扩展青霉 G 蛋白偶联受体预测及生物信息学分析

宋 聪[1]，宋文明[1,3]，孙 阳[1,3]，贾振华[1]，宋水山[1*]，关军锋[2*]

（1. 河北省科学院 生物研究所，河北 石家庄 050081；
2. 河北省农林科学院 生物技术与食品科学研究所，河北 石家庄 050051；3. 河北民族师范学院，河北 承德 067000）

摘 要 病原真菌的 G 蛋白偶联受体（G protein-coupled receptors，GPCRs）是病害防治的关键靶点之一。通过多种生物信息学方法，综合运用 DeePTMHMM、Phobius 和 TOPCONS 三种跨膜预测工具，以及 GPCRHMM、GPCRPipe、GPCR-Pen 和 GPCRPred 四个 GPCR 预测平台，从扩展青霉（*Penicillium expansum*）基因组序列中筛选了 36 个候选 GPCRs，通过拓扑结构、亚细胞定位和功能分析验证了该预测方法的可靠性，经系统进化树分类、保守结构域、理化性质等特征分析，确定这些候选 GPCRs 分布在 6 个真菌 GPCRs 类别中，具有真菌视紫红质受体样家族、真菌信息素交配因子 STE2、信息素 A 受体 STE3、RTA1 样蛋白等保守结构域，候选 GPCRs 多为中性或碱性蛋白，其中不稳定性蛋白占 44.44%，疏水性蛋白占 97.22%。本研究结果为进一步研究扩展青霉 GPCRs 的生物学功能、信号转导机制以及致病机制提供了重要的理论基础，也为开发新型抗真菌药物提供了潜在的靶点信息。

关键词 G 蛋白偶联受体（GPCRs）；扩展青霉（*Penicillium expansum*）；生物信息学；跨膜预测软件；保守结构域
中图分类号 Q93－31 **文献标识码** A **文章编号** 1005－7021(2026)02－0058－11
doi：10.3969/j.issn.1005－7021.2026.02.006

Prediction and Bioinformatics Analysis of G Protein-Coupled Receptors in *Penicillium expansum*

SONG Cong[1], SONG Wen-ming[1,3], SUN Yang[1,3], JIA Zhen-hua[1], SONG Shui-shan[1*], GUAN Jun-feng[2*]

(1. *Institute of Biology, Hebei Academy of Sciences, Shijiazhuang* 050081;
2. *Institute of Biotechnology and Food Science, Hebei Academy of Agricultural and Forestry Sciences, Shijiazhuang* 050051;
3. *Hebei Minzu Normal University, Chengde* 067000)

Abstract G protein-coupled receptors (GPCRs) in pathogenic fungi have become one of the critical targets for controlling pathogenic disease. In this study, 36 candidate GPCRs were identified from *Penicillium expansum* by bioinformatics approaches including three transmembrane prediction tools (DeepTMHMM, Phobius, and TOPCONS) and four GPCR prediction platforms (GPCRHMM, GPCRPipe, GPCR-Pen, and GPCRPred). The topological structure analysis, subcellular localization, and functional annotation demonstrated the reliability of the prediction strategy. Phylogenetic classification revealed that these candidate GPCRs were distributed across six distinct fungal GPCR classes, which are characterized by conserved domains such as fungal rhodopsin-like receptor families, pheromone mating factor STE2, pheromone A receptor STE3, and RTA1-like proteins. Physicochemical property analysis demonstrated that the candidate GPCRs predominantly exhibit neutral or alkaline proteins, with 44.44% classified as unstable proteins and 97.22% showing hydrophobic properties. These findings provide a critical theoretical foundation for elucidating the biological functions, signaling mechanisms, and pathogenic roles of GPCRs in *P. expansum*, and also highlighting

基金项目：河北省科学院科技计划项目（23A06，24A06，25A06）；国家梨产业技术体系项目（CARS-28-23）
作者简介：宋聪 女，副研究员，博士研究生。研究方向为植物病害生物防治。E-mail：songcong1982@126.com
* 通讯作者。宋水山，男，研究员，博士。研究方向为细胞生物学与分子生物学。E-mail：shuishans620@163.com
关军锋，男，研究员，博士，博士生导师。研究方向为植物生理与分子生物学。E-mail：junfeng-guan@263.net
收稿日期：2025-12-03

中国知网 https://www.cnki.net

potential targets for the development of novel antifungal molecules.

Keywords G protein-coupled receptors(GPCRs); *Penicillium expansum*; bioinformatics; transmembrane prediction software; conserved domain

扩展青霉(*Penicillium expansum*)隶属于子囊菌门(Ascomycota)散囊菌纲(Eurotiomycetes)散囊菌目(Eurotiales)青霉属(*Penicillium*),是一种对水果产业造成重大经济损失的病原真菌[1-2]。该真菌通过果实的果柄、萼筒、皮孔或伤口等途径侵入,在适宜环境下迅速导致果实内部腐烂,并产生刺鼻气味[3-4]。更为严重的是,扩展青霉还能分泌展青霉素(Patulin,PAT),这种毒素对人体具有明显的致癌、致畸、致突变作用,对消费者的生命健康构成了严重威胁[3,5-6]。G 蛋白偶联受体(G protein-coupled receptors,GPCRs)是一类具有七个跨膜 α 螺旋结构的蛋白质,其跨膜结构域由疏水性的氨基酸残基组成,这些残基能够嵌入细胞膜的磷脂双分子层中,从而稳定受体在膜内的位置,这种疏水性是 GPCRs 能够正确插入细胞膜并保持其天然构象的关键因素。GPCRs 具有细胞外 N 端配体结合域和 C 端胞质尾巴,多数 GPCRs 的长度为 200 ~ 1 000 个氨基酸[7-8],GPCRs 负责调节细胞与其周围环境之间的通信[9-10],并广泛存在于真核生物中。GPCRs 能够识别和结合各种配体,并通过与效应分子相互作用触发细胞内信号传导途径,将细胞外的信号传递到细胞内,从而调控多种生理和生化过程[11]。在真菌中,GPCRs 是感知和响应外界环境变化的关键因素,参与真菌的生长、发育、代谢和致病性等多种生理活动[12]。因此,GPCRs 作为潜在的抗真菌药物靶点,具有重要的研究价值[13]。通过靶向 GPCRs 开发新的抗真菌策略,能够有效防控真菌病害,减少农业经济损失并保障食品安全。目前,真菌 GPCRs 被分为 14 种类别。继 6 种经典真菌 GPCRs(Ⅰ~Ⅴ和Ⅸ类:信息素类受体Ⅰ、信息素类受体Ⅱ、碳源受体类、氮源受体类、cAMP 受体类和视紫红质蛋白)后,又报道了 8 种新型真菌类(Ⅵ~Ⅷ和 X ~ XIV 类:GrPK 类/RGS 域、大鼠生长激素释放因子受体类、mPR 类/PAQR、肺 7TM 超家族、GPCR89/ABA GPCR、C 类家族、DUF300 超家族/PsGPR11 和 Pth11 类)[14-15]。寻找真菌中 GPCRs 的方法通常有两种:一种是借助跨膜预测软件预测候选 GPCRs[16];另一种是通过模式生物和已知 GPCRs 蛋白同源比对及遗传关系分析方法寻找 GPCRs,并应用此方法进行了烟曲霉(*Aspergillus fumigatus*)[14,17]、禾谷炭疽菌(*Colletotrichum graminicola*)[18]、大豆疫霉(*Phytophthora sojae*)[19]、稻瘟病(*Pyricularia oryzae*)[20]和绿僵菌(*Metarhizium anisopliae*)[21]等GPCRs 的寻找和功能研究,然而,目前关于扩展青霉 GPCRs 的研究仍较少,限制了对这一重要病原真菌防治策略的发展。为了填补这一研究空白,本研究以三阶段的方式在扩展青霉基因组序列中鉴定了 36 个 GPCRs,结合基于不同算法的三种跨膜预测工具和四个 GPCR 预测平台,考虑 GPCRs 的固有特性,通过拓扑结构、亚细胞定位和功能分析验证了该鉴定方法的可靠性。同时,结合已知其他丝状真菌 GPCRs 的遗传关系分析,对鉴定的扩展青霉 36 个 GPCRs 进行分类,并对 GPCRs 的保守结构域、保守基序、理化性质及疏水性进行深入分析。通过预测和分析扩展青霉中的 GPCRs,可以更好地理解这些蛋白在真菌信号转导和致病性中的作用,为后续的功能研究、致病机制和药物靶点验证提供参考。此外,本研究设计的 GPCR 鉴定方法也可用于其他生物的 GPCRs 鉴定。

