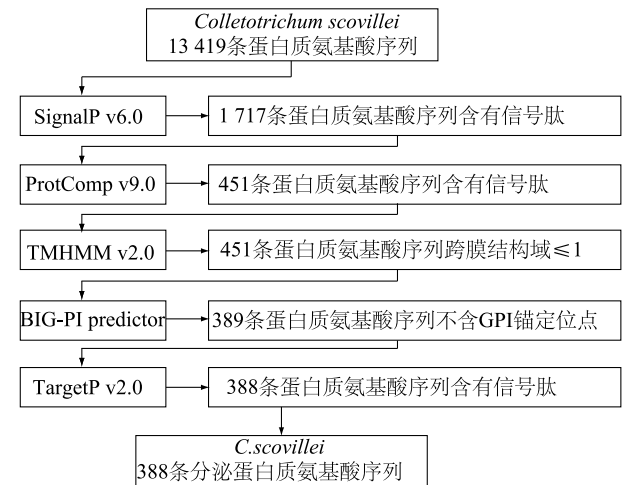
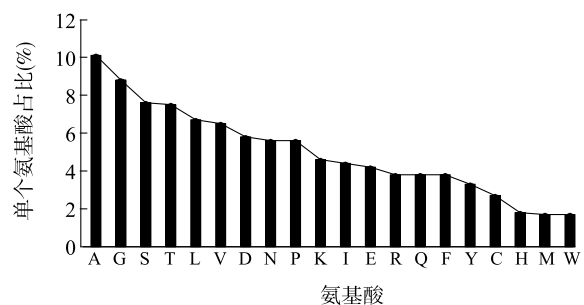


图 1 分泌蛋白鉴定流程图

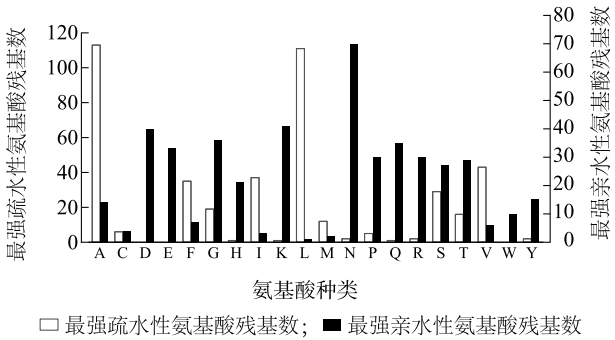
Fig.1 Flow chart for identification of secreted proteins



2.2.3 亲疏水性

运用 ProtScale 在线网站对分泌蛋白进行亲疏水性分析,结果表明辣椒炭疽病原菌 388 个分泌蛋白均为亲水性蛋白,高亲水性氨基酸残基排名前 3 的为天冬酰胺、赖氨酸和天冬氨酸,数量依次为 70 个、41 个和 40 个。高疏水性氨基酸

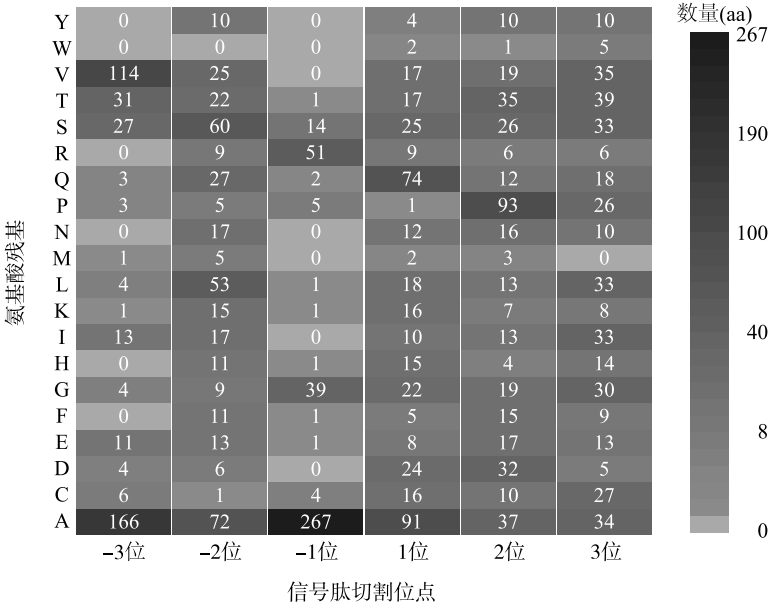
残基排名前3的为丙氨酸、亮氨酸、缬氨酸,数量依次为113个、111个和43个(图3)。



