

DOI: 10.3724/SP.J.1006.2021.04217

脯氨酸羟化酶 GhP4H2 在棉花纤维发育中的功能研究

高璐^{1,2} 许文亮^{1,*}

¹ 华中师范大学生命科学学院 / 遗传调控与整合生物学湖北省重点实验室, 湖北武汉 430079; ² 山西农业大学小麦研究所, 山西临汾 041000

摘要: 阿拉伯半乳糖蛋白(arabinogalactan proteins, AGP)在棉花纤维发育过程中发挥重要作用。AGP 由富含羟脯氨酸的主链蛋白和大量 II 型阿拉伯半乳糖(arabinogalactan, AG)侧链组成, 其合成过程要经历 2 次翻译后修饰, 首先是氨基酸主链上的脯氨酸被脯氨酸羟化酶(prolyl-4-hydroxylases, P4H)羟化, 随后在糖基转移酶(glycosyltransferases, GT)催化作用下将阿拉伯半乳糖或寡糖添加到羟脯氨酸残基上。我们前期利用 P4Hs 的抑制剂处理棉花离体胚发现, 棉纤维伸长受到抑制, 暗示 P4H 参与棉纤维生长发育过程。为深入研究 P4H 在棉纤维发育中的功能, 我们从棉花中分离鉴定了 1 个在棉纤维发育过程中高量表达的脯氨酸羟化酶基因 *GhP4H2*。本研究分别构建了 *GhP4H2* 过表达和 RNA interference (RNAi)载体, 通过农杆菌介导法转化棉花, 获得转基因植株, 发现过表达转基因棉花株系 T₁~T₃ 代成熟棉纤维变短, AGP 含量增加, AG 多糖侧链也发生变化。对过表达转基因植株和野生型植株棉纤维的转录组分析表明, *GhP4H2* 正调控包括 AGP 在内的细胞壁糖蛋白基因的表达。基于以上研究, 我们推测 *GhP4H2* 可能主要通过影响 AGP 糖链组分调控棉纤维生长发育。

关键词: 棉纤维; 阿拉伯半乳糖蛋白; 脯氨酸羟化酶

GhP4H2 encoding a prolyl-4-hydroxylase is involved in regulating cotton fiber development

GAO Lu^{1,2} and XU Wen-Liang^{1,*}

¹ Hubei Key Laboratory of Genetic Regulation and Integrative Biology / School of Life Sciences, Central China Normal University, Wuhan 430079, Hubei, China; ² Wheat Research Institute, Shanxi Agricultural University, Linfen 041000, Shanxi, China

Abstract: Arabinogalactan proteins (AGPs) perform crucial roles during cotton fiber development, and they are composed of a hydroxyproline (Hyp)-rich core protein and large type-II arabinogalactan (AG) moieties. AGPs are highly glycosylated proteins which involve two most important post-translational processes. First, the proline residues were hydroxylated by prolyl-4-hydroxylases (P4Hs), then the Hyps were substituted with large arabinogalactan polysaccharides or small arabinogalactan oligosaccharides by glycosyltransferases (GTs). Our previous work had shown that fiber elongation was repressed when *in vitro*-cultured cotton ovules were treated with inhibitors of P4Hs, suggesting that P4Hs were involved in cotton fiber development. In this study, *GhP4H2* was detected to be relatively highly expressed at fiber elongation stage. Overexpression and RNAi-silencing vectors of *GhP4H2* were constructed and transformed into cotton. Phenotypic analysis showed that *GhP4H2* overexpressing fibers were significantly shorter compared with the wild type from T₁ generation to T₃ generation. In addition, AGP content had increased in *GhP4H2* overexpression fibers. Immuno dot-blot analysis also showed that carbohydrate moieties of AGP had changed. Moreover, RNA-seq revealed that expression levels of genes encoding hydroxyproline (Hyp)-rich glycoproteins including AGPs were enhanced. Overall, these results indicated that *GhP4H2* may regulate fiber development primarily through affecting AGPs composition.

Keywords: cotton fiber; arabinogalactan proteins (AGPs); prolyl-4-hydroxylase (P4Hs)

本研究由国家自然科学基金项目(31970516, 31671735)资助。

This study was supported by the National Natural Science Foundation of China (31970516, 31671735).

* 通信作者(Corresponding author): 许文亮, E-mail: wenliangxu@mail.ccnu.edu.cn

第一作者联系方式: E-mail: 13613418576@163.com

Received (收稿日期): 2020-09-22; Accepted (接受日期): 2020-12-01; Published online (网络出版日期): 2020-12-29.

URL: <https://kns.cnki.net/kcms/detail/11.1809.S.20201228.1747.017.html>

棉花作为世界上最大的纤维作物和纺织工业原料,品质是其经济价值的重要考量因素。棉纤维是由棉花胚珠外珠被表皮层的单细胞发育而成,是自然界最长的单细胞之一。细胞壁作为棉纤维的重要组成部分,其合成过程直接决定着棉纤维的品质参数^[1],它是一个复杂的动态结构,由 90% 高分子量的多聚物和少量糖蛋白组成^[2]。虽然糖蛋白组分占比很少,却是细胞壁形态构成和功能发挥的重要因子。富含羟脯氨酸糖蛋白(hydroxyproline (Hyp)-rich glycoproteins, HRGP)代表细胞壁糖蛋白的一个大家族,根据其核心蛋白的糖基化程度,可分为高度糖基化的阿拉伯半乳聚糖蛋白(arabinogalactan proteins, AGP),中度糖基化的伸展蛋白(extensins, EXT)以及轻度糖基化的富含脯氨酸蛋白(proline-rich proteins, PRPs) 3 类^[3-5]。PRP 广泛存在于不同作物中且功能不等,如参与玉米胚的发育^[6-7]、响应大豆干旱和盐胁迫应答^[8]等。我们前期从棉花 cDNA 文库中分离了 3 个 PRP 家族基因,命名为 *GhPRP1*、*GhPRP2* 和 *GhPRP5*,发现 *GhPRP1* 在下胚轴和根中高量表达, *GhPRP2* 在棉花开花后 10 d (days post anthesis, DPA) 胚珠中优势表达^[9]。在拟南芥中表达 *GhPRP5* 基因增强了转基因拟南芥对盐和 ABA 的敏感性^[10]。*GhPRP5*-RNAi 转基因棉纤维的长绒和短绒均增长,酵母双杂交表明, *GhPRP5* 可形成同源和异源二聚体发挥功能,推测 *GhPRP5* 通过复杂的调控网络影响细胞壁的合成从而影响棉纤维的发育^[11]。AGP 可能是自然界中翻译后修饰程度最高的蛋白,其合成过程包括 N 末端信号肽序列的切割、由脯氨酸羟化酶(prolyl 4-hydroxylases, P4Hs)催化的脯氨酸残基的羟基化、GPI 锚的修饰和羟脯氨酸残基上的阿拉伯半乳聚糖糖基化。它是一类 O-糖基化蛋白, O-糖基化部位主要发生在含有丝氨酸(Ser)和羟脯氨酸(Hyp)的肽链中^[12],广泛参与细胞膨胀、细胞分化、生殖发育、体细胞胚胎发育、木质部分化、非生物胁迫应答、激素信号应答等过程^[5]。AGP 功能的发挥很大程度上依赖于占到整个分子量 90% 以上的糖侧链^[13],拟南芥 *mur1* 突变体中 AGP 分子的糖侧链缺失岩藻糖,使得根长明显变短^[14]。在 *N. alata* 中, TTS 蛋白参与花粉管的伸长,但去糖基化的 TTS 蛋白丧失该功能^[15]。近年来,已经筛选出大量编码 AGP 合成的基因,根据核心蛋白中氨基酸的组成, AGP 可分为典型性和非典型性两大类,成束蛋白样阿拉伯半乳聚糖蛋白(Fasciclin-like arabinogalactan proteins,