1 材料与方法

1.1 材料

全基因组数据:扩展青霉(*Penicillium expansum*) d1 全基因组数据来源于 NCBI 数据库(https://www.ncbi.nlm.nih.gov/datasets/genome/GCA_000769735.1/),下载其中蛋白质序列文件。

1.2 方法

1.2.1 GPCRs 获取策略 为了预测扩展青霉菌中的可能 GPCRs,本研究采用了一系列不同的生物信息学方法,首先对菌株蛋白质组进行功能注释及序列特征分析,探讨氨基酸长度与蛋白功能类群间的分布规律,基于机器学习和统计算法,使用三种预测软件进行七跨膜结构预测,4 个预测平台进行 GPCRs 预测,并通过不同软件、平台对

中国知网 https://www.cnki.net

筛选到的候选 GPCRs 进行拓扑结构、亚细胞定位和 GO 功能预测等验证，最终获得扩展青霉 GPCRs。整个 GPCRs 识别流程由四个关键步骤组成，见图 1。

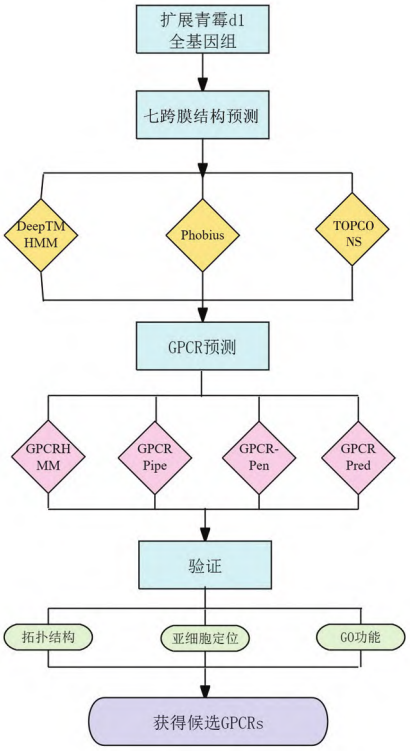


图 1 扩展青霉候选 GPCRs 获取方法流程图

Fig. 1 Flowchart detailing the computational analysis of candidate GPCRs in *Penicillium expansum*

1.2.2 长度筛选和跨膜检测 首先根据蛋白质序列的长度对大量蛋白质序列进行筛选。扩展青霉菌 *P. expansum* d1 的蛋白质组数据集包含 11 023个蛋白质序列，分别根据序列长度进行筛选，以弃除少于 200 个氨基酸残基的蛋白质。然后检查这些序列中跨膜螺旋的数量。使用 3 种不同的跨膜检测器：① DeepTMHMM[22]（https://services. healthtech. dtu. dk/services/DeepTMHMM-1.0/），一种基于隐马尔可夫模型（HMM）的膜蛋白拓扑预测方法；②Phobius[23]（https://phobius. sbc. su. se/index. html），一种基于 HMM 的信号肽预测器，同时检测跨膜蛋白拓扑结构；③TOPCONS[24]（https://topcons. cbr. su. se/pred/），一种基于复杂算法的膜蛋白拓扑预测工具。

1.2.3 GPCRs 预测 使用 4 个不同的 GPCR 预测平台对检测到的七跨膜序列进行鉴定，筛选出候选 GPCRs。4 个不同的 GPCR 预测平台分别为①GPCRHMM[25]（https://gpcrhmm. sbc. su. se/index. html），一种基于 HMM 的 GPCR 预测工具，会开启本地评分；②GPCRPipe（http://aias. biol. uoa. gr/GPCRpipe/search. php），它通过两个连续步骤检测 GPCRs。首先，使用 HMM，然后使用 39 个 Pfam 特征 HMMs 的库（专门针对不同的 GPCR 家族）。本研究选择只有在两种方法都确认的情况下才会预测 GPCR，这种选择会减少假阳性的数量；③GPCR-Pen[26]（https://gpcr. utep. edu/），本研究将预测算法限制在 Pfam 和 BLAST 上，以仅基于序列相似性获取结果，并获取这两种方法都确认的序列；④GPCRPred[27]（https://webs. iiitd. edu. in/raghava/gpcrpred/），一种基于支持向量机（SVM）的方法，使用默认参数。

1.2.4 GPCRs 验证 拓扑结构验证：根据 DeepTMHMM、TOPCONS 再次验证筛选到的 GPCRs 的 N 端是否位于细胞膜外侧；亚细胞定位：使用 WoLF PSORT（https://www. genscript. com/wolf-psort. html）预测蛋白在细胞的具体位置；功能验证：利用 PSIPRED（http://bioinf. cs. ucl. ac. uk/psipred/）工作台上的 FFPRED4 对候选 GPCRs 进行功能注释，输出数据包括生物过程、分子功能和细胞成分预测。

1.2.5 GPCRs 分类 将扩展青霉的候选 GPCRs 与在 NCBI 文献中已知的产黄青霉（*Penicillium rubens*）IBT 27055[14] 和构巢曲霉（*Aspergillus nidulans*）FGSC A4[28] 的 GPCRs 进行 BLAST 比对，亲缘关系相近的蛋白（相似性 >50%，E-value <1E-9），利用 MEGA 11 进行多重比对序列分析，构建系统进化树，使用邻接法（Neighbor-joining method，NJ），设置 Bootstrap 重抽样 1 000 次。并对预测的扩展青霉候选 GPCRs 进行分类。使用 iTools 进行美化。

