

基于 MNP 标记的暗褐脉柄牛肝菌菌株鉴定与群体遗传研究

孟国良¹, 李嘉嘉¹, 曹瑶², 李凡², 曾雨康²,
VLASENKO Vyacheslav Alexandrovich⁴, 李荣春^{2,3*}, 董彩虹^{1*}

1 中国科学院微生物研究所 微生物多样性与资源创新利用全国重点实验室, 北京 100101

2 云南菌视界生物科技有限公司 云南省级(珍稀食用菌)企业工程技术中心 昆明市级(食用菌)企业工程技术中心, 云南 昆明 650200

3 云南农业大学食用菌研究所, 云南 昆明 650201

4 俄罗斯科学院西伯利亚分院中央西伯利亚植物园, 俄罗斯 新西伯利亚 630090

摘要: 暗褐脉柄牛肝菌 *Phlebopus portentosus* 是牛肝菌中首个实现人工驯化栽培的物种, 近年来其产业化栽培技术逐步成熟, 规模持续扩大, 因此其品种精准鉴定与保护至关重要。本研究对 66 株不同产地的暗褐脉柄牛肝菌菌株进行全基因组重测序, 构建了包含 306 个位点的多核苷酸多态性(MNP)分子标记数据库, 基于单核苷酸多态性 SNP 和 MNP 数据的群体结构与亲缘关系分析显示供试菌株可划分为 2 个主要群体。继代培养不同代数菌株的数据表明 MNP 标记具有良好的稳定性和分辨力, 能实现菌株间的精准鉴定。通过对同一菌株不同批次的重测序数据分析, 建立了克隆菌株判定标准(血缘同源性 PI_HAT>0.95 且遗传距离 D<0.01); 当遗传相似度(GS)达到 100%时, 可将两样本判定为同一菌株。结合机器学习进一步筛选出对群体分化具有指示意义的关键标记位点, 为品种鉴定和功能基因研究奠定基础。本研究构建了暗褐脉柄牛肝菌分子标记数据库, 验证了其在遗传资源保护与分子育种中的应用潜力, 为品种权保护及种质创新提供科技支撑。

关键词: 暗褐脉柄牛肝菌; 群体结构; 多核苷酸多态性; 遗传相似度; 菌株鉴定

[引用本文]

孟国良, 李嘉嘉, 曹瑶, 李凡, 曾雨康, VLASENKO Vyacheslav Alexandrovich, 李荣春, 董彩虹, 2026. 基于 MNP 标记的暗褐脉柄牛肝菌菌株鉴定与群体遗传研究. 菌物学报, 45(3): 250231

Meng GL, Li JJ, Cao Y, Li F, Zeng YK, Vlasenko VA, Li RC, Dong CH, 2026. Strain identification and population genetic analysis of *Phlebopus portentosus* based on MNP markers. Mycosystema, 45(3): 250231

资助项目: 云南省(珍稀食用菌)企业技术中心专项资金(JSJ2023016KJ)

This work was supported by the Special Fund of Yunnan Province (Rare Edible Fungi) Enterprise Technology Center (JSJ2023016KJ).

*Corresponding authors. E-mails: DONG Caihong, dongch@im.ac.cn; LI Rongchun, rongchunli@126.com

Received: 2025-07-25; Accepted: 2025-08-19

Strain identification and population genetic analysis of *Phlebopus portentosus* based on MNP markers

MENG Guoliang¹, LI Jiajia¹, CAO Yao², LI Fan², ZENG Yukang²,
VLASENKO Vyacheslav Alexandrovich⁴, LI Rongchun^{2,3*}, DONG Caihong^{1*}

1 State Key Laboratory of Microbial Diversity and Innovative Utilization, Institute of Microbiology, Chinese Academy of Sciences, Beijing 100101, China

2 Yunnan Junshijie Biotechnology Co. Ltd., Yunnan (Rare Edible Fungi) Enterprise Technology Center, Kunming (Edible Fungi) Enterprise Technology Center, Kunming 650200, Yunnan, China

3 Institute of Edible Fungi, Yunnan Agricultural University, Kunming 650201, Yunnan, China

4 Central Siberian Botanical Garden of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk 630090, Russia

Abstract: *Phlebopus portentosus*, the first successfully domesticated and cultivated bolete species, has undergone rapid industrial expansion in recent years, highlighting the growing need for accurate strain identification and germplasm protection. In this study, whole-genome resequencing was performed on 66 strains of *P. portentosus* collected from various regions, resulting in the construction of a database of multiple nucleotide polymorphism (MNP) markers containing 306 loci. Population structure and kinship analyses based on both single nucleotide polymorphisms (SNPs) and MNPs consistently grouped the strains into two distinct genetic clusters. The MNP markers exhibited high resolution and stability across successive subculturing generations, enabling precise strain discrimination. Clonal identification criteria were established using replicate sequencing data of the same strains, with thresholds set at $PI_{HAT} > 0.95$ and genetic distance (D) < 0.01 , while a genetic similarity (GS) value of 100% within the MNP dataset was validated as a robust indicator of strain identity. Additionally, machine learning approaches were employed to identify key MNP loci indicative of population divergence, offering valuable markers for varietal authentication and functional genomics. This work established MNP molecular marker database for *P. portentosus* and validated the potential applications in genetic resource conservation and molecular breeding, providing scientific support for variety protection and germplasm innovation.

Keywords: *Phlebopus portentosus*; population structure; multiple nucleotide polymorphism; genetic similarity; strain identification

我国食用菌产业正加速由资源依赖型向技术驱动型转型,品种选育与创新已成为推动产业高质量发展的核心驱动力(鲍大鹏等 2022; 戴玉成 2022; 谷莉等 2025)。近年来,随着品种选育与引种工作的深入推进,品种命名混乱问题日益突出,“同物异名”“同名异种”等现象频发,不仅增加了生产管理难度,也制约了种质资源的有效利用与创新(孔祥会等 2019; 谷莉等 2025)。准确的品种鉴定有助于栽培者科学管理、育种者权益保护及科技成果转化,是实施知识产权保护的前提。自 1999 年起,我国陆续将 15 个食用菌种(属)纳入《植物新品种保护名录》,逐步建立了较为完善的品种权保护制度(吕志文等 2024; 谷莉等 2025)。但传统鉴定方法如对

峙培养、随机扩增多态性(random amplified polymorphic DNA, RAPD)、间隔重复序列扩增(inter-simple sequence repeat, ISSR)和同工酶电泳等,因主观性强、重复性差,已难以满足现代品种保护对精准性与客观性的要求(Lee *et al.* 2017; 刘飞等 2025)。

随着高通量测序技术的发展,基于 DNA 多态性的分子标记方法因其多样性丰富、稳定性高,逐渐成为品种鉴定的重要工具(Goodwin *et al.* 2016; Bayer *et al.* 2020)。其中,单核苷酸多态性(single nucleotide polymorphism, SNP)和插入缺失变异(insertion-deletion, InDel)是目前应用较多的两类标记(Zou *et al.* 2014; Li *et al.* 2023; Nizamani *et al.* 2023)。SNP 依赖全基因组

重测序(whole genome re-sequencing, WGS)技术筛选标记位点,具有多态性好、分辨率高等特点(Mammadov *et al.* 2012; 黄杰等 2015; 沈秀芬等 2020),但是存在数据分析难度大,假阳性率较高的问题(Eren *et al.* 2023; Kotlarz *et al.* 2024; 刘飞等 2025)。InDel 标记变异形式清晰、操作简便,分布密度高于 SSR 但低于 SNP,已应用于香菇等食用菌品种的鉴定(Xiang *et al.* 2016; Yamaguchi *et al.* 2023),然而,其准确性和开发效率受限于基因组重复序列、结构特征及检测方法的可靠性(Li 2014; Bennett *et al.* 2020)。尽管 SNP 和 InDel 为食用菌品种鉴定提供了新的技术支撑,但在实际应用中仍面临标准化不足与操作复杂等挑战,建立科学、高效、可推广的分子标记体系仍是亟需攻克的关键问题。

多核苷酸多态性(multiple nucleotide polymorphism, MNP)标记法作为一种整合多位点 SNP 的高通量分子检测手段,通过在同一 DNA 片段内组合多个 SNP,可在理论上产生多达 2^n 种等位基因型,显著提升多态性和分辨能力(彭海等 2020; 徐云碧等 2020; 贾定洪等 2024)。相比 SSR 和 SNP 标记, MNP 在多态性、检测准确性和重现性方面均具有优势,依托高通量测序和多重 PCR,克服了凝胶电泳和微阵列的局限,实现了流程简化、通量提升与成本降低(Ling *et al.* 2022; 吕志文等 2024)。近年来,该技术已在水稻、木薯、棉花等作物品种的精准鉴定中取得良好应用(彭海等 2020; 万人静等 2023; 艾莎等 2024)。在食用菌领域, MNP 标记同样显著提高了香菇 *Lentinula edodes*、金针菇 *Flammulina filiformis* 和刺芹侧耳 *Pleurotus eryngii* 等品种的鉴定准确性,为遗传多样性分析和品种保护提供了有力工具(Ling *et al.* 2022; 魏传正等 2023; 吕志文等 2024)。近期,研究开发了基于核心基因关联多核苷酸多态性(core gene-associated MNP, cgMNP)标记的新方法,筛选出的 84 个 cgMNP 标记不仅在双孢蘑菇 *Agaricus bisporus* 中实现了高分辨率菌株鉴定,还在金针菇 *F. filiformis*、酵母 *Saccharomyces cerevisiae* 等真菌中展现了良好适用性,为菌株水

平的精确区分提供了可靠手段(Liu *et al.* 2025)。

暗褐脉柄牛肝菌 *Phlebopus portentosus* (Berk. & Broome) Boedijn 是一种广泛分布于热带和亚热带地区的珍稀食用菌,作为已实现人工规模化栽培的牛肝菌类群,其成功驯化打破了传统依赖野生采集的局限,具有重要的经济和生态意义(李凡等 2024)。尽管我国已在其资源调查、人工栽培和组学研究等方面取得初步进展,并利用 ITS 序列和 SSR 分子标记分析了其遗传多样性,发现国内菌株存在一定程度的遗传分化(李凡等 2025),但由于种质资源分布广泛、生态多样,现有分子标记在菌株鉴定和群体划分中的分辨率和效率仍有限,难以满足优良菌株选育需求(曹昉等 2017; 高锋等 2021)。本研究以不同地理来源的暗褐脉柄牛肝菌菌株为材料,基于基因组重测序数据构建 MNP 分子标记位点数据库,并用于该物种的菌株鉴定和群体研究。

