

# Tissue-Specific Expression Profiling of Seedling Stage in Early-Maturity Mutant Induced by Carbon Ion Beam in Sweet Sorghum

Xicun Dong<sup>1,†</sup>, Xia Yan<sup>2,3,†</sup>, Wenjian Li<sup>1</sup>, Ruiyuan Liu<sup>1</sup> & Wenting Gu<sup>1</sup>

<sup>1</sup> Department of Radiobiology, Institute of Modern Physics, Chinese Academy of Sciences, Lanzhou, China

<sup>2</sup> Key Laboratory of Inland River Ecohydrology, Cold and Arid Regions Environmental and Engineering Research Institute, Chinese Academy of Sciences, Lanzhou, China

<sup>3</sup> Key Laboratory of Stress Physiology and Ecology in Cold and Arid Regions of Gansu Province, Cold and Arid Regions Environmental and Engineering Research Institute, Chinese Academy of Sciences, Lanzhou, China

Correspondence: Xicun Dong, Radiobiology Department, Institute of Modern Physics, Chinese Academy of Sciences, Lanzhou, China. Tel: 86-0931-496-9175. E-mail: dongxicun@impcas.ac.cn

Received: September 27, 2016

Accepted: October 27, 2016

Online Published: December 15, 2016

doi:10.5539/jas.v9n1p49

URL: <http://dx.doi.org/10.5539/jas.v9n1p49>

† These authors contributed equally to this work.

The research is financed by the STS project (KFJ-EW-ST-086) and Western Light Co-scholar (29Y406020) Program of the Chinese Academy of Sciences.

## Abstract

An early-maturity mutant KFJT-1 has been screened out after carbon ion irradiation in sweet sorghum (*Sorghum bicolor* (L.) Moench). In this study, tissue specific digital gene expression analysis was performed between the KFJT-1 mutant and the wild type KFJT-CK at seedling stage. The results showed that a total of 717, 2160 and 2,331 tags-mapped genes were differently expressed in roots, stems and leaves of young seedling, respectively. In KFJT-1, 557 (77.7%) genes were up-regulated and 160 (22.3%) genes were down-regulated in young root; 1,232 (57.0%) genes were up-regulated and 928 (43.0%) were down-regulated in young stem; and 1,577 (67.7%) genes were up-regulated and 754 (32.3%) genes were down-regulated in young leaf. Functional annotation revealed that most induced genes functioned as “binding”, “synthase activity”, “transferase” and “transporter activity” which involved in the biological processes of metabolic and response to stimulus. Surprisingly, the up-regulated genes in KFJT-1 were classified into four KEGG pathways: “alpha-Linolenic acid metabolism”, “flavonoid biosynthesis”, “inositol phosphate metabolism” and “fatty acid biosynthesis”, which related to the stress resistance and supported the outstanding agronomic traits of KFJT-1 in the process of plant growth and development. Among the DEGs, a critical photoreceptor from photoperiod pathway *PHYA* gene was significantly up-regulated in leaf and root of KFJT-1, suggesting the mutation could occur on the genomic upstream of *PHYA*. This work may provide helpful insights to further understand the mutation mechanism in sweet sorghum.

**Keywords:** tissue-specific expression profiling, young seedling, different expressed genes, sweet sorghum

## 1. Introduction

Sweet sorghum (*Sorghum bicolor* (L.) Moench) is a useful energy crop because of high photosynthetic efficiency, high biomass- and sugar- yielding (Billa et al., 1997). However, being a short-day plant, the grains cannot mature under long day condition. We previously isolated an early-maturity mutant KFJT-1 from wild type plants KFJT-CK by heavy ion beam irradiation. Resistance experiment showed that proline content was increased by 11.05% with drought stress, which showed that the tolerance of KFJT-1 to the stress is advantage to KFJT-CK (Dong & Li, 2012).

The biological effects of heavy-ion radiation encompass a wide range of alterations, including developmental abnormalities (Kranz, 1994), chromosomal aberrations (Kawat et al., 2001; Kikuchi et al., 2009; Wei et al., 2006) and genomic structural variation (Mei et al., 2011; Xu et al., 2006). Many studies has shown that the carbon ions beam induce more effective structural alterations in DNA than other radiation (Shikazono et al., 2005), sequence

analyses of radiation-induced mutations have been widely carried out in plants (Bruggenmann et al., 1996; Shikazono et al., 2000; Shikazono et al., 2003). These genetic variations could directly induce expression differentiation of the plenty of genes which involved in the biological processes. Nowadays, Digital gene expression (DGE) tag profiling has been widely utilized to monitor the differences in transcriptional level to elucidate the genome-wide expression profiling among different tissues and organs. It directly quantify the transcript abundance of the uniquely tagged corresponding genes with ultra-high-throughput sequencing of cDNA fragments (Hong et al., 2011), which could be conveniently detected for the organisms without prior annotations of genomic information, such as cotton (Wei et al., 2013), spruce (Albouyeh et al., 2010), *Brassica napus* (Jiang et al., 2013), and moss (Nishiyama et al., 2012).

This study aimed to gain comprehensive understanding of DGEs in KFJT-1 compared to KFJT-CK at seedling stage and improve our current understanding of the molecular mechanism of KFJT-1 induced by carbon ions beam.

## 2. Materials and Methods

### 2.1 Plant Materials and Growth Conditions

The dry seeds of equal size from KFJT-1 and KFJT-CK, which showed no moldy and lesion, were selected and placed in a 90 mm Petri dish containing double-layer wet filter paper, respectively. The seeds were germinated at  $25 \pm 2$  °C in a growth chamber under a 16 h light photoperiod provided by fluorescent light tubes ( $50 \mu\text{mol m}^{-2}\text{s}^{-1}$ ). Each genotype was replicated three times and 100 seeds were employed for each replication. After 30 days, the samples of roots, stems and leaves were harvested, respectively and quickly frozen in liquid nitrogen for RNA isolation.

### 2.2 RNA Isolation and Library Preparation for DGE

RNA extraction was performed according to the manufacturer's instructions of TRIzol reagent (Invitrogen, USA), followed by RNase-free DNase treatment (TaKaRa, Dalian, China). The total RNA was checked for quality and quantity using a Biophotometer Plus (Multiskan Spectrum, German), and a minimum of 6  $\mu\text{g}$  of total RNA was used for Illumina sequencing. The total RNA samples isolated from the three parallels tissues were pooled for libraries preparation, in which RC and RF, SC and SF, LC and LF represented the transcripts of roots, stems and leaves from control KFJT-CK ("C" characterized) and mutant KFJT-1 ("F" characterized), respectively.

Over 6  $\mu\text{g}$  from each total RNA samples were constructed as the DGE libraries using Illumina gene expression kit (Illumina Inc; San Diego, CA, USA) according to the manufacturer's protocol (version 2.1B), mRNA was purified using biotin-Oligo (dT) magnetic bead adsorption. The first- and second-strand cDNA synthesis was performed after the RNA was bound to the beads. The double stranded cDNA were digested with *NlaIII* to produce cohesive end. After purification with Dynabeads, and the digestion was ligated to GEX adapter 1 which contains *MmeI* restriction CATG site, and downstream 17 bp then cut with the *NlaIII*. The 21 bp tags containing adapter I were ligated to GEX adapter 2 to generate a tag library. These tag fragments were amplified by liner PCR for 15 cycles using PCR primers anneral to the adapter ends. The 85 bp amplicons were seperated on 6% TBE PAGE gel, purified and denatured to produce single strand molecules. These molecules were anchored to Solexa sequencing array and sequenced on Illumina GA II at BGI- Shenzhen, Shenzhen, China. Raw sequence data were generated by Illumina pipeline.