1.2.6 保守结构域、基序分析 将候选 GPCRs 蛋白序列集提交至 MEME 网站[29]（https://meme-suite. org/meme/tools/meme），分析序列保守基序，查找基序数量为 10，利用 NCBI 中 Batch-CD search 分析序列保守结构域信息。使用 TBtools 中的 One Step Build a ML Tree 插件构建系统进化树，Bootstrap 重抽样 1 000 次。利用 TBtools 中 Gene Structure View 功能完成可视化。使用默认参数在 HHpred[30]（http://toolkit. tuebin-

中国知网 https://www.cnki.net

gen. mpg. de/hhpred)中对基序进行功能注释。

1.2.7 蛋白质理化性质及疏水性预测 利用 Protparam[31] 程序(https://web. expasy. org/protparam)进行蛋白氨基酸含量、PI、不稳定系数及疏水性测定,其中亲水性(Grand average of hydropathicity)指数为正数是疏水蛋白,负数为亲水蛋白。

2 结果与分析

2.1 蛋白数量与功能分析

对扩展青霉全基因组编码的 11 023 个蛋白质进行序列特征解析显示,蛋白质氨基酸序列长度范围为 40 ~7 287 个氨基酸(aa)。氨基酸长度大于 200 小于等于 1 000 的蛋白个数为 8 450 个,占总蛋白数量的 76.66%(图 2A)。

对扩展青霉中 11 023 个蛋白功能归类分析发现,假定蛋白(Hypothetical protein)为 3 560 个,主要易化子超家族(Superfamily)486 个,转录因子(Transcription factor)210 个,三者数量最多,占比分别为 32.30%、4.41%、1.91%。其次分别为细胞色素 P450(Cytochrome P450)160 个,短链脱氢酶(Short-chain dehydrogenase)123 个,氨基酸/多胺转运蛋白 I(Amino acid/polyamine transporter I)70 个,醛缩酶型 TIM 桶(Aldolase-type TIM barrel)68 个,解旋酶 C 端(Helicase, C-terminal)68 个,糖苷水解酶(Glycoside hydrolase)66 个,共 555 个。除此之外,具有其他特征功能蛋白总数为 6 350 个,占比 57.61%(图 2B)。

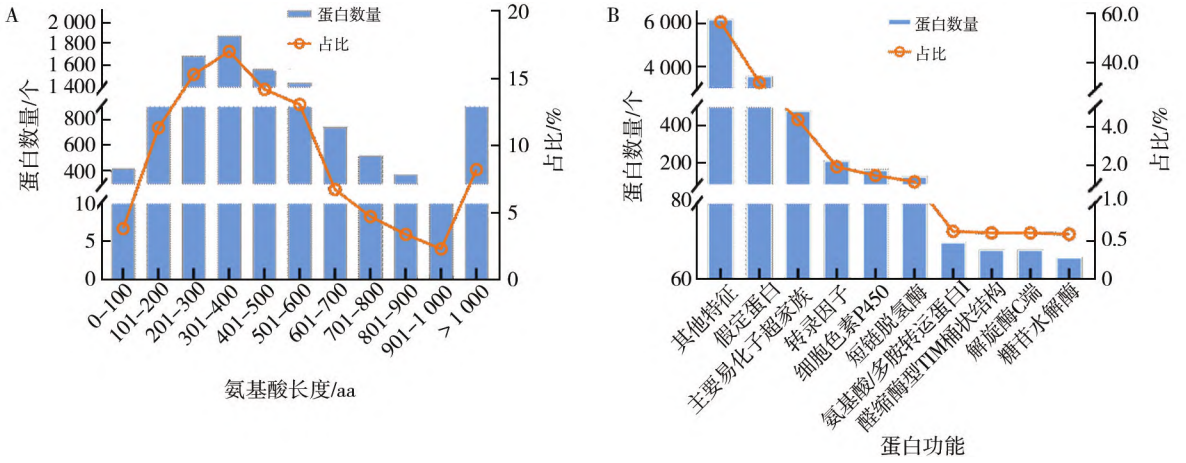


图 2 扩展青霉氨基酸长度分布和蛋白功能分析

Fig. 2 Distribution of amino acid length and functional analysis of proteins in *Penicillium expanum*

A:氨基酸长度分布;B:蛋白功能分析

A: Amino acid length; B: Protein function

2.2 扩展青霉跨膜蛋白数量分析

通过 DeepTMHMM 软件分析发现,扩展青霉中具有 7 个跨膜结构域的蛋白数量为 137 个,占比 1.24%;不具有跨膜结构域的蛋白数量为 8 712 个,占比 79.03%。同时,利用 Phobius 软件分析,得出该菌中具有典型七次跨膜螺旋结构的蛋白为 156 个,占比仅为 1.42%;不具有跨膜结构域的蛋白数量为 8 381 个,占比 76.03%。通过 TOPCONS 软件分析得出该菌中具有 7 个跨膜结构域的蛋白数量为 150 个,占比 1.36%,不具有跨膜结构域的蛋白数量为 8 947 个,占比 81.17%(图 3)。通过上述三种跨膜结构数量预测软件分

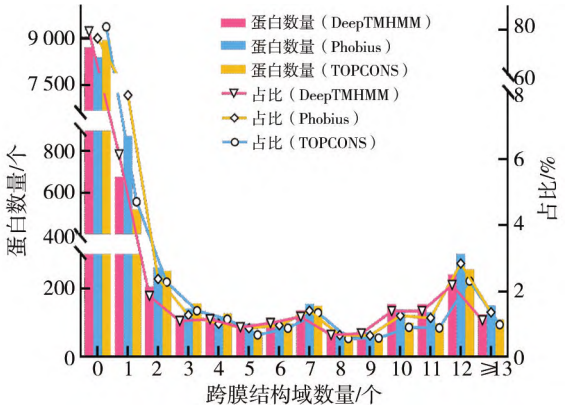


图 3 跨膜蛋白数量分析

Fig. 3 Analysis of transmembrane protein number

中国知网 https://www.cnki.net

析,明确不同预测软件因算法不同预测到的典型七次跨膜结构蛋白数量也不同。

对上述三种软件获得的具有 7 个跨膜结构域的 193 个蛋白进行分析发现,三种软件共同预测到的蛋白数量为 108 个,由 DeepTMHMM 和 Phobiu 共同预测到的七跨膜蛋白数为 11 个;由 DeepTMHMM 和 TOPCONS 共同预测到的七跨膜蛋白数也为 11 个;由 Phobius 和 TOPCONS 共同预测到的七跨膜蛋白数为 12 个。同时,三种软件分别特异识别预测的七跨膜蛋白数分别为 7、25、19 个(图 4A)。

为了保证本筛选策略的严谨性,选取三种跨膜预测软件均预测到的 108 个七跨膜蛋白进行后续分析,使用 4 个 GPCR 预测平台预测并鉴定这些序列中的 GPCRs。GPCRHMM、GPCRPipe、GPCR-Pen 和 GPCRPred 分别鉴定出 36、4、104、93 条序列为 GPCRs。GPCRHMM、GPCRPipe 和 GPCRPred 没有发现任何独有的 GPCR,但 11 个 GPCRs 仅由 GPCR-Pen 鉴定出。综上所述,4 个 GPCR 预测平台中,至少 3 个预测平台识别的 36 条序列被认为是假定的 GPCRs,并通过序列的详细分析进一步证实(图 4B)。

