

杨丹, 鲍军秀, 陈沛, 等. 棉花查尔酮异构酶基因家族的鉴定及表达分析[J]. 华中农业大学学报, 2023, 42(5): 94-104.  
DOI: 10.13300/j.cnki.hnlkxb.2023.05.011

## 棉花查尔酮异构酶基因家族的鉴定及表达分析

杨丹<sup>1</sup>, 鲍军秀<sup>1</sup>, 陈沛<sup>1</sup>, 王斐<sup>1,2</sup>, 汤寿伍<sup>3</sup>, 刘海峰<sup>3</sup>, 李鸿彬<sup>1,2,3</sup>

1. 石河子大学生命科学学院, 石河子 832003; 2. 绿洲城镇与山盆系统生态兵团重点实验室/新疆植物药资源  
利用教育部重点实验室, 石河子 832000; 3. 中国彩棉(集团)股份有限公司, 乌鲁木齐 830001

**摘要** 为研究陆地棉查尔酮异构酶CHI(chalcone isomerase)在彩色棉纤维发育及响应胁迫中的作用, 对陆地棉基因组中*GhCHI*基因家族进行鉴定, 开展蛋白理化性质、结构域、启动子顺式作用元件、蛋白质高级结构、系统进化分析, 利用转录组数据和实时荧光定量PCR分析其表达特征。结果显示, 从陆地棉基因组中鉴定获得12个*GhCHI*基因家族成员, 可分为2个亚家族, 具有典型的查尔酮超家族结构域, 编码201~452个氨基酸, 主要为亲水性蛋白, 二级结构主要以 $\alpha$ -螺旋和无规则卷曲为主。*GhCHI*基因启动子顺式作用元件类型主要包含光反应元件、激素响应元件及参与类黄酮生物合成调控的元件等。表达分析结果显示*GhCHI1*、*GhCHI2*、*GhCHI3*、*GhCHI4*与纤维发育密切相关, 尤其在彩色棉纤维发育后期的色素沉积着色时期表达水平较高; 这4个基因在叶、花托、雌蕊中也具有较高的表达水平; *GhCHI1*、*GhCHI2*、*GhCHI3*在热胁迫、盐胁迫和干旱胁迫下均受到显著的诱导表达。*GhCHI1*—*GhCHI4*蛋白互作分析结果显示, 它们可能和参与类黄酮合成的2-氧代戊二酸3-双加氧酶和类黄酮3'-单加氧酶等蛋白质存在相互作用。结果表明, *GhCHI1*、*GhCHI2*、*GhCHI3*和*GhCHI4*基因在彩色棉的纤维发育和胁迫应答中发挥重要作用。

**关键词** 天然彩色棉; 查尔酮异构酶; 基因家族; 纤维发育; 色素沉积

**中图分类号** S562 **文献标识码** A **文章编号** 1000-2421(2023)05-0094-11

随着社会发展和生活水平的不断提高, 人们的绿色消费观念不断加深, 对绿色纺织品的需求量不断增大。作为陆地棉的重要组成部分, 天然彩色棉纤维是绿色纺织品的重要原料来源。目前生产上天然彩色棉纤维主要包括棕色棉纤维和绿色棉纤维, 与白色棉纤维相比, 彩色棉纤维存在质量略差、着色不稳定、色泽不均匀等问题, 制约了天然彩色棉产品的应用和推广<sup>[1-2]</sup>。因此, 研究和分析彩色棉纤维色素沉着的分子机制, 对推动彩色棉的应用和产业发展具有重要的理论意义。前人研究结果表明, 天然彩色棉纤维中色素的沉积主要是由特定类型黄酮类化合物在纤维中的积累造成<sup>[3]</sup>, 如不同花青素及其衍生物<sup>[4]</sup>等, 通过对黄酮类化合物合成途径的修饰有望改变彩棉纤维的颜色和提高纤维质量<sup>[2]</sup>。

查尔酮异构酶(CHI)是黄酮类化合物合成途径的关键限速酶, 决定了植物细胞内黄酮类化合物的

产生<sup>[5]</sup>。查尔酮异构酶基因表达量的提升可以促进植物细胞内黄酮类次级代谢物质的合成, 在改良植物花色、提高药用成分、逆境胁迫应答中发挥重要作用<sup>[6-7]</sup>。植物受到干旱胁迫时黄酮类物质含量增加, 促进气孔关闭, 参与植物对干旱的响应过程<sup>[8]</sup>。在盐胁迫和低温胁迫下, 黄酮类物质含量增加, 可以清除植物细胞内产生的活性氧进而降低植物受到的伤害<sup>[9-10]</sup>。查尔酮异构酶的产生与活性由其编码基因*CHI*决定, 自1987年首次在豌豆中发现*CHI*基因以来, 在许多植物中已鉴定得到*CHI*基因<sup>[11]</sup>。*CHI*基因的表达与功能发挥受到多种因素的调控, 如植物激素等<sup>[12]</sup>。茉莉酸和生长素显著诱导海南龙血树*DcCHI*的诱导表达<sup>[13]</sup>; 脱落酸、赤霉素和水杨酸均能够诱导白木香愈伤组织中*AsCHI1*基因表达<sup>[14]</sup>。

目前, 关于棉花*GhCHI*基因已有部分研究, 宋成攀等<sup>[15]</sup>从陆地棉中克隆得到*GhCHI*基因, 利用病毒

收稿日期: 2023-04-24

基金项目: 新疆维吾尔自治区创新环境建设专项(2020D14035); 新疆生产建设兵团科技攻关与成果转化项目(2016AC017)

杨丹, E-mail: 1621220721@qq.com

通信作者: 李鸿彬, E-mail: lihb@shzu.edu.cn; 刘海峰, E-mail: cncclhf@163.com

诱导的基因沉默(VIGS)技术抑制该基因表达后,导致棉花对黄萎病的抗性显著降低,但对陆地棉 *GhCHI* 基因进行系统的全基因组鉴定和非生物胁迫下的表达分析还未见报道。本研究对陆地棉 *GhCHI* 基因家族进行系统分析,深入研究 *GhCHI* 在棉纤维发育和色素沉积中的作用,旨在为进一步利用基因工程技术培育棉花新品种奠定基础。

## 1 材料与方法

### 1.1 陆地棉 *GhCHI* 基因家族成员数据获取

从 Pfam 数据库(<https://pfam-legacy.xfam.org/>)下载棉花 CHI 隐马尔科夫模型 Chalcone (PF02431) 作为种子序列,从 CottonGen 数据库(<https://www.cottongen.org/>)和 CottonFGD 数据库(<https://cottonfgd.net/>)下载棉花基因组数据,对这些基因组数据进行分析筛选 *GhCHI* 基因。

### 1.2 陆地棉 *GhCHI* 基因家族生物信息分析

1) 陆地棉 *GhCHI* 基因家族成员鉴定及其理化性质分析。利用 TBtools 中的 Simple HMM Search 工具,确定陆地棉 *GhCHI* 基因家族成员的基因 ID,通过 TBtools 中的 Fasta Extract 工具提取陆地棉 *GhCHI* 基因家族成员的蛋白序列。将提取序列提交 Clustal Omega(<https://www.ebi.ac.uk/Tools/msa/clustala>)进行多序列比对;利用 ProtParam tool(<https://web.expasy.org/>)对鉴定蛋白结构域后的基因序列进行理化性质分析。利用 MEME(<http://alter-nate.meme-suite.org/tools/meme>)在线分析棉花 GhCHI1—GhCHI12 蛋白保守基序。利用 NCBI 中 BatchCD Search 工具分析 GhCHI1—GhCHI12 蛋白结构域并通过 TBtools 处理得到可视化图形。

