

## 基于转录组的黄精多糖代谢关键酶基因的筛选与验证

陶鹏, 刘应, 唐子惟, 尹彦棚, 周罗静, 陈胡兰, 高继海, 彭腾\*

(成都中医药大学药学院西南特色中药资源国家重点实验室, 中药材标准化教育部重点实验室,  
成都 611137)

**[摘要]** 目的:通过湖北黄精、滇黄精、多花黄精根茎的转录组学分析,对影响不同品种黄精多糖含量的关键酶基因进行筛选和验证,并进行氨基酸序列的深入分析,丰富黄精属植物的转录组数据并为其多糖生物合成机制和遗传改良提供参考。方法:使用了Illumina NovaSeq高通量测序平台对黄精转录组进行测序分析,根据Nr、基因本体(GO)、京都基因与基因组百科全书(KEGG)数据库的注释分析比较3种黄精转录组的差异,筛选出多糖代谢途径中的关键酶,对关键酶基因表达量进行聚类分析并与多糖含量进行相关性分析,然后对筛选出的8个关键酶基因进行实时荧光定量聚合酶链式反应(Real-time PCR)验证,结合测序结果进一步筛选出多糖生物合成关键酶基因进行同源基因序列分析,构建系统进化树,预测模体(motif)、保守结构域、蛋白序列等电点与相对分子质量,并使用同源建模法构建三维蛋白结构。结果:通过Nr数据库的注释发现3种黄精与芦笋的基因同源性最高,GO数据库注释结果表明这3种黄精在结合功能、催化活性、代谢过程和细胞组份上差异显著,而KEGG通路注释结果说明这3种黄精在淀粉与蔗糖代谢通路和半乳糖代谢通路上差异显著,根据聚类分析、相关性分析、Real-time PCR验证实验、表达量情况与氨基酸序列的结构与功能预测推测出显著影响不同品种黄精多糖含量的关键酶为 $\beta$ -果糖苷酶(sacA)。结论:sacA可能是黄精多糖含量差异的主要影响因素,也是黄精多糖主要为果聚糖的重要原因。

**[关键词]** 黄精; 转录组; 多糖;  $\beta$ -果糖苷酶; 果糖基转移酶

**[中图分类号]** R284.2; R285; R289; R287; R22; R2-031; R33; R24 **[文献标识码]** A **[文章编号]** 1005-9903 (2023)12-0157-11

**[doi]** 10.13422/j.cnki.syfjx.20230414 **[增强出版附件]** 内容详见 <http://www.syfjxzz.com> 或 <http://cnki.net>

**[网络出版地址]** <https://kns.cnki.net/kcms/detail/11.3495.R.20230313.0859.002.html>

**[网络出版日期]** 2023-03-13 15:01:27

### Transcriptome-based Screening and Validation of Key Enzyme Genes for *Polygonatum* Polysaccharide Metabolism

TAO Peng, LIU Ying, TANG Ziwei, YIN Yanpeng, ZHOU Luoqing, CHEN Hulan, GAO Jihai,  
PENG Teng\*

(School of Pharmacy, Chengdu University of Traditional Chinese Medicine, State Key Laboratory of Southwestern Chinese Medicine Resources, Key Laboratory of Standardization of Chinese Herbal Medicine, Ministry of Education, Chengdu 611137, China)

**[Abstract]** **Objective:** To screen and validate key enzyme genes affecting the polysaccharide content in different *Polygonatum* species and perform in-depth amino acid sequence analysis by transcriptomic analysis of *P. zanlanscianense*, *P. kingianum*, and *P. cyrtonema* rhizomes to enrich the transcriptome data of *Polygonatum* plants and provide references for polysaccharide biosynthesis mechanism and genetic improvement. **Method:** The *Polygonatum* transcriptome was sequenced and analyzed using the Illumina NovaSeq high-throughput

**[收稿日期]** 2022-11-20

**[基金项目]** 四川省科技厅重点项目(2020JDTD0022, 2021YFN0015, 2022YFS0444);成都市科技局技术创新研发项目(2019-YF05-02306-SN, 2021-YF05-02298-SN);达州市重点研发计划项目(20ZDYF0029)

**[第一作者]** 陶鹏,在读硕士,从事中药质量标准研究,E-mail:1326846882@qq.com

**[通信作者]** \*彭腾,教授,从事中药质量标准及大健康产品研发,E-mail:56649923@qq.com

sequencing platform, and the differences in the transcriptomes of the three *Polygonatum* species were compared and according to the annotations of Nr, Gene Ontology (GO), and Kyoto Encyclopedia of Genes and Genomes (KEGG) databases. The key enzymes in the polysaccharide metabolism pathway were screened, and the expression of key enzyme genes was clustered and correlated with the polysaccharide content. Finally, Real-time polymerase chain reaction (Real-time PCR) was performed to validate the eight key enzyme genes, and the key genes of polysaccharide biosynthesis were further screened for homologous gene sequence analysis in combination with sequencing results, followed by constructing phylogenetic trees, predicting motifs, conserved structural domains, protein sequence isoelectric points, and molecular weights, and constructing 3D protein structures by using homology modeling method. **Result:** The annotation of the Nr database revealed that three *Polygonatum* species had the highest gene homology with *Asparagus officinalis*. GO database annotation results showed that three *Polygonatum* species differed significantly in binding, catalytic activity, metabolic processes, and cellular components, while the KEGG pathway annotation results indicated that three *Polygonatum* species differed significantly in the starch and sucrose metabolic pathway and galactose metabolic pathway. According to clustering analysis, correlation analysis, Real-time PCR, expression profiles, and structural and functional predictions of amino acid sequences, the key enzyme significantly affecting the polysaccharide content in different *Polygonatum* species was inferred to be  $\beta$ -fructofuranosidase (sacA). **Conclusion:** SacA may be the main influencing factor for the difference in polysaccharide content of *Polygonatum*, and is also an important reason why *Polygonatum* polysaccharides are mainly fructans.

**[Keywords]** *Polygonatum*; transcriptome; polysaccharide;  $\beta$ -fructofuranosidase; fructosyltransferase

黄精为百合科植物黄精、多花黄精、滇黄精的干燥根茎<sup>[1]</sup>,药用历史悠久,始载于《名医别录》<sup>[2]</sup>,具有补气养阴、健脾益肾、润肺之功效,用于脾胃阴虚、肺虚燥咳、精血不足等证<sup>[3]</sup>,在全国多地均有种植,由于其喜湿好阴的习性,主要分布于长江流域一带及长江以南等地区<sup>[4-7]</sup>。主要化学成分有多糖、萜类、黄酮类、酚类、苯丙素类<sup>[8]</sup>及氨基酸<sup>[9]</sup>等,现代药理学研究表明黄精具有抗骨质疏松<sup>[10]</sup>、降血糖<sup>[11]</sup>、调节脂质代谢<sup>[12]</sup>、抗氧化<sup>[13]</sup>等作用。

湖北黄精(*Polygonatum zanlanscianense*)作为民间常用中药,分布广泛<sup>[14]</sup>,资源量大,在部分地区能作为黄精的替代品或混淆品使用<sup>[15]</sup>,然而目前对于湖北黄精相关研究较少,朱巧等<sup>[16]</sup>对包括湖北黄精在内的6种黄精属植物进行了简单重复序列标记(SSR)遗传变异分析,龙炳宏等<sup>[17]</sup>使用DNA条形码(DNAbarcode)技术对湖北黄精等6种黄精属植物进行了物种鉴定和遗传多样性分析,但湖北黄精的转录组学研究尚属空白。