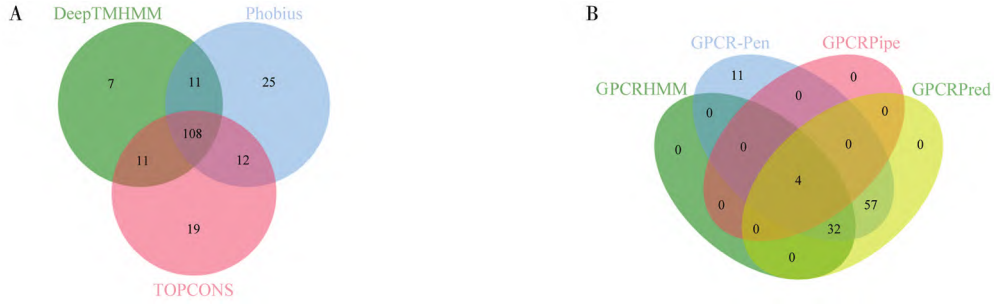


图 4 不同工具和平台对跨膜结构预测及 GPCRs 预测

Fig. 4 7-TM sequences predicted and GPCRs predicted with different software tools and platforms

A:不同软件(DeepTMHMM、Phobius、TOPCONS)预测的具有 7 个跨膜结构域的蛋白数量;B:4 个不同平台(GPCRHMM、GPCRPipe、GPCRPred 和 GPCR-Pen)预测的 GPCRs 数量

A: 7-TM predicted by different tools (DeepTMHMM, Phobius, and TOPCONS); B: GPCRs predicted by 4 different platforms (GPCRHMM, GPCRPipe, GPCRPred, and GPCR-Pen)

2.3 拓扑结构、亚细胞定位及蛋白功能预测

依据 DeepTMHMM、TOPCONS 的二维拓扑结构预测平台预测的 36 个蛋白 N 端均位于细胞膜外侧,WoLF PSORT 对亚细胞定位的预测结果也显示,所有候选 GPCRs 均定位于细胞膜上,符合 GPCR 典型特征。基于支持向量机(SVM)的可靠性评分,并结合基因本体(GO)注释的统计显著性概率,对所有注释结果进行了筛选与分析(表 1)发现,生物过程预测(Biological Process Predictions,BP)中有 13 个蛋白注释到了 GPCRs 信号通路(G-protein coupled receptor signaling pathway,GO:0007186),11 个蛋白注释到细胞表面受体信号通路(cell surface receptor signaling pathway,GO:0007166),10 个蛋白注释到转运(transport,GO:0006810),还有 2 个蛋白分别注释到膜转运(cation transmembrane transport,GO:0098655)和对刺激的响应(response to stimulus,GO:0050896)。分子功能预测(Molecular Function Prediction)中有 16 个蛋白注释到 GPCRs 活性(GO:0004930 G-protein coupled receptor activity),8 个蛋白注释到信号受体活性(GO:0038023 signaling receptor activity),6 个蛋白注释到受体活性(GO:0004872 receptor activity),还有 6 个蛋白分别注释到离子跨膜转运蛋白活性(ion transmembrane transporter activity,GO:0015075)、锌离子结合(zinc ion binding,GO:0008270)、跨膜信号传导受体活性(transmembrane signaling receptor activity,GO:0004888)和金属离子跨膜转运活性(metal ion transmembrane transporter activity,GO:0046873)。所有的候选蛋白均注释到细胞组分(Cellular Component Prediction)中的膜组成成分(integral component of membrane,GO:0016021)。

2.4 遗传关系分析及分类

本研究通过 BLAST 比对分析发现了与已报道的产黄青霉(*P. rubens*) IBT 27055[14] 和构巢曲

中国知网 https://www.cnki.net

霉(*A. nidulans*)FGSC A4[28]真菌 GPCRs 高度相似的候选 GPCRs 蛋白(E-value < 1E-09)并进行了分类(图 5)。其中Ⅰ、Ⅱ类 GPCR 为真菌信息素(Pheromone receptor)各 1 个,Ⅲ类为碳传感器(Carbon sensor)1 个,X 类为肺 7-TM 超家族(Lung 7-TM Superfamily)1 个,Ⅻ类为类 C 家族(Family C-like)1 个,XIV 类为 Pth11 样蛋白(Pth11-like)25 个,另有 6 个蛋白未确定分类。

表 1 扩展青霉候选 GPCRs 序列的基因本体注释

Table 1 Gene Ontology annotations of candidate GPCR sequences of *Penicillium expansum*

蛋白 ID	基因本体(GO)注释分类		
	生物过程	分子功能	细胞成分
KGO35835.1	GO:0007186 G-protein coupled receptor signaling pathway	GO:0001664 G-protein coupled receptor activity	GO: 0016021 integral component of membrane
KGO36060.1	GO:0007166 cell surface receptor signaling pathway	GO: 0004888 transmembrane signaling receptor activity	GO: 0016021 integral component of membrane
KGO37377.1	GO:0007166 cell surface receptor signaling pathway	GO:0004930 G-protein coupled receptor activity	GO: 0016021 integral component of membrane
KGO37443.1	GO:0006810 transport	GO:0015075 ion transmembrane transporter activity	GO: 0016021 integral component of membrane
KGO38035.1	GO:0007166 cell surface receptor signaling pathway	GO:0038023 signaling receptor activity	GO: 0016021 integral component of membrane
KGO38598.1	GO:0007186 G-protein coupled receptor signaling pathway	GO:0004872 receptor activity	GO: 0016021 integral component of membrane
KGO38905.1	GO:0007166 cell surface receptor signaling pathway	GO:0004930 G-protein coupled receptor activity	GO: 0016021 integral component of membrane
KGO38985.1	GO:0007186 G-protein coupled receptor signaling pathway	GO:0004930 G-protein coupled receptor activity	GO: 0016021 integral component of membrane
KGO39371.1	GO:0007186 G-protein coupled receptor signaling pathway	GO:0004930 G-protein coupled receptor activity	GO: 0016021 integral component of membrane
KGO39537.1	GO:0007166 cell surface receptor signaling pathway	GO:0004930 G-protein coupled receptor activity	GO: 0016021 integral component of membrane
KGO39543.1	GO: 0098655 cation transmembrane transport	GO:0004930 G-protein coupled receptor activity	GO: 0016021 integral component of membrane
KGO40765.1	GO:0007166 cell surface receptor signaling pathway	GO:0004930 G-protein coupled receptor activity	GO: 0016021 integral component of membrane
KGO40834.1	GO:0007166 cell surface receptor signaling pathway	GO:0038023 signaling receptor activity	GO: 0016021 integral component of membrane
KGO40853.1	GO:0007166 cell surface receptor signaling pathway	GO:0004872 receptor activity	GO: 0016021 integral component of membrane
KGO41132.1	GO:0006810 transport	GO:0038023 signaling receptor activity	GO: 0016021 integral component of membrane
KGO41145.1	GO:0006810 transport	GO:0015075 ion transmembrane transporter activity	GO: 0016021 integral component of membrane
KGO41364.1	GO:0006810 transport	GO:0046873 metal ion transmembrane transporter activity	GO: 0016021 integral component of membrane
KGO41636.1	GO:0007186 G-protein coupled receptor signaling pathway	GO:0004872 receptor activity	GO: 0016021 integral component of membrane
KGO41804.1	GO:0006810 transport	GO:0038023 signaling receptor activity	GO: 0016021 integral component of membrane
KGO41854.1	GO:0007186 G-protein coupled receptor signaling pathway	GO:0004872 receptor activity	GO: 0016021 integral component of membrane
KGO42481.1	GO:0007186 G-protein coupled receptor signaling pathway	GO:0004930 G-protein coupled receptor activity	GO: 0016021 integral component of membrane
KGO42528.1	GO:0007166 cell surface receptor signaling pathway	GO:0008270 zinc ion binding	GO: 0016021 integral component of membrane
KGO42745.1	GO:0007166 cell surface receptor signaling pathway	GO:0004930 G-protein coupled receptor activity	GO: 0016021 integral component of membrane
KGO44072.1	GO:0050896 response to stimulus	GO:0004930 G-protein coupled receptor activity	GO: 0016021 integral component of membrane
KGO44724.1	GO:0007186 G-protein coupled receptor signaling pathway	GO:0004930 G-protein coupled receptor activity	GO: 0016021 integral component of membrane

