

研究报告 Original Papers

谷子穗发育过程的基因表达动态分析

申慧敏¹, 程金金², 王亚敏¹, 吴年隆², 杨致荣², 赵雄伟^{1,2,*}, 王兴春^{1,2,*}¹山西农业大学生命科学学院, 山西晋中030801²山西农业大学杂粮功能基因组学中心, 山西晋中030801

*共同通信作者: 王兴春(wxingchun@sxau.edu.cn)、赵雄伟(xwzhao_sxau@outlook.com)

摘要: 穗是谷子(*Setaria italica*)产量和品质形成的基础, 理想的穗型是谷子遗传改良的重要目标。以新型C₄禾谷类模式植物超早熟小型谷子突变体*xiaomi*为对象, 利用RNA测序(RNA-Seq)技术分析了抽穗期、开花期和灌浆期3个发育时期穗中基因表达动态。结果表明, 开花期/抽穗期、灌浆期/开花期和灌浆期/抽穗期差异表达基因数分别为4 159、4 254和9 243个。其中, 光敏色素基因(*PHYTOCHROME*)、光周期枢纽基因(*CONSTANS*, *CO*)以及成花素基因(*FLOWERING LOCUS T*, *FT*)等在上述3个穗发育时期表达量差异显著, 而一些与淀粉合成和代谢相关的基因在谷子穗发育过程中表达量逐渐升高。此外, 还发现大量未知功能基因, 这些基因在不同穗发育时期表达量差异显著。这些基因主要涉及次生代谢物生物合成、苯丙素类化合物生物合成、植物病原菌互作、植物激素信号转导、脂肪酸合成、糖类物质合成和蜡质合成等代谢通路。本研究揭示了谷子穗发育的基因表达动态, 拟为深入解析穗发育调控的分子机制奠定基础。

关键词: 谷子; 穗发育; 抽穗期; 开花期; 灌浆期; 差异表达基因

Dynamic gene expression analysis during panicle development in foxtail millet

SHEN Huimin¹, CHENG Jinjin², WANG Yamin¹, WU Nianlong², YANG Zhirong², ZHAO Xiongwei^{1,2,*}, WANG Xingchun^{1,2,*}¹College of Life Sciences, Shanxi Agricultural University, Jinzhong, Shanxi 030801, China²Minor Crops Functional Genomics Center, Shanxi Agricultural University, Jinzhong, Shanxi 030801, China

*Co-corresponding authors: Wang XC (wxingchun@sxau.edu.cn), Zhao XW (xwzhao_sxau@outlook.com)

Abstract: Panicle is crucial for the formation of yield and quality in foxtail millet (*Setaria italica*), and an ideal panicle architecture is a desirable target trait for genetic improvement of foxtail millet. The dynamic gene expression profiles at heading, flowering and grain-filling stages were analyzed in *xiaomi*, a novel C₄ model plant for cereals using RNA sequencing (RNA-Seq) technology. The results show that there were 4 159, 4 254 and 9 243 differentially expressed genes at flowering/heading, grain-filling/flowering and grain-filling/heading stages, respectively. Among these genes, the phytochrome-related genes, photoperiod hub (*CONSTANS*, *CO*) and florigen *FLOWERING LOCUS T* (*FT*) were significantly differentially expressed in the above three panicle development stages. The expression levels of the genes involved in starch synthesis and metabolism increased gradually during panicle development. In addition, a large number of unknown

收稿 2020-12-14 修定 2021-03-17

资助 国家自然科学基金(31600289)和山西省研究生教育创新项目(2019BY067)。

functional genes were significantly differentially expressed in different panicle development stages, including those involved in secondary metabolite biosynthesis, phenylpropanoid biosynthesis, plant-pathogen interaction, plant hormone signal transduction, fatty acid synthesis, carbohydrate synthesis, and wax metabolic pathways. The gene expression dynamics during panicle development was dissected in foxtail millet. We wish this study will lay a foundation for further understanding of the molecular regulation mechanism of panicle development in millet species.

Key words: foxtail millet; panicle development; heading stage; flowering stage; grain-filling stage; differentially expressed gene

穗是禾谷类作物的生殖器官,也是作物产量构成的重要因素。近年来,一大批水稻(*Oryza sativa*)、小麦(*Triticum aestivum*)和玉米(*Zea mays*)等作物穗发育相关位点和基因被克隆。2009年,中国科学院遗传与发育生物学研究所傅向东课题组从超级水稻中分离出一个直立穗基因*DENSE AND ERECT PANICLE1 (DEP1)*,该基因在中国超级稻增产中发挥关键作用;*DEP1*能促进细胞分裂,使稻穗变密、枝梗数增加、每穗籽粒数增多,从而促使水稻增产(Huang等2009)。最近,李学勇课题组发现穗退化与部分不育基因*DEGENERATED PANICLE AND PARTIAL STERILITY 1 (DPS1)*在调控水稻体内活性氧(reactive oxygen species, ROS)的稳态、花药角质层形成和穗发育过程中起重要作用(Zafar等2020)。植物穗发育受到温度等环境条件的影响。温度较高时,水稻*pdf1 (panicle development failure 1)*穗起始分化正常,之后穗轴基部细胞开始死亡,最终整个穗坏死;但在相对低温的环境中,*pdf1*突变体可以正常抽穗(赵冬生等2020)。穗发育状态与作物产量紧密相关。小麦*AGOLD (ARGONAUTE 1d)*基因突变后,小麦的可育花粉数大大下降,严重影响了后期的每穗籽粒数(Feng等2017)。在玉米中,增强表达行粒数基因*KERNEL NUMBER PER ROW6 (KNR6)*可显著提高玉米杂交种籽粒产量,*KNR6*通过与Arf-GTPase激活蛋白(Arf-GTPase activating protein, AGAP)互作来调节花序发育,从而影响穗长和行粒数(Jia等2020)。因此,深入挖掘禾谷类植物穗发育相关基因不仅有助于穗发育生理机制的解析,而且对于作物高产分子育种具有重要的指导意义。

