

# 陆地棉 GST 基因家族全基因组分析

许磊, 陈文, 司国阳, 黄艺园, 林毅, 蔡永萍, 高俊山

安徽农业大学生命科学学院, 合肥 230036

**摘要:** 谷胱甘肽转移酶(glutathione-S-transferase, GST)是一种普遍存在的具有多功能的超家族蛋白, 在植物初次生代谢、逆境胁迫、胞间信号传递等方面具有重要作用; 同时, 作为配体其在植物激素代谢以及物质转运方面也发挥作用。为了解析陆地棉(*Gossypium hirsutum* L.) GST 基因家族的信息, 本研究对该基因家族成员的种类、进化关系、物理定位、基因结构和保守基序以及表达模式进行了分析。结果显示, 在陆地棉全基因组中共含有 70 个 GST 基因, 进化树和基因结构分析将该家族分为 U 族、F 族、T 族、Z 族、EF1B $\gamma$  族和 TCHQD 族。基因定位分析发现, 除了 AD/At2、AD/At4、AD/At5、AD/Dt5、AD/Dt10 号染色体上没有 GST 基因外, 其他染色体上都有 GST 基因, 并且在 AD/At9、AD/Dt7、AD/Dt12、AD/Dt13 这 4 条染色体上出现基因簇。对 F 族(Phi 类) 9 个 GST 基因进行荧光定量分析, 结果表明, 除 *GhGSTF1* 可能为假基因外, *GhGSTF2~9* 等 8 个基因在陆地棉根、茎、叶以及各个发育时期的纤维中均有表达; 结合生物信息学分析, 推测 *GhGSTF8* 可能参与原花青素/花青素的转运和积累; *GhGSTF4*、*GhGSTF6* 和 *GhGSTF9* 可能在调节陆地棉的生长和胁迫反应中起作用, 而 *GhGSTF2*、*GhGSTF3*、*GhGSTF5* 和 *GhGSTF7* 的功能还有待进一步研究。本研究为陆地棉 GST 基因家族的分子进化及功能研究提供了理论依据。

**关键词:** 陆地棉; 谷胱甘肽转移酶; 全基因组分析; 进化

## Genome-wide analysis of the GST gene family in *Gossypium hirsutum* L.

Lei Xu, Wen Chen, Guoyang Si, Yiyuan Huang, Yi Lin, Yongping Cai, Junshan Gao

College of life science, Anhui Agricultural University, Hefei 230036, China

**Abstract:** Glutathione-S-transferase (GST) is a ubiquitous multi-functional protein superfamily that plays important roles in plant primary and secondary metabolism, stress and intercellular signal transduction. Concomitantly, it also functions as a ligand in the metabolism of plant hormones and substance transport. In order to understand the GST gene family in upland cotton (*Gossypium hirsutum* L.), herein we analyzed the species, evolutionary relationship, physical location, gene structure, conserved motifs and expression patterns. We identified 70 GST genes in the whole genome of upland cotton, and divided them into U, F, T, Z, EF1B $\gamma$  and TCHQD groups by phylogenetic tree and gene

收稿日期: 2017-01-03; 修回日期: 2017-05-10

基金项目: 国家自然科学基金项目(编号: 31672497)和安徽省“115”产业创新团队项目(皖人才办[2011]2 号)资助[Supported by the National Natural Science Foundation of China (No.31672497) and Anhui Province “115” Industrial Innovation team(No.[2011]2)]

作者简介: 许磊, 硕士研究生, 专业方向: 细胞生物学。E-mail: 437185342@qq.com

通讯作者: 高俊山, 教授, 硕士生导师, 研究方向: 植物分子遗传学。E-mail: gaojsh@ahau.edu.cn

DOI: 10.16288/j.ycz.16-435

网络出版时间: 2017/6/13 15:15:37

URI: <http://kns.cnki.net/kcms/detail/11.1913.R.20170613.1515.002.html>

structure analyses. The gene mapping analysis indicated that the *GST* genes were on every chromosome except chromosome AD/At2, AD/At4, AD/At5, AD/Dt5 and AD/Dt10. Moreover, the *GST* gene cluster appeared on four chromosomes (AD/At9, AD/Dt7, AD/Dt12 and AD/Dt13). qRT-PCR assays showed that eight genes (*GhGSTF2-9*) were expressed in the root, stem, leave and fiber of different developmental stages while *GhGSTF1* might be a pseudogene. Combining qRT-PCR and bioinformatic analysis, we speculated that *GhGSTF8* might be involved in the transport and accumulation of proanthocyanidins/anthocyanins; *GhGSTF4*, 6 and 9 might play roles in regulating the growth and stress response of upland cotton; the function of *GhGSTF2*, 3, 5 and 7 remains to be further investigated. Our work provides a theoretical basis for further studies on the molecular evolution and function of the *GST* gene family in upland cotton.

**Keywords:** *Gossypium hirsutum* L.; glutathione-S-transferase (GST); genome-wide analysis; evolution

谷胱甘肽转移酶又称谷胱甘肽-S-转移酶(glutathione S-transferases, GST), 普遍存在于动物、植物和微生物中, 是一类由多个基因编码、具有多种功能的超家族酶, 通过谷胱甘肽与有害异源物质或氧化产物结合, 从而促进此类物质的代谢、区域化隔离或消除<sup>[1]</sup>。根据植物蛋白质的同源性和基因结构特征, GST 家族分为 F(Phi)、U(Tau)、T(Theta)、Z(Zeta)、L(Lambda)、DHAR、EF1B $\gamma$  和 TCHQD 8 个亚家族<sup>[2]</sup>。而 F 族和 U 族是植物所特有的, 与其他亚家族相比, 其成员最多, 含量也最为丰富。可溶性的 GST 主要分布于细胞质中, 少数分布在叶绿体、微体中, 也有少量存在于细胞核和质外体中<sup>[3]</sup>。植物 GST 最早在玉米(*Zea mays* L.)中被发现<sup>[4]</sup>, 随后在拟南芥(*Arabidopsis thaliana*)、大豆(*Glycine max*)、水稻(*Oryza sativa* L.)、烟草(*Nicotiana tabacum* L.)等植物中相继被发现<sup>[5-7]</sup>。

结构研究表明, GST 是一种球状二聚体蛋白, 一般通过 25~27 kDa 的 2 条亚基以同源或异源的方式聚合而成, 每个亚基又含有 2 个空间结构不同的基本结构域, 即位于 N 端的 GSH(glutathione)结合位点(G-site)和位于 C 端的连接 H 离子的作用位点(H-site)<sup>[8]</sup>。GST 对底物的特异选择性主要取决于这两个功能结构域, G 位点决定 GST 的种类, H 位点主要影响 GST 与底物的亲和力, G、H 结构域的不同决定了 GST 结构和功能的不同。

目前研究表明, GST 主要具有以下几个生物学功能: (1)解毒功能。GST 在植物体内能够催化诱导某些有害物质(内源或外来)与谷胱甘肽结合, 从而降低或消除有害物质的毒性<sup>[9]</sup>; (2)跨膜运输定位。在玉米中, 敲除 *BZ2* 基因, 可以导致花色素的转运受

阻, 从而只在细胞质中积累<sup>[10]</sup>。Pérez-Díaz 等<sup>[11]</sup>研究发现, 葡萄(*Vitis vinifera* L.)中 *VvGST3*、*VvGST4* 具有转运原花青素和花青素的功能。黄燕梅<sup>[12]</sup>对柿(*Diospyros Kaki* L.)原花青素前体跨膜及转运相关基因的研究表明, *DkGSTL* 能够转运原花青素前体; (3)作为结合蛋白或配体起作用。1994 年, Zettl 等<sup>[13]</sup>利用光亲和标记的方法, 证实 GST 是生长素和细胞分裂素的结合蛋白; (4)提高植物的抗胁迫能力。GST 利用 GSH 作为电子供体亲核攻击核酸和脂肪酸中的有机过氧化物, 使过氧化物还原生成毒性小的一元醇化合物, GSH 转化为次酸形式, 与另一个 GSH 分子相结合形成 GS-SG 二硫化物, 使植物耐受各种逆境的胁迫<sup>[14,15]</sup>。Chan 等<sup>[16]</sup>将大豆 *GmGSTL1* 转入拟南芥中, 发现转 *GmGSTL1* 拟南芥能够缓解盐胁迫症状; (5)参与植物的生长发育。GST 与植物的生长发育密切相关, 细胞分裂素、生长素等激素调控 Tau 类 GST 基因的表达, 进而在植物的生长和发育过程中发挥至关重要的作用<sup>[17]</sup>; (6)参与细胞信号传导。GST 作为胁迫信号蛋白参与信号传导, 对细胞进行调控<sup>[18]</sup>。

棉花(*Gossypium spp.*)是我国最重要的纤维作物, 品种繁多, 其中陆地棉(*G.hirsutum*)栽培最为广泛。目前陆地棉的全基因组测序已经完成<sup>[19]</sup>, 通过全基因组数据库能够对基因进行系统的筛选、鉴定和比较基因组学的研究。前人研究发现, *GST* 基因在植物的初级代谢、次级代谢、跨膜转运定位和胞间信号转导等过程中发挥重要作用, 其中 F 族(Phi 类) *GST* 基因参与原花青素、花青素等黄酮类物质的转运<sup>[20]</sup>, 但在陆地棉中还未见详细的报道。本研究通过生物信息学方法对陆地棉 *GST* 基因进行筛选、鉴定和分

析,同时对 Phi 类 GST 基因进行荧光定量分析,初步分析陆地棉中 Phi 类 GST 基因的功能,确定与陆地棉纤维色素转运相关的 GST 基因,为后续进一步解析陆地棉纤维色素合成机理提供理论基础。

