ORIGINAL PAPER

Global gene expression responses to waterlogging in roots of sesame (*Sesamum indicum* L.)

Linhai Wang · Yanxin Zhang · Xiaoqiong Qi · Donghua Li · Wenliang Wei · Xiurong Zhang

Abstract Waterlogging stress lowers yields in sesame (*Sesamum indicum* L.). A major component of waterlogging stress is the lack of oxygen available to submerged tissues. Although the morphology and physiology of plants grown under anaerobic conditions have been studied in detail, limited work has been done to elucidate adaptations at the molecular level. To gain comprehensive insight into how sesame responds to hypoxia at the genome level, we performed gene expression profiling at two time points during a 36-h period following hypoxic treatment using a whole-genome RNA-Seq analysis. We identified sets of significantly positively and negatively expressed genes (induced and repressed, respectively) in response to hypoxia with distinct temporal profiles. The genes that were affected were associated with glycolysis, nitrogen metabolism, starch and sucrose metabolism and plant hormone signal transduction and indicated the upregulation of particular pathways (glycolysis/glycogenesis) in the Kyoto Encyclopedia of Genes and Genomes. Moreover, significant changes in the expression of genes were found for pathways, including flavone and flavonol biosynthesis, steroid biosynthesis, photosynthesis, cysteine and methionine metabolism, glutathione metabolism, as well as phenylpropanoid biosynthesis, spliceosome, circadian rhythm. This study helps in elucidating the molecular mechanisms of waterlogging tolerance and provides a basis for the genetic engineering of sesame.

Introduction

Sesame (*Sesamum indicum* L.), a member of the Pedaliaceae family, is a diploid ($2n = 26$) dicotyledon. It is one of the most ancient oil seed crops, and it is grown widely in tropical and subtropical areas (Ashri 2010; Bedigian and Harlan 1986). Sesame is extremely susceptible to waterlogging and continuous heavy rains. When grown on soils with poor drainage, sesame is adversely affected and can suffer yield losses of greater than 30 % (in severe cases, 50–90 %). Field experiments on waterlogged sesame have recorded premature senescence resulting from leaf chlorosis, necrosis, defoliation, and reduced nitrogen fixation, leading to cessation of growth and reduced yield (Snowden and Wheeler 1993). Effects of waterlogging and growth regulators on the yield as well as morphological and physiological characteristics of sesame have been investigated. The growth rate, capsules per plant, biomass, seed yield, net photosynthesis (Pn), and chlorophyll content (Chl) are markedly reduced after plant waterlogging during various stages of growth (Sun et al. 2009). This can lead to fungal attack by *Fusarium oxysporum* and *Macrophomina phaseolina*, which induce the serious diseases Fusarium wilt and charcoal rot, respectively, in sesame (Liu et al. 1993). Field observations that have demonstrated changes in the growth rate, photosynthetic rate, and mineral nutrient content of waterlogged sesame plants indicate that waterlogging stress does more than just alter the energy

L. Wang · Y. Zhang · X. Qi · D. Li · W. Wei · X. Zhang (✉)
Key Laboratory of Biology and Genetic Improvement of Oil Crops, Ministry of Agriculture, Sesame Germplasm and Genetic Breeding Laboratory, Oil Crops Research Institute of the Chinese Academy of Agricultural Sciences (OCRI-CAAS), Wuhan 430062, People's Republic of China
e-mail: zhangxr@oilcrops.cn

metabolism of roots. Comprehensive evaluations of waterlogging tolerance of different sesame varieties and plant types have identified several germplasm lines with waterlogging tolerance, providing important material for genetic improvement (Sun et al. 2010).

Waterlogging results in decreased oxygen (hypoxia) supply to the plant rhizosphere due to the low diffusion rate of molecular oxygen in water (Christianson et al. 2010). This causes plants to switch from respiration to fermentative metabolism to maintain energy supply for continued metabolism and growth under anaerobic conditions. Therefore, a lack of oxygen in the root zone affects the maintenance of numerous cellular pathways (Drew 1997) and results in morphological, anatomical, and metabolic adaptations to flooding. Responses to low oxygen levels have been studied in *Arabidopsis thaliana* (Klok et al. 2002; Liu et al. 2005) and in many crops (Kreuzwieser et al. 2009; Christianson et al. 2010; Du et al. 2010), many of which show rapid changes in gene transcription, protein synthesis and degradation, and cellular metabolism (Bailey-Serres and Voesenek 2008). Global gene expression studies in *Arabidopsis*, rice, poplar, and cotton have revealed complex responses to low oxygen involving significant changes in 5–10 % of all genes assayed (Klok et al. 2002; Liu et al. 2005; Loreti et al. 2005; Lasanthi-Kudahettige et al. 2007; Kreuzwieser et al. 2009). In these studies, the genes affected were associated with cell wall growth and modification, hormone response, starch metabolism, and nitrogen metabolism. Furthermore, promoter elements and transcription factors involved in the regulation of anaerobically induced genes have been characterized (Yang et al. 2011; Banti et al. 2010), and some of these genes have been cloned, e.g., *XET* (xyloglucan endotransglycosylase) (Wang et al. 2004) and Arabidopsis *RAP2.2* (Related to AP2 2) (Hinz et al. 2010). Therefore, molecular approaches have sought to enhance anaerobic responses by introducing individual genes involved in signaling (hormones or transcription factors) or in metabolic pathways known to be important to anoxic tolerance. In sesame, although the morphology and physiology of plants grown under anaerobic conditions have been studied in detail, limited work has been done to elucidate adaptations at the molecular level. Only suppression subtractive hybridization (SSH) has been used to construct a cDNA library for isolation of waterlogging tolerance-related genes for basic material and energy metabolism, signal transduction, transcription regulation, and detoxification defense responses (Wang et al. 2010). *XET* gene involved in cell wall metabolism, which can be induced in roots of sesame under anoxia stress, has been cloned (Wang et al. 2004). However, the mechanisms associated with tolerance to flooding in sesame remain unclear.

