棉花SMXL基因家族的全基因组分析及表达分析

贾婷婷¹, 竺丽萍², 肖光辉^{2*}, 李鸿彬^{1*}

1. 石河子大学生命科学学院, 石河子 832003;

2. 陕西师范大学生命科学学院, 西安 710119

* 联系人, E-mail: lih@shzu.edu.cn; guanghuix@snnu.edu.cn

收稿日期: 2021-03-24; 接受日期: 2021-06-25; 网络版发表日期: 2021-12-03

国家自然科学基金面上项目(批准号: 32070549)、中国科协青年人才托举工程(批准号: 2021QNRC001)、陕西省科协青年托举人才计划(批准号: 20190205)、陕西省博士后项目(一等资助)(批准号: 2018BSHYDZZ76)、中央高校基本科研业务费重点项目(批准号: GK202002005)和棉花生物学国家重点实验室开放课题(批准号: CB2020A12)资助

摘要 独脚金内酯(strigolactones, SLs)是一种新型植物激素,在植物生长发育中发挥重要作用.目前在棉花中对SL研究仍较少,探究SL影响棉花生长发育及纤维品质形成的功能和机制具有重要意义. *SMXL*(*SUPPRESSOR OF MAX2 1-like*)是独角金内酯信号通路的关键基因,充当SL信号转导的阻遏物,关于棉花*SMXL*基因的研究尚未见报道.本文以陆地棉、海岛棉、亚洲棉及雷蒙德氏棉全基因组为基础,运用生物信息学方法,根据*SMXL*蛋白序列结构域特征鉴定到16个*GhSMXL*基因、16个*GbSMXL*基因、8个*GaSMXL*基因和8个*GrSMXL*基因.系统发育分析发现*SMXL*分为4个亚家族: *SMXL-I* (*SMXL1*, 2), *SMXL-II* (*SMXL6*, 7, 8), *SMXL-III* (*SMXL3*)和*SMXL-IV* (*SMXL4*, 5).对*SMXL*的进化特征分析发现,位于同一亚组间的*SMXL*成员具有相似的基因结构与蛋白结构.分析发现,片段复制是*SMXL*成员在棉花中扩张的主要方式.组织表达谱分析显示,在陆地棉中*GhSMXL1-A*, *GhSMXL1-D*, *GhSMXL7-1-A*, *GhSMXL7-1-D*, *GhSMXL8-1-A*, *GhSMXL8-1-D*在根、茎、叶、纤维及胚珠中优势表达;在亚洲棉中*GaSMXL8-1*, *GaSMXL1*, *GaSMXL7*在根、茎、叶和胚珠中优势表达.这些优势表达基因多数位于*SMXL-I*, *SMXL-II*亚家族,推测位于*SMXL-I*, *SMXL-II*亚家族中的多数基因可能是陆地棉或亚洲棉中在根、茎、叶和胚珠发育中发挥功能的关键基因.胁迫表达谱分析显示: *SMXL-II*和*SMXL-III*亚家族的基因(*GhSMXL8-1-A*, *GhSMXL8-1-D*, *GhSMXL8-2-A*, *GhSMXL8-2-D*, *GhSMXL3-1-A*, *GhSMXL3-1-D*, *GhSMXL3-2-A*, *GhSMXL3-3-A*和*GhSMXL3-3-D*)在盐、旱、低温、高温不同的胁迫模式下具有不同的响应方式,同一基因可能同时响应两种不同胁迫方式,例如*GhSMXL8-1-A*响应低温胁迫同时也响应盐胁迫,推测这些基因可能调控陆地棉非生物胁迫的响应.本研究为此后深入研究棉花*SMXL*基因功能奠定了基础.

关键词 棉花, *SMXL*, 进化特征, 组织表达, 非生物胁迫

棉花是世界上重要的经济作物和重要的天然纤维来源之一.棉纤维是天然纤维的主要来源^[1],也是纺织

工业的重要原料.随着全球气候变化和环境破坏,棉花的生长发育受到各种环境压力的影响,如干旱、盐

引用格式: 贾婷婷, 竺丽萍, 肖光辉, 等. 棉花*SMXL*基因家族的全基因组分析及表达分析. 中国科学: 生命科学, 2022, 52: 1868–1882

Jia T T, Zhu L P, Xiao G H, et al. Genome-wide identification and expression analysis of the *SMXL* gene family in cotton (in Chinese). Sci Sin Vitae, 2022, 52: 1868–1882, doi: 10.1360/SSV-2021-0023

碱、低温和高温等逆境胁迫^[2]。这些非生物胁迫通过影响植物生理和代谢反应来影响植物生长发育, 对作物造成不可逆的损伤甚至死亡, 最终影响作物产量。

独脚金内酯(strigolactones, SLs)是类胡萝卜素衍生的植物激素分子, 调节植物生长发育^[3]。有研究表明, 不同SMXL直系同源物具有独特的生理作用^[4-9]。在拟南芥中, SMXL1和SMXL2主要调控拟南芥根和根毛发育^[6,10], SMXL6, SMXL7和SMXL8是SMAX1-LIKE(SMXL)家族成员, 调节拟南芥枝条分支、叶片形状和侧根发育^[5,6,8,11], 而SMXL3, SMXL4和SMXL5是拟南芥韧皮部细胞形成的关键调节因子^[9,12]。在水稻中SMXL6(MAX2-LIKE6抑制剂), SMXL7和SMXL8是DWARF53(OsD53)的直系同源物, 依赖 α/β 水解酶超家族蛋白D14(α/β -hydrolase superfamily protein D14)和F-box蛋白MAX2(F-box protein MAX2)泛素化并被26S蛋白酶降解, 从而充当SL信号转导的阻遏物^[4,10,13]。SLs在根中促进初级根分生组织中细胞数量的增加, 抑制侧根和不定根的形成, 并促进根毛伸长^[14-16]; 在茎中调节茎的分支、叶片衰老及茎伸长^[17-19]; 在调节种子萌发、幼苗发育及干旱响应方面起到一定作用^[3,20,21]。与生长素、茉莉酸和赤霉素的信号机制类似, SL信号转导是基于激素激活的蛋白水解发挥作用。被子植物对SL的感知通过D14发挥作用。D14具有严格保守的Ser-His-Asp催化三联体, 是SL水解和信号传递功能所必需的^[22-24]。SL反应需要F-box蛋白MAX2/D3。GR24(SLs类似物)促进D14和MAX2/D3之间的相互作用, MAX2/D3增强了D14不稳定性^[12,25,26]。SLs在植物发育过程中最具代表性的作用是通过典型的D14-MAX2-D53信号机制^[13-20]。MAX2作用的直接靶点是MAX2抑制基因(*SMAX1*)或D53蛋白家族成员, 而D53蛋白家族与ClpB/热休克蛋白(HSP100)有关。水稻D53蛋白突变后具有*d3*和*d14*突变体表型, 包括分蘖数增多和SLs含量增多^[13,20]。随后在拟南芥中的研究表明, D53同源基因*SMAX1-LIKE6*(*SMXL6*), *SMXL7*和*SMXL8*的三个功能缺失突变体完全抑制了*max2*表型中所有SL相关表型^[4,5]。*SMXL2*也参与了*SMAX1*调控拟南芥幼苗的生长^[6]。SLs还参与拟南芥和水稻的干旱响应调节^[20,21,22], GR24已被证明可以增强野生型植物的耐旱性^[20]。此外, *SMXL6*, *SMXL7*和*SMXL8*参与的胁迫应答和SL途径相关, 调节植物对干旱胁迫的响应和适应。