转录组学是研究细胞中基因转录的情况及调控规律的学科,转录组测序技术作为转录组学研究中最重要的手段,已经广泛应用于植物学各领域的研究之中<sup>[18]</sup>,用于新基因的发掘、标记基因的开发、转录图谱的绘制和生物代谢途径的确定等<sup>[19]</sup>。随着测序技术的更迭,转录组学研究在黄精中的应用

逐渐增多,目前已有国内外学者对多花黄精、滇黄精、黄精的多糖、皂苷、黄酮类化合物代谢通路进行相关基因的发掘,也进行了部分候选基因的鉴定和验证<sup>[20-25]</sup>,但对于湖北黄精的相关研究未见报道,显著影响不同品种黄精多糖含量差异的关键酶尚不清楚,本文通过湖北黄精、多花黄精、滇黄精根茎的转录组学分析,填补了湖北黄精转录组学研究的空白,探索显著影响不同品种黄精多糖含量差异的关键酶,对关键酶基因的氨基酸序列进行了深入分析,为黄精属植物多糖合成机制和转录组学深入研究提供了重要参考。

## 1 材料

**1.1 药材** 所有黄精样品均采集自四川省成都市农林科学院黄精培育基地,采集时间2021年7月14日。样品清洗擦干表面水分后将迅速放入-80℃冰箱保存,所有样品生长周期为3年,长势良好形态相近且无病虫害现象,采用完全相同的培育条件,所采集样品经成都中医药大学龙飞教授鉴定。黄精样品分为3组,分别为S1(3批湖北黄精,P1~P3),S2(3批多花黄精,P4~P6),S3(3批滇黄精,P7~P9)。

**1.2 仪器及试剂** UV-3100PC型紫外分光光度计(上海美谱达公司);QuantStudio™ 1 Plus实时荧光定量聚合酶链式反应(Real-time PCR)仪(美国ABI公司);A300型PCR仪(杭州朗基科学仪器有限公

司);Centrifuge 5428 R型离心机、移液器(德国Eppendorf公司);Minip-2500型板式离心机(杭州米欧);MC-15K型手掌离心机(北京擎科生物科技有限公司);JY600C型电泳仪、JY04S-3C型凝胶成像仪(北京君意公司);V6RTAX-5型螺旋震荡仪(海门市其林贝尔仪器制造有限公司);Epoch型微量核酸定量仪(美国BioTek公司);HiPure Universal RNA Kit通用型RNA小提试剂盒(广州美基生物科技有限公司,批号RIK30-01),TIANSeq mRNA捕获试剂盒(天根生化科技优先公司,批号X0701);多糖多酚植物总RNA提取试剂盒、Real-time PCR酶、Goldenstar RT6 cDNA Synthesis Mix逆转录酶(北京擎科生物科技有限公司,批号分别为TSP412、TSE202、TSK314S);浓硫酸、蒽酮、无水乙醇、三氯甲烷、异丙醇等为分析纯级别国药试剂。

## 2 方法

**2.1 RNA的提取和浓度检测** 采用通用型RNA小提试剂盒提取新鲜黄精根茎的RNA,利用Qubit™4 Fluorometer测量RNA浓度并筛选合格样品,再采用TIANSeq mRNA捕获试剂盒捕获样品的mRNA进行后续处理。

**2.2 转录组测序组装** 基于边合成边测序(SBS)技术,使用Illumina NovaSeq高通量测序平台对cDNA文库进行测序,对Raw Data进行数据过滤,去除其中的接头序列及低质量Reads获得高质量的Clean Data,使用Trinity软件进行序列组装。

**2.3 Unigene的功能注释及差异表达基因的筛选与分析** 使用DIAMOND软件将拼接后的Unigene序列与Nr、Swiss-Prot、COG、KOG、eggNOG、京都基因与基因组百科(KEGG)等数据库比对,再用KOBAS得到Unigene在KEGG中的KEGG Orthology结果,InterProScan利用InterPro整合的数据库分析新基因的基因本体(GO)Orthology结果,预测完Unigene的氨基酸序列之后使用HMMER软件与Pfam数据库比对。对于有生物学重复的样本根据各样本Unigene表达量矩阵采用DEseq2<sup>[26]</sup>进行差异表达分析,采用了公认有效的Benjamini-Hochberg方法对原有假设检验得到的显著性P值( $P_{\text{value}}$ )进行校正,并最终采用校正后的P值,即错误发现率(FDR)作为差异表达基因筛选的关键指标,以降低对大量基因的表达值进行独立的统计假设检验带来的假阳性。在筛选过程中,将FDR<0.01且差异倍数(FC)≥2作为筛选标准。其中,FC表示两样品(组)间表达量的比值,对差异表达基因进行模式聚类、功能注

释及富集性分析。

## 2.4 多糖代谢关键酶基因的筛选和表达量分析

从多糖代谢相关通路中筛选出与多糖生物合成相关的关键酶基因进行样本间的差异统计,依据基因表达量绘制聚类热图,聚类方法为系统聚类法,距离指标选择Euclidean,标准化方向选择row,分析样本间的差异性,同时绘制标准化方向为column的热图,其他指标不变,以此来分析样本内不同关键酶基因的表达情况,并根据表达同种关键酶的差异表达基因表达量进行不同样本中总表达量的统计数据,绘制峰峦图。

**2.5 多糖含量测定及与多糖代谢关键酶的相关性分析** 采用紫外分光光度法按2020年版《中华人民共和国药典》(简称《中国药典》)<sup>[1]</sup>中黄精多糖测定方法对黄精根茎进行测定。首先进行标准曲线建立,精密称取无水葡萄糖对照品33 mg溶于100 mL量瓶,配制成 $0.33 \text{ g} \cdot \text{L}^{-1}$ 的葡萄糖对照品溶液。分别取0.2、0.3、0.4、0.5、0.6、0.7、0.8 mL对照品溶液于10 mL具塞试管中,加蒸馏水至2 mL,摇匀,试管置于冰水浴中,加入质量分数为0.2%蒽酮-硫酸溶液8 mL,摇匀,放冷后于100 °C沸水中保温10 min,取出放入冰水浴中10 min,取出后恢复至室温,于582 nm波长下测定吸光度。以对照品溶液的质量浓度为横坐标,吸光度为纵坐标,绘制标准曲线,然后按2020年版《中国药典》通则进行方法学考察。得到样品多糖含量后使用SPSS软件结合多糖代谢关键酶表达量进行相关性分析。

## 2.6 多糖代谢关键酶基因的Real-time PCR验证

筛选出表达量差异较大的多糖代谢关键酶基因,根据转录组测序获得的基因序列设计扩增引物,以18S rRNA为内参基因<sup>[27]</sup>,进行Real-time PCR验证,引物由擎科生物有限公司进行设计与合成,见表1。采用多糖多酚植物总RNA提取试剂盒提取样品RNA,纯度测定后以逆转录试剂盒进行逆转录扩增,扩增得到的cDNA产物4倍稀释后作为Real-time PCR模板,以擎科2×T5 Fast Real-time PCR Mix(SYBR Green I)进行扩增,扩增体系各组分分别为2×T5 Fast Real-time PCR Mix(SYBR Green I) 10 μL、10 μmol·L<sup>-1</sup> Primer F 1 μL、10 μmol·L<sup>-1</sup> Primer R 1 μL、Template(cDNA) 1 μL、ddH<sub>2</sub>O 7 μL,扩增程序为95 °C预变性2 min、循环数1;温度95 °C变性15 s;60 °C退火20 s;温度72 °C延伸20 s、循环数40;熔解阶段:温度95 °C变性时间15 s、循环数1。相对定量结果计算采用2<sup>-ΔΔC<sub>t</sub></sup>法。

