

DOI:10.12171/j.1000-1522.20210291

油橄榄III类过氧化物酶家族的进化和表达分析

薛 丽¹ 刘晓霞¹ 王晨合¹ 张建国^{1,2,3} 饶国栋^{1,2,3}

(1. 林木遗传育种国家重点实验室, 中国林业科学研究院林业研究所, 北京 100091; 2. 南京林业大学南方现代林业协同创新中心, 江苏 南京 210037; 3. 国家林业和草原局林木培育重点实验室, 中国林业科学研究院林业研究所, 北京 100091)

摘要:【目的】III类过氧化物酶(PRX)是一类植物特有的氧化还原酶,在植物生长发育和胁迫响应方面具有十分重要的作用。本研究拟探讨油橄榄III类 PRX 基因家族的进化和表达模式,旨在为油橄榄分子育种提供参考。【方法】利用生物信息学方法鉴定并分析了油橄榄III类 PRX 基因家族的系统发育关系、分组、染色体分布、复制基因、基序分布、基因结构图、顺式作用元件分布和在不同组织和不同非生物胁迫下的表达,并对部分成员基因进行了实时荧光定量 PCR 验证。【结果】(1)鉴定得到了 106 个 *OePRX* 基因,根据其拟南芥和毛果杨 PRX 蛋白序列的系统进化关系分为了 14 个组。(2)*OePRX* 基因不均匀地分布于 23 条染色体上;片段复制是该基因家族扩张的主要动力,且复制基因在进化过程中受到了较强的纯化选择;与拟南芥相比,油橄榄 PRX 基因与毛果杨 PRX 基因的亲缘关系更近。(3)同一组中 *OePRX* 蛋白的等电点、分子量、基序分布、基因结构、信号肽分布、在不同组织中的表达模式均比较保守;*OePRX* 基因启动子中含有较多的生长发育元件和激素应答元件;在油橄榄干旱胁迫和水涝胁迫下,38% 的 *OePRX* 基因显著差异表达。【结论】油橄榄 *OePRX* 基因家族明显的进化扩张和多变的表达模式暗示了其功能的复杂性,尤其是复制基因的新功能化对油橄榄在地中海地区的广泛分布具有十分重要的适应性意义。

关键词: 油橄榄; III类 PRXs; 复制基因; 新功能; 非生物胁迫

中图分类号: S718.46; S565.7 **文献标志码:** A **文章编号:** 1000-1522(2023)04-0036-14

引文格式: 薛丽, 刘晓霞, 王晨合, 等. 油橄榄III类过氧化物酶家族的进化和表达分析 [J]. 北京林业大学学报, 2023, 45(4): 36-49. Xue Li, Liu Xiaoxia, Wang Chenhe, et al. Evolution and expression analysis of the class III peroxidase family in olive[J]. Journal of Beijing Forestry University, 2023, 45(4): 36-49.

Evolution and expression analysis of the class III peroxidase family in olive

Xue Li¹ Liu Xiaoxia¹ Wang Chenhe¹ Zhang Jianguo^{1,2,3} Rao Guodong^{1,2,3}

(1. State Key Laboratory of Tree Genetics and Breeding, Research Institute of Forestry,
Chinese Academy of Forestry, Beijing 100091, China;

2. Collaborative Innovation Center of Sustainable Forestry in Southern China,
Nanjing Forestry University, Nanjing 210037, Jiangsu, China;

3. Key Laboratory of Tree Breeding and Cultivation, National Forestry and Grassland Administration,
Research Institute of Forestry, Chinese Academy of Forestry, Beijing 100091, China)

Abstract: [Objective] The class III peroxidase is a group of plant-specific oxidoreductases, and plays an important role in plant development and stress response. This paper presents a detailed overview of evolution and expression of olive class III PRX gene family, aiming to provide a reference for olive molecular breeding in the future. [Method] With some bioinformatic tools, we finished the identification of the olive class III PRX genes, completed the analyses of phylogenetic relationship, gene mapping on the

收稿日期: 2021-08-03 修回日期: 2021-09-17

基金项目: 中央级公益性科研院所专项(CAFYBB2021QC001)。

第一作者: 薛丽。主要研究方向: 林木遗传育种。Email: xueli8762@163.com 地址: 100091 北京市海淀区香山南路东小府 1 号中国林科院林业所。

责任作者: 饶国栋, 博士, 研究员。主要研究方向: 林木遗传育种。Email: rgd@caf.ac.cn 地址: 同上。

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

chromosomes, pairs of duplicates, motifs, gene structure, *cis*-acting elements and gene expression in different tissues or biotic stresses, and did a verification of RNA-seq by RT-qPCR. [Result] (1) 106 *OePRX* genes were obtained, and classified into 14 groups based on the phylogeny with *AtPRX* and *PtPRX*. (2) *OePRX* genes were unevenly located on 23 chromosomes, and segment replication was the main driving force for the expansion of gene family. Compared with *AtPRX*, *OePRX* had a closer relationship with *PtPRX*. (3) The characterization of pI, MW, motifs, gene structure, signal peptide, and expression in different tissues had the expected group-conserved patterns. The promoters of *OePRX* contained a variety of development and hormone elements; 38% of *OePRX* genes was differentially expressed in heat, drought and waterlogged stress. [Conclusion] The differential expansion and differential expression patterns may imply flexibility in neofunctionalization of duplicated class III peroxidase genes, which is of adaptive significance to the strong resistance of olive to a diversity of conditions, hence contributing to the importance of olive as a Mediterranean Basin staple food.

Key words: *Olea europaea*; class III peroxidase; replicator; new function; abiotic stress

油橄榄(*Olea europaea*)主要分布于地中海地区,是一种重要的木本油料作物。橄榄油中富含健康的不饱和脂肪酸、酚类化合物和维生素E等营养物质,不仅是地中海式饮食的重要成分,而且远销美国、墨西哥、巴西、印度、澳大利亚等国家^[1]。近年来,全球橄榄油需求量增加,产量却十分有限。要持续扩大油橄榄的适生区,增加橄榄油的产量,种质选育创制尤为关键。随着测序技术的飞速发展,油橄榄已完成了高质量基因组文件的组装和注释^[2],并获得了丰富的转录组数据,这为我们借助分子育种选育抗逆性强、产油率高的油橄榄奠定了基础。

过氧化物酶(peroxidase, PRX)是一类广泛存在于生物体内的酶,它通过将过氧化氢还原为水来催化底物的氧化。根据其蛋白序列和结构特征,PRX可分为2类,血红素类和非血红素类;血红素PRX又可进一步分为动物型和非动物型;非动物的PRX包含3个小类,I、II和III类,这3小类PRX具有非常相似的三维结构,但氨基酸序列同源性较低,且具有不同的功能和反应机制^[3]。I类PRX广泛分布于非动物性生命体内,其在原核生物中的分布表明其他2类PRX可能起源于I类PRXs^[4],I类PRX包含抗坏血酸PRX、细胞色素c PRX和过氧化氢酶,主要功能是消除过量的H₂O₂^[5-6]。II类PRX仅分布于真菌中,包含锰PRX、木质素PRX和其他PRX^[7],主要功能是降解木质素。III类PRX分布于陆生植物体内,是一个庞大的多基因家族,通常分泌进入细胞壁或其周围基质或液泡中。III类PRX广泛参与植物生命周期内的一系列的生理活动^[8],比如细胞壁形成^[9]、伤口愈合^[10]、生长素代谢^[11-12]等。另外,III类PRX通过调节多种活性氧水平响应生物胁迫和非生物胁迫^[12],因此,研究III类PRX基因功能对作物遗传育种具有十分重要的

作用。

目前,拟南芥(*Arabidopsis thaliana*)^[13]、水稻(*Oryza sativa*)^[14]、玉米(*Zea mays*)^[15]、毛果杨(*Populus trichocarpa*)^[16]、白梨(*Pyrus bretschneideri*)^[17]、葡萄(*Vitis vinifera*)^[18]、木薯(*Manihot esculenta*)^[19]和小麦(*Triticum aestivum*)^[20]等植物的III类PRX基因全基因组鉴定和分析已经完成,这些物种中有47~138个不等的III类PRX基因。单个III类PRX基因功能的研究也日益丰富,例如,*AtPRX4*、*AtPRX52*和*AtPRX72*拟南芥突变体木质素含量降低,木质素组成发生改变^[21-23]。烟草(*Nicotiana tabacum*)中过表达*AtPRX64*促进根系生长,减少活性氧和铝在根系中的积累,从而提高对铝胁迫的耐受性^[24];拟南芥中过表达*OsPRX38*促进胼胝质沉积、抗氧化酶活性和叶绿素含量上升,从而提高植物对砷胁迫的耐受性^[25]。然而,关于油橄榄III类PRX基因家族的全基因组鉴定及基因功能研究仍未见报道。本研究利用最近组装注释的油橄榄基因组和公共数据库中的转录组信息,完成了油橄榄III类PRX家族的鉴定、进化和表达分析。研究发现:油橄榄是目前已知III类PRX家族成员最多的双子叶植物,其复制基因的亚功能化和新功能化对油橄榄在地中海地区的广泛分布具有十分重要的意义。

1 材料与方法

1.1 油橄榄III类PRX基因家族成员的鉴定

油橄榄基因组文件下载于NGDC(<https://bigd.big.ac.cn/search/?dbId=gwh&q=Olea%20europaea&page=1>),III类PRX基因家族的保守结构域(PF00141)的HMM模型下载于pfam(<http://pfam.xfam.org/family/PF00141#tabview=tab6>)。利用PF00141的HMM模型,在油橄榄蛋白质注释文件(GWHAOPM000000000.

