Published in *Funct Integr Genomics* 2019 Sep;19(5):715-727. doi: 10.1007/s10142-019-00675-2

Comparative RNA-seq analysis of the drought sensitive lentil (*Lens culinaris*) root and leaf under short- and long-term water deficits

Hande Morgil[1¶], Mehmet Tardu[2¶], Gül Cevahir[1], İbrahim Halil Kavakli[2, 3*]

[1]Department of Biology, Istanbul University, 34134 Suleymaniye, Istanbul, Turkey

[2]Department of Chemical and Biological Engineering, Koc University, Rumelifeneri Yolu, Sariyer, Istanbul, Turkey

[3]Department of Molecular Biology and Genetics, Koc University, Rumelifeneri Yolu, Sariyer, Istanbul, Turkey

[¶] these authors are equally contributed

***Corresponding Author**

e-mail:hkavakli@ku.edu.tr

Tel: +90 212 338 1708

Keywords: Lentil, drought stress, roots, leaf, *de novo* RNA-seq

Abstract

Drought stress is one of the main environmental factors that effects growth and productivity of crop plants, including lentil. To gain insights into the genome-wide transcriptional regulation in lentil root and leaf under short- and long-term drought conditions, we performed RNA-seq on a drought sensitive lentil cultivar (*Lens culinaris* Medik. cv. Sultan). After establishing drought conditions, lentil, samples were subjected to *de novo* RNA-seq based transcriptome analysis. The 207,076 gene transcripts were successfully constructed by *de novo* assembly from the sequences obtained from root, leaf, and stems. Differentially expressed gene (DEG) analysis on these transcripts indicated that period of drought stress had a greater impact on the transcriptional regulation in lentil root. The numbers of DEGs were 2,915 under short-term drought stress while the numbers of DEGs were increased to 18,327 under long-term drought stress condition in the root. Further, Gene Ontology analysis revealed that following biological processes were differentially regulated in response to long-term drought stress: protein phosphorylation, embryo development seed dormancy, DNA replication, and maintenance of root meristem identity. Additionally, DEGs, play role in circadian rhythm and photoreception, were down-regulated suggesting that drought stress has a negative effect on the internal oscillators which may have detrimental consequences on plant growth and survival. Collectively, this study provides a detailed comparative transcriptome response of drought sensitive lentil strain under short- and long-term drought conditions in root and leaf. Our finding suggests that not only the regulation of genes in leaves are important but also genes regulated in roots are important and needs to be considered for improving drought tolerance in lentil.

INTRODUCTION

Drought is one of the major environmental factors that inhibit the growth and productivity of crop plants. Studies with different plants showed that drought stress result in up to 50% yield losses (Daryanto et al. 2016;Wang et al. 2003). In the near future, global climate change is likely to gradually increase the intensity of drought which may result in even more yield losses (Battisti and Naylor 2009;Dai 2013). Therefore, there is a need to understand which mechanisms by which crop plants are affected by drought stress, and how they can tolerate it.

Several studies have been conducted on various crop plants to understand the effect of drought stress at physiological, molecular and biochemical levels (Lamaoui et al. 2018). Plants exhibit morphological changes such as enhanced root formation, decreased leaf area, reduced stomata number and its conductance, increased leaf thickness, and leaf rolling to minimize water loss and maximizing the water uptake (Sicher et al. 2012). As drought conditions persist, plant produces wax to cover aerial part of it to minimize dehydration (Lee and Suh 2013). At biochemical level, plants produce osmolytes like proline and glycine betaine to protect various components of cell to minimize the effect of water loss (Lamaoui, et al. 2018). Genomics approaches such as microarray and RNA-seq are used to understand genome-wide regulation in various crop plants at the molecular level, and to relate with physiological and biochemical changes. Studies revealed a complex interaction network between different transcription factors to enable plants to produce the appropriate response. In addition, genomics studies also highlighted the importance

of transcriptional plasticity among different cultivars of the same crop plants (Kulkarni et al. 2017), which adds another layer of complexity and necessities further genome-wide studies.

Lentil is one of the legume crop plants whose productivity is affected by drought stress causing 6-54% yield loss (Johansen et al. 1994;Oweis et al. 2004;Sehgal et al. 2017;Stoddard et al. 2006). Several breeding programs have been initiated to obtain drought tolerant by breeding different genotype of the lentil lines (Singh et al. 2016). Recently, a genome-wide study has been performed on leaf tissues of drought-sensitive and resistance lentil lines to identify the affected biological processes (Singh et al. 2017). Currently, it is still not known how drought stress effects transcriptional regulation in different tissues of lentil under short- and long-term drought-stress conditions. Additionally, it is not known how transcriptional plasticity occurs in lentil depending on its genotype under short- and long-term drought conditions. Therefore, we took a genome-wide approach to understand the effect of drought period on transcriptional response and plasticity using drought-sensitive lentil (*Lens culinaris* Medik. cv. Sultan) root and leaf tissues. After *de novo* transcriptome assembly using RNA-seq reads, differentially expressed gene (DEG) analysis showed a drastic global transcriptional change occur in root when compared with leaf under long-term drought stress. The GO enrichment analysis of the DEGs revealed that genes involved in transcription, regulation of, transcription, response to abscisic acid, protein phosphorylation and protein ubiquitination were commonly regulated in first day stress conditions in both root and leaf. However, as stress prolonged the expression of additional genes playing role in processes of light response, circadian rhythm, and positive regulation of cell death were affected. Additionally, as duration of drought stress increased several genes having role in root development DNA repair, cell cycle DNA replication, sterol biosynthetic process, mismatch

repair, cell death response to salt stress, tyrosine kinase signaling pathways, and mitotic nuclear division were highly regulated in root tissues.

Our study indicated the complexity of lentil plant responses to short- and long-term drought stress in different tissues. These data will help us to understand mechanisms to drought response, may be used to improve productivity and yield of plants under drought stress.

MATERIALS AND METHODS

Determination of Optimum Drought Condition

Seeds of lentil (*Lens culinaris* Medik. cv. Sultan) were kindly provided by the Southeastern Anatolia Agricultural Research Institute of Turkey. The seeds were selected for size homogeneity, surface-sterilized for 2–3 min in 1% sodium hypochlorite, rinsed and soaked in distilled water. Seeds were then placed in 6 cm petri dishes and germinated by watering with 5 ml of dH_2O containing 0% (control), 5%, 20%, and 25% polyethylene glycol (PEG) (MW 6000) concentration corresponding to final osmotic potentials of 0.0, −0.10, −0.14, and −0.19 MPa, respectively. PEG, constitute water stress and lack of penetration to plant cells, was used to adjust osmotic potential and also exhibit no toxic effects to plant cells. The petri dishes were then sealed with parafilm to prevent evaporation and they were kept in a controlled climate chamber at a temperature of 25°C /18°C day/night; photoperiod was 16/8 h light/dark with 10% humidity. Seeds were regarded “germinated” when the radicle had extended for at least 3 mm. After 7 days of treatment, the 10-seeds in the control group (0% PEG) were completely germinated, drought dose were determined for the seedling survivability or medium with different concentrations of PEG. Seedling survivability percent was calculated as follows:

Seedling survivability % =100 x (number of seedlings which survived under the drought stress) / (total number of seedlings) (Idrissi et al. 2016). The data was obtained as three biological replicates with 10 seeds.