A、C、D、E、F、G、H、I、K、L、M、N、P、Q、R、S、T、V、W、Y 见图2注。

图3 分泌蛋白最强亲疏水性氨基酸残基分布图

Fig.3 Distribution of strongest hydrophilic and hydrophobic amino acid residues of secreted proteins



A、C、D、E、F、G、H、I、K、L、M、N、P、Q、R、S、T、V、W、Y 见图2注。

图4 分泌蛋白信号肽切割位点-3位到3位氨基酸分布图

Fig.4 Amino acid distribution from position -3 to position 3 of the signal peptide cleavage site of secreted proteins

2.4 辣椒炭疽病病原菌分泌蛋白的保守结构域

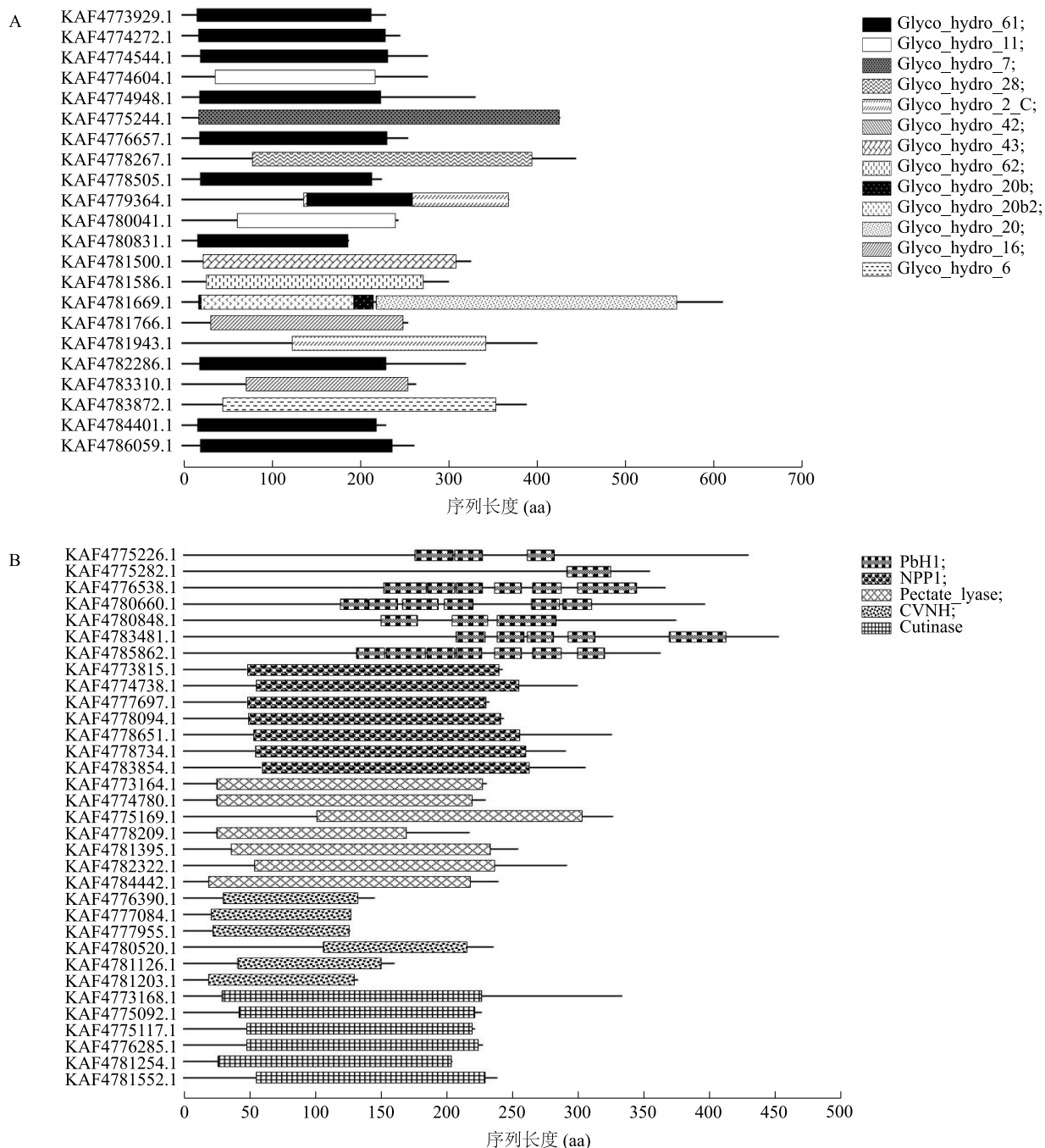
蛋白质具有高度保守的结构域,这些结构域决定了蛋白质的功能,对辣椒炭疽病病原菌中388个分泌蛋白进行保守结构域预测分析,结果发现有152个蛋白质具有明显的保守结构域,所占比例为39.18%。在辣椒炭疽病病原菌分泌蛋白结构域中,有22个蛋白质含有糖苷水解酶(GH)保守结构域,分别是Glyco_hydro_6、Glyco_hydro_7、Glyco_hydro_11、

