

木薯 PG 基因家族鉴定及表达分析

唐湘宁^{1,2,3}, 王晓彤^{1,2,3}, 汪路花^{2,3,4}, 李瑞梅^{2,3}, 姚 远^{2,3}, 王亚杰^{2,3},
郭建春^{2,3*}, 刘 姣^{2,3*}

1. 海南大学生命健康学院, 海南海口 570228; 2. 中国热带农业科学院热带生物技术研究所/热带作物生物育种全国重点实验室, 海南海口 571101; 3. 中国热带农业科学院三亚研究院, 海南三亚 572025; 4. 海南大学热带农林学院, 海南儋州 571737

摘 要: 多聚半乳糖醛酸酶 (polygalacturonase, PG) 在果胶的分解过程中起着关键作用, 从而导致细胞壁结构的变化, 参与植物的生长发育和应对逆境胁迫等多个阶段。尽管 PG 基因在其他植物中已有研究, 但在木薯中尚未见对该基因家族成员鉴定和功能研究的报道。本研究从木薯基因组中鉴定出 89 个 MePGs, 编码的氨基酸数量在 183~808 之间, 分子量介于 19.75~87.07 kDa 之间, 理论等电点 (pI) 在 4.64~9.71 之间, 大部分成员预测定位于细胞膜。染色体定位分析显示 MePGs 不均匀地分布于 17 条染色体上。基于进化关系, MePGs 被分为 A~G 共 7 个亚组, 亚组成员基因结构相似, 且存在明显的串联重复现象; 共线性分析结果表明, 相比拟南芥, 木薯 PG 家族基因的亲缘性与橡胶更接近。MePGs 启动子区域富含光响应、激素响应与逆境响应元件。MePGs 基因具有组织特异性, 与木薯生长发育有关。在木薯采后变质过程中, MePGs 表现出相似的表达模式, 特别是 MePG20、MePG21、MePG25、MePG64 和 MePG72, 均呈现先上升后下降的表达趋势, 表明这些基因可能在木薯块根采后变质过程中先响应逆境胁迫, 高表达以分解果胶, 随后随着细胞壁的水解完成而降低表达, 这表明 PG 基因在逆境胁迫中发挥重要作用。综上所述, MePG 基因家族成员可能通过片段复制和内含子减少的方式进化, 并可能通过感知不同种类的信号而行使不同的功能, 从而形成多样化的表达模式。本研究为进一步探索 MePGs 在木薯块根采后变质机制中的作用提供基础。

关键词: 木薯; PG 基因家族; 表达分析

中图分类号: S31 **文献标志码:** A

Identification and Expression Analysis of the PG Gene Family in Cassava

TANG Xiangning^{1,2,3}, WANG Xiaotong^{1,2,3}, WANG Luhua^{2,3,4}, Li Ruimei^{2,3}, Yao Yuan^{2,3}, WANG Yajie^{2,3},
GUO Jianchun^{2,3*}, LIU Jiao^{2,3*}

1. School of Life and Health Sciences, Hainan University, Haikou, Hainan 570228, China; 2. Institute of Tropical Biotechnology, Chinese Academy of Tropical Agricultural Sciences / National Key Laboratory of Tropical Crop Breeding, Haikou, Hainan 571101, China; 3. Sanya Research Institute, Chinese Academy of Tropical Agricultural Sciences, Sanya, Hainan 572025, China; 4. School of Tropical Agriculture and Forestry, Hainan University, Danzhou, Hainan 571737, China

Abstract: Polygalacturonase (PG) plays a crucial role in plant growth and development, and in stress responses, by participating in the degradation of pectin and altering cell wall structure. Although studies on PG genes have been conducted in various plants, there is a lack of research on the identification and functional characterization of the PG gene family in cassava (*Manihot esculenta* Crantz). This study identified a total of 89 members of the MePG family in the cassava genome, encoding proteins with 183 to 808 amino acids, molecular weights ranging from 19.75 to 87.07 kDa, and theoretical pI values between 4.64 and 9.71. Most of the family members are predicted to be localized in the cell membrane. Chromosome mapping analysis revealed that MePG family members were unevenly distributed across 17

收稿日期 2025-01-17; **接受日期** 2025-02-12

基金项目 海南省自然科学基金项目 (No. 323MS087); 中央级公益性科研院所基本科研业务费专项 (No. 1630052024008); 现代农业产业技术体系建设专项 (No. CARS-11-HNGJC)。

作者简介 唐湘宁 (1999—), 男, 硕士研究生, 研究方向: 木薯分子育种。*通信作者 (Corresponding author): 郭建春 (GUO Jianchun), E-mail: guojianchun@itbb.org.cn; 刘 姣 (LIU Jiao), E-mail: liujiao@itbb.org.cn。

chromosomes. Based on evolutionary relationships, *MePGs* were classified into seven subgroups (A to G), with similar gene structures within subgroups and evidence of tandem duplication. Comparative analysis indicated that the PG family genes in cassava were more closely related to rubber than to *Arabidopsis*. The promoter regions of *MePG* genes were enriched with elements responsive to light, hormones, and stress. *MePG* genes exhibit tissue specificity and were associated with cassava growth and development. During postharvest deterioration of cassava, *MePGs* displayed similar expression patterns, particularly *MePG20*, *MePG21*, *MePG25*, *MePG64* and *MePG72*, which showed an initial increase followed by a decrease in expression, suggesting that the genes may initially respond to stress, upregulate to break down pectin, and then downregulate as cell wall hydrolysis is completed. This implies a significant role for PG in stress responses. Collectively, the findings suggest that *MePG* gene family members may have evolved through segmental duplication and intron loss and may perform different functions by sensing various types of signals, leading to diverse expression patterns. This would study provide a foundation for further exploration of the role of *MePGs* in the postharvest deterioration mechanism of cassava roots.

Keywords: cassava; PG gene family; expression analysis

DOI: 10.3969/j.issn.1000-2561.2025.06.004

植物细胞壁是细胞的外部支撑结构, 主要由纤维素、半纤维素和果胶构成, 为细胞提供结构支持和保护功能^[1]。果胶在植物细胞壁中扮演多重角色, 包括维持细胞结构, 促进细胞间粘连, 增强细胞壁的延展性, 提供生长和发育信号以及参与植物的防御响应^[2]。

果胶的生物合成涉及多种糖基转移酶和蛋白质复合体^[3]。果胶由同源半乳糖醛酸 (homogalacturonan, HG)、鼠李半乳糖醛酸聚糖 I (rhamnogalacturonan I, RG-I)、鼠李半乳糖醛酸聚糖 II (rhamnogalacturonan II, RG-II)、木糖半乳糖醛酸聚糖 (xylogalacturonan, XGA) 等组成^[4-6], 其中 HG 约占果胶总量的 65%^[7]。

多聚半乳糖醛酸酶 (polygalacturonase, PG) 是一类催化果胶降解的酶, 是 HG 的主要酶之一, 参与细胞壁中果胶的降解, 影响细胞壁稳定性, 在植物各生长阶段以及不同组织中扮演着关键角色^[8]。

植物中 PG 蛋白通常包含 4 个保守的结构域。其中, I 域和 II 域分别包含核心序列 SPNTDG 和 GDDC, 这 2 个序列中包含 3 个天冬氨酸残基 (D), 这些残基参与构成酶的催化活性位点。III 域由序列 CGPGHG 组成, 其中的组氨酸 (H) 在催化过程中发挥重要作用。IV 域的序列为 RIK, 该结构域涉及与底物羧基端的离子相互作用。这些结构域的保守性和功能多样性是 PG 蛋白在植物细胞壁降解过程中发挥作用的基础^[9]。在这 4 个保守结构域中, 半胱氨酸残基 (C) 均有较好的保守性, 推测其可能参与维持蛋白质的三级结构^[10]。

木薯 (*Manihot esculenta* Crantz) 是以富含淀

粉的块根为主要利用形式的重要粮食和能源作物, 其块根在采收后极易发生生理性变质 (postharvest physiological deterioration, PPD)^[11]。木薯变质分为初发性变质和次发性变质 2 个阶段, 初发性变质取决于内源酶作用, 而次发性变质主要由微生物感染引起。在次发性变质过程中, 作为植物细胞第一道屏障的细胞壁结构发生变化, 微生物随之侵入引发次级变质, 继而发生严重的氧化损伤和细胞死亡, 最终导致整个块根腐烂^[12]。在这个过程中, 细胞壁结构的变化可能与 PG 家族基因的活性密切相关。

目前 PG 基因家族在许多物种中都进行了全基因组鉴定及分析, 如榴莲^[13]、甘薯^[14]和苹果^[15]等, 但在木薯中暂未见报道。本研究旨在木薯基因组数据库中筛选鉴定木薯 PG 家族成员, 对其理化性质、基因结构和启动子序列等进行分析, 并且分析其在木薯不同组织部位和块根采后变质过程中的表达情况, 以探讨 PG 基因在木薯块根采后变质过程中的作用。

1 材料与方法

1.1 材料

木薯种植基地位于海南省三亚市南滨农场, 年降水量 1400~1600 mm, 年平均气温 24~26℃。选择数棵生长成熟 (约 150 d) 的华南 8 号 (SC8) 木薯作为试验材料。将未损伤的木薯块根完整挖出, 置于实验室室温存放。每次选择完整的块根在 0、1、3、5、7、10、14、21 d 进行切片取样, 并存放在 -80℃ 冰箱中保存备用。