Hydroponic Assay and Drought-stress Treatment

Seeds were surface-sterilized for 5 min in 1% sodium hypochlorite and rinsed with dH_2O and then germinated on perlit medium with irrigated everyday by 50 ml Hoagland's nutrient solution (5 mM KNO_3, 5 mM $Ca(NO_3)_2$, 1 mM $MgSO_4$, 5 mM KH_2PO_4, 0.1 mM Fe-EDTA, 46 mM H_3BO_3, 4.5 mM $MnCl_2$, 3.8 mM $ZnSO_4$, 0.3 mM $CuSO_4$, 0.1 mM NH_4MoO_7). Seven days old seedlings were transferred to hydroponic medium (Conn et al. 2013). Control samples were kept in the Hoagland's nutrient solution for the entire hydroponic assay, the drought stress exposed samples were kept in the Hoagland's nutrient solution including 15% PEG 6000, for 5 days. All the experimental pods were aerated through the experiment by air pump and kept in a controlled climate chamber at a temperature of 25°C/18°C day/night; photoperiod was 16/8 h light/dark with 10% humidity.

Physiological Analysis

Samples were harvested at 24^{th} hours (1^{st} day) and 96^{th} hours (4^{th} day) after drought stress treatment in hydroponic medium. Following physiological parameters were determined as described in (Kabbadj et al. 2017): relative water content (RWC), total chlorophyll content, proline content, and lipid peroxidation.

The measurements of the chlorophyll fluorescence were carried out from lentil samples by Imaging-PAM Chlorophyll Fluorometer (Walz Effeltrich, Germany) as previously described in

(Duan et al. 2016). Briefly, sample plants were kept in the dark for 10 min and then exposed to the weak-measuring ray of the fluorescence imaging instrument where the minimal level of fluorescence (F_o) was obtained due to the open state of the photosystem II (PSII) centers. Then samples were exposed to a saturating light pulse with a photosynthetic photon flux density (PPFD) of 960 µmol photons m^{-2} s^{-1} for 0.5 s. Under this conditions PSII centers will go into closed state which resulted in the maximal level of fluorescence (F_m). After high-light exposure the fluorescence level were expected to decrease to a stable state. After that sample leaves were exposed to an actinic PPFD of 260 µmol photons m^{-2} s^{-1} to be able to calculate the transient fluorescence level, F_i, before a peak value of fluorescence, F_p. The following parameters were calculated: F_v/F_m, ϕ_{PSII},ϕ_{NPQ}, ϕ_{NO}, qP and ETR (Baker 2008;Barbagallo et al. 2003;Kramer et al. 2004) .

Preparation of cDNA Libraries for RNA-seq

Total RNA was isolated from 1 gr of root, stem and leaves of lentil grown in hydroponic culture as described above (see section Hydroponic Assay and Drought-stress Treatment), using TRIzol reagent (Invitrogen, USA) as described in (Sarayloo et al. 2017). BioAnalyzer (Agilent, USA) was used to measure the quality and the quantity of the isolated total RNAs. After that, total RNA samples were treated with RNase-free DNase I (Thermo Scientific, USA) at a concentration of 1 U/µg to eliminate the genomic DNA. Then, RNA-seq library preparation was performed using TruSeq Stranded RNA LT Kit (Illumina, USA) as previously described in (Lahens et al. 2014;Tardu et al. 2017;Tardu et al. 2016).The 100x2 paired-end sequencing of cDNA libraries for control and drought stress exposed samples (1st day and 4th day samples of root and leaf, and 1st day sample of stem) with two biological replicates were performed using

the Illumina HiSeq4000 sequencing platform (Illumina, USA).

De novo Transcriptome Assembly and Quality Assessment

The quality of the raw reads were checked using FastQC tool (http://www.bioinformatics.babraham.ac.uk/projects /fastqc/) (v0.11.5) and then Trimmomatic (v0.35) (Bolger et al. 2014) was used to remove the reads with adaptor contamination and low-quality reads (phred score < 20) with the argument -sliding window 4:20. The high quality reads were saved in fastq files and deposited in the GEO database at NCBI with the accession number: GSE115199. All of the downstream analyses were done using the high quality, clean data. *De novo* assembly was performed using Trinity tool (http://trinityrnaseq.sourceforge.net/) with the default parameters (Grabherr et al. 2011). To construct a *de novo* transcriptome assembly for lentil, we used all drought stress exposed samples together with their controls (no-stress samples) at the indicated time points for root and leaf (1^{st} day and 4^{th} day samples), and stem tissues (1^{st} day sample, which its RNA seq data were only used for *de novo* assembly to obtain increase coverage. No further analysis was carried out on 1^{st} day stem samples. After constructing *de novo* transcriptome assembly, the completeness of the assembly was assessed by CEGMA (Parra et al. 2007). Then, full-length transcript counting was done by BLASTX homology search (BLAST+ v2.3.0) with the arguments -evalue 1e-20 -num_threads 6 -max_target_seqs 1 -outfmt 6 against the UniProt database and running perl script analyze_blastPlus_topHit_coverage.pl from the Trinity package.

Differential Expression Analysis

Reads were mapped to the *de novo* transcriptome assembly using Bowtie aligner (Langmead et al. 2009) with default parameters. Then, RSEM method (Li and Dewey 2011) was used to

quantify transcript abundances in all samples. To assess the reproducibility between biological replicates of each sample, a Pearson's correlation analysis was applied to obtain the transcript-level R^2 using RSEM-computed counts. For count-based differential expression testing between root and leaf tissues at different time-points, edgeR (Robinson et al. 2010) was used with the following arguments: min_rowSum_counts of 10, false discovery rate (FDR) < 0.01 and the value of $|\log_2 \text{fold change}| \geq 1$.

Gene Ontology (GO) Assignments and KEGG Pathway annotations of the Transcripts in the Assembly

Annocript tool (Musacchia et al. 2015) was used to annotate trinity assembled *de novo* transcripts (https://github.com/frankMusacchia/Annocript/tree/master/GUIDE). We used the Swiss-Prot (SP) and UniRef90 (version: February 2016) for the blastp against proteins with the following arguments: word_size = 4; e-value = 10^{-5}; num_descriptions = 5; num_alignments = 5; threshold=18. For each sequence, the best hit was chosen. Rpsblast parameters, to identify domain composition of putative proteins in the Conserved Domains Database (CDD) (Marchler-Bauer et al. 2015), were: e-value = 10^{-5}; num_descriptions = 20; num_alignments = 20.