四倍体陆地棉(*Gossypium hirsutum*, (AD)₁)和海岛棉(*Gossypium barbadense*, (AD)₂)大约起源于100万~150万年前^[27]。其中四倍体中的A亚基因组的实际供体既不是A₁也不是A₂, 而是可能已灭绝的共同祖先A₀。A₀与D₅基因组大约在160万年前形成异源四倍体, 随后, A₀基因组在大约70万年前分化形成现存的A₁和A₂基因组^[28]。现今三个重要的二倍体棉种草棉(*Gossypium herbaceum*)^[28]、雷蒙德氏棉(*Gossypium raimondii*)^[29]和亚洲棉(*Gossypium arboreum*)^[30]以及两个四倍体棉种陆地棉^[31]和海岛棉^[31]全基因组数据库的完成, 为棉花功能基因组研究提供了全新的平台。根据已有研究表明, *SMXL*基因家族可能在非生物胁迫反应中发挥重要作用, 但其在棉花中相关的研究尚未见报道。本研究通过对棉花*SMXL*家族基因进行全基因组关联分析, 共鉴定到48个棉花*SMXL*家族成员。进化分析显示, 棉花*SMXL*被分为4个亚家族, 同一簇的*SMXL*成员有着相似的基因结构、蛋白结构域。此外, 本团队对陆地棉和亚洲棉不同组织、不同时期纤维转录组数据进行分析, 发现在亚洲棉中*GaSMXL8-1*, *GaSMXL1*, *GaSMXL7*主要参与胚珠、根、茎和叶的发育, 在陆地棉中*GhSMXL1-A*, *GhSMXL1-D*, *GhSMXL7-1-A*, *GhSMXL7-1-D*, *GhSMXL8-1-A*和*GhSMXL8-1-D*主要参与胚珠、纤维、根、茎和叶生长发育。此外, 对陆地棉的胁迫表达谱分析发现, *GhSMXL8-1-A*, *GhSMXL8-1-D*, *GhSMXL8-2-A*, *GhSMXL8-2-D*, *GhSMXL3-1-A*, *GhSMXL3-1-D*, *GhSMXL3-2-A*, *GhSMXL3-3-A*和*GhSMXL3-3-D*对非生物胁迫具有调控作用。本研究将为进一步研究棉花*SMXL*的功能奠定基础。

1 材料与方法

1.1 *SMXL*基因家族鉴定及系统发育分析

为了鉴定*SMXL*基因家族成员, 利用拟南芥(*Arabidopsis thaliana*)*SMXL*蛋白序列检索8个物种的基因组数据, 根据序列相似性获取这8个物种的*SMXL*序列。其中雷蒙德氏棉^[29]和亚洲棉^[30]基因组数据来源于MagenDB网站^[32](<http://magen.whu.edu.cn>), 陆地棉^[31]和海岛棉^[31]数据来源CottonFGD(<https://cottonfgd.org/>), 其余四种基因组文件下载于ePhytozome 2.1数据库(<http://www.phytozome.net>), 包括可可(*Theoroma*

cacao)、水稻(*Oryza sativa*)、二穗短柄草(*Brachypodium distachyon*)和三角叶杨(*Populus trichocarpa*)。利用BLAST软件中的Blastp程序(<https://blast.ncbi.nlm.nih.gov/Blast.cgi>),以拟南芥SMXL蛋白序列为查询序列,对8个物种的蛋白序列进行Blastp搜索, e 值设为 1.0×10^{-5} 。同时手动删除冗余转录本,鉴定四种棉花及其他四个物种SMXL家族成员。

进化树分析首先使用ClustalW 2.1软件对包括拟南芥在内的9个物种SMXL蛋白序列进行多序列比对,之后用MEGA 7软件对比对后的蛋白序列构建无根进化树,采用邻接法,设置循环数为1000次。

1.2 棉花SMXL基因结构和蛋白保守结构域分析

从陆地棉、海岛棉、亚洲棉和雷蒙德氏棉基因组注释文件中获取SMXL基因的结构信息,并使用基因结构显示在线服务器(Gene Structure Display Server, GSDS)(<http://gsds.cbi.pku.edu.cn/>)^[33],将获取到的棉花SMXL基因结构信息置入在线服务器,对SMXL基因外显子/内含子结构进行可视化分析。

将棉花SMXL成员蛋白序列置入在线程序Pfam(<http://pfam.xfam.org/>),分析候选基因的保守结构域。同时使用TBtools^[34](<https://github.com/CJ-Chen/TBtools>)对SMXL蛋白结构域进行可视化分析。使用在线ExPASy网站(http://web.expasy.org/compute_pi/)计算SMXL蛋白等电点(isoelectric point, IP)和分子量(molecular weight, Mw)。

1.3 棉花SMXL染色体定位和共线性分析

利用陆地棉、海岛棉、亚洲棉和雷蒙德氏棉注释文件,获取棉花SMXL家族所有成员的位置信息及所在染色体的全长信息,利用TBtools^[34](<https://github.com/CJ-Chen/TBtools>)进行染色体定位可视化分析。

准备陆地棉、海岛棉、亚洲棉和雷蒙德氏棉注释文件和基因组蛋白数据库,利用TBtools软件^[34]对陆地棉、海岛棉、亚洲棉和雷蒙德氏棉的全基因组蛋白库进行数据库构建。利用Blastp程序对蛋白数据进行全蛋白序列比对, e 值为 1.0×10^{-5} ,生成矩阵文件,再利用MCScan软件^[35]进行全基因组共线性分析。

1.4 基因表达分析

从MagenDB网站^[32](<http://magen.whu.edu.cn>)调取

已拼接完成的亚洲棉^[36]的转录组数据,从CottonFGD(<https://cottonfgd.org/>)调取已拼接完成的陆地棉不同组织表达转录组数据及非生物胁迫表达转录组数据。以 $\log_2(\text{FPKM}+1)$ 对转录组数据进行均一化。利用R(3.3.0)软件中的pheatmap程序绘制热图,并将基因表达模式可视化。