谷子(*Setaria italica*)是起源于中国的一种古老作物,广泛种植于干旱和半干旱地区,在旱作可持

续生态农业建设中具有重要作用。由于基因组较小(≈438 Mb),且具有抗旱耐瘠薄、光合效率高、易于室内大规模种植等特点,已成为C₄禾谷类研究的理想模式植物(Yang等2020; Bennetzen等2012; Doust等2009)。刁现民研究员课题组构建了甲基磺酸乙酯(ethyl methylsulfonate, EMS)诱变的‘豫谷1号’突变体库,从中筛选出一系列穗发育相关的突变体(Hussin等2020; 相吉山等2020; 薛红丽等2018; Xiang等2017; 李雯等2015)。其中,松散穗基因*LOOSE PANICLE1 (LPI)*编码一个WRKY转录因子,该基因突变后导致穗子松散、半矮秆、籽粒变大等异常表型(Xiang等2017);*sipaall*突变体除了表现为穗顶端败育外,谷穗长、谷穗宽、叶长、叶宽和千粒重等都大幅减小(薛红丽等2018);小穗突变体*sisp1*表现为穗子变小、株高降低、籽粒变大等表型(李雯等2015)。上述研究表明,穗发育相关基因往往具有多效性。谷子的祖先是青狗尾草(*Setaria viridis*),青狗尾草生长素运输基因*SvAUX-IN (SvAUX)*及其玉米同源基因*ZmAUX*突变后都导致穗分枝减少和穗松散的表型,暗示青狗尾草/谷子可以作为玉米等其他禾本科植物穗发育研究的模式植物(Huang等2017)。高通量测序技术的出现,为谷子功能基因组学研究提供了强有力的工具。最近,Wang等(2020)利用RNA测序(RNA sequencing, RNA-Seq)技术研究了谷子籽粒灌浆过程中基因表达情况,获得11 399个差异表达基因,为谷子籽粒发育相关基因的发掘奠定了基础。尽管如此,目前谷子穗发育相关研究仍局限在少数基因,穗发育过程的基因表达动态仍知之甚少。

最近,本课题组创制了一个超早熟小型谷子突变体*xiaomi*,其抽穗期仅为38 d左右,生育期仅为

2个月左右,与模式植物拟南芥相当,是禾本科穗发育机制研究的理想模式植物(Yang等2020)。为揭示谷子穗发育过程中基因表达的动态并发掘相关基因,本研究利用HiSeq X Ten (因美纳,美国加利福尼亚)测序平台分别对抽穗期、开花期和灌浆期的*xiaomi*穗子进行转录组测序,拟为穗发育调控分子机制的解析奠定基础。

1 材料与方法

1.1 实验材料

实验材料为本实验室创制的超早熟小型谷子[*Setaria italica* (L.) P. Beauv.]突变体*xiaomi*,其抽穗期仅为38 d左右,极大地方便了穗发育机制的研究(Yang等2020)。将*xiaomi*种植于山西农业大学杂粮分子育种团队实验田内(山西省晋中市太谷区,37°25'13"N 112°35'26"E),分别在抽穗期、开花期和灌浆期进行取样。每个样品至少随机取5株健康植株的主穗,3次生物学重复。所有样品立即置于液氮中冷冻保存。

1.2 总RNA提取、RNA-Seq文库构建及测序

总RNA提取采用RNAprep Pure试剂盒(天根生化科技有限公司,中国大连)进行,提取方法参照说明书。提取的总RNA用Agilent 2100生物分析仪(安捷伦,美国加利福尼亚)和琼脂糖凝胶电泳检测,合格的样品用于后续mRNA的分离纯化和测序文库的构建。首先,采用NEBNext Poly(A) mRNA磁性分离模块(纽英伦生物技术,美国马萨诸塞)分离纯化mRNA,将其打断成约200 nt的RNA短片段;然后,以片段化的mRNA为模版,合成cDNA第一链和互补链;最后,双链cDNA经末端修复和接头连

接后,用Agencourt AMPure XP beads (贝克曼库尔特,美国加利福尼亚)纯化,并进行PCR扩增以建立cDNA文库。构建的cDNA文库由北京百迈克生物科技有限公司使用Illumina HiSeq X Ten平台进行测序。

1.3 测序数据处理及差异表达基因分析

RNA-Seq得到的原始读段(raw reads)测序数据利用Trimmomatic (Bolger等2014)过滤获得干净读段(clean reads),利用hisat2 (Kim等2015)将干净读段比对到*xiaomi*参考基因组(<http://sky.sxau.edu.cn/MDSi.htm>; Yang等2020)。采用每百万读段的转录本数(transcripts per million reads, TPM; Li等2010)计算基因的表达量,以 $|\log_2(\text{差异倍数})| \geq 1$ 且校验后的*P*值(adjusted *P*-value, P_{adj}) ≤ 0.01 为阈值筛选差异表达基因。然后,利用*xiaomi*注释信息对差异表达基因进行功能注释,并进行基因本体(gene ontology, GO)和KEGG (Kyoto Encyclopedia of Genes and Genomes)通路富集分析。

1.4 实时定量PCR反应

采用宝日医生物技术(北京)有限公司的反转录试剂盒PrimeScript RT reagent Kit (货号RR047A)合成cDNA第一链, TB Green Premix Ex Taq II试剂盒 (货号RR820A)进行实时定量PCR (real-time quantitative PCR, RT-qPCR), 3次生物学重复。15 μL 反应体系包括: 7.5 μL 2 $\times$ Taq PCR Master Mix (含荧光染料)、各0.6 μL 10 $\mu\text{mol}\cdot\text{L}^{-1}$ 的上下游引物、3.3 μL 双蒸水(ddH₂O)和3 μL cDNA模板。反应条件为: 首先95℃预变性2 min, 然后采用两步法(95℃变性5 s、60℃退火和延伸30 s)扩增35个循环。RT-qPCR引物序列信息见表1。以内参基因*Si9g-*

表1 RT-qPCR用到的引物
Table 1 Primers for RT-qPCR

基因编号	正向引物序列(5'→3')	反向引物序列(5'→3')
<i>Si9g37480</i>	ATGGCCCGTACCAAGCAAACCT	GGATCTCACGGAGGGCAACAGT
<i>Si7g09640</i>	AGGAGGACATCTGCGAGGTCATC	AAGTAGCCGAGCGGGTTGGG
<i>Si5g03150</i>	TCTAGCACAAAGGAGGGCCTACTC	TTCACCTCCTTCTTGCCCATCACG
<i>Si4g06400</i>	TCTGCCTCTGCACCACCATCC	GTGCTTGCCGCAGGTGATGAG
<i>Si4g02980</i>	CCTGACCGTGAGCCCCTACTAC	GTCCATGCCGTTGACGATACCG

基因编号为*xiaomi*基因组编号(Yang等2020),下同。

37480为对照, 该基因在谷子的11个组织中表达量稳定。利用公式 $2^{-\Delta\Delta C_t}$ 计算基因相对表达量。

1.5 基因表达热图绘制

基因表达热图利用广州基迪奥生物科技有限公司云平台(<https://www.omicshare.com/>)进行。首先, 对基因的表达量TPM值做归一化处理, 将每个基因的表达量减去这个基因在所有样本中表达量的均值, 然后除以其标准差。这个处理也叫标准正态化, 或Z-score处理。然后利用R语言热图(pheatmap)包, 用表达量信息计算两两样本间的欧氏距离, 最后用欧式距离实现样本的聚类。

