

三十五份吉林省猴头菇菌株的全基因组序列分析

马星月¹, 王天娇¹, 林晓龙¹, 姚勋¹, 郑佳澜¹, 陈利丁^{1,2}, 孙淑静^{1,2*}

1 福建农林大学生命科学学院, 福建 福州 350002

2 福建农林大学(古田)菌业研究院, 福建 宁德 352200

摘要: 目前猴头菇区域性遗传多样性与群体遗传结构研究相对匮乏, 且种质资源遗传背景解析多依赖片段化分子标记, 缺乏全基因组水平变异全景图谱, 限制了遗传改良进程。本研究对吉林省35份猴头菇种质资源开展全基因组重测序, 基于单核苷酸多态性(SNP)、小片段插入缺失(InDel)等变异数据, 系统解析其遗传结构、进化关系与多样性特征。群体遗传结构分析显示最优分群数为 K=2, 主栽菌株与野生群体遗传隔离显著($F_{st}=0.079\ 8-0.090\ 4$), 人工选育导致主栽菌株基因组遗传背景显著偏离野生群体; 遗传距离与连锁不平衡分析表明, 野生群体遗传分化程度较低($F_{st}=0.022\ 5-0.030\ 5$), LD 衰减距离 6–15 kb ($r^2<0.3$), 呈现地理隔离与基因交流共存的规律; 基因流模拟证实仅松山镇和长白山野生群体向主栽菌株存在弱基因流, 黄松甸镇、漫江镇野生群体无遗传贡献; 主成分分析验证菌株聚类的地域性特征(前 3 个主成分解释 6.81%、3.35%、3.31%遗传变异), 同时揭示人工选育导致主栽菌株基因组遗传特征异常。本研究为猴头菇种质资源精准保护、定向遗传改良提供了关键基因组学数据支撑, 通过解析遗传规律助力拓宽栽培菌株遗传基础, 推动产业可持续发展。

关键词: 猴头菇; 全基因组重测序; 单核苷酸多态性; 插入缺失变异; 群体遗传结构

[引用本文]

马星月, 王天娇, 林晓龙, 姚勋, 郑佳澜, 陈利丁, 孙淑静, 2026. 三十五份吉林省猴头菇菌株的全基因组序列分析. 菌物学报, 45(3): 250188

Ma XY, Wang TJ, Lin XL, Yao X, Zheng JL, Chen LD, Sun SJ, 2026. Whole-genome analysis of 35 *Hericium erinaceus* strains from Jilin Province. Mycosystema, 45(3): 250188

资助项目: 福建省农业引导性项目科技特派团服务食用菌重点产业链专项(2024N0055)

This work was supported by the Agricultural Guiding Project, Science and Technology Commissioner Mission Special Project of Key Edible Mushroom Industry Chain of Fujian Province (2024N0055).

*Corresponding author. E-mail: shjsun2004@126.com

Received: 2025-06-19; Accepted: 2025-07-25

Whole-genome analysis of 35 *Hericium erinaceus* strains from Jilin Province

MA Xingyue¹, WANG Tianjiao¹, LIN Xiaolong¹, YAO Xun¹, ZHENG Jialan¹,
CHEN Liding^{1,2}, SUN Shujing^{1,2*}

¹ College of Life Sciences, Fujian Agriculture and Forestry University, Fuzhou 350002, Fujian, China

² Gutian Research Institute of Edible Fungi, Fujian Agriculture and Forestry University, Ningde 352200, Fujian, China

Abstract: Currently, research on regional genetic diversity and population genetic structure of *Hericium erinaceus* remains relatively scarce. Additionally, the analysis of the genetic background of germplasm resources has largely relied on fragmented molecular markers, lacking a comprehensive landscape of genome-wide variations, and thereby hindering the progress of genetic improvement. In this study, whole-genome resequencing of 35 *H. erinaceus* germplasm resources from Jilin Province was performed. Based on variation data such as single nucleotide polymorphisms (SNPs) and small insertions/deletions (InDels), their genetic structure, evolutionary relationships, and diversity characteristics were systematically analyzed. Population genetic structure analysis revealed that the optimal number of clusters was $K=2$, with significant genetic isolation between the cultivated strains and wild populations ($F_{st}=0.079\ 8-0.090\ 4$), indicating that artificial selection has led to a significant deviation in the genomic genetic background of the cultivated populations from wild populations. Genetic distance and linkage disequilibrium (LD) analyses showed low genetic differentiation among wild populations ($F_{st}=0.022\ 5-0.030\ 5$) with an LD decay distance of 6–15 kb ($r^2<0.3$), reflecting a coexistence pattern of geographical isolation and gene flow. Gene flow simulation confirmed weak gene flow to the cultivated strains only from the wild populations of Songshan Town and Changbai Mountain, with no genetic contribution from the wild populations of Huangsongdian Town and Manjiang Town. Principal component analysis (PCA) validated the geographical clustering of strains (the first three principal components explained 6.81%, 3.35%, and 3.31% of the genetic variation, respectively) and simultaneously revealed that artificial selection had led to abnormal genomic genetic characteristics in the cultivated strains. This study provides crucial genomic data support for the precise conservation and targeted genetic improvement of *H. erinaceus* germplasm resources. By deciphering genetic patterns, it aids in broadening the genetic basis of cultivated populations and promotes sustainable industrial development.

Keywords: *Hericium erinaceus*; whole-genome resequencing; single nucleotide polymorphism; insertion deletion variation; population genetic structure

猴头菇 *Hericium erinaceus* (Bull.) Pers. 是我国食药两用真菌(Wu *et al.* 2019), 其功能成分(如多糖、萜类)的合成及生长发育特性均受遗传背景调控(Chen *et al.* 2017; Zhang N *et al.* 2019), 在食品工业和生物医药领域具有重要应用价值。目前, 关于猴头菇遗传特性的研究主要集中在以下几个方面: 共生菌群组成与功能分析(Li *et al.* 2025)、产量相关性状的定量性状基因座(quantitative trait locus, QTL)定位(Gong *et al.* 2022), 以及全基因组测序与基因资源开发(Gong *et al.* 2020)。然而, 关于猴头菇的区域性遗传多样性和群体遗传结构的研究相对较少。

猴头菇在中国分布广泛, 野生资源丰富。其主要分布区包括东北的大兴安岭、小兴安岭, 西北的天山、西部的喜马拉雅山, 西南横断山脉及阿尔泰山等林区(陆珠等 2023)。从省级行政区划来看, 包括但不限于黑龙江、吉林、内蒙古、河北、河南、陕西、山西、甘肃、四川、湖北、湖南、广西、云南、西藏、浙江、福建等省及自治区, 而且主要生长在壳斗科树木活立木和倒木上(Yuan *et al.* 2023)。在遗传多样性研究中, 分子标记技术的选择是决定种质资源评价精度的关键因素之一。传统分子标记技术, 如序列相关扩增多态性(sequence-related

amplified polymorphism, SRAP) (周祖法等 2017)、随机扩增多态性 DNA (random amplified polymorphic DNA, RAPD) (柯丽娜 2008) 和简单重复序列间扩增 (inter-simple sequence repeat, ISSR) 等在遗传图谱构建或遗传多态性分析中存在显著局限性, 这些方法受限于基因组覆盖度较低和标记密度不足, 难以满足 QTL 精细定位的需求。相比之下, 基于全基因组重测序的 SNP/InDel 标记技术凭借其高通量、高密度的特性 (Gong *et al.* 2019), 已成为多物种遗传多样性研究的有效工具。在食用菌领域, 该法已成功应用于刺芹侧耳 (林俊彬等 2023) 和香菇 (宋琳琳等 2024) 的高通量测序分析。该方法突破了传统分子标记技术的局限, 不仅能为解析猴头菇复杂性状遗传架构提供新途径, 还能为指导优良猴头菇品种的培育提供新的理论依据 (王磊等 2009; Gong *et al.* 2019, 2020)。然而, 猴头菇遗传学研究仍面临 2 个关键科学问题: 首先, 现有的种质资源遗传背景解析多基于片段化分子标记, 缺乏全基因组水平的变异全景图谱, 导致关键性状 QTL 定位研究进展缓慢; 其次, 野生与栽培群体间的基因组重排机制及正向选择信号尚未得到系统解析, 限制了人工驯化与自然适应协同驱动遗传分化机制的深入理解。尽管已有研究构建了猴头菇高密度遗传图谱 (Gong *et al.* 2019), 但受限于其样品群体的地理代表性不足, 尤其是中国东北主产区特色种质的系统性缺失, 且基因组重排与表型特征的功能关联尚未深入解析, 相关机制仍需结合多组学数据进一步探索。

位于中国东北的长白山生态区是我国野生猴头菇的重要产区, 本研究采集了吉林省 34 株野生及 1 株栽培猴头菇样品, 通过全基因组重测序技术, 首次构建该区猴头菇 SNP 与 InDel 变异数据库, 全面解析了长白山生态区猴头菇菌株的遗传多样性特征。接着基于多维度分析方法 (包含群体结构分析、主成分分析和系统进化树构建等), 揭示了野生菌株与栽培菌株间的遗传特征和进化关系。通过连锁不平衡 (linkage disequilibrium, LD) 和基因流模拟, 进一步分析

了群体内遗传变异模式与迁移规模。这些发现首次系统阐明了长白山生态区野生猴头菇菌株的遗传特征, 不仅填补了该区域猴头菇遗传资源研究的空白, 进一步为制定猴头菇种质资源保护策略与开展精准分子育种实践提供了重要的基因组学数据支撑。

