

• 药材与资源 •

虎杖 UGT 基因家族鉴定及其在药材不同部位的表达分析

尚进¹, 陈丽娜¹, 胡碟¹, 邓爱平¹, 陈庆梅², 杜雄军³, 黄晓^{1,4*}, 余坤^{1,4*}

1. 湖北中医药大学药学院, 湖北 武汉 430065

2. 湖北省中医院, 湖北 武汉 430060

3. 湖北福人药业股份有限公司, 湖北 通城 437400

4. 湖北中医药大学 中药资源与化学湖北省重点实验室, 湖北 武汉 430065

摘要: 目的 系统解析虎杖 *Polygonum cuspidatum* UDP-糖基转移酶 (UDP-glycosyltransferases, UGT) 基因家族的组成及结构特征, 并分析其在药材不同组织中的表达特征, 以及在白藜芦醇和大黄素糖苷化过程中的潜在功能, 为阐明虎杖药效物质生物合成机制提供基础。方法 利用虎杖基因组数据对 UGT 基因家族进行全基因组鉴定, 分析其理化性质、染色体定位、基因结构、保守基序及顺式调控元件, 结合系统发育树、基因复制事件和蛋白互作网络揭示其进化特征; 结合转录组与代谢组数据进行相关性分析获取代表性基因, 并利用 qRT-PCR 进行验证。结果 虎杖全基因组中共鉴定出 UGT 基因家族成员 178 个, 分为 14 个亚类, 具有丰富的结构与功能多样性。成员在染色体上分布不均, 部分基因可能来源于串联和片段复制扩张。启动子区域顺式元件与胁迫、激素响应密切相关, 富含与发育、压力等响应调控相关元件。表达量相关性分析结果显示, *PcUGT065*、*PcUGT130* 等基因与白藜芦醇、大黄素及其糖苷类物质的合成显著相关, qRT-PCR 分析验证了转录组数据的可靠性。结论 虎杖 UGT 基因家族具有明显的结构保守性与表达差异性, 部分成员参与白藜芦醇与大黄素的糖基修饰, 为深入研究虎杖活性产物的生物合成机制及其分子育种提供了基础。

关键词: 虎杖; UGT 基因家族; 生物信息分析; 成员鉴定; 表达分析; 白藜芦醇; 大黄素

中图分类号: R282.12 文献标志码: A 文章编号: 0253-2670(2026)08-3110-11

DOI: 10.7501/j.issn.0253-2670.2026.08.023

Identification of UGT gene families in *Polygonum cuspidatum* and expression analysis in different parts of medicinal material

SHANG Jin¹, CHEN Lina¹, HU Die¹, DENG Aiping¹, CHEN Qingmei², DU Xiongjun³, HUANG Xiao^{1,4}, YU Kun^{1,4}

1. College of Pharmacy, Hubei University of Chinese Medicine, Wuhan 430065, China

2. Hubei Provincial Hospital of Traditional Chinese Medicine, Wuhan 430060, China

3. Hubei Furen Pharmaceutical Co., Ltd., Tongcheng 437400, China

4. Hubei Key Laboratory of TCM Resources and Chemistry, Hubei University of Chinese Medicine, Wuhan 430065, China

Abstract: Objective To systematically analyze the composition, structural characteristics of the UDP-glycosyltransferase (UGT) gene family in *Polygonum cuspidatum*, investigate its expression patterns across different root tissues, as well as its potential roles in the glycosylation of resveratrol and emodin, so as to provides a theoretical foundation for elucidating the biosynthetic mechanisms of active metabolites. **Methods** The UGT gene family of *P. cuspidatum* was identified at the genome-wide level based on transcriptome and genomic datasets. Analyses of physicochemical properties, chromosomal localization, gene structure, conserved motifs and cis-regulatory elements were then performed. The evolutionary characteristics were analyzed through phylogenetic reconstruction, gene duplication analysis and protein interaction network prediction. Correlation analysis integrating transcriptomic and metabolomic data was conducted, and six representative genes potentially involved in the glycosylation of resveratrol and emodin were selected for

收稿日期: 2025-11-02

基金项目: 湖北省自然科学基金资助项目 (2023AFD122); 湖北省高等学校优秀中青年科技创新团队计划 (T2022020)

作者简介: 尚进 (1999—), 硕士研究生, 研究方向为中药资源与品质研究。E-mail: 17860501868@163.com

*通信作者: 余坤, 教授, 博士生导师, 从事中药资源与品质研究。E-mail: yukun_hbucm@163.com

黄晓, 博士, 硕士生导师, 从事中药资源与品质研究。E-mail: hxiao0430@hbucm.edu.cn

validation using qRT-PCR. **Results** A total of 178 *UGT* genes were identified and grouped into 14 subfamilies, exhibiting substantial structural and functional diversity. The gene family members were unevenly distributed in chromosomes, and several genes were likely generated through tandem or segmental duplication events. Promoter analysis revealed that cis-acting elements were predominantly associated with stress and hormone responses, as well as developmental regulation. Expression correlation heatmaps indicated that genes such as *PcUGT065* and *PcUGT130* showed strong correlation with resveratrol, emodin, and their glycosylated derivatives. The qRT-PCR results further confirmed the reliability of transcriptome-based expression patterns. **Conclusion** The *UGT* gene family in *P. cuspidatum* exhibits clear structural conservation along with distinct expression divergence. Several members may participate in the modification of resveratrol and emodin. This study provides a theoretical basis for future investigations into the biosynthetic mechanisms of active metabolites in *P. cuspidatum* and offers insights for its molecular breeding.

Key words: *Polygonum cuspidatum* Sieb. et Zucc.; *UGT* gene family; bioinformatics analysis; member identification; expression analysis; resveratrol; emodin

虎杖来源于蓼科多年生草本植物虎杖 *Polygonum cuspidatum* Sieb. et Zucc. 的根茎和根, 具有利湿退黄、清热解毒、散瘀止痛、止咳化痰等功效^[1], 主要活性成分为蒽醌类、二苯乙烯类、黄酮类及其糖苷衍生物, 如白藜芦醇 (resveratrol)、大黄素 (emodin)、虎杖苷和大黄素-8-*O*-葡萄糖苷等^[2]。研究表明虎杖苷和大黄素-8-*O*-葡萄糖苷对湿热黄疸、风湿痹痛、痈肿疮毒等疾病的治疗具有良好的活性^[2]。糖基化修饰是虎杖活性成分生物合成的重要环节, 对其理化性质、稳定性和生物活性具有较大影响^[3]。糖基化反应通过糖基转移酶 (glycosyltransferases, GTs) 催化完成, 将激活供体上的糖基团转移至特定受体分子, 从而生成更加稳定的糖苷^[4]或糖酯^[5]。常见的糖基包括葡萄糖、鼠李糖、木糖和半乳糖^[6-7]。