2) 陆地棉 *GhCHI* 基因启动子分析及系统发育树构建。利用 TBtools 中的 GXF Sequences Extract 工具,从基因组文件中提取家族基因起始密码子上游 2 000 bp 的核酸序列,经 Fasta Stats 验证后,再用 Fasta Extract 提取启动子序列。将筛选后的陆地棉 *GhCHI* 基因家族的启动子序列提交至 PlantCARE(<https://bioinformatics.psb.ugent.be/webtools/plantcare/html/>)分析顺式作用元件组成,再利用 TBtools 的 Basic Biosequence View 进行可视化。利用 Mega 软件通过邻接法(Neighbor-Joining)构建陆地棉 GhCHI 与拟南芥 AtCHI 和水稻 OsCHI(蛋白序列来源于 JGI,<https://phytozome-next.jgi.doe.gov/>)系统发育树。

3) GhCHI 蛋白质结构预测及蛋白相互作用分析。利用 SOMPA([https://npsa-prabi.ibcp.fr/cgi-bin/npsa\\_automat.pl?page=npsa\\_sopma.html](https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_sopma.html))、SWISS-MODEL(<https://swissmodel.expasy.org/>)对陆地棉 GhCHI 蛋白序列进行二级结构、三级结构分析。利用 STRING 数据库(<https://cn.string-db.org/>)在线分析基因间的互作关系。

4) 陆地棉 *GhCHI* 基因家族成员在棉花不同组织、不同胁迫处理以及不同时期多种纤维中的表达分析。利用 R 包 Heatmap,以 Log<sub>2</sub>(FPKM+1)对棉花在开花后不同发育期的胚珠和纤维材料 0 d(胚珠)、5 d(胚珠+纤维)、10 d(纤维)、15 d(纤维)、20 d(纤维)转录组数据进行均一化,分析 *GhCHI* 基因在纤维发育不同时期的表达情况。参考 NCBI 数据库中经过 0、1、3、6、12 h 冷、热、盐、干旱胁迫处理的不同转录组数据分析 *GhCHI* 基因在棉花不同胁迫处理及棉花不同组织中的表达变化。

### 1.3 *GhCHI* 基因家族成员在棉花纤维发育中的表达分析

利用天根生化科技(北京)有限公司的多糖多酚植物总 RNA 提取试剂盒(DP441,北京)和 cDNA 第一链合成预混试剂(KR211,北京)提取棕色棉、白色棉和绿色棉不同发育时期的纤维样品(开花后 0、5、10、15 和 20 d)的总 RNA,反转录获得 cDNA。使用 Primer Premier 5.0 软件设计 *ChCHI* 基因特异性引物(表 1),并由生工生物工程(中国上海)股份有限公司合成。以棉花 *GhUBQ7* 作为内参基因,qRT-PCR 分析 *GhCHI* 基因在棉纤维发育过程中的表达特征,每个样本 3 次生物学重复,使用  $2^{-\Delta\Delta C_t}$  法计算相对表达量。

### 1.4 数据处理

用 Excel 2019 和 IBM SPSS Statistics 27 进行数据处理和差异显著性分析。

## 2 结果与分析

### 2.1 陆地棉 *GhCHI* 基因家族的鉴定与序列分析

从陆地棉基因组数据库中鉴定获得 12 个 *GhCHI* 基因家族成员(*GhCHI1*—*GhCHI12*),陆地棉 *GhCHI* 基因的开放阅读框由 1 013~7 460 个核苷酸组成,编码的蛋白质包含 201~451 个氨基酸、等电点为 4.86~9.38,GhCHI 蛋白主要为亲水蛋白, $\alpha$ -螺旋和无规则卷曲是 GhCHI 蛋白二级结构的主要形式

表 1 荧光定量 PCR 引物  
Table 1 Primers of qRT-PCR

基因名称/基因编号 Gene name/Gene ID	引物名称 Primer name	引物序列(5'-3') Primer sequence (5'-3')
<i>GhCHI1</i> /Gh_A13G0197	<i>GhCHI1</i> -F	CCACCAAAACCCTGTTCCCTT
	<i>GhCHI1</i> -R	ACTCCACCGCGCTCTTACCC
<i>GhCHI2</i> /Gh_D13G0211	<i>GhCHI2</i> -F	AGCCTCCTGGTTCCACCAAA
	<i>GhCHI2</i> -R	CACCGCGCTCTTACCCTTCC
<i>GhCHI3</i> /Gh_A05G3491	<i>GhCHI3</i> -F	CCTTTTCCTCCACAAATCAC
	<i>GhCHI3</i> -R	ATCTTCAGCTAGCACGTTCC
<i>GhCHI4</i> /Gh_D04G1926	<i>GhCHI4</i> -F	TTCACGACGGAGGGAAGGA
	<i>GhCHI4</i> -R	GCAAAGCTCAGCGGAGAGGG
<i>GhCHI5</i> /Gh_D13G1602	<i>GhCHI5</i> -F	CAAATTAACTCAAACCCATT
	<i>GhCHI5</i> -R	CAAAAACTTTCTCCCTGTAC
<i>GhCHI6</i> /Gh_A06G0722	<i>GhCHI6</i> -F	GTGGCAGTCTATCTTTTCTA
	<i>GhCHI6</i> -R	TGATGATCTTCATGGTTCTT
<i>GhCHI7</i> /Gh_D08G2001	<i>GhCHI7</i> -F	GAACCACCCGTTCTTCCAAC
	<i>GhCHI7</i> -R	GCTTTTTTTCCTCAACCCAA
<i>GhCHI8</i> /Gh_D06G0839	<i>GhCHI8</i> -F	CGACACATTCTTTGCTATC
	<i>GhCHI8</i> -R	GCATCCTTGACCAACTTCAC
<i>GhCHI9</i> /Gh_A08G2369	<i>GhCHI9</i> -F	TTGCTGTCGTCTCGCAGAAT
	<i>GhCHI9</i> -R	CGGGAAAGAAACCCACTCT
<i>GhCHI10</i> /Gh_D07G0160	<i>GhCHI10</i> -F	CTTCTGAGGTGCTTGTTGGT
	<i>GhCHI10</i> -R	AAAGTCATTGTGTTTGTTCG
<i>GhCHI11</i> /Gh_A07G2350	<i>GhCHI11</i> -F	CTTCTGAGGTGCTTGTTGGT
	<i>GhCHI11</i> -R	AAAGTCATTGTGTTTGTTCG
<i>GhCHI12</i> /Gh_D08G2008	<i>GhCHI12</i> -F	CTACCTTTTCCAATTTTCT
	<i>GhCHI12</i> -R	CTATATCCCCACTTCACACC
<i>GhUBQ7</i> /DQ116441	<i>GhUBQ7</i> -F	GAAGGCATTCCACCTGACCAAC
	<i>GhUBQ7</i> -R	CTTGACCTTCTTCTTCTGTGCTTG

(表 2)。 *GhCHI* 基因结构和保守结构域分析结果发现, *GhCHI6*、*GhCHI8* 和 *GhCHI10* 含有 6 个保守基序 (motif 1 ~ motif 6), 其他 *GhCHI* 含有 1~5 个数量不等的基序 (图 1); 不同 *GhCHI* 蛋白含有不同类型的查尔酮超家族结构域 (Chalcone\_3 superfamily、PLN02804 family protein、PLN02311 family protein、PLN03175 family protein 和 PLN03174 family protein) (图 2)。

