



研究报告

基于转录组解析淡玫红鹅膏环肽基因家族特征及其系统发育

占宁^{1,2} 王刚正² 张成花² 周小云^{*1} 邓旺秋^{*1,2} 李泰辉²

1 湖南农业大学生物科学技术学院 湖南 长沙 410128

2 广东省科学院微生物研究所 华南应用微生物国家重点实验室 广东省菌种保藏与应用重点实验室
广东 广州 510070

摘要:【背景】淡玫红鹅膏(*Amanita pallidorozea*)是鹅膏属檐托鹅膏组的一种剧毒鹅膏菌,其子实体内含有丰富的鹅膏环肽毒素,但其编码毒环肽和相关肽的基因家族“MSDIN”尚待深入研究。

【目的】探究淡玫红鹅膏中编码毒环肽及相关肽的基因家族成员的多样性、保守性及其系统发育。

【方法】采用 Illumina HiSeq 2000 平台对淡玫红鹅膏转录组进行测序,使用 TBLASTn 软件对 MSDIN 基因家族进行检索,并设计特异性引物进行 PCR 验证,通过生物信息学比对分析 MSDIN 基因家族成员的种类和序列构成,通过重建分子系统发育树了解其演化历程。【结果】通过对 MSDIN 基因家族的查找,从转录组数据获得了 60 条环肽编码基因,经 PCR 验证其可编码 32 条环肽,包括 α -鹅膏毒肽、 β -鹅膏毒肽和羧基二羟鬼笔毒肽。本研究报道了 8 条新环肽。分子系统发育树分析显示,鹅膏环肽分为鹅膏毒肽、鬼笔毒肽和未知功能环肽 3 大支。系统发育结合保守序列推测出 7 条潜在的新毒肽。【结论】淡玫红鹅膏具有丰富的鹅膏环肽资源,利用转录组测序能够系统挖掘鹅膏环肽新资源,为其整体结构解析奠定基础。

关键词: 鹅膏属, 剧毒蘑菇, 毒肽基因, 保守序列, 系统发育分析

Foundation items: Guangzhou Science and Technology Project (201804020018); National Natural Science Foundation of China (31970024); Science and Technology Project of Guangdong Province (2019B121202005, 2018B030324001, 2018B020205001)

***Corresponding authors:** ZHOU Xiaoyun: E-mail: xyzhou71@hotmail.com
DENG Wangqiu: Tel: 86-20-87137620; E-mail: dengwq@gdim.cn

Received: 03-03-2021; **Accepted:** 10-04-2021; **Published online:** 18-05-2021

基金项目: 广州市科技计划重点项目(201804020018); 国家自然科学基金(31970024); 广东省科技计划(2019B121202005, 2018B030324001, 2018B020205001)

***通信作者:** 周小云: E-mail: xyzhou71@hotmail.com

邓旺秋: Tel: 020-87137620; E-mail: dengwq@gdim.cn

收稿日期: 2021-03-03; **接受日期:** 2021-04-10; **网络首发日期:** 2021-05-18

Characteristics and phylogeny of the cyclic peptide gene family in *Amanita pallidorozea*: based on transcriptome analysis

ZHAN Ning^{1,2} WANG Gangzheng² ZHANG Chenghua² ZHOU Xiaoyun^{*1}
DENG Wangqiu^{*1,2} LI Taihui²

1 College of Bioscience and Biotechnology, Hunan Agricultural University, Changsha, Hunan 410128, China

2 Guangdong Provincial Key Laboratory of Microbial Culture Collection and Application; State Key Laboratory of Applied Microbiology Southern China; Institute of Microbiology, Guangdong Academy of Sciences, Guangzhou, Guangdong 510070, China

Abstract: [Background] *Amanita pallidorozea* is a lethal *Amanita* species in the sect. *Phalloideae*, containing abundant cyclic peptide toxins. However, its gene family (MSDIN) encoding the toxic peptides and related peptides needs further systematic study. [Objective] This paper aims to explore the diversity, conservation, and phylogeny of genes encoding toxic peptides and related peptides in *A. pallidorozea*. [Methods] Illumina HiSeq 2000 platform was used for the transcriptome sequencing of *A. pallidorozea* and TBLASTn for searching MSDIN genes. Specific primers were designed for PCR verification. The types and sequence composition of MSDIN genes were analyzed with the bioinformatics method, and the phylogenetic tree was reconstructed to clarify the evolution. [Results] A total of 60 genes encoding cyclic peptides were obtained from the transcriptome data, which encoded 32 cyclic peptides as verified by PCR, including α -amanitin, β -amanitin, and phalloidin. Among them, 8 new cyclic peptide sequences were reported for the first time in this study. Molecular phylogenetic analysis showed that all cyclic peptides were classified into three clades: amatoxins, phallotoxins, and cyclic peptides of unknown function. Seven potential and new toxic peptides were predicted according to phylogenetic analysis and conserved sequences. [Conclusion] *A. pallidorozea* boasts abundant cyclic peptides. Transcriptome sequencing can help discover new resources of cyclic peptides in *A. pallidorozea*, which lays the foundation for the overall structure analysis.

Keywords: *Amanita*, poisonous mushroom, toxic peptide gene, conserved sequence, phylogenetic analysis

淡玫红鹅膏(*Amanita pallidorozea* P. Zhang & Zhu L. Yang)是 2010 年发表并原产于我国的鹅膏新种^[1], 主要分布在中国东北、东部、西北、中部和西南部等地以及日本。淡玫红鹅膏是壳斗科(*Fagaceae*)和松科(*Pinaceae*)植物的外生菌根菌^[2], 其子实体尚不能进行室内栽培。淡玫红鹅膏是一种剧毒蘑菇, 曾在山东、贵州等地引发了 10 多起中毒事件, 导致 20 多人中毒, 5 人死亡^[3-5]。已有研究表明, 淡玫红鹅膏子实体中含有多种极毒的鹅膏环肽毒素, 包括 α -鹅膏毒肽(α -Amanitin, α -AMA)、 β -鹅膏毒肽(β -Amanitin, β -AMA)、二羟鬼笔毒肽(Phalloidin, PHD)和羧基二羟鬼笔毒肽(Phalloidin, PCD)等, 其毒肽总含量高达 8.27 mg/g-干重^[6]。

目前已报道的鹅膏肽类毒素有 22 种^[7], 这些鹅膏肽类毒素由基因家族编码。该基因家族编码的前体肽由引导肽、环肽编码区及识别序列组成, 环肽编码区决定了最终环肽的序列, 理论上该区域可以由任何氨基酸序列组成^[8]。鹅膏环肽的前体肽由 32-39 个氨基酸组成, 引导肽多数为相对保守的 10 个氨基酸, 中间环肽编码区为 7-10 个氨基酸, 识别序列一般为 15-17 个氨基酸, 引导肽和环肽编码区的最后一个氨基酸均为脯氨酸, 这 2 个位点为环肽合成时的水解位点; 因前体肽上游保守序列的前 5 个氨基酸序列通常为“MSDIN”, 所以该基因家族被称为 MSDIN 基因家族(图 1)^[9-10]。

随着组学测序技术的发展, 近年来基于基因

组和转录组测序对剧毒鹅膏 MSDIN 基因家族的研究成为鹅膏毒肽基因研究的热点^[11-14]。相关研究表明,不同剧毒鹅膏种类,其 MSDIN 基因家族具高度多样性和特异性^[9-15]。本研究采用 Illumina HiSeq 2000 平台对我国剧毒蘑菇淡玫红鹅膏转录组进行测序和组装,探究其编码毒环肽及相关肽的基因家族成员的多样性、保守性及其系统发

育,以期为其结构解析和中毒治疗奠定基础。

1 材料与方法

1.1 材料

1.1.1 样品采集

淡玫红鹅膏子实体新鲜样品(图 2)采集于山东泰山,将泥土清理干净后放入低温保存箱,然后运回实验室在-80 °C 超低温冰箱保存。鹅膏样品通

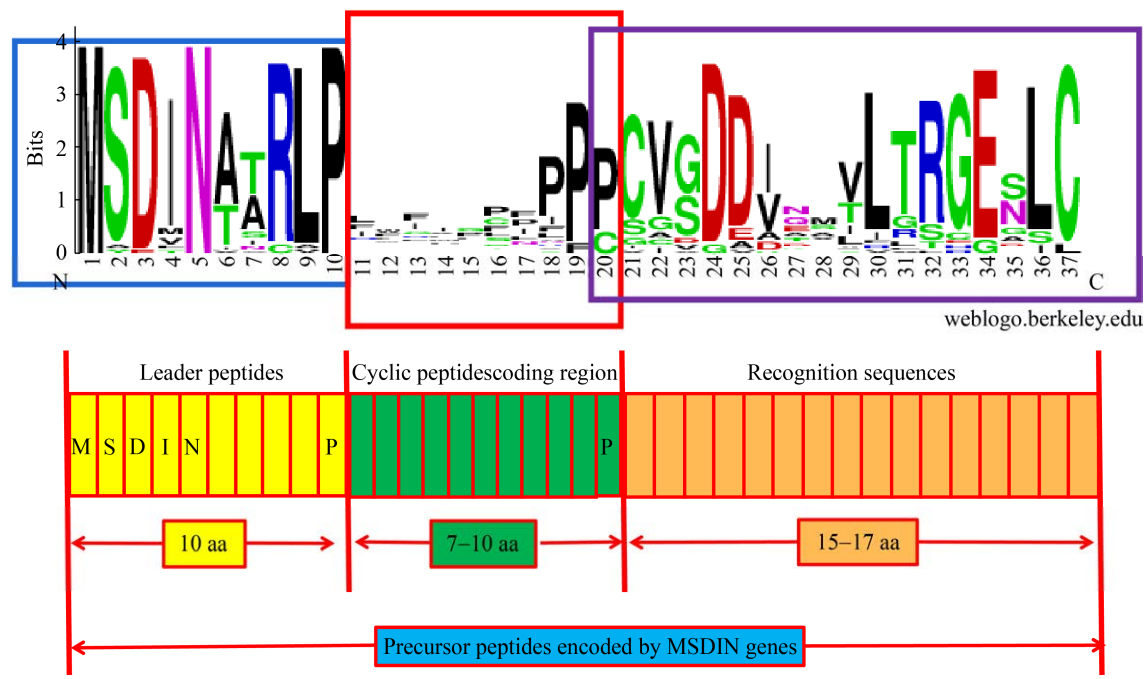


图 1 MSDIN 基因编码的前体肽结构示意图
Figure 1 Schematic diagram of precursor peptides encoded by MSDIN genes