依据序列相似性、底物特异性和催化特异性, 糖基转移酶可划分为 110 个不同的家族, 其中 GT1 基因成员数量最多且功能最为关键, 主要催化尿苷二磷酸 (uridine diphosphate, UDP)-糖基向蛋白质^[8]、抗生素^[9]、核酸^[10]、植物激素^[11]等受体分子转移, 因此该家族被称为 UDP-糖基转移酶 (UDP-glycosyltransferases, UGT)^[12]。UGT 基因家族广泛存在于植物体内, 能够识别多种酚羟基和苷元底物^[13], 在天然产物生物合成、植物激素代谢^[14]及抗逆应答^[15]等生理过程中发挥关键作用。尽管已有研究报道部分 UGT 参与白藜芦醇类或蒽醌类物质糖基化^[16], 但虎杖中 *UGT* 基因的组成、结构特征及其潜在功能尚未得到系统解析, 限制了对白藜芦醇和大黄素糖苷合成机制的深入理解。

本研究基于虎杖转录组与基因组数据, 从全基因组水平鉴定 *UGT* 家族成员, 分析其理化性质、染色体分布、基因结构、保守基序、顺式调控元件和

系统发育关系, 结合组织表达模式与蛋白互作网络预测关键功能基因, 并进一步预测与白藜芦醇及大黄素糖基化相关的候选 *UGT* 基因, 为后续功能验证与代谢工程调控提供基础。

1 材料

虎杖植株样品种植在湖北中医药大学虎杖种质资源圃, 经湖北中医药大学刘义梅教授鉴定为蓼科植物虎杖。选择长势一致, 无病虫害的虎杖植株 3 株, 挖取根茎及根后分别分离皮部、木部和髓部, 即根皮部 (RC, 包括根的周皮及形成层)、根本部 (RL, 为根的木质部)、根茎皮部 (RhC, 是根茎除去木部和髓部的部分, 包括根茎的周皮、皮层、形成层及韧皮部)、根茎木部 (RhL, 为根茎的木质部) 和根茎髓部 (RhM, 为根茎的髓部) 5 个部位, 共 15 个样品, 表面清洗去除杂质后立即放入液氮中速冻, 再转入 -80 °C 超低温冰箱保存, 用于后续基因定量分析。

2 方法

2.1 数据来源

本研究涉及物种的数据来源如下: 虎杖全基因组及转录组, 拟南芥蛋白序列 (<https://www.arabidopsis.org/>)、氨基酸序列和相应的注释文件 (<https://www.ncbi.nlm.nih.gov/datasets/genome/>)。本研究的虎杖数据 (PRJCA032970) 保存在中国国家基因组科学数据中心 (<https://bigd.big.ac.cn/>)。

2.2 UGT 家族成员的鉴定

以拟南芥 *UGT* (AtUGT) 基因家族的蛋白序列作为参考, 筛选虎杖全基因组数据, 使用 TBtools v2.376 软件中的 simple hmm search 程序, 利用 hmm 模型 (PF00201) 筛选获得 *UGT* 蛋白序列, 获得虎杖 *UGT* (PcUGT) 基因家族全长氨基酸序列。利用 The MEME Suite 5.5.5 (<https://meme-suite.org/meme/>)

tools/meme)在线分析 PcUGT 家族蛋白的基序,完成基序和基因的可视化。

2.3 PcUGT 蛋白的理化性质和亚细胞定位预测

利用在线工具 ExPasy ProtParam tool (<https://web.expasy.org/protparam/>)分析 PcUGT 蛋白的基本理化性质,包括相对分子质量、等电点和亲水性。利用 signalP 4.1 对蛋白质信号肽进行预测。通过 WoLF PSORT (<https://wolfsort.hgc.jp/>)在线预测蛋白亚细胞定位,参数均设置为默认值。

2.4 PcUGT 基因染色体定位

使用 TBtools v2.376 软件从虎杖基因组 gff 文件中提取 PcUGT 基因家族成员的染色体位置信息^[17],完成基因家族的基因重复事件分析,并进行绘图可视化。

2.5 系统发育树、种间共线性与保守结构域分析

从 TAIR (<https://www.arabidopsis.org/>)拟南芥数据库中下载完整的 *AtUGT* 基因家族氨基酸序列 117 个,与 PcUGT 基因家族鉴定进行相同处理。选取拟南芥和虎杖 UGT 基因家族保守结构域构建进化树。采用 DNAMAN 6.0 软件对 PcUGT 基因家族蛋白保守结构域进行多序列比对,利用 MEGA 11 进行多序列比对分析,邻位法构建系统进化树,设置重复次数为 1 000,其余参数选择默认值,构建虎杖与拟南芥 UGT 基因家族蛋白序列进化树。使用 MCSanX (<https://github.com/wyp1125/MCSanX>)软件分析虎杖与拟南芥种间的 UGT 基因共线性关系。

2.6 PcUGT 基因家族的基因结构和编码蛋白的保守基序分析

使用 MEME (<http://meme-suite.org/meme/tools/meme>)在线工具预测 PcUGT 家族蛋白序列的 motif 分布,目标基序数目设定为 10,每个基序宽 6~50 个氨基酸残基。使用 TBtools v2.376 完成基序和基因结构的可视化。

2.7 基因结构分析与启动子顺式作用元件的预测

利用 TBtools v2.376 软件对 *PcUGT* 基因的内含子和外显子特征进行分析^[18]。从基因组中提取基因上游 2 000 bp 启动子序列,并通过在线工具 PlantCARE (<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>)预测顺式作用元件,并分析了其应激和激素反应元件。数据的可视化由 TBtools v2.376 软件完成。

2.8 PcUGT 成员蛋白互作预测

利用 STRING 数据库 (<https://string-db.org/>)对

PcUGT 蛋白序列进行了互作预测。输入所有 PcUGT 蛋白的氨基酸序列,互作网络构建参数中,选择中等置信度 (≥ 0.4),并启用“实验验证”“数据库注释”“共表达”等多个证据来源;网络图绘制及后续分析使用 STRING 自带的分析工具。

在获得的互作网络基础上,进一步使用 Markov Cluster Algorithm (MCL)进行聚类分析,通胀参数设置为默认值 (3.0),用于识别可能存在的功能模块或蛋白复合体。同时对网络的平均节点度、聚类系数等拓扑属性进行了计算,以评估整体网络结构特征。

2.9 PcUGT 基因家族表达量差异分析

为研究 PcUGT 家族中参与白藜芦醇、虎杖苷、大黄素及大黄素-8-O-葡萄糖苷生物合成的关键酶基因,结合转录组与代谢组数据筛选出 44 个 *PcUGTs* 并绘制表达量分析热图,选取 8 个与 2 类物质合成高度相关的关键基因 (*PcUGT060*、*PcUGT065*、*PcUGT097*、*PcUGT106*、*PcUGT130*、*PcUGT132*、*PcUGT134* 和 *PcUGT137*),并采用实时荧光定量 PCR (qRT-PCR)方法对其在 RC、RL、RhC、RhL 和 RhM 5 个部位的表达量进行检测。