To gain comprehensive insight into how sesame responds to hypoxia at the genome level, we conducted gene expression profiling at two time points during a 16-h period following hypoxic treatment of plants using whole-genome RNA-Seq analysis. We identified sets of significantly induced and repressed genes in response to hypoxia whose expression showed distinct temporal profiles. Our analyses based on the gene ontology (GO; http://www.geneontology.org) terms for these responsive genes, together with pathway enrichment analysis (http://www.genome.jp/kegg/), suggest that hypoxia affects a broad spectrum of functional gene categories. This study helps to elucidate the molecular mechanisms of sesame waterlogging tolerance and provides a basis for the genetic engineering of sesame.

Materials and methods

Plant material and experimental setup

Experiments were performed with the sesame waterlogging-tolerant cultivar "3321" at the experimental station of the Oil Crops Research Institute, Chinese Academy of Agricultural Sciences, Wuhan, China. Sesame plants at anthesis were irrigated to above soil level so that the surface was covered with a thin layer of water, which was maintained for 16 h. After a further 9 h of treatment, the flooded plants were harvested, and roots were immediately frozen in liquid nitrogen and stored at −80 °C until further analysis. As controls, plants 15 h before the flooding were also harvested.

RNA isolation and cDNA synthesis

Total RNA was isolated using TRIzol reagent according to the manufacturer's instructions (Invitrogen). The concentration of total RNA was determined using an ultraviolet (UV) spectrophotometer, and RNA intactness was assessed on 1.0 % denaturing agarose gels. The RNA was subjected to RNA-Seq analysis at the Beijing Genomics Institute (BGI, Shenzhen, China). RNA quality and quantity was verified using a NanoDrop 1000 spectrophotometer and an Agilent 2100 Bioanalyzer prior to further processing at BGI. Total RNA was treated with DNase I prior to library construction, and then enriched using oligo-(dT) magnetic beads. Fragmentation buffer was used to cut the mRNA into short fragments (about 200 bp) then the first-strand cDNA was synthesized using random hexamer primers with the mRNA fragments as templates. Buffer, dNTPs, RNase H, and DNA polymerase I were added to synthesize the second strand. The double-strand cDNA was purified with the QiaQuick polymerase chain reaction (PCR) extraction kit and washed with EB buffer for end repair and single nucleotide A (adenine) addition. Finally, sequencing

adaptors were ligated to the fragments. The required fragments were purified via agarose gel electrophoresis and enriched by PCR amplification. The library products were prepared for sequencing analysis with an Illumina HiSeq™ 2000.

Aligning clean reads to the reference transcriptome data set

The cDNA library was sequenced using Solexa/Illumina RNA-Seq. The 49-bp raw reads were generated with the Illumina HiSeq™ 2000. Image deconvolution and quality value calculations were performed using the Illumina HCS 1.1 software. The raw reads were cleaned by removing the adaptor sequences, reads containing undefined nucleotides ('N's) and low-quality reads (the percentage of low-quality bases with a quality value ≤5 was more than 50 % in each read). Clean reads were then mapped to reference sesame transcriptome sequences using SOAPaligner/soap2 (Li et al. 2009). Mismatches of no more than two bases were allowed in the alignment.

Statistical analysis of gene expression level

Gene expression was calculated from the number of reads mapped to the reference sequence and for every gene. The expression level was calculated using the RPKM (Mortazavi et al. 2008) method (reads per kb per million reads) with the formula:

$$RPKM = \frac{C}{NL} \times 10^9,$$

where RPKM(A) is the expression of gene A, C is the number of reads uniquely aligned to gene A, N is the total number of reads uniquely aligned to all genes, and L is the number of bases on gene A. The RPKM method eliminates the influence of different gene lengths and sequencing discrepancies on the calculation of gene expression, making it possible to use these calculations to directly compare differences in gene expression among samples. If there is more than one transcript for a gene, the longest is used to calculate its expression level and coverage.

Identification of differentially expressed genes

Rigorous algorithms have been developed to identify differentially expressed genes between samples. The correlation of the detected count number between parallel libraries was assessed statistically by calculating the Pearson's correlation (P). In addition to the P value, false discovery rate (FDR) was used to determine the threshold P value in multiple tests. If we assume that we have selected R differentially expressed genes in which S genes

show real differential expression and V genes are false positives, the error ratio "$Q = V/R$" must remain below a cutoff (e.g., 1 %), and the FDR should be preset to a number no larger than 0.01 (Benjamini and Yekutieli 2001). We used FDR ≤ 0.001 and the absolute value of $\log_2$Ratio ≥ 1 as the threshold to judge the significance of differences in gene expression. More stringent criteria with a smaller FDR and a bigger fold-change value can be used to identify differentially expressed genes (DEGs).

