

# 枯草芽孢杆菌 XF-1 的碳水化合物活性酶类(CAZymes)蛋白预测与遗传关系分析

韩长志, 祝友朋, 许僖

(西南林业大学 生物多样性保护与利用学院; 云南省森林灾害预警与控制重点实验室, 昆明 650224)

**摘要:** 枯草芽孢杆菌 XF-1 对多种植物病原菌具有较好的抑制作用, 是农业生产上非常重要的生防菌之一。碳水化合物活性酶类蛋白(CAZymes)作为植物病原真菌、细菌中重要的蛋白, 在其生长和发育过程中发挥着重要作用。基于前期获得的枯草芽孢杆菌 XF-1 中 104 个分泌蛋白基本序列信息, 利用 CAT 预测程序对该菌中的 CAZymes 蛋白进行预测分析, 明确其含有 CAZymes 蛋白的数量是 58 个, 上述蛋白可以分为主要类别和复合类别两大类, 前者包括 26 个 GHs、21 个 CBMs 以及 2 个 CEs、3 个 PLs、2 个 GTs; 后者则包括 4 个复合蛋白类型 GH/CBM。该研究有效地实现了枯草芽孢杆菌 XF-1 中 CAZymes 的预测, 为进一步开展其功能研究打下坚实的理论基础。

**关键词:** 枯草芽孢杆菌 XF-1; 碳水化合物活性酶蛋白; 预测程序; 遗传关系

**中图分类号:** S436.35

**文献标志码:** A

碳水化合物活性酶类(Carbohydrate-Active Enzymes, CAZymes)作为植物病原真菌、细菌中重要的一类蛋白, 在植物病原菌的生长发育过程中、特别是侵染植物过程中发挥着重要的作用。近些年, 国内外学者关于植物病原菌中 CAZymes 的研究报道呈现较快的增长<sup>[1]</sup>, 主要涉及以下 6 大类, 糖苷水解酶(Glycoside Hydrolases, GHs)<sup>[2]</sup>、糖基转移酶(Glycosyl Transferases, GTs)<sup>[3]</sup>、多糖裂解酶(Polysaccharide Lyases, PLs)<sup>[4]</sup>、碳水化合物酯酶(Carbohydrate Esterases, CEs)<sup>[4]</sup>、辅助酶类家族(Auxiliary Activities, AAs)<sup>[5]</sup>以及碳水化合物绑定结构(Carbohydrate-Binding Modules, CBMs)<sup>[6]</sup>。

枯草芽孢杆菌(*Bacillus subtilis*)XF-1 不仅对十字花科根肿病菌具有较好的抑制效果<sup>[7]</sup>, 也对致病疫霉、玉米小斑病菌、稻瘟病菌等 21 种植物病原菌具有较好的抑制效果<sup>[8]</sup>。目前, 研究成果主要有: (1)XF-1 在大白菜根际的定殖能力研究<sup>[9-10]</sup>; (2)XF-1 中主要功能蛋白的纯化<sup>[11]</sup>以及粗蛋白的抑菌活性<sup>[12]</sup>、粗提物对 11 种植物病原菌抑制作用<sup>[13]</sup>、膏状悬浮剂增稠液<sup>[14]</sup>等; (3)利用不同培养基, 对 XF-1 产生脂肽抗生素种类和产量的影响方面<sup>[15]</sup>; (4)XF-1 所具有的抑菌机制<sup>[16, 17]</sup>等方面。前期, 基于其全基因组序列<sup>[18]</sup>, 通过 SignalP、ProtComp、TMHMM 等生物信息学在线分析程序, 明确其存在着 104 个分泌蛋白<sup>[19]</sup>, 然而, 尚未见对其 CAZymes 蛋白的预测报道。同时, 前人多对植物病原菌中的 CAZymes 进行分析研究<sup>[1]</sup>, 尚缺乏对于植物病原拮抗菌中 CAZymes 的研究报道。