2 实验结果

2.1 转录组测序及测序数据统计分析

谷子穗子发育主要包括抽穗期、开花期和灌

浆期3个时期; 抽穗期穗子较短, 开花期穗子长度与灌浆期相近(图1)。为解析谷子穗发育过程的基因表达情况, 我们对这三个发育时期的xiaomi穗子进行了转录组测序, 每个时期3次生物学重复。9个测序文库共获得611.86 M干净读段, 平均每个样品为67.98 M (表2)。经过与xiaomi参考基因组比对分析发现, 平均86%的干净读段可比对到参考基因组上。比对到参考基因组唯一位置和多处位置的比例分别为73.66%~86.18%和4.87%~7.04%; 比对到参考基因组正链和负链的比例分别为36.80%~43.39%和36.86%~43.47%, 二者比例相当(表2)。

为鉴定9个样品间的生物学重复关系, 对各样本间生物学重复数据进行皮尔逊相关系数(Pearson correlation coefficient, PCC)分析。结果表明, 抽穗期3个生物学重复间的相关性最高, PCC均大于

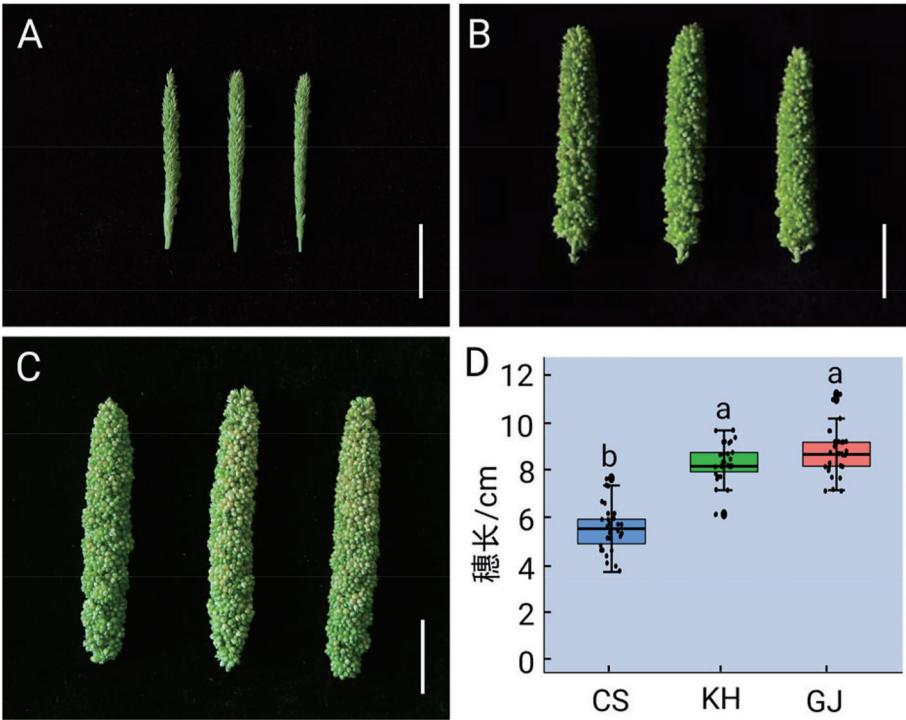


图1 三个发育时期的xiaomi穗子及穗长统计

Fig. 1 The xiaomi panicles and statistics of panicle length at three developmental stages

A: 抽穗期的xiaomi穗子; B: 开花期的xiaomi穗子; C: 灌浆期的xiaomi穗子; D: 不同发育时期的xiaomi穗长, CS、KH和GJ分别表示抽穗期、开花期和灌浆期。在箱形图中, 中线代表中位数, 底边和顶边分别代表第一和第三个四分位数, 上下触须分别为最大和最小值, 每个点代表一个穗长数据。箱体上相同的小写字母表示差异不显著, 不同的小写字母表示差异显著, $n=30$, $P<0.05$ 。标尺: 2 cm。

表2 测序数据统计
Table 2 Statistics of sequencing data

样品	干净读段数	匹配的读段数	唯一匹配的读段数	多匹配的读段数	匹配正链的读段数	匹配负链的读段数
CS-1	65 690 050	57 466 979	53 945 218	3 521 761	26 958 571	26 986 647
CS2	57 408 110	49 870 439	46 776 299	3 094 140	23 374 814	23 401 485
CS-3	94 010 158	80 683 568	75 359 925	5 323 643	37 659 782	37 700 143
KH-1	54 710 677	50 125 532	47 521 483	2 604 049	23 739 771	23 781 712
KH-2	59 071 640	51 313 652	48 150 907	3 162 745	24 059 219	24 091 688
KH-3	71 754 319	65 331 176	61 835 751	3 495 425	30 881 498	30 954 253
GJ-1	66 280 825	55 114 156	51 007 091	4 107 065	25 485 720	25 521 371
GJ-2	61 736 100	53 103 059	49 650 886	3 452 173	24 806 448	24 844 438
GJ-3	81 198 183	65 527 846	59 813 869	5 713 977	29 880 949	29 932 920

CS表示抽穗期, KH表示开花期, GJ表示灌浆期, 1、2、3代表3次生物学重复。下同。

0.986; 其次是灌浆期的3个生物重复, PCC均大于0.905; 开花期KH-1和KH-3样品间的PCC值较高, 为0.960; 样品KH-2与KH-1和KH-3之间的PCC稍低, 分别为0.753和0.848 (图2)。表明这3个时期的3个生物学重复间的相关性达到了强相关, 可进行差异基因表达分析。

2.2 差异表达基因分析与RT-qPCR验证

开花期与抽穗期相比, 共有4 159个差异表达基因, 其中上调基因2 544个, 下调基因1 615个; 灌浆期与开花期相比, 共有6 987个差异表达基因, 其中上调基因2 733个, 下调基因4 254个; 灌浆期与

抽穗期相比, 共有9 243个差异表达基因, 其中上调基因4 478个, 下调基因4 675个(图3-A和表3)。其中3个组共有的差异表达基因为1 991个, 特有的差异表达基因分别为384、699和2 143个(图3-B)。为分析测序结果的可靠性, 挑选了4个基因进行RT-qPCR验证。其中, *Si7g09640*基因编码一个花粉特异表达类C13蛋白(pollen-specific protein C13-like), 在开花期表达量最高; *Si5g03150*与胚胎发育后期富含类蛋白合成相关, 在开花期表达量最高; *Si4g-06400*基因编码一个蛋白酶抑制子/种子储存/LTP家族蛋白前体(protease inhibitor/seed storage/LTP family protein precursor), 在灌浆期表达量最高; *Si4g-02980*基因编码一个颗粒结合淀粉合成酶(granule bound starch synthase), 是直链淀粉合成的关键基因, 在灌浆期表达量最高。RT-qPCR结果表明, 4个基因的上调或下调倍数虽然与测序结果有差异, 但基因表达变化趋势与测序结果一致, 且符合不同穗发育时期的规律(图3-C~F)。

