

周豪楠,马喜玲,涂志伟,等. 基于转录组学的粉葛葛根素生物合成相关基因挖掘[J]. 江苏农业科学,2023,51(5):30-37.

doi:10.15889/j.issn.1002-1302.2023.05.004

基于转录组学的粉葛葛根素生物合成相关基因挖掘

周豪楠,马喜玲,涂志伟,李林桐,胡远锋,肖旭峰,罗莎,范淑英

(江西农业大学农学院,江西南昌 330000)

摘要:挖掘与粉葛葛根素生物合成的相关基因,为相关控制基因的验证及深入研究提供理论参考。对 6 个生长时期的粉葛块根进行转录组测序,在葛根素含量最高时期与最低时期初步筛选差异表达基因,对各时期葛根素及其上下游代谢物含量变化趋势与差异表达基因相对表达量进行相关性分析,最终挖掘出相关性较高的基因片段。结果表明,经 2 步筛选差异基因后,挖掘 11 个与粉葛葛根素生物合成相关的片段,共编码 3 个基因,分别为异黄酮羟化酶控制基因(*I2'H*)、异黄酮 7-O-葡萄糖苷转移酶控制基因(*IF7GT*)和异黄酮合酶控制基因(*IFSI*),其中 *I2'H*、*IF7GT*、*IFSI* 均与葛根素的生物合成呈现正向相关,即均正向调控葛根素的生物合成。说明 *I2'H* 与 *IF7GT* 所编码的酶主要负责调控大豆苷元向芒柄花素和 7-O-葡萄糖苷的合成,两者同为葛根素上游代谢物的支链代谢物,因此与葛根素的含量呈正向相关,而 *IFSI* 则通过调控 2,7,4'-三羟基异黄酮的生物合成与异甘草素向葛根素的直接合成,最终调控葛根素生物合成进程。

关键词:粉葛;葛根素;转录组;异黄酮羟化酶控制基因;异黄酮 7-O-葡萄糖苷转移酶控制基因;异黄酮合酶控制基因

中图分类号:S188 **文献标志码:**A **文章编号:**1002-1302(2023)05-0030-08

粉葛(*Pueraria montana* var. *thomsonii*)属豆科葛属,它在我国的栽培史可追溯至南梁(公元 502 年)以前,自此之后对粉葛的炮制方法也日新月异,期间出现过捣为末法、蒸焙法、炒法、煨法等^[1]。《神农本草经》中释义主消渴,起阴气,解诸毒;《本草纲目》中,释名:葛、鸡齐、鹿藿、黄斤,根部供食,鲜食可缓内热脉洪,可解酒,可升阳止泻。2001 年,葛属野生与栽培作物在我国栽培面积已达 40 万 hm^2 ,年终产可达 150 万 $\text{t}^{[2]}$ 。同年,郑水庆等对粉葛主栽的云南地区进行了植物资源调查,并将云南地区现有葛属作物细分为 9 类,根据其主要的药用价值黄酮类物质进行排名,以野葛药用价值最高,粉葛次之^[3]。纵观世界,葛属作物栽培主要分布于亚洲东南部^[4],因此相关的葛类研究样本主要来自中国、泰国、日本等地。时至今日,相关研究显示,粉葛的淀粉含量(鲜样)可高达 20%,且其中含有少量维生素与矿质元素,干物质中黄酮类次生代谢物及异黄酮类次生代谢物总量可达 500 $\text{mg}/\text{kg}^{[5-6]}$ 。因而在

新时代背景下粉葛具有了更深刻的药食兼用的含义,而其综合利用价值也相应被提升至新的维度。例如,利用粉葛株内丰富的异黄酮提取物生产靶向药物缓解心脑血管疾病^[7]和 利用粉葛淀粉含量高且含有微量矿质元素的性质生产保健食品等^[8]。

粉葛的研究兴起较早,一方面是对粉葛体内异黄酮类代谢物代谢合成机制的研究,有学者对粉葛黄酮类代谢物起始合成的控制片段基因作了全片段克隆并在大肠杆菌中过表达,通过比对最终鉴定出粉葛基因组内存在的查尔酮异构酶(CHI)合成基因^[9];伴随这一浪潮的兴起,He 等展开了针对粉葛异黄酮代谢通路的基因组研究,通过建立表达序列标签(EST)文库及相关片段的过表达(表达序列标签),成功从粉葛内鉴定出查尔酮合酶(CHS)、查尔酮异构酶(CHI)、查尔酮还原酶(CHR)、2'-羟基异黄酮脱水酶(2'HID)的控制片段基因^[10];羽健宾等成功从粉葛内克隆出了查尔酮合酶(*PtCHS*)基因的全片段,片段全长 1 179 bp,编码了 389 个氨基酸^[11];Li 等的研究表明,在粉葛块根内异黄酮的积累伴随着 *PIUGT1* 基因的高度表达,因此,他们认为该 7-O-葡萄糖苷转移酶的编码基因对异黄酮在粉葛体内的生物合成起关键控制作用^[12]。另一方面,关于粉葛药用价值的研究从未间断过,其中被关注最多的是粉葛块根内的葛根素,由于其具有广泛的药理特性^[13],在血管扩张^[14]、心脏保护^[15]、抗氧