2.2 棉花 GhCHI 的进化树分析

为了研究棉花 *GhCHI* 基因家族在进化过程中的亲缘关系, 选取拟南芥和水稻 CHI 蛋白与陆地棉 *GhCHI* 蛋白构建系统进化树。结果显示, 这些 CHI 可以分为 3 个组: Group 1 (*OsCHI2*、*OsCHI4*、*GhCHI3*、*GhCHI4*)、Group 2 (*AtCHI1*–*AtCHI8*、*AtCHI10*、*AtCHI12*)、Group 3 (*OsCHI1*、*OsCHI3*、*AtCHI9*、*GhCHI1*、*GhCHI2*、*GhCHI5*–*GhCHI12*)。 Group 1 含有 2

个棉花 CHI, Group 2 均为拟南芥 CHI, 都是具有物种特异性的分支, Group 3 中的 CHI 在棉花中具有更紧密的进化关系 (图 3)。

2.3 棉花 GhCHI 基因启动子顺式作用元件组成分析

选取棉花 *GhCHI* 基因起始密码子上游 2 000 bp 的启动子序列, 利用 PlantCARE 分析基因启动子序列中顺式作用元件的组成, 结果显示, 棉花 *GhCHI* 基因启动子中包含大量的光响应元件, 尤其是 *GhCHI10* 启动子含有 16 个光响应元件, 暗示了它们与光响应之间的密切联系。此外, 大多数 *GhCHI* 启动子还含有脱落酸响应元件、茉莉酸响应元件以及厌氧诱导顺式作用元件。 *GhCHI1*、*GhCHI6*、*GhCHI7*、*GhCHI9* 的启动子中均含有参与类黄酮生物合成基因调控的 MYB 结合位点, 推测这些基因在类黄酮化合物生物合成中发挥重要作用 (图 4)。

表 2 陆地棉 <i>GhCHI</i> 基因的理化性质和二级结构分析									
Table 2 Physicochemical property and predicted secondary structure analysis of <i>GhCHI</i> genes									
基因 ID Gene ID	基因名称 Gene name	核苷酸数 Nucleotide number	氨基酸数量 Amino acid number	等电点 Isoelectric point	亲水指数 Grand average of hydropathicity	α-螺旋/% α-Helix	延伸链/% Extended strand	β-折叠/% β-Turn	无规则卷曲/% Random coil
Gh_A13G0197	<i>GhCHI1</i>	2 750	227	5.02	−0.043	43.61	19.38	9.25	27.75
Gh_D13G0211	<i>GhCHI2</i>	3 097	227	4.92	−0.058	40.53	18.50	9.25	31.72
Gh_A05G3491	<i>GhCHI3</i>	1 056	216	4.86	−0.090	43.52	24.54	7.41	24.54
Gh_D04G1926	<i>GhCHI4</i>	1 013	203	4.98	−0.172	46.80	21.18	7.39	24.63
Gh_D13G1602	<i>GhCHI5</i>	1 575	278	9.13	−0.063	33.45	19.42	2.52	44.60
Gh_A06G0722	<i>GhCHI6</i>	2 881	432	8.52	−0.045	38.43	16.67	5.09	39.81
Gh_D08G2001	<i>GhCHI7</i>	2 410	297	9.13	−0.154	40.07	19.87	5.72	34.34
Gh_D06G0839	<i>GhCHI8</i>	2 867	432	8.71	−0.080	36.8	16.90	4.86	41.44
Gh_A08G2369	<i>GhCHI9</i>	7 460	451	9.38	−0.116	38.58	21.06	5.54	34.81
Gh_D07G0160	<i>GhCHI10</i>	2 557	439	6.95	−0.031	39.64	15.26	5.01	40.09
Gh_A07G2350	<i>GhCHI11</i>	6 264	243	7.00	−0.299	28.40	19.34	6.58	45.68
Gh_D08G2008	<i>GhCHI12</i>	1 056	201	8.90	0.131	42.79	23.38	6.47	27.36

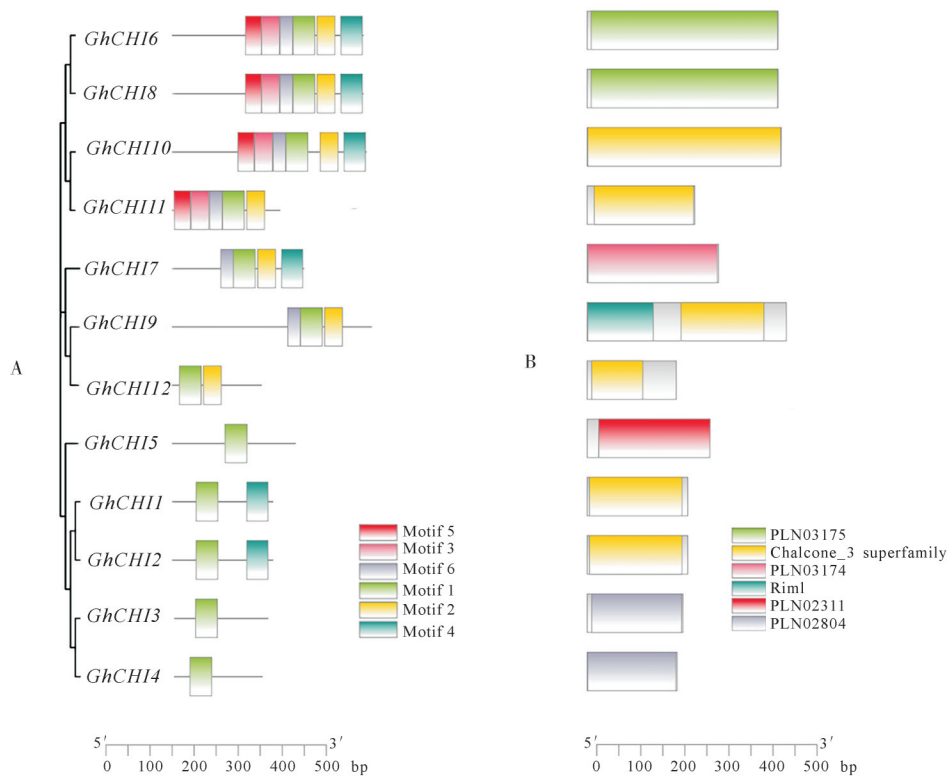


图 1 棉花 *GhCHI* 基因的基因结构(A)与保守基序(B)分析

Fig.1 Gene structure (A) and conserved motif (B) analysis of cotton *GhCHI* genes

2.4 *GhCHI* 基因参与彩色棉纤维发育的表达分析

选取彩色棉和白棉纤维不同发育时期(开花后 0、5、10、15、20 d)、不同组织、不同胁迫处理的转录组数据进行 *GhCHI* 基因的表达特征分析。结果显示, *GhCHI1*、*GhCHI2*、*GhCHI3*、*GhCHI4* 四个基因在棉纤维发育过程中具有较高的表达,在白色棉纤维快速伸长发育时期的 10 d 累积,在棕色棉和绿色棉纤维中主要在 10~20 d 高表达,表明这些基因在纤维发育的次生壁发育时期具有重要作用(图 5)。在棉花的不同组织中, *GhCHI1*、*GhCHI2*、*GhCHI3*、*GhCHI4*