进一步分析发现, 一些抽穗期相关基因在3个穗发育时期差异表达, 如受体相关基因光敏色素A (*PHYTOCHROME A*, *Si9g11420*)、光敏色素相关蛋白基因(*PHYTOCHROME-ASSOCIATED PROTEIN*, *Silg03830*和*Si3g08620*)、光周期途径的枢纽*CO*基因(*Silg07180*、*Silg23650*、*Si2g41970*、*Si4g10360*、*Si4g18240*、*Si7g16720*和*Si9g11970*), 以及成花素基因*FT* (*Si4g07330*和*Si4g17010*)等。一些与淀粉合成和代谢相关的基因在谷子穗发育过程的表达

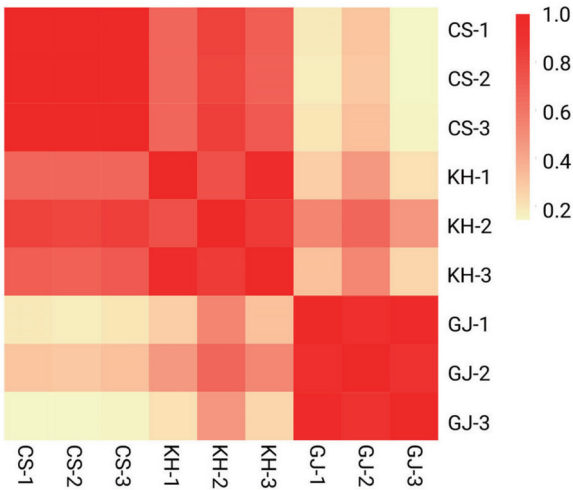


图2 九个样品间的皮尔逊相关系数分析
Fig. 2 Pearson correlation coefficients of the nine samples

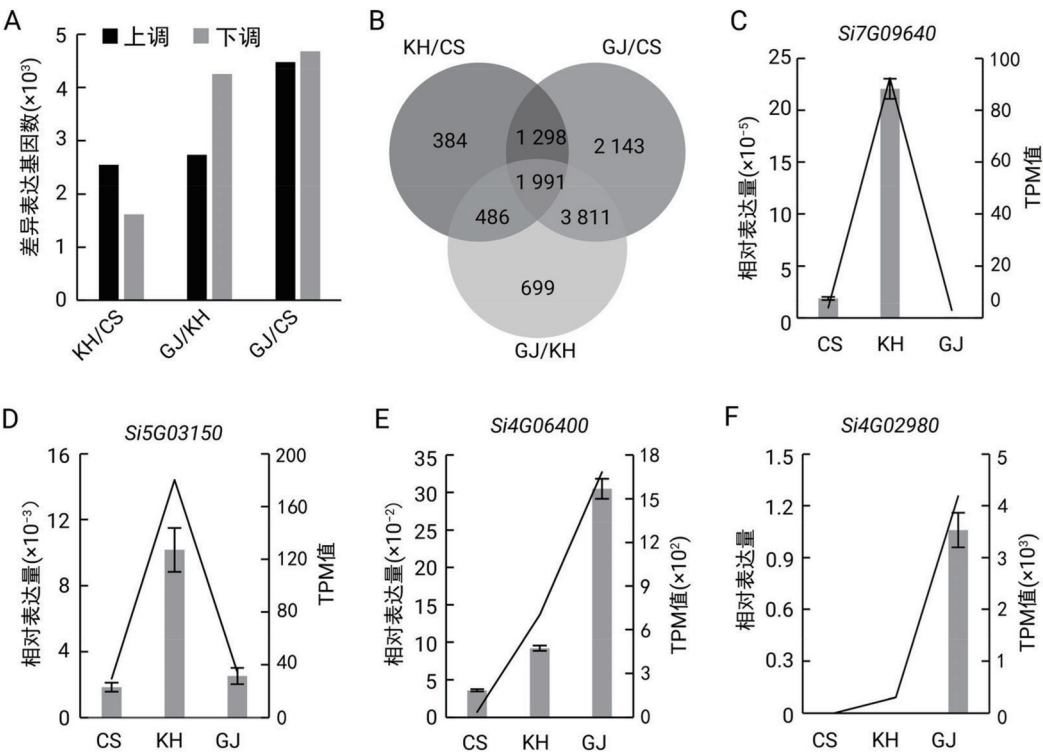


图3 谷子穗发育过程基因表达分析

Fig. 3 Gene expression analysis during panicle development in foxtail millet

A: 不同穗发育时期差异表达基因情况; B: 不同穗发育时期差异表达基因韦恩图; C~F: 不同穗发育时期差异表达基因 RT-qPCR验证, 柱形为RT-qPCR测定的基因相对表达量, 折线为RNA-Seq的基因表达量TPM值。

量逐渐升高, 且表达量差异显著(*Si0g16430*、*Si2g-34430*、*Si4g07100*、*Si6g13270*、*Si6g21130*和*Si7g-02800*)。

植物激素在穗发育过程中起着极其重要的作用。生长素早期响应基因(*SMALL AUXIN UP RNA*, *SUAR*: *Si4g03290*)在抽穗期表达量最高(TPM值 3 646), 而在开花期和灌浆期表达量迅速下降, TPM值分别为831和61; 类似地, 生长素响应基因(*INDOLE ACETIC ACID*, *IAA*: *Si1g03830*、*Si3g08620*、*Si3g37150*、*Si3g37160*、*Si4g15080*和*Si5g27630*)、生长素转运蛋白基因*AUX1* (*Si5g38740*、*Si8g05250*和*Si9g28470*)、茉莉酸信号途径关键调控因子基因(*JASMONATE ZIM-DOMAIN*, *JAZ*: *Si2g22440*、*Si2g-38670*和*Si3g04270*)等基因表达量的变化趋势也是随着穗发育逐渐降低。细胞分裂素受体基因(*ARABIDOPSIS HISTIDINE KINASE*, *AHK*: *Si9g11580*)

和乙烯响应转录因子在抽穗期和开花期表达量无显著差异, 而在灌浆期大幅下降; 相反, 脱落酸受体基因(*PYRABACTIN RESISTANCE 1-LIKE*, *PYL*: *Si4g-22990*、*Si5g37100*和*Si9g42890*)在抽穗期和开花期表达量较低, 在灌浆期表达量大幅提高。

