Transcriptome sequencing of okra (*Abelmoschus esculentus* L. Moench) uncovers differently expressed genes responding to drought stress

Denghong Shi^{1,3} · Jiyue Wang^{1,2} · Yu Bai^{1,3} · Yan Liu^{1,3}

Received: 25 January 2019 / Accepted: 13 August 2019 / Published online: 6 September 2019
© The Author(s) 2019

Abstract

Drought affects the growth and development of plants and can cause crop failure. As an important edible and medicinal plant, okra (*Abelmoschus esculentus* L. Moench) is very popular in China and shows strong resistance to drought. In this study, gene expression profiles of okra leaves were investigated under water-limited conditions at the seedling stage. A total of 94,769 unigenes with an average length of 1921 bp and an N90 of 981 bp were obtained through next-generation sequencing. Furthermore, 38.18% of total unigenes were annotated in NT, NR, COG, KEGG, and SwissProt databases. A total of 2276, 5626, 11,719, 14,892, 855, 4593, and 5170 differentially expressed genes (DEGs) in seven pairwise comparisons (A5 vs. A0; A7 vs. A0; A15 vs. A0; A20 vs. A0; A7 vs. A5; A15 vs. A7, and A20 vs. A15) were identified. Among these, 22,751 DEGs were up-regulated, and the number was slightly more than that of the down-regulated DEGs. The number of up-regulated DEGs among different samples increased with increasing drought stress. We identified that the majority of DEGs were related to “integral component of membrane”, “oxidation–reduction process”, and “metal ion binding” categories. Pathway analysis revealed that most DEGs were enriched in “Biosynthesis of secondary metabolites”, “Carbon metabolism”, and “Glycolysis/Gluconeogenesis pathways”. Additionally, a total of 17,456 simple sequence repeat motifs, 47,9831 single nucleotide polymorphisms, and 3746 transcription factors were identified in this study. RNA-Seq data were also verified by RT-qPCR. Our findings engender a further understanding of molecular mechanisms in response to drought stress.

Keywords Okra (*Abelmoschus esculentus* L.) · Transcriptome sequencing · Drought stress · Differential expression

Abbreviations

RNA-Seq Ribonucleic acid-sequencing
NGS Next generation sequencing
cDNA Complementary deoxyribonucleic acid

PCR Polymerase chain reaction
EB Ethambutol
GO Gene ontology
KEGG Kyoto encyclopedia of genes and genomes
CDSs Coding sequences
SSR Simple sequence repeat

Jiyue Wang and Yu Bai: co-first author.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s13562-019-00528-w>) contains supplementary material, which is available to authorized users.

✉ Jiyue Wang
acute2803764@126.com

- ¹ Guizhou Provincial Key Laboratory for Rare Animal and Economic Insect of the Mountainous Region, Guiyang 550005, China
- ² Graduate School, Guiyang University, Guiyang 550005, China
- ³ School of Biological and Environmental Engineering, Guiyang University, Guiyang 550005, China

Introduction

Okra (*Abelmoschus esculentus* L. Moench) belongs to the Malvaceae family and is one of the most important vegetables worldwide. It is used as both a food source and a medicinal resource (Dhankhar et al. 2005). Because it is rich in amino-acids, vitamins and mineral elements (iron, potassium, and calcium), okra is extremely popular today (Salameh 2014). Most okra cultivation is conducted exclusively in developing countries and particularly in

African countries (2.25 MT ha⁻¹). It is also grown in India, China, Turkey, and USA. India is the largest okra-planted country, with an area of over 0.5 million ha, followed by Nigeria and Sudan, which produce 6.3 million MT per year (Varmudy 2011; Faostat 2014). Extensive previous research has mostly focused on cropping systems (Anonymous 2015), nutritional components (Calisir et al. 2005), heredity and breeding (Kumar et al. 2017; Mishra et al. 2017), fresh maintenance (Tajiri 2011) and processing technology (Rai and Balasubramanian 2009). However, reports on the molecular mechanisms of okra are relatively few.

Studies have shown that drought-stress induce a series of physiological and morphological changes in plants, for example: growth inhibition and reduction in crop yields (Elagib 2014), enhancement of root systems and root-shoot ratios (Lu et al. 2014), regulation of the closure of stomata (Wan et al. 2009), disruption of photosynthesis (Zivcak et al. 2013), activation of respiration (Sperlich et al. 2016), accumulating compatible solutes and protective proteins (Ogbaga et al. 2016; Sperdouli and Moustakas 2012), and the strengthening of anti-oxidative enzyme activity (Bhaskara et al. 2015; Hu et al. 2015). These changes are controlled by gene expression, which is particularly affected by biotic and abiotic factors in the environment.

The RNA-sequencing (RNA-Seq) technique is considered to be an effective and feasible method for analyzing the gene expression profiles of plants and is an accepted method for studying a suite of traits such as drought resistance and salt resistance (Mortazavi et al. 2008; Shen et al. 2016; Wang et al. 2009). Previous studies have identified drought-tolerant genes in rice, wheat, and some other plant models, based on their genomic information. Recently, arid-resistance genes have been found in some non-model plants using RNA-Seq, such as *Populus trichocarpa* (Tang et al. 2015), *Gmelina arborea* (Rosero et al. 2011), *Camellia sinensis* (Paul et al. 2014), *Hordeum vulgare* var. *Nudum* (Zeng et al. 2016), *Glycine max* (Han et al. 2016), *Phormium tenax* (Bai et al. 2017), and *Pyrus betulaefolia* (Li et al. 2016).

Some studies on the transcriptomes of okra have been reported (Schafleitner et al. 2013; Zhang et al. 2017) using next-generation sequencing (NGS). However, owing to the lack of a reference genome, few drought-resistant genes have been identified in okra. In this study, we examined the transcriptome of okra under drought stress conditions to obtain the gene expression patterns of okra and to identify drought-tolerant genes. We detected the physiological and biochemical response of okra to drought stress. These results will further improve our understanding of drought-tolerant mechanisms in okra.

Materials and methods

Plant materials

‘Xianzhi’, the drought tolerant okra cultivar from the Zhenjiang province of China, was cultivated in a greenhouse in the Guiyang University in Guizhou province, China. Blades from seedling cuttings were used to extract total RNA.