## 1 材料和方法

### 1.1 实验材料及基因组数据库序列来源

棕色棉棕彩选 1 号是由本实验室选育的一种陆地棉自交系,种植于合肥市大杨镇安徽农业大学高新技术农业园试验基地,按照正常田间管理。陆地棉全基因组数据库来源于网站 <http://cgp.genomics.org.cn/>;拟南芥数据见参考文献[21]。

### 1.2 方法

#### 1.2.1 陆地棉 GST 基因鉴定

利用 DNATOOLS 软件建立一个以陆地棉全基因组氨基酸序列为数据源的本地数据库。以拟南芥 GST 结构域的氨基酸序列作为查询(query),与已建立的陆地棉氨基酸序列本地数据库进行 TblastN ( $E\text{-value}=0.001$ )序列比对,初步筛选得到 GST 家族基因序列。将筛选出的陆地棉候选 GST 核苷酸序列分别通过 Pfam PFAM(<http://pfam.wustl.edu/hmmsearch.shtml>) ( $E\text{-value}=0.001$ )<sup>[22]</sup>以及 SMART(<http://smart.embl-heidelberg.de>)<sup>[23]</sup>检验其是否含有 GST 结构域。利用 MEGA5.0 软件中 ClustalW 工具<sup>[24]</sup>对得到的 GST 基因序列进行多重序列比对,除去重复序列。

#### 1.2.2 陆地棉 GST 基因家族进化树的构建、命名及分类

通过 MEGA5.0 软件比对陆地棉 GST 氨基酸序列,使用邻接法(Neighbor-Joining method) (Bootstrap=1000)构建系统进化树,分析陆地棉 GST 基因家族的进化关系。利用最大似然法(Maximum Likelihood)(Bootstrap=1000)构建陆地棉和拟南芥 GST 基因家族进化树。以 Soranzo 等<sup>[25]</sup>提出的命名标准对陆地棉 GST 基因进行命名和分类。

#### 1.2.3 陆地棉 GST 基因家族氨基酸序列分析

为了分析陆地棉 GST 基因家族氨基酸序列的基本属性和特征,利用 ExPASy(<http://www.expasy.org>)

在线工具计算等电点、蛋白分子量等属性。利用 GS-DS(<http://gsds.cbi.pku.edu.cn>)<sup>[26]</sup>在线工具分析该基因家族的基本结构。

#### 1.2.4 陆地棉 GST 基因家族保守基序分析

利用 MEME(Multiple Expectation Maximization for Motif Elicitation)在线分析工具([http://meme.sdsc.edu/meme4\\_3\\_0/intro.html](http://meme.sdsc.edu/meme4_3_0/intro.html))<sup>[27]</sup>对陆地棉 GST 基因家族蛋白保守基序进行分析,基序的最大数值设为 20,基序的长度设置介于 6~200 之间。使用 Pfam 和 SMART 对保守基序进行功能注释。

#### 1.2.5 陆地棉 GST 基因家族染色体定位

从陆地棉基因组数据库中获得每个 GST 基因在染色体上的起始位置等相关信息,利用 MapInspect (<http://mapinspect.software.informer.com>)软件将这些基因在染色体上的物理位置绘制成图。通过 Circos 软件(<http://mkweb.bcgsc.ca/tableviewer/>)对陆地棉、雷蒙德氏棉(*G. raimondii*)及亚洲棉(*G. arboreum*) GST 基因进行共线性分析。

#### 1.2.6 陆地棉 Phi 类 GST 基因表达分析

根据筛选出的陆地棉 Phi 类 GST 基因序列设计荧光定量 PCR 引物(表 1)。利用 RNA 提取试剂盒(北京全式金生物公司)提取棕色棉棕彩选 1 号各组织及不同发育时期纤维的 RNA。利用荧光反转录试剂盒(北京全式金生物公司)进行 qRT-PCR 分析。qRT-PCR 体积为 20  $\mu\text{L}$ ,包括 10  $\mu\text{L}$  SYBR®Premix Ex Taq™ (2 $\times$ ), 2  $\mu\text{L}$  cDNA, 0.8  $\mu\text{L}$  上游和下游引物。反应程序为:50 2 min ;95 30 s ,95 5 s ,60 20 s , 40 个循环;最后再 72 10 min。以 UBQ7 基因作为内参,每个样品进行 3 次重复。相对表达水平应用  $2^{-\Delta\Delta\text{Ct}}$  法进行计算。热图利用 heat map 软件([https://en.wikipedia.org/wiki/Heat\\_map](https://en.wikipedia.org/wiki/Heat_map))绘制。

## 2 结果与分析

### 2.1 陆地棉 GST 基因家族鉴定

通过查询比对,利用 pfam 蛋白数据库得到陆地棉 GST 基因的结构域氨基酸序列。利用“DNATOOLS”软件,以结构域氨基酸序列为标准序列,对陆地棉全基因组进行 TblastN( $E\text{-value}=0.001$ )序列比

表 1 荧光定量 PCR 所用的引物序列

Table 1 Primer sequences used for qRT-PCR

基因名称	引物序列(5'→3')
<i>GhGSTF1</i>	F: TTGTAGGTATGTATGTGAA R: TCCAACCATGTGATCTATT
<i>GhGSTF2</i>	F: TTGTGCTAACCTGATAAATGAG R: AAGCCACCTTCTGGAAAG
<i>GhGSTF3</i>	F: TTCCTGTCACTCAAGATG R: CCTCAACTTCTAGCCAAT
<i>GhGSTF4</i>	F: CATGAATACAGCGACAAG R: CTTCCACTTCCTTCCATA
<i>GhGSTF5</i>	F: GGCATTAGAAGATGGACAT R: GGTTGAATTGGTGAGACT
<i>GhGSTF6</i>	F: ATGAATACGCCGACAAAG R: TCAACTTCCTTCCACAATT
<i>GhGSTF7</i>	F: TTCCTGTCACTCAAGATG R: CCTCAACTTCTAGCCAAT
<i>GhGSTF8</i>	F: ATAGAGGATGGCGACTTC R: TGCTCGTCTTCCAATGA
<i>GhGSTF9</i>	F: TTCATGTCAATCCTATTCTT R: CATCCATTCAATTACTTCTT

对, 初步筛选出 78 个候选基因。通过 MEGA5.0 软件提供的 ClustalW 工具进行多重序列比对, 以及 pfam 和 SMART 程序进行结构域分析, 删除重复序列和没有 GST 保守结构域的序列, 最终确定了 70 个陆地棉 GST 基因。参考 Soranzo 等<sup>[25]</sup>提出的命名标准及命名方式, 将本研究中的 GST 命名为 *GhGST*, 并根据陆地棉与拟南芥 GST 蛋白构建的符合进化树进行亚类命名。同时, 通过 ExPASy 在线工具对陆地棉 GST 的特征进行分析。结果显示, GST 编码 180~421 个氨基酸, 分别位于陆地棉 A、D 组的不同染色体上(表 2)。另外, 亚细胞定位预测显示, GST 主要位于细胞质中, 还有少数位于叶绿体和线粒体(表 2)。

## 2.2 陆地棉 GST 基因家族的进化分析

通过 MEGA5.0 软件对 70 个陆地棉和 57 个拟南芥 GST 蛋白序列构建了一个复合进化树, 并通过邻接法对陆地棉 GST 氨基酸序列进行了系统进化分析。结果显示陆地棉 GST 可聚为 6 个亚家族, 其成员中有 50 个属于 U 族、9 个属于 F 族、4 个属于 T 族、3 个属于 Z 族、2 个属于 EF1B $\gamma$  族、2 个属于 TCHQD

族, 但未发现 DHAR 族和 Lambda 族的 GST 成员(图 1, 图 2A)。每个亚家族中的成员个数与拟南芥中各亚家族成员个数具有相同的趋势, 即 U 族成员最多, F 族与 T 族次之, EF1B $\gamma$  族和 TCHQD 族最少(图 1)。而且陆地棉与拟南芥 GST 能够紧密聚类在一起, 未发现亚家族隔断现象, 说明陆地棉与拟南芥 GST 亲缘关系较近。同时, 利用 GSDS 在线工具对陆地棉 GST 家族成员的蛋白序列结构进行分析, 得到 1 个“外显子—内含子”结构图。其中, 陆地棉 70 个 GST 基因的内含子数量都在 10 个以内, 最多的有 9 个, 最少的只有 1 个(图 2B)。而且同一亚家族的各成员具有相同或者相近的基因结构, 如 Tau 类 GST 基因大部分含有 1 个内含子; Zeta 类 GST 基因含有 9 个内含子, 但 *GhGSTZ1* 含有 8 个内含子; Phi 类 GST 基因含有两个内含子; Theta 类 GST 基因含有 6 个内含子, 但 *GhGSTT2* 含有 7 个内含子(图 2B)。

为了比较陆地棉与其两个祖先亚洲棉和雷蒙德氏棉之间的进化关系, 根据 Dong 等<sup>[28]</sup>筛选的亚洲棉和雷蒙德氏棉 GST 基因家族, 对陆地棉、亚洲棉和雷蒙德氏棉 GST 构建了复合进化树。结果发现陆地棉、亚洲棉和雷蒙德氏棉 GST 基因家族也聚为 U 族、F 族、T 族、Z 族、EF1B $\gamma$  族和 TCHQD 族 6 个亚家族, 各亚家族成员能够紧密聚在一个分支上, 且未发现亚家族间互相隔断现象(图 3)。有趣的是, 亚洲棉和雷蒙德氏棉 GST 基因家族含有 DHAR 族和 Lambda 族两个亚家族, 而陆地棉中却没有, 推测可能是由于陆地棉在长期进化过程中逐渐丢失了 DHAR 族和 Lambda 族, 而其所发挥的功能可能被其他家族的 GST 所替代。