### 2.3 Sequence Annotation and DGEs Pathways Identification

Raw sequences were transformed into clean tags by filtering off adapter-only tags and low-quality tags as described (Li et al., 2013). All the clean tags were mapped to the reference sequences of *Sorghum bicolor* and only 1 bp of mismatch was considered. The remaining clean tags were designed as unambiguous clean tags. In order to compare the expression abundance among the samples, the number of unambiguous clean tags for each gene was calculated and then normalized to TPM (number of transcripts per million clean tags). The final assembled transcripts ( $\geq 100$  bp) were submitted for homology and annotation searches using Blast2GO software v2.4.4 (Wei et al., 2013). For BLASTX against the NR database, the threshold was set to E-value lower than  $10^{-5}$ . However, most of the gene information of sorghum was hypothetical or putative. Therefore, all the putative sorghum genes were BlastX with was performed against *Sorghum* genes. GO classification was achieved using WEGO software together with David Bioinformatics Resources 6.7 (<http://david.abcc.ncifcrf.gov/home.jsp>). Enzyme codes were extracted and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways were retrieved from KEGG web server (<http://www.genome.jp/kegg/>). We used a rigorous algorithm to identify differentially expressed genes between the KFJT-1 and KFJT-CK in this study. FDR ranking, FDR (False Discovery Rate) was used applied to adjust the p-value in multiple tests and analyses (Qin et al., 2011). The transcripts with at least two-fold

differences (absolute values of  $\log_2$  (Ratio)  $\geq 1$  with FDR < 0.001) were regarded as significantly different expressed genes.

#### 2.4 Real-Time Quantitative RT-PCR (qRT-PCR) Analysis

Real-time quantitative RT-PCR (qRT-PCR) analysis was used to verify the DGE results. The RNA samples used for the qRT-PCR assays were the same as in the DGE experiments. Gene specific primers were designed according to the reference unigene sequences using Primer Premier 5.0. Seven genes were selected from the DEGs for quantitative qRT-PCR assays. QRT-PCR was performed according to the manufacturer's specifications (). The following SYBR Green PCR cycling conditions were used: denaturation at 95 C for 10 s, followed by 40 cycles of 94 C for 5 s and 60 C for 20 s. The PCR experiments were performed using an iQ 5 Multicolor real-time PCR detection system (BioRAD, USA). Sorghum actin gene (forward: GCCGAGCGAGAAATTGTAAG and reverse: ATCATGGATGGCTGGAAGAG) was used as a normalizer. The relative gene expression levels were calculated using  $2^{-\Delta\Delta CT}$ .

### 3. Result

#### 3.1 Construction of Digital Gene Expression (DGE) Library for KFJT-1 and KFJT-CK at Seedling Stage

To obtain a global view of the tissue specific characteristics at the transcriptional level between KFJT-1 and KFJT-CK at seedling stage, total six DGE libraries from roots, stems and leaves were sequenced with Solexa/Illumina DGE analysis, respectively. Among the libraries, we got the total numbers of tags ranging from 5.8 to 6.7 million, which composing distinct tags with 125840 and 131973, 140208 and 156400, 147227 and 137977 in young roots, stems and leaves libraries for KFJT-CK and KFJT-1, respectively (Table 1). The number of the tags and unambiguous tags mapping to genes was almost the same for about 70%, for example in RC, 62208 distinct tags (72.07% of clean tag), 62089 unambiguous tags (71.99% of clean tag) was mapping to gene, which means that the tags was well matched to the specific genes. Those distinct tags matched to the genes occupied about 50% of the clean tags. The distribution of total clean tags and distinct clean tags over different tag-abundance categories were shown in Figures 1A and 1B. In terms of the total clean tags, the percentage of 2-5 copies ranged from 3.31-3.89%, 6-10 copies from 2.3-2.71%, 11-20 copies from 3.20-3.81%, 21-50 copies from 6.44-7.47%, 51-100 copies from 7.7-8.83%. The largest percentage was constituted by the copies over 100 which ranged from 73.58-76.90%. However, compared to the total clean reads, the distinct clean tags displayed different distribution among all the six DGE libraries. The largest proportion was constituted by the 2-5 copies for about 54%. The second largest part was the 6-10 copies which occupied about 14%, approximately 10% had copy numbers higher than 100. The smallest proportion was constituted by 51-100 copies for about 7%. The numbers of the tag-mapped genes or unambiguous tag-mapped gene were decreased sharply compared to the distinct tags. For example, the number of matched gene in RC was 15768, however, the distinct tags was 62208. Finally, 15786 (RC), 16007 (RF), 16179 (SC), 16780 (SF), 15378 (LC) and 15200 (LF) tag-mapped genes were generated against sorghum reference genome between KFJT-1 and KFJT-CK. The saturation analysis showed that the number of the genes was not increased proportionally with the number of sequences (total tag number) when the sequencing counts reached 4M (Figure S1). Thus, these tag-mapped genes were completely satisfying the further analysis.

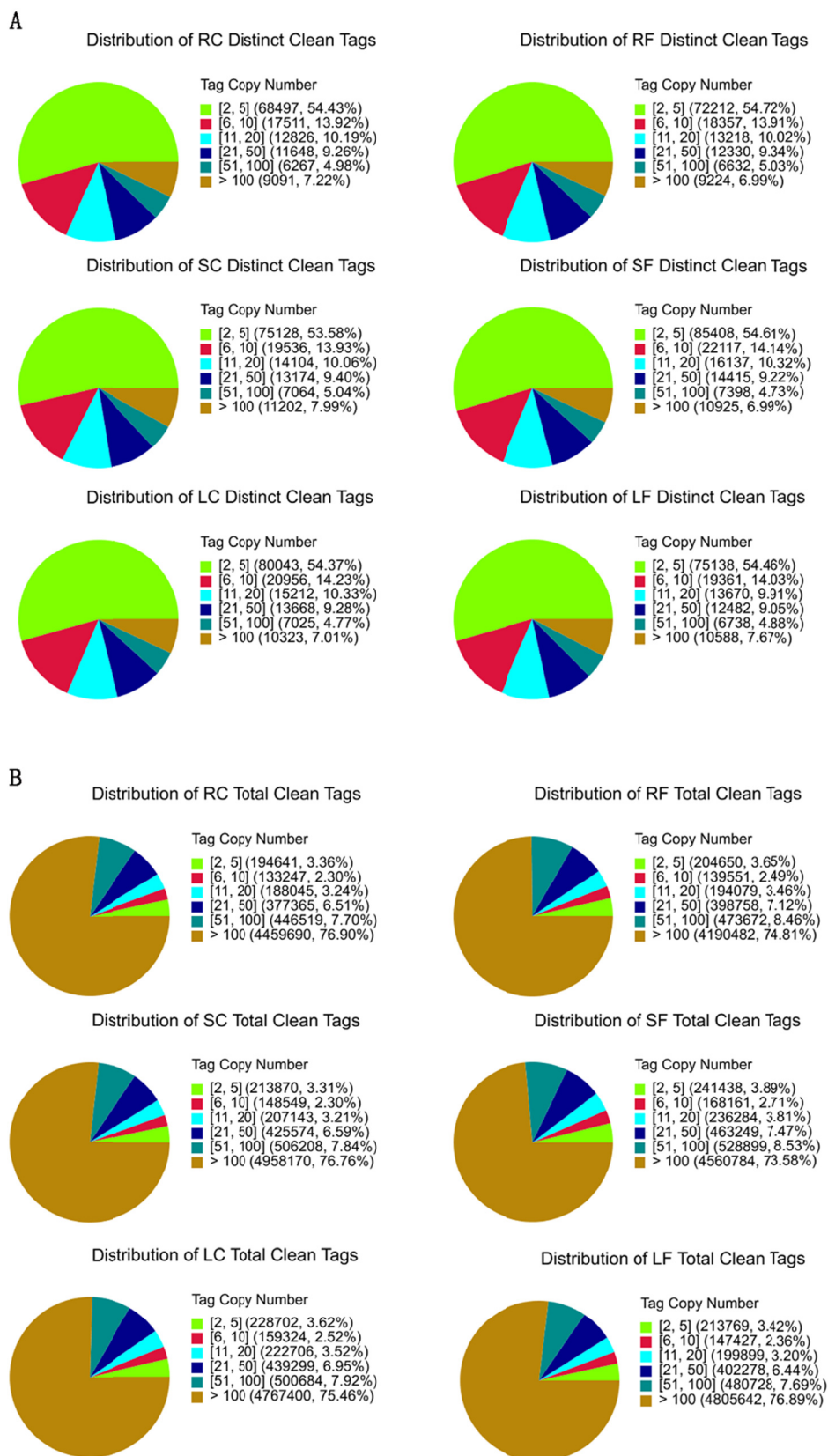


Figure 1. Distribution of distinct clean tags and total clean tags. (A) Distribution of distinct clean tags. (B) Distribution of total clean tags. RC, SC, LC, and RF, SF, LF represented roots, stems and leaves of wild type and mutant, respectively. Numbers within square brackets indicate the range of copy number for a specific category of tags. Numbers within parentheses indicate the types of tag in A or copy number of tags in B

