

地黄 UDP-鼠李糖:鼠李糖基转移酶基因家族的生物信息学和表达分析

周延清,李慧敏

(河南师范大学 生命科学学院,河南 新乡 453007)

摘要:尿苷二磷酸鼠李糖基转移酶能将活性鼠李糖基供体转移到其糖基化受体分子上,合成鼠李糖基化的化合物.它们结构多样,活性广泛,具有药理、维持细胞结构、信号转导和化学防御及调节植物体内激素水平等重要作用.迄今,已从多种植物中分离得到这类酶,但有关其基因家族的结构特性与在植物次生代谢产物毛蕊花糖苷生物合成中的作用研究较少.以地黄尿苷二磷酸鼠李糖基转移酶基因(*RgURT*)家族的4个成员 *RgURT1*~*RgURT4* 为材料,利用生物信息学对该家族的特征进行了分析,比较了 *RgURT* 在高、低毛蕊花糖苷地黄品种块根中的表达模式.结果表明,其 cDNA 大小为 645~1 422 bp;编码蛋白质分子量为 24.05~53.24 kDa,由 214~473 个氨基酸残基组成,属于糖基转移酶 GT-B 型超家族,C 端具有保守结构域盒(PSPG 盒),无信号肽,亚细胞定位在细胞质中,二级结构以无规则卷曲和 α -螺旋为主;*RgURT4* 与 *RgURT1*~*RgURT3* 分别聚类为 2 个簇,*RgURT4* 蛋白与其他 3 个 URT 蛋白序列之间的同源性低;RT-qPCR 分析表明,*RgURT* 基因在高毛蕊花糖苷地黄块根中的表达量高于低毛蕊花糖苷地黄块根中的表达量.研究首次揭示了地黄 *RgURT* 家族的特征,初步验证了 *RgURT* 与地黄毛蕊花糖苷合成相关,将为进一步研究毛蕊花糖苷合成提供了重要酶基因.

关键词:鼠李糖基转移酶;地黄;*RgURT* 家族的特征;PSPG 盒

中图分类号:Q94

文献标志码:A

文章编号:1000-2367(2024)06-0128-07

地黄是玄参科地黄属一种多年生草本药用植物,主要分布于中国、日本和韩国等国家.其根部(鲜地黄、生地黄和熟地黄)可用作蔬菜和药材,具有重要的药食价值和经济价值.随着不同组学技术与分子生物学技术在地黄研究中的广泛应用,大量的地黄基因被挖掘与报道^[1-3],越来越多地黄基因的未知功能有待解析,如地黄尿苷二磷酸鼠李糖基转移酶(UDP-rhamnose:rhamnosyltransferase)基因.

尿苷二磷酸鼠李糖基转移酶普遍存在于生物体中,将活性鼠李糖基供体转移到其糖基化受体分子上.含鼠李糖基取代的化合物结构多样,药理活性广泛,具有维持细胞结构、信号转导和化学防御及调节植物体内激素水平等重要作用.已有研究从多种植物中鉴定出鼠李糖基转移酶^[4-5].如,金荞麦 FtUGT79A15 具有催化槲皮素 3-O-葡萄糖苷转化为芦丁的鼠李糖基转移酶活性^[6];甘草皂素 2'-鼠李糖基转移酶 GuRhaGT 催化萜皂苷的 C-3 糖基配体的 2'-羟基的鼠李糖基化^[7];梁平柚鼠李糖基转移酶基因 *Cm1*, *2RhaT* 的时空表达与其黄酮酮的含量相关^[8];拟南芥黄酮醇 7-O-鼠李糖基转移酶基因 *AtUGT78D1* (*F3RhaT*)使类黄酮鼠李糖化^[9];拟巫山淫羊藿鼠李糖基转移酶基因 (*EpPF3RT*)使异戊基类黄酮鼠李糖化^[10];甜荞鼠李糖基转移酶基因 (*FeF3G6RhaT*)和桑树鼠李糖基转移酶基因参与芦丁生物合成^[11-12];拟南芥鼠李糖基转移酶 UGT89C1 (*F7RhaT*)具有转移鼠李糖基的作用^[13];半边莲的依赖尿苷二磷酸-鼠李糖的鼠李糖转移酶使花青素鼠李糖

收稿日期:2022-12-12;**修回日期:**2023-02-27.

基金项目:国家自然科学基金(31870312);河南省重点研发与推广专项(212102110405);河南省自然科学基金(182300410018);河南省高校科技创新团队支持计划项目(23IRTSTHN022).

作者简介(通信作者):周延清(1963-),男,河南邓州人,河南师范大学教授,博士,研究方向为植物遗传,E-mail:yqzhou@htu.edu.cn.

引用本文:周延清,李慧敏.地黄 UDP-鼠李糖:鼠李糖基转移酶基因家族的生物信息学和表达分析[J].河南师范大学学报(自然科学版),2024,52(6):128-134.(Zhou Yanqing, Li Huimin. Bioinformatics and expression analyses of UDP-rhamnose:rhamnosyltransferase family in *Rehmannia glutinosa* [J]. Journal of Henan Normal University (Natural Science Edition), 2024, 52(6): 128-134. DOI: 10.16366/j.cnki.1000-2367.2022.12.12.0001.)