Gene ontology analysis of DEGs

Gene ontology (GO) is an international standardized gene function 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. GO has three ontologies: molecular function, cellular component, and biological process. The basic unit of GO is the GO term. Every GO term belongs to one type of ontology. GO enrichment analysis provides all GO terms that are significantly enriched in DEGs as compared to the genome background, and filters the reads that correspond to biological functions. This method first maps all DEGs to GO terms in the database (http://www.geneontology.org/), calculating gene numbers for every term, then uses a hypergeometric test to find significantly enriched GO terms in DEGs as compared to the genome background. The formula is:

$$P = 1 - \sum_{i=0}^{m-1} \frac{\binom{M}{i}\binom{N-M}{n-i}}{\binom{N}{n}},$$

where N is the number of all genes with a GO annotation, n is the number of DEGs in N, M is the number of genes that are annotated to certain GO terms, and m is the number of DEGs in M (Yang et al. 2007). The calculated P value, after Bonferroni's correction, takes a corrected $P \leq 0.05$ as a threshold. GO terms fulfilling this condition are defined as significantly enriched in DEGs. This analysis recognizes the main biological functions of DEGs.

Pathway enrichment analysis of DEGs

Genes usually interact with each other to carry out specific biological functions. Pathway-based analysis helps to further understand gene biological functions. The Kyoto Encyclopedia of Genes and Genomes (KEGG) is the major public pathway-related database (Saldanha 2004) used to identify significantly enriched metabolic pathways or signal transduction pathways in DEGs as compared to whole-genome backgrounds. The formula used is the same as that

for GO analysis. Here, N is the number of all genes with a KEGG annotation, n is the number of DEGs in N, M is the number of all genes annotated to specific pathways, and m is the number of DEGs in M.

Real-time quantitative RT-PCR (qRT-PCR) analysis

To validate the repeatability and reproducibility of gene expression data obtained by RNA sequencing in sesame, we performed qRT-PCR on 10 differentially expressed genes with the total RNA used in RNA-Seq. RNA isolation and quality verification were performed as mentioned above. The first-strand cDNA fragment was synthesized from total RNA using Fernments Reverse transcription Kit (K1652). The quantitative RT-PCR was performed with a iQ^{TM}5 Real-Time PCR Detection Systems (Bio-Rad) in a final volume of 20 μl containing 12.5 μl 2× SYBR Green Real-time PCR Master Mix (TOYOBO), 2 μl of cDNA template, and 0.2 pM each primer. The thermal cycling conditions were as follows: 95 °C for 3 min; 40 cycles of 95 °C for 10 s, annealing temperature for 15 s, and 72 °C for 45 s; and 72 °C for 5 min.

Samples were run in triplicate on the same plate with a negative control that lacked cDNA. Positive controls were set for each sample using sesame gene ubiquitin-conjugating enzyme 9 (UBC9). PCR efficiency was determined by a series of 10-fold dilutions of cDNAs. The calculated efficiency of all primers was 0.9–1.0. The relative expression levels of genes were calculated using the $2^{-\Delta\Delta CT}$ method (Livak and Schmittgen 2001), which represents the difference of CT between the control UBC9 and the target gene products.

Results

RNA-Seq analysis

Solexa/Illumina RNA-Seq analysis was performed to identify genes involved in sesame's response to waterlogging. This approach generates absolute rather than relative gene expression measurements and avoids many of the inherent limitations of microarray analysis. This method produced 13.05 million and 12.98 million raw reads per library, respectively. After transforming the raw sequences into clean reads, approximately 12 million high-quality non-redundant reads were obtained in both groups (Fig. 1). The data sets represent the expressed sequences or transcriptome for each library. Matching the reads to genes is important to annotate sequences; it can reveal molecular events behind gene expression (Hegedus et al. 2009). Gene annotation was performed by read mapping analysis using the 86,222 non-redundant consensus sequences from
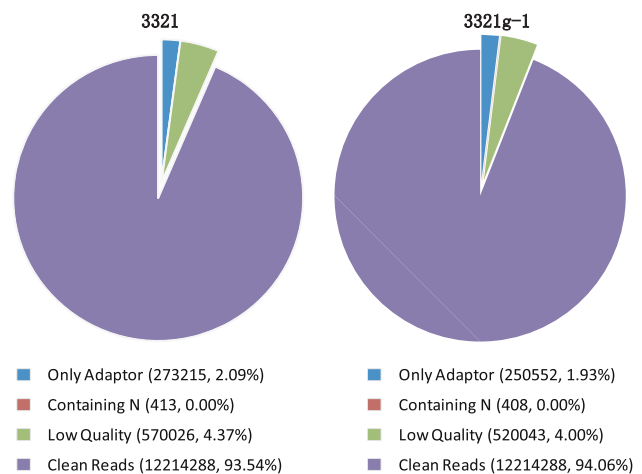

Fig. 1 Classification of raw reads, including clean reads, adaptor sequences, reads containing undefined nucleotides ('N's) and low-quality reads generated from Solexa/Illumina RNA-Seq. 3321 indicated sample of 15 h before the waterlogging stress, while 3321g-1 indicated sample of 9 h treatment; 0.00 % was the approximate value