Table 1. Statistics of DGE sequencing

Summary		Roots		Stems		Leaves	
		RC	RF	SC	SF	LC	LF
Raw data	Total	6091059	5890651	6746137	6493952	6694215	6603642
	Distinct tag	298254	305052	333748	343180	341594	328619
Clean tag	Total number	5799507	5601192	6459514	6198815	6318115	6249743
	Distinct tag number	125840	131973	140208	156400	147227	137977
All tags mapping to gene	Total % of clean tag	72.07%	71.15%	71.61%	71.10%	66.34%	68.92%
	Distinct tag number	62208	62402	66229	75306	68097	65630
	Distinct tag % of clean tag	49.43%	47.28%	47.24%	48.15%	46.25%	47.57%
Unambiguous tags mapping to gene	Total % of clean tag	71.99%	71.08%	71.54%	71.00%	66.29%	68.87%
	Distinct tag number	62089	62276	66098	75159	67968	65523
	Distinct tag % of clean tag	49.34%	47.19%	47.14%	48.06%	46.17%	47.49%
All tag-mapped genes	Number	15786	16007	16179	16780	15378	15200
	% of ref genes	53.61%	54.36%	54.94%	56.98%	52.22%	51.62%
Unambiguous tag-mapped genes	Number	15747	15963	16132	16730	15327	15166
	% of ref genes	53.47%	54.21%	54.78%	56.81%	52.05%	51.50%

### 3.2 Tissue-Specific Gene Expression in the Development of the Seedling between KFJT-1 and KFJT-CK

To compare differential expression patterns between KFJT-1 and KFJT-CK, we normalized tag distribution for gene expression level in each library to make an effective library size and extracted significance of differentially expressed transcripts (DETs) with  $FDR \leq 0.05$  and  $\log_2$  fold-change  $\geq 1$  by edgeR (Empirical analysis of Digital Gene Expression in R). The regulated genes were shown in Figure 2. The red dots and green dots represent transcripts higher or lower in abundance for more than two fold, and the blue dots represented the transcripts that differed less than two fold between the KFJT-1 and the wild type. In root, a total of 557 genes were up-regulated (77.7%, red dot in Figure 2A) and 160 genes were down-regulated (22.3%, green dot in Figure 2A). In stem, a total of 1,232 genes were up-regulated (57.0%, red dot in Figure 2B) and 928 genes were down-regulated (43.0%, green dot in Figure 2B). In leaf, total 1,577 genes were up-regulated (67.7%, red dot in Figure 2C) and 754 genes were down-regulated (32.3%, green dot in Figure 2C). An increasing trend in the number of differently expressed genes was observed in young stem and leaf compared to young root. The total numbers of the tags-mapped genes were 717, 2160 and 2331 in root, stem and leaf, respectively (Figure 3A). Venn analysis revealed that 66 genes were differently expressed in the all young root, stem and leaf. In addition, 219 genes were differently expressed both in leaf and root, 489 genes in leaf and stem, and 148 genes in root and stem. Numbers of 284, 1457 and 1557 genes were differentially expressed specific to the root, stem and leaf, respectively (Figure 3B). To be mentioned, about 99% unique tags were expressed within five-fold difference (red bar in Figure 4) between KFJT-1 and KFJT-CK, covered 99.85% in young root, 99.39% in young stem and 98.77% in young leaf, respectively. Only 0.1-1.01% of the DEGs over five folds was up-regulated (green bar in Figure 4), while 0.05-0.23% was down-regulated.

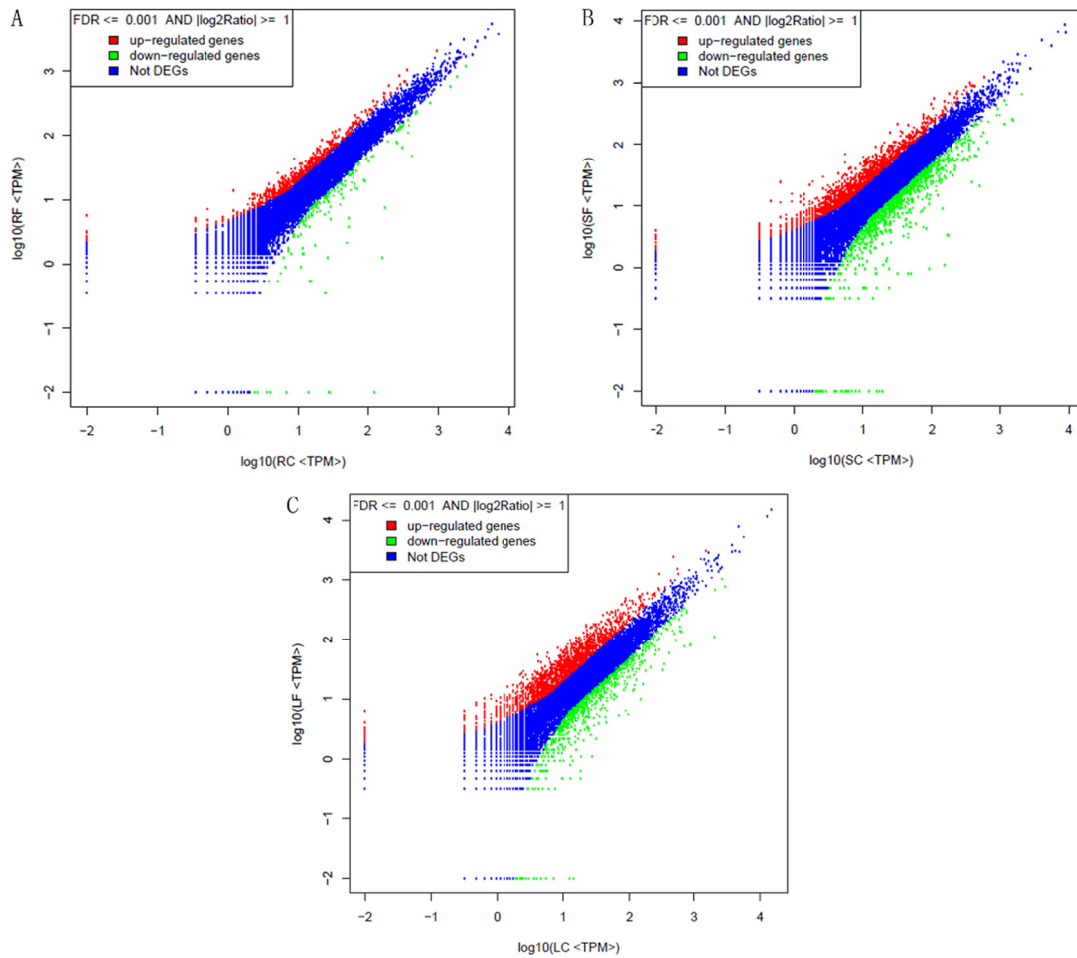


Figure 2. Comparison of gene expression between different libraries. Blue dots represent the transcripts with no significant expression. Red dots and green dots represent transcripts more abundant between early-maturity mutant and wild type in roots, stems and leaves, respectively. “FDR < 0.001” and “absolute value of log2 Ratio > 1” were used as the thresholds to judge the significance of gene expression difference

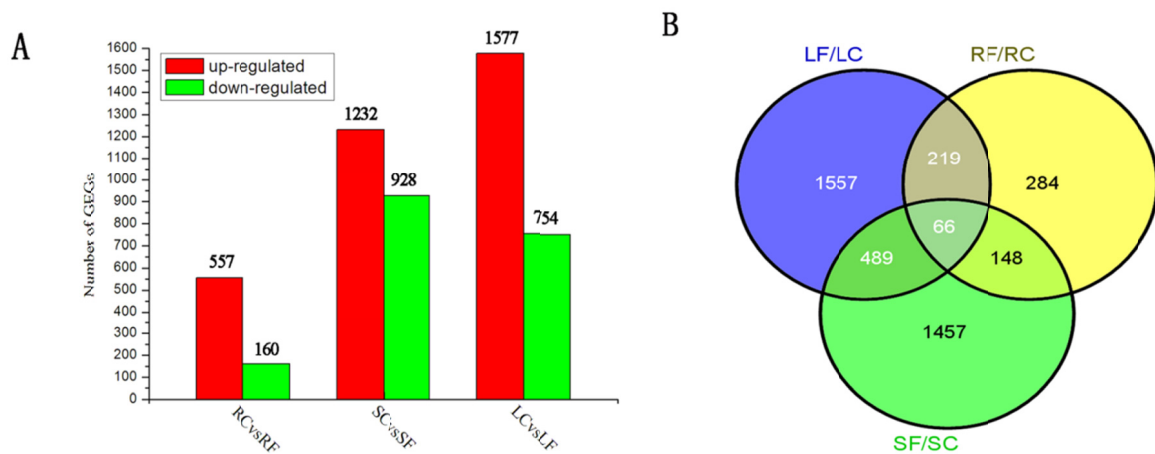


Figure 3. Distribution of differentially expressed genes. (A) Changes in gene expression profile of roots, stems and leaves. (B) Venn diagram to illustrate the number of genes regulated in the roots, stems and leaves between EM mutant and the wild type in sweet sorghum