化^[14].研究发现,许多天然产物的糖基化取代都表现出了显著的位置特异性,如,苯乙醇苷类化合物的结构具有结构多样性,但其分子中鼠李糖基取代反应都特异性地发生在葡萄糖 3-OH 位置,表明鼠李糖基转移酶催化鼠李糖基转移酶取代反应具有高度的位置特异性^[5].迄今,对植物 UDP-鼠李糖基转移酶研究发现,其主要修饰萜类和黄酮类分子,序列相似性比较低,C-端具有 PSPG 盒子,没有发挥催化功能的共同机制.因此,植物 UDP-鼠李糖基转移酶的挖掘与应用受到很大限制^[4].本团队在前期研究毛蕊花糖苷生物合成途径与分子机理过程中,推测出其合成过程中的鼠李糖化反应需要鼠李糖基转移酶,并且用三代转录组测序技术挖掘了一个地黄 *RgURT* 家族,包括 *RgURT1*~*RgURT4* (基因注册号:MN646769、MT767427、MT767428 和 MT767429),但是没有对该基因家族的特征与合成毛蕊花糖苷功能的报道^[15],以高、低毛蕊花糖苷地黄品种的块根为材料进行该酶基因的基因型依赖性表达分析,为进一步研究毛蕊花糖苷合成提供了重要酶基因,对于完全解析毛蕊花糖苷合成分子调控机制和促进毛蕊花糖苷的合成生物学研究以及分子辅助育种的分子标记研发具有重要意义.

1 材料与方法

1.1 基因家族成员的碱基序列

地黄全长转录组测序与数据见 NCBI SRA Database (accession number: PRJNA780233) <https://trace.ncbi.nlm.nih.gov/Traces/sra/?study=SRP346035>.鉴定出 *RgURT* 基因家族 4 个成员 *RgURT1*~*RgURT4*.

1.2 生物信息学分析

使用 NCBI 在线分析工具 ORF Finder 查找这些基因完整的开放阅读框(ORF);用 ProtParam 对这些基因编码的蛋白质进行理化性质分析;采用 TMHMM 2.0 和 NetPhos 3.1 Server 预测分析这些蛋白质的跨膜区和磷酸位点;应用 Signal P-5.0 预测分析这些蛋白质的氨基酸序列中的信号肽;用 SOPMA、SWISS-MODEL、NCBI 的 CD search 和 InterProScan 分别预测分析这些蛋白质的二级和三级结构及其结构域;利用在线工具 Blastn 查找与这些蛋白质同源性高的其他物种的蛋白序列;运用软件 DNAMAN 6.0 和 MEGA 6.0 软件对其进行同源性分析和系统进化树构建^[16].

1.3 地黄 *RgURT* 基因表达的荧光定量 PCR 测定

根据文献^[17],选定地黄品种北京 3 号和温 85-5 作为高、低毛蕊花糖苷地黄品种,其毛蕊花糖苷质量分数依次为 $0.53 \text{ mg} \cdot \text{g}^{-1}$ 和 $0.28 \text{ mg} \cdot \text{g}^{-1}$;从武陟县四大怀药研究所采挖其成熟期地黄植株各 10 株,取其块根,去除泥沙,蒸馏水冲洗干净后,用滤纸吸干表面,按品种分别取其新鲜块根,切成约 3 mm 厚的薄片,按品种不同株的切片混合,液氮研磨成粉末,用于提取 RNA,以地黄 *RgTIP41* 作为内参基因^[1],以上述 RNA 为模板,用附录表 S1 中引物对 *RgURT1*、*RgURT2* 与 *RgURT4* 进行基因型依赖性表达的 qRT-PCR 分析,引物由英潍捷基贸易有限公司合成.

1.4 统计分析

使用 Microsoft Excel 2010 进行基因在高、低毛蕊花糖苷含量的地黄品种中表达水平的统计与作图.

2 研究结果

2.1 *RgURT* 基因家族成员大小及其相似性和开放阅读框分析

RgURT 基因家族有 4 个成员 *RgURT1*~*RgURT4*,大小为 1 422 bp、963 bp、645 bp 和 1 416 bp.使用 NCBI 在线分析工具 ORF Finder 查找 *RgURT1*~*RgURT4* 基因完整的开放阅读框(ORF)分别由 1 419、960、642 和 1 413 个碱基组成(附录表 S2).*RgURT1*~*RgURT4* 的多重比对结果如图 1,发现 4 个基因之间的相似性为 59.24%,*RgURT1* 与 *RgURT2* 之间的相似性为 65.91%,*RgURT1* 与 *RgURT4* 之间的相似性为 42.72%,*RgURT2* 与 *RgURT3* 之间的相似性为 37.75%.