2 结果与分析

2.1 SMXL基因家族成员的鉴定

为了阐明SMXL基因家族进化关系,本研究鉴定并分析来自9个物种的SMXL基因,其中在陆地棉、亚洲棉、海岛棉和雷蒙德氏棉中共鉴定出48个SMXL基因,分别是16个*GhSMXL*基因、16个*GbSMXL*基因、8个*GaSMXL*基因和8个*GrSMXL*基因。同时在棉花近缘物种可可(*T. cacao*)中鉴定到6个*TcSMXL*成员,鉴定出12个杨树(*P. trichocarpa*)*PtSMXL*成员,二种单子叶植物二穗短柄草(*B. distachyon*)和水稻(*O. sativa*)中分别鉴定出4个*BdSMXL*成员和9个*OsSMXL*成员(网络版附表1)。对棉花SMXL基因家族成员进行理化性质分析,发现该家族蛋白序列长度范围为781~1122个氨基酸,相对分子质量的范围为87.156~123.714 kD,理论等电点范围为6.281~8.024,平均值为7.297,表明该家族的蛋白偏碱性(表1)。

2.2 SMXL基因家族系统进化及基因结构分析

为了研究棉花SMXL基因家族在进化过程中的亲缘关系,本研究选取了5种典型的单双子叶植物——水稻、二穗短柄草、拟南芥、可可、杨树和4种棉花——陆地棉、海岛棉、雷蒙德氏棉和亚洲棉的SMXL蛋白序列构建系统进化树。系统进化分析发现,并无物种特异性SMXL分支出现,9个物种中所有SMXL基因都可归到4个亚家族:SMXL-I(SMXL1, 2),SMXL-II(SMXL6, 7, 8),SMXL-III(SMXL3)和SMXL-IV(SMXL4, 5)中,这4个亚家族分别包含15, 32, 29和11个SMXL基因(图1)。研究还发现,SMXL基因在棉花之间具有更紧密的进化关系,且和其他5个物种比,可可SMXL基因与四种棉花SMXL基因具有更近的亲缘关系,而单子叶植物水稻和二穗短柄草SMXL基因和四种棉花SMXL基因的亲缘关系最远。

为了更好地理解SMXL基因家族在不同棉花中的

表 1 棉花 *SMXL* 基因家族成员信息Table 1 Information of *SMXL* gene family members in cotton

基因名称	染色体位置	基因长度(bp)	氨基酸数量(aa)	染色体定位	等电点(IP)	蛋白质分子量(kD)
GaSMXL3-2	Chr04	2346	781	95242458~95245014	7.692	87.256
GaSMXL1	Chr05	3174	1057	748814~752504	8.021	115.686
GaSMXL7	Chr05	3294	1097	8141840~8145931	6.859	120.874
GaSMXL8-1	Chr05	3366	1121	43831373~43835849	6.425	123.610
GaSMXL3-1	Chr09	2583	860	72503416~72506219	6.833	95.317
GaSMXL8-2	Chr10	3015	1004	617219~621650	8.049	111.242
GaSMXL3-3	Chr13	2574	857	977357~979930	6.817	88.262
GaSMXL4	Chr13	2769	922	121322521~121325670	7.633	103.048
GbSMXL1-A	A05	3168	1055	702090~705771	8.024	115.392
GbSMXL7-1-A	A05	3351	1116	8339616~8343711	6.784	122.845
GbSMXL8-1-A	A05	3369	1122	42056154~42060637	6.281	123.759
GbSMXL3-2-A	A05	2346	781	104306914~104309470	7.939	87.135
GbSMXL8-2-A	A10	3018	1005	110924219~110928660	8.286	111.286
GbSMXL3-1-A	A09	2583	860	67281364~67284167	6.833	95.365
GbSMXL3-3-A	A13	2388	795	888559~891131	6.689	89.160
GbSMXL4-A	A13	2769	922	110467899~110471049	8.124	103.066
GbSMXL3-2-D	D04	2346	781	2897103~2899659	7.960	87.437
GbSMXL1-D	D05	3168	1055	702090~705771	8.024	115.392
GbSMXL7-1-D	D05	3351	1116	7368329~7372426	6.618	122.798
GbSMXL8-1-D	D05	3369	1122	33847628~33852136	6.395	123.767
GbSMXL3-1-D	D09	2583	860	42975932~42978737	6.962	95.336
GbSMXL8-2-D	D10	3006	1001	66148734~66153240	8.334	110.726
GbSMXL3-3-D	D13	2388	795	744946~747524	6.728	89.203
GbSMXL4-D	D13	2769	922	58526390~58529536	7.402	102.950
GhSMXL1-A	A05	3168	1055	862452~866132	8.021	115.517
GhSMXL7-1-A	A05	3351	1116	8074588~8078682	6.784	122.859
GhSMXL8-1-A	A05	3366	1121	41609679~41614161	6.338	123.677
GhSMXL3-2-A	A05	2346	781	107654551~107657107	7.961	87.156
GhSMXL3-1-A	A09	2583	860	71312310~71315113	6.833	95.330
GhSMXL8-2-A	A10	3018	1005	114367315~114371756	8.239	111.428
GhSMXL3-3-A	A13	2388	795	957699~960272	6.751	89.172
GhSMXL4-A	A13	2769	922	107919772~107922922	8.124	103.048
GhSMXL3-2-D	D04	2346	781	3041228~3043784	7.965	87.403
GhSMXL1-D	D05	3168	1055	705343~709022	8.024	115.441
GhSMXL7-1-D	D05	3351	1116	7102903~7106979	6.618	122.772
GhSMXL8-1-D	D05	3366	1121	33448747~33453270	6.395	123.714
GhSMXL3-1-D	D09	2583	860	41468864~41471668	6.962	95.162
GhSMXL8-2-D	D10	3018	1005	66217138~66221651	7.945	111.045
GhSMXL3-3-D	D13	2388	795	837492~840070	6.728	89.203
GhSMXL4-D	D13	2769	922	61883371~61886519	7.518	102.807

(表1续1)

基因名称	染色体位置	基因长度(bp)	氨基酸数量(aa)	染色体定位	等电点(IP)	蛋白质分子量(kD)
GrSMXL3-1	Chr06	2583	860	40934216~40937669	6.962	95.293
GrSMXL1	Chr09	3168	1055	641385~645653	8.126	115.453
GrSMXL7	Chr09	3351	1116	6656769~6661861	6.618	122.716
GrSMXL8-1	Chr09	3366	1121	31564141~31569124	6.507	123.658
GrSMXL8-2	Chr11	3270	1089	62065572~62070199	7.65	111.174
GrSMXL3-2	Chr12	2346	781	2956774~2959871	7.574	87.250
GrSMXL3-3	Chr13	2388	795	724265~727292	6.665	89.293
GrSMXL4	Chr13	2769	922	56184994~56188589	7.402	103.055