本研究利用 CAT(CAZymes Analysis Toolkit, <http://cricket.ornl.gov/cgi-bin/cat.cgi>)<sup>[20]</sup>在线分析程序对 XF-1 中的 CAZymes 蛋白进行找寻, 同时, 对其遗传关系以及蛋白互作关系进行分析, 以期深入解析 CAZymes 蛋白在该菌发挥生防作用的功能研究提供的理论借鉴。

## 1 材料与方 法

### 1.1 枯草芽孢杆菌 XF-1 分泌蛋白序列

XF-1 中的 104 个分泌蛋白来源于前期的研究结果<sup>[19]</sup>。

收稿日期: 2017-05-01; 修回日期: 2018-02-20.

基金项目: 国家自然科学基金(31560211); 云南省森林灾害预警与控制重点实验室开放基金(ZK150004); 云南省研发经费投入补助项目(2016YB462&2017YB236); 云南省林学一流学科建设项目(51600625).

作者简介(通信作者): 韩长志(1981-), 男, 河北石家庄市人, 西南林业大学副教授, 博士, 研究方向为经济林木病害生物防治与真菌分子生物学, E-mail: hanchangzhi2010@163.com

## 1.2 CAZymes 预测

利用在线工具 CAT<sup>[20]</sup> 来开展预测工作.

## 1.3 CAZymes 遗传关系分析

对上述所获得的 CAZymes 蛋白序列 ClustalX<sup>[21]</sup> 进行多重比对分析,随后利用 MEGA 7.0 软件<sup>[22]</sup> 构建系统进化树.

## 1.4 CAZymes 相互作用关系分析

利用 STRING v.10<sup>[23]</sup> 在线数据库对 CAZymes 蛋白序列开展相互作用预测,探寻蛋白-蛋白相互作用关系.

# 2 结果与分析

## 2.1 枯草芽孢杆菌 XF-1 中含有 58 个 CAZymes

枯草芽孢杆菌 XF-1 中 58 个 CAZymes 中, GHs 的含量较多, 为 26 个, 所占比例为 44.83%, 其次是 CBMs, 为 21 个, 所占比例为 36.21%, 而 CEs, PLs, GTs 的数量分别为 2、3、2, 所占比例分别为 3.45%、5.17%、3.45%. 值得注意的是, 该菌中并没有辅助酶家族 AAs 得以预测, 这与生物发挥正常的功能明显存在差异, 有待今后通过试验得以验证(图 1). 此外, 复合类型只有一种类型, 为 GH/CBM 型, 其数量为 4 个.

## 2.2 枯草芽孢杆菌 XF-1 中含有 26 个 GHs

通过对上述糖基水解酶蛋白进行分析, 发现 GH23 含量较多, 为 6 个, 所占比例为 23.08%, 其次 GH3、GH18 均含有 3 个, 而 GH47、GH53、GH73、GH102 均含有 2 个(表 1).

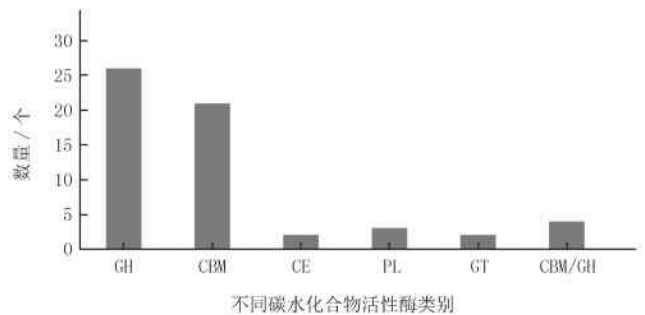


图 1 枯草芽孢杆菌 XF-1 中 58 个 CAZymes 分布情况

表 1 枯草芽孢杆菌 XF-1 中糖基水解酶蛋白的分布情况

不同类型	ID	数量/个
GH3	gi 449092877,gi 449092875,gi 449094570	3
GH13	gi 449092997	1
GH18	gi 449095853,gi 449093775,gi 449096266	3
GH19	gi 449092919	1
GH23	gi 449094895,gi 449096413,gi 449093056,gi 449095934,gi 449093031,gi 449095382	6
GH32	gi 449095140	1
GH43	gi 449095327	1
GH46	gi 449095125	1
GH47	gi 449096266,gi 449096307	2
GH53	gi 449095863,gi 449096049	2
GH68	gi 449095899	1
GH73	gi 449094916,gi 449095551	2
GH102	gi 449095649,gi 449094597	2
合计	—	26