## 2.3 陆地棉 GST 基因保守基序分析

陆地棉 GST 基因编码的氨基酸长度在 180~421 之间, 应用 MEME 在线软件对陆地棉 GST 蛋白保守基序进行分析, 共鉴定出 20 个保守基序(图 4)。通过 Pfam(*E*-value=0.001)以及 SMART 在线程序对 20 个保守基序进行功能分析, 结果发现 20 个保守基序中有 3 个保守基序为 GST-N 结构域, 且所有的陆地棉中都含有 GST-N 结构域(图 5)。这说明 GST-N 结构域是决定 GST 蛋白功能的主要结构域, 也暗示了 GST 基因在进化过程中的功能保守性。

表 2 陆地棉 70 个 GST 基因特征

Table 2 Characteristics of 70 GST genes in *G. hirsutum*

基因名称	编号	染色体	氨基酸长度(aa)	分子量(kDa)	等电点	亚细胞定位预测
<i>GhGSTU1</i>	CotAD_30124	AD/At1	220	25.3	5.69	Cyto
<i>GhGSTU2</i>	CotAD_46104	AD/At3	346	39.9	5.33	Chlo
<i>GhGSTU3</i>	CotAD_68466	AD/At6	220	25.9	6.46	Cyto
<i>GhGSTU4</i>	CotAD_18896	AD/At6	217	24.7	5.15	Cyto
<i>GhGSTU5</i>	CotAD_18897	AD/At6	225	25.8	5.75	Cyto
<i>GhGSTU6</i>	CotAD_18898	AD/At6	216	24.6	5.89	Cyto
<i>GhGSTU7</i>	CotAD_30502	AD/At9	223	25.7	5.30	Cyto
<i>GhGSTU8</i>	CotAD_30505	AD/At9	223	25.8	5.32	Cyto
<i>GhGSTU9</i>	CotAD_30504	AD/At9	223	25.9	5.89	Chlo
<i>GhGSTU10</i>	CotAD_30506	AD/At9	223	25.8	6.23	Cyto
<i>GhGSTU11</i>	CotAD_30488	AD/At9	224	26.2	5.91	Cyto
<i>GhGSTU12</i>	CotAD_30509	AD/At9	209	24.3	6.14	Cyto
<i>GhGSTU13</i>	CotAD_30501	AD/At9	195	22.9	6.72	Cyto
<i>GhGSTU14</i>	CotAD_66588	AD/At9	291	33.1	5.27	Cyto
<i>GhGSTU15</i>	CotAD_59520	AD/At10	220	25.0	5.66	Chlo
<i>GhGSTU16</i>	CotAD_49195	AD/At12	232	26.1	6.01	Cyto
<i>GhGSTU17</i>	CotAD_59017	AD/At13	224	25.5	5.27	Cyto
<i>GhGSTU18</i>	CotAD_41056	AD/Dt1	220	25.1	5.43	Cyto
<i>GhGSTU19</i>	CotAD_40770	AD/Dt2	227	26.1	6.18	Cyto
<i>GhGSTU20</i>	CotAD_76502	AD/Dt3	215	25.0	5.44	Cyto
<i>GhGSTU21</i>	CotAD_48229	AD/Dt6	220	25.7	6.45	Cyto
<i>GhGSTU22</i>	CotAD_03316	AD/Dt6	225	25.8	5.75	Cyto
<i>GhGSTU23</i>	CotAD_03315	AD/Dt6	191	21.7	5.13	Cyto
<i>GhGSTU24</i>	CotAD_03318	AD/Dt6	237	27.1	8.10	Cyto
<i>GhGSTU25</i>	CotAD_09371	AD/Dt7	219	25.0	6.53	Cyto
<i>GhGSTU26</i>	CotAD_09373	AD/Dt7	198	23.2	5.85	Cyto
<i>GhGSTU27</i>	CotAD_09374	AD/Dt7	219	25.4	5.57	Cyto
<i>GhGSTU28</i>	CotAD_09394	AD/Dt7	219	25.4	6.53	Cyto
<i>GhGSTU29</i>	CotAD_64376	AD/Dt9	221	25.7	5.54	Cyto
<i>GhGSTU30</i>	CotAD_21534	AD/Dt9	220	25.7	5.42	Cyto
<i>GhGSTU31</i>	CotAD_38703	AD/Dt12	220	25.4	5.69	Cyto
<i>GhGSTU32</i>	CotAD_43155	AD/Dt12	236	26.1	6.01	Chlo
<i>GhGSTU33</i>	CotAD_74979	AD/Dt12	234	25.8	5.87	Chlo
<i>GhGSTU34</i>	CotAD_74978	AD/Dt12	234	25.9	5.25	Cyto
<i>GhGSTU35</i>	CotAD_16661	AD/Dt13	301	34.4	4.98	Cyto
<i>GhGSTU36</i>	CotAD_13954	AD/Dt13	204	23.7	7.60	Cyto
<i>GhGSTU37</i>	CotAD_13974	AD/Dt13	219	25.4	7.61	Cyto
<i>GhGSTU38</i>	CotAD_13952	AD/Dt13	219	25.2	6.53	Cyto
<i>GhGSTU39</i>	CotAD_13955	AD/Dt13	219	25.3	7.63	Cyto
<i>GhGSTU40</i>	CotAD_70631	AD/Dt13	226	25.7	5.29	Cyto



续表

基因名称	编号	染色体	氨基酸长度(aa)	分子量(kDa)	等电点	亚细胞定位预测
<i>GhGSTU41</i>	CotAD_20991	AD/Dt13	220	24.4	5.06	Cyto
<i>GhGSTU42</i>	CotAD_11525	scaffold	224	26.5	5.93	Cyto
<i>GhGSTU43</i>	CotAD_11530	scaffold	236	27.8	6.47	Cyto
<i>GhGSTU44</i>	CotAD_11531	scaffold	224	26.2	6.19	Cyto
<i>GhGSTU45</i>	CotAD_11516	scaffold	224	26.4	6.15	Cyto
<i>GhGSTU46</i>	CotAD_11534	scaffold	223	25.7	5.59	Cyto
<i>GhGSTU47</i>	CotAD_11533	scaffold	223	25.6	6.32	Chlo
<i>GhGSTU48</i>	CotAD_04867	scaffold	228	26.5	6.01	Cyto
<i>GhGSTU49</i>	CotAD_34416	scaffold	222	24.7	5.27	Cyto
<i>GhGSTU50</i>	CotAD_07291	scaffold	240	26.8	5.87	Cyto
<i>GhGSTF1</i>	CotAD_70142	AD/At1	219	24.6	8.67	Chlo
<i>GhGSTF2</i>	CotAD_26375	AD/At7	219	24.7	5.55	Chlo
<i>GhGSTF3</i>	CotAD_65302	AD/At7	215	24.7	5.46	Chlo
<i>GhGSTF4</i>	CotAD_12984	AD/At8	215	24.1	6.44	Chlo
<i>GhGSTF5</i>	CotAD_50520	AD/At11	132	14.7	8.24	Chlo
<i>GhGSTF6</i>	CotAD_57509	AD/Dt4	212	24.1	6.19	Chlo
<i>GhGSTF7</i>	CotAD_05169	AD/Dt7	215	24.9	5.58	Mito
<i>GhGSTF8</i>	CotAD_38732	AD/Dt12	214	24.5	6.22	Cyto
<i>GhGSTF9</i>	CotAD_39225	scaffold	180	20.4	5.97	Cyto
<i>GhTCHQD1</i>	CotAD_57121	AD/At3	265	30.9	8.91	Mito
<i>GhTCHQD2</i>	CotAD_21529	AD/Dt9	265	30.8	9.34	Mito
<i>GhEF1By1</i>	CotAD_23490	AD/At8	421	47.7	6.97	Cyto
<i>GhEF1By2</i>	CotAD_14252	AD/Dt8	421	47.6	7.01	Cyto
<i>GhGSTT1</i>	CotAD_59092	AD/At10	250	28.6	8.93	Cyto
<i>GhGSTT2</i>	CotAD_39522	AD/At13	277	31.4	9.48	Cyto
<i>GhGSTT3</i>	CotAD_39881	AD/Dt8	250	28.3	9.49	Mito
<i>GhGSTT4</i>	CotAD_51881	scaffold	250	28.7	9.33	Cyto
<i>GhGSTZ1</i>	CotAD_73291	AD/At8	193	21.8	4.99	Cyto
<i>GhGSTZ2</i>	CotAD_72993	AD/At11	219	24.7	5.34	Cyto
<i>GhGSTZ3</i>	CotAD_15580	AD/Dt11	218	24.7	5.21	Cyto

注：AD：陆地棉 A 组、D 组染色体；At：陆地棉 A 组染色体；Dt：陆地棉 D 组染色体；Cyto：细胞质；Chlo：叶绿体；Mito：线粒体。

## 2.4 陆地棉 *GST* 基因染色体定位

根据陆地棉基因组数据库，利用 MapInspect 软件对陆地棉 *GST* 进行染色体定位分析。结果显示，U 族的 *GST* 基因分布广泛，主要分布在 AD/At6、AD/At9、AD/Dt6、AD/Dt7、AD/Dt12、AD/Dt13 6 条染色体上，且大部分处于染色体的中上部，只有少数几个处于下部。T 族的 *GST* 基因分布在 AD/At10、AD/At13、AD/Dt8 3 条染色体上，且都在染色体的

中下游。F 族的 *GST* 基因分布较为广泛，主要集中在染色体的中上游。Z 族的 *GST* 基因分布在 AD/At8、AD/At11、AD/Dt11 3 条染色体上，且在染色体的中上游。TCHQD 族的 *GST* 基因分布在 AD/At3 和 AD/Dt9 两条染色体上。EF1B $\gamma$  族的 *GST* 基因则分布在 AD/At8 和 AD/Dt8 两条染色体上(图 6)。其中，染色体 AD/At9 上有 8 个 *GST* 基因，染色体 AD/Dt13 上有 7 个 *GST* 基因，染色体 AD/Dt7、AD/Dt12 上分别有 5 个 *GST* 基因。根据串联复制的标准<sup>[29,30]</sup>，将