木薯 (*Manihot esculenta* Crantz) 基因组及注释数据下载于 NCBI 数据中心 (<https://www.ncbi.>

nlm.nih.gov/), 拟南芥[*Arabidopsis thaliana* (L.) Heynh.]和橡胶[*Hevea brasiliensis* (Willd. ex A. Juss.) Müll. Arg.]PG 基因家族序列下载于 NCBI (<https://www.ncbi.nlm.nih.gov/>)和 Phytozome(<https://phytozome-next.jgi.doe.gov/>) 数据库。

1.2 方法

1.2.1 木薯 PG 基因家族成员鉴定 通过 TBtools 软件获取木薯蛋白序列, 在拟南芥的蛋白序列中搜索与木薯 PG 蛋白相似的蛋白序列, 将这些蛋白放到 swissprot(<http://www.gpmaw.com/html/swissprot.html>) 数据库中进行验证, 筛选出木薯多聚半乳糖醛酸酶 (MePG) 家族基因的候选序列。通过 InterPro (<https://www.ebi.ac.uk/interpro/>) 在线软件和 TBtools 软件的 Advanced Hmmer Search 插件, 将候选序列与已知的 PG 保守结构域 (PF00295) 进行筛选, 以确定最终的 MePG 基因家族成员, 并对筛选出的基因进行重命名。

1.2.2 蛋白质特性分析 利用 TBtools 软件的 Protein Parameter Calc 插件对 MePG 蛋白的特性进行预测, 包括氨基酸数量、分子量 (MW) 和等电点 (pI)。通过 PSORT Prediction (<https://psort.hgc.jp/form.html>) 在线软件进行亚细胞定位的预测。

1.2.3 木薯 PG 蛋白多序列比对 利用 DNAMAN 软件对 PG 蛋白家族进行蛋白多序列比较。

1.2.4 木薯 PG 蛋白系统发育树的构建 使用 TBtools 软件的 One Step Build a ML Tree 插件创建 MePG 蛋白序列的系统发育树。

1.2.5 木薯 PG 家族基因结构和 motif 分析 利用 NCBI batch wab CD-search tool (<https://www.ncbi.nlm.nih.gov/Structure/bwrpsb/bwrpsb.cgi>) 在线软

件进行保守结构域分析, 通过 MEME(<https://me-me-suite.org/meme/tools/meme>) 在线软件分析 MePG 蛋白的保守基序; 利用 TBtools 软件进行可视化处理。

1.2.6 染色体分布和物种间共线性 使用 TBtools 软件确定 89 个 MePG 基因的染色体位置。使用 TBtools 软件的 One Step MCSanX 插件进行木薯与拟南芥和橡胶中的基因同源性分析。

1.2.7 MePG 基因家族启动子顺式作用元件分析及功能预测 利用 TBtools 软件提取 MePG 基因家族序列上游长度为 2000 bp 的启动子序列, 利用 PlantCARE (<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>) 软件进行预测分析, 并通过 TBtools 软件进行可视化分析。

1.2.8 MePG 家族成员的表达模式分析 根据本实验室的 SC8 木薯转录组数据, 分析 MePG 基因在各组织部位以及将 SC8 木薯采后变质 21 d 的 FPKM 值, 利用 ChiPlot (<https://www.chiplot.online/>) 在线软件绘制 MePG 基因的表达热图。选取 MePG20/21/25/64/72 基因, 设计 qRT-PCR 特异性引物 (表 1), 由北京擎科生物科技有限公司 (海口) 合成引物。使用 M5 HiPer Plant RNeasy Complex Mini Kit “繁可简” 多糖多酚植物 RNA 提取试剂盒 (北京聚合美生物科技有限公司) 提取块根 RNA, 使用 MonScript™ RTIII All-in-One Mix with dsDNase (莫纳生物科技有限公司) 进行逆转录。使用 Taq-HS SYBR® Green qPCR Premix (Universal) (江苏愚公生命科技有限公司) 和上海力康 HealForce 的 CG-05 仪器进行 qRT-PCR 验证。使用 2^{-ΔΔCt} 法计算基因的相对表达量。

表 1 木薯 MePG 及内参基因引物序列
Tab. 1 Primer sequences of MePG and reference genes in Cassava

基因名称 Gene name	上游引物序列 (5'-3') Forward primer sequence (5'-3')	下游引物序列 (5'-3') Reverse primer sequence (5'-3')
MePG20	CCCCAAATCTCCCTACACCG	ATGCCATATTCGTCCCAGCC
MePG21	GAGTCCCCTGTTTCTCCAGC	CGGAAGTGTCGTCTGTGGAA
MePG25	CTATGACCCAAATGCGCTGC	TCTTCTTGGCCTTAGCTGCC
MePG64	TGCTTGCAATGGCTAGGAGAG	GTCTTCTTTTGGTGGCGGTG
MePG72	TACAGCCGGCCTCATCTAGT	GCCGATGGTGTAAGGGGATT
Tublin	ATGCGGTTCTTGATGTTGTTC	TCGGTGAAGGGAATACAGAGA

2 结果与分析

2.1 MePG 基因家族全基因组鉴定与理化分析 利用生物信息学方法分析木薯基因组数据库, 共鉴定到 89 个 PG 基因。根据染色体位置分

布, 将 89 个 MePG 基因依次命名为 MePG1~MePG89。氨基酸序列分析表明, MePG 蛋白质的氨基酸数目为 183 (MePG75) ~808 (MePG22), 平均长度为 442 aa; 分子量为 19.75 (MePG75) ~

87.08 (MePG22) kDa, 平均值为 47.79 kDa; 理论等电点为 4.64 (MePG21)~9.71 (MePG32) (表 2)。亚细胞定位预测结果显示 82 个 MePG 家族成员定位于细胞膜, 其余成员可能分布在细胞壁、叶绿体、细胞质或细胞核中。

MePG 的氨基酸序列比对结果如图 1 所示, 89 个 MePGs 中有 33 个 MePGs 的氨基酸序列完整包含了 4 个典型的 PG 保守结构域 I (序列为 SPNTDGI)、II (序列为 GDDC)、III (序列为

CGPGHG)和IV(序列为 RIK), 如 MePG1、MePG2、MePG3、MePG4 等。部分 MePGs 的氨基酸序列存在丢失或缺少的现象, 其中 42 个 MePGs 的氨基酸序列缺少结构域 I, 19 个 MePGs 的氨基酸序列缺少结构域 II, 38 个 MePGs 的氨基酸序列缺少结构域 III, 10 个 MePGs 的氨基酸序列缺少结构域 IV。MePG17、MePG18、MePG20 的氨基酸序列同时缺少 4 个保守结构域, 并伴有氨基酸替换或缺失, 推测这些 PG 基因可能具有更高的进化程度。

表 2 木薯 PG 基因基本信息
Tab. 2 Basic informations of PG genes in cassava

基因名称 Gene name	基因 ID Gene ID	氨基酸数量 Number of amino acids	分子量 Molecular weight/kDa	等电点 pI	外显子数 Number of exons	亚细胞定位 Subcellular localization
MePG1	LOC110629774	393	41.48 698	8.82	4	细胞膜
MePG2	LOC110630512	392	41.39 796	8.82	4	
MePG3	LOC110631211	400	42.01 564	9.02	4	
MePG4	LOC110630006	384	40.54 294	8.91	4	
MePG5	LOC110607276	447	48.93 046	7.70	5	
MePG6	LOC110619718	468	50.84 848	4.85	6	
MePG7	LOC110613895	389	41.50 687	9.24	4	
MePG8	LOC110612992	394	42.24 033	8.92	4	
MePG9	LOC110614534	390	41.48 746	8.92	4	
MePG10	LOC110630199	467	51.35 530	6.29	6	
MePG11	LOC110617273	466	51.48 177	5.52	6	
MePG12	LOC110609136	449	48.51 313	6.41	5	
MePG13	LOC110608650	482	53.12 119	5.10	6	
MePG14	LOC110604904	390	42.35 668	9.38	4	
MePG15	LOC110610957	387	41.72 722	4.70	7	
MePG16	LOC110611892	493	55.19 641	9.00	5	
MePG17	LOC110611171	492	53.61 593	4.93	3	叶绿体/细胞质
MePG18	LOC110611158	497	53.31 323	5.87	4	细胞膜/细胞壁/细胞质
MePG19	LOC110610407	411	44.43 841	8.81	5	细胞膜
MePG20	LOC110610596	454	49.53 701	5.37	7	
MePG21	LOC110610961	483	52.21 556	4.64	6	
MePG22	LOC110612661	808	87.07 602	6.96	16	
MePG23	LOC110615930	501	54.91 085	5.31	6	
MePG24	LOC110616197	466	51.30 865	5.56	6	
MePG25	LOC110615380	496	54.22 802	5.96	6	
MePG26	LOC110615372	388	40.72 296	8.95	4	
MePG27	LOC110614253	392	41.88 369	9.10	4	
MePG28	LOC110615169	394	41.19 004	5.50	5	
MePG29	LOC110616253	392	41.51 907	8.90	4	
MePG30	LOC110616343	390	42.08 508	9.14	4	
MePG31	LOC110616346	314	33.05 354	9.35	4	
MePG32	LOC110614543	391	41.78 594	9.71	4	
MePG33	LOC110615373	393	42.02 998	8.61	4	

