

# 棘孢木霉糖苷水解酶 3 基因家族的生物信息学及表达模式分析

李聪, 王允\*

长江大学生命科学学院, 湖北 荆州 434025

李聪, 王允. 棘孢木霉糖苷水解酶 3 基因家族的生物信息学及表达模式分析[J]. 微生物学通报, 2023, 50(1): 1-12.

LI Cong, WANG Yun. Bioinformatics and expression analyses of glycoside hydrolase 3 genes in *Trichoderma asperellum*[J]. Microbiology China, 2023, 50(1): 1-12.

**摘要:**【背景】糖苷水解酶(glycoside hydrolase, GH) 3 基因家族成员主要编码胞外  $\beta$ -葡萄糖苷酶, 是纤维素降解中的关键酶。【目的】鉴定棘孢木霉 *GH3* 基因家族成员, 探究其在纤维素降解过程中转录水平的表达模式。【方法】通过生物信息学方法对棘孢木霉 *GH3* 基因家族成员进行鉴定, 对其基因结构、系统进化、蛋白理化性质、亚细胞定位及蛋白质三级结构进行分析, 并采用荧光定量 PCR 技术对纤维素诱导下转录水平的表达模式进行综合分析。【结果】棘孢木霉基因组共鉴定到 16 个 *GH3* 基因家族成员, 含有 1-8 个外显子, 编码蛋白质长度为 533-934 个氨基酸, 分子量为 57.82-101.91 kDa, 大多数为胞外蛋白。系统发育表明, 该基因家族成员可分为 4 组, 与里氏木霉的相似性较高。基因表达模式分析表明, 纤维素诱导下, 16 个 *GH3* 基因均有表达, 但不同成员在转录水平的表达存在差异。其中, 1 个基因呈组成型表达, 2 个基因表达下调, 13 个基因表达上调。棘孢木霉的胞外  $\beta$ -葡萄糖苷酶活力在纤维素诱导下明显提升, 与 *GH3* 基因家族成员在转录水平的整体表达模式相一致。【结论】棘孢木霉基因组共包含 16 个 *GH3* 基因家族成员, 而且多为亲水性稳定蛋白质, 有良好的商业开发潜力; 绝大部分 *GH3* 基因家族成员的转录量在纤维素诱导下呈上调, 与胞外  $\beta$ -葡萄糖苷酶活性变化相一致, 说明其在纤维素降解中发挥重要作用, 这为天然木质纤维原料的资源化利用和纤维素降解工程菌的构建提供了重要的理论依据。

**关键词:** 棘孢木霉; *GH3* 基因家族; 进化分析; 纤维素降解

资助项目: 国家自然科学基金(31601679)

This work was supported by the National Natural Science Foundation of China (31601679).

\*Corresponding author. E-mail: 1wangyun@yangtzeu.edu.cn

Received: 2022-05-12; Accepted: 2022-06-29; Published online: 2022-07-22

## Bioinformatics and expression analyses of glycoside hydrolase 3 genes in *Trichoderma asperellum*

LI Cong, WANG Yun\*

College of Life Sciences, Yangtze University, Jingzhou 434025, Hubei, China

**Abstract:** [Background] The glycoside hydrolase 3 gene (*GH3*) family mainly encodes extracellular  $\beta$ -glucosidases, which are the key enzymes in cellulose degradation. [Objective] To identify the *GH3* genes in *Trichoderma asperellum* (*TaGH3* genes) and investigate their transcriptional patterns. [Methods] The bioinformatics tools were employed to identify the *TaGH3* genes and analyze the gene structure, phylogenetic relationship, as well as the physicochemical properties, subcellular localization, and tertiary structure of the deduced proteins. Quantitative real-time polymerase chain reaction (RT-qPCR) was employed to investigate the transcriptional patterns of *TaGH3* genes in the presence of cellulose. [Results] A total of 16 *TaGH3* genes were identified, which contained 1–8 exons. The deduced *TaGH3* proteins had the length of 533–934 amino acid (aa) and the theoretical molecular weights of 57.82–101.91 kDa. The majority of *TaGH3* proteins were predicted to be extracellular. Phylogenetic analysis demonstrated that all the *TaGH3* proteins could be classified into 4 clades, which had higher homology with *Trichoderma reesei*. Cellulose affected the transcription of all the 16 *TaGH3* genes, while the transcription level varied among genes. Specifically, cellulose induced the constitutive expression of 1 gene, down-regulated the expression of 2 genes, and up-regulated the expression of 13 genes. Moreover, cellulose enhanced the extracellular  $\beta$ -glucosidase activity, which was consistent to the transcriptional pattern of most *TaGH3* genes. [Conclusion] The genome of *T. asperellum* carried 16 *GH3* genes, most of which encoded hydrophilic proteins with the potential for commercial exploitation. The transcription of the majority of *TaGH3* genes was up-regulated by cellulose, which was consistent with the enhanced  $\beta$ -glucosidase activity, indicating that these genes played a key role in cellulose degradation. This study systematically investigated the *GH3* genes in *T. asperellum*, providing new insights for the resource utilization of lignocellulosic materials and the construction of strains with high cellulase production.

**Keywords:** *Trichoderma asperellum*; *GH3* gene family; evolutionary analysis; cellulose degradation

植物中的木质纤维素是陆地生态系统中含量最丰富的光合固碳资源，主要组分包括40%–60%纤维素、20%–40%半纤维素、10%–25%木质素和少量的果胶<sup>[1]</sup>。植物来源木质纤维素资源可替代石油等化石原料所生产的燃料和化学用品，是缓解资源、能源、环境危机和促进人类社会可持续发展的重要途径。目前，木质纤维素材料主要依靠细菌和真菌分泌

的纤维素降解酶转化为可发酵的糖类物质，但纤维素酶较低的水解效率和较高的生产成本极大地限制了生物质能源的转化效率<sup>[2]</sup>。细菌所生产的纤维素酶多为胞内酶且酶活较低，工业纯化难度大，因此，丝状真菌所生产的纤维素酶广泛应用于工业生产<sup>[3-4]</sup>。

纤维素降解酶是一组包含多种水解酶的复合酶系，主要包括内切葡聚糖酶、外切葡聚糖

酶和  $\beta$ -葡萄糖苷酶; 其中, 内切葡聚糖酶通过作用于纤维素非结晶区, 随机切割糖苷键, 将纤维素长链降解为小分子纤维素或寡糖链; 外切葡聚糖酶又称作纤维二糖水解酶, 与内切葡聚糖酶协同作用并负责降解纤维素的结晶区, 以渐进的方式切割纤维素链的还原端和非还原端释放纤维二糖;  $\beta$ -葡萄糖苷酶可将纤维素内切酶和外切酶作用释放的纤维寡糖和纤维二糖水解为葡萄糖, 在纤维素的彻底降解过程中发挥重要的作用; 这 3 种酶通过协同作用, 对纤维素进行降解<sup>[5]</sup>。