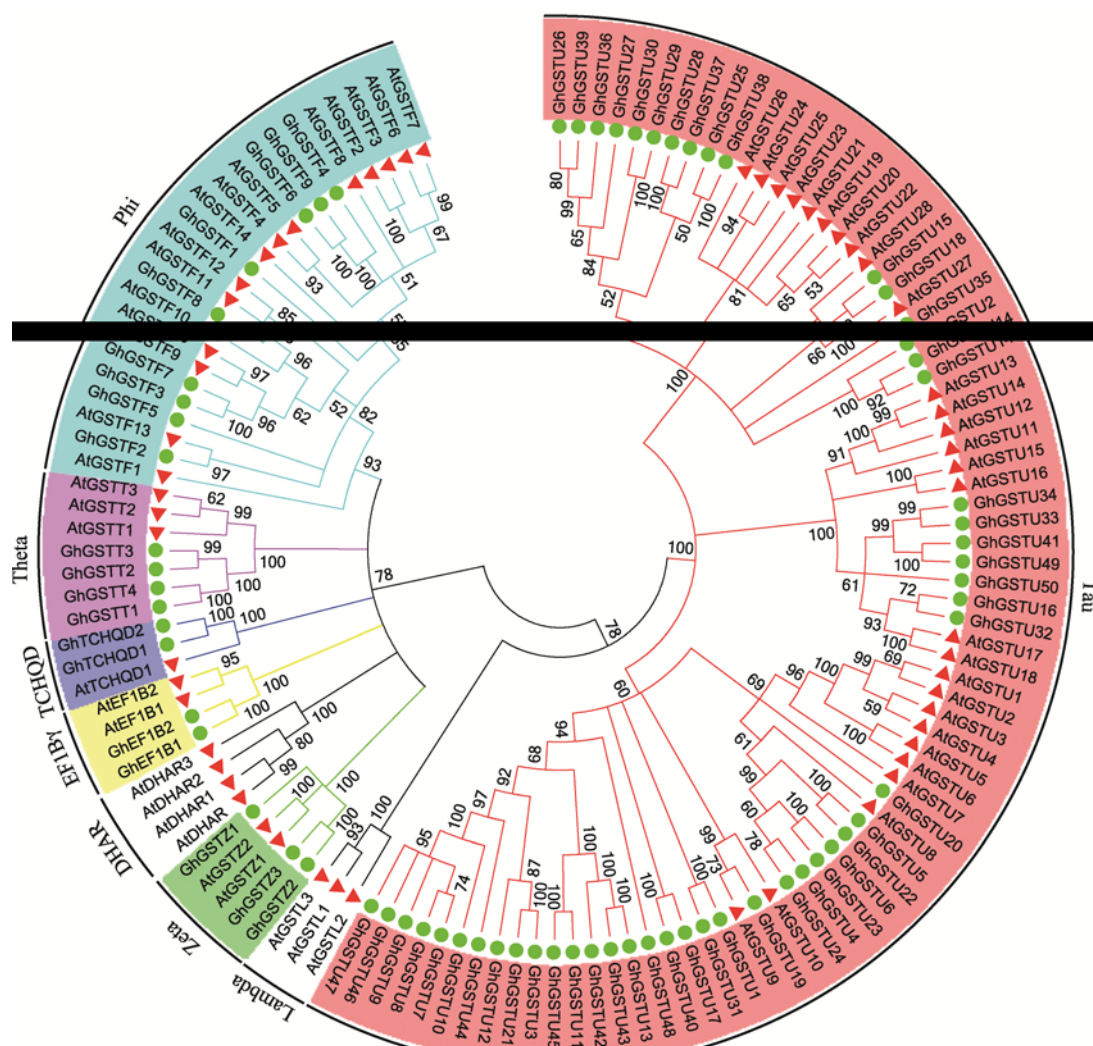


图 1 陆地棉与拟南芥 GST 基因家族的进化分析

Fig. 1 Evolutionary analysis of the GST gene family in *G. hirsutum* and *A. thaliana*

基因簇规定为两个相邻的 GST 基因之间最多有 5 个干预基因, 陆地棉中 GST 家族可分为 4 个基因簇, 且在多个节段出现复制事件, 基因簇和节段复制事件为陆地棉 GST 基因复制进化提供途径。

为了进一步了解 GST 的进化关系, 利用 Circos 软件对陆地棉、亚洲棉及雷蒙德氏棉的 Phi 类 GST 基因进行共线性分析。结果表明, 陆地棉, 雷蒙德氏棉和亚洲棉 GST 基因之间存在明显的共线性关系, 其中陆地棉与亚洲棉之间有 24 对旁系同源基因, 陆地棉与雷蒙德氏棉之间也存在 24 对旁系同源基因 (图 7)。这说明陆地棉 GST 基因复制半数来自于雷蒙德氏棉, 半数来自于亚洲棉。此外, 还有多个基因与亚洲棉和雷蒙德氏棉不存在复制片段, 说明陆地棉 GST 基因还存在独立复制事件, 这可能是导致陆

地棉 GST 基因数目加倍的一种方式。

## 2.5 Phi 类 GST 基因表达分析

研究表明, 在植物 GST 基因家族中与黄酮类物质转运积累相关的是 Phi 类亚家族<sup>[31]</sup>。为了分析陆地棉中与原花青素/花青素转运和积累相关的 GST 基因, 本文根据陆地棉 Phi 类 GST 基因序列设计荧光定量引物 (表 1), 对陆地棉 Phi 类 9 个 GST 基因在各组织器官 (根、茎、叶) 和各发育时期的纤维进行荧光定量 PCR 分析。结果显示, *GhGSTF1* 在根、茎、叶及纤维中都不表达, 推测其可能为假基因。*GhGSTF2* 在根、茎、叶及纤维中都显示较高的表达量, 在不同发育时期的纤维中表达呈现先上升后下降趋势, 在花后 15 d 时表达量最高, *GhGSTF3* 和 *GhGSTF7*

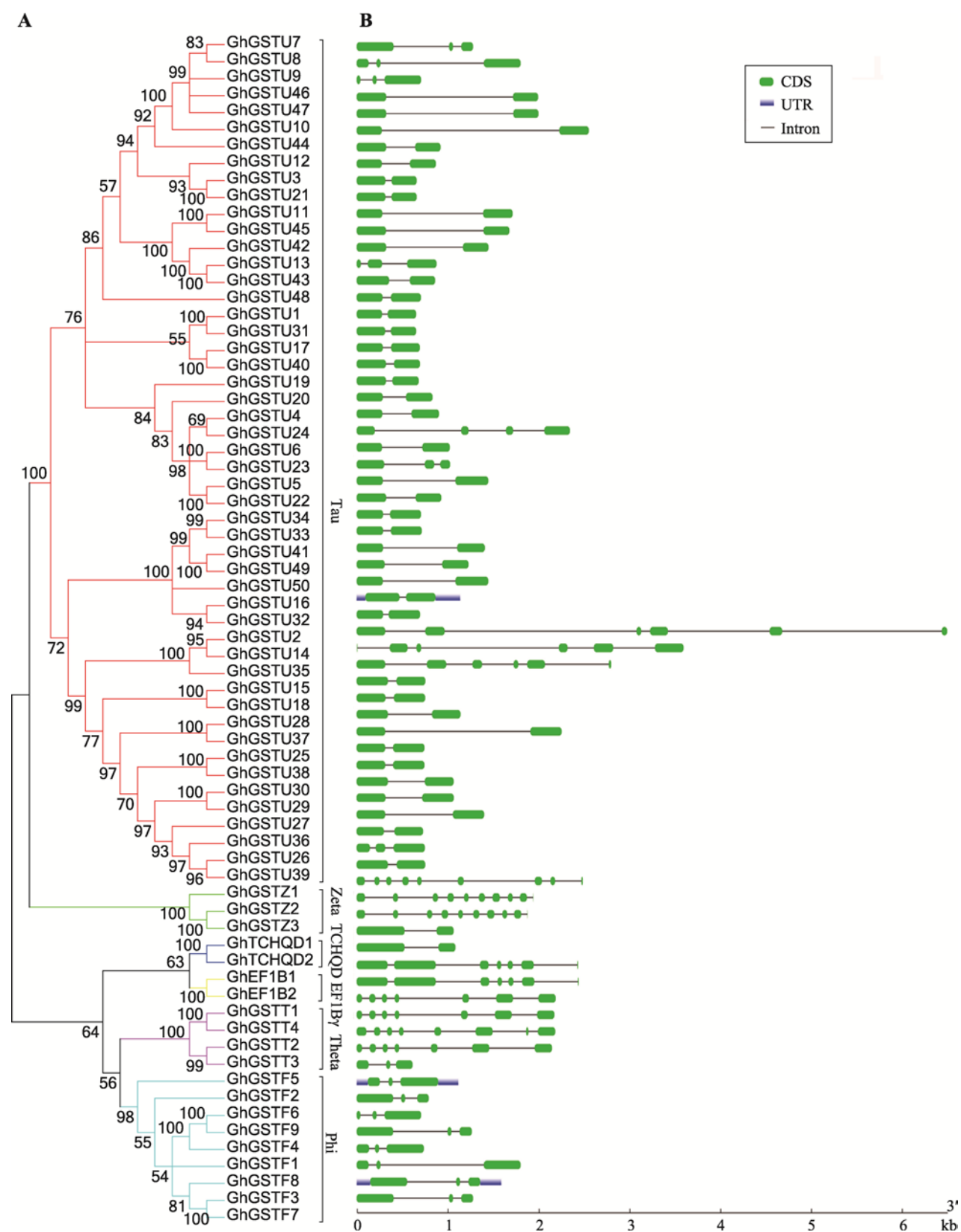


图 2 陆地棉 GST 的系统进化和基因结构分析

Fig. 2 The phylogenetic tree and protein structure of GST proteins in *G. hirsutum*