图 2 淡玫红鹅膏子实体
Figure 2 Basidiocarps of *Amanita pallidorosea*

过形态观察和内部转录间隔区(Internal Transcribed Spacer, ITS)序列比对进行鉴定。

1.1.2 主要试剂和仪器

真菌 DNA 小量提取试剂盒, Magen 公司; 全能型植物 RNA 提取试剂盒(DNase I), 北京康为世纪生物科技有限公司; Trizol 试剂, 上海碧云天生物技术有限公司; 三氯甲烷, 天津市百世化工有限公司; 异丙醇, 上海麦克林生化科技有限公司; 75%乙醇, 广东广试试剂科技有限公司; cDNA 反转录试剂盒、Tks Gflex™ DNA Polymerase Low DNA、pMD™19-T 载体、大肠杆菌 DH5 α 、SOC 培养基, TaKaRa 公司。PCR 扩增仪, Eppendorf 公司; 小型水平电泳系统, 北京六一生物科技有限公司; 凝胶成像系统, Bio-Rad 公司; Agilent 2100 生物分析仪, Agilent 公司; NanoDrop, Thermo Fisher Scientific 公司。

1.2 方法

1.2.1 总 RNA 的提取及质量检测

取出冻存的淡玫红鹅膏样品, 采用 Trizol 法提取总 RNA, 提取步骤按照 Invitrogen 公司提供的 Trizol Reagent 说明书进行。用 Agilent 2100 生物分析仪对 RNA 样品的浓度、28S/18S、RNA 完整值(RNA Integrity Number, RIN)/RNA 质量值(RNA Quality Number, RQN)进行检测; 用 NanoDrop 进行 $OD_{260/280}$ 检测。

1.2.2 构建 cDNA 文库及测序

用 mRNA 富集法对总 RNA 进行处理, 将带有 OligodT 的磁珠富集有 PolyA 尾的 mRNA, 使用打断缓冲液进行 mRNA 片段化, 用随机 N6 引物(Random Hexamers)进行反转录, 合成 cDNA 第二链形成双链 DNA, 末端补平, 5'端磷酸化, 3'端形成突出一个“A”的黏末端, 连接 3'端有凸出一个“T”的鼓泡状的接头, 通过特异引物进行 PCR 扩增, 热变性成单链, 用桥式引物将单链 DNA 环化得到单链环状 DNA 文库, 用 Illumina HiSeq 2000 序列分析仪进行测序。

1.2.3 测序原始数据处理及从头组装

测序得到的原始图像数据经碱基调取转化为原始序列数据, 即为 Raw Reads。使用软件 SOAPnuke 进行统计和 Trimmomatic 进行过滤, 去除包含接头的原始序列、未知碱基 N 含量大于 5% 的原始序列、质量值低于 10 的碱基占比大于 20% 的原始序列, 过滤后获得整洁数据, 通过无参考基因组短序列组装软件 Trinity^[16]进行从头组装。

1.2.4 转录组数据基本功能注释

在 7 大数据库中对照淡玫红鹅膏转录组的 Unigene 库进行基本功能注释, 包括 Kyoto Encyclopedia of Genes and Genomes (KEGG)注释、Cluster of Orthologous Groups of Proteins (COG)功能注释、Gene Ontology (GO)功能注释、Non-Redundant Protein Database (NR)、Nucleotide Sequence Database (NT)、SwissProt 以及 Interpro 注释等。通过 BLASTx 将 Unigene 库序列比对到 7 大蛋白数据库(E Value $<10^{-5}$), 以序列匹配度最高(E 值最低)的蛋白注释相应的 Unigene 库的蛋白。

1.2.5 MSDIN 基因家族检索及 PCR 验证

通过 TBLASTn 软件(<ftp://ftp.ncbi.nlm.nih.gov/blast/executables/blast+/2.9.0/>)在 Unigene 库编码的氨基酸序列中进行搜索, 得到符合条件的多肽氨基酸序列及 Unigene 编号, 由 Unigene 编号获得相应的 DNA 编码序列(Coding DNA Sequence, CDS)序列, 根据 CDS 序列在两边设计特异性引物(表 1)。挑取 100 mg 保存在 -80°C 超低温冰箱的淡玫红鹅膏子实体, 根据全能型植物 RNA 提取试剂盒(DNase I)的操作说明书进行 RNA 的提取, 利用 cDNA 反转录试剂盒将 RNA 反转录成 cDNA。采用 MSDIN 基因特异引物, 以淡玫红鹅膏 cDNA 为模板进行 PCR 扩增和测序, 验证 MSDIN 家族基因序列的正确性。PCR 反应体系($25\ \mu\text{L}$): cDNA 模板($200\ \text{ng}/\mu\text{L}$) $1\ \mu\text{L}$, 上、下游引物($10\ \mu\text{mol}/\text{L}$)各 $1\ \mu\text{L}$, PCR 聚合酶($1.25\ \text{U}/\mu\text{L}$) $0.5\ \mu\text{L}$, $2\times\text{Mix}$ $12.5\ \mu\text{L}$, ddH_2O $9\ \mu\text{L}$ 。PCR 反应条件(采用 Touchdown 方法): 98°C 3 min; 98°C 25 s, 65°C 20 s, 72°C 25 s,

8个循环; 98 °C 30 s, 58 °C 20 s, 72 °C 30 s, 18个循环; 98 °C 30 s, 50 °C 20 s, 72 °C 30 s, 18个循环; 72 °C 10 min。PCR产物经切胶回收后连接到 pMD™19-T 载体, 转化到大肠杆菌感受态细胞 DH5α 中, 挑选 3–5 个转化子送北京六合华大基因科技有限公司进行测序。

1.2.6 MSDIN 的系统发育树构建

如表 2 所示, 共收集了 79 条 MSDIN 基因的前体肽序列, 其中 36 条序列来源于 NCBI 数据库, 8 条来源于文献[11-12], 35 条来源于本研究。将所有 FASTQ^[17]格式的氨基酸序列转换成

FASTA 格式导入 MEGA 7.0.14^[18]软件中, 使用 ClustalW 软件进行序列比对, 采用最大似然法 (Maximum Likelihood, ML) 建树分析, Bootstrap 设置 1 200, 获得的系统树保存为 nwk 格式文件, 然后将 nwk 格式文件上传到 iTOL 网站上 (<https://itol.embl.de/>)。MEGA 中提供的替代模型测试选择 Jones-Taylor-Thornton 作为本次建树的最佳进化模型, 位点之间采用 Gamma 分布, 含有 Gaps 或缺失数据高于 5% 的位点则选择部分删除 (位点覆盖截止为 95%), 进化树上只显示高于 50% 的支持率。

表 1 淡玫红鹅膏转录组 MSDIN 基因序列验证的特异性引物

Table 1 The specific primers used in the validation of the MSDIN genes from *Amanita pallidorozea* transcriptome

多肽名 Peptides name	正向引物 Forward primers (5'→3')	反向引物 Reverse primers (5'→3')
A.pa 1 α-amanitin	ACGGCTTCGTGACATAACAGC	TGGTTCCGCATTCACCTTCCTA
A.pa 1 β-amanitin	TTTCAAACCTCCAAGAACCTC	TTTCAAACCTCCAAGAACCTC
A.pa “Phallotoxin”	CACCAAATGCTTTCCCACTG	TGTCTCCACCGACCGTTTAT
A.pa 1 Phalloidin	ATGGAATCATTACGTGGGT	CTCCGTGTCAGAACAAGATACTC
MSDIN-1	CCACCCAGATAGTAAGCGTCCAT	AGAAAGGCTACAACACGCGCAC
MSDIN-2	TTACGCCAATCTTTACGAAC	CTTGACCACGAACATAGTCCC
MSDIN-3	AATCCCAACTTGAATCCAAAGA	CAACGGAGTCAGTATGGTAGCC
MSDIN-4	TTTCAAACCTCCAAGAACCTC	GACTAAAGCACGCAGCACATA
MSDIN-5	TCAAAGATCCAAGCCTGATA	AGTTATTTTCGCTTCCTCCAT
MSDIN-6	GGACCGATAGACCGTGCATTTA	CCCTCCGACCTCAGAAACCTAC
MSDIN-8	GTAAGCCGATCATTAAACATG	CTTGACTGGGACAGATAAAA
MSDIN-9	CCCCAAACTTGAAACCTACA	TCAATGGTCCGAGCCTGATA
MSDIN-10	GCGTTGATAAGGTATTCCAGC	ATCCATGAACCAGGTATTTAG
MSDIN-11	CTCCCAACTCGAATCCAAGACC	GATGAAAGAAAGCGGCAAATCC
MSDIN-13	GGTCACGCTTGAGATCAT	AACGGATTCAACATAAAA
MSDIN-14	ACAGATAAAAAGGCGTCGGGGT	CGATGTAGGCTCAATGGTCCA
MSDIN-15	ATGAGTGGACTGATAGATCGTG	GAATCCTTCAGAAACCTACAGA
MSDIN-16	CCCTCACAATGTCTGACAT	R: CATAAGAGTGGACCGATAG
MSDIN-17	CAAGCCTCATATCGTGAACAA	GGGACGGTACAATATCCAGTT
MSDIN-18	GAAAACACACCACAAACAT	CTGGAACAGATAAAAAGGC
MSDIN-20	CCATCAGCTCTTCTACCCATA	CGTCAACAACAGTGTCCACTA
MSDIN-21	GTAGGACGACAGTGTAGCA	AACAGAGAACGAGAAAGAA
MSDIN-22	AAAAAAAGATAAAGAACA	GACTAGGACGTATAAAAG
MSDIN-23	CTTCCGAACCTGAATCCCTC	CGCATCACTAAACCCTCTTG
MSDIN-24	ACTCTCGCCACGAGTGAGGGT	TCTGCAATGGGCACGCTCAAG
MSDIN-25	GTATTTCAATGGCGCTCTACC	ATCCTACGGCTTCCTTCACTC
MSDIN-26	AATAGTGACTCAACTCCCTTTC	AACCGTTGTAACTGGACTGAA
MSDIN-27	ATGAAAAAAGAAGCTCGAAACC	TGAATCTCTCAGAAGCCTACAG
MSDIN-28	TGTGAGATGACAACCTTTGAATGGAT	CAGAAGCCTACAGCTACCTTTAC
MSDIN-29	GTCCAAGGGCTCGTCTGAAAT	GCCATGACCGGGACAGATAAA