For gene ontology assignments, annotated *de novo* transcripts were analyzed using Database for Annotation, Visualization and Integrated Discovery (DAVID, https://david.ncifcrf.gov/) (Dennis et al. 2003) with default parameters. The statistical significance is determined using the Benjamini correction; adjusted $p < 0.05$ was considered significant. In order to obtain KEGG Pathway annotations of the Transcripts in the Assembly**,** all the *de novo* assembled transcripts were aligned with GENES (http://www.genome.jp/kegg/genes.html) using BLAST algorithm (Blastx/Blastp 2.2.29+) based on KEGG database (http://www.genome.jp/kegg/).

Validation of RNA-seq Results Using qRT-PCR

To validate the RNA-seq results, qRT-PCR was performed for 12 representative DEGs. In briefly, total RNA was extracted using TRIzol reagent (Invitrogen, USA) followed by DNase I treatment. Then, 1 µg of total RNA was used for first-strand cDNA synthesis using random hexamer oligos. qRT-PCR was performed with gene-specific primers shown in **Table S1** using Luminaris HiGreen qRT-PCR Master Mix (Thermo Scientific, USA) on the CFX Connect Real-Time PCR Detection System (Biorad, USA). *β-tubulin* gene whose expression level didn't change across samples was used as an internal reference gene. The amplification programs were performed according to the following protocol: 95 ºC for 10 min; 40 cycles of 95 ºC for 15s, 58 ºC for 30s and 72 ºC for 30s, and followed by a thermal denaturing step to generate the melting curves. All reactions were performed in biological triplicates (each triplicate with two technical replicates), and the results were expressed relative to the transcription level of *β-tubulin* gene in each sample by using the $2^{-\Delta\Delta CT}$ method (Livak and Schmittgen 2001). The mRNA expression data analyzed by using GraphPad Prism version 7 (GraphPad software, USA). All data were presented as the means ± S.D. (n = 6) of the relative mRNA expression.

RESULTS AND DISCUSSION

Seedling Response to Drought-stress and Physiological Analysis

In this study, a drought-sensitive lentil cultivar (*Lens culinaris* Medik. cv. Sultan) was used to get better insight into the genome-wide transcriptional regulation of root and leaf tissues under short- and long-term drought conditions. For this purpose, lentil seeds were germinated at

different concentrations of PEG to determine the optimum drought concentration by calculating the survival rate (**Figure 1a**). Seeds were grown in the presence of the 15% PEG had survival rate of 60% compared to control ($p<0.0001$). Then, leaf samples were harvested after 24 hours (1st day; referred as short-term) and 96 hours (4th day; referred as long-term) and were subjected to the biochemical and physiological analysis to determine the level of drought stress. The relative water content (RWC), total chlorophyll (*Chl*), proline, and malonyldialdehyde (MDA) contents were measured as drought stress indicators. The RWC values of the samples taken from the 1st day ($p<0.01$ versus control) and 4th day ($p<0.005$ versus control) were significantly lower than control (**Figure 1b**). Expectedly, the RWC value of the long-term drought exposed seedlings were 35% lower than the seedlings exposed to short-term drought stress. The analysis of the *Chl* content of the samples indicated that there was significantly low *Chl* content in samples under short- ($p<0.01$ versus control) and long-terms ($p<0.005$ versus control) drought stress (**Figure 1c**). Another indicator of the drought stress is the lipid peroxidation level, which is determined by the level of the MDA content. MDA levels were significantly higher in stressed samples compared to control (**Figure 1d**). Finally, the amount of proline was measured in the samples, which are expected to increase under drought stress. As seen in **Figure 1e**, proline contents were significantly increased in both 1st day ($p<0.005$ versus control) and 4th day ($p<0.01$ versus control) drought-stress samples. All these results showed that drought stress had great impact on lentil.

Chlorophyll Fluorescence Imaging Analysis

The efficiency of photosystem II (F_v/F_m), operating efficiency of photosystem II (Φ_{PSII}), photochemical quenching (qP), and rate of photosynthetic electron transport (ETR) are expected to decrease while the efficiency of non-photochemical quenching (ϕ_{NPQ}) and non-regulated

energy dissipation (ϕ_{NO}) are expected to increase in plants under drought stress (Yao et al. 2018). Therefore, we measured all these parameters to ensure that the seedlings were under stress conditions. As expected, F_v/F_m, $\phi_{PSII,}$ and qP values were reduced under the drought conditions (**Table 1**). The ϕ_{NPQ} did not change under short-term drought condition whereas its level increased significantly under long-term drought condition ($p<0.05$). We estimated the fraction of open PSII reaction centers by measuring with modulated fluorescence qP. Results showed that long-term drought stress caused a significant decrease (~40%, $p< 0.01$) in the number of open PSII reaction centers when compared to the short-term drought stress (**Table 1**). Other photosynthetic parameters, ETR and ϕ_{NO}, were also monitored. The ETR decreased during the continued drought and showed a significant decline at the end of the 4th day ($p< 0.01$). We also observed increased ϕ_{NO} levels after long-term drought stress which indicates an increase the plant's inability to cope with radiation (**Table 1**).

Collectively, all physiological data suggested that the lentil seedlings were in drought conditions and RNA-seq study was conducted using leaves and root samples for understanding the effect of short- and long drought stress in lentil at transcrit level.

Sequencing and *De novo* Assembly Statistics

After collecting samples from seedlings of lentil root, stem and leaf, grown in hydroponic conditions, total RNAs extraction was followed by mRNA isolation. Then, RNA-seq libraries were prepared and sequenced using Illumina platform.

Sequencing resulted in ~60 G bp raw output. Before *de novo* transcriptome assembly of the samples with Trinity tool (http://trinityrnaseq.sourceforge.net/) (Zhao et al. 2011), raw data were subjected to analysis to eliminate low quality reads (phred score <20) and short reads with length <20 nucleotide. After filtering, a total of 568,571,434 paired-end high quality clean reads with an average length of 64 bp were obtained. *De novo* assembly generated 207,076 transcripts with a mean length of 950 nt. At the gene level, assembly resulted in 124,151 genes with N50 length of 1,638 nt (**Table 2**). Length distribution analyses of the transcripts revealed that 94,449 transcripts (45.61%) have the length between 500 to 2,500 nts. Additionally, the shortest transcript has the 201 nts whereas the longest transcript has 12,367 nts. When compared with other published data (Singh, et al. 2017;Sudheesh et al. 2016), the number of transcripts in our *de novo* transcriptome assembly was significantly higher possibly due to the usage of different *de novo* assemblers (Xie et al. 2014). Additionally, using RNA-seq data from different lentil tissues including root, stem and leaf both under short-term and long-term drought stress conditions may enable us to get a more complete *de novo* transcriptome assembly with a high number of transcripts.

Evaluation of Quality and Accuracy of *L. culinaris de novo* Transcriptome Assembly

CEGMA tool (Parra et al. 2007) was used to determine quality and correctness of *de novo* assembled transcriptome. This analysis revealed that the 246 out of 248 core eukaryotic proteins (corresponds to 99.2% coverage) were completely covered in the assembled transcriptome. Then, all reads were mapped back to assembled transcriptome using Bowtie (Langmead et al. 2009) to obtain mapping percentage of each RNA-seq sample. The 92-96% of all reads were successfully mapped back to the assembled transcriptome, which indicates correctness of the assembly and

enabled us to perform further analysis. Additionally, full-length transcript analysis was performed to evaluate the quality of our transcriptome assembly. Results indicated that top matching proteins, which had at least 80% of their sequences included in the assembled transcripts, composed of 54% (9,394/17,574) of all matched proteins (**Table S2**).

