

• 药材与资源 •

西伯利亚白刺 *L*-半乳糖-1-磷酸磷酸酶基因克隆与序列分析高卫东^{1,3}, 杜雨晴¹, 张林凤¹, 蒋路园¹, 刘 园⁴, 邱德有¹, 陈贵林^{2*}, 杨艳芳^{1*}

1. 中国林业科学研究院林业研究所, 国家林业和草原局林木培育重点实验室, 北京 100091

2. 内蒙古大学牧草与特色作物生物学教育部重点实验室, 呼和浩特 内蒙古 010020

3. 国家林业和草原局国家公园(自然保护地)发展中心, 北京 100013

4. 北京市海淀区圆明园管理处, 北京 100193

摘要: **目的** 克隆西伯利亚白刺 *Nitraria sibirica* 维生素 C 合成过程中 *L*-半乳糖-1-磷酸磷酸酶 (*L*-galactose-1-phosphate phosphatase, *GPP*) 编码基因。**方法** 利用生物信息学方法对西伯利亚白刺三代转录组数据进行分析, 共鉴定出 2 个 *GPP* 基因, 并从西伯利亚白刺中成功克隆得到 2 条 *NsGPP* 基因序列, 进一步利用生物信息学方法对其理化性质、蛋白结构、保守序列以及系统进化等进行分析, 并结合 RNA-seq 数据分析其在盐胁迫下的表达模式, 利用 qRT-PCR 实验检测其在不同组织中的表达水平。**结果** 从西伯利亚白刺中成功鉴定克隆出 2 个 *NsGPP* 基因, 2 个 *NsGPP* 基因所编码蛋白均含有保守的 FBPase/IMPase/glpX-like (FIG) 结构域, 且与其他植物 *GPP* 蛋白具有较高的相似性。利用多个物种的 *GPP* 蛋白序列构建系统进化树, 结果显示西伯利亚白刺与灌木或半木质化草本植物亲缘关系更近。Motif 分析发现 *NsGPP* 蛋白与拟南芥等 *GPP* 蛋白较为相似, 但也存在差异, 并且 2 个 *NsGPP* 蛋白之间也存在差异。此外, *NsGPP1* 与 *NsGPP2* 的蛋白三维结构之间也存在微小差别。RNA-seq 分析表明, 不同浓度 NaCl 处理下, *NsGPP1* 的表达逐步升高的趋势, 而 *NsGPP2* 的表达量则较低, 且在盐胁迫后其表达水平逐渐降低。qRT-PCR 实验结果显示, 白刺 *GPP* 基因的表达存在组织表达特异性, 且在同一组织中 *NsGPP1* 的表达水平显著高于 *NsGPP2*。**结论** 从西伯利亚白刺中成功克隆到 2 个 *NsGPP* 基因, 其编码蛋白均含有保守的 FIG 结构域, 且与拟南芥等草本植物的 *GPP* 蛋白更为相似。*NsGPP* 基因在不同组织中表达水平具有显著差异, 其中 *NsGPP1* 可能参与西伯利亚白刺维生素 C 生物合成并参与白刺抵抗盐胁迫。为进一步解析白刺维生素 C 生物合成以及耐盐分子机制提供了理论依据。

关键词: 西伯利亚白刺; *L*-半乳糖-1-磷酸磷酸酶基因; 基因鉴定; 生信分析; 基因表达**中图分类号:** R286.12 **文献标志码:** A **文章编号:** 0253-2670(2024)06-2048-09**DOI:** 10.7501/j.issn.0253-2670.2024.06.026Cloning and sequence analysis of *L*-galactose-1-phosphate phosphatase gene from *Nitraria sibirica*GAO Weidong^{1,3}, DU Yuqing¹, ZHANG Linfeng¹, JIANG Luyuan¹, LIU Yuan⁴, QIU Deyou¹, CHEN Guilin², YANG Yanfang¹

1. Research Institute of Forestry, Chinese Academy of Forestry, State Key Laboratory of Forest Genetics and Breeding, Key Laboratory of Forest Cultivation, National Forestry and Grassland Administration, Beijing 100091, China

2. Key Laboratory of Forage and Endemic Crop Biology (Inner Mongolia University) Ministry of Education, Hohhot 010020, China

3. National Forestry and Grassland Authority National Park (Nature Reserve) Development Centre, Beijing 100013, China

4. Beijing Haidian District Yuanmingyuan Management Office, Beijing 100193, China

Abstract: Objective The aim of this study is to clone the gene encoding the *L*-galactose-1-phosphate phosphatase (*GPP*), an**收稿日期:** 2023-08-05**基金项目:** 牧草与特色作物生物学教育部重点实验室开放课题资助 (FECBOF2021010)**作者简介:** 高卫东, 博士, 主要研究方向为苗木水分生理生态。Tel: (010)84239798 E-mail: sxtzgwd2008@163.com***通信作者:** 杨艳芳, 博士, 副研究员, 主要研究方向为林木次生代谢与调控。Tel: (010)62889638 E-mail: echoyyf@caf.ac.cn

陈贵林, 博士, 教授, 主要研究方向为药用植物化学与资源利用。Tel: (0471)4992944 E-mail: guilinch@aliyun.com

important rate-limiting step in vitamin C biosynthesis in *Nitraria sibirica*. **Methods** Basing on the third-generation transcriptome data of *N. sibirica* in NCBI public database, *GPP* genes were identified and cloned. With the bioinformatics methods, the structure, chemical properties, conserved sequence and phylogenetic evolution were analyzed. Moreover, the expression pattern of *NsGPP* genes under salt stress was analyzed via RNA-seq data, and the expression levels in different tissues were analyzed by qRT-PCR. **Results** The results showed that two proteins encoded by the *NsGPP1/2* genes were similar to the other plant GPP proteins, which have typical conserved FBPase/IMPase/glpX-like (FIG) domain. Moreover, from an evolutionary perspective, *NsGPPs* may be more closely related to GPPs of shrubs or robust herbs. At the same time, with motif analysis, a few differences between *NsGPP* and GPP of herbaceous were found. In addition, the three-dimensional structures of *NsGPP1* and *NsGPP2* were also slightly different. With the transcriptome data, the expression level of *NsGPP1* was high and gradually increased under different concentrations of NaCl treatment, while the expression level of *NsGPP2* was low and gradually decreased. According to the experimental data of qRT-PCR, *NsGPP1/2* were specifically expressed in different tissues in *N. sibirica*. Moreover, the expression level of *NsGPP1* was significantly higher than that of *NsGPP2*. **Conclusion** In this study, two *NsGPP* genes were identified and cloned from *N. sibirica*, the structures of these two proteins were relatively conservative and closely related to GPP of herbaceous plants. Furthermore, they were expressed in fruits, stems and leaves, and had obvious tissue specificity. Among them, *NsGPP1* may be involved in vitamin C biosynthesis and salt tolerance in *N. sibirica*. This study is helpful to further study of the molecular mechanism of vitamin C biosynthesis and salt tolerance in *N. sibirica*.