1 材料与amp;方法

1.1 菌株及测序数据来源

本研究所用暗褐脉柄牛肝菌菌株采集自中国云南、广东及福建等地区,基因组重测序数据均由本研究完成,具体信息详见表 1。

1.2 DNA 提取、重测序及数据质量控制

用于基因组测序的菌株接种于 M1 液体培养基(马铃薯 200 g/L,葡萄糖 20 g/L,酵母提取物 2 g/L,硫酸镁 1 g/L,磷酸二氢钾 1 g/L),28 °C、120 r/min 振荡培养 10 d,收集菌丝体,用无菌去离子水冲洗两次。经滤纸吸干多余水分后,菌丝体迅速在液氮中速冻,并储存于-80 °C 备用。

基因组 DNA 的提取参照改良的 CTAB 法(傅俊生 2007),使用 Nanodrop 微量分光光度计和 1.0%琼脂糖凝胶电泳评估 DNA 的浓度和完整性。样品经质检合格后,委托武汉未来组生物科技有限公司进行测序,测序平台为 MGI。原始测序数据使用 fastp (Chen *et al.* 2018)进行质控分析,去除接头和低质量序列得到 clean data,并进行碱基测序准确率(Q30)分析,测序数据 Q30>85%的样品用于后续分析。

表 1 暗褐脉柄牛肝菌菌株及测序数据信息

Table 1 Information of strains and sequencing data of *Phlebopus portentosus*

菌株 Strain	地点 Sampling location	碱基测序准确率 Q30	测序数据量 Total bases/G	Reads 成功比对映射率 Reads mapping rate/%	平均深度 Mean depth
1701	景洪市 Jinghong	90.07	4.14	86.19	113.17
1703	昆明市 Kunming	96.62	7.42	90.10	211.01
1706	深圳市 Shenzhen	96.72	8.74	91.06	249.52
1707	广州市 Guangzhou	96.17	4.78	94.11	141.39
1708	广州市 Guangzhou	92.48	5.99	92.10	174.20
1709	南宁市 Nanning	96.00	6.80	95.33	204.34
1711	景洪市 Jinghong	94.08	4.26	93.04	125.75
1712	个旧市 Gejiu	94.32	6.81	91.15	194.02
1713	景洪市 Jinghong	92.74	5.08	86.57	135.59
1714	景洪市 Jinghong	93.71	6.59	89.97	184.72
1715	景洪市 Jinghong	91.73	4.24	94.30	126.67
1716	景洪市 Jinghong	94.45	7.68	86.79	208.59
1717	景洪市 Jinghong	92.44	7.44	86.74	197.10
1718	景洪市 Jinghong	95.08	5.26	84.82	138.52
1719	景洪市 Jinghong	94.72	5.14	88.54	143.07
1720	景洪市 Jinghong	95.20	7.68	92.73	223.28
1721	景洪市 Jinghong	92.71	4.56	91.49	131.96
1722	普洱市 Pu'er	93.17	3.08	95.31	93.21
1723	普洱市 Pu'er	93.72	5.53	88.59	154.66
1724	普洱市 Pu'er	95.21	9.65	90.79	273.56
1725	普洱市 Pu'er	94.63	4.99	94.06	146.88
1726	普洱市 Pu'er	92.89	3.74	88.53	101.55
1727	普洱市 Pu'er	93.17	4.83	96.13	147.19
1728	普洱市 Pu'er	94.60	4.39	91.89	127.05
1729	景洪市 Jinghong	93.61	6.79	93.93	195.63
1730	普洱市 Pu'er	92.91	4.40	94.26	131.20
1731	普洱市 Pu'er	94.53	3.80	93.08	111.43
1732	昆明市 Kunming	93.61	4.52	92.93	130.77
1733	昆明市 Kunming	93.67	7.14	85.27	184.92
1735	楚雄市 Chuxiong	94.21	4.26	94.43	126.42
1736	昆明市 Kunming	95.09	5.71	92.56	165.76
1737	禄丰市 Lufeng	94.92	4.39	92.32	127.40
1738	禄丰市 Lufeng	93.54	4.17	93.50	123.09
1739	曲靖市 Qujing	94.28	4.95	83.60	127.79
1740	个旧市 Gejiu	95.20	9.08	89.60	251.31
1741	普洱市 Pu'er	94.91	6.17	91.44	173.70
1742	文山市 Wenshan	94.93	4.87	91.71	140.91
1743	香格里拉市 Shangri-La	94.65	6.13	96.59	187.10
1744	楚雄市 Chuxiong	92.84	5.81	90.36	165.55
1745	普洱市 Pu'er	95.77	3.87	88.91	107.61
1746	普洱市 Pu'er	93.59	4.58	93.12	135.21
1747	普洱市 Pu'er	96.26	5.38	93.56	160.04
1748	普洱市 Pu'er	93.41	3.97	93.48	117.69
1749	普洱市 Pu'er	94.59	6.05	92.20	175.19

(待续)

(续表 1)

菌株 Strain	地点 Sampling location	碱基测序准确率 Q30	测序数据量 Total bases/G	Reads 成功比对映射率 Reads mapping rate/%	平均深度 Mean depth
1750	福州市 Fuzhou	93.63	4.40	92.11	127.53
1751	景洪市 Jinghong	96.01	5.70	91.45	166.73
1752	普洱市 Pu'er	94.44	4.88	92.05	141.16
1753	文山市 Wenshan	92.92	4.02	88.56	111.84
1763	文山市 Wenshan	96.67	5.76	93.69	170.63
1764	文山市 Wenshan	94.36	7.51	93.24	219.78
1765	文山市 Wenshan	94.42	5.28	89.39	148.35
1766	文山市 Wenshan	96.28	5.23	91.27	149.79
1767	文山市 Wenshan	95.90	4.85	94.19	143.28
1768	文山市 Wenshan	96.20	5.72	95.57	172.69
1769	文山市 Wenshan	96.28	4.82	94.09	143.48
1770	文山市 Wenshan	96.39	5.68	96.18	170.39
1771	文山市 Wenshan	96.55	6.03	94.25	179.56
1772	文山市 Wenshan	96.29	5.60	90.92	160.82
1773	昆明市 Kunming	96.14	5.20	93.33	148.30
1774	福州市 Fuzhou	96.33	8.98	91.35	257.78
1775	福州市 Fuzhou	96.56	4.82	94.30	143.68
1776	景洪市 Jinghong	96.67	6.22	97.62	188.32
1777	景洪市 Jinghong	96.60	6.06	93.41	178.78
1778	福州市 Fuzhou	96.39	4.52	94.48	135.90
1779	福州市 Fuzhou	94.72	8.89	90.96	252.89

1.3 SNP 变异位点检测

以菌株 PP78 (通过原生质体单核化技术从菌株 1701 分离获得的单倍体菌株)的染色体级别基因组作为参考基因组(PRJNA1269926), 使用 Bowtie2 v2.5.1 (Langmead & Salzberg 2012)、Samtools v1.17 (Li *et al.* 2009)构建参考序列索引, GATK v4.5.0.0 (McKenna *et al.* 2010)构建参考序列字典文件。使用 Bowtie2 v2.5.1 和 Samtools v1.17 将测序数据比对至参考基因组生成 bam 文件, 并利用 Samtools v1.17 检测测序深度和比对率。

使用 Picard 标记 PCR 重复, 使用 GATK v4.5.0.0 的 AddOrReplaceReadGroups 添加样本信息, HaplotypeCaller 模块对每个样本每个位点进行分型, 产生 gvcf 文件, 所有样本数据经 CombineGVCFs 合并, 并通过 GenotypeGVCFs 利用合并后的 gvcf 文件进行变异检测, 使用 GATK SelectVariants 筛选出 SNP 数据。使用 VariantFiltration 对 SNP 位点进行过滤: $QD < 2.0$ || $MQ < 40.0$ || $FS > 60.0$ || $SOR > 3.0$ || $MQRankSum <$

-12.5 || $ReadPosRankSum < -8.0$, 利用 SelectVariants 提取过滤后的 SNP。使用 BCFtools v1.8 (Danecek *et al.* 2021)对变异位点进行基于深度和质量的过滤, 筛选条件为 “ $FMT/DP < 50$ || $FMT/DP > 300$ || $FMT/GQ < 30$ ”。进一步去除所有样本中基因型均缺失的位点, 最终构建高质量的 SNP 数据集。

1.4 基于 SNP 的群体结构及亲缘关系分析

为进行群体结构划分及相关分析, 首先使用 VCFtools v0.1.16 (Danecek *et al.* 2011)过滤缺失率大于 5%和最小等位基因频率(minor allele frequency, MAF)小于 0.01 的 SNP 位点。随后通过 PLINK v1.90 (Purcell *et al.* 2007)将 VCF 文件转为 PLINK 格式, 并使用参数 indep-pairwise 50 10 0.2 进行连锁不平衡(linkage disequilibrium, LD)筛选, 剔除高连锁不平衡($r^2 > 0.2$)的位点, 获得适用于群体结构分析的低 LD 的 SNP 子集。

使用 PLINK v1.90 对过滤后的 SNP 数据进行格式转换, 使用主成分分析(principal component analysis, PCA), 计算前 20 个主成分对群体结构分析, 使用 R v4.2.2 对 PCA 结果进行可视化展

示; 使用 ADMIXTURE v1.3.0 (Alexander & Lange 2011)在预设 K 值范围内运行群体遗传结构推断,每个 K 值重复 50 次以评估模型稳定性,并基于最小交叉验证误差(CV error)确定最优 K 值, R v4.2.2 进行可视化; 使用 VCFtools v0.1.16 将 SNP 数据转化为 PHYLIP 格式,通过 PhyloSuite v1.2.3 (Zhang *et al.* 2020)中集成的 ModelFinder v1.6.1 (Kalyaanamoorthy *et al.* 2017)筛选最佳替代模型,并基于 RAxML-NG (Kozlov *et al.* 2019)使用最大似然法(maximum likelihood, ML)构建系统发育树,设定 1 000 次 bootstrap 重复以评估节点支持率, FigTree v1.4.3 进行可视化。

使用 Plink v1.90 计算血缘同源性(identity by descent, IBD)及遗传距离(genetic distance, D),以分析不同菌株间的亲缘关系和遗传差异,从而判断是否属于同一菌株。

1.5 MNP 标记库的构建及遗传相似性分析

使用 TBtools v2.2.10 (Chen *et al.* 2020)中的 SSR Miner 识别基因组的 SSR 区域,利用 Python 脚本筛选没有 SSR、没有连续 SNP (连续 3 个及以上位点含有 SNP)的数据,以 125 bp 为窗口、平移距离为 20 bp, 计算出窗口内多态性信息量(polymorphism information content, PIC)及区分度(discrimination power, DP), 筛选“SNP>2, DP>0.2, PIC>0.5”的窗口, 位点之间的间距大于 50 kb, 且每个窗口内包含 2–10 个 SNP 位点, 优先保留高 DP 值的窗口, 相同 DP 值下保留 PIC 值更高的窗口, 若仍相同, 优先 SNP 数更多的窗口。将满足上述条件的窗口定义为一个 MNP 位点。使用 TBtools v2.2.10 绘制 MNP 位点染色体分布图。