2.3 辣椒炭疽病病原菌分泌蛋白信号肽的特征

对辣椒炭疽病病原菌分泌蛋白信号肽开展分析,结果发现辣椒炭疽病病原菌的分泌蛋白信号肽以19 aa居多,所占比例为15.21%,信号肽长度多集中于17 aa至21 aa,所占比例为60.82%,组成信号肽的氨基酸残基数量最多的为丙氨酸,占比21.92%,其次为亮氨酸,占比16.85%。此外,在信号肽切割位点的-3位、-2位、-1位、1位、2位和3位,氨基酸残基数量最多的分别为丙氨酸、丙氨酸、丙氨酸、丙氨酸、脯氨酸和苏氨酸,所占比例为42.78%、18.56%、68.81%、23.45%、23.97%和10.05%(图4)。由此可见辣椒炭疽病病原菌分泌蛋白信号肽切割位点属于A-X-A型,与课题组前期对禾谷炭疽病病原菌^[20]、希金斯炭疽病病原菌^[21]等分泌蛋白具有的信号肽酶切位点类型相同。

Glyco_hydro_16、Glyco_hydro_20、Glyco_hydro_28、Glyco_hydro_42、Glyco_hydro_43、Glyco_hydro_61、Glyco_hydro_62、Glyco_hydro_20b、Glyco_hydro_20b2、Glyco_hydro_2_C,尤以含有Glyco_hydro_61保守结构域的蛋白质最多,数量为10个(图5)。前人研究发现,碳水化合物活性酶是植物病原真菌分泌蛋白中的一类,而糖苷水解酶作为碳水化合物活性酶中重要的一类蛋白质可以降解植物细胞壁中的各

种糖苷键,促进真菌进入植物细胞内部,从而引发植物病害,根据其序列的差异程度进行分类,可分为 173 个家族^[22]。



A: 糖苷水解酶保守结构域; B: PbH1、NPP1、Pectate_lyase、CVNH 和 Cutinase 结构域。

图 5 辣椒炭疽病原菌中分泌蛋白保守结构域

Fig.5 Conserved domains of secreted proteins in pepper anthracnose pathogens

有 7 个分泌蛋白含有 PbH1 保守结构域, PbH1 (平行 β -螺旋重复序列) 是存在于许多碳水化合物裂解酶中的基序, 例如果胶裂解酶和鼠李糖半乳糖醛酸酶, 其功能与蛋白质互作相关; 有 7 个分泌蛋白