中国知网 https://www.cnki.net

续表 1

蛋白 ID	基因本体(GO)注释分类		
	生物过程	分子功能	细胞成分
KGO45253.1	GO:0007186 G-protein coupled receptor signaling pathway	GO:0004930 G-protein coupled receptor activity	GO: 0016021 integral component of membrane
KGO45330.1	GO:0007186 G-protein coupled receptor signaling pathway	GO:0004930 G-protein coupled receptor activity	GO: 0016021 integral component of membrane
KGO45386.1	GO:0006810 transport	GO:0004930 G-protein coupled receptor activity	GO: 0016021 integral component of membrane
KGO45922.1	GO:0007186 G-protein coupled receptor signaling pathway	GO:0004930 G-protein coupled receptor activity	GO: 0016021 integral component of membrane
KGO45933.1	GO:0007186 G-protein coupled receptor signaling pathway	GO:0038023 signaling receptor activity	GO: 0016021 integral component of membrane
KGO46318.1	GO:0006810 transport	GO:0038023 signaling receptor activity	GO: 0016021 integral component of membrane
KGO46483.1	GO:0006810 transport	GO:0004872 receptor activity	GO: 0016021 integral component of membrane
KGO47822.1	GO:0006810 transport	GO:0008270 zinc ion binding	GO: 0016021 integral component of membrane
KGO47949.1	GO:0006810 transport	GO:0038023 signaling receptor activity	GO: 0016021 integral component of membrane
KGO48848.1	GO:0007186 G-protein coupled receptor signaling pathway	GO:0004872 receptor activity	GO: 0016021 integral component of membrane
KGO48990.1	GO:0004871 cell surface receptor signaling pathway	GO:0038023 signaling receptor activity	GO: 0016021 integral component of membrane

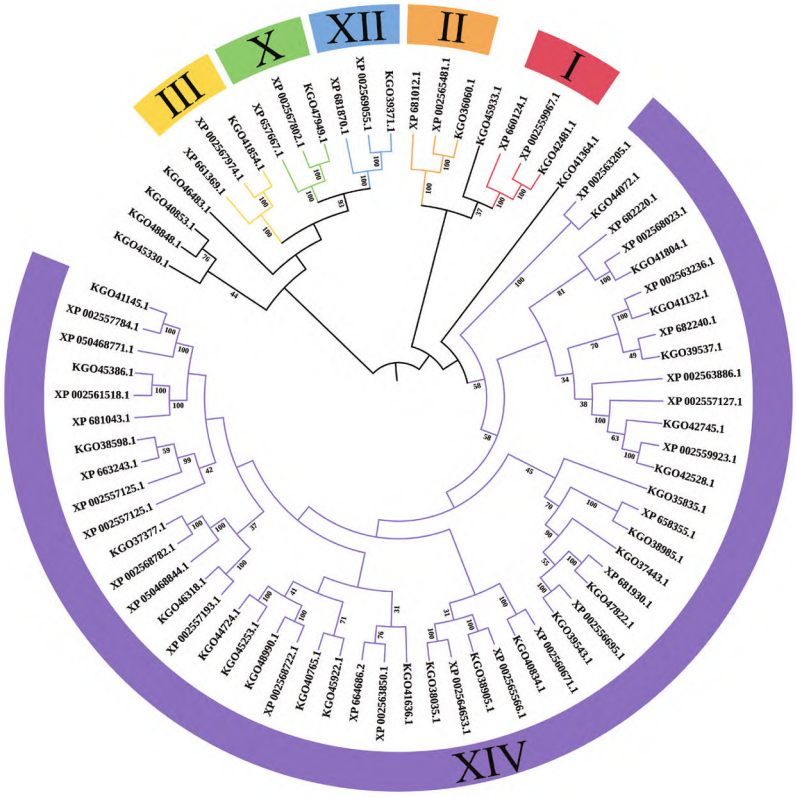


图 5 候选 GPCRs 遗传分析及类型分组

Fig. 5 Genetic analysis and type grouping of candidate GPCRs

2.5 保守结构域、Motif 基序分析

NCBI 保守结构域分析显示,分别有 1 个真菌信息素交配因子 STE2 和 1 个信息素 A 受体 STE3、2 个 RTA1 样蛋白,以及 7-tm GPCRs 超家

中国知网 https://www.cnki.net

族、DUF3112、G 蛋白偶联葡萄糖受体调节 Gpa2 C 端、PQ 循环重复等结构，均为 GPCRs 保守结构域，其余候选 GPCRs 均属于真菌视紫红质受体样家族（Fungi rhodopsin），该家族是 GPCRs 超家族中最大的组成部分，包括光敏感视紫红质和生物胺、脂质、核苷酸、气味分子、肽激素等各种配体的受体。在 Motif 结构中，多数 XIV 类 GPCRs 含有 Motif 1、Motif 2、Motif 3、Motif 4、Motif 5、Motif 6、Motif 7，这些 Motif 通过 HHpred 预测并被归类于 pfam 数据库中的 PF20684.2 分类，即真菌视紫红质结构域（Fungal rhodopsin domain）。未获得分类的蛋白中 KGO40853.1 和 KGO48848.1 有 RTA1 样保守结构域，KGO46483.1 有 PQ-loop 结构（图 6）。