参照总 RNA 提取试剂盒提取虎杖各部位组织样品的 RNA,利用反转录试剂盒合成 cDNA,于 $-20\text{ }^{\circ}\text{C}$ 冰箱中保存备用。用 Primer Premier 5.0 软件和 DNAMAN 6.0 设计 8 个 *PcUGTs* 的 qRT-PCR 引物 (表 1),按照试剂盒说明书进行荧光定量 PCR,分析 PcUGT 家族成员的组织表达模式,以 EF-1 α 为内参基因^[19],利用 $2^{-\Delta\Delta Ct}$ 法计算每个基因的相对表达量。每组设置 3 个生物重复,表达量均为 3 次重复的平均值,使用 *t* 检验比较样本平均数差异是否具有统计学意义 ($P < 0.05$)。

3 结果与分析

3.1 PcUGT 家族成员鉴定及理化性质分析

共筛选出 178 条 *UGT* 基因序列 ($E\text{-value} < 1 \times 10^{-5}$),并依次对 *PcUGT* 进行编号,从 *PcUGT001*~*PcUGT178*,对其氨基酸理化性质及亚细胞定位分析发现,PcUGT 基因家族氨基酸数目为 69~987 aa,平均氨基酸数为 422.07 个,蛋白质相对分子质量为 7 920~110 380,理论等电点 (PI) 为 4.66~9.77,其中有 16 个蛋白等电点大于 7,其余 162 个蛋白均小于 7,偏酸性;不稳定系数为 12.93~69.71,脂肪族氨基酸指数在 57.55~113.27,亲水性总平均值为 $-0.452\sim 0.759$,有 40 个蛋白的疏水性大于 0,其余

表 1 qRT-PCR 引物序列

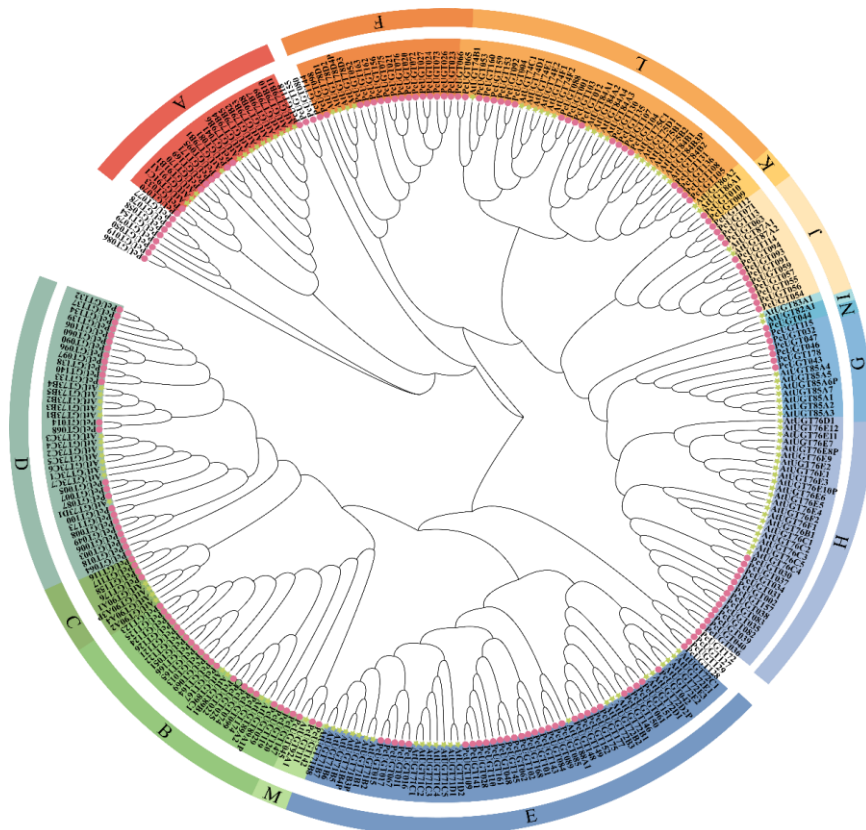
Table 1 qRT-PCR primer sequences

引物	序列 (5'-3')
<i>PcUGT060-F</i>	GGGCTTGTAATAAAAGGGTGG
<i>PcUGT060-R</i>	TGAAAGAAGAACCCCCTTCTTC
<i>PcUGT065-F</i>	CACAACTTCCCAATGACTTCC
<i>PcUGT065-R</i>	CTCTTAATCTCCTCTCCACC
<i>PcUGT097-F</i>	GGTGTTCGAGGCGGAGAAGAAG
<i>PcUGT097-R</i>	GCGGAGTTGGACGGTAGAGATG
<i>PcUGT106-F</i>	GCCAATGTTCGCCGAGCAATTC
<i>PcUGT106-R</i>	GCAGCACTCCTCCTCATCTCCT
<i>PcUGT130-F</i>	GGCGACCACGAGCACAAAGTT
<i>PcUGT130-R</i>	TCTCCTCCACGAACTCCTCCAA
<i>PcUGT132-F</i>	GGAGCGGATCAGAAGTGCATT
<i>PcUGT132-R</i>	TGAACAGCCTTCCAAGCCAAC
<i>PcUGT134-F</i>	GACGAACATTGGTTGCCGAGTG
<i>PcUGT134-R</i>	GCTGCTCTGCGAACATTGGC
<i>PcUGT137-F</i>	CCTTCTCCGAATCCGACAATGTTG
<i>PcUGT137-R</i>	GCGATCAGCTTCAGGGCTTTTATG
<i>EF-1α-F</i>	GTCATCCCTGATTGATTACGC
<i>EF-1α-R</i>	GTGGGCAATAAAGCCAAGAC

138 个蛋白都是亲水性蛋白。亚细胞定位预测分析发现,多数 UGT 蛋白 (90/178, 50.56%) 定位于叶绿体,其次是细胞质 (50/178, 28.10%),少数定位于细胞核、内质网、质膜等部位。

3.2 *PcUGT* 家族成员系统发育及种间同源性分析

为了解 *PcUGT* 基因家族成员的同源进化关系,对虎杖中 178 个 *PcUGT* 基因成员与拟南芥 117 个 *AtUGT* 基因成员构建系统进化树 (图 1)。参考 *AtUGT* 基因家族分类依据,对 *PcUGT* 基因进行分类,结果显示,大部分 *PcUGT* 基因能与拟南芥既有的分类体系进行归类,但同时发现有 3 个独立的分支未能与拟南芥中任何已知类别聚类,这些差异可能反映了虎杖与拟南芥在 UGT 家族基因的进化过程中存独立的演化路径,预示这 3 个分支可能具有不同的功能特性或调控机制。虎杖与拟南芥共线性分析结果显示与系统发育树预测一致 (图 2),表明 UGT 基因在不同物种间具有保守性。



○-虎杖; ☆-拟南芥; A~N-表示 UGT 基因不同亚家族。

○-*P. cuspidatum*; ☆-*Arabidopsis thaliana*; A~N-represent different subgroups of the UGT.