Table 1 Statistics of the distinct reads mapped to reference database

	3321		3321g-1	
	Read number	Percentage	Read number	Percentage
Total reads	12,214,287	100.00	12,214,287	100.00
Total base pairs (bp)	598,500,063	100.00	598,500,063	100.00
Total mapped reads	7,570,363	61.98	8,580,510	70.25
Perfect match	5,256,020	43.03	6,569,708	53.79
≤2 bp mismatch	2,314,343	18.95	2,010,802	16.46
Unique match	6,568,842	53.78	7,736,004	63.34
Multi-position match	1,001,521	8.20	844,506	6.91
Total unmapped reads	4,643,924	38.02	3,633,777	29.75

3321 indicated sample of 15 h before the waterlogging stress, while 3321g-1 indicated sample of 9 h treatment

RNA-Seq-based transcriptome analysis as a reference transcript database (Wei et al. 2011). A total of 61.98 and 70.25 % of all distinct reads could be mapped to the entire reference database (sense or antisense) of each group, respectively (Table 1), and approximately 53.78 and 63.34 % of the clean reads could be uniquely mapped to the reference sequences, respectively. These reads covered 56,177 and 55,065 of the 86,222 non-redundant consensus
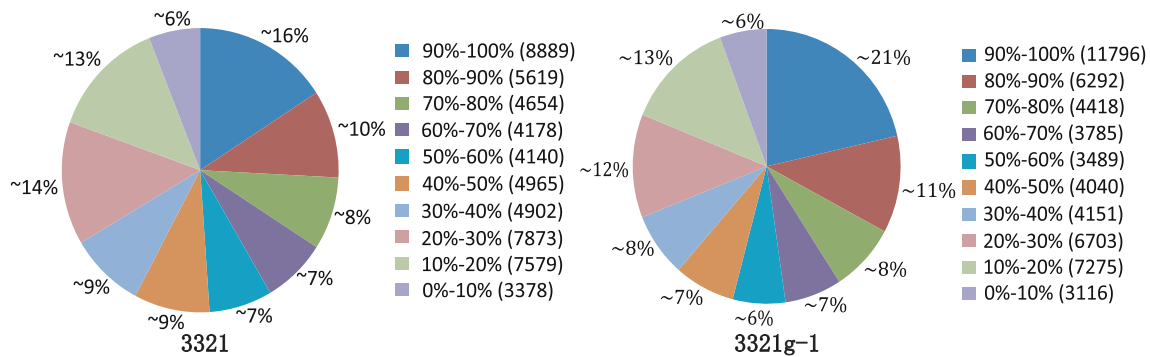
Fig. 2 Distributions of gene coverage values of 3321 and 3321g-1. 3321 indicated sample of 15 h before the waterlogging stress, while 3321g-1 indicated sample of 9 h treatment

sequences, respectively. The distributions of gene coverage values based on these reads from each library were shown in Fig. 2.

Changes in gene expression profiles during different treatments

The differentially expressed genes between the two libraries provide a clue as to the molecular events related to waterlogging. The expression abundance of read-mapped genes in the data sets was calculated using the RPKM (Mortazavi et al. 2008) method. Soap2 software was used to map all measured reads to the corresponding assembled consensus sequences (Li et al. 2009). We first normalized the read density measurement and then used FDR < 0.001 and the absolute value of |log$_2$Ratio| ≥ 1 as a threshold to judge the statistical significance of the gene expression. Exposure to root hypoxia altered the transcript levels of more 13,307 root genes during the experiment (Table 2 lists the partial DEGs). Of these, more than 60.86 % (8,099) were upregulated and 39.14 % (5,208) were downregulated, with 823 genes (6 % of the DEGs) showing a 10-fold change in expression with an adjusted $P < 0.001$. Of these, 406 (49 %) genes were upregulated and 417 (51 %) were downregulated in response to waterlogging stress. Genes with altered expression spanned a wide variety of regulatory and metabolic processes.

Hypoxia-perturbed pathways and functional categories

Functional assignments were defined by GO terms (http://www.geneontology.org/), which provide broad functional classifications for genes and gene products representing their corresponding biological process, molecular function, and cellular localization; pathway assignments were derived from KEGG (http://www.genome.ad.jp/kegg). A wide spectrum of physiological processes was affected by low-oxygen stress, as evidenced by an overrepresentation

of the corresponding GO terms. In the biological processes category, metabolic processes (76.30 %) were dominant, followed by cellular processes (67.10 %), primary metabolic processes (58.30 %), cellular metabolic processes (56.40 %), and macromolecule metabolic processes (38.80 %). With regard to molecular function, 61 % of unigenes were assigned to catalytic activity, followed by binding (60.5 %), transferase activity (22.6 %), and ion binding (19.7 %). In the cellular component category, cell parts (98.8 %) and cells (98.8 %) were dominant, followed by intracellular (75.7 %) and intracellular parts (74.6 %). The GO terms identified as being significantly overrepresented were consistent among the molecular function, biological process, and cellular component classes and agreed with KEGG pathway analysis; e.g., enrichment of glycolysis, nitrogen metabolism, starch and sucrose metabolism and plant hormone signal transduction glycolysis, indicated upregulation of a particular KEGG pathway (glycolysis/glycogenesis). In addition, significant changes in the expression of genes were found for pathways, including flavone and flavonol biosynthesis, steroid biosynthesis, photosynthesis, cysteine and methionine metabolism, glutathione metabolism, as well as phenylpropanoid biosynthesis, spliceosome, circadian rhythm. Table 2 lists the partial DEGs and their KEGG pathway analysis.