Key words: *Nitraria sibirica* Pall.; *GPP* gene; gene identification; bioinformatics; gene expression

西伯利亚白刺 *Nitraria sibirica* Pall. 为蒺藜科白刺属落叶灌木，是我国分布最为广泛的白刺种，因其具有良好的耐盐碱、耐干旱等特性，多被种植在盐碱或生态脆弱地区^[1]。西伯利亚白刺除拥有极强的抗逆性外，在其叶、果、枝中还含有丰富的对人体有益的营养物质和微量元素等^[2-3]，是优质的医疗保健、食品工业原材料。但现阶段针对白刺的研究多集中在其抗性方面，而对其营养价值、活性成分的研究较少。

维生素 C 也称为抗坏血酸 (ascorbic acid, AsA)，是一类在生物体内合成的己糖内酯化合物，在植物体内与生长发育、次生代谢产物及激素的合成密切相关^[4]，如在植物细胞分裂、光合作用、乙烯和赤霉素的合成以及抗氧化和清除自由基等植物抗氧化胁迫中起着非常重要作用^[5-8]。研究表明，维生素 C 能够促进人体对锌和铁的吸收^[9]，还对心血管病、癌症、糖尿病、白内障及衰老等病症具有良好的预防功效^[10-11]。维生素 C 为人体所必需，却在人体内无法合成、储存，必须从饮食中获取。因此，阐明植物中维生素 C 的合成途径，提高植物体内维生素 C 含量，对于人体健康有着重要意义。

研究认为植物维生素 C 合成途径主要有 *L*-半乳糖途径、*D*-半乳糖醛酸途径、*L*-古洛糖途径以及葡萄糖醛酸途径^[12]。*L*-半乳糖途径被认为是高等植物中维生素 C 生物合成的主要途径^[13-14]。目前，植物中 *L*-半乳糖合成途径所涉及的反应酶和相应的基因已全部被鉴定出^[15-16]。其中，*L*-半乳糖-1-磷酸

酶 (*L*-galactose-1-phosphate phosphatase, GPP) 在该途径中起着重要作用，其能够催化 *L*-半乳糖-1-磷酸生成 *L*-半乳糖，即维生素 C 合成所必需的前体物质。Conklin 等^[17]研究发现拟南芥 *Arabidopsis thaliana* *GPP* 缺失突变体中维生素 C 的含量仅为野生型的 50%。还有研究证实在苹果 *Malus domestica* Borkh 叶片和番茄 *Solanum lycopersicon* L. 果实中，*GPP* 基因是维生素 C 合成积累的关键基因^[18-19]。因此，本研究以西伯利亚白刺为研究对象，基于其三世代测序数据鉴定出 *NsGPP* 基因，并根据序列信息设计特异引物，从西伯利亚白刺叶片中克隆得到 *NsGPP1* 和 *NsGPP2* 基因，利用生物信息学对其进行分析，并结合转录组数据分析其在盐胁迫中的表达模式，同时利用 qRT-PCR 分析其不同组织中的表达水平，为进一步揭示白刺中维生素 C 生物合成的分子机制提供参考。

1 材料与方法

1.1 西伯利亚白刺 RNA 的提取及 cDNA 的合成

西伯利亚白刺果实于 2022 年 8 月采摘自内蒙古自治区阿拉善盟吉兰泰镇，采摘后立即用液氮速冻，带回实验室后存放于 -80 °C 冰箱保存。西伯利亚白刺茎、叶组织于 2023 年 5 月取自中国林业科学研究院温室种植的组培苗，取样后，液氮速冻，-80 °C 冰箱保存。使用 RNA Easy Fast Plant Tissue Kit 试剂盒 (北京天根生化科技有限公司) 提取样品总 RNA，利用 FastKing RT Kit (With gDNase) 反转录试剂盒 (北京天根生化科技有限公司) 合成 cDNA，于 -20 °C 冰箱保存。

1.2 西伯利亚白刺 GPP 基因鉴定

通过本地化 Blast 程序,以拟南芥 VTC4 蛋白(登录号:NP_186936)作为 query,与西伯利亚白刺转录组数据(登录号:SRR18000662)进行比对,并利用 CD-HIT 软件默认参数进行去冗余,之后将去冗余得到的蛋白序列利用 CD-search (<https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>)进行结构域预测,筛选结构域完整的蛋白序列进行后续分析。

1.3 NsGPP 基因的克隆

本研究从西伯利亚白刺转录组数据库中共鉴定出两个 NsGPP 基因(SRR18000662.115405.1 和 SRR18000662.126374.1)。依据 2 个基因的序列信息,利用 Primer 5 软件进行引物设计(NsGPP1-F: 5'-TCATTGCTTTCCTTCAC-3'; NsGPP1-R: 5'-AACACCTGTCCTTTTCC-3'; NsGPP2-F: 5'-TACCATCAGGTTCCATT-3'; NsGPP2-R: 5'-CGATAACACCTGTCCTTTT-3')。

PCR 反应体系为: 2 μ L cDNA, 0.5 μ L 引物, 10 μ L 2 $\times$ M5 HiPer Plus Taq HiFi PCR mix (with blue dye) (北京聚合美生物科技有限公司), 7 μ L Nuclease-free ddH₂O。PCR 反应程序为: 95 $^{\circ}$ C 预变性 3 min; 94 $^{\circ}$ C 变性 30 s, 59 $^{\circ}$ C 退火 30 s, 72 $^{\circ}$ C 延伸 45 s, 共 35 个循环; 最后 72 $^{\circ}$ C 延伸 5 min。通过 1.2% 琼脂糖凝胶对 PCR 产物进行电泳分离,利用 M5 HiPure Gel Extraction Kit 胶回收试剂盒(北京聚合美生物科技有限公司)回收目的片段。将胶回收产物与 pTOPO-TA 载体连接后转化大肠杆菌菌株 DH5 α ,挑取单克隆进行菌液 PCR 鉴定,对条带正确的阳性克隆菌液进行测序(中美泰和生物技术(北京)有限公司)。

基于菌液测序的结果,使用 NCBI 网页工具 ORF Finder 寻找测序序列完整的开放阅读框(open reading frame, ORF),随后利用 DNAMAN 6.0 对目的基因序列与转录组数据对应序列进行比对。