图 1 来自9个物种的87个SMXL基因家族成员的系统进化树。SMXL序列选自陆地棉(*G. hirsutum*, Gh)、海岛棉(*G. barbadense*, Gb)、亚洲棉(*G. arboreum*, Ga)、雷蒙德氏棉(*G. raimondii*, Gr)、拟南芥(*A. thaliana*, At)、可可(*T. cacao*, Tc)、杨树(*P. trichocarpa*, Pt)、二穗短柄草(*B. distachyon*, Bd)和水稻(*O. sativa*, Os)

Figure 1 Phylogenetic tree of 87 SMXL genes from 9 species. The SMXL sequences were selected from *G. hirsutum* (Gh), *G. barbadense* (Gb), *G. arboreum* (Ga), *G. raimondii* (Gr), and *A. thaliana* (At), *T. cacao* (Tc), *P. trichocarpa* (Pt), *B. distachyon* (Bd), and *O. sativa* (Os)

进化关系, 本研究对四种棉花SMXL蛋白结构域构建无根系统发育树. 分析表明, 棉花SMXL蛋白结构域可分为4个亚簇, 每个亚簇由6个成员组成, 分别包含1个亚洲棉(A基因组)基因、1个雷蒙德氏棉(D基因组)基因、2个陆地棉基因和2个海岛棉基因. 结果表明, SMXL基因家族在二倍体杂交形成四倍体的进化过程中没有发生基因扩张和基因丢失(图2A, 网络版附表2).

为了分析四种棉花SMXL基因序列的结构特征, 利用GSDS软件对SMXL基因家族的外显子/内含子结构进行可视化分析, 发现不同亚组间具有相似的基因结构(图2B). 多数SMXL基因包含3个外显子, 2~3个内含子. 这些结果表明, 棉花SMXL基因家族在进化上是非常保守的.

2.3 棉花SMXL基因家族成员蛋白结构域分析

为研究棉花SMXL家族蛋白序列的结构特征, 本研究提取了四种棉花SMXL家族蛋白序列, 使用在线程序Pfam(<http://pfam.xfam.org>)分析蛋白结构域, 利用Tbtools软件进行可视化分析(图3), 发现SMXL基因家族主要包含Clp_N与AAA_2两个结构域. Clp_N是Clp蛋白酶氨基端结构域^[37], AAA_2是ATP酶中AAA结构域, 在细胞中起着多种作用, 包括细胞周期调控、蛋白质水解和解聚等^[38]. 由图3可知, 大部分SMXL家族蛋白含有1~2个Clp_N结构域; 同时, 还发现AAA_2结构域是SMXL-II亚家族(SMXL6, 7, 8)所特有的.

2.4 SMXL基因家族成员分布、共线性及基因复制分析

根据基因组注释文件中获取的棉花SMXL基因家族所在染色体位置信息, 对SMXL基因家族成员在染色体分布进行可视化分析, 获得染色体定位图(图4). 在亚洲棉中8个GaSMXL基因主要分布在5条染色体上, 分别是Chr04, Chr05, Chr09, Chr10和Chr13, 其余染色体均无GaSMXL基因分布. 其中, Chr05染色体含有3个GaSMXL基因(GaSMXL1, GaSMXL7, GaSMXL8-1), Chr13含有2个GaSMXL基因(GaSMXL3-3, GaSMXL4), Chr4, Chr9, Chr10染色体分别含有1个GaSMXL基因(分别为GaSMXL3-2, GaSMXL3-1, GaSMXL8-2). 雷蒙德氏棉GrSMXL同样分布在5条染色体上, 分别为Chr06, Chr09, Chr11, Chr12和Chr13. 其中, Chr09染色

图2 SMXL成员在四个棉花中亲缘关系、基因结构分析. A: 四种棉花SMXL基因系统发育树. Gh, 陆地棉; Gb, 海岛棉; Gr, 雷蒙德氏棉; Ga, 亚洲棉. B: 棉花SMXL基因结构分析. 绿色表示外显子(exon)区域, 横线表示内含子(intron)区域

Figure 2 Genetic relationship and gene structure analysis of SMXL members in four cottons. A: Phylogenetic trees of four cotton SMXL genes. Gh, *G. hirsutum*; Gb, *G. barbadense*; Gr, *G. raimondii*; Ga, *G. arboreum*. B: Structural analysis of SMXL gene in cotton. The green is the exon region, and the horizontal line is the intron region

体含有3个GrSMXL基因(GrSMXL1, GrSMXL7, GrSMXL8-1), Chr13染色体含有2个GrSMXL基因(GrSMXL3-3, GrSMXL4), Chr06, Chr11, Chr12染色体

图 3 SMXL家族成员蛋白结构域分析. A: 四种棉花SMXL基因系统发育树; B: 蛋白结构域分布. 绿色代表Clp_N结构域, 黄色代表AAA_2结构域

Figure 3 Structural domain analysis of SMXL gene family members. A: Phylogenetic trees of SMXL genes in cotton; B: distribution of protein domains. Green represents the Clp_N domain and yellow represents the AAA_2 domain

分别含有1个*GrSMXL*基因(分别为*GrSMXL3-1*, *GrSMXL3-2*, *GrSMXL3-2*). 在陆地棉中, 共含有16个*GhSMXL*基因. 8个*GhSMXL*家族成员分布在A亚基因

组的4条染色体上, 分别是A05, A09, A10和A13, 其中A05染色体含有4个*GhSMXL*基因(*GhSMXL1-A*, *GhSMXL7-1-A*, *GhSMXL8-1-A*, *GhSMXL3-2-A*), A13染

图 4 SMXL家族成员在陆地棉(*G. hirsutum*, Gh)、海岛棉(*G. barbadense*, Gb)、亚洲棉(*G. arboreum*, Ga)和雷蒙德氏棉(*G. raimondii*, Gr)中染色体定位分析

Figure 4 Chromosomal localization analysis of SMXL gene family members in *G. hirsutum* (Gh), *G. barbadense* (Gb), *G. arboreum* (Ga), and *G. raimondii* (Gr)