Protein.faa)中进行 HMM 模型搜索(hmmsearch),得到 120 条蛋白序列。利用拟南芥的 73 条 PRX 蛋白序列^[13]在油橄榄基因组文件(GWHAOPM000000000.Protein.faa)中进行本地化搜索(blastp), *E* 值设置为 10^{-6} , 得到 119 条序列。2 种方法得到的共有序列为 119 条。将 119 条序列上传至 Pfam(<http://pfam.xfam.org/search#tabview=tab1>)和 NCBI-CDD(<https://www.ncbi.nlm.nih.gov/Structure/bwrpsb/bwrpsb.cgi>)数据库中, 预测其是否含有 III 类 PRX 基因家族的保守结构域从而进行筛选。Pfam 鉴定结果表明, 119 条序列均含有 PF00141 结构域; NCBI-CDD 鉴定结果表明, 13 条序列含有不完整的分泌类过氧化物酶保守结构域(secretory_peroxidase, cd00693)。因此, 最后选定 106 条序列作为油橄榄 III 类 PRX 基因成员。利用 Prot-Param(https://web.expasy.org/compute_pi/)预测等电点和分子量; 利用 SignalP-5.0 Server(<http://www.cbs.dtu.dk/services/SignalP/>)预测信号肽 Sec/SPI 的存在概率及长度, Signal peptide likelihood 值大于 0.5 的序列视为含有信号肽, 反之认为不含有信号肽; 利用 Wolfpsort(<https://wolfpsort.hgc.jp/>)预测亚细胞定位。

1.2 PRX 蛋白系统进化树的构建

利用 TBtools v1.095 中的 Fasta Extract 从油橄榄基因组文件中提取得到 106 个 *OePRX* 的蛋白序列, 从前人的论文中下载得到拟南芥和毛果杨 PRX 蛋白序列^[16], 其中, 拟南芥 73 条, 毛果杨 93 条。利用 MEGA7.0 中的 ClustalW 比对 272 条全长的 PRX 蛋白序列, 并使用邻接法构建系统进化树, 设置替代模型为 Poisson model, 设置缺失数据处理方式为 Pairwise deletion, 设置 bootstrap 值为 1 000。

1.3 *OePRX* 染色体定位和复制基因分析

利用 TBtools v1.095 中的 Fasta Stats 从油橄榄基因组文件中获取 23 条染色体的长度和 *OePRX* 基因家族成员的位置信息, 利用 MapChart 绘制 *OePRX* 基因家族在 23 条染色体上的分布, 需要注意的是, *OePRX* 基因家族成员中 100 个基因分布于染色体上, 但 6 个基因分布于 Contig 上, 这 6 个基因未在图中展示。利用 TBtools v1.095 中的 One Step MCScanX 分析 *OePRX* 基因家族中的复制基因; 利用 MEGA7.0 中的 Muscle (Codons) 比对复制基因对的编码区(CDS), 再利用 KaKs_Calculator2.0 计算复制基因对的非同义替换(Ka)和同义替换(Ks)值, 设置计算方式为 YN。拟南芥基因组文件下载自 http://ensembl.gramene.org/Arabidopsis_thaliana/Info/Index, 毛果杨基因组文件下载自 <https://phytozome.jgi.doe.gov>。利用同样的方法分析油橄榄和拟南芥及毛果杨 PRX 之

间的复制基因, 并计算 Ka 和 Ks 值。利用 TBtools v1.095 中的 Advanced Circos 绘制油橄榄 PRX 基因线性图; 用 Multiple Synteny Plot 绘制油橄榄与拟南芥和毛果杨 PRX 基因的线性图。

1.4 *OePRX* 基因序列分析

利用 TBtools v1.095 中的 Fasta Extract 从油橄榄基因组文件提取 *OePRX* 基因家族成员的 DNA 序列和 CDS 序列, 再利用 GSDS2.0(<http://gsds.gao-lab.org/>)绘制其基因结构图。利用在线网站 MEME(<https://meme-suite.org>)预测 *OePRX* 基因的保守基序, 预测基序的个数设置为 10, 其他参数默认。利用 TBtools v1.095 中的 Fasta Extract 从油橄榄基因组文件中获取 106 个 *OePRX* 基因上游 1 500 bp 的序列作为启动子, 并将其提交至 PlantCARE(<http://bioinformatics.psb.ugent.be>)预测启动子上的顺式作用元件, 最后对结果进行分析整理。

1.5 *OePRX* 基因表达模式分析

利用 NCBI(<https://www.ncbi.nlm.nih.gov/>)中油橄榄不同组织^[26](根、茎、叶、花、果和分生组织)、油橄榄枝条响应干旱胁迫及复水^[27]和油橄榄叶片响应水涝胁迫的 RNA-seq 数据(未发表)分析 *OePRX* 基因家族的表达模式。使用 sra explorer 获取测序数据的 FTP 地址, 使用 fastp 去除接头和低质量的 reads, 利用 hisat2 比对并使用 htseq-count 获取基因的原始的读长数目(reads counts), 使用 DESeq2 进行差异表达分析。自行编写 R 脚本将原始读长数目转换为 tpm(每千个碱基的转录每百万映射读取的转录本, 均一化后的 read counts 值)。差异基因鉴定标准为 $FDR \leq 0.05$, 并且 $|\log_2(\text{Fold Change})| \geq 2$ (Fold Change 为差异倍数)。取 $\log_2(\text{tpm} + 1)$, 利用 R 语言绘制热图。

1.6 实时荧光定量 PCR

油橄榄品种“科罗莱卡”种植于中国林科院林木遗传育种国家重点实验室温室, 选择 3 株长势均一的 3 年生油橄榄幼苗, 剪取叶片、茎、根和果实, 于液氮中带回实验室后冻存于 $-80\text{ }^{\circ}\text{C}$ 冰箱内。利用 MiniBEST 植物 RNA 提取试剂盒(TaKaRa, 9769)提取不同组织的总 RNA; 利用 PrimeScript™ RT Master Mix(Takara, RR036A)反转录合成 cDNA; 利用 Primer Premier5.0 设计引物; 利用 TB Green® Premix Ex Taq™ II (Tli RNaseH Plus)(Takara, RR820A)在 Roche Light Cycler 480 上进行实时荧光定量 PCR(qRT-PCR), 反应体系如下: 总体系 20 μL , 包括 TB Green Premix Ex Taq II (Tli RNaseH Plus) (2 $\times$)10 μL , cDNA 2 μL , 正、反向引物各 0.8 μL , 去离子水 6.4 μL 。反应程序为 2 步法, 详见试剂盒说明

书。以 *Actin*^[28]为内参基因,利用 $2^{-\Delta\Delta CT}$ 法计算基因相对表达量。

2 结果与分析

2.1 *OePRX* 基因家族成员的鉴定

血红素过氧化物酶(haem peroxidase)的保守结构域是 PF00141, 这类酶大都含有 2 个 α 结构域, 中间嵌有血红素基团^[29]。利用 PF00141 的 HMM 模型和 73 条 *AtPRX* 蛋白在油橄榄蛋白质序列文件中分别进行 hmmsearch 和 blastp, 将 2 种方法得到的结果取交集, 得到了 119 条候选 PRX。植物III类过氧化物酶特有的结构域是分泌类过氧化物酶保守结构域(secretory peroxidase, cd00693), 这类酶含有 4 个保守的二硫桥和 2 个保守的钙离子结合位点^[30]。根据 NCBI-CDD 预测结果, 删除分泌类过氧化物酶保守结构域不完整的 13 条序列, 最后得到了 106 条油橄

榄III类 PRX 蛋白序列(表 1)。根据其在染色体上的分布情况, 将其命名为 *OePRX1-OePRX106*。*OePRX* 蛋白的氨基酸个数最小为 240, 最大为 509; 分子量最小为 26.55 kDa, 最大为 58.00 kDa; 等电点变化范围为 4.45 ~ 9.75, 其中 74 个 *OePRX* 蛋白(69.8%)大于等于 7。通过在线预测发现: 75 个 *OePRX* 蛋白(70.8%)含有 Sec/SPI 信号肽, 信号肽长度变化范围为 19 ~ 32 个氨基酸; 值得注意的是, 组 14 中的 10 个 *OePRX* 蛋白均不含有信号肽。亚细胞定位预测结果显示: 含有信号肽的 *OePRX* 蛋白大多位于叶绿体、质膜、液泡膜和内质网(50 个 *OePRX* 蛋白, 占含有信号肽的 *OePRX* 蛋白总数的 66.7%), 也有一部分位于细胞外和细胞质(25 个 *OeORXs* 蛋白, 占含有信号肽的 *OePRX* 总数的 33.3%); 不含信号肽的 *OePRX* 蛋白(31 个 *OePRX* 蛋白)中, 11 个 *OePRX* 蛋白位于细胞外和细胞质, 其余位于叶绿体、内质

表 1 油橄榄III类 PRX 基因的命名及其理化性质

Tab. 1 Nomenclature and physicochemical properties of class III PRX genes from olive