2.2 *RgURT* 酶家族成员间的多重比对与理化性质及亚细胞定位

RgURT1~*RgURT4* 依次由 473、320、214 和 471 个氨基酸残基组成(附录表 S2),相似性为 55.58%,*RgURT1* 与 *RgURT2* 之间相似性为 66.25%,*RgURT4* 与 *RgURT1*~*RgURT3* 之间的相似性为 41%、28%

和 21%(图 2).其蛋白质的分子量、等电点、不稳定性、信号肽和二级结构等理化性质见表 S2,亚细胞定位数据见附表 S3;RgURT1 的二级结构预测如图 3,4 个 RgURT 都很有可能定位于细胞质.



图1 RgURT基因家族的多重比对分析

Fig.1 Multiple sequence alignment of RgURT gene family



图2 RgURT1~RgURT4之间基于其氨基酸序列的多重比对

Fig.2 The homologies among RgURT1~RgURT4 based on their amino acid sequences

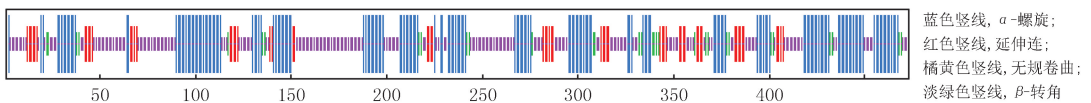


图3 SPOMA预测RgURT1的二级结构

Fig.3 Secondary structures of RgURT1 predicted by SPOMA

2.3 RgURT 蛋白质家族的三级结构预测

用 SWISS-MODEL(<https://swissmodel.expasy.org/>)在线软件分析 RgURT 蛋白质的空间结构,比对结果显示序列同源性(sequence identity)在 40% 以上时,则待预测蛋白与模板蛋白结构可能属于同一家族,即为同源蛋白,则同源建模方法可用于预测该蛋白三维结构,预测模型可信度高.然后,根据 GMQE 值(全球性模型质量估测)及 QMEAN 值评价同源建模的结果.GMQE 值在 0~1 之间,越接近 1 则建模质量越好,

QMEAN 值区间为-4~0,越接近 0 则匹配度越好(附录表 S4)。4RgURT 的单体三维结构如图 4。从图 4 看,RgURT4 的单体三维结构与其他差异较大,其他三者相似性较大,图中颜色同图 3。

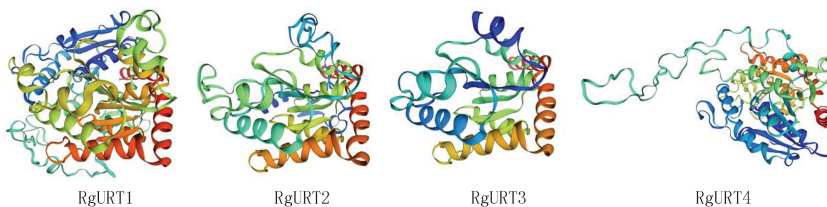
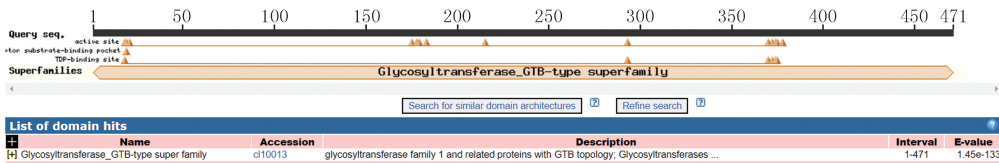


图4 RgURT蛋白质的单体三级结构比较

Fig.4 Monomer 3D structures of RgURTs

2.4 地黄 RgURT 酶家族属性与保守结构区域

利用 NCBI CD search 和 InterProScan 预测 RgURT 蛋白的家族,发现其属于糖基转移酶 GTB 型超家族(图 5)。将 RgURT 蛋白与同源蛋白进行序列多重比对和 C 末端保守区域预测分析结果表明,4 个 RgURT 蛋白在 352~395 位都有一个由 44 个氨基酸残基组成的 UDP-糖基转移酶(UDP-glycosyltransferase)保守结构区域,即 C 末端保守区域 PSPG 盒(plant secondary product glycosyltransferase box)-HCGMNS(图 6)。



1~471, the amino acid residue number of RgURT4. Glycosyltransferase-GTB-type superfamily, GTB型糖基转移酶超家族. Domain, 结构域.