染色体含有2个*GhSMXL*基因(*GhSMXL4-A*, *GhSMXL3-3-A*), A09和A10分别含有1个*GhSMXL*基因(分别为*GhSMXL3-1-A*, *GhSMXL8-2-A*)。8个*GhSMXL*家族成员分布在D亚基因组的5条染色体上, 分别为D04, D05, D09, D10和D13, 其中D05染色体含有3个*GhSMXL*基因(*GhSMXL1-D*, *GhSMXL7-1-D*, *GhSMXL8-1-D*), D13含有2个*GhSMXL*基因(*GhSMXL3-3-D*, *GhSMXL4-D*), D04, D09, D10染色体分别含有1个*GhSMXL*基因(分别为*GhSMXL3-2-D*, *GhSMXL3-1-D*, *GhSMXL8-2-D*)。在海岛棉中, 16个*GbSMXL*基因分布在9条染色体上, 与陆地棉具有相似的基因分布。8个*GbSMXL*家族成员分布在A亚基因组的4条染色体上, 分别是A05, A09, A10和A13, 其中A05染色体含有4个*GbSMXL*基因(*GbSMXL1-A*, *GbSMXL7-1-A*, *GbSMXL8-1-A*, *GbSMXL3-2-A*), A13染色体仅含有2个*GbSMXL*基因(*GbSMXL3-3-A*, *GbSMXL4-A*), A09和A10分别含有1个*GbSMXL*基因(分别为*GbSMXL3-1-A*, *GbSMXL8-2-A*); 8个*GbSMXL*家族成员分布在D亚基因组的5条染色体上, 分别是D04, D05, D09, D10和D13, 其中D05与D13分别包含3个*GbSMXL*(*GbSMXL1-D*, *GbSMXL7-1-D*, *GbSMXL8-1-D*)和2个*GbSMXL*基因(*GbSMXL3-3-D*, *GbSMXL4-D*), D09和D10染色体分别含有1个*GbSMXL*基因(分别为*GbSMXL3-1-D*, *GbSMXL8-2-D*)。

串联和片段DNA复制是驱动基因家族扩张的主要驱动力, 并促进整个基因组的进化。为了探究SMXL基因扩张的潜在驱动力, 本研究分析了SMXL基因的复制事件, 发现片段复制可能是SMXL基因在棉花中扩张的主要原因(网络版附表3)。为了观察二倍体与四倍体棉花之间进化关系, 分析了二倍体与四倍体棉花之间的共线性, 结果表明, 16个陆地棉*GhSMXL*基因与8个亚洲棉*GaSMXL*基因存在共线关系, 与8个雷蒙德氏棉*GrSMXL*基因同样存在共线关系(图5A)。海岛棉16个*GbSMXL*基因与8个亚洲棉*GaSMXL*基因存在共线关系, 与8个雷蒙德氏棉*GrSMXL*基因同样存在共线关系(图5B)。共线性分析表明, 来自二倍体棉种(亚洲棉或雷蒙德氏棉)的每个SMXL基因在其对应四倍体棉种(陆地棉或海岛棉)的At和Dt亚基因组中都有两个同源基因。

2.5 棉花SMXL基因家族不同组织表达模式及响应非生物胁迫表达模式分析

为研究陆地棉*GhSMXL*成员在棉花生长发育过程

中的作用, 本研究对*GhSMXL*与*GbSMXL*家族成员的表达模式进行分析。如图6A所示, 陆地棉中棉花SMXL成员的组织表达谱显示, 所有*GhSMXL*成员在根和茎中表达量都很高, 且*GhSMXL1-A*, *GhSMXL1-D*, *GhSMXL7-1-A*, *GhSMXL7-1-D*, *GhSMXL8-1-A*, *GhSMXL8-1-D*在根、茎、叶、纤维及胚珠中均有较高表达, 同时还发现*GhSMXL3-3-A*, *GhSMXL3-3-D*, *GhSMXL3-2-A*, *GhSMXL3-2-D*, *GhSMXL8-2-A*, *GhSMXL4-A*, *GhSMXL4-D*随着纤维的起始及伸长表达量逐渐降低。这些结果表明, 陆地棉SMXL在根和茎的发育中起重要作用, 且*GhSMXL3-3-A*, *GhSMXL3-3-D*, *GhSMXL3-2-A*, *GhSMXL3-2-D*, *GhSMXL8-2-A*, *GhSMXL4-A*, *GhSMXL4-D*可能与胚珠的发育和纤维的起始过程相关。亚洲棉*GaSMXL*的表达谱显示, *GaSMXL8-1*, *GaSMXL1*, *GaSMXL7*在胚珠、根、茎和叶中表达较高, *GaSMXL1*在10DPA纤维中表达较高, *GaSMXL3-2*在10DPA胚珠中表达较高, 而*GaSMXL3-1*, *GaSMXL3-3*, *GaSMXL8-2*和*GaSMXL4*在胚珠、纤维、根、茎和叶中表达均不高(图6B)。这些优势表达基因多数位于SMXL-I和SMXL-II亚家族, 推测位于SMXL-I和SMXL-II亚家族中的多数基因可能是陆地棉或亚洲棉中在根、茎、叶和胚珠发育中发挥功能的关键基因。

为了探究陆地棉*GhSMXL*基因家族在非生物胁迫中的潜在作用, 本研究分析了陆地棉*GhSMXL*基因在盐、旱、低温、高温胁迫下的表达模式(图7)。结果显示, 和对照组相比, 在低温处理条件下, *GhSMXL8-1-A*, *GhSMXL8-1-D*和*GhSMXL8-2-A*在处理1和3 h后表达显著增高, *GhSMXL8-2-D*和*GhSMXL3-2-A*在处理3 h后表达显著增高, *GhSMXL3-1-A*和*GhSMXL3-1-D*在处理6 h后表达显著增高, *GhSMXL3-3-D*在处理1 h后表达显著增高; 在高温处理条件下, *GhSMXL8-2-A*和*GhSMXL3-3-D*在处理1 h后表达显著增高, *GhSMXL3-2-A*在处理1和3 h后表达显著增高; 在盐胁迫处理条件下, *GhSMXL8-1-A*在处理3和6 h后表达显著增高, *GhSMXL8-1-D*在处理1和3 h后表达显著增高, *GhSMXL3-1-A*在处理6 h后表达显著增高, *GhSMXL3-2-A*在处理3 h后表达显著增高, *GhSMXL3-3-D*在处理1 h后表达显著增高; 在干旱胁迫条件下, *GhSMXL8-1-D*在处理1 h后表达显著增高, *GhSMXL3-1-A*, *GhSMXL3-1-D*和*GhSMXL3-3-A*在处理6 h后表达显著

图 5 棉花SMXL基因共线性分析. A: 陆地棉、亚洲棉和雷蒙德氏棉SMXL共线性分析; B: 海岛棉、亚洲棉和雷蒙德氏棉SMXL共线性分析

Figure 5 Collinearity analysis of SMXL genes. A: Collinearity analysis of SMXL genes in *G. hirsutum*, *G. arboreum*, and *G. raimondii*; B: collinearity analysis of SMXL genes in *G. barbadense*, *G. arboreum*, and *G. raimondii*