续表 2 木薯 PG 基因基本信息
Tab. 2 Basic informations of PG genes in cassava (continued)

基因名称 Gene name	基因 ID Gene ID	氨基酸数量 Number of amino acids	分子量 Molecular weight/kDa	等电点 pI	外显子数 Number of exons	亚细胞定位 Subcellular localization
MePG34	LOC110617356	340	36.07 284	4.98	5	细胞膜
MePG35	LOC110607361	392	41.58 553	6.11	4	
MePG36	LOC110619169	437	48.09 009	9.28	9	
MePG37	LOC110618420	392	41.65 042	9.03	4	
MePG38	LOC110618724	383	40.21 470	8.74	4	
MePG39	LOC110619762	394	41.39 830	5.62	4	
MePG40	LOC110618995	392	41.26 276	8.89	4	
MePG41	LOC110619752	394	41.13 693	5.37	4	
MePG42	LOC122723995	329	35.71 986	8.41	3	
MePG43	LOC122724002	344	37.27 057	8.85	5	
MePG44	LOC110621204	466	52.02 487	6.42	3	
MePG45	LOC110621781	465	51.36 716	6.14	9	
MePG46	LOC110621657	474	51.37 488	6.80	6	
MePG47	LOC110622452	453	50.32 548	5.14	9	
MePG48	LOC110607362	342	37.23 980	8.60	5	
MePG49	LOC122724619	387	41.11 911	5.76	5	
MePG50	LOC110623231	392	41.57 151	6.32	4	
MePG51	LOC110623246	598	65.50 002	8.78	6	
MePG52	LOC122724630	394	41.91 553	5.71	5	
MePG53	LOC122724631	285	30.54 989	5.85	3	
MePG54	LOC110623259	379	40.56 998	5.35	5	
MePG55	LOC122724638	392	41.56 052	6.60	4	
MePG56	LOC110608437	508	55.12 757	7.46	6	
MePG57	LOC110623288	469	50.19 411	7.07	5	
MePG58	LOC122724983	497	53.53 649	8.72	7	
MePG59	LOC110625530	452	49.36 066	8.83	10	
MePG60	LOC110625554	787	84.47 748	5.21	18	
MePG61	LOC122724997	566	61.42 298	5.27	14	
MePG62	LOC110625492	312	33.93 008	6.59	8	
MePG63	LOC110628585	491	53.31 959	5.15	9	
MePG64	LOC110627693	519	56.49 224	8.42	7	
MePG65	LOC110627266	509	55.87 234	8.39	5	
MePG66	LOC110630542	465	51.30 474	8.92	8	细胞膜/细胞壁
MePG67	LOC110630402	482	52.58 219	5.40	11	
MePG68	LOC110629701	515	56.27 061	6.42	7	
MePG69	LOC110599771	493	54.18 174	5.40	6	细胞膜
MePG70	LOC110631264	474	52.19 344	9.11	6	
MePG71	LOC110599772	477	52.25 114	8.62	4	
MePG72	LOC110602494	454	49.47 686	5.26	6	叶绿体
MePG73	LOC110601489	405	43.89 332	9.52	5	
MePG74	LOC110601493	473	50.55 794	6.45	7	
MePG75	LOC110604004	183	19.75 071	9.59	2	
MePG76	LOC122722067	392	41.59 960	6.33	4	
MePG77	LOC110603533	506	54.85 671	6.83	3	

续表 2 木薯 PG 基因基本信息
Tab. 2 Basic informations of PG genes in cassava (continued)

基因名称 Gene name	基因 ID Gene ID	氨基酸数量 Number of amino acids	分子量 Molecular weight/kDa	等电点 pI	外显子数 Number of exons	亚细胞定位 Subcellular localization
MePG78	LOC110604137	491	52.64 941	5.52	4	细胞膜/细胞质
MePG79	LOC110603022	492	55.42 656	9.25	5	细胞膜
MePG80	LOC110605462	395	42.59 880	9.27	7	
MePG81	LOC122722202	366	40.12 109	9.04	3	
MePG82	LOC122722215	390	41.87 305	9.26	4	
MePG83	LOC110605622	480	51.99 343	6.00	7	
MePG84	LOC122722428	393	42.32 833	8.79	4	
MePG85	LOC122722349	395	42.58 388	9.08	4	
MePG86	LOC122722437	346	37.66 107	8.60	3	细胞膜/细胞壁
MePG87	LOC122722439	767	83.84 571	9.09	10	细胞膜
MePG88	LOC110623260	392	41.55 958	6.60	4	
MePG89	LOC110608409	733	79.40 345	9.19	9	

2.2 MePG 基因家族进化树分析

为研究木薯 PG 基因的进化关系，本研究采用邻接法（NJ），构建木薯和拟南芥（*Arabidopsis thaliana*）的 PG 基因家族成员蛋白序列（68 个 AtPG，89 个 MePG）进化树。如图 2 所示，2 个物种的 157 个 PG 蛋白被分为 7 个亚族（Group A~Group G）。A 至 G 亚族分别包含 9、6、21、27、16、5、5 个木薯 PG。2 个物种在 A 亚族、B 亚族和 E 亚族中的基因数量大致相同，2 个物种的 PG 蛋白在每个亚族中不均匀分布。亚族的划分基于蛋白序列的相似度和进化关系，不同亚族的基因通常具有不同的功能。

2.3 MePG 基因保守基序及基因结构分析

在基因家族进化过程中，通常伴有外显子的缺失和增加，为了理解木薯 PG 基因的特性，通过 TBtools 软件对木薯 89 个 MePG 基因的外显子和内含子结构进行分析，并绘制基因结构图。结果表明（图 3A），同一亚族中的 PG 基因具有相似的基因结构，其内含子的数量接近，而不同亚族（如 Group F 和 Group G）包含基因之间内含子数量差异较大。所有 MePG 基因中外显子数量最少为 3 个，而位于 Group F 中的 MePG60 外显子的数量最多（18 个）。以上结果表明，在木薯 PG 基因进化过程中，外显子的丢失、获得与 PG 基因的功能多样密切相关，这是因为基因家族内部的基因在分子功能上是冗余的，它们可能经历了进化选择过程，基因中断由内含子插入是真核基因的特征之一。

木薯 PG 的保守基序分析结果如图 3B 所示，亲缘关系越近的成员其基序结构越相似。大部分 MePGs 都包含有 Motif 10 和 Motif 6，且均以 N→C 端按序排列。其中 Motif 1 的序列包含“SPNTDGI”（保守结构域 I）与“GDDC”（保守结构域 II），Motif 3 和 Motif 5 的序列分别包含“CGPGHG”（保守结构域 III）、“RIK”（保守结构域 IV）。保守结构域 I 包含 1 个天冬氨酸残基（D），该残基参与酶的催化活性位点，直接影响 PG 酶的催化效率，缺少该结构域的基因可能催化活性较低或功能不同。保守结构域 II 同样包含天冬氨酸残基，参与底物的结合和催化过程，缺少该结构域的基因可能无法有效结合果胶底物，导致酶活性降低。保守结构域 III 中的组氨酸（H）在催化过程中发挥重要作用，尤其是在果胶的水解过程中，缺少该结构域的基因可能在果胶降解过程中效率较低。保守结构域 IV 可能与底物羧基端的离子相互作用有关，影响 PG 酶与果胶的结合能力，缺少该结构域的基因可能无法有效结合果胶，导致酶活性下降。

从进化关系上看，亲缘关系较近的同亚族的 PG 基因在基因结构组成和蛋白结构域分布上也具有较高保守性。如图 3C 所示，68 个（76.4%）木薯 PG 家族成员包含 1 个共同的结构域 PL-6 superfamily，16 个（17.9%）木薯 PG 家族成员包含 1 个共同的结构域 Pgu1 superfamily，4 个（4.4%）木薯 PG 家族成员包含结构域 Pectate_lyase_3 superfamily，但 MePG17 并不包含这