## 2.3 枯草芽孢杆菌 XF-1 中含有 21 个 CBMs

通过对上述碳水化合物结合结构蛋白进行分析, 发现 CBM50 较多, 为 11 个, 其次 CBM5 为 3 个, 上述蛋白所占比例为 66.67%, 其余涉及的 GH9, GH12, GH26 等, 数量均为 1 个(表 2).

表 2 枯草芽孢杆菌 XF-1 中碳水化合物结合结构蛋白的分布情况

不同类型	ID	数量/个
CBM5	gi 449094167,gi 449096300,gi 449093730	3
CBM9	gi 449094519	1
CBM12	gi 449096307	1
CBM26	gi 449092997	1
CBM32	gi 449094223	1
CBM50	gi 449094597,gi 449095649,gi 449093642,gi 449093639,gi 449095788,gi 449094095,gi 449095701,gi 449095320,gi 449095914,gi 449093028,gi 449094570	11
CBM57	gi 449094223	1
CBM63	gi 449094554	1
CBM66	gi 449095140	1
合计	—	21

2.4 枯草芽孢杆菌 XF-1 中含有 CEs,PLs,GTs 的数量分别是 2、3、2 个

通过对上述碳水化合物酯酶蛋白进行分析,发现含有的 2 类,分别为 CE4、CE12,其 ID 分别为 gi|449093505、gi|449096375.同时,通过对上述多糖裂解酶蛋白进行分析,发现只含有 PL1 和 PL3 两种家族,数量分别为 2、1 个,PL1 的 ID 分别为 gi|449094556、gi|449093465,PL3 的 ID 为 gi|449095948.此外,通过对上述糖基转移酶蛋白进行分析,发现 2 种 GTs,分别为 GT2、GT39,ID 分别为 gi|449094531、gi|449096448.

2.5 枯草芽孢杆菌 XF-1 中含有 4 个 GH/CBM 复合类型蛋白

通过对 4 个复合类型蛋白进行分析,结果显示,上述复合类型蛋白种类均属于 GH/CBM 复合类型,具体为:GH102/CBM50, GH32/CBM66, GH13/CBM26, GH3/CBM50,其 ID 分别为 gi|449094597、gi|449095140、gi|449092997、gi|449094570.

2.6 遗传关系分析

通过对上述所获得的 CAZymes 蛋白进行遗传关系分析,结果表明,58 个 CAZymes 可以分为 5 个大类,分别为 GHs, GH/CBM, CBMs, GH/PL 以及 GH/CE(图 2).然而,通过 CAT 预测分析,上述 CAZymes 蛋白仅存在一种复合类型 GH/CBM,而不具有 GH/PL, GH/CE 两类.遗传关