增高. 根据系统发育树分析发现(图2A), 这些基因(*GhSMXL8-1-A*, *GhSMXL8-1-D*, *GhSMXL8-2-A*, *GhSMXL8-2-D*, *GhSMXL3-1-A*, *GhSMXL3-1-D*, *GhSMXL3-2-A*, *GhSMXL3-3-A*和*GhSMXL3-3-D*)主要位于SMXL-II和SMXL-III亚家族, 推测位于SMXL-II和SMXL-III亚家族基因可能调控陆地棉非生物胁迫的响应. 但是随着胁迫处理时间的增长, 多数基因的表达显著降低, 如在低温胁迫6 h时, *GhSMXL8-1-D*和*GhSMXL4-D*的表达显著下降, 低温胁迫12 h时, *GhSMXL8-1-A*, *GhSMXL8-1-D*, *GhSMXL8-2-A*, *GhSMXL8-2-D*, *GhSMXL3-1-A*, *GhSMXL3-1-D*, *GhSMXL3-2-A*, *GhSMXL3-3-D*, *GhSMXL3-2-D*,

*GhSMXL4-D*和*GhSMXL7-1-D*表达都显著降低, 其余处理, 如高温处理和干旱均出现随着处理时间增长, 多数基因表达显著降低. 推测长时间的胁迫处理可能对植物造成损伤进而对SMXL家族基因的表达起到抑制作用.

3 讨论

SMXL在植物生长发育和生理代谢过程的各个方面发挥着重要作用^[10]. 目前大部分关于SMXL的研究都是以拟南芥和水稻为研究对象, 棉花SMXL基因功能或家族分析并未见报道. 本研究以四种棉花为研究

图 6 SMXL基因在不同组织的表达分析. A: 16个*GhSMXL*基因在根、茎、叶以及纤维发育不同时期的表达水平 heatmap; B: 8个*GaSMXL*基因在根、茎、叶、胚珠以及纤维发育不同时期的表达水平 heatmap. 使用 $\text{Log}_2(\text{FPKM}+1)$ 方式对转录组数据进行均一化. 右上角显示SMXL家族基因在不同组织中的表达量变化范围, 红色越深代表表达量高, 蓝色越深代表表达量越低

Figure 6 Expression analysis of SMXL genes in different tissues. A: Expression levels of 16 *GhSMXL* genes at root, stem, leaf and fiber development. The data is presented as a heat map. B: Expression levels of 8 *GaSMXL* genes at root, stem, leaf, ovule and fiber development. $\text{Log}_2(\text{FPKM}+1)$ method was used to homogenize transcriptome data. The upper right corner shows the variation range of the expression of SMXL family genes in different tissues. The darker red means the higher expression level, and the darker blue means the lower expression level

对象, 系统分析了棉花SMXL家族的序列结构特征和进化关系以及陆地棉*GhSMXL*成员的组织表达谱及胁迫处理表达谱. 棉花SMXL成员蛋白序列分析显示, 75%的SMXL家族成员具有Clp-N基序, 这一蛋白基序是核苷三磷酸水解酶超家族的特征, 可能负责SL感知的信号成分^[6,13,20,22]. SMXL-II (SMXL6, 7, 8)亚家族还含有AAA_2特异结构域(图3), 主要调控细胞周期、蛋白质水解等^[38].

系统进化分析发现, 棉花SMXL成员在进化树中被分为四个亚组: SMXL-I (SMXL1, 2), SMXL-II (SMXL6, 7, 8), SMXL-III (SMXL3), SMXL-IV (SMXL4, 5)(图1). 其他物种SMXL, 如拟南芥和水稻等, 在进化过程中也被分为四组^[5,39]. 本研究结果和其他物种的SMXL分组一致, 说明棉花中没有特异进化的SMXL成员. 对棉花SMXL成员的基因结构分析及

蛋白保守结构域分析发现, 位于相同亚组的SMXL家族成员具有相似的外显子-内含子结构(图2)及相似的蛋白保守结构域(图3), 推测处于同一亚组的*GhSMXL*成员可能存在相似的生物学功能. 在陆地棉的组织表达谱中, *GhSMXL1-A*, *GhSMXL1-D*, *GhSMXL7-1-A*, *GhSMXL7-1-D*, *GhSMXL8-1-A*, *GhSMXL8-1-D*在根、茎、叶、纤维及胚珠中均有较高表达, 在亚洲棉的组织表达谱中, *GaSMXL8-1*, *GaSMXL1*, *GaSMXL7*在胚珠、根、茎和叶中表达较高, 这些优势表达基因主要位于SMXL-I和SMXL-II亚家族, 可进一步推测同一亚组间可能发挥相似的功能.

环境的波动促使生物不断进化以适应环境的变化^[40]. 物种进化过程中多倍体化是一种主要的进化方式^[41]. 四倍体陆地棉主要是由两种二倍体棉花亲本杂交形成, 陆地棉基因在进化过程中随着环境的变化也

图 7 陆地棉*GhSMXL*成员在低温胁迫、高温胁迫、盐胁迫以及干旱胁迫条件下基因表达分析. 对于每个处理, 从左往右分别为处理1, 3, 6, 12 h. 使用Log₂(FPKM+1)方式对转录组数据进行均一化. 右上角浅蓝色表示对照组(CK), 橙色表示低温处理, 灰色表示高温处理, 黄色表示盐处理, 深蓝色表示干旱处理. *, $P < 0.05$; **, $P < 0.01$

Figure 7 Gene expression analysis of *GhSMXL* genes under cold, hot, salt, and drought stress (PEG) in *G. hirsutum*. For each treatment, 1, 3, 6, and 12 h are respectively processed from left to right. Log₂(FPKM+1) method was used to homogenize transcriptome data. In the upper right corner, light blue represents control (CK), orange represents cold, gray represents hot, yellow represents salt and dark blue represents drought (PEG). *, $P < 0.05$; **, $P < 0.01$

会发生一系列变化, 其中就包括基因复制. 基因复制是基因组进化的基本过程, 在基因的复制过程中会引入突变, 其结果主要有两种, 一种是由于突变而丧失功能, 另一种是适应突变而产生新的功能^[42]. 对SMXL家族的复制基因进行分析, 发现片段复制是SMXL基因在棉花中扩展的主要原因(图4). 共线性分析表明, 来自二倍体棉种(亚洲棉或雷蒙德氏棉)的每个SMXL基因

在其对应四倍体棉种(陆地棉或)的At和Dt亚基因组中都有两个同源基因, 没有发生基因的扩张, 表明在二倍体棉花向四倍体棉花进化过程中SMXL基因家族功能较为保守, 可能在棉花生长发育中起着重要作用(图5).