表 2 用于系统发育分析的 MSDIN 前体肽序列信息表

Table 2 The information of MSDIN precursor peptide sequences used in phylogenetic analyses

前体肽序列 Peptide sequences	多肽名称 Peptide names	物种名 Species	序列来源 Sequences source	缩写 Abbreviations
MSDINATRLPIWGIGCNPCVGDEVAAILTRGEALC	α -amanitin	<i>A. pallidrosea</i>	本研究 This study	A.pa 1 α -AMA
MSDINATRLPIWGIGCNPCVGDEVAALLTRGEALC	α -amanitin	<i>A. pallidrosea</i>	本研究 This study	A.pa 2 α -AMA
MSDINATRLPIWGIGCNPCVGDEVTALITRGEALC	α -amanitin	<i>A. pallidrosea</i>	本研究 This study	A.pa 3 α -AMA
MSDINATRLPIWGIGCNPCVGDEVTALITRGEALC	α -amanitin	<i>A. pallidrosea</i>	MN264221	A.pa 4 α -AMA
MSDINATRLPIWGIGCNPCVGDEVAALLTRGEALC	α -amanitin	<i>A. pallidrosea</i>	MN264222	A.pa 5 α -AMA
MSDINATRLPIWGIGCNPCVGDDVTLLTRGEALC	α -amanitin	<i>A. bisporigera</i>	[11]	A.b α -AMA
MSDINATRLPIWGIGCNPCVGDEVAALLTRGEALC	α -amanitin	<i>A. phalloides</i>	[11]	A.ph α -AMA
MSDINATRLPIWGIGCNPCVGDDVTSVLTRGEALC	α -amanitin	<i>A. exitilis</i>	MN264220	A.e α -AMA
MSDINATRLPIWGIGCNPSVGDEVTALLTSGEALC	α -amanitin	<i>A. fuliginea</i>	MN318174	A.f α -AMA
MSDINATRLPIWGIGCNPCVGDEVTALLTRGEALC	α -amanitin	<i>A. fuligineoides</i>	AHB18723	A.fs α -AMA
MSDINATRLPIWGIGCNPSVGDEVTALLASGEALC	α -amanitin	<i>A. rimosa</i>	MN318183	A.r 1 α -AMA
MSDINSTRLPWGIGCNPSVGDEVTALLTRGEALC	α -amanitin	<i>A. rimosa</i>	MN264223	A.r 2 α -AMA
MSDINATRLPIWGIGCNPCVGDEVAALLTRGEALC	α -amanitin	<i>A. subfuliginea</i>	MN318187	A.sf α -AMA
MSDINATRLPIWGIGCNPCVGDEVAALLTRGEALC	α -amanitin	<i>A. subjunquillea</i>	MH142175	A.sj α -AMA
MSDINATRLPIWGIGCNPSVGDEATALLTRGEALC	α -amanitin	<i>A. virosa</i>	MN272412	A.v α -AMA
MSDINATRLPIWGIGDCPCVGDDVTAVLTRGE	β -amanitin	<i>A. pallidrosea</i>	本研究 This study	A.pa 1 β -AMA
MSDINATRLPIWGIGDCPCVGDDVTAVLTCGEALC	β -amanitin	<i>A. pallidrosea</i>	MN318179	A.pa 2 β -AMA
MSDINATRLPIWGIGDCPCVGDDVTAVLTRGEALC	β -amanitin	<i>A. pallidrosea</i>	MN264228	A.pa 3 β -AMA
MSDINATRLPIWGIGDCPCVGDEVTALLTRGEALC	β -amanitin	<i>A. phalloides</i>	[11]	A.ph 1 β -AMA
MSDINATRLPIWGIGDCPCVGDDVTALLTRGEALC	β -amanitin	<i>A. phalloides</i>	[11]	A.ph 2 β -AMA
MSDINATRLPIWGIGDCPCVGDDVTALLTRGEALC	β -amanitin	<i>A. exitilis</i>	AGW83701	A.e β -AMA
MSDINATRLPIWGIGDCPCVGDDVTALITRGEALC	β -amanitin	<i>A. fuliginea</i>	MN264226	A.f 1 β -AMA
MSDINATRLPIWGIGDCPCVGEDVAALITRGEALC	β -amanitin	<i>A. fuliginea</i>	MN318173	A.f 2 β -AMA
MSDINATRLPIWGIGDCPCVGDDVTALLTRGEALC	β -amanitin	<i>A. molliuscula</i>	MN264227	A.m β -AMA
MSDINATRLPIWGIGDCPCVGDDVAALATRGEALC	β -amanitin	<i>A. rimosa</i>	MN318182	A.r 1 β -AMA
MSDINATRLPIWGIGDCPCVGDDVAALITRGEALC	β -amanitin	<i>A. rimosa</i>	AHB18720	A.r 2 β -AMA
MSDINATRLPIWGIGDCPCVGDEVTALLTRGEALC	β -amanitin	<i>A. fuligineoides</i>	AHB18717	A.fs β -AMA
MSDINATRLPIWGIGDCPCVSDDVAALLTRGEALC	β -amanitin	<i>A. subfuliginea</i>	MN318186	A.sf β -AMA
MSDINATRLPIWGIGDCPCIGDDVTALLTRGEALC	β -amanitin	<i>A. subjunquillea</i>	MH142177	A.sj 1 β -AMA
MSDINATRLPIWGIGDCPCVGDEVTALLTRGEALC	β -amanitin	<i>A. subjunquillea</i>	MH142176	A.sj 2 β -AMA
MSDINATRLPIWGIGDCPCVGDDVA AFLTRGEALC	β -amanitin	<i>A. subpallidrosea</i>	MN264230	A.sp β -AMA
MSDVNATRLPAWLVDPCPCVGDDINRLLTRGENLC	Phallacidin	<i>A. pallidrosea</i>	本研究 This study	A.pa 1 PHA
MSDVNATRLPAWLVDPCPCVGDDINRLLTRGENLC	Phallacidin	<i>A. pallidrosea</i>	MN264233	A.pa 2 PHA
MSDINATRLPAWLVDPCPCVGDDVNRLLTRGESLC	Phallacidin	<i>A. bisporigera</i>	ABW87771	A.b PCD
MSDINATRLPAWLVDPCPCVGDDINRLLTRGENLC	Phallacidin	<i>A. phalloides</i>	[11]	A.ph PCD
MSDINATRLPAWLVDPCPCVGDDVNRLLTRGESLC	Phallacidin	<i>A. exitilis</i>	MN264231	A.e PCD
MSDINATRLPAWLVDPCPCVGDDVNRLLRGENLC	Phallacidin	<i>A. fuliginea</i>	MN264232	A.f PCD
MSDINATRVPAWLVDPCPCVGDDISRLLTRGENLC	Phallacidin	<i>A. rimosa</i>	[12]	A.r PCD
MSDINATRLPAWLVDPCPCVGDDVNRLITRGENLC	Phallacidin	<i>A. subfuliginea</i>	MN318188	A.sf PCD
MSDINATRLPAWLVDPCPCVGDDINRLLTRGENLC	Phallacidin	<i>A. subjunquillea</i>	MN264234	A.sj PCD

(待续)

(续表 2)

