

DOI:10.12171/j.1000-1522.20240385

千年桐物种特异 *PGK3-TPS10-CYP726* 基因簇 挖掘及其进化和表达特征

王银行^{1,2} 高 暝² 赵耘霄² 汪阳东² 李 伟¹ 陈益存²

(1. 林木遗传育种国家重点实验室, 东北林业大学, 黑龙江 哈尔滨 150040; 2. 中国林业科学研究院亚热带林业研究所, 浙江 杭州 311400)

摘要:【目的】基于全基因组挖掘高质量物种特异基因簇, 为深入分析物种性状和遗传改良提供新的思路。【方法】通过基因组、转录组、蛋白数据联合分析, 利用生物信息手段进行高质量基因簇预测与筛选, 对目标基因簇进行基因组共线性、基因组结构、系统进化和转录表达分析。【结果】(1) 预测了千年桐高可信度 *PGK3-TPS10-CYP726* 基因簇, 该基因簇中共包括 13 个酶编码基因, 具体包括: 3-磷酸甘油醛脱氢酶 3(*GAPCP3*)、CHD4 细胞色素 P45071D(*CYP71D55*)、细胞色素 P450726A(*CYP726A*)、萜类合酶 10(*TPS10*)、磷酸甘油酸激酶 3(*PGK3*) 等。(2) *PGK3-TPS10-CYP726* 基因簇具有良好的共表达和共通路特征, 在千年桐基因组中展示了良好的独特性和完整性。(3) *PGK3-TPS10-CYP726* 基因簇有约一半基因在双子叶植物中存在较高的保守性, 但在单子叶植物中存在基因缺失的现象; *PGK3* 基因一直保留于单双子叶植物和苔藓植物中, 其序列在进化过程中高度保守; 推测 *PGK3-TPS10-CYP726* 基因簇起源于早期基部类群, 并在千年桐物种分化后形成。(4) 千年桐物种特异 *PGK3-TPS10-CYP726* 基因簇在根部特异高表达, 在枯萎病病原菌感染后的早、中期的千年桐根木质部呈现持续显著高表达。【结论】首次鉴定了千年桐物种特异基因簇 *PGK3-TPS10-CYP726*, 并分析了其进化起源和表达特征, 为解析千年桐抗枯萎病等特有性状提供了新的视角。

关键词: 千年桐; 基因簇; 物种特异; 进化起源; 细胞色素 P450; 磷酸甘油酸激酶 3; 抗枯萎病

中图分类号: S794.3 **文献标志码:** A **文章编号:** 1000-1522(2025)02-0001-09

引文格式: 王银行, 高暝, 赵耘霄, 等. 千年桐物种特异 *PGK3-TPS10-CYP726* 基因簇挖掘及其进化和表达特征 [J]. 北京林业大学学报, 2025, 47(2): 1-9. Wang Yinhang, Gao Ming, Zhao Yunxiao, et al. Identification of *Vernicia montana* species-specific *PGK3-TPS10-CYP726* gene cluster and its evolution and expression characteristics [J]. Journal of Beijing Forestry University, 2025, 47(2): 1-9.

Identification of *Vernicia montana* species-specific *PGK3-TPS10-CYP726* gene cluster and its evolution and expression characteristics

Wang Yinhang^{1,2} Gao Ming² Zhao Yunxiao² Wang Yangdong² Li Wei¹ Chen Yicun²

(1. State Key Laboratory of Tree Genetics and Breeding, Northeast Forestry University, Harbin 150040, Heilongjiang, China;

2. Research Institute of Subtropical Forestry, Chinese Academy of Forestry, Hangzhou 311400, Zhejiang, China)

Abstract: [Objective] Identification of high-quality species-specific gene clusters at the genome-wide level provides new ideas for in-depth analysis of species traits and genetic improvement. [Method] By combining genomic, transcriptome, and protein data analysis, high-quality gene cluster prediction and screening were carried out using bioinformatics methods. Genomic collinearity, genomic structure, systematic evolution, and transcriptional expression analysis were performed on target gene clusters.

收稿日期: 2024-11-18 修回日期: 2024-12-13

基金项目: 国家自然科学基金面上项目(31971685), 浙江省科技创新领军人才项目(2022R52028)。

第一作者: 王银行。主要研究方向: 林木遗传育种。Email: 2022110449@nefu.edu.cn 地址: 150040 黑龙江省哈尔滨市香坊区和兴路 26 号东北林业大学。

责任作者: 陈益存, 研究员。主要研究方向: 林木遗传育种。Email: chenyc@caf.ac.cn 地址: 311400 浙江省杭州市中国林业科学研究院亚热带林业研究所。

本刊网址: <http://j.bjfu.edu.cn>; <http://journal.bjfu.edu.cn>

[Result] (1) The *PGK3-TPS10-CYP726* gene cluster with high confidence was predicted, including 13 enzyme coding genes, i.e. glyceraldehyde 3-phosphate dehydrogenase 3 (*GAPCP3*), CHD4 cytochrome P45071D (*CYP71D55*), cytochrome P450726A (*CYP726A*), terpenoid synthase 10 (*TPS10*), phosphoglycerate kinase 3 (*PGK3*), etc. (2) *PGK3-TPS10-CYP726* gene cluster had good co-expression and common pathway characteristics, and showed good uniqueness and integrity in genome of *Vernicia montana*. (3) About half of *PGK3-TPS10-CYP726* gene cluster was highly conserved in dicotyledonous plants, but it was absent in monocotyledonous plants. The *PGK3* gene had been preserved in monocotyledonous plants and bryophytes, and its sequence was highly conserved during evolutionary process. It was speculated that the *PGK3-TPS10-CYP726* gene cluster originated from the early basal taxa and formed after *V. montana* species differentiation. (4) The specific *PGK3-TPS10-CYP726* gene cluster of *V. montana* species was specifically highly expressed in roots, and the xylem of root showed a sustained and significant high expression in the early and middle stages after being infected by *Fusarium* wilt pathogen. **[Conclusion]** The specific gene cluster *PGK3-TPS10-CYP726* is identified for the first time, and its evolutionary origin and expression characteristics are analyzed, which provides a new perspective for analyzing the unique traits of *V. montana* such as its resistance to *Fusarium* wilt.