FLAs)作为典型性 AGP 的代表,广泛存在于拟南芥^[16]、水稻^[17]、小麦^[18]、毛果杨^[19]、大麻^[20]、陆地棉^[21]、海岛棉^[22]等多种植物中。我们之前从棉花中分离了 19 个编码 FLA 的基因,其中的 3 个(*GhFLA1/2/4*)在 10 DPA 纤维中优势表达。我们进一步的研究发现,过表达 *GhFLA1* 可以促进纤维伸长,而抑制 *GhFLA1* 表达减缓了纤维起始和伸长。免疫组学分析显示, *GhFLA1* 影响 AGP 组分尤其是 AG 多糖侧链的合成^[21]。随后,我们又分离了糖基转移酶基因 GT31 家族成员 *GhGalT1*,发现其参与 AGP (主要是 FLA) II 型 AG 糖链 β -1,3-半乳糖苷链的合成,通过调控 *GhFLA* 的表达,导致转基因株系棉纤维细胞壁的形态结构发生明显变化^[23]。综上, AGP 多糖侧链在棉纤维发育过程中发挥重要的作用,但对于催化脯氨酸羟基化的脯氨酸羟化酶在棉纤维发育过程中的作用还不清楚。

P4H 是普遍存在于动物和植物中的一种 2-氧化戊二酸双加氧酶,主要定位于内膜系统,其功能的发挥需要 Fe^{2+} 、 α -酮戊二酸、 O_2 等的参与^[24]。动物等高级生物中 P4H 分为两大类,影响胶原蛋白合成的 C-P4H 和缺氧诱导 HIF-P4H。C-P4H 是胶原蛋白合成的关键因子,这类 P4H 以 $\alpha_2\beta_2$ 四聚体结构发挥功能, β -基序形成二硫键异构酶, α -基序则形成催化亚基,催化底物氨基酸序列-X-Pro-Gly-中的 Pro 羟基化^[25-26]。目前植物中关于 P4H 的报道还较少,与动物中 C-P4H 不同的是其只含有 α 催化亚基,催化底物中位于 Ala、Gln、Hyp、Pro、Ser、Thr 和 Val 之后的 Pro 羟基化,但位于其他氨基酸之后的第 1 个 Pro 不能被羟基化。P4H 催化的底物 EXT 中的 Pro 为连续的 2~4 个单元,并且始终毗邻 Ser(Ser-(Pro)₂₋₄),羟基化后形成 Ser-(Hyp)₂₋₄。当 AGP 作为催化底物时, AP/PA/SP/TP 重复序列是其羟基化的敏感位点。而 PRPs 的 PPVX [KT]、KKPCPP 和 PPV 序列以及其他富含 Pro 的嵌合蛋白 XP_nY 基序是其催化的靶序列。光谱测定和生物信息学方法预测到了拟南芥中 AtPRP4 (富含 32.5% Pro)、At1g09750、At1g31580 (含 27%~28% Pro)、At3g08030、At2g10940 (含 6%~8% Pro)催化底物序列 Pro 的定位,证明了 Pro 羟基化序列的多样性和特异性^[27]。莱茵衣藻中共有 10 个 *P4H* 基因,但只有 *Cr-P4H1* 的功能得到证实,缺陷型 *Cr-P4H1* 导致细胞壁的合成受阻^[28]。

棉花是世界上最重要的经济作物之一,目前关

于棉花中 *GhP4H* 的报道还尚属空白。P4H 功能的发挥离不开保守残基与 2-氧戊二酸盐(2-oxoglutarate)和 Fe^{2+} 的结合。3,4-二羟基苯甲酸乙酯(ethyl-3,4-dihydroxybenzoate, EDHB)和 α,α -联吡啶(α,α -Dipyridyl, DP)是 P4H 酶的 2 种抑制剂, EDHB 可以特异结合 2-oxoglutarate, DP 则通过螯合 Fe^{2+} 抑制 P4H 酶活^[29]。我们前期用不同浓度梯度的 EDHB 和 DP 处理棉花离体胚珠后, 棉纤维伸长受到严重抑制, 且抑制程度与抑制剂浓度呈正相关, 说明 *GhP4H* 参与棉纤维生长发育^[12]。随后, 我们从棉花基因组中识别了 30 个可能编码 P4H 的基因, 许多 P4H 家族基因在纤维发育的不同阶段高量表达, 进一步表明这些基因可能在纤维发育过程中起作用。为研究 P4H 在棉纤维发育过程中的功能, 本研究克隆了一个在棉纤维伸长期高量表达的 P4H 家族基因 *GhP4H2*, 通过农杆菌介导法转化棉花, 获得 *GhP4H2* 转基因棉花植株, 旨在了解 P4H 在棉纤维发育过程中的功能, 为棉纤维发育的分子机制提供新的视野。

1 材料与方法

1.1 植物材料

本研究选用陆地棉(*Gossypium hirsutum* L.) Corker 312 品种为棉花遗传转化受体。1/2 MS 培养基培养棉花无苗苗 5~6 d, 将棉花下胚轴切成 10 mm 左右的小段, 并用含有 *GhP4H2* 过量表达载体或 RNAi 载体的农杆菌 LBA4404 悬浮液浸染 15 min, 然后将外植体转移到共培养基上(不含任何抗生素)在 28℃ (16 h 光照/8 h 黑暗)下共培养 2 d, 后续详细继代培养步骤参照 Qin 等^[23]的方法。获得的转基因幼苗转移至温室土壤中用于目的基因的鉴定和遗传分析。