收稿日期:2022-06-16

基金项目:江西省现代农业产业技术体系-葛栽培岗资助项目(编号:050014-9021107274)。

作者简介:周豪楠(1997—),男,江苏南通人,硕士研究生,主要从事粉葛栽培及代谢机理研究。E-mail:zhnsrm@126.com。

通信作者:范淑英,教授,主要从事蔬菜栽培生态与环境调控研究。E-mail:fansy12@126.com。

化^[16]、抑制乙醇吸收^[17]等领域皆有应用。现如今的研究进展^[18-19]中认可度较高的粉葛生物合成通路有 2 个(图 1),一是大豆苷元在一个糖苷转移酶(GT)的催化下最终在大豆苷元 8 号 C 原子位增加一个糖苷,该通路的葛根素生物合成主要受到大豆苷元含量的控制;另一个是大豆苷元的上游产物 2,7,4'-3-羟基异黄酮在 GT 的催化下形成三羟基异黄酮 8-碳葡萄糖苷,随后在脱水酶(HID)的作用下形成葛根素。

目前,研究报道均集中在异黄酮类代谢物在粉葛块根内的生物合成机制,但关于粉葛葛根素生物合成的机制方面却鲜有报道。为进一步研究葛根

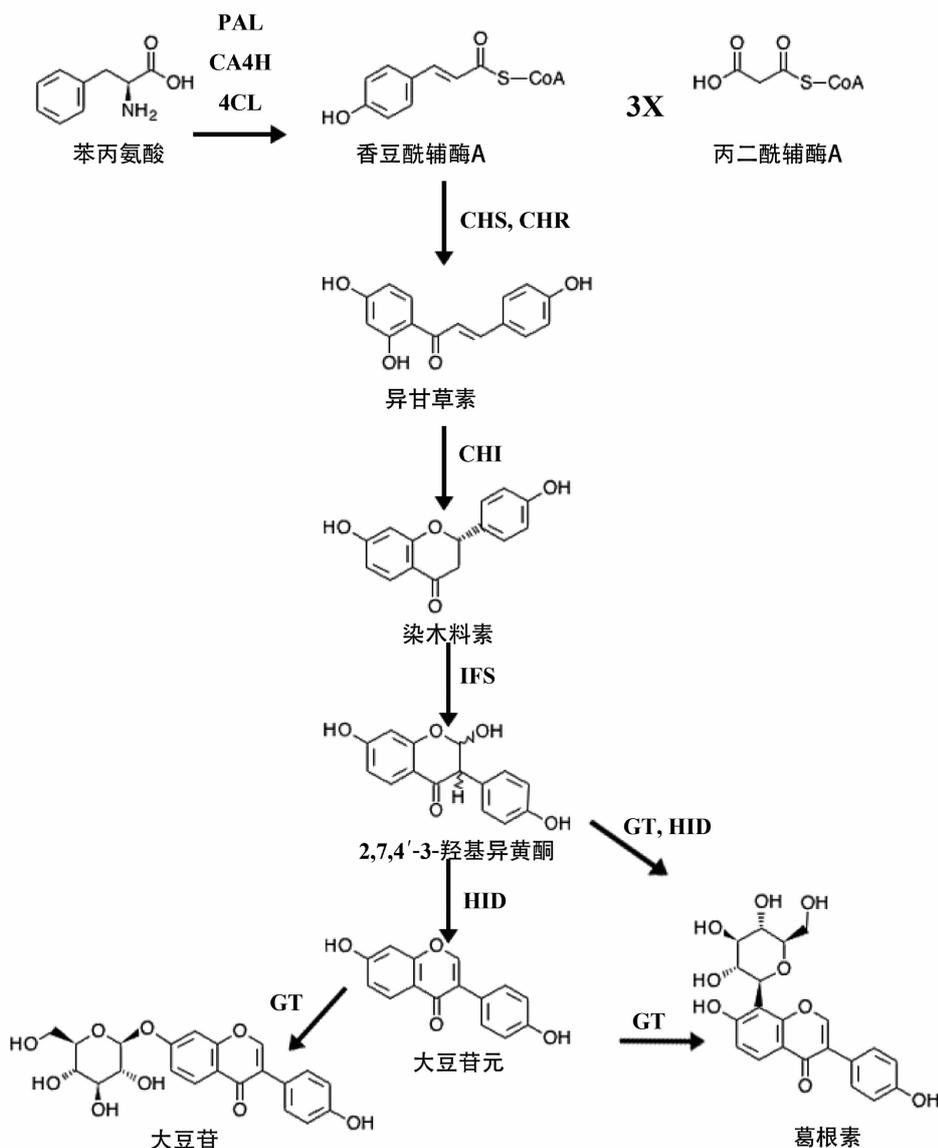
素生物合成的机制,试验通过对 6 个生长时期的粉葛块根进行转录组测序,与各时期目标代谢物含量进行联合分析,从而挖掘出与葛根素生物合成最为相关的基因,以期为后期遗传转化试验的验证和最终探索葛根素生物合成的机制提供理论依据。

1 材料与方法

1.1 试验材料

1.1.1 试验样本 江西主栽粉葛品种赣葛一号。

1.1.2 试验器材 高效液相色谱仪 LC-20AT(日本 SHIMADZU 公司)、反相色谱柱 C₁₈ 4.6 mm × 250 mm(日本 SHIMADZU 公司)。



PAL—苯丙氨酸解氨酶; CA4H—肉桂酸 4-羟化酶; 4CL—4-香豆酸 CoA 连接酶; CHS—查尔酮合酶; CHR—查尔酮还原酶; CHI—查尔酮异构酶; IFS—2-羟基异黄酮合酶; HID—2-羟基异黄酮脱水酶; GT—未知糖苷转移酶

图1 目前公认的葛根素合成通路

1.2 试验方法

1.2.1 样本获取 样本于2021年栽培于江西省宜春市奉新县棚下试验基地,基肥为1 800 kg/hm² 有机肥料,定植后于每月的第1天施600 kg/hm² 液体肥(氮、磷、钾含量均为20%),于2021年6—11月每月的30日收取供试样本,并于2021年12月开展试验(编号及取样说明见表1)。