To validate RNA-seq results, 12 DEGs from leaf and root samples, subjected to short- and long-drought stress, were randomly selected and their transcription levels were measured by qRT-PCR. Both qRT-PCR and RNA-seq results exhibited similar transcriptional level for each gene, indicating that RNA-seq results were reliable and that the results obtained using both methods were comparable for all conditions (**Figure S1**).

Functional Annotation of Trinity Assembled Transcripts

Trinity tool was used to calculate the number of possible open reading frames (ORFs) of assembled transcripts (Grabherr, et al. 2011). Our results revealed that 72.8% of assembled transcripts (150,680 out of 207,076) possess ORFs. Annocript pipeline was used to annotate these ORFs and to determine putative long non-coding RNAs (lncRNA) of the assembled transcripts (Musacchia, et al. 2015). This analysis identified 96,549 and 138,111 hits in the Swiss-prot and UniRef databases, respectively. Additionally, the 8730 possible lncRNAs ($p<0.05$ and maximum length of ORF: 100) were predicted. Identified transcripts of *L. culinaris* by *de novo* assembly displayed a significant similarity with those in *Medicago truncatula* (34.6%), followed by those in *Cicer arietinum* (18.3%) and *Papilionoideae* (12.9%) (**Figure 2 and Table S3**). Our result is consistent with previously published results where they have shown high similarity between lentil and *C.arietinum* and *M. truncatula* (Singh, et al. 2017;Sudheesh, et

al. 2016). Our result is different than the others where our analysis revealed that *L. culinaris* Medik. cv Sultan has the third highest homology with *Papilionoideae* (12.9%).

Gene Ontology Annotation and Kyoto Encyclopedia of Genes and Genomes Pathway Mapping

In order to understand the biological significance of the transcripts in the transcriptome assembly, gene ontology (GO) term enrichment analysis was performed with DAVID tool (Dennis, et al. 2003). GO term assignments were used to classify *de novo* assembled transcripts into three categories: biological process (BP), molecular function (MF) and cellular component (CC). GO analysis for the 87,298 of *L. culinaris* transcripts (42.2%) were performed, and GO assignments were plotted by WEGO (http://wego.genomics.org.cn/) (**Figure 3**). CC category contains 14 functional classes, classes for cell and cell part accounted for the major proportion with 54,248 (62.1%) and 53,935 (61.8%) transcripts, respectively. In the MF assignments, classes for catalytic activity with 44,114 transcripts (50.5%) and binding with 44,003 transcripts (50.3%) were mostly enriched. In the BP category, classes for cellular process and metabolic process possess highest number of transcripts, with cellular process 41,589 (47.6%) and metabolic process 35,237 (40.4%) transcripts, respectively. In addition, 11,692 (13.4%) transcripts were classified under pigmentation and 9,824 (11.3%) transcripts were classified under response to stimulus in the GO assignments.

To fully uncover the biological pathways that *L. culinaris* possess, all the assembled sequences were assigned with Kyoto Encyclopedia of Genes and Genomes **(**KEGG) orthology (KO)

identifiers using KAAS (http://www.genome.jp/tools/kaas/) with the single-directional best hit method, and subsequently mapped to pathways using the KEGG API (http://www.kegg.jp/kegg/rest/keggapi.html). As a result, 25,829 transcripts were assigned to 399 different KEGG pathways (**Table S4**). The pathways with the highest number of unique transcripts were metabolic pathways (848 members), biosynthesis of secondary metabolites (389 members), biosynthesis of antibiotics (213 members), microbial metabolism in diverse environments (168 members), ribosome (127 members), carbon metabolism (107 members), biosynthesis of amino acids (107 members) and oxidative phosphorylation (104 members) (**Figure 4**).

Identification of Differentially Expressed Genes (DEGs) in Leaf and Root under Drought-Stress

To identify the transcriptional changes at genome-wide level under short (24 hours, 1st day) and long-term drought (96 hours, 4th day) conditions in leaf and root tissues, DEG analyses have been performed. We, initially, carried out Pearson's correlation test to evaluate the reproducibility among biological replicates using the expression values (counts per million). After observing a strong correlation ($R^2 > 0.95$) between the biological replicates of each sample (**Figure S2**), edgeR tool was applied to identify the differentially expressed genes (DEGs) under different drought conditions in both tissues. All identified DEGs were given in the **Table S5**. In the 1st day of drought condition, the transcriptional regulations were significantly altered in leaf tissues when compared with root tissues. The numbers of the DEGs were 6,949 in leaf while the numbers of the DEGs were 2,915 in the root (**Figure 5a**). Among these 2,725 and 1,745 genes

were up-regulated whereas 4,224 and 1,170 genes were down-regulated in leaf and roots, respectively (**Figure 5a**). A similar analysis was carried on the RNA-seq data exposed to long-term drought conditions. The numbers of DEGs in root were 18,327, which were remarkably high compared with number of DEGs under short-term drought condition (**Figure 5A**). On the other hand, the numbers of the DEGs in leaf were 8,306, which were slightly higher compared with the number of DEGs identified under short-term drought stress. These data suggested that long-term drought stress in lentil had great impact on the gene regulation in root of lentil. Additionally, the comparison of numbers of DEGs in leaf and root samples (6,949 DEGs vs 2,915 DEGs) under short-term drought stress showed that leaf adapt to the drought stress more quickly than the root tissues in *L. culinaris* and produces a more rapid response at the transcriptional level against to water loss. Our results are consistent with the results of a previously reported study, where it was shown that there were twice the number of DEGs in the roots than in the leaves in peach, subjected to drought stress for 16 days (Ksouri et al. 2016). In addition, another study performed on a drought tolerance sunflower where plants were exposed to 24 h of drought stress in the presences of the 15% PEG 6000 revealed that the number of DEGs in leaf tissue is significantly higher than the root tissue after 24 hours drought condition (Liang et al. 2017).

Comparison of DEGs between Root and Leaf under Short- and Long-Terms Drought Conditions

After the identification of DEGs for each sample, we determined common and sample specific DEGs in both tissues under two drought conditions (**Figure 5b**). The analysis showed that 195

DEGs were commonly regulated in all condition and tissues. We performed a GO term enrichment analysis for 195 DEGs and identified significantly enriched biological processes and molecular function assignments (**Figure 6a-c, Table S6**). The most effected biological processes were regulation of transcription (19 DEGs), DNA-templated transcription (18 DEGs), response to abscisic acid (9 DEGs), and response to water deprivation (7 DEGs) (**Figure 6a, Table S6**). The root is the primary organ that responds at early stages of water deficit to mediated by abscisic acid (ABA) (Jiang and Hartung 2008). This may produce an immediate response through root to initiate ABA signaling which may result in root-to-shoot signaling and in the partial or complete stomatal closure to reduce transpiration. Therefore, ABA signaling pathway is heavily regulated throughout all drought conditions. At gene level, for example, a transcription factor called MYB108 was shown to be act as a negative regulator of ABA-induced cell death under biotic and abiotic stress in Arabidopsis (Mengiste et al. 2003). Another example, a receptor-like kinase FERONIA which plays a crucial role in plant response to small molecule hormones (e.g., auxin and ABA) (Chen et al. 2016). In addition, the most affected molecular function and cellular components were protein binding, transcription factor activity and plasma membrane, respectively (**Figure 6a, Table S6**). In fact, analysis of the most effected genes in each category indicated their role in drought stress specifically in ABA signaling in various studies (Shinozaki and Yamaguchi-Shinozaki 2007).