1.2 棉花基因组 DNA 提取

取棉花幼苗 2 g 左右幼叶, 利用 CTAB 法提取基因组 DNA^[23]。

1.3 载体构建

设计带有 *Xba* I 和 *Sal* I 酶切位点的 *GhP4H2* OE 引物(表 1), 利用 Primer STAR 高保真酶扩增 *GhP4H2* 开放阅读框序列, 连接到克隆载体 pSK 上, 测序验证没有碱基突变后构建到表达载体 pBI121 上获得 pBI121-*GhP4H2* 过表达载体。

两端酶切位点为 *Bam* H I 和 *Xba* I 的 L1 引物(表 1)扩增正义链, 酶切位点为 *Sac* I 和 *Not* I 的 L2 引物(表 1)扩增反义链, 正义链和反义链中间有一个

GhTUA 基因内含子(Intron), 两端酶切位点分别为 *Xba* I 和 *Not* I。将 *GhP4H2* 正义链-Intron-反义链序列构建到 pBI121 质粒中, 获得 pBI121-*GhP4H2* RNAi 载体^[23]。

1.4 RNA 提取和实时定量 RT-PCR

利用 TianGen RNA 试剂盒提取棉花 RNA。逆转录获得 cDNA, 将各个样品的 cDNA 稀释到浓度相当后对基因进行实时定量 RT-PCR 分析^[23]。棉花泛素基因(*Ubiquitin1*, *GhUBI1*)作为内参, 本研究所用基因 RT-PCR 引物见表 1。

1.5 棉纤维 AGP 的分离、纯化及定量分析

按照 Qin 等^[23]的方法, 取 10 g 冷冻干燥的 10 DPA 纤维, 加入液氮研磨成粉末后, 加入 10 mL 缓冲液(50 mmol L⁻¹ Tris-Cl, 10 mmol L⁻¹ pH 8.0 EDTA, 0.1% β -巯基乙醇, 1% (w/v) Triton X-100), 4℃ 孵育 3 h, 14,000×g 离心 10 min, 取上清加入 10 mL 乙醇, 4℃ 孵育过夜。离心弃上清, 5 mL 50 mmol L⁻¹ Tris-Cl 重悬沉淀, 冷冻干燥过夜。350 μ L 1% NaCl 重悬样品, 加入 350 μ L 的 β -Yariv, 混匀, 4℃ 过夜。14,000×g 离心 1 h, 1% NaCl 洗涤沉淀 3 次, 甲醇洗涤沉淀 2 次去除 β -Yariv, 静置干燥, 加入适量二甲基亚砜至沉淀完全溶解, 加入微量亚硫酸钠, 再加入适量超纯水至溶液呈亮黄色, 溶液过 PD-10 柱脱盐, 所得即为 AGP 溶液。

AGP 浓度测定按 Qin 等^[23]的方法, 配制浓度为 1% 琼脂糖, 0.02% 叠氮钠(NaN_3), 0.15 mol L⁻¹ NaCl 以及 10 μ g mL⁻¹ β -Yariv 的溶液 20 mL, 加热至沸腾, 待琼脂糖完全融化后, 将混合液倒入培养皿中, 冷却。用直径为 2 mm 的玻璃吸管对冷却的凝胶打孔。取 1 μ L AGP 点样, 均匀点入每个孔中, 并选择浓度为 0.1~0.6 μ g μ L⁻¹ 的阿拉伯树胶作为标样对照。将平板置于黑暗潮湿的环境中过夜。测量每个扩散色圈的直径, 并计算色圈面积, 根据标样绘制标准曲线并计算出各个待测样品的浓度。

1.6 Immuno Dot blot 分析

根据 Qin 等^[23]的方法进行 Dot blot 分析野生型与转基因棉花株系中 AGP 的差异。分别利用 JIM8 和 JIM13 作为一抗, 碱性磷酸酶标记的山羊抗大鼠 IgG(H+L)作为二抗进行免疫组织化学反应。

1.7 转录组分析

收取野生型和 *GhP4H2* 过表达株系 5 DPA 纤维, 利用 Spectrum plant total RNA Kit 试剂盒提取总 RNA, 纯化后检测 RNA 的完整性。将完整性好的

表 1 本研究所用引物
Table 1 Primers used in this study

引物 Primer	正向序列 Forward sequence (5'–3')	反向序列 Reverse sequence (5'–3')
Gh_D12G1409RT	TAGAACCCAAGCTGGTCAGG	TCGATTTAGACCCTGTCGATG
Gh_D01G1917RT	AGATCTCATGCGCCGTCA	AGGCAAAGAAAGAAACAAGTGTG
Gh_D07G2484RT	TTTGCCTTTTCTGTGAGCA	GCTGGTTGGGATGAACACTT
Gh_D10G0598RT	CTATGAGTGTGATTGGCGTACAGGTAA	AATCCGTGGTCACGTTATATTTGG
Gh_D08G2475RT	TAACCCACCAAGTGGGAAAC	TGAGCTTAAGGGTGGTGCAT
Gh_A07G0433RT	TGAAGAAGACGAAGAACCATCACCA	AGAGCCAGAGCTTTCAACTGATGT
Gh_D05G0131RT	TTAAGCCACCCATTGACCTC	TCCTGAAAGAACCGGACAAC
Gh_D12G2178RT	TCTTTTCCTGGTGAGTGCT	AACGGCAGCTTCAAGATCAG
GhGalT1RT	TCCCTCCTCACCTTCGCCATTG	CCTTGACGATCAGAAGGCATCCAC
GhGalT4RT	GATGATACTTTGAAAATCGTTGCT	ATTTTCTTCAAAGTTTCACCGCT
GhGalT6RT	TCAGGAGTATGTACCCCAGCTTGC	TGCAAATGAAGTCACATGTCGAC
GhGalT7RT	CCTCACCAGATCAACAGCCCTCT	CCGTTGAAGGCCCTGATGATC
GhP4H2RT	ATCGGAATGTGCAAAAGAAGG	AGCCAGGAAGTTCTGCAGTT
GhUBI1RT	CTGAATCTTCGCTTTCACGTTATC	GGGATGCAAATCCTTCGTGAAAAC
GhP4H2 RNAiL1	CTTGATCCGTGCAAAAGAAGGGAATTG	GGGTCTAGAAATAGCCAGGAAGTTCTG
GhP4H2RNAiL2	GGGGAGCTCGTGCAAAAGAAGGGAATTG	CTTGCGGCCCAATAGCCAGGAAGTTCT
GhP4H2OE	GGGTCTAGAATGGCTATTGAGAGGATTT	CTTGTCGACTCAACATACTTTGCAGCTT

RNA 进行 RAN-seq (北京诺禾致源基因公司)。GO (Gene ontology)预测差异基因功能并分类, 以 *P* 值为依据, 筛选出差异表达基因进行 RT-PCR 验证。

2 结果与分析

2.1 GhP4H2 蛋白的系统进化分析和基因表达谱

利用 Clustalx、MEGA7.0 (<http://www.megasoftware.net/>)对拟南芥 P4H 与棉花 GhP4H2 进行系统进化树分析(图 1-A)发现, GhP4H2 与 AtP4H2、AtP4H4 处