采用 Python 脚本提取各菌株的 MNP 基因型, 将菌株两两比对, 统计 MNP 基因型相同的数量(n)和 MNP 标记总数(N), 计算两菌株间的遗传相似度(genetic similarity, GS), $GS=n/N \times 100\%$ 。使用 R v4.2.2 对结果进行可视化。

1.6 测序数据量对 MNP 识别准确率的影响

以菌株 1738 为测试菌株, 委托武汉未来组生物科技有限公司进行测序, 测序平台为 MGI, 测序数据量为 38 G。使用 GATK v4.5.0.0 随机抽

取测序 reads 构建 0.5、0.6、0.7、0.8、0.9、1、2–15、20、25、35 G 共 23 个处理。将各处理下识别到的 MNP 位点基因型与 MNP 标记数据库中菌株 1738 的参考基因型进行逐位比对, 统计各处理下 MNP 识别的准确率(准确识别的 MNP 数/MNP 标记库中 MNP 标记总数)。

1.7 继代培养菌株 MNP 稳定性

选取菌株 1738、1741 不同继代次数的菌株, 时间间隔分别为 3 个月(第 3 代)和 9 个月(第 8 代), 进行 DNA 提取与重测序。所有测序数据经过严格的质量控制, 进行 SNP 变异位点分析, 并与 MNP 标记数据库中相应菌株的 MNP 标记进行比对。通过比较不同继代周期下的 MNP 基因型, 评估 MNP 标记的稳定性, 并分析其对基因型变化的潜在影响。

1.8 MNP 位点注释和功能分类分析

基于全基因组注释文件(GFF), 利用 BEDTools (Quinlan & Hall 2010)将已鉴定的 MNP 位点与基因结构区域进行比对, 判断其是否位于基因间区或基因区域(包括外显子和内含子)。对于定位于基因区域的 MNP 位点, 提取其对应的蛋白, 进一步使用功能注释数据库 Pfam (Finn *et al.* 2014)、InterProScan v5.0 (Jones *et al.* 2014)和 eggNOG-mapper v2 (Cantalapiedra *et al.* 2021)进行功能注释。

1.9 机器学习构建预测模型

基于构建的 MNP 标记库, 结合群体划分和 MNP 位点对应基因的功能注释, 建立集成化机器学习分析框架。首先将 MNP 原始基因型数据(如 0|0、0|1、1|1、0|2 等)转化为数值编码(0–9), 生成特征矩阵。整合群体标签和功能注释类别, 采用随机森林构建监督式学习模型, 通过交叉验证筛选关键功能型 MNP 位点, 建立群体划分与功能分类的预测模型。

2 结果与分析

2.1 暗褐脉柄牛肝菌群体结构分析

本研究对不同产地来源的 66 个暗褐脉柄牛肝菌菌株进行重测序分析, 测序数据 Q30 在 90.07%–96.72% (表 1), 可用于后续分析。以高质

量的暗褐脉柄牛肝菌基因组(PRJNA1269926)为参考, 66个菌株的重测序数据共筛选到2 394 282个SNP变异位点, 通过过滤及剔除高连锁不平衡位点后, 共得到20 738个高质量SNP位点。

基于SNP数据采用主成分分析(PCA)、ADMIXTURE以及系统发育树分析对群体结构进行了全面解析。PCA结果显示, 66个菌株明显分为2个遗传群体(图1A)。PC1解释了主要的遗传变异(52.26%), 两组样本呈现明显分离, 表明存在显著的群体结构。2个群体几乎没有重叠, 表明在SNP水平差异较大。每组内部样本分布紧密, 反映出较高的群体一致性。部分样本(如1701、1703、1708、1770等)偏离主群体, 可能存在独特的遗传背景或进化路径。基于ADMIXTURE的群体结构分析, 通过交叉验证误差(CV error)评估不同K值下模型的拟合效果, 结果表明K=2时CV error最小(图1B), 即该模型在此K值下拟合最佳。基于K=2的ADMIXTURE结构图进一步确认66个菌株可分为2个群体(图1C), 这一结果与SNP分析构建的系统发育树分支结构相吻合, 显示出相似的群体划分(图1D)。整合3种分析方法, 均证实本次研究所收集的暗褐脉柄牛肝菌菌株主要分为2个类群。

2.2 暗褐脉柄牛肝菌菌株亲缘关系

为评估样本之间的亲缘关系及克隆菌株的存在, 首先计算了所有样本对之间的遗传距离($D=1-DST$)与血缘同源性(identity by descent, IBD)即PI_HAT值, 并构建了对称热图以全面展示其基因型相似性与聚类结构(图2A)。聚类结果显示样本之间的PI_HAT值为0.00–0.99, 遗传距离D值为0.00–0.35, 大多数样本之间存在一定的遗传距离, 呈现清晰的群体分化结构, 与群体结构分析结果相一致(图1)。热图中少数样本对具有较高的PI_HAT与较小的D值, 提示存在潜在的克隆样本。

为进一步明确识别同一菌株的判定标准, 增加相同菌株(1738、1741)不同批次的重测序数据(A1738、A1741), 绘制了PI_HAT值与遗传距离D值的散点图(图2B)。结果表明, 绝大多数样本

对的PI_HAT值较低($PI_HAT < 0.3$)且遗传距离较大($D > 0.15$), 符合非克隆、非亲缘个体的预期。而部分样本对(红色点)集中分布在 $PI_HAT > 0.95$ 且遗传距离接近0 ($D < 0.01$)的位置, 菌株1738、1741的不同批次重测序数据同样与这些样本集中在一起, 表明位于该区域的样本对菌株(1701–1708、1721–1728、1733–1734、1736–1737、1753–1771–1772)均属于高度一致的克隆菌株, $PI_HAT > 0.95$, $D < 0.01$ 可作为有效区分是否为同一菌株的标准。

2.3 暗褐脉柄牛肝菌MNP分子标记数据库构建

选取过滤后的高质量SNP位点, 进一步筛选没有SSR、没有连续SNP的数据, 共得到17 253个SNP位点用于MNP分子标记数据库的构建, 通过层级筛选, 成功筛选出306个MNP分子标记位点, 包含1 475个SNP位点。所有MNP分子标记在基因组上分布较为均匀(图3A), 平均每个MNP分子标记含有4.82个SNP, MNP分子标记的PIC值分布范围为0.50–0.74, 平均值为0.63(图3B), MNP标记在不同个体间的变异性较大, 能够提供足够的信息用于群体遗传分析。DP值分布范围为0.21–0.96, 平均值为0.76(图3C), 表示这些MNP标记具有较好的检测深度和覆盖度, 能够准确反映不同个体之间的遗传差异。

2.4 基于MNP分子标记数据库的暗褐脉柄牛肝菌菌株间遗传相似度分析

基于MNP分子标记数据库对66个暗褐脉柄牛肝菌菌株进行遗传相似度分析。结果表明, 该方法可将菌株划分为2个群体(图4), 划分结果与基于SNP的群体结构高度一致。此外, 在亲缘关系分析中判定为克隆菌株的样本(如1701–1708、1721–1728、1733–1734、1736–1737、1753–1771–1772), 在本分析中相似度均为100%, 进一步验证了其一致性。除克隆菌株外, 其余样本的相似度分布在12.5%–52.1%之间, 其中Group A各菌株间相似度范围为19.6%–48.6%, 集中于25.0%–40.0%; Group B为22.5%–52.1%, 集中于30.0%–45.0%。不同群体间的相似度集中在12.5%–32.0%, 不同群体之间的相似度明显低于群体内部。

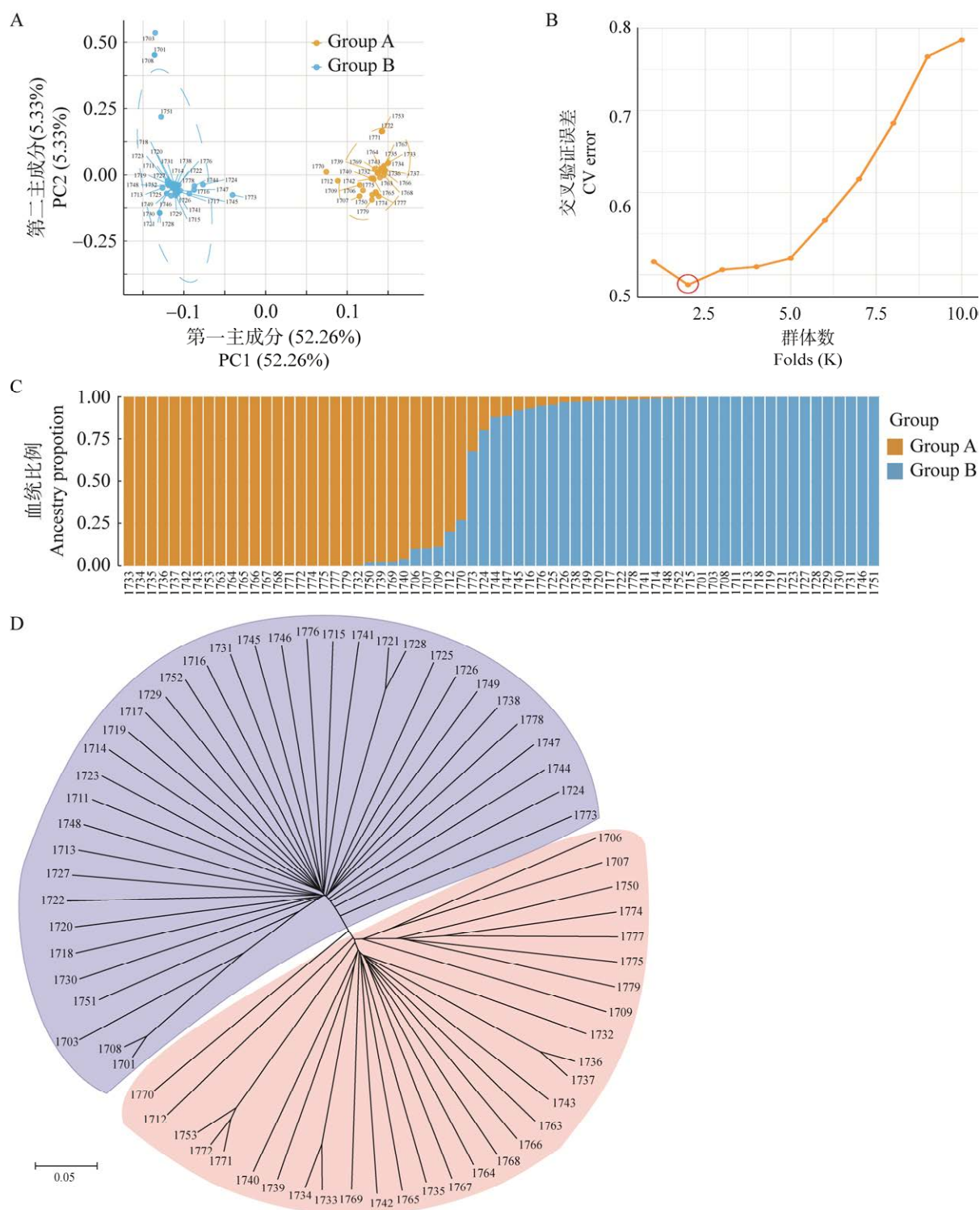


图 1 暗褐脉柄牛肝菌群体结构分析 A: 基于主成分分析(PCA)的群体结构分析; B: 交叉验证误差评估不同群体数(K值)下模型拟合效果, 红色圈代表最佳拟合; C: 基于 K=2 的 ADMIXTURE 群体结构分析; D: 基于 SNP 的系统发育分析, 蓝色和红色分别代表不同群体

Fig. 1 Population structure analysis of *Phlebotomus portentosus*. A: Population structure analysis based on principal component analysis (PCA); B: Cross-validation error evaluating model fit across different numbers of populations (K-values); The red circle indicates the optimal K-value with the best model fit; C: Population structure analysis based on ADMIXTURE with K=2; D: Phylogenetic analysis based on SNPs; Blue and pink representing different populations.