Key words: *Vernicia montana*; gene cluster; species-specific; evolutionary origin; CYP450; PGK3; resistance to *Fusarium* wilt

基因簇(gene cluster)通常指在染色体上某个区域内成簇出现的彼此生物学功能相关的非同源基因的集合。这一概念最早起源于细菌等原核生物中的操纵子,即可以在一个启动子的作用下一起转录或进行反式剪接,翻译出同一个代谢通路中的不同的酶,组成一个完整的代谢通路来合成次生代谢产物^[1]。类似于原核生物的操纵子,基因簇普遍存在于植物等真核生物中,在玉米(*Zea mays*)、水稻(*Oryza sativa*)、小麦(*Triticum aestivum*)、大麦(*Hordeum vulgare*)、燕麦(*Avena sativa*)、拟南芥(*Arabidopsis thaliana*)、番茄(*Solanum lycopersicum*)、马铃薯(*Solanum tuberosum*)等 20 多种植物的基因组中已鉴定到超过 28 种功能型次生代谢基因簇,这些基因簇分别参与合成苯并萜类、噁唑啉酮类和生物碱类等代谢产物,发挥多种生物胁迫和非生物胁迫的作用^[2]。植物基因簇的长度主要分布在 33 ~ 284 kbp,大多数为 100 kbp 左右,基因簇内代谢基因的个数为 4 ~ 18,如番茄中的甾族糖苷生物碱 α -tomatine 合成基因簇包含 12 个基因,其中代谢基因 6 个(占比 50%)^[3-4]。在进化过程中,基因簇可以通过基因复制、重组、插入和删除等事件改变其结构,进而影响生物体的适应性和功能特性。基因簇的研究对于理解基因的演化、基因组的组织和重要性状的调控机制具有重要意义。物种特异基因簇(species-specific gene cluster)是指那些在特定物种中存在,而在其他物种中不存在的基因簇,这些基因簇在该物种中具有独特的遗传特征和功能。这些基因簇通常与物种的适应性、生理功能和代谢途径密切相关,是物种特

有性状的基础。

基因组、转录组和代谢组等技术的快速发展使得研究者们利用算法和程序去批量挖掘代谢基因簇成为可能^[5-6]。基于大量已经发表的植物基因组和转录组信息,开发出一些系统的算法来有效识别植物的代谢基因簇,这些算法大多是运用典型的基因簇起始酶和修饰酶的组合形式来检索筛选。Schlöpfer 等^[7]开发了一种新的预测软件 PlantClusterFinder,分析了 18 种植物,包括 1 种绿藻、2 种低等植物和 15 种高等植物,检测到 11 969 个候选的代谢基因簇,其中 1 700 个以上的基因簇可能与特定代谢物的产生有关。PlantSMASH^[8]和 PhytoClust^[9]等算法也相继被开发出来,二者使用更为精确的隐马尔科夫模型鉴定不同的生物合成酶,并结合基因组位置来预测候选基因簇。另外,当预测模型中结合了基因共表达数据时,候选基因簇的数量会大大减少,预测的准确性也会显著提高^[10]。

植物生物合成基因簇通常包括某个终端次生代谢产物合成通路上编码催化反应酶的基因,称为信号基因(signature gene),例如萜烯合成酶(terpene synthase, TPS)编码基因。同时,还附加 2 个或更多基因编码通路下游的修饰酶,亦称剪裁酶^[11],包括细胞色素 P450 加氧酶(cytochrome P450 enzyme, CYP450)、2-氧戊二酸酯依赖性双加氧酶、甲基转移酶、酰基转移酶和糖基转移酶等。基因簇具有鲜明的特点:共表达、共调控。以玉米的 *Bx* 基因簇为例,该基因簇中的 *Bx1* 基因是玉米中色氨酸合成酶 α -亚基的同源基因,可将 3-甘油磷酸转化为吲哚,而

Bx2 ~ Bx5 均为 CYP71C 家族基因, 可逐步催化修饰吡啶转化为 DIBOA(丁布前体物质)^[3]。番茄单萜合成酶基因簇则由编码 TPS、异戊二烯转移酶(CPT, *cis-prenyl transferase*)和 CYP 等酶的非同源基因组成, 其中 TPS 基因编码信号酶, 合成单萜骨架, CPT 和 CYP 基因编码剪裁酶, 催化修饰中间产物形成单萜目标化合物^[12-13]。Guo 等^[14]分析了罂粟(*Papaver somniferum*)基因组, 发现天然药物诺司卡品和吗啡的生物合成基因簇也具有类似的信号和修饰基因, 并且发现在诺司卡品和燕麦根皂苷的生物合成基因簇中, 基因排序与其在代谢通路上的反应步骤存在明显的共线性关系。

油桐(*Vernicia* spp.)是我国传统的重要化工油料树种, 已有几千年的栽培历史, 由其种子提炼而得的桐油是一种优质干性油, 广泛应用于工业、农业、渔业、军事和医药等领域。三年桐(*V. fordii*)是我国油桐属的主栽品种, 具有生长快、结实早、产量高的特征, 但规模化种植面临严重的枯萎病危害。油桐枯萎病又称“桐瘟”, 是由尖孢镰刀菌(*Fusarium oxysporum*)引起的致死性土传病害^[15]。迄今油桐枯萎病已在我国 8 个省份 90 多个县市爆发, 对现有的近百万公顷油桐林产生了严重威胁^[16]。相比而言, 同属的千年桐(*V. montana*)具有高抗枯萎病能力, 但是生长缓慢、结实晚、产油量和质量低。因而, 挖掘千年桐的物种特异性抗病基因簇, 探究千年桐的抗枯萎病病原菌的分子机制具有十分重要的意义。

本研究拟通过综合分析基因组、转录组等多组织多水平表型和遗传数据, 结合同种属物种联合分析, 首次挖掘大戟科(Euphorbiaceae)油桐属(*Vernicia*)千年桐响应枯萎病病原菌的高质量的物种特异基因簇, 并对该基因簇的结构特征、表达特征、物种特异性、进化起源等进行了系统分析。研究结果旨在探索千年桐的物种特异性抗病基因簇, 为油桐抗病性遗传基础的深入理解提供新的见解, 揭示千年桐在应对枯萎病病原菌侵染过程中的关键分子机制, 为培育高抗病性油桐品种提供重要的基因资源, 推动油桐枯萎病防控措施的科学化和精准化。