1.4 NsGPP 蛋白理化性质分析

利用 ExPASy 在线网站(<https://www.expasy.org/>)中的 ProtParam tool (<https://web.expasy.org/protparam/>)对 NsGPP 蛋白的氨基酸数量、相对分子质量、等电点以及疏水性进行分析。

1.5 NsGPP 蛋白序列比对及系统进化发育分析

利用 DNAMAN 6.0 对 NsGPP 蛋白序列进行比对。从 NCBI 数据库下载拟南芥等植物 GPP 家族蛋白序列。利用 MEGA X 软件对 NsGPP 基因编码的

蛋白和拟南芥等植物 GPP 蛋白进行比对,并采用邻接法(neighbor-joining, NJ)法构建进化树,参数为默认值,重复 1 000 次。

1.6 NsGPP 蛋白保守基序分析

利用 MEME (<http://meme-suite.org/tools/meme>)在线工具对 NsGPP 蛋白以及拟南芥等植物 GPP 蛋白进行保守基序(Motif)分析^[20],最大 motif 检索数为 10,其余为默认参数。

1.7 西伯利亚白刺 NsGPP 蛋白结构分析

利用 phyre2 在线网站(<http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index>)对 NsGPP 蛋白三级结构进行建模。

1.8 NsGPP 基因盐胁迫下表达分析

从 NCBI 基因表达数据库中下载西伯利亚白刺盐胁迫的 GEO 数据(SRP127697),利用 Blast 比对寻找 NsGPP 同源序列,获得其 FPKM 数据,分析 2 个 NsGPP 基因的表达水平。

1.9 NsGPP 基因组织表达分析

为检测 NsGPP 基因在不同组织中的表达水平,以西伯利亚白刺果实、茎、叶的 cDNA 为模板,利用 SuperRe-al PreMix Plus (SYBR Green)试剂盒(北京天根生化科技有限公司)在 Roche LightCycler[®] 480II 荧光定量 PCR 仪上进行 qRT-PCR 反应,每个样品设置 3 个生物学重复以及 3 次技术重复。反应程序: 95 $^{\circ}$ C 酶激活 15 min, 95 $^{\circ}$ C 变性 10 s, 55 $^{\circ}$ C 退火 30 s, 72 $^{\circ}$ C 延伸 32 s, 循环次数 40,以 Actin^[21]为内参基因,利用 2^{- $\Delta\Delta C_t$} 法^[22]计算基因的相对表达量。相关引物见表 1,利用 SPSS 软件进行显著性分析。

2 结果与分析

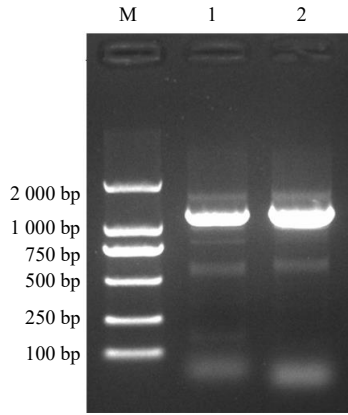
2.1 西伯利亚白刺 NsGPP 基因克隆及其编码蛋白理化性质分析

利用生物信息学的方法,从西伯利亚白刺转录组数据中鉴定出 2 条 GPP 基因序列,依据这 2 条序列信息,设计特异引物进行 RT-PCR 扩增(图 1),

表 1 qRT-PCR 引物

Table 1 Primers for qRT-PCR

引物名称	引物序列(5'-3')
Actin-F	GGAATCCACGAGACCACCTACA
Actin-R	GATTGATCCTCCGATCCAGACA
NsGPP1-F-E	TTTTGCGTGTTTGATAG
NsGPP1-R-E	TCTTTAGCTTCTCCGTC
NsGPP2-F-E	GCACCATCGGATCTT
NsGPP2-R-E	TTCTTTGTTTGCACTC



M-Marke, 1-*NsGPP1* PCR 产物, 2-*NsGPP2* PCR 产物。
M-Marker, 1-product of *NsGPP1* PCR, 2-product of *NsGPP2* PCR.