表1 样本编号说明

样本编号	取样时期	取样部位
GP6R	6月30日	块根中间部
GP7R	7月30日	块根中间部
GP8R	8月30日	块根中间部
GP9R	9月30日	块根中间部
GP10R	10月30日	块根中间部
GP11R	11月30日	块根中间部

1.2.2 葛根素、大豆苷元、大豆苷含量测定 测定方法参考 Cherdshewasart 等的方法^[20]。

1.2.3 转录组测序 无参考基因转录组测序由北京百迈克生物科技有限公司完成。

1.2.4 RT-qPCR 验证 使用华越洋 RNA 提取试剂盒提取总 RNA,使用 TaKaRa RT-PCR kit 试剂盒对目标片段进行相对表达量测定。

1.3 数据分析与制图

使用 SPSS 21.0 软件进行数据分析,使用 GGplot2、Origin 2022 pro 软件进行绘图。

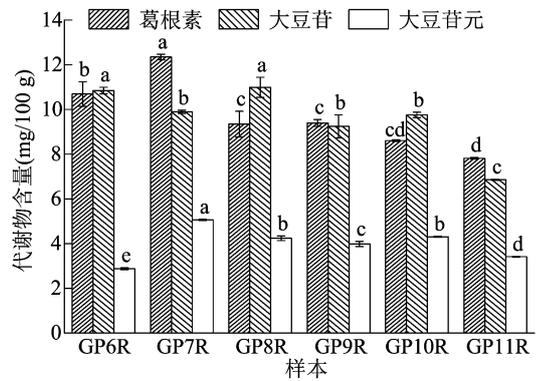
2 结果与分析

2.1 粉葛块根内葛根素、大豆苷元、大豆苷的含量

由图2可知,粉葛块根内葛根素含量总体呈现先增后降的趋势,GP7R时期粉葛块根内葛根素含量最高,达到了12.35 mg/100 g(鲜质量);GP8R与GP9R时期葛根素含量整体无明显变化,但均显著低于GP6R与GP7R时期,在GP10R与GP11R时期时,粉葛块根内葛根素含量持续下降。大豆苷和大豆苷元作为葛根素生物合成途径上游2,7,4'-3-羟基异黄酮的分支产物,其含量变化也总体呈现先增后降的规律,不同的是大豆苷含量在GP8R时期达到峰值,而大豆苷元含量则与葛根素含量变化趋势大体一致。

2.2 粉葛转录组测序结果分析

以 Illumina 测序平台对6个生长时期粉葛块根样本进行全长转录组测序,获得全长非嵌合序列



柱上不同小写字母表示不同生长期间的指标差异显著($P < 0.05$)

图2 不同生长期粉葛块根内葛根素、大豆苷元、大豆苷的含量

177 535 条。对全长非嵌合的 Unigene 进行聚类得到高质量一致 Unigenes 共 75 943 个,对高质量一致 Unigenes 合并进行去冗余分析得到 42 211 个 Unigenes,对去冗余后的 Unigenes 进行功能注释后共有 40 650 个 Unigenes 得到注释(表2)。各样本 GC 含量均高于 44%,Q30 均高于 91%(表3)。

表2 注释的转录本数量统计

注释数据库	注释 Unigene 数量(个)	占去冗余后总 Unigenes 比例(%)
同源群集(COG)	17 612	41.72
基因组学(GO)	33 348	79.00
京都基因和基因组百科全书(KEGG)	18 136	42.97
欧洲核糖体(KOG)	25 845	61.23
蛋白质家族的集合(Pfam)	34 015	80.58
注释蛋白序列(Swiss-Prot)	30 214	71.58
直系同源蛋白分组比对(eggNOG)	38 750	91.80
非冗余蛋白质序列(nr)	40 547	96.06
总计	40 650	96.30

表3 转录组测序质量分析结果

样本编号	有效序列(个)	有效总碱基数(个)	GC 含量(%)	Q30(%)
GP6R	22 422 956	6 708 859 928	46.11	91.67
GP7R	23 738 717	7 098 074 924	44.71	91.21
GP8R	23 824 312	7 121 215 508	44.95	91.32
GP9R	25 310 363	7 544 586 228	45.19	91.33
GP10R	25 538 543	7 625 692 636	44.70	91.46
GP11R	22 338 901	6 691 182 460	44.53	91.37

2.3 差异表达基因 GO 和 KEGG 通路注释分析

根据试验测得的各月份间粉葛块根内葛根素含量变化数据,选取葛根素含量变化最大的2个生长时期:7月和11月,根据转录本中被注释的 Unigenes 表达量的 FPKM(每千个碱基的转录每百

万映射读取的片段)值,以 \log_2 差异倍数(FC) ≥ 1 和错误发现率(FDR) < 0.1 为标准,筛选出 961 个差异 Unigenes,其中上调表达的共 549 个,下调表达的共 412 个。根据 GO 注释,这 961 个差异 Unigenes

主要被注释至生物过程中的细胞过程、代谢过程和单生物过程中。根据 KEGG 的注释,这 961 个差异 Unigenes 也主要被注释在代谢过程的氨基酸生物合成、淀粉与糖代谢途径和碳代谢中(图 3 和图 4)。

