



不同营养方式真菌中分泌蛋白数量及其功能对比研究

覃悦¹, 韩长志^{1,2*}

¹西南林业大学生物多样性保护学院, 云南 昆明 650224

²云南省森林灾害预警与控制重点实验室, 云南 昆明 650224

摘要: 【目的】分泌蛋白在植物病原真菌致病过程中发挥着重要的作用, 前人多以单种菌株分泌蛋白预测分析, 尚未见多种类型真菌中分泌蛋白的预测及对比研究报道。【方法】本研究基于多种不同营养方式真菌的全基因组序列, 根据分泌蛋白所具有的基本特征, 采用在线分析程序(主要包括SignalP v5.0、ProtComp v9.0等)对模式生物、死体营养型真菌、半活体营养型真菌及活体营养型真菌共19种菌株的分泌蛋白开展预测及功能分析。【结果】在上述真菌约13万条蛋白序列中, 分泌蛋白所占比例为0.74%–4.83%, 其中, 活体营养型真菌所具有的分泌蛋白数量占比最高, 平均为3.51%, 而死体营养型真菌和模式生物平均比例最低, 平均为1.36%。同时, 活体营养型真菌具有的功能种类最多, 为433种, 其次为半活体营养型真菌, 为266种, 而模式生物具有的功能种类最少, 为100种, 其中, 功能为假设蛋白和非特征蛋白的蛋白数量最多。【结论】该研究为深入解析分泌蛋白在实现不同营养方式真菌的致病机制方面提供理论依据。

关键词: 真菌, 营养方式, 分泌蛋白, 功能, 对比分析

植物病原真菌依据其生活方式可以分为活体营养型真菌、半活体营养型真菌及死体营养型真菌^[1]。活体营养型真菌仅能生活在植物细胞中, 不能在死体细胞中获取营养; 而死体营养型真菌则主要从死体细胞中获取营养; 半活体营养型真菌在获取营养物质方面则介于上述两种营养型真菌之间, 既能在侵染初期作为寄生物在寄主上

生活, 又能在侵染后期作为腐生物从死体植物组织上获取营养^[2]。植物与病原真菌互作过程中, 病原真菌主要通过其分泌系统将分泌蛋白输入到植物组织中, 这些蛋白通过与植物中具有防卫反应的相关蛋白发生多层次的免疫作用, 进而形成类似于“军备竞赛”的植物与病原真菌的互作反应^[3–5]。植物与病原真菌的互作理论研究主要涉

基金项目: 国家自然科学基金(31960314); 云南省高层次人才培养支持计划“青年拔尖人才专项”; 云南省应用基础研究计划(2018FG001-028)

*通信作者。Tel: +86-871-63862918; E-mail: hanchangzhi2010@163.com

收稿日期: 2021-03-26; 修回日期: 2021-05-20; 网络出版日期: 2021-09-28

及植物免疫理论、病原效应蛋白致病理论等^[6]。作为真(卵)菌中效应蛋白的重要筛选来源,分泌蛋白在实现病原真菌侵染植物、定殖细胞以及对植物造成危害等方面发挥着重要作用^[7-9]。

近些年,学术界对全基因组序列已经公布的植物病原菌(包括真菌、细菌以及卵菌等)中的分泌蛋白开展了诸多研究工作^[7],明确了包括稻瘟病菌^[10]、香蕉枯萎病菌^[11]、核盘菌^[12]、灰霉菌^[13]、致病疫霉^[14]、大丽轮枝菌^[15-16]、尖孢镰刀菌^[17]、柑橘黄龙病菌^[18]、甘薯长喙壳^[19]、小麦叶锈菌^[20]、黄单胞菌^[21]以及细菌性黑斑病菌^[22]等诸多植物病原菌中的分泌蛋白情况。上述研究对于进一步开展病原菌致病机制的解析及蛋白功能的研究提供了重要的理论支撑,然而,由于缺乏对比分析不同种类真菌或不同种类物种之间分泌蛋白的数量及功能情况,制约着学术界对于分泌蛋白在不同真菌中发挥作用的认知,影响着学术界对于不同营养方式植物病原真菌致病机制的解析工作。

因此,为了更好地明确不同种真菌之间分泌蛋白数量及功能情况,本研究选择模式生物、死体营养型真菌、半活体营养型真菌及活体营养型真菌等 19 种具有全基因组序列的真菌作为研究对象,根据真菌中分泌蛋白的典型特征,利用多种生物信息学分析工具,从信号肽、亚细胞定位、跨膜结构、锚定蛋白等方面进行层层筛选,明确上述真菌中的分泌蛋白数量及功能情况,同时,对不同营养方式真菌中分泌蛋白数量及功能展开对比研究,以期为进一步开展不同营养方式真菌中分泌蛋白功能解析及致病机制研究工作奠定坚实的理论基础,并为后续开展生物学试验验证打下基础。

1 材料和方法

1.1 蛋白序列来源

通过美国国立卫生院数据库(NCBI)下载 19 种具有不同营养方式真菌的全基因组序列,共获取了 233362 个蛋白序列,上述真菌所含有蛋白数量范围为 6002-27347 个,平均每个真菌含有 12282 个蛋白序列,真菌含有序列大小范围为 12157105-140802721 Mb (表 1)。

1.2 分泌蛋白预测

根据分泌蛋白的基本特征,在 Linux 系统下利用 SignalP v5.0^[23]、ProtComp v9.0^[24]、TMHMM v2.0^[25]分别预测分析蛋白信号肽和预测分析亚细胞定位,同时利用 Python 脚本对数据进行筛选,排除具有跨膜结构的蛋白。再利用 big-PI predictor 预测分析 GPI 锚定蛋白等,排除具有 GPI 的锚定蛋白^[26]。

1.3 分泌蛋白功能分析

利用 Microsoft Excel 2010 统计和分析分泌蛋白功能分布情况,并采用 Origin 2019 软件制图。

2 结果和分析

2.1 分泌蛋白预测

2.1.1 信号肽预测:基于真菌中分泌蛋白所具有的典型特征(在 N 端含有信号肽、无跨膜结构域、无 GPI 锚定位点且不定位于胞内细胞器等)^[7],对 19 种不同营养方式真菌开展分泌蛋白预测分析,结果表明,通过 SignalP v5.0 预测,明确具有信号肽的蛋白数量为 311-2495 个,所占比例为 5.18%-13.39%。其中, *F. oxysporum* 具有信号肽的蛋白数量最多,为 2021 个,而 *S. cerevisiae*