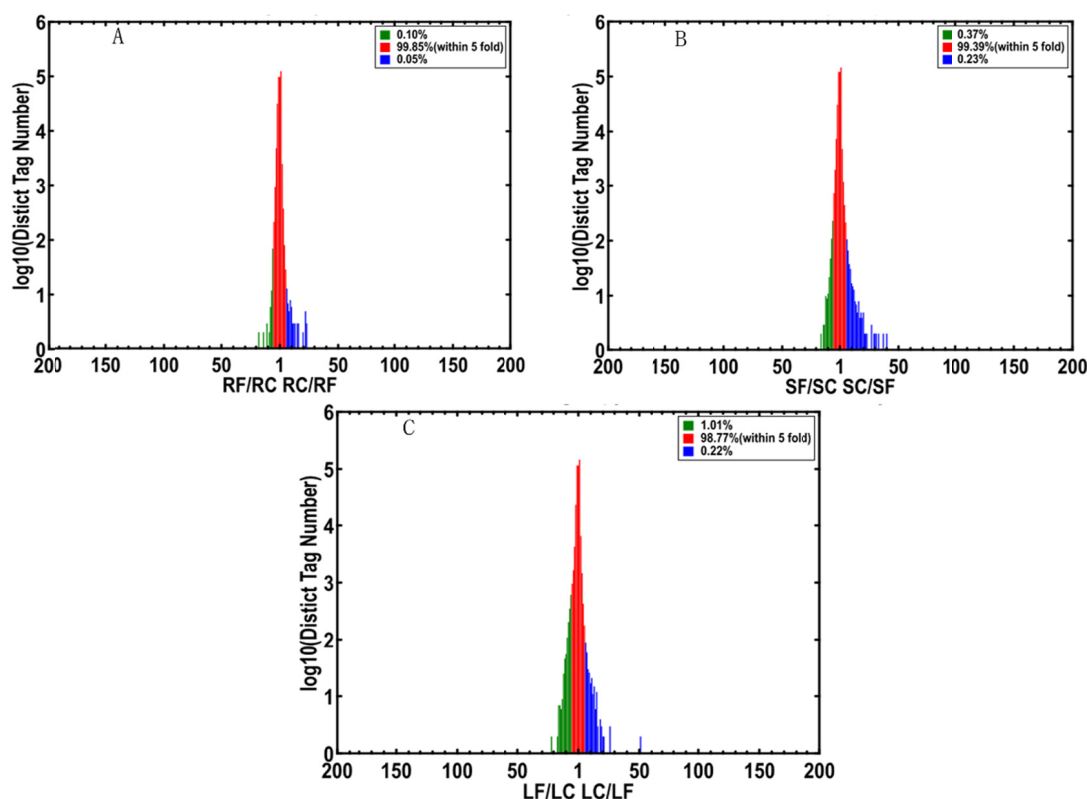


Figure 4. Tags with different expression in early-maturity compared to wild type sample. Red region represents the differentially expressed tags with differentia expressionless than 5 folds. Blue and green region represent the up- and down-regulated tags for more than 5 folds, respectively

### 3.3 Functional Analysis of the Differentially Expressed Genes at Seedling Stage

Gene ontology (GO) assignments were used to classify the functions of differentially expressed genes with at least two-fold differences (Figure 5). According to biological process, 193 genes were related to “small molecule metabolic process” with p-value lower than 0.05. Accordingly, 55 genes related to “structural molecule activity” for molecular function, and total 1131 genes cluster into main 9 category were over-represented with p-value lower than 0.05 (Table 2). Regarding to regulated genes was about 99% within five fold, we used genes with expression level over five fold as more strict criteria to assign GO annotation on DAVID Bioinformatics Resource 6.7. Among the DEGs over five folds, 331 genes were related to “metabolic process” and 148 related to “response to stimulus”, respectively. The genes response to stimulus including “cold”, “water deprivation”, “salt”, “wounding” and “osmotic” stress. Furthermore, the term “fatty acid biosynthetic process” which involve in stress resistance was also enriched, which suggested that KFJT-1 mutants were more adaptable to stress environment than KFJT-CK since seedling stage. “Lignin biosynthetic process” also enriched to support high compact stem structure in KFJT-1. Genes involved flowering time were also enriched such as “response to far red light”, “pollen development” including “auxin transport”, which consist with the short growth cycle of KFJT-1 compared to KFJT-CK. Specially, in young leaf, “polysaccharide metabolic process”, “carboxylic acid biosynthetic process” and “developmental maturation” founded which suggested that KFJT-1 fixed higher carbon than KFJT-CK. In young root, genes involved in “nitrate metabolic process” were high enriched, which suggested that KFJT-1 harbor more potential than KFJT-CK in the nitrogen fixation to facilitate development in barren soil. In stem, besides of genes involved in stress response, “sugar mediated signaling” and “response to carbohydrate stimulus” were enriched which might be correlated with the high sugar content in KFJT-1. According to the molecular function, those genes were related to the terms “binding”, “synthase activity”, “transferase” and “transporter activity”, which facilitate to the biological process above.

Table 2. List of first twenty pathways for DEGs

Table 2A. List of first twenty pathways in root

Pathway term (RC vs RF)	Pathway ID	DEGs tested	P value	Q value
Glutathione metabolism	ko00480	15 (3.23%)	0.0000858	0.009091637
Sulfur metabolism	ko00920	5 (1.08%)	0.009928955	0.442890563
Glycolysis/Gluconeogenesis	ko00010	12 (2.59%)	0.01419812	0.442890563
Stilbenoid, diarylheptanoid and gingerol biosynthesis	ko00945	14 (3.02%)	0.02643276	0.442890563
Sesquiterpenoid and triterpenoid biosynthesis	ko00909	4 (0.86%)	0.02738994	0.442890563
Biosynthesis of secondary metabolites	ko01110	82 (17.67%)	0.02848042	0.442890563
Ribosome	ko03010	17 (3.66%)	0.02924749	0.442890563
Vitamin B6 metabolism	ko00750	3 (0.65%)	0.03542163	0.469336598
Cysteine and methionine metabolism	ko00270	7 (1.51%)	0.04333564	0.510397538
Selenocompound metabolism	ko00450	3 (0.65%)	0.05098186	0.540407716
Limonene and pinene degradation	ko00903	9 (1.94%)	0.06588606	0.564560846
Glycine, serine and threonine metabolism	ko00260	6 (1.29%)	0.06739679	0.564560846
Spliceosome	ko03040	16 (3.45%)	0.07073317	0.564560846
Amino sugar and nucleotide sugar metabolism	ko00520	10 (2.16%)	0.07456464	0.564560846
ABC transporters	ko02010	9 (1.94%)	0.1039357	0.671309224
Peroxisome	ko04146	7 (1.51%)	0.1053809	0.671309224
Caffeine metabolism	ko00232	1 (0.22%)	0.1076628	0.671309224
Carotenoid biosynthesis	ko00906	9 (1.94%)	0.1189165	0.700286056
Flavonoid biosynthesis	ko00941	12 (2.59%)	0.1282596	0.715553558
Proteasome	ko03050	4 (0.86%)	0.1386467	0.73482751