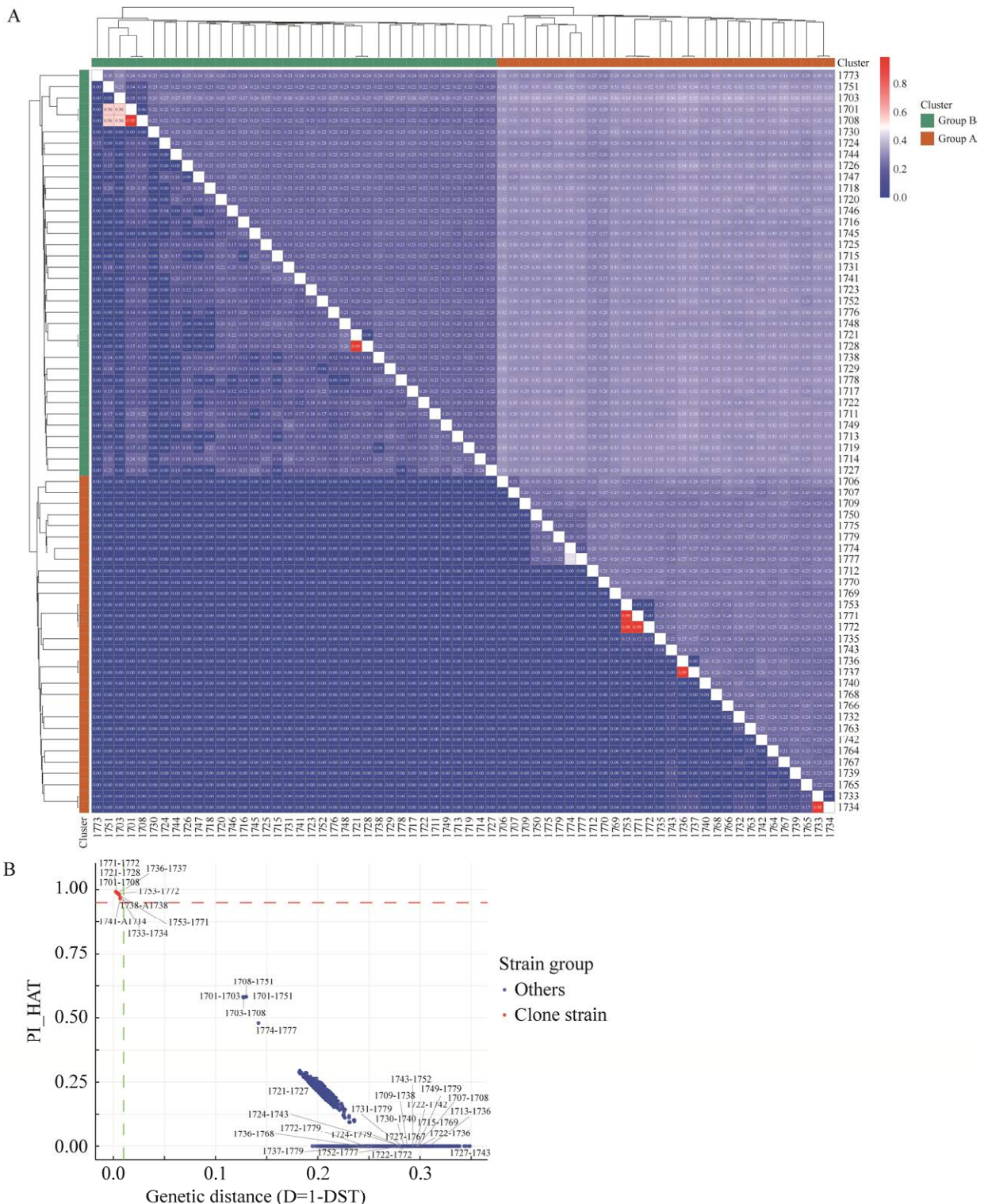


图 2 暗褐脉柄牛肝菌亲缘关系与遗传距离分析 A: 亲缘关系与遗传距离对称性热图, 上三角为样本间遗传距离配对, 下三角为样本间亲缘关系配对; B: 亲缘关系与遗传距离配对关系图, 红色虚线代表 $PI_HAT > 0.95$, 绿色虚线代表 $D < 0.01$, A1738、A1741 代表同一菌株的不同测序批次

Fig. 2 Analysis of kinship and genetic distance among *Phlebobus portentosus* strains. A: Symmetric heatmap of pairwise kinship coefficients (Proportion of alleles identical by descent, PI_HAT , lower triangle) and genetic distances (D , upper triangle); B: Scatter plot showing the relationship between pairwise kinship coefficients (PI_HAT) and genetic distances. The red dashed line represents $PI_HAT > 0.95$, and the green dashed line represents $D < 0.01$. A1738 and A1741 refer to different sequencing batches of the clone strain.

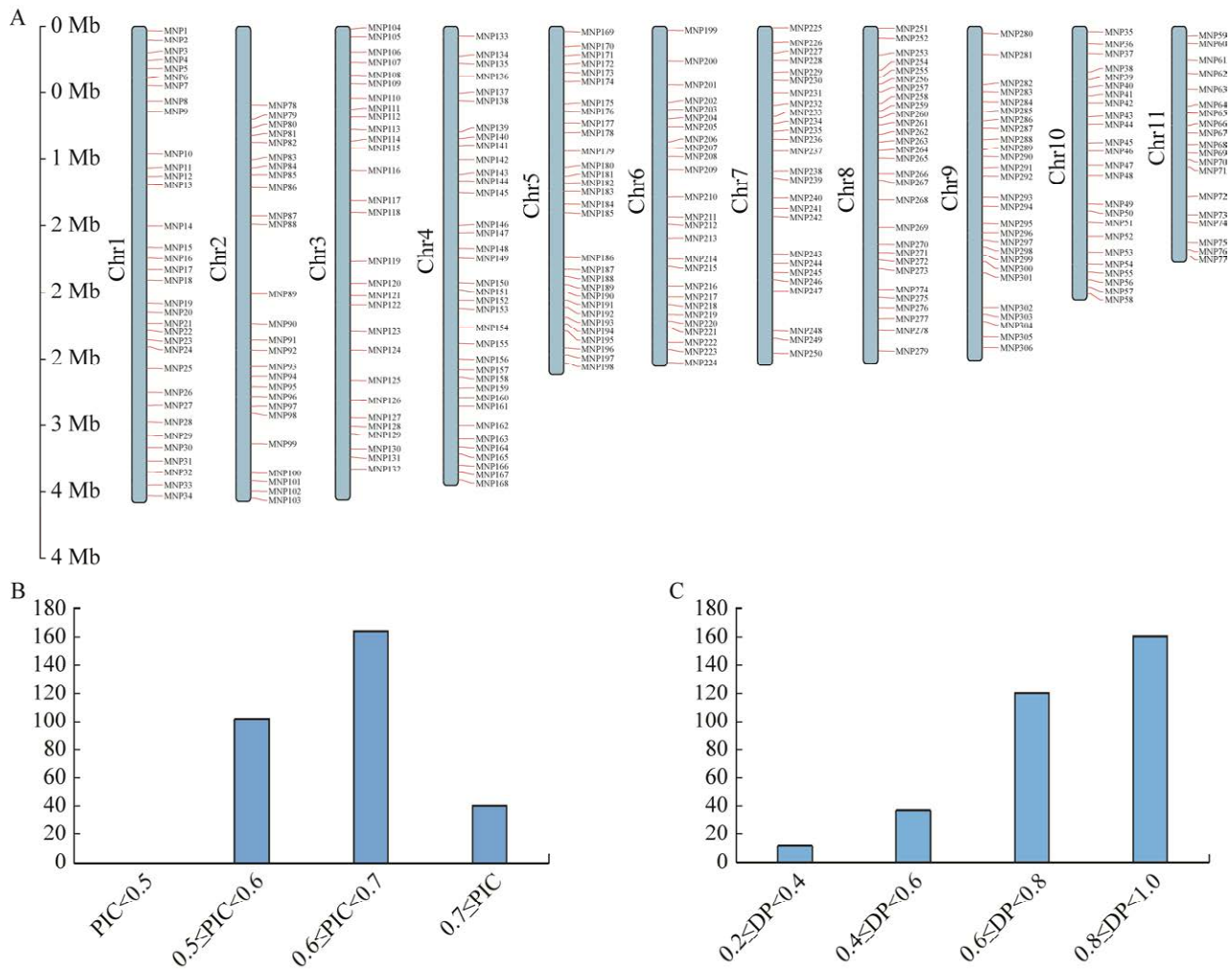


图 3 暗褐脉柄牛肝菌 MNP 分子标记数据库的构建 A: MNP 分子标记在染色体上的分布; B: MNP 分子标记在 66 个暗褐脉柄牛肝菌菌株中的多态性信息量; C: MNP 分子标记在 66 个暗褐脉柄牛肝菌菌株中的区分度
Fig. 3 Construction of the *Phlebotomus portentosus* MNP molecular marker database. A: Distribution of MNP molecular markers on the chromosomes; B: Polymorphic information content (PIC) analysis of MNP molecular markers in 66 *P. portentosus* strains; C: Discriminatory power (DP) analysis of MNP molecular markers in 66 *P. portentosus* strains.