表 1. 不同营养方式真菌的基本信息
Table 1. Basic information of different fungus

Type	Name	Assembly ID	Total number of protein	Total sequence length/Mb
Model organisms	<i>Saccharomyces cerevisiae</i>	GCF_000146045.2	6002	12.16
	<i>Aspergillus nidulans</i>	GCF_000149205.2	9556	30.28
	<i>Neurospora crassa</i>	GCF_000182925.2	10812	41.10
Necrotrophic	<i>Botrytis cinerea</i>	GCF_000143535.2	13703	42.63
	<i>Parastagonospora nodorum</i>	GCF_000146915.1	15994	37.21
	<i>Sclerotinia sclerotiorum</i>	GCF_000146945.2	14490	38.46
	<i>Verticillium dahliae</i>	GCF_000150675.1	10535	33.90
	<i>Verticillium alfalfae</i>	GCF_000150825.1	10237	32.86
	<i>Bipolaris maydis</i>	GCF_000354255.1	12705	32.93
	Hemibiotrophic	<i>Pyricularia oryzae</i>	GCF_000002495.2	12989
<i>Colletotrichum graminicola</i>		GCF_000149035.1	12020	51.64
<i>Fusarium oxysporum</i>		GCF_000149955.1	27347	61.39
<i>Zymoseptoria tritici</i>		GCF_000219625.1	10941	39.69
<i>Fusarium graminearum</i>		GCF_000240135.3	13313	36.46
<i>Colletotrichum higginsianum</i>		GCF_001672515.1	14650	50.72
Biotrophic		<i>Sporisorium reilianum</i>	GCA_900162835.1	6682
	<i>Blumeria graminis</i>	GCA_900519115.1	8347	140.80
	<i>Melampsora larici-populina</i>	GCF_000204055.1	16257	101.13
	<i>Ustilago maydis</i>	GCF_000328475.2	6782	19.66

所具有信号肽的蛋白数量最少, 为 311 个; 同时, 就具有信号肽的蛋白数量占全部蛋白的比例而言, *P. oryzae* 所具有信号肽的蛋白数量占比最高, 为 13.39%, 而 *S. cerevisiae* 所具有信号肽的蛋白数量占比最低, 为 5.18% (图 1)。

2.1.2 亚细胞定位分析: 进一步通过 ProtComp v9.0 预测软件对上述菌株中具有信号肽的蛋白进行亚细胞定位分析, 结果显示, 定位在细胞外的蛋白数量为 72–731 个, 所占总蛋白数量比例为 0.85%–5.09%, 所占信号肽蛋白数量比例为 14.45%–43.28% (图 2)。其中, *M. larici-populina* 所具有的细胞外蛋白数量最多, 为 731 个, *S. cerevisiae* 所具有的细胞外蛋白数量最少, 为 72 个; *B. graminis* 所具有的细胞外蛋白数量所占比例最高, 所占总蛋白和信号肽蛋白数量比例分

别为 5.09% 和 43.28%, *S. sclerotiorum* 所具有的细胞外蛋白数量所占比例最高, 所占总蛋白和信号肽蛋白数量比例分别为 0.85% 和 14.45%, 其余蛋白分别定位在细胞质膜、胞内细胞器和细胞质。进一步分析发现, 4 种活体营养型真菌具有信号肽的胞外蛋白数量所占总蛋白数量平均所占比例最高为 3.67%, 6 种半活体营养型真菌为 2.76%, 6 种死体营养型真菌为 1.94%, 3 种模式生物为 1.56%, 除 *B. maydis* 的胞外蛋白数量所占总蛋白数量比例比 *U. maydis* 高 0.08% 外, 其余 5 种死体营养型真菌及 3 种模式生物均比活体营养型真菌胞外蛋白数量所占比例低。同时, 除 *P. oryzae* 和 *P. nodorum* 胞外蛋白数量所占信号肽蛋白数量比例为 30.88% 和 28.26% 外, 4 种活体营养型真菌胞外蛋白数量所占信号肽蛋白数量

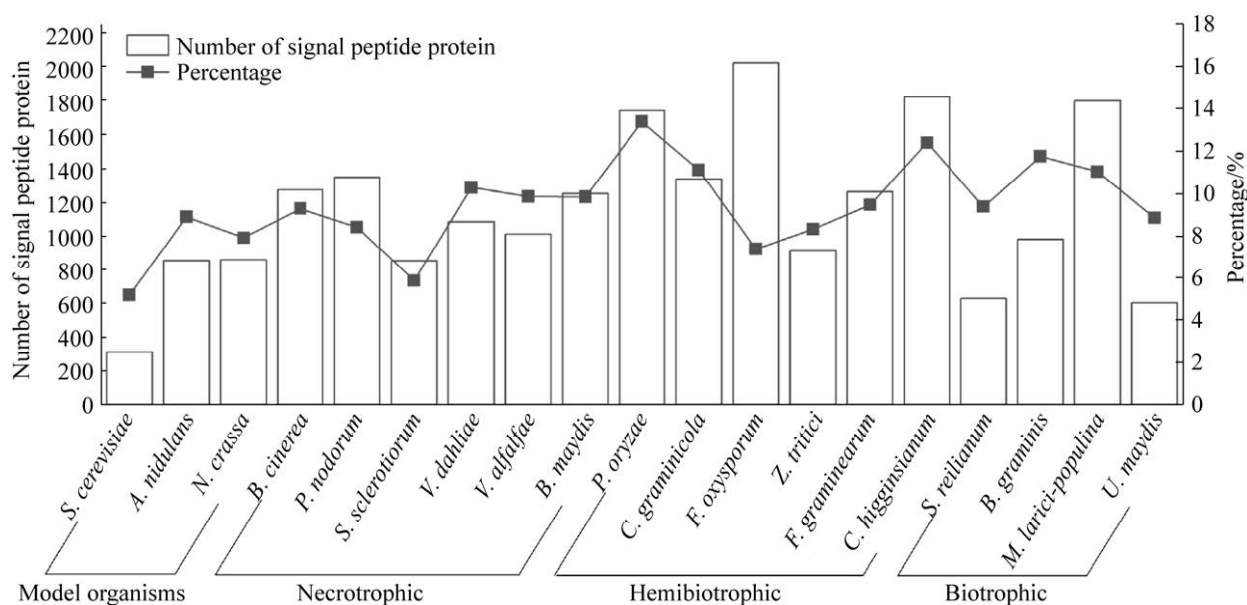


图 1. 不同营养方式真菌中具有信号肽蛋白的数量分布情况
Figure 1. Distribution of signal peptide proteins in different fungus.