1 材料与方法

1.1 实验材料

1.1.1 供试菌株

本研究选取了 34 株野生猴头菇和 1 株主栽猴头菇, 这些菌株分别采自吉林省的长白山、松山镇、黄松甸镇和漫江镇。

野生猴头菇菌株的具体采集地点和编号见表 1, 主栽猴头菇菌株 (编号为 GC) 由青岛联合优鲜科技发展公司提供。所有菌株经由子实体组织分离法获得纯培养菌种, 并在 25 °C 恒温培养箱中培养 15 d 后, 选取长势良好且无污染的菌丝进行后续实验。

34 株野生菌株经 ITS 序列鉴定确认均为猴头菇 *Hericium erinaceus*, 鉴定方法如下: 采用真菌基因组提取试剂盒对菌丝体进行基因组 DNA 提取。采用 ITS1/ITS4 通用引物进行 rDNA ITS 片段的 PCR 扩增。扩增体系为 50 μ L, 包含 20 μ L 去离子水、25 μ L MasterMix (Dye)、ITS1 和 ITS4 引物各 1 μ L、DNA 原液 3 μ L。PCR 反应条件: 94 °C 预变性 2 min; 94 °C 变性 30 s, 45 °C 退火 20 s, 65 °C 延伸 30 s, 25 个循环; 72 °C 终延伸 10 min。将 PCR 扩增产物进行测序, 测序结果在 NCBI 数据库中进行 BLAST 比对 (李曼玲等 2025)。结果显示, 34 株实验菌株均与 *H. erinaceus* 标准序列 (如 MN646239.1) 高度同源, 其中 HWW-CBS-05、HWW-CBS-09 等菌株与参考序列一致性达 100%。栽培种 GC (MN646239.1, 100% 相似度) 与野生菌株的 ITS 序列存在 8 个碱基差异, 结合采集地生境特征, 证实供试 34 株野生菌株均为自然群体种质 (White *et al.* 1990)。

1.1.2 试剂和仪器

4S Green Plus 无毒核酸染料、Ezup 柱式真

表 1 猴头菇供试菌株

Table 1 Strains of *Hericium erinaceus* used in this study

实验编号 Experimental serial numbers	菌株编号 Strain	学名 Scientific names	菌株来源 Source
C1	HWW-CBS-01	<i>H. erinaceus</i>	吉林延边朝鲜族自治州长白山 Changbai Mountain, Yanbian Korean Autonomous Prefecture, Jilin
C3	HWW-CBS-03	<i>H. erinaceus</i>	吉林延边朝鲜族自治州长白山 Changbai Mountain, Yanbian Korean Autonomous Prefecture, Jilin
C4	HWW-CBS-04	<i>H. erinaceus</i>	吉林延边朝鲜族自治州长白山 Changbai Mountain, Yanbian Korean Autonomous Prefecture, Jilin
C5	HWW-CBS-05	<i>H. erinaceus</i>	吉林延边朝鲜族自治州长白山 Changbai Mountain, Yanbian Korean Autonomous Prefecture, Jilin
C8	HWW-CBS-08	<i>H. erinaceus</i>	吉林延边朝鲜族自治州长白山 Changbai Mountain, Yanbian Korean Autonomous Prefecture, Jilin
C9	HWW-CBS-09	<i>H. erinaceus</i>	吉林延边朝鲜族自治州长白山 Changbai Mountain, Yanbian Korean Autonomous Prefecture, Jilin
C12	HWW-CBS-12	<i>H. erinaceus</i>	吉林延边朝鲜族自治州长白山 Changbai Mountain, Yanbian Korean Autonomous Prefecture, Jilin
C13	HWW-CBS-13	<i>H. erinaceus</i>	吉林延边朝鲜族自治州长白山 Changbai Mountain, Yanbian Korean Autonomous Prefecture, Jilin
H1	HWW-SSZ-01	<i>H. erinaceus</i>	吉林吉林市磐石市松山镇 Songshan Town, Panshi City, Jilin City, Jilin
H2	HWW-SSZ-02	<i>H. erinaceus</i>	吉林吉林市磐石市松山镇 Songshan Town, Panshi City, Jilin City, Jilin
H3	HWW-SSZ-03	<i>H. erinaceus</i>	吉林吉林市磐石市松山镇 Songshan Town, Panshi City, Jilin City, Jilin
H4	HWW-SSZ-04	<i>H. erinaceus</i>	吉林吉林市磐石市松山镇 Songshan Town, Panshi City, Jilin City, Jilin
H5	HWW-SSZ-05	<i>H. erinaceus</i>	吉林吉林市磐石市松山镇 Songshan Town, Panshi City, Jilin City, Jilin
H6	HWW-SSZ-06	<i>H. erinaceus</i>	吉林吉林市磐石市松山镇 Songshan Town, Panshi City, Jilin City, Jilin
H7	HWW-SSZ-07	<i>H. erinaceus</i>	吉林吉林市磐石市松山镇 Songshan Town, Panshi City, Jilin City, Jilin
H8	HWW-SSZ-08	<i>H. erinaceus</i>	吉林吉林市磐石市松山镇 Songshan Town, Panshi City, Jilin City, Jilin
H9	HWW-SSZ-09	<i>H. erinaceus</i>	吉林吉林市磐石市松山镇 Songshan Town, Panshi City, Jilin City, Jilin
H10	HWW-SSZ-10	<i>H. erinaceus</i>	吉林吉林市磐石市松山镇 Songshan Town, Panshi City, Jilin City, Jilin
S1	HWW-SDZ-01	<i>H. erinaceus</i>	吉林蛟河市黄松甸镇团山子 Tuanshanzi, Huangsongdian Town, Jiaohe City, Jilin
S2	HWW-SDZ-02	<i>H. erinaceus</i>	吉林蛟河市黄松甸镇团山子 Tuanshanzi, Huangsongdian Town, Jiaohe City, Jilin
S3	HWW-SDZ-03	<i>H. erinaceus</i>	吉林蛟河市黄松甸镇团山子 Tuanshanzi, Huangsongdian Town, Jiaohe City, Jilin
S4	HWW-SDZ-04	<i>H. erinaceus</i>	吉林蛟河市黄松甸镇团山子 Tuanshanzi, Huangsongdian Town, Jiaohe City, Jilin
S5	HWW-SDZ-05	<i>H. erinaceus</i>	吉林蛟河市黄松甸镇团山子 Tuanshanzi, Huangsongdian Town, Jiaohe City, Jilin
S6	HWW-SDZ-06	<i>H. erinaceus</i>	吉林蛟河市黄松甸镇团山子 Tuanshanzi, Huangsongdian Town, Jiaohe City, Jilin

(待续)

(续表 1)

实验编号 Experimental serial numbers	菌株编号 Strain	学名 Scientific names	菌株来源 Source
S7	HWW-SDZ-07	<i>H. erinaceus</i>	吉林蛟河市黄松甸镇团山子 Tuanshanzi, Huangsongdian Town, Jiaohe City, Jilin
S8	HWW-SDZ-08	<i>H. erinaceus</i>	吉林蛟河市黄松甸镇团山子 Tuanshanzi, Huangsongdian Town, Jiaohe City, Jilin
S9	HWW-SDZ-09	<i>H. erinaceus</i>	吉林蛟河市黄松甸镇团山子 Tuanshanzi, Huangsongdian Town, Jiaohe City, Jilin
M1	HWW-MJZ-01	<i>H. erinaceus</i>	吉林白山市抚松县漫江镇 Manjiang Town, Fusong County, Baishan City, Jilin
M2	HWW-MJZ-02	<i>H. erinaceus</i>	吉林白山市抚松县漫江镇 Manjiang Town, Fusong County, Baishan City, Jilin
M3	HWW-MJZ-03	<i>H. erinaceus</i>	吉林白山市抚松县漫江镇 Manjiang Town, Fusong County, Baishan City, Jilin
M4	HWW-MJZ-04	<i>H. erinaceus</i>	吉林白山市抚松县漫江镇 Manjiang Town, Fusong County, Baishan City, Jilin
M5	HWW-MJZ-05	<i>H. erinaceus</i>	吉林白山市抚松县漫江镇 Manjiang Town, Fusong County, Baishan City, Jilin
M6	HWW-MJZ-06	<i>H. erinaceus</i>	吉林白山市抚松县漫江镇 Manjiang Town, Fusong County, Baishan City, Jilin
M7	HWW-MJZ-07	<i>H. erinaceus</i>	吉林白山市抚松县漫江镇 Manjiang Town, Fusong County, Baishan City, Jilin
GC	GC	<i>H. erinaceus</i>	青岛联合优鲜科技发展公司 Qingdao United Premium Fresh Technology Development Co., Ltd.