图 1 拟南芥和虎杖 UGTs 系统发育进化树

Fig. 1 Phylogenetic tree of UGTs in *A. thaliana* and *P. cuspidatum*

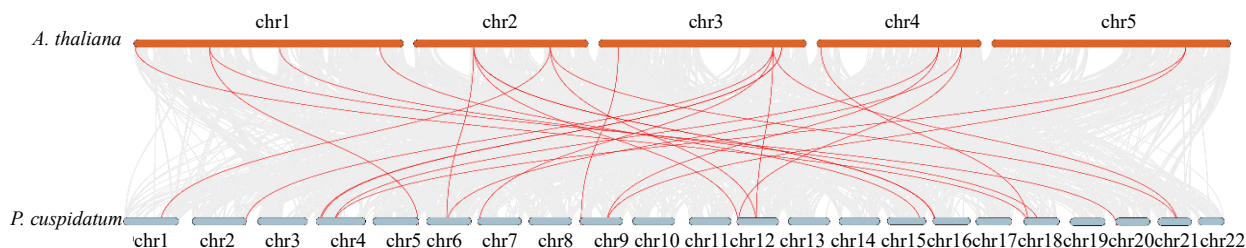


图 2 虎杖与拟南芥基因组间 UGT 家族基因共线性分析

Fig. 2 Collinear prediction of UGT family genes between *A. thaliana* and *P. cuspidatum*

3.3 PcUGT 基因家族的染色体分布、基因结构和基因复制分析

PcUGT 基因家族在其 22 条染色体上的分布不均衡 (图 3), Chr3、Chr4、Chr15、Chr16 为 UGT 基因富集的染色体。Chr16 中 UGT 基因数量最多 (17 个), 其次是 Chr3 (13 个)、Chr4 (13 个)、Chr15 (12 个), Chr17 只包含了 1 个 UGT 基因, Chr15、Chr16 上定位的基因密集且呈簇状排列, 整体来看, PcUGT 基因覆盖于 22 条染色体, 在某些特定的染

色体上存在着明显的聚集趋势。进一步观察系统发育亚群在染色体上的分布发现, 多数 UGT 基因呈簇状排列, 同一亚群成员倾向于分布在同一染色体或相邻基因区域, 例如在 Chr3、Chr6、Chr16 上的簇集中, 表明这些基因可能存在串联重复。部分亚群成员分布较为分散, 如 E 组成员分别定位于 Chr2、Chr13 等染色体, H 组成员分别定位于 Chr1、Chr3、Chr9 等染色体, 推测可能源于较早期的复制或基因重排事件。