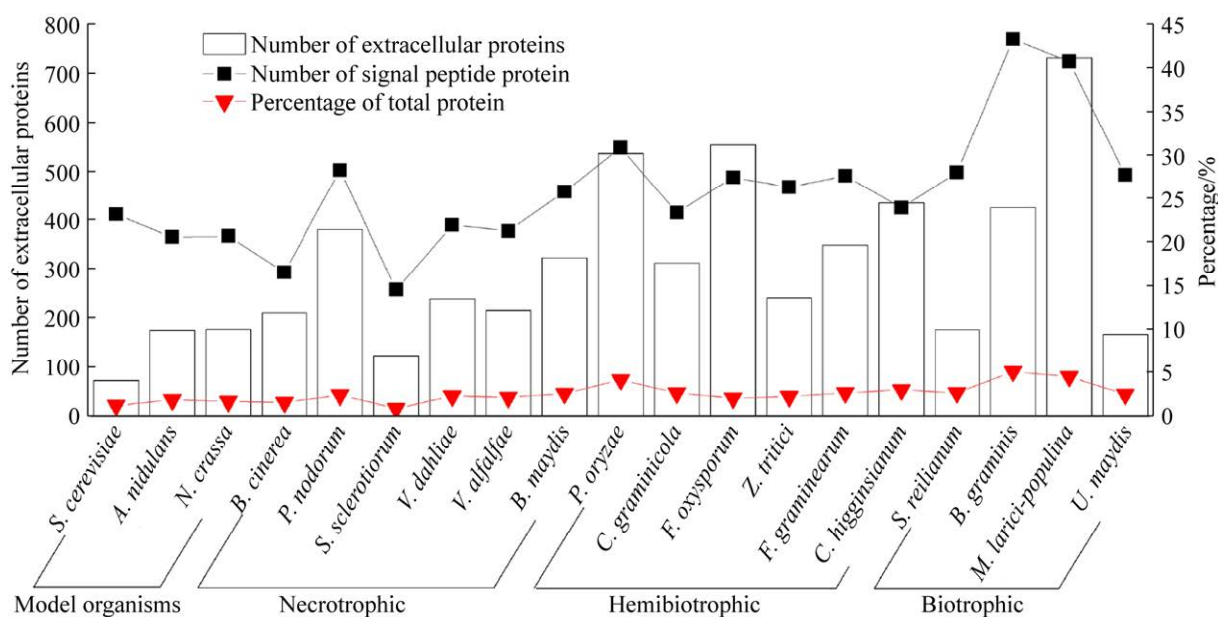


图 2. 不同营养方式真菌胞外蛋白分布情况
Figure 2. Distribution of subcellular location of different fungus.

比例均比其余菌株高, 平均为 34.94%, 而 6 种尸体营养型真菌及 3 种模式生物平均所占比例最低, 分别为 21.35% 和 21.45%。

2.1.3 跨膜结构域筛选: 利用 TMHMM v2.0 以

及 Phobius 等软件对上述所获得的蛋白进一步筛选, 结果表明, 上述所获得的蛋白中仅小部分蛋白具有 2 个及以上跨膜结构域, 蛋白数量均为 5 个以下, 其余大多数蛋白均不具有跨膜结构

域(图 3-A)。接下来,利用 big-PI predictor 对上述不具有跨膜结构域的蛋白进一步筛选,结果表明,有 1.93%–17.14%的蛋白具有 GPI 锚定蛋白,其中 4 种活体营养型真菌具有 GPI 锚定蛋白的平均所占比例最少,为 3.64%,3 种模式生物平均所占比例最高,为 11.13% (图 3-B),其余 82.86%–98.07%的蛋白均无 GPI 锚定蛋白。通过上述分析,最终明确了不同营养方式真菌中具有典型特征的分泌蛋白数量,范围为 58–712 个,

占总蛋白数量的 0.74%–4.83%,其中尤以 *M. larici-populina* 所具有的分泌蛋白数量最多,为 712 个,所占总蛋白数量比例高达 4.38%,而以 *B. graminis* 所具有的分泌蛋白数量所占总蛋白数量的比例最高,为 4.83%,所具有的分泌蛋白数量为 403 个(图 3-C)。

2.2 不同营养方式真菌中分泌蛋白对比

2.2.1 分泌蛋白数量对比:

为了进一步明确分泌蛋白数量与营养方式之间的关系,按照死体营养

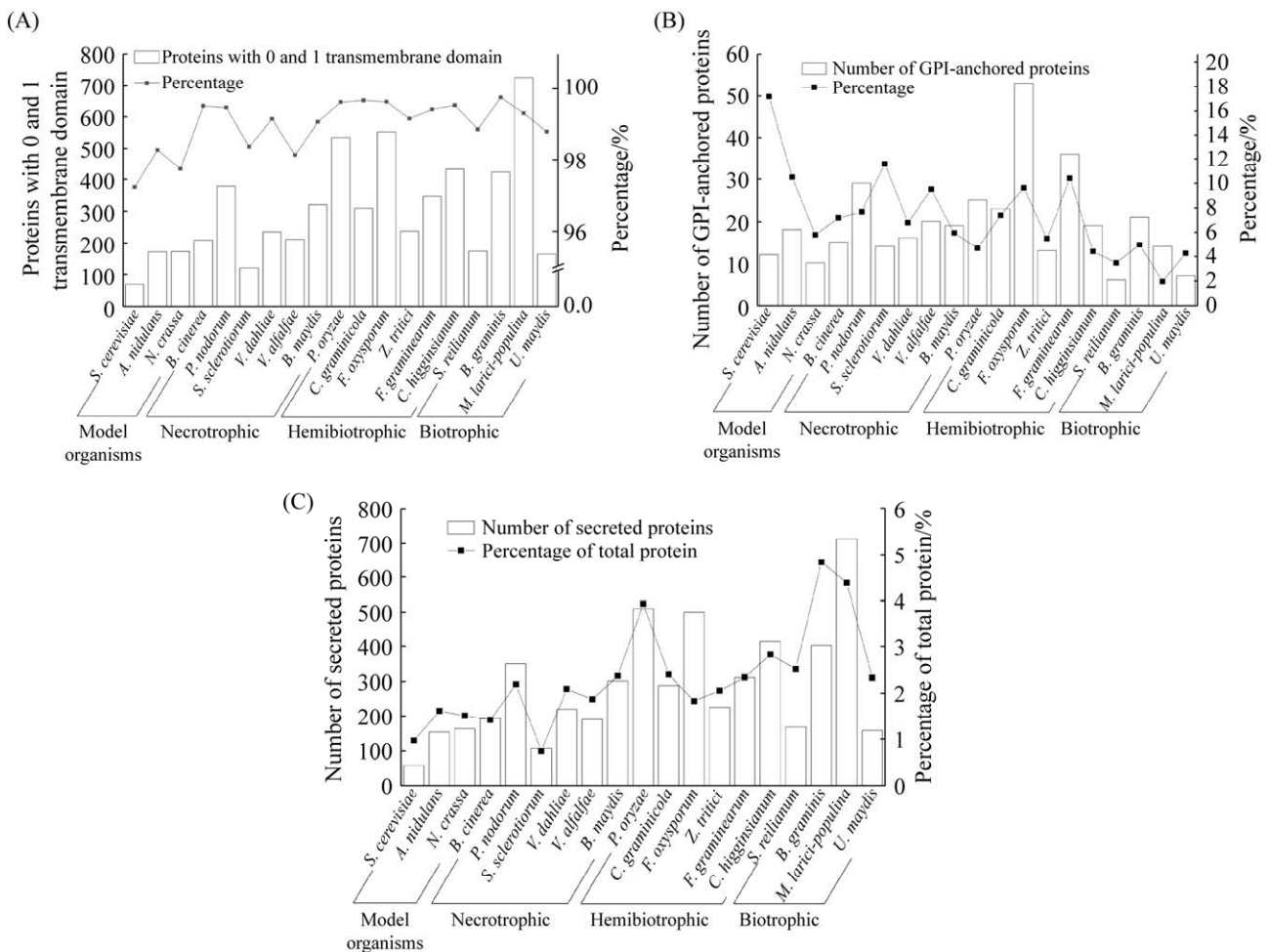


图 3. 不同营养方式真菌中具有跨膜结构域及 GPI 锚定蛋白的分布情况

Figure 3. Distribution of transmembrane domains and GPI-anchored proteins in different fungus. A: distribution of transmembrane domains; B: distribution of GPI-anchored proteins; C: distribution of secreted proteins.

型真菌、半活体营养型真菌、活体营养型真菌进行分析, 结果发现, 4 种活体营养型真菌所具有的分泌蛋白数量所占总蛋白数量的平均比例最高, 为 3.51%, 其次为 6 种半活体营养型真菌, 平均比例为 2.56%, 而 6 种死体营养型真菌和 3 种模式生物平均比例最低, 分别为 1.78% 和 1.36% (图 4)。由此可见, 不同营养方式真菌中所具有的分泌蛋白有所不同, 而造成以上菌株分泌蛋白含量的差异还有待进一步研究。

2.2.2 分泌蛋白功能对比: 对上述真菌中分泌蛋白具有的功能进行分析, 结果显示, 上述 19 种菌株的 5427 个分泌蛋白共具有 861 种功能, 具有共同功能的蛋白种类有 152 种, 蛋白数量为 4718 个, 其中有 31 种功能具有 5 个及 5 个以上的

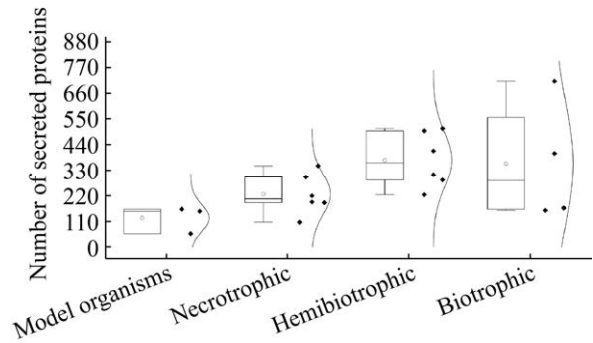


图 4. 不同营养方式真菌中分泌蛋白数量对比分析
Figure 4. Comparative analysis of secreted proteins in different fungus.

蛋白, 所具有的蛋白数量为 4418 个; 非共同功能的蛋白种类有 709 种, 蛋白数量为 709 个(图 5)。

进一步对不同营养型真菌中分泌蛋白所具有的功能分析, 结果显示, 不同营养型真菌中分

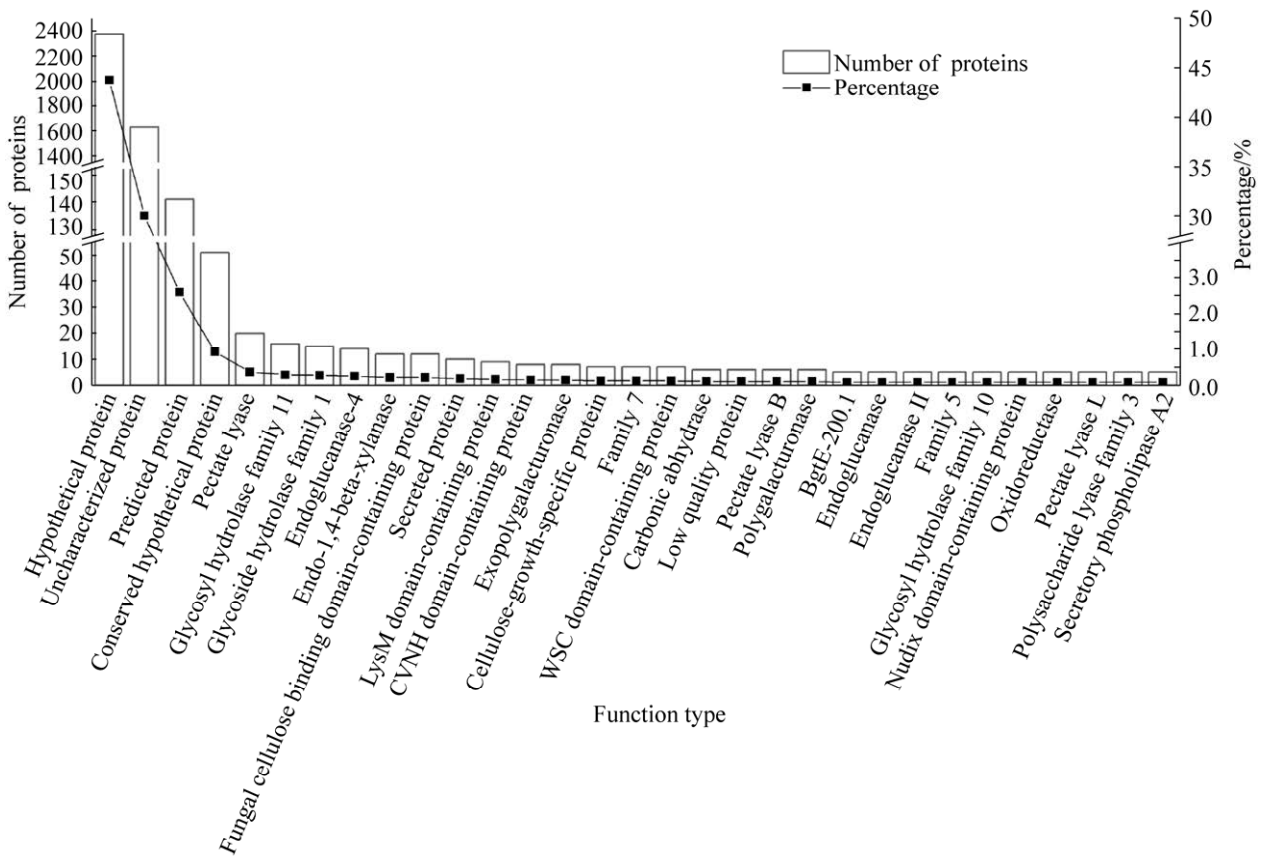


图 5. 不同营养方式真菌中分泌蛋白功能整体分布情况
Figure 5. Distribution of secretory protein function in different fungus.