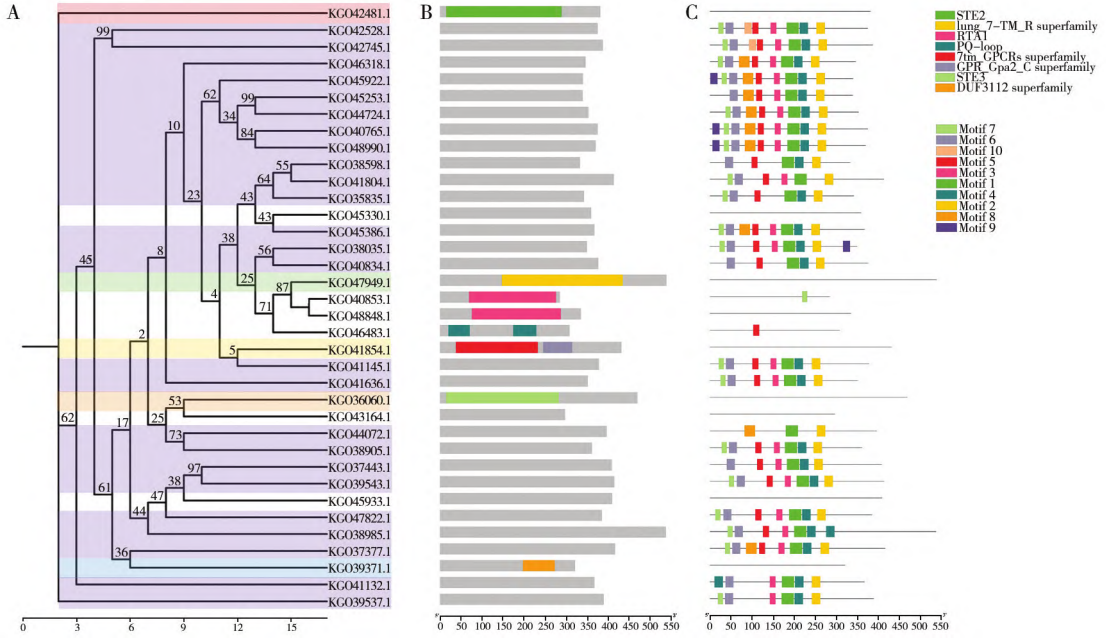


图 6　保守结构域和 Motif 分析

Fig. 6　The conserved domain and motif analysis

A：进化树；B：保守结构域；C：保守基序

A: Phylogenetic tree; B: Conserved domain; C: Conserved motif distribution

2.6　蛋白质理化性质及疏水性预测

对扩展青霉 36 个候选 GPCRs 蛋白的理论等电点（pI）分析显示，KGO41145.1 的理论 pI 最高，为 9.54；KGO41364.1 的理论 pI 最低，为 4.74。进一步以 1.00 为间隔单位对全候选蛋白进行分类统计（图 7A），结果表明，候选 GPCRs 蛋白的理论 pI 主要分布在 8.01 ~ 9.00 和 9.01 ~ 10.00 区间，对应数量分别为 16 个和 12 个，合计占比 77.78%；其次分布在 6.01 ~ 7.00 和 7.01 ~ 8.00 区间，共包含 6 个蛋白，占比 16.67%；而分布在 4.01 ~ 5.00 和 5.01 ~ 6.00 区间的蛋白数量最少，均仅为 1 个。上述结果表明，候选蛋白中碱性蛋白占主导地位。

对扩展青霉中 36 个候选 GPCRs 蛋白的不稳定性系数分析显示（图 7B），KGO41854.1 的不稳定系数最高，为 51.63；KGO40853.1 的不稳定系数最低，为 29.44。进一步统计表明，不稳定系数 <40 的候选蛋白共 20 个，占比 55.56%。不稳定性系数 ≥40 的候选蛋白数量共 16 个，占比 44.44%。上述结果表明，不稳定蛋白与稳定蛋白在候选 GPCRs 中的比例相对接近。

对上述 36 个候选 GPCRs 的疏水性分析表明（图 7C），亲水性蛋白（亲水性值 <0）仅 1 个（KGO37443.1），其余 35 个均为疏水性蛋白（亲水性值 >0），占比 97.22%。进一步统计疏水性最强的氨基酸残基种类发现，主要分布为 5 种，其中亮氨酸（Leu）、异亮氨酸（Ile）、丙氨酸（Ala）为优势残基，对应蛋白数量分别为 23、6 和 3 个，占比依次为 63.89%、16.67% 和 8.33%。

中国知网 https://www.cnki.net

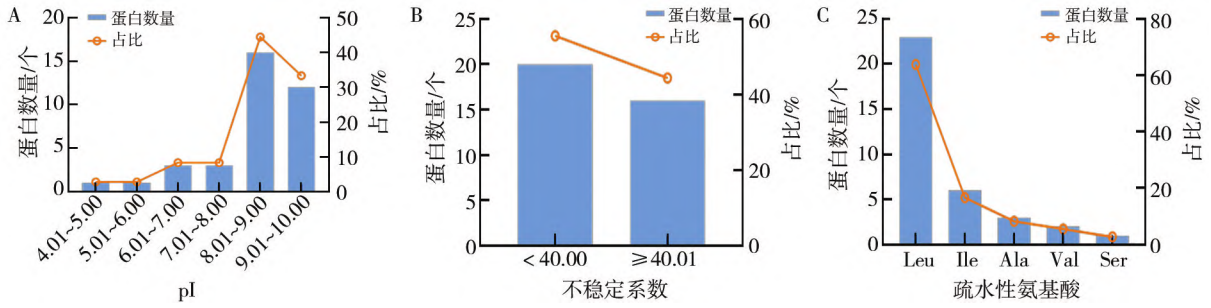


图 7 候选 GPCRs 基本性质分析

Fig. 7 Basic property analysis of candidate GPCRs

3 讨 论

本研究采用一种严谨的生物信息学流程,成功识别了扩展青霉中 36 个潜在的 G 蛋白偶联受体(GPCRs)。根据 Trzaskowski 等[32]的研究,GPCR 的特征结构存在七段 25 ~ 35 个连续氨基酸残基。为了维持热力学上的稳定,跨膜蛋白需要折叠和包装,使得蛋白质表面环和螺旋的平均长度分别为 19 和 26 个氨基酸残基。因此,包含蛋白质编码 DNA 序列的 7 个跨膜域的最小长度应为 296 个碱基。本研究使用了大于 200 个碱基的序列长度。