表1 引物序列

Table 1 Primer sequence

ID	名称	5'-3'	长度/bp
内参	18SrRNA	上游 AACGGCTACCACATCCAAGG 下游 ACCAGACTTGCCCTCCAATG	
c242090.graph_c0	HK	上游 AAGCCTGAATCCTGGTGAGC 下游 TGCATGGAAGACACAACAGGT	998
c255269.graph_c0	scrK	上游 AGGAGCCTGCGTTAAGTTGG 下游 AGCACAGTGAAGCGAACACT	1 577
c248696.graph_c0	sacA	上游 CATACTCAGCCCATTCCC 下游 GTCCCTTAATCGGTCAGGGC	1 437
c165359.graph_c0	sacA	上游 CGGCCAAGAGTGTGAAGACA 下游 GCAACAGGGCAAATCAAAGC	376
c243229.graph_c1	sacA	上游 TACACGCGAGAAATGGTGCT 下游 AGGTCATCAAGGGCCAATCG	799
c232351.graph_c2	USP	上游 CTTGTGAGAGATGCCTTGGAGA 下游 CACCTTTGGCGATTGAGCTT	460
c220186.graph_c0	UGE	上游 GCTGGGAGTGAAGGCAAAGA 下游 GAACTTTCTGAGGCCGGTGG	601
c252538.graph_c0	MPI	上游 TGTGCGAGCTGGACTTACTC 下游 CCGAAGGTAGCAAGCAGTGT	1 811

2.7 关键酶基因的氨基酸序列特征分析 将筛选验证后得到的关键基因结合转录组测序结果,通过美国国家生物技术信息中心(NCBI, <https://www.ncbi.nlm.nih.gov/orffinder/>)在线工具预测开放阅读框并挑选长度>1 000 bp的基因进行氨基酸序列比对,筛选下载同源基因使用MEGA11软件采用邻接法构建系统进化树分析亲缘演化关系,使用MEME(<http://meme-suite.org/tools/meme>)网站工具预测氨基酸序列 motif,使用NCBI在线CDD(<https://www.ncbi.nlm.nih.gov/Structure/bwrpsb/bwrpsb.cgi>)工具进行氨基酸序列保守结构域预测,使用TBtools软件对结果进行整合,绘制三图合一的氨基酸序列特征图。使用ProtParam(<https://web.expasy.org/protparam/>)在线软件预测蛋白序列等电点和相对分子质量,使用DNAMAN软件对氨基酸序列进行比对分析,再使用SWISS-MODEL(<https://www.swissmodel.expasy.org/>)进行三维蛋白结构构建,分析蛋白结构差异。

### 3 结果与分析

3.1 转录组测序及差异表达基因注释结果 共完成14个样品的转录组测序,共获得121.59 Gb Clean Data,各样品 Clean Data均达到3.75 Gb, Q30碱基百

分比在88.91%及以上,组装共得到100 195条Unigene,获得39 860条Unigene的注释结果,其中长度在1 kb以上的Unigene有27 647条,Unigene的N50为1 260条,组装完整性较高,Unigene在各数据库中的注释结果见表2,从中筛选出需要进行后续分析的样品共9个,涉及3个黄精品种,其差异表达基因在不同数据库中注释到的基因数量见表3。

表2 不同数据库注释的Unigene数量

Table 2 Number of Unigene annotated in different databases 条

数据库	注释数字	300<=长度<1000	长度>=1000
COG	8 535	2 266	5 441
GO	30 021	11 474	14 538
KEGG	23 537	8 435	12 399
KOG	18 684	6 555	10 101
Pfam	22 355	7 013	13 603
Swissprot	20 601	6 816	11 674
TrEMBL	36 930	14 543	17 526
eggNOG	29 864	11 176	15 221
nr	38 369	15 456	17 666
All	39 860	16 213	17 927

3.2 Nr数据库物种分布统计 在Nr数据库中,大部分匹配到的Unigene与芦笋(54.49%)的同源性最

表3 不同数据库注释的差异表达基因数量

Table 3 Number of differentially expressed genes annotated in different databases

数据库	S1与S2比较	S1与S3比较	S2与S3比较
COG	878	984	666
GO	3 453	3 587	2 868
KEGG	2 848	2 929	2 345
KOG	2 140	2 267	1 689
Pfam	2 692	2 883	2 182
Swissprot	2 462	2 666	1 989
TrEMBL	4 375	4 478	3 675
eggNOG	3 636	3 718	3 032
Nr	4 406	4 490	3 699
All	4 569	4 616	3 835

接近,油棕(5.89%)和海枣(5.06%)次之,同时与芭蕉科类的植物也具有一定同源性,见表4和增强出版附加材料。

表4 基因同源性对照

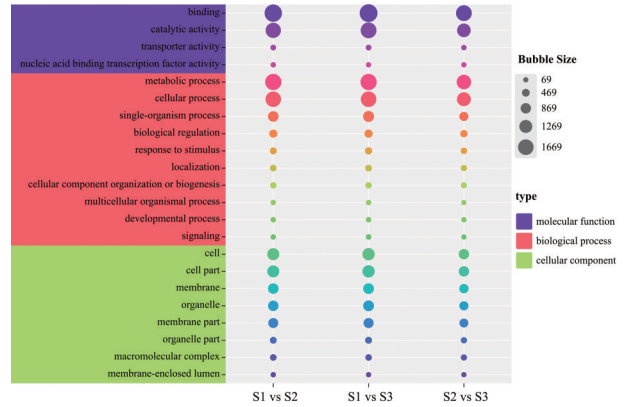
Table 4 Gene homology control

物种名称	同源性比例/%
芦笋( <i>Asparagus officinalis</i> )	54.49
油棕( <i>Elaeis guineensis</i> )	5.89
海枣( <i>Phoenix dactylifera</i> )	5.06
外瓶霉属( <i>Exophiala aquamarina</i> )	1.39
菠萝( <i>Ananas comosus</i> )	1.25
小果野蕉( <i>Musa acuminata</i> )	1.22
铁皮石斛( <i>Dendrobium catenatum</i> )	1.11
象腿蕉属( <i>Ensete ventricosum</i> )	0.99
葡萄( <i>Vitis vinifera</i> )	0.98
野蕉( <i>Musa balbisiana</i> )	0.85
other(其他)	26.78

**3.3 差异表达基因统计分析** 筛选出S1与S2差异表达基因共有4 569个,上调基因2 387个(52.24%),下调基因2 182个(47.76%),S1与S3差异表达基因共有4 616个,上调基因2 714个(58.80%),下调基因1 902个(41.20%),S2与S3差异表达基因共有3 835个,上调基因2 226个(58.04%)下调基因1 609个(41.96%),3种黄精差异表达基因的差异分布情况,见增强出版附加材料。

**3.4 差异表达基因GO分类及富集分析** 根据Unigene在GO数据库中的比对结果,S1与S2、S1与S3、S2与S3分别有3 453、3 587、2 868个差异表达

基因注释到生物过程、细胞组分、分子功能3个GO分类的60个亚类之中,选择各样本组差异表达基因数目排名前20的GO功能绘制气泡图,共涉及到22个GO功能见图1,以结合功能、催化活性、代谢过程和细胞组分差异最显著。