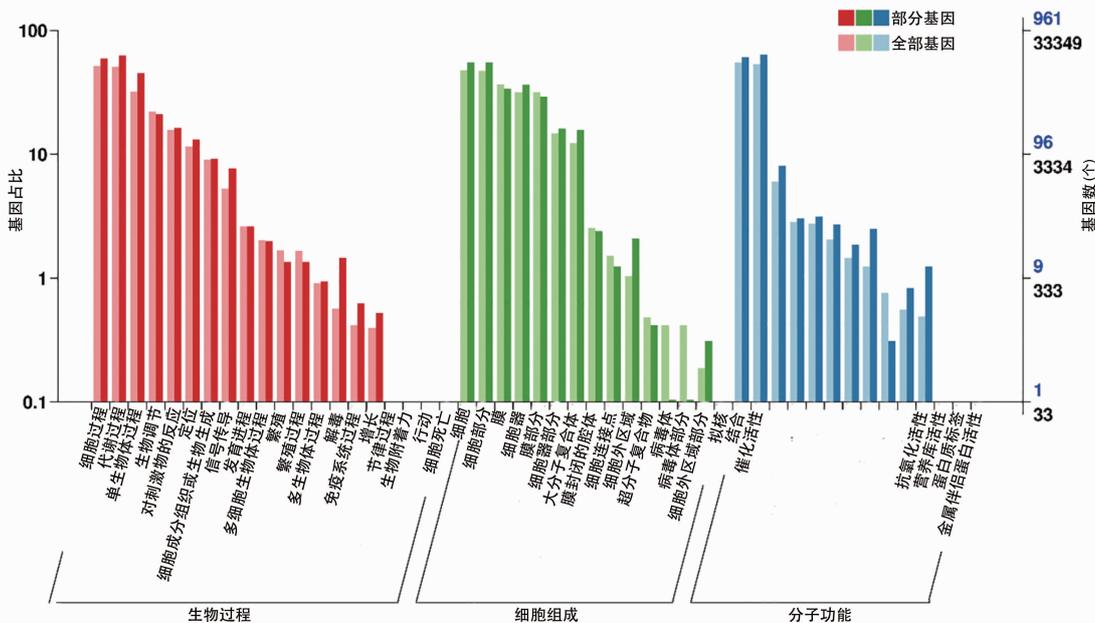


图3 差异表达 Unigenes 的 GO 注释结果(GP7R 和 GP11R)

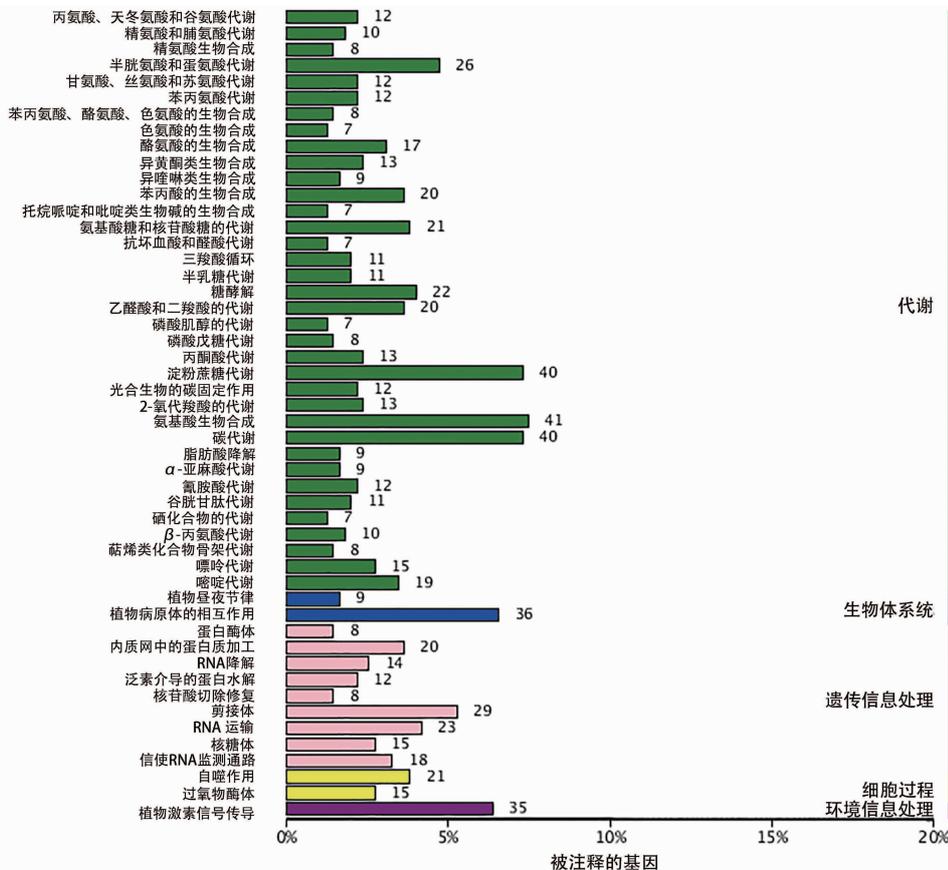


图4 差异表达 Unigenes 的 KEGG 注释(GP7R 和 GP11R)

2.4 差异表达基因 GO 和 KEGG 富集分析

初步筛选出的 961 个差异表达基因在 GO 数据库中主要被富集在生物过程的磷信号转导系统和蔗糖代谢中(图 5);在 KEGG 的富集分析中选择了

Q 值(矫正后的 P 值)低于 1 的注释片段进行富集,并按 Q 值由低到高进行排序,结果显示,可信度较高的富集通路(图 6)分别为异黄酮类物质生物合成、乙醛酸和二羧酸代谢、半胱氨酸-蛋氨酸代谢。