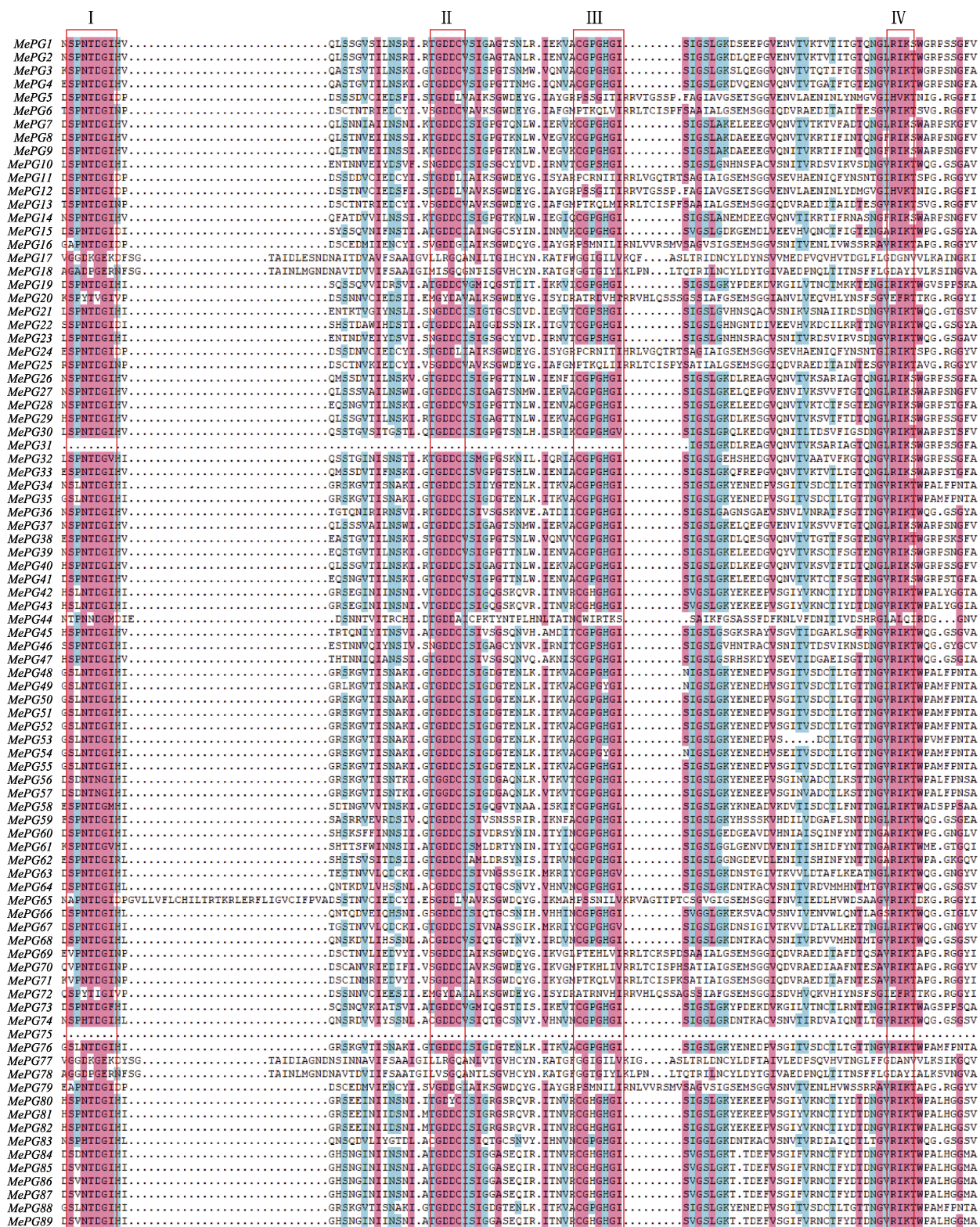


图 1 木薯 PG 氨基酸的多序列比对

Fig. 1 Multiple sequence alignment of PG amino acid in cassava

3 种结构域中的任何一种。

2.4 MePG 染色体定位分析

用 TBtools 软件中的染色体定位工具对 89 个

MePG 基因在染色体上的位置分析发现, 89 个 MePG 基因不均匀地分布在 17 条染色体上, 其中第 1、5 和 9 号染色体上的基因数量最多, 各含有

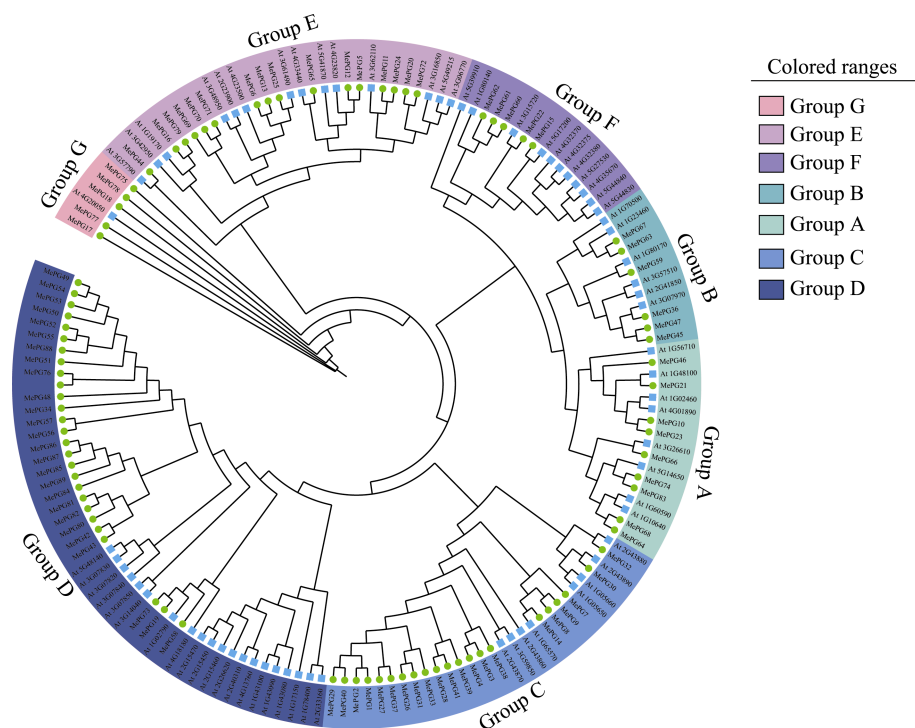
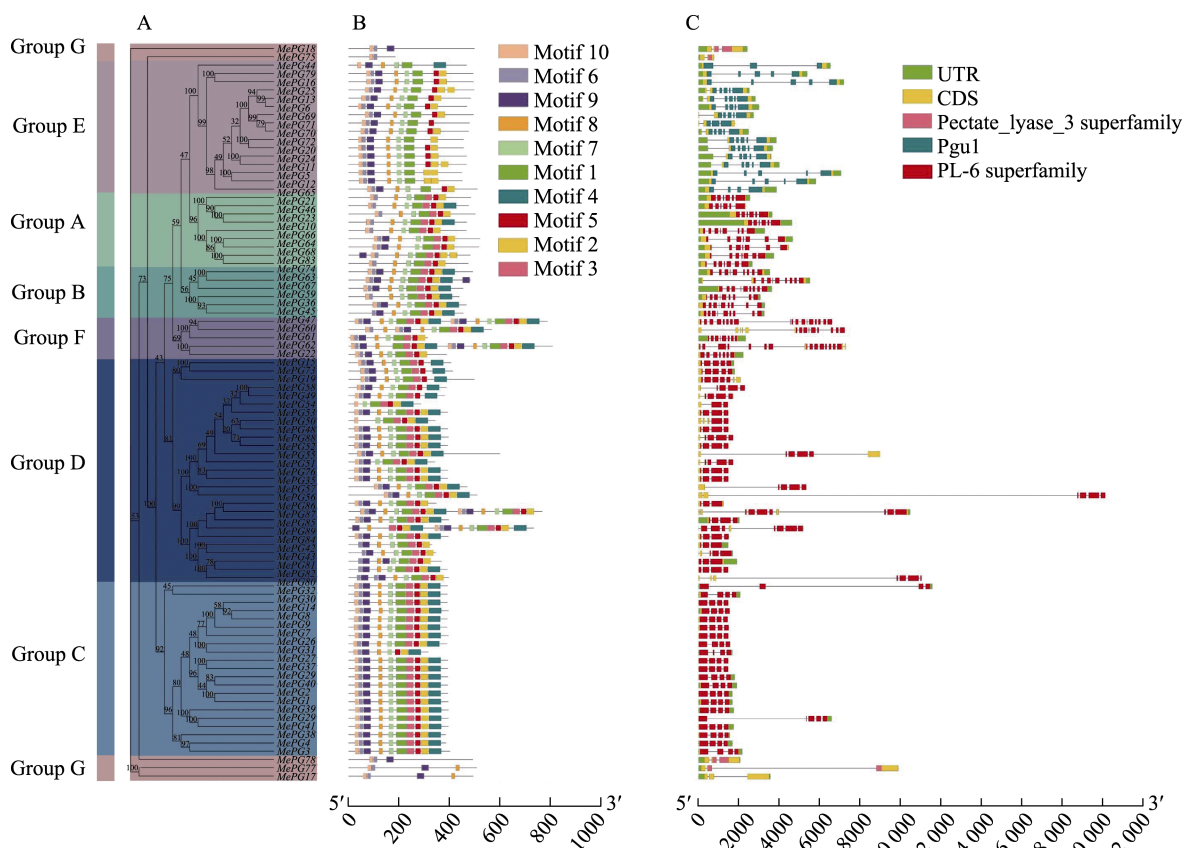


图 2 木薯和拟南芥 PG 蛋白的系统进化树

Fig. 2 Phylogenetic tree of the PG gene family in cassava and *Arabidopsis*

A: MePGs 系统发育树; B: MePGs 保守基序分布; C: MePGs 基因结构。

A: Phylogenetic tree of MePGs; B: Distribution of conserved motifs in MePGs; C: Gene structure of MePGs.