在纤维素降解过程中, 纤维二糖不能被很多发酵微生物直接利用<sup>[6]</sup>, 而且还会成为非竞争性抑制物, 反馈抑制纤维素酶的催化活性<sup>[7]</sup>。 $\beta$ -葡萄糖苷酶能够将纤维寡糖水解为葡萄糖, 提高糖平台中可发酵糖的得率, 并可通过降低反馈抑制促进纤维素的酶解<sup>[6,8]</sup>。另外,  $\beta$ -葡萄糖苷酶还可以直接用于天然多糖的降解, 也可作为风味酶用于果酒增香或脱苦、水解大豆异黄酮、生产低聚龙胆糖、分解乳糖、生产天然色素等领域<sup>[9]</sup>。所以,  $\beta$ -葡萄糖苷酶具有重要的理论和实用研究价值。编码丝状真菌中  $\beta$ -葡萄糖苷酶的主要包括 *GHI* 和 *GH3* 基因家族, 但 *GHI* 基因家族编码的  $\beta$ -葡萄糖苷酶主要为胞内糖苷酶, 而 *GH3* 基因家族主要编码胞外  $\beta$ -葡萄糖苷酶, 因此 *GH3* 基因家族编码的蛋白更利于满足工业需求<sup>[10-11]</sup>。

棘孢木霉(*Trichoderma asperellum*)是我国 2005 年新记录的可有效抑制植物病害的木霉菌<sup>[12]</sup>。研究表明该菌种可表达分泌多种纤维素酶, 具有良好的木质纤维素降解潜力<sup>[13-14]</sup>。同时, 研究表明棘孢木霉也是一种生物防治菌种, 对多种植物病原菌有着良好的拮抗作用<sup>[15]</sup>。在此期间, *GH3* 基因家族编码的  $\beta$ -葡萄糖苷酶发挥着重要作用, 但目前对于棘孢木霉 *GH3* 基因家族成员的研究较少。

里氏木霉、嗜热毁丝菌、黑曲霉和粗糙脉孢菌是应用范围较广的纤维素降解菌种, 通常用于木质纤维素降解酶的生产 and 研究<sup>[16]</sup>。本研究利用生物信息学方法在基因组范围内对棘孢木霉和这 4 种真菌的 *GH3* 基因家族成员进行筛选鉴定, 系统分析了其进化关系; 并对棘孢木霉的 *GH3* 基因家族成员进行了进一步研究, 探讨了其家族成员的理化性质及基因结构, 对纤维素降解过程中 *GH3* 基因家族在转录水平的表达模式进行分析, 以期为推动棘孢木霉在生物能源领域的应用奠定研究基础。

## 1 材料与方法

### 1.1 主要试剂、仪器及培养基

真菌 RNA 提取试剂盒, Omega Bio-Tek 公司; 反转录试剂盒, 南京诺唯赞生物科技股份有限公司; 2 $\times$ TransStart Green qPCR SuperMix, 北京全式金生物科技股份有限公司; 4-硝基苯基- $\beta$ -D-葡萄糖苷, 上海麦克林生化科技有限公司。反转录 PCR 仪和 NanoDrop 8000, Thermo Fisher Scientific 公司; 荧光定量 PCR 仪, Bio-Rad 公司; 多功能酶标仪, PerkinElmer 公司。PDB 培养基, 北京酷来博科技有限公司; 2%纤维素培养基<sup>[17]</sup>(g/L): 微晶纤维素 20.00, 酵母浸提物 2.00, 蛋白胨 2.00,  $K_2HPO_4$  1.00,  $MgSO_4 \cdot 7H_2O$  0.50,  $KH_2PO_4$  0.46。

### 1.2 *GH3* 基因家族成员鉴定及系统发育进化分析

从 NCBI 数据库(<https://www.ncbi.nlm.nih.gov/>)下载棘孢木霉(*T. asperellum*) CBS 433.97、嗜热毁丝菌(*Thermothelomyces thermophilus*) ATCC 42464、黑曲霉(*Aspergillus niger*) CBS 554.65、粗糙脉孢菌(*Neurospora crassa*) OR74A 和里氏木霉(*Trichoderma reesei*) QM6a 基因组数据, 它们的基因组登录号分别为 PRJNA207877、

PRJNA79339、PRJNA715116、PRJNA132 和 PRJNA225530。利用 Pfam 数据库(<http://pfam.xfam.org/>)下载的 GH3 蛋白保守结构域的隐马尔科夫模型(PF00933)，对 5 种真菌的 GH3 基因家族成员进行筛选鉴定。利用在线数据库 SMART ([http://smart.embl-heidelberg.de/smart/set\\_mode.cgi?NORMAL=1](http://smart.embl-heidelberg.de/smart/set_mode.cgi?NORMAL=1))和 NCBI-CDD (<https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi?>)对筛选的 GH3 基因家族成员进行二次鉴定。利用 MUSCLE 软件对 5 种真菌的 GH3 基因家族成员进行多序列比对，然后通过在线软件 iTOL (<https://itol.embl.de/>)绘制系统发育树。

### 1.3 生物信息学分析

棘孢木霉 GH3 家族成员的基因结构通过 GSDS2.0 (<http://gsds.gao-lab.org/>)进行分析<sup>[18]</sup>。GH3 蛋白的保守基序通过 MEME 软件(<https://meme-suite.org/meme/tools/meme>)进行分析，之后通过 TBtools 进行可视化<sup>[19-20]</sup>。利用在线软件 ExPASy (<https://web.expasy.org/protparam/>)对 GH3 家族的蛋白序列进行理化性质分析；利用 SWISS-MODEL 软件(<https://swissmodel.expasy.org/>)进行同源建模，分析 GH3 蛋白序列三维结构特性，预测蛋白质的空间模型。通过在线数据库 WOLF SPORT (<https://wolfpsort.hgc.jp/>)、CELLO (<http://cello.life.nctu.edu.tw/>)、YLoc (<https://abi-services.informatik.uni-tuebingen.de/yloc/webloc.cgi>)、Euk-mPLoc (<http://www.csbio.sjtu.edu.cn/bioinf/euk-multi/>)和 Predotar (<https://urgi.versailles.inra.fr/Tools/Predotar>)对 GH3 家族成员进行亚细胞定位预测分析<sup>[21]</sup>。

### 1.4 GH3 基因家族成员在转录水平的表达模式分析

棘孢木霉(*T. asperellum*) T4 菌株为本实验室保藏菌种，分离于黑龙江省密山市裴德镇土壤中，同时保藏于中国普通微生物菌种中心，保藏号为 CGMCC3.14975。菌株接种于 PDB 培

培养基，于 28 °C、200 r/min 摇床培养 2 d 后进行抽滤，称取等量湿重的菌丝体重悬于 150 mL 的 2%纤维素培养基和 PDB 培养基中，28 °C、180 r/min 振荡培养<sup>[17]</sup>。接种后于 1、2、3、5 和 7 d 收集培养的菌丝体进行 RNA 提取，以 PDB 培养基培养的菌丝体为对照。

使用实时荧光定量 PCR (quantitative real-time polymerase chain reaction, RT-qPCR)分析 GH3 基因家族成员在纤维素降解过程中的转录情况。利用真菌 RNA 提取试剂盒提取棘孢木霉 RNA，并通过反转录试剂盒获得 cDNA。利用 1%的琼脂糖凝胶电泳检测 RNA 质量，用 NanoDrop 8000 对 RNA 和 cDNA 进行定量。利用 NCBI 网站设计特异性引物(表 1)，以 *actin* 为内参基因，采用荧光定量 PCR 仪进行实验。RT-qPCR 反应体系(20  $\mu$ L)：2 $\times$ TransStart Green qPCR SuperMix 10  $\mu$ L，cDNA 模板(20 ng/ $\mu$ L) 5  $\mu$ L，正、反引物(10  $\mu$ mol/L) 0.4  $\mu$ L，ddH<sub>2</sub>O 4.2  $\mu$ L。反应条件：95 °C 3min；95 °C 10 s，60 °C 30s，45 个循环。绘制熔解曲线，采用 2<sup>- $\Delta\Delta$ C<sub>T</sub></sup> 法进行数据分析，每个样品设置 3 次生物学重复<sup>[22-23]</sup>。