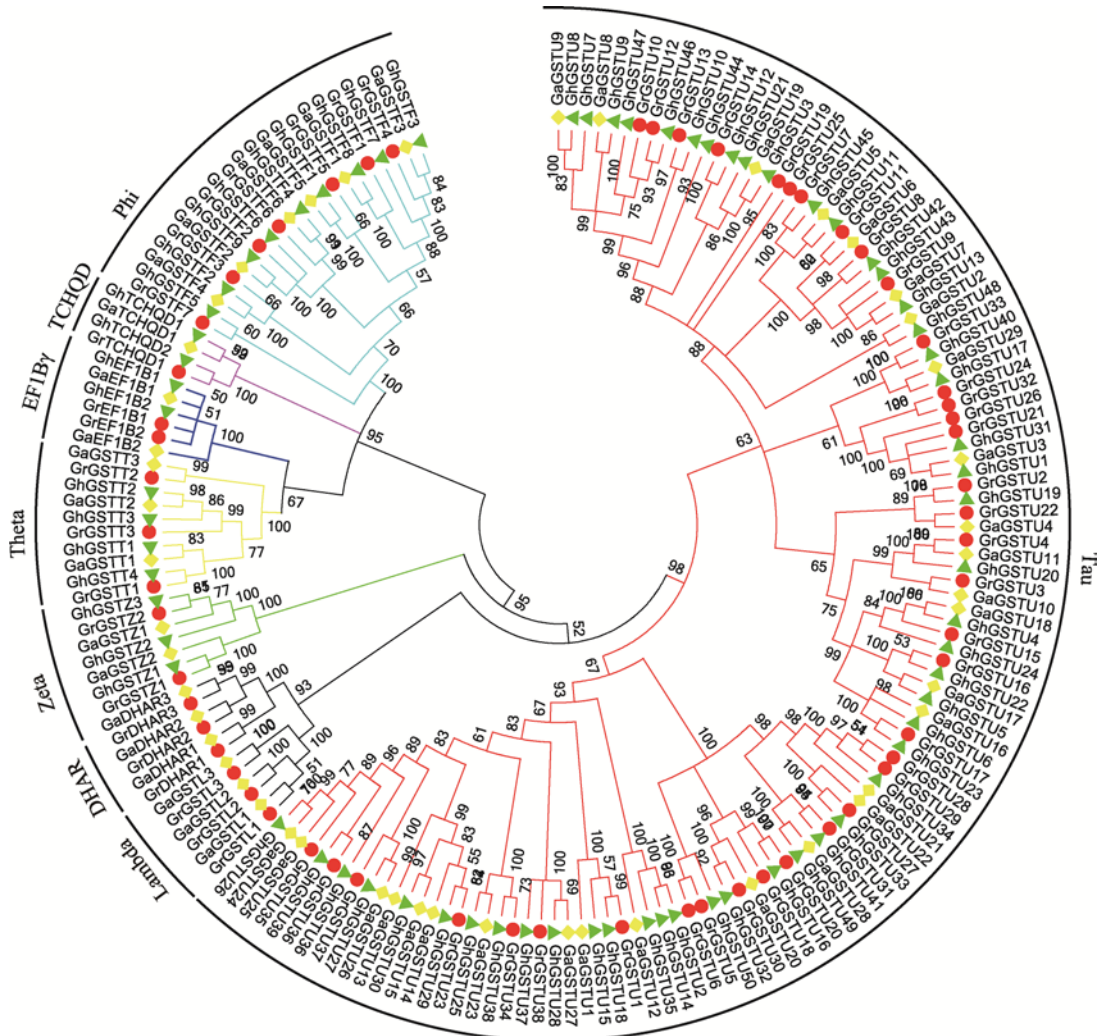


图 3 陆地棉与亚洲棉、雷蒙德氏棉 GST 基因家族的进化分析

Fig. 3 Evolutionary analysis of the GST gene family in *G. hirsutum*, *G. raimondii* and *G. arboreum*

在根、茎、叶组织中几乎不表达或有少量表达,但在不同发育时期的纤维中表达呈现先上升后下降趋势,且花后 15 d 的表达量最高;而 *GhGSTF5* 在根、茎、叶组织中几乎不表达,在不同发育时期的纤维中表达量也是呈现先上升后下降的趋势,但表达量最高在花后 25 d; *GhGSTF8* 在根中基本上不表达,在茎和叶中都有较高的表达水平,在纤维发育前期表达量较低,花后 25 d 时达到最大,之后又降低; *GhGSTF4* 在根、茎、叶组织中有少量表达,在不同发育时期的纤维中表达先升高后下降,花后 25 d 时表达量最大; *GhGSTF6* 和 *GhGSTF9* 表达趋势相近,在根、茎、叶中表达水平较高,在纤维发育的各个时期表达水平先上升后下降,在 25 d 时表达水平达

到最高(图 8)。

根据系统发育树分析(图 1), *GhGSTF2*、*GhGSTF3*、*GhGSTF5* 和 *GhGSTF7* 与拟南芥 *AtGSTF13*、*AtGSTF9*、*AtGSTF10* 在同一分支上, Kiyosue 等<sup>[32]</sup>研究发现 *AtGSTF10*(ERD13)的表达受干旱的诱导,推测 *GhGSTF2*、*GhGSTF3*、*GhGSTF5* 和 *GhGSTF7* 的表达可能也受干旱的诱导,它们的功能还有待进一步研究。 *GhGSTF8* 与拟南芥的 *AtGSTF11*、*AtGSTF12* 聚在一个分支上, *AtGSTF12*(即 *AtTT19*)不仅可以转运花青素,还能参与原花青素的转运与积累<sup>[33]</sup>,推测 *GhGSTF8* 也可能参与陆地棉纤维中原花青素/花青素的转运和积累。 *GhGSTF4*、*GhGSTF6* 和 *GhGSTF9* 与拟南芥的 *AtGSTF2*、*AtGSTF6*、

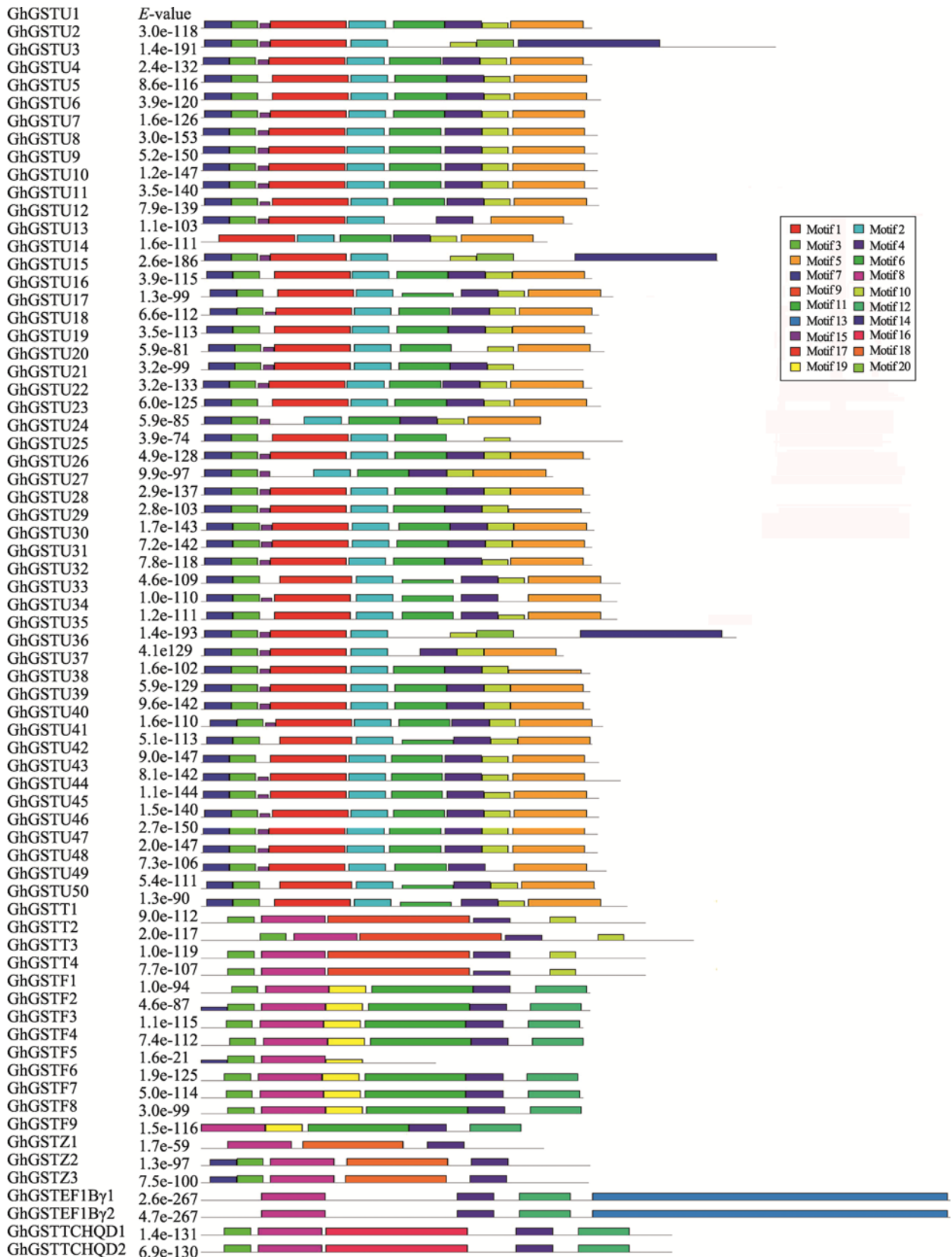


图 4 陆地棉 GST 蛋白保守基序

Fig. 4 Conservative motifs of GST proteins in *G. hirsutum*



图 5 陆地棉 GST 蛋白保守基序分布

**Fig. 5** The distribution of conservative motifs of GST proteins in *G. hirsutum*

1 : Motif 1 ; 2 : Motif 8 ; 3 : Motif 17.