于同一个亚支, 亲缘关系最近。

为研究 *GhP4H2* 基因在棉花中的表达模式, 提取棉花的根、下胚轴、子叶、真叶、花瓣、花药、15 d 胚珠及不同发育时期纤维的 RNA, 逆转录后进行实时定量 RT-PCR 分析。*GhP4H2* 基因在子叶、真叶、花药中高量表达, 在纤维发育的不同时期(0~20 DPA)均有较高水平的表达, 且在 5 DPA 和 9 DPA 纤维中保持较高的表达量(图 1-B), 暗示 *GhP4H2* 可能参与棉纤维伸长过程。

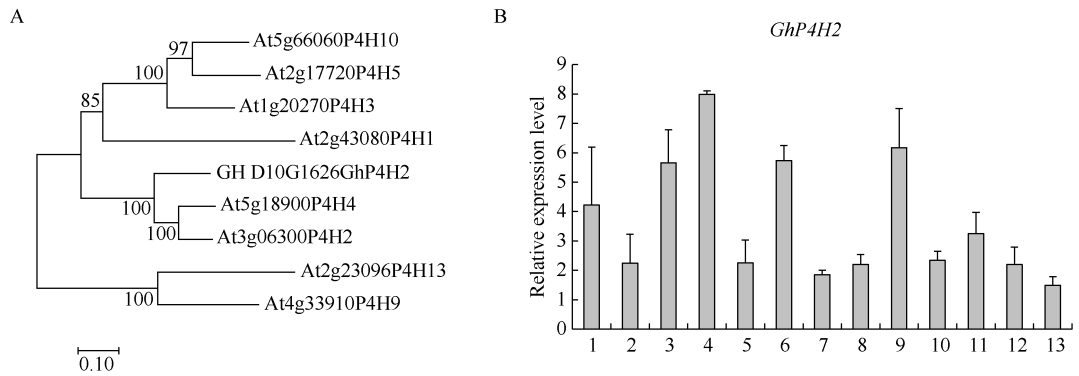


图 1 GhP4H2 的系统进化分析和 *GhP4H2* 基因表达谱

Fig. 1 Phylogenetic analysis of GhP4H2 and expression profile of *GhP4H2*

A: 棉花 GhP4H2 与拟南芥 GhP4H 蛋白系统进化关系; B: *GhP4H2* 基因在棉花各组织的表达谱。1: 根; 2: 下胚轴; 3: 子叶; 4: 真叶; 5: 花瓣; 6: 花药; 7: 15 DPA 胚珠; 8: 0 DPA 纤维(含胚珠); 9: 3 DPA 纤维(含胚珠); 10: 5 DPA 纤维; 11: 9 DPA 纤维; 12: 15 DPA 纤维; 13: 20 DPA 纤维。DPA: 开花后天数。误差线代表标准误差。

A: phylogenetic relationship of GhP4H2 and *Arabidopsis* P4Hs; B: expression analysis of *GhP4H2* in different cotton tissues. 1: root; 2: hypocotyl; 3: cotyledon; 4: leaves; 5: petals; 6: anthers; 7: 15 DPA ovule; 8: 0 DPA fiber (with ovule); 9: 3 DPA fiber (with ovule); 10: 5 DPA fiber; 11: 9 DPA fiber; 12: 15 DPA fiber; 13: 20 DPA fiber. DPA: days post anthesis. Error bar represents the standard deviation.

2.2 转基因棉花株系中 *GhP4H2* 的表达分析

为研究 *GhP4H2* 在棉纤维生长发育过程中的功能, 本研究分别构建了由 35S 启动子驱动的表达 (overexpression, OE) 和 RNA interference (RNAi) 载体, 通过农杆菌介导法转化棉花下胚轴, 获得转基因棉花植株。取不同转基因株系 5 DPA 和 10 DPA 纤维进行表达量鉴定表明, *GhP4H2* 基因在 RNAi (Ri) 株系 L16 和 L28 中表达量显著低于 WT, 在过表达株系 L10 和 L38 中表达量显著增加 (图 2)。我们后续研究选择 RiL16、RiL28、OEL10、OEL38 这 4 个转基因株系进行表型分析。

2.3 *GhP4H2* 转基因棉花成熟纤维表型分析

分析 T₁~T₃ 代成熟纤维表型变化发现, 与野生型相比, *GhP4H2* 过表达株系的成熟棉纤维长度显著变短, 而 RNAi 株系没有明显变化。脱绒后比较不同株系种子形态大小, 未发现明显差异 (图 3), 表明 *GhP4H2* 主要影响纤维发育而对胚珠影响较小。

2.4 *GhP4H2* 转基因棉纤维 AGP 含量分析

考虑到 AGP 基因数目占了棉纤维 HRGP 基因

的绝大多数^[23], 我们推测 AGP 可能是 P4H 的首选底物。为研究 *GhP4H2* 对 AGP 合成的影响, 本研究提取了 10 DPA 纤维 AGP 进行定量检测。以 0.1~0.6 $\mu\text{g } \mu\text{L}^{-1}$ 浓度梯度的阿拉伯树胶作对照, 绘制折线图并计算各转基因株系中 AGP 样品浓度 (图 4)。RiL28 中 AGP 含量明显低于 WT, RiL16 中 AGP 含量有轻微减少; 而过表达株系 OEL38 中含量显著增加, OEL10 中 AGP 含量有轻微增加, 但未见显著差异。

2.5 *GhP4H2* 影响 AGPs 多糖侧链的合成

为研究 *GhP4H2* 对 AGP 糖侧链合成的影响, 本研究用不同的抗体进行了 Dot blot 分析。其中, JIM8、JIM13 能够与 AGP 糖侧链抗原决定簇特异性结合, JIM8 特异结合阿拉伯半乳糖侧链, JIM13 特异结合 $\beta\text{-GlcA-(1,3)-}\alpha\text{-Gal-(1,2)-Rha}$ 侧链。由图 5 可知, RNAi 株系杂交信号与 WT 相比减弱, 并且 JIM8 杂交信号弱于 JIM13。过表达株系则呈现相反趋势。说明转基因株系中 *GhP4H2* 影响了 AGP 糖侧链的组成。