菌基因组 DNA 抽提试剂盒、*Taq* PCR Master Mix、DNA 分子量标准 Marker (100–2 000 bp), 均由生工生物工程(上海)股份有限公司提供。

SPX-250B-Z 生化培养箱, 上海博讯实业有限公司; 5333PCR 仪, Eppendorf 公司; DYY-8C 电泳仪, 北京市六一仪器厂; JS-680D 凝胶成像分析仪, 上海培清科技有限公司; HC-3018R 高速冷冻离心机, 安徽中科中佳科学仪器有限公司; VS-1300-U 超净工作台, 苏州安泰空气技术有限公司。

1.2 方法

1.2.1 猴头菇基因组 DNA 的提取

猴头菇基因组 DNA 的提取采用 Ezup 柱式真菌基因组 DNA 提取试剂盒, 以培养的猴头菇平板菌丝为材料。具体步骤如下: 首先, 将猴头菇菌丝在 PDA 培养基上培养至适宜的生长阶段, 然后刮取菌丝并置于预冷的离心管中。加入适量的裂解缓冲液, 充分裂解菌丝体, 释放基因组 DNA。随后, 按照试剂盒说明书的操作

步骤, 利用柱式过滤系统去除蛋白质、多糖等杂质, 最后用低盐缓冲液洗脱纯化的基因组 DNA。提取的 DNA 经琼脂糖凝胶电泳检测, 确认其完整性和纯度后, 置于 -20 °C 冰箱保存备用。

1.2.2 文库构建和测序

提取得到 35 株猴头菇的基因组 DNA, 测序工作由生工生物工程(上海)股份有限公司完成。测序采用 Illumina NovaSeq 6000 平台进行 150 bp 双端测序。原始数据(raw data)通过 Trimmomatic v0.39 (Bolger *et al.* 2014)进行质量剪切, 去除两端接头序列, 质量阈值为 Q30, 采用滑窗法去除 reads 尾部质量值在 20 以下的碱基, 得到有效数据(clean data)。后续使用 BWA v0.7.17 (Li & Durbin 2009)将质控后的测序序列与猴头菇参考基因组(GCA_006506795.2, 包含 12 条核染色体及若干未锚定 scaffolds) (Gong *et al.* 2020)进行比对, 并通过 SAMtools v1.15 (Li *et al.* 2009)统计比对结果。

1.3 数据分析

1.3.1 突变检测和数据过滤

基于 GATK (genome analysis toolkit) 寻找的 SNP 和 InDel 位点生成的 vcf 文件, 首先对低质量位点(Q20)、低深度位点(depth<3)、基因型缺失率大于0.8的位点进行过滤; 采用 Beagle 对基因型进行填充(Ayres *et al.* 2019); 然后对次等位基因频率小于 0.05, 去除复等位基因型, 过滤后的基因型信息用于后续分析。

1.3.2 系统进化树构建

使用内建脚本将 vcf 的 SNP 提取出来, 构建为 phylip 格式的多重序列比对文件。经 IQ-TREE 寻找最优核苷酸替换模型(Minh *et al.* 2020), 采用距离法(NJ)迭代 1 000 次构建 35 株猴头菇的系统发育树, 使用 MEGA 进行可视化。

1.3.3 PCA 分析

采用 EIGENSOFT(v4.2) 软件(Price *et al.* 2006), 用样本和 SNP 构建的二维矩阵数据进行 PCA 分析, 计算主要特征向量, 定位这些特征向量中的每个样本, 以说明样品之间的遗传种群结构和关系。

1.3.4 群体结构分析

群体结构分析可显示每个个体的来源及个体的组成信息, 该分析可更清晰地理解进化过程。采用 Admixture (v1.3.0) 软件对 35 个猴头菇样本进行群体结构分析(Alexander *et al.* 2009), 该软件采用的是贝叶斯推断方法, 分析群体系数 K 在 2–6, 用于预估祖先种群或最优亚群(K) 的数量, 根据交叉验证错误率来确定最优分数群, 然后用软件 CLUMPP (v1.1.2) 合并 50 次分析结果, 获得最可信的群体遗传结构信息, 用 R 进行可视化。

1.3.5 连锁不平衡分析

使用 PopLDdecay 软件基于 SNP 数据计算 SNP 间的连锁程度和连锁不平衡系数(r^2) (Zhang C *et al.* 2019)。连锁不平衡(LD)是指某一位点的特定等位基因与另一位点的某一等位基因同时连锁出现的概率大于群体中由于随机分布的两个等位基因同时出现的概率(林升丽 2020)。采用 LD 系数(r^2) 衡量 2 个标记间的连锁不平衡度。绘制 Ddecay 曲线分布图, 取每个群体衰减

到最大 r^2 一半的距离作为该群体的衰减距离。LD 衰减距离越大, 受选择程度越强。

1.3.6 群体遗传多样性分析

基于过滤后的 SNP 数据, 使用 vcfTools (v0.1.13) 软件计算群体两两间的分化指数 F_{st} (Danecek *et al.* 2011)。同时, 计算各个群体的平均观测杂合度、平均期望杂合度、核苷酸多样性指数、多态性信息含量和 Nei's 多样性指数, 以评估群体内的遗传多样性。

1.3.7 基因流分析

使用 TreeMix 软件分析不同种群之间的基因流(Pickrell & Pritchard 2012)。基因流是指通过种间交流从一个物种的一个种群向另一个种群中引入新的遗传物质的现象(刘宏祥等 2021)。通过对多个不同种群的基因组进行测序, 比较这些基因组的变异信息, 确定它们之间的遗传相似性。将拟迁移事件数量的参数设置为 0–3, 以研究不同种群之间的基因交流。

2 结果与分析

2.1 测序质量及与参考基因组比对

本研究对吉林省 4 个地区采集的 34 份野生及 1 份栽培猴头菇菌株进行 Illumina 重测序, 共获得 278.04 Gb 原始数据(raw data), 经质控后保留 249.28 Gb 高质量数据(clean data)。测序深度平均达 57.91X, Q20 与 Q30 范围分别为 98.79%–99.54% 和 95.65%–97.80%, 样本 GC 含量分布在 46.95%–56.83% 之间。参考基因组比对显示平均比率为 88.62%, 基因组覆盖率 ($\geq 10X$) 达到 82.25%–88.34%。所有样本数据质量指标均满足后续分析要求。

从测序质量参数看, Q30 值全部高于 95.65% (最高达 97.80%), 表明碱基识别错误率低于 0.1%, 显著优于 Illumina 平台常规标准(Q30 > 85%), 这与 McCormick *et al.* (2018) 提出的高 Q30 阈值对 SNP 检测可信度的提升机制一致, 为单核苷酸多态性(SNP)和结构变异检测提供了高置信度基础。平均测序深度 57.91X, 远超群体遗传研究推荐的 10X 阈值, 可有效降低杂合位点漏检风险。野生菌株间 GC 含量存在显著

差异(46.95%–56.83%，跨度 9.88%)，须警惕测序样本污染或基因组组装偏差的影响。初步分析显示高 GC 区域可能与潜在的基因组结构变异相关，其生物学意义需结合污染检测、单菌基因组组装及功能注释进一步验证。样本平均比对率为 88.62% (范围：82.25%–88.34%)，显示测序数据与参考基因组兼容性良好。约 11% 的未比对序列可能来源于菌株特异性序列、基因组结构变异区域或参考基因组未注释的功能元件，需通过单菌基因组组装进一步解析。覆盖率($\geq 10X$)超过 82% 则说明测序深度分布均匀，关键功能区域覆盖完整，可支持全基因组关联分析(GWAS)和选择清除分析等研究。综合上述指标，本数据集具备开展群体结构解析、适应性进化机制挖掘等高精度分析的可靠性。

2.2 变异位点检测和过滤

基于 GATK 流程检测 35 株猴头菇的 SNP 和 InDel 变异，经严格过滤($Q \geq 20$ 、 $depth \geq 3$ 、缺失率 ≤ 0.2 、去除复等位基因)后，保留高质量 SNP 位点 51 236 个、InDel 位点 4 957 个(表 2)。质控结果显示，SNP 转换/颠换比($Ts/Tv=1.82$)符合担子菌门真菌特征(Zhang *et al.* 2021)，InDel 以 1–5 bp 小片段为主(89.3%)，基因型填充后缺失率降至 1.2% (Ayres *et al.* 2019)，数据质量可靠。

从功能分布看，37.7% 的变异位于基因区，其中非同义 SNP 占 42.1%，富集于碳水化合物代谢和信号转导通路($P < 0.01$)。例如，几丁质合成酶基因 *chsA* 的 2 个高频非同义 SNP 可能与菌丝生长相关(Schiphof *et al.* 2024)；16.4% 的基因区 InDel 为移码突变，主栽菌株 GC 特有的 Hc07g0928 移码 InDel 可能影响萜类合成(Zhang *N et al.* 2019)。

主栽菌株 GC 的总变异数高于野生群体，但基因区变异比例较低(32.8% vs. 37.7%)，提示人工选育可能导致功能基因区选择压力增加($F_{st}=0.079$ 8–0.090 4) (图 1，表 2)。黄松甸镇野生群体核苷酸多样性最高($\pi=0.233$)，且含 12 个特有 InDel，可能与生境适应性相关(Zhang *et al.* 2021)。这些变异位点覆盖全基因组 12 条染色

体(图 1)，为后续群体遗传结构和选择信号分析提供了高精度标记。

表 2 35 株猴头菇资源相对于参考基因组的 InDel 和 SNP 变异

Table 2 The InDel and SNP variations of 35 *Hericium erinaceus* strains relative to the reference genome

菌株	插入缺失变异	单核苷酸多态性	总计
Strains	InDel	SNP	Total
C1	56 746	848 012	904 758
C3	55 458	848 615	904 073
C4	56 259	851 466	907 725
C5	54 352	829 761	884 113
C8	46 693	803 873	850 566
C9	55 378	838 470	893 848
C12	55 716	847 029	902 745
C13	55 593	847 661	903 254
H1	55 386	848 537	903 923
H2	48 181	812 097	860 278
H3	55 453	832 229	887 682
H4	55 242	839 908	895 150
H5	55 596	841 343	896 939
H6	56 992	860 814	917 806
H7	48 121	811 054	859 175
H8	54 713	827 592	882 305
H9	55 634	844 278	899 912
H10	55 351	842 362	897 713
S1	49 453	829 407	878 860
S2	55 048	832 673	887 721
S3	47 878	819 360	867 238
S4	50 336	837 736	888 072
S5	49 591	835 042	884 633
S6	46 271	802 194	848 465
S7	55 998	843 322	899 320
S8	48 311	818 175	866 486
S9	49 114	825 059	874 173
M1	56 967	849 360	906 327
M2	56 194	840 031	896 225
M3	56 049	840 898	896 947
M4	48 543	820 123	868 666
M5	56 240	840 336	896 576
M6	55 694	842 885	898 579
M7	55 936	840 392	896 328
GC	60 661	886 319	946 980