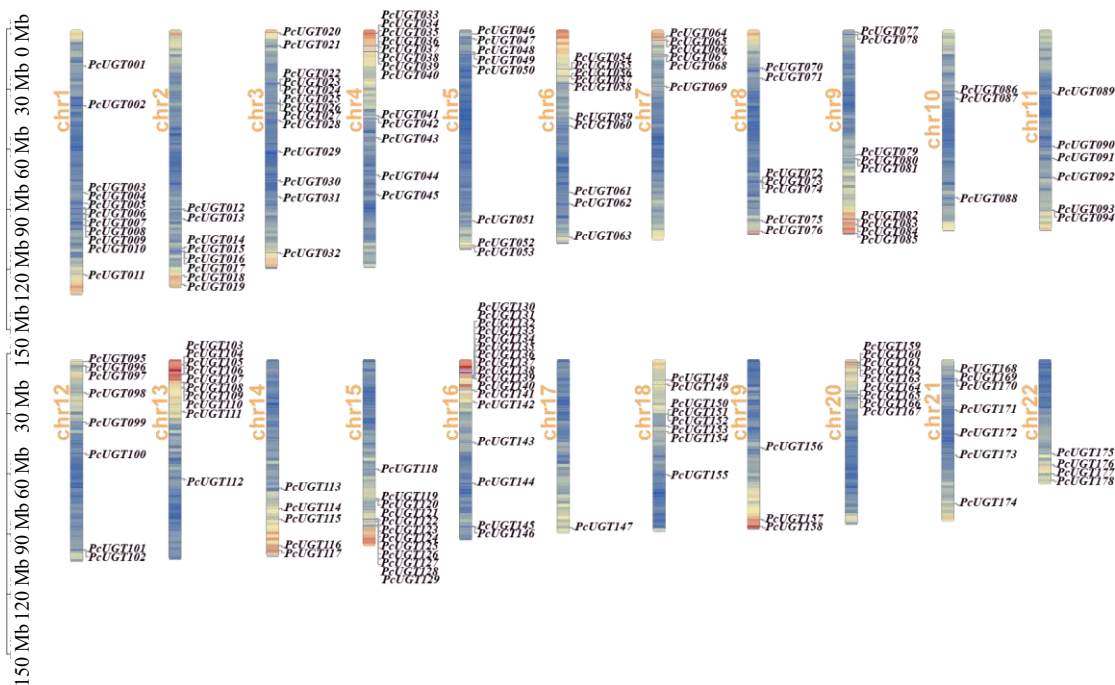


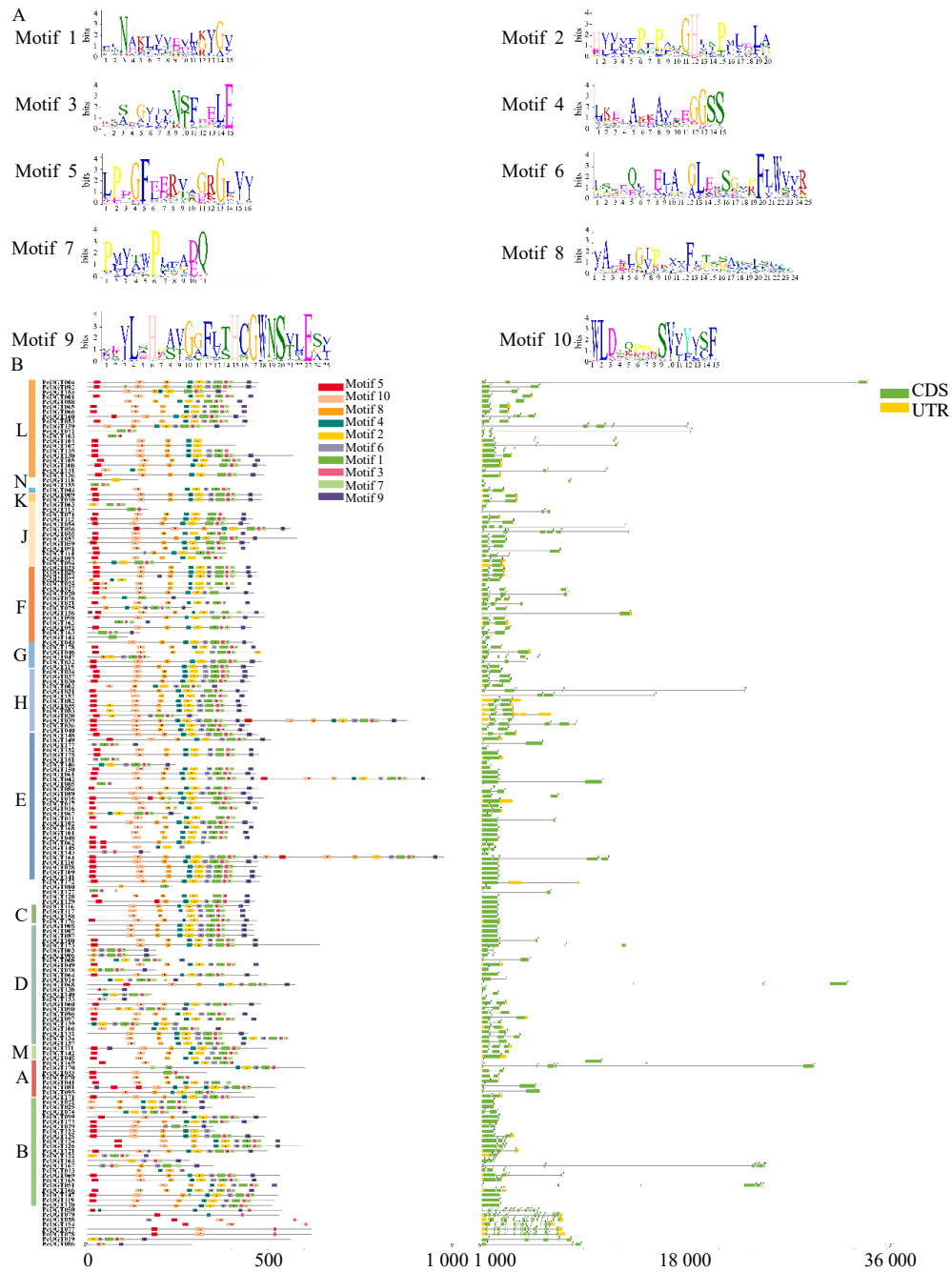
图 3 PcUGT 基因家族成员在染色体上的定位与串联重复

Fig. 3 Chromosomal location and tandem duplication of PcUGT gene family members

3.4 PcUGT 基因家族的基因结构和保守基序分析

在 PcUGT 蛋白中鉴定了 10 个保守基序, Motif7 和 Motif9 构成了虎杖糖基转移酶 PSPG box 的主要保守基序^[20] (图 4-A)。有 151 个 PcUGT 的蛋白序列在 C 端拥有 Motif7 或 Motif9, 具有典型的糖基

转移功能。14 个 PcUGT 的蛋白序列在 C 端拥有 Motif3, 6 个 PcUGT 在 C 端拥有 Motif2, 2 个 PcUGT 在 C 端拥有 Motif1, 5 个 PcUGT 在 C 端拥有 Motif6 (图 4-B)。多数 PcUGT 蛋白均包含 5~10 个 Motif, 表明该家族在结构上具有较高的保守性。内含子和



A-PcUGT 家族的保守基序; B-PcUGT 基因家族结构域分析及基因结构。

A-conserved motifs of PcUGT gene family; B-analysis of the domain structure and gene structure of the PcUGT gene family.

图 4 PcUGT 家族保守基序与结构域分布图

Fig. 4 Conserved motif and domain distribution map of PcUGT family

外显子结构的多样性通常在基因家族的进化中发挥关键作用，并为支持系统发育类群提供了额外的证据，基因结构分析发现，各成员之间基因结构相似，157 个成员含有 1~3 个外显子，113 个成员不含内含子，H 亚群的成员多数含有 2~4 个内含子。整体来看，PcUGT 家族的基因结构在同一亚群中相

对一致，而在不同亚群间呈现出显著的差异。这种结构上的多样性不仅为系统发育关系提供了支持证据，也为推测不同亚群基因的功能演化方向提供了线索。

3.5 PcUGT 基因家族启动子顺式作用元件分析

将预测的顺式作用原件归为 4 种类型：发育响

应 (development responsive), 压力响应 (stress responsive), 激素响应 (hormone responsive) 和光响应 (light responsive) [21], 大多数 UGT 基因启动子区域都具有这 4 种类别的顺式作用元件 (图 5-A)。光响应元件在 UGT 基因家族成员的启动子区域是最为普遍的, 共有 2 772 个, 推测光在诱导 *PcUGTs* 表达中发挥重要作用。激素响应元件的数量有 1 893 个, 压力响应元件的数量有 932 个, 发育响应元件的数量 545 个 (图 5-B)。通过对启动子区序列的分析, 可以更好地了解基因的调控机制, 同时对 *PcUGTs* 启动子区调控元件的预测, 也有助于对该基因的功能进行更深层次的研究。

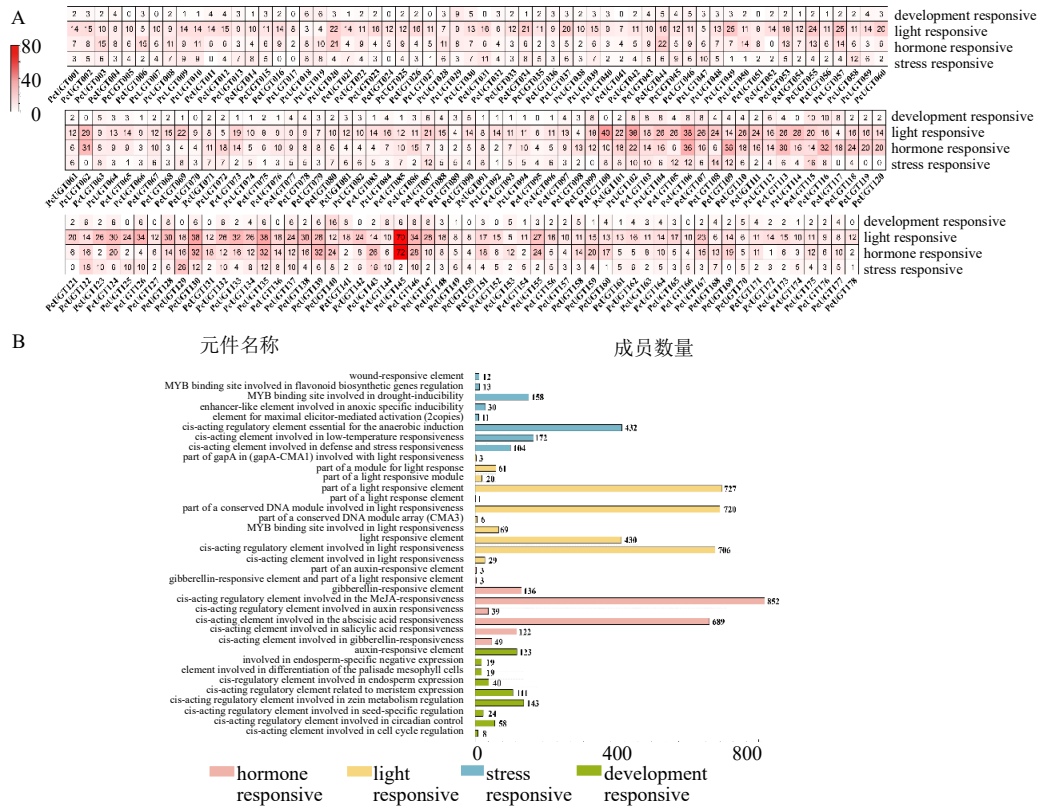
3.6 PcUGT 成员蛋白互作预测

PcUGTs 家族各成员之间的蛋白互作网络见图 6, 互作的平均节点度为 3.4, 有 50 个 PcUGT 蛋白参与互作, 但 7 个蛋白断开连接, 只有 43 个 PcUGT 蛋白之间建立互作关系, 其中 PcUGT099~PcUGT079 互作得分最高, 达到 0.784, 其次是 PcUGT086~PcUGT079、PcUGT079~PcUGT044 都