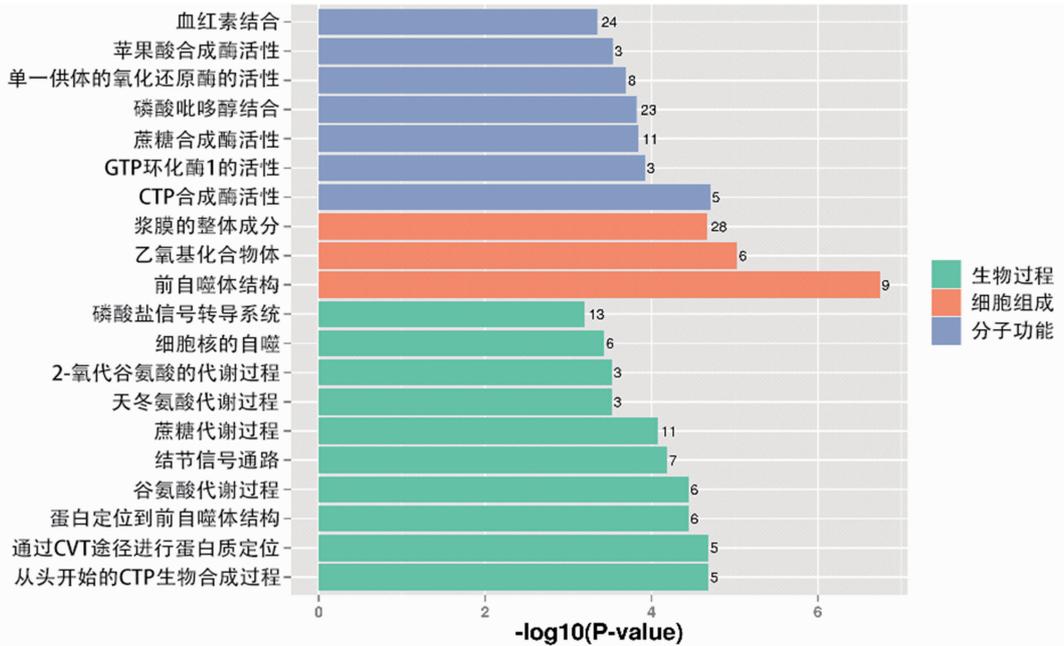


图5 差异表达 Unigenes GO 富集通路(GP7R 和 GP11R)

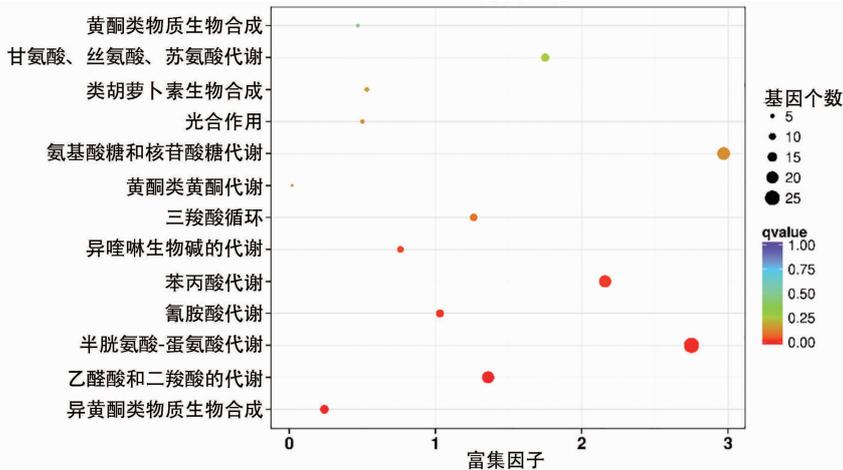


图6 差异表达 Unigenes KEGG 富集通路(GP7R 和 GP11R)

2.5 差异表达基因相对表达量分析与筛选

从上述初步筛选的差异表达基因中筛选出异黄酮代谢通路差异表达基因共 13 个,而后根据上述粉葛块根内葛根素、大豆苷、大豆苷元含量随月份变化的趋势规律,结合转录本所测得的异黄酮代谢通路的 13 个差异 Unigenes 的相对表达量 FPKM 值(归一化处理),筛选出各月份 FPKM 值变化规律与代谢物产物含量数据呈正向相关的 Unigenes 共 11 个,无与代谢物含量变化呈负向相关的 Unigene

(相对表达量变化趋势见图 7)。

2.6 代谢物与差异表达基因表达量相关性分析

图 8 反映出所筛选出的 11 个差异基因与 3 种代谢产物的相关性,11 个基因片段均与葛根素的生物合成呈现正相关,各个片段与粉葛葛根素、大豆苷、大豆苷元的含量变化具有显著相关性。