1 数据来源与研究方法

1.1 数据来源

1.1.1 基因组数据和蛋白数据

以大戟科油桐属代表物种千年桐为研究对象, 课题组测序并组装了千年桐基因组、蛋白质注释数据, NCBI 登录号为 PRJNA1147434。为了研究基因簇的进化演变和特异性, 我们从多个数据库(如 NCBI、Ensembl、Phytozome 等)获取了多个物种的

基因组数据和注释文件。这些物种包括拟南芥(*Arabidopsis thaliana*, GCF_000001735.4)、番茄(*Solanum lycopersicum*, GCA_036512215.2)、毛果杨(*Populus trichocarpa*, GCF_000002775.4)、甜橙(*Citrus sinensis*, GCF_000317415.1)、蓖麻(*Ricinus communis*, GCF_019578655.1)、麻风树(*Jatropha curcas*, GCF_014843425.1), 以及其他基部类群植物, 如葡萄(*Vitis vinifera*, GCF_000003745.3)、水稻(*Oryza sativa* subsp. *japonica*, GCF_034140825.1)等。这些数据为我们深入了解基因簇在不同物种中的演化与特异性特征提供了重要的参考依据。

1.1.2 转录组数据

课题组在贵州省独山县分别采集千年桐和三年桐不同组织, 具体包括主根、侧根、根尖、茎、侧枝、雄花、雌花、腺体、种仁、种皮、嫩叶、老叶、小果、果实不同时期共 14 个组织 42 个样本, 送由上海美吉公司测序, 基于 PacBio 平台测序获取了大戟科油桐属的千年桐和三年桐的转录组数据, 并将其上传到 NCBI, 登录号为 PRJNA1146716。

1.2 研究方法

1.2.1 基因簇挖掘

基因簇的挖掘包括数据准备、基因簇预测和基因簇筛选 3 个步骤。

首先, 基因组注释数据、蛋白数据、转录组数据的准备。所有的数据需要经过预处理, 蛋白数据需要通过 E2P2 蛋白酶预测软件进行酶预测, 预测之后的结果数据通过桌面版 Pathway Tools 软件^[17]构建全基因组代谢网络, 参考 Metacyc V23.0 数据库, 对大戟科物种构建 PlantCyc 代谢通路数据库。PlantCyc 是一个全面的植物代谢途径数据库, 旨在容纳全方位的植物通路和酶^[18]。

其次, 基因簇的预测。结合上述注释数据、代谢通路数据, 基于生物信息软件 PlantClusterFinder 进行千年桐的全基因组基因簇预测, PlantClusterFinder 1.3 是一款在 Linux 系统下基于 Matlab 环境运行的 Java 软件, 采用更为精确的隐马尔可夫模型来鉴别不同的生物合成酶, 同时结合基因组位置及其部分特征和构成规律^[5]。

最后, 高质量基因簇的筛选。对预测的基因簇进一步进行共表达、共通路的分析, 筛选候选基因簇, 此时基因簇数量显著减少, 预测准确性显著提高。共表达分析我们采用两种标准, 一种是基于皮尔森相关性检验的 P 值; 另一种则是基于预测簇内基因的皮尔森相关性系数^[19]的分布, 通过转录组数据对各基因样本的定量结果(FPKM), 对基因簇簇内基因进行两两之间进行皮尔森相关性检验, 并建立

共表达热图。共通路判断标准为簇内基因是否参与了同一代谢通路,若是,判定为簇内基因共通路。

1.2.2 共线性分析

为了研究筛选出的基因簇的结构及其演化,采用 Tbttools^[20]的 One step MCScanX-Super Fast 模块,对其进行物种微共线性分析。MCScanX 是一款专用于基因组和蛋白质数据分析的软件,特别适用于共线性分析和系统发育研究。One step MCScanX-Super Fast 模块集成了 MCScanX 的完整功能,能够整合不同物种的基因组信息,通过线性分析揭示物种间的遗传关系和进化历史。具体选取进化过程中的多个物种,包括三年桐、千年桐、蓖麻、麻风树、杨树、拟南芥、甜橙等双子叶物种和葡萄、水稻等单子叶物种,一对一构建共线性,观测筛选共线基因,再通过 Text Merge for MCScanX 模块合并所有数据,锚定共线染色体,可视化多物种共线性结果。

1.2.3 系统进化分析

选取筛选出的基因簇上的高度共线保守基因序列,以及从 GenBank 中下载物种(小立碗癣 *Physcomitrella patens*、菠萝 *Ananas comosus*、小麦 *Triticum aestivum*、水稻、葡萄)的候选共线基因序列,使用 MUSCLE Wrapper 软件进行比对,通过 TrimAL Wrapper 剪切软件处理蛋白数据,使用 MEGA11^[21]工具构建系统进化树,自展支持值(bootstrap) > 60% 的显示在各个进化分支节点上。

1.2.4 表达特征分析

1.2.4.1 组织特异性表达分析

为进一步探讨筛选出的基因簇的组织表达特征,利用转录组数据,分析了该基因簇在千年桐主根、侧根、根尖、茎、侧枝、雄花、雌花、腺体、种仁、种皮、嫩叶、老叶、小果、果实不同时期 14 个组织共 42 个样本中的组织特异表达规律。

1.2.4.2 响应枯萎病病原菌感染的表达分析

选择上述组织特异表达的基因,利用转录组数据(PRJNA810877)对千年桐和三年桐根木质部在枯萎病病原菌感染后的不同时间点(早期、中期和后期)的表达模式进行特异性分析,了解千年桐和三年桐在抗病机制上的潜在差异和共性,挖掘抗枯萎病的关键基因。

2 结果与分析

2.1 千年桐 *PGK3-TPS10-CYP726* 特异基因簇结构特征

通过 PlantClusterFinder 预测和严格筛选,最终得到千年桐一个高质量的萜类基因簇。该基因簇长度为 173 kbp,位于千年桐 h1tg0000031 染色体 220 346 ~ 1 058 525 bp 之间,共包括 13 个酶编码基因和 1 个假基因(*Pseudoogene*),酶具体包括:3-磷酸甘油醛脱氢酶 3(glyceraldehyde 3-phosphate dehydrogenase, GAPCP3), 2 个未知蛋白(uncharacterized protein)、染色体重塑蛋白 4(chromatin remodeling 4, CHD4), 细胞色素 P450 71D(cytochrome P450 71D, CYP71D455)、细胞色素 P450 726A(CYP726A27)、萜类合酶 10(terpene synthase 10, TPS10)、细胞色素 P450 726A25(CYP726A25)、磷酸甘油酸激酶 3(phosphoglycerate kinase 3, PGK3)、萜类合酶 10(TPS10)、细胞色素 P450 726A25(CYP726A25)、抗坏血酸氧化酶(L-ascorbate oxidase, AO)、磷酸甘油酸激酶 3(PGK3) (图 1)。TPS10 构成了次生代谢物的“信号”酶骨架, CYP71D455 和 CYP726A27 作为典型的“裁剪”酶修饰,因此将该基因簇命名为 *PGK3-TPS10-CYP726* 基因簇。