为 0.784; 互作得分最低的为 PcUGT154、PcUGT174 和 PcUGT176, 只有 0.400。同时按照 MCL 聚类, 平均聚类系数为 0.306。

3.7 PcUGT 基因家族的表达模式分析

虎杖中的白藜芦醇、虎杖苷、大黄素及大黄素-8-O-葡萄糖苷具有重要的生物活性, 为进一步探究 PcUGT 基因家族在黄酮苷合成途径中的调控作用, 结合转录组与代谢组数据进行了综合筛选与分析, 挖掘 44 个潜在关键功能基因, 并绘制表达量相关性热图 (图 7)。由图可见, 虎杖苷含量同基因 *PcUGT056*、*PcUGT060*、*PcUGT079*、*PcUGT089*、*PcUGT130* 相关性较强; 白藜芦醇含量同基因 *PcUGT039*、*PcUGT056*、*PcUGT060*、*PcUGT073*、*PcUGT093*、*PcUGT106* 相关性较强; 大黄素-8-O-葡萄糖苷含量同基因 *PcUGT060* 相关性较强; 大黄素含量同基因 *PcUGT016*、*PcUGT031*、*PcUGT039*、*PcUGT051*、*PcUGT056*、*PcUGT060*、*PcUGT065*、*PcUGT073*、*PcUGT093*、*PcUGT143* 相关性较强。



A-PcUGT 基因家族启动子区域的顺式作用元件热图, B-PcUGT 启动子顺式作用元件分布。

A-heatmap of cis-acting elements in the promoter region of PcUGT gene family, B-distribution of cis-acting elements of PcUGT promoters.

图 5 PcUGT 基因家族顺式作用元件图谱

Fig. 5 Cis-acting element map of PcUGT gene family

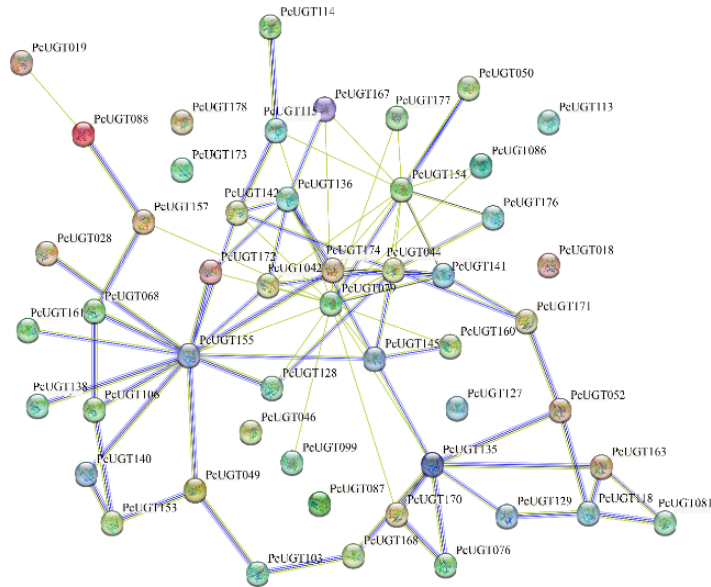


图6 PcUGT 家族成员蛋白互作网络

Fig. 6 Protein interactions network among members of PcUGT family

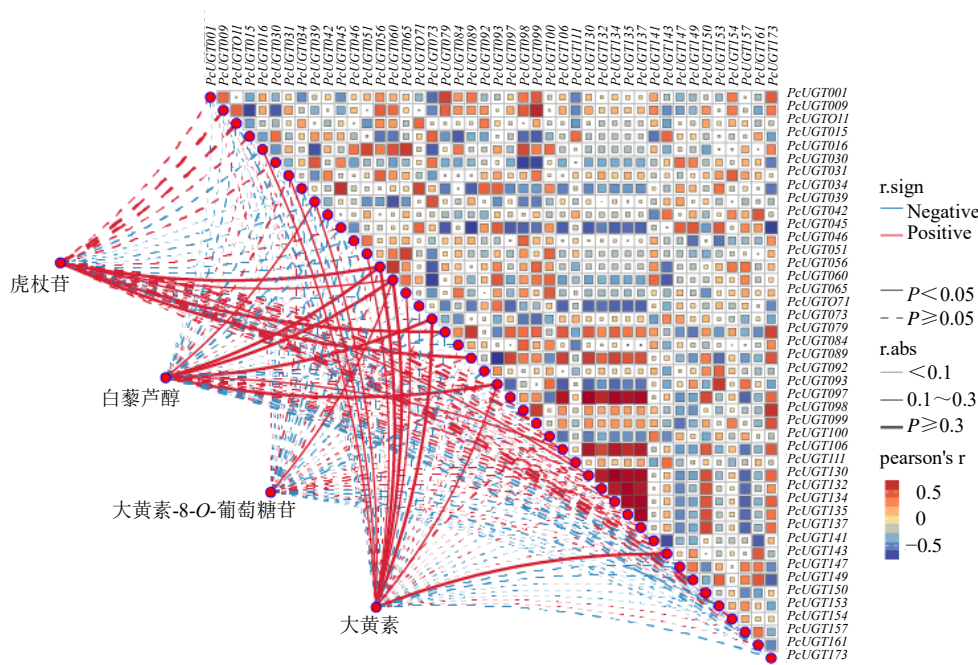


图7 部分 PcUGT 基因在不同部位中的表达水平的相关性及其与 4 种成分含量的相关性

Fig. 7 Correlation of expression levels of some PcUGT genes in different parts and their correlation with contents of four components

已有报道指出 AtUGT84 亚家族与白藜芦醇及大黄素糖基化高度相关^[22]，在源性分析中 PcUGT065、PcUGT130、PcUGT135 等 UGT 基因与 AtUGT84 同属于一个亚家族，结合相关性分析的结果，提示 PcUGT065、PcUGT130 可能具有白藜芦醇与大黄素糖基化的功能。此外，对 PcUGT130 在内的 8 个 PcUGT 基因进行 qRT-PCR 验证（图 8），

通过相对表达量得出，有 6 个基因在不同部位间的表达存在显著差异，其中，PcUGT060 在 RC 部位表达高于 RhM ($P < 0.01$)；PcUGT065 在 RC 部位表达显著高于 RL 和 RhL ($P < 0.001$)；PcUGT097 在 RL 部位表达显著高于 RhM ($P < 0.05$)；PcUGT106 在 RC 部位的表达显著高于 RL 和 RhL ($P < 0.05, 0.01$)；PcUGT132 在 RL 部位表达显著