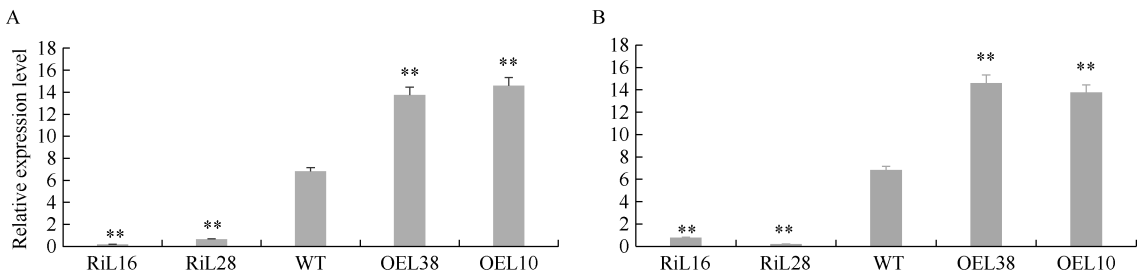


图 2 *GhP4H2* 在野生型和转基因棉花株系中的表达分析

Fig. 2 Expression analysis of *GhP4H2* in wild type and transgenic cotton lines

A: *GhP4H2* 在不同株系 5 DPA 纤维中表达分析; B: *GhP4H2* 在不同株系 10 DPA 纤维中表达分析。RiL16、RiL28 表示 2 个独立的 *GhP4H2* RNAi 株系; WT 表示野生型; OEL10、OEL38 表示 2 个独立的 *GhP4H2* 过表达株系。**表示在 0.01 水平差异显著。

A and B: quantitative RT-PCR analysis of *GhP4H2* expression in 5 DPA (A) and 10 DPA (B) fibers from independent transgenic cotton lines and wild type. RiL16 and RiL28 represent two independent *GhP4H2*-RNAi lines; WT represents the wild type; OEL10 and OEL38 represent two independent *GhP4H2* overexpression lines. ** means significant difference at the 0.01 probability level.

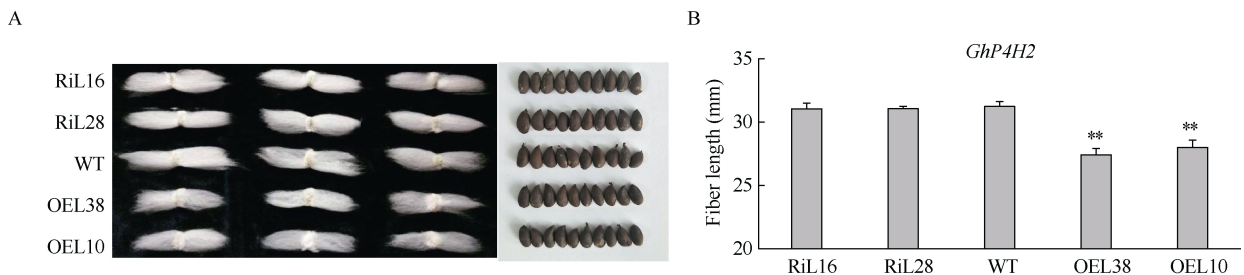


图 3 *GhP4H2* 转基因与野生型棉花成熟纤维表型分析

Fig. 3 Comparison of mature fibers from transgenic cotton lines and wild type

A: *GhP4H2* 转基因棉花与 WT 成熟种子和纤维表型比较; B: *GhP4H2* 转基因棉花与 WT 成熟纤维长度比较。**表示在 0.01 水平差异显著。株系名称缩写同图 2。

A and B: measurement and statistical analysis of mature fiber length and seed size of the transgenic cotton plants and wild type. ** means significant difference at the 0.01 probability level. Abbreviations of lines name are the same as those given in Fig. 2.

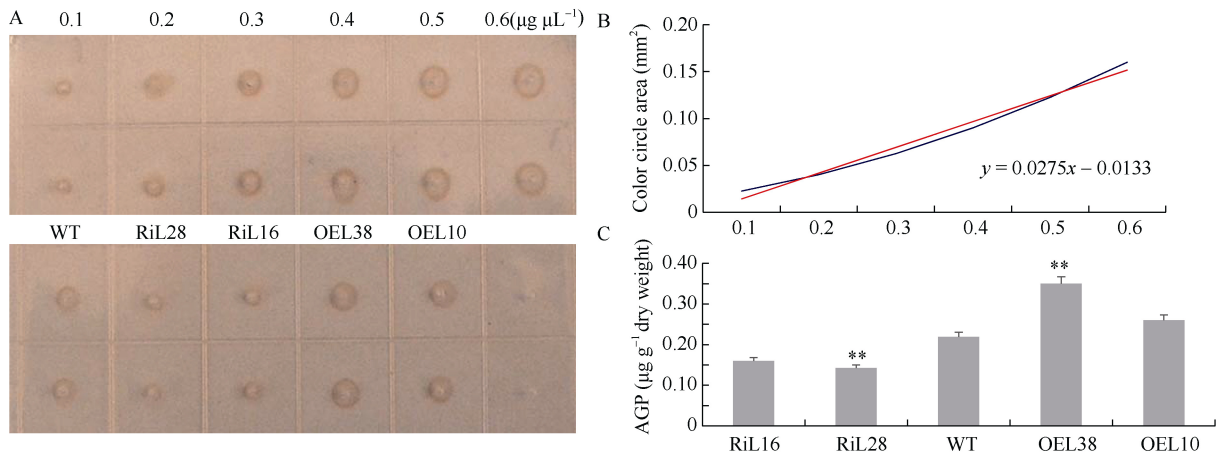


图 4 *GhP4H2* 转基因与野生型棉花 AGP 定量分析

Fig. 4 AGP content in fibers of *GhP4H2* transgenic cotton lines and wild type

A: 不同样品 AGP 特异反应色圈。B: 阿拉伯树胶标准曲线; 横坐标代表 AGP 浓度, 纵坐标代表对应的色圈面积。C: 不同样品 AGP 含量。0.1~0.6 ($\mu\text{g } \mu\text{L}^{-1}$): 阿拉伯树胶浓度。**表示在 0.01 水平差异显著。株系名称缩写同图 2。

A: halos of different samples from transgenic lines and wild type. B: standard curve; Abscissa: the AGPs concentration; Ordinate: the area of the corresponding halos. C: the AGPs content of different samples from transgenic lines and wild type. 0.1–0.6 ($\mu\text{g } \mu\text{L}^{-1}$): gum arabic concentration. ** means significant difference at the 0.01 probability level. Abbreviations of lines name are the same as those given in Fig. 2.

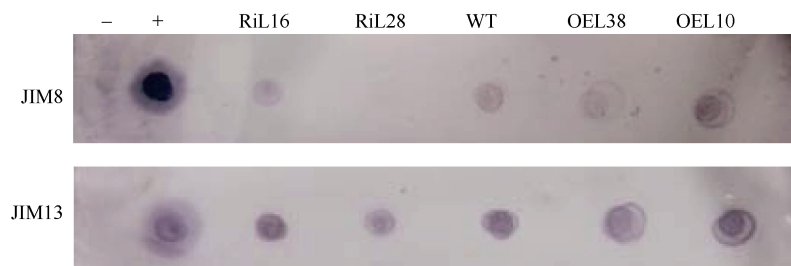


图 5 AGP 糖链抗原决定簇在 *GhP4H2* 转基因与野生型棉花纤维中的分布与丰度

Fig. 5 Distribution and abundance of AGP epitopes in *GhP4H2* transgenic cotton and wild type

–: PBS 阴性对照。+: 1 mol L⁻¹ arabic gum 阳性对照。株系名称缩写同图 2。

–: PBS negative control. +: 1 mol L⁻¹ arabic gum positive control. Abbreviations of lines name are the same as those given in Fig. 2.