图 1 *NsGPP1/2* 基因 PCR 产物琼脂糖凝胶电泳图

Fig. 1 PCR amplification of *NsGPP1/2*

A

<i>NsGPP1</i>	ATGGGCAGATCATTGCTTCCCTTACAGACATTCTCTCAGATTTCTCAGATACCCAGATCACTTGCACCATCGGATCTCCACCGATGTCAAATCTGAGCCACC	0
<i>NsGPP2</i>	ATGGGCAGATCATTGCTTCCCTTACAGACATTCTCTCAGATTTCTCAGATACCCAGATCACTTGCACCATCGGATCTCCACCGATGTCAAATCTGAGCCACC	110
consensus	ATGGGCAGATCATTGCTTCCCTTACAGACATTCTCTCAGATTTCTCAGATACCCAGATCACTTGCACCATCGGATCTCCACCGATGTCAAATCTGAGCCACC	
<i>NsGPP1</i>	TGCAAAATCTCTATCTGGTATCGAAAGTTTCGATTTCGATGCGCAAAACAAAGAAATATCTACAAAGCGGTTCGTCTGAGAATAAGCAGTACCCGAAAT	4
<i>NsGPP2</i>	TGCAAAATCTCTATCTGGTATCGAAAGTTTCGATTTCGATGCGCAAAACAAAGAAATATCTACAAAGCGGTTCGTCTGAGAATAAGCAGTACCCGAAAT	220
consensus	TGCAAAATCTCTATCTGGTATCGAAAGTTTCGATTTCGATGCGCAAAACAAAGAAATATCTACAAAGCGGTTCGTCTGAGAATAAGCAGTACCCGAAAT	
<i>NsGPP1</i>	GGCCAAAACGACTGCACAAATCCGCCAATCAGCTCATCAACTTCCTGAGTCTCCCTTAAACTGCCCTGAGCTTCTGATGACCCCTGTTAATAAGCTCTGTGAT	114
<i>NsGPP2</i>	GGCCAAAACGACTGCACAAATCCGCCAATCAGCTCATCAACTTCCTGAGTCTCCCTTAAACTGCCCTGAGCTTCTGATGACCCCTGTTAATAAGCTCTGTGAT	330
consensus	GGCCAAAACGACTGCACAAATCCGCCAATCAGCTCATCAACTTCCTGAGTCTCCCTTAAACTGCCCTGAGCTTCTGATGACCCCTGTTAATAAGCTCTGTGAT	
<i>NsGPP1</i>	ATACATATAGCGGCTTACCGACTTGTCTACAGACACAGACAAAAATAGCAGCGTCTATATTACAAGTTGTCAAAAGAAATTTGGAGATCACCTTATCTTGAGA	224
<i>NsGPP2</i>	ATACATATAGCGGCTTACCGACTTGTCTACAGACACAGACAAAAATAGCAGCGTCTATATTACAAGTTGTCAAAAGAAATTTGGAGATCACCTTATCTTGAGA	440
consensus	ATACATATAGCGGCTTACCGACTTGTCTACAGACACAGACAAAAATAGCAGCGTCTATATTACAAGTTGTCAAAAGAAATTTGGAGATCACCTTATCTTGAGA	
<i>NsGPP1</i>	AGACGGTGGAAATATAGGAGACTCTCATCTGATATCTCTGTCGATGACCTTTAGATGCAACAACCAATTTTGACATGCTTACCAAGCTTTGACAGTTTCCCTGG	334
<i>NsGPP2</i>	AGACGGTGGAAATATAGGAGACTCTCATCTGATATCTCTGTCGATGACCTTTAGATGCAACAACCAATTTTGACATGCTTACCAAGCTTTGACAGTTTCCCTGG	550
consensus	AGACGGTGGAAATATAGGAGACTCTCATCTGATATCTCTGTCGATGACCTTTAGATGCAACAACCAATTTTGACATGCTTACCAAGCTTTGACAGTTTCCCTGG	
<i>NsGPP1</i>	GAGTTCCTTATCAGCAAAACCCAGCTGCTGCTGACAGTGTACAGTTCCTTGGTCTCCATGTTTGGAAATACAGTATATTTTCTGCAACCCAGCTTGTGAGCATTC	444
<i>NsGPP2</i>	GAGTTCCTTATCAGCAAAACCCAGCTGCTGCTGACAGTGTACAGTTCCTTGGTCTCCATGTTTGGAAATACAGTATATTTTCTGCAACCCAGCTTGTGAGCATTC	660
consensus	GAGTTCCTTATCAGCAAAACCCAGCTGCTGCTGACAGTGTACAGTTCCTTGGTCTCCATGTTTGGAAATACAGTATATTTTCTGCAACCCAGCTTGTGAGCATTC	
<i>NsGPP1</i>	TGTAATCGCCAGAAAGATTATCTGAGTCAAACTGATAAGCTGACCGGATCACTTCTGTAACCGGCTTTGGATATGAACATGATGACGATGGCCACTAACATAGAATT	554
<i>NsGPP2</i>	TGTAATCGCCAGAAAGATTATCTGAGTCAAACTGATAAGCTGACCGGATCACTTCTGTAACCGGCTTTGGATATGAACATGATGACGATGGCCACTAACATAGAATT	770
consensus	TGTAATCGCCAGAAAGATTATCTGAGTCAAACTGATAAGCTGACCGGATCACTTCTGTAACCGGCTTTGGATATGAACATGATGACGATGGCCACTAACATAGAATT	
<i>NsGPP1</i>	ATTTAAACCAATTTACTGACGTGACCGGGGTACCAAGACTGGCTGCTGACGGTGGATATGTCCTATGGCCCTAGGGATCTGAGAAGCTTACTCGGAATATCGTC	664
<i>NsGPP2</i>	ATTTAAACCAATTTACTGACGTGACCGGGGTACCAAGACTGGCTGCTGACGGTGGATATGTCCTATGGCCCTAGGGATCTGAGAAGCTTACTCGGAATATCGTC	880
consensus	ATTTAAACCAATTTACTGACGTGACCGGGGTACCAAGACTGGCTGCTGACGGTGGATATGTCCTATGGCCCTAGGGATCTGAGAAGCTTACTCGGAATATCGTC	
<i>NsGPP1</i>	TAAACCGTGGGATATGGCTGCTGCTGTTCTGTAAGCAGCAGCTGGAATCCTTAGTCGATGATGCGCGGAAATTTGCTGTTTATAGATCTTTTCTGTA	774
<i>NsGPP2</i>	TAAACCGTGGGATATGGCTGCTGCTGTTCTGTAAGCAGCAGCTGGAATCCTTAGTCGATGATGCGCGGAAATTTGCTGTTTATAGATCTTTTCTGTA	990
consensus	TAAACCGTGGGATATGGCTGCTGCTGTTCTGTAAGCAGCAGCTGGAATCCTTAGTCGATGATGCGCGGAAATTTGCTGTTTATAGATCTTTTCTGTA	
<i>NsGPP1</i>	TCTAACCGCATCTGCAACAGCTTTTGAGAGGATAGGCCAGGACGCGCAAGCTAAAGACCAAGGAATCAATTTCTCTGTTATGCTTTAAGCCAGAGATTACCA	884
<i>NsGPP2</i>	TCTAACCGCATCTGCAACAGCTTTTGAGAGGATAGGCCAGGACGCGCAAGCTAAAGACCAAGGAATCAATTTCTCTGTTATGCTTTAAGCCAGAGATTACCA	1100
consensus	TCTAACCGCATCTGCAACAGCTTTTGAGAGGATAGGCCAGGACGCGCAAGCTAAAGACCAAGGAATCAATTTCTCTGTTATGCTTTAAGCCAGAGATTACCA	
<i>NsGPP1</i>	TAAGTACATTG	896
<i>NsGPP2</i>	TAAGTACATTG	1112
consensus	TAAGTACATTG	