含有 NPP1 保守结构域, NPP1 渗入拟南芥植物叶片会导致发病相关基因的转录、乙烯产生、胼胝质沉淀和 HR 样细胞死亡等; 有 7 个分泌蛋白含有 Pectate_lyase 保守结构域, Pectate_lyase 果胶裂解酶是一种

胞外酶,由果胶诱导,负责植物组织的浸渍和软腐,与植物病害有关(图 5)。

有 6 个分泌蛋白含有 CVNH 保守结构域,CVNH 是一种常见的糖类结合结构域,通常与糖类分子结合,参与许多生物过程,包括细胞附着、信号传导和免

疫反应等;有 6 个分泌蛋白含有 Cutinase 保守结构域,Cutinase 脂质酶属于水解酶,在植物病害发生中发挥重要作用,能促进真菌渗透植物角质层(图 5B)。辣椒炭疽病病原菌分泌蛋白还含有纤维素酶、酪氨酸酶、肽酶等结构域,详见表 2。

表 2 分泌蛋白其他保守结构域

Table 2 Other conserved domains of secreted proteins

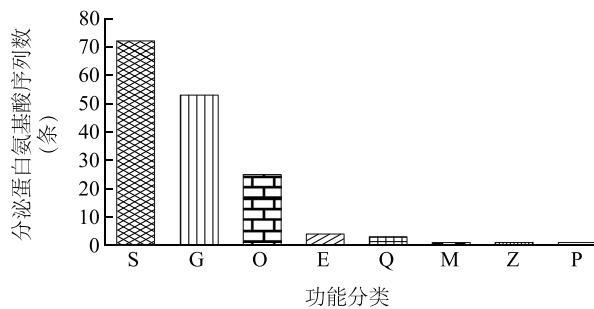
保守结构域	蛋白质氨基酸序列数	保守结构域	蛋白质氨基酸序列数	保守结构域	蛋白质氨基酸序列数
4LE4 D	5	Abhydrolase_3	1	fn3_3	1
Cellulase	4	SUN	1	FTP	1
SCP	4	Alpha_L_fucos	1	Glyco_10	1
transmembrane_domain	4	Asp	1	Glyco_32	1
Tyrosinase	4	Aspzincin_M35	1	GMC_oxred_C	1
ChtBD1	3	Bys1	1	GMC_oxred_N	1
DUF1996	3	Carb_anhydrase	1	His_Phos_2	1
fCBD	3	CarboxypepD_reg	1	HRXXH	1
Lactonase	3	CBM-like	1	HsbA	1
Peptidase_M28	3	COesterase	1	Lipase_GDSL	1
AltA1	2	Cu-oxidase	1	Lipocalin_2	1
Amb_all	2	Cu-oxidase_2	1	Lipocalin_9	1
CFEM	2	Cu-oxidase_3	1	LPMO_10	1
Cupin_1	2	d1crua_	1	MRJP	1
DUF3129	2	d1eo9b_	1	p450	1
DUF4360	2	d1jer_	1	PBPb	1
Exo_endo_phos	2	d1qlka_	1	Peptidase_M35	1
Hce2	2	d1scjb_	1	Peptidase_M36	1
Hydrophobin_2	2	d2dnja_	1	Peptidase_M43	1
PDB:4LE5 D	2	d2ple_	1	Peptidase_S10	1
Cellulase	2	Dioxygenase_C	1	Peptidase_S41	1
Peptidase_M20	2	DPBB_1	1	Phosphoesterase	1
Polysacc_deac_1	2	DUF3328	1	Pollen_allerg_1	1
SGL	2	EFh	1	Polysacc_lyase	1
SnoaL_2	2	EGF	1	Ribonuclease	1
WSC	2	Enterotoxin_a	1	Str_synth	1
Zn_pept	2	EthD	1	Trm112p	1
ZnMc	2	ETS	1		

