

引用格式: 刘晨, 樊鸿瑞, 周明兵, 等. 山核桃 *PAL* 基因家族的全基因组鉴定及表达模式分析[J]. 浙江农林大学学报, 2025, 42(4): 725–735. LIU Chen, FAN Hongrui, ZHOU Mingbing, *et al.* Genome-wide identification and expression profiling of *PAL* gene family in *Carya cathayensis*[J]. *Journal of Zhejiang A&F University*, 2025, 42(4): 725–735.

山核桃 *PAL* 基因家族的全基因组鉴定及表达模式分析

刘 晨^{1,2}, 樊鸿瑞^{1,2}, 周明兵³, 周小红^{1,2}

(1. 浙江农林大学 森林食物资源挖掘与利用全国重点实验室, 浙江 杭州 311300; 2. 浙江农林大学 全省经济林品质调控与加工利用重点实验室, 浙江 杭州 311300; 3. 浙江农林大学 竹子研究院, 浙江 杭州 311300)

摘要: 【目的】苯丙氨酸解氨酶 (*PAL*) 作为酚类物质合成代谢的起始酶, 对植物的生长发育及环境适应起着至关重要的作用。鉴定山核桃 *Carya cathayensis* *PAL* (*CcPALs*) 基因家族成员, 分析家族成员在愈伤组织增殖褐化过程中的表达模式, 为挖掘能够抑制褐化、促进再生的 *PAL* 基因提供理论依据。【方法】以山核桃全基因组数据为基础, 通过比对筛选出包含 *PAL* 保守基序的 *PAL* 家族成员。运用生物信息学方法, 对山核桃 *PAL* 蛋白的理化性质、基因结构等特征进行预测分析, 并与多个物种的 *PALs* 序列共同构建发育进化树。同时, 通过转录组测序, 分析继代后愈伤组织中 *CcPALs* 基因的表达变化, 并制作表达热图。【结果】山核桃共有 6 条 *PAL* 蛋白序列, 长度为 657~760 个氨基酸。这些蛋白均属于酸性、稳定且亲水的蛋白, 主要定位在细胞核和内质网上, 蛋白二级结构以 α -螺旋与无规则卷曲为主。进化关系分析表明: *CcPAL1*~*CcPAL4* 聚类在双子叶植物分支, 而 *CcPAL5* 和 *CcPAL6* 则聚类在裸子植物分支。表达分析显示: *CcPALs* 基因在愈伤组织继代后呈现出不同的表达模式。【结论】*CcPALs* 成员之间具有较高的保守性, 但 *CcPAL5* 和 *CcPAL6* 在进化关系上与裸子植物 *PAL* 更为接近, 可能与 *CcPAL1*~*CcPAL4* 在次生代谢物合成过程中的功能存在较大差异。在愈伤组织增殖褐化过程中, 进化树上不同分支的 *CcPALs* 表达模式各异, 可能控制着酚类物质的代谢流, 进而影响山核桃的再生进程。图 7 表 4 参 37

关键词: 山核桃; 苯丙氨酸解氨酶 (*PAL*); 愈伤; 基因鉴定; 基因家族; 表达分析

中图分类号: S722.3 文献标志码: A 文章编号: 2095-0756(2025)04-0725-11

Genome-wide identification and expression profiling of *PAL* gene family in *Carya cathayensis*

LIU Chen^{1,2}, FAN Hongrui^{1,2}, ZHOU Mingbing³, ZHOU Xiaohong^{1,2}

(1. National Key Laboratory for Development and Utilization of Forest Food Resources, Zhejiang A&F University, Hangzhou 311300, Zhejiang, China; 2. Provincial Key Laboratory for Non-wood Forest and Quality Control and Utilization of Its Products, Zhejiang A&F University, Hangzhou 311300, Zhejiang, China; 3. Bamboo Industry Institute, Zhejiang A&F University, Hangzhou 311300, Zhejiang, China)

Abstract: [Objective] Phenylalanine ammonia-lyase (*PAL*), the initial enzyme in the biosynthesis of phenolic substances, is of crucial significance in plant growth, development, and environmental adaptation. The present study aims to identify the members of the *PAL* gene family in *Carya cathayensis* (*CcPALs*) and analyze their expression patterns during callus proliferation and browning. This analysis may provide a theoretical basis for the manipulation of *CcPALs* to inhibit browning and promote regeneration. [Method] Based on the whole-genome data of *C. cathayensis*, *PAL* family members containing conserved *PAL* motifs were screened. The physicochemical properties and gene-structure characteristics of *CcPAL* proteins were predicted and analyzed.

收稿日期: 2024-10-24; 修回日期: 2025-03-05

基金项目: 浙江省自然科学基金资助项目 (LQ21C160003, LZ24C160002)

作者简介: 刘晨 (ORCID: 0009-0003-9980-4315), 从事林木再生研究。E-mail: 2528546736@qq.com。周小红 (ORCID: 0000-0001-5049-2849), 副教授, 博士, 从事林木再生研究。E-mail: 20190024@zafu.edu.cn

A phylogenetic tree was constructed using PAL sequences from multiple species. Additionally, the expression changes of *CcPALs* in callus after subculture were examined through transcriptomic analysis. [Result] There were six PALs in *C. cathayensis* in total. The amino-acid lengths ranged from 657 to 760. These proteins were all acidic, stable, and hydrophilic, mainly localized in the nucleus and endoplasmic reticulum. The secondary structure of the protein was dominated by α -helix and random coil. Phylogenetic analysis revealed that CcPAL1–CcPAL4 cluster in the dicotyledonous plant branch, while CcPAL5 and CcPAL6 cluster in the gymnosperm branch. Expression analysis showed that the *CcPALs* displayed different expression patterns after callus subculture. [Conclusion] CcPALs exhibit high conservation. However, CcPAL5 and CcPAL6 were more closely related to gymnosperm PALs, potentially indicating significant functional differences from CcPAL1–CcPAL4 in secondary metabolism. During the callus proliferation and browning process, the expression patterns of *CcPALs* in different branches vary. This variation may potentially regulate the metabolic flux of phenolic substances and thus influence the regeneration process of *C. cathayensis*. [Ch, 7 fig. 4 tab. 37 ref.]