泌蛋白的功能种类数量不同,活体营养型真菌所具有的种类数量居多,其次为半活体营养型真菌、死体营养型真菌、模式生物。具体而言,在4种模式生物375个分泌蛋白中,共具有100种功能,其中假设蛋白有236个,预测蛋白有25个,非特征蛋白有13个,其余功能仅具有1或2个蛋白(图6-A);在6种死体营养型真菌的1363个分泌蛋白中,共具有128种功能,其中假设蛋白有844个,非特征蛋白有133个,预测蛋白有116个,保守假设蛋白有50个,糖苷水解酶家族-1蛋白有15个,内切葡聚糖酶-4蛋白有9个,果胶酸裂解酶B蛋白有6个,其余功能仅具有1至5个蛋白(图6-B);在6个半活体营养型真菌2248个分泌蛋白中,共具有266种功能,其中假设蛋白有1153个,非特征蛋白有662个,糖基水解酶家族-1蛋白有16个,果胶酸裂解酶蛋白有15个,含CVNH结构域的蛋白、真菌纤维素结合域蛋白、分泌蛋白各有8个,LysM结构域的蛋白含有7个,其余功能仅具有1至6个蛋白(图6-C);4活体营养型真菌的1441个分泌蛋白中共具有433种功能,其中非特征蛋白有821个,假设蛋白有142个,糖苷水解酶家族-7蛋白有7个,BgtE-200.1蛋白和糖苷水解酶家族-5蛋白各有5个,BgtE-59.1蛋白和效应家族蛋白各有4个,其余功能仅具有1至3个蛋白(图6-D)。

3 讨论

近年来,国内外学者对植物病原菌中分泌蛋白和碳水化合物活性酶、效应分子等的研究已经

成为植物病理学研究的重点和难点^[7]。随着诸多植物病原菌全基因组序列的释放,为学术界深入开展上述病原菌致病基因的研究提供了重要的数据支撑。前期,包括本研究小组在内的诸多国内外学者对植物生物(如真菌^[10,15,17,27-28]和卵菌^[29])分泌蛋白进行了大量的生物信息分析和预测工作,以及对枯草芽孢杆菌^[30]、细菌性黑斑病菌等植物病原细菌的分泌蛋白展开了预测工作^[22]。各类植物中的不同真菌和细菌及卵菌中所含分泌蛋白的数量和所占比例是不完全相同的^[26]。其中,植物病原真菌分泌蛋白占总蛋白的比例约为3.65%–9.58%^[31-32],且不同活体营养型病原菌之间存在着一定差异;不同卵菌中含有的分泌蛋白占总蛋白的比例约为2.96%–4.01%^[31];不同细菌中分泌蛋白占总蛋白的比例约为5.05%–5.41%^[22]。在褐环乳牛肝菌^[33]、粘绿木霉^[34]、里氏木霉^[35]等非病原菌中分泌蛋白占总蛋白的比例为1.79%–3.89%。由于真菌测序费用相对较高,学术界选择同一种真菌的不同个体方面开展全基因组测序的报道不多,而对于同一种细菌的不同个体则开展了较多的测序报道。同时,前期本研究小组选择黄单胞菌^[36]、细菌性黑斑病菌^[22]等同一种细菌不同个体之间的分泌蛋白开展预测研究,明确了黄单胞菌和细菌性黑斑病菌等同一种细菌中不同个体之间所含的分泌蛋白在氨基酸及组成、信号肽特征方面的差异并不大,且分泌蛋白所占比例也较为相似。此外,学术界对植物病原真菌中分泌蛋白的预测及分析多针对于单种或者多种真菌开展研究工作,尚未见到关于同一种真菌不同个体之间分泌蛋白的预测及对比分析,有待于今后进一步深入研究。

