

# 高粱 SBP-box 基因家族全基因组鉴定及表达分析

常建忠<sup>1</sup>, 闫凤霞<sup>2</sup>, 乔麟轶<sup>3</sup>, 郑军<sup>3</sup>, 张福耀<sup>2</sup>, 柳青山<sup>2</sup>

1. 山西省农业科学院旱地农业研究中心, 农业部黄土高原作物基因资源和种质创制重点实验室, 太原 030031;
2. 山西省农业科学院高粱研究所, 高粱遗传与种质创新山西省重点实验室, 晋中 030600;
3. 山西省农业科学院作物科学研究所, 太原 030031

**摘要:** SQUAMOSA PROMOTER BINDING PROTEIN box (SBP-box) 基因家族编码一类绿色植物特有的转录因子, 其功能涉及作物遗传改良的许多方面, 如产量、株型、抗逆性等, 具有重要的实际应用价值。本研究通过生物信息学方法从高粱(*Sorghum bicolor* L.) 全基因组中鉴定出 18 个 SBP-box 基因, 分布于 9 条染色体上, 其中 8 个基因位于基因组重复区域。系统发育分析表明高粱 SBP-box 基因家族可分为 6 个亚家族, 其中 *SbSBP12*、*SbSBP3* 和 *SbSBP15* 分别与玉米 *ZmLG1*、*ZmTGA1* 和 *ZmUB2/3* 直系同源。基于 RNA-seq 数据的表达分析发现高粱 SBP-box 基因在花序原基中表达量最高, *SbSBP9* 和 *SbSBP17* 为花序原基特异表达基因, *SbSBP5*、*SbSBP8* 和 *SbSBP18* 等基因受外源 ABA 和 PEG 胁迫上调表达, 表明 SBP-box 基因可能参与高粱对非生物逆境的响应。本研究为高粱 SBP-box 家族重要基因的克隆提供了参考, 相关基因可作为高粱遗传改良的候选基因。

**关键词:** 高粱; SBP-box 基因家族; 系统进化分析; 表达分析

## Genome-wide identification and expression analysis of SBP-box gene family in *Sorghum bicolor* L.

Jianzhong Chang<sup>1</sup>, Fengxia Yan<sup>2</sup>, Linyi Qiao<sup>3</sup>, Jun Zheng<sup>3</sup>, Fuyao Zhang<sup>2</sup>, Qingshan Liu<sup>2</sup>

1. Dryland Agriculture Research Center, Shanxi Academy of Agricultural Sciences/Key Laboratory of Crop Gene Resources and Germplasm Enhancement on Loess Plateau, Ministry of Agriculture, Taiyuan 030031, China;
2. Sorghum Institute of Shanxi Academy of Agricultural Sciences/Shanxi Key Laboratory of Sorghum Genetic and Germplasm Innovation, Jinzhong 030600, China;
3. Institute of Crop Science, Shanxi Academy of Agricultural Sciences, Taiyuan 030031, China

**Abstract:** SQUAMOSA PROMOTER BINDING PROTEIN-box (SBP-box) family genes encoding plant-specific transcription factors are involved in many aspects of crop genetic improvement such as yield, plant-type and stress-resistance.

收稿日期: 2016-01-17; 修回日期: 2016-03-28

基金项目: 现代农业产业技术体系建设专项(编号: CARS-06-01-01), 山西省财政支农项目(编号: 2014ZYFZ-09), 山西省农业科学院重点项目(编号: YZD1401)和山西省农业科学院博士研究基金项目(编号: YBSJJ1407)资助[Supported by the Earmarked Fund for Modern Agro-industry Technology Research System (No. CARS-06-01-01), Financial Supporting Agriculture Project in Shanxi Province (No. 2014ZYFZ-09), Key Project of Shanxi Academy of Agricultural Science (No. YZD1401), and Doctoral Science Foundation of Shanxi Academy of Agricultural Science (No. YBSJJ1407)]

作者简介: 常建忠, 博士, 助理研究员, 研究方向: 作物逆境分子生物学。E-mail: cjzyfx@163.com

闫凤霞, 硕士, 助理研究员, 研究方向: 高粱种质资源创新。E-mail: yfxcjz@126.com

常建忠和闫凤霞为并列第一作者。

通讯作者: 柳青山, 硕士, 研究员, 研究方向: 高粱种质资源创新。E-mail: 673673@126.com

DOI: 10.16288/j.ycz.16-008

网络出版时间: 2016/4/20 13:52:59

URI: <http://www.cnki.net/kcms/detail/11.1913.R.20160420.1352.006.html>

The SBP-box gene family have important practical applications. In this study, 18 SBP-box genes were identified from the reference genome of sorghum (*Sorghum bicolor* L.) using bioinformatics. These genes distributed on nine chromosomes while eight of them located in the segmental duplication region. Phylogenetic reconstruction resulted in six subfamilies of SBP-box genes, among which *SbSBP12*, *SbSBP3* and *SbSBP15* are orthologous to *ZmLG1*, *ZmTGA1* and *ZmUB2/3* in corn, respectively. RNA-seq data analysis indicated that *SbSBP*-box genes show the highest expression level in primordial inflorescences. Moreover, *SbSBP9* and *SbSBP17* exhibited a tissue specific expression in primordial inflorescences. The expression levels of *SbSBP5*, *SbSBP8* and *SbSBP18* were increased in response to exogenous ABA and PEG, indicating that *SbSBP*-box genes are involved in the defense response against abiotic stresses in sorghum. This research provides references for cloning important genes in *SbSBP*-box gene family. Genes identified in this study could be considered as candidate genes for genetic improvement of sorghum.

**Keywords:** *Sorghum bicolor* L.; SBP-box gene family; phylogenetic analysis; expression analysis

转录因子(Transcription factor, TF)是一类能特异地与 DNA 分子结合并通过激活或抑制下游靶基因的转录来调节基因表达的蛋白分子<sup>[1,2]</sup>。高等植物转录因子在生长发育、形态建成、次生代谢和抗逆反应等多种过程发挥着重要作用。SQUAMOSA PROMOTER BINDING PROTEIN box(SBP-box)基因也称 SQUAMOSA PROMOTER BINDING PROTEIN LIKE(SPL)基因, 编码一类绿色植物特有的转录因子。1996 年德国马普研究所 Klein 等<sup>[3]</sup>从金鱼草(*Antirrhinum majus*)花序 cDNA 表达文库中筛选得到 *SBP1* 和 *SBP2* 基因, 这两个基因通过与 *SQUAMOSA* 基因启动子区域结合, 参与调控花的早期发育。SBP-box 蛋白均含有一个约 76 个氨基酸残基长度、高度保守的 SBP 结构域, 该结构域包含 2 个锌指结构(Zinc-finger domain)<sup>[4]</sup>及高度保守的双向核定位信号(Nuclear localization signal, NLS)<sup>[5]</sup>。自发现以来, SBP-box 基因家族在拟南芥(*Arabidopsis thaliana*)<sup>[6]</sup>、水稻(*Oryza sativa*)<sup>[7]</sup>、玉米(*Zea mays*)<sup>[8]</sup>、小麦(*Triticum aestivum*)<sup>[9]</sup>和番茄(*Solanum lycopersicum*)<sup>[10]</sup>等植物中均有报道。SBP-box 转录因子的功能涉及植物生长发育的许多重要过程, 包括植物胚胎发生<sup>[11]</sup>、叶片发育<sup>[12,13]</sup>、间隔期长度<sup>[14]</sup>、营养生长到生殖生长的更替<sup>[15,16]</sup>、育性维持<sup>[17]</sup>、次生产物合成<sup>[18]</sup>。SBP-box 基因家族还参与植物的胁迫应答过程, 如衣藻(*C. reinhardtii*)SBP 转录因子 CRR1 在维持其体内铜稳态平衡过程中发挥重要作用<sup>[19]</sup>。在作物育种方面, SBP-box 基因家族的许多功能都具有重要的利用价值, 如 *ZmLG1* 控制玉米叶片中叶舌的发育, 其突变