### 1.5 $\beta$ -葡萄糖苷酶活力测定

收集菌株在 2%纤维素培养基及 PDB 培养基中培养 1、2、3、5 和 7 d 的上清液，置于 4 °C、8 000 r/min 离心 10 min，取上清进行  $\beta$ -葡萄糖苷酶活力测定，以 1 mg/mL 的 4-硝基苯基- $\beta$ -D-葡萄糖苷(pNPG)溶液为反应底物<sup>[11]</sup>。吸取 100  $\mu$ L 酶液与 50  $\mu$ L 底物进行混匀，50 °C 反应 1 h 后加入 150  $\mu$ L 的 10% Na<sub>2</sub>CO<sub>3</sub> 终止反应，多功能酶标仪中测定 420 nm 处的吸光度值<sup>[24-25]</sup>。

## 2 结果与分析

### 2.1 棘孢木霉 GH3 基因家族成员的生物信息学鉴定及其编码蛋白的理化性质分析

依据棘孢木霉基因组数据检索其 GH3 基因

家族成员, 去除冗余序列后, 经过 SMART 和 NCBI-CDD 在线数据库比对校对, 共鉴定得到 16 个棘孢木霉 GH3 家族的蛋白(表 2)。棘孢木霉的基因组序列被组装在 419 个 scaffolds

上, 其中棘孢木霉的 GH3 基因家族成员分布在 1-25 号 scaffolds 上, 依据其在 scaffolds 上的相对位置, 16 个 GH3 家族成员被命名为 TaGH3-1-TaGH3-16。

表 1 RT-qPCR 所用引物

Table 1 Primers for RT-qPCR

引物名称 Primers	引物序列 Primer sequence (5'→3')	产物大小 Size (bp)	引物名称 Primers	引物序列 Primer sequence (5'→3')	产物大小 Size (bp)
<i>TaActinF</i>	AGGCAACCTTCTCGCCAACG	256	<i>TaGH3-9F</i>	CGTCAATGCTGCTTCCGATG	174
<i>TaActinR</i>	TCGCTTCTCGACAATGCCAACT		<i>TaGH3-9R</i>	TTGTCAACCAGGCGGATAACC	
<i>TaGH3-1F</i>	TTGACAGCCTTGTCGAGACC	85	<i>TaGH3-10F</i>	GCCAAATACGCTTCGCTGAG	153
<i>TaGH3-1R</i>	TCCCAGTTCTCGTTGACAGC		<i>TaGH3-10R</i>	TGGGCCAATAACAGCGACAT	
<i>TaGH3-2F</i>	GCACGGCCATATTCAACGAC	168	<i>TaGH3-11F</i>	TGAAGGACAACCCGAAGACG	102
<i>TaGH3-2R</i>	TGTTACTGATCCGCCACC		<i>TaGH3-11R</i>	TCGTACAAGCCAGGGTTTCC	
<i>TaGH3-3F</i>	GCTGACCTTGCTCGAGAGTT	131	<i>TaGH3-12F</i>	ATCCAAGGAGCTAATCGGCG	141
<i>TaGH3-3R</i>	TTAACCGTCGCTGAGACACC		<i>TaGH3-12R</i>	GGACTTCCCCTCAGACAAGC	
<i>TaGH3-4F</i>	TCACCAAGACGGAGATGCAC	94	<i>TaGH3-13F</i>	AACGCGCTGGTCGATATTCT	135
<i>TaGH3-4R</i>	TCATCGGCGGTATAAGTGCC		<i>TaGH3-13R</i>	CGTAAACAGCCCCTCACTGT	
<i>TaGH3-5F</i>	CCGCACTGCTGGATACTCAT	119	<i>TaGH3-14F</i>	TCGAATGGGCAACCTCCTTC	156
<i>TaGH3-5R</i>	GCCCACCGAGTAAGGTATCG		<i>TaGH3-14R</i>	CATTGATGTTGGGCGCGTAG	
<i>TaGH3-6F</i>	TGTTATGGCTGTGGAGGCTG	105	<i>TaGH3-15F</i>	CGCAACATGGAATAGGCAGC	189
<i>TaGH3-6R</i>	ATGCCGTTCTCATAGCCAG		<i>TaGH3-15R</i>	ATGGTCTGAGAGCCAGCAAC	
<i>TaGH3-7F</i>	GATCCTGTGCTTCAGGCTGT	114	<i>TaGH3-16F</i>	AGCTAATCCGCAAACAGGCT	142
<i>TaGH3-7R</i>	CATGCGGTACATCTCCTGCT		<i>TaGH3-16R</i>	TTGTCTCCGTTGTAGGCACC	
<i>TaGH3-8F</i>	CAGACACCTTTGAGGAGCGT	113			
<i>TaGH3-8R</i>	AACATCGGCCTCCTTACCT				

表 2 棘孢木霉 GH3 蛋白的理化性质分析

Table 2 Physicochemical characters of GH3 proteins identified in *Trichoderma asperellum*

名称 Name	氨基酸 Amino acids	分子量 $M_w$ (kDa)	等电点 pI	不稳定系数 Instability index	脂肪系数 Aliphatic index	亲水性平均系数 GRAVY	基因组定位 Genomic location
TaGH3-1	884	94.72	5.54	29.83	75.61	-0.260	Scaffold1: 664 481-667 205
TaGH3-2	781	84.15	5.13	35.50	78.54	-0.208	Scaffold1: 3290 843-3 293 717
TaGH3-3	771	83.93	5.17	31.89	83.81	-0.134	Scaffold2: 2 900 558-2 903 190
TaGH3-4	836	91.1	5.59	26.32	78.65	-0.255	Scaffold3: 118 379-121 033
TaGH3-5	934	101.91	5.79	36.37	81.90	-0.229	Scaffold3: 383 661-386 880
TaGH3-6	932	102.03	6.64	38.64	98.58	-0.056	Scaffold3: 1 013 511-1 016 752
TaGH3-7	893	96.09	5.02	31.21	75.60	-0.248	Scaffold5: 1 218 063-1 221 133
TaGH3-8	533	57.82	5.53	42.51	91.93	-0.127	Scaffold6: 1 943 431-1 945 432
TaGH3-9	763	83.59	5.77	28.81	81.85	-0.281	Scaffold11: 202 562-205 201
TaGH3-10	872	92.62	4.77	31.01	80.94	-0.048	Scaffold11: 217 842-220 634
TaGH3-11	865	92.78	5.60	31.34	75.31	-0.311	Scaffold16: 107 054-109 883
TaGH3-12	830	90.52	5.85	34.45	87.57	-0.097	Scaffold16: 366 080-368 681
TaGH3-13	733	76.67	5.71	29.28	83.49	-0.040	Scaffold16: 402 753-405 123
TaGH3-14	795	87.03	5.44	35.26	80.92	-0.290	Scaffold16: 398 760-401 402
TaGH3-15	889	96.83	5.73	35.96	80.71	-0.338	Scaffold18: 107 287-109 956
TaGH3-16	814	88.57	5.12	33.73	80.98	-0.216	Scaffold25: 64 952-67 449