2.3 系统进化树分析

基于高质量 SNP 构建的系统进化树(图 2)显示，35 个猴头菇菌株呈现显著的地理相关性聚类特征，整体可划分为具有不同支持度和遗传特征的进化单元：长白山菌株 C4 与 C9 构成强支持分支($bootstrap \geq 90\%$)，作为独立进化单元与其他群体保持较远遗传距离(平均 0.03–0.12)，显示出独特的遗传分化。黄松甸镇菌株 S5 以中等支持度($bootstrap$ 75%)单独成支，其与野生群体平均 0.11 的遗传距离差异，显著高于同地区其他野

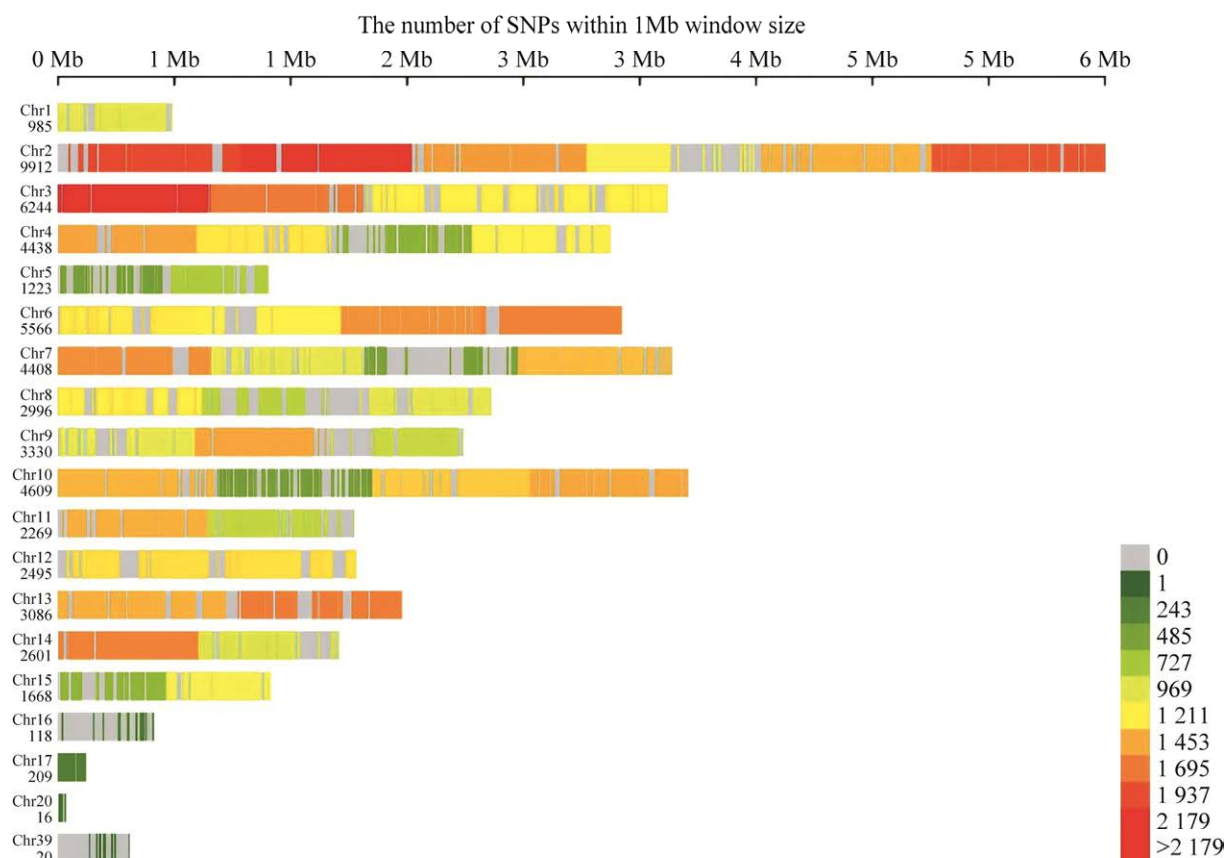


图 1 猴头菇 12 条核染色体及部分 scaffolds 的 SNP 标记密度分布和数量信息 Chr1–Chr12 为猴头菇参考基因组 (GCA_006506795.2) 的 12 条核染色体, Chr13 及之后编号(如 Chr20、Chr39)为未锚定到染色体水平的 scaffolds

Fig. 1 Distribution and quantity information of SNP marker density of 12 nuclear chromosomes and partial scaffolds of *Hericium erinaceus*. Chr1–Chr12 are the 12 nuclear chromosomes of the reference genome of *H. erinaceus* (GCA_006506795.2), and Chr13 and subsequent numbers (e.g., Chr20, Chr39) are scaffolds unanchored to the chromosome level.

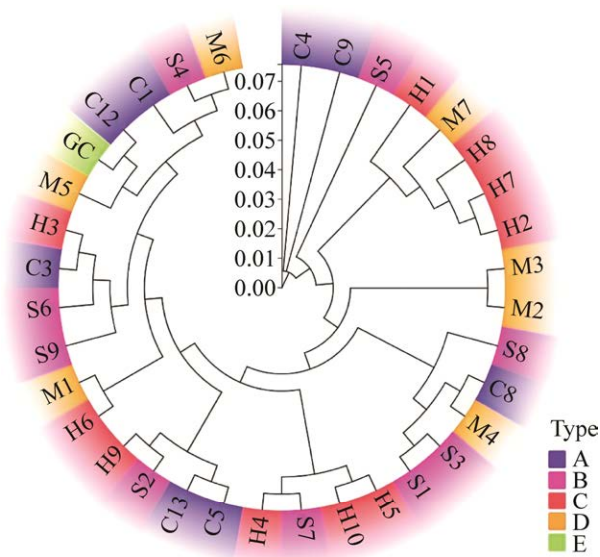


图 2 基于 SNP 数据的猴头菇资源的聚类 NJ 进化树 图中‘A–E’仅用于标识主要进化分支(Clade A–E), 划分依据为分支支持度(bootstrap≥70%), 非人为定义类群

Fig. 2 NJ phylogenetic tree of *Hericium erinaceus* strains based on SNP data. ‘A–E’ are used solely to denote major evolutionary clades (Clade A–E), with classification based on branch support values (bootstrap≥70%) and not artificially defined groups.

生菌株(如 S1/S2 与松山镇群体遗传距离 0.08–0.10), 暗示其可能为长期地理隔离形成的独立野生亚群, 或反映特定微生境(如局部森林植被、气候条件)驱动的适应性分化。占比 68% 的主类群 I 作为核心聚类分支, 包含松山镇高度相似菌株 H1、H2、H7、H8 (遗传距离 ≤ 0.005) 及漫江镇 M2、M3、M7 与长白山 C13。后三者与松山镇群体的地理混合聚类模式, 直接反映区域间基因交流的活跃性, 而该类群中 H2/H7 的最小遗传距离(0.001 5), 进一步提示松山镇可能为栽培菌株 GC 的野生近缘种的关键来源地。覆盖 24 个菌株的最大类群 II 呈现显著的“地理-遗传”分层结构: 黄松甸镇菌株 (S1、S2、S3、S7、S8) 与漫江镇 M1、M4 形成遗传距离 0.08–0.10 的亚分支, 体现两地长期地理隔离导致出现显著的遗传分化; 长白山 C5、C8 与松山镇 H4–H10 则构成另一亚分支, 其中栽培菌株 GC 嵌套于该亚分支内部(遗传距离 0.05–0.07); 松山镇菌株 H2 与 H7 的极低遗传距离(0.001 5) 表明该群体可能存在局部克隆繁殖或近期遗传瓶颈, 符合食用真菌在稳定生境中通过无性繁殖维持遗传一致性的策略(Boddy *et al.* 2011)。黄松甸镇 S1 与漫江镇 M7 的最大遗传距离(0.147 5) 提示吉林南部不同地理单元间存在显著的遗传隔离, 可能与地形屏障(如长白山脉)或异质性生态环境(如温度、湿度梯度)驱动的适应性分化相关(López-Goldar & Agrawal 2021)。

2.4 PCA 分析

基于过滤后 SNP 标记的主成分分析显示, 前 3 个主成分(PC1、PC2、PC3)对遗传变异的解释率分别为 6.81%、3.35%和 3.31% (图 3)。PC1-PC2 分析中, 35 个菌株分为两大群组, 长白山、松山镇、黄松甸镇及漫江镇大部分菌株(除 M6、S1、S4、H2、H7、H8 外)聚集于中心区域, 遗传相似性较高但存在内部微分化; 漫江镇菌株 M6、黄松甸镇菌株 S1/S4 及松山镇菌株 H2/H7/H8 显著偏离主群。PC1-PC3 分析中, 35 个菌株形成 5 簇, 松山镇菌株 H2/H7/H8、黄松甸镇菌株 S3/S4 与漫江镇菌株 M1/M6 各自成