Confirmation of tag-mapped genes by qRT-PCR

To confirm the reliability of Solexa/Illumina sequencing technology, we used qRT-PCR to validate the expression levels of 10 differentially expressed genes randomly selected from Table 2. Gene-specific primer pairs (Table 3) were designed according to the 10 gene sequences using primer3-blast program available online (NCBI, USA). The results showed the expressed patterns of the 10 genes were consistent each other in the two methods (Fig. 3) with a high correlation ($R^2 = 0.89$), indicating the authenticity of these differentially expressed genes.

Table 2 Partial differentially expressed genes in 3321 and 3321g-1 with annotation

Gene ID	Description	3321-RPKM	3321g-1-RPKM	log2Ratio(3321g-1/3321)	Up–down regulation (3321g-1/3321)	Pathway
Unigene15054_All	Chloroplastic glutamine synthetase	0.001	13.354	13.705	Up	Flavone and flavonol biosynthesis
Unigene21702_All	Gamma-glutamyl transpeptidase	0.001	11.173	13.448	Up	Cysteine and methionine metabolism
Unigene27494_All	Receptor-like protein kinase	0.001	13.284	13.697	Up	Steroid biosynthesis
Unigene28905_All	16-hydroxytabersonine O-methyltransferase	0.001	24.717	14.593	Up	Photosynthesis–antenna proteins
Unigene32938_All	Carbonic anhydrase	0.001	21.544	14.395	Up	Carbon fixation in photosynthetic organisms
Unigene33532_All	Glucose/ribitol dehydrogenase family protein	0.001	15.473	13.917	Up	Benzoxazinoid biosynthesis
Unigene33542_All	Non-hemic monooxygenase	0.001	13.978	13.771	Up	Flavonoid biosynthesis
Unigene38353_All	Peroxidase	0.001	32.942	15.008	Up	Starch and sucrose metabolism
Unigene41407_All	Beta-glucosidase 24	0.001	14.896	13.863	Up	Cyanoaminoacid metabolism
Unigene45015_All	Oxygen-evolving enhancer protein 1	0.001	16.378	13.999	Up	Starch and sucrose metabolism
Unigene56503_All	Cytochrome P450	0.001	18.292	14.159	Up	Carotenoid biosynthesis
Unigene58698_All	Sedoheptulose-1,7-bisphosphatase	0.001	11.217	13.453	Up	Fatty acid biosynthesis
Unigene60241_All	Desacetoxyvindoline 4-hydroxylase	0.001	13.485	13.719	Up	Nitrogen metabolism
Unigene61376_All	Similar to flavanone 3-hydroxylase-like protein	0.001	16.959	14.050	Up	Nitrogen metabolism
Unigene84221_All	L-Ascorbate oxidase precursor	0.001	28.931	14.820	Up	Plant hormone signal transduction
Unigene65781_All	Fasciated ear2	0.001	14.083	13.782	Up	Phenylalanine metabolism
Unigene73565_All	Auxin-induced beta-glucosidase	0.001	12.429	13.601	Up	Starch and sucrose metabolism
Unigene14464_All	–	25.754	0.001	−14.653	Down	Phenylalanine, tyrosine and tryptophan biosynthesis
Unigene14934_All	Unnamed protein product	4.060	0.001	−11.987	Down	Mismatch repair
Unigene15995_All	Chlorophyll a–b binding protein 4	4.114	0.001	−12.006	Down	Photosynthesis–antenna proteins
Unigene17298_All	Transcription initiation factor IIF, alpha subunit family protein, expressed	6.963	0.001	−12.765	Down	Basal transcription factors
Unigene19448_All	ATP binding/kinase/protein kinase/protein serine/threonine kinase/ubiquitin-protein ligase	23.467	0.001	−14.518	Down	Apoptosis
Unigene26815_All	Myo-inositol oxidase	3.034	0.001	−11.567	Down	Ascorbate and aldarate metabolism
Unigene28434_All	Putative late blight resistance protein homolog R1A-4	5.890	0.001	−12.524	Down	Plant–pathogen interaction
Unigene3168_All	ARR7 (RESPONSE REGULATOR 7); transcription regulator/two-component response regulator	10.086	0.001	−13.300	Down	Circadian rhythm–plant
Unigene37681_All	Putative cellulose synthase-like protein D5	4.951	0.001	−12.273	Down	Starch and sucrose metabolism

Table 2 continued

Gene ID	Description	3321-RPKM	3321g-1-RPKM	log2Ratio(3321g-1/3321)	Up–down regulation (3321g-1/3321)	Pathway
Unigene42588_All	Three-deoxy-D-manno-octulosonic-acid transferase, N-terminal	3.397	0.001	−11.730	Down	Lipopolysaccharide biosynthesis
Unigene52912_All	17.7 kDa class I small heat-shock protein	7.225	0.001	−12.819	Down	Purine metabolism
Unigene59069_All	DEAH box RNA helicase	9.822	0.001	−13.262	Down	Spliceosome
Unigene61388_All	Malate dehydrogenase precursor	5.736	0.001	−12.486	Down	Citrate cycle (TCA cycle)
Unigene72580_All	TPA: class III peroxidase 55 precursor	9.307	0.001	−13.184	Down	Phenylpropanoid biosynthesis
Unigene7827_All	ATP binding/protein binding	5.619	0.001	−12.456	Down	RNA degradation