2.2 千年桐 *PGK3-TPS10-CYP726* 特异基因簇共表达分析

根据基因簇上所包含的骨架酶(TPS10)和修饰酶(CYP71D455 和 CYP726A27)类型,推断所得到的 *PGK3-TPS10-CYP726* 基因簇为典型的萜类基因簇。结合转录组数据,对该基因簇的所有基因进一步进行共表达分析发现,簇内基因中 6 个基因的表达量不高,表达相关性不显著,因而选取其余 8 个基因进行皮尔森相关性分析和后续的表达特征分析。

共表达热图(图 2)显示, *TPS10* 与 *CYP726A27* 基因之间存在显著的共表达关系,这一结果符合萜烯类基因簇构成的基本要求,表明 *PGK3-TPS10-CYP726* 基因簇具有良好的簇内基因共表达特征(图 2)。推测这一基因簇可能在抗病过程中发挥协同作用,有效地参与萜烯类物质的合成,并在植物的

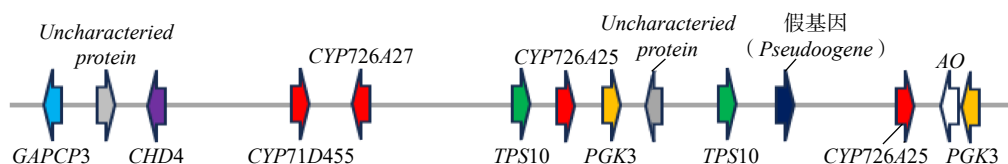
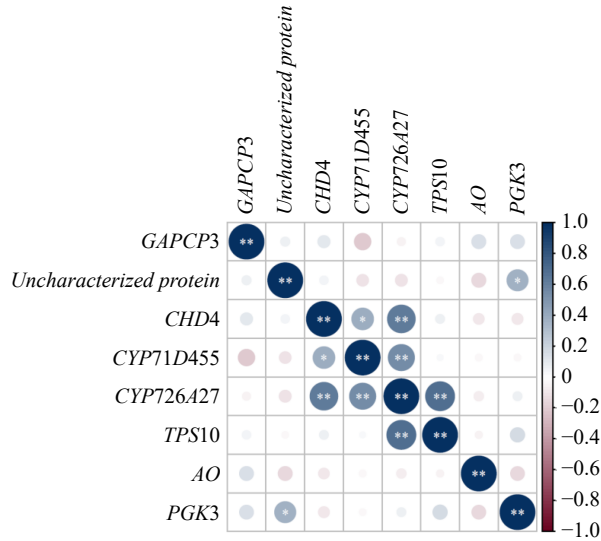


图 1 千年桐 *PGK3-TPS10-CYP726* 基因簇基因组结构

Fig. 1 Genome structure of *PGK3-TPS10-CYP726* gene cluster in *Vernicia montana*



*表示在 0.05 水平上显著相关, **表示在 0.01 水平上极显著相关。

图 2 千年桐 *PGK3-TPS10-CYP726* 特异基因簇共表达热图
Fig. 2 Co-expression heatmap of *V. montana* specific *PGK3-TPS10-CYP726* gene cluster

防御反应中起到重要的作用。

2.3 千年桐 *PGK3-TPS10-CYP726* 特异基因簇共线性分析

在种间进行拟南芥、水稻、番茄、葡萄等多个物种的共线性分析时, 结果(图 3)显示, *PGK3-TPS10-CYP726* 基因簇在千年桐中表现出独特性和完整性。千年桐的 *PGK3-TPS10-CYP726* 基因簇具有完整的基因排列和特有的结构特征, 而在拟南芥、水稻、番茄和葡萄等物种中, 虽然也能找到多个与千年桐基因簇相关的共线基因, 但这些基因簇表现出明显的差异。这些物种中的共线基因未能形成与千年桐完全相同的结构, 并且存在多个关键基因缺失, 尤其是千年桐特有的基因。这些差异可能与物种间的

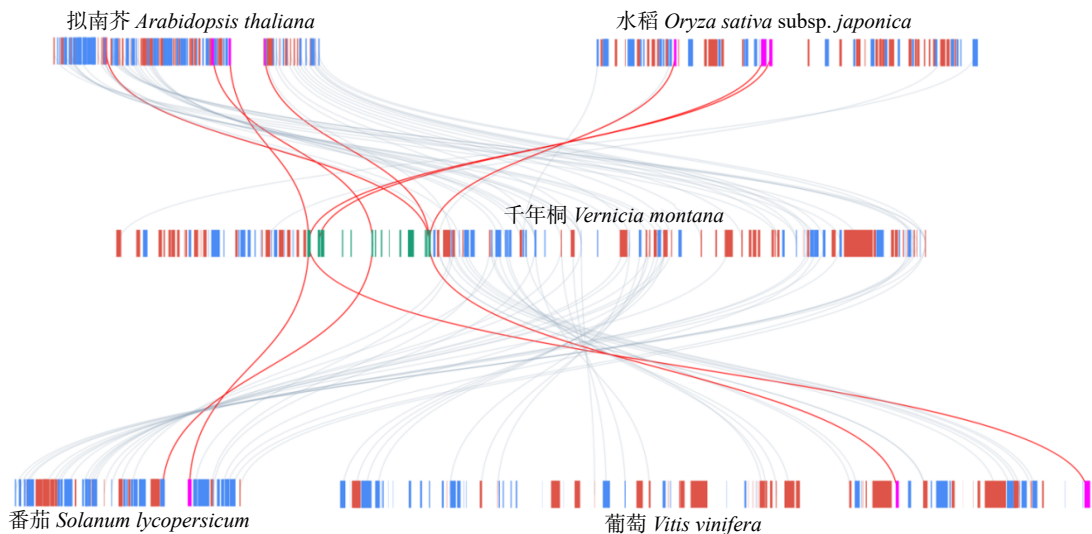
亲缘关系及其进化历史密切相关。

进一步选择与千年桐同为油桐属的三年桐进行共线性分析发现, 三年桐和千年桐共享的共线基因数量最多(图 4), 表明这两者之间存在较为密切的亲缘关系。然而, 三年桐与千年桐的 *PGK3-TPS10-CYP726* 基因簇存在显著差异, 三年桐基因簇缺失了 6 条关键基因。这一缺失表明, 三年桐的 *PGK3-TPS10-CYP726* 基因簇并未能完全保留千年桐的基因排列和结构特征, 可能受到进化过程中的基因重排或突变的影响。尽管在三年桐基因组中发现了一些与千年桐 *PGK3-TPS10-CYP726* 基因簇相关的同源基因, 但由于缺少了几个核心基因或存在功能缺失, 这些同源基因未能发挥与千年桐相似的生物学作用, 从而限制了其基因簇的形成和功能表现。