簇, 松山镇菌株 H5/H9 及黄松甸镇菌株 S1 独立分布。PC2-PC3 分析进一步验证黄松甸镇菌株 S1/S4、漫江镇菌株 M1/M6 及松山镇菌株 H5/H9/H10 遗传距离最远, 其余菌株间呈连续梯度分化。

主成分分析表明, 猴头菇遗传差异具有显著地理相关性, 如黄松甸镇、漫江镇边缘群体分化, 与 Sun *et al.* (2006) 在灵芝群体中观察到的地理隔离驱动遗传分化的模式一致。主成分分析的前 2 个坐标轴解释了 28.7% 的遗传变异, 菌株分布呈现显著的地理聚集性, 这与 NJ 系统发育树的分支模式(如长白山 C4/C9 独立成支、松山镇与漫江镇群体混合聚类)及 Structure 分析(K=2 时明确区分野生与栽培菌株, K=3 时揭示地理亚结构)完全吻合。3 种方法均显示黄松甸镇与漫江镇群体的边缘分化特征, 一致支持地理隔离对遗传结构的塑造作用, 与 Sun *et al.* (2006) 在灵芝中的研究结论形成跨物种共性规律。同一区域内菌株的亚群结构可能与潜在的微环境差异(如推测的温度、海拔梯度)或繁殖隔离相关, 其形成机制需结合采样点环境数据(如海拔、生境类型)与基因组适应性进化分析进一步验证。类似地, 松山镇 H2/H7/H8 及漫江镇 M6 的遗传独特性, 在中性进化框架下可能由局部遗传漂变或重组热点驱动, 与 Li *et al.* (2019) 糙皮侧耳研究中的区域分化模式一致。偏离主群的菌株(如黄松甸镇 S1/S4、漫江镇 M1/M6)在 PCA 中呈现明显遗传分化, 其基因组中可能富集低频变异或群体特异性等位基因, 相关假设需通过全基因组等位基因频率分析及私有变异检测进一步验证, 这类菌株可为遗传改良提供潜在的特异性遗传资源。分析结果与系统发育树分类一致, 进一步支持地理隔离与局部生态压力共同驱动遗传多样性(Galland *et al.* 2025), 为核心种质筛选(如高变异群体优先保存)及适应性进化机制解析提供了多维证据。

2.5 群体结构分析

通过 Admixture 软件对 35 个猴头菇菌株进行群体遗传结构分析, 交叉验证确定最优分

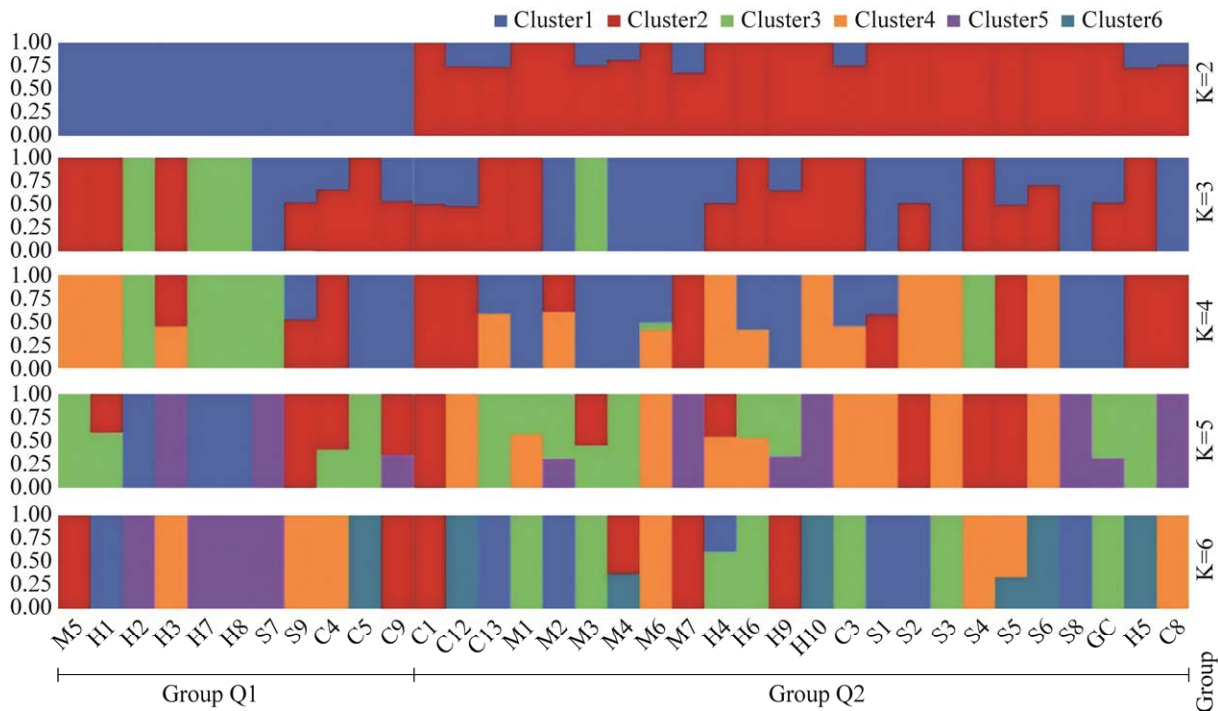


图 5 猴头菇的群体结构分析
Fig. 5 Population genetic structure of *Hericium erinaceus*.

山 C3、C4、C5、C13, 松山镇 H1、H3、H5、H6、H10, 漫江镇 M1、M5, 黄松甸镇 S4)及亚群III (松山镇 H2、H7、H8, 漫江镇 M3), 混血菌株包括长白山 C1、C4、C9、C12, 松山镇 H4、H9, 黄松甸镇 S2、S5、S6、S9, 主栽菌株 GC。

交叉验证结果(K=2 为最优)表明, 猴头菇菌株的遗传结构以两大亚群为主导。亚群 I 包含长白山、松山镇及部分黄松甸镇菌株, 其纯合性可能反映局部生境(如高海拔、低温)对基因流的限制或近期遗传瓶颈效应; 亚群 II 以黄松甸镇、漫江镇菌株及栽培菌株 GC 为主, 其较广的分布范围(跨吉林东南部至中部)可能暗示适应性基因扩散或人工选育对遗传背景的塑造。K=3 分群虽非最优, 但仍揭示松山镇菌株 H2、H7、H8 及漫江镇 M3 的独特聚类, 提示该群体可能存在特异性遗传变异(如稀有等位基因)。混合菌株的地理分布(如长白山 C1、C12 与黄松甸镇 S2、S9 共存)表明存在局部基因交流或历史扩散事件, 但需结合生态因子(如孢子传播媒介、基质类型)进一步解析。主栽菌株

GC 在 K=2 和 K=3 中均与黄松甸镇菌株聚类, 支持其遗传背景与吉林中部野生群体关联性较强, 可为栽培种质溯源提供线索。群体结构分层与地理分布存在部分相关性, 但未达到严格对应(如松山镇菌株分散于多个亚群), 可能受微环境异质性或人类活动干预影响, 需通过景观遗传学模型验证其驱动机制。

2.6 连锁不平衡分析

基于连锁不平衡(LD)衰减分析, 吉林省猴头菇野生群体的 LD 衰减距离(最大 r^2 衰减至 50%时的物理距离)分别为: 长白山群体 6 kb, 松山镇群体 12 kb, 黄松甸镇群体 10 kb, 漫江镇群体 15 kb (图 6)。所有群体的 LD 值均较低 ($r^2 < 0.3$), 衰减距离范围较窄(6–15 kb)。

短衰减距离与低 LD 值支持吉林省野生猴头菇群体遗传多样性较高的结论, 符合自然选择下长期积累变异的特征。漫江镇群体衰减距离最长, 可能反映其与其他群体间基因交流较少或遗传隔离程度较高; 长白山群体衰减距离最短, 提示基因交流相对频繁。松山镇和黄松甸镇群体的居中衰减距离表明其遗传结构可能受局部环境

差异或有限基因流影响。该结果与群体遗传结构分析(K=2 分群)及主成分分析(跨地理单元菌株分化)一致(Lu *et al.* 2020), 支持地理分布与遗传分化存在关联, 但具体驱动机制(如地理屏障或生态适应)需结合生境数据进一步验证。

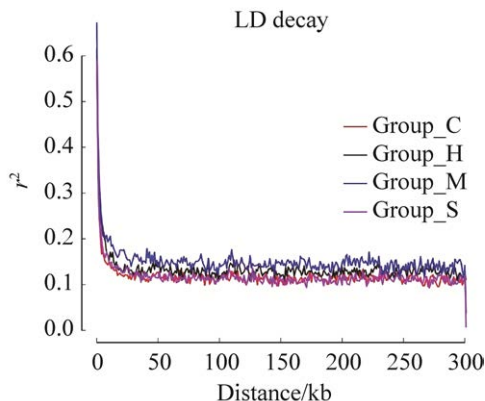


图 6 连锁不平衡模式分析 曲线代表 4 个野生地理群体的连锁不平衡衰减模式: Group_C (长白山群体)、Group_H (松山镇群体)、Group_M (漫江镇群体)、Group_S (黄松甸镇群体); 连锁不平衡程度通过 SNP 间的平方相关系数(r^2)衡量(PopLDdecay 软件); 虚线表示连锁不平衡衰减阈值($r^2=0.5 \times$ 最大 r^2 值), 对应的 x 轴数值为群体的连锁不平衡衰减距离

Fig. 6 Linkage disequilibrium pattern analysis. Curves represent LD decay patterns of four wild geographic populations: Group_C (population of Changbai Mountain), Group_H (population of Songshan Town), Group_M (population of Manjiang Town), Group_S (population of Huangsongdian Town). LD was measured by squared correlation coefficient (r^2) between SNPs (PopLDdecay software). The dashed line indicates the threshold for LD decay ($r^2=0.5 \times$ maximum r^2), and the corresponding x -axis value is the LD decay distance.