MSDINATRLPAWLVDPCVGGDDINRLLTRGENLC	Phallacidin	<i>A. virosa</i>	MN318199	A.v PCD
MSDINTTCLPAWLATCPCVTGDDVNPTLTCGESLC	Phalloidin	<i>A. phalloides</i>	[11]	A.ph 1 PHD
MSDINASRLPAWLATCPCVGGDDVNPTLSRGESLC	Phalloidin	<i>A. phalloides</i>	[11]	A.ph 2 PHD
MSDINASRLPAWLATCPCIGDDVNPTITRGESLC	Phalloidin	<i>A. fuliginea</i>	MN264249	A.f PHD
MSDINASRLPAWLATCPCAGDDVNPTLTRGESLC	Phalloidin	<i>A. pallidorozea</i>	MN318180	A.pa PHD
MSDINATRLPAWLATCPCVGGDDVNPTLSRGESLC	Phalloidin	<i>A. subjunquillea</i>	MN318191	A.sj 1 PHD
MSDINATRLPAWLATCPCVGGDDVNPTLTRGESLC	Phalloidin	<i>A. subjunquillea</i>	MN264250	A.sj 2 PHD
MSDINASRFPALWATCPCVGGDDVNPTIARGESLC	Phalloidin	<i>A. subpallidorozea</i>	MN272407	A.sp PHD
MSDINATRLPAWLATCPCVGGDDVNPTLARGESLC	Phalloidin	<i>A. virosa</i>	MN318198	A.v PHD
MSDINATCLPAWLVTCPVGGDDVNPTLTRGESLC	“Phallotoxin”	<i>A. pallidorozea</i>	本研究 This study	A.pa “phallotoxin”
MSDINATRLPLGRPELPCVGGDDVNYILVSG	未知功能的环肽 Cyclopeptide of unknown function	<i>A. pallidorozea</i>	本研究 This study	MSDIN-1
MSDINAIRAPLPIFSLNPCVGGDDVEVLRSEGLS	未知功能的环肽 Cyclopeptide of unknown function	<i>A. pallidorozea</i>	本研究 This study	MSDIN-2
MSDINAARLPNLFVWIPPCISDDINSTLTRGESLC	未知功能的环肽 Cyclopeptide of unknown function	<i>A. pallidorozea</i>	本研究 This study	MSDIN-3
MSDINATRLPHFPPLGLQPCAGDVNDLTLFRGE	未知功能的环肽 Cyclopeptide of unknown function	<i>A. pallidorozea</i>	本研究 This study	MSDIN-4
MSDINTARLPVFFMPPFIPPCVSDDIQMVLTTRGENLC	未知功能的环肽 Cyclopeptide of unknown function	<i>A. pallidorozea</i>	本研究 This study	MSDIN-5
MSDINTTRLPYMWDHHLPCASDDIQMVFTTRGENLC	未知功能的环肽 Cyclopeptide of unknown function	<i>A. pallidorozea</i>	本研究 This study	MSDIN-6
MSDINATRLPNWHAGPTRPPCVADDVSLTLTRGESLC	未知功能的环肽 Cyclopeptide of unknown function	<i>A. pallidorozea</i>	本研究 This study	MSDIN-7
MADINASRLPLNILPFHLPPCVSDDATSTLTRGESLC	未知功能的环肽 Cyclopeptide of unknown function	<i>A. pallidorozea</i>	本研究 This study	MSDIN-8
MYDINTTRLPHFFNLTPPCVGGDDIEMVLTTRGESLC	未知功能的环肽 Cyclopeptide of unknown function	<i>A. pallidorozea</i>	本研究 This study	MSDIN-9
MSDLNATRLPFNLFRFPYPCIGDDSGSVLTLGEGLC	未知功能的环肽 Cyclopeptide of unknown function	<i>A. pallidorozea</i>	本研究 This study	MSDIN-10
MSDINASRLPFFPEVGFPCVGGDDTNPILTRGGSL	未知功能的环肽 Cyclopeptide of unknown function	<i>A. pallidorozea</i>	本研究 This study	MSDIN-11
MSDINAARLPVYMILFPSVGGDDIDVVLGRGENLC	未知功能的环肽 Cyclopeptide of unknown function	<i>A. pallidorozea</i>	本研究 This study	MSDIN-12

(待续)

(续表 2)

MSDINTARLPMHILAPPPCVSDDIEMVLTRGESLC	未知功能的环肽 Cyclopeptide of unknown function	<i>A. pallidorozea</i>	本研究 This study	MSDIN-13
MSDINATRLPDPRLIPGGSDDVDSALTRGESLC	未知功能的环肽 Cyclopeptide of unknown function	<i>A. pallidorozea</i>	本研究 This study	MSDIN-14
MSDMNVVRLPISDPTAYPCVGDDIQAVLGRGESLC	未知功能的环肽 Cyclopeptide of unknown function	<i>A. pallidorozea</i>	本研究 This study	MSDIN-15
MSDINTARLPEFIVFGIFPCVGDDIQTVLTRGEDLC	未知功能的环肽 Cyclopeptide of unknown function	<i>A. pallidorozea</i>	本研究 This study	MSDIN-16
MSDINATRLPHQHHPFVPGGSDDVGSTLTRGESLC	未知功能的环肽 Cyclopeptide of unknown function	<i>A. pallidorozea</i>	本研究 This study	MSDIN-17
MSDVNATRLPMAFPEFLACVGDDVNHTLTRGENLC	未知功能的环肽 Cyclopeptide of unknown function	<i>A. pallidorozea</i>	本研究 This study	MSDIN-18
MSDINTARLPLLFMPPFIPPCVSDDIQMVLTRGENLC	未知功能的环肽 Cyclopeptide of unknown function	<i>A. pallidorozea</i>	本研究 This study	MSDIN-19
MSDINTARLPFVIIPPFIFPCVSDDIEMVLTRGEGLC	未知功能的环肽 Cyclopeptide of unknown function	<i>A. pallidorozea</i>	本研究 This study	MSDIN-20
MSDINTARLPFDFIPSSLPPCVSDDIQMVLTRGEGLC	未知功能的环肽 Cyclopeptide of unknown function	<i>A. pallidorozea</i>	本研究 This study	MSDIN-21
MSDINTARLPLGILPPIPPCVSDDIHMVLTRGEDLC	未知功能的环肽 Cyclopeptide of unknown function	<i>A. pallidorozea</i>	本研究 This study	MSDIN-22
MSDINTIRVPFPWTGPFVPCVGDDVGSVLTHGESL	未知功能的环肽 Cyclopeptide of unknown function	<i>A. pallidorozea</i>	本研究 This study	MSDIN-23
MSDINASRLPYFFGFTPSVGDDVNPTLTRGES	未知功能的环肽 Cyclopeptide of unknown function	<i>A. pallidorozea</i>	本研究 This study	MSDIN-24
MSDINAARLPLIYMILFSPVGDDIDVVLGRGENLC	未知功能的环肽 Cyclopeptide of unknown function	<i>A. pallidorozea</i>	本研究 This study	MSDIN-25
MSETNAARLPTIHLFSAPSVGDDIEVVLGRGENLC	未知功能的环肽 Cyclopeptide of unknown function	<i>A. pallidorozea</i>	本研究 This study	MSDIN-26
MSDINANRLPLVSIGSIPCVGDDAAANILRSCESIC	未知功能的环肽 Cyclopeptide of unknown function	<i>A. pallidorozea</i>	本研究 This study	MSDIN-27
MSDINANRLPFVPFSFIPICIGDDAAASILRSCESIC	未知功能的环肽 Cyclopeptide of unknown function	<i>A. pallidorozea</i>	本研究 This study	MSDIN-28
MSDMNIACLPIFIVFIIPPCVSDDIQTVLTRGEKLC	未知功能的环肽 Cyclopeptide of unknown function	<i>A. pallidorozea</i>	本研究 This study	MSDIN-29

2 结果与分析

2.1 RNA 样品的浓度、纯度及完整性

根据武汉华大医学检验所有限公司对构建 cDNA 文库的要求, 合格 RNA 样品应达到的指标有 4 点: RNA 浓度 $>40\text{ ng}/\mu\text{L}$ 、 $OD_{260/280}>2.0$ 、 $RIN/RQN>6.5$ 和 $28S/18S>1.0$ 。RNA 样品用 Agilent 2100 生物分析仪对 RNA 样品的浓度、28S/18S 和 RIN/RQN 进行检测, 其值分别为 $316\text{ ng}/\mu\text{L}$ 、1.5 和 8.0; 用 NanoDrop 进行 $OD_{260/280}$ 检测, 测得 $OD_{260/280}$ 为 2.13, 以上指标均符合构建 cDNA 文库的要求。

2.2 淡玫红鹅膏转录组测序与 *De novo* 组装

对淡玫红鹅膏进行转录组测序, 共获得 30.65 Mb 的高质量 Read, Q20 为 97.47%。使用 Trinity 软件对 Read 进行从头组装, 最终得到 30 497 个 Unigene, 总长度为 50 362 992 bp, 平均长度为 1 651 bp, N50 长度为 2 532 bp。

2.3 淡玫红鹅膏转录组 Unigene 的基本功能注释

组装完毕后, 使用 BLASTx^[19]对 Unigene 进行 NT、NR、COG、KEGG 及 SwissProt 注释, 使用 BLAST2GO^[20]及 NR 注释结果进行 GO 注释, 使用 InterProScan^[21]进行 InterPro 注释, 注释结果显示共有 30 497 个 Unigene, 其中有 22 571 个 Unigene 得到了注释, 将 Unigene 对比到 NR 数据库中有 21 743 个 Unigene 得到了注释, 其占比 71.30%; 对比到 NT 数据库中有 8 666 个基因得到了注释, 占比 28.42%; 对比到 SwissProt 数据库中有 14 571 个 Unigene 得到了注释, 占比 47.78%; 对比到 KEGG 数据库中有 15 567 个 Unigene 得到了注释, 占比 51.04%; 对比到 COG 数据库中有 11 100 个 Unigene 得到了注释, 占比 36.40%; 对比到 InterPro 数据库中有 14 272 个 Unigene 得到了注释, 占比 46.80%; 对比到 GO 数据库中有 6 022 个 Unigene 得到了注释, 占比 19.75%; 将所有比对上 NR 数据库的 Unigene, 利用 BLAST2GO 软件获得 GO 注释信息, 并统计注释到 GO 一级功能群 3 个类: 生物过程(Biological Process)、细

胞组成(Cellular Component)、分子功能(Molecular Function) (图 3), 生物过程类群包括 19 个二级类群, 其中代谢过程、细胞过程和单一生物过程这 3 个类群包含的 Unigene 数较多, 复制、发育过程、生殖过程、多组织过程、多细胞生物过程、运动、生长和生物黏附这 8 个类群中包括的 Unigene 数较少。细胞组成类群包括 11 个二级类群, 其中细胞和细胞部分这 2 个类群中含有的 Unigene 数最多, 拟核、胞外区、胞外区部分这 3 个类群包含的 Unigene 数最少。分子功能类群包括 14 个二级类群, 其中催化活性、结合这 2 个类群中被注释的 Unigene 数最多, 蛋白质结合转录因子活动、受体活性、鸟苷酸交换因子活性、金属伴侣活性和蛋白质标签这 5 个类群被注释到的 Unigene 数最少。