Table 2B. List of first twenty pathways in stem

Pathway term (SC/SF)	Pathway ID	DEGs tested	P value	Q value
Fatty acid metabolism	ko00071	16 (1.19%)	0.000612223	0.07346681
Ribosome	ko03010	44 (3.28%)	0.005398256	0.27306652
Taurine and hypotaurine metabolism	ko00430	6 (0.45%)	0.006826663	0.27306652
Other types of O-glycan biosynthesis	ko00514	5 (0.37%)	0.01560009	0.3743676
Regulation of autophagy	ko04140	14 (1.04%)	0.01770056	0.3743676
Amino sugar and nucleotide sugar metabolism	ko00520	26 (1.94%)	0.02162359	0.3743676
Carotenoid biosynthesis	ko00906	25 (1.86%)	0.02409221	0.3743676
Phenylalanine, tyrosine and tryptophan biosynthesis	ko00400	11 (0.82%)	0.02495784	0.3743676
Glycerophospholipid metabolism	ko00564	23 (1.71%)	0.02836514	0.37820187
Porphyrin and chlorophyll metabolism	ko00860	11 (0.82%)	0.03293996	0.39527952
Peroxisome	ko04146	18 (1.34%)	0.03923865	0.41932133
Ubiquinone and other terpenoid-quinone biosynthesis	ko00130	10 (0.74%)	0.05126163	0.41932133
Butanoate metabolism	ko00650	8 (0.6%)	0.05197998	0.41932133
Pyruvate metabolism	ko00620	17 (1.27%)	0.0522923	0.41932133
alpha-Linolenic acid metabolism	ko00592	15 (1.12%)	0.05507743	0.41932133
C5-Branched dibasic acid metabolism	ko00660	2 (0.15%)	0.05590951	0.41932133
Lysine degradation	ko00310	9 (0.67%)	0.06190139	0.43469967
Tyrosine metabolism	ko00350	13 (0.97%)	0.06520495	0.43469967
beta-Alanine metabolism	ko00410	7 (0.52%)	0.07620636	0.47301936
Biosynthesis of secondary metabolites	ko01110	212 (15.79%)	0.07883656	0.47301936



Table 2C. List of first twenty pathways in stem

Pathway term (LC/LF)	Pathway ID	DEGs tested	P value	Q value
Ribosome	ko03010	52 (3.52%)	0.000502993	0.06035914
Amino sugar and nucleotide sugar metabolism	ko00520	30 (2.03%)	0.007302478	0.31563592
Regulation of autophagy	ko04140	16 (1.08%)	0.007890898	0.31563592
Glyoxylate and dicarboxylate metabolism	ko00630	15 (1.02%)	0.01449332	0.4347996
Nicotinate and nicotinamide metabolism	ko00760	4 (0.27%)	0.02350194	0.47046309
Riboflavin metabolism	ko00740	9 (0.61%)	0.02450513	0.47046309
Glutathione metabolism	ko00480	23 (1.56%)	0.02744368	0.47046309
Selenocompound metabolism	ko00450	6 (0.41%)	0.04639775	0.65290427
beta-Alanine metabolism	ko00410	8 (0.54%)	0.04896782	0.65290427
Protein processing in endoplasmic reticulum	ko04141	45 (3.05%)	0.07631821	0.74382171
Alanine, aspartate and glutamate metabolism	ko00250	14 (0.95%)	0.07877682	0.74382171
Pentose phosphate pathway	ko00030	11 (0.75%)	0.07967838	0.74382171
Peroxisome	ko04146	18 (1.22%)	0.08199375	0.74382171
Thiamine metabolism	ko00730	4 (0.27%)	0.09704446	0.74382171
alpha-Linolenic acid metabolism	ko00592	15 (1.02%)	0.1034048	0.74382171
Endocytosis	ko04144	21 (1.42%)	0.104978	0.74382171
Carbon fixation in photosynthetic organisms	ko00710	15 (1.02%)	0.1090852	0.74382171
Spliceosome	ko03040	41 (2.78%)	0.1177536	0.74382171
Glycine, serine and threonine metabolism	ko00260	13 (0.88%)	0.1213056	0.74382171
Cutin, suberine and wax biosynthesis	ko00073	15 (1.02%)	0.127228	0.74382171

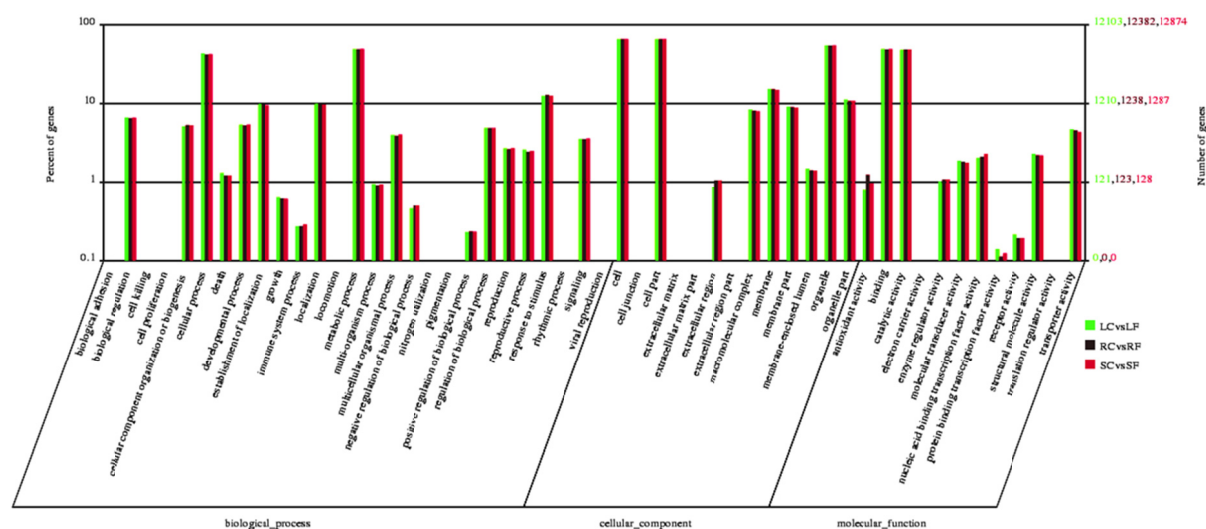


Figure 5. Gene ontology classification. The results are the percentage of each GO category of genes in terms of biological process, cellular component and molecular function. The y-axis indicates the percentage of a specific category of genes and number of genes in that main category

### 3.4 Metabolism Pathway Involved in the Development of Seedling from the DEGs

To shed more lights into the functional roles of DEGs between KFJT-1 and KFJT-CK, biological metabolic pathways were investigated by the enrichment analysis of DEGs among the different tissues samples at seedling stage. In young roots, it is revealed that 106 metabolic pathways were affected by DEGs (Table S1), in young stems were 120 (Table S2), in young leaves were 120 (Table S3). With hypergeometric test, there is only 9 pathway was significantly enriched in young root with the p-value lower than 0.05. The most enriched pathway was glutathione metabolism (ko00480) with p-value adjustment (q-value lower than 0.05, Table 2A). Similarly, there were 11 pathways in young stems with p value lower than 0.05. The most enriched pathway in stems was the

fatty acid metabolism (ko00071) (q-value less than 0.05, Table 2B). However, it was hard to judge which pathway was the most significantly enriched with q-value control in young leaves (q-value less than 0.05, Table 2C). With genes over five fold change, four KEGG term involved in the stress tolerance were enriched: “alpha-Linolenic acid metabolism”, “flavonoid biosynthesis”, “inositol phosphate metabolism” and “fatty acid biosynthesis”, which also suggested that KFJT-1 should have advantage of stress resistance in the process of plant growth and development.

In this study, the circadian rhythm (ko04712) was over-represented at seedling stage. In young root (Figure 6A), the genes involved in the circadian rhythm were mostly up-regulated including (PHYA, PHYB, TOC1, APR3, GI, LHY, CCA1 and WNK1). Whereas, in young stem (Figure 6B), most of the genes were down-regulated including (PHYA, PHYB, ARP7, CHS, TOC1, CCA1). In young leaf, six genes were involved, three of which were up-regulated (PHYA, TOC1 and APR3) which marked with red borders, the other three genes were down-regulated (WNK1 and CK2 $\alpha$ , CK2 $\beta$ ) which marked with green borders in the Figure 6C. The gene GI which is a typical disruption of the PHYB signal transduction pathway was also slightly up-regulated in young root.