3321 indicated sample of 15 h before the flooding, while 3321g-1 indicated sample of 9 h treatment
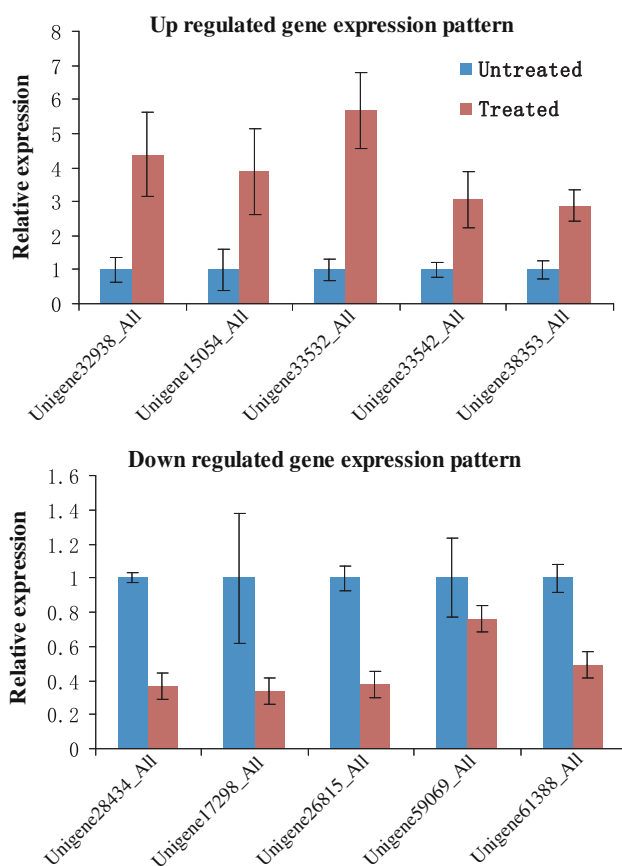
Fig. 3 Real-time PCR validations of the upregulated and downregulated genes characterized by RNA-Seq

Discussion

Global analysis of the sesame transcriptome will facilitate the identification of systemic gene expression and regulatory mechanisms under waterlogging conditions (Lee et al. 2006; Wu et al. 2006). In this study, we performed transcriptome profiling under waterlogging and normal conditions to identify genes that are differentially expressed in this treatment and in the control. Using RNA-Seq deep sequencing (Morrissy et al. 2009), we obtained a direct digital readout of cDNAs and a large range of genes from the libraries. Thus, the present study represents a comprehensive analysis of the sesame transcriptome under waterlogging conditions. Global gene transcription in stressed root tissue was significantly altered, which is in line with the results for *Arabidopsis* hypoxia, and rice and poplar waterlogging (Klok et al. 2002; Liu et al. 2005; Loreti et al. 2005; Lasanthi-Kudahettige et al. 2007; Kreuzwieser et al. 2009). As found in experiments on poplar (Kreuzwieser et al. 2009), waterlogging led to increased gene expression of glycolysis, fermentation, and some catabolic pathways and to decreased expression of synthesis pathways, cell wall activity, and secondary metabolism-associated genes. Gene transcription responses to waterlogging in cotton and poplar resemble that seen in *Arabidopsis* subjected to hypoxic gas mixtures (Klok et al. 2002; Branco-Price et al. 2005; Liu et al. 2005; Loreti et al. 2005), reinforcing that, at least initially, the major factor in waterlogging stress is a lack of oxygen.

In waterlogged sesame plants, many genes with potential roles in carbon and energy metabolism were identified as having a significant transcriptional response to the stress. The most notable examples were genes involved in glycolysis, fermentation, mitochondrial electron transport and starch synthesis pathway (Table 2). Waterlogging also affects nitrogen metabolism in plants. Waterlogging treatments of cotton increased total nitrogen content in aerial portions of the plants and this increase was influenced by soil type (Soomro and Waring 1987). Given the need to maintain balance between carbon and nitrogen levels in the cell and the role of the mitochondrial electron transport chain in this (Quesada et al. 2000; Noctor et al. 2004), it is not surprising that major changes in carbon metabolism caused by waterlogging stress should be accompanied by changes in nitrogen metabolism.