We then aimed to identify specific biological processes mediated by DEGs in root and leaf tissues under short- and long-term drought conditions. GO term enrichment analysis revealed 1,159 DEGs in root under both drought conditions (**Figure 5b**). The major significantly (p<0.05) up-regulated biological processes were transcription, regulation of transcription, and protein

ubiquitination while major significantly (p<0.05) down-regulated biological processes were protein phosphorylation, transport, and tyrosine kinase signaling pathway under both short- and long term drought conditions in root (**Figure 7a, Table S6**). On the other hand, 1,699 DEGs were specifically expressed under short-term drought condition while 17,148 DEGs were expressed under the long-term drought condition (**Figure 5b**). It is surprising to see very high number of the DEGs in the 4th day of the drought condition. To see how all these DEGs were distributed among the biological processes we again performed GO term enrichment analysis. As can be seen in **Figure 7a**, some of the up- and down-regulated processes (such as protein transport, response to ABA, RNA splicing, flower development, nitrate transport, sterol biosynthesis) in the 1st day of drought condition were disappeared and instead the new biological process were seems to be up- and down-regulated in the 4th day of drought condition (such as protein phosphorylation, mismatch repair, development of the seed dormancy, cell division, and DNA replication maintenance of the DNA repeat element). A very similar trend was observed in comparison of the DEGs between 1st day and 4th day leaf samples. The majority of the DEGs and biological processes (transcription, regulation of transcription, activated signaling pathway, and response to water deprivation) were common under both short- and long-term drought conditions (**Figure 7a, Table S6**). Additionally, some of the pathways were specifically regulated under long-term drought conditions. The following biological processes were up-regulated: response to cold, response to salt stress and ubiquitin-dependent protein degradation while these biological processes were down-regulated: chloroplast organization, circadian rhythm, positive regulation of cell death and cellular response to light stimulus (**Figure 7a, Table S6**). We also compared our results with a previously published result by Singh et al. (2017). Although our and their studies were carried out under different conditions, the genes that play role in the following

processes were seems to be commonly regulated in the leaf: Response to abscisic acid DNA-templated transcription, plant type cell wall, cell wall organization, chloroplast. The other affected processes were different from each other in both stdies. All these results suggested that different sets of gene were being regulated under different drought conditions, which may reflect transcriptional plasticity or experimental differences that performed under different conditions. Additionally, as drought stress persists the transcription of the genes are heavily regulated in roots when compared with leaf.

Previous studies highlight the importance of the circadian rhythm in drought stress (Legnaioli et al. 2009). Our results are consistent with previous studies where we showed circadian rhythms in both tissues are being most effective under either condition in the lentil. For example, one of the core clock proteins named as CRYPTOCHROMEs are a part of core clock protein in mammals (Kavakli et al. 2017) and function as the photoreceptors that sense environmental cues, such as irradiance, day-night transition, photoperiods, and light quality for optimal growth and development in plant (Liu et al. 2011). Also, Lysine-specific demethylase JMJ30, involved in the control of flowering time by demethylating H3K36me2 at the FT locus and repressing its expression, acts within the central clock (Jones et al. 2010). In addition, a number of genes encodes circadian-associated factors have been regulated such as APRR7, REVEILLE 1 FBX2c, GIGANTEA, TOC2, and Photoropin-2 (Adams et al. 2015;Cha et al. 2017;Lu et al. 2017).

CONCLUSION

The current study reports the transcriptome analysis of leaves and roots related to short- and long-terms drought stress at the seedling stage in lentil using next generation sequencing approach. Due to lack of a reference genome for lentil, we first constructed a transcriptome

assembly and then transcripts were annotated and mapped to the KEGG pathways. At the beginning of the drought stress, root and leaf produces very initial response by regulating the transcription of the several genes required to cope with drought stress. The persistence of the stress causes a huge transcriptional regulation at genome-wide level in the roots of the lentil. The additional 11, 201 genes being up-regulated possibly regulate processes to minimize the effect of the stress in the root. All these results suggest that genes specifically express in roots which can be further utilized and serve as new resources for future genetic and functional genomics studies for drought tolerance in lentil.

AUTHOR CONTRIBUTIONS

HM and MT carried out experiments. MT performed analysis. MT, HM and GC contribute writing the manuscript. IHK designed experiments and writing the manuscript.

FUNDING

This work was supported by an Istanbul Development Agency grant (ISTKA-TR/14/EVK/0039) to IHK, by a Istanbul University BAP project (21207, 22583) to GC and IHK.

ACKNOWLEDGEMENTS

We would like to thanks Southeastern Anatolia Agricultural Research Institute providing us seed of *Lens culinaris* Medik. cv. Sultan.