Key words: *Carya cathayensis*; phenylalanine ammonia-lyase (PAL); callus; gene identification; gene family; expression analysis

苯丙氨酸解氨酶 (PAL) 是植物次生代谢途径中的关键酶, 催化苯丙氨酸转化成反式肉桂酸^[1], 进而通过苯丙烷代谢途径调控木质素、黄酮类、香豆素等化合物的合成, 对植物生长发育和环境适应至关重要。PAL 基因家族成员在不同植物中的数量差异显著, 如拟南芥 *Arabidopsis thaliana*^[2] 和烟草 *Nicotiana tabacum* 各有 4 个^[3], 水稻 *Oryza sativa* 有 9 个^[4], 核桃 *Juglans regia* 有 12 个^[5], 这些基因可能具有不同的表达模式和功能冗余。拟南芥 *AtPAL1* 和 *AtPAL2* 影响花青素和原花青素的合成, 且存在功能冗余^[6], 而四重突变体 (*pal1/2/3/4*) 表现出生长和发育缺陷^[2]。在梨 *Pyrus bretschneideri* 中, *PbPAL1* 影响石细胞和木质素合成^[7]。此外, PAL 基因在植物防御反应中也发挥重要作用。火炬松 *Pinus taeda* 的 PAL 在生物和非生物胁迫下的表达模式, 表明其参与植物防御反应^[8]。高粱 *Sorghum bicolor* 蚜虫侵害的研究发现: 甘蔗蚜虫诱导 *SbPAL* 基因表达, 通过正向调控水杨酸介导的防御信号通路抑制蚜虫种群的生长和繁殖^[9]。由此可见, 深入研究 PAL 基因家族的功能, 对于提升植物生长、抗病抗虫能力以及次生代谢物产量具有重要意义。

山核桃 *Carya cathayensis* 为胡桃科 Juglandaceae 山核桃属 *Carya* 落叶乔木, 种植历史悠久^[10]。山核桃次生代谢物丰富, 具有重要的营养价值和观赏价值, 果实为世界四大名优坚果之一^[11]。然而, 在离体培养过程中, 山核桃由于酚类物质如黄酮类、木质素及单宁等大量积累, 存在严重的褐化问题^[12]。PAL 作为调控这些酚类物质合成的关键酶, 其结合位点和表达模式对控制致褐物质合成至关重要。通过对欧芹 *Petroselinum crispum* 中苯丙氨酸解氨酶 (PcPAL) 的疏水结合口袋进行改造, 可以提高 PcPAL 对苯丙氨酸和肉桂酸单取代衍生物的活性和选择性^[13]。而将 PcPAL 的底物特异性调节剂残基 L134、F137、I460 印记到拟南芥 *AtPAL* 和多变鱼腥藻 *Anabaena variabilis* 的 AvPAL 上, 可揭示不同 PALs 在底物特异性上的差异, 并为底物特异性的调节提供更多信息^[14]。通过研究不同 PAL 成员的结构功能和底物偏好, 找到控制褐化物质合成的关键 PAL 基因, 对于山核桃再生体系和遗传转化体系的建立具有重要作用。本研究基于山核桃全基因组测序和转录组测序的结果, 筛选山核桃 PAL (*CcPAL*) 基因家族成员, 研究 CcPAL 编码蛋白的理化性质和结构特点, 为研究 CcPAL 在山核桃酚类物质合成以及愈伤组织增殖与褐化、组织分化和器官再生方面的作用机制提供参考。

1 材料与方法

1.1 材料

山核桃愈伤组织由浙江农林大学玲珑山山核桃试验林#34 山核桃树的未成熟合子胚诱导而来, 果实采集时间为 2021 年 7 月中上旬。跟踪山核桃愈伤组织增殖褐化过程, 选择了继代后 8 个时间点 (继代后 0、3、7、10、14、17、21 和 24 d) 的愈伤组织用于转录组测序。

1.2 CcPAL 基因家族成员鉴定

为鉴定 *CcPAL* 基因家族成员, 从拟南芥信息资源中心数据库 (TAIR)(www.arabidopsis.org) 中获取 4 条 *AtPAL* 蛋白序列, 在美国国家生物技术信息中心 (NCBI) 获得 *JrPAL* (JX069977.1)^[15] 基因序列, 分别作为 query 序列检索山核桃 *PAL* 成员, 进行交叉验证。通过 TBtools 软件^[16] 在山核桃基因组数据^[17] 和转录组数据中进行 BLAST 比对, 设定比对期望值 $E < 10^{-10}$, 筛选山核桃 *PAL* 同源序列。通过 InterPro (www.ebi.ac.uk/interpro) 网站^[18] 确认 *PAL* 蛋白的特征结构域 PF00221, 进一步验证候选基因家族成员。根据山核桃染色体组装结果, 将各成员依次命名为 *CcPAL1*~*CcPAL6*。

1.3 CcPAL 家族蛋白理化性质分析及亚细胞定位预测

利用 ProtParam (<https://web.expasy.org/protparam/>) 网站^[19] 在线预测 *CcPALs* 蛋白的氨基酸长度、相对分子量、等电点、脂肪系数及总平均亲水系数等理化性质。同时通过 DeepLoc2.1 (<https://services.healthtech.dtu.dk/services/DeepLoc-2.1/>) 网站^[20] 对 *CcPAL* 家族蛋白的亚细胞分布位置进行预测分析。

1.4 CcPAL 家族蛋白二级结构和三级结构分析

通过 SOPMA (<https://npsa-prabi.ibcp.fr/>) 网站^[21] 和 Phyre2 (<http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index>) 网站^[22] 对 *CcPALs* 的二级结构和三级结构进行预测分析和可视化。

1.5 CcPAL 基因家族保守结构域和保守基序分析

利用山核桃基因组注释文件, 使用 TBtools 软件 Visualize Gene Structure 模块可视化 *CcPALs* 基因结构图; 通过 NCBI 网站中的 Batch CD-Search 工具 (<https://www.ncbi.nlm.nih.gov/Structure/bwrpsb/bwrpsb.cgi>) 获得 *CcPALs* 蛋白保守结构域, 使用 TBtools 软件 Visualize NCBI CDD Domain Pattern 模块可视化结构域; 通过 MEME (<https://meme-suite.org>) 网站^[23] 分析 *CcPALs* 蛋白的保守基序, 并使用 TBtools 软件 Visualize MEME/MAST Motif Pattern 模块绘制图像。

1.6 CcPAL 家族系统进化树分析

利用 MEGA 11^[24] 对山核桃、拟南芥、水稻 *Oryza sativa*、毛果杨 *Populus trichocarpa*、核桃、杉木 *Cunninghamia lanceolata*、海岸松 *Pinus pinaster* 和火炬松 *Pinus taeda* 等 8 个物种 42 条 *PAL* 蛋白序列进行比对, 采用邻接法 (NJ), 步长设为 1 000, 其余参数默认, 构建系统进化树。

1.7 转录组测序

收集#34 山核桃树继代后 0、3、7、10、14、17、21 和 24 d (对应 A、B、C、D、E、F、G 和 H 样本) 的愈伤组织样品, 每个时间点 3 个重复。样品冻存于 -80°C 冰箱, 研磨粉碎, 参考 MonPure™ Plant RNA Kit (polysaccharide & polyphenol-rich)(Monad, cat.MI10101S) 高多糖多酚 RNA 提取试剂盒产品说明书的方法, 提取山核桃愈伤总 RNA。质检合格后委托杭州联川生物技术股份有限公司构建 cDNA 文库, 对 24 个样品 RNA 进行有参转录组分析。由杭州联川生物技术股份有限公司进行下机原始数据的质控, 浙江农林大学提供山核桃基因组数据^[17] 进行测序比对, 之后对基因或转录本进行组装、每千个碱基的转录每百万映射读取的片段 (FPKM) 定量以及差异基因分析。