蛋白理化性质分析表明, 棘孢木霉 *GH3* 基因家族成员所编码的蛋白质序列长度为 533–934 个氨基酸, 分子量为 57.82–101.91 kDa, 其中 TaGH3-8 的氨基酸数目最少, 分子量仅为 57.82 kDa; 而 TaGH3-5 的氨基酸数目最多, 分子量高达 101.91 kDa。棘孢木霉的 *GH3* 蛋白均为酸性蛋白, 其中 TaGH3-6 蛋白的等电点高达 6.64; 而 TaGH3-10 蛋白的等电点最低, 仅为 4.77。棘孢木霉 *GH3* 蛋白的不稳定系数范围为 26.32–42.51, 其中绝大多数蛋白的不稳定系数低于 40, 表明该基因家族编码的蛋白质多属于稳定蛋白。但 TaGH3-8 的不稳定系数高达 42.51, 表明该蛋白属于不稳定蛋白质。棘孢木霉 *GH3* 蛋白的脂肪系数范围为 75.31–98.58。其中, TaGH3-11 的脂肪系数最低, 表明其热稳定性最差; 而 TaGH3-6 的脂肪系数最高, 表明该蛋白热稳定性较优。同时, 棘孢木霉 *GH3* 蛋白的亲水性平均系数均小于 0, 表明该物种全

部的 *GH3* 蛋白均为亲水性蛋白。另外, 利用 5 种不同数据库对棘孢木霉 *GH3* 基因家族成员的亚细胞定位进行预测(表 3), 除少部分蛋白定位于细胞质、细胞膜及膜性细胞器外, 其余绝大部分蛋白均为胞外蛋白。

## 2.2 棘孢木霉 *GH3* 基因家族的系统发育分析

在棘孢木霉、粗糙脉孢菌、黑曲霉菌、里氏木霉和嗜热毁丝菌 5 个物种中共鉴定到 66 个 *GH3* 家族的基因。其中, 棘孢木霉包含 16 个, 黑曲霉包含 17 个, 里氏木霉包含 13 个, 嗜热毁丝菌和粗糙脉孢菌各鉴定到 10 个, 所有成员依据染色体相对位置进行命名。将 66 个 *GH3* 基因家族成员进行多序列比对, 构建系统发育树(图 1)。结果表明, 5 个物种的 *GH3* 家族成员可聚为 4 个类群。第 IV 类群的 *GH3* 家族成员最多, 5 个物种的大部分 *GH3* 蛋白均聚类于此类群, 其中可分为 5 个亚群, 棘孢木霉和黑曲霉的

表 3 棘孢木霉 *GH3* 蛋白的亚细胞定位预测分析

Table 3 Predicted subcellular localization of *GH3* proteins in *Trichoderma asperellum*

基因名称 Gene name	Predotar	Euk-mPLoc	YLoc	WoLF PSORT	CELLO
<i>TaGH3-1</i>	胞外 Extra	胞外 Extra	胞外 Extra	胞外 Extra	胞外 Extra
<i>TaGH3-2</i>	–	细胞膜 Mem	胞外 Extra	胞外 Extra	胞外 Extra
<i>TaGH3-3</i>	–	过氧化物酶体 Pex	胞外 Extra	胞外 Extra	胞外 Extra
<i>TaGH3-4</i>	–	过氧化物酶体 Pex	细胞质 Cyto	细胞质 Cyto	细胞质 Cyto
<i>TaGH3-5</i>	线粒体 Mito	胞外 Extra	线粒体 Mito	胞外 Extra	胞外 Extra
<i>TaGH3-6</i>	–	细胞质 Cyto	细胞质 Cyto	细胞膜 Mem	细胞膜/细胞质 Mem/Cyto
<i>TaGH3-7</i>	–	胞外 Extra	胞外 Extra	胞外 Extra	胞外 Extra
<i>TaGH3-8</i>	–	–	线粒体 Mito	细胞质 Cyto	细胞质 Cyto
<i>TaGH3-9</i>	胞外 Extra	–	胞外 Extra	胞外 Extra	细胞质 Cyto
<i>TaGH3-10</i>	胞外 Extra	胞外 Extra	胞外 Extra	胞外 Extra	胞外 Extra
<i>TaGH3-11</i>	胞外 Extra	胞外 Extra	胞外 Extra	胞外 Extra	胞外 Extra
<i>TaGH3-12</i>	–	过氧化物酶体 Pex	细胞质 Cyto	细胞质 Cyto	–
<i>TaGH3-13</i>	胞外 Extra	胞外 Extra	胞外 Extra	胞外 Extra	胞外 Extra
<i>TaGH3-14</i>	胞外 Extra	胞外 Extra	胞外 Extra	胞外 Extra	胞外 Extra
<i>TaGH3-15</i>	线粒体 Mito	–	线粒体 Mito	胞外 Extra	细胞质 Cyto
<i>TaGH3-16</i>	–	细胞膜 Mem	胞外 Extra	胞外 Extra	胞外 Extra

–: 未预测到; Extra: 胞外; Mito: 线粒体; Mem: 细胞膜; Pex: 过氧化物酶体; Cyto: 细胞质

–: Unpredicted; Extra: Extracellular matrix; Mito: Mitochondrion; Mem: Membrane; Pex: Peroxisome; Cyto: Cytoplasm.