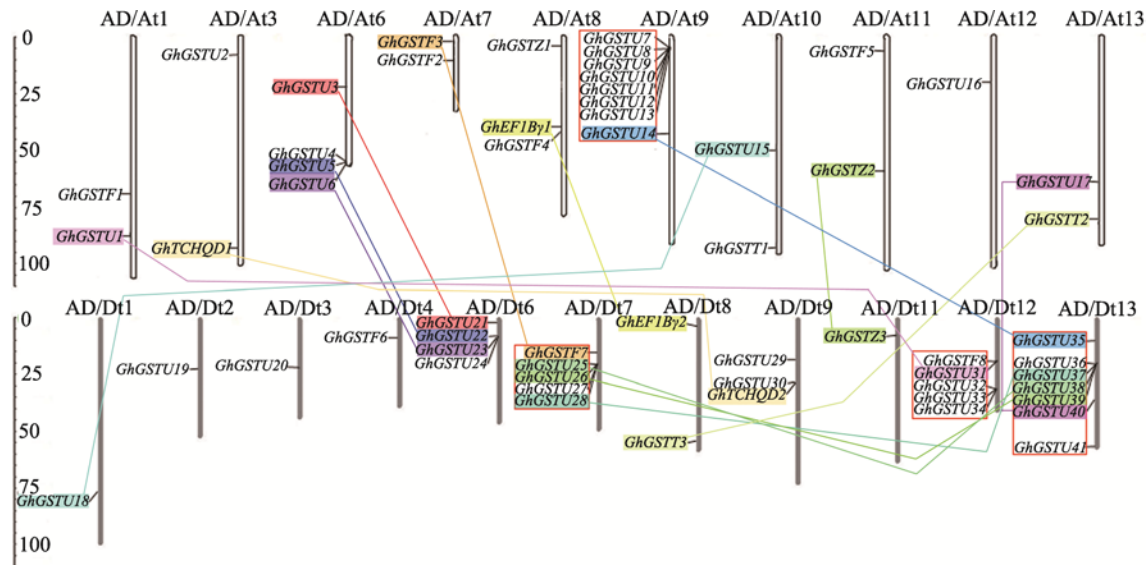


图 6 陆地棉 GSTs 染色体定位

**Fig. 6** Chromosome localization of GST genes in *G. hirsutum*

AD/At : 陆地棉 A 组染色体 ; AD/Dt : 陆地棉 D 组染色体。

*AtGSTF8* 等基因聚在同一分支上, *AtGSTF2*、*AtGSTF6* 具有转运类黄酮的功能, 并且参与槲皮素、山奈酚和生长素的转运和积累, 调节植物生长和抗胁迫反应<sup>[34]</sup>, 推测 *GhGSTF4*、*GhGSTF6* 和 *GhGSTF9* 可能与槲皮素、山奈酚和生长素结合参与调节陆地棉的生长和抗胁迫反应。

3 讨 论

本研究在基因组水平上对陆地棉 GST 基因家族进行了鉴定和分析, 共鉴定出 70 个 *GST* 成员, 通

过与拟南芥以及亚洲棉和雷蒙德氏棉的 *GST* 家族进行系统进化分析, 将陆地棉 *GST* 家族分成 U 族、Z 族、F 族、T 族、EF1B $\gamma$  族和 TCHQD 族 6 个亚家族。Dixon 等<sup>[5]</sup>在拟南芥中发现了 54 个 *GST*, 被分为 U 族、F 族、Z 族、L 族、T 族、EF1B $\gamma$  族、DHAR 族和 TCHQD 族 8 个亚家族, 而 Kumar 等<sup>[6]</sup>在水稻中发现了 82 个 *GST*, 被分为 F 族、U 族、T 族、Z 族、L 族以及 EF1B $\gamma$  6 个亚家族。Lan 等<sup>[35]</sup>在哺乳动物中发现 *GST* 有 6 个亚家族, 分别为 Alpha、Kappa、Mu、Pi、Sigma 和 Omega。F 和 U 族是植物 *GST* 基



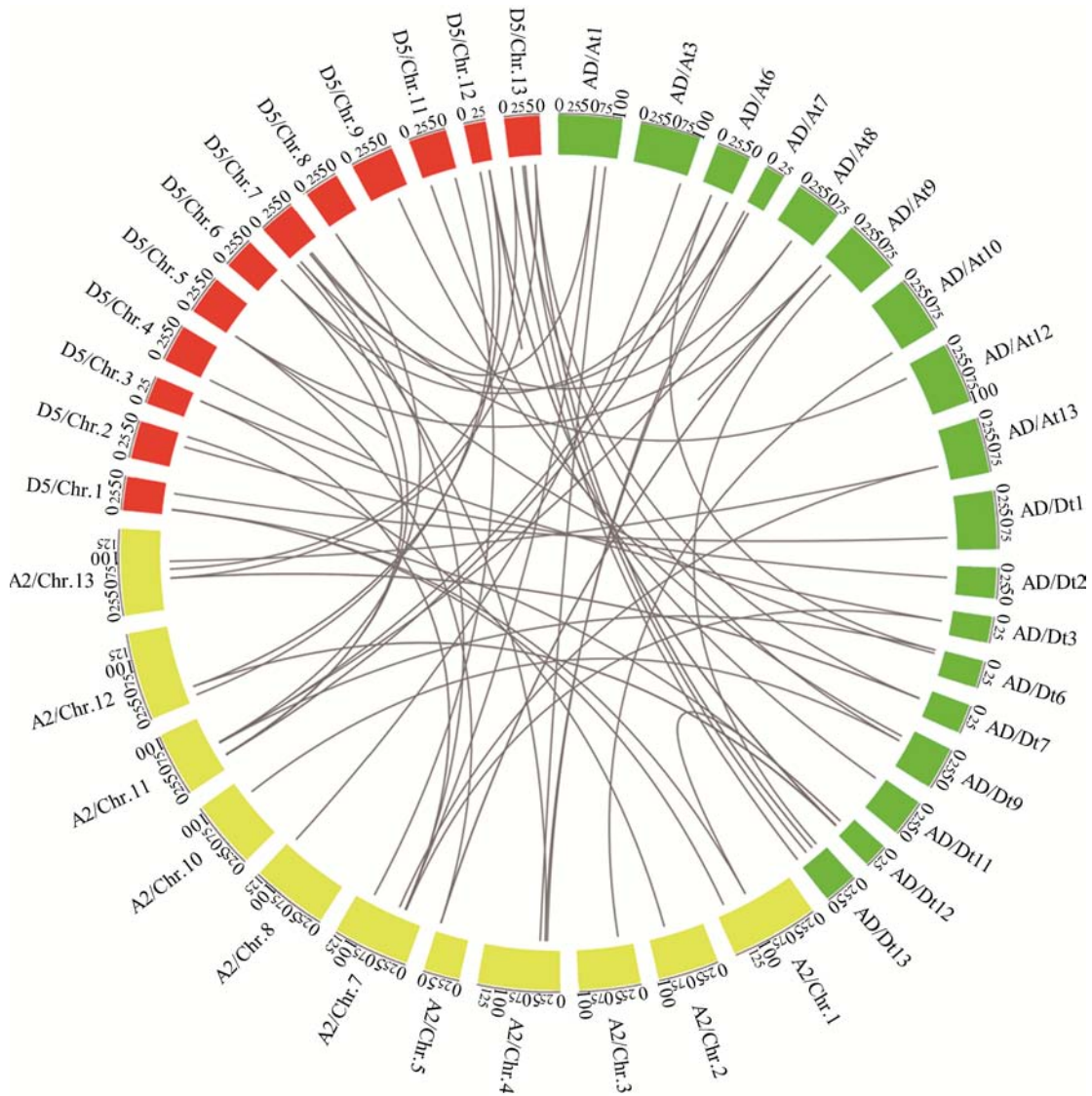


图 7 陆地棉 *GSTs* 共线性分析

**Fig.7 Collinearity analysis of *GST* genes in *G. hirsutum***

AD/At: 陆地棉 A 组染色体; AD/Dt: 陆地棉 D 组染色体; A2/Chr.: 亚洲棉染色体; D5/Chr.: 雷蒙德氏棉染色体。

因家族中特有的两个亚家族,可能的原因是,与动物相比,高等植物 *GST* 除具有解毒作用外,还具有缓解逆境胁迫、抗氧化作用、转运黄酮类物质、调节植物生长和发育等特有的功能<sup>[36-38]</sup>。

通过系统进化分析,将陆地棉 *GST* 基因家族分为 U 族、F 族、Z 族、T 族、EF1B $\gamma$  族和 TCHQD 族,这与拟南芥的分类相近,但却没有发现 DHAR 族和 lambda 族。可能原因是 DHAR 族和 lambda 族 *GST* 在长期进化历程中逐渐丢失。通过与拟南芥 *GST* 基因家族的复合进化树分析,发现 lambda 族与 U 族聚在一个分支上,而 DHAR 族与 Z 族和 EF1B $\gamma$  族也聚

在一个分支上,这个现象表明,陆地棉的 U 族以及 Z 族、EF1B $\gamma$  族部分成员可能具有 Lambda 族和 DHAR 族相似的功能。*GST* 的种类不同,其功能存在一定的差别。有研究表明,对于植物中特有的 U 族和 F 族,当植物受到生物和非生物胁迫时, U 族和 F 族 *GST* 才被诱导产生<sup>[1]</sup>。而 T 族 *GST* 具有解毒氧化脂类的作用, Z 族和 EF1B $\gamma$  族 *GST* 主要功能是参与酪氨酸分解代谢和真核翻译延伸因子  $\gamma$  亚基的编码。这些结果说明植物的生长发育是通过不同亚家族的 *GST* 协同参与调节。