名称 Name	ID	内含子相位 Intron phase	蛋白质长度 Protein length/aa	分子量 Molecular mass/kDa	等电点 Isoelectric point	信号肽 Signal peptide	预测的亚细胞定位 Predicted subcellular localization
OePRX1	EVM0022477	0	344	38.40	6.10	无 No	胞外 Extracellular
OePRX2	EVM0032926	0	266	29.38	8.97	无 No	细胞质 Cytosol
OePRX3	EVM0020804	0	327	35.65	9.21	有 Yes	叶绿体 Chloroplast
OePRX4	EVM0047797	100	327	35.92	8.24	无 No	细胞核 Nucleus
OePRX5	EVM0009292	100	316	34.40	8.72	有 Yes	液泡膜 Vacular membrane
OePRX6	EVM0060238	0	330	35.97	8.89	无 No	叶绿体 Chloroplast
OePRX7	EVM0022620	12001000	276	30.39	8.66	无 No	细胞质 Cytosol
OePRX8	EVM0017319	1200201	279	30.58	7.70	无 No	细胞质 Cytosol
OePRX9	EVM0047458	20010	250	27.43	5.67	无 No	细胞质 Cytosol
OePRX10	EVM0057886	0	319	34.48	6.98	有 Yes	叶绿体 Chloroplast
OePRX11	EVM0047784	0	339	37.00	7.48	有 Yes	质膜 Plasma membrane
OePRX12	EVM0039807	20	358	38.84	7.51	有 Yes	叶绿体 Chloroplast
OePRX13	EVM0011450	0	317	34.35	9.14	有 Yes	叶绿体 Chloroplast
OePRX14	EVM0019223	1	309	35.27	9.41	有 Yes	胞外 Extracellular
OePRX15	EVM0012728	1	320	36.91	9.20	有 Yes	细胞质 Cytosol
OePRX16	EVM0000231	0	336	37.28	7.08	有 Yes	叶绿体 Chloroplast
OePRX17	EVM0024793	0	322	35.21	8.38	有 Yes	叶绿体 Chloroplast
OePRX18	EVM0042519	0	321	34.96	8.76	有 Yes	叶绿体 Chloroplast
OePRX19	EVM0059982	0	321	34.90	8.77	有 Yes	叶绿体 Chloroplast
OePRX20	EVM0019685	0	333	36.73	5.68	无 No	质膜 Plasma membrane
OePRX21	EVM0058087	0	324	35.74	8.53	有 Yes	细胞质 Cytosol
OePRX22	EVM0027348	0	323	35.02	5.82	有 Yes	液液泡膜 Vacular membrane
OePRX23	EVM0060830	0	331	36.34	9.41	有 Yes	胞外 Extracellular
OePRX24	EVM0015621	0	247	27.39	8.88	无 No	细胞核 Nucleus
OePRX25	EVM0010445	12101221	327	35.56	9.26	无 No	叶绿体 Chloroplast

表 1(续)

Tab.1 (continued)

名称 Name	ID	内含子相位 Intron phase	蛋白质长度 Protein length/aa	分子量 Molecular mass/kDa	等电点 Isoelectric point	信号肽 Signal peptide	预测的亚细胞定位 Predicted subcellular localization
OePRX26	EVM0006524	0	338	37.53	6.45	有 Yes	胞外 Extracellular
OePRX27	EVM0042768	22	304	32.63	6.58	有 Yes	叶绿体 Chloroplast
OePRX28	EVM0048359	0	355	37.59	8.35	有 Yes	叶绿体 Chloroplast
OePRX29	EVM0028664	0	317	34.49	9.09	有 Yes	液泡膜 Vacular membrane
OePRX30	EVM0035621	0	327	34.85	5.51	有 Yes	胞外 Extracellular
OePRX31	EVM0048816	100	324	36.67	6.26	有 Yes	胞外 Extracellular
OePRX32	EVM0003637	1002	350	39.07	8.83	有 Yes	胞外 Extracellular
OePRX33	EVM0019330	0	333	36.44	9.48	有 Yes	液泡膜 Vacular membrane
OePRX34	EVM0033281	2100102	273	30.15	7.73	无 No	细胞质 Cytosol
OePRX35	EVM0015060	10100000	319	35.73	6.63	无 No	叶绿体 Chloroplast
OePRX36	EVM0017946	0	329	37.29	8.48	有 Yes	细胞质 Cytosol
OePRX37	EVM0021366	—	331	36.79	9.34	有 Yes	叶绿体 Chloroplast
OePRX38	EVM0060527	0	342	37.64	6.21	有 Yes	细胞质 Cytosol
OePRX39	EVM0014494	0	321	35.37	6.18	有 Yes	胞外 Extracellular
OePRX40	EVM0042196	0	319	34.12	9.41	有 Yes	叶绿体 Chloroplast
OePRX41	EVM0015681	0	346	38.94	8.86	有 Yes	内质网 Endoplasmic reticulum
OePRX42	EVM0047610	0	327	35.27	8.80	有 Yes	叶绿体 Chloroplast
OePRX43	EVM0047452	20002	379	41.95	8.87	无 No	叶绿体 Chloroplast
OePRX44	EVM0061929	0	319	34.23	4.45	有 Yes	胞外 Extracellular
OePRX45	EVM0054999	0	291	32.27	5.44	无 No	胞外 Extracellular
OePRX46	EVM0048606	0	328	35.67	6.51	有 Yes	液泡膜 Vacular membrane
OePRX47	EVM0037058	20	315	34.41	7.55	无 No	内质网 Endoplasmic reticulum
OePRX48	EVM0022847	20	413	45.85	9.56	无 No	叶绿体 Chloroplast
OePRX49	EVM0020385	20	373	41.16	9.35	无 No	叶绿体 Chloroplast
OePRX50	EVM0045080	0	317	34.55	9.02	有 Yes	叶绿体 Chloroplast
OePRX51	EVM0044622	0	320	35.53	9.54	有 Yes	叶绿体 Chloroplast
OePRX52	EVM0021425	0	324	36.14	6.17	有 Yes	胞外 Extracellular
OePRX53	EVM0045494	0	327	35.52	8.71	有 Yes	液泡膜 Vacular membrane
OePRX54	EVM0020915	20	306	33.46	8.90	有 Yes	液泡膜 Vacular membrane
OePRX55	EVM0031491	0	314	34.07	9.32	有 Yes	叶绿体 Chloroplast
OePRX56	EVM0005944	2	292	31.79	9.45	有 Yes	叶绿体 Chloroplast
OePRX57	EVM0015088	0	317	34.42	9.13	有 Yes	胞外 Extracellular
OePRX58	EVM0050632	0	314	34.04	9.25	有 Yes	叶绿体 Chloroplast
OePRX59	EVM0028088	0	317	34.23	8.99	有 Yes	胞外 Extracellular
OePRX60	EVM0025336	0	326	34.42	7.58	无 No	叶绿体 Chloroplast
OePRX61	EVM0012469	0	317	34.68	9.33	有 Yes	胞外 Extracellular
OePRX62	EVM0049574	0	340	37.08	6.25	有 Yes	胞外 Extracellular
OePRX63	EVM0050273	10	240	26.55	6.22	无 No	叶绿体 Chloroplast
OePRX64	EVM0050817	0	310	34.15	8.46	有 Yes	叶绿体 Chloroplast
OePRX65	EVM0009492	0	322	34.85	4.99	有 Yes	叶绿体 Chloroplast
OePRX66	EVM0047271	1200201	333	36.96	8.64	无 No	质膜 Plasma membrane

表 1(续)
Tab.1 (continued)