注:GO功能分别为结合(binding)、催化活性(catalytic activity)、转运蛋白活性(transporter activity)、核酸结合转录因子活性(nucleic acid binding transcription factor activity)、代谢过程(metabolic process)、细胞过程(cellular process)、单一生物体过程(single-organism process)、生物调控(biological regulation)、应激性(response to stimulus)、定位(localization)、细胞成分组织或生物发生(cellular component organization or biogenesis)、多细胞生物体过程(multicellular organismal process)、发育过程(developmental process)、信号(signaling)、细胞(cell)、细胞部分(cell part)、膜(membrane)、细胞器(organelle)、膜部分(membrane part)、细胞器部分(organelle part)、大分子复合体(macromolecular complex)、膜封闭腔体(membrane-enclosed lumen)

图1 差异表达基因GO功能分类统计

Fig. 1 GO functional classification statistics of differentially expressed genes

**3.5 KEGG通路统计分析** S1与S2共有1 425个差异表达基因被注释到了KEGG的129条通路中,其余为S1与S3(1 428个,128条),S2与S3(1 175个,127条)。主要有细胞过程(cellular process)、环境信息处理(environmental information processing)、遗传信息处理(genetic information processing)、代谢(metabolism)、有机系统(organic system)5大类,差异基因数目及比例见表5,不同品种间在新陈代谢相关的通路差异基因数目最多。

**3.6 KEGG通路富集分析** 对样品组差异表达基因进行KEGG通路富集,选择富集到的前20条通路进行分析,见增强出版附加材料。从各样品组差异表达基因富集的前20条通路来看,S1与S2比较与多糖代谢有关的通路为淀粉与蔗糖代谢通路,S2与S3比较为淀粉与蔗糖代谢通路与半乳糖代谢通路。

表5 差异表达基因KEGG通路统计

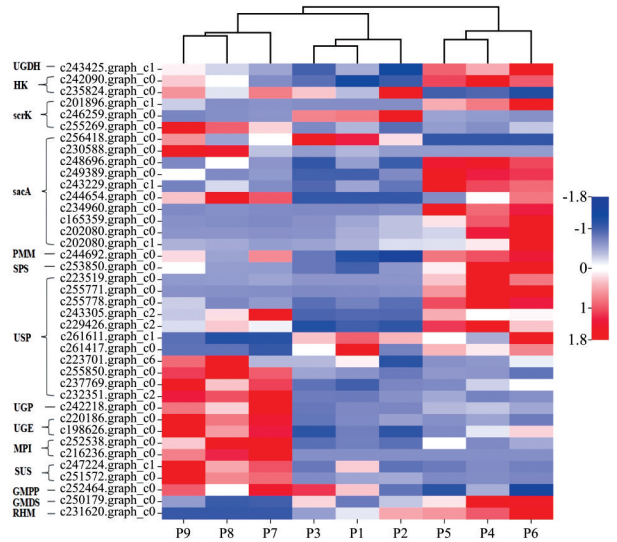
Table 5 Statistics of differentially expressed gene in KEGG pathway

通路名称	S1与S2		S1与S3		S2与S3	
	数目/条	比例/%	数目/条	比例/%	数目/条	比例/%
细胞过程	60	3.92	57	3.81	42	3.24
环境信息处理	92	6.01	86	5.75	87	6.72
遗传信息处理	445	29.07	382	25.53	345	26.64
新陈代谢	803	52.45	834	55.75	679	52.43
有机系统	131	8.56	137	9.16	142	10.97

以上结果说明,这3种黄精多糖代谢在这两条通路上差异性最显著,在后续多糖代谢的研究中可作为主要研究通路。

**3.7 多糖代谢关键酶差异表达基因分析** 黄精中含有果糖、葡萄糖、甘露糖、半乳糖、阿拉伯糖、鼠李糖<sup>[28-29]</sup>等单糖,这些单糖经过糖基转移酶等一系列酶合成黄精多糖,本次研究结果表明3种黄精在多糖代谢途径基因表达上有较大差异,在淀粉与蔗糖代谢途径(ko00500)、氨基糖和核苷酸糖代谢途径(ko00520)、果糖和甘露糖代谢途径(ko00051)、半乳糖代谢途径(ko00052)中共鉴定出以下与多糖合成相关的关键酶,UDP-葡萄糖脱氢酶(UGDH)、己糖激酶(HK)、果糖激酶(scrK)、β-果糖苷酶(sacA)、磷酸甘露糖异构酶(PMM)、蔗糖磷酸合成酶(UPS)、UDP糖焦磷酸化酶(USP)、UTP-葡萄糖-1-磷酸尿苷转移酶(UGP)、UDP-葡萄糖醛酸4-外聚酶(UGE)、甘露糖-6-磷酸异构酶(MPI)、蔗糖合成酶(SUS)、GDP-甘露糖焦磷酸化酶(GMPP)、GDP甘露糖4,6-脱水酶(GMDS)、UDP-鼠李糖合成酶(RHM)<sup>[30-32]</sup>。根据DEseq2处理后的基因表达量数据绘制聚类热图,见图2,结果显示3种黄精之间差异显著,可根据这些关键酶基因的表达量差异将黄精进行初步分类,而同种黄精之间差异并不显著,多花黄精UGDH、PMM、SPS、sacA、RHM、GMDS的基因表达量显著高于湖北黄精与滇黄精,滇黄精MPI、UGP、SUS、UGE、GMPP的基因表达量显著高于多花黄精与湖北黄精,图3表明在同一样品内sacA基因c249389.graph\_c0、c248696.graph\_c0、c243299.graph\_c0的表达量显著高于大部分酶基因,热图结果均为相对表达量情况,根据图4总表达量结果综合分析推测出不管是种间还是种内,sacA均可能是影响黄精多糖含量最主要的酶类。

**3.8 多糖含量测定结果与多糖代谢关键酶的分析** 由吸光度数据计算得到线性回归方程为  $Y=$



注:绝对值越大表示差异性越大,0表示无差异性,红色代表高于平均表达量,蓝色代表低于平均表达量(图3同)

图2 多糖代谢关键酶基因表达量聚类热图

Fig. 2 Clustering Heat map of gene expression of key enzymes for polysaccharide metabolism

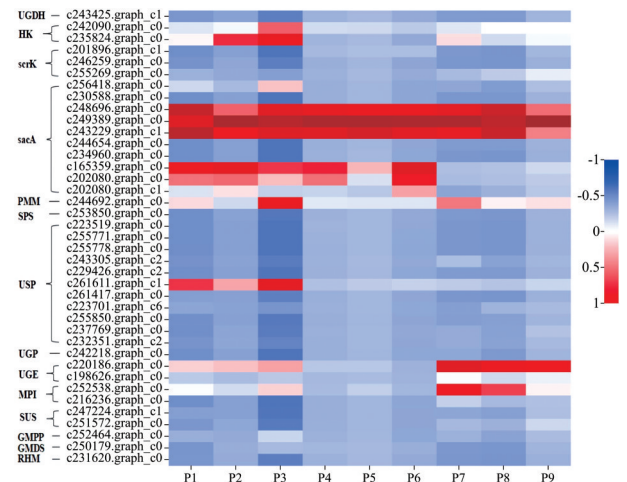


图3 多糖代谢关键酶基因表达量聚类热图

Fig. 3 Clustering Heat map of gene expression of key enzymes for polysaccharide metabolism