1.8 山核桃愈伤组织继代不同时期 CcPALs 表达分析

通过联川生物云平台 (<https://www.omicstudio.cn/home>) 分析山核桃愈伤组织的 *CcPALs* 基因表达量, 并制作热图。

选择继代后 0、7、14 和 21 d 山核桃愈伤组织进行实时荧光定量 PCR (RT-qPCR) 表达量验证。采用多糖多酚 RNA 提取试剂盒 (DR0407050, 浙江易思得生物科技有限公司) 分别提取愈伤组织的总 RNA。通过反转录试剂盒 (RR036A, 北京宝日医生物技术有限公司) 分别合成 cDNA, 置于 -80°C 冰箱备用。将愈伤组织的 cDNA 质量浓度调整到 $100\text{ ng}\cdot\mu\text{L}^{-1}$, 采用 NovoStart® SYBR qPCR SuperMix Plus(E096, 苏州近岸蛋白质科技股份有限公司) 在 CFX96 Touch 荧光定量 PCR 仪 (BIO-RAD, 美国) 上检测 *CcPALs* 基因的表达情况。 β -actin 为内参基因, 引物信息见表 1, 每个样品 3 个生物学重复, 通过 GraphPad Prism 软件进行单因素方差分析 (one-way ANOVA) 和制图。

2 结果与分析

2.1 CcPAL 基因家族鉴定及理化特征分析

通过生物信息学分析发现: 无论是以拟南芥 *AtPALs* 作为 query 序列检索, 还是以亲缘关系更近的

表 1 实验中使用的 RT-qPCR 引物

Table 1 RT-qPCR primers used in the experiment		
名称	上游引物(5'→3')	下游引物(5'→3')
<i>β-actin</i>	ATGGCCGATGCTGAGGATATT	CACCAGTGTGTCGAGGTCTGC
<i>CcPAL1</i>	ATGGGTTGCCTTCGAATCTC	CCAGGGAGTTCACATCTTGG
<i>CcPAL2</i>	AGGTTACTCGGGCATCAGAT	GCCTATCAAAAGACCAGCGA
<i>CcPAL3</i>	CTGCAGCGAAACCTATCCAT	TGGCAGTGTTC AAGTTCCTC
<i>CcPAL4</i>	CAGTGC GACATACCCATTGA	TTCCTCAAAGACCCCGATCT
<i>CcPAL5</i>	GCTTGGGGTAGAGGTAGCAA	GCATTGGAGAAGTCACCGTG
<i>CcPAL6</i>	TCAGGGGATCTTGTACCGTT	GCAGAAATCCCAGCTCTTGT

核桃 *JrPALs* 进行检索，都是鉴定到 6 个 *CcPALs* 蛋白，分别命名为 *CcPAL1* (CCA0546S0150)、*CcPAL2* (CCA0589S0228)、*CcPAL3* (CCA0589S0229)、*CcPAL4* (CCA1013S0012)、*CcPAL5* (CCA1591S0016) 和 *CcPAL6* (CCA1591S0017)。分析 *CcPAL* 家族成员的氨基酸组成和理化性质发现：*CcPALs* 蛋白含有的氨基酸数量为 657~760 个；6 个 *CcPALs* 家族成员中亮氨酸 (Leu)、丙氨酸 (Ala)、甘氨酸 (Gly) 和谷氨酸 (Glu) 含量最为丰富；分子量为 71.87~83.88 kDa；等电点为 6.01~6.41，均小于 7.00，属于酸性蛋白；所有 *CcPALs* 蛋白的不稳定系数均小于 40，稳定性较高；6 个 *CcPALs* 蛋白平均亲水性系数均小于 0，均属于亲水蛋白，其中 *CcPAL6* 亲水性最强 (表 2)。

表 2 CcPALs 蛋白理化性质

Table 2 Physicochemical properties of CcPALs protein							
蛋白质	氨基酸数量/个	主要氨基酸组分(相对占比/%)	理论等电点	分子量/kDa	不稳定系数	脂肪指数	平均亲水性系数
CcPAL1	713	Leu(10.7)、Ala(9.8)、Gly(7.7)	6.41	77.54	32.96	91.85	-0.149
CcPAL2	657	Leu(11.4)、Ala(9.1)、Glu(7.3)	6.07	71.87	32.62	95.65	-0.114
CcPAL3	708	Leu(11.0)、Ala(9.2)、Gly(7.5)	6.23	77.34	32.40	92.75	-0.189
CcPAL4	710	Leu(10.3)、Ala(9.9)、Glu(8.0)	6.10	77.50	37.69	91.56	-0.160
CcPAL5	712	Leu(10.1)、Ala(8.6)、Glu(7.7)	6.01	78.57	34.77	90.27	-0.174
CcPAL6	760	Leu(10.3)、Ala(8.8)、Glu(7.5)	6.22	83.88	35.75	89.08	-0.214

说明：Leu. 亮氨酸; Ala. 丙氨酸; Gly. 甘氨酸; Glu. 谷氨酸。

对 *CcPALs* 蛋白进行亚细胞定位预测发现：除 *CcPAL5* 定位在细胞核和内质网上外，*CcPAL1*、*CcPAL2*、*CcPAL3*、*CcPAL4* 和 *CcPAL6* 可能在细胞核、线粒体和内质网均有定位 (表 3)，说明 *CcPALs* 参与了基因转录调控、氧化磷酸化和蛋白质的加工等环节，在细胞生长周期中发挥着作用。

对 *CcPALs* 的 6 个蛋白进行二级结构和三级结构预测分析发现： α -螺旋与无规则卷曲在二级结构中占比较高。其中 *CcPAL4* 的 α -螺旋占比最高 (56.48%)，*CcPAL5* 的延伸链结构、 β -转角占比最高 (9.13%、6.88%)；*CcPAL6* 的无规则卷曲占比最高 (33.29%)(表 4)。进一步对 6 个 *CcPAL* 蛋白进行了三级结构模型预测 (图 1)，结果显示编码蛋白三级结构主要以 α -螺旋为主要结构元件。

2.2 CcPALs 基因结构及蛋白结构域分析

对 *CcPALs* 进行基因结构分析发现：*CcPAL1*~*CcPAL4* 都各有 2 个外显子和 1 个内含子，结构相似，而 *CcPAL5* 和 *CcPAL6* 均只有 1 个外显子，且基因中不包含内含子序列 (图 2A)。6 条 *CcPALs* 基因编码序列 (CDS) 长度为 1 974~2 283 bp。