2.4 千年桐 *PGK3-TPS10-CYP726* 特异基因簇的系统进化

共线性分析结果(图 5a)显示: 在双子叶植物(三年桐、蓖麻、毛果杨、拟南芥)中, *PGK3-TPS10-CYP726* 基因簇的基因表现出较高的保守性, 至少有一半的基因保持了共线性。然而, 在单子叶植物(甜橙、葡萄、水稻)中, 基因的缺失现象明显。这一变化与被子植物的系统发育起源相吻合, 进一步支持了千年桐 *PGK3-TPS10-CYP726* 特异性基因簇可能是在早期基部类群的不断进化过程中形成的推论, 部分新基因的加入使得这一基因簇具备了独特的特性。

在系统进化分析中, *PGK3-TPS10-CYP726* 基因簇中 *PGK3* 为高度保守的共线基因, 因而选择 *PGK3* 进行进一步的系统进化分析, 构建系统发育进化树, 结果(图 5b)显示, *PGK3* 在双子叶、单双子叶、苔藓代表物种进化过程中一直保留, 其序列在整



绿色代表千年桐 *PGK3-TPS10-CYP726* 特异基因簇上的基因, 部分未共线的绿色基因为千年桐特有。

图 3 特异基因簇 *PGK3-TPS10-CYP726* 多物种共线性分析
Fig. 3 Multi-species collinearity analysis of specific gene cluster *PGK3-TPS10-CYP726*

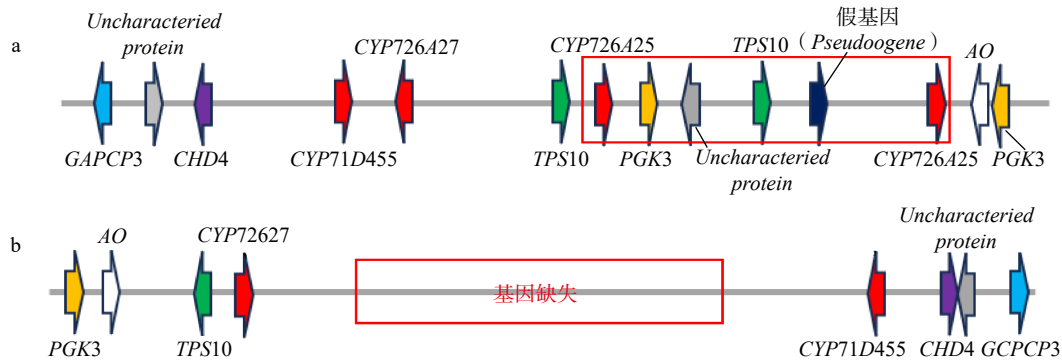


图4 千年桐(a)特异性基因簇与三年桐(b)同源基因差异性对比

Fig. 4 Comparison of difference between specific gene clusters of *V. montana* (a) and homologous genes of *V. fordii* (b)

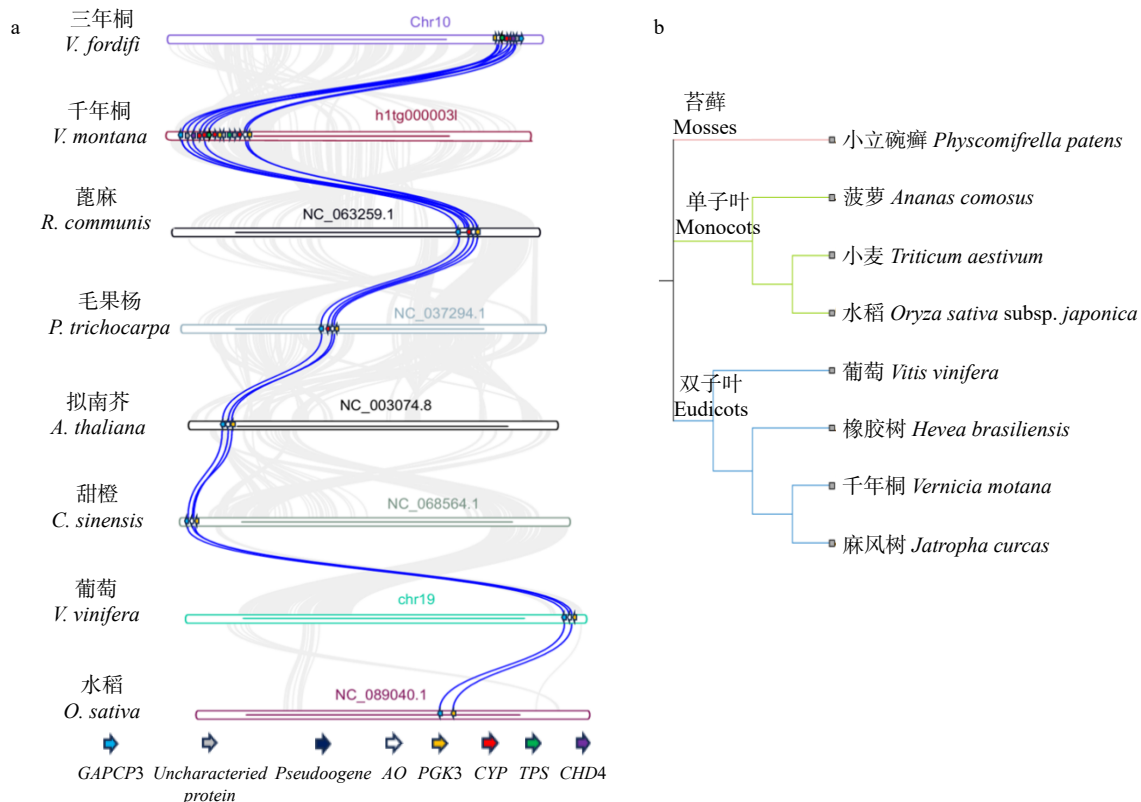


图5 千年桐特异基因簇 *PGK3-TPS10-CYP726* 的进化(a)和 *PGK3* 基因系统进化树(b)