0.031 7X+0.050 3 ( $R^2=0.999 1$ ) 线性范围 (0.003 3~0.036 3  $g \cdot L^{-1}$ ),方法学考察结果,精密度、稳定性、重复性、加样回收率、平均加样回收率分别为0.19%、0.39%、0.75%、2.95%、104.62%。P1~P9得到多糖质量分数测定结果分别为9.76%、8.65%、9.01%、19.63%、17.94%、19.45%、12.63%、11.28%、13.66%,其中多花黄精含量最高,滇黄精含量次之,湖北黄精含量最低。使用SPSS软件将各样品多糖代谢关键酶表达量与多糖含量进行相关性分析,结果表明UGDH、sacA、PMM、SPS、scrK、USP、GMPP的表达量与样本多糖含量成显著正相关,见表6。

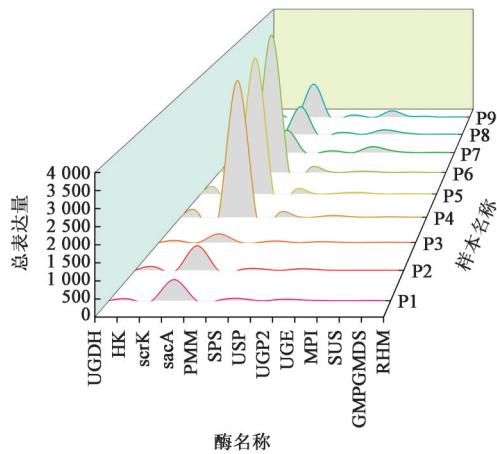


图4 总表达量峰峦

Fig. 4 Total expression peaks map

**3.9 Real-time PCR 验证结果** 以18SrRNA为内参基因,对8个筛选出的多糖代谢关键酶基因进行Real-time PCR验证,分析目的基因在不同品种中的表达差异,与转录组测序结果进行比较,验证实验结果可靠性。结果显示sacA、MPI在不同品种中的基因表达量差异与测序结果具有一致性,而在相关性分析中并未发现MPI与多糖含量存在显著相关性,无法推测出MPI对黄精多糖含量有显著影响,故未作进一步的研究分析。见图5。

**3.10 scaA及其同源基因序列分析** 根据前文中得

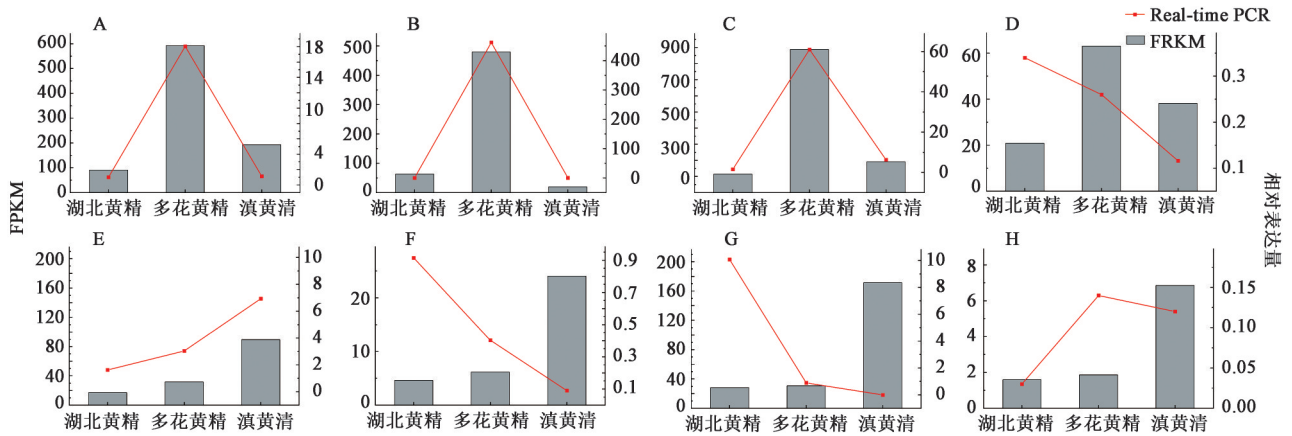
到的研究结果进一步从黄精转录组的sacA基因中筛选出6个开放阅读框长度>1 000 bp的关键基因进行深入分析,对其同源基因构建系统进化树,并预测其motif区段和保守结构域,结果见图6,主要将这些基因归属于4个家族,依据亲缘关系主要分为2大类,与motif预测结果具有一致性。其中c230588.graph\_c0、c256418.graph\_c0、c244654.graph\_c0同属于DUF247家族,与深圳拟兰*Apostasia shenzhenica*、黄连*Coptis chinensis*、杨柳*Populus deltoides*、独叶草*Kingdonia uniflora*等植物有较近的亲缘关系,c234960.graph\_c0、c249389.graph\_c0、c248696.graph\_c0同属于糖苷水解酶家族(包括GH32\_Fruct1-like、Glyco\_32),并具有糖苷水解酶家族的功能结构域,与芦笋*Asparagus officinalis*、太医龙舌兰*Agave tequilana*、洋葱*Allium cepa*、小麦*Triticum aestivum*、冰草*Agropyron cristatum*、芦荟*Aloe vera*等植物有较近的亲缘关系,并且比对结果显示c249389.graph\_c0、c248696.graph\_c0的氨基酸序列分别与果糖1-果糖基转移酶XP\_020240888.1(芦笋)、6(G)-果糖基转移酶ABS72186.1(太医龙舌兰)有较高匹配率(>85%),c234960.graph\_c0与β-果糖苷酶蛋白KAH7668350.1(参薯)有较高匹配率(69%)。对