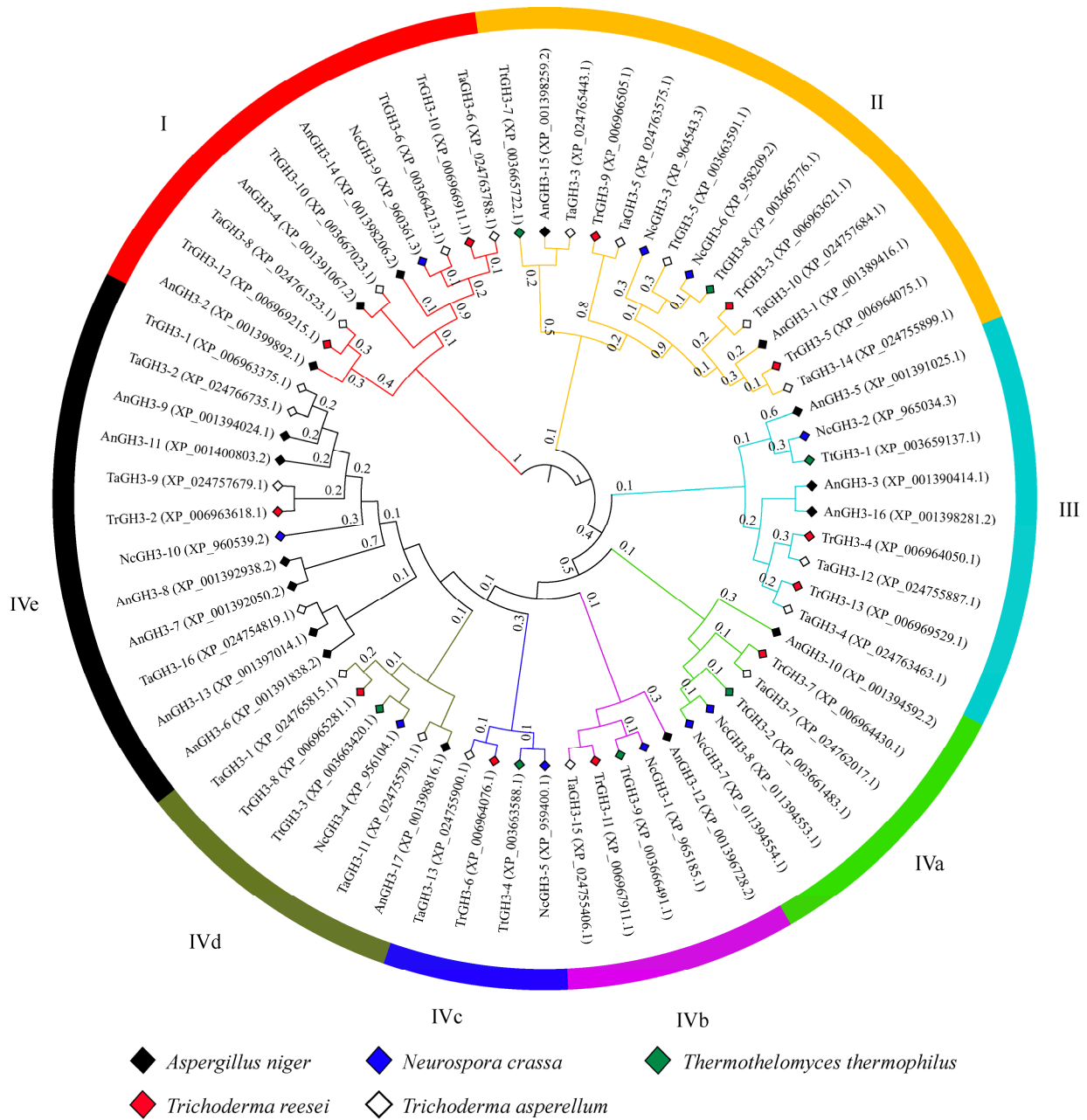


图 1 不同物种 GH3 蛋白的系统发育树 分支长度代表遗传距离；各分支上数字为自举法检验循环 1 000 次时的置信度

Figure 1 Phylogenetic tree of GH3 proteins in different species. Branch length represented the distance of evolution; The numbers for the branches of the tree were bootstrap confidence out of 1 000 replications.

GH3 蛋白多分布于 IVe 亚群。第 II 类群包含 14 个成员，为 GH3 蛋白成员第二多类群；其次为第 I 类群的 10 个成员和第 III 类群的 9 个成员，这 2 个类群中黑曲霉 GH3 家族成员居多。

### 2.3 棘孢木霉 GH3 基因家族成员的保守基序及基因结构分析

对棘孢木霉 GH3 家族成员进行保守基序分析，结果表明 GH3 家族成员包含 3-10 个保守

基序, 而且所有的 GH3 蛋白均含有保守基序 6 (图 2)。各亚类群之间的保守基序有所差异, 第 I 类群和第 II 类群的保守基序数量较少, 大部分成员含有保守基序 2、3 和 5。第 III 类群和第 IV 类群的保守基序数量较多, 所有的成员均含有保守基序 1、2、3、5、8 和 9。基因结构分析表明, 除 TaGH3-14 和 TaGH3-15 外, 其余成员均含有 1-7 个内含子和 2-8 个外显子 (图 3A)。利用 SWISS-MODEL 对 TaGH3 蛋白进行同源建模, 结果显示 16 个 GH3 家族成员同样可分为 4 个构型 (图 3B)。其中, TaGH3-1、TaGH3-8、TaGH3-10、TaGH3-13 和 TaGH3-14 具有相似的三角形蛋白构型, TaGH3-3、TaGH3-5 和 TaGH3-6 的蛋白构型形似四边形, TaGH3-4 和 TaGH3-12 的三级结构近似圆形, 其余家族成员具有相似的哑铃型三级结构。

## 2.4 棘孢木霉 GH3 基因家族成员在转录水平的表达模式及 $\beta$ -葡萄糖苷酶活力分析

在 2% 纤维素培养基中培养 1-7 d 后, *TaGH3* 家族的基因在转录水平上的表达情况如图 4 所示。结果表明, 接种于 2% 纤维素培养基后, *TaGH3-9* 的转录量波动不大, *TaGH3-4* 和 *TaGH3-6* 转录下调; 其余所有基因均表现为转录上调, 并在第 3 天和第 5 天达到峰值, 第 7 天后的转录量有所下降。其中, *TaGH3-3*、*TaGH3-7*、*TaGH3-10*、*TaGH3-11*、*TaGH3-13* 和 *TaGH3-14* 的转录量明显高于其他基因, 而且 *TaGH3-13* 的转录量最高。*TaGH3-10* 的转录量在第 5 天达到峰值, 其余 5 个基因均是在第 3 天达到峰值。

将棘孢木霉在 PDB 培养基和 2% 纤维素培养基中培养 7 d, 对其胞外  $\beta$ -葡萄糖苷酶的活性进行测定, 结果如图 5 所示。总而言之, 胞

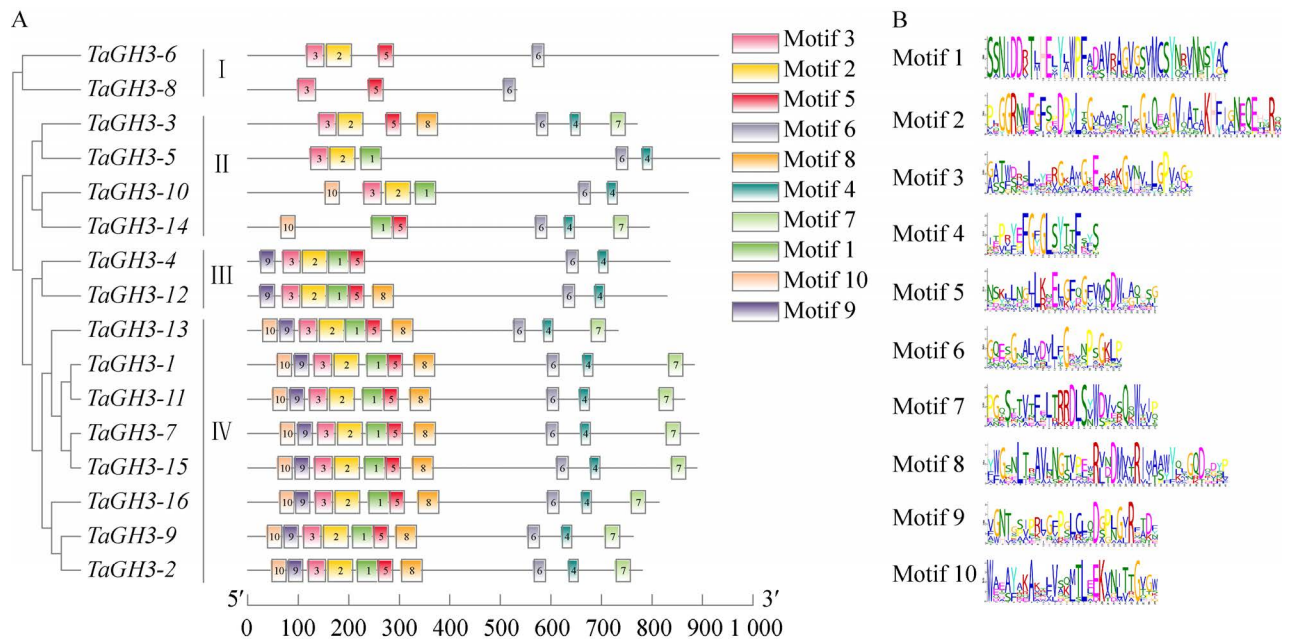


图 2 TaGH3 蛋白的保守基序分析 A: TaGH3 基因家族蛋白的基序分布. B: 10 种保守基序的氨基酸组成

Figure 2 The analysis of conserved motifs of TaGH3 proteins. A: The motif pattern of TaGH3 family. B: The detailed information of ten motifs.