2.5 MNP 分子标记数据库的准确性和稳定性验证

为验证 MNP 分子标记数据库的准确性与稳定性,本研究从不同维度进行评估。首先,在不同测序数据量(0.5 G–35 G)下, MNP 识别准确度随数据量增加而显著提高,在数据量达到 4 G (测序深度约为 59.77×)时,准确度达到 100% (图 5A)。该结果表明,对于暗褐脉柄牛肝菌样本,重测序数据量达到 4 G,即可确保 MNP 分子标记的准确性和稳定性。

其次,评估了继代培养对 MNP 分子标记稳定性的影响。选取菌株 1738 和 1741,在第 0 代、第 3 代和第 8 代进行 3 次独立测序,间隔时间为

0–9 个月。分析结果表明,不同继代次数样本 MNP 识别结果高度一致,未发现基因型变化 (图 5B),表明 MNP 分子标记在短期继代过程中具有较好的遗传稳定性,适用于菌株鉴定和长期遗传监测。

2.6 MNP 分子标记位点注释和功能分类分析

对经过验证的 MNP 分子标记数据库中的 306 个位点进行注释与功能分类分析,结果表明,这些 MNP 位点在基因组中呈现多样化分布。其中,49% (150 个位点)定位于基因间区,3% (10 个位点)处于基因区与间区交界的复合区域;而其余 48% (146 个位点)位于基因区内

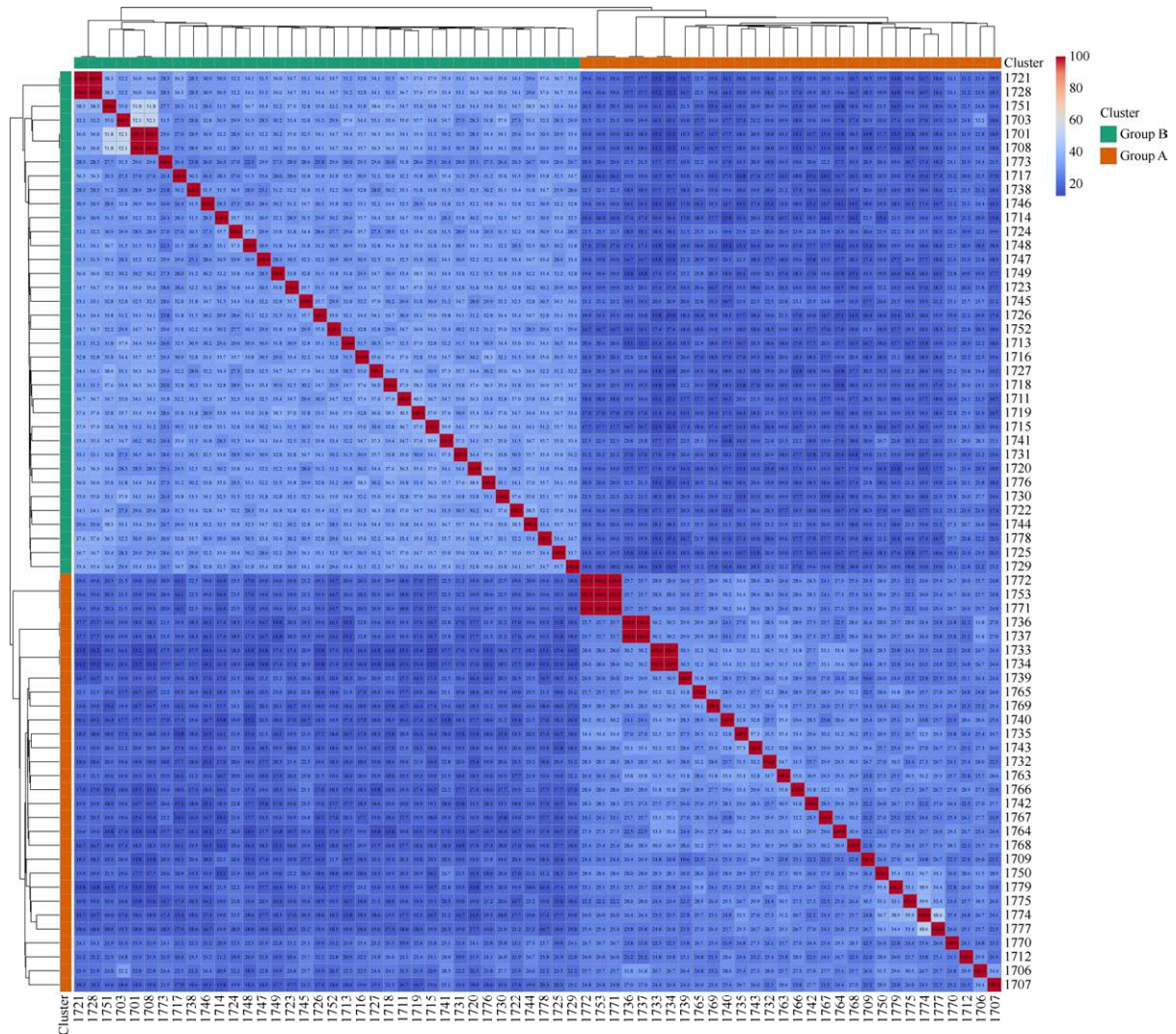


图 4 基于 MNP 分子标记数据库的菌株间遗传相似性分析
 Fig. 4 Genetic similarity analysis among strains based on the MNP molecular marker database.

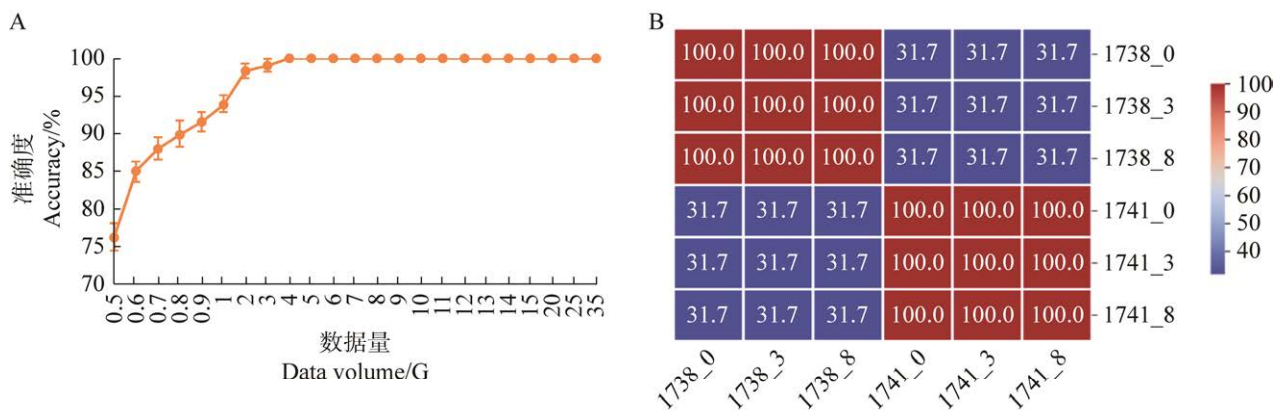


图 5 MNP 分子标记数据库的准确性与稳定性验证 A: 测序数据量对 MNP 标记检测的准确度和稳定性的影响; B: 继代培养对 MNP 标记检测的准确度和稳定性的影响
 Fig. 5 Validation of the accuracy and stability of the MNP molecular marker database. A: Impact of sequencing data volume on the accuracy and stability of MNP marker detection; B: Impact of successive subculture on the accuracy and stability of MNP marker detection.

部(图 6A)。在基因区内的位点中, 35% (106 个位点)位于外显子区域, 6% (19 个位点)位于内含子区域, 另有 7% (21 个位点)同时分布于外显子与内含子上(图 6A)。这些结果表明, MNP 位点在基因组中分布广泛, 既包括对基因表达可能具有直接影响的编码区, 也涵盖了潜在调控区域如基因间区和基因边界区。这种分布特征提示 MNP 标记在遗传研究、功能基因定位及分子育种等领域具有广泛的应用潜力。

对位于基因区域的 MNP 位点进行功能注释。共有 156 个涉及基因区域的 MNP 位点, 其中 33 个位点(21%)对应的基因功能被标注为假定蛋白, 其余的 MNP 位点所对应的基因功能大致可分为 5 类: 代谢相关酶类、基因表达与信号转导调控、膜运输与细胞结构、核糖体功能与蛋

白质加工、转座元件与可动遗传因子(图 6B)。

2.7 融合 MNP 和机器学习的群体划分及功能分类的预测模型

基于 MNP 分子标记数据库(包括 306 个 MNP 标记位点), 结合 1 475 个基因型数据和 156 个位点的功能注释, 以及 SNP 与 MNP 位点群体划分的结果(其中 30%的数据用于训练), 本研究通过机器学习方法成功构建了暗褐脉柄牛肝菌的群体划分模型, 准确率达到了 100%, 且精确率、召回率和 F1-score 指标均表明其分类性能极高(图 7A), 验证了 MNP 分子标记在群体分类中的高效性和可行性。同时, 基于模型分析, 分别筛选出了对群体分化具有显著影响的具有功能注释的 MNP 位点(图 7B)及关键 MNP 中的特征 SNP 基因型(图 7C)。通过分析与这些特征 SNP

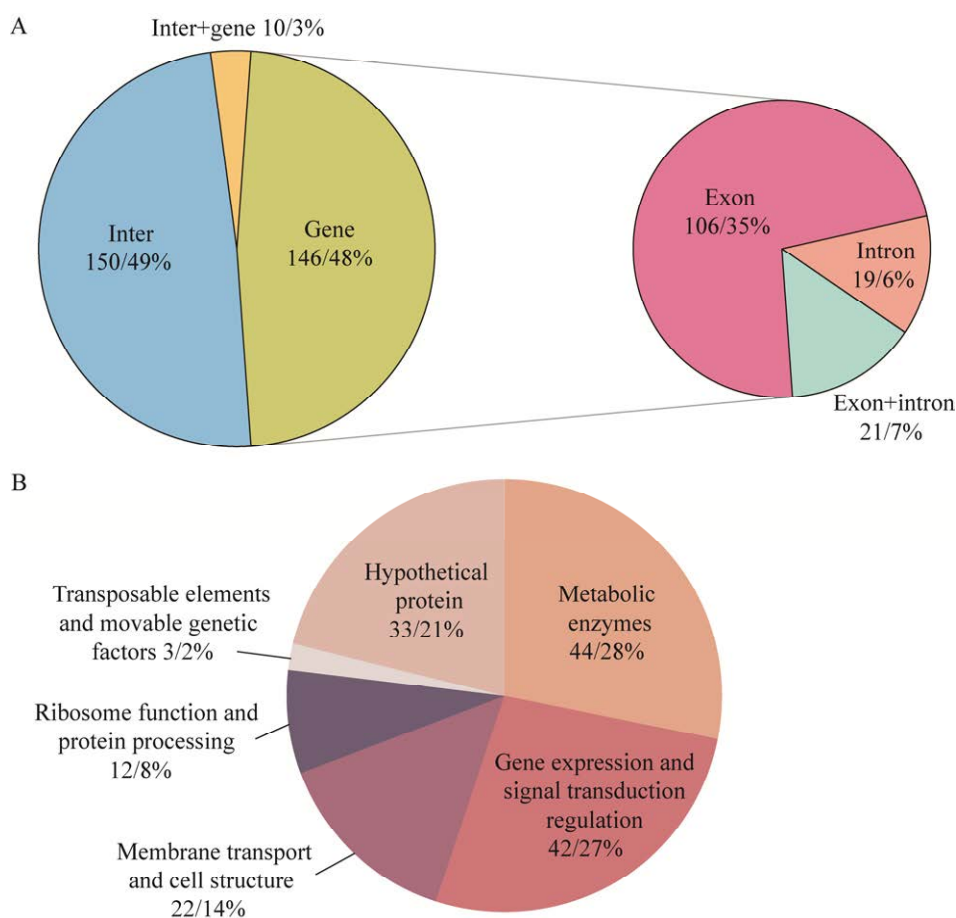


图 6 MNP 分子标记位点注释与功能分类分析 A: MNP 分子标记位点在基因区域和非基因区域的分布; B: 位于基因区域的 MNP 分子标记位点的功能分类

Fig. 6 Annotation and functional classification analysis of MNP molecular marker loci. A: Distribution of MNP molecular marker loci in gene regions or intergenic regions; B: Functional classification of MNP molecular marker loci located within gene regions.