通过 NCBI 的 Conserved Domain Search 对山核桃 *PAL* 家族氨基酸序列进行蛋白功能分析 (图 2B) 显示：PLN02457 (phenylalanine ammoni-lyase) 结构域在 *CcPALs* 中高度保守，它是 Lyase_I_like 超家族的一部分。对 *CcPAL* 家族蛋白进行保守基序分析，6 个 *CcPALs* 蛋白都含有 10 个保守基序，且蛋白结构大体一致，说明 *CcPALs* 成员之间具有较高的保守性 (图 2C 和图 2D)。

2.3 CcPAL 家族蛋白进化关系及保守基序分析

MIO 结构域在氨基酸的转化过程中通过消除氨形成不饱和产物，与 *PAL* 的催化活性密切相关。它

表 3 CcPALs 蛋白亚细胞定位预测

Table 3 Subcellular localization prediction of CcPALs protein in *C. cathayensis*

蛋白质	预测定位	定位信号	膜类型
CcPAL1	细胞核、线粒体、内质网	核定位信号	外周、脂锚定
CcPAL2	细胞核、线粒体、内质网	核定位信号	外周、脂锚定
CcPAL3	细胞核、线粒体、内质网	核定位信号	外周、脂锚定
CcPAL4	细胞核、线粒体、内质网	核定位信号	外周、脂锚定
CcPAL5	细胞核、内质网	核定位信号	外周、脂锚定
CcPAL6	细胞核、线粒体、内质网	核定位信号	外周、脂锚定

表 4 CcPALs 蛋白二级结构

Table 4 Secondary structure of CcPALs protein

蛋白质	占比/%			
	α-螺旋	延伸链结构	β-转角	无规则卷曲
CcPAL1	53.02	8.70	6.73	31.56
CcPAL2	55.71	8.37	6.09	29.83
CcPAL3	55.37	7.49	6.50	30.65
CcPAL4	56.48	7.46	5.49	30.56
CcPAL5	55.34	9.13	6.88	28.65
CcPAL6	51.97	8.29	6.45	33.29

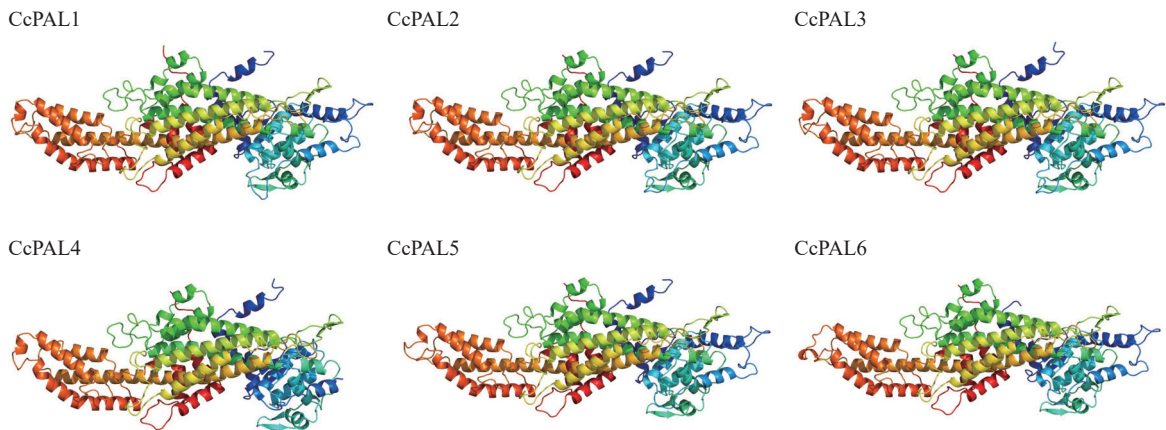


图 1 CcPALs 蛋白质三级结构
Figure 1 Tertiary structure of CcPALs protein

包含 PAL 家族的保守序列 (GTITASGDLVPLSYIAG) 和丙氨酸-丝氨酸-甘氨酸 (Ala-Ser-Gly) 三肽活性中心。多重序列比对结果 (图 3) 显示：CcPALs 所有序列都含有 MIO 结构域 (如 CcPAL1 的 195~211 个氨基酸) 和 Core Domain 中的 PAL 活性位点 (如 CcPAL1 上的³⁴⁸Tyr 和⁴⁹¹Gly)，说明 6 个 CcPALs 都具有解氨酶催化活性。而 MIO 结构域在核桃 JrPAL2、JrPAL3、JrPAL6 和 JrPAL9 中完全缺失，在 JrPAL8 中缺失了 4 个氨基酸，且 JrPAL1、JrPAL2、JrPAL3、JrPAL6、JrPAL7 和 JrPAL9 与其他 PAL 在蛋白长度上表现出很大差异，缺乏明显的主要结构域 PRK09367。另外 Shielding Domain 中的 Thr 残基被认为是潜在的磷酸化位点 (CcPAL1 中⁵⁴⁶Thr)。

构建 *PAL* 基因家族系统进化树 (图 4) 发现：*PAL* 家族蛋白主要分成 3 个支。Group I 主要包含双子叶植物的 *PAL* 蛋白 (用橙色表示)，包括拟南芥 AtPAL1~AtPAL4、毛果杨 PtPAL1~PtPAL5，核桃 JrPAL1、JrPAL2-2、JrPAL3、JrPAL5 和 JrPAL7 以及山核桃 CcPAL1~CcPAL4；Group II 则由单子叶植物的 *PAL* 蛋白组成 (用绿色表示)，如水稻 OsPAL1~OsPAL9；Group III 由裸子植物 *PAL* 蛋白 (海岸松 PpPAL1 和 PpPAL2，火炬松 PtaPAL1 以及杉木 CIPAL1~CIPAL3) 和 2 条 CcPAL (CcPAL5、CcPAL6)、7 条 JrPAL (JrPAL2-1、JrPAL2-3、JrPAL4-1、JrPAL4-2、JrPAL6、JrPAL8 和 JrPAL9) 构成 (用蓝色表示)。系统进化树分析结果显示：单子叶和双子叶植物的 *PAL* 家族成员在进化过程中呈现明显的分歧。山核桃的 CcPAL5 和 CcPAL6 与裸子植物的 *PAL* 聚在 Group III，表明在进化上这 2 个 *PAL* 蛋白与裸子植物同源性较高，功能上可能更相似。这一发现为进一步探索 *PAL* 基因家族的功能分化和进化提供了新的视角。

2.4 山核桃愈伤组织增殖过程中 *CcPALs* 基因的表达分析

基于 Illumina NovaSeq6000 测序平台构建文库。山核桃愈伤组织增殖褐化过程中 8 个时间点构建了 24 个 cDNA 文库。各样本 Q30 碱基百分比 (碱基质量为 30 的碱基在总碱基数量中所占的比例) 均高于