B

<i>NsGPP1</i>	NRGSLSFTEIPLRFSHIRSLAPSLPHRCQLSPANSLSGYRKVSLSAKRRRLCTKAVLSENKQYPR	38
<i>NsGPP2</i>	NRGSLSFTEIPLRFSHIRSLAPSLPHRCQLSPANSLSGYRKVSLSAKRRRLCTKAVLSENKQYPR	110
consensus	NRGSLSFTEIPLRFSHIRSLAPSLPHRCQLSPANSLSGYRKVSLSAKRRRLCTKAVLSENKQYPR	
<i>NsGPP1</i>	ITYKGLTELVTLDKNSAAIICVVKKNFGHLILGEECGIIGDSSSDYLVCIDPLDGTTFNAHGYPFAVSVGLYQGNPAAAVVEFVGGPMCVNTRI	148
<i>NsGPP2</i>	ITYKGLTELVTLDKNSAAIICVVKKNFGHLILGEECGIIGDSSSDYLVCIDPLDGTTFNAHGYPFAVSVGLYQGNPAAAVVEFVGGPMCVNTRI	220
consensus	ITYKGLTELVTLDKNSAAIICVVKKNFGHLILGEECGIIGDSSSDYLVCIDPLDGTTFNAHGYPFAVSVGLYQGNPAAAVVEFVGGPMCVNTRI	
<i>NsGPP1</i>	CNGCKIHYVSTKVERSLVITGFGYEHDDAWATNIELFKCFIDVSRGVRRLGAAAVDCHVALGIVEGYVEYRLKPVEMAAAGLVVVEEAGGIVSRDGGKFCVFRSVLV	258
<i>NsGPP2</i>	CNGCKIHYVSTKVERSLVITGFGYEHDDAWATNIELFKCFIDVSRGVRRLGAAAVDCHVALGIVEGYVEYRLKPVEMAAAGLVVVEEAGGIVSRDGGKFCVFRSVLV	330
consensus	CNGCKIHYVSTKVERSLVITGFGYEHDDAWATNIELFKCFIDVSRGVRRLGAAAVDCHVALGIVEGYVEYRLKPVEMAAAGLVVVEEAGGIVSRDGGKFCVFRSVLV	
<i>NsGPP1</i>	SNGLIHKLRLERIGPATKLTKEGIFSLWFKPEDYHTD	297
<i>NsGPP2</i>	SNGLIHKLRLERIGPATKLTKEGIFSLWFKPEDYHTD	369
consensus	SNGLIHKLRLERIGPATKLTKEGIFSLWFKPEDYHTD	

图中黑色代表相同氨基酸。

Black indicates identical amino acid.

图 2 西伯利亚白刺 GPP 基因 ORF 核苷酸序列 (A) 和蛋白序列比对 (B)

Fig. 2 Nucleotide sequences alignment (A) and protein sequences alignment (B) of *NsGPPs*

经胶回收, 克隆以及测序, 得到 2 个基因完整的 ORF 序列。2 个基因 ORF 全长分别为 897 bp 和 1113 bp, 与转录组预测序列长度一致且相似性达到 99%, 进而分别将其命名为 *NsGPP1* 和 *NsGPP2*。*NsGPP1* 和 *NsGPP2* 编码的蛋白分别由 298 和 370 个氨基酸残基组成, 其相对分子质量为 32 220 和 40 451, 等电点分别为 5.39 和 7.10, GRAVY 值为-0.015 与-0.118, 表明 2 个 *NsGPP* 蛋白均为亲水蛋白。

2.2 序列比对分析

将 2 个 *NsGPP* 基因的 ORF 核苷酸序列进行比对, 发现 *NsGPP1* 和 *NsGPP2* 基因同源性较高, 两者除了在 5'端存在明显差异 (*NsGPP2* 比 *NsGPP1* 基因长 216 bp) 之外, 在后半部分同源性较高的序列中, 还存在 6 个差异碱基 (图 2-A)。蛋白序列比对结果如图 2-B 所示, 2 个 *NsGPP* 蛋白序列的相

似性为 80%，NsGPP2 蛋白 N 端明显长于 NsGPP1 蛋白；2 个蛋白存在 2 个氨基酸差异，NsGPP1 蛋白序列的第 145 位与第 276 位氨基酸分别为甘氨酸与丙氨酸，而 NsGPP2 蛋白序列中与其对应位置上的氨基酸则分别为天冬氨酸与谷氨酸。利用拟南芥、

油菜 *Brassica napus* L. 以及苹果、枣树 *Zizyphus jujuba* (L.) Lam. 等不同物种的 GPP 蛋白序列（下载自 Genbank 数据库）与 2 个 NsGPP 蛋白序列进行多序列比对。由图 3 可以看出，NsGPP 与其他植物 GPP 蛋白具有较高的同源性，且均含有完整的



图中红色下划线代表 FIG 保守结构域；黑色代表相同氨基酸；红色和蓝色代表同源的氨基酸残基；图中所示各物种蛋白质的登录号如下：AtGPP1 (NM_111155_拟南芥)，AtGPP2 (NM_001125086_拟南芥)，AtGPP3 (NM_001035549_拟南芥)，BrGPP (JN157615_小白菜)，MoGPP (AB924376_辣木)，CsGPP (KC505202_茶树)，AdGPP (AY787585_美味猕猴桃)，MdGPP (AY787586_苹果)，CrGPP (XM_001697943_莱茵衣藻)，SmGPP (KU356959_茄)，CtGPP (AB981061_柑桔)，ZjGPP (KJ739593_枣)，CuGPP (HQ224949_柑橘)，PpGPP (AB457584_桃)。The red underline represent conservative subdomains; Black indicates identical amino acid; Red and blue indicate homologue amino acid residues; The protein accession numbers of each species in the figure are as follows: AtGPP1 (NM_111155_ *Arabidopsis thaliana*), AtGPP2 (NM_001125086_ *Arabidopsis thaliana*), AtGPP3 (NM_001035549_ *Arabidopsis thaliana*), BrGPP (JN157615_ *Brassica rapa* subsp. *Chinensis*), MoGPP (AB924376_ *Moringa oleifera* (horseradish tree)), CsGPP (KC505202_ *Camellia sinensis*), AdGPP (AY787585_ *Actinidia deliciosa*), MdGPP (AY787586_ *Malus × domestica*), CrGPP (XM_001697943_ *Chlamydomonas reinhardtii*), SmGPP (KU356959_ *Solanum melongena*), CtGPP (AB981061_ *Citrus tamurana*), ZjGPP (KJ739593_ *Zizyphus jujuba*), CuGPP (HQ224949_ *Citrus unshiu*), PpGPP (AB457584_ *Prunus persica*).

图 3 西伯利亚白刺与其他植物 GPP 蛋白多序列比对

Fig. 3 Sequence alignment of GPP proteins of *N. sibirica* and other plants

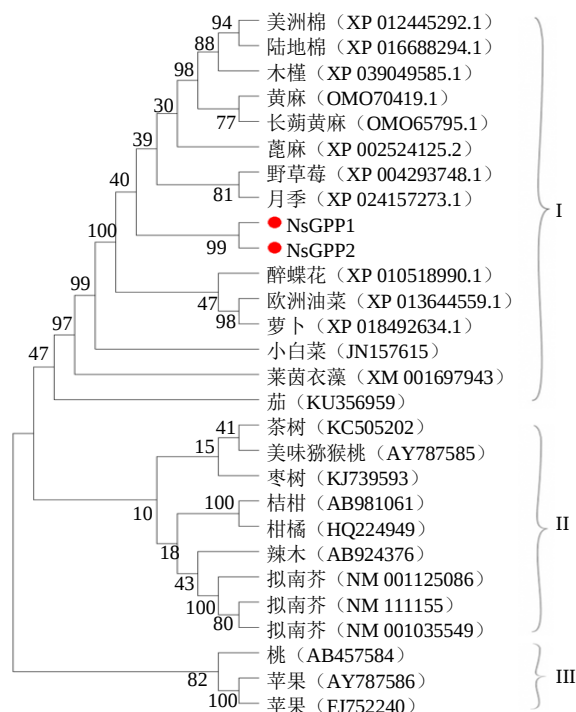
FBPase/IMPase/glpX-like(FIG)保守结构域, 可见其属于 FIG 家族。