表6 多糖含量与酶表达量相关性数据

Table 6 Correlation data between polysaccharide content and enzyme expression

相关性	多糖含量	UGDH	HK	scrK	sacA	PMM	SPS	USP	UGP2	UGE	MPI	SUS	GMPP	GMDS	RHM
多糖含量	1	0.896 <sup>2)</sup>	0.104	0.696 <sup>1)</sup>	0.957 <sup>2)</sup>	0.922 <sup>2)</sup>	0.908 <sup>2)</sup>	0.696 <sup>1)</sup>	0.011	-0.075	-0.039	-0.173	0.741 <sup>1)</sup>	-0.544	0.638
UGDH	0.896 <sup>2)</sup>	1	-0.142	0.823 <sup>2)</sup>	0.881 <sup>2)</sup>	0.834 <sup>2)</sup>	0.833 <sup>2)</sup>	0.893 <sup>2)</sup>	-0.082	-0.102	-0.040	-0.145	0.626	-0.583	0.638
HK	0.104	-0.142	1	0.157	0.061	0.103	0.176	-0.296	0.254	0.373	0.097	0.310	0.082	-0.080	-0.014
scrK	0.696 <sup>1)</sup>	0.823 <sup>2)</sup>	0.157	1	0.581	0.718 <sup>1)</sup>	0.707 <sup>1)</sup>	0.738 <sup>1)</sup>	0.171	0.352	0.150	0.355	0.389	-0.285	0.300
sacA	0.957 <sup>2)</sup>	0.881 <sup>2)</sup>	0.061	0.581	1	0.801 <sup>2)</sup>	0.886 <sup>2)</sup>	0.692 <sup>1)</sup>	-0.224	-0.317	-0.210	-0.400	0.790 <sup>1)</sup>	-0.741 <sup>1)</sup>	0.790 <sup>1)</sup>
PMM	0.922 <sup>2)</sup>	0.834 <sup>2)</sup>	0.103	0.718 <sup>1)</sup>	0.801 <sup>2)</sup>	1	0.793 <sup>1)</sup>	0.658	0.334	0.180	0.230	0.041	0.575	-0.253	0.401
SPS	0.908 <sup>2)</sup>	0.833 <sup>2)</sup>	0.176	0.707 <sup>1)</sup>	0.886 <sup>2)</sup>	0.793 <sup>1)</sup>	1	0.759 <sup>1)</sup>	-0.153	-0.190	-0.233	-0.270	0.885 <sup>2)</sup>	-0.603	0.769 <sup>1)</sup>
USP	0.696 <sup>1)</sup>	0.893 <sup>2)</sup>	-0.296	0.738 <sup>1)</sup>	0.692 <sup>1)</sup>	0.658	0.759 <sup>1)</sup>	1	-0.116	-0.191	-0.134	-0.218	0.580	-0.528	0.685 <sup>1)</sup>
UGP2	0.011	-0.082	0.254	0.171	-0.224	0.334	-0.153	-0.116	1	0.860 <sup>2)</sup>	0.858 <sup>2)</sup>	0.718 <sup>1)</sup>	-0.454	0.577	-0.570
UGE	-0.075	-0.102	0.373	0.352	-0.317	0.180	-0.190	-0.191	0.860 <sup>2)</sup>	1	0.790 <sup>1)</sup>	0.961 <sup>2)</sup>	-0.532	0.579	-0.693 <sup>1)</sup>
MPI	-0.039	-0.040	0.097	0.150	-0.210	0.230	-0.233	-0.134	0.858 <sup>2)</sup>	0.790 <sup>1)</sup>	1	0.640	-0.567	0.411	-0.612
SUS	-0.173	-0.145	0.310	0.355	-0.400	0.041	-0.270	-0.218	0.718 <sup>1)</sup>	0.961 <sup>2)</sup>	0.64	1	-0.578	0.603	-0.724 <sup>1)</sup>
GMPP	0.741 <sup>1)</sup>	0.626	0.082	0.389	0.790 <sup>1)</sup>	0.575	0.885 <sup>2)</sup>	0.580	-0.454	-0.532	-0.567	-0.578	1	-0.570	0.832 <sup>2)</sup>
GMDS	-0.544	-0.583	-0.080	-0.285	-0.741 <sup>1)</sup>	-0.253	-0.603	-0.528	0.577	0.579	0.411	0.603	-0.570	1	-0.864 <sup>2)</sup>
RHM	0.638	0.638	-0.014	0.300	0.790 <sup>1)</sup>	0.401	0.769 <sup>1)</sup>	0.685 <sup>1)</sup>	-0.570	-0.693 <sup>1)</sup>	-0.612	-0.724 <sup>1)</sup>	0.832 <sup>2)</sup>	-0.864 <sup>2)</sup>	1

注: <sup>1)</sup>P<0.05, <sup>2)</sup>P<0.01



注:柱形图代表目的基因的FPKM值,折线表示目的基因的Real-time PCR相对表达量

图5 8个关键酶基因表达量与Real-time PCR相对表达量对比

Fig. 5 Expression of 8 key enzyme genes compared with Real-time PCR relative expression

6个关键基因蛋白序列进行理化性质预测,见表7。结果显示DUF247家族的3个基因不稳定系数在41.38以上,属于不稳定型蛋白,而糖基转移酶家族的3个基因属于稳定型蛋白,DNAman比对结果表明这3个基因氨基酸序列与芦笋和太匮龙舌兰的同源序列一致性为64.6%,中间有多段保守序列(黑色标注

部分),对这3个基因进行同源建模进一步分析蛋白结构,见图7。虽然这3个基因表现出氨基酸序列多样性,但其蛋白质空间结构仍然保守,具有相似的结构和功能,均有4~5个 $\alpha$ 螺旋,而c248696.graph\_c0相较于其他2个基因在左边缺少一部分 $\beta$ 折叠结构,3个基因的螺旋结构图见增强出版附加材料。

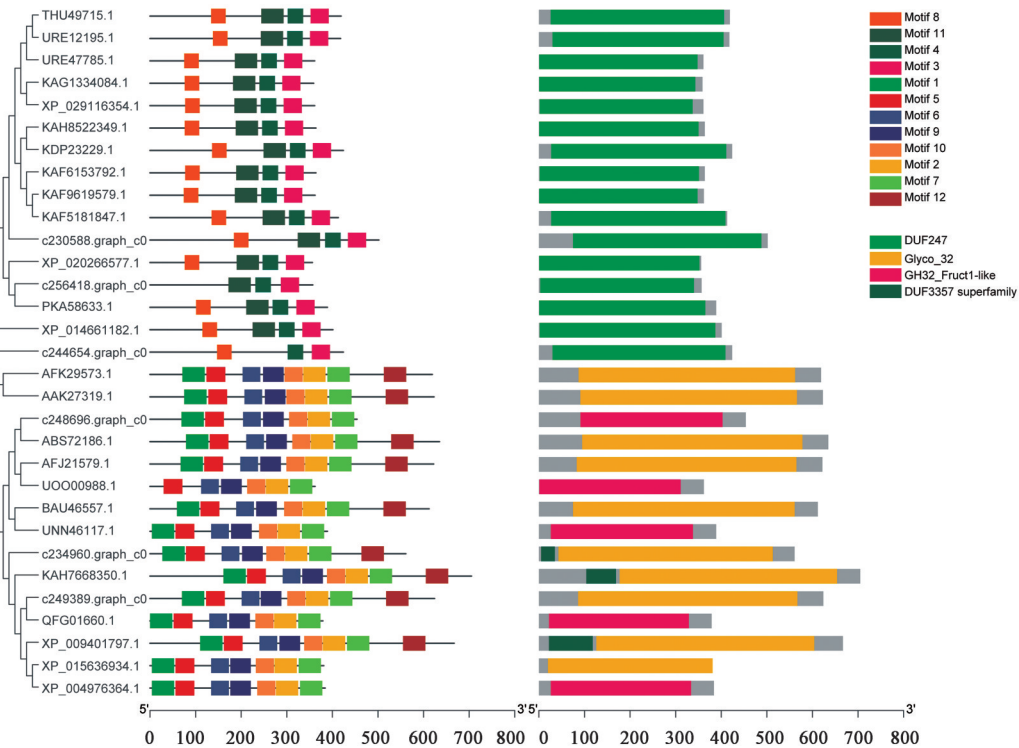


图6 关键基因的氨基酸序列分析

Fig. 6 Amino acid sequence analysis of key genes

#### 4 讨论

多糖作为黄精的主要药效成分,具有抗疲劳、抗氧化、降血糖等重要功效,2020年版《中国药典》规定其含量不低于7%<sup>[1]</sup>,由于不同品种黄精多糖含

量差异较大,同时环境差异也会造成相应影响<sup>[33-34]</sup>,为了控制环境变量对实验结果的影响,本文采集相同产地同等培育条件下的3种黄精进行实验,探究影响不同品种黄精多糖含量的关键酶,对湖北黄

表7 关键基因蛋白序列理化性质

Table 7 Physicochemical properties of key gene protein sequences

基因ID	等电点	相对分子质量/Da	不稳定系数	带正电荷的残基总数/个	带负电荷的残基总数/个
c248696.graph_c0	5.37	50 091.46	34.38	39	52
c234960.graph_c0	8.22	63 861.57	30.90	71	68
c249389.graph_c0	5.15	69 809.03	34.20	47	70
c256418.graph_c0	5.96	41 520.03	41.38	40	47
c244654.graph_c0	5.38	49 165.02	52.03	49	61
c230588.graph_c0	7.07	57 396.87	42.08	61	62