2.7 筛选后的差异表达基因注释

根据 KEGG 的注释(表 4),11 个 Unigenes 有 6 个编码了异黄酮 2'-羟化酶控制基因(K13260)、2

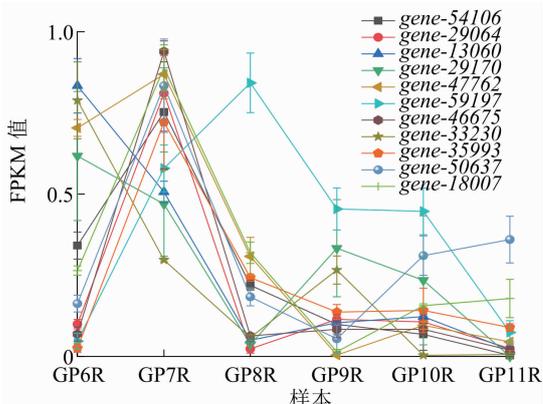


图7 筛选差异表达 Unigenes 表达量变化趋势

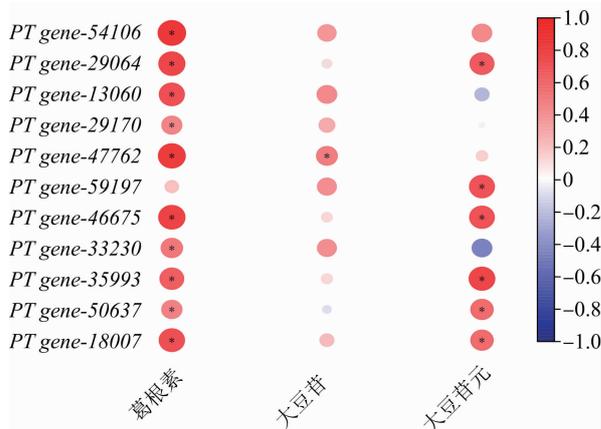


图8 代谢产物与差异基因表达量相关性热力图
*表示显著相关(P<0.05)

个编码了异黄酮 7-O-葡萄糖基转移酶控制基因 (K13263)、3个编码了 2-羟基异黄酮合酶控制基因 (K13257), 结合注释可知, 在粉葛的生长过程中 *I2'H*、*IF7GT* 与 *IFS1* 均与葛根素的生物合成存在正

向相关性。

2.8 RT-qPCR 验证

选取转录本筛选后具有代表性的 4 个差异表达基因片段 (表 5) 进行相对表达量验证, 相对表达量的表达趋势与转录本的 FPKM 值趋势一致 (图 9), 说明转录本数据可靠。

3 讨论与结论

以 KEGG 所标注的代谢通路为参考, 在试验所测得的 3 种代谢物的含量变化中可知, 3 种代谢物的含量总体均呈先增加后降低的趋势, 结合该变化趋势可从 961 个初筛的差异基因片段中筛选出 11 个与葛根素生物合成呈正向相关的差异基因片段, 经过注释后共编码 3 个基因, 分别为 *I2'H*、*IFS1*、*IF7GT*。*IF7GT* 的主要作用为在大豆苷的 7 号氧原子位添加一个糖苷^[21], 负责控制大豆苷元向大豆苷的合成, 试验数据表明, *IF7GT* 与 3 种代谢物均呈现正向调控的趋势, 由此推断当该控制基因表达量增加时, 控制 8 号碳原子位糖苷转移的控制基因的表达量也会相应增加或降低, 最终呈现正向或负向调控葛根素的生物合成, 而目前关于 8 号碳原子位的糖苷转移酶的控制基因有研究表明是 *PIUGT43* 基因^[22]; 前人的研究表明, *I2'H* 可正向调控丙二酰糖苷的积累与苯丙素类物质的合成^[23], 且在异黄酮代谢通路中, *I2'H* 主要作用在大豆苷元向 2-羟基芒柄花素的生物合成中^[24], 属于大豆苷元向下游代谢物合成的关键控制酶的编码基因, 结合 *I2'H* 在粉葛各个生长期内的表达量变化规律和大豆苷元与大

表 4 差异表达 Unigenes 的注释情况

KEGG 基因名称	KEGG 编号	中文名称	名称缩写	被注释的 Unigene 数量 (个)
isoflavone 2'-hydroxylase	K13260	异黄酮 2'-羟化酶	<i>I2'H</i>	6
isoflavone 7-O-glucosyltransferase	K13263	异黄酮 7-O-葡萄糖基转移酶	<i>IF7GT</i>	2
2-hydroxyisoflavanone synthase	K13257	2-羟基异黄酮合酶	<i>IFS1</i>	3

表 5 差异表达 Unigenes 引物序列

基因 ID	正向引物序列 (5'→3')	反向引物序列 (5'→3')
<i>Pt gene - 13060</i>	GTTGAGGTGGTTTGATTTTGATGG	GTCCTTTGATGATTTGATCCGCTGTA
<i>Pt gene - 59197</i>	TGGAGTFTTTCCTCGATACTTTGC	AGACCTCATCCCCGAGCCTTT
<i>Pt gene - 33230</i>	TATAAACTCCTCCTGGAATCCG	TCCCCGACACCATCTCTCAT
<i>Pt gene - 50637</i>	GCCACAGCACCCAAAACGT	GGAGGCACCAGAAGTGAAGTAGA
<i>Pt gene - actin</i>	TTGTTTACTACTCCAGTGGCTCC	TCCAACCCATTCACTTCTCCA