体叶片直立、株型紧凑, 是作物株型育种的重要种质资源; *OsSPL14* 控制水稻穗分支和分蘖, 在水稻生殖生长阶段过表达 *OsSPL14* 可促进穗分支, 提高籽粒产量<sup>[7]</sup>。miRNAs 是真核生物中一类内源性非编码 RNA, 作为一种转录后调控因子在基因表达调控中发挥重要作用<sup>[20]</sup>。研究表明 miR156/157 靶基因就是 SBP-box 基因, 如拟南芥 *AtSPL3* 能被 miR156 识别, 并依靠降解或翻译抑制来调控基因表达<sup>[21,22]</sup>。过半数的 SBP-box 基因受 miR156/157 调控, 如水稻 19 个 SBP-box 基因中有 11 个基因含 miR156 识别位点<sup>[23]</sup>, 番茄 15 个 SBP-box 基因中有 10 个基因具有 miR156 识别位点<sup>[10]</sup>。Arazi 等<sup>[24]</sup>研究发现小立陶宛蕨(*Physcomitrella patens*)中也存在 miR156 和 SBP-box 基因的调控关系, 可见 miR156/157 与 SBP-box 基因这种特异关系存在已久。

高粱(*Sorghum bicolor* L.)是世界第 5 大禾谷类作物, 具有抗旱、耐盐碱和耐贫瘠等多重抗性, 是我国重要的旱地粮食作物。高粱基因组测序工作于 2009 年顺利完成<sup>[25]</sup>, 这使得从全基因组水平揭示高粱重要基因家族的功能成为可能。SBP-box 基因在植物株型、产量、抗逆性等方面发挥重要作用, 因此分离鉴定 SBP-box 基因家族中的重要基因可以为作物遗传改良提供候选基因。本研究采用生物信息学方法分离了高粱 SBP-box 基因家族, 分析了该家族成员的序列特征、染色体位置分布、基因结构等, 并利用公共数据库中的 RNA-seq 数据对该家族成员在不同组织和环境下的表达模式进行了研究, 为高粱 SBP-box 基因家族的深入研究及重要基因的克隆提供重要信息。

## 1 材料和方法

## 1.1 实验材料

高粱 SBP-box 基因序列、蛋白、CDS、mRNA 序列来自 JGI(<http://genome.jgi-psf.org/Sorbi1>)和 NCBI(<http://www.ncbi.nlm.nih.gov/>)数据库 ; 水稻、拟南芥和番茄 SBP-box 蛋白序列分别来自 RGAP 数据库(<http://rice.plantbiology.msu.edu/index.shtml>)、TAIR 数据库(<http://www.Arabidopsis.org>)和 SolGenomics Network (<https://solgenomics.net/>) ; 玉米和小立陶宛蕨(*Physcomitrella patens*)SBP-box 蛋白均下载于 NCBI ; 高粱 RNA-seq 表达数据来自 NCBI GEO (Gene Expression Omnibus)和 SAR(Sequence Read Archive)数据库。

## 1.2 高粱 SBP-box 基因家族成员识别、基因结构分析和保守域检测

在 Pfam 数据库(<http://pfam.sanger.ac.uk>)下载 SBP-box 基因家族的隐马尔科夫模型文件,并用 HMM EMIT 软件获取高粱 SBP-box 基因家族的蛋白保守序列,利用该序列检索高粱 JGI(<http://genome.jgi-psf.org/Sorbi1>)和 NCBI(<http://www.ncbi.nlm.nih.gov/>)数据库, $E$  值设为  $1e^{-10}$ ,Blast 检索 the Conserved Domain Database(CCD, <http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>),参数默认,除去无 SBP-box 典型结构域的序列,获得高粱 SBP-box 基因家族序列。利用 Gene Structure Display Server(<http://gsds.cbi.pku.edu.cn/index.php>)分析高粱 SBP-box 基因结构。

### 1.3 高粱 SBP-box 蛋白基序分析和系统发育树的构建

利用 MEME 软件分析高粱 SBP-box 蛋白保守基序<sup>[26]</sup>, 基序长度范围为 10~50 个氨基酸残基, 基序最大发现数目为 10 个, 其他参数为默认值。采用 WebLogo<sup>[27]</sup>分析残基保守性, 参数默认。通过 Clustal X 对高粱、拟南芥、水稻、番茄及部分玉米和小立陶宛藓 SBP-box 蛋白进行多序列联配, 参数为默认。用 ME-GA4.1 软件采用邻接法构建系统进化树<sup>[28]</sup>, 具体参数为设置 Test of Phylogeny: Bootstrap Method; plication: 00; del/Method: Poisson model; Rates among Sites: Uniform rates; Gaps: Pairwise deletion, 其他参数默认。

#### 1.4 高粱 SBP-box 基因的定位和重复

利用高粱 SBP-box 基因序列,通过 BlastN 检索 NCBI 高粱基因组数据库(CM000760-CM00769),获得每个基因的染色体位置信息;高粱基因组重复片段信息来自 SYMAP 数据库<sup>[29]</sup>。利用 Circos 工具对高粱染色体重复片段信息及 SBP-box 基因的染色体位置进行可视化<sup>[30]</sup>。

### 1.5 高粱 SBP-box 基因启动子分析

截取高粱 SBP-box 基因家族成员起始密码子上游 2000 bp 序列作为启动子序列,将序列提交到 PLACE26.0(<http://www.dna.affrc.go.jp/PLACE/signalscan.html>)和 PlantCARE(<http://bioinformatics.psb.gent.eeoolsplantcare/html>)数据库分析启动子的顺式作用元件。

### 1.6 高粱 SBP-box 基因表达分析

高粱转录组测序(RNA-seq)数据来自 RNA-seq 数据集 1<sup>[31]</sup>(登录号为: GSE30249)和 RNA-seq 数据集 2<sup>[32]</sup>。RNA-seq 数据集 1 包括高粱叶片(Leaves)、原生花序(Primordial inflorescences)、花序(Inflorescences)、花药(Anther)、雌蕊(Pistil)、种子(Seed)、胚(Embryo)、胚乳(Endosperm); RNA-seq 数据集 2 包括不同处理条件下的根(Root)和茎(Shoot)。下载以上数据集中 RNA-seq 数据, 利用 MATLAB(R2008B)提取 SBP-box 基因的表达量(Fragments per kilobase of exon per million fragments mapped, FPKM)值, 对 SBP-box 基因表达量进行标准化, 用 R 软件中的 pheatmap 函数绘制基因表达热图。

## 2 结果与分析

### 2.1 高粱 SBP-box 基因家族的分离、染色体分布及重复

利用含 75 个氨基酸残基的 SBP-box 蛋白保守基序对高粱 JGI 数据库进行检索,同时以“SBP-box”为关键词,检索 NCBI 数据库,对得到的 SBP-box 蛋白序列进行保守结构域检测,最终获得 18 个 SBP-box 基因(表 1)。根据高粱基因组信息,将高粱 SBP-box 基因定位在高粱 9 条染色体上(第 8 号染色体除外)(图 1),高粱 SBP-box 基因在染色体上成簇分布,都位于染色体两端,其中第 2、4 和 7 号染色体上最多,

表 1 高粱 SBP-box 基因家族信息

Table 1 The information of SBP-box gene family in *Sorghum bicolor* L.