图 3 MePG 基因的基因结构和保守基序分析

Fig. 3 Analysis of the gene structure and conserved motifs of MePG genes

11 个 *MePG* 基因，而第 4、6 号染色体上的基因数量最少，仅含有 1 个 *MePG* 基因（图 4）。这种不均匀的分布可能反映了基因复制事件和染色体重排。*MePG88* 和 *MePG89* 由于基因组未完全

组装，仅定位到基因组的特定区域。*MePG* 基因在染色体上的分布模式表明，这些 *PG* 基因可能通过片段复制和串联复制的方式进化，且不同染色体上的基因可能具有不同的功能或表达模式。

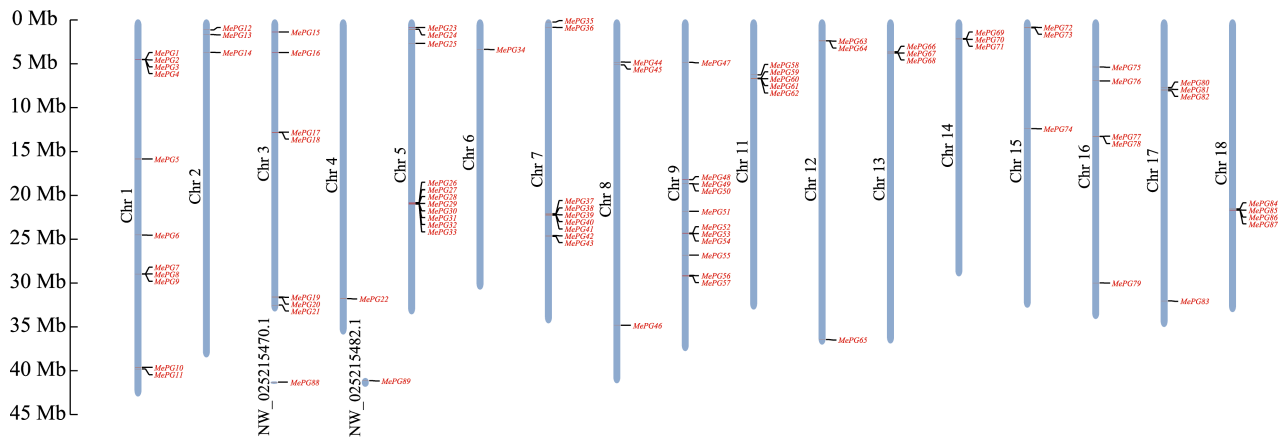


图 4 木薯 PG 基因家族染色体定位
Fig. 4 Chromosome localization map of cassava PG gene family

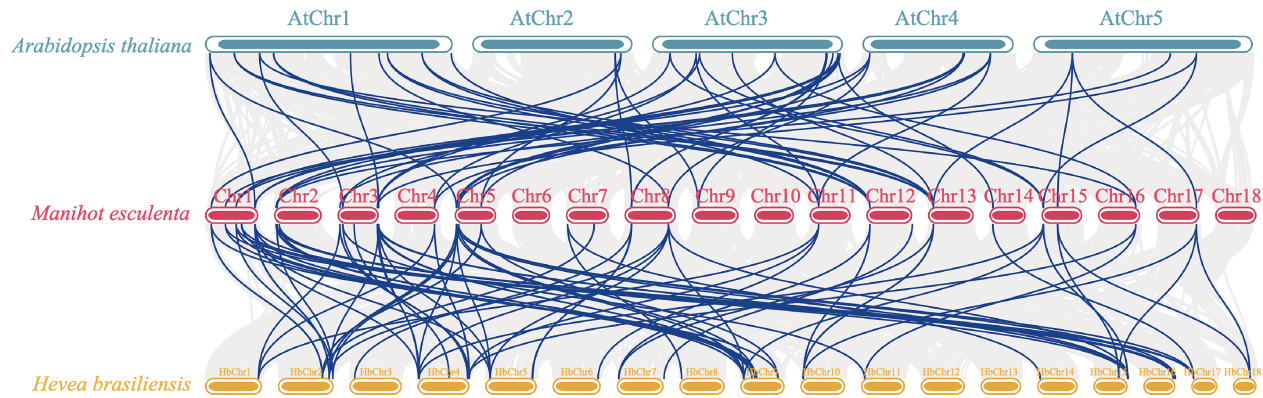
2.5 *MePG* 基因的物种间共线性分析

为了进一步了解木薯 *PG* 基因家族的系统发育机制，本研究分析了木薯与拟南芥、橡胶的种间共线性关系。如图 5 所示，在木薯与拟南芥共有 51 对 *PG* 同源基因对，木薯与橡胶共有 70 对 *PG* 同源基因对，表明木薯与橡胶的亲缘关系更近。

2.6 *MePG* 基因启动子顺式作用元件分析

为了弄清 *MePG* 基因的潜在功能和调控机制，本研究截取 89 个木薯 *PG* 基因上游 2000 bp 的启动子序列，利用 PlantCARE 软件对 *MePGs* 启动子区域的顺式作用元件进行分析。结果（图

6）显示，大部分基因都包含与植物激素和非生物胁迫响应相关的作用元件，这些元件与木薯的生长发育息息相关。所含顺式作用元件中，最多的是光响应元件（671 个），其次是茉莉酸甲酯响应元件（185 个）和脱落酸响应元件（140 个），其余作用元件除损伤响应元件（6 个）和栅栏叶肉细胞分化响应元件（4 个）外，数量分布范围在 20~68 个之间。不同 *MePG* 基因家族成员启动子区域所含的顺式作用元件类型差异较大，大部分 *MePGs* 的启动子区域含有光响应元件，表明这些基因可能受到光信号的调控，尤其是在光合作用活跃的组织中。茉莉酸甲酯和脱落酸响应元件



蓝色线条代表木薯和拟南芥、橡胶的共线性 *PG* 基因对，灰色线条表示木薯和拟南芥、橡胶的同源基因。
Blue lines represent the collinear *PG* gene pairs between cassava, arabidopsis and rubber, while gray lines indicate the orthologous genes between cassava, arabidopsis and rubber.

图 5 木薯与拟南芥、橡胶的 *PG* 基因种间共线性分析
Fig. 5 Cross-species collinearity analysis of *PG* genes between cassava and arabidopsis, rubber

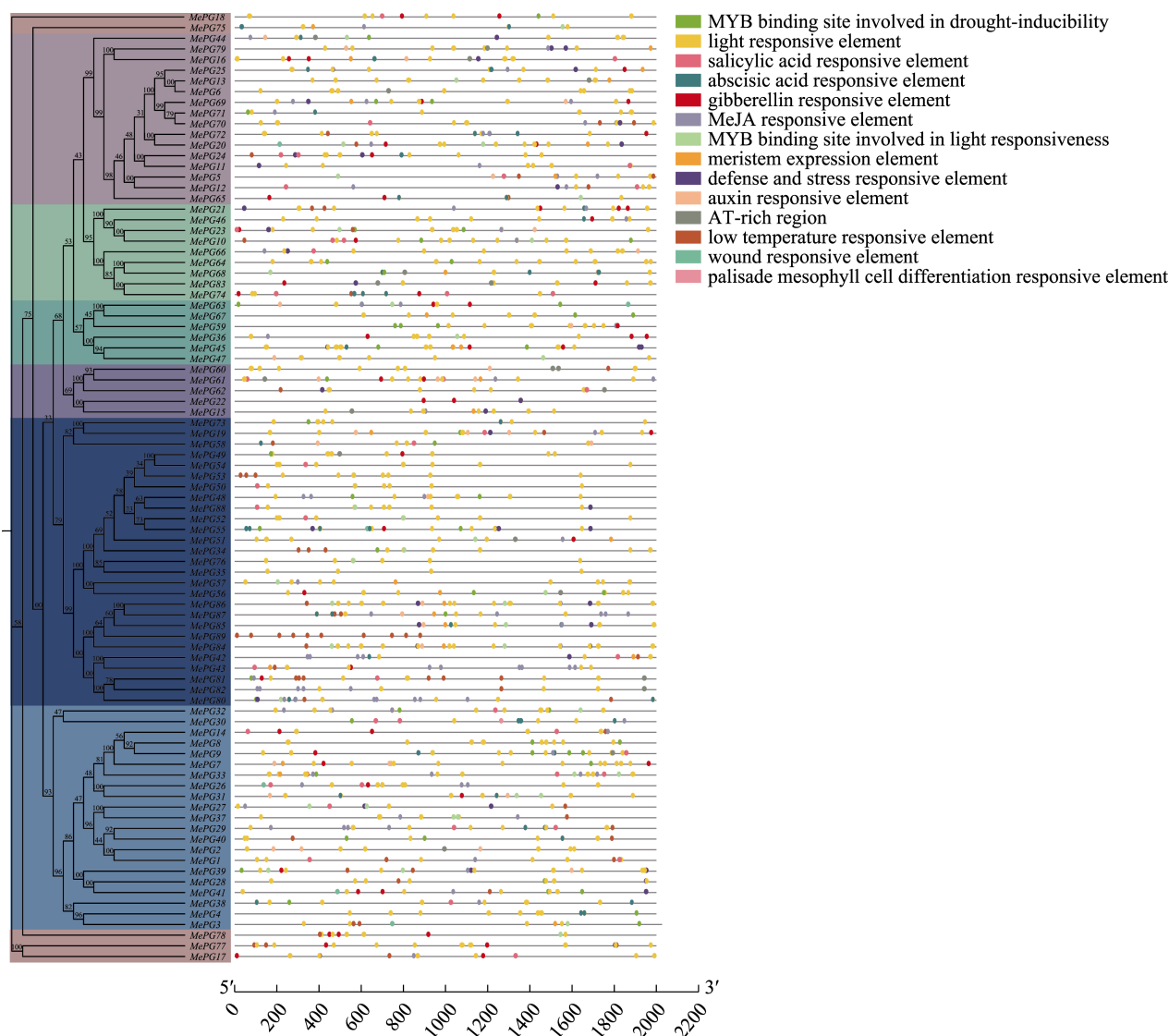


图 6 木薯 PG 家族基因启动子区顺式作用元件分析

Fig. 6 Analysis of Cis-acting elements in promoter region of cassava PG family genes