Drought stress treatment

Okra plants were cultivated for 35 days at 70% humidity and a temperature of 25 ± 2 °C. Firstly, plants were planted by direct seeding into a basin containing nutrient-enriched soil. After sprouting, the seedlings were watered every 2 days, then after an adaptation period of 2 weeks, dehydration treatment was applied to all plants. Leaves were collected from the seedlings after 0, 5, 7, 15, and 20 days of dehydration treatment. They were then kept in liquid nitrogen for RNA extraction and stored at 80 °C in an ultra-low temperature refrigerator. The five treatments were marked as A0, A5, A7, A15, and A20 (taking samples under different drought conditions), respectively. Each treatment was repeated three times. A total of 15 samples were collected and used in RNA-Seq.

RNA-Seq

Total RNA was extracted from the blade of each sample using the Trizol kit (Takara Bio Inc., Beijing, China). The A260/A280 ratios of all RNA samples were varied in the range of 1.8–2.1. The quality of the RNA sample was determined with an Agilent 2100 Bioanalyzer, and samples without degradation were selected for RNA-Seq.

Beads with oligo (dT) were used to isolate poly (A) mRNA. A fragmentation buffer was added to digest mRNA into short fragments. The first-strand cDNA was synthesized with random hexamer-primers, taking these short mRNA fragments as templates. The second-strand cDNA was obtained using buffer, dNTPs, RNase H, and DNA polymerase I. Purification was conducted with a QIAquick PCR purification kit (Takara Bio Inc. Beijing, China) and resolved with EB buffer. For end reparation and poly (A) addition, these short fragments were then connected to sequencing adaptors. Those suitable fragments were selected as templates for amplification with PCR. The constructed library was qualified with an Agilent 2100 Bioanalyzer and ABI StepOnePlus real-time PCR System. Finally, the library was sequenced on the Illumina HiSeq 2500 system.

Before the bioinformatics analysis was performed, quality control was required in order to detect whether the data were suitable. Filtering the raw data is needed to remove unwanted noise from the data. Therefore, low-quality reads, adaptor sequences and reads with an excess of N were removed by using in-house Perl scripts. We defined the baseline as a sequencing quality of no more than 10 and ratios of unknown bases more than 10% as low quality, which would be filtered out. After filtering, the clean reads were used for all subsequent analyses.

De novo assembly and functional annotation

All clean reads from the 15 samples were mixed and assembled from scratch using Trinity (<http://trinityrnaseq.sourceforge.net/>). All assembled clean data were aligned to transcripts using Bowtie2 with default settings (<http://bowtie-bio.sourceforge.net/Bowtie2/index.shtml>). The sequence assembled by Trinity is called a transcript. Then, Tgicl (<http://www.tigr.org/tdb/tgi/software/>) was used to remove the redundancy and further splicing in order to obtain the final unigenes. Unigenes obtained by removing redundancy and splicing were divided into two classes: clusters and singletons. We used RSEM (<http://deweylab.biostat.wisc.edu/RSEM>) to assess the gene expression amount of assembled transcripts. The sequences obtained by RNA-Seq were aligned to NR (non-redundant protein sequences) in NCBI (<ftp://ftp.ncbi.nlm.nih.gov/blast/db>), Swiss-Prot (<http://ftp.ebi.ac.uk/pub/databases/swissprot>), TrEMBL (<http://www.ebi.ac.uk/interpro>), KEGG (Kyoto Encyclopedia of Genes and Genomes database) (<http://www.genome.jp/kegg>), and COG (clusters of orthologous groups) (<http://www.ncbi.nlm.nih.gov/COG>) database by Blastx (BLAST, basic local alignment search tool). They were aligned to the Nt (nucleotide sequence; <ftp://ftp.ncbi.nlm.nih.gov/blast/db>) database by Blastn (e value < 1E^{-5}) to annotate and classify the function of genes.

Gene expression analysis

The differentially expressed genes (DEGs) among the six samples were analysed using DESeq2 (<http://www.biocconductor.org/packages/release/bioc/html/DESeq2.html>). The threshold of log₂ fold-change ≥ 1 and the false discovery rate (FDR) < 0.05 in multiple tests was used to define the DEGs. The log₂ fold-change ≥ 1 means a fold-change value of gene expression levels ≥ 2 .

Gene Ontology (GO) is an internationally standardized gene functional classification system that offers a dynamically-updated controlled vocabulary and a strictly defined concept to comprehensively describe properties of genes and their products in any organism. The GO terms fulfilling this condition were defined as significantly enriched GO

terms in DEGs. This analysis was able to recognize the main biological functions of DEGs. Different genes usually cooperate with each other to exercise their biological functions. The pathway-based analysis helps us to further understand the biological functions of genes. The KEGG database is a major public pathway-related database. Pathway enrichment analysis identifies significantly enriched metabolic pathways or signal transduction pathways based on DEGs.

Coding sequences (CDSs) predication

The tool ESTScan (<http://www.ch.embnet.org/software/ESTScan2.html>) was used to predict the potential protein-coding sequences.

SSR and SNP/InDel analysis

Simple Sequence Repeat (SSR), also known as microsatellite DNA, is a segment of DNA composed of 1–6 nucleotides in the genome, repeated several times. The SSR unigene sequences can be detected using MicroSatellites (MISA). The length of the sequence from SSR repeated units before and after on the genes were screened. Those SSR whose sequence was less than 150 bp were retained. The primer design of detected SSR was carried out with Primer 3 (default parameters). Representative Microsatellites DNA (unit size/minimum number of repeats) was described as 1/12, 2/6, 3/5, 4/5, 5/4, and 6/4. The maximum number of bases interrupting 2 SSRs in a compound microsatellite was 100. Single-nucleotide polymorphism (SNP), also called mutational site was referred to in our study as RNA sequence variation occurring when a single nucleotide (A, T, C, or G) differed among samples. Insertion-deletion (InDel) was defined as an insertion or deletion of small fragments occurring in a sample, relative to the reference genome. We used the BWA tool to align the clean reads of each sample to the transcript. The SNPs/InDels from each sample were identified using GATK. The SNPs/InDels were annotated and counted after deep filtering.

TF analysis

Transcription factor (TF), also known as trans-action factor, has been found to be abundant in plants, animals, fungi and bacteria. Firstly, the Hidden Markov Model (HMM) profiles of a known transcription factor were collected from various databases, such as PlantTFDB (Jin et al. 2014), AnimalTFDB (Zhang et al. 2015), FTFD (Park et al. 2008) and DBD (Sarah and Sarah 2006), and published literature. Secondly, the protein sequence was aligned to the HMM profiles of known transcription factors using hmmsearch.