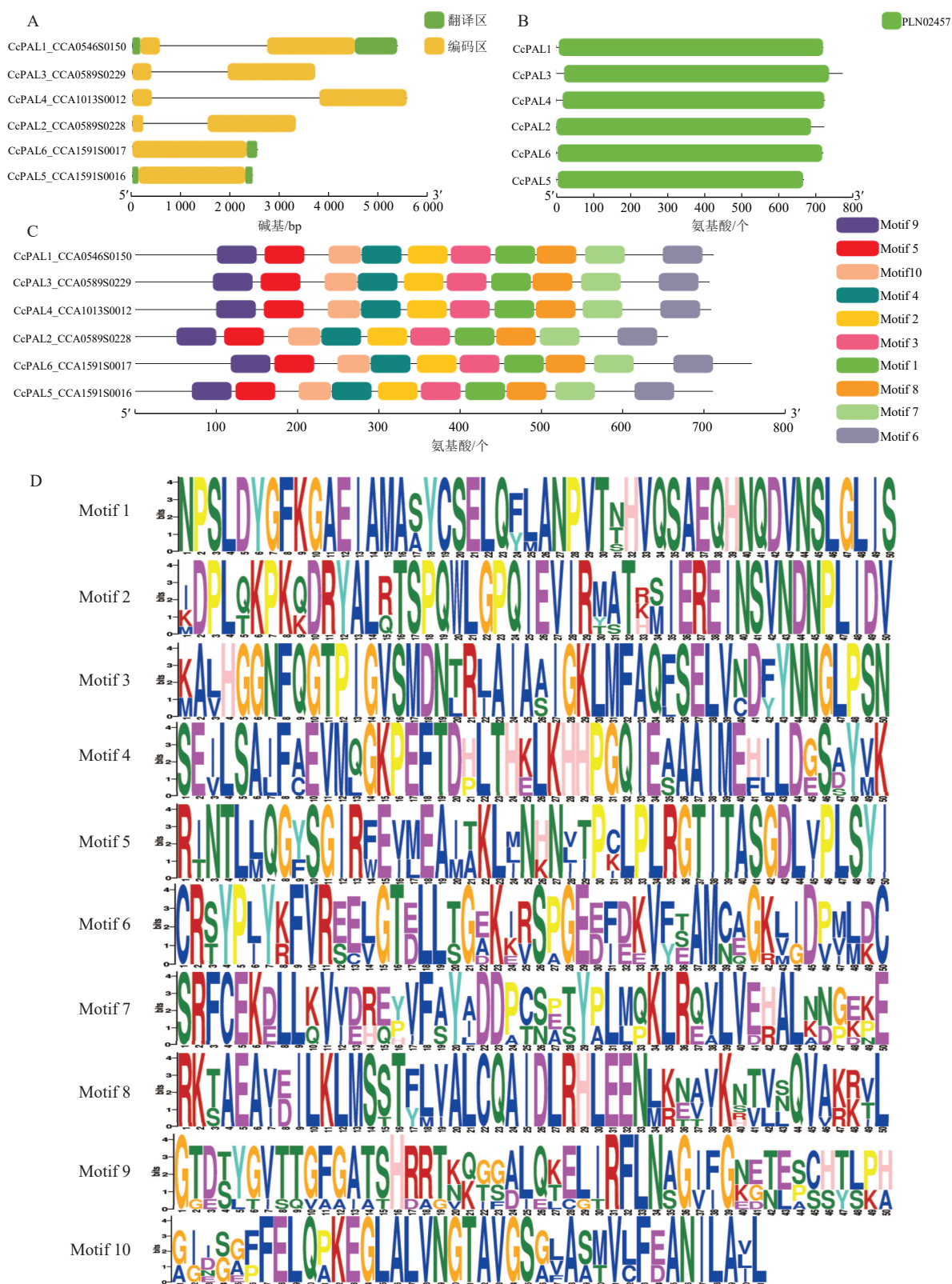


图2 *CcPAL* 家族的基因结构 (A)、保守蛋白 (B)、保守基序 (C) 和基序标识分析 (D)

Figure 2 Analysis of gene structure (A), conserved proteins (B), conserved motifs (C), and motif logos (D) of the *CcPAL* family

96%, GC 碱基占总碱基的 44%~45%, 符合正态分布。测序样本间转录本的主成分分析 (PCA) 显示: 主成分 1、主成分 2 的贡献率分别为 74.11% 和 10.92%, 每组 3 个样本间比较聚集, 各自组内重复性较好, 可用于后续分析 (图 5)。通过转录组测序, 分析了山核桃愈伤组织在继代培养过程中从增殖到褐化阶段 *CcPALs* 基因的表达模式, 并根据 FPKM 值绘制了热图 (图 6)。结果显示: 聚类在同一分支的

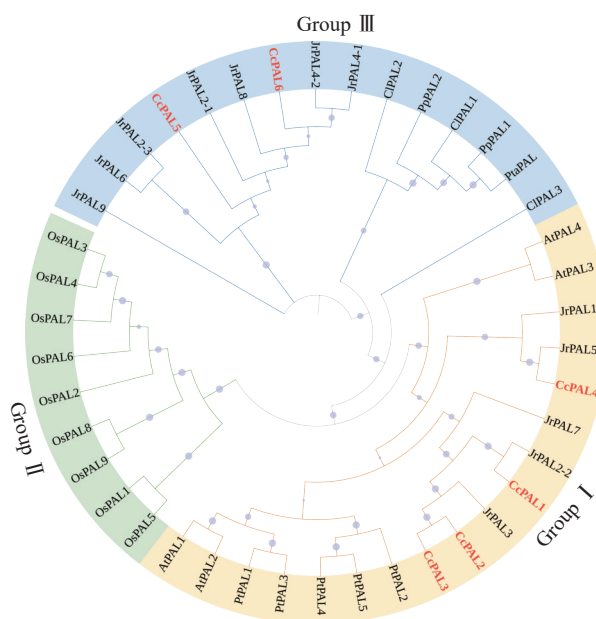


图4 8个物种PAL家族蛋白的进化发育树
Figure 4 Phylogenetic tree of PAL family proteins in 8 species

CcPAL6 整体表达量较高,且在继代后0 d表达量最高,说明继代后愈伤组织受伤期间 *CcPAL5* 和 *CcPAL6* 已经开始响应表达; *CcPAL1*~*CcPAL4* 基因在愈伤组织开始增殖(继代后7 d)和出现褐化(继代后21 d)时表达量较高,说明2个不同分支下的 *PAL* 基因对山核桃愈伤组织的环境适应和细胞增殖褐化有不同的影响(图7)。