2.3 差异基因表达趋势分析

为研究谷子穗发育过程差异基因在抽穗期、开花期和灌浆期的表达情况, 分析3个穗发育时期差异表达基因的表达趋势, 结果发现共可以分为8组, 同一组的基因有相同的表达趋势(图4)。其中, 在开花期和灌浆期表达量均显著下调的基因有2 536个(图4-A), 均显著上调的基因有2 777个(图4-H), 而另外一些基因在这三个时期表现出先升高后降低或先降低后升高的趋势(图4-C和F), 还有一部分表现出逐渐升高或下降的趋势(图4-B、D、E和G)。

上述结果表明在谷子穗发育过程有大量基因

表3 谷子3个穗发育时期差异表达的基因

Table 3 Differential expressed genes during the three panicle development stages in foxtail millet

差异表达时期	基因编号	log ₂ (差异倍数)	基因功能
开花期/抽穗期	<i>Si7g06850</i>	-5.99	编码肉桂醇脱氢酶6类蛋白
	<i>Si5g09280</i>	-5.13	编码吩嗪生物合成类结构域蛋白
	<i>Si9g35840</i>	-4.68	未知
	<i>Si0g01280</i>	-4.39	未知
	<i>Si0g01230</i>	-4.11	未知
	<i>Si9g32150</i>	20.44	未知
	<i>Si4g14640</i>	20.77	未知
	<i>Si3g16710</i>	21.08	编码雌蕊特异性伸展蛋白样蛋白
	<i>Si8g18850</i>	21.83	编码类α-玉米醇溶蛋白PMS1
	<i>Si9g29210</i>	23.11	未知
灌浆期/开花期	<i>Si4g16520</i>	-13.87	编码非特异性脂质转运蛋白3类蛋白
	<i>Si7g06600</i>	-12.43	编码4-丁香酸酯CoA连接酶
	<i>Si7g19980</i>	-11.90	编码蓝铜蛋白
	<i>Si6g17580</i>	-11.67	编码类萌发素8-14-类蛋白
	<i>Si9g46040</i>	-11.51	编码富含脯氨酸蛋白
	<i>Si2g33360</i>	10.15	编码蛋白磷酸酶2C
	<i>Si6g03550</i>	10.30	编码防御素类蛋白2
	<i>Si9g29690</i>	10.37	未知
	<i>Si7g09890</i>	10.98	编码36.4 kDa富含脯氨酸蛋白
	<i>Si6g03540</i>	12.86	编码防御素类蛋白21
灌浆期/抽穗期	<i>Si4g16520</i>	-14.70	未知
	<i>Si7g06600</i>	-13.89	编码锌指蛋白4
	<i>Si9g46040</i>	-13.26	编码漆酶-15
	<i>Si6g17580</i>	-13.23	未知
	<i>Si7g19980</i>	-12.65	编码枯草杆菌蛋白酶抑制剂2B亚型X1
	<i>Si9g32150</i>	24.81	编码E3泛素蛋白连接酶EL5
	<i>Si3g11320</i>	25.04	编码锌指蛋白4
	<i>Si8g18850</i>	25.84	未知
	<i>Si4g14640</i>	26.17	未知
	<i>Si9g29210</i>	27.03	编码叶绿体FAF类蛋白

由于篇幅限制, 本表仅列出上调/下调倍数前5的基因。如需要全部结果请联系通信作者。

表达量发生改变。由于转录因子是影响基因表达的最主要因素, 我们进一步分析了谷子穗发育时期转录因子的表达情况。差异表达的基因中有803个编码转录因子。其中, 开花期/抽穗期、灌浆期/开花期、灌浆期/抽穗期分别有316、504和693个差异表达的转录因子基因。其中, 在抽穗期和开花期表达量下调的转录因子较多, 分别有420和309个; 而在灌浆期表达量上调的转录因子较多, 为417个(图5)。

2.4 谷子穗发育相关差异表达基因KEGG通路分析

对抽穗期、开花期和灌浆期3个穗发育时期间的差异表达基因进行KEGG分析, 结果表明抽穗期/开花期、灌浆期/开花期和灌浆期/抽穗期3个组间的差异基因分别被注释到104、118和125条KEGG代谢通路中, 其中有100条是3个对照组间共同注释到的代谢通路。以Q值(由P值经过多重检验校正得到)≤0.05为筛选标准, 3个对照组间的397、806和633个差异表达基因分别被注释到17、22和

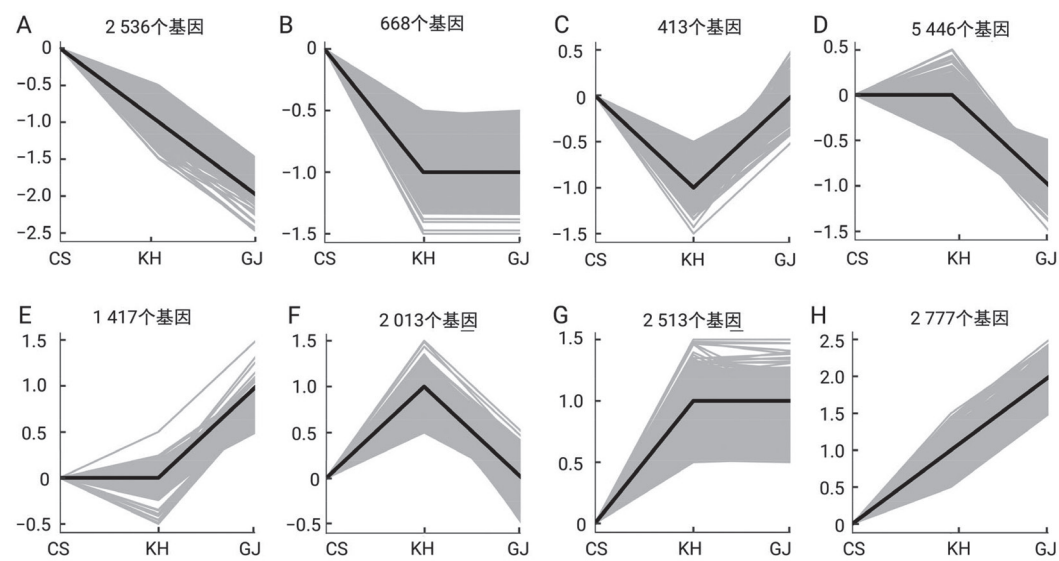


图4 穗发育过程差异表达基因的表达趋势分析

Fig. 4 Expression trend analysis of differentially expressed genes during panicle development