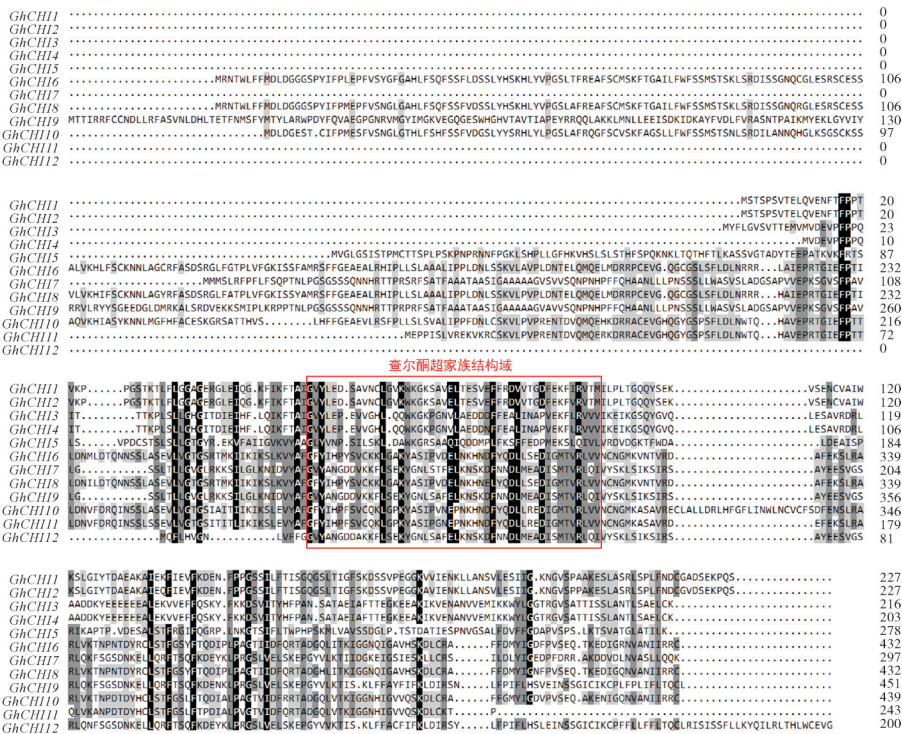


图 2 棉花 GhCHI 的多序列比对和保守结构域分析

Fig.2 Multiple sequence alignment and conserved domain analysis of GhCHI

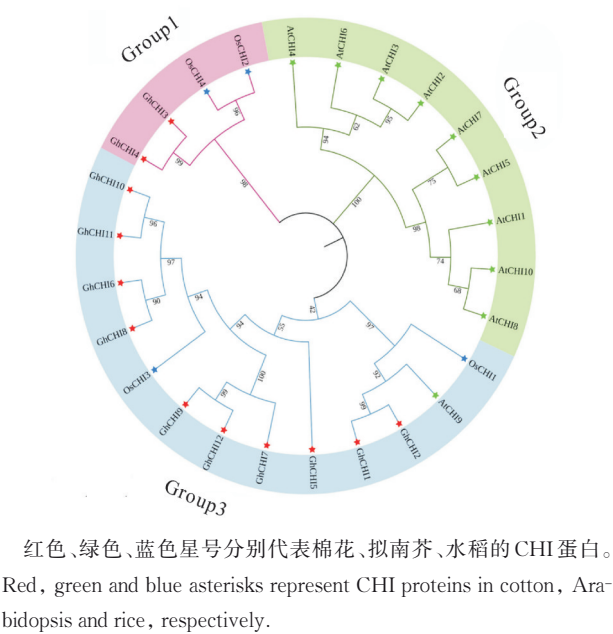


图 3 棉花与拟南芥、水稻 *CHI* 基因系统进化分析

Fig.3 Phylogenetic tree of *CHI* gene family members of cotton, rice and *Arabidopsis*

四个基因具有显著的累积表达,尤其是在叶和花器官中的花托及雌蕊中具有较高的表达水平(图 6),暗示了这些基因在器官发育中的重要作用。在冷、热、盐和干旱 4 种非生物胁迫处理下, *GhCHI1*、*GhCHI2*、

*GhCHI3*、*GhCHI4* 四个基因表现出较高的表达水平,其中 *GhCHI1*、*GhCHI2*、*GhCHI3* 基因显著受到热、盐和干旱的诱导表达(图 7),表明这些基因在非生物胁迫响应过程中具有重要作用。

*GhCHI* 基因在棕色、白色和绿色棉纤维的不同发育时期(开花后 0、5、10、15、20 d)的表达情况如图 8 所示,结果显示, *GhCHI1*—*GhCHI4* 基因在棕色、白色和绿色棉花纤维发育中均显著积累; *GhCHI5*—*GhCHI10* 基因在棉花纤维发育中表达量较低, qRT-PCR 结果与转录组表达数据具有较高的一致性,表明 *GhCHI1*—*GhCHI4* 基因在棉纤维发育中具有重要作用。

## 2.5 GhCHI 蛋白质三级结构预测及蛋白质互作分析

棉花 GhCHI 蛋白三级结构分析显示,  $\alpha$ -螺旋、无规则卷曲、延伸链以及  $\beta$ -折叠在 GhCHI 蛋白中均有分布,且以  $\alpha$ -螺旋和无规则卷曲为主,不同的蛋白含有不同数目的螺旋-转角-螺旋结构(图 9)。利用 STRING 数据库在线分析可与纤维中高表达的 GhCHI1、GhCHI2、GhCHI3、GhCHI4 相互作用的蛋白(图 10A),结果显示,能与 4 个 GhCHI 蛋白发生互作的 5 个蛋白分别属于 2-氧代戊二酸 3-双加氧酶家族