References

Adams S, Manfield I, Stockley P, Carre IA (2015) Revised morning loops of the arabidopsis circadian clock based on analyses of direct regulatory interactions. Plos One 10.
Baker NR (2008) Chlorophyll fluorescence: a probe of photosynthesis in vivo. Annu Rev Plant Biol 59:89-113.
Barbagallo RP, Oxborough K, Pallett KE, Baker NR (2003) Rapid, noninvasive screening for perturbations of metabolism and plant growth using chlorophyll fluorescence imaging. Plant Physiol 132:485-493.
Battisti DS, Naylor RL (2009) Historical warnings of future food insecurity with unprecedented seasonal heat. Science 323:240-244.
Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114-2120.
Cha JY, Kim J, Kim TS, Zeng Q, Wang L, Lee SY, Kim WY, Somers DE (2017) GIGANTEA is a co-chaperone which facilitates maturation of ZEITLUPE in the Arabidopsis circadian clock. Nat Commun 8.
Chen J, Yu F, Liu Y, Du CQ, Li XS, Zhu SR, Wang XC, Lan WZ, Rodriguez PL, Liu XM, Li DP, Chen LB, Luan S (2016) FERONIA interacts with ABI2-type phosphatases to facilitate signaling cross-talk between abscisic acid and RALF peptide in Arabidopsis. P Natl Acad Sci USA 113:E5519-E5527.
Conn SJ, Hocking B, Dayod M, Xu B, Athman A, Henderson S, Aukett L, Conn V, Shearer MK, Fuentes S, Tyerman SD, Gilliham M (2013) Protocol: optimising hydroponic growth systems for nutritional and physiological analysis of Arabidopsis thaliana and other plants. Plant Methods 9:4.
Dai AG (2013) Increasing drought under global warming in observations and models. Nat Clim Change 3:52-58.
Daryanto S, Wang LX, Jacinthe PA (2016) Global synthesis of drought effects on maize and wheat production. Plos One 11.
Dennis G, Jr., Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, Lempicki RA (2003) DAVID: Database for annotation, visualization, and integrated discovery. Genome Biol 4:P3.
Duan Z, Kong F, Zhang L, Li W, Zhang J, Peng L (2016) A bestrophin-like protein modulates the proton motive force across the thylakoid membrane in Arabidopsis. J Integr Plant Biol 58:848-858.
Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng QD, Chen ZH, Mauceli E, Hacohen N, Gnirke A, Rhind N, di Palma F, Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N, Regev A (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol 29:644-U130.
Idrissi O, Udupa SM, De Keyser E, McGee RJ, Coyne CJ, Saha GC, Muehlbauer FJ, Van Damme P, De Riek J (2016) Identification of quantitative trait loci controlling root and shoot traits associated with drought tolerance in a lentil (*Lens culinaris* Medik.) recombinant inbred line population. Front Plant Sci 7:1174.
Jiang F, Hartung W (2008) Long-distance signalling of abscisic acid (ABA): the factors regulating the intensity of the ABA signal. J Exp Bot 59:37-43.
Johansen C, Baldev B, Brouwer JB, Erskine W, Jermyn WA, Lijuan L, Malik BA, Miah AA, Silim SN (1994) Biotic and abiotic stress constraining productivity of cool-season food legumes in Asia, Africa and Oceania. Curr Plant Sci Biot 19:175-194.
Jones MA, Covington MF, DiTacchio L, Vollmers C, Panda S, Harmer SL (2010) Jumonji domain protein JMJD5 functions in both the plant and human circadian systems. P Natl Acad Sci USA 107:21623-21628.

Kabbadj A, Makoudi B, Mouradi M, Pauly N, Frendo P, Ghoulam C (2017) Physiological and biochemical responses involved in water deficit tolerance of nitrogen-fixing *Vicia faba*. PLoS One 12:e0190284.
Kavakli IH, Baris I, Tardu M, Gul S, Oner H, Cal S, Bulut S, Yarparvar D, Berkel C, Ustaoglu P, Aydin C (2017) The Photolyase/Cryptochrome family of proteins as DNA repair enzymes and transcriptional repressors. Photochem Photobiol 93:93-103.
Kramer DM, Johnson G, Kiirats O, Edwards GE (2004) New Fluorescence Parameters for the Determination of QA Redox State and Excitation Energy Fluxes. Photosynth Res 79:209.
Ksouri N, Jimenez S, Wells CE, Contreras-Moreira B, Gogorcena Y (2016) Transcriptional responses in root and leaf of *Prunus persica* under drought stress using RNA sequencing. Front Plant Sci 7:1715.
Kulkarni M, Soolanayakanahally R, Ogawa S, Uga Y, Selvaraj MG, Kagale S (2017) Drought Response in Wheat: Key Genes and Regulatory Mechanisms Controlling Root System Architecture and Transpiration Efficiency. Front Chem 5:106.
Lahens NF, Kavakli IH, Zhang R, Hayer K, Black MB, Dueck H, Pizarro A, Kim J, Irizarry R, Thomas RS, Grant GR, Hogenesch JB (2014) IVT-seq reveals extreme bias in RNA sequencing. Genome Biol 15:R86.
Lamaoui M, Jemo M, Datla R, Bekkaoui F (2018) Heat and drought stresses in crops and approaches for their mitigation. Front Chem 6.
Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol 10:R25.
Lee SB, Suh MC (2013) Recent advances in cuticular wax biosynthesis and its regulation in Arabidopsis. Mol Plant 6:246-249.
Legnaioli T, Cuevas J, Mas P (2009) TOC1 functions as a molecular switch connecting the circadian clock with plant responses to drought. EMBO J 28:3745-3757.
Li B, Dewey CN (2011) RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics 12.
Liang C, Wang W, Wang J, Ma J, Li C, Zhou F, Zhang S, Yu Y, Zhang L, Li W, Huang X (2017) Identification of differentially expressed genes in sunflower (*Helianthus annuus*) leaves and roots under drought stress by RNA sequencing. Bot Stud 58:42.
Liu H, Liu B, Zhao C, Pepper M, Lin C (2011) The action mechanisms of plant cryptochromes. Trends Plant Sci 16:684-691.
Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(T)(-Delta Delta C) method. Methods 25:402-408.
Lu H, McClung CR, Zhang C (2017) Tick tock: Circadian regulation of plant innate immunity. Annu Rev Phytopathol 55:287-311.
Marchler-Bauer A, Derbyshire MK, Gonzales NR, Lu S, Chitsaz F, Geer LY, Geer RC, He J, Gwadz M, Hurwitz DI, Lanczycki CJ, Lu F, Marchler GH, Song JS, Thanki N, Wang Z, Yamashita RA, Zhang D, Zheng C, Bryant SH (2015) CDD: NCBI's conserved domain database. Nucleic Acids Res 43:D222-226.
Mengiste T, Chen X, Salmeron J, Dietrich R (2003) The BOTRYTIS SUSCEPTIBLE1 gene encodes an R2R3MYB transcription factor protein that is required for biotic and abiotic stress responses in Arabidopsis. Plant Cell 15:2551-2565.
Musacchia F, Basu S, Petrosino G, Salvemini M, Sanges R (2015) Annocript: a flexible pipeline for the annotation of transcriptomes able to identify putative long noncoding RNAs. Bioinformatics 31:2199-2201.
Oweis T, Hachum A, Pala M (2004) Lentil production under supplemental irrigation in a Mediterranean environment. Agr Water Manage 68:251-265.