2.3 系统进化发育分析

利用 NsGPP 蛋白序列和其他植物的 GPP 蛋白序列, 构建系统进化树。结果如图 4 所示, NsGPP 蛋白与陆地棉 *Gossypium hirsutum* L.、长蒴黄麻 *Corchorus olitorius* L.、月季 *Rosa chinensis* Jacq.和茄 *Solanum melongena* L.、萝卜 *Raphanus sativus* L.等灌木或半木质化草本植物的 GPP 蛋白聚在一起, 而与主要由木本植物枣 *Ziziphus jujuba* (L.) Lam.、柑橘 *Citrus unshiu* Marc.、辣木 *Moringa oleifera* Lam.组成的 II 亚家族和由苹果 *Malus domestica* (Suckow) Borkh.、桃 *Prunus persica* (L.) Batsch 组成的 III 亚家族关系较远。因此, 从进化上来讲, 西伯利亚白刺可能与灌木或半木质化草本植物的亲缘关系更近。

2.4 NsGPP 蛋白保守基序分析

利用 MEME 在线网站对 2 个 NsGPP 蛋白以及拟南芥、油菜和辣木的 GPP 蛋白进行 motif 分析(图 5), 结果发现 NsGPP1 和 NsGPP2 与其他植物的 GPP 蛋白较为保守, 均含有 motif 1、2、5。同时, NsGPP 与其他植物的 GPP 蛋白还存在着差异, 如只有 NsGPP1 和 NsGPP2 中含有 motif8, 且两者均不含有 motif 6 和 9; motif 6 存在于拟南芥和辣木的 GPP 蛋白中, motif9 存在于拟南芥、油菜和辣木 GPP 蛋



红色字体为西伯利亚白刺 GPP 蛋白; 黑色字体为其他植物 GPP 蛋白; 由登录号和拉丁名组成。

Red font is GPP protein of *N. sibirica*; Black font is the GPP protein of other plants, which is composed of accessing number and Latin name.

图 4 西伯利亚白刺与其他植物 GPP 蛋白系统进化分析
Fig. 4 Phylogenetic analysis of GPP proteins of *N. sibirica* and other plants

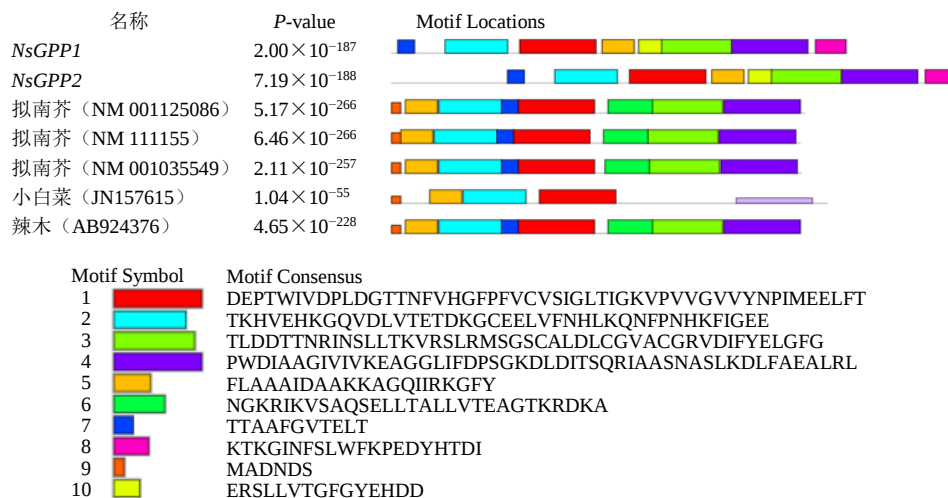


图 5 白刺及其他植物 GPP 蛋白 motif 分析

Fig. 5 Motif analysis of GPP proteins of *N. sibirica* and other plants

白中, 而油菜中的 GPP 蛋白缺少 motif 3、4、6、7 和 8。由此可见, GPP 蛋白在不同物种存在较高的保守性, 但也存在基序数目以及类型上的差异。

2.5 NsGPP 蛋白三级结构分析

利用 phyre2 在线软件 (<http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index>)对 NsGPP 基因编码的蛋白进行三级结构特征分析。从图 6 中可

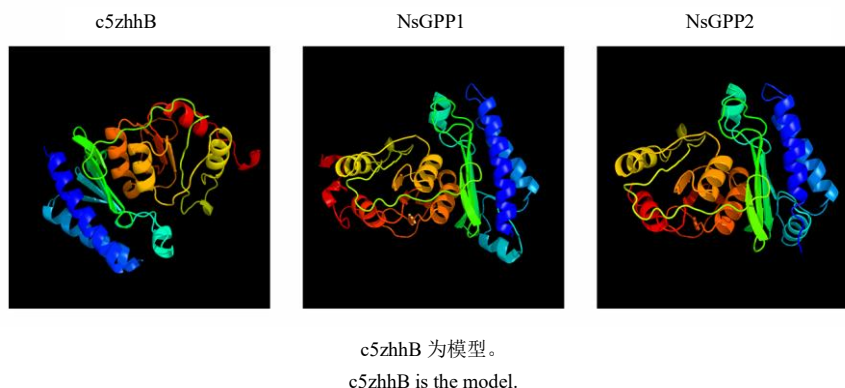


图 6 NsGPP 蛋白的三级结构分析

Fig. 6 3D structures of NsGPP proteins

以看出, NsGPP 蛋白与模型结构非常相似, 但也存在微小差异, 如 NsGPP2 中蓝色螺旋长度明显低于 NsGPP1 和模型蛋白, NsGPP1 蛋白比 NsGPP2 多了 2 个红色 β -折叠。

2.6 NsGPP 基因的表达模式分析

通过转录组数据, 对 2 个 *NsGPP* 基因在不同浓度 NaCl 处理下的西伯利亚白刺叶片中的表达水平进行分析。结果显示, *NsGPP1* 在 0 mmol/L (对照)、100 mmol/L 和 400 mmol/L 盐处理下的表达量均明显高于 *NsGPP2*; 而 *NsGPP2* 在对照以及盐处理下的表达值均显著低于 *NsGPP1*, 且 *NsGPP2* 在盐处理下的表达水平无显著差异 (图 7)。由此可见, *NsGPP1* 能够响应盐胁迫, 盐胁迫可能促进叶片中维生素 C 的合成。