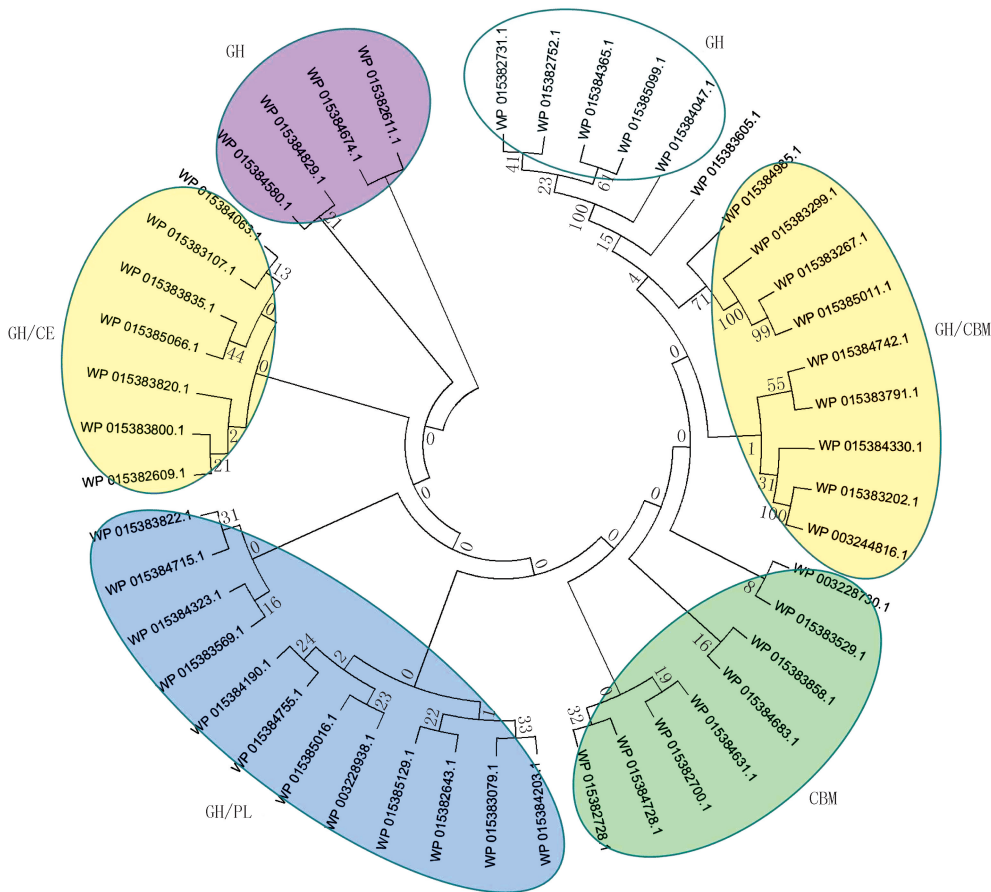


图 2 遗传关系分析

系结果反映出蛋白之间在序列方面存在着相似性,上述 GH/PL,GH/CE 蛋白彼此之间是否存在功能方面的相似性,有待于后续进一步开展研究.

### 2.7 枯草芽孢杆菌中 CAZymes 蛋白相互关系分析

蛋白质相互作用网络已经成为系统生物学重要的研究方向之一,本研究通过 STRING v.10 在线数据分析,由于 XF-1 菌株并未在该数据库中,选择编号为 168 的枯草芽孢杆菌菌株进行后续分析,明确大部分 CAZymes 蛋白具有相互作用关系,该结果为进一步解析 XF-1 菌株中 CAZymes 蛋白互作用关系提供重要的理论参考(图 3).