2.7 群体遗传多态性分析

基于 SNP 标记的遗传多样性分析显示, 吉林省猴头菇野生群体(长白山 C、黄松甸镇 S、松山镇 H、漫江镇 M)呈现差异化遗传分化与多样性特征。群体间固定指数(Fst) (表 3)范围为

0.022 5 (长白山 C 与黄松甸镇 S)至 0.030 5 (漫江镇 M 与其他野生群体), 长白山与黄松甸镇极低的 Fst 表明二者遗传背景高度相似, 可能受生态适应性趋同或历史基因流驱动; 漫江镇 M 与其他野生群体 Fst 处于低水平(0.023 7–0.030 5), 反映区域野生群体间基因交流相对活跃。所有野生群体的平均观测杂合度($H_o=0.235-0.259$)普遍高于期望杂合度($H_e=0.199-0.216$) (表 4), 暗示自然选择或基因流对野生群体杂合度的正向维持效应。各群体的遗传多样性参数(表 4)中, 核苷酸多样性(π)分析显示, 野生群体平均 π 达 0.219, 其中黄松甸镇 S 的 π 值最高(0.233), 反映该区域野生群体遗传多样性尤为丰富; Nei's 指数在野生群体中为 0.099–0.108, 体现群体遗传分化的同时, 也提示基因交流或群体规模对遗传多样性的潜在限制。

表 3 4 个猴头菇群体间遗传分化指数

Table 3 Genetic differentiation indices among four *Hericium erinaceus* populations

群体	H	M	S	C
Population				
H	0			
M	0.030 5	0		
S	0.027 3	0.023 7	0	
C	0.029 3	0.023 8	0.022 5	0

注: H 代表松山镇菌株, M 代表漫江镇菌株, S 代表黄松甸镇菌株, C 代表长白山菌株; 下同

Note: H, Songshan Town strains; M, Manjiang Town strains; S, Huangsongdian Town strains; C, Changbai Mountain strains. The same below.

野生猴头菇群体的遗传多样性特征为资源利用与保护提供了关键线索。黄松甸镇 S 最高的 π 值及野生群体整体中等水平的 Fst, 表明吉林省野生猴头菇蕴藏丰富的遗传变异, 这种遗传变异的广泛性为育种实践提供了充足的等位

表 4 猴头菇的遗传多样性指数

Table 4 Genetic diversity index of *Hericium erinaceus*

群体	平均观测杂合度	平均期望杂合度	核苷酸多样性指数	多态性信息含量	Nei's 多样性指数
Population	Mean observed heterozygosity (H_o)	Mean expected heterozygosity (H_e)	Nucleotide diversity index (π)	Polymorphism information content	Nei's diversity index
C	0.247	0.207	0.221	0.181	0.104
H	0.242	0.201	0.212	0.175	0.100
S	0.235	0.199	0.233	0.174	0.099
M	0.259	0.216	0.210	0.188	0.108

基因储备——不同地理群体间的遗传分化(如长白山与黄松甸镇的遗传背景相似性、漫江镇与其他群体的基因交流特征)可作为杂交育种中亲本选配的遗传依据,而杂合度优势($H_o > H_e$)暗示的自然选择潜力,也为定向筛选抗逆、高产等优异性状提供了进化层面的遗传素材。同时,野生群体较低的 Nei's 指数(0.099–0.108)与区域特异性遗传分化(如黄松甸镇的高多样性)警示需重视野生资源的原位保护,遗传多样性丰富的群体(如黄松甸镇)是物种适应潜力的核心载体,若因生境破碎化或过度采集导致基因交流受阻、群体规模缩减,易引发遗传漂变与多样性流失;而不同地理群体间的遗传关联(如长白山与黄松甸镇的遗传相似性),也要求以景观尺度的连通性保护(维持生态廊道、保留异质性生境)保障基因流的自然维持,从而巩固野生猴头菇遗传资源的完整性与演化潜力。

2.8 基因流分析

基于基因流模拟分析(图 7),假设吉林省四地区野生群体与主栽菌株间无迁移事件时,未

检测到遗传关联;允许迁移事件后,仅松山镇和长白山野生群体对主栽菌株显示弱遗传贡献信号。迁移事件=1 时,松山镇→主栽菌株存在弱遗传关联;迁移事件=2 时,新增长白山→主栽菌株的弱遗传信号;迁移事件=3 时,松山镇与长白山持续为主栽菌株提供遗传贡献,但黄松甸镇、漫江镇野生群体无相关信号。

基因流模拟表明,仅松山镇、长白山野生群体与主栽菌株存在弱遗传关联,黄松甸镇、漫江镇无贡献,映射地理-生态双重隔离:长白山脉地形破碎化阻碍孢子风媒扩散,黄松甸“人工干扰次生林”与漫江“原始针阔混交林”的生境异质性(宿主、微环境差异)可能切断了与栽培菌株祖先的遗传交流;松山镇-长白山因地理邻近、历史引种(如早期驯化材料采集)可能为主栽菌株提供了遗传物质来源。野生群体间(如黄松甸与漫江)的遗传分化,也指向次级地形/生境隔离对基因流的限制。这种格局既是野生多样性分化的驱动,也为实践指路:保护松山镇长白山野生资源作为潜在育种材

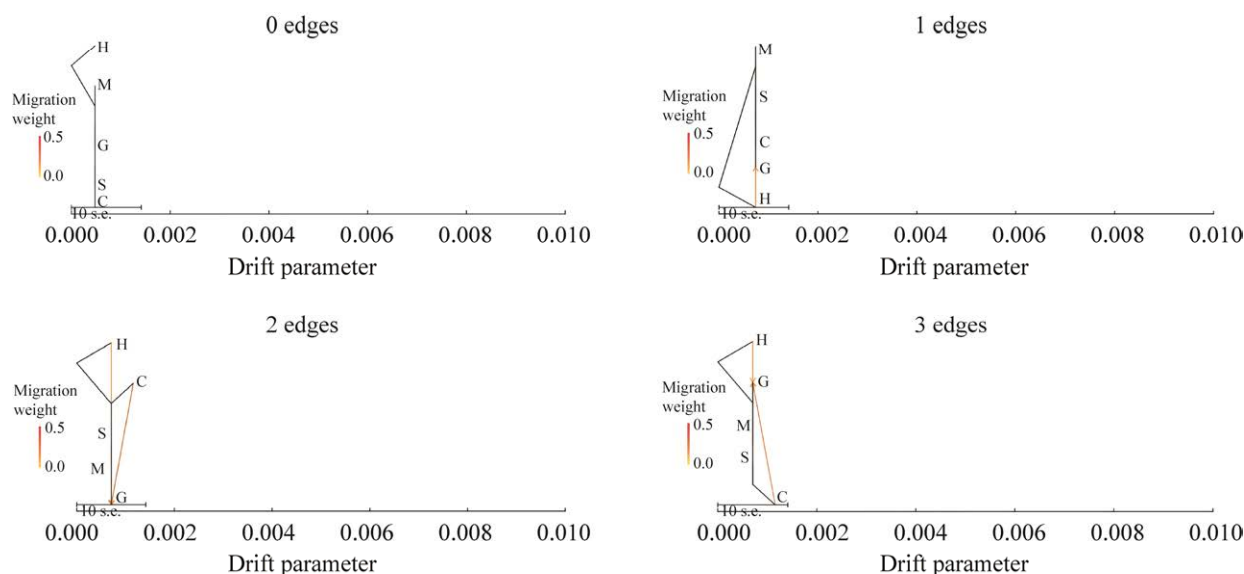


图 7 猴头菇基因流分析 图中的箭头代表群体间发生的基因渗透或迁移事件,即基因流动的方向;热图表示基因渗入的权重,其中箭头颜色越深表示权重越大;每个分支代表一个群体,水平分支的长度表示每个分支已发生的遗传漂变百分比,横坐标代表漂变参数;C代表长白山菌株,H代表松山镇菌株,S代表黄松甸镇菌株,M代表漫江镇菌株,G代表主栽菌株

Fig. 7 Gene flow analysis of *Hericium erinaceus*. Arrows indicate gene introgression or migration events between populations, representing the direction of gene flow. The heatmap illustrates the weight of gene introgression, with darker arrow colors indicating higher weights. Each branch represents a population, and the horizontal branch lengths correspond to the percentage of genetic drift experienced by each branch. The x-axis denotes the genetic drift parameter. Abbreviations: C, Changbai Mountain strains; H, Songshan Town strains; S, Huangsongdian Town strains; M, Manjiang Town strains; G, the cultivated strain.