基因	长度(bp)	内含子	氨基酸	位点	基因组位置	染色体
<i>SbSBP1</i>	6225	10	969	Sb01g002530	2045981-2052205	1
<i>SbSBP2</i>	2947	2	388	Sb02g028420	63596720-63599666	2
<i>SbSBP3</i>	4359	2	464	Sb02g029300	64428707-64433065	2
<i>SbSBP4</i>	3803	1	247	Sb02g034180	68669877-68673679	2
<i>SbSBP5</i>	5856	10	800	Sb03g011920	13662259-13668114	3
<i>SbSBP6</i>	3911	2	403	Sb03g044160	71458293-71462203	3
<i>SbSBP7</i>	3672	3	481	Sb04g003175	2977779-2981450	4
<i>SbSBP8</i>	3320	3	337	Sb04g004940	4754287-4757606	4
<i>SbSBP9</i>	3331	2	445	Sb04g005180	5039031-5042361	4
<i>SbSBP10</i>	8163	2	325	Sb05g017510	43088257-43096419	5
<i>SbSBP11</i>	3526	2	400	Sb06g024630	53706223-53709748	6
<i>SbSBP12</i>	3285	2	418	Sb06g031290	59700952-59704236	6
<i>SbSBP13</i>	5708	3	1095	Sb07g026220	61420520-61426227	7
<i>SbSBP14</i>	5604	10	435	Sb07g027420	62468266-62473869	7
<i>SbSBP15</i>	3873	2	409	Sb07g027740	62690521-62694393	7
<i>SbSBP16</i>	13303	9	864	Sb09g020110	49199333-49212635	9
<i>SbSBP17</i>	2758	2	444	Sb10g026200	55599173-55601930	10
<i>SbSBP18</i>	3720	4	440	Sb10g029190	59020738-59024457	10

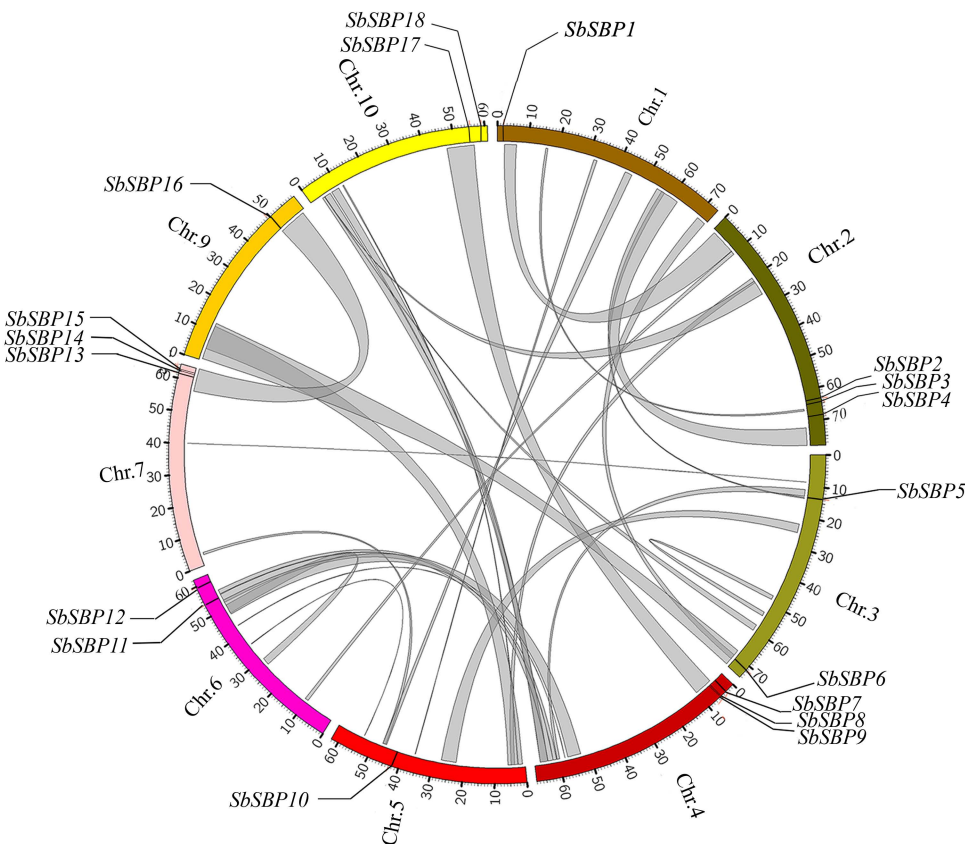


图 1 高粱 SBP-box 基因的染色体定位

Fig. 1 Chromosome location of SBP-box genes in *Sorghum bicolor* genome

均为 3 个;第 1、5、9 号染色体上最少,均为 1 个。根据基因在染色体上的位置,将高粱 18 个 SBP-box 基因分别命名为 *SbSBP1*~*SbSBP18*,其中 8 个基因位于已知的基因组重复区域,分别是 *SbSBP6*、*SbSBP8*、*SbSBP9*、*SbSBP11*、*SbSBP13*、*SbSBP14*、*SbSBP15* 和 *SbSBP17*。*SbSBP9* 和 *SbSBP17* 位于片段重复区,系旁系同源基因,其他 6 个基因在重复区只有 1 个拷贝。

## 2.2 高粱 SBP-box 基因序列分析

基于高粱 SBP-box 蛋白序列构建的进化树,将高粱 18 个 SBP-box 成员分为 6 组(图 2)。各组间基因结构差异较大,如第 I 组 *SbSBP16* 序列最长,为 13 303 bp,而第 II 组 *SbSBP17* 最短,仅 2758 bp;高粱 SBP-box 家族各成员内含子数目差异较大,其中同属第 I 组的 *SbSBP1*、*SbSBP5* 和 *SbSBP14* 内含子数目最多,为 10 个,而第 II 组 *SbSBP4* 仅 1 个内含子。各组内基因内含子数目相对保守,但也有差别,如第 III 组 *SbSBP3* 和 *SbSBP13* 分别含有 2 和 3 个内含子。miR 位点分析表明高粱 7 个 SBP-box 基因具有 miR156 识别位点

(图 3),均属于第 I 或 II 组,占整个基因家族的 38.9%,且均位于基因第 3 或 4 外显子。

## 2.3 高粱 SBP-box 蛋白保守序列分析

对高粱 18 个 SBP-box 蛋白序列进行基序扫描,共发现 8 个 Motif,基于 Motif 可将高粱 SBP-box 蛋白分为 6 组(G1~G6)(图 4A),Motif1、Motif2 和 Motif4 是各组共有基序,其中 Motif1 和 Motif2 分别是 SBP 结构域中的 2 个锌指结构, Motif1 组成为  $CX_2CX_3HX_{11}C(CCHC)$ , Motif2 组成为  $CX_4CX_{16}CX_2H(CCCH)$  或  $CX_4CX_{16}CX_2C(CCCC)$ ,在 Motif1C 端还有一个双向核定位序列(NLS),一致序列为  $K(R/K)X_{11}RRR(K/R)$ (图 4B)。高粱 SBP-box 蛋白高度保守的锌指结构和 NLS 对于特异识别和启动核基因的表达具有重要作用。除共有基序外,各组基序还具有一定特异性,如 Motif6 和 Motif7 分别仅在 G1 和 G5 中存在, Motif5 只在 G2 和 G5 中存在,此类 motif 可能与 SBP-box 基因功能的多样性有关。