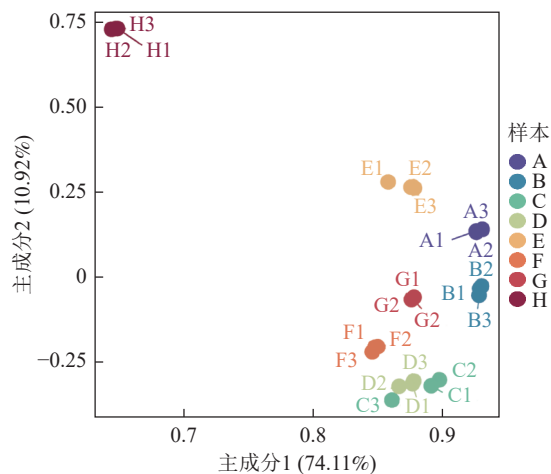


图5 测序样本间转录本PCA分析
Figure 5 PCA analysis of transcripts between sequencing samples

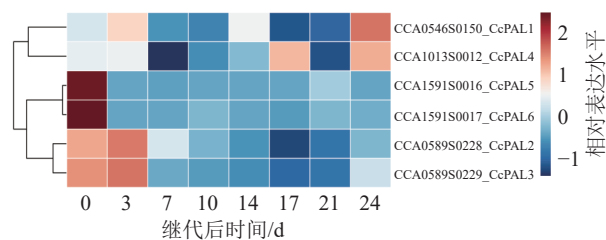
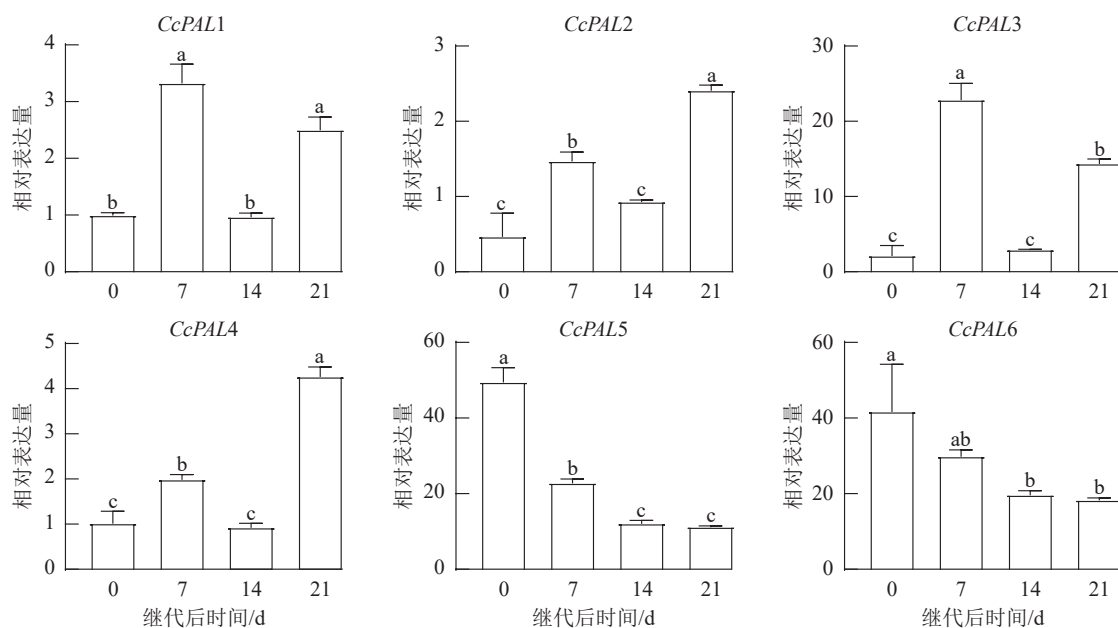


图6 山核桃愈伤组织增殖过程中 *CcPALs* 基因表达热图
Figure 6 Heat map of *CcPALs* expression during the proliferation of callus in *C. cathayensis*



不同小写字母表示不同时间差异显著 ($P < 0.05$)。

图7 继代后不同时间下山核桃愈伤组织中的 *CcPALs* 基因表达

Figure 7 Expression of *CcPALs* gene in the *C. cathayensis* callus at different time points after subculture

3 讨论

CcPALs 蛋白的理化性质、基因结构分析结果与其他植物中的 PAL 蛋白具有一定的相似性和关联性^[15, 25-26]。在功能鉴定的 PAL 基因中都包含保守的 GTITASGDLVPLSYIA 特征基序、丙氨酸-丝氨酸-甘氨酸 (Ala-Ser-Gly) 三肽活性中心^[27-28]和 PAL 活性位点氨基酸^[29-30], 而且在序列的屏蔽域中存在着磷酸化位点, 即 Thr 残基^[31]。所有 CcPAL 蛋白的不稳定系数小于 40, 属于稳定蛋白。这一特性对于山核桃在生长过程中维持 PAL 活性至关重要。在其他植物中, 稳定的 PAL 蛋白也有助于植物稳定地参与次生代谢途径。例如, 在火炬松中, *PAL* 基因在植物防御反应中发挥着重要作用, 其稳定性可能有助于火炬松在面对生物和非生物胁迫时保持代谢的稳定^[8]。此外, CcPAL 蛋白平均亲水性系数均小于 0, 均属于亲水蛋白。亲水性质可能使得这些蛋白更容易与其他分子相互作用, 参与到各种代谢过程中。

基因结构方面, *CcPAL1~CcPAL4* 与 *CcPAL5* 和 *CcPAL6* 的外显子和内含子结构不同。这种差异可能反映了不同的进化路径和分化功能。在拟南芥中, 不同的 *PAL* 成员也可能具有不同的基因结构和表达模式^[2, 6]。例如, 拟南芥中的 *PAL1* 和 *PAL2* 在黄酮类生物合成中有冗余作用, 四重突变体表现出矮小和不育, 这暗示着不同的 *PAL* 基因在功能上可能存在差异, 而这种差异可能与基因结构的不同有关。*CcPAL1~CcPAL4* 与 *CcPAL5* 和 *CcPAL6* 的基因结构差异可能也导致了它们在山核桃生长发育和代谢过程中的不同功能。

基因复制是进化创新和功能分化的主要机制^[32]。核桃中存在 12 个 *PAL* 基因, 彼此间蛋白结构差异较大, 部分功能结构域缺失; 而山核桃中存在 6 个 *PAL* 基因, 成员间氨基酸数量相当, 结构较为完整, 存在个别氨基酸突变, 说明 *PAL* 在核桃中发生了重大的变异和较多的基因复制活动, 在山核桃中 *PAL* 更为保守。裸子植物的 *PAL* 基因亚群与被子植物的亚群是姐妹群, 这可能意味着这些基因编码的 PAL 同工酶具有相似的功能, 或者受到相似的发育调控机制的调控^[8]。根据已测序物种的基因序列构建包含裸子植物、单子叶植物和双子叶植物的 PAL 系统发育树, 结果显示: CcPAL5、CcPAL6 和裸子植物 PAL 聚在一起, 而 CcPAL1~CcPAL4 和双子叶植物聚在一起且具有高自展值, 说明 CcPAL5、CcPAL6 与 CcPAL1~CcPAL4 可能有不同的进化起源。其他 PAL 基因在双子叶植物和单子叶植物分化后呈单系, 表明基因复制在双子叶植物和单子叶植物分化后 PAL 基因的进化中很重要^[33-35]。