2.6 *GhP4H2* 影响棉纤维细胞壁相关基因的表达

为研究 *GhP4H2* 在棉纤维发育过程中的调控通路, 取野生型和过表达株系 OEL38 的 5 DPA 纤维进行转录组测序分析。本研究主要关注细胞壁相关基因。与 WT 相比, OEL38 株系 202 个上调差异表达基因中与细胞壁合成相关的有 15 个(图 6-A), 其中编码 HRGP 的基因有 8 个, 占到细胞壁基因的 53.3%, *Gh_D01G1917*、*Gh_D12G1409* 属于 AGP 类, *Gh_D07G2484*、*Gh_D10G0598* 属于 EXT 类, *Gh_D08G2475*、*Gh_A07G0433*、*Gh_D05G0131*、*Gh_D12G2178* 属于 PRP 类。因此推测这些编码 HRGP 的基因很可能是 *GhP4H2* 作用的靶基因。选择这 8 个基因进行 RT-PCR 表达验证(图 6-B)发现, 这 8 个基因在 OEL38 中表达量均显著升高, 说明 *GhP4H2* 可能通过催化这些底物来调控棉纤维发育。

2.7 *GhP4H2* 影响糖基转移酶基因(*GhGalTs*)的表达

我们之前的研究显示, *GhGalTs* 参与棉纤维发育过程^[23]。*GhGalTs* 位于 *GhP4H* 下游, 为研究 *GhP4H2* 对 *GhGalT* 表达的影响, 本研究选择 *GhGalT* 家族部分基因进行表达分析(图 7)。选择的 7 个基因中有 4 个 *GhGalT* 基因(*GhGalT1*、*GhGalT4*、*GhGalT6* 和 *GhGalT7*)在 *GhP4H2* 过表达株系中的表达量发生变化, 表明 *GhP4H2* 可能影响了 AGP 糖基化过程, 这可能是 AGP 糖侧链发生改变的原因。

3 讨论

先前有研究发现, *AtP4H1* 参与拟南芥低氧应答, 过量表达 *AtP4H1* 致使拟南芥发生根毛增长, 毛状体缺失, 种子变小等变化, 同时导致与缺氧应答相关基因表达量升高。转录组数据显示, *AtP4H1* 还参

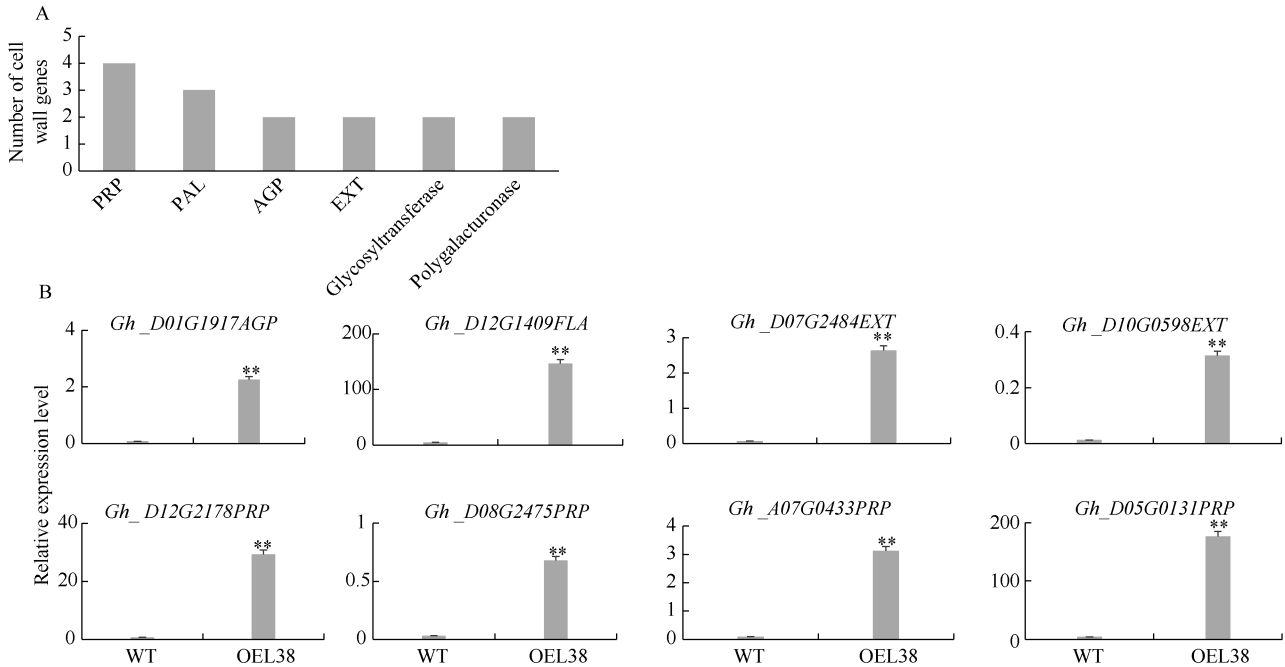


图 6 细胞壁相关差异表达基因分析
Fig. 6 Analysis of differentially expressed genes related cell wall
A: 细胞壁相关基因的主要类别; B: 糖蛋白相关基因的表达验证。PRP: 富含脯氨酸蛋白; AGP: 阿拉伯半乳糖蛋白; EXT: 伸展蛋白; PAL: 苯丙氨酸解氨酶。**表示在 0.01 水平差异显著。株系名称缩写同图 2。
A: major classes of the upregulated cell wall genes in transgenic fibers; B: expression analysis of genes encoding glycoproteins. PRP: proline-rich protein; AGP: arabinogalactan protein; EXT: extensin; PAL: phenylalanine ammonia-lyase. ** means significant difference at the 0.01 probability level. Abbreviations of lines name are the same as those given in Fig. 2.