Fig. 5 Evolution of *PGK3-TPS10-CYP726* gene cluster in *V. montana* (a) and *PGK3* phylogenetic tree (b)

个进化过程中一直高度保守。

2.5 千年桐物种特异 *PGK3-TPS10-CYP726* 基因簇组织特异性表达分析

为进一步探讨 *PGK3-TPS10-CYP726* 基因簇的组织表达特征,我们分析了该基因簇在千年桐 14 个组织共 42 个样本中的表达规律。结果(图 6)显示,发现 *PGK3-TPS10-CYP726* 基因簇都在千年桐根组织中有明显的表达特异性。因而,接下来选择根来探究千年桐物种特异基因簇 *PGK3-TPS10-CYP726* 响应病原菌侵染的表达模式。

2.6 千年桐物种特异基因簇 *PGK3-TPS10-CYP726* 响应枯萎病原菌侵染的表达模式

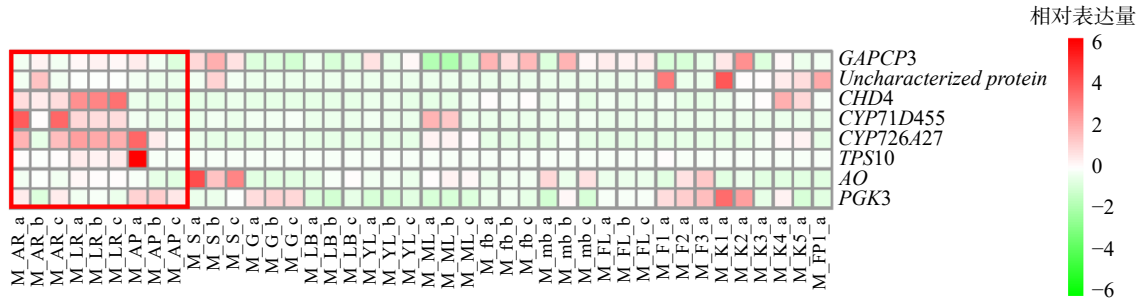
与同属的三年桐相比,千年桐具有显著的抗枯

萎病能力。已有研究结果表明千年桐的根木质部对枯萎病起到了关键的抗病作用^[22],因此我们将根组织表达特异基因簇延伸到根木质部中进行研究。在根木质部中,千年桐特异基因簇 *PGK3-TPS10-CYP726* 在侵染后的早期、中期和后期均出现高表达,但主要集中在中后期;而三年桐响应病原菌侵染主要在未侵染期和侵染早期有显著的高表达(图 7)。

3 讨论

3.1 千年桐 *PGK3-TPS10-CYP726* 萜类基因簇特异性和抗枯萎病表达特征

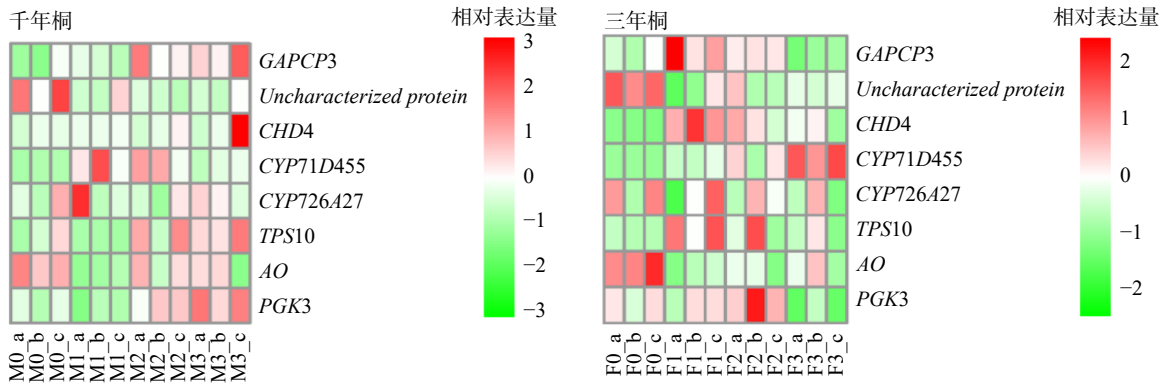
本研究中, *PGK3-TPS10-CYP726* 基因簇在千年桐中展现了独特的结构特征。从种间共线性分析来



AR. 主根; LR. 侧根; AP. 根尖; S. 茎; G. 腺体; LB. 枝条; YL. 嫩叶; ML. 老叶; fbf. 雄花; mbf. 雌花; FL. 小果; F1. 果实; K. 种仁; FP. 种皮。

图 6 千年桐 *PGK3-TPS10-CYP726* 基因簇组织表达模式

Fig. 6 Tissue expression patterns of *PGK3-TPS10-CYP726* gene cluster in *V. montana*



M0 和 F0 为千年桐和三年桐的未侵染时期, M1 和 F1 为千年桐和三年桐的侵染初期, M2 和 F2 为千年桐和三年桐的侵染中期, M3 和 F3 为千年桐和三年桐的侵染后期。

图 7 *PGK3-TPS10-CYP726* 基因簇在千年桐和三年桐根木质部中响应病原菌侵染的表达模式

Fig. 7 Expression patterns of *PGK3-TPS10-CYP726* gene cluster in root xylem of *V. montana* and *V. fordii* in response to pathogen infection

看, 与拟南芥、水稻、番茄和葡萄等物种相比, 千年桐的这一基因簇结构完整且特异。从种内共线性分析来看, 与同一泡桐属的三年桐相比, 千年桐的 *PGK3-TPS10-CYP726* 基因簇也展现出独特性。千年桐和三年桐在基因簇的结构上存在明显的差异, 三年桐存在共线基因存在倒位现象, 且缺少了部分千年桐特异的基因(图 4)。千年桐萜类基因簇的特殊性, 意味着这个基因簇可能承担着特定的生物学功能, 如抗病性、代谢途径调控等, 这可能是千年桐比三年桐抗枯萎病能力更高的原因。