## 2.4 SBP-box 蛋白的系统进化分析

本研究所有 SBP-box 蛋白可分为 6 个亚家族

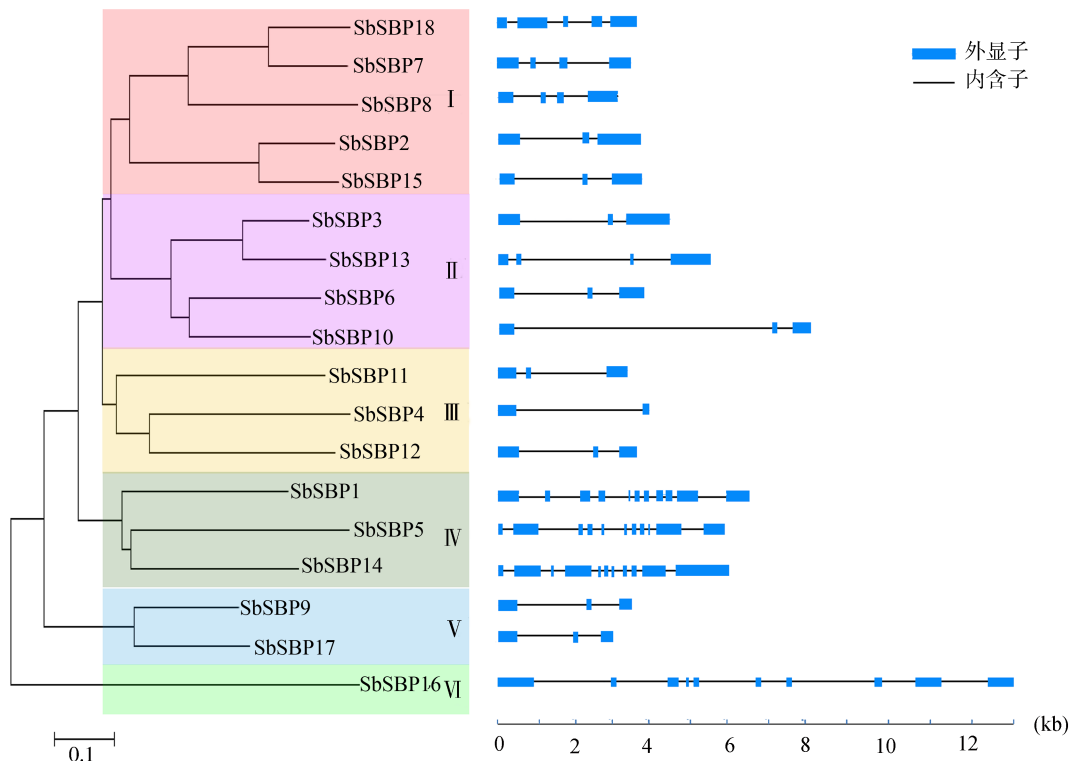


图 2 高粱 SBP-box 基因聚类及结构分析

Fig. 2 Phylogenetic and intron-exon structure analysis of SBP-box genes in *Sorghum bicolor* L.

名称	序列(5'→3')	
<i>SbSBP2</i>	GACTCCAGCTGTGCTCTCTCTTTCTGTCAACTCA-GCCAT	40
<i>SbSBP3</i>	GACTCAGATTGTGCTCTCTCTTTCTGTCAATCCCCGCCA	40
<i>SbSBP6</i>	AACGGCAACTGTGCTCTCTCTTTCTGTCAACAACCCGAC	41
<i>SbSBP7</i>	GATCTTCAGCGTGCTCTCTCTTTCTGTCAACAGTTCGG	40
<i>SbSBP10</i>	GGCCCCGAGTGTGCTCTCTCTTTCTGTCAATCGTC-GCTGC	40
<i>SbSBP13</i>	GACTCGGAATGTGCTCTCTCTTTCTGTCAAGCGCCAGCCA	40
<i>SbSBP15</i>	--CTCTAGCTGTGCTCTCTCTTTCTGTCAACTCA-GCCAT	38
<i>SbimiR156a</i>	-----GTGCTCACTCTCTTTCTGTCA	20
<i>SbimiR156b</i>	-----GTGCTCACTCTCTTTCTGTCA	20
<i>SbimiR156c</i>	-----GTGCTCACTCTCTTTCTGTCA	20
<i>SbimiR156d</i>	-----TGTGCTCTCTCTCTTTCTGTCA	21
<i>SbimiR156e</i>	-----GTGCTCGTCTCTTTCTGTCA	20
一致序列	gtgctc ctctctttctgtca	

图 3 miR156 互补序列与高粱 SBP-box 基因多序列比对

Fig. 3 Multiple sequence alignment of the miR156 complementary sequence and SbSBP-box genes

除 *SbSBP7* 和 *SbSBP13* 的 miRNA 互补序列位于第 4 外显子，其他都位于第 3 外显子。

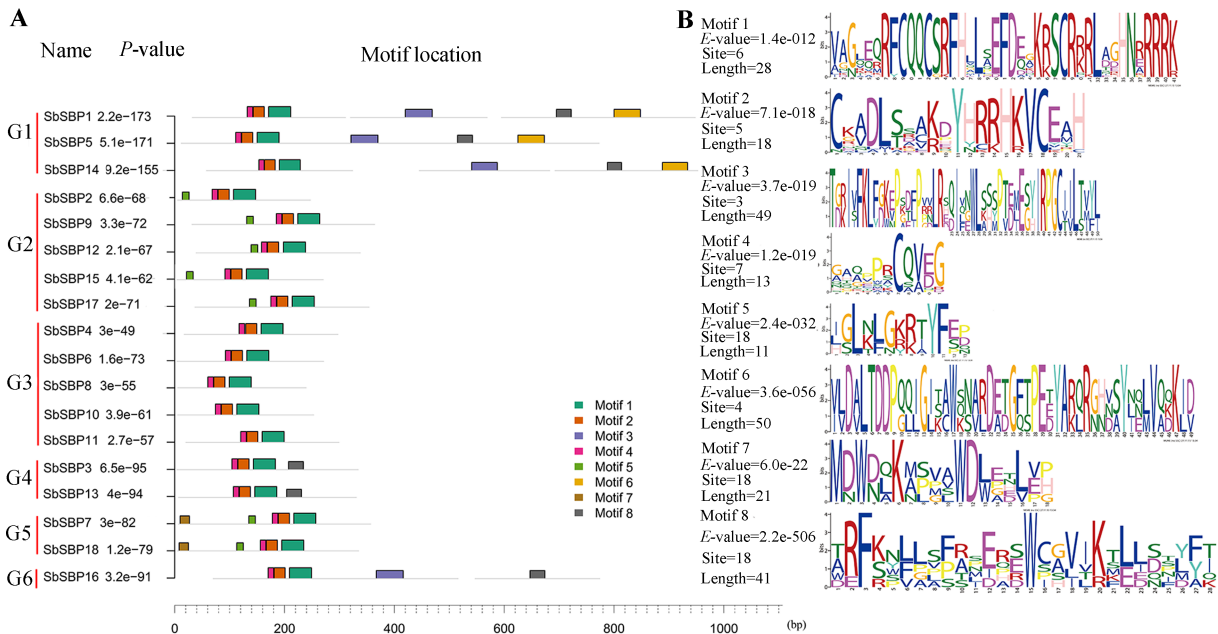


图 4 高粱 SBP-box 蛋白保守基序分析

Fig. 4 Analysis of conserved motifs in SbSBP-box proteins

A :高粱各 SBP-box 蛋白基序分布。不同颜色的方块表示不同基序 ,Motif1 和 motif2 分别为高粱 SBP-box 保守域的 2 个锌指结构。B :各保守基序序列分析。

(图 5)，各亚家族同时含有拟南芥和水稻的 SBP-box 蛋白，说明 SBP-box 基因早于单双子叶植物分化。高粱 SBP-box 基因具有 1 对旁系同源基因，为 *SbSBP9*/*SbSBP17*；与水稻 SBP-box 基因形成 9 对直系同源基因，分别为 *SbSBP1*/*OsSPL6*、*SbSBP5*/*OsSPL1*、*SbSBP2*/*OsSPL17*、*SbSBP7*/*SbSBP18*、*SbSBP4*/*OsSPL13*、*SbSBP10*/*OsSPL19*、*SbSBP16*/*OsSPL9*、*SbSBP17*/*OsSPL10*、*SbSBP11*/*OsSPL7* 和 *SbSBP12*/*OsSPL8*，其