关于内含子的起源,目前有两种假说:一种是

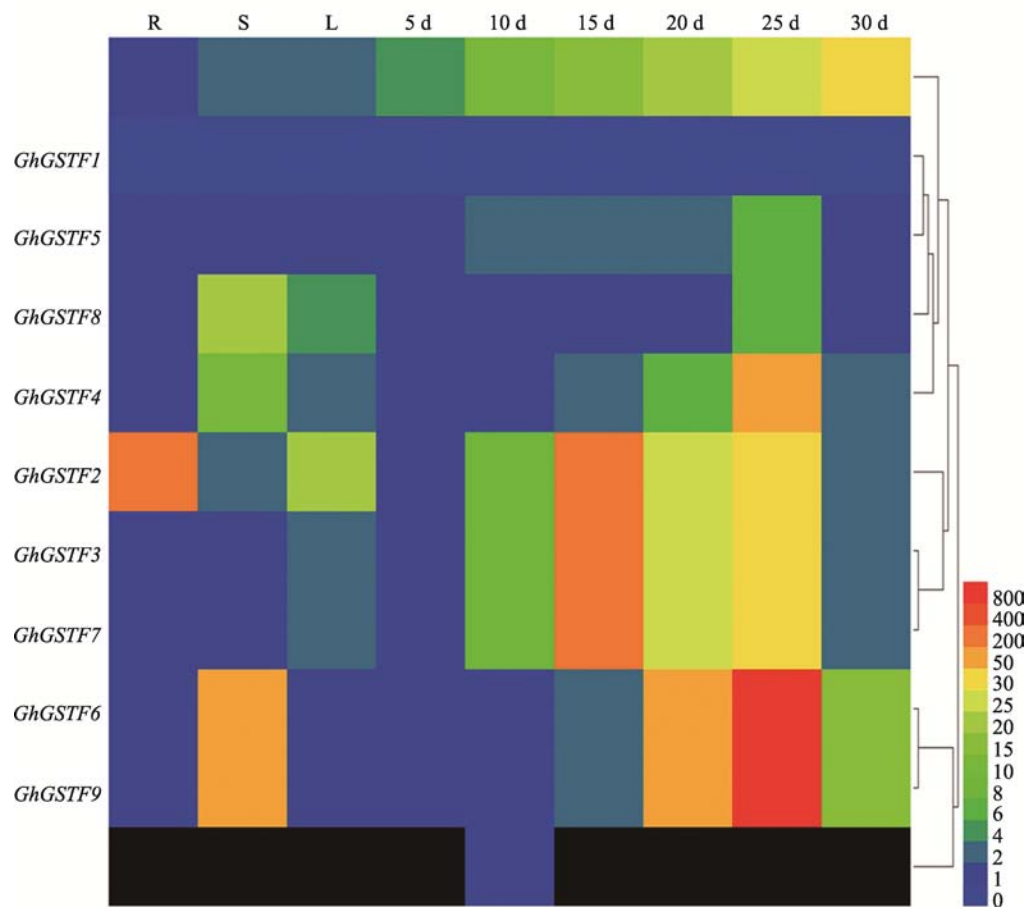


图 8 陆地棉 *GSTF* 基因的表达模式  
**Fig. 8 Expression patterns of *GSTF* genes in *G. hirsutum***  
R: 根; S: 茎; L: 叶; 5 d~30 d: 花后天数。

基因本身就不连续，内含子一开始就存在于基因内部；另一种是基因本身是连续的，在进化的过程中由于转座子的插入而产生内含子。曹运鹏等<sup>[39]</sup>通过对砀山酥梨(*Pyrus bretschneideri* Rehd.)4CL 基因家族研究表明，最早出现的基因内含子数目最少，随着复制的不断进行，基因成员不断增多且内含子数目逐渐增多；同时还指出基因结构中内含子的起源有的是自身内部本身存在，有的是随着基因复制过程中转座子的插入而产生或增多。本研究通过对陆地棉 *GST* 的结构和保守基序分析，发现 U 族 *GST* 基因大部分含有一个内含子，Z 族 *GST* 基因大多数含有 9 个内含子，F 族 *GST* 基因含有两个内含子。因此，由各亚家族成员的数目以及基因中内含子数目分析推测，陆地棉 *GST* 家族中可能最先出现的是 U 族和 F 族，随着 *GST* 家族进化的发生，*GST* 家族

不断产生其他亚家族。

对陆地棉 *GST* 基因染色体定位分析显示，*GST* 基因在染色体上分布广泛，且不均匀，主要分布在 AD/At9、AD/Dt7、AD/Dt12、AD/Dt13 4 条染色体上，且 AD/At9 染色体上分布最多，有 8 个 *GST* 基因。但在 AD/At2、AD/At4、AD/At5、AD/Dt5、AD/Dt10 染色体上没有发现 *GST* 基因。研究发现 陆地棉 *GST* 成员最多的一类亚家族为 U 族，也是染色体上形成基因簇现象最多的一类亚家族，共有 6 组基因簇，推测 U 亚家族产生时期可能最早，之后衍生出其他亚家族。而其他 3 个亚家族 *GST* 基因都各自分散在不同染色体上，没有出现基因簇，其原因可能与这些亚家族成员的数目及出现的时间有关。陆地棉 *GST* 共线性分析表明，陆地棉中除了多倍体化使 *GST* 基因数目加倍外，还存在独立复制事件使 *GST*



基因数目加倍。

GST 蛋白结构的差异决定了其功能的多样性。GST 最主要的功能是解毒, 其次是次生代谢物的转运<sup>[31]</sup>。前人研究表明, Phi(F)类 GST 基因功能主要是对各种有害化合物的解毒、次生代谢物的生物合成和/或运输以及作为过氧化物酶除去氧化应激和脂质过氧化作用的产物<sup>[40,41]</sup>。在 GST 所有亚家族中, 发现参与黄酮类物质运输最多的是 Phi(F)类 GST。Kitamura 等<sup>[42]</sup>研究表明, 拟南芥 *AtTT19* 参与原花青素/花青素的转运积累, 且 *AtTT19* 属于 F 族 GST(即 *GSTF* 类)<sup>[5]</sup>; Dixon 等<sup>[34]</sup>研究发现, 拟南芥 *AtGSTF2*、*AtGSTF6* 参与类黄酮(槲皮素、山奈酚和生长素)的转运和积累, 调节植物生长和抗胁迫反应; 金雪花等<sup>[43]</sup>从瓜叶菊(*Pericallis hybrida*)中获得 *ScGST3* 基因, 该基因与拟南芥 *AtTT19* 亲缘关系较近, 并且在含有花青素的组织中表达, 在花青素含量极少的组织中表达极弱, 推测 *ScGST3* 基因参与花青素的转运和积累。通过对陆地棉 Phi 类基因进行荧光定量分析以及复合进化树分析, 发现 *GhGSTF8* 与拟南芥 *AtTT19* 聚在一个分支上, *AtTT19* 不仅可以转运花青素, 还能参与原花青素的转运与积累<sup>[33]</sup>, 推测陆地棉 F 族中 *GhGSTF8* 可能与原花青素/花青素的转运和积累有关, 而 *GhGSTF4*、*GhGSTF6* 和 *GhGSTF9* 与 *AtGSTF2*、*AtGSTF6* 聚为一类, 推测 *GhGSTF4*、*GhGSTF6* 和 *GhGSTF9* 可能参与槲皮素、山奈酚和生长素的转运和积累。这为后续相关的功能研究提供理论基础。

## 参考文献(References):

- [1] Nutricati E, Miceli A, Blando F, De Bellis L. Characterization of two *Arabidopsis thaliana* glutathione S-transferases. *Plant Cell Rep*, 2006, 25(9): 997–1005. [DOI]
- [2] Jain M, Ghanashyam C, Bhattacharjee A. Comprehensive expression analysis suggests overlapping and specific roles of rice glutathione S-transferase genes during development and stress responses. *BMC Genomics*, 2010, 11: 73. [DOI]
- [3] Dixon DP, Davis BG, Edwards R. Functional divergence in the glutathione transferase superfamily in plants identification of two classes with putative functions in redox homeostasis in *Arabidopsis thaliana*. *J Biol Chem*, 2002, 277(34): 30859–30869. [DOI]
- [4] Shimabukuro RH, Swanson HR, Walsh WC. Glutathione conjugation: atrazine detoxication mechanism in corn. *Plant Physiol*, 1970, 46(1): 103–107. [DOI]
- [5] Dixon DP, Laphorn A, Edwards R. Plant glutathione transferases. *Genome Biol*, 2002, 3: 3004.1. [DOI]
- [6] Kumar S, Asif MH, Chakrabarty D, Tripathi RD, Dubey RS, Trivedi PK. Differential expression of rice lambda class GST gene family members during plant growth, development, and in response to stress conditions. *Plant Mol Biol Rep*, 2013, 31(3): 569–580. [DOI]
- [7] McGonigle B, Keeler SJ, Lau SMC, Koeppe MK, O'Keefe DP. A genomics approach to the comprehensive analysis of the glutathione S-transferase gene family in soybean and maize. *Plant Physiol*, 2000, 124(3): 1105–1120. [DOI]
- [8] Edwards R, Dixon DP, Walbot V. Plant glutathione S-transferases: enzymes with multiple functions in sickness and in health. *Trends Plant Sci*, 2000, 5(5): 193–198. [DOI]
- [9] Kampranis SC, Damianova R, Atallah M, Toby G, Kondi G, Tsiachlis PN, Makris AM. A novel plant glutathione S-transferase/peroxidase suppresses Bax lethality in yeast. *J Biol Chem*, 2000, 275(38): 29207–29216. [DOI]
- [10] Goodman CD, Casati P, Walbot V. A multidrug resistance-associated protein involved in anthocyanin transport in *Zea mays*. *Plant Cell*, 2004, 16(7): 1812–1826. [DOI]
- [11] Pérez-Díaz R, Madrid-Espinoza J, Salinas-Cornejo J, González-Villanueva E, Ruiz-Lara S. Differential roles for VviGST1, VviGST3, and VviGST4 in proanthocyanidin and anthocyanin transport in *Vitis vinifera*. *Front Plant Sci*, 2016, 7: 1166. [DOI]
- [12] HUANG YM. Expression analyses and functional verification of proanthocyanidin precursor transmembrane genes MATE and GSTJN persimmon fruit[D]. Wuhan: Huazhong Agricultural University, 2015.  
黄燕梅. 柿原花青素前体跨膜及转运相关基因 MATE、GST 的表达分析及功能验证[学位论文]. 武汉: 华中农业大学, 2015. [DOI]
- [13] Zettl R, Schell J, Palme K. Photoaffinity labeling of *Arabidopsis thaliana* plasma membrane vesicles by 5-azido-[7-3H] indole-3-acetic acid: identification of a glutathione S-transferase. *Proc Natl Acad Sci USA*, 1994, 91(2): 689–693. [DOI]
- [14] Roxas VP, Smith RK Jr, Allen ER, Allen RD. Overexpression of glutathione-S-transferase/glutathioneperoxidase enhances the growth of transgenic tobacco seedlings during stress. *Nat Biotechnol*, 1997, 15(10): 988–991. [DOI]
- [15] Cummins I, Cole DJ, Edwards R. A role for glutathione S-transferases functioning as glutathione peroxidases in resistance to multiple herbicides in black-grass. *Plant J*,