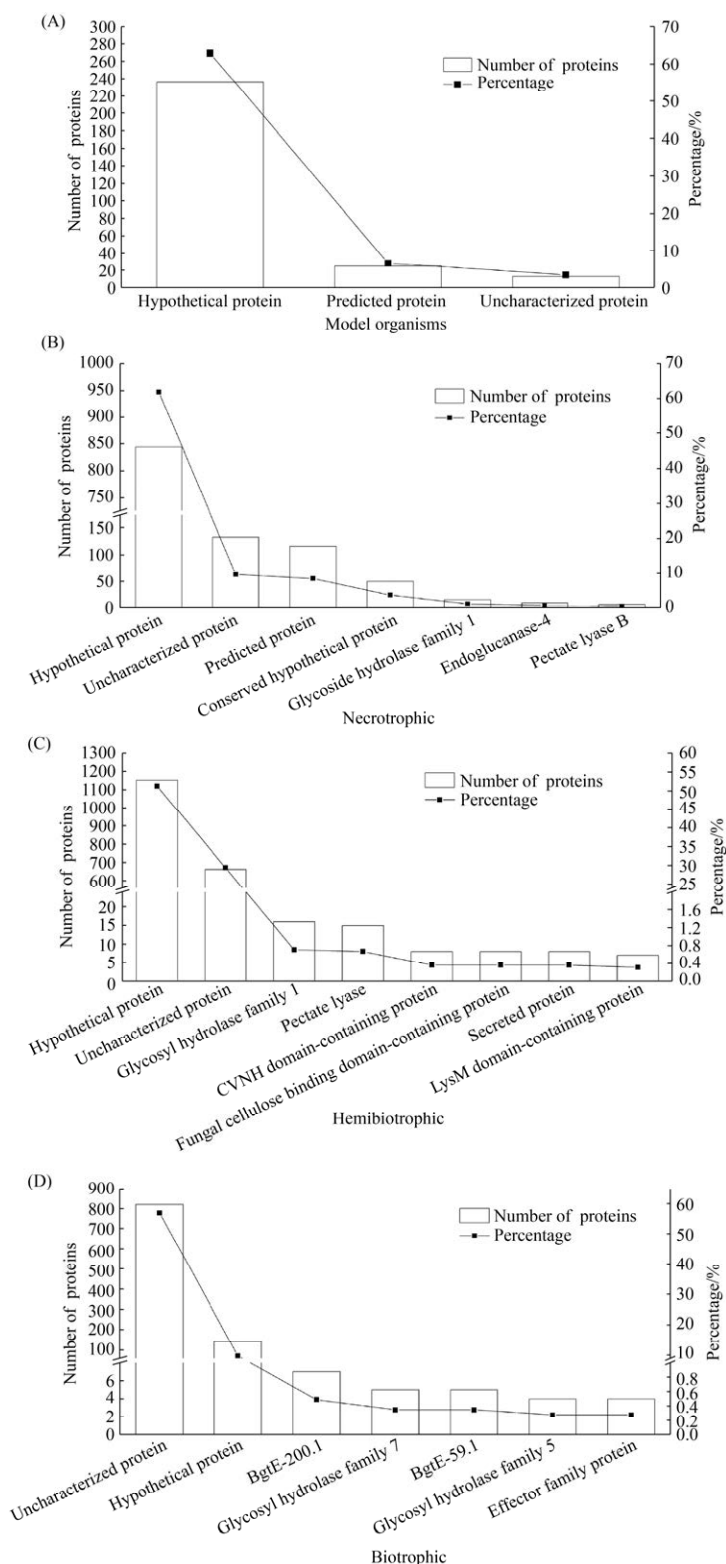


图 6. 不同营养方式真菌中分泌蛋白功能对比分析

Figure 6. Comparative analysis of functions of secreted proteins in different fungus. A: model organisms; B: necrotrophic; C: hemibiotrophic; D: biotrophic.

本研究通过对模式生物、死体营养型真菌、半活体营养型真菌及活体营养型真菌共 19 种真菌的分泌蛋白开展预测及分析,明确分泌蛋白数量为 58–712 个,占总蛋白数量的 0.74%–4.83%,活体营养型真菌所具有的分泌蛋白数量所占总蛋白数量的平均比例高于半活体营养型、死体营养型真菌以及模式生物,且死体营养型真菌和模式生物分泌蛋白数量所占总蛋白数量的平均比例仅为 1.78% 和 1.36%。活体营养型真菌具有的功能种类最多,其次为半活体营养型真菌,而模式生物具有的功能种类最少,同时功能为假设蛋白和非特征蛋白的蛋白含量最多,在模式生物和死体营养型真菌中具有预测蛋白功能的蛋白也具有较高的含量。

一般而言,植物病原真菌在实现成功侵染植物过程中,经历了克服植物固有防御反应、基础免疫反应以及系统免疫等一系列防卫反应,从而实现了对寄主植物的侵染、定殖以及系统操控等。分泌蛋白作为病原真菌克服寄主植物免疫反应过程中重要的“武器”,在实现上述过程中特别是基础免疫反应方面发挥着重要的作用。本研究发现不同营养方式真菌在分泌蛋白数量方面存在着较大的差异性,活体营养型真菌中分泌蛋白数量明显高于半活体、死体营养型真菌,同时,不同营养方式个体中也存在着差异,对于活体营养型真菌而言,其分泌蛋白所占比例最小的为 *U. maydis*, 数值为 2.33%, 最大为 *B. graminis*, 数值为 4.83%; 而就半活体营养型真菌而言,其分泌蛋白所占比例最小的为 *F. oxysporum*, 数值为 1.83%, 最大为 *P. oryzae*, 数值为 3.93%; 对于死体营养型真菌而言,其分泌蛋白所占比例最小的为 *S. sclerotiorum*, 数值为 0.74%, 最大为

B. maydis, 数值为 2.37%。此外,不同寄主植物中所具有的固有防御反应、基础免疫反应以及系统免疫不尽相同,不同营养方式植物病原真菌在克服上述防卫反应方面也具有不同的强度特征。通过上述分析,明确不同营养方式植物病原真菌之间乃至同一营养方式植物病原真菌不同个体之间,在分泌蛋白数量及比例方面均存在着一定的差异性,因此,推测活体营养型真菌需要更多数量的分泌蛋白及更大表达量的分泌蛋白,从而实现其对于寄主植物免疫系统的操控作用。同时,由于存在着基因组注释程度的差异,可能导致分析结果受到影响,本研究仅明确分泌蛋白数量及功能方面存在差异性,尚未明确分泌蛋白表达量方面是否存在差异性,有待于今后进一步明确。此外,结合前人的研究成果,明确植物病原真菌、卵菌以及细菌等不同物种之间,甚至在同一群体不同物种之间以及同一属真菌不同种之间,所含有的分泌蛋白比例存在着一定的差异,这为未来学术界进一步开展不同物种、不同群体以及同一物种不同种之间分泌蛋白的功能共性和差异化研究提供了重要的参考。

参 考 文 献

- [1] Zhu YP, Han CZ. Differences in gene clusters of synthesizing secondary metabolites of plant pathogenic fungi with different nutritional types. *Journal of Huazhong Agricultural University*, 2020, 39(6): 37–43. (in Chinese)
祝友朋, 韩长志. 不同营养类型植物病原真菌次生代谢产物合成基因簇的差异性研究. *华中农业大学学报*, 2020, 39(6): 37–43.
- [2] 许志刚. 普通植物病理学. 第 4 版. 北京: 高等教育出版社, 2009.
- [3] Jones JDG, Dangl JL. The plant immune system. *Nature*, 2006, 444(7117): 323–329.