名称 Name	ID	内含子相位 Intron phase	蛋白质长度 Protein length/aa	分子量 Molecular mass/kDa	等电点 Isoelectric point	信号肽 Signal peptide	预测的亚细胞定位 Predicted subcellular localization
OePRX67	EVM0005889	210122	509	58.00	8.73	无 No	细胞核 Nucleus
OePRX68	EVM0026529	0	331	36.43	9.50	有 Yes	液泡膜 Vacular membrane
OePRX69	EVM0038965	20	310	33.90	9.23	有 Yes	质膜 Plasma membrane
OePRX70	EVM0024841	0	328	35.72	8.37	有 Yes	叶绿体 Chloroplast
OePRX71	EVM0047959	3	314	34.07	6.94	有 Yes	叶绿体 Chloroplast
OePRX72	EVM0059042	0	339	37.43	5.56	有 Yes	质膜 Plasma membrane
OePRX73	EVM0022158	0	318	34.18	6.88	有 Yes	叶绿体 Chloroplast
OePRX74	EVM0027175	0	326	36.14	9.75	有 Yes	叶绿体 Chloroplast
OePRX75	EVM0057531	20	287	30.73	8.89	有 Yes	液泡膜 Vacular membrane
OePRX76	EVM0032201	0	293	32.34	8.75	无 No	叶绿体 Chloroplast
OePRX77	EVM0059919	0	339	37.61	8.36	有 Yes	液泡膜 Vacular membrane
OePRX78	EVM0011789	0	355	39.51	8.06	无 No	质膜 Plasma membrane
OePRX79	EVM0024970	—	329	36.54	9.19	有 Yes	叶绿体 Chloroplast
OePRX80	EVM0047604	100	339	37.82	8.30	无 No	细胞质 Cytosol
OePRX81	EVM0061300	0	330	36.94	5.94	有 Yes	叶绿体 Chloroplast
OePRX82	EVM0042525	0	298	32.64	8.62	有 Yes	胞外 Extracellular
OePRX83	EVM0045066	0	318	34.03	9.15	有 Yes	叶绿体 Chloroplast
OePRX84	EVM0029808	0	354	39.38	5.68	无 No	胞外 Extracellular
OePRX85	EVM0032932	1022010	247	27.30	5.65	无 No	细胞质 Cytosol
OePRX86	EVM0038819	2001020	250	27.72	5.44	无 No	细胞质 Cytosol
OePRX87	EVM0002745	200	321	34.58	4.77	无 No	细胞核 Nucleus
OePRX88	EVM0038873	200	326	34.71	8.30	有 Yes	液泡膜 Vacular membrane
OePRX89	EVM0026260	2	348	38.22	9.56	有 Yes	质膜 Plasma membrane
OePRX90	EVM0044542	0	296	32.61	8.14	有 Yes	胞外 Extracellular
OePRX91	EVM0046395	0	318	34.24	9.54	有 Yes	叶绿体 Chloroplast
OePRX92	EVM0042595	0	316	33.95	7.51	有 Yes	叶绿体 Chloroplast
OePRX93	EVM0047351	0	320	35.08	5.48	有 Yes	叶绿体 Chloroplast
OePRX94	EVM0045920	0	332	35.75	4.63	有 Yes	胞外 Extracellular
OePRX95	EVM0044144	0	325	35.63	8.75	无 No	叶绿体 Chloroplast
OePRX96	EVM0048820	0	314	35.88	8.61	有 Yes	细胞质 Cytosol
OePRX97	EVM0058650	0	350	38.67	6.94	有 Yes	叶绿体 Chloroplast
OePRX98	EVM0026875	0	371	40.68	6.31	无 No	叶绿体 Chloroplast
OePRX99	EVM0019820	0	340	36.88	6.59	有 Yes	胞外 Extracellular
OePRX100	EVM0018143	20020200000	394	43.59	9.28	无 No	叶绿体 Chloroplast
OePRX101	EVM0029423	0	343	37.01	8.94	有 Yes	叶绿体 Chloroplast
OePRX102	EVM0036633	0	342	37.59	9.17	有 Yes	胞外 Extracellular
OePRX103	EVM0041887	1	312	33.33	5.97	有 Yes	叶绿体 Chloroplast
OePRX104	EVM0032297	0	327	35.71	8.57	有 Yes	胞外 Extracellular
OePRX105	EVM0042746	0	302	33.53	8.96	有 Yes	胞外 Extracellular
OePRX106	EVM0051394	0	343	37.01	8.94	有 Yes	叶绿体 Chloroplast

网、细胞核和质膜。

2.2 III类 PRX 基因的进化扩张

III类 PRX 基因是一类植物特有的 PRX 基因, 莱茵衣藻(*Chlamydomonas reinhardtii*)、小立碗藓(*Physcomitrella patens*)、江南卷柏(*Selaginella moellendorffii*)^[31]、拟南芥^[13]、毛果杨^[16]、玉米^[15]和水稻^[14]中分别有 6、31、62、73、93、119 和 138 个 III 类 PRX 基因, 说明 III 类 PRX 基因家族在进化过程中不断发生扩张, 这可能与 III 类 PRX 基因参与了植物广泛的生理过程有关。如图 1, 利用邻接法构建了包含 73 个 AtPRX、93 个 PtPRX 和 106 个 OePRX 蛋白的系统进化树, 然后用同样的方法构建了仅包含 106 个 OePRX 蛋白的系统进化树, 根据 3 种植物 PRX 蛋白的系统进化关系和 106 个 OePRX 蛋白之间的系统进化关系, 将 III 类 PRX 基因家族分为 14 个组。组 1、2、4 聚为 1 支, 组 6、7 聚为 1 支, 组 9 单独聚为 1 支, OePRX 和 PtPRX 基因在这 3 支中均发生了明显的扩张; 组 3、10 聚为 1 支, 组 12 单独 1 支, OePRX 和 PtPRX 基因在这 2 支中均发生了轻微的收缩。另外, OePRX 基因在组 13 中发生了扩张, 而 PtPRX 基因发生了收缩; PtPRX 基因在组 5、8、11 中发生了扩张, OePRX 基因在组 5、11 中发生了收缩。另外发现: 组 1~13 中, OePRX、PtPRX、AtPRX 基因均有分布; 组 14 中仅包含 10 个 OePRX 基因, 是一个油橄榄特有的组; 这说明组 1~13 中, 油橄榄、毛果杨、拟南芥有共同的 III 类 PRX 基因的祖先基因, 组 14 中的 OePRX 基因发生了明显的遗传分化。

2.3 OePRX 基因的染色体定位和复制基因分析

2.3.1 OePRX 基因在染色体上的不均匀分布

为了分析 OePRX 基因家族在进化过程中的扩张, 绘制了 106 个 OePRX 基因在 23 条染色体上的分布图(图 2)。结果发现: 100 个 OePRX 基因不均匀

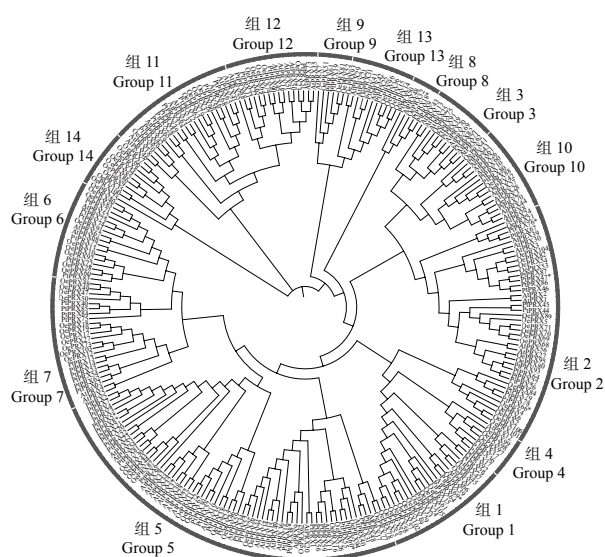


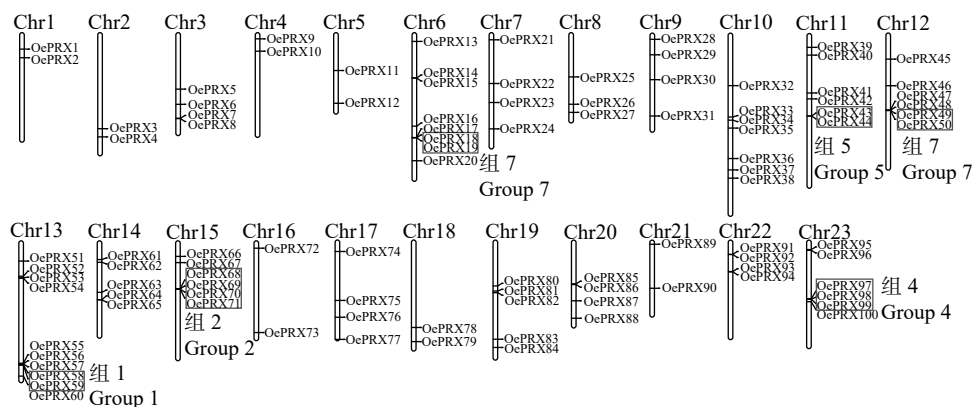
图 1 油橄榄、毛果杨和拟南芥 III 类 PRX 蛋白的系统进化关系

Fig. 1 Phylogeny of class III PRX proteins from olive, black cottonwood and *Arabidopsis thaliana*

地定位于 23 条染色体上, 6 个 OePRX 基因定位于 6 个不同的 contig 上。第 13 条染色体(Chr13)上 OePRX 基因最多, 数量为 10(9.4%); 其次是第 6 条染色体, 数量为 8(7.5%); 第 10 条上有 7 个 OePRX 基因(6.6%), 第 11、12、15、23 条上各有 6 个 OePRX 基因; 第 14、19 条上各有 5 个 OePRX 基因; 第 3、7、9、17、20、22 条上各有 4 个 OePRX 基因; 第 8 条上有 3 个 OePRX 基因; 第 1、2、4、5、16、18、21 条上各有 2 个 OePRX 基因。

2.3.2 OePRX 共线性分析

油橄榄经历了 2 次木犀科(Oleaceae)特有的全基因组复制事件, 分别发生在约 28 和 59 百万年前^[1]。为了挖掘 OePRX 基因家族的进化扩张与油橄榄全基因组复制事件的关系, 利用 TBtools v1.095 中的 One Step MCScanX 分析了 OePRX 基因家族内的复制基因, 发现 OePRX 基因家族中共 62 个 OePRX 基



串联复制基因簇用方框标记。Tandem replication gene cluster is marked with a box.