Fig. 1 All-UUnigene length distribution from 15 samples. The Y-axis represents the number of unigenes

Validation of RNA-Seq

To validate the expression levels of DEGs, quantitative real-time PCR (qRT-PCR) was performed. For DEGs, specific primers were designed in conserved domains (Table S1). Total RNA from the leaves under different drought stress conditions (A0, A5, A7, A15, and A20) was extracted using the Trizol kit (Invitrogen, USA). First-strand cDNA synthesis was performed using PrimeScript RT Reagent Kit with gDNA Eraser (Takara) and PCR amplification using PrimeSTAR HS DNA Polymerase (Takara). Quantitative real-time PCR (qRT-PCR) was performed using the commercial kit (AccuPower® 2X Greenstar qPCR Master Mix) (Bioneer Inc., Korea). A total of 20 μ L reaction mixture consisting of 10 μ L of 2 $\times$ Master Mix, 0.5 μ L of each primer (10 μ M), 2 μ L of cDNA template and 7 μ L RNA-free H₂O. β -Actin was assigned as internal control genes for qRT-PCR. Relative amounts of gene expression were calculated using the $2^{-\Delta C_t}$ method. Non-template controls and melting curve analyses were performed for each gene during each run. Each sample was repeated three times.

Measurement of physiological parameters

Leaves were collected from seedlings after 0, 5, 7, 15, and 20 days of dehydration treatment for enzyme activity measurements. Photosynthetic rate (Pn), stomatal conductance (Gs) and CO₂ concentration (Ci) were determined from 9:00 a.m. to 12:00 a.m. using an LI-6400

portable photosynthetic analyzer. The chlorophyll content of the leaves in the middle of the main stem was determined by UV-2450 ultraviolet spectrophotometry. The soluble protein content was determined using Coomassie's brilliant blue staining assay. The anthrone colorimetric assay was used to determine the soluble sugar content. The indanone assay was used to determine free proline content. Malondialdehyde content (MDA) was determined using thiobarbituric acid chromatometry.

Results

RNA sequencing and read assembly

A total of 751,403,932 reads with a length of 250 bp were obtained from the 15 samples. All reads produced from this study were deposited in the National Center for Biotechnology Information (NCBI) and can be downloaded from the Short Read Archive (SRA accession: PRJNA514435) Sequence Database using the accession number. There were 94,769 unigenes with an N90 of 918 bp generated after assembling all the clean reads and filtering out low-grade reads ($p < 2$) and redundancy by using TGICL (Table S2). The mean value of percent GC-content was 40.67%. Total mapped reads among the 15 samples were from 68.47 to 74.75% and the unique mapped reads of samples varied in the range of 19.38–21.54%. The unigenes with lengths of > 3000 bp (17,630) accounted for 19.94%, while unigenes with lengths of 2300–3000 bp

(5689) accounted for 6.44%. There were 29 unigenes with lengths of > 20,000 bp. The longest and shortest unigenes found in the 15 samples were 23,634 bp (1) and 150 bp (1997), respectively (Fig. 1).

Functional annotation of unigenes

All the assembled unigenes with an E-value of less than $1e^{-5}$ were aligned to five databases by Blast. These databases were NT, NR, COG, KEGG, and SwissProt. The GO annotations were carried out using Blast2GO with default parameters, based on the NR annotation results. A total of 74,981, 64,239, 58,398, 75,166 and 33,467 transcripts were detected in the NR, SwissProt, KEGG, NT and COG databases, respectively. There were 30,989 transcripts possessing homologous sequences in these databases, while approximately 16% of transcripts were not identified in them (Fig. 2). Venn diagrams (Fig. 2) showed that 30,989 unigenes were shared in the same five databases. Only nine and ten unigenes were uniquely annotated in KEGG and COG, respectively. There were 74,981 unigenes found in the NR database. Based on NR annotations, the three top hit species were *Gossypium raimondii* (56.66%), *Theobroma cacao* (15.93), and *Gossypium arboreum* (15.09) (Fig. 3). Regarding the similarity distribution,

approximately 40% of the annotated sequence had higher similarity than 80%.

Based on GO annotations, 58,219 unigenes were classified into 64 categories. Most unigenes were categorized in “biological process”, followed by “cellular process”, and “molecular function” (Fig. 4). There were 1, 479, 68 unigenes grouped into “biological process”, which were annotated within 23 classes. Among the transcripts categorized in “biological process”, 32,509 were involved in “metabolic process” accounting for its largest proportion, followed by 20,157 in “single-organism process”, and 11,220 in “biological regulation”. Among the unigenes that were categorized in “molecular function”, 30,895 and 29,378 were mostly involved in “binding” and “catalytic activity”, respectively. Within cellular components, the assignments were mostly related to “cell part” (28,590), “membrane” (23,042), and “organelle” (20,460).

All these unigenes were aligned against the COG database to check the completeness of *Abelmoschus esculentus* L. transcriptome library and to assess the effectiveness of the annotation process. In total, 3,3467 unigenes (35.31%) were grouped into 25 COG categories (Fig. 5). The largest category was “general function prediction only” containing 11,889 (12.55%), followed by “transcription” (6380; 6.73%), “replication, recombination and repair” (5, 860; 6.18%), “signal transduction mechanisms” (5199; 5.49%) and “posttranslational modification, protein turnover and chaperones” (5022; 5.30%). Only a few unigenes were assigned to “extracellular structures” (45) and “nuclear structure” (9). The unigenes assigned to “function unknown” accounted for 3.03% (2868).

To further understand the function of gene products in *A. esculentus* under drought conditions, 58,398 unigenes were classified into 134 KEGG pathway items (Fig. 6) and 52,242 unigenes belonged to “metabolic pathways”. Among these, the majority (44.67% of 58,398) belonged to “global and overview maps” (23,338). There were 15,009 (25.70% of 58,398) unigenes assigned to “genetic information processing”. Among these, 71.46% (10,725 of 15,009) were related to “translation” (6233) and “folding, sorting and degradation” (4492). Interestingly, 3679 unigenes were involved in “environmental information processing”.

Fig. 2 Venn diagram of functional annotation aligned to NR, NT, Swissprot, KEGG, and COG. The Numbers in the figure correspond to the number of uUnigenes on the annotation. Different databases are distinguished by different colors. Overlapping circles indicates the number of shared expressed genes. shared

Prediction of coding sequences (CDSs)

According to functional annotation, 76,328 CDSs were found by BLASTx and 1180 CDSs were predicted by ESTScan. The length distribution of the CDSs (A) and its predictable peptide sequences (B) are shown in Fig. S1.