Table 3 Ten pairs of primers designed for each target gene for qRT-PCR

Gene ID	Forward primer (5′–3′)	Reverse primer (5′–3′)
Up-regulated		
Unigene32938_All	GGCTTTCTGAGGCCCAAGCACC	TGGGGGCACCAAATTTGCCACA
Unigene15054_All	GGTGACTGGAACGGTGCAGGA	GCTCCTTGTGGCGAAGCGAGAG
Unigene33532_All	GAGAAGAGGGGAATGGAAGC	CCGGTAACTCCAGCATTGTT
Unigene33542_All	ACACAGCAGCAAGCCAGGTGT	TGGCCAGGTGATGTGCTGCAA
Unigene38353_All	GTGCCCGCAGCGCTTCTCAG	TGCCAGGGCACACAGCTTCC
Down-regulated		
Unigene28434_All	TGCGAACGCTGAAAGCCCCA	AGGAGGCCGCCAATCGCAAC
Unigene17298_All	TCGTGATTAACGGGGTGGGGT	TGCCGACTGAGAACCCTCAAGC
Unigene26815_All	CGACGTTGCCCTCAGCTGGT	ACTCTTCCCATCCCTCACCGCA
Unigene59069_All	TGTGGCAGCACGTGGATTGGA	TGCCAGTTTTCCCTGCTCTCCCT
Unigene61388_All	TTGGACGTGGAGGAGCTGAAGAGG	TCCAATCCGACCCTGTTGATAGCCC

Through systematically exploring sets of differentially expressed genes for high-level biological categories using GO and pathway assignments, we confirmed much of what is currently known for low-oxygen stress responses and, more importantly, obtained new insights into processes that were not previously associated with the anaerobic response. This study provides comprehensive gene expression profiling for hypoxia-perturbed transcriptional networks in plants. This information represents a basis for formulating working models and establishing testable hypotheses to identify the underlying low-oxygen sensing mechanisms, elucidate signal transduction pathways, and further characterize hypoxia-responsive genes. Together with information from independent approaches, such as proteomics, metabolite profiles, and use of reverse genetics resources, our results help elucidate low-oxygen stress responses and will aid in the effort to develop flood-tolerant crops.

Author contribution L. Wang contributed to the experimental design and management, data analysis, and manuscript preparation and revision. Y. Zhang assisted with the experimental design and manuscript organization. X. Qi prepared plant materials for qRT-PCR, data analysis, and reviewed the manuscript. D. Li contributed to tissue collection, DNA extraction. W. Wei participated in tissue collection and RNA isolation. X. Zhang designed and managed the experiments, organized and reviewed the manuscript. All authors have read and approved the final manuscript.

References

Ashri A (2010) Sesame breeding. In: Janick J (ed) Plant breeding reviews, vol 16. Wiley, Oxford. doi:10.1002/9780470650110.ch5

Bailey-Serres J, Voesenek LA (2008) Flooding stress: acclimations and genetic diversity. Annu Rev Plant Biol 59:313–339. doi:10.1146/annurev.arplant.59.032607.092752

Banti V, Mafessoni F, Loreti E, Alpi A, Perata P (2010) The heat-inducible transcription factor HsfA2 enhances anoxia tolerance in Arabidopsis. Plant Physiol 152(3):1471–1483. doi:10.1104/pp.109.149815

Bedigian D, Harlan J (1986) Evidence for cultivation of sesame in the ancient world. Econ Bot 40(2):137–154. doi:10.1007/bf02859136

Benjamini BY, Yekutieli D (2001) The control of the false discovery rate in multiple testing under dependency. Ann Stat 29(4):1165–1188. doi:10.1214/aos/1013699998

Branco-Price C, Kawaguchi R, Ferreira RB, Bailey-Serres J (2005) Genome-wide analysis of transcript abundance and translation in Arabidopsis seedlings subjected to oxygen deprivation. Ann Bot 96(4):647–660. doi:10.1093/aob/mci217

Christianson JA, Llewellyn DJ, Dennis ES, Wilson IW (2010) Global gene expression responses to waterlogging in roots and leaves of cotton (Gossypium hirsutum L.). Plant Cell Physiol 51(1):21–37. doi:10.1093/pcp/pcp163

Drew MC (1997) Oxygen deficiency and root metabolism: injury and acclimation under hypoxia and anoxia. Annu Rev Plant Physiol Plant Mol Biol 48:223–250. doi:10.1146/annurev.arplant.48.1.223

Du H, Zhang Z, Li J (2010) Isolation and functional characterization of a waterlogging-induced promoter from maize. Plant Cell Rep 29(11):1269–1275. doi:10.1007/s00299-010-0913-x

Hegedus Z, Zakrzewska A, Agoston VC, Ordas A, Racz P, Mink M, Spaink HP, Meijer AH (2009) Deep sequencing of the zebrafish transcriptome response to mycobacterium infection. Mol Immunol 46(15):2918–2930. doi:10.1016/j.molimm.2009.07.002

Hinz M, Wilson IW, Yang J, Buerstenbinder K, Llewellyn D, Dennis ES, Sauter M, Dolferus R (2010) Arabidopsis RAP2.2: an ethylene response transcription factor that is important for hypoxia survival. Plant Physiol 153(2):757–772. doi:10.1104/pp.110.155077

Klok EJ, Wilson IW, Wilson D, Chapman SC, Ewing RM, Somerville SC, Peacock WJ, Dolferus R, Dennis ES (2002) Expression

profile analysis of the low-oxygen response in Arabidopsis root cultures. Plant Cell 14(10):2481–2494. doi:10.1105/tpc.004747

Kreuzwieser J, Hauberg J, Howell KA, Carroll A, Rennenberg H, Millar AH, Whelan J (2009) Differential response of gray poplar leaves and roots underpins stress adaptation during hypoxia. Plant Physiol 149(1):461–473. doi:10.1104/pp.108.125989