2.7 NsGPP 基因组织表达分析

为进一步分析 *NsGPP* 基因在白刺不同组织中的表达水平, 本研究利用 qRT-PCR 分析了 *NsGPP1* 和 *NsGPP2* 在西伯利亚白刺果实、茎与叶中的表达

情况。结果如图 8 显示, *NsGPP1* 在果实、茎和叶中的表达量均明显高于 *NsGPP2* 的表达量 ($P < 0.05$), 且其在茎和叶中的表达量无显著差异; 而 *NsGPP2* 在叶片中的表达量显著高于其在茎中的表达量。此外, 在叶中, *NsGPP1* 表达量接近于 *NsGPP2* 的 2 倍, 这也与盐处理转录组数据中未经过盐处理的叶片中 2 个基因表达的结果基本一致, 验证了转录组数据结果的准确性。

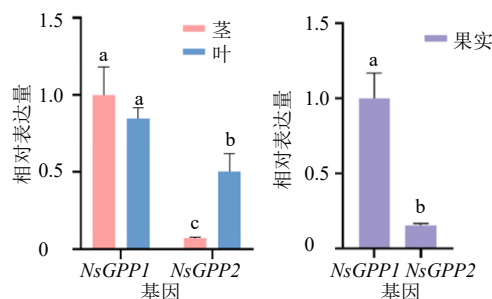
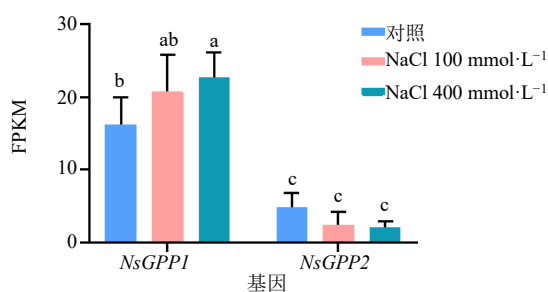


图 8 NsGPP 在不同组织中的相对表达水平

Fig. 8 Relative gene expression of NsGPP in different tissues



不同字母之间表示处理间存在显著性差异 ($*P < 0.05$), 下图同。
Different letter indicates significant between different treatments ($*P < 0.05$).

图 7 不同浓度 NaCl 处理下 *NsGPP1/2* 的表达量Fig. 7 Expression levels of *NsGPP 1/2* under different concentrations of NaCl treatment

3 讨论

L-半乳糖途径是植物合成维生素 C 的主要途径, 而 GPP 蛋白是维生素 C 合成中的关键限速酶^[4, 12], 目前已在多种植物中克隆得到 *GPP* 基因^[23-24]。本研究从西伯利亚白刺中克隆出 2 个 *GPP* 基因, 其 cDNA 序列均含有完整的 ORF, 且与其他植物的同源基因的编码序列长度相近, 所编码的 GPP 蛋白与其他植物的 GPP 蛋白序列具有较高的相似性, 表明所获得的基因为 *GPP* 基因。

GPP 蛋白具有 FIG 保守结构域, 本研究中的蛋白多序列比对结果显示, 2 个 NsGPP 蛋白均具有完整的 FIG 结构域 (图 3)。Motif 分析结果显示,

NsGPP1 与 NsGPP2 蛋白具有一致的 motif 组成(图 5)。此外发现, NsGPP 蛋白不仅与拟南芥等其他植物存在相似性, 同时也存在差异之处(图 5)。蛋白三维结构分析表明, NsGPP2 比 NsGPP1 的蓝色螺旋短小并且缺少 β -折叠(图 6)。蛋白结构与其功能密切相关, 由此可以推测, 本研究所获得的 NsGPP 基因所编码蛋白很有可能与其他植物的 GPP 蛋白具有类似功能, 为白刺维生素 C 合成中的重要一环, 但由于 NsGPP 蛋白与其他植物 GPP 蛋白存在结构上的显著差异, 乃至 NsGPP1 与 NsGPP2 蛋白之间的差异, 可能导致它们在功能上也会存在差异。在转录水平, NsGPP1 与 NsGPP2 在不同组织中均有表达, 其中 NsGPP1 在茎中高表达, NsGPP2 在叶片中表达量最高, 且在 3 种不同组织中, NsGPP1 的表达量均显著高于 NsGPP2 的表达量(图 8), 由此可见, 白刺 NsGPP1 与 NsGPP2 基因可能具有不同的功能。而为确定 NsGPP1 和 NsGPP2 功能上两者之间存在哪些差异, 还需要进一步研究来揭示。

进化分析结果显示, 西伯利亚白刺 GPP 蛋白与多数草本植物 GPP 蛋白进化关系较近, 与苹果等木本植物进化较远(图 4)。分析原因, 可能是由于西伯利亚白刺自身为小灌木, 身形匍散, 与半木质化草本植物更为接近。此外, 由于 NsGPP 与月季等 GPP 更为相近, 可以通过月季等植物 GPP 蛋白功能进一步了解其功能。

适度盐胁迫下提高了猕猴桃中的维生素 C 含量^[25], 而维生素 C 可以清除植物体内活性氧^[26], 增强植物的抗盐性^[25]。众所周知, 白刺具有很强的耐盐性, 本研究也发现 NsGPP1 在盐胁迫(100 mmol/L 或 400 mmol/L NaCl)下表达量均高于对照, 这与上述前人研究结果较为一致, 表明其可能在西伯利亚白刺抵抗盐胁迫过程中起到重要作用。NsGPP2 在盐胁迫后表达量无显著变化, 且其在正常生长条件下表达量也显著低于 NsGPP1, 因此推测其可能不响应盐胁迫, 且在维生素 C 合成过程不如 NsGPP1 作用重要。前人研究发现, 猕猴桃 AceGPP 基因表达并未随着 NaCl 浓度增加而改变, 反而是 AceGMP 基因表达量显著提高^[25]。也有人发现, 在水稻中转化拟南芥维生素 C 合成基因, 结果发现 GGP、GME 或 GDH 对水稻叶片维生素 C 积累贡献最大, 且转化 GGP 基因的转基因水稻幼苗在盐胁迫下维生素 C 含量最高, 抗盐能力最高^[27]。