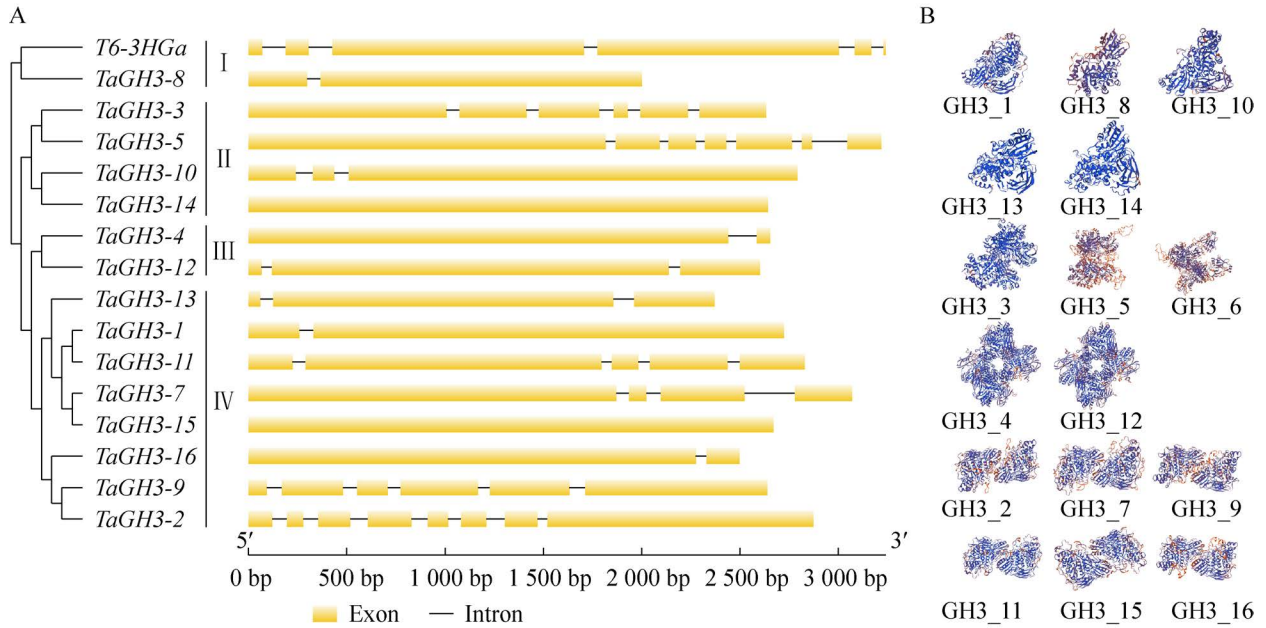


图 3 *TaGH3* 家族的基因结构分析及蛋白质三级结构预测 A: *TaGH3* 家族的基因结构. B: 预测的蛋白质三维结构

Figure 3 The gene structure and tertiary structure prediction of *TaGH3* gene family. A: The gene structure of *TaGH3* genes. B: The predicted tertiary structure of *TaGH3* proteins.

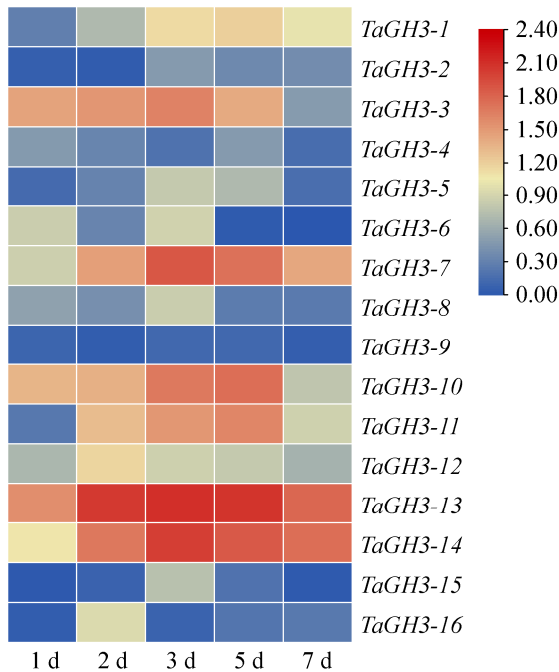


图 4 *TaGH3* 基因家族在 2%纤维素培养基中的转录模式

Figure 4 Transcription patterns of *TaGH3* genes family on 2% cellulose medium.

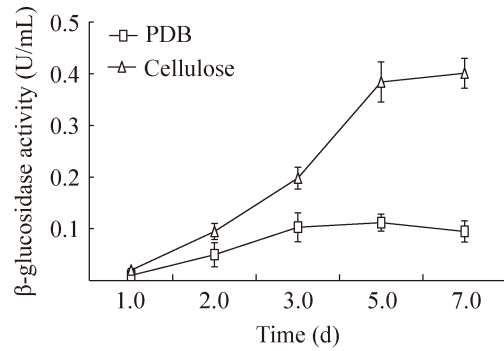


图 5 棘孢木霉的胞外  $\beta$ -葡萄糖苷酶活力分析

Figure 5 The analysis of extracellular  $\beta$ -glucosidase activity in *Trichoderma asperellum*.

外  $\beta$ -葡萄糖苷酶活力随培养时间延长呈上升趋势, 而且棘孢木霉在 2%纤维素培养基中具有更高的  $\beta$ -葡萄糖苷酶活力。在培养的前 3 天, 2 种培养基中的胞外  $\beta$ -葡萄糖苷酶活力差距较小, 但其酶活差距在第 5 天和第 7 天差距较大, 其酶活变化趋势与 *TaGH3* 基因的转录模式基本一致。

### 3 讨论与结论

木质纤维素类生物质是丰富的可再生资源,可在纤维素酶的作用下降解为单体糖继而用以制取乙醇等生物燃料<sup>[26]</sup>。纤维素酶在能源、食品、纺织、洗涤剂、造纸等众多领域也有着十分广阔的应用前景<sup>[1,14]</sup>。木霉菌是研究纤维素降解机制的模式菌株,在纤维素类物质降解生产中具有重要的地位。木霉中 *GH3* 基因家族成员主要编码胞外  $\beta$ -葡萄糖苷酶,不仅在纤维素的彻底降解过程中发挥重要作用,同样可以用于食品风味改良、生产活性单体成分等领域<sup>[9]</sup>。