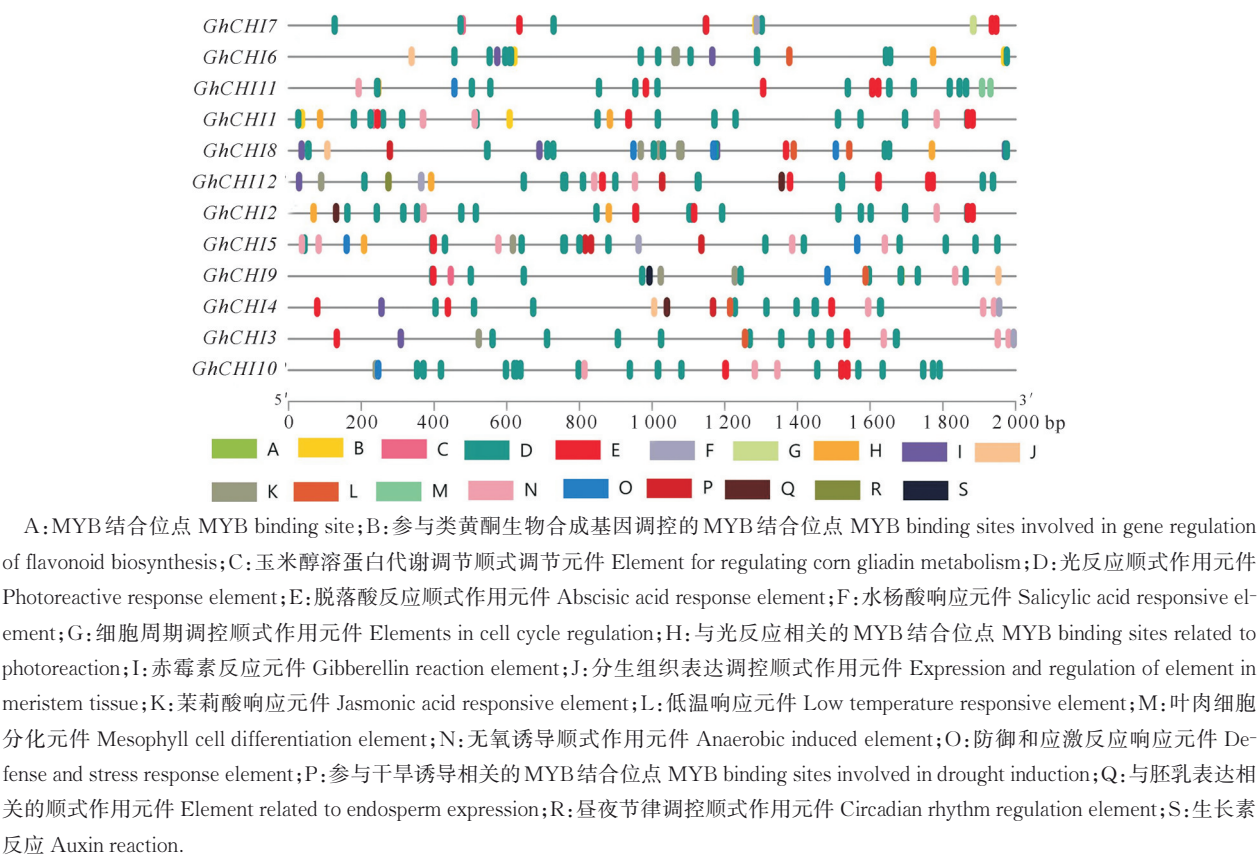


图4 *GhCHI*基因启动子顺式作用元件分析  
Fig.4 Constituent analysis of *cis*-elements of *GhCHI* promoters

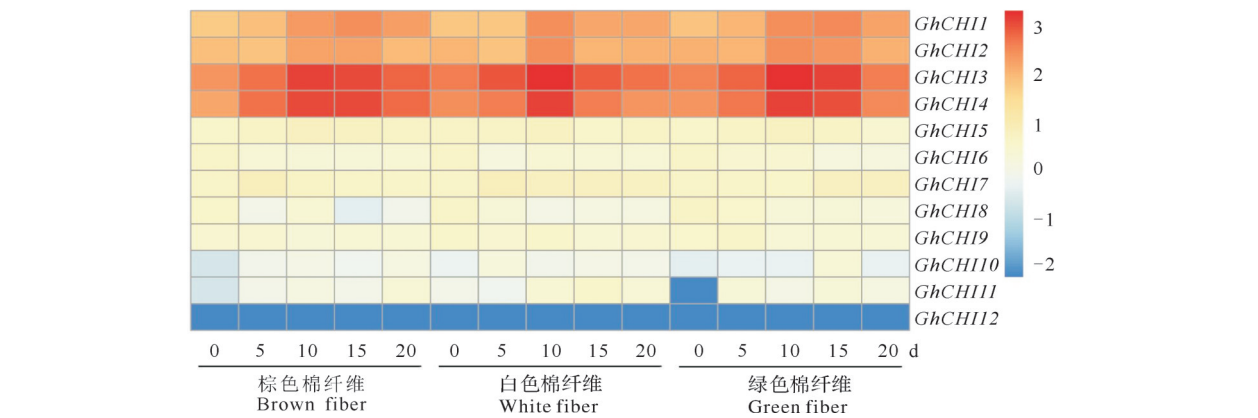


图5 *GhCHI*基因参与纤维发育的表达分析  
Fig.5 Expression analysis of *GhCHI* genes involved in cotton fiber development

和类黄酮 3'-单加氧酶家族,这 2 类酶是类黄酮合成途径中的关键酶,与类黄酮的合成密切相关。利用 qRT-PCR 分析查尔酮异构酶 *GhCHI1* (Gh\_A13G01 97)、2-氧代戊二酸 3-双加氧酶 *GhAN3* (Gh\_D12G05 66) 和类黄酮 3'-单加氧酶 *GhCYP75B2* (Gh\_D12G17 98) 在棕色棉纤维不同发育时期中的表达特征(图 10B),发现 *GhAN3* 和 *GhCYP75B2* 与 *GhCHI1* 的表