丝状真菌 GPCRs 预测研究多采用模式生物同源比对策略,但该方法存在假阳性和遗漏现象。研究表明,基于构巢曲霉、粗糙脉孢霉(*Neurospora crassa*)及禾谷镰刀菌(*Fusarium graminearum*)的 GPCRs 蛋白为参考序列,联合跨膜拓扑结构预测,可鉴定 116 个候选 GPCRs,却难以规避非特异性匹配导致的偏差[33]。曹淑琳[18]通过整合转录组动态表达谱与多算法拓扑验证,将预测结果优化至 85 个 GPCRs。蛋白跨膜结构域预测软件主要基于 ANN 和 HMM 算法,不同软件预测结果存在偏差。韩长志等[34]通过分析 223 个真菌 GPCRs 蛋白序列,发现 TOPCONS 预测到 158 个蛋白序列具有 7 个跨膜结构域,TMHMM 预测仅 115 个蛋白具有 7 个跨膜结构域。这表明 GPCR 蛋白跨膜结构域预测方法需进一步优化,也强调了多种预测方法结合使用的必要性。本研究使用的 GPCR 找寻方法是高度保守的,除了使用 3 种跨膜预测软件预测 7 个跨膜螺旋结构外,还使用了 4 个不同的平台预测 GPCRs,通过“hmmsearch”将蛋白质与 GPCR Pfam 隐马尔可夫模型(HMMs)数据库进行比较。单一生物信息学工具在跨膜预测中可能出现错误,尽管所有工具均基于类似的 HMM 架构,但预测过程中所涉及的方法和途径存在显著差异。此外,由于不同平台采用的交叉验证方法和后台数据集不同,其准确性也各有不同。因此,综合运用 3 种不同的跨膜预测软件、4 个 GPCRs 预测平台,不仅能够提供准确的预测结果,还能有效降低假阳性或遗漏出现的概率,从而显著增强 GPCR 检测流程的可靠性和准确性。为了验证预测准确性,本研究通过二维拓扑结构验证扩展青霉 36 个蛋白 N 端是否位于细胞膜外,亚细胞定位预测确定蛋白在细胞的具体定位。在 GO 功能注释中,生物过程预测将候选 GPCRs 注释到“G 蛋白偶联受体信号通路”(GO:0007186)和“细胞表面受体信号通路”(GO:0007166),确认候选 GPCRs 在信号转导中的核心作用,支持其跨膜受体特性;分子功能预测将候选 GPCRs 注释到“G 蛋白偶联受体活性”(GO:0004930)和“离子跨膜转运活性”(GO:0015075),明确蛋白的配体结合、信号传递或物质运输功能,验证其作为 GPCR 的功能标签;细胞组分预测所有候选蛋白均注释到“膜整合成分”(GO:0016021),进一步证实其定位于细胞膜,符合 GPCR 的典型亚细胞定位特征。同时,注释到“转运”(GO:0006810)的蛋白可能参与营养吸收或毒素分泌,与致病性相关;注释到“对刺激的响应”(GO:0050896)的蛋白可能介导环境信号感知。这些结果为后续靶向基因敲除或功能实验提供优先级依据。GO 功能注释通过标准化功能标签,将候选 GPCRs 的预测结果与已知生物学过程、分子机制关联,增强数据可信度,并为解析其致病机制和药物靶点筛

中国知网 https://www.cnki.net

选提供功能层面的理论支撑。进一步通过 NCBI 保守结构域和 MEME 基序预测验证候选 GPCRs 的保守结构域和基序。可以认为,本研究使用的方法因具有更高的过滤标准而更准确和精确,应用的生物信息学流程可用于从各类生物体基因组中识别 GPCRs。

通过同源性比对分析并构建系统进化树,鉴定出的 36 个 GPCRs 分属于 14 个类别中的 6 类中,其中 6 个蛋白未获得分类,其功能还需进一步研究。对候选 GPCRs 的理化性质分析发现,这些蛋白多为中性或碱性蛋白,不稳定蛋白与稳定蛋白数目差别不大,疏水性蛋白占较大比例,这些特征与典型的膜蛋白属性相吻合。

随着基因组学和分子生物学的快速发展,扩展青霉全基因组序列的公布为深入研究其生物学特性、侵染过程、细胞信号转导和致病基因提供了宝贵的信息资源[35]。GPCRs 的典型结构为含有七跨膜螺旋结构域的蛋白,其在细胞信号转导过程中发挥关键作用,并影响真菌的一系列生理生化功能。例如,在真菌的生长发育过程中,GPCRs 能够感知外界环境信号,如营养物质、pH、激素、蛋白质、离子、疏水表面和光等,并将这些信号传递到细胞内,从而调节真菌的发育、代谢、毒力和毒素合成等[36-39]。在真菌的致病性方面,GPCRs 也参与其中,它可以帮助真菌识别宿主细胞表面的信号分子,进而启动一系列致病相关的信号通路,使真菌能够成功侵染宿主并在宿主体内生存繁殖[40-42]。

本研究组前期研究发现,阿氏芽孢杆菌(*Bacillus aryabhattai*)挥发性活性物质能影响扩展青霉菌丝生长、孢子萌发及次级代谢产物的合成,并引起扩展青霉转录组的变化[43],其中多个 GPCRs 相关基因显著上调(结果未发表)。下一步将对 GPCRs 在响应芽孢杆菌挥发性活性物质时的表达水平和功能进行深入研究,这将有利于更好理解 GPCRs 的生物学功能及其在外界信号传导过程中发挥的作用,为开发有效的抗真菌策略提供理论基础[44]。综上所述,病原真菌扩展青霉 GPCRs 的预测和生物信息学分析,为深入理解扩展青霉的生物学特性、信号转导机制以及致病机制提供了理论依据,同时也为新型抗真菌药物的研发提供参考。

参考文献:

[1] Yu L, Qiao N, Zhao J, et al. Postharvest control of *Penicillium expansum* in fruits: A review[J]. Food Bioscience, 2020, 36: 100633.

[2] Xu M, Yang Q, Serwah Boateng NA, et al. Ultrastructure observation and transcriptome analysis of *Penicillium expansum* invasion in postharvest pears[J]. Postharvest Biology and Technology, 2020, 165: 111198.

[3] Mahato DK, Kamle M, Sharma B, et al. Patulin in food: A mycotoxin concern for human health and its management strategies[J]. Toxicon, 2021, 198: 12-23.

[4] Godana EA, Yang Q, Zhang X, et al. Biotechnological and biocontrol approaches for mitigating postharvest diseases caused by fungal pathogens and their mycotoxins in fruits: A review [J]. Journal of Agricultural and Food Chemistry, 2023, 71 (46): 17584-17596.

[5] Li B, Chen Y, Zong Y, et al. Dissection of patulin biosynthesis, spatial control and regulation mechanism in *Penicillium expansum*[J]. Environmental Microbiology, 2019, 21(3): 1124-1139.

[6] Jimdjio CK, Xue H, Bi Y, et al. Effect of ambient pH on growth, pathogenicity, and patulin production of *Penicillium expansum*[J]. Toxins (Basel), 2021, 13(8): 1-13.

[7] Kundu A, Jaiswal N, Rao U, et al. Stringent in-silico identification of putative G-protein-coupled receptors (GPCRs) of the entomopathogenic nematode *Heterorhabditis bacteriophora*[J]. Journal of Nematology, 2023, 55(1): 20230038.

[8] Yadav DK, Srivastava GP, Singh A, et al. Proteome-wide analysis reveals G protein-coupled receptor-like proteins in rice (*Oryza sativa*)[J]. Plant Signaling & Behavior, 2024, 19 (1): 2365572.

[9] Insel PA, Sriram K, Gorr MW, et al. GPCRomics: An approach to discover GPCR drug targets[J]. Trends in Pharmacological Sciences, 2019, 40(6): 378-387.

[10] Dondapati SK, Stech M, Zemella A, et al. Cell-free protein synthesis: A promising option for future drug development[J]. Biodrugs, 2020, 34(3): 327-348.

[11] El-Defrawy MMH, Hesham AEL. G-protein-coupled receptors in fungi[M]. Switzerland: Fungal Biotechnology and Bioengineering, 2020: 37-126.