Fig. 3 The tTop hit species distribution. The X axis represents the number of uUnigenes, and the Y axis represents the number of NR species

Fig. 4 GO classification. The horizontal axis is the type of GO function, the vertical axis on the right is the number of uUnigenes annotated to the corresponding GO function, and the vertical axis on the left is the percentage of the total number of uUnigenes

Fig. 5 COG function classification. The X axis represents the number of uUnigenes, and the Y axis represents the COG function classification

Gene expression

We calculated the expression levels of the unigenes from the 15 samples by using the FPKM method (fragments per kilobases per million fragments), which is currently a commonly used methodology. To ensure the reliability of the experimental data, we carried out three replicates for each treatment. The results of a correlation analysis showed that the correlation index between these replicates ranged from 0.86 to 0.98 (Fig. S2), suggesting that the sequencing data from RNA-Seq was reliable and provided an important guarantee for subsequent analysis.

Analysis of differential gene expression

An analysis of differential gene expression was conducted to determine the expression levels of genes, which were significantly altered in different samples or under different experimental conditions. The DEGs were obtained through pairwise comparison between different treatment durations (0, 5, 7, 15, and 20 days of dehydration). Seven sample-

pairs, being A5 versus A0, A7 versus A0, A15 versus A0, A20 versus A0, A7 versus A5, A15 versus A7 and A20 versus A15 were chosen in this study. There were, respectively, 2276, 5626, 11,719, 14,892, 855, 4593, and 5170 DEGs identified in these pairwise comparisons (Fig. 7), with most being shared among the seven pairs. The results indicated that the number of DEGs between each sample and the blank control (A0) increased with the increasing degree of drought stress. However, the number of DEGs among different samples (A5 vs. A7, A7 vs. A15, A15 vs. A20.) was significantly less than that between each sample and A0 (A5 vs. A0, A7 vs. A0, A15 vs. A0, and A20 vs. A0). For example, a total of 5626 DEGs were found in A7 versus A0, considerably more than the number of DEGs in A7 versus A5 (855). Furthermore, with increased drought stress, the number of up-regulated DEGs among different samples also increased. The number of down-regulated genes was significantly higher in A7 conditions than in A5 conditions. However, the number of up-regulated genes were significantly higher in A15 conditions than in A7 conditions.

Fig. 6 KEGG Pathway Item. The X axis represents the number of uGenes, and the Y axis represents the KEGG function classification

Fig. 7 Statistics of differentially expressed genes from seven sample-pairs, being namely A5 versus A0, A7 versus A0, A15 versus A0, A20 versus A0, A7 versus A5, A15 versus A7, and A20 versus A15. Blue: up-regulated gene, Red: down-regulated gene

According to the enrichment analysis, DEGs in A5 versus A0, A7 versus A0, A15 versus A0, A20 versus A0, A7 versus A5, A15 versus A7, and A20 versus A15 were classified into 16 cellular component categories, ten molecular function and 23 biological process categories (Figs. S3–S9). The largest category in cellular components was “integral component of membrane”. There were 544, 1286, 2425, 3147, 186, 1042, and 1244 DEGs enriched in A5 versus A0, A7 versus A0, A15 versus A0, A20 versus A0, A7 versus A5, A15 versus A7, and A20 versus A15, respectively. The “oxidation–reduction process” was the largest category of the biological process, with 288, 511, 867, 1139, 82, 452, and 404 DEGs enriched in A5 versus A0, A7 versus A0, A15 versus A0, A20 versus A0, A7 versus A5, A15 versus A7, and A20 versus A15, respectively. In the molecular function group, the largest category was “metal ion binding” in A5 versus A0 (117), A15 versus A0 (490), A20 versus A0 (635), A15 versus A7 (231), and A20 versus A15 (239) and “DNA binding” in A7 versus A0 (317) and A7 versus A5 (59).

Based on the KEGG enrichment analysis of DEGs, 127, 132, 133, 133, 113, 132, and 131 metabolic pathways were obtained from DEGs in seven sample-pairs, which were, respectively, A5 versus A0, A7 versus A0, A15 versus A0, A20 versus A0, A7 versus A5, A15 versus A7, and A20

versus A15. Most DEGs were involved in “global and overview maps” and “carbohydrate metabolism”. The Enriched Pathways of DEGs from seven sample-pairs (above mentioned) are shown in Figs. S10–S16. The top three enriched pathways in each comparison pair are indicated in Table 1. Among the seven sample-pairs, the top-three enriched pathways were mainly involved in “biosynthesis of secondary metabolites”, “carbon metabolism”, and “glycolysis/gluconeogenesis”. The number of up-DEGs in the top-three enriched pathways from A5 versus A0, A7 versus A0, and A7 versus A5 were more than that of down-DEGs.

SSRs, SNP/InDel and TF in the transcriptome library

There were 17,456 SSR motifs identified using MISA, from 94,769 examined sequences, including 2716 sequences containing more than 1 SSR. The total number of examined sequences was 182,133,171 bp. Specifically, tri-nucleotide repeats (39.21%) were the most abundant motif in all examined SSR, followed by mono-nucleotide (27.36%), di-nucleotide (26.45%), penta-nucleotide (2.65%), hexa-nucleotide (2.42%), and quad-nucleotide (2.11%; Fig. 8). The tri-, mono-, and di-nucleotides

Table 1 Top-three enriched pathways of differentially expressed genes (DEGs) from the seven comparison

Pairs	Pathway	Up-DEGs annotation in term	Down-DEGs annotation in term
A5 versus A0	Biosynthesis of secondary metabolites	132	183
	Carbon metabolism	23	33
	Glycolysis/gluconeogenesis	15	28
A7 versus A0	Biosynthesis of secondary metabolites	305	374
	Plant hormone signal transduction	100	157
	Carbon metabolism	61	73
A15 versus A0	Biosynthesis of secondary metabolites	657	515
	Carbon metabolism	192	124
	Starch and sucrose metabolism	96	111
A20 versus A0	Biosynthesis of secondary metabolites	957	707
	Carbon metabolism	280	175
	Plant-pathogen interaction	178	259
A7 versus A5	Metabolic pathways	51	122
	Biosynthesis of secondary metabolites	35	79
	Carbon metabolism	8	27
A15 versus A7	Biosynthesis of secondary metabolites	423	204
	Carbon metabolism	139	44
	Glycolysis/gluconeogenesis	51	25
A20 versus A15	Biosynthesis of secondary metabolites	364	242
	Carbon metabolism	103	55
	Glycolysis/gluconeogenesis	49	38