本研究通过比较基因组学分析,鉴定到黑曲霉基因组中的 *GH3* 基因家族成员 17 个,里氏木霉基因组中 *GH3* 基因家族成员 13 个,粗糙脉孢菌和嗜热毁丝菌基因组中 *GH3* 基因家族成员各 10 个。研究表明,同源基因数目的变化与其生物学功能分化密切相关<sup>[27-28]</sup>。棘孢木霉基因组中包含 16 个 *GH3* 基因家族成员,其基因家族的拷贝数发生显著扩张,可能与其纤维素降解能力密切相关。对 5 个物种的 *GH3* 基因家族成员进行系统发育分析,结果表明,不同类群中 5 个物种的 *GH3* 基因家族成员均有分布。系统发育树中亲缘关系越近,其具有相似蛋白功能的可能性也越大。这 5 种真菌均常用于纤维素酶的工业生产和科学研究,相似的功能分化可能与其较为相近的亲缘关系有所关联<sup>[1,14]</sup>。在各类群中,棘孢木霉的 *GH3* 基因家族成员与里氏木霉的亲缘关系更近,这可能与它们同为木霉属菌种有关。但棘孢木霉 *GH3* 基因家族成员与嗜热毁丝菌的亲缘关系较远,这可能与 2 个菌种不同的培养方式密切相关。Tambor 等研究表明,嗜热毁丝菌具有较高的耐热性,可在 pH 值较高的环境中稳定表达多种纤维素酶,其分泌的纤维素降解酶系在

70 °C 高温仍具有较高活性<sup>[29]</sup>。

研究发现,棘孢木霉的 *GH3* 家族成员大部分为胞外蛋白质,同时,棘孢木霉 *GH3* 家族蛋白的亲水性系数均小于 0,表明该物种全部的 *GH3* 家族成员均为亲水性蛋白。根据“相似相溶”的原则,在进行蛋白提取时可采用偏碱性的稀盐和缓冲系统水溶液来提取。另外,前人研究表明,木质纤维素中的木质素会降低纤维素酶对纤维素的可及性,因此酶解之前常需要进行适当的预处理<sup>[30]</sup>。由于棘孢木霉 *GH3* 家族成员均为酸性蛋白质,表明利用棘孢木霉对木质纤维素降解之前可采用酸预处理法<sup>[31]</sup>。

基因表达模式是基因功能的重要体现,是筛选关键基因的重要途径<sup>[32]</sup>。接种于 2% 纤维素培养基后,棘孢木霉的 16 个 *GH3* 基因家族成员在转录水平上的表达均有所上调,但基因间的表达模式有较大差异。总体而言,在纤维素的诱导下,棘孢木霉大部分的 *GH3* 基因家族成员的转录量在第 5 天达到峰值,而胞外  $\beta$ -葡萄糖苷酶活力却在第 5、7 天均保持较高的活性,这可能与基因表达调控的滞后性密切相关<sup>[33]</sup>。纤维素诱导下, *TaGH3-13* 的转录量最高,其编码的蛋白质稳定性和耐热性较好,而且在 5 个数据库注释下均定位于细胞外基质,利于工业分离和纯化,具有良好的研究价值和商业开发潜力。

综上所述,本研究利用生物信息学方法对 *GH3* 基因家族成员进行鉴定、分析,并对纤维素降解过程中棘孢木霉 *GH3* 基因家族成员在转录水平的表达模式进行分析,为进一步研究 *GH3* 家族蛋白的生物学功能提供了科学依据和理论支撑。

### 致谢

感谢武汉大学侯丽秀博士对本论文的建议和修改。

## REFERENCES

- [1] JØRGENSEN H, VIBE-PEDERSEN J, LARSEN J, FELBY C. Liquefaction of lignocellulose at high-solids concentrations[J]. *Biotechnology and Bioengineering*, 2007, 96(5): 862-870.
- [2] LI DP, FENG L, LIU KR, CHENG Y, HOU N, LI CY. Optimization of cold-active CMCase production by psychrotrophic *Sphingomonas* sp. FLX-7 from the cold region of China[J]. *Cellulose*, 2016, 23(2): 1335-1347.
- [3] 王晓涛, 魏佩玲, 胡波, 宫平. 纤维素降解酶研究进展[J]. 草食家畜, 2019(3): 13-18.  
WANG XT, WEI PL, HU B, GONG P. Research progress on cellulose degrading enzymes[J]. *Grass-Feeding Livestock*, 2019(3): 13-18 (in Chinese).
- [4] 曲音波. 木质纤维素降解酶系的基础和技术研究进展[J]. 山东大学学报(理学版), 2011, 46(10): 160-170.  
QU YB. Progress in basic and technological research of enzyme system for lignocellulosics biodegradation[J]. *Journal of Shandong University (Natural Science)*, 2011, 46(10): 160-170 (in Chinese).
- [5] GUSAKOV AV. Alternatives to *Trichoderma reesei* in biofuel production[J]. *Trends in Biotechnology*, 2011, 29(9): 419-425.
- [6] 刘国栋, 高丽伟, 曲音波. 青霉生产木质纤维素降解酶系的研究进展[J]. 生物工程学报, 2021, 37(3): 1058-1069.  
LIU GD, GAO LW, QU YB. Progress in the production of lignocellulolytic enzyme systems using *Penicillium* species[J]. *Chinese Journal of Biotechnology*, 2021, 37(3): 1058-1069 (in Chinese).
- [7] van DYK JS, PLETSCHKE BI. A review of lignocellulose bioconversion using enzymatic hydrolysis and synergistic cooperation between enzymes—Factors affecting enzymes, conversion and synergy[J]. *Biotechnology Advances*, 2012, 30(6): 1458-1480.
- [8] SINGHANIA RR, PATEL AK, PANDEY A, GANANSOUNOU E. Genetic modification: a tool for enhancing beta-glucosidase production for biofuel application[J]. *Bioresource Technology*, 2017, 245: 1352-1361.
- [9] 周庆新, 戴炳业, 陈蕾蕾, 刘孝永, 裘纪莹, 陈相艳. 瑞氏木霉中  $\beta$ -葡萄糖苷酶基因功能研究进展[J]. 中国农业科技导报, 2014, 16(2): 74-78.  
ZHOU QX, DAI BY, CHEN LL, LIU XY, QIU JY, CHEN XY. Progress on Function Studies of  $\beta$ -glucosidase gene in *Trichoderma reesei*[J]. *Journal of Agricultural Science and Technology*, 2014, 16(2): 74-78 (in Chinese).
- [10] EZEILO UR, LEE CT, HUYOP F, ZAKARIA II, WAHAB RA. Raw oil palm frond leaves as cost-effective substrate for cellulase and xylanase productions by *Trichoderma asperellum* UC1 under solid-state fermentation[J]. *Journal of Environmental Management*, 2019, 243: 206-217.
- [11] 周林芳, 江波, 张涛, 李淑华. 糖苷水解酶第 3 家族  $\beta$ -葡萄糖苷酶的研究进展[J]. 食品工业科技, 2017, 38(14): 330-335.  
ZHOU LF, JIANG B, ZHANG T, LI SH. Research progress of  $\beta$ -glucosidases of glycoside hydrolase family 3[J]. *Science and Technology of Food Industry*, 2017, 38(14): 330-335 (in Chinese).
- [12] 章初龙, 徐同. 我国河北、浙江、云南及西藏木霉种记述[J]. 菌物学报, 2005, 24(2): 184-192.  
ZHANG CL, XU T. Records of trichoderma species from Hebei, Zhejiang, Yunan and Tibet of China[J]. *Mycosystema*, 2005, 24(2): 184-192 (in Chinese).
- [13] 赵鹏. 棘孢木霉 T4 株 *task1* 基因过表达及功能的初步分析[D]. 哈尔滨: 哈尔滨工业大学硕士学位论文, 2014.  
ZHAO P. Preliminary analysis of Task1 gene from *Trichoderma asperellum* T4 over expression and function[D]. Harbin: Master's Thesis of Harbin Institute of Technology, 2014 (in Chinese).
- [14] SHANMUGAM S, KRISHNASWAMY S, CHANDRABABU R, VEERABAGU U, PUGAZHENDHI A, MATHIMANI T. Optimal immobilization of *Trichoderma asperellum* laccase on polymer coated  $\text{Fe}_3\text{O}_4/\text{SiO}_2$  nanoparticles for enhanced biohydrogen production from delignified lignocellulosic biomass[J]. *Fuel*, 2020, 273: 117777.
- [15] SINGH G, TIWARI A, GUPTA A, KUMAR A, HARIPRASAD P, SHARMA S. Bioformulation development via valorizing silica-rich spent mushroom substrate with *Trichoderma asperellum* for plant nutrient and disease management[J]. *Journal of Environmental Management*, 2021, 297: 113278.
- [16] 高丽伟. 草酸青霉木质纤维素降解酶转录激活机制解析及高产菌株构建[D]. 济南: 山东大学博士学位论文, 2018.  
GAO LW. Studies on the mechanisms for the transcriptional activation of lignocellulolytic enzyme genes in *Penicillium oxalicum* and the construction of enzyme high-producing strains[D]. Jinan: Doctoral Dissertation of Shandong University, 2018 (in Chinese).
- [17] 蔡英丽. 香菇木质纤维素降解酶在不同碳源上的表达模式分析[D]. 武汉: 华中农业大学博士学位论文,