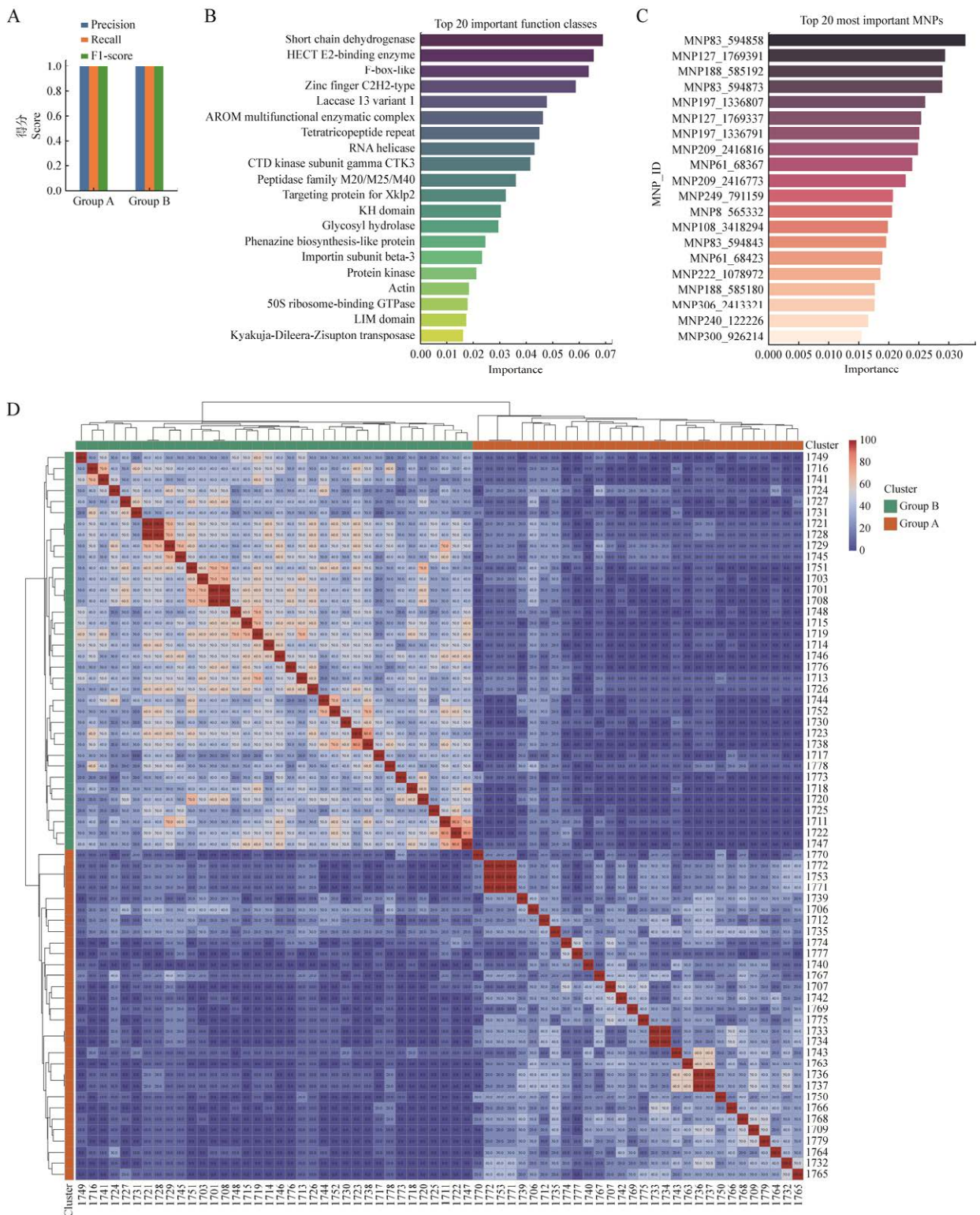


图7 融合 MNP 和机器学习的群体划分及功能分类的预测模型 A: 模型分类效果评估; B: 影响群体划分的关键 MNP 位点的功能类型; C: 影响群体划分的关键 MNP 中的特征 SNP 基因型; D: 基于 10 个 MNP 位点的菌株间遗传相似度分析

Fig. 7 Population classification and functional categorization prediction model integrating MNP and machine learning. A: Model classification performance evaluation; B: Key MNP loci influencing population classification; C: Key functional SNP types influencing population classification; D: Genetic similarity analysis among strains based on 10 MNP loci.

基因型对应的 10 个 MNP 位点(包括 MNP8、MNP61、MNP83、MNP108、MNP127、MNP188、MNP197、MNP209、MNP222 和 MNP249), 共计 49 个特征 SNP 基因型的遗传相似度, 结果进一步证实了影响显著的关键 MNP 位点可以有效地用于暗褐脉柄牛肝菌的群体划分(图 7D)。

3 讨论

3.1 MNP 分子标记在菌株精准鉴定中的优势

MNP 标记通过整合同一基因组区域内多个紧密 SNP 位点, 形成高多态性的复合型标记, 显著提高了分辨率和鉴别效率, 已广泛应用于水稻、棉花等作物鉴定(彭海等 2020; 徐云碧等 2020), 在多种食用菌中也展现出良好实用性(吕志文等 2024; 刘飞等 2025)。相比植物中需 DNA 提取、多重 PCR 和高通量测序验证的复杂流程, 食用菌因基因组小、重复序列少, 无需引物设计或 PCR 扩增, 仅通过全基因组重测序即可直接比对数据库, 完成鉴定并实时更新基因型, 显著提高效率和准确性(魏传正等 2023; 吕志文等 2024; 王森等 2025)。

除香菇仍采用多重 PCR 筛选 501 个位点并构建数据库外(Ling *et al.* 2022), 其他食用菌多通过重测序直接比对数据库完成鉴定(魏传正等 2023; 吕志文等 2024; 王森等 2025)。如金针菇筛选出 428 个位点, 验证了灵敏性和稳定性(吕志文等 2024), 茯苓 *Wolfiporia hoelen*、刺芹侧耳等也取得了类似成果(魏传正等 2023; 王森等 2025)。本研究构建了包含 306 个位点的暗褐脉柄牛肝菌 MNP 数据库, 平均每个位点含 4.82 个 SNP, PIC 值达 0.63, 显著高于传统 SSR 和 ITS 标记, 成功区分 66 个菌株, 并在不同测序深度和不同继代条件下保持良好重现性, 验证了其在种质管理和菌株监测中的应用价值。综上所述, MNP 标记凭借高多态性、简便流程和优异重现性, 已成为食用菌精准鉴定的重要工具, 有望推动遗传多样性研究和资源利用的标准化与高效化。

3.2 食用菌菌株鉴定中遗传相似度阈值的判定依据

已有研究对遗传相似度(GS)在植物和食用菌品种鉴定中的阈值设定提出了较为明确的参

考标准(彭海等 2020; Ling *et al.* 2022; 刘飞等 2025)。《植物品种鉴定 MNP 标记法》(GB/T 38551-2020)规定: GS<96.0%为不同品种, 96.0%-99.0%为近似品种, GS≥99.0%为极近似或相同品种。Ling *et al.* (2022)在香菇研究中进一步提出, GS>99.5%即可认定为同一品种。在金针菇、刺芹侧耳、茯苓等中发现部分菌株 GS=100%, 并直接将此作为同一菌株的判定依据(魏传正等 2023; 吕志文等 2024; 王森等 2025)。这类判断多基于目标物种长期无性繁殖、基因组变异率极低的特性, 从而推断基因型一致性。

本研究在此基础上进一步优化了同一菌株的判定方法。通过人为设置的同一菌株不同批次重测序数据结合其他菌株数据开展亲缘关系分析, 引入 PI_HAT 值与遗传距离(D 值)的联合判定标准(Purcell *et al.* 2007; Chen 2010), 明确了克隆菌株的界限条件, 即成对菌株 PI_HAT>0.95 且 D<0.01。随后将该判定结果与 MNP 数据库分析结果对应, 确认 GS=100%可作为暗褐脉柄牛肝菌同一菌株的判断标准。

需要指出的是, GS=100%作为判定标准是建立在当前有限菌株资源基础上的经验结果。根据本研究 66 个菌株的遗传相似度数据, 所有不同菌株间的 GS 值均低于 60%, 尚未观察到高度相似但非同源的菌株情况。然而, 随着资源的进一步拓展, 未来仍可能出现 GS 极高但未达 100%, 但为同一菌株的情况发生, 须避免盲目以 GS=100%作为唯一标准。

此外, 品种鉴别不仅取决于基因序列, 还受到表观遗传、转录调控等因素影响。因此, 在 MNP 分析结果存在争议或边界模糊的情况下, 建议辅以农艺性状观察, 实现更加全面、精准的品种或菌株判定。

3.3 MNP 标记助力群体结构分析和遗传分化研究

基于分子标记的群体遗传学研究已广泛应用于真菌遗传多样性与种质资源的评估。暗褐脉柄牛肝菌的相关研究过去多依赖于 ITS 序列、RAPD 和 SSR 等传统标记, 虽然在揭示地理来源相关的遗传差异方面取得一定进展, 但因分辨率和重复性不足, 难以准确解析菌株间的亲缘关系(曹旻等 2017; 高锋等 2021; 李凡等 2025)。

本研究基于全基因组重测序和高质量 SNP 位点,结合 PCA、ADMIXTURE 和系统发育树,明确划分出 2 个主要遗传群体,群体内一致性显著高于群体间,揭示了清晰的遗传分化格局。

在此基础上,引入 MNP 标记作为新型分析工具,对群体结构进行交叉验证。结果表明,MNP 标记在划分群体、识别克隆样本及揭示群体间遗传差异方面与 SNP 数据高度一致,能够准确反映遗传多样性分布特征。