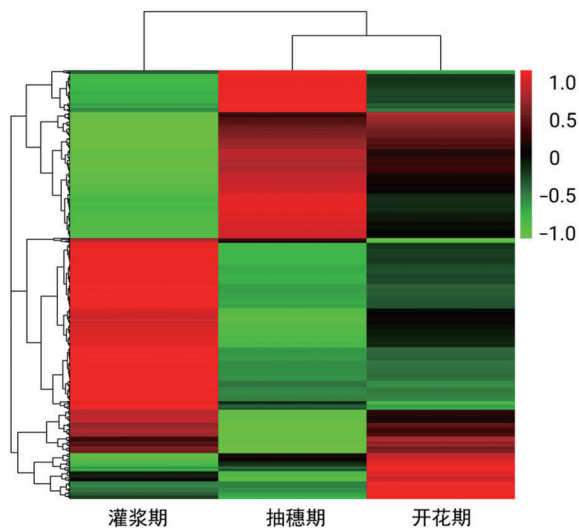


图5 差异表达转录因子基因表达热图

Fig. 5 Heat map of differentially expressed transcription factor genes

29条代谢通路中, 其中次生代谢物生物合成、苯丙素类化合物生物合成、植物病原菌互动、植物激素信号转导、脂肪酸合成、糖类物质合成、角质、丝氨酸和蜡质等代谢通路富集程度较高(表4)。

2.5 GO分析

对抽穗期、开花期和灌浆期3个穗发育时期的

差异表达基因进行GO富集分析, 结果表明抽穗期/开花期参与生物过程、分子功能和细胞组分的基因分别有2 180、2 393和2 362个; 灌浆期/开花期参与生物过程、分子功能和细胞组分的基因分别有3 749、3 955和4 074个; 灌浆期/抽穗期参与生物过程、分子功能和细胞组分的基因分别有4 998、5 285和5 398个(表5~7)。抽穗期/开花期的差异表达基因被富集到3 635个GO条目中, 其中生物过程、分子功能和细胞组分富集的基因分别占63.28%、27.89%和8.83%; 灌浆期/抽穗期的差异基因被富集到5 383个GO条目中, 富集到生物过程、分子功能和细胞组分的基因分别占60.38%、29.91%和9.71%; 灌浆期/开花期差异表达的基因被富集到4 697个GO条目中, 其中富集到生物过程、分子功能和细胞组分的基因分别占61.85%、28.72%和9.43%。此外, 生物过程中有2 115个GO条目是3个比较组间共同富集到的通路, 其中催化活性的负调节(GO:0043086)、蛋白磷酸化(GO:0006468)、外部囊状结构组织(GO:0045229)、细胞壁修饰(GO:0042545)、碳水化合物代谢过程(GO:0005975)和分子功能负调节(GO:0044092)等GO条目富集程度较高; 分子功能中有884个GO条目是3个比较组间共同富集到的通路, 其中氧化还原酶活性(GO:

表4 差异表达基因KEGG富集条目

Table 4 KEGG classification terms of differential expression genes

差异表达时期	KEGG途径	基因数	Q值
开花期/抽穗期	次生代谢物的生物合成	196 (867)	4.46×10 ⁻¹⁹
	脂肪酸延长	29 (46)	4.71×10 ⁻¹⁴
	代谢途径	264 (1 496)	1.08×10 ⁻¹¹
	角质、琥珀和蜡的生物合成	19 (30)	2.27×10 ⁻⁹
	氨基糖和核苷酸糖代谢	40 (122)	6.42×10 ⁻⁸
	苯丙素类化合物生物合成	65 (263)	3.40×10 ⁻⁷
	植物-病原菌互作	55 (208)	3.40×10 ⁻⁷
	戊糖、葡萄糖醛酸转换	17 (50)	9.29×10 ⁻⁴
	二萜生物合成	10 (22)	1.70×10 ⁻³
	吡啶生物碱生物合成	13 (36)	2.68×10 ⁻³
	次生代谢物的生物合成	305 (867)	7.83×10 ⁻²⁶
	代谢途径	449 (1 496)	8.58×10 ⁻²⁴
灌浆期/开花期	苯丙素类化合物生物合成	120 (263)	1.02×10 ⁻¹⁸
	类黄酮生物合成	23 (38)	3.96×10 ⁻⁶
	淀粉和蔗糖的代谢	42 (102)	6.54×10 ⁻⁵
	角质、琥珀和蜡的生物合成	18 (30)	7.61×10 ⁻⁵
	戊糖、葡萄糖醛酸转换	25 (50)	8.66×10 ⁻⁵
	氨基糖和核苷酸糖代谢	47 (122)	9.56×10 ⁻⁵
	脂肪酸延长	23 (46)	1.61×10 ⁻⁴
	芪类、二苯基庚酮和姜酚生物合成	14 (23)	4.53×10 ⁻⁴
	次生代谢物的生物合成	370 (867)	1.32×10 ⁻²¹
	代谢途径	568 (1 496)	1.85×10 ⁻²¹
	角质、琥珀和蜡的生物合成	25 (30)	2.78×10 ⁻⁸
	苯丙素类化合物生物合成	119 (263)	7.93×10 ⁻⁸
灌浆期/抽穗期	脂肪酸延长	31 (46)	1.25×10 ⁻⁶
	植物激素信号转导	111 (260)	8.74×10 ⁻⁶
	氨基糖和核苷酸糖代谢	60 (122)	1.72×10 ⁻⁵
	谷胱甘肽代谢	48 (92)	2.18×10 ⁻⁵
	β-氨基丙酸代谢	20 (32)	9.23×10 ⁻⁴
	淀粉和蔗糖的代谢	47 (102)	1.47×10 ⁻³

由于篇幅限制, 本表仅列出富集度前10的KEGG条目。如需要全部结果请联系通信作者。基因数一栏, 括号前为该途径差异表达的基因数目, 括号内为谷子基因组中该途径总的基因数目; 表5~7同。

0016705)、催化活性(GO:0003824)、单加氧酶活性(GO:0004497)、血红素结合(GO:0020037)、酶抑制剂活性(GO:0004857)和葡糖基转移酶活性(GO:0046527)等GO条目富集程度较高; 细胞组分中有301个GO条目是3个比较组间共同富集到的通路, 其中膜结合囊泡(GO:0031988)、细胞质囊泡(GO:0031410)、非原质体(GO:0048046)、细胞壁(GO:0005618)、胞外区(GO:0005576)和质膜的固

定组分(GO:0046658)等GO条目富集程度较高。

3 讨论

谷子穗发育过程主要包括抽穗期、开花期、灌浆期和成熟期等4个时期, 其中前3个时期对穗形态发育和产量影响较大。本研究利用RNA-Seq技术检测了这三个时期差异表达的基因, 初步揭示了谷子穗发育过程的基因表达动态图谱。

表5 开花期/抽穗期富集的生物过程、分子功能和细胞组分GO条目
Table 5 GO terms enrichment for biological processes, molecular function, and cellular component of flowering stage/heading stage