山核桃愈伤组织增殖过程中不同的 *CcPALs* 基因表现出不同的表达模式, 这与其他植物在生长发育或胁迫响应过程中 *PAL* 基因的差异表达现象类似^[36]。在拟南芥中, *PAL1* 和 *PAL2* 在黄酮类生物合成中有冗余作用, 而本研究的山核桃中, *CcPAL2* 和 *CcPAL3* 之间、*CcPAL5* 和 *CcPAL6* 之间表达趋势相同, 可能彼此间功能也存在冗余。在继代后 3 d, *CcPAL5* 和 *CcPAL6* 表达量显著上调, 说明其可能参与机械损伤后抗逆相关黄酮类物质的合成, 提高愈伤组织的环境适应能力。*CcPAL4* 在褐化出现时 (继代后 17 d) 上调表达, 可能参与了致褐物质的合成。这种差异表达可能是由于不同的 *PAL* 基因对不同的信号和环境因素做出了响应, 或者在不同的代谢途径中发挥了特定的作用^[37]。

4 结论

本研究表明: CcPAL 蛋白具有酸性、稳定、亲水特性, 可能参与基因转录调控等生物过程。系统进化分析显示: CcPAL5 和 CcPAL6 与裸子植物关系更近。不同 *CcPAL* 基因在愈伤组织增殖中表现出不同表达模式, 影响组织增殖和褐化。研究结果为理解山核桃生长发育和次生代谢物合成提供了重要信息, 并为 *CcPAL* 在山核桃酚类物质合成、愈伤组织增殖与褐化、组织分化和器官再生中的作用机制研究奠定了基础, 为山核桃遗传改良和产业发展提供了理论支持。

5 参考文献

- [1] SAITO K, YONEKURA-SAKAKIBARA K, NAKABAYASHI R, *et al.* The flavonoid biosynthetic pathway in *Arabidopsis*: structural and genetic diversity [J]. *Plant Physiology and Biochemistry*, 2013, 72: 21-34.
- [2] HUANG Junli, GU Min, LAI Zhibing, *et al.* Functional analysis of the *Arabidopsis PAL* gene family in plant growth, development, and response to environmental stress [J]. *Plant Physiology*, 2010, 153(4): 1526-1538.

- [3] REICHERT A I, HE Xianzhi, DIXON R A. Phenylalanine ammonia-lyase (PAL) from tobacco (*Nicotiana tabacum*): characterization of the four tobacco PAL genes and active heterotetrameric enzymes [J]. *Biochemical Journal*, 2009, **424**(2): 233–242.
- [4] 曾嘉丽, 欧阳林娟, 刘家林, 等. 水稻 PAL 基因的全基因组分析及胁迫表达研究[J]. 基因组学与应用生物学, 2018, **37**(9): 3881–3888.
- ZENG Jiali, OUYANG Linjuan, LIU Jialin, *et al.* Whole genome analysis and stress expression research of PAL gene in rice [J]. *Genomics and Applied Biology*, 2018, **37**(9): 3881–3888.
- [5] YAN Feng, LI Huaizhu, ZHAO Peng. Genome-wide identification and transcriptional expression of the PAL gene family in common walnut (*Juglans regia* L.) [J/OL]. *Genes*, 2019, **10**(1): 46 [2024-10-01]. DOI: [10.3390/genes10010046](https://doi.org/10.3390/genes10010046).
- [6] ROHDE A, MORREEL K, RALPH J, *et al.* Molecular phenotyping of the pal1 and pal2 mutants of *Arabidopsis thaliana* reveals far-reaching consequences on phenylpropanoid, amino acid, and carbohydrate metabolism [J]. *The Plant Cell*, 2004, **16**(10): 2749–2771.
- [7] LI Guohui, SONG Cheng, MANZOOR M A, *et al.* Functional and kinetics of two efficient phenylalanine ammonia lyase from *Pyrus bretschneideri* [J/OL]. *BMC Plant Biology*, 2023, **23**(1): 612 [2024-10-01]. DOI: [10.1186/s12870-023-04586-0](https://doi.org/10.1186/s12870-023-04586-0).
- [8] BAGAL U R, LEEBENS-MACK J H, LORENZ W W, *et al.* The phenylalanine ammonia lyase (PAL) gene family shows a gymnosperm-specific lineage [J/OL]. *BMC Genomics*, 2012, **13**(suppl 3): S1 [2024-10-01]. DOI: [10.1186/1471-2164-13-S3-S1235](https://doi.org/10.1186/1471-2164-13-S3-S1235).
- [9] PANT S, HUANG Yinghua. Genome-wide studies of PAL genes in sorghum and their responses to aphid infestation [J/OL]. *Scientific Reports*, 2022, **12**(1): 22537 [2024-10-01]. DOI: [10.1038/s41598-022-25214-1](https://doi.org/10.1038/s41598-022-25214-1).
- [10] 谢林峰, 凌晓晓, 黄圣妍, 等. 临安区山核桃林地土壤水解酶活性空间分布特征及土壤肥力评价[J]. 浙江农林大学学报, 2022, **39**(3): 625–634.
- XIE Linfeng, LING Xiaoxiao, HUANG Shengyan, *et al.* Spatial distribution characteristics of soil hydrolase activities and soil fertility evaluation of *Carya cathayensis* forests in Lin'an District [J]. *Journal of Zhejiang A&F University*, 2022, **39**(3): 625–634.
- [11] 张璐璐, 贾桂民, 叶建丰, 等. 浙江临安市核桃干腐病发生发展规律[J]. 浙江农林大学学报, 2013, **30**(1): 148–152.
- ZHANG Lulu, JIA Guimin, YE Jianfeng, *et al.* Frequency of *Carya cathayensis* canker disease in Lin'an City, Zhejiang Province [J]. *Journal of Zhejiang A&F University*, 2013, **30**(1): 148–152.
- [12] LIU Chen, FAN Hongrui, ZHANG Jiaqi, *et al.* Combating browning: mechanisms and management strategies in *in vitro* culture of economic woody plants [J/OL]. *Forestry Research*, 2024, **4**: e032 [2024-10-01]. DOI: [10.48130/forres-0024-0026](https://doi.org/10.48130/forres-0024-0026).
- [13] NAGY E Z A, TOROK S D, LANG P A, *et al.* Mapping the hydrophobic substrate binding site of phenylalanine ammonia-lyase from *Petroselinum crispum* [J]. *ACS Catalysis*, 2019, **9**(9): 8825–8834.
- [14] TOMOIAGĂ R B, TOROK S D, FILIP A, *et al.* Phenylalanine ammonia-lyases: combining protein engineering and natural diversity [J]. *Applied Microbiology and Biotechnology*, 2023, **107**(4): 1243–1256.
- [15] XU Feng, DENG Guang, CHENG Shuiyuan, *et al.* Molecular cloning, characterization and expression of the phenylalanine ammonia-lyase gene from *Juglans regia* [J]. *Molecules*, 2012, **17**(7): 7810–7823.
- [16] CHEN Chengjie, CHEN Hao, ZHANG Yi, *et al.* TBtools: an integrative toolkit developed for interactive analyses of big biological data [J]. *Molecular Plant*, 2020, **13**(8): 1194–1202.
- [17] HUANG Youjun, XIAO Lihong, ZHANG Zhongren, *et al.* The genomes of pecan and Chinese hickory provide insights into *Carya* evolution and nut nutrition [J/OL]. *GigaScience*, 2019, **8**(5): giz036 [2024-10-01]. DOI: [10.1093/gigascience/giz036](https://doi.org/10.1093/gigascience/giz036).
- [18] PAYSAN-LAFOSSE T, BLUM M, CHUGURANSKY S, *et al.* InterPro in 2022 [J]. *Nucleic Acids Research*, 2023, **51**(D1): D418–D427.
- [19] GASTEIGER E, HOOGLAND C, GATTIKER A, *et al.* Protein identification and analysis tools on the ExPASy server [M]// WALKER J M. *The Proteomics Protocols Handbook*. Totowa N J: Humana Press, 2005: 571–607.
- [20] THUMULURI V, ARMENTEROS J J A, JOHANSEN A R, *et al.* DeepLoc 2.0: multi-label subcellular localization prediction using protein language models [J]. *Nucleic Acids Research*, 2022, **50**(W1): W228–W234.
- [21] GEORJON C, DELÉAGE G. SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments [J]. *Computer Applications in the Biosciences*, 1995, **11**(6): 681–684.