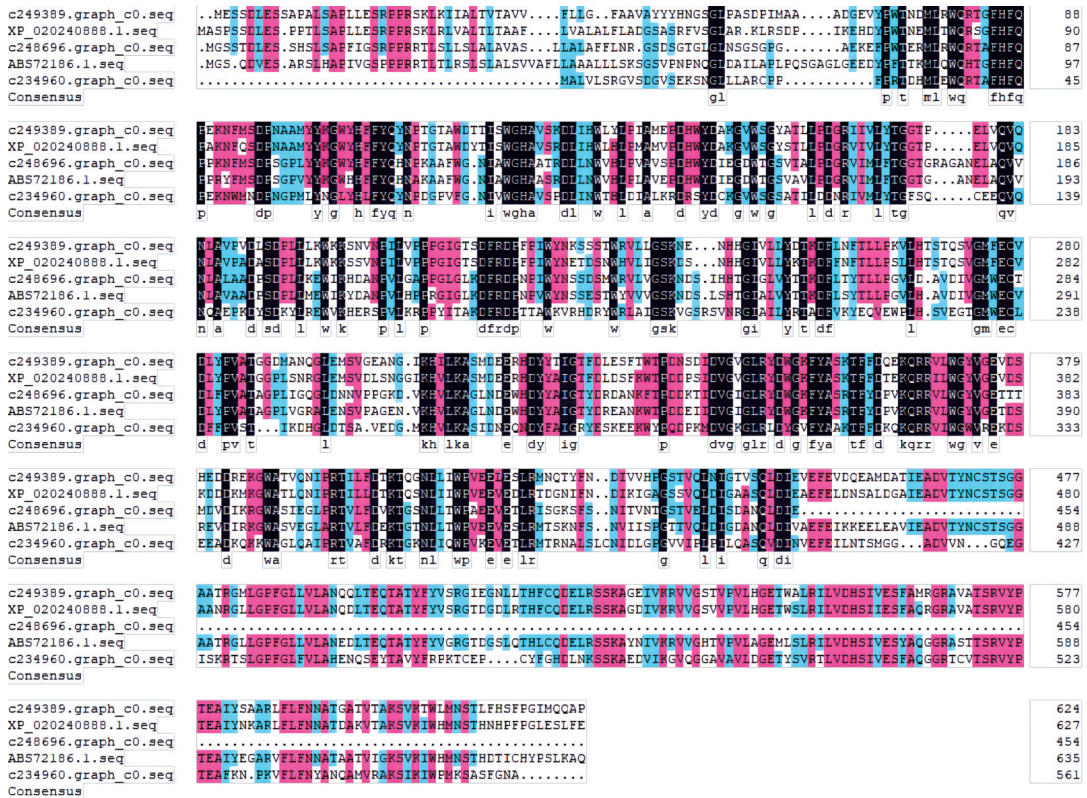


图7 DNAMAN 比对结果

Fig. 7 DNAMAN matching results

精、滇黄精、多花黄精的根茎转录组数据进行系统分析,并对测序结果进行质量评估,确保了后续数据分析的可靠性,主要结合了Nr、GO、KEGG 3大数据库对基因进行注释及分析。

Nr数据库分析结果表明黄精与芦笋等植物的基因同源性较高,经GO与KEGG富集分析、文献研究筛选、热图分析与Real-time PCR验证实验结合差异基因的数目和表达量分析推测出sacA是影响湖北黄精、多花黄精、滇黄精多糖含量的主要关键酶,经KEGG代谢通路图发现在4条与多糖代谢相关的通路中仅在淀粉与蔗糖代谢通路(ko00500)与半乳糖代谢通路(ko0052)中发现该种酶,同时富集分析结果也显示3种黄精在这2条通路上差异性显

著,也可证明该结论的可靠性,而相关性分析结果表明UGDH、PMM、SPS、scrK、USP、GMPP这几种酶对多糖含量也有显著影响,但是Real-time PCR验证实验结果并未证明这个结论准确性,同时从其表达量来看也远低于sacA,所以本文对sacA基因进行了更深入地研究分析。

对sacA基因氨基酸序列进行同源基因系统进化树构建、motif区段与保守结构域预测,进一步证明黄精中的3个sacA基因与果糖基转移酶的结构与功能极其相似。ZHANG等<sup>[35]</sup>从多花黄精中分离得到了2种多糖,分别为果聚糖DP1和半乳聚糖PP1,理化性质和结构鉴定表明DP1组成为果糖和葡萄糖,峰面积百分比分别为96.29%与3.71%,表

明DP1主要由大量果糖通过不同类型的糖苷键连接而成,而PP1主要由半乳糖通过糖苷键连接而成,从分离纯化得率来看DP1(11.40%)含量远高于PP1(0.89%)。ZHAO等<sup>[36]</sup>的研究结果也表明黄精中的多糖主要由果糖及半乳糖组成,果聚糖含量显著高于其他类型多糖,由于聚糖不是像蛋白质和核酸那样从基因组编码的模板直接合成的,其保真度取决于糖基转移酶的特异性,这些糖基转移酶通过一次添加一个糖来合成复杂的聚糖,为了形成新的糖苷键,糖基转移酶与受体底物结合,并定向一个特定的羟基,通常是许多羟基中的一个,以攻击供体糖的异构碳。最近几个具有结合受体底物的糖基转移酶的晶体结构显示,这些酶具有共同的核心结构,作为支架的功能,在其上插入可变环,以赋予底物特异性并正确引导亲核羟基,受体结合位点组装的不同方法表明,这些环状区域的持续进化为生物学中观察到的各种糖结构的组装提供了模板<sup>[37]</sup>。宋春丽等<sup>[38]</sup>以pET28a为载体实现了黑曲霉来源的果糖基转移酶基因 $fru$ 在大肠埃希菌BL21(DE3)中的异源表达,并研究分析了纯化后重组果糖基转移酶的酶学性质,大幅度提高果糖基转移酶的产量,且酶活力也有所提高,有利于扩大工业上低聚果糖生产规模,提高产量。结合2个相关代谢通路中 $sacA$ 的上游及下游化合物来看,推测黄精的 $sacA$ 基因也可将蔗糖与棉子糖等作为底物催化生成更为复杂的果聚糖。

转录组学研究可以从基因的转录调控及表达对品种差异做出合理的阐释,本文通过 $sacA$ 高表达基因的结构与功能分析推测出 $sacA$ 是不同品种黄精多糖含量差异的主要原因,同时揭示了 $sacA$ 是黄精多糖主要为果聚糖的关键影响因素,可进一步根据 $sacA$ 高表达基因序列设计引物进行质粒构建,通过异源表达<sup>[39]</sup>验证其合成果聚糖的功能,为黄精多糖合成机制的深入研究奠定基础,也为黄精属植物的遗传改良和质量评价提供了重要参考。

[利益冲突] 本文不存在任何利益冲突。

#### [参考文献]

- [1] 国家药典委员会. 中华人民共和国药典:一部[M]. 北京:中国医药科技出版社,2020:1088.
- [2] 陶弘景. 名医别录[M]. 尚志钧,校注. 北京:人民卫生出版社,1986:23.
- [3] 雷高明,冯卫生,冯云霞,等. 酒黄精饮片质量标准研究[J]. 中药材,2011,34(9):1346-1348.
- [4] 崔阔澍,肖特,李慧萍,等. 我国黄精种质资源研究进