功能分类	GO条目	基因数	Q值
生物过程	GO:0043086, 催化活性负调节	58 (174)	2.33×10^{-9}
	GO:0044092, 分子功能负调节	60 (185)	2.33×10^{-9}
	GO:0006468, 蛋白磷酸化	245 (1 428)	1.67×10^{-4}
	GO:0071554, 细胞壁组织或生物发生	95 (449)	1.67×10^{-4}
分子功能	GO:0006026, 氨基聚糖分解代谢过程	12 (20)	2.09×10^{-4}
	GO:0004553, 水解酶活性, 水解邻糖基化合物	122 (439)	2.60×10^{-13}
	GO:0016798, 水解酶活性, 作用于糖基键	131 (493)	4.95×10^{-13}
	GO:0030599, 果胶酯酶活性	36 (76)	1.81×10^{-10}
	GO:0003824, 催化活性	1 626 (11 234)	4.39×10^{-8}
	GO:0016746, 转移酶活性, 转移酰基	94 (381)	1.80×10^{-7}
细胞组分	GO:0016023, 细胞质、膜结合囊泡	948 (4 403)	1.21×10^{-82}
	GO:0031410, 细胞质囊泡	948 (4 404)	1.21×10^{-82}
	GO:0031988, 膜结合囊泡	948 (4 411)	2.13×10^{-82}
	GO:0031982, 囊泡	948 (4 416)	3.20×10^{-82}
	GO:0005576, 胞外区	195 (693)	2.05×10^{-27}

由于篇幅限制, 本表仅列出富集度前5的条目, 如需要全部结果请联系通信作者; 表6和7同。

表6 灌浆期/开花期富集的生物过程、分子功能和细胞组分GO条目
Table 6 GO Terms enrichment for biological processes, molecular function, and cellular component of filling stage/flowering stage

功能分类	GO条目	基因数	Q值
生物过程	GO:0044092, 分子功能负调节	81 (185)	3.17×10^{-8}
	GO:0043086, 催化活性负调节	77 (174)	3.17×10^{-8}
	GO:0007018, 微管运动	39 (65)	3.17×10^{-8}
	GO:0006928, 细胞或亚细胞组分的运动	44 (81)	1.31×10^{-7}
	GO:0045229, 外部囊状结构组织	103 (268)	2.68×10^{-7}
分子功能	GO:0016705, 氧化还原酶活性, 作用于配对供体, 与分子氧结合或还原	192 (521)	3.03×10^{-12}
	GO:0004553, 水解酶活性, 水解邻糖基化合物	165 (439)	2.03×10^{-11}
	GO:0016798, 水解酶活性, 作用于糖基键	176 (493)	3.92×10^{-10}
	GO:0003777, 微管马达运动	38 (58)	4.00×10^{-10}
	GO:0008017, 微管结合	47 (84)	2.96×10^{-9}
	GO:0031988, 膜结合囊泡	1 341 (4 411)	2.23×10^{-53}
细胞组分	GO:0031982, 囊泡	1 342 (4 416)	2.23×10^{-53}
	GO:0031410, 细胞质囊泡	1 338 (4 404)	2.91×10^{-53}
	GO:0016023, 细胞质、膜结合囊泡	1 337 (4 403)	3.50×10^{-53}
	GO:0005576, 胞外区	277 (693)	7.07×10^{-27}

表7 灌浆期/抽穗期富集的生物过程、分子功能和细胞组分GO条目

Table 7 GO Terms enrichment for biological processes, molecular function, and cellular component of filling stage/heading stage

功能分类	GO条目	基因数	Q值
生物过程	GO:0007018, 微管运动	52 (65)	8.62×10^{-14}
	GO:0006928, 细胞或亚细胞组分的运动	56 (81)	1.75×10^{-10}
	GO:0009719, 内源性刺激反应	419 (1 090)	6.36×10^{-9}
	GO:0009725, 激素应答	404 (1 046)	6.36×10^{-9}
	GO:0080167, karrikin响应	81 (152)	2.33×10^{-7}
分子功能	GO:0003777, 微管马达运动	47 (58)	6.51×10^{-13}
	GO:0008017, 微管结合	58 (84)	4.01×10^{-11}
	GO:0015631, 微管蛋白结合	60 (90)	1.13×10^{-10}
	GO:0016757, 转移酶活性, 转移糖基	306 (737)	1.15×10^{-10}
	GO:0016705, 氧化还原酶活性, 作用于配	227 (521)	4.42×10^{-10}
	对供体, 与分子氧结合或还原		
细胞组分	GO:0031988, 膜结合囊泡	1 605 (4 411)	7.16×10^{-35}
	GO:0031982, 囊泡	1 606 (4 416)	7.16×10^{-35}
	GO:0031410, 细胞质囊泡	1 601 (4 404)	8.64×10^{-35}
	GO:0016023, 细胞质、膜结合囊泡	1 600 (4 403)	8.96×10^{-35}
	GO:0005576, 胞外区	313 (693)	4.48×10^{-19}

抽穗期是受多基因调控的数量性状, 目前大量基因已被克隆。本研究发现受体光敏色素A、光敏色素相关蛋白、光周期途径的枢纽*CO*基因以及成花素*FT*基因等在抽穗开花期差异表达(表3)。这些基因可能与抽穗开花调控有关, 在谷子生态适应性遗传改良中有潜在的应用价值。小米中富含多种碳水化合物、蛋白质、脂肪、维生素和矿物质, 其中碳水化合物的质量分数约为74.61%, 淀粉的质量分数平均为71%。本研究表明, 颗粒结合淀粉合成酶基因(*Si4g02980*)、可溶性淀粉合成酶基因(*Si0g16430*)、淀粉靶向蛋白基因(*Si1g09890*)、 α -淀粉酶样蛋白基因(*Si3g02050*)、糖原/淀粉合成酶基因(*Si4g07100*)、异淀粉酶1基因(*Si6g21130*)、极限糊精酶基因(*Si7g02800*)和蔗糖合酶3基因(*Si9g40320*)在灌浆期表达量显著升高(图2-F和表3), 这些基因在谷子灌浆期起着重要的作用, 可能与产量性状紧密相关。而KEGG富集结果表明, 次生代谢物、苯丙素类化合物、脂肪酸、糖类物质、角质、丝氨酸和蜡质等物质也参与谷子穗发育的调控。