萜类化合物可通过激活植物的免疫反应和调节代谢途径, 显著增强植物对枯萎病的抗性。在烟草中, 外源施用二萜类化合物能够激活植物的防御机制, 增强根部免疫反应, 从而有效抑制病原的生长, 并减少病害的扩散^[23]。茶树精油中的萜烯类化合物可直接抑制镰刀菌(*Fusarium*)的生长和诱导植物系统性抗性^[24]。在岷江百合(*Lilium regale*)与尖孢镰刀菌(*Fusarium oxysporum*)的相互作用过程中发现, 萜类基因簇可与其他次生代谢产物共同作用, 提升植物的免疫和抗病能力^[25]。本研究中, 千年桐的 *PGK3-TPS10-CYP726* 特异萜类基因簇在病原菌侵

染的根木质部区域具有显著的表达, 与同属物种三年桐相比, 千年桐展现出明显的病原菌响应差异。已有研究^[22]表明, 病原菌的菌丝无法侵入千年桐的根木质部, 这一特性与千年桐特有的抗性机制密切相关。推测在枯萎病病原菌侵染过程中, 千年桐 *PGK3-TPS10-CYP726* 特异萜类基因簇可能通过调控植物的代谢途径和生理反应, 增强了植物的免疫反应, 阻止了病原菌的扩散和入侵, 在抗病过程中发挥了关键作用。相比之下, 三年桐的 *PGK3-TPS10-CYP726* 基因簇同源基因表达较弱或缺失, 导致其表现出较差的病原菌响应能力和免疫防御效果。

3.2 *PGK3* 基因在千年桐中的功能

磷酸甘油酸激酶(PGK)是一种催化 ATP 形成 ADP 的酶, 与植物能量代谢、碳分配和细胞应答机制紧密相关^[26]。在高等植物中, *PGK* 基因主要参与光合碳还原和叶绿体糖酵解过程。在烟草、甘蓝型油菜和拟南芥中, *PGK* 基因在不同组织中存在差异表达^[27]。Massange-Sánchez 等^[28]在拟南芥中发现了 *AtPGK1*、*AtPGK2* 和 *AtPGK3* 3 个 *PGK* 基因, 其中 *AtPGK1* 仅在于光合组织的叶绿体中表达, *AtPGK2* 在光合和非光合细胞的叶绿体/质体中表达, 而

AtPGK3 在根中表达。本研究中, *PGK3* 基因在多个物种中具有较高的同源性, 且其序列在进化过程中表现出较强的保守性。千年桐 *PGK3* 主要高表达富集于根组织中(图 6), 且能响应枯萎病病原菌侵染(图 7), 推测 *PGK3* 基因在千年桐抗病过程中发挥重要作用, 可帮助千年桐在根部建立强有力的免疫屏障, 促进抗病相关物质的合成, 进而增强植株的整体免疫防御能力。

3.3 预测框架对物种的普遍适用性

我们当前构建的千年桐萜类基因簇预测框架, 主要依托于大戟科油桐的基因组数据和已知的萜类合成途径。该框架的核心优势在于能够针对性地整合基因组特征与萜烯类基因簇的结构, 从而实现对大戟科萜类基因簇的高精度预测。研究人员^[7]先前已经对 18 个不同科属的物种进行了基因簇预测, 均获得了良好的结果。但是, 框架能否广泛适用于其他物种, 仍取决于不同物种间基因簇的进化差异以及萜烯合成基因的保守性。我们认为, 尽管此预测框架专为大戟科油桐物种设计, 但其基本方法和概念(如基因簇的识别与鉴定)在一定程度上也能够适用于其他物种, 尤其是与大戟科基因组相似的植物类群。值得注意的是, 在将框架应用于其他物种时, 仍需根据具体物种的基因组特征进行适当的调整和优化, 以进一步提升预测的准确性。

4 结 论

本研究结合转录组、基因组与蛋白质组数据, 首次建立了针对千年桐萜类基因簇的预测流程, 并成功筛选出 *PGK3-TPS10-CYP726* 这一特异性基因簇, 填补了千年桐基因簇预测研究的空白。主要结论如下:

(1) 预测了千年桐高可信度 *PGK3-TPS10-CYP726* 基因簇, 该基因簇中共包括 13 个酶编码基因, 具体包括: 3-磷酸甘油醛脱氢酶 3(*GAPCP3*), *CHD4* 细胞色素 P45071D(*CYP71D55*)、细胞色素 P450726A(*CYP726A*)、萜类合酶 10(*TPS10*)、磷酸甘油酸激酶 3(*PGK3*)等。

(2) *PGK3-TPS10-CYP726* 基因簇具有良好的共表达和共通路特征, 在千年桐基因组中展示了良好的独特性和完整性。

(3) *PGK3-TPS10-CYP726* 基因簇有约一半基因在双子叶植物中存在较高的保守性, 但在单子叶植物中存在基因缺失的现象; 其中 *PGK3* 基因一直保留于单双子叶植物和苔藓植物中, 其序列在进化过程中高度保守; 推测 *PGK3-TPS10-CYP726* 基因簇起源于早期基部类群, 并在千年桐物种分化后形成。

(4) 千年桐物种特异 *PGK3-TPS10-CYP726* 基因簇在根部特异高表达, 在枯萎病病原菌侵染后的早中期的千年桐根木质部呈现持续显著高表达。

本研究基于生物信息学工具 *PlantClusterFinder*, 预测并挖掘出了千年桐特异萜类 *PGK3-TPS10-CYP726* 基因簇, 为其他植物物种的基因簇分析提供了可借鉴的研究框架; 系统分析了千年桐萜类基因簇在应对枯萎病过程中表达模式的动态变化, 为揭示油桐抗枯萎病的分子响应机制奠定了研究基础。研究结果还将为大戟科植物在抗病性进化过程中的基因演化和适应机制提供理论支持, 推动相关领域的基础研究和应用研究发展, 有助于提高油桐产业的可持续性和生产效益, 保障桐油产业的稳定发展。