Parra G, Bradnam K, Korf I (2007) CEGMA: a pipeline to accurately annotate core genes in eukaryotic genomes. Bioinformatics 23:1061-1067.
Robinson MD, McCarthy DJ, Smyth GK (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26:139-140.
Sarayloo E, Tardu M, Unlu YS, Simsek S, Cevahir G, Erkey C, Kavakli IH (2017) Understanding lipid metabolism in high-lipid-producing *Chlorella vulgaris* mutants at the genome-wide level. Algal Res 28:244-252.
Sehgal A, Sita K, Kumar J, Kumar S, Singh S, Siddique KHM, Nayyar H (2017) Effects of drought, heat and their interaction on the growth, yield and photosynthetic function of lentil (*Lens culinaris* Medikus) genotypes varying in heat and drought sensitivity. Front Plant Sci 8.
Shinozaki K, Yamaguchi-Shinozaki K (2007) Gene networks involved in drought stress response and tolerance. J Exp Bot 58:221-227.
Sicher RC, Timlin D, Bailey B (2012) Responses of growth and primary metabolism of water-stressed barley roots to rehydration. J Plant Physiol 169:686-695.
Singh D, Singh CK, Taunk J, Tomar RS, Chaturvedi AK, Gaikwad K, Pal M (2017) Transcriptome analysis of lentil (*Lens culinaris* Medikus) in response to seedling drought stress. BMC Genomics 18:206.
Singh D, Singh CK, Tomar RSS, Taunk J, Singh R, Maurya S, Chaturvedi AK, Pal M, Singh R, Dubey SK (2016) Molecular Assortment of Lens Species with Different Adaptations to Drought Conditions Using SSR Markers. Plos One 11.
Stoddard FL, Balko C, Erskine W, Khan HR, Link W, Sarker A (2006) Screening techniques and sources of resistance to abiotic stresses in cool-season food legumes. Euphytica 147:167-186.
Sudheesh S, Verma P, Forster JW, Cogan NO, Kaur S (2016) Generation and Characterisation of a Reference Transcriptome for Lentil (*Lens culinaris* Medik.). Int J Mol Sci 17.
Tardu M, Bulut S, Kavakli IH (2017) MerR and ChrR mediate blue light induced photo-oxidative stress response at the transcriptional level in *Vibrio cholerae*. Sci Rep 7:40817.
Tardu M, Dikbas UM, Baris I, Kavakli IH (2016) RNA-seq analysis of the transcriptional response to blue and red light in the extremophilic red alga, *Cyanidioschyzon merolae*. Funct Integr Genomic 16:657-669.
Wang W, Vinocur B, Altman A (2003) Plant responses to drought, salinity and extreme temperatures: towards genetic engineering for stress tolerance. Planta 218:1-14.
Xie Y, Wu G, Tang J, Luo R, Patterson J, Liu S, Huang W, He G, Gu S, Li S, Zhou X, Lam TW, Li Y, Xu X, Wong GK, Wang J (2014) SOAPdenovo-Trans: de novo transcriptome assembly with short RNA-Seq reads. Bioinformatics 30:1660-1666.
Yao JN, Sun DW, Cen HY, Xu HX, Weng HY, Yuan F, He Y (2018) Phenotyping of Arabidopsis drought stress response using kinetic chlorophyll fluorescence and multicolor fluorescence imaging. Front Plant Sci 9: 603
Zhao QY, Wang Y, Kong YM, Luo D, Li X, Hao P (2011) Optimizing de novo transcriptome assembly from short-read RNA-Seq data: a comparative study. BMC Bioinformatics 12 Suppl 14:S2.

Table 1. Effect of drought stress (DS) on the chlorophyll fluorescence parameters (F_v/F_m) and changes in the balance between light capture and energy use in *L. culinaris* seedlings.

Sample	F_v/F_m	Φ_{PSII}	Φ_{NPQ}	Φ_{NO}	qP	ETR
1st day control	0,81±0,01	0,42±0,02	0,22±0,02	0,26±0,01	0,75±0,01	57,58±2,10
1st day DS	0,79±0,01*	0,39±0,01	0,24±0,04	0,34±0,09	0,56±0,02	49,62±1,15
4th day control	0,82±0,00	0,41±0,01	0,23±0,04	0,27±0,00	0,79±0,05	51,37±2,35
4th day DS	0,58±0,03**	0,19±0,02*	0,32±0,02*	0,57±0,04**	0,31±0,04**	25±1,85**

All measurements are represented as mean values ± standard deviation of three biological replicates. Tukey test was performed to test for significant differences. An asterisk represents significantly different means of the same parameter between control and drought-stress samples. (*$p < 0.05$), (**$p < 0.01$).

Table 2. Sequencing and assembly statistics of *Lens culinaris* Medik. cv. Sultan transcriptome.

Item	Value
Total number of clean reads after filtering	568,571,434
Mean length of trimmed read pairs (nt)	64
Total number of assembled transcripts	207,076
Assembled transcripts size in Mbp	196,5
Mean transcript length (nt)	950
Median contig length (nt)	550
N50 value of transcripts (nt)	1,638
Longest transcript length (nt)	12,367
Total number of genes	124,151

Figure Legends

Figure 1. (**A**) *L. culinaris* seedling survivability percentage under different drought-stress conditions (0% (control), 15%, 20% and 25% PEG). Errors bars represent standard deviation (SD) (n=30). Physiological analysis of *L. culinaris* leaves with respect to changes in (**B**) relative water content, (**C**) total chlorophyll (*Chl*) content, (**D**) malondialdehyde content, and (**E**) proline content in control (no-stress) and drought-stress (+) samples. Assays were done after 24 hours (1^{th} day) and 96 hours (4^{th} day) of stress treatment. Means, standard errors of three biological and three technical replicates. The data were analyzed using a two-way ANOVA analysis of variance (Brown-Forsythe Test). Asterisks indicate significant differences at *p<0.05, **p<0.01, ***p<0.005 versus control samples.

Figure 2: UniProt Top-Hit species distribution of transcripts.

Figure 3: Gene ontology (GO) functional classification of transcripts in *L. culinaris.* The visualization of GO distribution was done with WEGO tool (http://wego. genomics.org.cn/) using GO level 2 descriptions for three main categories (biological process, molecular function, cellular component).

Figure 4: The top 10 assigned KEGG pathways with the highest transcript numbers. Numbers in the pie-chart represent the number of transcripts in the assigned KEGG pathways.

Figure 5: (A) Number of differentially expressed genes (DEGs) under drought-stress in leaf and root samples of *L. culinaris* at 1^{st} and 4^{th} days. **(B**) Venn diagram shows the number of DEGs across leaf and root tissues after 1^{st} and 4^{th} days of drought-stress and the overlap between each set of genes.

Figure 6: Significantly enriched (p<0.05) GO term analysis of common 195 DEGs; **(A)** biological processes, **(B)** molecular functions and **(C)** cellular compartments.

Figure 7: Significantly enriched gene ontology (GO) biological processes (BP) in leaf and root tissues of *L. culinaris*. Each panel shows significantly enriched GO biological processes in **(A)** 1^{st} day and 4^{th} day samples in leaf, and **(B)** 1^{st} day and 4^{th} day samples of root. Asterisks indicate significantly regulated BP terms both in 1^{st} day and 4^{th} day samples. Number of enriched genes for each BP terms are represented both with blue and red colors. Down-regulated genes were represented with blue color whereas up-regulated genes were represented with red color.

SUPPLEMENTARY MATERIAL

Table S1: qRT-PCR forward and reverse primer sequences.

Table S2: Transcripts were aligned to Uniprot protein database using BlastX

Table S3: Species distribution *Medicago truncatula*

Table S4: KEGG pathways

Table S5: Identified DEGs in all samples.

Table S6: Significantly enriched gene ontology (GO) biological process terms.