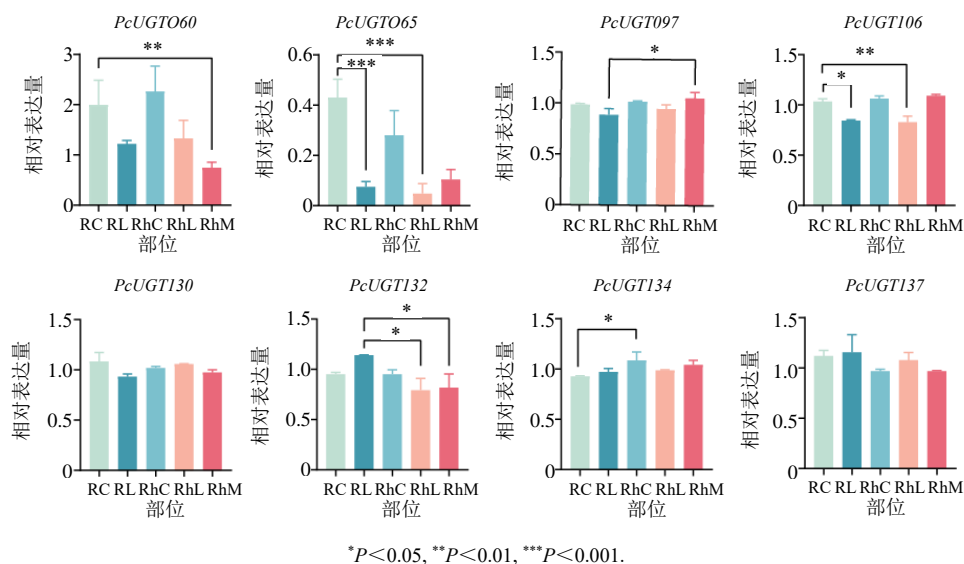


图 8 8 个 *PcUGT* 基因在不同部位的表达水平

Fig. 8 Expression levels of eight *PcUGT* genes in different parts

高于 RhL 和 RhM ($P < 0.05$); *PcUGT134* 在 RhC 部位表达最高, RC 显著降低 ($P < 0.05$)。相比之下, *PcUGT130* 和 *PcUGT137* 在各部位间的表达差异不显著。这展现了与转录组一致的趋势, 这些结果验证了转录组数据的可靠性, 同时也表明 *PcUGT065*、*PcUGT130* 很可能是参与虎杖中白藜芦醇与大黄素糖苷类物质生物合成的关键酶基因。

4 讨论

UGT 基因家族作为调控植物次生代谢网络的关键节点, 在多酚、黄酮、蒽醌等活性成分的糖基化修饰中发挥核心作用^[12, 23]。关于糖基化修饰的研究, 主要集中在黄酮类化合物, 在苜蓿^[24]、金柑^[25]等植物中均有相关报道, 这些发现为进一步揭示虎杖活性成分, 尤其是白藜芦醇和大黄素类物质的糖基化修饰机制提供了重要的分子基础。

随着基因组学等多组学技术的发展, UGT 基因已在多种植物中得到鉴定和研究, 如大豆^[26]、水稻^[27]、石榴^[28]等。本研究首次从全基因组水平系统鉴定并解析了虎杖 178 个 *PcUGT* 基因, 相较于石榴 (32 个)、大豆 (215 个) 与水稻 (194 个), *PcUGT* 基因家族成员数量适中, 表明其在进化过程中既保持了一定的保守性, 也可能存在特异性的扩张。本研究构建了虎杖和拟南芥的 UGT 基因系统进化树, 将 *PcUGT* 基因家族分为 14 个亚族 (A~N), 在家族成员数量上虎杖与拟南芥等植物的 UGT 基因亚族上有类似规律^[11]。在 178 个 *PcUGTs* 中, 有 120 个含有完整的 PSPG 保守结构域, 表明其具备糖基

转移酶活性。同时, 不同亚类之间在基因结构和保守基序组合上存在差异, 提示这些基因可能在功能上发生了分化。因此, 从全基因组水平解析 *PcUGT* 基因家族, 对于揭示其药效成分的形成机制具有重要意义。

基因复制事件分析结果表明, 串联重复是 *PcUGT* 基因家族扩张的主要驱动力, 染色体定位显示 178 个 *PcUGT* 基因成簇分布, 不均匀的定位在 22 条染色体上, 这一特征与其他物种表现相似^[29]。*PcUGT* 蛋白的理化性质与亚细胞定位预测结果显示, 多数 *PcUGT* 蛋白定位于叶绿体, 表明叶绿体可能是次生代谢物空间分布与积累的重要场所, 通过进一步的顺式作用元件分析表明, *PcUGT* 基因的表达受到多种调控元件的影响, 其中光响应元件最为普遍, 因此可以推测, 光信号可能是影响 *PcUGT* 基因转录调控的主要因素, 这与其在叶绿体中的亚细胞定位预测结果也相符合, 这一结果与葡萄、水稻等物种的研究结论相一致^[30]。

白藜芦醇、虎杖苷、大黄素及大黄素-8-O-葡萄糖苷是虎杖药材质量评价的重要指标, 分别属于苯丙素类和蒽醌类 2 条关键代谢途径, 能够充分代表虎杖主要活性成分的代谢特征。由于糖基化在这些活性成分的稳定性、溶解性及生物活性形成中发挥关键作用, 解析 UGT 基因与这些化合物合成的关系对于理解虎杖药效成分的代谢调控具有重要意义。转录组与代谢组联合分析是用来解析生理机制、寻找关键基因的有效方法, 根据已报道的 UGT

家族功能分析, *AtUGT84* 被证实参与白藜芦醇与大黄素生物合成^[22], 本研究通过相关性热图分析筛选得到 *PcUGT065*、*PcUGT130* 等 14 个与 4 个成分合成相关的基因, 且发现 *PcUGT065*、*PcUGT130* 与 *AtUGT84* 同属于一个亚家族, 可能高度参与白藜芦醇、大黄素及其糖苷类物质合成, 这些候选基因的 qRT-PCR 表达趋势与相应代谢物的积累模式高度一致, 进一步增强了其在虎杖药效成分生物合成中潜在功能的可信度。结合基因结构、保守基序以及蛋白互作网络的信息, 这些候选 *UGT* 基因极有可能作为虎杖苯丙素类及蒽醌类活性物质糖苷化途径中的关键酶基因。上述证据从多个层面共同支持了这些 *UGT* 基因的潜在功能, 不仅为阐明虎杖药效物质的形成机制提供了新的分子证据, 也为未来利用合成生物学和代谢工程手段提高药材品质、实现药效成分的高效积累奠定了基础。

综上所述, 本研究展示了 *PcUGT* 家族的图谱, 揭示其结构多样性、进化保守性与调控潜力, 并提出与白藜芦醇、虎杖苷、大黄素及大黄素-8-*O*-葡萄糖苷密切相关的候选 *UGT* 基因群。本研究将虎杖的药效基础与基因表达相结合, 对推动虎杖及相关药材的药效物质挖掘、品质提升及新药开发具有重要意义。