料库，对黄松甸等高多样性“隔离群体”，需通过主动引种或就地保护维持遗传潜力，避免资源流失。

3 讨论

本研究基于 Illumina NovaSeq 6000 平台对吉林省 4 个地理群体(长白山、松山镇、黄松甸镇、漫江镇)的 34 份野生猴头菇菌株及 1 份主栽菌株(GC)进行全基因组重测序，系统解析了其遗传多样性特征、群体结构及进化动态。通过平均测序深度 57.91X 的高质量数据(Q30>95.65%，覆盖率 $\geq 10X$)，共获得 727 739 个 InDel 位点和 6 244 946 个 SNP 位点，经严格质控筛选出 56 193 个高质量变异位点，覆盖全基因组关键功能区域。相较于传统分子标记技术(如 SRAP、RAPD)，全基因组重测序技术展现出显著优势：周祖法等(2017)基于 SRAP 标记仅能解析 46 个多态性位点，而本研究通过单核苷酸分辨率实现了 5.2 Mb 结构变异的精准识别，检测效率提升近 3 个数量级；与沈颖越等(2020)在金针菇重测序研究中获取的 5.8 万个 SNP 相比，本研究的变异密度更高，且首次定位了猴头菇 2 号染色体高 GC 区域的重组热点，为功能基因挖掘提供了高精度遗传图谱。

群体遗传结构分析显示，吉林省猴头菇野生群体间遗传分化程度较低，遗传分化指数(F_{st})表明不同地理群体间存在频繁的基因交流或生态适应性趋同。然而，主栽菌株 GC 与野生群体间分化显著， F_{st} 值高达 0.079 8–0.090 4，提示人工选育导致栽培菌株基因组显著偏离野生遗传背景。Admixture 分析进一步验证了这一结论：交叉验证确定最优分群数为 $K=2$ ，主栽菌株 GC 独立成簇，与野生群体遗传隔离明显； $K=3$ 分群虽非最优，但仍揭示松山镇菌株 H2、H7、H8 及漫江镇 M3 的独特聚类，暗示局部适应性进化可能形成特异性遗传单元。主成分分析(PCA)与系统进化树结果高度一致，松山镇菌株 H2、H7、H8 及漫江镇 M3 在 PC1–PC3 空间中显著偏离主群，遗传距离峰值达 0.147 5，进一步佐证其遗传独特性。基因流

模拟分析表明，野生与栽培菌株间基因交流受限。当允许迁移事件时，仅检测到松山镇和长白山野生群体向主栽菌株 GC 的弱基因流，而黄松甸镇、漫江镇群体无贡献，可能受地理隔离(如吉林南部与主栽区距离)或生态适应性差异(如温度耐受性)限制。这一结果与 Xiao *et al.* (2016)在香菇中观察到的人工选育加速遗传分化的模式一致，表明高强度性状筛选(如高产、形态定向选择)可能削弱自然基因流的有效性。连锁不平衡(LD)衰减分析进一步支持遗传多样性特征：吉林省野生群体 LD 衰减距离较短(6–15 kb)，LD 值普遍较低($r^2 < 0.3$)，其中漫江镇衰减距离最长(15 kb)，反映其遗传隔离程度较高；长白山衰减距离最短(6 kb)，暗示开放生境(如风媒孢子扩散通道)促进基因交流。野生群体核苷酸多样性(π)平均为 0.219，黄松甸镇群体最高($\pi=0.233$)，可能与生境异质性驱动突变积累相关。猴头菇种质资源的遗传改良需以基因组功能标记开发为核心突破口。本研究基于全基因组重测序技术，系统鉴定了覆盖全基因组的 56 193 个高质量 SNP 位点，较传统 RAPD 标记(王磊等 2009)显著提升了变异检测的密度与分辨率，为关键性状关联分析奠定了数据基础。然而，与香菇(宋琳琳等 2024)、金针菇等物种相比，当前猴头菇种质资源库的遗传背景仍显单一，未来研究需整合我国西南、俄罗斯远东等地理单元的野生种质资源，通过定向杂交引入稀有等位基因以拓宽遗传基础。例如，黄松甸镇野生群体展现出最高的核苷酸多样性($\pi=0.233$)，长白山群体与黄松甸镇群体间遗传分化程度最低($F_{st}=0.022 5$)，二者可作为核心亲本材料，通过群体间杂交打破栽培菌株的遗传瓶颈，恢复因人工选育丢失的适应性变异。

本研究首次揭示了东北主产区野生猴头菇的基因组重排现象，与 Roda *et al.* (2021)提出的结构变异理论框架形成互补，为解析其环境适应性进化机制提供了新视角。例如，推测 2 号染色体 SNP 位点富集区与萜类合成酶基因簇(HE-TPS)共定位，参与低温适应性调控；重组热点与 GC 偏向基因转换(gBGC)的关联性(Lassalle

et al. 2015)进一步揭示了自然选择与中性进化共同塑造基因组结构的复杂性。此外, 主栽菌株 GC 与野生群体间的高 Fst 值(>0.07)及基因流受限特征, 凸显了野生种质资源在抗逆性引入中的不可替代性。这一结论与崔筱等(2020)在糙皮侧耳驯化研究中提出的“野生-栽培基因渗入”策略不谋而合, 为猴头菇遗传改良提供了明确方向。

综上所述, 本研究通过全基因组重测序技术系统解析了吉林省猴头菇种质资源的遗传多样性格局。然而, 当前猴头菇遗传资源仍面临两大挑战: 一是主栽品种遗传背景单一化导致的抗逆性退化风险; 二是野生种质资源收集与评价体系的不完善。未来研究需从三方面突破: 其一, 扩大样本采集范围, 涵盖我国西南高山、俄罗斯远东等多样性热点区域; 其二, 整合转录组、代谢组数据, 阐明关键性状(如多糖合成、低温适应)的分子调控网络; 其三, 构建基于基因组选择的分子育种平台, 加速优良等位基因聚合。这些工作将推动猴头菇产业从“经验育种”向“精准设计育种”转型, 为食用菌遗传资源可持续利用提供范式参考。

作者贡献

马星月: 论文构思及撰写、数据管理; 王天娇: 数据验证、论文审核; 林晓龙、郑佳澜: 提供实验材料和菌种; 姚勋: 数据管理; 陈利丁、孙淑静: 指导论文写作。

利益冲突声明

该研究不存在任何潜在利益冲突的商业或财务关系。

[REFERENCES]

- Alexander DH, Novembre J, Lange K, 2009. Fast model-based estimation of ancestry in unrelated individuals. *Genetics*, 181(3): 1065-1073
- Ayres DL, Cummings MP, Baele G, Darling AE, Lewis PO, Swofford DL, Huelsenbeck JP, Lemey P, Rambaut A, Suchard MA, 2019. BEAGLE 3: improved performance, scaling, and usability for a high-performance computing library for statistical phylogenetics. *Systematic Biology*, 68(6): 1052-1061
- Boddy L, Crockatt ME, Ainsworth AM, 2011. Ecology of *Hericium cirrhatum*, *H. coralloides* and *H. erinaceus* in the UK. *Fungal Ecology*, 4(2): 163-173
- Bolger AM, Lohse M, Usadel B, 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15): 2114-2120
- Chen J, Zeng X, Yang YL, Xing YM, Zhang Q, Li JM, Ma K, Liu HW, Guo SX, 2017. Genomic and transcriptomic analyses reveal differential regulation of diverse terpenoid and polyketide secondary metabolites in *Hericium erinaceus*. *Scientific Reports*, 7(1): 10151
- Cui X, Liu Q, Duan YK, Kang YC, Kong WW, Zhang YT, Yuan RQ, Kong WL, 2020. Genetic diversity analysis of *Pleurotus ostreatus* germplasm resources. *Journal of Henan Agricultural Sciences*, 49(3): 138-144 (in Chinese)
- Danecek P, Auton A, Abecasis G, Albers CA, Banks E, de Pristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, McVean G, Durbin R, 2011. The variant call format and VCF tools. *Bioinformatics*, 27(15): 2156-2158
- Galland LM, Faske TM, Osuna-Mascaró C, Bisbing SM, Parchman TL, 2025. Geography and environment shape spatial genetic variation and predict climate maladaptation across isolated and disjunct populations of *Pinus muricata*. *Molecular Ecology*, <https://doi.org/10.1111/mec.17638>
- Gong WB, Song XY, Xie CL, Zhou YJ, Zhu ZH, Xu C, Peng YD, 2022. Detection of quantitative trait loci underlying fruiting body and yield-related traits in *Hericium erinaceus*. *Scientia Horticulturae*, 293: 110729
- Gong WB, Wang YH, Xie CL, Zhou YJ, Zhu ZH, Peng YD, 2020. Whole genome sequence of an edible and medicinal mushroom, *Hericium erinaceus* (Basidiomycota, Fungi). *Genomics*, 112(3): 2393-2399
- Gong WB, Xie CL, Zhou YJ, Zhu ZH, Wang YH, Peng YD, 2019. A resequencing-based ultradense genetic map of *Hericium erinaceus* for anchoring genome sequences and identifying genetic loci associated with monokaryon growth. *Frontiers in Microbiology*, 10: 3129
- Ke LN, 2008. Molecular polymorphism in germplasm resource of *Tricholoma giganteum masseei*, *Stropharia rugoso-annulata* Farlow, *Hericium erinaceus*, *Agrocybe cylindracea*. MS Thesis, Fujian Agriculture and Forestry University, Fuzhou. 1-164 (in Chinese)
- Lassalle F, Périán S, Bataillon T, Nesme X, Duret L, Daubin V, 2015. GC-content evolution in bacterial genomes: the biased gene conversion hypothesis expands. *PLoS Genetics*, 11(2): e1004941
- Li H, Durbin R, 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, 25(14):