已有研究发现, 拟南芥SMXL-I, SMXL-II, SMXL-IV亚家族中的SMXL基因在基因结构和分子功能上是一致的, 但SMXL-III亚家族功能尚未完全确

定^[4,6,20,32]。拟南芥中SMXL- I (SMXL1, 2)亚家族分支已被证明参与Karrikins(KARs)通路, 促进种子萌发和幼苗的光形态发生^[6]。在陆地棉中SMXL- I 亚家族成员*GhSMXL1-A*和*GhSMXL1-D*在胚珠中有较高的表达, 推测其可能在胚珠发育中发挥作用。在拟南芥和水稻中, 处于进化树分支SMXL- II 的家族成员(*AtSMXL6*, 7, 8和*OsD53*)能够参与独脚金内酯通路及响应干旱胁迫^[4,20,43]。通过分析棉花SMXL基因在盐、干旱、低温

和高温胁迫下的表达模式, 发现SMXL- II, SMXL- III亚家族成员对非生物胁迫具有调控作用, 例如*GhSMXL8-1-D*和*GhSMXL3-1-A*同时响应低温胁迫、盐胁迫和干旱胁迫, *GhSMXL8-2-A*, *GhSMXL3-2-A*和*GhSMXL3-3-D*在响应低温胁迫同时也响应高温胁迫, 这些基因在棉花响应非生物胁迫中似乎扮演着正调节因子的角色。本研究结果为进一步研究SMXL基因家族在抵御非生物胁迫方面提供了理论基础。

参考文献

- Gai J Y. Monographs of Crop Breeding (Second Edition) (in Chinese). Beijing: China Agriculture Press, 2010. 415–452 [盖钧镒. 作物育种学各论(第二版). 北京: 中国农业出版社, 2010, 415–452]
- Parida A K, Dagaonkar V S, Phalak M S, et al. Alterations in photosynthetic pigments, protein and osmotic components in cotton genotypes subjected to short-term drought stress followed by recovery. *Plant Biotechnol Rep*, 2007, 1: 37–48
- Kohlen W, Charnikhova T, Liu Q, et al. Strigolactones are transported through the xylem and play a key role in shoot architectural response to phosphate deficiency in nonarbuscular mycorrhizal host *Arabidopsis*. *Plant Physiol*, 2011, 155: 974–987
- Soundappan I, Bennett T, Morffy N, et al. SMXL1-LIKE/D53 family members enable distinct max2-dependent responses to strigolactones and karrikins in *Arabidopsis*. *Plant Cell*, 2015, 27: 3143–3159
- Wang L, Wang B, Jiang L, et al. Strigolactone signaling in *Arabidopsis* regulates shoot development by targeting D53-Like SMXL repressor proteins for ubiquitination and degradation. *Plant Cell*, 2015, 27: 3128–3142
- Stanga J P, Morffy N, Nelson D C. Functional redundancy in the control of seedling growth by the karrikin signaling pathway. *Planta*, 2016, 243: 1397–1406
- Stanga J P, Smith S M, Briggs W R, et al. *SUPPRESSOR OF MORE AXILLARY GROWTH2 1* controls seed germination and seedling development in *Arabidopsis*. *Plant Physiol*, 2013, 163: 318–330
- Bennett T, Liang Y, Seale M, et al. Strigolactone regulates shoot development through a core signalling pathway. *Biol Open*, 2016, 5: 1806–1820
- Wu Y Y, Hou B H, Lee W C, et al. DCL2- and RDR6-dependent transitive silencing of *SMXL4* and *SMXL5* in *Arabidopsis dcl4* mutants causes defective phloem transport and carbohydrate over-accumulation. *Plant J*, 2017, 90: 1064–1078
- Villacéija-Aguilar J A, Hamon-Josse M, Carbonnel S, et al. SMXL1/SMXL2 regulate root and root hair development downstream of KAI2-mediated signalling in *Arabidopsis*. *PLoS Genet*, 2019, 15: e1008327
- Wallner E S, López-Salmerón V, Belevich I, et al. Strigolactone- and karrikin-independent SMXL proteins are central regulators of phloem formation. *Curr Biol*, 2017, 27: 1241–1247
- Zhou F, Lin Q, Zhu L, et al. Corrigendum: D14-SCFD3-dependent degradation of D53 regulates strigolactone signalling. *Nature*, 2016, 532: 402
- Jiang L, Liu X, Xiong G, et al. DWARF 53 acts as a repressor of strigolactone signalling in rice. *Nature*, 2013, 504: 401–405
- Agusti J, Herold S, Schwarz M, et al. Strigolactone signaling is required for auxin-dependent stimulation of secondary growth in plants. *Proc Natl Acad Sci USA*, 2011, 108: 20242–20247
- Ruyter-Spira C, Kohlen W, Charnikhova T, et al. Physiological effects of the synthetic strigolactone analog GR24 on root system architecture in *Arabidopsis*: another belowground role for strigolactones? *Plant Physiol*, 2011, 155: 721–734
- Rasmussen A, Mason M G, De Cuyper C, et al. Strigolactones suppress adventitious rooting in *Arabidopsis* and pea. *Plant Physiol*, 2012, 158: 1976–1987
- Gomez-Roldan V, Fermas S, Brewer P B, et al. Strigolactone inhibition of shoot branching. *Nature*, 2008, 455: 189–194
- Umehara M, Hanada A, Yoshida S, et al. Inhibition of shoot branching by new terpenoid plant hormones. *Nature*, 2008, 455: 195–200
- Waters M T, Gutjahr C, Bennett T, et al. Strigolactone signaling and evolution. *Annu Rev Plant Biol*, 2017, 68: 291–322
- Ha C V, Leyva-González M A, Osakabe Y, et al. Positive regulatory role of strigolactone in plant responses to drought and salt stress. *Proc Natl*