[12] Velazhahan V, Mccann BL, Bignell E, et al. Developing novel antifungals: Lessons from G protein-coupled receptors[J]. Trends in Pharmacological Sciences, 2023, 44(3): 162-174.

[13] Cheng L, Xia F, Li Z, et al. Structure, function and drug discovery of GPCR signaling[J]. Molecular Biomedicine, 2023, 4(46): 1-39

[14] Martin JF, Van Den Berg MA, Ver Loren Van Themaat E, et al. Sensing and transduction of nutritional and chemical signals in filamentous fungi: Impact on cell development and secondary

中国知网 https://www.cnki.net

metabolites biosynthesis[J]. Biotechnology Advances, 2019, 37(6): 107392.
[15] Mohri M, Moghadam A, Burketova L, et al. Rysanek P. genome-wide identification of the opsin protein in *Leptosphaeria maculans* and comparison with other fungi (pathogens of *Brassica napus*)[J]. Frontiers in Microbiology, 2023, 14: 1193892.
[16] 喻红稠, 韩长志. 禾谷炭疽菌中候选 G 蛋白偶联受体蛋白的找寻[J]. 科学技术与工程, 2022, 22(1): 103-109.
[17] Grice CM, Bertuzzi M, Bignell EM. Receptor-mediated signaling in *Aspergillus fumigatus*[J]. Frontiers in Microbiology, 2013, 4(26): 1-12.
[18] 曹淑琳. 禾谷镰刀菌候选 G 蛋白偶联受体基因和 AMD1 基因的功能研究[D]. 咸阳:西北农林科技大学, 2017.
[19] 郝玉娟. 大豆疫霉 G 蛋白 α 亚基互作蛋白的筛选及功能分析[D]. 南京:南京农业大学, 2013.
[20] 刘淑集. 稻瘟病菌两个假定 G 蛋白偶联受体的功能分析[D]. 福州:福建农林大学, 2007.
[21] 周荣. 罗伯茨绿僵菌中 G 蛋白偶联受体基因的敲除及功能分[D]. 合肥:安徽农业大学, 2015.
[22] Hallgren J, Tsirigos KD, Pedersen MD, et al. DeepTMHMM predicts alpha and beta transmembrane proteins using deep neural networks[EB/OL]. bioRxiv, 2022, 4(8): 487609.
[23] Käll L, Krogh A, Sonnhammer ELL. A combined transmembrane topology and signal peptide prediction method[J]. Journal of Molecular Biology, 2004, 338(5): 1027-1036.
[24] Tsirigos KD, Peters C, Shu N, et al. The TOPCONS web server for consensus prediction of membrane protein topology and signal peptides[J]. Nucleic Acids Research, 2015, 43(W1): W401-407.
[25] Wistrand M, Kall L, Sonnhammer EL. A general model of G protein-coupled receptor sequences and its application to detect remote homologs[J]. Protein Science, 2006, 15(3): 509-521.
[26] Begum K, Mohl JE, Ayivor F, et al. GPCR-PEnDB: A database of protein sequences and derived features to facilitate prediction and classification of G protein-coupled receptors[J]. Database (Oxford), 2020,2020:baaa087.
[27] Bhasin M, Raghava GP. GPCRpred: An SVM-based method for prediction of families and subfamilies of G-protein coupled receptors[J]. Nucleic Acids Research, 2004, 32: W383-389.
[28] Cabrera IE, Pacentine IV, Lim A, et al. Global analysis of predicted G protein-coupled receptor genes in the filamentous fungus, *Neurospora crassa*[J]. G3 (Bethesda), 2015, 5(12): 2729-2743.
[29] Bailey TL, Johnson J, Grant CE, et al. The MEME suite[J]. Nucleic Acids Research, 2015, 43(W1): W39-49.
[30] Zimmermann L, Stephens A, Nam SZ, et al. A completely reimplemented MPI bioinformatics toolkit with a new HHpred server at its core[J]. Journal of Molecular Biology, 2018, 430(15): 2237-2243.
[31] Wilkins MR, Gasteiger E, Bairoch A, et al. Protein identification and analysis tools in the ExPASy server[J]. Methods in Molecular Biology, 1999, 112: 531-552.
[32] Trzaskowski B, Latek D, Yuan S, et al. Action of molecular switches in GPCRs--theoretical and experimental studies[J]. Current Medicinal Chemistry, 2012, 19(8): 1090-1109.
[33] Ma LJ, Van Der Does HC, Borkovich KA, et al. Comparative genomics reveals mobile pathogenicity chromosomes in *Fusarium*[J]. Nature, 2010, 464(7287): 367-373.
[34] 韩长志, 薄淑文. 植物病原丝状真菌 G 蛋白偶联受体的预测方法对比研究[J]. 福建农林大学学报(自然科学版), 2023, 52(5): 607-616.
[35] Luciano D, Keller P, Jurick M. *Penicillium expansum*: Biology, omics, and management tools for a global postharvest pathogen causing blue mould of pome fruit[J]. Molecular Plant Pathology, 2020, 21(11): 1391-1404.
[36] Filho A, Brancini GTP, De Castro PA, et al. *Aspergillus fumigatus* G-protein coupled receptors GprM and GprJ are important for the regulation of the cell wall integrity pathway, secondary metabolite production, and virulence[J]. mBio, 2020, 11(5): e02458-02420.
[37] Hinterdobler W, Beier S, Monroy AA, et al. The G-protein coupled receptor GPR8 regulates secondary metabolism in *Trichoderma reesei*[J]. Frontiers in Bioengineering and Biotechnology, 2020, 8: 558996.
[38] Gao J, Xu X, Huang K, et al. Fungal G-protein-coupled receptors: A promising mediator of the impact of extracellular signals on biosynthesis of ochratoxin A[J]. Frontiers in Microbiology, 2021, 12: 631392.
[39] Moon H, Han KH, Yu JH. Upstream regulation of development and secondary metabolism in *Aspergillus* species[J]. Cells, 2022, 12(2): 1-19.
[40] Yang Y, Huang P, Ma Y, et al. Insights into intracellular signaling network in *Fusarium* species[J]. International Journal of Biological Macromolecules, 2022, 222(Pt A): 1007-1014.
[41] Park HS, Kim MJ, Yu JH, et al. Heterotrimeric G-protein signalers and RGSs in *Aspergillus fumigatus*[J]. Pathogens, 2020, 9(902): 1-9.
[42] Jiang C, Xia A, Xu D, et al. Sensing host and environmental cues by fungal GPCRs[J]. Current Opinin in Plant Biology, 2024, 82: 102667.
[43] Song C, Zhang Y, Zhao Q, et al. Volatile organic compounds produced by *Bacillus aryabhattai* AYG1023 against *Penicillium expansum* causing blue mold on the Huangguan pear[J]. Microbiological Research, 2024, 278: 127531.
[44] Serrano-Marin J, Reyes-Resina I, Martinez-Pinilla E, et al. Natural compounds as guides for the discovery of drugs targeting G-protein-coupled receptors[J]. Molecules, 2020, 25(5060): 1-14.

中国知网 https://www.cnki.net