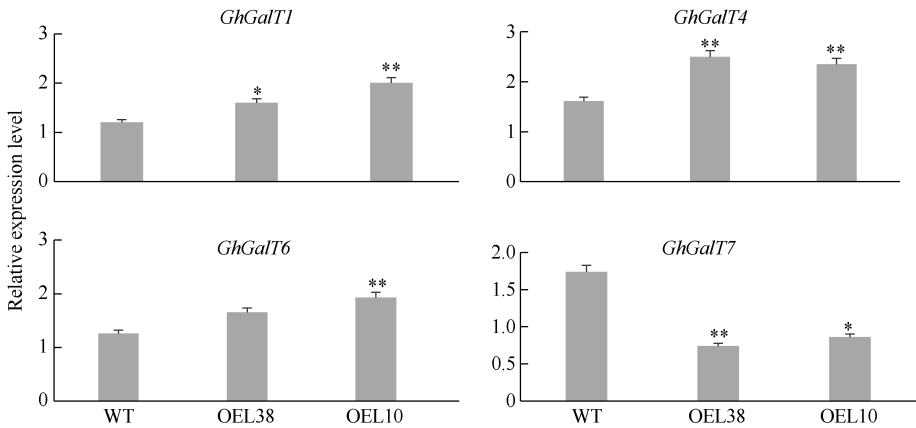


图 7 部分 *GhGalT* 基因表达分析
Fig. 7 Expression analysis of several *GhGalT* genes
*、**分别表示在 0.05 和 0.01 水平差异显著。株系名称缩写同图 2。
* and ** mean significant differences at the 0.05 and 0.01 probability levels, respectively. Abbreviations of lines name are the same as those given in Fig. 2.

与光合作用、JA 信号转导、植物激素调控等过程^[30]。另外, 有研究表明, AtP4H5 分别通过与 AtP4H2 和 AtP4H13 形成异源二聚体发挥 Pro 羟基化功能, 并且优先羟基化 EXT, 三者共同调控拟南芥根毛的伸长^[31]。敲除 *SIP4H1* 会促进番茄根的伸长和叶子的发育, 免疫组学分析显示, 与 JIM8 特异性结合的 AGP

糖链组分缺失^[32]。拟南芥 *Atp4h2*、*Atp4h5* 和 *Atp4h13* 突变体中, 根毛相对于野生型都变短, DP 和 EDHB 处理也会抑制根毛的伸长^[31]。前期我们分别用 DP 和 EDHB 处理棉花离体胚发现, 棉纤维的生长发育受到了明显的抑制, 推测 *GhP4Hs* 参与调控棉纤维的生长发育^[12]。本研究克隆了 *GhP4H2* 基因获得过

表达和 RNAi 不同株系转基因棉花, 对 $T_1 \sim T_3$ 代成熟棉铃进行表型分析发现, *GhP4H2*-RNAi 株系长绒与 WT 相比并没有明显变化, 这可能是由于基因功能冗余的原因。而 *GhP4H2* OE 株系的成熟纤维明显短于 WT。Dot blot 分析显示, JIM8 和 JIM13 的杂交信号均增强。我们前期证实 *GhFLA1* 促进棉纤维的生长发育, 免疫组学分析显示, *GhFLA1* 转基因株系 AGP 侧链糖组分发生不同程度的变化, 其中过表达株系中 JIM13 杂交信号减弱, 说明 *GhFLA1* 过表达株系中与 JIM13 特异结合的 AGP 表位抗原决定簇减少^[21]。随后我们发现过量表达 *GhGalT1* 的棉纤维伸长受到抑制, 经免疫组学分析, JIM8 和 JIM13 的杂交信号均增强, 表明过量表达 *GhGalT1* 会增加与 JIM8 和 JIM13 特异结合的 AGP 表位抗原决定簇丰度^[23]。本研究与 *GhGalT1* 过表达棉纤维取得的结果一致, 我们推测 *GhP4H2* 过表达转基因棉纤维变短的原因与 *GhGalT1* 类似。而在 RNAi 株系中 AGP 变化的程度可能不足以明显影响纤维发育。前期我们的结果显示, *GhGalT1* 正调控 *FLA* 基因的表达^[23], 本研究依据转录组及 RT-PCR 结果, 证实筛选到的 8 个表达 AGP、EXT、PRP 的差异基因受 *GhP4H2* 调控, 可能是其下游靶基因, 后续又发现 *GhP4H2* 影响下游 *GalTs* 的表达, 暗示 *GhP4H2* 不仅仅影响 HRGP 羟基化, 对参与糖基化的糖基转移酶基因的表达也有影响, 而 *GhP4H2* 转基因株系中 AGP 的变化很可能正是由于 *GhP4H2* 与 *GhGalTs* 协同调控引起。但过表达株系成熟纤维表型差异是否与 AGP 核心蛋白变化有关? AGP 可能是 *GhP4H2* 优先催化的底物, EXT 或 PRP 也是其作用底物吗? 这些问题都有待深入研究。

4 结论

本研究克隆了一个脯氨酸羟化酶基因 *GhP4H2*, 该基因通过影响 *HRGP* 基因和糖基转移酶基因的表达来影响 AGP 的合成。过表达株系成熟纤维变短与 AGP 含量增加和糖侧链组分改变相关, 而 RNAi 株系成熟纤维长度未发生明显变化, 可能是由于基因功能冗余导致。推测过表达基因株系出现的表型差异是由 AGP 糖侧链组分变化所致, *GhP4H2* 影响棉花纤维发育的分子机制还需进一步研究。

References

[1] Haigler C H, Betancur L, Stiff M R, Tuttle J R. Cotton fiber: a powerful single-cell model for cell wall and cellulose research.