图 2 III 类油橄榄 PRX 基因在染色体上的分布图

Fig. 2 Chromosomal distribution of PRX genes in class III olive

因参与了基因复制, 占 *OePRX* 基因总数的 58.5%。如图 2 所示: 15 个基因(14.2%)参与了串联复制, 共得到了 6 个串联复制基因簇, 6 个基因簇分布于不同的染色体上, 分别是第 6、11、12、13、15、23 条染色体。如图 3 所示: 50 个基因(47.1%)参与了片段复制, 共得到了 34 对片段复制基因, 34 对基因分布于 19 条染色体上, 第 2、21、22 条上未分布, 说明片段复制是油橄榄III类 *PRX* 进化扩张的主要驱动力。值得注意的是, 3 个基因(*OePRX43*、*OePRX58*、*OePRX68*)既参与了片段复制, 也参与了串联复制。

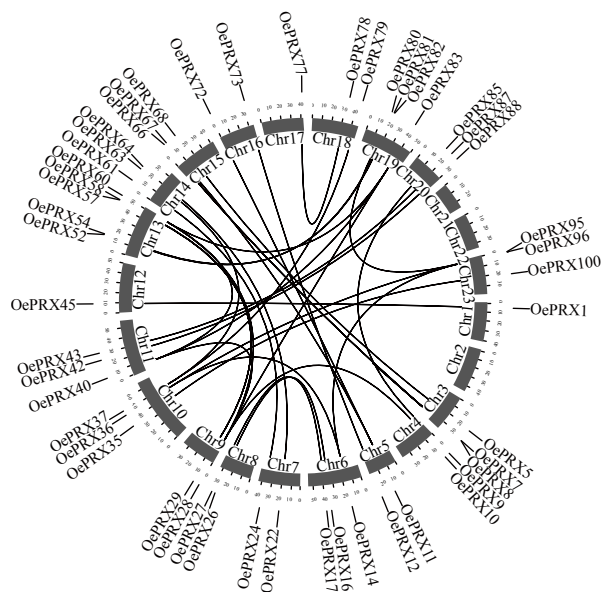


图3 油橄榄III类 *PRX* 的共线性分析
Fig. 3 Collinearity analysis of olive type III *PRX* genes

2.3.3 拟南芥、毛果杨和油橄榄 *PRX* 基因同线性分析及 Ka、Ks 分析

为了进一步探究 *OePRX* 基因家族的进化扩张, 绘制了 *OePRX* 与 *AtPRX* 和 *PtPRX* 基因的同线性图(图 4)。结果表明: 油橄榄和拟南芥之间有 40 对 III 类 *PRX* 复制基因; 油橄榄和毛果杨之间有 76 对 III 类 *PRX* 复制基因。并注意到, 25 个 *OePRX*(23.6%) 基因在拟南芥和毛果杨中均有同源基因, 推测在拟南芥、油橄榄和毛果杨分化前, 双子叶植物可能有过一次基因复制事件, 这与 Tuskan 等^[32]的研究一致。另外, 31 个 *OePRX*(29.2%) 基因在拟南芥中有直系同源基因, 46 个 *OePRX*(43.4%) 基因在毛果杨中有直系同源基因, 推测拟南芥与油橄榄的分化可能早于毛果杨与油橄榄的分化。

2.3.4 复制基因的 Ka、Ks 和 Ka/Ks 值分析

用 KaKs_Calculator2.0 计算复制基因的 Ka、Ks 和 Ka/Ks 值(图 5)。在 *OePRX* 基因家族的 6 个串联复制基因簇和 34 对片段复制基因中, 1 个串联复制基因簇(*OePRX97*、*OePRX98* 和 *OePRX99*)和 5 对片

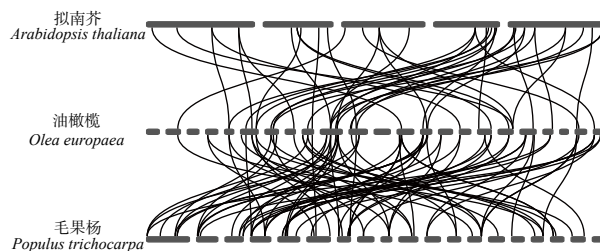


图4 油橄榄、毛果杨和拟南芥III类 *PRX* 的共线性分析
Fig. 4 Collinearity analysis of class III *PRX* in olive, black cottonwood and *Arabidopsis thaliana*

段复制基因(*OePRX8*/*OePRX66*、*OePRX16*/*OePRX26*、*OePRX24*/*OePRX61*、*OePRX36*/*OePRX96*、*OePRX35*/*OePRX100*)的 Ks 在 0.2 ~ 0.3 之间, 推测这些基因的复制可能发生于 28 百万年前的油橄榄全基因组复制事件^[1]。另外, 发现油橄榄复制基因对、*Oe-At* 复制基因对、*Oe-Pt* 复制基因对的 Ka/Ks 均小于 1, 说明 III 类 *PRX* 在双子叶植物的进化过程中, 受到纯化选择的作用。油橄榄片段复制基因对的 Ka/Ks 小于串联复制基因对的 Ka/Ks, 说明纯化选择在片段复制基因中更强。

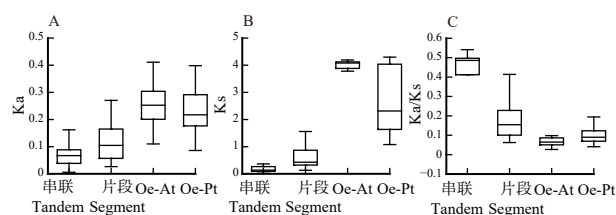


图5 油橄榄III类 *PRX* 各复制基因对的 Ka、Ks 分析
Fig. 5 Ka, Ks analyses of olive class III *PRX* duplicated genes

2.4 *OePRX* 基因序列分析

2.4.1 *PRX* 基因基序分布和基因结构分析

如图 6A, 用邻接法构建了 106 个 *OePRX* 蛋白的系统进化树, *OePRX* 蛋白家族分为 14 个组, 组 1、2、5、7、14 中的 *OePRX* 蛋白明显多于其他组。106 个 *OePRX* 蛋白平均约有 324 个氨基酸, 使用 MEME 分析了每一个 *OePRX* 蛋白的基序分布(图 6B), 基序个数设置为 10, 每个基序为 6 ~ 50 个氨基酸。发现 46 个 *OePRX* 蛋白(43.4%)含有 10 个基序, 34 个 *OePRX* 蛋白(32.1%)含有 9 个基序, 剩下的 *OePRX* 蛋白含有 3 ~ 8 个基序, 说明 *OePRX* 蛋白的基序分布整体比较保守。另外还注意到, 组 1 ~ 13 中 *OePRX* 蛋白含有 6 ~ 10 个基序 *PRX*, 组 14 的 *OePRX* 蛋白含有 3 ~ 5 个基序, 进一步说明组

件, 在 64 个 *OePRX* 基因的启动子中有分布; ARE、GC-motif 均为厌氧胁迫应答的顺式作用元件, 其中 ARE 分布于 76 个 *OePRX* 基因的启动子中, 说明 *OePRX* 基因广泛参与激素代谢过程, 调控油橄榄的生长发育, 且能响应逆境胁迫, 增强油橄榄抗逆性。

2.5 *OePRX* 表达模式分析

2.5.1 不同器官或组织中 *OePRX* 的差异表达

为了研究 *OePRX* 基因家族的表达模式, 分析了 106 个 *OePRX* 在根、茎、叶、花、果和分生组织中的转录组数据, 发现在 106 个 *OePRX* 基因中, 61 个 *OePRX* 基因有表达, *tpm* 在 1~846 之间; 其余 *OePRX* 基因的 *tpm* 小于 1, 可认为其不表达。如图 8 所示: 根据表达模式的相似性, 106 个 *OePRX* 基因在不同组织中的表达量聚为 8 个模块(module)。18 个基因(17.0%, 模块 A、B 和 C)在不同组织内普遍表达, 其中 7 个基因(模块 A)的整体表达量较高, 5 个基因(模块 B)表达量次之, 6 个基因(模块 C)表达量较低。13 个基因(12.3%, 模块 D、E、F 和 G)在 1 个或多个特定的组织内表达, 其中, 5 个基因(模块 E 和 F)在除果实之外的其他组织内表达, 2 个基因(模块 E)在茎、叶、分生组织中表达量更高, 7 个基因(模块 D)主要在花中表达, 1 个基因 *OePRX72*(模块 G)主要在根中表达。值得注意的是, 75 个基因(70.8%, 模块 H)几乎不表达或者仅在特定的组织中具有极低的表达。

对比图 8 中 *OePRX* 基因的进化分组和表达模块发现: 同一组中的基因表达模式相似, 比如组 3、

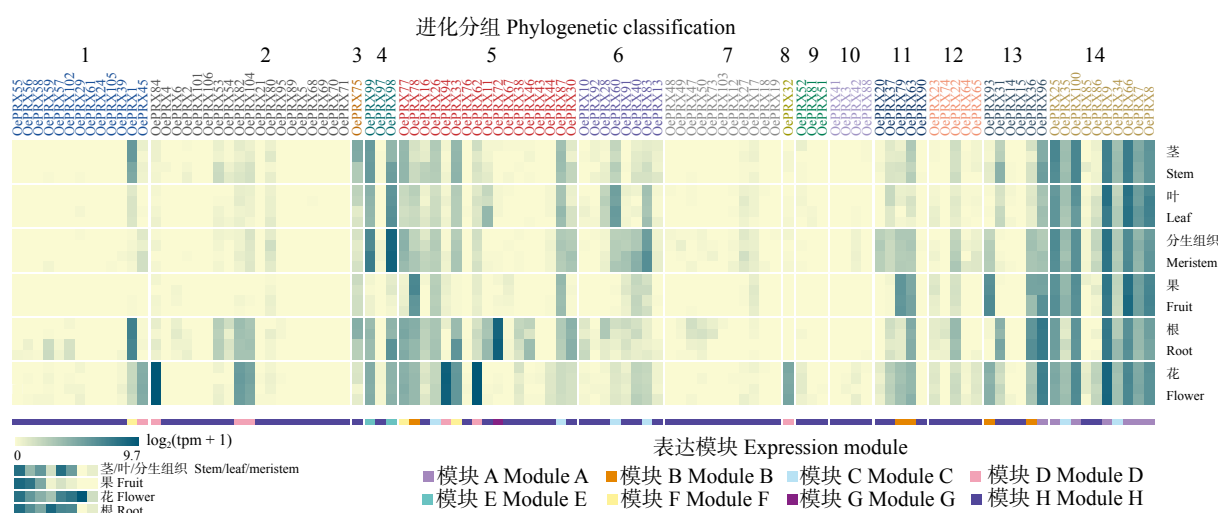
7、8、9、10 和 12 中的基因均有单一的基因表达模块; 也有一些组中的基因表达模式的变化幅度比较小, 比如组 2、4、6 和 11 中的基因均聚类为 2 个表达模块。然而, 组 5 中的基因聚类为 6 个不同的表达模块, 组 1、13、14 中的基因均聚类为 3 个不同的表达模块。尽管如此, 在组织间普遍表达类模块 A、B 和 C 的基因其进化关系更为相近, 在特定组织表达类模块 D、E、F 和 G 的基因也是如此。另外, 值得注意的是, 模块 H 中几乎不表达或者仅在特定的组织表达的 75 个基因均匀分布于各个组。