- Acad Sci USA*, 2014, 111: 851–856
- 21 Haider I, Andreo-Jimenez B, Bruno M, et al. The interaction of strigolactones with abscisic acid during the drought response in rice. *J Exp Bot*, 2018, 69: 2403–2414
 - 22 Hamiaux C, Drummond R S M, Janssen B J, et al. DAD2 is an α/β hydrolase likely to be involved in the perception of the plant branching hormone, strigolactone. *Curr Biol*, 2012, 22: 2032–2036
 - 23 Nakamura H, Xue Y L, Miyakawa T, et al. Molecular mechanism of strigolactone perception by DWARF14. *Nat Commun*, 2013, 4: 2613
 - 24 Waters M T, Scaffidi A, Moulin S L Y, et al. A *Selaginella moellendorffii* ortholog of KARRIKIN INSENSITIVE2 functions in *Arabidopsis* development but cannot mediate responses to karrikins or strigolactones. *Plant Cell*, 2015, 27: 1925–1944
 - 25 Zhao J, Wang T, Wang M, et al. DWARF3 participates in an SCF complex and associates with DWARF14 to suppress rice shoot branching. *Plant Cell Physiol*, 2014, 55: 1096–1109
 - 26 Zhao L H, Zhou X E, Yi W, et al. Destabilization of strigolactone receptor DWARF14 by binding of ligand and E3-ligase signaling effector DWARF3. *Cell Res*, 2015, 25: 1219–1236
 - 27 Endrizzi J, Turcotte E, Kohel J. Genetics, cytology and evolution of *Gossypium*. *Adv Genet*, 1985, 23: 271–375
 - 28 Huang G, Wu Z, Percy R G, et al. Genome sequence of *Gossypium herbaceum* and genome updates of *Gossypium arboreum* and *Gossypium hirsutum* provide insights into cotton A-genome evolution. *Nat Genet*, 2020, 52: 516–524
 - 29 Paterson A H, Wendel J F, Gundlach H, et al. Repeated polyploidization of *Gossypium* genomes and the evolution of spinnable cotton fibres. *Nature*, 2012, 492: 423–427
 - 30 Du X, Huang G, He S, et al. Resequencing of 243 diploid cotton accessions based on an updated A genome identifies the genetic basis of key agronomic traits. *Nat Genet*, 2018, 50: 796–802
 - 31 Hu Y, Chen J, Fang L, et al. *Gossypium barbadense* and *Gossypium hirsutum* genomes provide insights into the origin and evolution of allotetraploid cotton. *Nat Genet*, 2019, 51: 739–748
 - 32 Wang D, Fan W, Guo X, et al. MaGenDB: a functional genomics hub for Malvaceae plants. *Nucleic Acids Res*, 2020, 48: D1076–D1084
 - 33 Hu B, Jin J, Guo A Y, et al. GSDS 2.0: an upgraded gene feature visualization server. *Bioinformatics*, 2015, 31: 1296–1297
 - 34 Chen C, Chen H, Zhang Y, et al. TBtools: An integrative toolkit developed for interactive analyses of big biological data. *Mol Plant*, 2020, 13: 1194–1202
 - 35 Wang Y, Tang H, Debarry J D, et al. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res*, 2012, 40: e49
 - 36 Wang K, Wang D, Zheng X, et al. Multi-strategic RNA-seq analysis reveals a high-resolution transcriptional landscape in cotton. *Nat Commun*, 2019, 10: 4714
 - 37 Zhou Y, Tao J, Yu H, et al. Hcp family proteins secreted via the type VI secretion system coordinately regulate *Escherichia coli* K1 interaction with human brain microvascular endothelial cells. *Infect Immun*, 2012, 80: 1243–1251
 - 38 Kedzierska S. Structure, function and mechanisms of action of ATPases from the AAA superfamily of proteins (in Polish). *Postepy Biochem*, 2006, 52: 330–338
 - 39 Moturu T R, Thula S, Singh R K, et al. Molecular evolution and diversification of the SMXL gene family. *J Exp Bot*, 2018, 69: 2367–2378
 - 40 Guo H, Jiao Y, Tan X, et al. Gene duplication and genetic innovation in cereal genomes. *Genome Res*, 2019, 29: 261–269
 - 41 Liu S, Liu Y, Yang X, et al. The Brassica oleracea genome reveals the asymmetrical evolution of polyploid genomes. *Nat Commun*, 2014, 5: 3930
 - 42 Lan X, Pritchard J K. Coregulation of tandem duplicate genes slows evolution of subfunctionalization in mammals. *Science*, 2016, 352: 1009–1013
 - 43 Liang Y, Ward S, Li P, et al. SMAX1-LIKE7 signals from the nucleus to regulate shoot development in *Arabidopsis* via partially EAR motif-independent mechanisms. *Plant Cell*, 2016, 28: 1581–1601

Genome-wide identification and expression analysis of the *SMXL* gene family in cotton

JIA TingTing¹, ZHU LiPing², XIAO GuangHui² & LI HongBin¹

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

² College of Life Sciences, Shaanxi Normal University, Xi'an 710119, China

Strigolactones (SLs) is a novel plant hormone, which plays an important role in plant growth and development. At present, there are still few studies on SL in cotton. Therefore, it is necessary to explore the function and mechanism of SL on cotton growth and fiber quality. *SMXLs* (*SUPPRESSOR OF MAX2 1*-likes) are key genes of the SL pathway and act as repressors of SL signal transduction. Until now, there are no studies about *SMXLs* gene in cotton. In this study, 16 *GhSMXLs*, 16 *GbSMXLs*, 8 *GaSMXLs*, and 8 *GrSMXLs* were identified in *G. raimondii*, *G. arboreum*, *G. hirsutum*, and *G. barbadense*. Phylogenetic analysis revealed that *SMXLs* were divided into four subfamilies: *SMXL*-I (*SMXL*1, 2), *SMXL*-II (*SMXL*6, 7, 8), *SMXL*-III (*SMXL*3), and *SMXL*-IV (*SMXL*4, 5). The evolutionary characteristics showed that *SMXL* members located in the same subgroup had similar gene and protein structures. Fragment replication was the main reason for *SMXL* member expansion in cotton. Tissue expression profile analysis showed that *GhSMXL1-A*, *GhSMXL1-D*, *GhSMXL7-1-A*, *GhSMXL7-1-D*, *GhSMXL8-1-A*, and *GhSMXL8-1-D* were dominant expressions in roots, stems, leaves, fibers and ovules. *GaSMXL8-1*, *GaSMXL1*, and *GaSMXL7* were mainly expressed in roots, stems, leaves and ovules. Interestingly, most of the dominant genes were located in the *SMXL*-I and *SMXL*-II subfamilies, indicating that these genes might be the key genes in root, stem, leaf and ovule development. Stress expression profile analysis showed that the *SMXL*-II and *SMXL*-III subfamily genes (*GhSMXL8-1-A*, *GhSMXL8-1-D*, *GhSMXL8-2-A*, *GhSMXL8-2-D*, *GhSMXL3-1-A*, *GhSMXL3-1-D*, *GhSMXL3-2-A*, *GhSMXL3-3-A*, and *GhSMXL3-3-D*) have different expression patterns under different salt, drought, low temperature and high temperature treatment. Furthermore, some genes respond to different stress treatment, for example, *GhSMXL8-1-A* responds to both low temperature stress and salt stress, implying that these genes may regulate the response of abiotic stress in cotton. This study lays the foundation for further study of cotton *SMXLs* gene function.

cotton, *SMXL*, evolutionary characteristics, tissue expression, abiotic stress

doi: [10.1360/SSV-2021-0023](https://doi.org/10.1360/SSV-2021-0023)