- Front Plant Sci*, 2012, 3: 104.
- [2] Minorsky P V. The wall becomes surmountable. *Plant Physiol*, 2002, 128: 345–353.
- [3] Ellis M, Egelund J, Schultz C J, Bacic A. Arabinogalactan-proteins: key regulators at the cell surface? *Plant Physiol*, 2010, 153: 403–419.
- [4] Showalter A M. Structure and function of plant cell wall proteins. *Plant Cell*, 1993, 5: 9–23.
- [5] Showalter A M, Keppler B, Lichtenberg J, Gu D, Welch L R. A bioinformatics approach to the identification, classification, and analysis of hydroxyproline-rich glycoproteins. *Plant Physiol*, 2010, 153: 485–513.
- [6] Estanyol J M, Avita L R, Puigdomenech P. A maize embryo-specific gene encode a proline-rich and hydrophobic protein. *Plant Cell*, 1992, 4: 413–423.
- [7] Estanyol J M, Puigdomenech P. Development and hormonal regulation of genes coding for proline-rich proteins in female inflorescences and kernels of maize. *Plant Physiol*, 1998, 116: 485–494.
- [8] He C Y, Zhang J S, Chen S Y. A soybean gene encoding a proline-rich protein is regulated by salicylic acid, an endogenous circadian rhythm and by various stresses. *Theor Appl Genet*, 2002, 104: 1125–1131.
- [9] 许文亮, 黄耿青, 王秀兰, 邵付菊, 汪虹, 李学宝. 两个棉花 *HyPRP* 基因的分子鉴定与初步表达分析. *作物学报*, 2007, 33: 1146–1153.
- Xu W L, Huang G Q, Wang X L, Tai F J, Wang H, Li X B. Molecular characterization and expression analysis of two genes encoding hybrid proline-rich proteins in cotton. *Acta Agron Sin*, 2007, 33: 1146–1153 (in Chinese with English abstract).
- [10] 张德静, 秦丽霞, 李龙, 饶玥, 李学宝, 许文亮. 异源表达棉花 *GhPRP5* 基因增强了拟南芥对盐和 ABA 的敏感性. *作物学报*, 2013, 39: 563–569.
- Zhang D J, Qin L X, Li L, Rao Y, Li X B, Xu W L. Expression of cotton *GhPRP5* gene in *Arabidopsis* enhances susceptibility to ABA and salt stresses. *Acta Agron Sin*, 2013, 39: 563–569 (in Chinese with English abstract).
- [11] Xu W L, Zhang D J, Wu Y F, Qin L X, Huang G Q, Li J, Li L, Li X B. Cotton *PRP5* gene encoding a proline-rich protein is involved in fiber development. *Plant Mol Biol*, 2013, 82: 353–365.
- [12] 秦丽霞, 李学宝, 许文亮. 植物阿拉伯半乳糖蛋白 AG 糖链的合成. *植物生理学报*, 2018, 54: 1263–1271.
- Qin L X, Li X B, Xu W L. Synthesis of plant arabinogalactan protein AG sugar chain. *Plant Physiol J*, 2018, 54: 1263–1271 (in Chinese with English abstract).
- [13] Shpak E, Leykam J F, Kieliszewski M J. Synthetic genes for glycoprotein design and the elucidation of hydroxyproline-O-glycosylation. *Proc Natl Acad Sci USA*, 1999, 96: 14736–14741.
- [14] Van Hengel A J, Roberts K. Fucosylated arabinogalactan-proteins are required for full root cell elongation in *Arabidopsis*. *Plant J*, 2002, 32: 105–113.
- [15] Wu H M, Wong E, Ogdahl J, Cheung A Y. A pollen tube growth promoting arabinogalactan protein from *Nicotiana glauca* is similar to the tobacco TTS protein. *Plant J*, 2000, 22: 165–176.
- [16] Schultz C J, Rumsewicz M P, Johnson K L, Jones B J, Gaspar Y M, Bacic A. Using genomic resources to guide research

- directions. The arabinogalactan protein gene family as a test case. *Plant Physiol*, 2002, 129: 1448–1463.
- [17] Ma H, Zhao J. Genome-wide identification, classification, and expression analysis of the arabinogalactan protein gene family in rice (*Oryza sativa* L.). *J Exp Bot*, 2010, 61: 2647–2668.
- [18] Faik A, Abouzouhair J, Sarhan F. Putative fasciclin-like arabinogalactan-proteins (FLA) in wheat (*Triticum aestivum*) and rice (*Oryza sativa*): identification and bioinformatic analyses. *Mol Genet Genomics*, 2006, 276: 478–494.
- [19] Showalter A M, Keppler B D, Liu X, Lichtenberg J, Welch L R. Bioinformatic identification and analysis of hydroxyproline-rich glycoproteins in populus trichocarpa. *BMC Plant Biol*, 2016, 16: 229.
- [20] Guerriero G, Mangeot P L, Legay S, Behr M, Lutts S, Siddiqui K S, Hausman J F. Identification of fasciclin-like arabinogalactan proteins in textile hemp (*Cannabis sativa* L.): in silico analyses and gene expression patterns in different tissues. *BMC Genomics*, 2017, 18: 741.
- [21] Huang G Q, Gong S Y, Xu W L, Li W, Li P, Zhang C J, Li D D, Zheng Y, Li F G, Li X B. A fasciclin-like arabinogalactan protein, GhFLA1, is involved in fiber initiation and elongation of cotton. *Plant Physiol*, 2013, 161: 1278–1290.
- [22] Liu H, Shi R, Wang X, Pan Y, Li Z, Yang X, Zhang G, Ma Z. Characterization and expression analysis of a fiber differentially expressed fasciclin-like arabinogalactan protein gene in sea island cotton fibers. *PLoS One*, 2013, 8: e70185.
- [23] Qin L X, Chen Y, Zeng W, Li Y, Gao L, Li D D, Bacic A, Xu W L, Li X B. The cotton β -galactosyltransferase 1 (GalT1) that galactosylates arabinogalactan proteins participates in controlling fiber development. *Plant J*, 2017, 89: 957–971.
- [24] Marzol E, Borassi C, Bringas M, Sede A, Garcia D R R, Capece L, Estevez J M. Filling the gaps to solve the extensin puzzle. *Mol Plant*, 2018, 11: 645–658.
- [25] Kivirikko K I, Myllyharju J. Prolyl 4-hydroxylases and their protein disulfide isomerase subunit. *Matrix Biol*, 1998, 16: 357–368.
- [26] Kivirikko K I, Pihlajaniemi T. Collagen hydroxylases and the protein disulfide isomerase subunit of prolyl-4-hydroxylases. *Adv Enzymol Relat Areas Mol Biol*, 1998, 72: 325–398.
- [27] Durufle H, Herve V, Balliau T, Zivy M, Dunand C, Jamet E. Proline hydroxylation in cell wall proteins: is it yet possible to define rules? *Front Plant Sci*, 2017, 8: 1802.
- [28] Keskiö K, Hieta R, Sormunen R, Myllyharju J. *Chlamydomonas reinhardtii* has multiple prolyl 4-hydroxylases, one of which is essential for proper cell wall assembly. *Plant Cell*, 2007, 19: 256–269.
- [29] Koski M K, Hieta R, Böllner C, Kivirikko K I, Myllyharju J, Wierenga R K. The active site of an algal prolyl 4-hydroxylase has a large structural plasticity. *J Biol Chem*, 2007, 282: 37112–37123.
- [30] Asif M H, Trivedi P K, Misra P, Nath P. Prolyl-4-hydroxylase (*AtP4H1*) mediates and mimics low oxygen response in *Arabidopsis thaliana*. *Funct Integr Genomics*, 2009, 9: 525–535.
- [31] Velasquez S M, Ricardi M M, Poulsen C P, Oikawa A, Dilokpimol A, Halim A, Mangano S, Juarez S P D, Marzol E, Salter J D S, Dorosz J G, Borassi C, Moller S R, Buono R, Ohsawa Y, Matsuo K, Otegui M S, Scheller H V, Geshi N, Petersen B L, Iusem N D, Estevez J M. Complex regulation of prolyl-4-hydroxylases impacts root hair expansion. *Mol Plant*, 2015, 8: 734–746.
- [32] Fragkostefanakis S, Sedeek K E M, Raad M, Zaki M S, Kalaitzis P. Virus induced gene silencing of three putative prolyl 4-hydroxylases enhances plant growth in tomato (*Solanum lycopersicum*). *Plant Mol Biol*, 2014, 85: 459–471.