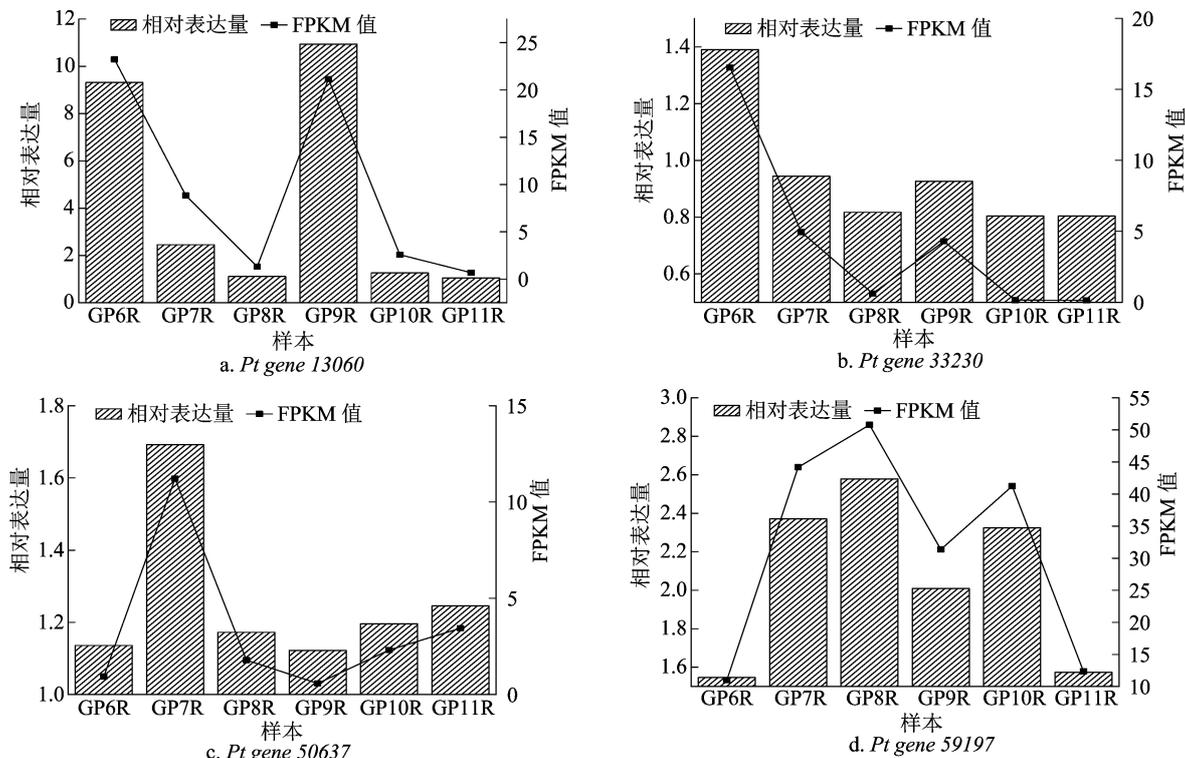


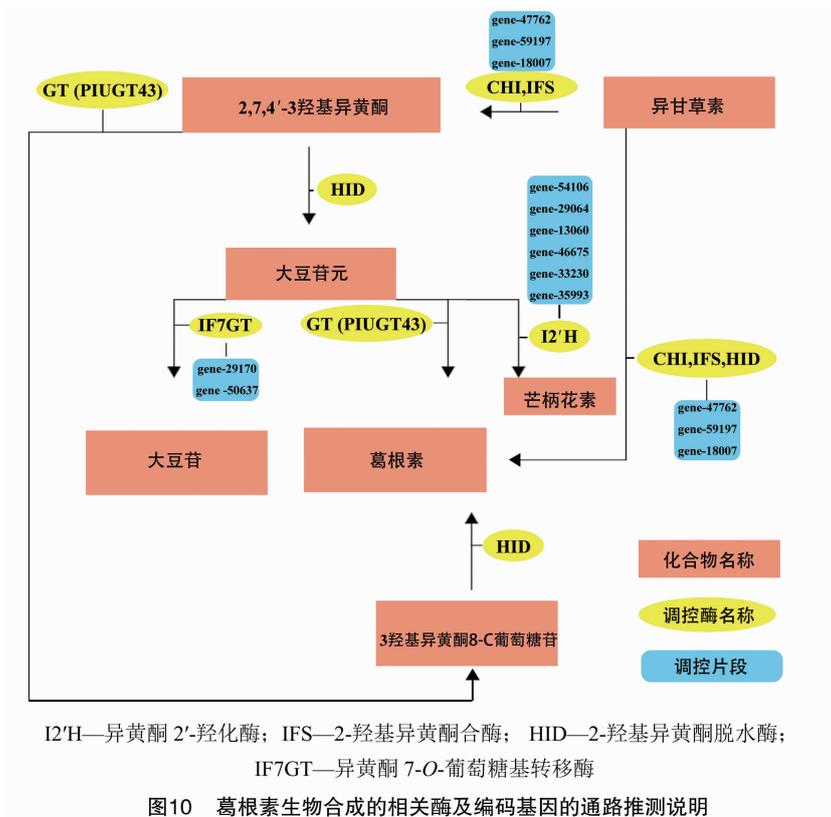
图9 差异表达 Unigenes 相对表达量验证结果

豆苷的含量变化规律可从侧面证明伴随大豆苷元含量上下起伏,其下游代谢物的合成关键控制酶的活性也相应发生变化,由此可推测控制葛根素合成的糖苷转移基因的表达量也会相应增加或降低,因而 *I2'H* 的表达量可从侧面反映葛根素含量的变化,且两者呈正相关;同理,作为大豆苷元向下代谢的支链产物大豆苷,其关键控制酶的编码基因 *IF7GT* 也呈现出了相同的正向调控规律,因此 *IF7GT* 也可从侧面反映出葛根素含量的变化,且两者呈现正相关性;*IFSI* 则可通过葛根素合成的 2 种通路正向调控葛根素的生物合成,分别为控制其上游代谢物甘草素向 2,7,4'-3-羟基异黄酮的转化^[25] 和控制异甘草素代谢合成为葛根素,以最终正向调控葛根素的合成。

试验结果(图 10)表明,在粉葛中 6 个基因片段编码了 *I2'H*,并正向调控了大豆苷元向支链代谢物芒柄花素的生物合成而与葛根素的生物合成存在正向相关;2 个基因片段编码了 *IF7GT*,并正向调控了大豆苷元向大豆苷的生物合成进程而与同为支链产物的葛根素的生物合成存在正向相关;3 个基因片段编码了 *IFSI*,正向调控了 2,7,4'-3-羟基异黄酮的生物合成与异甘草素向葛根素的直接合成而最终表现为调控葛根素的生物合成。