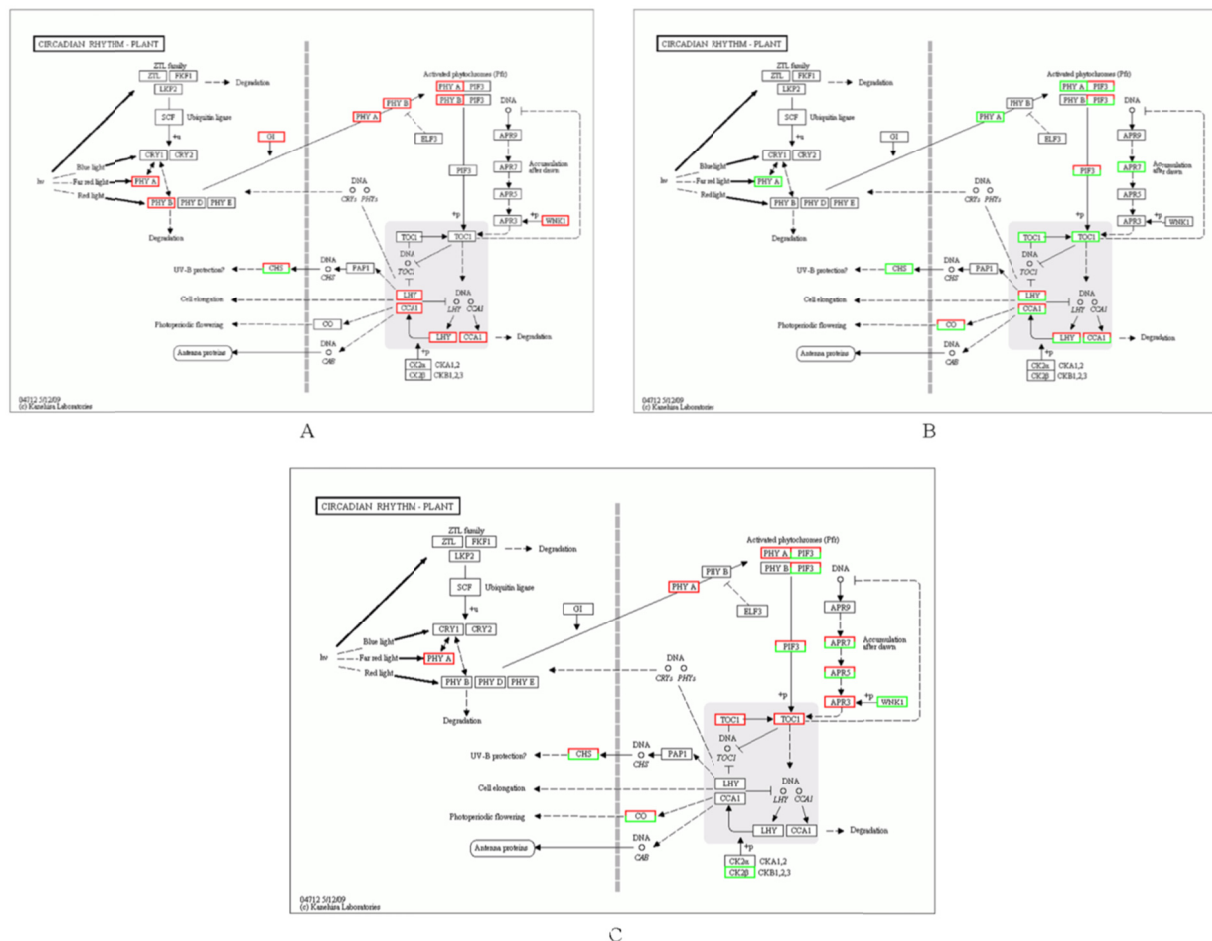


Figure 6. Change in gene expression of the circadian rhythm pathway in roots, stems and leaves at seedlings stage between early maturity mutant and wild type in sweet sorghum. Genes that up-regulated are marked with red borders while genes that down-regulated are marked with green borders. Genes that did not change are marked with black borders. A: roots, B: stems, C: leaves

Regarding the high sugar content in stems at maturation stage in the future, we checked the starch and sucrose metabolism pathway (ko00010) in different tissues at seedling stage. In young root, 8 genes were up-regulated which involved in the synthesis of  $\beta$ -D-Fructose,  $\alpha$ -D-Glucose, pectin and trehalose; three genes were down-regulated which involved in the  $\alpha$ - and  $\beta$ -D-Glucose (Figure 7A). In young stem, 9 genes were up-regulated (Figure 7B). 7 genes involved in the synthesis of sugar such as pectin, sucrose,  $\alpha$ - and  $\beta$ -D-Glucose, D-Xylose, whereas, the up-regulated gene SS4 was for the decrease of the ADP-glucose which result in the less starch

synthesis. The other gene was for the pectate degradation. 9 genes were down-regulated in this pathway, too. In young leaf, 11 genes were up-regulated, 7 genes were down-regulated. These genes involved in accumulation of glucose and sucrose, or limit their consumption (Figure 7C). Those genes were functioned to increase sugar content or reduce its consumption. For example in young stem, the down-regulated gene 3.2.1.26 decreased ADP-glucose, whereas, the up-regulated gene 2.4.1.21 accelerated the ADP-glucose consumption, which limited synthesis of the starch. For the sucrose, the up-regulated gene 2.4.1.13 increase the sucrose, whereas, the down-regulated gene 3.2.1.26 limited its consumption. The process was similar sugars, which resulted in the high sugar content.

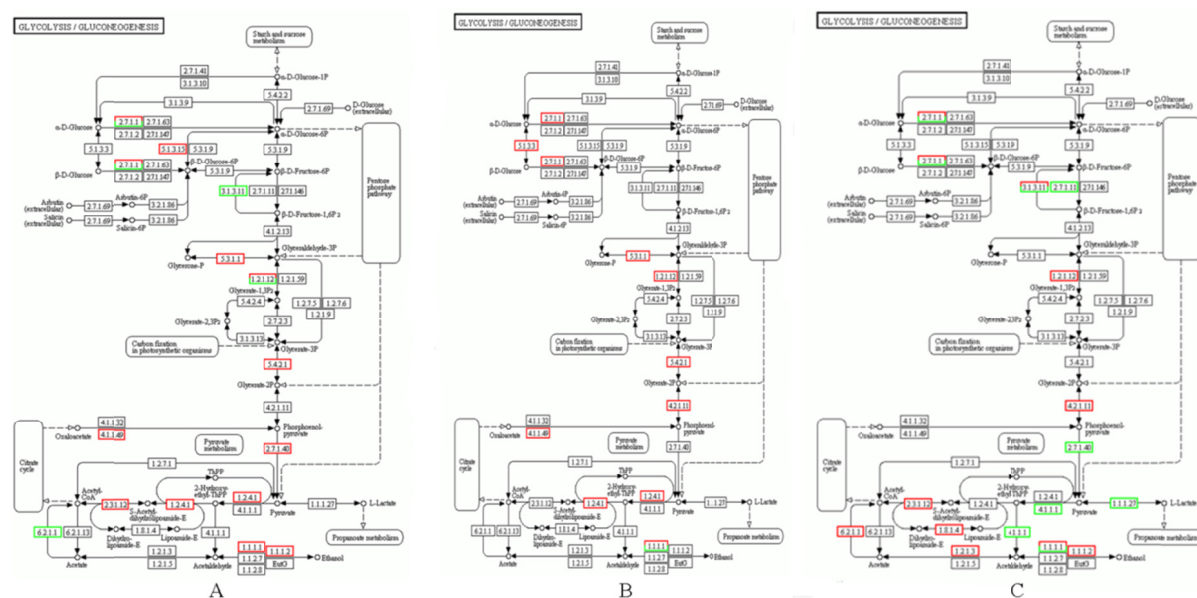


Figure 7. Change in gene expression of the Glycolysis pathway in roots, stems and leaves at seedlings stage between early maturity mutant and wild type in sweet sorghum. Genes that up-regulated are marked with red borders while genes that down-regulated are marked with green borders. Genes that did not change are marked with black borders. A: roots, B: stems, C: leaves

### 3.5 Confirmation of Differentially Expressed Genes by qRT-PCR

Seven genes were selected for qRT-PCR analysis to validate the DGE data. Expression of six genes detected by qRT-PCR matched the DGE data. Moreover, the fold-changes obtained by DGE were generally greater than those obtained by qRT-PCR (Figure 8). The corresponding primers are listed in Table S4.