中 *SbSBP12* 同时与玉米叶舌发育基因 *ZmLG1* 直系同源；*SbSBP3* 和 *SbSBP15* 分别与功能较明确的 *ZmTGA1* 和 *ZmUB2/3* 直系同源。

### 2.5 高粱 SBP-box 基因启动子分析

高粱 SBP-box 基因启动子区域所含顺式元件主要为 2 类，一类是与植株生长发育有关的元件，如生长素响应元件(S000024、S000026、S000270、S000370 等)，赤霉素响应元件(S000298、S000420)；

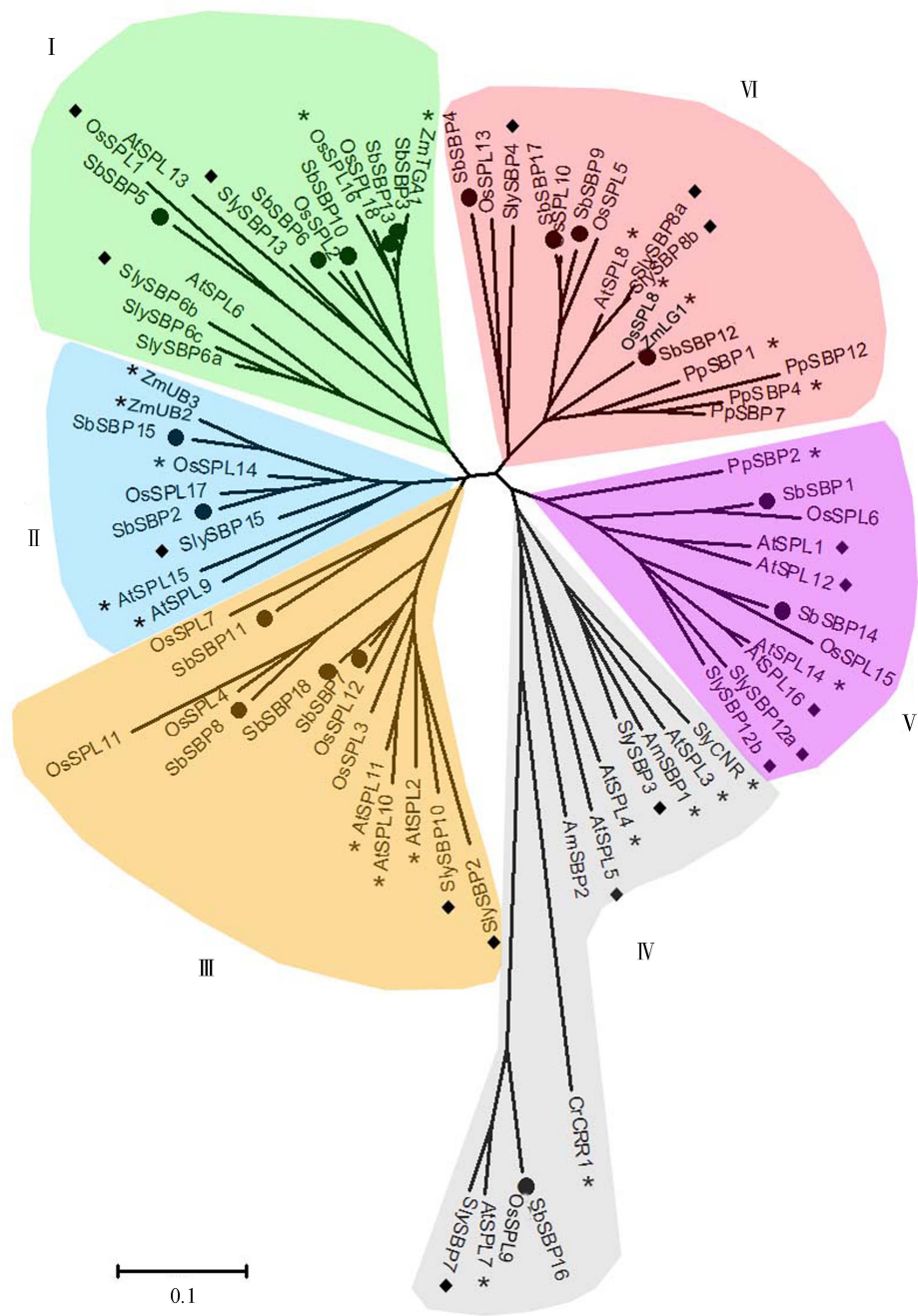


图 5 高粱 SBP-box 蛋白的系统进化分析

**Fig. 5 Phylogenetic analysis of SBP-boxproteins in *Sorghum bicolor* L.**

星号和实心菱形分别表示基因功能和表达数据已知，实心圆圈表示高粱 SBP-box 蛋白；基因前缀 Am、At、Cr、Os、Pp、Sb、Sly、Zm 分别代表 *Antirrhinum majus*、*Arabidopsis thaliana*、*C. reinhardtii*、*Oryza sativa*、*Physcomitrella patens*、*Sorghum bicolor*、*Solanum lycopersicum*、*Zea mays*。

组织特异性表达元件 ,胚乳(S000277、S000353 等)、叶片(S000166)、根毛(S000512)、种子(S000154)、花粉(S000245)等；另一类是逆境响应相关元件，如 ARE(S000022 等)，ABRE(S000013、S000026 等)、DRE(S0000176 等)、LTRE(S000153 等)、MYB 结合位点(S000408)、CuRE(S000493)等。以上部分元件在 18 个基因家族成员中均有分布，如生长素响应元件 S000024、铜响应元件 S000493；某些元件只在个

别基因中存在,如生长素响应元件 S000026、抗氧化反应元件 S000022 及 ABA 响应元件 S000013 分别仅在 *SbSBP14*、*SbSBP6* 和 *SbSBP10* 中存在。

## 2.6 高粱 SBP-box 基因表达分析

根据 RNA-seq 数据集 1(图 6A)对高粱 SBP-box 基因进行表达分析,结果 18 个 SBP-box 基因聚为 5 组,第 1 组 *SbSBP16* 在各组织中均有表达,在幼叶(播种后 25 d)中优势表达;第 2 组 *SbSBP10* 和 *SbSBP11* 均在胚中优势表达;第 3 组所有 SBP-box 基因均在花序原基中优势表达,其中 *SbSBP9* 和 *SbSBP17* 在花序原基中特异表达,*SbSBP11* 在花序原基和胚中特异表达;第 4 组 SBP-box 基因主要在花序原基和雌蕊中表达,其中 *SbSBP15* 在种子(授粉后 5 d)中表达较高;第 5 组在花序原基和花序中表达量相当,其中 *SbSBP5* 和 *SbSBP8* 分别在胚和胚乳中优势表达。根据 RNA-seq 数据集 2(图 6B)可将高粱 SBP-box 基因分成 4 类,第 1 类主要在茎中表达,而第 2 类主要在根中表达,第 3 类 *SbSBP5* 和第 4 类 *SbSBP16* 在不同处理条件下的根茎中均由较高表达;此外,高粱部分 SBP-box 基因受 ABA 和 PEG 诱导上调表达,如 *SbSBP5*、*SbSBP8* 和 *SbSBP18*。