Fig. 8 The statistics statistics of SSRs. The horizontal axis have shown shows the length and the specific type of repeat, and the vertical axis have shown shows the number of SSRs

(92.81%) represented the majority, while the remainder accounted for less than 8%. In total, 232 different types of SSR motifs were identified in this study. The Mono-, di-, tri-, tetra-, penta-, and hexa-nucleotide repeats had 2, 4, 10, 23, 64, and 129 types, respectively. The five most abundant types were A/T (4373), AT/AT (2517), AAC/GTT (1955), AG/CT (1635), and AAT/AAT (918) (Table S3). On the other hand, ACAG/CTGT, AAAAG/CTTTT, and AAAATC/ATTTTG were the most frequent quad-, penta-, and hexa-nucleotide motif types, but their proportions were very small. Interestingly, only 26 SSR motifs belonged to CG/GC types. The most mono-nucleotide motifs ranged from 12 to 16 repeats and among them, two had 27 repeats, which were (A) 27 and (T) 27. Furthermore, we found that most di-nucleotide motifs ranged from six to seven repeats, while more than half of the tri-nucleotide motifs ranged from three to four repeats. There was a majority of penta-nucleotide and hexa-nucleotide motifs with four repeats. Although quad-nucleotide motifs with five repeats accounted for 73.71% (272 out of 369), there was also one with 32 repeats, namely (ATCC/ATGG)₃₂.

Fig. 9 The statistics Statistics of SNPs/InDel. This is the average of three replications from each sample. Blue: total SNPs, Orange: total InDels

The analysis of SNPs/InDel (Fig. 9) showed a total of 479,831 SNPs in A7, which was the most abundant among the five samples. A total of 12 type SNPs were identified, which were all intergenic variations. There were a large proportion of A → G, G → T, G → A, and T → G type transitions, while the others were still quite low. A total of 80,132 InDels were also identified, which is significantly

A total of 3746 TFs distributed in 53 transcriptional factors families were obtained from the five samples. The five top gene families were bHLH (413), MYB (288), MYB-like (269), C2H2 (255), and bZIP (217; Fig. 10), while a few TFs were involved in the STAT, ZF-HD, and RAV families. In addition, 183, 493, 738, 923, 203, 358, and 345 differentially expressed TFs (DETFs) were identified in seven pairwise comparisons, including A5 versus A0, A7 versus A0, A15 versus A0, A20 versus A0, A7 versus A5, A15 versus A7, and A20 versus A15. Similarly to the results of differentially expressed genes, the number of DETFs between each sample and the blank control (A0) increased with the increasing degree of drought stress. However, the number of DETFs among different samples was significantly lower than that between each sample and the blank control (A0), with the exception

Confirmation of DEGs using quantitative real-time PCR (qRT-PCR)

Physiological indicators of okra under drought stress

To determine the physiological and biochemical mechanisms of okra responding to drought stress, eight related

Fig. 11 The qRT-PCR results of 10 ten DEGs. β -Actin was used as an internal control for normalization. The Y-axis indicates gene expression levels

Table 2 The pPhotosynthetic traits of leaf leaves in okra under different drought stress treatments

T	TC (mg L ⁻¹)	Pn ($\mu\text{mol m}^{-2} \text{S}^{-1}$)	Gs (mol H ₂ O m ⁻² S ⁻¹)	Ci ($\mu\text{mol CO}_2 \text{mol}^{-1}$)	SP (mg g ⁻¹)	PRO (mg g ⁻¹)	SS (mg g ⁻¹)	MDA (nmol g ⁻¹)
A0	0.43bA	18.28a*A#	0.39aA	116eD	321.53eD	0.11dC	0.13cB	232.14eE
A5	0.55aA	14.72bB	0.30bA	172dC	399.31dC	0.28cB	0.18cB	398.42dD
A7	0.57aA	12.16cB	0.21cA	223cB	439.37cB	0.49bA	0.28bA	756.28cC
A15	0.30cB	6.79dC	0.08 dB	280bA	558.72bA	0.59aA	0.43aA	1170.12bB
A20	0.22cB	5.77dC	0.05 dB	296aA	582.38aA	0.63aA	0.46aA	1350.52aA

All eight of the critical physiological and biochemical indexes were investigated at five time points (0, 5, 7, 15, and 20 days)

TC total chlorophyll, Pn photosynthetic rate, Gs stomatal conductance, Ci CO₂ concentration, SP soluble protein, PRO proline, SS soluble sugar, MDA malonaldehyde

*Lowercase letters in the same column represent significant differences at the level of $p < 0.05$

#Capital letters in the same column represent significant differences at the level of $p < 0.01$

indexes including photosynthetic rate, stomatal conductance, CO₂ concentration, chlorophyll content, soluble protein, proline, soluble sugar and malonaldehyde (MDA) contents were detected in this study. They all exhibited significant changes under different drought stress treatments (Table 2). Specifically, the stomatal conductance and net photosynthesis rates in the blades were reduced with increasing the intensity of water shortage. The chlorophyll content in the blades increased initially and then decreased. The CO₂ concentration, soluble protein, proline, soluble sugar, and malondialdehyde contents of okra also decreased when increasing the intensity of water shortage. In particular, MDA levels rapidly increased with increasing drought intensity, exhibiting significant differences among the different treatments.

Discussion

Extensive research has demonstrated that water-limited conditions have significant effects on the growth, development and yield formation of okra. Understanding the genetics of traits and genes related to drought tolerance would enable the development of some new drought-tolerant cultivars using genome-editing techniques or molecular breeding (Mahdavi et al. 2018). Plants often resist abiotic stress by altering their physiological and metabolic processes and gene expression patterns. In this study, we used the RNA-Seq method to reveal gene expressions accompanying the physiological response of okra blades to drought stress. The results showed that the average total mapped reads of transcriptome were $> 70\%$, with approximately 60% of transcripts longer than 1000 bp (N50), indicating a successful assembly is available for subsequent bioinformatics analysis. A large number of differentially expressed genes responding to drought stress

were identified in our study with a tendency for the number of DEGs to increase with increasing drought intensity. This tendency implies that the higher the degree of drought, the more obvious the effect on gene expression in plants. Similar results were also reported on other plants, such as *Camellia oleifera* (Dong et al. 2017; Xia et al. 2014) and *Cicer arietinum* L. (Mahdavi et al. 2018).