- 1754-1760
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 2009. The sequence alignment/map format and SAMtools. *Bioinformatics*, 25(16): 2078-2079
- Li J, Liu XB, Zhao ZW, Yang ZL, 2019. Genetic diversity, core collection and breeding history of *Pleurotus ostreatus* in China. *Mycoscience*, 60(1): 14-24
- Li JH, Xu SY, Yang ZW, Li D, Zhuang SW, Chen JQ, Qian X, Sun SJ, 2025. Differences analysis of symbiotic bacterial community composition between wild and cultivated strains from *Hericium erinaceus*. *Scientia Horticulturae*, 339: 113868
- Li ML, Zhu AH, Ma GY, Qu Z, Lu XH, Ma HX, Liu ZD, 2025. Biological characteristics and domestication of wild *Pleurotus giganteus*. *Mycosystema*, 44(3): 240255 (in Chinese)
- Lin JB, Yan JJ, Han X, Zhao J, Gan Y, Miao RY, Feng RC, Li F, Lai XF, Tong ZJ, Gan BC, 2023. Genetic diversity and trait analysis of *Pleurotus eryngii* strains based on whole-genome resequencing. *Journal of Southwest Agricultural University*, 36(11): 2311-2318 (in Chinese)
- Lin SL, 2020. Genome-wide and transcriptome-wide association analysis unraveling the genetic mechanisms of natural variation in thousand-seed weight of rapeseed. MS Thesis, Huazhong Agricultural University, Wuhan. 1-98 (in Chinese)
- Liu HX, Wang J, Song WT, Zhu CH, Tao ZY, Xu WJ, Zhang SJ, Li HF, 2021. Genetic evaluation of conservation effects in four populations of Taihu geese. *Journal of Henan Agricultural Sciences*, 50(1): 134-141 (in Chinese)
- López-Goldar X, Agrawal AA, 2021. Ecological interactions, environmental gradients, and gene flow in local adaptation. *Trends in Plant Science*, 26(8): 796-809
- Lu SJ, Dong LD, Fang C, *et al.* (more than 20 authors), 2020. Stepwise selection on homeologous *PRR* genes controlling flowering and maturity during soybean domestication. *Nature Genetics*, 52(4): 428-436
- Lu Z, Wang Y, Ren ZM, Ji SJ, Wang H, Yang Y, Yu YS, 2023. Biological characteristics and cultivation patterns of a wild *Hericium erinaceus* strain. *Special Wild Economic Animal and Plant Research*, 45(5): 28-32, 41 (in Chinese)
- McCormick RF, Truong SK, Sreedasyam A, Jenkins J, Shu S, Sims D, Kennedy M, Amirebrahimi M, Weers BD, McKinley B, Mattison A, Morishige DT, Grimwood J, Schmutz J, Mullet JE, 2018. The *Sorghum bicolor* reference genome: improved assembly, gene annotations, a transcriptome atlas, and signatures of genome organization. *Plant Journal*, 93(2): 338-354
- Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, Lanfear R, 2020. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Molecular Biology and Evolution*, 37(5): 1530-1534
- Pickrell JK, Pritchard JK, 2012. Inference of population splits and mixtures from genome-wide allele frequency data. *PLoS Genetics*, 8(11): e1002967
- Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D, 2006. Principal components analysis corrects for stratification in genome-wide association studies. *Nature Genetics*, 38(8): 904-909
- Roda E, Priori EC, Ratto D, de Luca F, di Iorio C, Angelone P, Locatelli CA, Desiderio A, Goppa L, Savino E, Bottone MG, Rossi P, 2021. Neuroprotective metabolites of *Hericium erinaceus* promote neuro-healthy aging. *International Journal of Molecular Sciences*, 22(12): 6379
- Schiphof K, Kawauchi M, Tsuji K, Yoshimi A, Tanaka C, Nakazawa T, Honda Y, 2024. Functional analysis of basidiomycete specific chitin synthase genes in the agaricomycete fungus *Pleurotus ostreatus*. *Fungal Genetics and Biology*, 172: 103893
- Shen YY, Jin QL, Cai WM, Fan LJ, Feng WL, Song TT, 2020. Genetic diversity and population structure analysis of *Flammulina velutipes* strains based on resequencing data. *Mycosystema*, 39(6): 1016-1028 (in Chinese)
- Song LL, Chen HZ, Wu LL, Li C, Meng L, Kong WL, 2024. Whole-genome sequence analysis of 23 *Lentinula edodes* germplasm resources based on resequencing. *China Cucurbits and Vegetables*, 37(3): 28-34 (in Chinese)
- Sun SJ, Gao W, Lin SQ, Zhu J, Xie BG, Lin ZB, 2006. Analysis of genetic diversity in *Ganoderma* population with a novel molecular marker SRAP. *Applied Microbiology and Biotechnology*, 72(3): 537-543
- Wang L, Su HY, Teng CY, Han LY, Xin CQ, 2009. Phylogenetic study of cultivated *Hericium erinaceus* strains in China. *Food Science*, 30(23): 270-273 (in Chinese)
- White TJ, Bruns TD, Lee SB, Taylor JW, 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. *PCR Protocols: a guide to methods and applications*. Academic Press, New York. 315-322
- Wu F, Zhou LW, Yang ZL, Bau T, Li TH, Dai YC, 2019. Resource diversity of Chinese macrofungi: edible, medicinal and poisonous species. *Fungal Diversity*, 98: 1-76
- Xiao Y, Cheng XJ, Liu J, Li C, Nong WY, Bian YB, Cheung MK, Kwan HS, 2016. Population genomic analysis uncovers environmental stress-driven selection and adaptation of *Lentinula edodes* population in China. *Scientific Reports*, 6: 36789

- Yuan Y, Bian LS, Wu YD, Chen JJ, Wu F, Liu HG, Zeng GY, Dai YC, 2023. Species diversity of pathogenic wood-rotting fungi (Agaricomycetes, Basidiomycota) in China. *Mycology*, 14(3): 204-226
- Zhang C, Li M, Zhang Q, Liu B, Dong Y, 2019. PopLDdecay: a fast and effective tool for linkage disequilibrium decay analysis based on variant call format data. *Bioinformatics*, 35(6): 1083-1085
- Zhang JC, Shen N, Li C, Xiang XJ, Liu GL, Gui Y, Patev S, Hibbett DS, Barry K, Andreopoulos W, Lipzen A, Riley R, He GF, Yan M, Grigoriev IV, Kwan HS, Cheung MK, Bian YB, Xiao Y, 2021. Population genomics provides insights into the genetic basis of adaptive evolution in the mushroom-forming fungus *Lentinula edodes*. *Journal of Advanced Research*, 38: 91-106
- Zhang N, Tang ZF, Zhang J, Li X, Yang ZQ, Yang C, Zhang ZF, Huang ZX, 2019. Comparative transcriptome analysis reveals the genetic basis underlying the biosynthesis of polysaccharides in *Hericiium erinaceus*. *Botanical Studies*, 60(1): 15
- Zhou ZF, Lu N, Song JL, Wang YF, Li HB, Wei HL, 2017. Evaluation of genetic relationships and screening of strain-specific molecular markers in *Hericiium erinaceus*. *Journal of Fungal Research*, 15(3): 188-194 (in Chinese)
- [附中文参考文献]**
- 崔筱, 刘芹, 段亚魁, 康源春, 孔维威, 张玉亭, 袁瑞奇, 孔维丽, 2020. 平菇种质资源的遗传多样性分析. *河南农业科学*, 49(3): 138-144
- 柯丽娜, 2008. 金福菇等四种食用菌种质资源的分子标记多态性研究. 福建农林大学硕士论文, 福州. 1-164
- 李曼玲, 朱安红, 马国营, 屈直, 卢雪花, 马海霞, 刘振东, 2025. 野生巨大侧耳生物学特性及驯化栽培. *菌物学报*, 44(3): 240255
- 林俊彬, 严俊杰, 韩星, 赵金, 甘颖, 苗人云, 冯仁才, 李斐, 赖学飞, 仝宗军, 甘炳成, 2023. 基于全基因组重测序的杏鲍菇菌株遗传多样性与性状分析. *西南农业学报*, 36(11): 2311-2318
- 林升丽, 2020. 全基因组和全转录组关联分析解析油菜千粒重自然变异的遗传机制. 华中农业大学硕士论文, 武汉. 1-98
- 刘宏祥, 王健, 宋卫涛, 朱春红, 陶志云, 徐文娟, 章双杰, 李慧芳, 2021. 太湖鹅 4 个群体保种效果的遗传评价. *河南农业科学*, 50(1): 134-141
- 陆珠, 王月, 任梓铭, 纪淑娟, 王欢, 杨阳, 于延申, 2023. 一株野生猴头菇的生物学特性及栽培模式研究. *特产研究*, 45(5): 28-32, 41
- 沈颖越, 金群力, 蔡为明, 范丽军, 冯伟林, 宋婷婷, 2020. 基于重测序信息的金针菇菌株资源遗传多样性及群体结构分析. *菌物学报*, 39(6): 1016-1028
- 宋琳琳, 陈红芝, 毋柳柳, 李畅, 孟丽, 孔维丽, 2024. 基于重测序的 23 份香菇种质资源全基因组序列分析. *中国瓜菜*, 37(3): 28-34
- 王磊, 宿红艳, 滕长英, 韩立亚, 辛成齐, 2009. 中国猴头菇栽培菌株的系统进化研究. *食品科学*, 30(23): 270-273
- 周祖法, 陆娜, 宋吉玲, 王燕飞, 李海波, 魏海龙, 2017. 猴头菇菌株亲缘关系评价和菌株特异性分子标记筛选. *菌物研究*, 15(3): 188-194