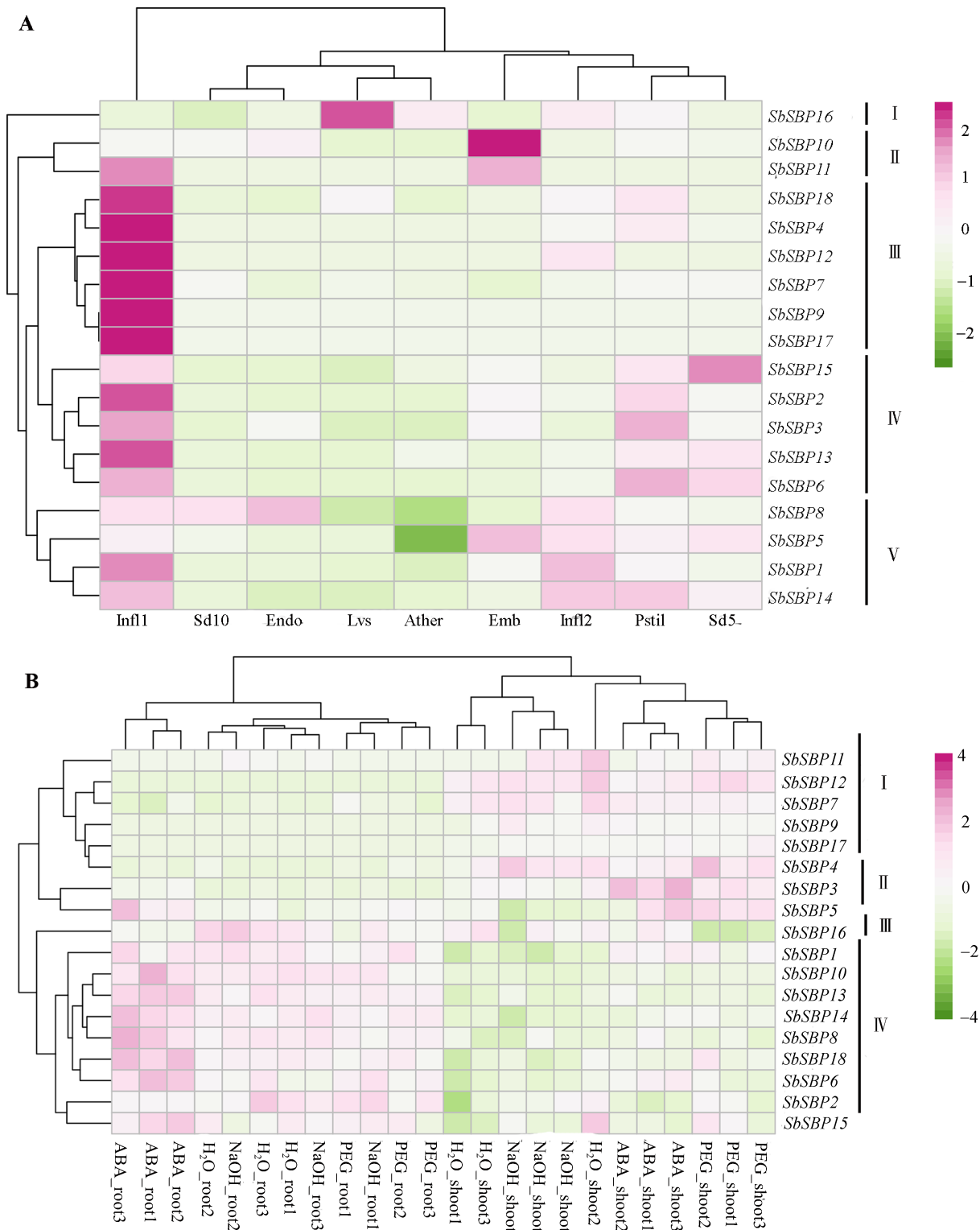
## 3 讨论

高粱属于禾本科高粱属,是 C4 单子叶植物,生物产量高,经济效益大,被誉为“高能作物”<sup>[33]</sup>。高粱基因组为二倍体,有 10 条染色单体,高粱品种 BTx623 基因组测序于 2009 年初完成<sup>[23]</sup>,其基因组大小约 730 Mb,高粱参考基因组序列的破译,加速了高粱基因组学研究及重要基因克隆。

SBP-box 转录因子家族是一类植物特有的转录因子<sup>[34]</sup>,该基因家族成员数目的物种差异性较大,如拟南芥和水稻中分别为 16<sup>[35]</sup>和 18<sup>[23]</sup>个,而玉米中多达 31 个<sup>[8]</sup>。全基因组基因家族研究表明高粱、水稻和拟南芥基因家族大小和数目相对保守<sup>[23]</sup>。本研究在高粱中鉴定出 18 个 SBP-box 基因,其基因家族大小与水稻、拟南芥近似,该结果与全基因组研究结果一致<sup>[25]</sup>。然而,高粱与玉米 SBP-box 基因数目差异较大,可能由于玉米与高粱分离后经历了大规模的全基因组复制事件<sup>[36]</sup>。基因组或基因重复是基

因组变异和基因家族扩张的主要动力,但在随后的进化过程中可能会发生基因丢失<sup>[37]</sup>。高粱 18 个 SPB-box 基因中有 8 个基因位于片段重复区,但只有 *SbSBP9*/*SbSBP17* 成对存在,另外 6 个基因在重复区仅存在 1 个拷贝,另外 1 个拷贝可能在进化过程中丢失,而水稻 18 个 SBP-box 基因中至少有 4 对基因位于片段重复区<sup>[23]</sup>。拟南芥 16 个 SBP-box 基因中有 3 对位于片段重复区,可见物种间的同源基因在进化中承受着不同的选择压。

本文利用高粱、水稻、拟南芥及部分已知功能的玉米、小立陶宛藜的 SBP-box 蛋白序列构建系统进化树,旨在探究不同物种 SBP-box 蛋白的进化关系,同时根据已知信息预测高粱全基因组 SBP-box 基因的功能。进化树结果显示,SBP-box 蛋白可分为 6 个亚家族,每个亚家族功能侧重点不同,本研究主要讨论 3 类。第一类是作物产量相关基因,包括第 1 亚家族的 *ZmTGA1* 控制玉米果壳发育,玉米与大刍草果壳的巨大差异是由 *ZmTGA1* 突变所致<sup>[38]</sup>; *OsSPL16* 可正向调控细胞增殖,控制水稻籽粒大小、形状和品质<sup>[39]</sup>;高粱 *SbSBP3* 和 *SbSBP13* 与这两个基因属于同一进化分支,可能与籽粒发育有关。第二类是株型相关基因,如第 2 亚家族 *ZmUB2/3* 的双突变体植株雄穗分支数减少<sup>[40]</sup>; *OsSPL14* 点突变可致水稻分蘖减少、抗倒伏性增强<sup>[7]</sup>;第 3 亚家族中的 *OsSPL8* 和 *ZmLG1* 直系同源,它们都与植株叶舌发育有关<sup>[13,41]</sup>,其无叶舌突变体叶片直立、株型紧凑是作物株型育种的重要种质资源。与这类基因亲缘关系较近的高粱基因包括 *SbSBP3*、*SbSBP15* 和 *SbSBP12* 其中 *SbSBP3* 与 *ZmTGA1*、*SbSBP15* 与 *ZmUB2/3*、*SbSBP12* 与 *OsSPL8* 和 *ZmLG1* 直系同源,可据此推断这 3 个基因可能参与高粱株型建成。第三类是植物胁迫应答基因,如第 4 亚家族中的 *CrC RRI*<sup>[19]</sup>和 *AtSPL7*<sup>[40]</sup>以及第 5 亚家族中的 *PpSBP2*<sup>[43]</sup>在维持植物体内铜稳态方面发挥重要作用,铜离子是植物正常生命活动所必需的微量元素;第 6 亚家族中的 *AtSPL14* 与拟南芥对真菌类毒素伏马菌素 B1 (Fumonisin B1)的响应有关<sup>[44]</sup>,与这几个基因关系较近的高粱基因为 *SbSBP1*、*SbSBP14* 和 *SbSBP16*,应重点关注其在植物胁迫响应方面的功能。另外,高粱 SBP-box 基因启动子区还存在 DRE(S0000176 等)、