图5 RgURT4家族预测

Fig.5 RgURT4 family prediction

	H C G M N S																																											
AtUGT78D1	W	A	P	Q	V	E	L	K	H	E	A	M	G	V	N	V	T	H	C	G	W	N	S	V	L	E	S	V	S	A	G	V	P	M	I	G	R	P	I	L	A	D	N	
SGT3	W	A	P	Q	Q	L	I	L	K	H	Q	A	I	G	F	L	T	H	C	G	W	N	S	I	L	E	S	L	A	V	G	V	P	L	I	T	W	P	L	F	S	D	N	
Cs1, 6RhaT	W	V	Q	Q	Q	L	I	L	R	H	E	S	V	G	C	Y	V	C	H	S	G	F	S	S	V	T	E	A	V	I	S	D	C	Q	L	V	L	L	P	L	K	G	D	Q
RgURT1	M	A	P	Q	L	E	I	L	G	H	G	S	I	G	A	F	L	T	H	C	G	M	S	S	I	V	E	G	L	M	F	G	N	P	L	I	M	L	P	F	L	V	D	Q
RgURT2	M	A	P	Q	L	E	I	L	G	H	G	S	I	G	A	F	L	T	H	C	G	M	S	S	I	V	E	G	L	M	F	G	N	P	L	I	M	L	P	F	L	V	D	Q
RgURT3	M	A	P	Q	L	E	I	L	G	H	G	S	I	G	A	F	L	T	H	C	G	M	S	S	I	I	E	A	L	V	L	G	K	P	L	V	M	L	P	F	F	V	E	Q
RgURT4	M	A	P	Q	R	D	I	L	G	H	R	S	I	G	S	L	F	H	A	G	M	A	S	I	I	E	T	M	I	Y	G	H	C	L	V	L	L	P	F	V	I	A	Q	
PETUNIA RhaT	M	V	Q	Q	Q	N	I	L	A	H	S	S	V	G	C	Y	C	V	H	A	G	F	S	S	V	I	E	A	L	V	N	D	C	Q	V	V	M	L	P	Q	K	G	D	Q
PummeIo RhaT	M	V	P	Q	A	K	I	L	R	H	G	S	I	G	F	L	S	H	C	G	M	G	S	V	V	E	G	M	V	F	G	V	P	I	I	G	V	P	M	A	Y	E	Q	

图6 RgURT蛋白质C端保守区域PSPG盒分析

Fig.6 Analysis of PSPG box at the C-end of RgURT proteins

2.5 构建系统进化树

通过 NCBI 的 Blastn 搜索与地黄 RgURT 基因编码氨基酸序列相似性较高的植物物种的同源基因,利用 MEGA6.0 软件以 NJ 算法进行分析,构建了相关植物 URT 蛋白的系统进化树(图 7)。其中,RgURT4 蛋白(Group1)、芝麻与独脚金 URT 蛋白聚为一枝。其中,地黄和独脚金同属玄参科植物,亲缘关系近;RgURT1、RgURT2 和 RgURT3 聚在一起(Group2),与茶(Camellia sinensis)等植物 URT 所在聚类分支聚类在一起。其中,RgURT4 氨基酸序列与芝麻 Sesamum indicum (XP_011092852.1)、独脚金 Striga asiatica (GER45226.1)、一串红 Salvia splendens (TEY36988.1)、紫花风铃木 Handroanthus impetiginosus (PIN20168.1)等物种的相似性分别为 75%、70%、68%和 65%(图 7)。

2.6 RgURT 基因家族的基因型依赖性表达分析

RgURT 基因家族中,RgURT1、RgURT2 和 RgURT4 在高、低毛蕊花糖苷地黄品种块根中在转录水平上表达量被测定与比较。基于用 $2^{-\Delta\Delta Ct}$ 法计算所得的基因相对表达水平值,用柱状图展现 3 个基因在高、低毛蕊花糖苷地黄品种块根中转录产生的 mRNA 水平(图 8)。结果表明,三者在高毛蕊花糖苷地黄中的表达

量高于其在低毛蕊花糖苷地黄中的表达量,与两地黄品种的毛蕊花糖苷含量相关。

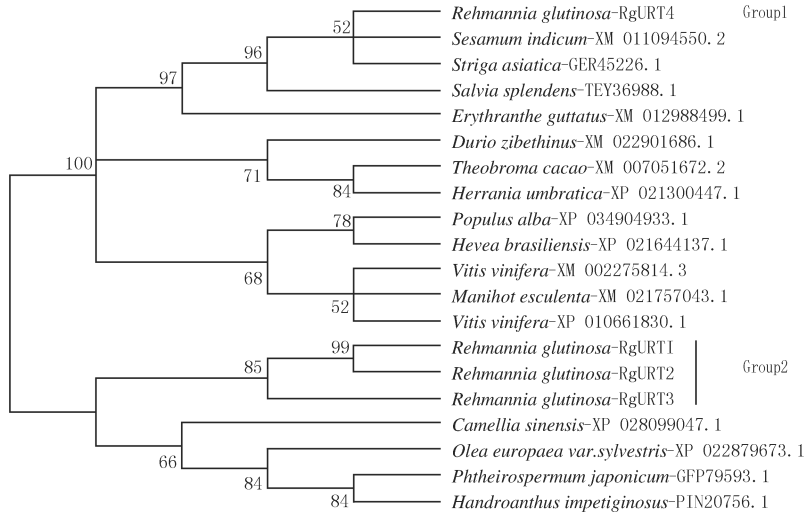


图7 RgURT与植物URT蛋白的系统进化树

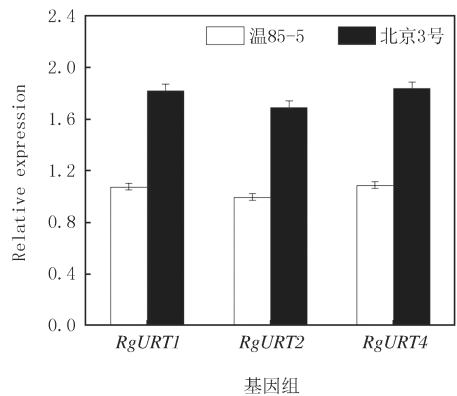
Fig. 7 The phylogenetic tree of RgURTs and other known plant species URT proteins