- 1999, 18(3): 285–292. [DOI]
- [16] Chan C, Lam HM. A putative lambda class glutathione S-transferase enhances plant survival under salinity stress. *Plant Cell Physiol*, 2014, 55(3): 570–579. [DOI]
- [17] Jepson I, Lay VJ, Holt DC, Bright SWJ, Greenland AJ. Cloning and characterization of maize herbicide safener-induced cDNAs encoding subunits of glutathione S-transferase isoforms I, II and IV. *Plant Mol Biol*, 1994, 26(6): 1855–1866. [DOI]
- [18] Loyall L, Uchida K, Braun S, Furuya M, Frohnmeyer H. Glutathione and a UV light-induced glutathione S-transferase are involved in signaling to chalcone synthase in cell cultures. *Plant Cell*, 2000, 12(10): 1939–1950. [DOI]
- [19] Zhang TZ, Hu Y, Jiang WK, Fang L, Guan XY, Chen JD, Zhang JB, Saski CA, Scheffler BE, Stelly DM, Hulse-Kemp AM, Wan Q, Liu BL, Liu CX, Wang S, Pan MQ, Wang YK, Wang DW, Ye WX, Chang LJ, Zhang WP, Song QX, Kirkbride RC, Chen XY, Dennis EB, Llewellyn DJ, Peterson DG, Thaxton P, Jones DC, Wang Q, Xu XY, Zhang H, Wu HT, Zhou L, Mei GF, Chen SQ, Tian Y, Xiang D, Li XH, Ding J, Zuo QY, Tao LN, Liu YC, Li J, Lin Y, Hui YY, Cao ZS, Cai CP, Zhu XF, Jiang Z, Zhou BL, Guo WZ, Li RQ, Chen ZJ. Sequencing of allotetraploid cotton (*Gossypium hirsutum* L. acc. TM-1) provides a resource for fiber improvement. *Nat Biotechnol*, 2015, 33(5): 531–537. [DOI]
- [20] Kitamura S, Akita Y, Ishizaka H, Narumi I, Tanaka A. Molecular characterization of an anthocyanin-related glutathione S-transferase gene in cyclamen. *J Plant Physiol*, 2012, 169(6): 636–642. [DOI]
- [21] Wagner U, Edwards R, Dixon DP, Mauch F. Probing the diversity of the *Arabidopsis* glutathione S-transferase gene family. *Plant Mol Biol*, 2002, 49(5): 515–532. [DOI]
- [22] Finn RD, Bateman A, Clements J, Coghill P, Eberhardt RY, Eddy SR, Heger A, Hetherington K, Holm L, Mistry JL, Sonnhammer EL, Tate J, Punta M. Pfam: the protein families database. *Nucleic Acids Res*, 2014, 42(D1): D222–D230. [DOI]
- [23] Letunic I, Doerks T, Bork P. SMART 6: recent updates and new developments. *Nucleic Acids Res*, 2009, 37(Suppl 1): D229–D232. [DOI]
- [24] Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG, Clustal W and Clustal X version 2.0. *Bioinformatics*, 2007, 23(21): 2947–2948. [DOI]
- [25] Soranzo N, Gorla MS, Mizzi L, De Toma G, Frova C. Organisation and structural evolution of the rice glutathione S-transferase gene family. *Mol Genet Genomics*, 2004, 271(5): 511–521. [DOI]
- [26] Bo H, Jin JP, Guo AY, Zhang H, Luo JC, Gao G. GSDS 2.0: an upgraded gene feature visualization server. *Bioinformatics*, 2015, 31(8): 1296–1297. [DOI]
- [27] Bailey TL, Johnson J, Grant CE, Noble WS. The MEME suite. *Nucleic Acids Res*, 2015, 43(W1): W39–W49. [DOI]
- [28] Dong YT, Li C, Zhang Y, He QL, Daud MK, Chen JH, Zhu SJ. Glutathione S-transferase gene family in *Gossypium raimondii* and *G. arboreum*: comparative genomic study and their expression under salt stress. *Front Plant Sci*, 2016, 7: 139. [DOI]
- [29] Zhao YX, Cai MJ, Zhang XB, Li YR, Zhang JH, Zhao HL, Kong F, Zheng YL, Qiu FZ. Genome-wide identification, evolution and expression analysis of *mTERF* gene family in maize. *PLoS One*, 2014, 9(4): e94126. [DOI]
- [30] Qiao LY, Zhang XJ, Han X, Zhang L, Li X, Zhan HX, Ma J, Luo PG, Zhang WP, Cui L, Li XY, Chang ZJ. A genome-wide analysis of the *auxin/indole-3-acetic acid* gene family in hexaploid bread wheat (*Triticum aestivum* L.). *Front Plant Sci*, 2015, 6: 770. [DOI]
- [31] Zhang YZ, Zhang F, Wang LY, Wei K, Cheng H. Plant glutathione S-transferases: roles in flavonoid accumulation. *Plant Physiol J*, 2015, 51(11): 1815–1820. 张亚真, 张芬, 王丽鸳, 韦康, 成浩. 植物谷胱甘肽转移酶在类黄酮累积中的作用. *植物生理学报*, 2015, 51(11): 1815–1820. [DOI]
- [32] Kiyosue T, Yamaguchi-Shinozaki K, Shinozaki K. Characterization of two cDNAs (ERD11) and (ERD13) for dehydration-inducible genes that encode putative glutathione S-transferases in *Arabidopsis thaliana* L. *FEBS Lett*, 1993, 335(2): 189–192. [DOI]
- [33] Kitamura S, Shikazono N, Tanaka A. *TRANSPARENT TESTA 19* is involved in the accumulation of both anthocyanins and proanthocyanidins in *Arabidopsis*. *Plant J*, 2004, 37(1): 104–114. [DOI]
- [34] Dixon DP, Sellars JD, Edwards R. The *Arabidopsis* phi class glutathione transferase *At GSTF2*: binding and regulation by biologically active heterocyclic ligands. *Biochem J*, 2011, 438(1): 63–70. [DOI]
- [35] Lan T, Yang ZL, Xue Y, Liu YJ, Wang XR, Zeng QY. Extensive functional diversification of the *Populus glutathione S-transferase* supergene family. *Plant Cell*, 2009, 21(12): 3749–3766. [DOI]

- [36] Moons A. Regulatory and functional interactions of plant growth regulators and plant glutathione *S*-transferases (GSTs). *Vitam Horm*, 2005, 72: 155–202. [DOI]
- [37] Marrs KA. The functions and regulation of glutathione *S*-transferases in plants. *Annu Rev Plant Phys Plant Mol Biol*, 1996, 47(1): 127–158. [DOI]
- [38] Frova C. Glutathione transferases in the genomics era: new insights and perspectives. *Biomol Eng*, 2006, 23(4): 149–169. [DOI]
- [39] Cao YP, Fang Z, Li SM, Yan CC, Ding QQ, Cheng X, Lin Y, Guo N, Cai YP. Genome-wide identification and analyses of *4CL* gene families in *Pyrus bretschneideri* Rehd. *Hereditas (Beijing)*, 2015, 37(7): 711–719.  
曹运鹏, 方志, 李姝妹, 闫冲冲, 丁庆庆, 程曦, 林毅, 郭宁, 蔡永萍. 砀山酥梨 *4CL* 基因家族的全基因组鉴定与分析. *遗传*, 2015, 37(7): 711–719. [DOI]
- [40] Munyampundu JP, Xu YP, Cai XZ. Phi class of glutathione *S*-transferase gene superfamily widely exists in nonplant taxonomic groups. *Evol Bioinform*, 2016, 12: 59–71. [DOI]
- [41] Edwards R, Dixon DP. Plant glutathione transferases. *Methods Enzymol*, 2005, 401: 169–186. [DOI]
- [42] Kitamura S, Matsuda F, Tohge T, Yonekura-Sakakibara K, Yamazaki M, Saito K, Narumi I. Metabolic profiling and cytological analysis of proanthocyanidins in immature seeds of *Arabidopsis thaliana* flavonoid accumulation mutants. *Plant J*, 2010, 62(4): 549–559. [DOI]
- [43] Jin XH. The study of anthocyanidins biosynthesis pathway in *senecio cruentus* based on the high-throughput RNA sequencing technique[D]. Beijing: Beijing Forestry University, 2013.  
金雪花. 基于高通量测序的瓜叶菊花青素苷合成途径研究[学位论文]. 北京: 北京林业大学, 2013. [DOI]

(责任编辑: 李付广)