2.5 辣椒炭疽病病原菌分泌蛋白的 COG 功能注释

利用 eggNOG-mapper v2 在线数据库对辣椒炭疽病病原菌的 388 条分泌蛋白氨基酸序列进行功能

分类,得到注释信息的有 360 条。通过 COG 功能对 360 条序列分类,去除 197 条未能分类的序列,剩下的 163 条蛋白质氨基酸序列被分为 8 类:S 类(功能

未知类) 72 条, G 类(碳水化合物转运代谢类) 53 条, O 类(蛋白质翻译后修饰和伴侣蛋白类) 25 条, E 类(氨基酸转运代谢类) 4 条, Q 类(次生代谢物的生物合成、转运和代谢类) 3 条, Z 类(细胞支架)、M 类(细胞壁/膜/包膜生物合成)、P 类(无机离子转运代谢) 各 1 条(图 6)。除此之外, 有的分泌蛋白氨基酸序列可以归属于两类功能, 如序列 KAF4775845.1 可归属于 E 类和 O 类, KAF4775908.1 可归属于 I 类和 Q 类, KAF4779278.1 可归属于 O 类和 P 类。



S:功能未知类;G:碳水化合物转运代谢类;O:蛋白质翻译后修饰和伴侣蛋白类;E:氨基酸转运代谢类;Q:次生代谢物的生物合成、转运和代谢类;M:细胞壁/膜/包膜生物合成类;Z:细胞支架类;P:无机离子转运代谢类。