3 讨论

鼠李糖基转移酶或者 URT 广泛存在于植物界中,参与不同种植物次生代谢产物的生成,在植物的结构组成及多种生理功能中发挥重要的作用^[3,5,16,18-19].它们的基因具有多样化的基因名^[6-13],如拟南芥的 *AtUGT78D1* (*F3RhaT*) 和 *UGT89C1* (*F7RhaT*)^[9,13]、梁平柚的 *Cm1, 2RhaT*^[8] 和地黄的 *RgURT*^[15],但是都具有编码 RhaT 的功能.从系统进化树构建结果看,4 个 *RgURT* 分为 2 个亚组(图 7);从蛋白质结构看,*RgURT* 家族结构比较保守,属于糖基转移酶 GT-B 型亚家族(图 5),具有 UGT 家族的植物次生产物糖基转移酶结构域(PSPG-box).该结构域是多种植物 RhaT 的 C 端的一个保守结构域^[4,19].与其他物种 RhaT 的 PSPG-box 相比,*RgURT* 家族的 PSPG-box 都具有完整的,但是,有一定数量的氨基酸差异(图 6);从基因家族成员大小

看,*RgURT* 基因(表 S1)与其他植物种的 RhaT 的大小均不同^[4,8,20-21].其中,*RgURT1* 和 *RgURT4* 与其他植物种的 RhaT 的大小相近,*RgURT2* 和 *RgURT3* 比其他植物种的 RhaT 小,同源性分析表明,*RgURT* 基因家族成员之间的相似性不高(图 2),但是,与芝麻和独脚金等植物的 URT 之间的相似性比较高(图 7);从亚细胞定位分析看,*RgURT* 家族成员均定位于细胞质(表 S3).这一结果与菊花 CiRhaT-GD4x 的定位结果一致^[22],而与多数植物 RhaT 定位于高尔基体与内质网的结果不一致^[21,23].综上所述,*RgURT* 基因及其编码酶与其他已知植物物种中 URT 基因及其编码酶 URT 既有进化上的保守性又有其独特性,预示其既有 RhaT 的功能又可能具有一定的独特生物学功能。

地黄和其他许多植物体内毛蕊花糖苷的合成途径是一个复杂的代谢过程,涉及很多代谢物与酶^[3].URT 被认为在毛蕊花糖苷的合成途径的下游步骤中起作用,即 3,4-二羟基酪醇葡萄糖至脱咖啡酰基毛蕊花糖苷的步骤和脱鼠李糖基毛蕊花糖苷至毛蕊花糖苷的步骤^[16-18].目前,URT 在毛蕊花糖苷合成过程中的作用研



误差线代表三重重复平均值的标准差,地黄 *RgTIP41* 基因因为内参基因. 基因 (*RgURT2*) 相对表达量最低的值作为 1.

图8 *RgURT* 基因家族在高、低毛蕊花糖苷地黄品种块根中的表达分析
Fig. 8 The expression analysis of *RgURT* gene family in two high and low acteoside level varieties of *R. glutinosa*

究鲜有报道,例如一个基于同源克隆的 *RgURT* 基因片段在高毛蕊花糖苷地黄品种的表达量比其在低毛蕊花糖苷地黄品种的表达量高,可能调控地黄毛蕊花糖苷合成^[22]; *RgURT1* 和 *RgURT2* 在地黄幼苗中的表达受茉莉酸甲酯的诱导^[15],与其他地黄毛蕊花糖苷合成相关酶基因的表达对茉莉酸甲酯的应答反应相似^[24]。本研究通过荧光定量 PCR 分析 *RgURT1*、*RgURT2* 和 *RgURT4* 在高、低毛蕊花糖苷地黄品种块根中的表达,发现它们在高毛蕊花糖苷地黄品种的表达量比在低毛蕊花糖苷地黄品种的表达量高(图 8),可能与地黄品种块根毛蕊花糖苷的含量相关。这一结果与地黄毛蕊花糖苷合成途径上游关键基因与甘蔗中蔗糖合成酶基因的基因型依赖性表达相似^[24-26]。因此,这 3 个 *RgURT* 基因很可能是调控地黄毛蕊花糖苷合成的基因,其编码的 RgURT 酶可能催化 UDP-鼠李糖中的鼠李糖基与其两种底物中葡萄糖的 3C 上的羟基之间形成鼠李糖苷键,合成脱咖啡酰基毛蕊花糖苷和毛蕊花糖苷,但它们催化毛蕊花糖苷合成的作用有待进一步验证。