达趋势高度一致,且均在开花后 10~20 d 后的纤维中具有较高的表达水平,表明这些基因可能在棉纤维发育后期的类黄酮化合物合成中发挥重要作用。

3 讨 论

查尔酮异构酶 CHI 是植物细胞内调控类黄酮化合物合成的关键酶,在植物生长发育和响应环境胁迫过程中发挥着重要作用。目前已在红花(*Cartha-*

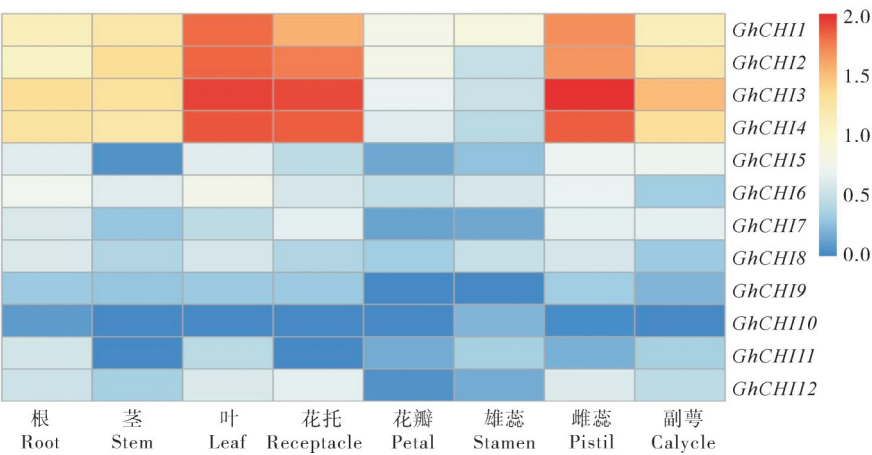


图 6 *GhCHI* 基因组织特异性表达分析

Fig.6 Expression analysis of *GhCHI* genes in different cotton tissues

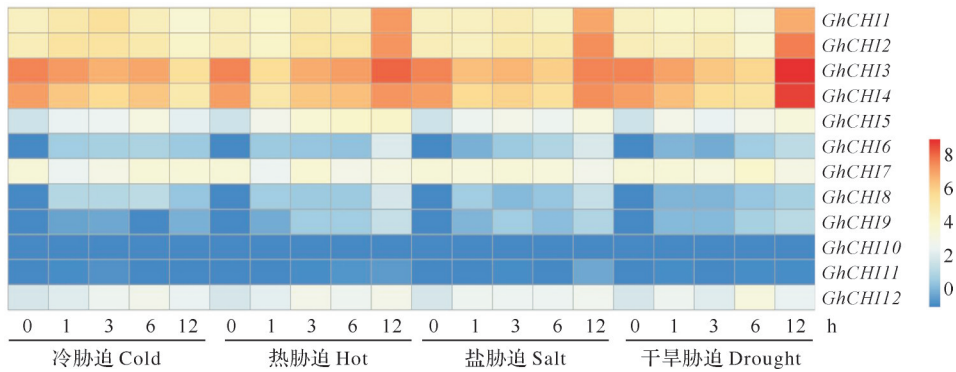


图 7 *GhCHI* 基因响应非生物胁迫的表达分析

Fig.7 Expression analysis of *GhCHI* genes in response to abiotic stress

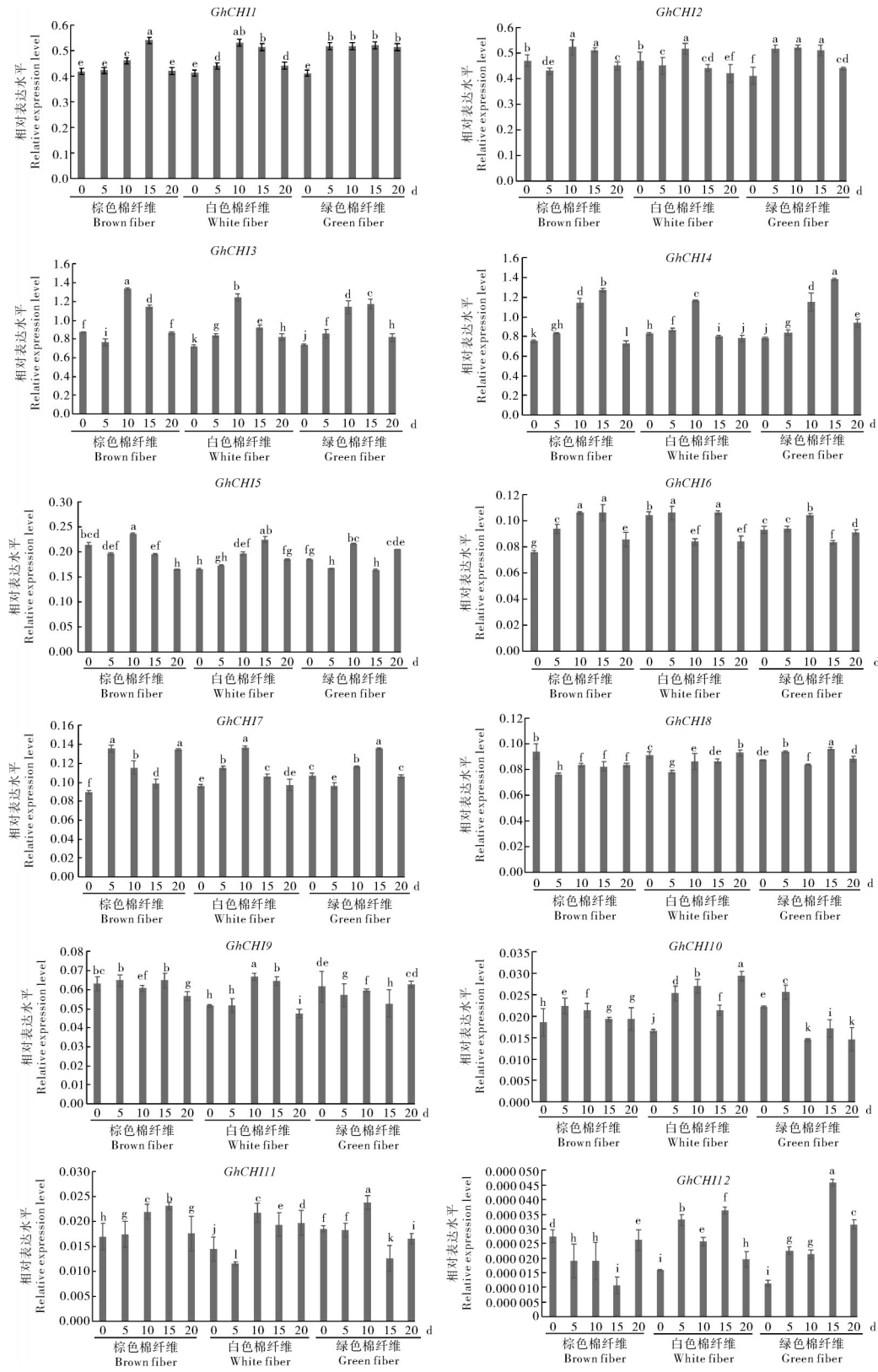
*mus tinctorius*)<sup>[16]</sup>、水稻 (*Oryza sativa*)<sup>[17]</sup>、桑树 (*Morus alba*)<sup>[18]</sup>、甘草 (*Glycyrrhiza uralensis*)<sup>[19]</sup> 等植物中系统分析了 *CHI* 基因家族, 这些研究主要集中在花色改良、药用成分提升等方面。本研究对陆地棉查尔酮异构酶 *CHI* 基因家族进行了鉴定, 获得 12 个棉花 *GhCHI* 基因 (表 2), 这些 *GhCHI* 均具有典型的查尔酮超家族保守结构域 (图 2), 表明了它们作为催化酶发挥调控类黄酮化合物合成的重要作用。

*CHI* 基因的表达与植物发育和逆境胁迫响应密切相关。如甘薯中 *CHI* 基因的表达在盐和干旱胁迫下显著累积<sup>[20]</sup>; 桃树 *CHI* 基因的表达在低温胁迫下显著上升, 促进了细胞内类黄酮的增加并抵御低温<sup>[21]</sup>。本研究中, 3 个棉花 *GhCHI* 基因 (*GhCHI1*、*GhCHI2*、*GhCHI3*) 的表达显著受到热、盐和干旱胁迫的诱导 (图 7), 表明了这 3 个基因在响应环境胁迫中的重要作用。*GhCHI* 基因启动子顺式作用元件预测发现多个与干旱胁迫响应密切相关的 MYB 结合元件和激素响应元件 (图 4), 暗示了棉花 *GhCHI* 基因响应胁迫的功能发挥可能受到激素和转录因子的

调控。

*GhCHI* 基因在天然彩色棉纤维发育后期的次生壁加厚时期具有较高的表达 (图 5、图 8), 棉纤维次生壁合成起始于 16~19 d, 在次生壁加厚期, 纤维素大量合成, 这一时期也是色素开始沉积的重要时期, 暗示了 *CHI* 及其调控的类黄酮化合物合成在彩色棉纤维着色过程中的重要作用。类黄酮 3'-单加氧酶和 2-氧代戊二酸 3-双加氧酶在类黄酮合成中发挥着关键作用<sup>[22]</sup>, *GhCHI* 可能与 2-氧代戊二酸 3-双加氧酶、类黄酮 3'-单加氧酶存在相互作用 (图 10A), 其中 *GhCHI1* 的表达趋势与类黄酮合成基因的表达趋势高度一致 (图 10B), 在茄子和桃树 2 个物种中也发现类黄酮含量随 *CHI* 基因的表达的变化而发生相同趋势的改变<sup>[23]</sup>, 表明 *CHI* 基因在调控细胞类黄酮生物合成中可能具有重要作用。类黄酮 3'-单加氧酶和 2-氧代戊二酸 3-双加氧酶可以促进花器官的色素积累<sup>[24]</sup>, *GhCHI* 在彩色棉纤维发育后期的色素沉积阶段具有显著的优势表达, 表明 *GhCHI* 基因在植物生长发育过程中尤其是色素沉积阶段发挥着重要作





不同小写字母表示组间具有显著差异。Different lowercase letters indicate significant difference between different groups ( $P < 0.05$ ).