**图 6** 高粱 SBP-box 基因表达的聚类分析

**Fig. 6** Heat map of SBP-box gene expression in *Sorghum bicolor* L.

A : 高粱 SBP-box 基因在 9 个组织中的表达分析。分别为花序原基(Infl1)、花序(Infl2)、授粉后 5 d 的种子(Sd5)、授粉后 10 d 的种子(Sd10)、播种 25 d 后的叶片(Lvs)、花药(Anther)、雌蕊(Pistil)、胚(Emb)和胚乳(Endo)。B : 高粱 SBP-box 基因在不同处理条件下在根(Root)和茎(Shoot)中的表达分析。

LTRE(S000153)等元件,推测 SBP-box 可能参与植物对干旱和低温等逆境的响应。

通过 RNA-seq 数据分析发现高粱 SBP-box 基因在花序原基中表达量显著高于其他组织,如第 6 个 SBP-box 基因在花序原基(花出现前 10 d)中具有较高表达量,但在花序出现后其表达量显著下降,表明这些基因可能参与花序原基的形成,促进高粱从营养生长转变为生殖生长。与葡萄和番茄 SBP-box 基因类似<sup>[10,45]</sup>,该组内的旁系同源基因 *SbSBP9-SbSBP17*,位于染色体片段重复区,它们具有相似的表达模式,即在花序原基中特异表达。第 V 组在花序原基和花序中表达量相当,可能参与高粱营养到生殖生长的转变,同时具有促进开花的作用。高粱 SBP-box 基因受 ABA 诱导上调表达(图 6B),如 *SbSBP5*、*SbSBP8* 和 *SbSBP18* 等,ABA 广泛参与植物对如干旱、低温、渗透胁迫等非生物逆境胁迫的响应<sup>[46]</sup>,因此这类基因可能参与 ABA 依赖性的逆境响应。

本研究利用从 NCBI 中下载的 RNA-seq 数据进行表达分析,发现 *SbSBP16* 在两个表达分析聚类中均独立于高粱其他 SBP-box 成员,单独归为一组,且都在检测的组织中有较高表达;*SbSBP14* 在 2 个数据集中的表达总量都位于最高水平;花序原基特异表达基因 *SbSBP9* 和 *SbSBP17* 在数据集 2 的根茎中均未检测到表达,可见,两个数据集分析的表达结果具有一定重复性,提高了结果的可靠性。

## 参考文献(References):

- [1] Cawley S, Bekiranov S, Ng HH, Kapranov P, Sekinger EA, Kampa D, Piccolboni A, Sementchenko V, Cheng J, Williams AJ, Wheeler R, Wong B, Drenkow J, Yamanaka M, Patel S, Brubaker S, Tammanna H, Helt G, Struhl K, Gingeras TR. Unbiased mapping of transcription factor binding sites along human chromosomes 21 and 22 points to widespread regulation of noncoding RNAs. *Cell*, 2004, 116(4): 499–509. [DOI]
- [2] Yang ZR, Wang XC, Li XM, Yang CD. Advance on the study of transcription factors in higher plants. *Hereditas (Beijing)*, 2004, 26(3): 403–408.  
杨致荣, 王兴春, 李西明, 杨长登. 高等植物转录因子的研究进展. *遗传*, 2004, 26(3): 403–408. [DOI]
- [3] Klein J, Saedler H, Huijser P. A new family of DNA binding proteins includes putative transcriptional regulators of the *Antirrhinum majus* floral meristem identity gene *SQUAMOSA*. *Mol Gen Genet*, 1996, 250(1): 7–16. [DOI]
- [4] Yamasaki K, Kigawa T, Inoue M, Tateno M, Yamasaki T, Yabuki T, Aoki M, Seki E, Matsuda T, Nunokawa E, Ishizuka Y, Terada T, Shirouzu M, Osanai T, Tanaka A, Seki M, Shinozaki K, Yokoyama S. A novel zinc-binding motif revealed by solution structures of DNA-binding domains of *Arabidopsis* SBP-family transcription factors. *J Mol Biol*, 2004, 337(1): 49–63. [DOI]
- [5] Birkenbihl RP, Jach G, Saedler H, Huijser P. Functional dissection of the plant-specific SBP-domain: overlap of the DNA-binding and nuclear localization domains. *J Mol Biol*, 2005, 352(3): 585–596. [DOI]
- [6] Cardon GH, Höhmann S, Nettesheim K, Saedler H, Huijser P. Functional analysis of the *Arabidopsis thaliana* SBP-box gene *SPL3*: a novel gene involved in the floral transition. *Plant J*, 1997, 12(2): 367–377. [DOI]
- [7] Miura K, Ikeda M, Matsubara A, Song XJ, Ito M, Asano K, Matsuoka M, Kitano H, Ashikari M. *OsSPL14* promotes panicle branching and higher grain productivity in rice. *Nat Genet*, 2010, 42(6): 545–549. [DOI]
- [8] Hultquist JF, Dorweiler JE. Feminized tassels of maize *mop1* and *ts1* mutants exhibit altered levels of miR156 and specific SBP-box genes. *Planta*, 2008, 229(1): 99–113. [DOI]
- [9] Liu X, Zhang B, Mao XG, Li A, Sun MR, Jing RL. Cloning of tae-MIR156 precursor gene and sequence polymorphisms of tae-miR156 targeted *TaSPL17*. *Hereditas (Beijing)*, 2014, 36(6): 592–602.  
刘霞, 张斌, 毛新国, 李昂, 孙美荣, 景蕊莲. 小麦 tae-MIR156 前体基因的克隆及其靶基因 *TaSPL17* 多态性分析. *遗传*, 2014, 36(6): 592–602. [DOI]
- [10] Salinas M, Xing SP, Höhmann S, Berndtgen R, Huijser P. Genomic organization, phylogenetic comparison and differential expression of the SBP-box family of transcription factors in tomato. *Planta*, 2012, 235(6): 1171–1184. [DOI]
- [11] Nodine MD, Bartel DP. MicroRNAs prevent precocious gene expression and enable pattern formation during plant embryogenesis. *Genes Dev*, 2010, 24(23): 2678–2692. [DOI]
- [12] Usami T, Horiguchi G, Yano S, Tsukaya H. The more and smaller cells mutants of *Arabidopsis thaliana* identify novel roles for *SQUAMOSA PROMOTER BINDING PROTEIN-LIKE* genes in the control of heteroblasty. *Development*, 2009, 136(6): 955–964. [DOI]
- [13] Moreno MA, Harper LC, Krueger RW, Dellaporta SL, Freeling M. *Liguleless1* encodes a nuclear-localized