附录见电子版(DOI:10.16366/j.cnki.1000-2367.2022.12.12.0001)。

参 考 文 献

- [1] LI J J, ZHU J L, LI H M, et al. The effects of NAA on the tuberous root yield and quality of *Rehmannia glutinosa* and its regulatory mechanism by transcriptome and metabolome profiling[J]. *Current Issues in Molecular Biology*, 2022, 44(8): 3291-3311.
- [2] MA L G, DONG C M, SONG C, et al. De novo genome assembly of the potent medicinal plant *Rehmannia glutinosa* using nanopore technology[J]. *Computational and Structural Biotechnology Journal*, 2021, 19: 3954-3963.
- [3] ZHOU Y Q, ZHU J L, SHAO L Y, et al. Current advances in acteoside biosynthesis pathway elucidation and biosynthesis[J]. *Fitoterapia*, 2020, 142: 104495.
- [4] 徐珍珍. 人参 UDP-鼠李糖基转移酶的优化改造与应用[D]. 开封: 河南大学, 2022.
XU Z Z. Optimization and application of UDP-rhamnose glycosyltransferase from *denseng*[D]. Kaifeng: Henan University, 2022.
- [5] 闫雅如, 齐博文, 莫婷, 等. 鼠李糖基转移酶研究进展[J]. *有机化学*, 2018, 38(9): 2281-2295.
YAN Y R, QI B W, MO T, et al. Research progress of rhamnosyltransferase[J]. *Chinese Journal of Organic Chemistry*, 2018, 38(9): 2281-2295.
- [6] XU H T, JIANG Z Q, LIN Z M, et al. FtUGT79A15 is responsible for rutinoylation in flavonoid diglycoside biosynthesis in *Fagopyrum tataricum*[J]. *Plant Physiology and Biochemistry*, 2022, 181: 33-41.
- [7] WANG Z L, ZHOU J J, HAN B Y, et al. GuRhaGT, a highly specific saponin 2'-O-rhamnosyltransferase from *Glycyrrhiza uralensis*[J]. *Chemical Communications*, 2022, 58(34): 5277-5280.
- [8] 张军. 梁平柚鼠李糖基转移酶基因 *Cm1*, *2RhaT* 的克隆及表达分析[D]. 重庆: 西南大学, 2012.
ZHANG J. Cloning and expression analysis of rhamnose glycosyltransferase genes *CM1* and *2RhaT* from Liangping pomelo[D]. Chongqing: Southwest University, 2012.
- [9] GU N, QIU C, ZHAO L G, et al. Enhancing UDP-rhamnose supply for rhamnosylation of flavonoids in *Escherichia coli* by regulating the modular pathway and improving NADPH availability[J]. *Journal of Agricultural and Food Chemistry*, 2020, 68(35): 9513-9523.
- [10] FENG K P, CHEN R D, XIE K B, et al. A regiospecific rhamnosyltransferase from *Epimedium pseudowushanense* catalyzes the 3-O-rhamnosylation of prenylflavonols[J]. *Organic & Biomolecular Chemistry*, 2018, 16(3): 452-458.
- [11] KOJA E, OHATA S, MARUYAMA Y, et al. Identification and characterization of a rhamnosyltransferase involved in rutin biosynthesis in *Fagopyrum esculentum* (common buckwheat)[J]. *Bioscience, Biotechnology, and Biochemistry*, 2018, 82(10): 1790-1802.
- [12] LI D, CHEN G, MA B, et al. Metabolic profiling and transcriptome analysis of mulberry leaves provide insights into flavonoid biosynthesis[J]. *Journal of Agricultural and Food Chemistry*, 2020, 68(5): 1494-1504.
- [13] ZONG G N, FEI S, LIU X, et al. Crystal structures of rhamnosyltransferase UGT89C1 from *Arabidopsis thaliana* reveal the molecular basis of sugar donor specificity for UDP- β -l-rhamnose and rhamnosylation mechanism[J]. *The Plant Journal*, 2019, 99(2): 257-269.
- [14] HSU Y H, TAGAMI T, MATSUNAGA K, et al. Functional characterization of UDP-rhamnose-dependent rhamnosyltransferase involved in anthocyanin modification, a key enzyme determining blue coloration in *Lobelia erinus*[J]. *The Plant Journal*, 2017, 89(2): 325-337.
- [15] 周延清, 邵露露, 李慧敏, 等. 地黄全长转录组分析及毛蕊花糖苷生物合成下游途径关键酶基因的挖掘与解析[J]. *植物生理学报*, 2022, 58(1): 197-206.
ZHOU Y Q, SHAO L Y, LI H M, et al. Full-length transcriptome analysis and the mining and elucidation of key enzyme genes of downstream acteoside biosynthesis pathway in *Rehmannia glutinosa*[J]. *Plant Physiology Journal*, 2022, 58(1): 197-206.
- [16] 周延清, 邵露露, 郭萌萌, 等. 地黄 *C3H* 基因的克隆及生物信息学分析[J]. *广西植物*, 2020, 40(9): 1281-1287.
ZHOU Y Q, SHAO L Y, GUO M M, et al. Cloning and bioinformatics analysis of *C3H* gene in *Rehmannia glutinosa*[J]. *Guihaia*, 2020, 40(9): 1281-1287.
- [17] CHAI M, DONG C M, JIANG D H, et al. Determination of catalpol and verbascoside of different cultivars *Rehmannia glutinosa* Libosch