2.5.2 实时荧光定量 PCR 验证

运用实时荧光定量 PCR 验证了 6 个基因(*OePRX1*、*OePRX8*、*OePRX33*、*OePRX36*、*OePRX66* 和 *OePRX99*) 在叶片、茎、根和果实中的表达量(图 9)。运用 Primer Premier5.0 设计引物, 如表 2 所示。RNA-seq 数据和实时荧光定量结果均表明, *OePRX1* 和 *OePRX36* 基因在根中表达量最高, 在果实中表达量最低; *OePRX8* 和 *OePRX66* 基因在叶片和果实中表达量较高, 在茎和根中表达量较低; *OePRX33* 基因在茎和根中表达量较高, 在叶和果中表达量较低; *OePRX99* 基因在茎中表达量最高, 在果实中表达量最低。

2.6 非生物胁迫下 *OePRX* 基因的差异表达

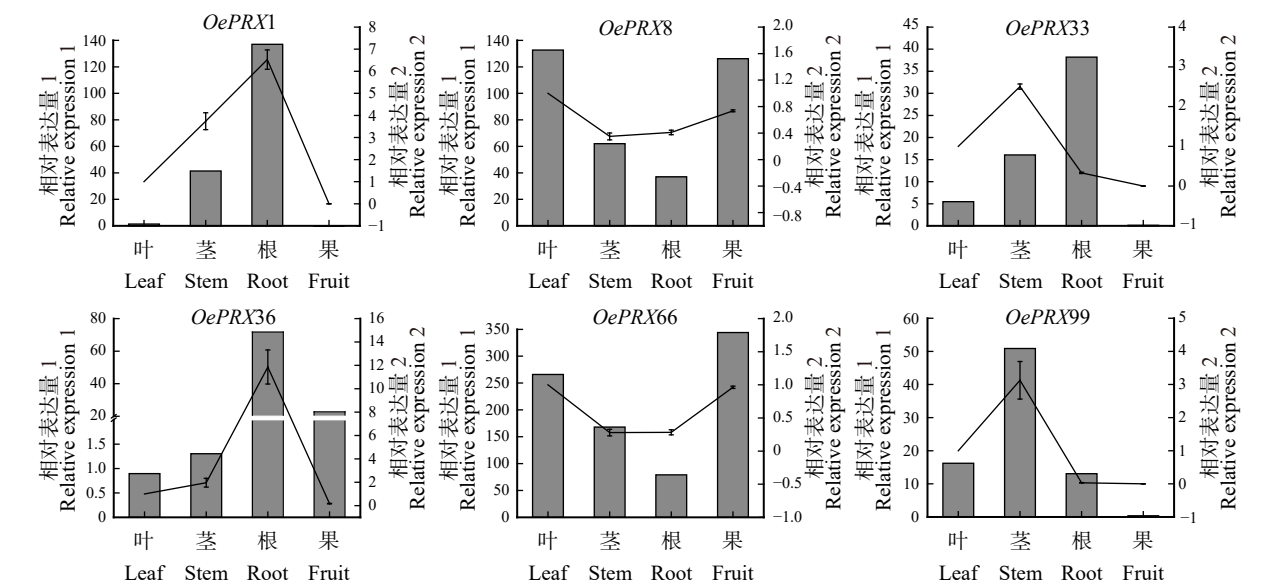
为了研究油橄榄III类 *PRX* 基因对非生物胁迫的响应, 利用 RNA-seq 数据分析了油橄榄响应干热胁迫和水涝胁迫过程中差异表达的 *OePRX* 基因。如图 10 所示: 干热胁迫下, 39 个 *OePRX* 差异表达, 占



模块 A. 在不同组织中均高表达; 模块 B. 在不同组织中均有较高表达; 模块 C. 在不同组织中均较低表达; 模块 D. 在花中特异高表达; 模块 E. 在除果外的组织中高表达; 模块 F. 在除果外的组织较高表达. 模块 G. 在根中特异高表达; 模块 H. 在各组织中低表达或几乎不表达。Module A, high expression in all tissues; module B, relative high expression in all tissues; module C, relative low expression in all tissues; module D, high specific expression in flower; module E, high expression in all tissues except fruit; module F, relative high expression in all tissues except fruit; module G, high specific expression in root; module H, low or no expression in all tissues.

图 8 油橄榄III类 *PRX* 基因在不同器官或组织中的表达模式

Fig. 8 Expression patterns of olive class III *PRX* genes in different organs or tissues



柱状图为运用 RNA-seq 得到的 tpm 值(相对表达量 1)。折线图为运用 RT-qPCR 得到的 $2^{-\Delta\Delta CT}$ 值(相对表达量 2)。Bars indicate tpm value measured by RNA-seq (relative expression 1). Lines indicate $2^{-\Delta\Delta CT}$ measured by RT-qPCR (relative expression 2).

图 9 RT-qPCR 验证转录组数据
Fig. 9 RT-qPCR validations of RNA-seq data

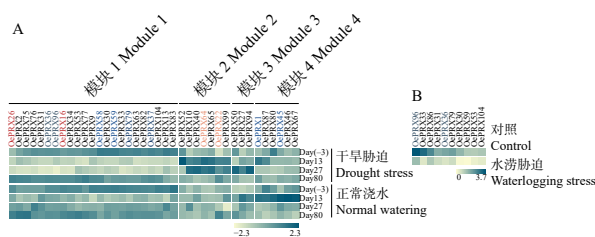
表 2 6 个油橄榄 III 类 PRX 基因 RT-qPCR 的引物

Tab. 2 Primers of six olive class III PRX genes used for RT-qPCR

基因名称 Gene name	引物(5'—3') Primer sequence (5'—3')	基因产物长度 Gene product length/bp
<i>OePRX1</i>	F: TCCAGGAGTTGTTTCTTGTC R: CCTTCTTCGGTCTTTTCTTCC	114
<i>OePRX8</i>	F: AAGGGCGCATCTCTGAAAG R: AGAAGGCATCTTCATCCTTAGC	200
<i>OePRX33</i>	F: GGAAGCATAATAAGCGAGAAG R: TTATTGGAGCCACTCAAACCTG	235
<i>OePRX36</i>	F: ACAAACGCCACAAGAACACTG R: CACTCCCTCTCTAAAGCCTCC	195
<i>OePRX66</i>	F: GATTCGCAATGAGGAGGAGTA R: AGCAACAACACCAGCAAGC	140
<i>OePRX99</i>	F: GCTCGGGCTTTTAGAATCATC R: GAAGGAAGGTTTGCCAGTGTT	204

比 36.8%；水涝胁迫下，10 个 *OePRX* 差异表达，占比 9.4%。基于差异表达基因在干旱胁迫和复水过程中的表达趋势，将其分为 4 个模块：模块 1，23 个基因(21.7%)的表达在干旱胁迫过程中显著下调，复水后显著上调；模块 2，7 个基因(6.6%)的表达在干旱胁迫过程中显著上调，复水后显著下调；模块 3，3 个基因(2.8%)的表达在高温和前期干旱时变化不显著，在后期干旱时显著上调，复水后显著下调；模块 4，6 个基因(5.7%)的表达在高温胁迫下显著上调，解除高温胁迫后，显著下调。另外，与正常浇水的油橄榄相比，在第 13 天，高温和早期干旱处理的油橄榄中有 8 个基因(7.5%)显著上调，分别是：*OePRX31*、

OePRX36、*OePRX45*、*OePRX46*、*OePRX63*、*OePRX82*、*OePRX83* 和 *OePRX96*。在第 27 天，13 个基因(12.3%)显著上调，分别是：*OePRX2*、*OePRX16*、*OePRX26*、*OePRX31*、*OePRX33*、*OePRX36*、*OePRX63*、*OePRX76*、*OePRX79*、*OePRX82*、*OePRX96* 和 *OePRX104*；4 个基因(3.8%)显著下调，分别是 *OePRX40*、*OePRX50*、*OePRX65* 和 *OePRX99*；在第 80 天，*OePRX* 基因表达无显著变化。水涝胁迫处理 90 d 后，油橄榄叶片中 10 个 *OePRX* 基因(9.4%)的表达显著下调，猜测这些差异表达的 III 类 PRX 在油橄榄响应非生物胁迫中具有十分重要的作用。PRX 除干旱和水淹胁迫外，我国油橄榄常常面临低温和营养等胁迫，因此有



Day(-3). 干旱处理前 3 天; Day13. 干旱处理第 13 天且伴随高温; Day27. 干旱处理第 27 天; Day80. 复水后 52 天。模块 1. 干旱期间基因表达下调, 复水后上调; 模块 2. 干旱期间基因表达上调, 复水后下调; 模块 3. 后期干旱期间基因表达上调, 后期干旱及复水后下调; 模块 4. 仅在高温下基因表达上调, 高温胁迫后, 基因表达下调。复制基因名称的颜色与系统进化分组的颜色相对应。Day(-3), 3 d before drought treatment; Day13, 13th day of drought treatment with high temperature; Day27, 27th day after drought treatment; Day80, 52 d after re-irrigation. Module 1, gene expression is down regulated during drought and up regulated after rehydration; module 2, gene expression is upregulated during drought and downregulated after rehydration; module 3, gene expression is upregulated during late drought, and downregulated after late drought and rehydration; module 4, gene expression is upregulated only under high temperature, and downregulated after high temperature stress. The color of the replicated gene name corresponds to the color of the phylogenetic grouping.