- [4] Zhang J, Coaker G, Zhou JM, Dong XN. Plant immune mechanisms: from reductionistic to holistic points of view. *Molecular Plant*, 2020, 13(10): 1358–1378.
- [5] van Wersch S, Tian L, Hoy R, Li X. Plant NLRs: the whistleblowers of plant immunity. *Plant Communications*, 2020, 1(1): 100016.
- [6] Zhang J, Dong SM, Wang W, Zhao JH, Chen XW, Guo HS, He GC, He ZH, Kang ZS, Li Y, Peng YL, Wang GL, Zhou XP, Wang YC, Zhou JM. Plant immunity and sustainable control of pests in China: Advances, opportunities and challenges. *Scientia Sinica: Vitae*, 2019, 49(11): 1479–1507. (in Chinese)
张杰, 董莎萌, 王伟, 赵建华, 陈学伟, 郭惠珊, 何光存, 何祖华, 康振生, 李毅, 彭友良, 王国梁, 周雪平, 王源超, 周俭民. 植物免疫研究与抗病虫绿色防控: 进展、机遇与挑战. *中国科学: 生命科学*, 2019, 49(11): 1479–1507.
- [7] Han CZ, Xu X. Advance in functional research of secreted protein and CAZymes in plant pathogenic filamentous fungus. *Journal of Nanjing Forestry University: Natural Sciences Edition*, 2017, 41(5): 152–160. (in Chinese)
韩长志, 许僖. 植物病原丝状真菌分泌蛋白及 CAZymes 的研究进展. *南京林业大学学报: 自然科学版*, 2017, 41(5): 152–160.
- [8] Yang DW, Li SP, Cui HT, Zou SH, Wang W. Molecular genetic mechanisms of interaction between host plants and pathogens. *Hereditas*, 2020, 42(3): 278–294. (in Chinese)
杨德卫, 李生平, 崔海涛, 邹声浩, 王伟. 寄主植物与病原菌免疫反应的分子遗传基础. *遗传*, 2020, 42(3): 278–294.
- [9] Xia YQ, Ma ZC, Qiu M, Guo BD, Zhang Q, Jiang HB, Zhang BY, Lin YC, Xuan MR, Sun L, Shu HD, Xiao JH, Ye WW, Wang Y, Wang YM, Dong SM, Tyler BM, Wang YC. N-glycosylation shields *Phytophthora sojae* apoplast effector PsXEG1 from a specific host aspartic protease. *Proceedings of the National Academy of Sciences of the United States of America*, 2020, 117(44): 27685–27693.
- [10] Chen JS, Zheng SQ, Zheng W, Zhou J, Lu GD, Wang ZH. Prediction for secreted proteins from *Magnaporthe grisea* genome. *Scientia Agricultura Sinica*, 2006, 39(12): 2474–2482. (in Chinese)
陈继圣, 郑士琴, 郑武, 周洁, 鲁国东, 王宗华. 全基因组预测稻瘟菌的分泌蛋白. *中国农业科学*, 2006, 39(12): 2474–2482.
- [11] He YQ, Yan R, Meng G, Yang WJ, Wang ZZ, Li YF, Nie YF. Genome-scale prediction and analysis of secreted proteins and effectors in *Fusarium oxysporum* f. sp. *cubense* race 1. *Acta Phytopathologica Sinica*, 2020, 50(2): 129–140. (in Chinese)
何艳秋, 颜瑞, 蒙姑, 杨文静, 王振中, 李云锋, 聂燕芳. 香蕉枯萎病菌 1 号小种分泌蛋白与效应子的预测与分析. *植物病理学报*, 2020, 50(2): 129–140.
- [12] 纪旭. 核盘菌分泌蛋白 SsCFEM1 的功能研究. 吉林大学硕士学位论文, 2020.
- [13] Yan HL, Zhu ZC, Zhang WB, Du PX, Zhang C, Zhao WJ, Li WM. Identification and prokaryotic expression of an Sec path protein EAMY_3046 of *Erwinia amylovora*. *Plant Quarantine*, 2020, 34(4): 1–6. (in Chinese)
严海璘, 朱宗财, 张王斌, 杜培秀, 张超, 赵文军, 李为民. 梨火疫病病菌 Sec 依赖分泌蛋白 EAMY_3046 的鉴定及其原核表达. *植物检疫*, 2020, 34(4): 1–6.
- [14] Zhou XG, Hou SM, Chen DW, Tao N, Ding YM, Sun ML, Zhang SS. Genome-wide analysis of the secreted proteins of *Phytophthora infestans*. *Hereditas*, 2011, 33(7): 125–133. (in Chinese)
周晓罡, 侯思名, 陈铎文, 陶南, 丁玉梅, 孙茂林, 张绍松. 马铃薯晚疫病病菌全基因组分泌蛋白的初步分析. *遗传*, 2011, 33(7): 125–133.
- [15] Tian L, Chen JY, Chen XY, Wang JN, Dai XF. Prediction and analysis of *Verticillium dahliae* VdLs.17 secretome. *Scientia Agricultura Sinica*, 2011, 44(15): 3142–3153. (in Chinese)
田李, 陈捷胤, 陈相永, 汪佳妮, 戴小枫. 大丽轮枝菌 (*Verticillium dahliae* VdLs.17) 分泌组预测及分析. *中国农业科学*, 2011, 44(15): 3142–3153.
- [16] Xu M, Chen JY, Ma XF, Wang XY, Guo W, Dai XF. Analysis on pathogenic function of a specific secreted protein in the high virulent *Verticillium dahliae* strain VDG1. *Cotton Science*, 2015, 27(4): 337–345. (in Chinese)
徐明, 陈捷胤, 马雪峰, 王新艳, 郭维, 戴小枫. 高毒力大丽轮枝菌 VDG1 特异分泌蛋白 HSSP 致病相关功能分析. *棉花学报*, 2015, 27(4): 337–345.
- [17] Zhu YP, Cai WY, Han CZ. Prediction of secreted proteins of *Fusarium oxysporum* based on whole genome sequence. *Journal of Henan Normal University: Natural Science Edition*, 2019, 47(2): 92–97. (in Chinese)
祝友朋, 蔡旺芸, 韩长志. 基于全基因组序列的尖孢镰刀菌分泌蛋白预测及其特征分析. *河南师范大学学报: 自然科学版*, 2019, 47(2): 92–97.