随着数据量和计算能力的提升,机器学习在真菌遗传学研究中的应用日益广泛,成为辅助菌株分类和功能预测的重要手段(Kõljalg *et al.* 2013; Walker & Clardy 2021; Riedling *et al.* 2024)。研究人员通过整合泛组学特征,利用机器学习挖掘多药耐药诊断标志物,提升了耐药真菌的早期、快速、精准诊断能力(Delavy *et al.* 2020; Seyer *et al.* 2022; Thorn & Xu 2025)。基于深度学习的模型也用于大型真菌分类,评估其从图像中识别真菌的性能,为生物多样性研究和生态保护提供了新见解(Picek *et al.* 2022; Korkmaz *et al.* 2025)。这些实践推动了分子标记在真菌育种和知识产权保护中的应用,验证了机器学习在解析复杂群体结构中的价值。

本研究基于暗褐脉柄牛肝菌 MNP 数据库和功能注释信息,结合 SNP 与 MNP 的群体划分结果,构建了机器学习分类模型,筛选出关键遗传标记位点。这些代表性 MNP 标记不仅准确识别群体差异,也为功能基因关联和快速菌株筛选提供技术支撑。该策略有效缩减所需标记数量,降低实验和计算成本,同时保持较高分类精度,为资源受限或需快速决策的应用场景提供了可行方案,并为构建高效、可推广的真菌群体分类体系开辟了新思路。

3.4 MNP 分子标记在菌株改良的应用前景

本研究构建的 MNP 分子标记数据库显示,306 个位点中 40.19% (123 个位点)定位于具有明确功能注释的基因区域,包括代谢调控、信号转导、膜运输及蛋白质加工等多种核心生物学过程。这一结果表明,相较于传统 SNP 标记,MNP 标记不仅在多态性方面具备更高的分辨力,还在

功能层面提供了更丰富的参考信息,为解析菌株间功能差异奠定了基础。

在食用菌育种实践中,通常倾向于选择遗传距离较远的亲本菌株进行杂交,以期在后代中产生更多的基因重组和性状分化,从而培育出具有优良表型的新种质(Sonnenberg *et al.* 2016; Huang *et al.* 2024)。然而,若仅依赖遗传距离进行亲本配对,往往难以兼顾关键功能位点一致性,可能增加育种目标性状分离和不稳定的风险(鲍大鹏 2020)。基于本研究构建的 MNP 分子标记数据库及功能注释结果,通过将功能相似度与遗传距离相结合,并结合性状关联标记位点,可实现更精准的亲本筛选策略。该方法有助于在保持功能位点相对稳定的同时,充分挖掘性状相关的遗传多样性,从而提高育种效率并增强预测性。

本研究提出的策略不仅为暗褐脉柄牛肝菌等大型真菌新品系的定向改良提供了理论依据,也为分子标记辅助育种的实践应用奠定了方法学基础。未来,随着更多 MNP 位点功能注释的深入以及表型数据的持续积累,结合机器学习模型和群体结构分析,有望在更大规模上实现高效的种质创新与性状改良。

致谢

感谢广东省科学院微生物研究所邓旺秋、王刚正,广西农业科学院微生物研究所祁亮亮提供暗褐脉柄牛肝菌菌株资源。

作者贡献

孟国良:数据分析、论文构思与撰写;李嘉嘉、曾雨康、Vlasenko Vyacheslav Alexandrovich:数据处理;曹瑶:文献调查、提供实验材料;李凡:提供实验材料;李荣春:论文修改;董彩虹:论文构思与修改。

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]

- Ai S, Li S, Fang ZW, Li L, Li TT, Gao LF, Chen LH, Xiao HF, Wan RJ, Yan DZ, Wu XT, Peng H, Han RX, Zhou JF, 2024. Development and application of cotton MNP marker for fingerprint construction. *Acta Agronomica Sinica*, 50(9): 2267-2278 (in Chinese)
- Alexander DH, Lange K, 2011. Enhancements to the ADMIXTURE algorithm for individual ancestry estimation. *BMC Bioinformatics*, 12(1): 246
- Bao DP, 2020. Scientific problems in crossbreeding of edible fungi. *Acta Edulis Fungi*, 27(4): 1-24 (in Chinese)
- Bao DP, Zou G, Pei XD, Yang RH, Li ZP, Tan Q, 2022. Discussion on the path toward high quality modernization of edible fungi industry in China. *Acta Edulis Fungi*, 29(6): 103-110 (in Chinese)
- Bayer PE, Golicz AA, Scheben A, Batley J, Edwards D, 2020. Plant pan-genomes are the new reference. *Nature Plants*, 6(8): 914-920
- Bennett EP, Petersen BL, Johansen IE, Niu YY, Yang Z, Chamberlain CA, Met Ö, Wandall HH, Frödin M, 2020. Indel detection, the 'Achilles heel' of precise genome editing: a survey of methods for accurate profiling of gene editing induced indels. *Nucleic Acids Research*, 48(21): 11958-11981
- Cantalapiedra CP, Hernández-Plaza A, Letunic I, Bork P, Huerta-Cepas J, 2021. eggNOG-mapper v2: functional annotation, orthology assignments, and domain prediction at the metagenomic scale. *Molecular Biology and Evolution*, 38(12): 5825-5829
- Cao Y, He MX, Gao F, Liu J, Zhang CX, Xu XJ, 2017. Genetic diversity analysis of *Phlebopus portentosus* based on SSR markers. *Northern Horticulture*, 8: 111-114 (in Chinese)
- Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He YH, Xia R, 2020. TBtools: an integrative toolkit developed for interactive analyses of big biological data. *Molecular Plant*, 13(8): 1194-1202
- Chen S, Zhou Y, Chen Y, Gu J, 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*, 34(17): i884-i890
- Chen WM, 2010. Robust relationship inference in genome-wide association studies. *Bioinformatics*, 26(22): 2867-2873
- Dai YC, 2022. New trend of edible and medicinal fungi research in China—based on publications in *Mycosystema* during 2000–2021. *Journal of Fungal Research*, 20: 141-156 (in Chinese)
- Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, McVean G, Durbin R, 2011. The variant call format and VCFtools. *Bioinformatics*, 27(15): 2156-2158
- Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM, Li H, 2021. Twelve years of SAMtools and BCFtools. *Gigascience*, 10(2): giab008
- Delavy M, Cerutti L, Croxatto A, Prod'homme G, Sanglard D, Greub G, Coste AT, 2020. Machine learning approach for *Candida albicans* fluconazole resistance detection using matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Frontiers in Microbiology*, 10: 3000
- Eren KK, Çınar E, Karakurt HU, Özgür A, 2023. Improving the filtering of false positive single nucleotide variations by combining genomic features with quality metrics. *Bioinformatics*, 39(12): btad694
- Finn RD, Bateman A, Clements J, Coggill P, Eberhardt RY, Eddy SR, Heeger A, Hetherington K, Holm L, Mistry J, Sonnhammer ELL, Tate J, Punta M, 2014. Pfam: the protein families database. *Nucleic Acids Research*, 42(D1): D222-D230
- Fu JS, 2007. Study on cross breeding and establishment of molecular genetic markers of *Volvariella volvacea*. MS Thesis, Fujian Agriculture and Forestry University, Fuzhou. 1-49 (in Chinese)
- Gao F, He MX, Zhang CX, Liu J, Xu XJ, Cao Y, 2021. Exploitation of SSR markers with high polymorphism from *Phlebopus portentosus*. *Tropical Agricultural Science & Technology*, 44(1): 24-28 (in Chinese)
- Goodwin S, McPherson JD, McCombie WR, 2016. Coming of age: ten years of next-generation sequencing technologies. *Nature Reviews Genetics*, 17(6): 333-351
- Gu L, Gao Q, Yan D, Song S, Wang SX, Liu Y, Zong XC, Fan YY, 2025. Research status of edible fungi breeding in China. *Vegetables*, 6: 27-42 (in Chinese)
- Huang C, Sabaratnam V, Aimi T, 2024. Genetics, genomics, and breeding of edible mushrooms in Asia. *Frontiers in Microbiology*, 15: 1390845
- Huang J, Tan Q, Bao DP, 2015. Location of important functional genes of *Lentinula edodes* by using the SNP-CAPS molecular markers. *Acta Agriculturae Shanghai*, 31(2): 13-17 (in Chinese)
- Jia DH, Wang B, He XL, Liu X, 2024. Exploitation of InDel, SV and abundance SNP markers in *Flammulina filiformis* based on genome resequencing. *Mycosystema*, 43(7): 230228 (in Chinese)
- Jones P, Binns D, Chang HY, Fraser M, Li WZ, McAnulla C, McWilliam H, Maslen J, Mitchell A, Nuka G, Pesseat S, Quinn AF, Sangrador-Vegas A, Scheremetjew M, Yong SY, Lopez R, Hunter S, 2014. InterProScan 5: genome-scale protein function classification. *Bioinformatics*, 30(9): 1236-1240
- Kalyaanamoorthy S, Minh BQ, Wong TKF, Haeseler A, Jermin LS, 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature Methods*, 14(6): 587-589
- Köljal U, Nilsson RH, Abarenkov K, *et al.* (more than 20 authors), 2013. Towards a unified paradigm for sequence-based identification of fungi. *Molecular*