图8 *GhCHI*基因在棉花纤维发育中的表达特征分析

Fig.8 Expression analysis of *GhCHI* genes in cotton fiber development



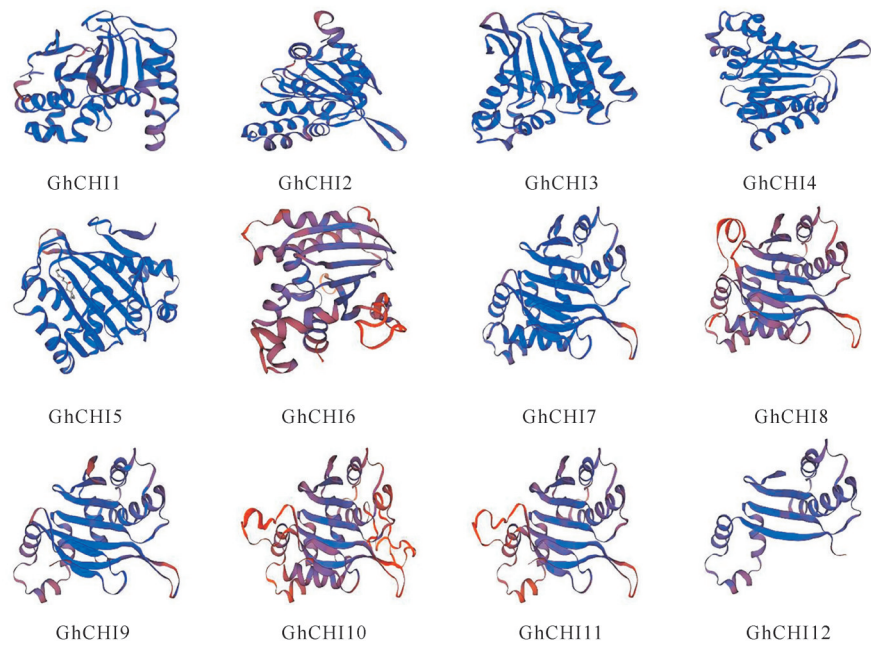
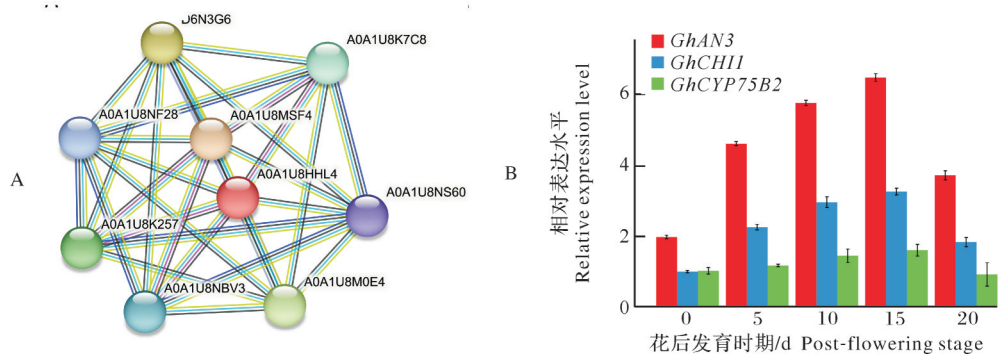


图 9 GhCHI 蛋白三级结构分析  
Fig.9 Tertiary structure analysis of GhCHI proteins



A0A1U8MSF4;GhCHI1;A0A1U8HHL4;GhCHI2;D6N3G6;GhCHI3;A0A1U8M0E4;GhCHI4;A0A1U8K257;X1 亚型 2-氧代戊二酸 3-双加氧酶 X1 subtype 2-oxoglutarate 3-dioxygenase (Gh\_D11G1775);A0A1U8K7C8;2-氧代戊二酸 3-双加氧酶 2-Oxoglutarate 3-dioxygenase (Gh\_A11G1616);A0A1U8NBV3;X1 亚型 2-氧代戊二酸 3-双加氧酶 X1 subtype 2-oxoglutarate 3-dioxygenase (Gh\_D12G0566);A0A1U8NF28;类黄酮 3'-单加氧酶 Flavonoids 3'-monooxygenase (Gh\_D12G1798);A0A1U8NS60;类黄酮 3'-单加氧酶 Flavonoids 3'-Monooxygenase (Gh\_A12G2650).

图 10 GhCHI 蛋白质相互作用分析(A)及互作蛋白的表达分析(B)  
Fig.10 GhCHI protein interaction analysis (A) and interaction protein expression analysis(B)

用。*GhCHI*影响纤维着色的功能及其受植物激素和转录因子调控的分子机制仍不清楚,*GhCHI*的遗传功能验证及调控机制解析是值得继续深入研究的重要内容,对于利用基因工程技术培育优异棉花新材料具有重要的研究意义和参考价值。

参考文献References

[1] LIU H F, LUO C, SONG W, et al. Flavonoid biosynthesis controls fiber color in naturally colored cotton[J/OL] PeerJ, 2018, 6: e4537 [2023-04-24]. <https://doi.org/10.7717/peerj.4537>.

[2] CHEN M Y, ZHANG T T, HE L, et al. Qualitative analysis of cotton fiber pigment composition[J]. Textile research journal, 2021, 91(3/4):456-463.

[3] MCKHANN H I, HIRSCH A M. Isolation of chalcone synthase and chalcone isomerase cDNAs from alfalfa (*Medicago sativa* L.); highest transcript levels occur in young roots and root tips[J]. Plant molecular biology, 1994, 24(5):767-777.

[4] GAO J F, SHEN L, YUAN J L, et al. Functional analysis of *GhCHS*, *GhANR* and *GhLAR* in colored fiber formation of