Lasanthi-Kudahettige R, Magneschi L, Loreti E, Gonzali S, Licausi F, Novi G, Beretta O, Vitulli F, Alpi A, Perata P (2007) Transcript profiling of the anoxic rice coleoptile. Plant Physiol 144(1):218–231. doi:10.1104/pp.106.093997

Lee JJ, Hassan OS, Gao W, Wei NE, Kohel RJ, Chen XY, Payton P, Sze SH, Stelly DM, Chen ZJ (2006) Developmental and gene expression analyses of a cotton naked seed mutant. Planta 223(3):418–432. doi:10.1007/s00425-005-0098-7

Li R, Yu C, Li Y, Lam TW, Yiu SM, Kristiansen K, Wang J (2009) SOAP2: an improved ultrafast tool for short read alignment. Bioinformatics 25(15):1966–1967. doi:10.1093/bioinformatics/btp336

Liu J, Tu L, Xu R, Zheng Y (1993) The relationship between the waterlogging resistance and the genotypes and the vigor of root system in sesame (Sesamum indicum L.). Acta Agric Bor Sin 8(3):82–86 (in Chinese)

Liu F, VanToai T, Moy LP, Bock G, Linford LD, Quackenbush J (2005) Global transcription profiling reveals comprehensive insights into hypoxic response in Arabidopsis. Plant Physiol 137(3):1115–1129. doi:10.1104/pp.104.055475

Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2-$\Delta\Delta$CT method. Methods 25(4):402–408. doi:10.1006/meth.2001.1262

Loreti E, Poggi A, Novi G, Alpi A, Perata P (2005) A genome-wide analysis of the effects of sucrose on gene expression in Arabidopsis seedlings under anoxia. Plant Physiol 137(3):1130–1138. doi:10.1104/pp.104.057299

Morrissy AS, Morin RD, Delaney A, Zeng T, McDonald H, Jones S, Zhao Y, Hirst M, Marra MA (2009) Next-generation tag sequencing for cancer gene expression profiling. Genome Res 19(10):1825–1835. doi:10.1101/gr.094482.109

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. doi:10.1038/nmeth.1226

Noctor G, Dutilleul C, De Paepe R, Foyer CH (2004) Use of mitochondrial electron transport mutants to evaluate the effects of redox state on photosynthesis, stress tolerance and the integration of carbon/nitrogen metabolism. J Exp Bot 55(394):49–57. doi:10.1093/jxb/erh021

Quesada A, Gomez-Garcia I, Fernandez E (2000) Involvement of chloroplast and mitochondria redox valves in nitrate assimilation. Trends Plant Sci 5(11):463–464

Saldanha AJ (2004) Java Treeview-extensible visualization of microarray data. Bioinformatics 20(17):3246–3248. doi:10.1093/bioinformatics/bth349

Snowden RED, Wheeler BD (1993) Iron toxicity to fen plant species. J Ecol 81(1):35–46

Soomro AW, Waring SA (1987) Effect of temporary flooding on cotton growth and nitrogen nutrition in soils with different organic-matter levels. Aust J Agric Res 38(1):91–99. doi:10.1071/AR9870091

Sun J, Zhang XR, Zhang YX, Wang LH, Huang B (2009) Effects of waterlogging on leaf protective enzyme activities and seed yield of sesame at different growth stages. Chin J Appl Environ Biol 15(6):790–795. doi:10.3724/SP.J.1145.2009.00790

Sun J, Zhang XR, Zhang YX, Wang LH, Li DH (2010) Evaluation of yield characteristics and waterlogging tolerance of sesame germplasm with different plant types after waterlogging. J Plant Genet Resour 11(2):139–146. doi:CNKI:SUN:ZWYC.0.2010-02-006

Wang WQ, Zheng YZ, Mei HX, Zhang JS, Zhang FS (2004) Anaerobic inducible expression of xyloglucan endotransglycosylase (XET) gene in sesame and w heat roots. J Agric Biotechol 12(3):258–263 (in Chinese)

Wang LH, Lv ZP, Zhang YX, Li DH, Lv HX, Zhang XY, Zhang XR (2010) Isolating waterlogging tolerance related genes from sesame (Sesamum indicum L.) by suppression subtractive hybridization. Chin J Oil Crop Sci 32(4):485–490 (in Chinese)

Wei WL, Qi XQ, Wang LH, Zhang YX, Hua W, Li DH, Lv HX, Zhang XR (2011) Characterization of the sesame (Sesamum indicum L.) global transcriptome using Illumina paired-end sequencing and development of EST-SSR markers. BMC Genomics 12:451

Wu Y, Machado AC, White RG, Llewellyn DJ, Dennis ES (2006) Expression profiling identifies genes expressed early during lint fibre initiation in cotton. Plant Cell Physiol 47(1):107–127. doi:10.1093/pcp/pci228

Yang E, Maguire T, Yarmush ML, Berthiaume F, Androulakis IP (2007) Bioinformatics analysis of the early inflammatory response in a rat thermal injury model. BMC Bioinform 8:10. doi:10.1186/1471-2105-8-10

Yang CY, Hsu FC, Li JP, Wang NN, Shih MC (2011) The AP2/ERF transcription factor AtERF73/HRE1 modulates ethylene responses during hypoxia in Arabidopsis. Plant Physiol 156(1):202–212. doi:10.1104/pp.111.172486