According to the GO analysis, most of the DEGs were involved in “integral component of membrane” “oxidation–reduction process”, “metal ion binding” and “DNA binding” from the seven pairs mentioned earlier. This suggests that DEGs in this cellular component, biological process, and molecular function groups would be associated with drought tolerance in plants. The expression of late embryogenesis abundant (LEA) proteins, heat shock proteins (HSP), and dehydrins were induced under water deficit and all of them may display a key role in the stabilization of membranes (Harfouche et al. 2014). These proteins are the products of DEGs involved in “integral component of membrane”. Some genes that display roles in oxidation–reduction process were identified in angiosperms and in gymnosperms under drought stress, such as fatty acid desaturases (FADs) genes (Yasemin et al. 2018). Furthermore, DNA binding proteins, such as CDC5 in *Arabidopsis*, have been shown to control many biological processes, including development and adaptations to abiotic stress (Zhang et al. 2013). Genes often play a role in certain biological processes by interacting with each other. The DEGs involved in metal ion binding process or DNA binding process are likely to enhance the drought resistance of plants through the regulation of expression of downstream target genes or accumulation of microRNAs.

Moreover, Pathway analysis can help us to better understand the genetic function and metabolism of plants. We found that the majority of DEGs were enriched in the “biosynthesis of secondary metabolites” pathway. Secondary metabolites were induced with biotic and abiotic stresses, which affect plant growth and development (Fox et al. 2017). The involvement of secondary metabolites in response to drought stress is extremely complicated and depends on various parameters (Niinemets 2016). Water deficiency could damage plant secondary metabolism, thereby interfering with the normal growth, leading to chlorosis, or even death. We also found that 3679 unigenes were mapped to “environmental information processing” based on the KEGG pathway. Among these, 3188 unigenes were related to “signal transduction”, implying that signal transduction processing may affect plants in response to adverse environmental stress (Dong et al. 2017).

As classic genetic markers, SSRs are also known as microsatellites, which are extensively distributed in plant genomes, while SNP is defined as a type of polymorphism caused by the variation of a single nucleotide at the

genome level. These molecular markers are widely used in crop genetic breeding. We found that the proportion of tri-, mono- and di-nucleotide (92.81%) was the largest among the SSRs examined under drought stress. This result is consistent with the SSRs distributions examined in alfalfa (Wang et al. 2013), barrel medic (Gupta and Prasad 2009), chickpea (Choudhary et al. 2009), castor bean (Qiu et al. 2010), peanut (Liang et al. 2009), and sweet potato (Wang et al. 2011). According to these SSR markers, we designed 62,463 primer pairs, which were involved in 12,693 unigenes. Their subsequent verification will be carried out using the PCR method.

We also found many differently expressed transcriptional factors in our study, which are widely implied in drought resistance. Several studies (Agarwal et al. 2016; Yoshida et al. 2015; Liu et al. 2016) have reported that TFs were considered to have vital roles in accommodating and resisting drought conditions. We identified abundant TFs belonging to the bHLH (413), MYB (288), MYB-like (269), C2H2 (255), and bZIP (217) families, which is consistent with reports on kabuli chickpea (*Cicer arietinum* L.) (Mahdavi et al. 2018) and tea oil camellia (*Camellia oleifera*) (Dong et al. 2017). Following this, we would like to select several key TFs as candidates for functional verification.

Our results showed that the stomatal conductance and net photosynthesis rate in blades were reduced and the CO₂ concentration, soluble protein, proline, soluble sugar, and malondialdehyde contents of okra were decreased with increasing drought stress intensity. The chlorophyll content in the blades increased initially and then decreased, which is consistent with Rabert's (Rabert et al. 2013) report on okra, and also consistent with the reports on rice (Reddy et al. 1998) and *Zea mays* (Jiang and Zhang 2002). The chlorophyll content, stomatal conductance, net photosynthesis rate, and CO₂ concentration are a series of important indexes, reflecting the photosynthesis in plants. Our results suggest that drought stress significantly inhibited photosynthesis in the leaves of okra, and the incidence was correlated with the degree of drought, which is consistent with Altaf's results (Altaf et al. 2015). Free amino acid proline, soluble protein, and soluble sugar are the most common osmolytes used in maintaining high turgor for drought adjustment (Krasensky and Jonak 2012). Accumulation of proline soluble sugar and soluble protein may counteract the damage induced by water shortages. Furthermore, MDA is a highly active lipid peroxide that can affect the balance of an active oxygen metabolism system (Song et al. 2016). The MDA contents in okra increased under drought stress, indicating that the peroxidation in cells was high and the cytoplasmic membrane would be impaired. Drought stress can increase intracellular reactive oxygen species (ROS) to explode, thus causing intracellular

metabolic disorders. Plants may resist the damage induced by drought stress by improving their antioxidant capacity.

The response of plants to drought stress is very complex because of the interactions between drought stress factors and the physiological and biochemical indexes affecting plant growth and development. Therefore, the relationship between gene expression and metabolism deserves further study in order to fully reveal the physiological and molecular mechanisms of drought tolerance in plants.

Conclusion

Our findings reveal the differential gene expression profiles of okra in response to drought stress and corroborate other findings showing water deficits can significantly alter the expression of some critical genes and transcription factors involved in secondary metabolism. These results may help explain the underlying drought resistance mechanisms in plants.

Acknowledgements This study was financed by the National Natural Science Foundation of China (No. 30971794), the Science and Technology Foundation of Guizhou Province, China (No. QKHJC[2017]1004), the Training Program for High-level Innovative Talents of Guizhou Province (No. QKHRC[2016]4020), the Program for First-class Discipline Construction in Guizhou Province (No. QJKYF[2017]85), the National Innovation Training Program for Undergraduate (Nos. 2018109760007, 2018109760008), the Innovative Training Program for Undergraduate in Guizhou Province (Nos. 2018520777, 2018520778, 2018520831, 2018520861), Foundation of Guiyang science and technology bureau (GYU-KYZ[2019-2020]PT13-03), and the Project for Major Research by Innovation Group in Guizhou Education Department (No. QJHKY[2018]027).