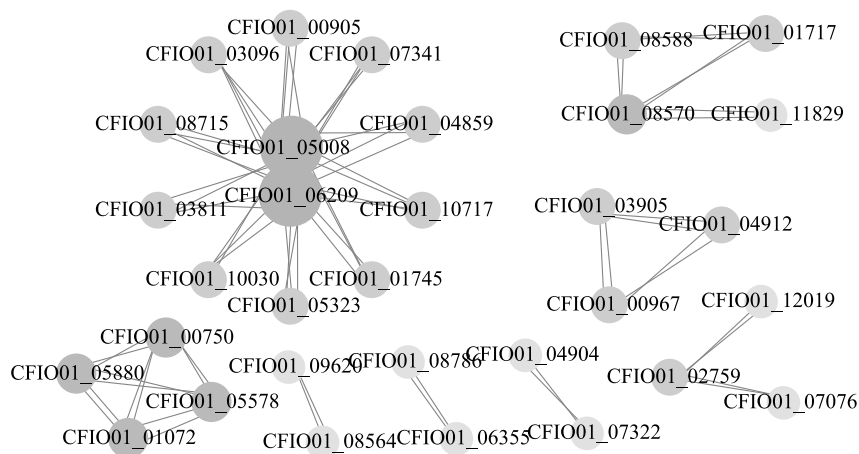
图 6 分泌蛋白 COG 功能分类

Fig.6 Functional classification of secreted proteins by COG

2.6 辣椒炭疽病原菌分泌蛋白的互作

通过 STRING 在线软件分析,在辣椒炭疽病原菌 388 个分泌蛋白中有 32 个分泌蛋白存在互作关系,其余 356 个分泌蛋白不存在互作关系。

进一步分析,CFIO01_05008 和 CFIO01_06209 属于多聚半乳糖醛酸内切酶,与 CFIO01_03096、CFIO01_10030、CFIO01_04859、CFIO01_08715、CFIO01_05323、CFIO01_07341、CFIO01_03811、CFIO01_10717、CFIO01_00905、CFIO01_01745 这 10 个蛋白质均有互作关系,而 CFIO01_05008 和 CFIO01_06209 之间无互作关系,CFIO01_03096、CFIO01_10030 等 10 个蛋白质之间也无互作关系。STRING 网站注释信息显示,CFIO01_10030、CFIO01_08715、CFIO01_05323、CFIO01_00905 为果胶裂解酶,CFIO01_03096 为果胶裂解酶 C,CFIO01_01745 为果胶裂解酶 F,CFIO01_04859、CFIO01_07341 为多糖裂解酶 3 家族成员,CFIO01_03811、CFIO01_10717 为糖苷水解酶 28 家族成员。同时,同类别的蛋白质之间也存在互作关系,CFIO01_01717 属于糖苷水解酶 7 家族成员,与属于糖苷水解酶 6 家族成员 CFIO01_08570 和糖苷水解酶 5 家族成员 CFIO01_08588 之间有互作关系,CFIO01_08570 与糖苷水解酶 61 家族成员 CFIO01_11829 之间也有互作关系(图 7)。



图中蛋白质节点的大小代表该蛋白质互作关系的多少,节点越大,则该蛋白质互作关系越多。

图 7 辣椒炭疽病原菌分泌蛋白的互作情况

Fig.7 Interaction of secreted proteins of pepper anthracnose pathogens

通过分泌蛋白互作关系图可见,CFIO01_05008 和 CFIO01_06209 注释信息和互作结果比较相似,本研究使用 NCBI 蛋白质数据库 Blastp 比对再次验

证发现,这两个蛋白质均为内多聚半乳糖醛酸酶。内多聚半乳糖醛酸酶是植物病原真菌分泌的一种酶,属于聚半乳糖醛酸酶 (Polygalacturonase, PG) 家

族,主要功能包括降解植物细胞壁中的果胶,破坏细胞壁完整性以促进病原菌侵染,并可能参与真菌的营养吸收过程^[23]。

3 讨论

植物病原菌中分泌蛋白研究已成为植物病理学研究的热点^[24-25],植物病原真菌分泌蛋白具有纤维素酶、蛋白酶等酶活性,可以分解植物细胞壁的主要成分,如纤维素等,从而破坏植物细胞壁的完整性,有利于病原菌的侵入^[26]。研究分泌蛋白有助于筛选植物病原菌中的关键致病基因,揭示病原菌的致病机理,对控制病害具有非常重要的作用。前期本课题组对核桃炭疽病病原菌(*C. gloeosporioides*)中的18 412个蛋白质用生物信息学方法进行预测得到293个分泌蛋白,占总蛋白质数量的1.59%,核桃炭疽病病原菌分泌蛋白占比低于辣椒炭疽病病原菌的分泌蛋白占比,究其原因,可能是核桃属于核果类木本植物,而辣椒是草本植物,二者的生长形态和生长环境不同,造成病原菌分泌蛋白数量有差异,后续将继续开展其差异性研究。

本研究参考前期课题组对禾谷炭疽病病原菌等炭疽菌属真菌采用的ProtComp在线预测软件,对辣椒炭疽病病原菌1 717个含有信号肽的蛋白质进行亚细胞定位预测分析,除了定位到胞外的415个蛋白质和其他细胞器的626个蛋白质外,有676个蛋白质亚细胞定位预测结果为膜结合细胞外受体蛋白。膜结合细胞外受体蛋白主要在细胞膜上起作用,如参与细胞信号传导等,不同于胞外蛋白,在筛选分泌蛋白时,视为不具有分泌蛋白分泌到胞外的特征。但随着生物信息学的发展,源源不断的生物信息学软件涌现出来,例如预测蛋白质的亚细胞定位就有ProtComp、WoLF PSORT、Cell-PLoc等十多个分析软件,本研究并未再用其他亚细胞定位软件对辣椒炭疽病病原菌的蛋白质进行验证。

另外,在进行蛋白质结构域和蛋白质互作分析时,得到糖苷水解酶家族蛋白质注释和蛋白质互作关系信息,前人研究结果表明,糖苷水解酶参与了植物病原菌在寄主植物中营养吸收和组织侵染等重要生物学过程,具体功能如何,有待后续进一步研究。