参考文献:

- [1] 廖晶晶,钟凌云,湛瑞林,等. 葛根、粉葛炮制历史沿革及现代炮制比较研究[J]. 中华中医药杂志,2021,36(11):6643-6650.
- [2] 杨明毅,史劲松,孙晓明. 葛根的综合利用及深加工[J]. 常德师范学院学报(自然科学版),2001,13(1):74-75,8.
- [3] 郑水庆,曾明. 云南葛属药用植物资源调查[J]. 中草药,2002,33(8):755-756.
- [4] 杨旭东,王爱勤,何龙飞. 葛根种质资源及其开发利用研究进展[J]. 中国农学通报,2014,30(24):11-16.
- [5] 刘云,张瑶,和润喜. 葛根及葛根食品的研究与开发现状[J]. 中国林副特产,2010(1):94-97.
- [6] 陆宁,宛晓春. 葛根总黄酮、淀粉的提取及应用[J]. 食品工业科技,1998,19(1):33-34,32.
- [7] 陈文杰. 葛根研究进展[J]. 中国中医药现代远程教育,2009,7(1):6-8.
- [8] 邵兰兰,赵燕,杨有仙,等. 葛根异黄酮、淀粉的提取及葛产品开发研究进展[J]. 食品工业科技,2012,33(6):452-455.
- [9] Terai Y, Fujii I, Byun S H, et al. Cloning of chalcone-flavanone isomerase cDNA from *Pueraria lobata* and its overexpression in *Escherichia coli* [J]. Protein Expression and Purification, 1996, 8(2):183-190.
- [10] He X Z, Blount J W, Ge S J, et al. A genomic approach to isoflavone biosynthesis in kudzu (*Pueraria lobata*) [J]. Planta, 2011, 233(4):843-855.
- [11] 羽健宾,李钰婷,张静,等. 粉葛查尔酮合成酶基因 *PtCHS* 的克隆与植物表达载体构建[J]. 分子植物育种,2021,19(4):1143-1149.



[12] Li J, Li Z B, Li C F, et al. Molecular cloning and characterization of an isoflavone 7-O-glucosyltransferase from *Pueraria lobata* [J]. *Plant Cell Reports*, 2014, 33(7): 1173-1185.

[13] Zhou Y X, Zhang H, Peng C. Puerarin: a review of pharmacological effects [J]. *Phytotherapy Research: PTR*, 2014, 28(7): 961-975.

[14] Liu H L, Zheng X L, Zhong X L, et al. Puerarin inhibits vascular calcification of uremic rats [J]. *European Journal of Pharmacology*, 2019, 855: 235-243.

[15] Wang Q, Wu T X, Chen X Y, et al. Puerarin injection for unstable angina pectoris [J]. *The Cochrane Database of Systematic Reviews*, 2006(3): CD004196.

[16] Jeon Y, Lee J, Lee Y, et al. Puerarin inhibits inflammation and oxidative stress in dextran sulfate sodium-induced colitis mice model [J]. *Biomedicine & Pharmacotherapy*, 2020, 124: 109847.

[17] Han Q, Chen K J, Su C L, et al. Puerarin loaded PLGA nanoparticles: optimization processes of preparation and anti-alcohol intoxication effects in mice [J]. *AAPS PharmSciTech*, 2021, 22(6): 217.

[18] Suntichaikamolkul N, Tantisuwanchikul K, Prombutara P, et al. Transcriptome analysis of *Pueraria candollei* var. *mirifica* for gene discovery in the biosyntheses of isoflavones and miroestrol [J]. *BMC Plant Biology*, 2019, 19(1): 581.

[19] Wang X, Li S T, Li J, et al. De novo transcriptome sequencing in *Pueraria lobata* to identify putative genes involved in isoflavones biosynthesis [J]. *Plant Cell Reports*, 2015, 34(5): 733-743.

[20] Cherdshewasart W, Subtang S, Dahlan W, et al. Major isoflavonoid contents of the phytoestrogen rich herb *Pueraria mirifica* in comparison with *Pueraria lobata* [J]. *Journal of Pharmaceutical and Biomedical Analysis*, 2007, 43(2): 428-434.

[21] Noguchi A, Saito A, Homma Y, et al. A UDP-glucose: isoflavone 7-O-glucosyltransferase from the roots of soybean (*Glycine max*) seedlings. Purification, gene cloning, phylogenetics, and an implication for an alternative strategy of enzyme catalysis [J]. *The Journal of Biological Chemistry*, 2007, 282(32): 23581-23590.

[22] Wang X F, Li C, Zhou C, et al. Molecular characterization of the C-glucosylation for puerarin biosynthesis in *Pueraria lobata* [J]. *The Plant Journal*, 2017, 90(3): 535-546.

[23] Chen J, Yuan H, Zhang L, et al. Cloning, expression and purification of isoflavone-2'-hydroxylase from *Astragalus membranaceus* Bge. var. *mongolicus* (Bge.) Hsiao [J]. *Protein Expression and Purification*, 2015, 107: 83-89.

[24] Uchida K, Akashi T, Aoki T. The missing link in leguminous pterocarpin biosynthesis is a dirigent domain-containing protein with isoflavanol dehydratase activity [J]. *Plant & Cell Physiology*, 2017, 58(2): 398-408.

[25] Gupta O P, Dahuja A, Sachdev A, et al. Cytosine methylation of isoflavone synthase gene in the genic region positively regulates its expression and isoflavone biosynthesis in soybean seeds [J]. *DNA and Cell Biology*, 2019, 38(6): 510-520.