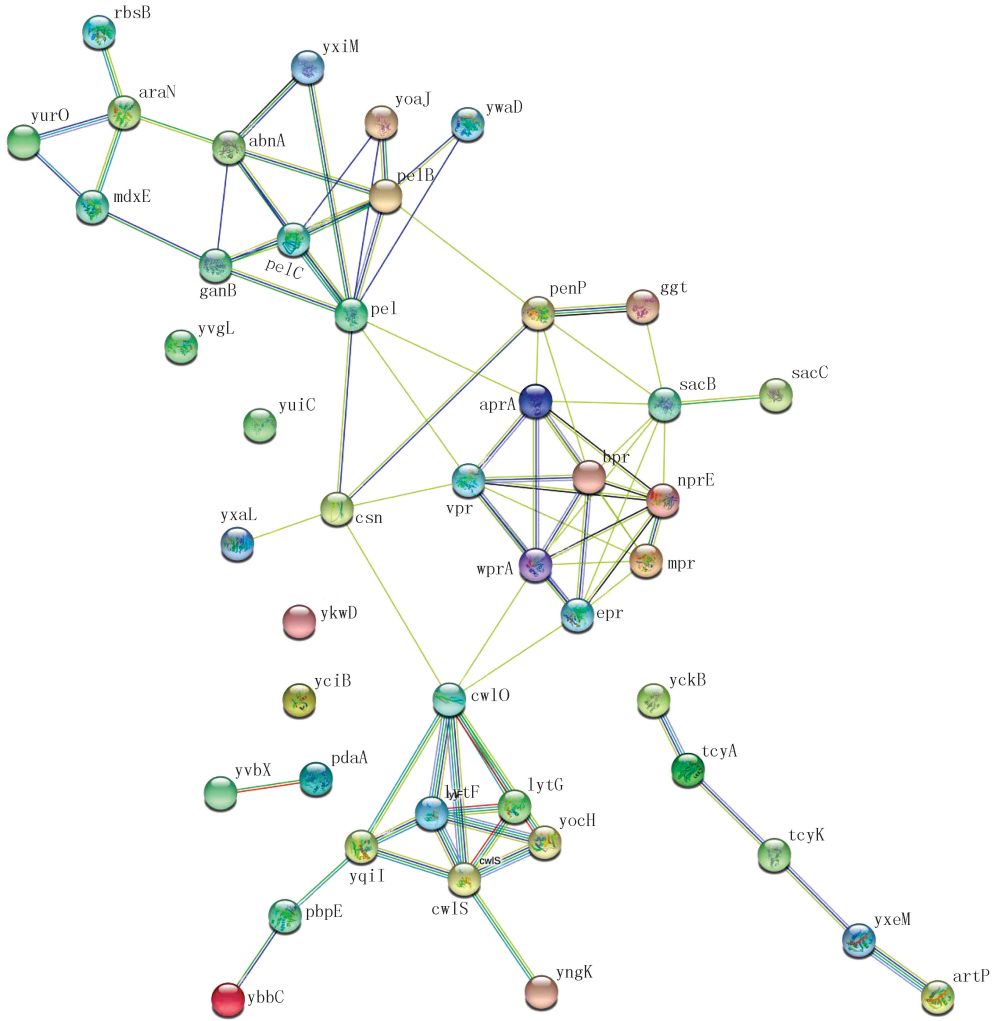


图3 枯草芽孢杆菌 XF-1 中 CAZymes 蛋白相互关系分析

## 3 讨论

目前,碳水化合物结构(功能)活性酶类数据库(Carbohydrate-Active enZYmes Database),作为专业描述结构相关的酶催化或者是碳水化合物绑定(或者是功能域)模块的重要数据库<sup>[24]</sup>,为进一步深入开展包括植物病原在内的真菌、细菌中胞外酶细胞壁降解酶活性基因的注释与聚类分析提供了极大便利<sup>[20]</sup>.截至 2016 年 5 月 18 日,该数据库中含有 GHs,GTs,PLs,CEs,AAAs,CBMs 类别的家族分别有 135、98、24、16、13 以及 74 个.

本研究主要基于前期通过全基因组序列预测所获得分泌蛋白,依据现有网站所获得 CAZymes,是否为枯草芽孢杆菌 XF-1 中的全部,均有待于今后进一步开展后续试验进行验证.同时,随着目前大量植物病原细

菌基因组 CAZymes 的逐步完成<sup>[25-28]</sup>,通过比较基因组学分析病原菌的寄生性与基因组变异的关系已经成为研究热点<sup>[29]</sup>.一般而言,CAZymes 蛋白作为植物病原菌侵染寄主植物过程中突破寄主细胞的第一道屏障——植物细胞壁的关键因素<sup>[30]</sup>.然而,枯草芽孢杆菌 XF-1 是多种植物病原菌的拮抗菌,显然不是植物病原菌,其内所含有的 CAZymes 所发挥的功能与植物病原菌显然不同,即同一类别蛋白在不同菌中所具有的功能不尽相同,有待进一步开展后续试验研究.