图 10 油橄榄III类 *PRX* 基因响应干热和水涝胁迫的差异表达模式
Fig. 10 Differential expression patterns of class III *PRX* genes in olive oil in response to dry heat and waterlogging stresses

必要对其进行更为丰富而广泛地探索, 进一步研究III类 *PRX* 基因在油橄榄响应非生物胁迫过程中的功能。

3 讨论与结论

III类 *PRX* 基因通过氧化特定小分子、调节活性氧水平来促进木质化、细胞伸长、响应生物胁迫和非生物胁迫^[33-34]。为了调控基因表达的时空特异性、蛋白分布的特定性和过氧化物酶活性的多样性, 高等植物进化得到了庞大的III类 *PRX* 基因家族^[35]。在高等植物的进化过程中, III类 *PRX* 基因家族通过全基因组复制事件、片段复制、串联复制和转座子复制等方式不断扩张。前人研究表明: 拟南芥中有 73 个 *PRX* 基因, 9 对片段复制基因, 10 个串联复制基因簇^[13]; 毛茛杨中有 93 个 *PRX* 基因, 由 60 ~ 65 百万年前的全基因组复制事件产生了 14 对复制基因, 共有 37 个串联复制基因簇^[16]; 白梨中有 94 个 *PRX* 基因, 26 对片段复制基因, 1 个串联复制基因簇^[17]; 水稻中有 138 个 *PRX* 基因, 多于 20 对的片段复制基因和 6 个串联复制基因簇^[14]; 玉米中有 119 个 *PRX* 基因, 16 对片段复制基因和 12 个串联复制基因簇^[15]。本研究中, 在双子叶木本植物油橄榄中鉴定得到了 106 个III类 *PRX* 基因, 其中, 62 个 *OePRX* (58.5%) 参与了基因复制, 共产生了 6 个串联复制基因簇和 34 对片段复制基因。据此猜想, 片段复制可

能是高等植物 *PRX* 基因家族进化扩张的主要方式。

基因复制事件是基因组进化过程中基因功能冗余和分化的主要来源^[36]。复制基因的功能具有相关性, 且其功能的分化与其表达模式的分化有关。例如: 三倍体小麦中的同源基因 *VRN-A1*、*VRN-B1* 和 *VRN-D1* 是调节开花时间的重要因子, 不同抽穗期的小麦中, 这 3 个同源基因在不同组织中的表达量不同^[37]。本研究发现, 许多 *OePRX* 基因在不同组织中的表达模式与其在拟南芥或其他模式物种中的表达模式相似, 说明 *PRX* 的基因功能在不同植物之间具有保守性。*AtPRX33* 和 *AtPRX34* 基因是一对高度同源(95%)的III类 *PRX* 基因, 但是其启动子和内含子序列却高度分化。半定量 RT-PCR 结果表明, 这 2 个基因主要在拟南芥幼苗的根部表达, 并且光照对其表达具有促进作用。与野生型拟南芥相比, *AtPRX33* 基因突变体的根变短, *AtPRX33* 和 *AtPRX34* 基因双突变体的根更短, 而 *AtPRX34* 基因过表达拟南芥幼苗的根显著变长^[38]。本研究发现: *OePRX* 基因家族中的 6 个串联复制基因簇和 34 个片段复制基因的分子量、等电点、信号肽的分布和长度以及基序分布均比较保守; 但 2 个串联复制基因簇(33.3%)和 22 对片段复制基因(64.7%)的内含子位置发生分化; 1 个串联复制基因簇(16.7%)和 11 对片段复制基因(32.4%)在不同组织中的表达模式发生分化。因此, 有必要利用分子生物学手段深入研究油橄榄III类 *PRX* 基因家族内复制基因的亚功能和新功能。

为了研究单个III类 *PRX* 基因在逆境下的特定功能, 大量III类 *PRX* 转基因植株已经获得。过表达甜土豆 (*Ipomoea batatas*) *swpa4 PRX* 基因的烟草 (*Nicotiana tabacum*) 幼苗和过表达 *AtPRX3* 基因的拟南芥幼苗^[39]的耐盐性和耐旱性增强。过表达 *AtPRX22*、*AtPRX39* 和 *AtPRX69* 基因的拟南芥对油菜素类固醇不敏感, 而突变体的耐寒性增强^[40]。同样地, 2 个长春花 (*Catharanthus roseus*) *PRX* 基因在烟草中的异源转化提高了其在脱水胁迫下的发芽率, 同时提高了其耐旱性^[41]。辣椒 (*Capsicum annuum*) *CpPRX2* 的转基因拟南芥幼苗提高了其体内 H_2O_2 水平的积累, 同时提高了其抗病性、耐旱性和耐盐性^[42]。菜豆 (*Phaseolus vulgaris*) 过氧化物酶 *FBP1* 反义表达的拟南芥中, *AtPRX33* 和 *AtPRX34* 基因的表达量降低, 体内氧化应激受到抑制, 抗病性减弱^[43]。本研究发现: 干热胁迫和水涝胁迫下油橄榄中 40 个 *OePRX* 基因(37.7%)显著差异表达, 且差异表达模式多样(图 10), 猜测其他 *OePRX* 基因可能参与另外的生物胁迫或非生物胁迫, 需要对多种胁迫下的油橄榄幼苗开展更为广泛的研究, 锁定响应胁迫的关

键基因,从而深入挖掘其功能。Wu等^[44]研究表明:经历了多次多倍化复制的植物保留了许多胁迫相关的基因家族,且得到的复制基因通常是胁迫相关通路的关键因子。本研究结果表明:与正常生长的油橄榄相比,干热胁迫下油橄榄中串联复制基因簇(*OePRX58*和*OePRX59*)和3对片段复制基因(*OePRX16*和*OePRX26*,*OePRX36*和*OePRX96*,*OePRX37*和*OePRX79*)的表达量均显著降低,片段复制基因对(*OePRX22*和*OePRX64*)的表达量均显著上升,片段复制基因对(*OePRX1*和*OePRX45*)的表达量仅在高温下显著上升;水涝胁迫下,片段复制基因对(*OePRX36*和*OePRX96*)的表达量均显著降低,猜测这些复制基因之间具有较强的功能相关性,可能是研究油橄榄非生物胁迫响应通路的关键候选基因。

综上所述,油橄榄*OePRX*基因家族明显的进化扩张和多变的表达模式暗示了其功能的复杂性,尤其是复制基因的亚功能化和新功能化对油橄榄在地中海地区的广泛分布具有十分重要的适应性意义。