4 结论

本研究通过对辣椒炭疽病病原菌全基因组序列

预测及分析,明确了其分泌蛋白数量为388个,占总蛋白质数量的2.89%,分泌蛋白的氨基酸数量主要分布于55~350 aa,大多为小分子蛋白质,组成分泌蛋白的氨基酸主要是丙氨酸、甘氨酸、丝氨酸。分泌蛋白信号肽长度集中于17 aa至21 aa,所占比例为60.82%,信号肽切割位点为A-X-A型,组成信号肽的氨基酸残基数量最多的是丙氨酸和亮氨酸,占比分别为21.92%和16.85%,疏水性氨基酸占比高能够帮助信号肽在蛋白质合成过程中与细胞质膜上的信号识别,更利于其穿透质膜,运输到胞外实现对寄主植物的侵染。辣椒炭疽病病原菌分泌蛋白大多为稳定蛋白,通过理论等电点分析可知,酸性蛋白占总分泌蛋白的70.10%。用ProtScale软件预测388个分泌蛋白均为亲水性蛋白,亲、疏水性氨基酸残基数量最多的分别为天冬酰胺、丙氨酸。

对分泌蛋白进行保守结构域分析,结果发现有152个蛋白质具有明显的保守结构域,预测到GH、PbH1、NPP1、CVNH、Cutinase等结构域。前人研究结果表明,脂质酶、纤维素酶和果胶酶能够分解植物表面的角质素、纤维素和果胶,从而帮助病原菌侵入寄主组织和细胞,这些保守结构域是辣椒炭疽病病原菌侵染的关键因子。利用eggNOG-mapper对分泌蛋白进行COG功能分类,可分为8类,除去功能未知的S类,剩余最多的分类是碳水化合物转运代谢类(G类),这一类别大多是果胶酸裂解酶、纤维素酶等酶类。分泌蛋白互作关系分析结果显示,有32个分泌蛋白存在互作关系,有不同类别蛋白质间的互作如多聚半乳糖醛酸酶和多糖裂解酶互作,也有同类蛋白质间的互作如糖苷水解酶7家族成员与糖苷水解酶6家族成员和糖苷水解酶5家族成员互作。以上结果表明,辣椒炭疽病病原菌分泌蛋白中含有大量降解植物细胞壁成分的酶类物质,为后续辣椒炭疽病病原菌致病机理的研究提供了理论支持。

参考文献:

- [1] 林巧,辛竹琳,孔令博,等.我国辣椒产业发展现状及育种应对措施[J].中国农业大学学报,2023,28(5):82-95.
- [2] 张子峰.我国辣椒产业发展现状、主要挑战与应对之策[J].北方园艺,2023(14):153-158.
- [3] 韩长志,尹青晓,祝友朋.核桃炭疽病研究进展[J].经济林研究,2023,41(4):1-11.
- [4] 周志成,孙海,肖仲久,等.辣椒炭疽病病原分离鉴定及对杀