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Agarwal G, Garg V, Kudapa H, Doddamani D, Pazhamala LT, Khan AW (2016) Genome-wide dissection of AP2/ERF and HSP90 gene families in five legumes and expression profiles in chickpea and pigeonpea. *Plant Biotechnol J* 14(7):1563–1577
- Altat R, Hussain K, Maryam U et al (2015) Effect of different levels of drought on growth, morphology and photosynthetic pigments of lady finger (*Abelmoschus esculentus*). *World J Agric Sci* 11(4):198–201
- Anonymous (2015) Indian horticulture database 2014, National Horticulture Board, Ministry of Agriculture, Government of India, Gurgaon, New Delhi
- Bai Z, Wang T, Wu Y, Wang K, Liang Q, Pan Y, Jiang B, Zhang L, Liu G, Jia Y, Liu Q (2017) Whole-transcriptome sequence analysis of differentially expressed genes in *Phormium tenax* under drought stress. *Sci Rep* 7:41700
- Bhaskara G, Yang T, Verslues P (2015) Dynamic proline metabolism: importance and regulation in water limited environments. *Front Plant Sci* 6:484
- Calisir S, Ozcan M, Haciseferogullari H, Yildiz M (2005) A study on some physico-chemical properties of Turkey okra (*Hibiscus esculentus* L.) seeds. *J Food Eng* 68:73–78
- Choudhary S, Sethy NK, Shokeen B, Bhatia S (2009) Development of chickpea EST-SSR markers and analysis of allelic variation across related species. *Theor Appl Genet* 118:591–608
- Dhankhar S, Dhankhar B, Yadava R (2005) Inheritance of resistance to yellow vein mosaic virus in an interspecific cross of okra (*Abelmoschus esculentus*). *Indian J Agric Sci* 75(23):87–89
- Dong B, Wu B, Hong W, Li X, Li Z, Xue L (2017) Transcriptome analysis of the tea oil camellia (*Camellia oleifera*) reveals candidate drought stress genes. *PLoS ONE* 12(7):e0181835
- Elagib N (2014) Development and application of a drought risk index for food crop yield in Eastern Sahel. *Ecol Indic* 43:114–125
- Faostat (2014) <http://faostat.fao.org/>. Accessed 8 Feb 2017
- Fox H, Doron-Faigenboim A, Kelly G, Bourstein R, Attia Z, Zhou J, Moshe Y, Moshelion M, David-Schwartz R (2017) Transcriptome analysis of *Pinus halepensis* under drought stress and during recovery. *Tree Physiol* 38:423–441. <https://doi.org/10.1093/treephys/tpx137>
- Gupta S, Prasad M (2009) Development and characterization of genic SSR markers in *Medicago truncatula* and their transferability in leguminous and non-leguminous species. *Genome* 52:761–771
- Han BC, Wei W, Mi XC, Ma KP (2016) De novo sequencing and comparative analysis of *Schima superba* seedlings to explore the response to drought stress. *PLoS ONE* 11(12):e0166975
- Harfouche A, Meilan R, Altman A (2014) Molecular and physiological responses to abiotic stress in forest trees and their relevance to tree improvement. *Tree Physiol* 34:1181–1198
- Hu Y, Wang B, Hu T, Chen H, Li H (2015) Combined action of an antioxidant defence system and osmolytes on drought tolerance and post-drought recovery of *Phoebe zhennan* S. Lee saplings. *Acta Physiol Plant* 37(4):1–13
- Jiang M, Zhang J (2002) Water stress-induced abscisic acid accumulation triggers the increased generation of reactive oxygen species and up-regulates the activities of antioxidant enzymes in maize leaves. *J Exp Bot* 53(379):2401–2410
- Jin J, Zhang H, Kong L, Gao G, Luo J (2014) PlantTFDB 3.0: a portal for the functional and evolutionary study of plant transcription factors. *Nucl Acids Res* 42(Database issue):D1182–D1187
- Krasensky J, Jonak C (2012) Drought, salt, and temperature stress-induced metabolic rearrangements and regulatory networks. *J Exp Bot* 63(4):1593–1608
- Kumar M, Sharma V, Naresh R (2017) Screening of microsatellite markers for genetic diversity assessment and conservation of germplasm in okra (*Abelmoschus*). *Int J Curr Microbiol Appl Sci* 6(6):509–520
- Li KQ, Xu XY, Huang XS (2016) Identification of differentially expressed genes related to dehydration resistance in a highly drought-tolerant pear, *Pyrus betulaeifolia*, as through RNA-Seq. *PLoS ONE* 11(2):e0149352
- Liang X, Chen X, Hong Y, Liu H, Zhou G et al (2009) Utility of EST-derived SSR in cultivated peanut (*Arachis hypogaea* L.) and *Arachis* wild species. *BMC Plant Biol* 9:35