统计注释上 KEGG 数据库第一水平层级和第二水平层级的 Unigene, 如图 4 所示, 第一水平层级分为细胞进程、环境信息处理、遗传信息处理、人类疾病和新陈代谢 5 大类。在第二水平层级上, Unigene 有 21 种功能被注释到, 细胞进程包括细胞的生长与死亡、运输和分解代谢这 2 个二级类群, 分别有 733 个和 1 324 个 Unigene; 环境信息处理包括膜运输和信号转导, 分别包括了 70 个和 672 个 Unigene; 遗传信息处理包括折叠、分类和降解、复制和修复、转录、翻译, 共 4 个二级类群; 人类疾病包括抗药性、内分泌和代谢疾病, 注释到这 2 个类群的 Unigene 个数极少, 仅有 2 个和 28 个; 新陈代谢包括氨基酸代谢、其他次生代谢产物的生物合成、碳水化合物代谢、能量代谢、全局和总览图、多糖生物合成与代谢、脂质代谢、辅助因子和维生素的代谢、其他氨基酸的代谢、萜类化合物和聚酮类化合物的代谢、核苷酸代谢, 共 11 个二级类群, 其中全局和总览图包含的 Unigene 数量最多, 有 4 453 个, 萜类化合物和聚酮类化合物的代谢这一类群包含的 Unigene 数量最少, 仅有 142 个。

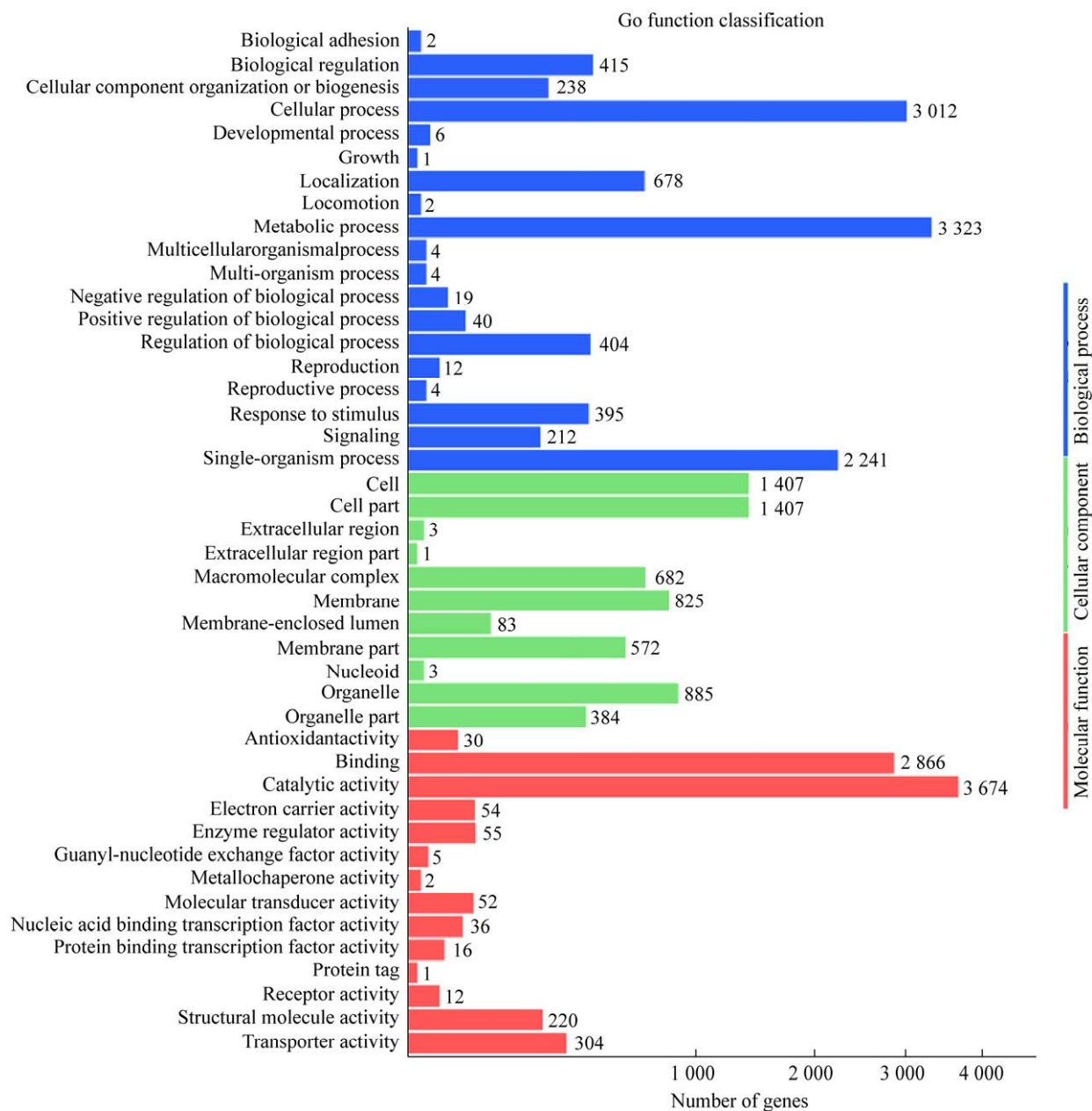


图3 淡玫红鹅膏转录组的GO功能分类

Figure 3 GO functional classification of *Amanita pallidrosea* transcriptome

注：红色：分子功能；绿色：细胞组成；蓝色：生物进程

Note: Red: Molecular function; Green: Cell component; Blue: Biological process

2.4 MSDIN 基因家族多样性分析

利用 TBLASTn 工具手动检索转录组数据中 MSDIN 基因，共检索到 60 条相关序列，可编码 32 条不同环肽序列(表3)，所有序列均设计特异性引物并进行了 PCR 验证。结果显示，淡玫红鹅膏

MSDIN 家族编码了 32 种环肽，包括 3 种已知毒肽： α -鹅膏毒肽(IWGIGCNP)、 β -鹅膏毒肽(IWGIGCDP)和羧基二羟鬼笔毒肽(AWLVDGP)；29 种未知环肽，其中包含 21 种已报道的肽类序列和 8 种未报道的序列。研究发现编码新肽的核心

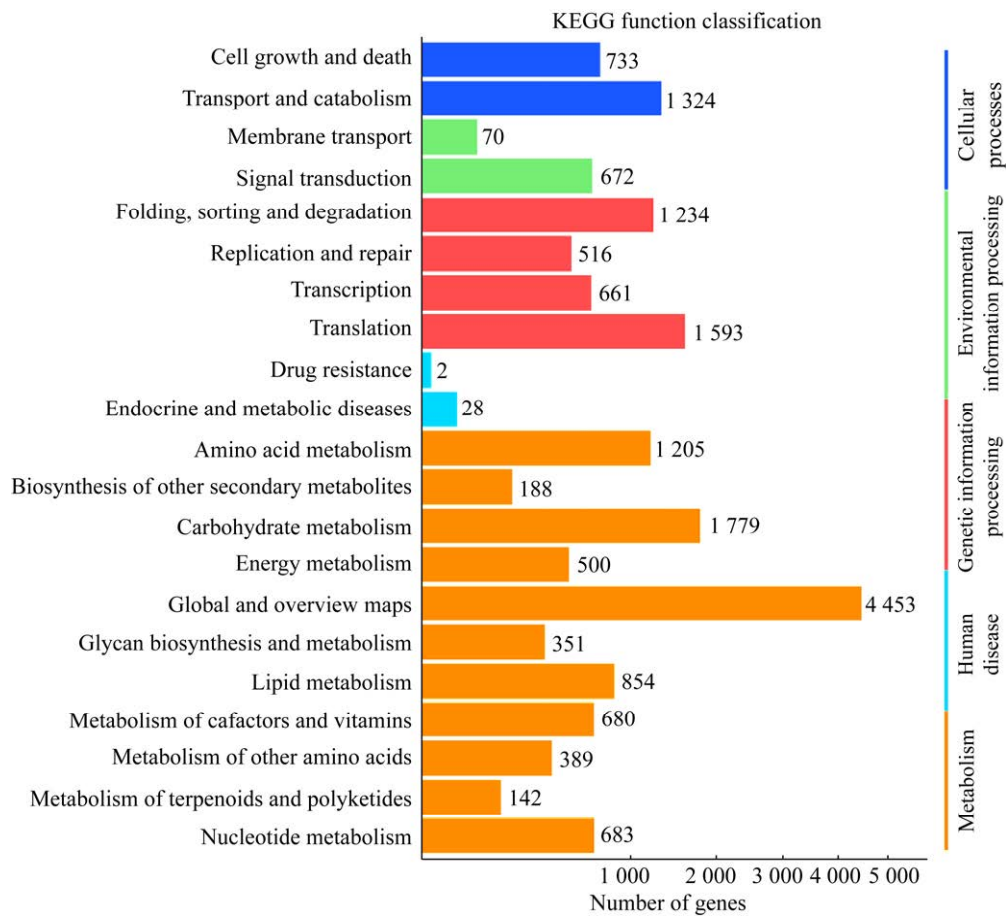


图 4 淡玫红鹅膏转录组的 KEGG 功能分类
Figure 4 KEGG functional classification of *Amanita pallidorozea* transcriptome

注: 蓝色: 细胞进程; 绿色: 环境信息处理; 红色: 遗传信息处理; 浅蓝色: 人类疾病; 橙红色: 新陈代谢
Note: Blue: Cellular processes; Green: Environmental information processing; Red: Genetic information processing; Light blue: Human disease; Orange red: Metabolism

序列包括 1 种七肽“YFFGFTP”、3 种八肽“FVPFSFIP”“LIYMILFP”“LVSIGSIP”、1 种九肽“IFIVFPIPP”, 以及 3 种十肽“FDFIPSSLPP”“LGILPPIPPP”“LLFMPPFIPP”。