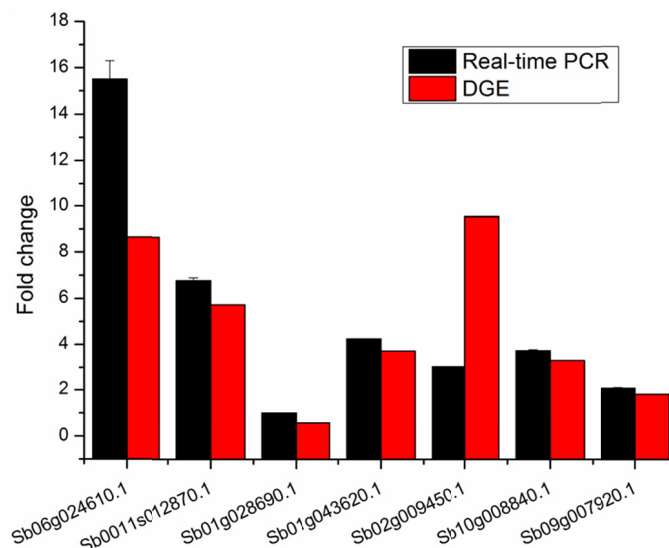


Figure 8. Quantitative PCR validations of tag-mapped genes of sweet sorghum at seedling stage

#### 4. Discussion

It is well known that ion beams can have dramatic mutagenic effects (Matuo et al., 2006; Kawata et al., 2004) through penetrating the seed coat and the cell envelope to induce strand breaking in the DNA (Abe et al., 2002; Liu et al., 2008). We previously analyze the agronomic character, physiological change and its genetic polymorphism of KFJT-1 induced by carbon ions irradiation (Dong & Li, 2012). In this study, we developed and tested a framework for analysis of short-read, sequence-based expression profiles using Illumina DGE technology and the first assembled sorghum reference genome (Paterson et al., 2009). A sequencing depth of 5.8 to 6.7 million tags per library was reached (Table 1). Our results demonstrated that 20- to 21-nucleotide DGE tags can be used to successfully resolve genome-wide expression profiles in sweet sorghum and detect differences in transcript abundance over a broad dynamic range with sufficient sequencing depth to cover the transcriptome (Figure S1). About 70% of clean tags mapped to 55% of reference genes in the six libraries from the three different tissues due to two main reasons: Firstly, reference gene annotations in sorghum were still not completely finished and may contain some mis-annotations. Second, *Nla*III site that is required for detection by DGE technology is contained by only 88% of reference genes (Hegedus et al., 2009), which means that some clean tags was not identified. However, this study provided much valuable information of the different expressed genes induced by carbon ion beam irradiation between KFJT-1 and KFJT-CK. Those unmapped tags could represent information for the novel genes which could be developed in the future.

In addition, we identified 717, 2160 and 2331 tags-mapped genes were differently expressed in roots, stems and leaves between KFJT-1 and KFJT-CK at seedling stage, respectively. More regulated genes enriched in young leaf and stem (the number of young stems and leaves were over three times the number of young roots) and most of them were up-regulated (*i.g.* 1577 up-regulated and 754 down regulated in young leaf), which suggested that the genes were activated by the carbon ion beam irradiation during the development of the seedling for KFJT-1. 66 regulated genes were shared in young roots, stems and leaves, which we suspected that the regulator elements in the promoter regions could be mutated by carbon ion beam irradiation. However, due to the genome sequence was not performed in KFJT-1 yet, the relationship between the mutation sites and the different expressed genes were not easy to be concluded. Further experiment or analysis on the genome sequences should be conducted to clarify the effect of carbon ion beam irradiation.

Regarding to the characteristic of KFJT-1, the expression of the circadian rhythm genes was analyzed at seedling stage. The genes *PHYA*, *PHYB*, *TOC1*, *APR3*, *GI*, *LHY*, *CCA1* and *WNK1*, *CK2 $\alpha$*  and *CK2 $\beta$*  genes was differently expressed in three tissue in KFJT-1 compare to KFJT-CK. In *Arabidopsis*, when the *CK2* was over-expressed, the period of rhythmic expression of the *CCA1/LHY* genes was shorted and caused early flowering in both long-and short-day conditions (Sugano et al., 1999). However, in early flower mutant KFJT-1, the *CK2 $\alpha$*  and *CK2 $\beta$*  were all down-regulated in young leaf. This suggested that there were no strong connection between the *CK2* and *CCA1/LHY* in sorghum. The gene *TOC1* has dual role in the control of circadian and

photomorphogenic responses in *Arabidopsis* (Mas et al., 2003), mostly it was negatively regulated by CCA1/LHY, and positively regulated by the WNK1 gene (Wang & Tobin, 1998). Interestingly, the genes all above were assigned to late flower, which was contradicted against the early flower phenotypic of KFJT-1. This might be contributed to the complex regulatory module for photoperiodic flowering signaling. We proposed that the earlier flower characteristic of KFJT-1 was related to the up-regulation of PHYA in young root and leaf. Previous study indicated that PHYA are the major day-length sensors in *Arabidopsis* (Mockler et al., 2003) and thought to promote flowering. The PHYA mutant flowers significantly later than the wild type in response to day-length extensions with a far-red-enriched white light (Johnson et al., 1994), differed between long-days and short-days (Weller et al., 1997). Interestingly, the PHYA is tissue-specific expressed in KFJT-1. Previous study on *Arabidopsis* suggested that sucrose could depress expression of PHYA (Dijkwel et al., 1997). Thus, we conferred that the down-regulated PHYA in young stem in KFJT-1 was due to high content of sucrose which mutual adjusted by the gene 3.2.1.26 and 2.4.1.13.

## 5. Conclusion

To our best knowledge, this study gained comprehensive understanding of DGEs between KFJT-1 and KFJT-CK, which was the first genome-wide effort to investigate the transcription dynamics of sweet sorghum induced by carbon ion beam at seedling stage. Furthermore, this work provides some useful information to develop function genes for the industry process of energy crop by carbon ion beam.

## References

- Abe, T., Matsuyama, T., Sekido, S., Yamaguchi, I., Yoshida, S., & Kameya, T. (2002). Chlorophyll-deficient mutants of rice demonstrated the deletion of a DNA fragment by heavy-ion irradiation. *Journal of Radiation Research*, 43, 157-161. <http://dx.doi.org/10.1269/jrr.43.S157>
- Albouyeh, R., Farzaneh, N., Bohlmann, J., & Ritland, K. (2010). Multivariate analysis of digital gene expression profiles identifies a xylem signature of the vascular tissue of white spruce (*Picea glauca*). *Tree Genet. Genomes*, 6, 601-611. <http://dx.doi.org/10.1007/s11295-010-0275-0>
- Billa, E., Koullas, D. P., & Monties, B. (1997). Structure and composition of sweet sorghum stalk components. *Ind Crops Prod.*, 6, 297-302. [http://dx.doi.org/10.1016/S0926-6690\(97\)00031-9](http://dx.doi.org/10.1016/S0926-6690(97)00031-9)
- Bruggenmann, E., Handwerger, K., Essex, C., & Storz, G. (1996). Analysis of fast neutron-generated mutants at the *Arabidopsis thaliana* HY4 locus. *Plant J.*, 10, 755-760. <http://dx.doi.org/10.1046/j.1365-3113X.1996.10040755.x>
- Dijkwel, P. P., Huijster, C., Weisbeek, P. J., Chua, N. H., & Smeekens, S. C. (1997). Sucrose control of phytochrome A signaling in *Arabidopsis*. *Plant Cell*, 9, 583-595. <http://dx.doi.org/10.1105/tpc.9.4.583>
- Dong, X. C., & Li, W. J. (2012). Biological features of an early-maturity mutant of sweet sorghum induced by carbon ions irradiation and its genetic polymorphism. *Advances in Space Research*, 50, 496-501. <http://dx.doi.org/10.1016/j.asr.2012.04.028>
- Hegedus, Z. A., Agoston, V. C., Ordas, A., Racz, P., & Mink, M. (2009). Deep sequencing of the zebrafish transcriptome response to mycobacterium infection. *Mol. Immunol.*, 46, 2918-2930. <http://dx.doi.org/10.1016/j.molimm.2009.07.002>
- Hong, L. Z., Li, J., Schmidt, K. A., Warren, W. C., & Barsh, G. S. (2011). Digital gene expression for non-model organisms. *Genome Res.*, 21, 1905-1915. <http://dx.doi.org/10.1101/gr.122135.111>
- Jiang, J. J., Shao, Y. L., Du, K., Ran, L. P., Fang, X. P., & Wang, Y. P. (2013). Use of digital gene expression to discriminate gene expression differences in early generations of resynthesized *Brassica napus* and its diploid progenitors. *BMC Genomics*, 14, 72. <http://dx.doi.org/10.1186/1471-2164-14-72>
- Johnson, E., Bradley, M., Harberd, N. P., & Whitelam, G. C. (1994). Photoresponses of Light-Grown phyA Mutants of *Arabidopsis* (Phytochrome A Is Required for the Perception of Daylength Extensions). *Plant Physiol.*, 105, 141-149. <http://dx.doi.org/10.1104/pp.105.1.141>
- Kawata, T., Ito, H., George, K., Durante, M., Furusawa, Y., Wu, H., & Cucinotta, F. A. (2001). G2-chromosome aberrations induced by high-LET radiations. *Advances in Space Research*, 27, 383-391. [http://dx.doi.org/10.1016/S0273-1177\(01\)00006-0](http://dx.doi.org/10.1016/S0273-1177(01)00006-0)
- Kawata, T., Ito, H., Uno, T., Saito, M., Yamamoto, S., Furusawa, Y., ... Cucinotta, F. A. (2004). G2 chromatid damage and repair kinetics in normal human fibroblast cells exposed to low- or high-LET radiation. *Cytogenetic and Genome Research*, 104, 211-215. <http://dx.doi.org/10.1159/000077491>