利益冲突 所有作者均声明不存在利益冲突

参考文献

- [1] 中国药典 [S]. 二部. 2025: 257.
- [2] Kuo C H, Chen B Y, Liu Y C, *et al.* Production of resveratrol by piceid deglycosylation using cellulase [J]. *Catalysts*, 2016, 6(3): 78-89.
- [3] Wu F, Wang Y, Mei Q X, *et al.* UGTs-mediated metabolic interactions contribute to enhanced anti-inflammation activity of Jinhongtang [J]. *J Ethnopharmacol*, 2023, 304: 116016.
- [4] Tan D P, Wang J M, Cao L G, *et al.* UDP-glycosyltransferases play a crucial role in the accumulation of alkaloids and sesquiterpene glycosides in *Dendrobium nobile* [J]. *Arab J Chem*, 2023, 16(5): 104673.
- [5] 陈雨彤, 张倩, 陈静, 等. 知母糖基转移酶基因 *AaUGT2* 克隆、表达及特性研究 [J]. *中国生物化学与分子生物学报*, 2025, 41(12): 1850-1859.
- [6] Gervay J, Nguyen T N, Hadd M J. Mechanistic studies on the stereoselective formation of glycosyl iodides: First characterization of β -*d*-glycosyl iodides [J]. *Carbohydr Res*, 1997, 300(2): 119-125.
- [7] Yang S D, An X N, Gu G F, *et al.* Novel dTDP-l-rhamnose synthetic enzymes (RmlABCD) From *Saccharothrix syringae* CGMCC 4.1716 for one-pot four-enzyme synthesis of dTDP-l-rhamnose [J]. *Front Microbiol*, 2021, 12: 772839.
- [8] Wang C, Qi R P, Xu Z Q. Glycosyl radical-based synthesis of C-glycoamino acids and C-glycopeptides [J]. *Chemistry*, 2023, 29(21): e202203689.
- [9] Kitamura K, Shigeta M, Maezawa Y, *et al.* Preparation of *L*-vancosamine-related glycosyl donors [J]. *J Antibiot*, 2013, 66(3): 131-139.
- [10] Fan W J, Han P H, Feng Q Y, *et al.* Nucleic acid substrate-independent DNA polymerization on the exosome membrane: A mechanism study and application in exosome analysis [J]. *Anal Chem*, 2022, 94(4): 2172-2179.
- [11] Yang J F, Ma L, Jiang W B, *et al.* Comprehensive identification and characterization of abiotic stress and hormone responsive glycosyl hydrolase family 1 genes in *Medicago truncatula* [J]. *Plant Physiol Biochem*, 2021, 158: 21-33.
- [12] Ma H P, Chen J H, Lin F, *et al.* Functional characterization of *PcUGT73BF6* from *Polygonum Cuspidatum* and the facilitation of emodin catalysis via site-directed mutagenesis [J]. *J Agric Food Chem*, 2025, 73(22): 13540-13554.
- [13] 狄少康, 尹青岗, 夏亚迎, 等. 大豆类黄酮糖基转移酶基因 *UGT73C19* 的功能研究 [J]. *中国农业科学*, 2019, 52(20): 3507-3519.
- [14] 焦阳, 马英, 张振, 等. 乙烯利对苦瓜幼苗内源激素代谢及相关调控基因表达的影响 [J]. *西北植物学报*, 2022, 42(7): 1180-1188.
- [15] Yang F, Zhang L, Zhang X, *et al.* Genome-wide investigation of UDP-glycosyltransferase family in Tartary buckwheat (*Fagopyrum tataricum*) [J]. *BMC Plant Biol*, 2024, 24(1): 249.
- [16] 魏永赞. 葡萄白藜芦醇糖基转移酶功能鉴定及转录调控机制解析 [D]. 北京: 中国科学院大学, 2021.
- [17] Chen C J, Chen H, Zhang Y, *et al.* TBtools: An integrative toolkit developed for interactive analyses of big biological data [J]. *Mol Plant*, 2020, 13(8): 1194-1202.
- [18] 张苛苛, 谭宇萍, 徐欢欢, 等. 菘蓝 4-香豆酸: 辅酶 A 连接酶全基因家族的鉴定及表达分析 [J]. *中国中药杂志*, 2024, 49(2): 361-369.
- [19] Wu Q W, Wei M, Feng L F, *et al.* Rhamnosyltransferases involved in the biosynthesis of flavone rutinosides in *Chrysanthemum* species [J]. *Plant Physiol*, 2022, 190(4): 2122-2136.
- [20] Hoffmann T D, Kurze E, Liao J R, *et al.* Genome-wide identification of UDP-glycosyltransferases in the tea plant (*Camellia sinensis*) and their biochemical and

- physiological functions [J]. *Front Plant Sci*, 2023, 14: 1191625.
- [21] Liu X J, Gong Q, Zhao C N, *et al.* Genome-wide analysis of cytochrome P450 genes in *Citrus clementina* and characterization of a CYP gene encoding flavonoid 3'-hydroxylase [J]. *Hortic Res*, 2022, 10(2): uhac283.
- [22] Lee C, Hong W J, Jung K H, *et al.* *Arachis hypogaea* resveratrol synthase 3 alters the expression pattern of UDP-glycosyltransferase genes in developing rice seeds [J]. *PLoS One*, 2021, 16(1): e0245446.
- [23] 邹妍, 王力, 王梅, 等. 分子结构对辛烯基琥珀酸淀粉酯荷载多酚行为影响的研究 [J]. *食品研究与开发*, 2019, 40(10): 50-55.
- [24] 孙志康, 王娜, 孟颖颖, 等. 蒺藜苜蓿糖基转移酶基因 SMALL AND EMERALD1 的克隆和功能研究 [J]. *生物技术进展*, 2021, 11(2): 182-189.
- [25] 谭嫣, 马小迪, 王久照, 等. 基于转录组测序的金柑类黄酮糖基转移酶基因的初步分析 [J]. *园艺学报*, 2019, 46(4): 664-676.
- [26] Yin Q G, Shen G A, Di S K, *et al.* Genome-wide identification and functional characterization of UDP-glycosyltransferase genes involved in flavonoid biosynthesis in *Glycine max* [J]. *Plant Cell Physiol*, 2017, 58(9): 1558-1572.
- [27] Ko J H, Kim B G, Kim J H, *et al.* Four glucosyltransferases from rice: cDNA cloning, expression, and characterization [J]. *J Plant Physiol*, 2008, 165(4): 435-444.
- [28] 徐会丽, 张太奎, 苑兆和. 石榴 UGT 基因家族的比较和进化分析 [J]. *经济林研究*, 2017, 35(4): 130-135.
- [29] Zhao X Q, Feng Y Y, Ke D, *et al.* Molecular identification and characterization of UDP-glycosyltransferase (UGT) multigene family in pomegranate [J]. *Horticulturae*, 2023, 9(5): 540.
- [30] 王媛媛, 刘云清, 徐晶宇, 等. 葡萄 UDP-糖基转移酶家族的全基因组表征、进化和表达分析 [J]. *食品科学*, 2025, 46(14): 134-146.

[责任编辑 时圣明]