- Gossypium hirsutum* L.[J/OL].BMC plant biology, 2019, 19 (1): 455 [2023-04-24]. <https://doi.org/10.1186/s12870-019-2065-7>.
- [5] 麻新妍,南春利,薛永常.植物查尔酮异构酶的结构与功能研究概述[J].生物学教学,2022,47(1):2-4.MA X Y, NAN C L, XUE Y C. Review on the structure and function of plant chalcone isomerase[J].Biology teaching, 2022, 47(1):2-4 (in Chinese).
- [6] 薛永常,麻新妍.黄芩查尔酮异构酶基因的全长cDNA克隆及生物信息学分析[J].中国野生植物资源,2022,41(6):50-53,96.XUE Y C, MA X Y. Cloning and bioinformatic analysis of chalcone isomerase gene full-length cDNA from *Scutellaria baicalensis*[J].Chinese wild plant resources, 2022, 41(6):50-53,96(in Chinese with English abstract).
- [7] NISHIHARA M, NAKATSUKA T, YAMAMURA S. Flavonoid components and flower color change in transgenic tobacco plants by suppression of chalcone isomerase gene[J]. FEBS Letters, 2005, 579(27):6074-6078.
- [8] 李国庆.黄酮醇在小麦抗旱节水中的功能分析[D].济南:山东大学, 2021. LI G Q. Function analysis of flavonols in drought resistance and water saving of wheat[D]. Ji'nan: Shandong University, 2021 (in Chinese with English abstract).
- [9] 宁亚茹,韩民利,张晓东,等.盐胁迫对黄蜀葵不同部位总黄酮含量及抗氧化活性的影响[J].江苏农业科学,2021,49(19):196-200.NING Y R, HAN M L, ZHANG X D, et al. Effects of salt stress on total flavonoids content and antioxidant activity in different parts of *Abelmoschus manihot*[J]. Jiangsu agricultural sciences, 2021, 49(19):196-200 (in Chinese).
- [10] 宋静武,殷德松,赵弟广,等.核桃叶片内多酚黄酮类成分对低温胁迫的响应[J].河北林果研究,2017,32(1):34-41. SONG J W, YIN D S, ZHAO D G, et al. The response of polyphenol flavonoids in *Juglans regia* leaves to low temperature stress[J]. Hebei journal of forestry and orchard research, 2017, 32(1):34-41 (in Chinese with English abstract).
- [11] 李军,李洪清,李美茹.大花美人蕉查尔酮异构酶基因的cDNA克隆和序列分析[J].植物生理学通讯,2006,42(3):449-453. LI J, LI H T, LI M R. Cloning and sequence analysis of cDNA of chalcone isomerase gene from *Canna generalis* Bailey[J]. Plant physiology communications, 2006, 42(3):449-453 (in Chinese with English abstract).
- [12] 段怡红.陆地棉开花期相关基因 *GhMYB44* 的功能研究[D].保定:河北农业大学,2020. DUAN Y H. Functional identification of flowering-stage related gene *GhMYB44* from *Gossypium hirsutum* [D]. Baoding: Hebei Agricultural University, 2020 (in Chinese with English abstract).
- [13] ZHU J H, ZHAO W, LI R S, et al. Identification and characterization of chalcone isomerase genes involved in flavonoid production in *Dracaena cambodiana*[J/OL].Frontiers in plant science, 2021, 12: 616396 [2023-04-24]. <https://doi.org/10.3389/fpls.2021.616396>.
- [14] 丁宁,海燕,王晓晖,等.白木香查尔酮异构酶基因的克隆鉴定与表达分析[J].药学报,2021,56(2):630-638.DING N, HAI Y, WANG X H, et al. Cloning and expression analysis of chalcone isomerase from *Aquilaria sinensis* [J].Acta pharmaceutica sinica, 2021, 56(2):630-638 (in Chinese with English abstract).
- [15] 宋成攀,夏松波,王孝刚,等.嫁接陆地棉查尔酮合成酶与查尔酮异构基因的克隆及表达分析[J].湖北农业科学,2017,56(23):4616-4623.SONG C P, XIA S B, WANG X G, et al. Cloning and expression analysis of chalcone synthase and chalcone isomerase genes in grafted upland cotton [J].Hubei agricultural science, 2017, 56(23):4616-4623 (in Chinese with English abstract).
- [16] 任超翔,唐小慧,何雯,等.红花查尔酮异构酶基因的克隆及表达分析[J].天然产物研究与开发,2018,30(9):1521-1525,1574.REN C X, TANG X H, HE W, et al. Cloning and expression analysis of chalcone isomerase gene in safflower[J]. Natural product research and development, 2018, 30(9):1521-1525,1574 (in Chinese with English abstract).
- [17] PARK S I, PARK H L, BHOO S H, et al. Biochemical and molecular characterization of the rice chalcone isomerase family [J/OL]. Plants, 2021, 10(10):2064 [2023-04-24]. <https://doi.org/10.3390/plants10102064>.
- [18] CHAO N, WANG R F, HOU C, et al. Functional characterization of two chalcone isomerase (CHI) revealing their responsibility for anthocyanins accumulation in mulberry [J]. Plant physiology and biochemistry, 2021, 161:65-73.
- [19] 张晓冬,尹彦超,周姗,等.光果甘草查尔酮异构酶基因的克隆及序列分析[J].生物技术通讯,2017,28(4):478-484. ZHANG X D, YIN Y C, ZHOU S, et al. Cloning and sequence analysis of chalcone isomerase gene from *Glycyrrhiza glabra* L. [J].Letters in biotechnology, 2017, 28(4):478-484 (in Chinese with English abstract).
- [20] 史敏莉,唐锐敏,张毅,等.甘薯 lncRNA TCONS\_00074371 及其靶基因 CHI 的克隆及生物信息学分析[J].山西农业科学,2022,50(12):1599-1607. SHI M L, TANG R M, ZHANG Y, et al. Cloning and bioinformatics analysis of sweet potato lncRNA TCONS\_00074371 and its target gene CHI [J].Journal of Shanxi agricultural science, 2022, 50(12):1599-1607 (in Chinese with English abstract).
- [21] 周平,林志楷,郭瑞,等.低温处理对桃树叶基因表达及类黄酮合成代谢的影响[J].农业生物技术学报,2021,29(7):1283-1294.ZHOU P, LIN Z K, GUO R, et al. Effects of low temperature treatment on gene expression and flavonoids biosynthesis metabolism in peach (*Prunus persica*) leaves [J]. Journal of agricultural biotechnology, 2021, 29(7):1283-1294 (in Chinese with English abstract).
- [22] WANG Y, SHI Y, LI K, et al. Roles of the 2-oxoglutarate-dependent dioxygenase superfamily in the flavonoid pathway: a review of the functional diversity of F3H, FNS I, FLS, and

- LDOX/ANS[J/OL].Molecules, 2021, 26(21): 6745 [2023-04-24].<https://doi.org/10.3390/molecules26216745>.
- [23] LV L L, FENG X F, LI W, et al. High temperature reduces peel color in eggplant (*Solanum melongena*) as revealed by RNA-seq analysis[J].Genome, 2019, 62(7): 503-512.
- [24] ZHAO D Q, TAO J, HAN C X, et al. Flower color diversity revealed by differential expression of flavonoid biosynthetic genes and flavonoid accumulation in herbaceous peony (*Paeonia lactiflora* Pall.) [J]. Molecular biology reports, 2012, 39(12): 11263-11275.

## Identification and expression analysis of chalcone isomerase gene family in cotton

YANG Dan<sup>1</sup>, BAO Junxiu<sup>1</sup>, CHEN Pei<sup>1</sup>, WANG Fei<sup>1,2</sup>,  
TANG Shouwu<sup>3</sup>, LIU Haifeng<sup>3</sup>, LI Hongbin<sup>1,2,3</sup>

1.College of Life Sciences, Shihezi University, Shihezi 832003, China;

2.Key Laboratory of Oasis Town and Mountain Basin System Ecology,

Xinjiang Production and Construction Group /Key Laboratory of Xinjiang Phytomedicine Resource and Utilization, Ministry of Education, Shihezi 832000, China;

3.China Colored-Cotton (Group) Co., Ltd., Urumqi 830001, China

**Abstract** To study the function of chalcone isomerase (CHI) gene in fiber development and response to stress in colored cotton, the physicochemical properties, conserved domain, *cis*-acting elements in the promoter of CHI gene, advanced structures and phylogeny of proteins were analyzed, transcriptome data and quantitative real-time PCR (qRT-PCR) were used to analyze its expression profile. The results showed that 12 members of *GhCHI* gene family were identified from the genome of upland cotton, which can be divided into two subfamilies with typical chalcone superfamily domains, encoding 201-452 amino acids, mainly hydrophilic proteins, and secondary structures mainly composed of  $\alpha$ -helix and irregular curls. The types of *cis*-acting elements in the promoter of *GhCHI* gene mainly include light response elements, hormone response elements, and elements involved in the regulation of flavonoid biosynthesis. The results of analyzing expression profile showed that *GhCHI1*, *GhCHI2*, *GhCHI3*, and *GhCHI4* were closely related to fiber development, especially with high level of expression during the period of pigmentation and coloring at the later stage of fiber development in colored cotton. These four genes also had high level of expression in leaves, receptacles, and pistils. *GhCHI1*, *GhCHI2*, and *GhCHI3* were significantly induced to express under heat stress, salt stress, and drought stress. The results of qRT-PCR showed that *GhCHI1-GhCHI4* were significantly enriched in fiber growth, especially in colored cotton. The results of analyzing the interaction of *GhCHI1-GhCHI4* protein showed that they may interact with proteins such as 2-oxoglutarate 3-dioxygenase and flavonoid 3'-monooxygenase involved in the flavonoid synthesis. It is indicated that *GhCHI1*, *GhCHI2*, *GhCHI3*, and *GhCHI4* genes play important roles in fiber development and stress response in colored cotton.

**Keywords** natural colored cotton; chalcone isomerase; gene family; fiber development; pigmentation

(责任编辑:葛晓霞)