- Kikuchi, S., Saito, Y., Ryuto, H., & Fukunishi, N. (2009). Effects of heavy-ion beams on chromosomes of common wheat *Triticumaestivum*. *Mutation Research*, 669, 63-66. <http://dx.doi.org/10.1016/j.mrfmmm.2009.05.001>
- Kranz, A. R. (1994). Heavy ion and cosmic radiation effects in different targets of the *Arabidopsis* seed. *Acta Astronaut*, 33, 201-210. [http://dx.doi.org/10.1016/0094-5765\(94\)90126-0](http://dx.doi.org/10.1016/0094-5765(94)90126-0)
- Li, J. Q., Wang, L. H., Zhan, Q. W., Liu Y. L., Fu, B. S., & Wang, C. M. (2013). Sorghum *bmr6* mutant analysis demonstrates that a shared MYB1 transcription factor binding site in the promoter links the expression of genes in related pathways. *Funct. Integr. Genomics*, 13, 445-453. <http://dx.doi.org/10.1007/s10142-013-0335-2>
- Liu, B. M., Wu, Y. J., Xu, X., Song, M., Zhao, M., & Fu, X. D. (2008). Plant height revertants of dominant semidwarf mutant rice created by low-energy ion irradiation. *Nuclear Instruments and Methods in Physics Research B*, 266, 1099-1104. <http://dx.doi.org/10.1016/j.nimb.2008.02.045>
- Más, P., Alabadi, D., Yanovsky, M. J., Oyama, T., & Kay, S. A. (2003). Dual role of TOC1 in the control of circadian and photomorphogenic responses in *Arabidopsis*. *Plant Cell*, 15, 223-236. <http://dx.doi.org/10.1105/tpc.006734>
- Matuo, Y., Nishijima, S., Hase, Y., Sakamoto, A., Tanaka, A., & Shimizu, K. (2006). Specificity of mutations induced by carbon ions in budding yeast *Saccharomyces cerevisiae*. *Mutation Research*, 602, 7-13. <http://dx.doi.org/10.1016/j.mrfmmm.2006.07.001>
- Mei, M., Deng, H., Lu, Y., Zhuang, C., Liu, Z., Qiu, Q., ... Yang, T. C. (2011). Mutagenic effects of heavy ion radiation in plants. *Advanced Space Research*, 14, 363-372. [http://dx.doi.org/10.1016/0273-1177\(94\)90489-8](http://dx.doi.org/10.1016/0273-1177(94)90489-8)
- Mockler, T. C., Yang, Y., Yu, H., Parikh, D., Cheng, Y. X., Dolan, S., & Lin, C. T. (2003). Regulation of photoperiodic flowering by *Arabidopsis* photoreceptors. *Proc. Natl. Acad. Sci. USA*, 100, 2140-2145. <http://dx.doi.org/10.1073/pnas.0437826100>
- Nishiyama, T., Miyawaki, K., Ohshima, M., Thompson, K., & Nagashima, A. (2012). Digital gene expression profiling by 5-end Sequencing of cDNAs during reprogramming in the moss *Physcomitrella patens*. *PLoS One*, 7, 36471. <http://dx.doi.org/10.1371/journal.pone.0036471>
- Paterson, A. H., Bowers, J. E., Bruggmann, R., Doreen, W., Peter, W., Klaus, M., ... Daniel, R. (2009). The sorghum bicolor genome and the diversification of grasses. *Nature*, 457, 551-556. <http://dx.doi.org/10.1038/nature07723>
- Qin, Y. F., Fang, H. M., Tian, Q. N., Liang, F., Zhang, Y. Z., & Zhang, S. T. (2011). Transcriptome profiling and digital gene expression by deep-sequencing in normal/regenerative tissues of planarian *Dugesia japonica*. *Genomics*, 97, 364-371. <http://dx.doi.org/10.1016/j.ygeno.2011.02.002>
- Shikazono, N., Suzuki, C., Kitamura, S., Watanabe, H., Tano, S., & Tanaka, A. (2005). Analysis of mutations induced by carbon ions in *Arabidopsis thaliana*. *Journal of Experimental Botany*, 56, 587-596. <http://dx.doi.org/10.1093/jxb/eri047>
- Shikazono, N., Tanaka, A., Watanabe, H., & Tano, S. (2000). Rearrangements of the DNA in carbon ion-induced mutants of *Arabidopsis thaliana*. *Genetics*, 157, 379-387.
- Shikazono, N., Yokota, Y., Kitayama, S., Chihiro, S., & Hiroshi, W. (2003). Mutation rate and novel *tt* mutants of *Arabidopsis thaliana* induced by carbon ions. *Genetics*, 163, 1449-1455.
- Sugano, S., Andronis, C., Ong, M. S., Green, R. M., & Tobin, E. M. (1999). The protein kinase CK2 is involved in regulation of circadian rhythms in *Arabidopsis*. *Proc. Natl. Acad. Sci. USA*, 96, 12362-12366. <http://dx.doi.org/10.1073/pnas.96.22.12362>
- Wang, Y., & Tobin, E. M. (1998). Constitutive expression of the *CIRCADIAN CLOCK ASSOCIATED 1 (CCA1)* gene disrupts circadian rhythms and suppresses its own expression. *Cell*, 93, 1207-1217. [http://dx.doi.org/10.1016/S0092-8674\(00\)81464-6](http://dx.doi.org/10.1016/S0092-8674(00)81464-6)
- Wei, L. J., Yang, Q., Xia, H. M., Furusawa, Y., Guan, S. H., Xin, P., & Sun, Y. Q. (2006). Analysis of cytogenetic damage in rice seeds induced by energetic heavy ions on-ground and after spaceflight. *Journal of Radiation Research*, 47, 273-278. <http://dx.doi.org/10.1269/jrr.0613>

- Wei, M. M., Song, M. Z., Fan, S. L., & Yu, S. X. (2013). Transcriptomic analysis of differentially expressed genes during anther development in genetic male sterile and wild type cotton by digital gene-expression profiling. *BMC Genomics*, *14*, 97. <http://dx.doi.org/10.1186/1471-2164-14-97>
- Weller, J. L., Murfet, I. C., & Reid, J. B. (1997). Pea mutants with reduced sensitivity to far-red light define an important role for phytochrome A in day-length detection. *Plant Physiol.*, *114*, 1225-1236. <http://dx.doi.org/10.1104/pp.114.4.1225>
- Xu, J. L., Wang, J. M., Sun, Y. Q., Wei, L. J., Luo, R. T., Zhang, M. X., & Li, Z. K. (2006). Heavy genetic load associated with the subspecific differentiation of japonica rice (*Oryza sativa* ssp. *japonica* L.). *Journal of Experimental Botany*, *57*, 2815-2824. <http://dx.doi.org/10.1093/jxb/erl046>

### Notes

Please download the supplementary files (Tables S1-S4 and Figure S1) at <http://ccsenet.org/journal/index.php/jas/article/view/63224/35097>

### Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).