- by high performance liquid chromatography[J].China Journal of Chinese Materia Medica,28(5):690-692.
- [18] 翟君叶,成旭,孙泽敏,等.毛蕊花糖苷的生物合成研究进展[J].中国生物工程杂志,2021,41(5):94-104.
ZHAI J Y,CHENG X,SUN Z M,et al.Current advances in biosynthesis of acteoside[J].China Biotechnology,2021,41(5):94-104.
- [19] 张桂芝,林继山,李燕洁,等.植物激素糖基化修饰研究进展[J].植物学报,2014,49(5):515-523.
ZHANG G Z,LIN J S,LI Y J,et al.Research advances in the glycosylation of plant hormones[J].Chinese Bulletin of Botany,2014,49(5):515-523.
- [20] 卢晓玲,何铭,张凯旋,等.苦荞鼠李糖基转移酶 FtF3GT1 基因的克隆与转化毛状根研究[J].作物杂志,2020(5):33-40.
LU X L,HE M,ZHANG K X,et al.Study on the cloning and transformation of rhamnose transferase FtF3GT1 gene in Tartary buckwheat[J].Crops,2020(5):33-40.
- [21] UEHARA Y,TAMURA S,MAKI Y,et al.Biochemical characterization of rhamnosyltransferase involved in biosynthesis of pectic rhamnogalacturonan I in plant cell wall[J].Biochemical and Biophysical Research Communications,2017,486(1):130-136.
- [22] 张丹丹.地黄毛蕊花糖苷生物合成和糖醇解途径相关关键酶基因的挖掘和表达分析[D].新乡:河南师范大学,2018.
ZHANG D D.Discovery and expression of key verbascoside biosynthesis-and glycolysis-associated enzyme genes in *Rehmannia glutinosa* [D].Xinxiang:Henan Normal University,2018.
- [23] TAKENAKA Y,KATO K,OGAWA-OHNISHI M,et al.Pectin RG-I rhamnosyltransferases represent a novel plant-specific glycosyltransferase family[J].Nature Plants,2018,4(9):669-676.
- [24] WANG F Q,ZHI J Y,ZHANG Z Y,et al.Transcriptome analysis of salicylic acid treatment in *Rehmannia glutinosa* hairy roots using RNA-seq technique for identification of genes involved in acteoside biosynthesis[J].Frontiers in Plant Science,2017,8:787.
- [25] ZHOU Y Q,WANG X N,WANG W S,et al.De novo transcriptome sequencing-based discovery and expression analyses of verbascoside biosynthesis-associated genes in *Rehmannia glutinosa* tuberous roots[J].Molecular Breeding,2016,36(10):139.
- [26] NOMAN M,SIDDIQUE I,SALEEM B,et al.In silico dissection and expression analysis of sucrose synthase gene family in sugarcane[J].Sugar Tech,2022,24(6):1766-1777.

Bioinformatics and expression analyses of UDP-rhamnose: rhamnosyltransferase family in *Rehmannia glutinosa*

Zhou Yanqing, Li Huimin

(College of Life Sciences, Henan Normal University, Xinxiang 453007, China)

Abstract: Rhamnosyltransferases can transfer active rhamnosyl donors to their rhamnosylated receptor molecules to synthesize rhamnosylated compounds with diverse structures and extensive activities. They play important roles in pharmacology, maintaining cell structure, signal transduction, chemical defense and regulating hormone levels in plants. So far, they have been isolated from many plants. However, few studies have been conducted on the structural characteristics of their gene families and their role in the biosynthesis of acteoside in plants. In this study, we used bioinformatics technology to analyze the structural characteristics of URT gene family in *Rehmannia glutinosa* (*RgURT1-RgURT4*), and their expression patterns in higher acteoside level variety and lower acteoside level variety were compared using its four members as materials. The results showed that their cDNA sizes were from 645 to 1 422 bp, encoded from 214 to 473 amino acid residue-composed proteins with molecular weights from 24.05 to 53.24 kDa. They belonged to the GT-B superfamily of glycosyltransferases, had no signal peptide and a conserved domain box (PSPG box) at the C-terminal, and were localized in the cytoplasm. Their secondary structures mainly were random coil and α -Helix; *RgURT4* and *RgURT1-RgURT3* were clustered into two different groups, respectively, and the homology between *RgURT4* protein and *RgURT1-RgURT3* was low; RT-qPCR analysis showed that the expressions of three *RgURT* genes in the tuberous root of the high acteoside variety were higher than that in the low acteoside variety. This study revealed the structural features of *RgURT* family of *Rehmannia glutinosa*, and preliminarily verified the correlation between *RgURT* gene family and acteoside biosynthesis, and will provide candidate genes for further study of acteoside biosynthesis.