- [18] 刘雪禄. 柑橘黄龙病菌分泌蛋白 Clsp33 致病机理初步研究. 西南大学硕士学位论文, 2019.
- [19] 张治萍, 郝志刚, 李迎宾, 罗来鑫, 李健强. *Ceratocystis fimbriata* 全基因组中分泌蛋白的预测. 中国植物病理学会 2019 年学术年会论文集. 成都, 2019: 141.
- [20] 张悦, 齐悦, 韦杰, 李建嫒, 杨文香, 刘大群. 小麦叶锈菌分泌蛋白的预测. 中国植物病理学会 2019 年学术年会论文集. 成都, 2019: 151.
- [21] Han CZ, Wang J. Bioinformatics analysis of type VI secretion proteins in *Xanthomonas campestris* pv. raphani 756C. *Journal of Huazhong Agricultural University*, 2016, 35(4): 42–48. (in Chinese)
韩长志, 王娟. 黄单胞菌 *Xanthomonas campestris* pv. raphani 756C 中 VI 型分泌蛋白的生物信息学分析. 华中农业大学学报, 2016, 35(4): 42–48.
- [22] Zhu YP, Liu HL, Han CZ. Prediction of secretory protein in walnut bacterial black spot pathogen. *Journal of Nanjing Forestry University: Natural Sciences Edition*, 2019, 43(3): 17–22. (in Chinese)
祝友朋, 刘宏莉, 韩长志. 基于全基因组序列的核桃细菌性黑斑病菌分泌蛋白的预测及特征分析. 南京林业大学学报: 自然科学版, 2019, 43(3): 17–22.
- [23] Almagro Armenteros JJ, Tsirigos KD, Sønderby CK, Petersen TN, Winther O, Brunak S, von Heijne G, Nielsen H. SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nature Biotechnology*, 2019, 37(4): 420–423.
- [24] Klee EW, Ellis LB. Evaluating eukaryotic secreted protein prediction. *BMC Bioinformatics*, 2005, 6(1): 1–7.
- [25] Möller S, Croning MDR, Apweiler R. Evaluation of methods for the prediction of membrane spanning regions. *Bioinformatics*, 2001, 17(7): 646–653.
- [26] Han CZ. Prediction for candidate effector proteins from *Phytophthora cinnamomi* genome. *Journal of Nanjing Forestry University: Natural Sciences Edition*, 2015, 39(2): 69–74. (in Chinese)
韩长志. 全基因组预测樟疫霉的候选效应分子. 南京林业大学学报: 自然科学版, 2015, 39(2): 69–74.
- [27] Han CZ. Prediction for secreted proteins from *Colletotrichum graminicola* genome. *Biotechnology*, 2014, 24(2): 36–41. (in Chinese)
韩长志. 全基因组预测禾谷炭疽菌的分泌蛋白. 生物技术, 2014, 24(2): 36–41.
- [28] Yu QL, Ma L, Liu L, Yang J, Su Y, Wang YY, Zhu YY, Li CY. Primary analysis of host-targeting-motif harbored secreted proteins in genome of *Fusarium graminearum*. *Biotechnology Bulletin*, 2008(1): 160–165, 180. (in Chinese)
于钦亮, 马莉, 刘林, 杨静, 苏源, 王云月, 朱有勇, 李成云. 禾谷镰刀菌基因组中含寄主靶向模体分泌蛋白功能的初步分析. 生物技术通报, 2008(1): 160–165, 180.
- [29] Liu CL, Li JD. Recent advances on the development of fungicides used for control of oomycete diseases. *Pesticides*, 2000, 39(8): 1–3. (in Chinese)
刘长令, 李继德. 卵菌纲病害用杀菌剂的开发进展. 农药, 2000, 39(8): 1–3.
- [30] Han CZ. Prediction algorithm on secretory proteins of *Bacillus subtilis* XF-1. *Fujian Journal of Agricultural Sciences*, 2016, 31(4): 384–388. (in Chinese)
韩长志. 全基因组预测枯草芽孢杆菌 XF-1 的分泌蛋白. 福建农业学报, 2016, 31(4): 384–388.
- [31] Chen XY, Chen JY, Xiao HL, Gui YJ, Li L, Dai XF. The comparative analysis of secreted CAZymes in phytopathogenic fungi with different lifestyle. *Acta Phytopathologica Sinica*, 2014, 44(2): 163–172. (in Chinese)
陈相永, 陈捷胤, 肖红利, 桂月靖, 李蕾, 戴小枫. 植物病原真菌寄生性与分泌蛋白组 CAZymes 的比较分析. 植物病理学报, 2014, 44(2): 163–172.
- [32] Chen QG, Wang CJZ, Yang M, Zhou EX. Prediction of candidate effectors from the genome of *Colletotrichum higginsianum*. *Chinese Journal of Tropical Crops*, 2015, 36(6): 1105–1111. (in Chinese)
陈琦光, 王陈骄子, 杨媚, 周而勋. 希金斯刺盘孢全基因组候选效应分子的预测. 热带作物学报, 2015, 36(6): 1105–1111.
- [33] Han CZ, Liu Y, Ren XM. Prediction of secretory protein of *Suillus luteus*. *Jiangsu Agricultural Sciences*, 2017, 45(23): 45–48. (in Chinese)
韩长志, 刘艳, 任雪敏. 全基因组预测褐环乳牛肝菌的分泌蛋白. 江苏农业科学, 2017, 45(23): 45–48.
- [34] Han CZ, Xu X. Prediction and genetic relationship of CAZymes proteins from *Trichoderma virens* Gv29-8 genome. *Journal of Huazhong Agricultural University*, 2017, 36(5): 39–44. (in Chinese)
- [35] 韩长志, 许僖. 粘绿木霉 Gv29-8 的碳水化合物活性酶类蛋白预测及遗传关系分析. 华中农业大学学报, 2017, 36(5): 39–44.

[36] Han CZ, Zhu YP, Xu X. Prediction for CAZymes from *Trichoderma reesei* QM6a genome. *Southwest China Journal of Agricultural Sciences*, 2018, 31(4): 705–710. (in Chinese)

韩长志, 祝友朋, 许僖. 全基因组预测里氏木霉 QM6a 的碳水化合物活性酶类蛋白. *西南农业学报*, 2018, 31(4): 705–710.

[1] Qin Y, Zhu YP, Han CZ. Prediction and characteristic

analysis of *Xanthomonas campestris* secretory protein based on whole genome sequence. *Jiangsu Journal of Agricultural Sciences*, 2021, 37(1): 53–59. (in Chinese)

覃悦, 祝友朋, 韩长志. 基于全基因组序列的黄单胞菌分泌蛋白质预测及其特征分析. *江苏农业学报*, 2021, 37(1): 53–59.

Prediction and characteristic analysis of secreted protein in different fungi

Yue Qin¹, Changzhi Han^{1,2*}

¹ College of Biodiversity Conservation, Southwest Forestry University, Kunming 650224, Yunnan Province, China

² The Key Laboratory of Forest Disaster Warning and Control of Yunnan Province, Kunming 650224, Yunnan Province, China

Abstract: [Objective] Secreted proteins play an important role in the pathogenic process of phytopathogenic fungi. Predecessors mostly used a single strain to predict and analyze secreted proteins. There are no reports on the prediction and comparative studies of secreted proteins in multiple types of fungi. [Methods] This study is based on the whole genome sequence of fungi with a variety of different nutrition methods. According to the characteristics of the secreted protein, online analysis programs such as SignalP v5.0 and ProtComp v9.0 were used to analyze the secreted proteins of 19 fungi, which including model organisms, hemibiotrophic, necrotrophic and biotrophic. [Results] The results showed that among the about 130000 protein sequences of the above-mentioned fungi, the proportion of secreted proteins was 0.74% to 4.83%. Among them, the secreted proteins of biotrophic accounted for the highest proportion, with an average of 3.51%. The average ratio of hemibiotrophic and model organisms was the lowest, with an average of 1.78% and 1.36%. At the same time, biotrophic have the most types of functions, 433 species, followed by necrotrophic, with 266 species, and model organisms have the least functional types, at 100, among which functions are hypothetical protein and non-characteristic. [Conclusion] This research has laid a solid theoretical foundation for the in-depth analysis of secreted proteins in realizing the pathogenic mechanism of fungi with different nutritional methods.

Keywords: fungi, nutrition mode, secreted protein, function, comparative analysis

(本文责编: 李磊)

Supported by the National Natural Science Foundation of China (31960314), by the Yunnan Provincial High-level Talent Training Support Program “Youth Top-notch Talent” Special Project and by the Yunnan Applied Basic Research Program (2018FG001-028)

*Corresponding author. Tel: +86-871-63862918; E-mail: hanchangzhi2010@163.com

Received: 26 March 2021; Revised: 20 May 2021; Published online: 28 September 2021