同时还有研究发现 GPP 在番茄果实成熟过程中维持高水平表达, 但只是维持了组成型的高表达, 并不响应乙烯、高温等非生物胁迫^[12]。由此可见, 物种在漫长进化中衍生出了较大差异, 而遗传背景迥异的物种在维生素 C 合成以及抗盐机制上也会存在显著差异。

本研究从西伯利亚白刺中鉴定出 2 个 NsGPP 基因, 发现 NsGPP 基因编码蛋白均具有保守的 FIG 结构域, 且与其他植物 GPP 蛋白具有较高的相似性, 推测 NsGPP 可能与其他植物的 GPP 蛋白具有相似的功能。此外, NsGPP1 与 NsGPP2 蛋白在三维结构上也存在差异, 暗示两者可能在功能上存在差异。盐胁迫与组织表达分析进一步证实, NsGPP1 基因不仅在西伯利亚白刺果实、茎和叶中维生素 C 生物合成过程中起到重要作用, 并且也极有可能在西伯利亚白刺耐盐胁迫过程中起到积极作用。本研究为进一步探索白刺维生素 C 生物合成以及耐盐分子机制提供理论基础。

利益冲突 所有作者均声明不存在利益冲突

参考文献

- [1] 李焕勇. 西伯利亚白刺对盐胁迫的分子与生理代谢响应机制研究 [D]. 北京: 中国林业科学研究院, 2018.
- [2] 李永丽, 王海美, 波拉提·马卡比力, 等. 西伯利亚白刺叶化学成分及其抑制 α -葡萄糖苷酶活性的研究 [J]. 中南药学, 2021, 19(5): 827-830.
- [3] Abuduwaili A, Mutailifu P, Nuerxiti R, et al. Structure and biological activity of polysaccharides from *Nitraria sibirica* pall fruit [J]. *Food Biosci*, 2021, 40: 100903.
- [4] 刘永立, 胡海涛, 兰大伟. 维生素 C 的生物合成及其基因调控研究进展 [J]. 果树学报, 2006, 23(3): 431-436.
- [5] Davey M W, Gilot C, Persiau G, et al. Ascorbate biosynthesis in *Arabidopsis* cell suspension culture [J]. *Plant Physiol*, 1999, 121(2): 535-543.
- [6] Ivanov B N. Role of ascorbic acid in photosynthesis [J]. *Biochemistry*, 2014, 79(3): 282-289.
- [7] Smirnov N. Ascorbate biosynthesis and function in photoprotection [J]. *Philos Trans R Soc Lond B Biol Sci*, 2000, 355(1402): 1455-1464.
- [8] Arrigoni O, De Tullio M C. Ascorbic acid: Much more than just an antioxidant [J]. *Biochim Biophys Acta*, 2002, 1569(1/2/3): 1-9.
- [9] Frossard E, Bucher M, Mchler F, et al. Potential for increasing the content and bioavailability of Fe, Zn and Ca in plants for human nutrition [J]. *J Sci Food Agric*, 2000, 80(7): 861-879.

- [10] Davey M, Montagu M, Inzé D, *et al.* Plant L-ascorbic acid: Chemistry, function, metabolism, bioavailability and effects of processing [J]. *J Sci Food Agric*, 2000, 80: 825-860.
- [11] 张鹏, 张瑞妮, 李飞, 等. 碳酸氢钠联合维生素 C 对早期糖尿病肾病患者 PCI 造影剂肾病的预防作用 [J]. 心血管康复医学杂志, 2019, 28(3): 333-336.
- [12] 安华明, 刘卫, 杨曼, 等. 高等植物维生素 C 合成途径主要参与基因研究进展 [J]. 山地农业生物学报, 2015, 34(6): 9-14.
- [13] Wheeler G L, Jones M A, Smirnoff N. The biosynthetic pathway of vitamin C in higher plants [J]. *Nature*, 1998, 393(6683): 365-369.
- [14] 王维鹏. 拟南芥 *GDPMPase* 和 *GALPase* 基因转化生菜研究 [D]. 上海: 上海交通大学, 2011.
- [15] Lu M, An H M, Li L L. Genome survey sequencing for the characterization of the genetic background of *Rosa roxburghii* tratt and leaf ascorbate metabolism genes [J]. *PLoS One*, 2016, 11(2): e0147530.
- [16] 罗丽华. 刺梨 GGP 基因在拟南芥中的表达及其效应 [D]. 贵阳: 贵州大学, 2018.
- [17] Conklin P L, Saracco S A, Norris S R, *et al.* Identification of ascorbic acid-deficient *Arabidopsis thaliana* mutants [J]. *Genetics*, 2000, 154(2): 847-856.
- [18] Li M J, Ma F W, Guo C M, *et al.* Ascorbic acid formation and profiling of genes expressed in its synthesis and recycling in apple leaves of different ages [J]. *Plant Physiol Biochem*, 2010, 48(4): 216-224.
- [19] Ioannidi E, Kalamaki M S, Engineer C, *et al.* Expression profiling of ascorbic acid-related genes during tomato fruit development and ripening and in response to stress conditions [J]. *J Exp Bot*, 2009, 60(2): 663-678.
- [20] Bailey T L, Boden M, Buske F A, *et al.* MEME SUITE: Tools for motif discovery and searching [J]. *Nucleic Acids Res*, 2009, 37: W202-W208.
- [21] Livak K J, Schmittgen T D. Analysis of relative gene expression data using real-time quantitative PCR and the 2^{-ΔΔC_T} method [J]. *Methods*, 2001, 25(4): 402-408.
- [22] 燕青, 王爽, 张洋, 等. 西伯利亚白刺 *NsCBL2* 基因克隆及表达分析 [J]. 安徽农业大学学报, 2022, 49(4): 540-546.
- [23] 陈莹莹, 赵智慧, 卜娇迪, 等. 枣 L-半乳糖-1-磷酸磷酸酶 cDNA 克隆及生物信息学分析 [J]. 河北农业大学学报, 2014, 37(6): 33-37.
- [24] 郭春苗, 李明军, 马锋旺, 等. 苹果 L-半乳糖-1-磷酸磷酸酶 cDNA 克隆及其表达 [J]. 中国农业科学, 2011, 44(4): 762-770.
- [25] 汪文杰, 刘晓莹, 成畅, 等. 盐和干旱胁迫及光质对猕猴桃叶片维生素 C 含量与合成基因表达的影响 [J]. 果树学报, 2022, 39(2): 203-210.
- [26] Padh H. Cellular functions of ascorbic acid [J]. *Biochim Biol Cell*, 1990, 68(10): 1166-1173.
- [27] 刘如如. 转基因水稻中过量表达维生素 C 代谢相关基因的研究 [D]. 扬州: 扬州大学, 2013.

[责任编辑 时圣明]