## 4 结 论

本研究基于前期所获得的 104 个分泌蛋白,利用 CAT 在线预测程序,明确枯草芽孢杆菌 XF-1 中含有 58 个 CAZymes,涉及主要类别和复合类别两大类,前者包括 26 个 GHs、21 个 CBMs 以及 12 个 CEs、3 个 PLs、2 个 GTs;后者包括 4 个 GH/CBM.同时,明确上述蛋白在遗传关系方面具有较好的对应关系,该研究为深入开展枯草芽孢杆菌 XF-1 中的 CAZymes 的功能研究打下了坚实的理论基础.

## 参 考 文 献

- [1] 韩长志,许億.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.
- [2] Bourne Y, Henrissat B. Glycoside hydrolases and glycosyltransferases: families and functional modules[J]. Current opinion in structural biology, 2001, 11(5): 593-600.
- [3] Campbell J A, Davies G J, Bulone V, et al. A classification of nucleotide-diphospho-sugar glycosyltransferases based on amino acid sequence similarities[J]. Biochem J, 1997, 326: 929-939.
- [4] Lombard V, Bernard T, Rancurel C, et al. A hierarchical classification of polysaccharide lyases for glycogenomics[J]. Biochem J, 2010, 432(3): 437-444.
- [5] Levasseur A, Drula E, Lombard V, et al. Expansion of the enzymatic repertoire of the CAZy database to integrate auxiliary redox enzymes[J]. Biotechnol Biofuels, 2013, 6(1): 41.
- [6] Boraston A B, Bolam D N, Gilbert H J, et al. Carbohydrate-binding modules: fine-tuning polysaccharide recognition[J]. Biochem J, 2004, 382(3): 769-781.
- [7] 何月秋,熊国如,范成明.防治十字花科根肿病的生物制剂及其应用[P]. In: CN101416641, 2009-04-29.
- [8] 熊国如,赵更峰,范成明,等.生防菌株 XF-1 的鉴定和抑菌谱的测定[J].云南农业大学学报,2009,24(2):190-194.
- [9] 刘庆丰,熊国如,王志远,等.枯草芽孢杆菌 XF-1 在大白菜根际定殖能力研究[C].中国植物病理学会 2009 年学术年会[A].中国云南昆明,2009.
- [10] 刘庆丰,熊国如,毛自朝,等.枯草芽孢杆菌 XF-1 的根围定殖能力分析[J].植物保护学报,2012,39(5):425-430.
- [11] 赵静,熊国如,周惠萍,等.枯草芽孢杆菌 XF-1 防治十字花科根肿病的功能蛋白纯化[C].中国植物病理学会 2009 年学术年会[A].中国云南昆明,2009.
- [12] 熊国如,张翠英,刘庆丰,等.枯草芽孢杆菌 XF-1 粗蛋白抑菌活性初步研究[J].植物保护,2009,35(4):92-95.
- [13] 贾媛,孙睿揆,吴毅歆,等.枯草芽孢杆菌 XF-1 提取物对 11 种植物病原菌的抑制作用[J].安徽农业大学学报,2011,38(5):753-756.
- [14] 孙海霞,孙睿揆,吴毅歆,等.枯草芽孢杆菌 XF-1 膏状悬浮剂增稠剂的筛选[J].现代农药,2011,10(5):24-27.
- [15] 李兴玉,毛自朝,吴毅歆,等.不同培养基对 *Bacillus subtilis* XF-1 产脂肽抗生素的影响[J].内蒙古农业大学学报(自然科学版),2014(6):49-53.
- [16] 赵静,熊国如,王志远,等.枯草芽孢杆菌 XF-1 中脱乙酰几丁质酶在大肠杆菌中的表达及其抑菌活性[J].中国生物防治学报,2011,27(4):504-509.
- [17] Li X Y, Wang Y H, He Y Q. Diversity and active mechanism of fengycin-type cyclopeptides from *Bacillus subtilis* XF-1 against *Plasmodiophora brassicae*[J]. Journal of microbiology and biotechnology, 2013, 23(3): 313-321.
- [18] Guo S, Mao Z, Wu Y, et al. Genome Sequencing of *Bacillus subtilis* Strain XF-1 with High Efficiency in the Suppression of *Plasmodiophora brassicae*[J]. Genome Announc, 2013, 1(2): e0006613.
- [19] 韩长志.全基因组预测枯草芽孢杆菌 XF-1 分泌蛋白[J].福建农业学报,2016,31(4):93-97.
- [20] 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.
- [21] Thompson J D, Gibson T J, Higgins D G. Multiple sequence alignment using ClustalW and ClustalX[J]. Current Protocols in Bioinformatics, 2002, 2(1): 1-3.
- [22] Tamura K, Stecher G, Peterson D, et al. MEGA6: molecular evolutionary genetics analysis version 6.0[J]. Molecular biology and evolution, 2013, 30(12): 2725-2729.