- 2017.
- CAI YL. Expression profiling analysis of lignocellulose degrading enzymes in *Lentinula edodes* grown on different carbon sources[D]. Wuhan: Doctoral Dissertation of Huazhong Agricultural University, 2017 (in Chinese).
- [18] HU B, JIN JP, GUO AY, ZHANG H, LUO JC, GAO G. GSDS 2.0: an upgraded gene feature visualization server[J]. *Bioinformatics*, 2015, 31(8): 1296-1297.
- [19] CHEN CJ, CHEN H, ZHANG Y, THOMAS HR, FRANK MH, HE YH, XIA R. TBtools: an integrative toolkit developed for interactive analyses of big biological data[J]. *Molecular Plant*, 2020, 13(8): 1194-1202.
- [20] BAILEY TL, BODEN M, BUSKE FA, FRITH M, GRANT CE, CLEMENTI L, REN JY, LI WW, NOBLE WS. MEME Suite: tools for motif discovery and searching[J]. *Nucleic Acids Research*, 2009, 37(suppl\_2): W202-W208.
- [21] YU CS, LIN CJ, HWANG JK. Predicting subcellular localization of proteins for Gram-negative bacteria by support vector machines based on n-peptide compositions[J]. *Protein Science: a Publication of the Protein Society*, 2004, 13(5): 1402-1406.
- [22] 唐家全, 郝大志, 李婷婷, 张慧, 陈捷. 棘孢木霉菌对钠胁迫的生理响应机制[J]. *微生物学通报*, 2021, 48(1): 23-34.
- TANG JQ, HAO DZ, LI TT, ZHANG H, CHEN J. Physiological response of *Trichoderma asperellum* to Na<sup>+</sup> stress[J]. *Microbiology China*, 2021, 48(1): 23-34 (in Chinese).
- [23] SCHMITTGEN TD, LIVAK KJ. Analyzing real-time PCR data by the comparative CT method[J]. *Nature Protocols*, 2008, 3(6): 1101-1108.
- [24] 姚卫蓉, 丁霄霖. pNPG 法测定纤维素酶系中  $\beta$ -葡萄糖苷酶[J]. *微生物学通报*, 1998, 25(3): 182-183.
- YAO WR, DING XL. Determination of  $\beta$ -glucosidase in cellulase system by pNPG method[J]. *Microbiology China*, 1998, 25(3): 182-183 (in Chinese).
- [25] D'INCECCO N, BARTOWSKY E, KASSARA S, LANTE AN, SPETTOLI P, HENSCHKE P. Release of glycosidically bound flavour compounds of Chardonnay by *Oenococcus oeni* during malolactic fermentation[J]. *Food Microbiology*, 2004, 21(3): 257-265.
- [26] ZHAO XL, ZHENG ZH, CAI YF, ZHAO YB, ZHANG Y, GAO YH, CUI ZJ, WANG XF. Accelerated biomethane production from lignocellulosic biomass: pretreated by mixed enzymes secreted by *Trichoderma viride* and *Aspergillus* sp.[J]. *Bioresource Technology*, 2020, 309: 123378.
- [27] 沈海默, 郑琪, 胡薇. 日本血吸虫组蛋白基因家族分子进化分析[J]. *国际医学寄生虫病杂志*, 2015, 42(2): 67-73.
- SHEN HM, ZHENG Q, HU W. Molecular evolution of the histone gene family in *Schistosoma japonicum*[J]. *International Journal of Medical Parasitic Diseases*, 2015, 42(2): 67-73 (in Chinese).
- [28] 刘聪, 杜国英, 唐磊, 高天, 唐祥海, 莫照兰, 茅云翔. 紫菜腐霉激发子基因家族特征及其在感染过程中的作用[J]. *微生物学通报*, 2022, 49(1): 139-152.
- LIU C, DU GY, TANG L, GAO T, TANG XH, MO ZL, MAO YX. Characteristics of elicitor gene family of *Pythium porphyrae* and its role in infection[J]. *Microbiology China*, 2022, 49(1): 139-152 (in Chinese).
- [29] TAMBOR JH, REN HN, USHINSKY S, ZHENG Y, RIEMENS A, ST-FRANCOIS C, TSANG A, POWLOWSKI J, STORMS R. Recombinant expression, activity screening and functional characterization identifies three novel endo-1,4- $\beta$ -glucanases that efficiently hydrolyse cellulosic substrates[J]. *Applied Microbiology and Biotechnology*, 2012, 93(1): 203-214.
- [30] MOONEY CA, MANSFIELD SD, TOUHY MG, SADDLER JN. The effect of initial pore volume and lignin content on the enzymatic hydrolysis of softwoods[J]. *Bioresource Technology*, 1998, 64(2): 113-119.
- [31] BORGES DG, BARALDO A Jr, FARINAS CS, de LIMA CAMARGO GIORDANO R, TARDIOLI PW. Enhanced saccharification of sugarcane bagasse using soluble cellulase supplemented with immobilized  $\beta$ -glucosidase[J]. *Bioresource Technology*, 2014, 167: 206-213.
- [32] 李培谦, 冯宝珍. 尖孢镰孢菌果胶裂解酶基因家族鉴定及侵染表达模式分析[J]. *微生物学通报*, 2021, 48(8): 2774-2783.
- LI PQ, FENG BZ. Characterization of pectate lyase gene family in *Fusarium oxysporum* f. sp. *lycopersici* genome and expression mode during inoculation[J]. *Microbiology China*, 2021, 48(8): 2774-2783 (in Chinese).
- [33] MAEDA YT, SANO M. Regulatory dynamics of synthetic gene networks with positive feedback[J]. *Journal of Molecular Biology*, 2006, 359(4): 1107-1124.