- Liu L, Zhang Z, Dong J, Wang T (2016) Overexpression of MtWRKY76 increases both salt and drought tolerance in *Medicago truncatula*. *Environ Exp Bot* 123:50–58
- Lu J, Yuan J, Yang J, Yang ZY (2014) Responses of morphology and drought tolerance of *Sedum lineare* to watering regime in green roof system: a root perspective. *Urban For Urban Gree* 13(4):682–688
- Mahdavi Mashaki K, Garg V, Nasrollahnezhad Ghomi A, Kudapa H, Chitkineni A, Zaynali Nezhad K (2018) RNA-Seq analysis revealed genes associated with drought stress response in kabuli chickpea (*Cicer arietinum* L.). *PLoS ONE* 13(6):e0199774
- Mishra G, Singh B, Seth T, Singh A, Halder J, Krishnan N, Tiwari S, Singh P (2017) Biotechnological advancements and begomovirus management in okra (*Abelmoschus esculentus* L.): status and perspectives. *Front Plant Sci* 17(8):360
- Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B (2008) Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods* 5(7):621–628
- Niinimets Ü (2016) Uncovering the hidden facets of drought stress: secondary metabolites make the difference. *Tree Physiol* 36:129–132
- Ogbaga C, Stepien P, Dyson B, Rattray N, Ellis D (2016) Biochemical analyses of sorghum varieties reveal differential responses to drought. *PLoS ONE* 11(5):e0154423
- Park J, Park J, Jang S, Kim S, Kong S, Choi J, Ahn K, Kim J, Lee S, Kim S, Park B, Jung K, Kim S, Kang S, Lee YH (2008) FTFD: an informatics pipeline supporting phylogenomic analysis of fungal transcription factors. *Bioinformatics* 24(7):1024–1025
- Paul A, Jha A, Bhardwaj S, Singh S, Shankar R, Kumar S (2014) RNA-seq-mediated transcriptome analysis of actively growing and winter dormant shoots identifies non-deciduous habit of evergreen tree tea during winters. *Sci Rep* 4:5932
- Qiu LJ, Yang C, Tian B, Yang JB, Liu AH (2010) Exploiting EST databases for the development and characterization of EST-SSR markers in castor bean (*Ricinus communis* L.). *BMC Plant Biol* 110:278
- Rabert G, Rajasekar M, Manivannan P (2013) Effect of triazole fungicide on biochemical and antioxidant enzymes activity in okra (*Abelmoschus esculentus* L.) plant under drought stress. *Int J Agric Food Sci* 3(3):100–107
- Rai D, Balasubramanian S (2009) Qualitative and textural changes in fresh okra pods (*Hibiscus esculentus* L.) under modified atmosphere packaging in perforated film packages. *Food Sci Technol Int* 15(15):131–138
- Reddy C, Vajranabhan S, Kumar U (1998) Lipid peroxidation as a mechanism of stress tolerance in upland rice (*Oryza sativa* L.) Calli. *Ind J Plant Physiol* 3(1):68–70
- Rosero C, Argout X, Ruiz M, Teran W (2011) A drought stress transcriptome profiling as the first genomic resource for white teak—Gamhar—(*Gmelina arborea* Roxb) and related species. *BMC Proc* 5(S7):178
- Salameh N (2014) Genetic diversity of okra (*Abelmoschus esculentus* L.) landraces from different agro-ecological regions revealed by AFLP analysis. *Am Eur J Agric Environ Sci* 14:155–160
- Sarah K, Sarah A (2006) DBD: a transcription factor prediction database. *Nucl Acids Res* 34(Database issue):D7–D81
- Schaffleitner R, Kumar S, Lin C, Hegde S, Ebert A (2013) The okra (*Abelmoschus esculentus*) transcriptome as a source for gene sequence information and molecular markers for diversity analysis. *Gene* 17:27–36
- Shen J, Xiao Q, Qiu H, Chen C, Chen H (2016) Integrative effect of drought and low temperature on litchi (*Litchi chinensis* Sonn.) floral initiation revealed by dynamic genome-wide transcriptome analysis. *SciRep* 6:32005
- Song H, Wang P, Li C, Han SY, Lopez-Baltazar J, Zhang XY et al (2016) Identification of lipoxygenase (LOX) genes from legumes and their responses in wild type and cultivated peanut upon *Aspergillus flavus* infection. *Sci Rep* 6:35245
- Sperdouli I, Moustakas M (2012) Interaction of proline, sugars, and anthocyanins during photosynthetic acclimation of *Arabidopsis thaliana* to drought stress. *J Plant Physiol* 169(6):577–585
- Sperlich D, Barbata A, Ogaya R, Sabaté S, Peñuelas J (2016) Balance between carbon gain and loss under long-term drought: impacts on foliar respiration and photosynthesis in *Quercus ilex* L. *J Exp Bot* 67(3):821–833
- Tajiri T (2011) Effect of the storage temperatures on changes of food physical properties and composition of Okra pods (*Hibiscus esculentus* L.). Part I. Studies on keeping quality of Okra pods (*Hibiscus esculentus* L.). *J Jpn Inst Cold Chain* 12(3):81–86
- Tang S, Dong Y, Liang D, Zhang Z, Ye CY, Shuai P et al (2015) Analysis of the drought stress-responsive transcriptome of black cottonwood (*Populus trichocarpa*) using deep RNA sequencing. *Plant Mol Biol Rep* 33(3):424–438
- Varmud V (2011) Marketing survey needed to boost okra exports. Department of economic Vive Kananda College, Puttur Karnataka
- Wan J, Griffiths R, Ying J, McCourt P, Huang Y (2009) Development of drought-tolerant canola (*Brassica napus* L.) through genetic modulation of ABA-mediated stomatal responses. *Crop Sci* 49(5):1539–1554
- Wang Z, Gerstein M, Snyder M (2009) RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet* 10(1):57–63
- Wang Z, Li J, Luo Z, Huang L, Chen X (2011) Characterization and development of EST-derived SSR markers in cultivated sweetpotato (*Ipomoea batatas*). *BMC Plant Biol* 11:139
- Wang Z, Yan HW, Fu XN, Li XH, Gao HW (2013) Development of simple sequence repeat markers and diversity analysis in alfalfa (*Medicago sativa* L.). *Mol Biol Rep* 40:3291–3298
- Xia E, Jiang J, Huang H, Zhang L, Zhang H, Gao L (2014) Transcriptome analysis of the oil-rich tea plant, *Camellia oleifera*, reveals candidate genes related to lipid metabolism. *PLoS ONE* 9(8):e104150
- Yasemin CA, Unel Necdet Mehmet, Baloglu Mehmet Cengiz, Ulu Ferhat, Can Tevfik Hasan, Cetinkaya Rahmi (2018) Comparative identification and evolutionary relationship of fatty acid desaturase (FAD) genes in some oil crops: the sunflower model for evaluation of gene expression pattern under drought stress. *Biotechnol Biotechnol Equip*. <https://doi.org/10.1080/13102818.2018.1480421>
- Yoshida T, Fujita Y, Maruyama K, Mogami J, Todaka D, Shinozaki K (2015) Four Arabidopsis AREB/ABF transcription factors function predominantly in gene expression downstream of SnRK2 kinases in abscisic acid signaling in response to osmotic stress. *Plant Cell Environ* 38(1):35–49
- Zeng X, Bai L, Wei Z, Yuan H, Wang Y, Xu Q et al (2016) Transcriptome analysis revealed the drought-responsive genes in Tibetan hulless barley. *BMC Genom* 17:386
- Zhang S, Xie M, Ren G, Yu B (2013) CDC5, a DNA binding protein, positively regulates posttranscriptional processing and/or transcription of primary microRNA transcripts. *PANC* 110(43):17588–17593
- Zhang HM, Liu T, Liu CJ, Song S, Zhang X, Liu W, Jia H, Xue Y, Guo AY (2015) AnimalTFDB 2.0: a resource for expression, prediction and functional study of animal transcription factors. *Nucl Acids Res* 43(D1):D76–D81
- Zhang S, Qiu S, Zheng Y, Zhang S, Wu S, He Y, Zheng K (2017) The purple transcriptome as a source for gene sequence information.

J Nucl Agric Sci 31(4):643–663 (in Chinese with English abstract)

Zivcak M, Brestic M, Balatova Z, Drevenakova P, Olsovska K, Kalaji HM (2013) Photosynthetic electron transport and specific photoprotective responses in wheat leaves under drought stress.

Photosynth Res 117(1–3):529–546. <https://doi.org/10.1007/s11120-013-9885-3> PMID: 23860828

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.