- [22] KELLEY L A, MEZULIS S, YATES C M, *et al.* The Phyre2 web portal for protein modeling, prediction and analysis [J]. *Nature Protocols*, 2015, **10**(6): 845–858.
- [23] BAILEY T L, JOHNSON J, GRANT C E, *et al.* The MEME suite [J]. *Nucleic Acids Research*, 2015, **43**(W1): W39–W49.
- [24] TAMURA K, STECHER G, KUMAR S. MEGA11: molecular evolutionary genetics analysis version 11 [J]. *Molecular Biology and Evolution*, 2021, **38**(7): 3022–3027.
- [25] OKADA T, MIKAGE M, SEKITA S. Molecular characterization of the phenylalanine ammonia-lyase from *Ephedra sinica* [J]. *Biological & Pharmaceutical Bulletin*, 2008, **31**(12): 2194–2199.
- [26] WANNER L A, LI Guoqing, WARE D, *et al.* The phenylalanine ammonia-lyase gene family in *Arabidopsis thaliana* [J]. *Plant Molecular Biology*, 1995, **27**(2): 327–338.
- [27] 施圆圆, 张声祥, 赵德蕊, 等. 异叶天南星苯丙氨酸解氨酶基因的克隆与蛋白结构分析[J]. *植物科学学报*, 2019, **37**(2): 221–229.
- SHI Yuanyuan, ZHANG Shengxiang, ZHAO Derui, *et al.* Gene cloning and structure characterization of phenylalanine ammonia-lyase from *Arisaema heterophyllum* [J]. *Plant Science Journal*, 2019, **37**(2): 221–229.
- [28] LIN Weiping, LIU Ao, WENG Caihong, *et al.* Cloning and characterization of a novel phenylalanine ammonia-lyase gene from *Inonotus baumii* [J]. *Enzyme and Microbial Technology*, 2018, **112**: 52–58.
- [29] DONG Chunjuan, CAO Ning, ZHANG Zhigang, *et al.* Phenylalanine ammonia-lyase gene families in cucurbit species: structure, evolution, and expression [J]. *Journal of Integrative Agriculture*, 2016, **15**(6): 1239–1255.
- [30] SHANG Qingmao, LI Liang, DONG Chunjuan. Multiple tandem duplication of the phenylalanine ammonia-lyase genes in *Cucumis sativus* L. [J]. *Planta*, 2012, **236**(4): 1093–1105.
- [31] DONG Chunjuan, SHANG Qingmao. Genome-wide characterization of phenylalanine ammonia-lyase gene family in watermelon (*Citrullus lanatus*) [J]. *Planta*, 2013, **238**(1): 35–49.
- [32] YANG Ji, HUANG Jinxia, GU Hongya, *et al.* Duplication and adaptive evolution of the *Chalcone* synthase genes of *Dendranthema* (Asteraceae) [J]. *Molecular Biology and Evolution*, 2002, **19**(10): 1752–1759.
- [33] WU Zhihua, GUI Songtao, WANG Shuzhen, *et al.* Molecular evolution and functional characterisation of an ancient phenylalanine ammonia-lyase gene (*NnPAL1*) from *Nelumbo nucifera*: novel insight into the evolution of the PAL family in angiosperms[J/OL]. *BMC Evolutionary Biology*, 2014, **14**: 100[2024-10-01]. DOI: [10.1186/1471-2148-14-100](https://doi.org/10.1186/1471-2148-14-100).
- [34] HAMBERGER B, ELLIS M, FRIEDMANN M, *et al.* Genome-wide analyses of phenylpropanoid-related genes in *Populus trichocarpa*, *Arabidopsis thaliana*, and *Oryza sativa*: the *Populus* lignin toolbox and conservation and diversification of angiosperm gene families [J]. *Canadian Journal of Botany*, 2007, **85**(12): 1182–1201.
- [35] LI Guohui, WANG Han, CHENG Xi, *et al.* Comparative genomic analysis of the *PAL* genes in five Rosaceae species and functional identification of Chinese white pear[J/OL]. *PeerJ*, 2019, **7**: e8064[2024-10-01]. DOI: [10.7717/peerj.8064](https://doi.org/10.7717/peerj.8064).
- [36] AMJAD M, WANG Yuexia, HAN Shiming, *et al.* Genome wide identification of phenylalanine ammonia-lyase (*PAL*) gene family in *Cucumis sativus* (cucumber) against abiotic stress[J/OL]. *BMC Genomic Data*, 2024, **25**(1): 76[2024-10-01]. DOI: [10.1186/s12863-024-01259-1](https://doi.org/10.1186/s12863-024-01259-1).
- [37] RASOOL F, UZAIR M, NAEEM M K, *et al.* Phenylalanine ammonia-lyase (*PAL*) genes family in wheat (*Triticum aestivum* L.): genome-wide characterization and expression profiling[J/OL]. *Agronomy*, 2021, **11**(12): 2511[2024-10-01]. DOI: [10.3390/agronomy11122511](https://doi.org/10.3390/agronomy11122511).