2.5 MSDIN 多肽序列保守性分析

2.5.1 淡玫红鹅膏 MSDIN 前体肽序列保守性分析
将得到的 35 条 MSDIN 基因的前体肽序列进行 WebLogo (<http://weblogo.berkeley.edu/logo.cgi>) 比对, 分析各位点的氨基酸相对保守程度 (图 5)。经统计分析, 前体肽序列可划分为上游保守区(1-10 号位点)、核心肽区(11-20 号位点)和下游保守区(21-37 号位点)。上游保守序列有

14 种不同的类型, 分别为“MSDINATRLP”“MSDINAIRAP”“MSDINTARLP”“MSDINAARLP”“MSDVNATRLP”“MSDINATCLP”“MSDINTTRLP”“MADINASRLP”“MYDINTTRLP”“MSDINTIRVP”“MSDLNATRLP”“MSDMNVVRLP”“MSDINANRLP”和“MSDMNIACLP”。在上游保守区中的 M、N、P 这 3 个氨基酸的保守性最高, 达到了 100%。第 2 号位点上保守的氨基酸为 S, 其保守度为 94.2% (33/35)。第 3 号位点上保守的氨基酸为 D, 保守度为 97.1% (34/35)。第 4 号位点保守的氨基酸为 I, 保守度为 82.8% (29/35)。第 6 号位点上氨基酸 A 和 T 较为保守, 保守度分别为 65.7% (23/35) 和 28.5% (10/35)。第 7 号位点上氨基酸 T 和 A 较

表 3 淡玫红鹅膏 MSDIN 基因家族编码的前体肽

Table 3 Precursor peptides encoded by MSDIN gene family in *Amanita pallidrosea*

编号 No.	上游保守序列 Upstream conserved sequence	核心肽序列 Core polypeptide	下游保守序列 Downstream conservative sequence	名称 Name	备注 Notes
A.pa-1	MSDINATRLP	IWGICNP	CVGDEVAAILTRGEALC*	A.pa 1 α -amanitin	未报道 Not reported
A.pa-2	MSDINATRLP	IWGICNP	CVGDEVAALLTRGEALC*	A.pa 2 α -amanitin	[22]
A.pa-3	MSDINATRLP	IWGICNP	CVGDEVTALITRGEALC*	A.pa 3 α -amanitin	[15]
A.pa-4	MSDINATRLP	IWGICDP	CVGDDVTAVLTRGE	A.pa 1 β -amanitin	[23]
A.pa-5	MSDINATCLP	AWLVTCP	CVGDDVNPTLTRGESLC*	A.pa “Phallotoxin”	[23]
A.pa-6	MSDVNATRLP	AWLVDCP	CVGDDINRLLTRGENLC*	A.pa 1 Phallacidin	[15]
A.pa-7	MSDINATRLP	LGRPESLP	CVGDDVNYILVSG	MSDIN-1	[11,13,23]
A.pa-8	MSDINAIRAP	LPIFSLNP	CVGDDVEVLRRSEGLS*	MSDIN-2	[13]
A.pa-9	MSDINAARLP	NLFVWIPP	CISDDINSTLTRGESLC*	MSDIN-3	[13,23]
A.pa-10	MSDINATRLP	HPFPLGLQP	CAGDVDNLTFRGE	MSDIN-4	[12-13,23]
A.pa-11	MSDINTARLP	VFFMPPFIPP	CVSDDIQMVLTRGENLC*	MSDIN-5	[13,23]
A.pa-12	MSDINTTRLP	YMWDHHLPP	CASDDIQMVFTTRGENLC*	MSDIN-6	[23]
A.pa-13	MSDINATRLP	NWHAGPTRPP	CVADDVSLTLTRGESLC*	MSDIN-7	[13,23]
A.pa-14	MADINASRLP	LNILPFHLPP	CVSDDATSTLTRGESLC*	MSDIN-8	[13,23]
A.pa-15	MYDINTTRLP	HFFNLTPP	CVGDDIEMVLTRGESLC*	MSDIN-9	[13]
A.pa-16	MSDLNATRLP	FNLFRFPYP	CIGDDSGSVLTLGEGLC*	MSDIN-10	[13,23]
A.pa-17	MSDINASRLP	FFPEVGFFP	CVGDDTNPILTRGGSL	MSDIN-11	[13,23]
A.pa-18	MSDINAARLP	LVYMILFP	SVGDDIDVVLGRGENLC*	MSDIN-12	[13,23]
A.pa-19	MSDINTARLP	MHILAPPP	CVSDDIEMVLTRGESLC*	MSDIN-13	[13-14,23]
A.pa-20	MSDINATRLP	DPRLLIP	GGSDDVDSALTRGESLC*	MSDIN-14	[23]
A.pa-21	MSDMNVVRLP	ISDPTAYP	CVGDDIQAVLGRGESLC*	MSDIN-15	[11-13,23]
A.pa-22	MSDINTARLP	EFIVFGIFP	CVGDDIQTVLTRGEDLC*	MSDIN-16	[13,23]
A.pa-23	MSDINATRLP	HQHHPFVP	GGSDDVGSTLTRGESLC*	MSDIN-17	[23]
A.pa-24	MSDVNATRLP	MAFPEFLA	CVGDDVNHTLTRGE	MSDIN-18	[11,23]
A.pa-25	MSDINTARLP	FVIIPPFIFP	CVSDDIEMVLTRGEGLC*	MSDIN-19	[13-14]
A.pa-26	MSDINTIRVP	FPWTGPFVP	CVGDDVGSVLTHGESL	MSDIN-20	[14,23]
A.pa-27	MSETNAARLP	TIHLFSAP	SVGDDIEVVLGRGENLC*	MSDIN-21	[13]
A.pa-28	MSDINTARLP	LLFMPPFIPP	CVSDDIQMVLTRGENLC*	MSDIN-22	未报道 Not reported
A.pa-29	MSDINTARLP	FDFIPSSLPP	CVSDDIQMVLTRGEGLC*	MSDIN-23	未报道 Not reported
A.pa-30	MSDINTARLP	LGILPPIPPP	CVSDDIHMVLTRGEDLC*	MSDIN-24	未报道 Not reported
A.pa-31	MSDINASRLP	YFFGFTP	SVGDDVNPTLTRGES	MSDIN-25	未报道 Not reported
A.pa-32	MSDINAARLP	LIYMILFP	SVGDDIDVVLGRGENLC*	MSDIN-26	未报道 Not reported
A.pa-33	MSDINANRLP	LVSIGSIP	CVGDDAAANILRSCESIC*	MSDIN-27	未报道 Not reported
A.pa-34	MSDINANRLP	FVPFSFIP	CIGDDAAASILRSCESIC*	MSDIN-28	未报道 Not reported
A.pa-35	MSDMNIACLP	IFIVFPIPP	CVSDDIQTVLTRGEK*	MSDIN-29	未报道 Not reported

注：*：终止密码子；“Phallotoxin”：一种新的七肽，类似于 Phallotoxin 环肽，含有 Trp-Cys 结构^[23]

Note：*：Indicates a stop codon; “Phallotoxin”：A new heptapeptide, similar to phallotoxin cyclic peptide, contains Trp-Cys structure^[23]

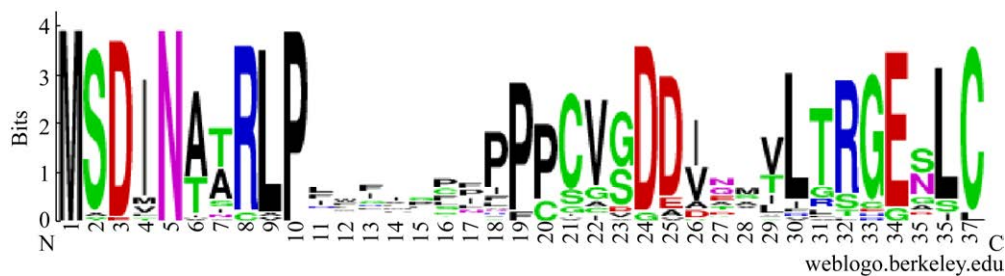


图 5 淡玫红鹅膏 35 个 MSDIN 家族成员的 WebLogo 比对
Figure 5 WebLogo alignment of 35 MSDIN family members in *Amanita pallidorosea*
注: 氨基酸的字母高度表示氨基酸的保守性程度
Note: The letter height of amino acids represents the degree of conservation

为保守, 保守度分别为 42.8% (15/35)和 34.2% (12/35)。第 8 号位点上保守的氨基酸为 R, 保守度为 94.2% (33/35)。第 9 号位点上保守的氨基酸为 L, 保守度为 94.2% (33/35)。核心区的序列呈现出多变性, 在核心区最长的序列由 10 个氨基酸组成, 最短的序列由 7 个氨基酸组成, 35 条序列中包括 3 个七肽、18 个八肽、7 个九肽和 7 个十肽。在下游保守区中第 24 号、25 号、30 号、32 号、33 号、34 号、36 号和 37 号位点的 D、D、L、R、G、E、L 和 C 氨基酸保守性较高, 其余位点保守性较低, 在第 27 号位点就出现了 N、Q、E、A、T、G、D 这 7 种氨基酸的可能。

2.5.2 剧毒鹅膏菌 α -鹅膏毒肽前体肽序列保守性分析

对 10 种已报道的剧毒鹅膏菌数据中收集到的 12 条 α -鹅膏毒肽的前体肽序列和本研究中的 3 条 α -鹅膏毒肽前体肽序列利用 WebLogo 网站比对进行保守性分析, 结果如图 6 所示, 在上游保守

区(1-10 位点)仅 6 号位点的氨基酸保守度不是 100%, 在该位点上出现了 2 种氨基酸分别为 A 和 S, 但 A 氨基酸保守度仍然较高, 为 93.3% (14/15)。在下游保守区(19-35 号位点), 主要在 23-30 号位点上保守性较低, 其保守性在 60.0%-86.6%; 除此之外, 19 号位点保守性也较低, 其 C 氨基酸的保守度为 73.3% (11/15)。