- 菌剂敏感性测定[J]. 华南农业大学学报, 2023, 44(3): 430-437.
- [5] DIAO Y Z, ZHANG C, LIU F, et al. *Colletotrichum* species causing anthracnose disease of chili in China[J]. *Persoonia*, 2017, 38: 20-37.
- [6] 王 妮, 尹显慧, 彭丽娟, 等. 辣椒炭疽病原鉴定及其杀菌剂毒力测定[J]. 植物保护, 2019, 45(4): 216-223.
- [7] 荆卓琼, 郭致杰, 孙 倩, 等. 甘肃省辣椒炭疽病原菌鉴定[J]. 北方园艺, 2024(15): 1-7.
- [8] 赵 美, 郑文博, 皮 磊, 等. 基于生物信息学的胶孢炭疽菌效应分子的筛选[J]. 分子植物育种, 2018, 16(18): 5927-5933.
- [9] JOSHI R. A review on *Colletotrichum* spp. virulence mechanism against host plant defensive factors[J]. *Journal of Medicinal Plants Studies*, 2018, 6(6): 64-67.
- [10] PRASANTH C N, VISWANATHAN R, MALATHI P, et al. Comparative transcriptome analysis of candidate secretory effector proteins from *Colletotrichum falcatum* infecting sugarcane[J]. *Agri Gene*, 2019, 13: 100089.
- [11] INOUE Y, VY T T P, SINGKARAVANIT-OGAWA S, et al. Selective deployment of virulence effectors correlates with host specificity in a fungal plant pathogen[J]. *New Phytologist*, 2023, 238(4): 1578-1592.
- [12] 臧 睿, 宋璐璐, 尹新明, 等. 葡萄座腔菌(*Botryosphaeria dothidea*)全基因组分泌蛋白的预测及功能分析[J]. 植物病理学报, 2021, 51(4): 559-571.
- [13] 邢启凯, 王欣芳, 彭军波, 等. 可可毛色二孢菌全基因组非经典分泌蛋白的预测及致病相关性分析[J]. 植物病理学报, 2024, 54(1): 102-115.
- [14] 夏雄飞, 陈云芳, 覃 悦, 等. 不同类型植物病原物中分泌蛋白及 CAZymes 对比研究[J]. 云南农业大学学报(自然科学), 2022, 37(6): 949-956.
- [15] 覃 悦, 韩长志. 不同营养方式真菌中分泌蛋白数量及其功能对比研究[J]. 微生物学报, 2021, 61(12): 4106-4117.
- [16] HUO J F, WANG Y F, HAO Y J, et al. Genome sequence resource for *Colletotrichum scovillei*, the cause of anthracnose disease of chili[J]. *Molecular Plant-Microbe Interactions*, 2021, 34(1): 122-126.
- [17] TEUFEL F, ARMENTEROS J J A, JOHANSEN A R, et al. SignalP 6.0 predicts all five types of signal peptides using protein language models[J]. *Nature Biotechnology*, 2022, 40(7): 1023-1025.
- [18] CHEN C J, WU Y, LI J W, et al. TBtools-II: a 'one for all, all for one' bioinformatics platform for biological big-data mining[J]. *Molecular Plant*, 2023, 16(11): 1733-1742.
- [19] 薄淑文, 喻红稠, 覃 悦, 等. 炭疽属植物病原真菌分泌蛋白的预测及比较分析[J]. 福建农林大学学报(自然科学版), 2023, 52(6): 747-754.
- [20] 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.
- [21] 韩长志. 全基因组预测希金斯炭疽菌的候选效应分子[J]. 生物技术, 2015, 25(6): 546-551.
- [22] 肖 会, 韩长志. 植物病原丝状真菌中糖苷水解酶的研究进展[J]. 江苏农业科学, 2024, 52(8): 1-7.
- [23] CHEN Y Z, YU J, LIN H T, et al. *Phomopsis longanae* Chi-induced Longan pulp breakdown and softening in relation to cell wall polysaccharides disassembly[J]. *Postharvest Biology and Technology*, 2022, 186: 111837.
- [24] 禄锦鹏, 秦春秀, 李 潇, 等. 可可毛色二孢 AM2As 全基因组分泌蛋白预测及功能分析[J]. 南方农业学报, 2023, 54(11): 3136-3155.
- [25] 唐 伟, 张成玲, 马居奎, 等. 基于基因组预测和分析甘薯间座壳菌(*Diaporthe batatas*)分泌蛋白中效应因子[J]. 江苏农业学报, 2023, 39(3): 665-673.
- [26] 韩长志, 许 僊. 植物病原丝状真菌分泌蛋白及 CAZymes 的研究进展[J]. 南京林业大学学报(自然科学版), 2017, 41(5): 152-160.

(责任编辑: 黄克玲)