在 *MePGs* 的启动子区域中非常丰富,表明这些基因可能在木薯逆境响应中发挥重要作用,脱落酸响应元件可能在干旱或盐胁迫条件下调控 *MePG* 基因的表达。部分 *MePG* 基因(如 *MePG3*、*MePG20*、*MePG26* 等)的启动子区域含有损伤响应元件,表明这些基因可能在机械损伤或病原体侵染时被激活,参与细胞壁的修复和防御过程。如赤霉素、生长素等激素响应元件也存在于部分 *MePG* 基因的启动子区域,表明这些基因可能在木薯的生长发育过程中发挥调控作用。

2.7 *MePG* 基因表达模式分析

本研究分析了 69 个 *MePG* 基因在木薯愈伤组织、体胚、嫩叶、成熟叶、茎、须根、块根、块根韧皮部和木质部等 9 个组织器官中的表达模

式,结果如图 7 所示。根据 *MePGs* 的表达谱可将其分为 3 个组: A 组, 21 个基因在大多数组织中显示出高表达水平; B 组, 37 个基因在特定组织中高度转录; C 组, 11 个基因在大多数组织中显示出低表达水平。B 组中的 *MePGs* 基本只在愈伤组织和体胚中表达,而 A 组中的 *MePG20*、*MePG72*、*MePG25*、*MePG16*、*MePG21*、*MePG5* 等基因在块根中表达量较高。此外, *MePG15*、*MePG31* 为愈伤组织特异表达基因, *MePG8*、*MePG9*、*MePG32* 为体胚特异表达基因。

目前研究表明 PG 参与细胞壁结构分解,木薯块根采后变质中细胞壁结构的变化可能与 PG 家族基因的活性密切相关。根据本实验室前期对木薯果胶和 PG 活性的研究,发现同时期 PG 活性

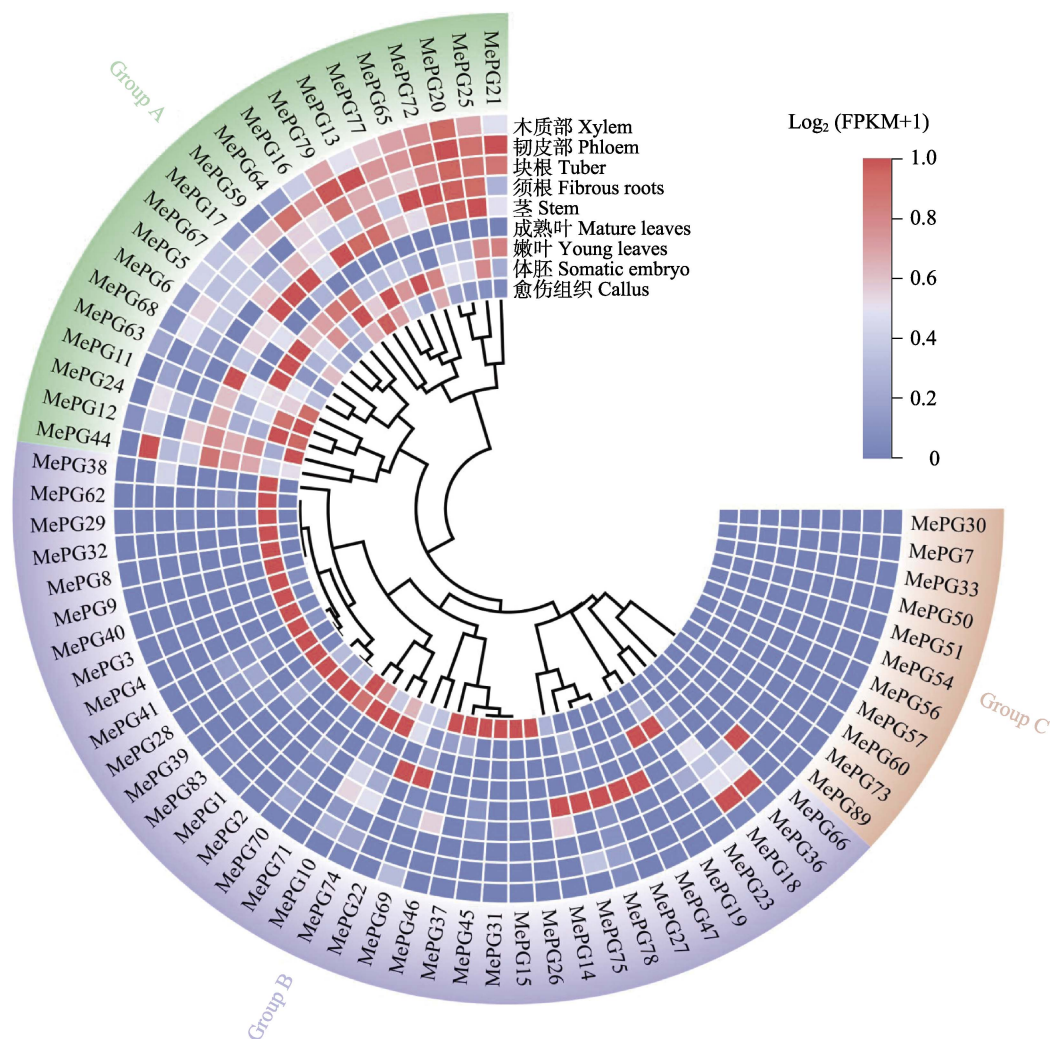


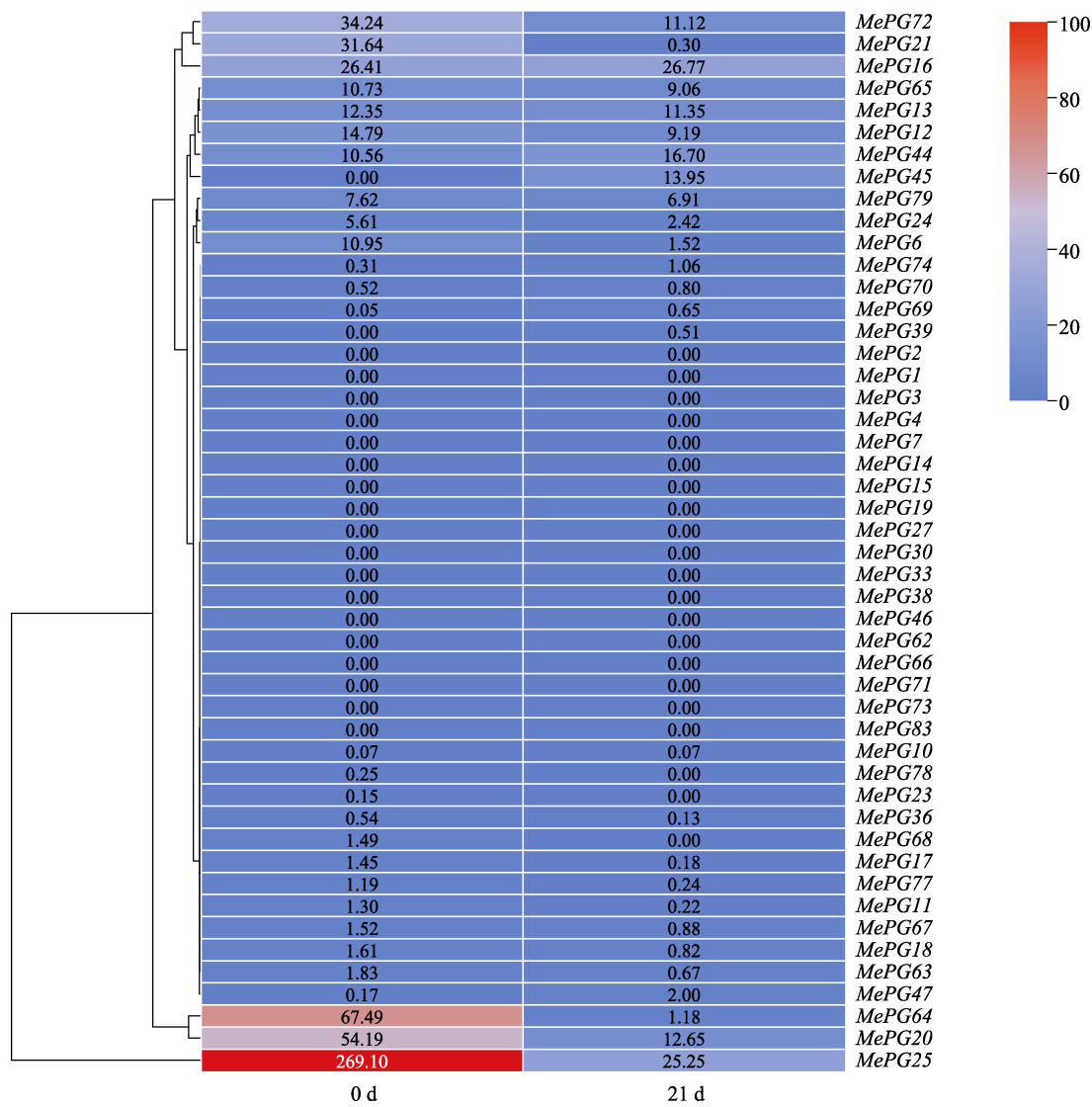
图 7 木薯 PG 基因在不同组织中的表达模式
Fig. 7 Expression patterns of cassava PG genes in different tissues