- [23] Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J et al. STRING v10: protein - protein interaction networks, integrated over the tree of life[J]. Nucleic acids research, 2014; gku1003.
- [24] 韩长志. 基于全基因组测序的禾谷炭疽菌中碳水化合物酶类蛋白预测[J]. 河南师范大学学报(自然科学版), 2016, 44(4): 118-124.
- [25] 张武, 潘伟, 马金田, 等. 菜豆根瘤菌 CFN42 全基因组非经典分泌蛋白的预测与分析[J]. 大理学院学报, 2015, 14(6): 45-48.
- [26] 潘宏, 高天一, 吴明德, 等. 西瓜果斑病菌 T2SS 分泌蛋白的筛选和功能预测[J]. 华中农业大学学报, 2015, 34(3): 27-35.
- [27] 张武, 马金田, 潘伟, 等. 菜豆根瘤菌 *Rhizobium etli* CFN42 全基因组含信号肽分泌蛋白的鉴定与分析[J]. 基因组学与应用生物学, 2014, 33(5): 961-969.
- [28] 肖文超, 范会云, 白亭亭, 等. 香蕉细菌性软腐病菌 XJ8-3-3 基因组中 ORFs 的信号肽及分泌蛋白功能预测分析[J]. 果树学报, 2014, 31(6): 1057-1064.
- [29] Schmidt S M, Panstruga R. Pathogenomics of fungal plant parasites; what have we learnt about pathogenesis[J]. Curr Opin Plant Biol, 2011, 14(4): 392-399.
- [30] van den Brink J, de Vries R P. Fungal enzyme sets for plant polysaccharide degradation[J]. Appl Microbiol Biotechnol, 2011, 91(6): 1477-1492.

## Prediction for CAZymes proteins from *Bacillus subtilis* XF-1 genome

Han Changzhi, Zhu Youpeng, Xu Xi

(College of Biodiversity Conservation and Utilization; The Key Laboratory of Forest Disaster Warning and Control of Yunnan Province, Southwest Forestry University, Kunming 650224, China)

**Abstract:** *Bacillus subtilis* XF-1 has inhibitory effects on a variety of plant pathogens, which is an important bio-control agents on agricultural production. Carbohydrate active enzymes protein (CAZymes) was a large class of important proteins of plant pathogenic fungi and bacteria, which plays an important role in the growth and development process. Based on 104 secreted proteins obtained preliminary data, CAZymes protein was predicted with CAT forecasting procedures. And 58 CAZymes proteins were found in *B. subtilis* XF-1, which divided into major categories and composite category, and the former includes 26 GH, 21 CBM, 15 AA, 2 CE, 2 PL and 2 GT; the latter includes 4 GH/CBM. Through the above bioinformatics analysis, this research lay a solid theoretical foundation for further study its function.

**Keywords:** *Bacillus subtilis* XF-1; Carbohydrate active enzymes protein; prediction algorithm; genetics relationship

[责任编辑 王凤产]