参 考 文 献

- [1] Hoskisson P A. Streptomyces in nature and medicine: the antibiotic makers[J]. *Microbe Magazine*, 2008, 3(1): 151.
- [2] Boycheva S, Daviet L, Wolfender J L, et al. The rise of operon-like gene clusters in plants[J]. *Trends in Plant Science*, 2014, 19(7): 447-459.
- [3] Frey M, Schullehner K, Dick R, et al. Benzoxazinoid biosynthesis, a model for evolution of secondary metabolic pathways in plants.[J]. *Phytochemistry*, 2009, 70(15-16): 1645-1651.
- [4] Field B, Osbourn A E. Metabolic diversification: independent assembly of operon-like gene clusters in different plants[J]. *Science*, 2008, 320: 543-547.
- [5] Chen W, Gong L, Guo Z, et al. A novel integrated method for large-scale detection, identification, and quantification of widely targeted metabolites: application in the study of rice metabolomics[J]. *Molecular Plant*, 2013, 6(6): 1769-1780.
- [6] Medema M H, Kottmann R, Yilmaz P, et al. Minimum information about a biosynthetic gene cluster[J]. *Nature Chemical Biology*, 2015, 11(9): 625-631.
- [7] Schläpfer P, Zhang P, Wang C, et al. Genome-wide prediction of metabolic enzymes, pathways, and gene clusters in plants[J]. *Plant Physiology*, 2017, 173(4): 2041-2059.
- [8] Kautsar S A, Suarez D H G, Blin K, et al. plantiSMASH: automated identification, annotation and expression analysis of plant biosynthetic gene clusters[J]. *Nucleic Acids Research*, 2017, 45(W1): 55-63.
- [9] Töpfer N, Fuchs L M, Aharoni A. The phytoClust tool for metabolic gene clusters discovery in plant genomes[J]. *Nucleic Acids Research*, 2017, 45(12): 7049-7063.
- [10] Luo J. Metabolite-based genome-wide association studies in plants[J]. *Current Opinion in Plant Biology*, 2015, 24: 31-38.
- [11] Boycheva S, Daviet L, Wolfender J L, et al. The rise of operon-like gene clusters in plants[J]. *Trends in Plant Science*, 2014, 19: 447-459.
- [12] Matsuba Y, Nguyen T T H, Wiegert K, et al. Evolution of a complex locus for terpene biosynthesis in *Solanum*[J]. *The Plant*

- Cell, 2013, 25(6): 2022–2036.
- [13] 吕海舟, 刘琬菁, 何柳, 等. 植物次生代谢基因簇研究进展 [J]. 植物科学学报, 2017, 35(4): 609–621.
- Lü H Z, Liu W J, He L, et al. Advances on the study of gene clusters involved in plant secondary metabolism[J]. Plant Science Journal, 2017, 35(4): 609–621.
- [14] Guo L, Winzer T, Yang X, et al. The opium poppy genome and morphinan production[J]. Science, 2018, 362: 343–347.
- [15] Chen Y, Yin H, Gao M, et al. Comparative transcriptomics atlases reveals different gene expression pattern related to *Fusarium* wilt disease resistance and susceptibility in two *Vernicia* species[J]. Frontiers in Plant Science, 2016, 7: 1974.
- [16] 王嘉, 梁晓洁, 高暝. 千年桐根部黄酮类化合物生物合成对枯萎病菌感染的响应 [J]. 林业科学, 2022, 58(2): 159–170.
- Wang J, Liang X J, Gao M. Response of flavonoids biosynthesis in roots of *Vernicia montana* to *Fusarium* wilt infection[J]. Scientia Silvae Sinicae, 2022, 58(2): 159–170.
- [17] Karp P D, Midford P E, Billington R, et al. Pathway tools version 23.0 update: software for pathway/genome informatics and systems biology[J]. Briefings in Bioinformatics, 2021, 22(1): 109–126.
- [18] Hawkins C, Ginzburg D, Zhao K, et al. Plant metabolic network 15: a resource of genome-wide metabolism databases for 126 plants and algae[J]. Journal of Integrative Plant Biology, 2021, 63(11): 1888–1905.
- [19] Usadel B, Obayashi T, Mutwil M, et al. Co-expression tools for plant biology: opportunities for hypothesis generation and caveats[J]. Plant, Cell & Environment, 2009, 32(12): 1633–1651.
- [20] Chen C, Wu Y, Li J, et al. TBtools-II: a “one for all, all for one” bioinformatics platform for biological big-data mining[J]. Molecular Plant, 2023, 16(11): 1733–1742.
- [21] Tamura K, Stecher G, Kumar S. MEGA11: molecular evolutionary genetics analysis version 11[J]. Molecular Biology and Evolution, 2021, 38(7): 3022–3027.
- [22] Zhang Q, Wu L, Yin H, et al. D6 protein kinase in root xylem benefiting resistance to *Fusarium* reveals infection and defense mechanisms in tung trees[J]. Horticulture Research, 2021, 8: 240.
- [23] Seo S, Gomi K, Kaku H, et al. Identification of natural diterpenes that inhibit bacterial wilt disease in tobacco, tomato and *Arabidopsis*[J]. Plant and Cell Physiology, 2012, 53(8): 1432–1444.
- [24] Dalio R J D, Maximo H J, Roma-Almeida R, et al. Tea tree oil induces systemic resistance against *Fusarium* wilt in banana and *Xanthomonas* infection in tomato plants[J]. Plants, 2020, 9(9): 1137.
- [25] Deng J, Che X, Gu Y, et al. Integrated multi-omics investigation revealed the importance of phenylpropanoid metabolism in the defense response of *Lilium regale* Wilson to *Fusarium* wilt[J]. Horticulture Research, 2024, 7(11): 140.
- [26] Rosa-Téllez S, Anoman A D, Flores-Tornero M, et al. Phosphoglycerate kinases are co-regulated to adjust metabolism and to optimize growth[J]. Plant Physiology, 2018, 176(2): 1182–1198.
- [27] 郭楠, 赵敬会, 高天姝, 等. 甘蓝型油菜 PGK 基因的克隆与表达分析 [J]. 西北植物学报, 2014, 34(11): 2188–2193.
- Guo N, Zhao J H, Gao T S, et al. Cloning and expression analysis of PGK gene in *Brassica napus*[J]. Acta Botanica Boreali-Occidentalia Sinica, 2014, 34(11): 2188–2193.
- [28] Massange-Sánchez J A, Casados-Vázquez L E, Juárez-Colunga S, et al. The phosphoglycerate kinase (PGK) gene family of maize (*Zea mays* var. *B73*)[J]. Plants, 2020, 9(12): 1639.

(责任编辑 赵田芸
责任编辑 康向阳)