2.5.3 剧毒鹅膏菌 β -鹅膏毒肽前体肽序列保守性分析

对 16 条 β -鹅膏毒肽前体肽序列(15 条序列来源于已报道的 10 种剧毒鹅膏, 1 条来源于本研究)进行保守性分析, 结果如图 7 所示, 上游保守区和核心肽区的 18 个氨基酸保守性高达 100%, 在下游保守区中绝大部分氨基酸位点都有 2 种或 2 种以上的氨基酸变化, 其中 28 号位点最不保守, 有 4 种氨基酸的可能, 分别是 L、I、A 和 T, 其保守度分别为 75.0% (12/16)、12.5% (2/16)、6.2% (1/16)、6.2% (1/16)。



图 6 剧毒鹅膏菌 15 条 α -鹅膏毒肽序列的 WebLogo 比对
Figure 6 WebLogo alignment of 15 α -amanitin sequences in lethal *Amanita*

2.5.4 剧毒鹅膏菌羧基二羟鬼笔毒肽序列保守性分析

收集到 10 条羧基二羟鬼笔毒肽前体肽序列, 将序列导入至 WebLogo 网站, 得到比对结果如图 8 所示。所有序列上游保守区和核心区除 4 号和 9 号位点保守性分别为 80.0% 和 90.0% 外, 其他氨基酸的保守性为 100%; 下游保守区有 12 个位点的保守性为 100%, 第 23、24、27、28、32 这 5 个氨基酸位点的保守性较低, 分别为 60.0%、90.0%、90.0%、90.0% 和 80.0%。

2.6 淡玫红鹅膏 MSDIN 家族前体肽聚类分析

将 26 条鹅膏毒肽前体肽序列、8 条二羟鬼笔毒肽(PHD)前体肽序列、9 条羧基二羟鬼笔毒肽(PCD)前体肽序列, 以及本研究获得的淡玫红鹅膏 35 条 MSDIN 家族前体肽序列构建系统发育树(图 9)。结果显示, 系统树共分为 3 个大支: 鹅膏毒肽、鬼笔毒肽和未知功能鹅膏环肽。来源于不同剧毒鹅膏菌的 α -鹅膏毒肽、 β -鹅膏毒肽、MSDIN-27 和 MSDIN-28 聚在一起形成鹅膏毒肽分支。二羟鬼笔毒肽、羧基二羟鬼笔毒肽, 以及

MSDIN-1、MSDIN-4、MSDIN-11、MSDIN-18 和 MSDIN-24 聚为鬼笔毒肽分支; 其中未知功能的鹅膏环肽 MSDIN-1、MSDIN-4 和 MSDIN-18 与羧基二羟鬼笔毒肽聚成了一个小支; MSDIN-11、MSDIN-24、A.pa-5 和二羟鬼笔毒肽聚集在一个小分支内, 其他未知肽类序列聚为一个分支。

3 讨论与结论

在淡玫红鹅膏中已报道 MSDIN 基因有 32 条不同核心肽的序列^[13,23], 本研究新增了未报道的 8 条核心肽序列。总计 40 条不同的核心肽序列中, 4 条为已知毒肽: α -鹅膏毒肽、 β -鹅膏毒肽、羧基二羟鬼笔毒肽和二羟鬼笔毒肽等, 其核心肽序列分别为“IWGIGCNP” “IWGIGCDP” “AWLVDCP” 和 “AWLATCP”, 其余的 36 条核心肽序列为未知环肽。目前报道的淡玫红鹅膏核心肽有 5 种七肽、21 种八肽、6 种九肽、7 种十肽和 1 种十一肽, 其中八肽占比最高为 52.5% (21/40), 其次为十肽, 占比为 17.5% (7/40), 最少的是十一肽, 仅占比 2.5% (1/40)。

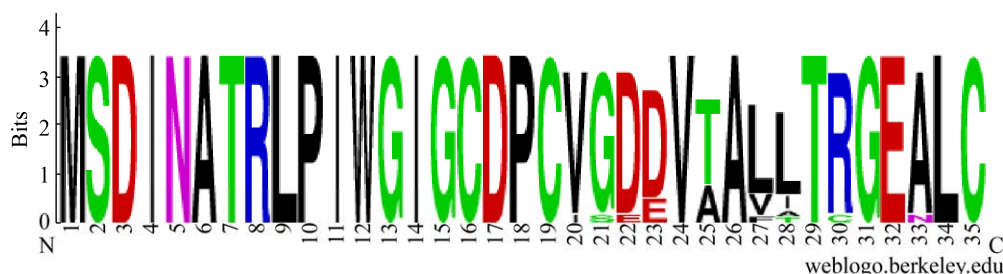


图 7 剧毒鹅膏菌 16 条 β -鹅膏毒肽序列的 WebLogo 比对

Figure 7 WebLogo alignment of 16 β -amanitin sequences in lethal *Amanita*



图 8 剧毒鹅膏菌 10 条羧基二羟鬼笔毒肽序列的 WebLogo 比对

Figure 8 WebLogo alignment of 10 phalloidin sequences in lethal *Amanita*



体肽中6号位点与7号位点的氨基酸起着同样的作用。在核心区最短的多肽为七肽，最长的多肽为十肽。在下游保守区中26、27、28、29和35号位点并不保守，有可能这些位点不是环肽合成中的关键位点。 α -鹅膏蕈毒肽前体肽 WebLogo 对比图上游保守区仅只有一个6号位点的氨基酸不保守，

下游保守区有较多的氨基酸不保守。 β -鹅膏毒肽前体肽 WebLogo 对比显示上游保守区的 10 个氨基酸保守度为 100%，下游保守区和 α -鹅膏毒肽一样，有较多的氨基酸不保守，核心肽区只有 17 号位点与 α -鹅膏毒肽的核心区不一样。根据 α -鹅膏毒肽、 β -鹅膏毒肽和羧基二羟鬼笔毒肽 WebLogo 对比图的结果可知， β -鹅膏毒肽保守度最高，其次为羧基二羟鬼笔毒肽，最不保守的为 α -鹅膏毒肽。

将 18 条 α -Amanitin、14 条 β -Amanitin、9 条 Phallacidin 和 9 条“Phallotoxin”的碱基序列及前体肽氨基酸序列进行比对(图 10)。结果显示，

α -Amanitin 和 β -Amanitin 前体肽碱基序列包含 108 个碱基，Phallacidin 和“Phallotoxin”前体肽碱基序列包含 105 个碱基。在 18 条 α -Amanitin 碱基序列比对中有 69 个碱基位点相同，碱基序列保守度为 63.8% (69/108)，其保守性最低；14 条 β -Amanitin 碱基序列比对中有 71 个碱基位点相同，碱基序列保守度为 65.7% (71/108)；9 条 Phallacidin 碱基序列比对中有 79 个碱基位点相同，碱基序列保守度为 75.2% (79/105)，其保守性最高；9 条“Phallotoxin”碱基序列比对中有 69 个碱基位点相同，碱基序列保守度为 65.7% (69/105)。后续研究中可尝试根据毒肽相同碱基位点以及密

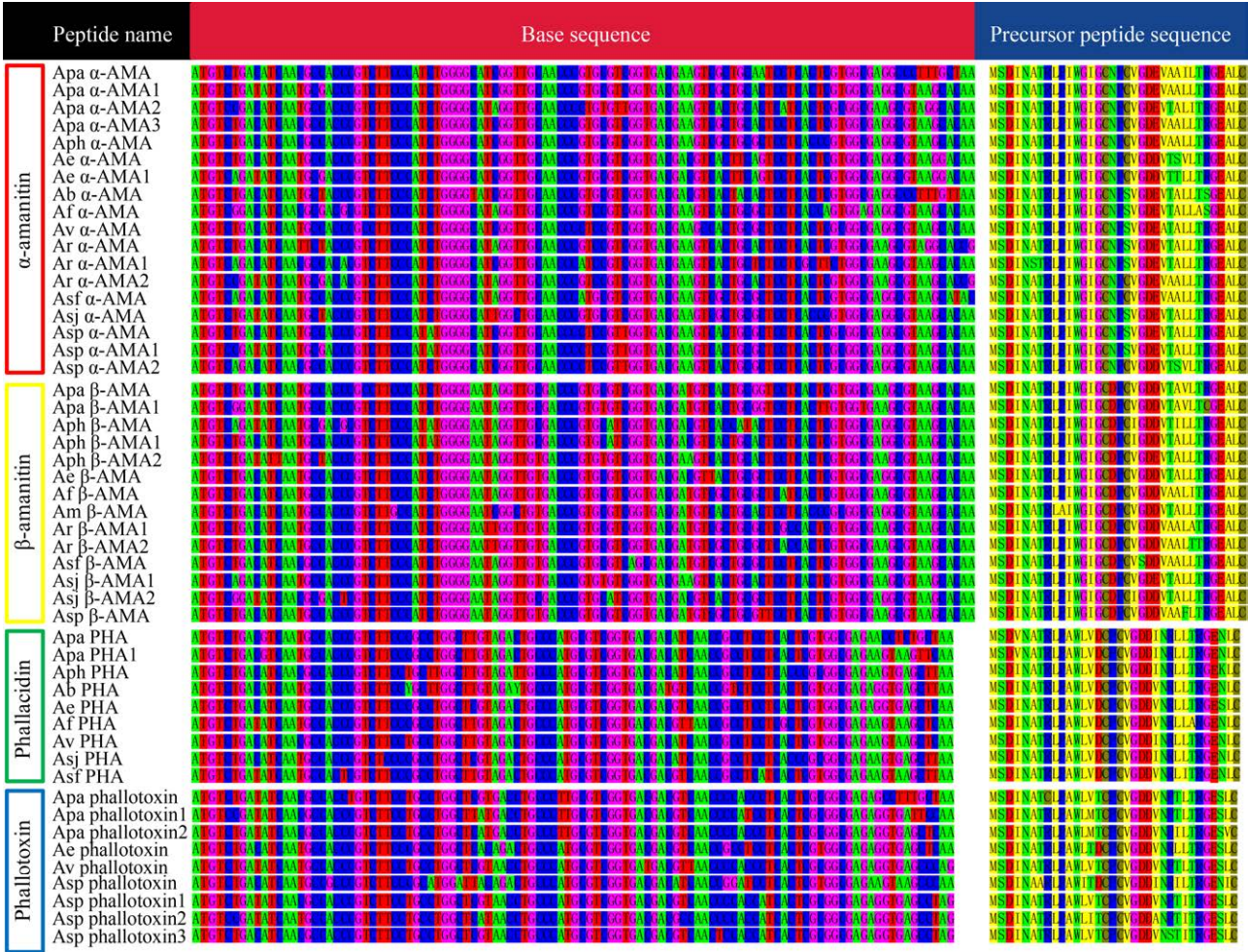


图 10 α -Amanitin、 β -Amanitin、Phallacidin 和“Phallotoxin”序列的比对
Figure 10 Sequence alignment of α -amanitin, β -amanitin, phallacidin and “phallotoxin”