低的木薯细胞壁可以保留更完整的酯化果胶^[16]。

本研究进一步对 SC8 木薯块根采后变质过程中 *MePGs* 的表达情况进行分析发现：一部分 *MePGs* 在采后变质中几乎不表达；一部分 *MePGs* 的表达量无明显变化，如 *MePG16*、*MePG65*、*MePG13* 和 *MePG12* 等；*MePG20*、*MePG21*、*MePG25*、*MePG64*、*MePG72* 在木薯块根采后变质 21 d 的表达量下调（图 8）。

如图 9 所示，选择 8 个时间段的木薯采后块根进行取样，在每颗块根中部选择大小一致位置作为取样点，切片并迅速研磨提取 RNA。可以发现，SC8 木薯整根块根在采后前 3 d 的薯块颜色、硬度以及韧皮部形态变化不大，3~5 d 开始发生明显褐化、薯块变软等现象。而到 14 d，薯块褐变面积明显增大。到 21 d 时，薯块褐变颜色加深，薯块木质部出现开裂，薯块进一步变软。

进一步验证 *MePG* 在块根采后变质过程中的表达模式。选择 *MePG20*、*MePG21*、*MePG25*、*MePG64*、*MePG72* 进行 qPCR 分析，结果如图 10 所示。在 SC8 变质 21 d 的过程中，*MePG20* 基因除第 1 天有短暂上升外，后面时间段基本呈下调的表达趋势，在第 2 周（7~14 d）的表达量无显著差异，第 21 天的表达量最低；*MePG21* 基因在第 5 天的表达量升高至最高，之后随着变质时间的延长其表达量持续降低；*MePG25* 基因表达量在前 3 d 短暂降低，在 5~7 d 升高，7 d 后其表达量开始随着变质时间降低；*MePG64* 基因表达量随变质时间的延长出现降低-升高-降低-升高-迅速降低的变化；*MePG72* 基因在第 1 天的表达量升至最高，3 d 后降低，然后在第 5 天升高，后续随着变质时间持续降低。5 个 *MePG* 基因的相对表达量变化模式与转录组相符。



图中标注值为 FPKM 原始数据。
The annotated values in the figure are the raw FPKM data.
图 8 SC8 木薯块根在采后 0 d 和 21 d 的 PG 基因表达模式
Fig. 8 PG gene expression patterns in cassava storage roots of SC8 at 0 and 21 days post-harvest

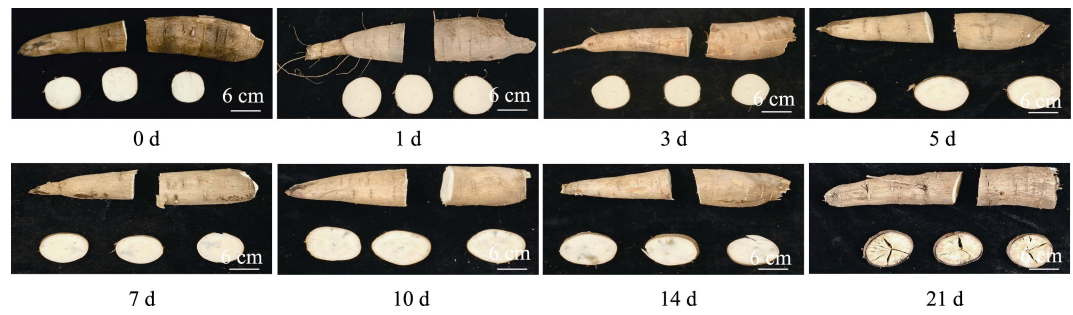


图 9 SC8 木薯块根采后变质 21 d
Fig. 9 SC8 cassava root deteriorates after 21 days of harvesting

3 讨论

作为果胶修饰酶的关键成员，植物 PG 基因家族的功能已受到广泛关注。PG 基因在植物不同发育阶段和组织中的表达模式表明其在细胞壁水解中发挥重要作用。本研究在木薯基因组中鉴定出 89 个 MePG 基因家族成员，这些成员在理化

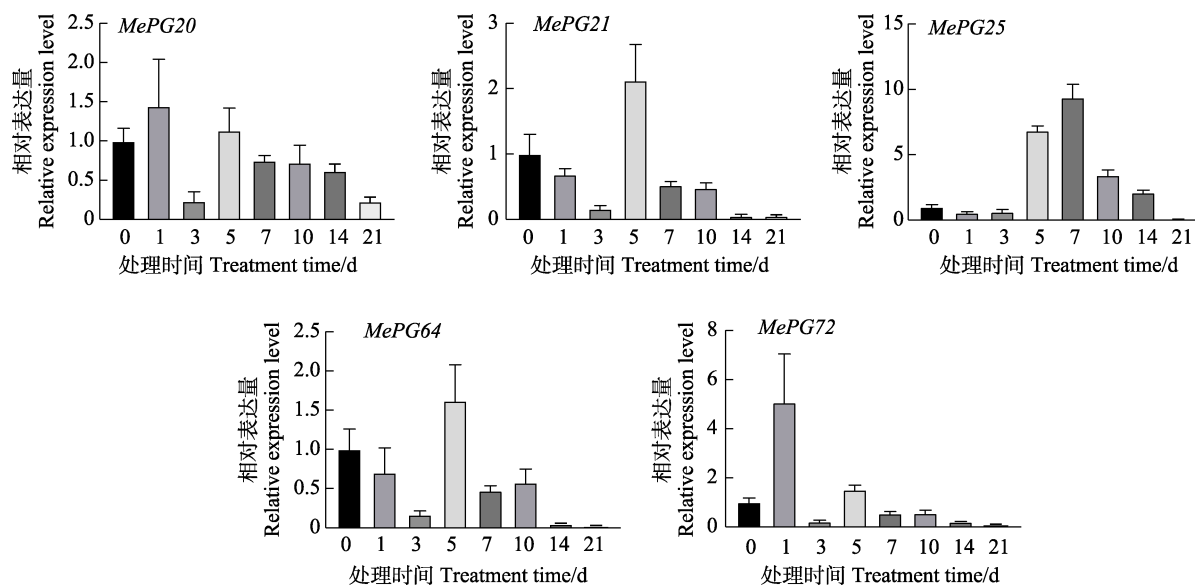


图 10 MePGs 在采后变质过程中的表达分析

Fig. 10 Expression analysis of MePG genes during cassava postharvest deterioration

性质如肽链长度、相对分子量和等电点等方面表现出显著差异，与甘薯 PG 基因家族的研究结果相似^[14]。多种高等植物中均已鉴定出 PG 基因家族成员，如拟南芥（68 个）、榴莲（57 个）^[13]、甘薯（103 个）^[14]、苹果（85 个）^[15]、桃（84 个）^[17]、水稻（44 个）^[18]、甘蓝（99 个）^[19]、小麦（113 个）^[20]、莲雾（25 个）^[21]等。系统进化分析显示，木薯与拟南芥、榴莲、甘薯和桃等的 PG 基因家族均分为 7 个亚族，亚族内基因结构和保守基序相似。木薯中 PG 基因的数量与苹果、桃接近，而与甘薯、水稻、小麦、莲雾等有较大差异，这可能反映了不同植物物种在进化过程中基因复制事件和环境适应性选择的影响。

植物激素如水杨酸、脱落酸、赤霉素、生长素、独脚金内酯、茉莉酸甲酯等在植物逆境响应中起到信号传导作用，介导逆境胁迫下的代谢物质合成^[22]。木薯与榴莲中 PG 基因家族启动子含有相似的激素与逆境响应元件^[13]，推测多数 PG 蛋白能够对非生物应激做出反应，特别是脱落酸和茉莉酸甲酯，作为植物逆境生长和发育的重要调控激素，其顺式作用元件在 MePG 基因家族中非常丰富，表明其可能在调节木薯生长和发育中发挥重要作用，并参与抗病过程^[22]。此外，MePG3、MePG20、MePG26、MePG41、MePG55 和 MePG63 的启动子区域存在损伤响应元件，这可能是研究植物机械损伤和病原体侵染的关键点。本研究关注的几个在木薯采后过程中表达量