- 展[J]. 江苏农业科学,2021,49(11):35-39.
- [5] 董琦,钱红月,王萍,等. 江西省黄精资源调查与分析[J]. 亚太传统医药,2021,17(12):13-15.
- [6] 倪天宇,罗晓朦,张春椿,等. 不同生境下多花黄精化学成分比较[J]. 中成药,2020,42(11):2948-2953.
- [7] 夏从龙. 云南道地药材——滇黄精[J]. 大理大学学报,2022,7(2):102.
- [8] 任洪民,邓亚玲,张金莲,等. 药用黄精炮制的历史沿革、化学成分及药理作用研究进展[J]. 中国中药杂志,2020,45(17):4163-4182.
- [9] 陈祖海,曾岳明,李永和,等. 多花黄精‘丽精1号’氨基酸组成及营养价值分析[J]. 中药材,2021,44(9):2153-2156.
- [10] LIU J, LI T, CHEN H, et al. Structural characterization and osteogenic activity *in vitro* of novel polysaccharides from the rhizome of *Polygonatum sibiricum* [J]. Food Function, 2021, doi: 10.1039/d1fo00938a.
- [11] CHAI Y, LUO J, BAO Y. Effects of *Polygonatum sibiricum* saponin on hyperglycemia, gut microbiota composition and metabolic profiles in type 2 diabetes mice [J]. Biomed Pharmacother, 2021, doi: 10.1016/j.biopha.2021.112155.
- [12] LIU B, TANG Y, SONG Z, et al. *Polygonatum sibiricum* F. delarochae polysaccharide ameliorates HFD-induced mouse obesity via regulation of lipid metabolism and inflammatory response [J]. Mol Med Rep, 2021, doi: 10.3892/MMR.2021.12140.
- [13] 王婧,陶爱恩,杨燕,等. 滇黄精中多糖的分离与抗氧化活性研究[J]. 中草药,2021,52(16):4789-4796.
- [14] 张雪可,张虹,章鹏飞,等. 湖北黄精潜在分布区预测及生态适宜性研究[J]. 中国农业科技导报,2021,23(8):185-192.
- [15] 宋歌,夏娇,肖强. 湖北黄精叶片挥发油GC-MS分析及抗氧化活性测定[J]. 湖北民族大学学报:自然科学版,2022,40(4):374-379,384.
- [16] 朱巧,邓欣,张树冰,等. 黄精属6种植物的SSR遗传差异分析[J]. 中国中药杂志,2018,43(14):2935-2943.
- [17] 龙炳宏,蒋向辉,宋荣,等. DNA条形码在黄精属药用植物鉴定与遗传多样性分析中的应用[J]. 植物科学学报,2022,40(4):533-543.
- [18] 蒋费涛,王书平,祁俊生,等. 转录组学技术及其在植物系统学上的研究进展[J]. 现代盐化工,2020,47(4):14-17.
- [19] 王飞,刘林波,高天歌,等. 转录组学在牧草上的应用进展[J]. 草业科学,2019,36(2):402-413.
- [20] LI D, WANG Q, CHEN S, et al. De novo assembly and

- analysis of *Polygonatum cyrtoneuma* Hua and identification of genes involved in polysaccharide and saponin biosynthesis [J]. BMC Genomics, 2022, 23(1):195.
- [21] FENG T, JIANG Y, JIA Q, et al. Transcriptome analysis of different sections of Rhizome in *Polygonatum sibiricum* Red. and mining putative genes participate in polysaccharide biosynthesis [J]. Biochemical genetics, 2022, doi: 10. 1007/s10528-022-10183-x.
- [22] 廖荣俊, 杨阳, 叶碧欢, 等. 多花黄精根茎的转录组分析与甾体皂苷生物合成相关基因发掘[J]. 中国中药杂志, 2020, 45(7):1648-1656.
- [23] 祝明珠, 俞年军, 王秋丽, 等. 基于多花黄精转录组的多糖及薯蓣皂苷生物合成路径研究[J]. 中国中药杂志, 2020, 45(1):85-91.
- [24] 肖韵铮, 韩世明, 秦昭, 等. 滇黄精转录组测序及类黄酮合成相关基因的分析[J]. 河南农业大学学报, 2020, 54(6):931-940.
- [25] 叶碧欢, 杨阳, 朱杰丽, 等. 基于比较转录组学的多花黄精黄酮类化合物合成基因表达分析[J]. 食品与生物技术学报, 2022, 41(4):84-92.
- [26] MICHAEL L, WOLFGANG H, SIMON A. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2 [J]. Genome Biol, 2014, 15(12):550.
- [27] 吕煜梦, 张舒婷, 王雪晶, 等. 多花黄精几丁质诱导赤霉素应答基因(CIGR)克隆及其功能[J]. 应用与环境生物学报, 2020, 26(2):255-263.
- [28] 何连军, 干雅平, 吕伟德, 等. 高效阴离子交换色谱-脉冲安培检测法测定多花黄精多糖的单糖组成[J]. 中草药, 2017, 48(8):1671-1676.
- [29] 杜泽飞, 陶爱恩, 夏从龙, 等. 基于PMP-HPLC和化学计量学的黄精基原物种多糖差异分析[J]. 中国实验方剂学杂志, 2019, 25(15):25-29.
- [30] 叶香媛, 周文彬. 植物果糖激酶研究进展[J]. 科学通报, 2021, 66(22):2820-2831.
- [31] 赵历强, 赵德蕊, 单春苗, 等. 多花黄精果糖激酶和GDP-甘露糖焦磷酸化酶的基因克隆及酶结构性质[J]. 生物学杂志, 2022, 39(4):1-6.
- [32] WANG C, PENG D, ZHU J, et al. Transcriptome analysis of *Polygonatum cyrtoneuma* Hua: Identification of genes involved in polysaccharide biosynthesis [J]. Plant Methods, 2019, doi: 10. 1186/s13007-019-0441-9.
- [33] 苏为耿, 赵永丰, 吴兴兴, 等. 不同种植模式及种植年限对滇黄精产量及品质的影响[J]. 西部林业科学, 2022, 51(5):113-117, 126. .
- [34] 高秋美, 任丽华, 米真如, 等. 不同光照强度对多花黄精生长及光合特性的影响[J]. 山东农业科学, 2021, 53(6):44-47.
- [35] ZHANG J, CHEN H, LUO L, et al. Structures of fructan and galactan from *Polygonatum cyrtoneuma* and their utilization by probiotic bacteria [J]. Carbohydr Polym, 2021, doi:10. 1016/j. carbpol. 2021. 118219.
- [36] ZHAO P, LI X, WANG Y, et al. Characterisation and saccharide mapping of polysaccharides from four common *Polygonatum* spp. [J]. Carbohydrate Polymers, 2020, doi: 10. 1016/j. carbpol. 2020. 115836.
- [37] KELLEY M, ROBERT H. Emerging structural insights into glycosyltransferase-mediated synthesis of glycans. [J]. Nat Chem Biol, 2019, doi: 10. 1038/s41589-019-0350-2.
- [38] 宋春丽, 杨海泉, 马道程, 等. 果糖基转移酶在大肠杆菌中的重组表达及酶学性质[J]. 生物学杂志, 2018, 35(6):11-15.
- [39] NING Y, XU Y, JIAO B, et al. Application of gene knockout and heterologous expression strategy in fungal secondary metabolites biosynthesis [J]. Marine Drugs, 2022, 20(11):705.

[责任编辑 顾雪竹]