码子的简并性来设计通用引物, 以便在鹅膏菌中进行这类毒素基因的快速检测。

MSDIN 基因家族聚类分析表明具有类似功能的鹅膏环肽可聚为一类, 基于本研究系统发育树推测: MSDIN-27 和 MSDIN-28 这 2 个未知肽功能类似, 其与鹅膏毒肽的关系较近, 可能为鹅膏毒肽的新毒素; MSDIN-1、MSDIN-4 和 MSDIN-18 这 3 个未知肽功能类似, 其关系与 Phalloidin 的关系较近; MSDIN-11 和 MSDIN-24 这 2 个未知肽功能类似, 其关系与 Phalloidin 的关系较近; 因此 MSDIN-1、MSDIN-4、MSDIN-11 和 MSDIN-24 可能为鬼笔毒肽的新毒素。在其他未知肽中, MSDIN-3、MSDIN-8、MSDIN-14 和 MSDIN-17 这 4 个未知肽关系比较接近, 可能有类似的功能; MSDIN-2、MSDIN-12、MSDIN-15、MSDIN-25 和 MSDIN-26 这 5 个未知肽的关系较近, 可能功能类似; MSDIN-4、MSDIN-5、MSDIN-6、MSDIN-9、MSDIN-13、MSDIN-16、MSDIN-19、MSDIN-20、MSDIN-22 和 MSDIN-29 这 10 个未知肽关系比较接近, 这些未知肽功能类似。但所有这些推测, 后续都需进行环肽提取分离和结构鉴定才能最终确定。

目前基于鹅膏菌基因组和转录组测序, 从 DNA 和 RNA 层面对鹅膏环肽 MSDIN 基因家族的研究已有一定报道, 但是在蛋白质组或代谢组层面的研究还未见报道。后续研究可基于基因层面的分析结果, 利用高分辨质谱等手段来寻找未知功能的环肽, 并对其结构和功能进行解析, 以便挖掘和获得更多可利用的环肽资源。

致谢: 感谢广东省科学院微生物研究所贺勇在实验过程中的大力指导和帮助。感谢济南市疾病预防控制中心曹若明副主任技师、广州市疾病预防控制中心刘于飞主任医师, 以及泰安市疾病预防控制中心相关人员在样品采集过程中的帮助。

REFERENCES

- [1] Zhang P, Chen ZH, Xiao B, Tolgor B, Bao HY, Yang ZL. Lethal amanitas of East Asia characterized by morphological and molecular data[J]. Fungal Diversity, 2010, 42(1): 119-133
- [2] Cai Q, Cui YY, Yang ZL. Lethal *Amanita* species in China[J]. Mycologia, 2016, 108(5): 993-1009
- [3] Cao RM, Liu Z, Zhang J, Liu SQ, Sun YB, Zhang J, Liu SH, Liu LZ, Cao XL, Jia TH. Disposal report of a case of mushroom poisoning caused by *Amanita pallidorosea*[J]. Journal of Shandong University: Health Sciences, 2011, 49(12): 160-162,164 (in Chinese)
曹若明, 刘仲, 张济, 刘守钦, 孙延斌, 张军, 刘素华, 刘岚铮, 曹小丽, 贾堂宏. 一起淡玫红鹅膏菌中毒事件处置报告[J]. 山东大学学报: 医学版, 2011, 49(12): 160-162,164
- [4] Sun YB, Cao RM, Liu SH, Zhang J, Liu Z, Cao XL. Investigation on a case of poisoning caused by *Amanita pallidorosea*[J]. Preventive Medicine Tribune, 2012, 18(10): 778-779,782 (in Chinese)
孙延斌, 曹若明, 刘素华, 张军, 刘仲, 曹小丽. 一起误食淡玫红鹅膏菌引起的中毒调查[J]. 预防医学论坛, 2012, 18(10): 778-779,782
- [5] Deng WQ, Li TH, Zhang M, Xiao ZD, He Y, Xu JY. Analysis of common poisonous species of *Amanita* and their poisoning cases in South China[J]. Mycosystema, 2020, 39(9): 1750-1758 (in Chinese)
邓旺秋, 李泰辉, 张明, 肖正端, 贺勇, 徐隽彦. 华南常见毒鹅膏菌及其中毒事件分析[J]. 菌物学报, 2020, 39(9): 1750-1758
- [6] Tang SS, Zhou Q, He ZM, Luo T, Zhang P, Cai Q, Yang ZL, Chen J, Chen ZH. Cyclopeptide toxins of lethal amanitas: compositions, distribution and phylogenetic implication[J]. Toxicon, 2016, 120: 78-88
- [7] Wieland T. Peptides of Poisonous *Amanita* Mushrooms[M]. Berlin, Heidelberg: Springer Berlin Heidelberg, 1986
- [8] Luo H. Advances of omics research on poisonous mushrooms[J]. Mycosystema, 2019, 38(12): 2087-2098 (in Chinese)
罗宏. 毒蘑菇的组学研究[J]. 菌物学报, 2019, 38(12): 2087-2098
- [9] Hallen HE, Luo H, Scott-Craig JS, Walton JD. Gene family encoding the major toxins of lethal *Amanita* mushrooms[J]. Proceedings of the National Academy of Sciences of the United States of America, 2007, 104(48): 19097-19101
- [10] Luo H, Hong SY, Sgambelluri RM, Angelos E, Li X, Walton JD. Peptide macrocyclization catalyzed by a prolyl oligopeptidase involved in α -amanitin biosynthesis[J]. Chemistry & Biology, 2014, 21(12): 1610-1617
- [11] Pulman JA, Childs KL, Sgambelluri RM, Walton JD. Expansion and diversification of the MSDIN family of cyclic peptide genes in the poisonous agarics *Amanita phalloides* and *A. bisporigera*[J]. BMC Genomics, 2016, 17(1): 1038
- [12] He Y, Zhang CH, Deng WQ, Li TH, Zhou XY. De novo transcriptome sequencing of highly toxic *Amanita rimosa*

- transcriptome and analyses of poison-peptide-related gene family[J]. *Genomics and Applied Biology*, 2019, 38(6): 2603-2612 (in Chinese)
- 贺勇, 张成花, 邓旺秋, 李泰辉, 周小云. 剧毒裂皮鹅膏转录组测序与毒肽相关基因家族分析[J]. *基因组学与应用生物学*, 2019, 38(6): 2603-2612
- [13] Luo H, Cai Q, Lüli Y, Li X, Sinha R, Hallen-Adams HE, Yang ZL. The MSDIN family in amanitin-producing mushrooms and evolution of the prolyl oligopeptidase genes[J]. *IMA Fungus*, 2018, 9: 225-242
- [14] He Y, Zhang CH, Deng WQ, Zhou XY, Li TH, Li CH. Transcriptome sequencing analysis of the MSDIN gene family encoding cyclic peptides in lethal *Amanita fuligineoides*[J]. *Toxicon*, 2020, 183: 61-68
- [15] Li P, Deng WQ, Li TH. The molecular diversity of toxin gene families in lethal *Amanita* mushrooms[J]. *Toxicon*, 2014, 83: 59-68
- [16] Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng QD, et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome[J]. *Nature Biotechnology*, 2011, 29(7): 644-652
- [17] Cock PJA, Fields CJ, Goto N, Heuer ML, Rice PM. The Sanger FASTQ file format for sequences with quality scores, and the Solexa/Illumina FASTQ variants[J]. *Nucleic Acids Research*, 2010, 38(6): 1767-1771
- [18] Kumar S, Stecher G, Tamura K. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets[J]. *Molecular Biology and Evolution*, 2016, 33(7): 1870-1874
- [19] Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool[J]. *Journal of Molecular Biology*, 1990, 215(3): 403-410
- [20] Conesa A, Götz S, García-Gómez JM, Terol J, Talón M, Robles M. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research[J]. *Bioinformatics*, 2005, 21(18): 3674-3676
- [21] Quevillon E, Silventoinen V, Pillai S, Harte N, Mulder N, Apweiler R, Lopez R. InterProScan: protein domains identifier[J]. *Nucleic Acids Research*, 2005, 33(Web Server issue): W116-W120
- [22] Li P, Deng WQ, Li TH, Song B, Shen YH. Illumina-based *de novo* transcriptome sequencing and analysis of *Amanita exitialis* basidiocarps[J]. *Gene*, 2013, 532(1): 63-71
- [23] He ZM, Long P, Fang F, Li SN, Zhang P, Chen ZH. Diversity of MSDIN family members in amanitin-producing mushrooms and the phylogeny of the MSDIN and prolyl oligopeptidase genes[J]. *BMC Genomics*, 2020, 21(1): 440