Keywords: rhamnosyltransferase; *Rehmannia glutinosa*; characteristics of *RgURT* gene family members; PSPG box

附 录

表 S1 qRT-PCR 所用引物及其序列

Tab. S1 Primers and their sequences of qRT-PCR

引物名称	正向引物序列(5'-3')	反向引物序列(5'-3')
<i>RgTIP41</i>	TGGCTCAGAGTTGATGGAGTGCT	CTCTCCAGCAGCTTTCTCGGAGA
<i>RgURT1</i>	TGGATGTCCGAACCGAAGA	CAACGAGGAAACGAGTTAAGTCA
<i>RgURT2</i>	AGGAAGTGAAGTTGTGCTTAGTAG	CGGCTGGCTTTCTGAGAG
<i>RgURT4</i>	TTCATGTCTCTTTTCTGTCAACTCC	GTGGCTTCGGCATTTC

表 S2 *RgURT* 基因家族成员及其编码 RgURTs 的特性

Tab. S2 Characteristics of *RgURT* gene family members and their encoding RgURTs

基因	<i>RgURT1</i>	<i>RgURT2</i>	<i>RgURT3</i>	<i>RgURT4</i>
碱基数(nt)	1 422	963	645	1 416
开放阅读框(nt)	1 419	960	642	1 413
编码酶	UDP-rhamnose:rhamnosyltransferase(EC:2.4.1.-)			
氨基酸数/个	473	320	214	471
相对分子质量/kDa	53.24	35.57	24.05	53.09
等电点	5.28	4.85	5.23	5.79
带正电残基数(Asp+Glu)	65	49	30	61
带负电残基数(Arg+Lys)	46	31	23	53
不稳定系数	33.93	31.18	29.8	37.7
总平均亲水性	-1.09	-0.228	-0.19	-0.094
脂肪系数	93.21	90.5	95.56	98.96
跨膜区	15~37,142~164	无	无	7~29
信号肽	无	无	无	无
磷酸化位点	丝氨酸 19 个, 苏氨酸 10 个, 酪氨酸 5 个	丝氨酸 14 个, 苏氨酸 5 个, 酪氨酸 5 个	丝氨酸 10 个, 苏氨酸 2 个, 酪氨酸 2 个	丝氨酸 16 个, 苏氨酸 7 个, 酪氨酸 5 个
二级结构	无规卷曲 43.55%, α 螺旋 37.00%, 延伸链 12.90%, β 转角 6.55%	无规卷曲 38.75%, α 螺旋 35.62%, 延伸链 15.31%, β 转角 10.31%	无规卷曲 34.11%, α 螺旋 36.92%, 延伸链 17.29%, β 转角 11.68%	无规卷曲 43.52%, α 螺旋 37.58%, 延伸链 13.80%, β 转角 5.10%

注:*RgURT1*~*RgURT4*,尿苷二磷酸鼠李糖基转移酶基因.*RgURT1*~*RgURT4* 为基因编码的蛋白质(酶).

表 S3 RgURT 蛋白质的亚细胞定位预测

Tab. S3 Subcellular localization prediction of RgURT proteins

RgURT1		RgURT2		RgURT3		RgURT4	
细胞器	可能性/%	细胞器	可能性/%	细胞器	可能性/%	细胞器	可能性/%
CP	34.8	CP	39.1	CP	47.8	CP	47.8
MC	21.7	Nuclear	30.4	Nuclear	21.7	MC	30.4
Golgi	13.0	MC	21.7	MC	21.7	Nuclear	13.0
ER	8.7	Golgi	4.3	Golgi	4.3	Vacuolar	4.3
Nuclear	8.7	PS	4.3	PS	4.3	ER	4.3
Vesicle	4.3	—	—	—	—	—	—
Vacuolar	4.3	—	—	—	—	—	—
EC	4.3	—	—	—	—	—	—

注:CP,细胞质;MC,线粒体;Golgi,高尔基体;ER,内质网;Nuclear,细胞核;Vesicle,储泡;Vacuolar,液泡;EC,Extracellular,细胞外;PS,过氧化物酶体;—,无定位.

表 S4 预测蛋白质三维结构的评价参数

Tab. S4 Evaluation parameters for predicting the three-dimensional structure of proteins

蛋白质	RgURT1	RgURT2	RgURT3	RgURT4
Template		glycosyltransferase		
序列同源性	39.91%	49.06%	51.79%	37.02%
GMQE	0.70	0.67	0.78	0.68
QMEAN	-3.50	-2.32	-1.80	-3.05
Biounit oligo State		RgURT monomers		