- protein required for induction of ligules and auricles during maize leaf organogenesis. *Genes Dev*, 1997, 11(5): 616–628. [DOI]
- [14] Kawakatsu T, Itoh J, Miyoshi K, Kurata N, Alvarez N, Veit B, Nagato Y. *PLASTOCHRON2* regulates leaf initiation and maturation in rice. *Plant Cell*, 2006, 18(3): 612–625. [DOI]
- [15] Wahl V, Ponnu J, Schlereth A, Arrivault S, Langenecker T, Franke A, Feil R, Lunn JE, Stitt M, Schmid M. Regulation of flowering by trehalose-6-phosphate signaling in *Arabidopsis thaliana*. *Science*, 2013, 339(6120): 704–707. [DOI]
- [16] Huijser P, Schmid M. The control of developmental phase transitions in plants. *Development*, 2011, 138(19): 4117–4129. [DOI]
- [17] Xing SP, Salinas M, Garcia-Molina A, Höhmann S, Berndtgen R, Huijser P. *SPL8* and miR156-targeted *SPL* genes redundantly regulate *Arabidopsis* gynoecium differential patterning. *Plant J*, 2013, 75(4): 566–577. [DOI]
- [18] Gou JY, Felippes FF, Liu CJ, Weigel D, Wang JW. Negative regulation of anthocyanin biosynthesis in *Arabidopsis* by a miR156-targeted *SPL* transcription factor. *Plant Cell*, 2011, 23(4): 1512–1522. [DOI]
- [19] Kropat J, Tottey S, Birkenbihl RP, Depege N, Huijser P, Merchant S. A regulator of nutritional copper signaling in *Chlamydomonas* is an SBP domain protein that recognizes the GTAC core of copper response element. *Proc Natl Acad Sci USA*, 2005, 102(51): 18730–18735. [DOI]
- [20] Voinnet O. Origin, biogenesis, and activity of plant microRNAs. *Cell*, 2009, 136(4): 669–687. [DOI]
- [21] Gandikota M, Birkenbihl RP, Höhmann S, Cardon GH, Saedler H, Huijser P. The miRNA156/157 recognition element in the 3' UTR of the *Arabidopsis* SBP box gene *SPL3* prevents early flowering by translational inhibition in seedlings. *Plant J*, 2007, 49(4): 683–693. [DOI]
- [22] Wu G, Poethig RS. Temporal regulation of shoot development in *Arabidopsis thaliana* by miR156 and its target *SPL3*. *Development*, 2006, 133(18): 3539–3547. [DOI]
- [23] Xie KB, Wu CQ, Xiong LZ. Genomic organization, differential expression, and interaction of *SQUAMOSA* promoter-binding-like transcription factors and micro-RNA156 in rice. *Plant Physiol*, 2006, 142(1): 280–293. [DOI]
- [24] Arazi T, Talmor-Neiman M, Stav R, Riese M, Huijser P, Baulcombe DC. Cloning and characterization of micro-RNAs from moss. *Plant J*, 2005, 43(6): 837–848. [DOI]
- [25] Paterson AH, Bowers JE, Bruggmann R, Dubchak I, Grimwood J, Gundlach H, Haberer G, Hellsten U, Mitros T, Poliakov A, Schmutz J, Spannagl M, Tang HB, Wang XY, Wicker T, Bharti AK, Chapman J, Feltus FA, Gowik U, Grigoriev IV, Lyons E, Maher CA, Martis M, Narechania A, Otillar RP, Penning BW, Salamov AA, Wang Y, Zhang LF, Carpita NC, Freeling M, Gingle AR, Hash CT, Keller B, Klein P, Kresovich S, McCann MC, Ming R, Peterson DG, Mehboob ur R, Ware D, Westhoff P, Mayer KF, Messing J, Rokhsar DS. The *Sorghum bicolor* genome and the diversification of grasses. *Nature*, 2009, 457(7229): 551–556. [DOI]
- [26] Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren JY, Li WW, Noble WS. MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res*, 2009, 37(Suppl. 2): W202–W208. [DOI]
- [27] Crooks GE, Hon G, Chandonia JM, Brenner SE. WebLogo: a sequence logo generator. *Genome Res*, 2004, 14(6): 1188–1190. [DOI]
- [28] Tamura K, Dudley J, Nei M, Kumar S. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol*, 2007, 24(8): 1596–1599. [DOI]
- [29] Soderlund C, Nelson W, Shoemaker A, Paterson A. SyMAP: A system for discovering and viewing syntenic regions of FPC maps. *Genome Res*, 2006, 16(9): 1159–1168. [DOI]
- [30] Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA. Circos: an information aesthetic for comparative genomics. *Genome Res*, 2009, 19(9): 1639–1645. [DOI]
- [31] Davidson RM, Gowda M, Moghe G, Lin H, Vaillancourt B, Shiu SH, Jiang N, Robin Buell C. Comparative transcriptomics of three Poaceae species reveals patterns of gene expression evolution. *Plant J*, 2012, 71(3): 492–502. [DOI]
- [32] Dugas DV, Monaco MK, Olsen A, Klein RR, Kumari S, Ware D, Klein PE. Functional annotation of the transcriptome of *Sorghum bicolor* in response to osmotic stress and abscisic acid. *BMC Genomics*, 2011, 12: 514. [DOI]
- [33] Luo H, Zhang LM, Xia Y, Wu XY, Wang C, Liu ZQ, Jing HC. An update on genome research of biofuel sorghum (*Sorghum bicolor*). *Sci Technol Rev*, 2015, 33(16): 17–26. 罗洪, 张丽敏, 夏艳, 吴小园, 王聪, 刘智全, 景海春. 能源植物高粱基因组研究进展. 科技导报, 2015, 33(16): 17–26. [DOI]
- [34] Guo AY, Zhu QH, Gu XC, Ge S, Yang J, Luo JC.

- Genome-wide identification and evolutionary analysis of the plant specific SBP-box transcription factor family. *Gene*, 2008, 418(1–2): 1–8. [DOI]
- [35] Cardon G, Höhmann S, Klein J, Nettekheim K, Saedler H, Huijser P. Molecular characterisation of the *Arabidopsis* SBP-box genes. *Gene*, 1999, 237(1): 91–104. [DOI]
- [36] Swigoňová Z, Lai JS, Ma JX, Ramakrishna W, Llaca V, Bennetzen JL, Messing J. Close split of sorghum and maize genome progenitors. *Genome Res*, 2004, 14(10A): 1916–1923. [DOI]
- [37] Semon M, Wolfe KH. Consequences of genome duplication. *Curr Opin Genet Dev*, 2007, 17(6): 505–512. [DOI]
- [38] Preston JC, Wang H, Kursel L, Doebley J, Kellogg EA. The role of *teosinte glume architecture (tga1)* in coordinated regulation and evolution of grass glumes and inflorescence axes. *New Phytol*, 2012, 193(1): 204–215. [DOI]
- [39] Wang SK, Wu K, Yuan QB, Liu XY, Liu ZB, Lin XY, Zeng RZ, Zhu HT, Dong GJ, Qian Q, Zhang GQ, Fu XD. Control of grain size, shape and quality by *OsSPL16* in rice. *Nat Genet*, 2012, 44(8): 950–954. [DOI]
- [40] Chuck GS, Brown PJ, Meeley R, Hake S. Maize *SBP-box* transcription factors *unbranched2* and *unbranched3* affect yield traits by regulating the rate of lateral primordia initiation. *Proc Natl Acad Sci USA*, 2014, 111(52): 18775–18780. [DOI]
- [41] Lee J, Park JJ, Kim SL, Yim J, An G. Mutations in the rice *liguleless* gene result in a complete loss of the auricle, ligule, and laminar joint. *Plant Mol Biol*, 2007, 65(4): 487–499. [DOI]
- [42] Yamasaki H, Hayashi M, Fukazawa M, Kobayashi Y, Shikanai T. *SQUAMOSA* promoter binding protein-like7 is a central regulator for copper homeostasis in *Arabidopsis*. *Plant Cell*, 2009, 21(1): 347–361. [DOI]
- [43] Nagae M, Nakata M, Takahashi Y. Identification of negative cis-acting elements in response to copper in the chloroplastic iron superoxide dismutase gene of the moss *Barbula unguiculata*. *Plant Physiol*, 2008, 146(4): 1687–1696. [DOI]
- [44] Stone JM, Liang XW, Nekl ER, Stiers JJ. *Arabidopsis AtSPL14*, a plant-specific SBP-domain transcription factor, participates in plant development and sensitivity to fumonisin B1. *Plant J*, 2005, 41(5): 744–754. [DOI]
- [45] Hou HM, Li J, Gao M, Singer SD, Wang H, Mao LY, Fei ZJ, Wang XP. Genomic organization, phylogenetic comparison and differential expression of the SBP-box family genes in grape. *PLoS One*, 2013, 8(3): e59358. [DOI]
- [46] Dalal M, Tayal D, Chinnusamy V, Bansal KC. Abiotic stress and ABA-inducible Group 4 *LEA* from *Brassica napus* plays a key role in salt and drought tolerance. *J Biotechnol*, 2009, 139(2): 137–145. [DOI]

(责任编辑: 张宪省)