植物激素, 如生长素、茉莉酸、细胞分裂素和

脱落酸等在植物穗发育过程中起着极其重要的作用。在水稻中, 生长素输入载体基因*LAX1*发生突变后导致水稻的穗部枝梗数和穗粒数减少(Komatsu等2001)。细胞分裂素在灌浆前期促进籽粒灌浆的速率和增加籽粒的质量, 在灌浆后期延缓籽粒灌浆时间。水稻粒数基因*GRAIN NUMBER1 (GNI)*编码细胞分裂素氧化酶, 降低该基因表达量后可提高籽粒产量(Ashikari等2005)。谷子穗发育早期生长素响应基因以及生长素转运蛋白、细胞分裂素受体等基因表达量较高, 而在穗发育后期表达量迅速下降, 这可能与生长素和细胞分裂素在穗发育早期穗快速生长以及穗部枝梗发育的调控有关。

转录因子在禾本科穗发育过程中起着极其重要的作用。水稻SBP-box家族转录因子*SPL6*在穗发育过程中呈现顶端高水平表达样式, 其功能缺失导致顶端小穗细胞衰老退化和穗“秃顶”性状产生(Wang等2018)。最近, 刁现民课题组发现谷子MADS-box转录因子*SiMADS34*调控一级枝梗的发育, 该基因突变后一级分枝明显变长, 导致穗子变粗(Hussin等2021)。我们也发现谷子抽穗期、开花期和灌浆期有803个转录因子差异表达, 包括MADS-

box、WRKY、MYB和AP2等。深入揭示穗发育过程差异表达转录因子的功能将有利于解析谷子穗发育的转录调控网络。

总之, 谷子穗发育是一个极其复杂的过程, 受到众多基因的调控。谷子抽穗期、开花期和灌浆期的基因表达动态图谱可为后续筛选和鉴定穗发育关键基因提供参考。未来可以利用CRISPR/Cas9基因编辑技术在*xiaomi*中研究这些基因的功能, 深入揭示穗发育调控的分子机制。

参考文献(References)

- Ashikari M, Sakakibara H, Lin S, et al (2005). Cytokinin oxidase regulates rice grain production. *Science*, 309 (5735): 741–745
- Bennetzen JL, Schmutz J, Wang H, et al (2012). Reference genome sequence of the model plant *Setaria*. *Nat Biotechnol*, 30 (6): 555–561
- Bolger AM, Lohse M, Usadel B (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, 30 (15): 2114–2120
- Doust AN, Kellogg EA, Devos KM, et al (2009). Foxtail millet: a sequence-driven grass model system. *Plant Physiol*, 149 (1): 137–141
- Feng N, Song G, Guan J, et al (2017). Transcriptome profiling of wheat inflorescence development from spikelet initiation to floral patterning identified stage-specific regulatory genes. *Plant Physiol*, 174 (3): 1779–1794
- Huang P, Jiang H, Zhu C, et al (2017). *Sparse panicle1* is required for inflorescence development in *Setaria viridis* and maize. *Nat Plants*, 3 (5): 17054
- Huang X, Qian Q, Liu Z, et al (2009). Natural variation at the *DEP1* locus enhances grain yield in rice. *Nat Genet*, 41 (4): 494–497
- Hussin SH, Wang H, Tang S, et al (2021). *SiMADS34*, an E-class MADS-box transcription factor, regulates inflorescence architecture and grain yield in *Setaria italica*. *Plant Mol Biol*, 105 (4–5): 419–434
- Jia H, Li M, Li W, et al (2020). A serine/threonine protein kinase encoding gene *KERNEL NUMBER PER ROW6* regulates maize grain yield. *Nat Commun*, 11 (1): 988
- Kim D, Langmead B, Salzberg SL (2015). HISAT: a fast spliced aligner with low memory requirements. *Nat Methods*, 12 (4): 357–360
- Komatsu M, Maekawa M, Shimamoto K, et al (2001). The *LAX1* and *FRIZZY PANICLE 2* genes determine the inflorescence architecture of rice by controlling rachis-branch and spikelet development. *Dev Biol*, 231 (2): 364–373
- Li B, Ruotti V, Stewart RM, et al (2010). RNA-Seq gene expression estimation with read mapping uncertainty. *Bioinformatics*, 26 (4): 493–500
- Li W, Zhi H, Zhang S, et al (2015). Morphological effect and genomic mapping of *Si-SP1* (small panicle 1) in foxtail millet. *J Plant Genet Resour*, 16 (3): 581–587 (in Chinese with English abstract) [李雯, 智慧, 张硕等(2015). 谷子 *Si-SP1* 小穗突变基因的遗传分析和定位. 植物遗传资源学报, 16 (3): 581–587]
- Wang QL, Sun AZ, Chen ST, et al (2018). SPL6 represses signalling outputs of ER stress in control of panicle cell death in rice. *Nat Plants*, 4 (5): 280–288
- Wang T, Song H, Li P, et al (2020). Transcriptome analysis provides insights into grain filling in foxtail millet (*Setaria italica* L.). *Int J Mol Sci*, 21 (14): 5031
- Xiang J, Tang S, Zhi H, et al (2017). *Loose Panicle1* encoding a novel WRKY transcription factor, regulates panicle development, stem elongation, and seed size in foxtail millet [*Setaria italica* (L.) P. Beauv.]. *PLOS One*, 12 (6): e0178730
- Xiang JS, Zhi H, Yong H, et al (2020). Phenotypic analysis of panicle development mutants of foxtail millet [*Setaria italica* (L.) P. Beauv.]. *Acta Agric Boreali-Occident Sin*, 29 (11): 1603–1614 (in Chinese with English abstract) [相吉山, 智慧, 永辉等(2020). 几个谷子穗发育突变体的表型分析. 西北农业学报, 29 (11): 1603–1614]
- Xue HL, Yang JJ, Tang S, et al (2018). Morphological characterization and gene mapping of a panicle apical abortion mutant (*sipaa 1*) in foxtail millet. *Sci Agric Sin*, 51 (9): 1627–1640 (in Chinese with English abstract) [薛红丽, 杨军军, 汤沙等(2018). 谷子穗顶端败育突变体 *sipaa1* 的表型分析和基因定位. 中国农业科学, 51 (9): 1627–1640]
- Yang Z, Zhang H, Li X, et al (2020). A mini foxtail millet with an *Arabidopsis*-like life cycle as a *C₄* model system. *Nat Plants*, 6 (9): 1167–1178
- Zafar SA, Patil SB, Uzair M, et al (2020). *DEGENERATED PANICLE AND PARTIAL STERILITY 1 (DPS1)* encodes a cystathionine β -synthase domain containing protein required for anther cuticle and panicle development in rice. *New Phytol*, 225 (1): 356–375
- Zhao DS, Ding AQ, Feng GN, et al (2020). Identification of a panicle development mutant dependent on temperature in rice. *Plant Physiol J*, 56 (10): 2149–2158 (in Chinese with English abstract) [赵冬生, 丁蔼秋, 封功能等(2020). 一个受温度调控的水稻穗发育突变体. 植物生理学报, 56 (10): 2149–2158]