- Ecology, 22(21): 5271-5277
- Kong XH, Yao FJ, Wang P, 2019. Method of edible fungi germplasm resources evaluation and application in cultivar breeding. *Edible Fungi of China*, 38(12): 8-10 (in Chinese)
- Korkmaz AF, Ekinci F, Altaş Ş, Kumru E, Güzel MS, Akata I, 2025. A deep learning and explainable AI-based approach for the classification of discomycetes species. *Biology*, 14(6): 719
- Kotlarz K, Mielczarek M, Biecek P, Guldbbrandtsen B, Szyda J, 2024. Exploring the impact of sequence context on errors in SNP genotype calling with whole genome sequencing data using AI-based autoencoder approach. *NAR Genomics and Bioinformatics*, 6(3): lqae131
- Kozlov AM, Darriba D, Flouri T, Morel B, Stamatakis A, 2019. RAXML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics*, 35(21): 4453-4455
- Langmead B, Salzberg S, 2012. Fast gapped-read alignment with Bowtie2. *Nature Methods*, 9(4): 357-359
- Lee HY, Moon S, Shim D, Hong CP, Lee Y, Koo CD, Chung JW, Ryu H, 2017. Development of 44 novel polymorphic SSR markers for determination of shiitake mushroom (*Lentinula edodes*) cultivars. *Genes*, 8: 109
- Li F, Cao Y, Jin ZS, Dong CH, Li RC, 2024. Research progress of biological characteristics and cultivation techniques of *Phlebobus portentosus*. *Journal of Fungal Research*, <https://doi.org/10.13341/j.jfr.2024.1781> (in Chinese)
- Li F, Cao Y, Yang AX, Luo XY, Lei HT, Meng GL, Dong CH, Li RC, 2025. Genetic diversity of *Phlebobus portentosus* germplasm resources. *Mycosystema*, 44(5): 4-17 (in Chinese)
- Li H, 2014. Toward better understanding of artifacts in variant calling from high-coverage samples. *Bioinformatics*, 30(20): 2843-2851
- 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 YD, Luo XB, Peng X, Jin YY, Tan HP, Wu LJ, Li JW, Pei Y, Xu XH, Zhang WP, 2023. Development of SNP and InDel markers by genome resequencing and transcriptome sequencing in radish (*Raphanus sativus* L.). *BMC Genomics*, 24(1): 445
- Ling YY, Zhang MZ, Ling ZL, Cao B, Wu XP, Peng H, Wang ZR, Zhao RL, 2022. Evolutionary relationship and a novel method of efficient identification of *Lentinula edodes* cultivars in China. *Mycosphere*, 13(2): 56-85
- Liu F, Cai ZX, Kang WY, Chen WZ, Lu YP, Chen MY, Zhao RL, 2025. A new method for constructing high-resolution phylogenomic topologies using core gene-associated MNP markers: a case study from *Agaricus bisporus*. *Microbial Biotechnology*, 18(2): e70070
- Liu F, Zhang MZ, Cao B, Ling YY, Zhao RL, 2025. Application and prospects of mnp molecular markers for precise identification of edible mushrooms. *Journal of Fungal Research*, 23(3): 247-253 (in Chinese)
- Lü ZW, Yang H, Wei CZ, Ye XY, He XY, Tao YX, Xie BG, Xie LY, 2024. Establishment of MNP molecular identification technology for *Flammulina filiformis* strains. *Mycosystema*, 43(1): 22-38 (in Chinese)
- Mammadov J, Aggarwal R, Buyyarapu R, Kumpatla S, 2012. SNP markers and their impact on plant breeding. *International Journal of Plant Genomics*, 1: 728398
- McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernysky A, Garimella K, Altshuler D, Gabriel S, Daly M, DePristo MA, 2010. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, 20(9): 1297-1303
- Nizamani MM, Zhang Q, Muhac-Ud-Din G, Wang Y, 2023. High-throughput sequencing in plant disease management: a comprehensive review of benefits, challenges, and future perspectives. *Phytopathology Research*, 5(1): 1-17
- Peng H, Fang ZW, Li L, *et al.* (more than 20 authors), 2020. Identification of plant varieties-MNP marker method. GB/T38551—2020. 1-332 (in Chinese)
- Picek L, Šulc M, Matas J, Heilmann-Clausen J, Jeppesen TS, Lind E, 2022. Automatic fungi recognition: deep learning meets mycology. *Sensors*, 22(2): 633
- Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklarb P, Bakker PIW, Daly MJ, Sham PC, 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. *The American Journal of Human Genetics*, 81(3): 559-575
- Quinlan AR, Hall IM, 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*, 26(6): 841-842
- Riedling O, Walker AS, Rokas A, 2024. Predicting fungal secondary metabolite activity from biosynthetic gene cluster data using machine learning. *Microbiology Spectrum*, 12(2): e03400-23
- Seyer Cagatan A, Taiwo Mustapha M, Bagkur C, Sanlidag T, Ozsahin DU, 2022. An alternative diagnostic method for *C. neoformans*: preliminary results of deep-learning based detection model. *Diagnostics*, 13(1): 81
- Shen XF, Zhang LJ, Zhang MY, Shang XD, Li Y, Yu HL, 2020. InDel marker and its application and prospection in research of edible fungi. *Acta Agriculturae Shanghai*, 36(3): 136-140 (in Chinese)
- Sonnenberg ASM, Gao W, Lavrijssen B, Hendrickx P, Sedaghat-Tellgerd N, Foulongne-Oriol M, Kong WS, Schijlen EGWM, Baars JJP, Visser RGF, 2016. A detailed analysis of the recombination landscape of the button mushroom *Agaricus bisporus* var. *bisporus*. *Fungal Genetics and Biology*, 93: 35-45

- Thorn V, Xu J, 2025. From patterns to prediction: machine learning and antifungal resistance biomarker discovery. *Canadian Journal of Microbiology*, 71: 1-3
- Walker AS, Clardy J, 2021. A machine learning bioinformatics method to predict biological activity from biosynthetic gene clusters. *Journal of Chemical Information and Modeling*, 61(6): 2560-2571
- Wan RJ, Li Q, Zhou XC, Li L, Li TT, Zhou JF, Li S, Peng H, Zhang WX, Fang ZW, 2023. Application of cassava MNP markers in variety identification. *Chinese Journal of Tropical Crops*, 44(12): 2417-2423 (in Chinese)
- Wang M, Wang Q, Dong CH, 2025. Genetic diversity analysis and strain identification of *Wolfiporia hoelen* based on MNP markers. *Mycosystema*, 44(11): 250120 (in Chinese)
- Wei CZ, Wang M, Zhang P, Liu F, Yan JJ, Xie BG, Deng YJ, Xie LY, 2023. Identification of *Pleurotus eryngii* strains by MNP makers based on next-generation sequencing. *Acta Edulis Fungi*, 30(1): 1-9 (in Chinese)
- Xiang X, Li C, Li L, Bian YB, Kwan HS, Nong WY, Cheung MK, Yang XY, 2016. Genetic diversity and population structure of Chinese *Lentinula edodes* revealed by InDel and SSR markers. *Mycological Progress*, 15: 1-13
- Xu YB, Yang QN, Zheng HJ, Xu YF, Sang ZQ, Guo ZF, Peng H, Zhang C, Lan HF, Wang YB, Wu KS, Tao JJ, Zhang JN, 2020. Genotyping by target sequencing (GBTS) and its applications. *Scientia Agricultura Sinica*, 53(15): 2983-3004 (in Chinese)
- Yamaguchi S, Sakamoto H, Tsukiyama T, Tanesaka E, 2023. Development of codominant DNA markers (InDel) for enokitake mushroom *Flammulina velutipes* cv. "Hatsuyuki". *Journal of Crop Research*, 68: 61-64
- Zhang D, Gao F, Jakovlić I, Zou H, Zhang J, Li WX, Wang GT, 2020. PhyloSuite: an integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies. *Molecular Ecology Resources*, 20(1): 348-355
- Zou XL, Shi C, Austin RS, Merico D, Munholland S, Marsolaris F, Navabi A, Crosby WL, Pauls KP, Yu KF, Cui YH, 2014. Genome-wide single nucleotide polymorphism and insertion-deletion discovery through next-generation sequencing of reduced representation libraries in common bean. *Molecular Breeding*, 33: 769-778
- [附中文参考文献]
- 艾莎, 李莎, 方治伟, 李论, 李甜甜, 高利芬, 陈利红, 肖华锋, 万人静, 闫多子, 武星廷, 彭海, 韩瑞玺, 周俊飞, 2024. 棉花 MNP 标记位点开发及其在 DNA 指纹图谱构建中的应用. *作物学报*, 50(9): 2267-2278
- 鲍大鹏, 2020. 食用菌杂交育种中的科学问题. *食用菌学报*, 27(4): 1-24
- 鲍大鹏, 邹根, 裴晓东, 杨瑞恒, 李正鹏, 谭琦, 2022. 中国食用菌产业实现高质量现代化发展的路径探讨. *食用菌学报*, 29(6): 103-110
- 曹昉, 何明霞, 高锋, 刘静, 张春霞, 徐欣景, 2017. 基于 SSR 标记的暗褐网柄牛肝菌遗传多样性分析. *北方园艺*, 8: 111-114
- 傅俊生, 2007. 草菇杂交育种研究及其分子遗传标记的建立. 福建农林大学硕士论文, 福州. 1-49
- 戴玉成, 2022. 中国食药真菌研究发展的新趋势. *菌物研究*, 20: 141-156
- 高锋, 何明霞, 张春霞, 刘静, 许欣景, 曹昉, 2021. 暗褐网柄牛肝菌高多态性 SSR 标记的开发. *热带农业科技*, 44(1): 24-28
- 谷莉, 高琪, 严冬, 宋爽, 王守现, 刘宇, 宗宪春, 范阳阳, 2025. 我国食用菌育种研究现状. *蔬菜*, 6: 27-42
- 黄杰, 谭琦, 鲍大鹏, 2015. 运用 SNP-CAPS 分子标记定位香菇重要功能基因的研究. *上海农业学报*, 31(2): 13-17
- 孔祥会, 姚方杰, 王鹏, 2019. 食用菌种质资源评价方法及在品种选育上的应用实践. *中国食用菌*, 38(12): 8-10
- 贾定洪, 王波, 何晓兰, 刘询, 2024. 基于基因组重测序发掘金针菇 InDel、SV 及丰度 SNP 标记. *菌物学报*, 43(7): 230228
- 李凡, 曹瑶, 金卓诗, 董彩虹, 李荣春, 2024. 暗褐网柄牛肝菌生物学特性及栽培技术研究进展. *菌物研究*, <https://doi.org/10.13341/j.jfr.2024.1781>
- 李凡, 曹瑶, 杨爱霞, 罗祥英, 雷洪涛, 孟国良, 董彩虹, 李荣春, 2025. 暗褐脉柄牛肝菌种质资源遗传多样性. *菌物学报*, 44(5): 4-17
- 刘飞, 张明哲, 曹槟, 凌云燕, 赵瑞琳, 2025. MNP 分子标记在食用菌品种精准鉴定中的应用与前景分析. *菌物研究*, 23(3): 247-253
- 吕志文, 杨环, 魏传正, 叶新怡, 何小钰, 陶永新, 谢宝贵, 谢路昱, 2024. 金针菇 MNP 分子标记菌株鉴别技术建立. *菌物学报*, 43(1): 22-38
- 彭海, 方治伟, 李论, 等(20 位作者以上), 2020. 植物品种鉴定 MNP 标记法, GB/T38551—2020. 1-332
- 沈秀芬, 章炉军, 张美彦, 尚晓冬, 李玉, 于海龙, 2020. InDel 标记及其在食用菌研究中的应用与展望. *上海农业学报*, 36(3): 136-140
- 万人静, 李琼, 周新成, 李论, 李甜甜, 周俊飞, 李莎, 彭海, 章伟雄, 方治伟, 2023. 木薯 MNP 标记在品种鉴定中的应用. *热带作物学报*, 44(12): 2417-2423
- 王淼, 汪琦, 董彩虹, 2025. 基于 MNP 分子标记的茯苓种质遗传多样性分析与菌株鉴定. *菌物学报*, 44(11): 250120
- 魏传正, 王朦, 张鹏, 刘芳, 严俊杰, 谢宝贵, 邓优锦, 谢路昱, 2023. 基于二代测序技术的 MNP 标记鉴别刺芹侧耳菌株. *食用菌学报*, 30(1): 1-9
- 徐云碧, 杨泉女, 郑洪建, 许彦芬, 桑志勤, 郭子锋, 彭海, 张丛, 蓝昊发, 王蕴波, 吴坤生, 陶家军, 张嘉楠, 2020. 靶向测序基因型检测(GBTS)技术及其应用. *中国农业科学*, 53(15): 2983-3004