参 考 文 献

- [1] Unver T, Wu Z, Sterck L, et al. Genome of wild olive and the evolution of oil biosynthesis[J/OL]. Proceedings of the National Academy of Sciences of the United States of America, 2017 [2017-10-09]. DOI: [10.1073/pnas.1708621111](https://doi.org/10.1073/pnas.1708621111).
- [2] Rao G, Zhang J, Liu X, et al. De novo assembly of a new *Olea europaea* genome accession using nanopore sequencing[J/OL]. Horticulture Research, 2021 [2021-04-01]. DOI: [10.1038/s41438-021-00498-y](https://doi.org/10.1038/s41438-021-00498-y).
- [3] Mathé C, Barre A, Jourda C, et al. Evolution and expression of class III peroxidases[J]. Archives of Biochemistry and Biophysics, 2010, 500(1): 58-65.
- [4] Passardi F, Bakalovic N, Teixeira F K, et al. Prokaryotic origins of the non-animal peroxidase superfamily and organelle-mediated transmission to eukaryotes[J]. Genomics, 2007, 89(5): 567-579.
- [5] Shigeoka S, Ishikawa T, Tamoi M, et al. Regulation and function of ascorbate peroxidase isoenzymes[J]. Journal of Experimental Botany, 2002, 53(372): 1305-1319.
- [6] Skulachev V P. Cytochrome c in the apoptotic and antioxidant cascades[J]. FEBS Letters, 1998, 423(3): 275-280.
- [7] Ruiz-Dueñas F J, Camarero S, Pérez-Boada M, et al. A new versatile peroxidase from *Pleurotus*[J]. Biochemical Society Transactions, 2001, 29(Pt2): 116-122.
- [8] Passardi F, Cosio C, Penel C, et al. Peroxidases have more functions than a Swiss army knife[J]. Plant Cell Reports, 2005, 24(5): 255-265.
- [9] Passardi F, Penel C, Dunand C. Performing the paradoxical: how plant peroxidases modify the cell wall[J]. Trends in Plant Science, 2004, 9(11): 534-540.
- [10] Allison S D, Schultz J C. Differential activity of peroxidase isozymes in response to wounding, gypsy moth, and plant hormones in northern red oak (*Quercus rubra* L.)[J]. Journal of Chemical Ecology, 2004, 30(7): 1363-1379.
- [11] Gazaryan I G, Lagrimini L M, Ashby G A, et al. Mechanism of indole-3-acetic acid oxidation by plant peroxidases: anaerobic stopped-flow spectrophotometric studies on horseradish and tobacco peroxidases[J]. The Biochemical Journal, 1996, 313(Pt3): 841-847.
- [12] Almagro L, Gómez Ros L V, Belchi-Navarro S, et al. Class III peroxidases in plant defence reactions[J]. Journal of Experimental Botany, 2008, 60(2): 377-390.
- [13] Tognolli M, Penel C, Greppin H, et al. Analysis and expression of the class III peroxidase large gene family in *Arabidopsis thaliana*[J]. Gene, 2002, 288(1): 129-138.
- [14] Passardi F, Longet D, Penel C, et al. The class III peroxidase multigenic family in rice and its evolution in land plants[J]. Phytochemistry, 2004, 65(13): 1879-1893.
- [15] Wang Y, Wang Q, Zhao Y, et al. Systematic analysis of maize class III peroxidase gene family reveals a conserved subfamily involved in abiotic stress response[J]. Gene, 2015, 566(1): 95-108.
- [16] Ren L L, Liu Y J, Liu H J, et al. Subcellular relocation and positive selection play key roles in the retention of duplicate genes of *Populus* class III peroxidase family[J]. The Plant Cell, 2014, 26(6): 2404-2419.
- [17] Cao Y, Han Y, Meng D, et al. Structural, evolutionary, and functional analysis of the class III peroxidase gene family in chinese pear (*Pyrus bretschneideri*) [J/OL]. Frontier in Plant Science, 2016 [2016-12-09]. DOI: [10.3389/fpls.2016.01874](https://doi.org/10.3389/fpls.2016.01874).
- [18] Xiao H, Wang C, Khan N, et al. Genome-wide identification of the class III *POD* gene family and their expression profiling in grapevine (*Vitis vinifera* L.) [J/OL]. BMC Genomics, 2020. [2020-06-29]. DOI: [10.1186/s12864-020-06828-z](https://doi.org/10.1186/s12864-020-06828-z).
- [19] Wu C, Ding X, Ding Z, et al. The class III peroxidase (*POD*) gene family in cassava: identification, phylogeny, duplication, and expression[J/OL]. International Journal of Molecular Sciences, 2019 [2019-06-03]. DOI: [10.3390/ijms20112730](https://doi.org/10.3390/ijms20112730).
- [20] Yan J, Su P, Li W, et al. Genome-wide and evolutionary analysis of the class III peroxidase gene family in wheat and *Aegilops tauschii* reveals that some members are involved in stress responses[J/OL]. BMC Genomics, 2019 [2019-08-22]. DOI: [10.1186/s12864-019-6006-5](https://doi.org/10.1186/s12864-019-6006-5).
- [21] Fernández-Pérez F, Pomar F, Pedreño M A, et al. The suppression of *AtPrx52* affects fibers but not xylem lignification in *Arabidopsis* by altering the proportion of syringyl units[J]. Physiologia Plantarum, 2015, 154(3): 395-406.
- [22] Fernández-Pérez F, Vivar T, Pomar F, et al. Peroxidase 4 is involved in syringyl lignin formation in *Arabidopsis thaliana*[J]. Journal of Plant Physiology, 2015, 175(1): 86-94.
- [23] Herrero J, Fernández-Pérez F, Yebra T, et al. Bioinformatic and functional characterization of the basic peroxidase 72 from *Arabidopsis thaliana* involved in lignin biosynthesis[J]. Planta,

- 2013, 237(6): 1599–1612.
- [24] Wu Y, Yang Z, How J, et al. Overexpression of a peroxidase gene (*AtPrx64*) of *Arabidopsis thaliana* in tobacco improves plant's tolerance to aluminum stress[J]. *Plant Molecular Biology*, 2017, 95(1): 157–168.
- [25] Kidwai M, Dhar Y V, Gautam N, et al. *Oryza sativa* class III peroxidase (*OsPRX38*) overexpression in *Arabidopsis thaliana* reduces arsenic accumulation due to apoplastic lignification[J]. *Journal of Hazardous Materials*, 2019, 362(15): 383–393.
- [26] Ramírez-Tejedo J A, Jiménez-Ruiz J, Leyva-Pérez M d l O, et al. Gene expression pattern in olive tree organs (*Olea europaea* L.) [J/OL]. *Genes*, 2020 [2020–05–12]. DOI: [10.3390/genes11050544](https://doi.org/10.3390/genes11050544).
- [27] Tsamir-Rimon M, Ben-Dor S, Feldmesser E, et al. Rapid starch degradation in the wood of olive trees under heat and drought is permitted by three stress-specific beta amylases[J]. *New Phytologist*, 2021, 229(3): 1398–1414.
- [28] Dastkar E, Soleimani A, Jafari H, et al. Differential expression of genes in olive leaves and buds of ON- versus OFF-crop trees[J]. *Scientific Reports*, 2020, 10(1): 1–13.
- [29] Li H, Poulos T L. Structural variation in heme enzymes: a comparative analysis of peroxidase and P450 crystal structures[J]. *Structure*, 1994, 2(6): 461–464.
- [30] Lu S, Wang J, Chitsaz F, et al. CDD/SPARCLE: the conserved domain database in 2020[J]. *Nucleic Acids Research*, 2019, 48(D1): D265–D268.
- [31] Oliveira R A d C, de Andrade A S, Imparato D O, et al. Analysis of *Arabidopsis thaliana* redox gene network indicates evolutionary expansion of class III peroxidase in plants[J/OL]. *Scientific Reports*, 2019 [2019–10–31]. DOI: [10.1038/s41598-019-52299-y](https://doi.org/10.1038/s41598-019-52299-y).
- [32] Tuskan G A, DiFazio S, Jansson S, et al. The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray)[J]. *Science*, 2006, 313: 1596–1604.
- [33] 芮伟康. 钙与过氧化物酶在梨石细胞合成中的关系 [D]. 南京: 南京农业大学, 2017.
- Rui W K. Effects of calcium and peroxidase involved in stone cells formation in pear fruit[D]. Nanjing: Nanjing Agricultural University, 2017.
- [34] 薛亚莉. 水稻III类过氧化物酶基因 *OsPER2* 和 *OsPER4* 在细胞壁合成中的功能分析 [D]. 武汉: 华中农业大学, 2021.
- Xue Y L. Functional analysis of rice class III peroxidases genes *OsPER2* and *OsPER4* in cell wall synthesis[D]. Wuhan: Huazhong Agricultural University, 2021.
- [35] Teufel A I, Johnson M M, Laurent J M, et al. The many nuanced evolutionary consequences of duplicated genes[J]. *Molecular Biology and Evolution*, 2018, 36(2): 304–314.
- [36] Shigeto J, Tsutsumi Y. Diverse functions and reactions of class III peroxidases[J]. *New Phytologist*, 2016, 209(4): 1395–1402.
- [37] Xu S, Chong K. Remembering winter through vernalisation[J]. *Nature Plants*, 2018, 4(12): 997–1009.
- [38] Passardi F, Tognolli M, de Meyer M, et al. Two cell wall associated peroxidases from *Arabidopsis* influence root elongation[J]. *Planta*, 2006, 223(5): 965–974.
- [39] Kim Y H, Kim C Y, Song W K, et al. Overexpression of sweetpotato *swpa4* peroxidase results in increased hydrogen peroxide production and enhances stress tolerance in tobacco[J]. *Planta*, 2008, 227(4): 867–881.
- [40] Kim B H, Kim S Y, Nam K H. Genes encoding plant-specific class III peroxidases are responsible for increased cold tolerance of the brassinosteroid-insensitive 1 mutant[J]. *Molecules and Cells*, 2012, 34(6): 539–548.
- [41] Kumar S, Jaggi M, Sinha A K. Ectopic overexpression of vacuolar and apoplastic *Catharanthus roseus* peroxidases confers differential tolerance to salt and dehydration stress in transgenic tobacco[J]. *Protoplasma*, 2012, 249(2): 423–432.
- [42] Choi H W, Kim Y J, Lee S C, et al. Hydrogen peroxide generation by the pepper extracellular peroxidase CaPO2 activates local and systemic cell death and defense response to bacterial pathogens[J]. *Plant Physiology*, 2007, 145(3): 890–904.
- [43] Daudi A, Cheng Z, O'Brien J A, et al. The apoplastic oxidative burst peroxidase in *Arabidopsis* is a major component of pattern-triggered immunity[J]. *The Plant Cell*, 2012, 24(1): 275–287.
- [44] Wu S, Han B, Jiao Y. Genetic contribution of paleopolyploidy to adaptive evolution in angiosperms[J]. *Molecular Plant*, 2020, 13(1): 59–71.

(责任编辑 范娟 赵田芸
责任编辑 蒋湘宁)