显著变化的基因如 MePG20、MePG21 和 MePG25 等，其启动子区域富含光响应元件、茉莉酸甲酯响应元件、脱落酸响应元件、低温响应元件以及损伤响应元件。这些元件的存在表明木薯 PG 基因可能在光调控的生长发育过程中发挥作用，并在多种逆境胁迫条件下被激活，参与细胞壁降解和防御机制。后续研究应该关注这些基因在逆境胁迫条件下的表达调控机制，以及其在木薯块根采后变质中的具体功能。

PG 家族成员在不同组织部位具有不同的表达模式，部分基因在采后变质过程中的表达量呈变化显著，表明这些 PG 基因可能响应其他生长机制的调控。MePG20、MePG21、MePG25、MePG64、MePG72 等基因。在采后变质过程中的表达量先上升后下降，表明 PG 酶活性可能在果实成熟软化的某个阶段达到高峰，随后下降，这与葡萄成熟中的 PG 活性变化趋势相似^[23]。这一过程中涉及到复杂的激素和酶的相互作用，以及细胞壁结构的变化。MePGs 的表达变化说明其在植物细胞壁的降解中扮演重要角色，在木薯采后阶段响应逆境胁迫中起到水解果胶改变细胞壁的作用。

MePG20、MePG21、MePG25、MePG64、MePG72 等基因在采后变质过程中表达量变化显著，表明其可能在木薯块根采后变质中起重要作用。qPCR 结果进一步验证了这些基因的表达模式，表明其在木薯块根细胞壁软化和变质过程中

发挥关键作用。

综上所述,本研究从木薯全基因组中鉴定出 89 个 *MePG* 基因,并分析其理化性质、保守结构域、保守基序、顺式作用元件等和其在不同组织部位及块根采后变质过程的表达模式。本研究初步认定 *MePG* 可能在木薯块根响应采后变质具有重要作用,为进一步研究木薯 *PG* 基因对逆境响应机制奠定基础。

参考文献

- [1] COSGROVE D J. Structure and growth of plant cell walls[J]. *Nature Reviews Molecular Cell Biology*, 2024, 25: 340-358.
- [2] JOBERT F, YADAV S, ROBERT S, ROBERT S. Auxin as an architect of the pectin matrix[J]. *Journal of Experimental Botany*, 2023, 74(22): 6933-6949.
- [3] ATMODJO M A, HAO Z, MOHNEN D. Evolving views of pectin biosynthesis[J]. *Annual Review of Plant Biology*, 2013, 64(1): 747-779.
- [4] MOHNEN D. Pectin structure and biosynthesis[J]. *Current Opinion in Plant Biology*, 2008, 11(3): 266-277.
- [5] LI L, GAO X L, LIU J G, CHITRAKAR B, WANG B, WANG Y C. Hawthorn pectin: extraction, function and utilization[J]. *Current Research in Food Science*, 2021, 4: 429-435.
- [6] LINARES-GARCÍA J A, RAMOS-RAMÍREZ E G, SALAZAR-MONTOYA J A. Viscoelastic properties and textural characterisation of high methoxyl pectin of hawthorn (*Crataegus pubescens*) in a gelling system[J]. *International Journal of Food Science & Technology*, 2015, 50(6): 1484-1493.
- [7] WOLF S, MOUILLE G, PELLOUX J. Homogalacturonan methyl-esterification and plant development[J]. *Molecular Plant*, 2009, 2(5): 851-860.
- [8] 郑秋萍, 林育钊, 李美玲, 王慧, 林河通. 果实采后软化的影响因素及抑制技术研究进展[J]. *亚热带农业研究*, 2019, 15(4): 262-270.
ZHENG Q P, LIN Y Z, LI M L, WANG H, LIN H T. Research progress on the influencing factors and inhibiting technologies of postharvest fruit softening[J]. *Subtropical Agriculture Research*, 2019, 15(4): 262-270. (in Chinese)
- [9] LU L, HOU Q C, WANG L L, WANG L L, ZHANG T Y, ZHAO W, YAN T W, ZHAO L N, LI J P, WAN X Y. Genome-wide identification and characterization of polygalacturonase gene family in maize (*Zea mays* L.)[J]. *International Journal of Molecular Sciences*, 2021, 22(19): 10722.
- [10] 葛廷. 柑橘 *PG* 基因的鉴定及其与幼果脱落的相关性分析[D]. 重庆: 西南大学, 2020.
GE T. Identification of Citrus *PG* genes and their roles in fruitlet abscission[D]. Chongqing: Southwest University, 2020. (in Chinese)
- [11] 张鹏. 能源木薯种质资源面临的问题与解决策略[J]. *生物技术*, 2008(5): 25-30.
ZHANG P. Problems and solutions of energy cassava germplasm resources[J]. *Biotechnology & Business*, 2008(5): 25-30. (in Chinese)
- [12] LIU S, ZAINUDDIN I M, VANDERSCHUREN H, JAMES D, BEECHING J R. RNAi inhibition of feruloyl CoA 6'-hydroxylase reduces scopoletin biosynthesis and post-harvest physiological deterioration in cassava (*Manihot esculenta* Crantz) storage roots[J]. *Plant Molecular Biology*, 2017, 94(1/2): 185-195.
- [13] 刘玲玲. 榴莲 *PG* 基因家族鉴定及其裂果的果胶变化分析[D]. 合肥: 安徽农业大学, 2022.
LIU L L. Identification of *PG* gene family in durian and pectin changes during fruit dehiscence[D]. Hefei: Anhui Agricultural University, 2022. (in Chinese)
- [14] HE P, ZHANG J, LV Z, CUI P, XU X M, GEORGE M S, LU G Q. Genome-wide identification and expression analysis of the polygalacturonase gene family in sweetpotato[J]. *BMC Plant Biology*, 2023, 23(1): 300.
- [15] 陈鸿飞, 邵红霞, 樊胜, 马娟娟, 张东, 韩明玉. 苹果全基因组多聚半乳糖醛酸酶基因家族的鉴定及进化分析[J]. *园艺学报*, 2016, 43(10): 1863-1877.
CHEN H F, SHAO H X, FAN S, MA J J, ZHANG D, HAN M Y. Identification and phylogenetic analysis of the polygalacturonase gene family in apple[J]. *Acta Horticulturae Sinica*, 2016, 43(10): 1863-1877. (in Chinese)
- [16] 周杨骄. 木薯 *MePMEI1* 对植物细胞壁果胶的影响及其响应重金属 Pb、Cd 胁迫的功能研究[D]. 海口: 海南大学, 2023.
ZHOU Y J. Effect of cassava *MePMEI1* on plant cell wall pectin and its function in response to heavy metal Pb and Cd stresses[D]. Haikou: Hainan University, 2023. (in Chinese)
- [17] 霍如雪, 刘振宁, 杨青, 王光全. 桃 *PG* 基因家族的鉴定与分析[J]. *江苏农业科学*, 2016, 44(6): 33-40.
HUO R X, LIU Z N, YANG Q, WANG G Q. Identification and analysis of *PG* gene family in peach[J]. *Jiangsu Agricultural Sciences*, 2016, 44(6): 33-40. (in Chinese)
- [18] 郑银珍. 水稻 *PLL* 和 *PG* 基因家族的鉴定及功能研究[D]. 武汉: 武汉大学, 2018.
ZHENG Y Z. Genomic identification of *PLL* and *PG* gene families and functional analysis in rice[D]. Wuhan: Wuhan University, 2018. (in Chinese)
- [19] LYU M, IFTIKHAR J, GUO R, WU B H, CAO J S. Patterns of expansion and expression divergence of the polygalacturonase gene family in *Brassica oleracea*[J]. *International Journal of Molecular Sciences*, 2020, 21(16): 5706.
- [20] YE J, YANG X, YANG Z, NIU F Q, CHEN Y R, ZHANG L

- L, SONG X Y. Comprehensive analysis of polygalacturonase gene family highlights candidate genes related to pollen development and male fertility in wheat (*Triticum aestivum* L.)[J]. *Planta*, 2020, 252(2): 31-48.
- [21] 陈迪飞, 魏秀清, 许玲, 许家辉, 曾黎辉. 莲雾 PG 基因家族全基因组鉴定及表达分析[J]. *果树学报*, 2022, 39(4): 548-563.
- CHEN D F, WEI X Q, XU L, XU J H, ZENG L H. Genome-wide identification and expression analysis of PG gene family in wax apple [*Syzygium samarangense* (Bl.) Merr. et Perry][J]. *Journal of Fruit Science*, 2022, 39(4): 548-563. (in Chineses)
- [22] CANNON S B, MITRA A, BAUMGARTEN A, YOUNG N D, MAY G. The roles of segmental and tandem gene duplication in the evolution of large gene families in *Arabidopsis thaliana*[J]. *BMC Plant Biology*, 2004, 4(1): 4.
- [23] 海龙飞, 栗温新, 李志谦, 李猛, 陈超阳, 魏红丽, 邹东方, 何畅, 冯建灿, 叶霞. 软/硬肉葡萄果实细胞壁结构、组分及降解酶活性的变化[J]. *果树学报*, 2023, 40(4): 690-698.
- HAI L F, LI W X, LI Z Q, LI M, CHEN C Y, WEI H L, ZOU D F, HE C, FENG J C, YE X. Variations in cell wall microstructure and components and activities of their degradation enzymes in grapes with soft or hard textures[J]. *Journal of Fruit Science*, 2023, 40(4): 690-698. (in Chineses)