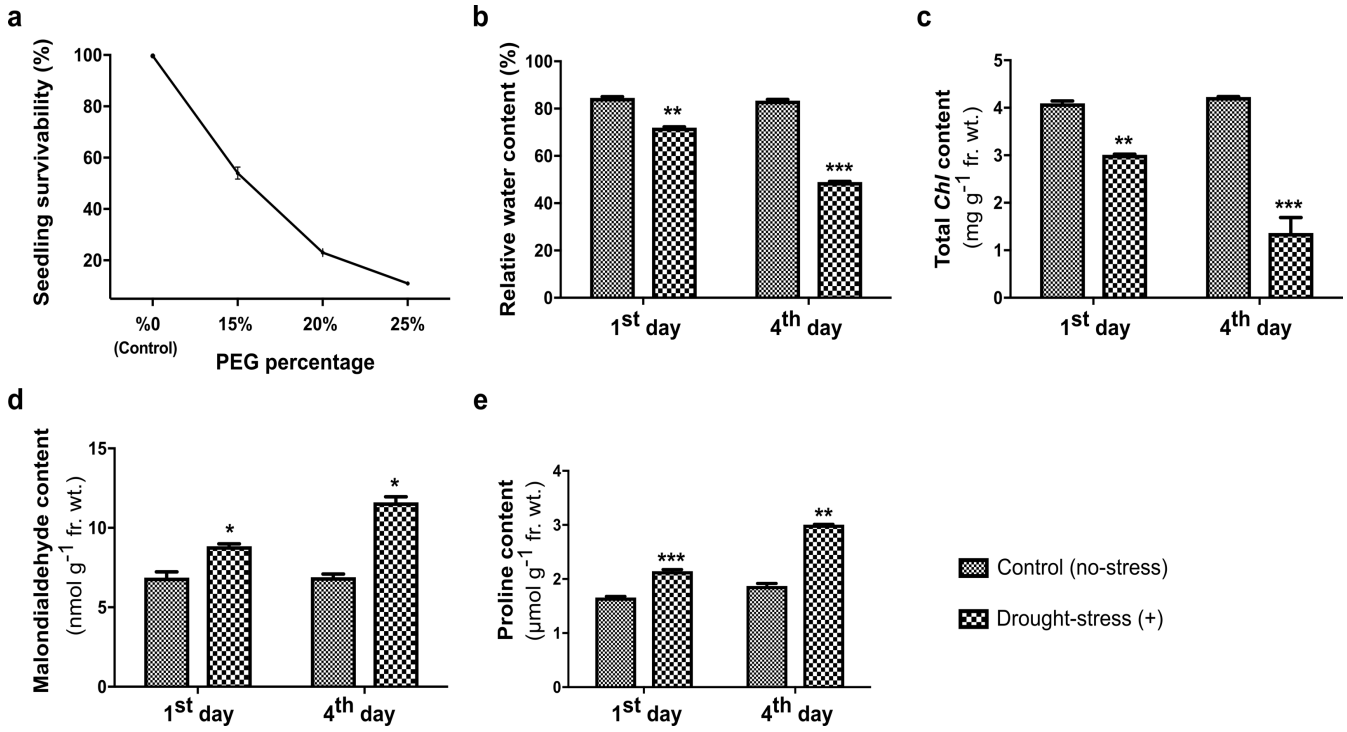

a
Seedling survivability (%)
100
80
60
40
20
%0
(Control)
15%
20%
25%
PEG percentage
b
Relative water content (%)
100
80
60
40
20
0
**
***
1st day
4th day
c
Total Chl content
(mg g-1 fr. wt.)
5
4
3
2
1
0
**
***
1st day
4th day
d
Malondialdehyde content
(nmol g-1 fr. wt.)
15
10
5
0
*
*
1st day
4th day
e
Proline content
(µmol g-1 fr. wt.)
4
3
2
1
0
***
**
1st day
4th day
Control (no-stress)
Drought-stress (+)

Figure 1

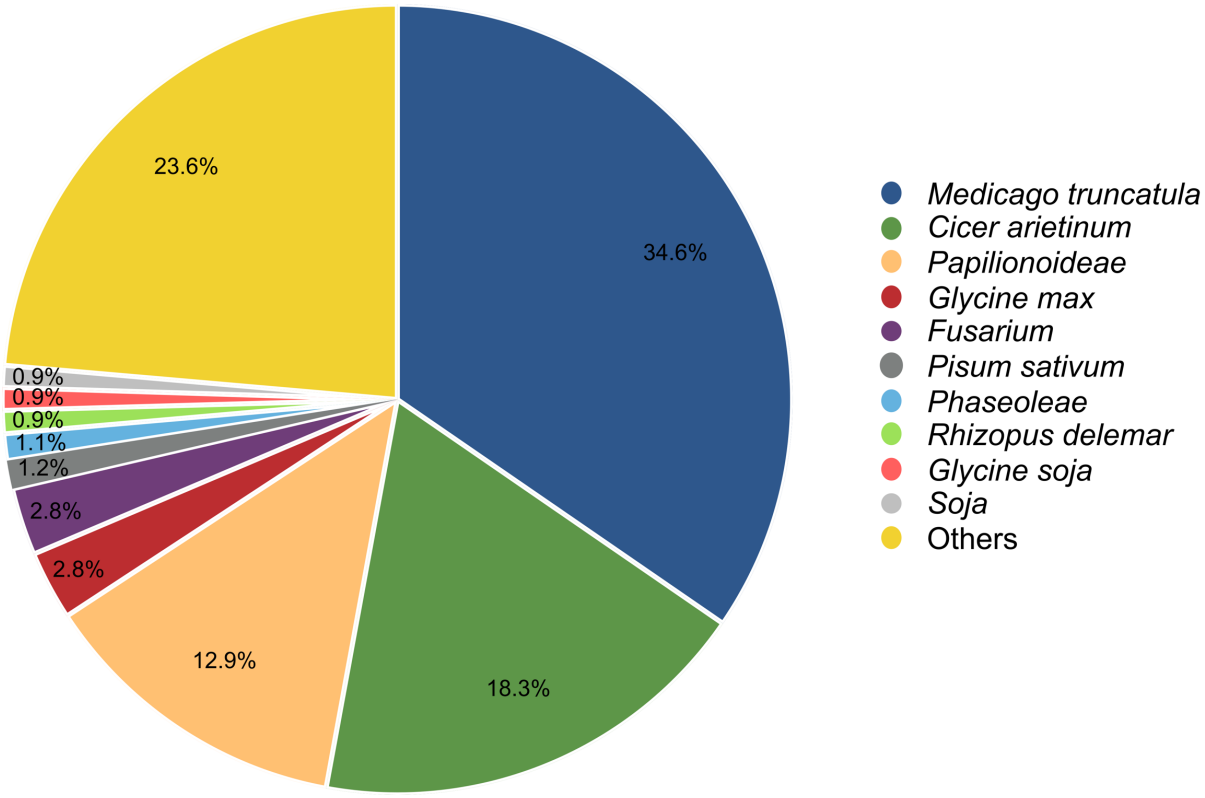

34.6%
18.3%
12.9%
2.8%
2.8%
1.2%
1.1%
0.9%
0.9%
0.9%
23.6%
Medicago truncatula
Cicer arietinum
Papilionoideae
Glycine max
Fusarium
Pisum sativum
Phaseoleae
Rhizopus delemar
Glycine soja
Soja
Others

Figure 2

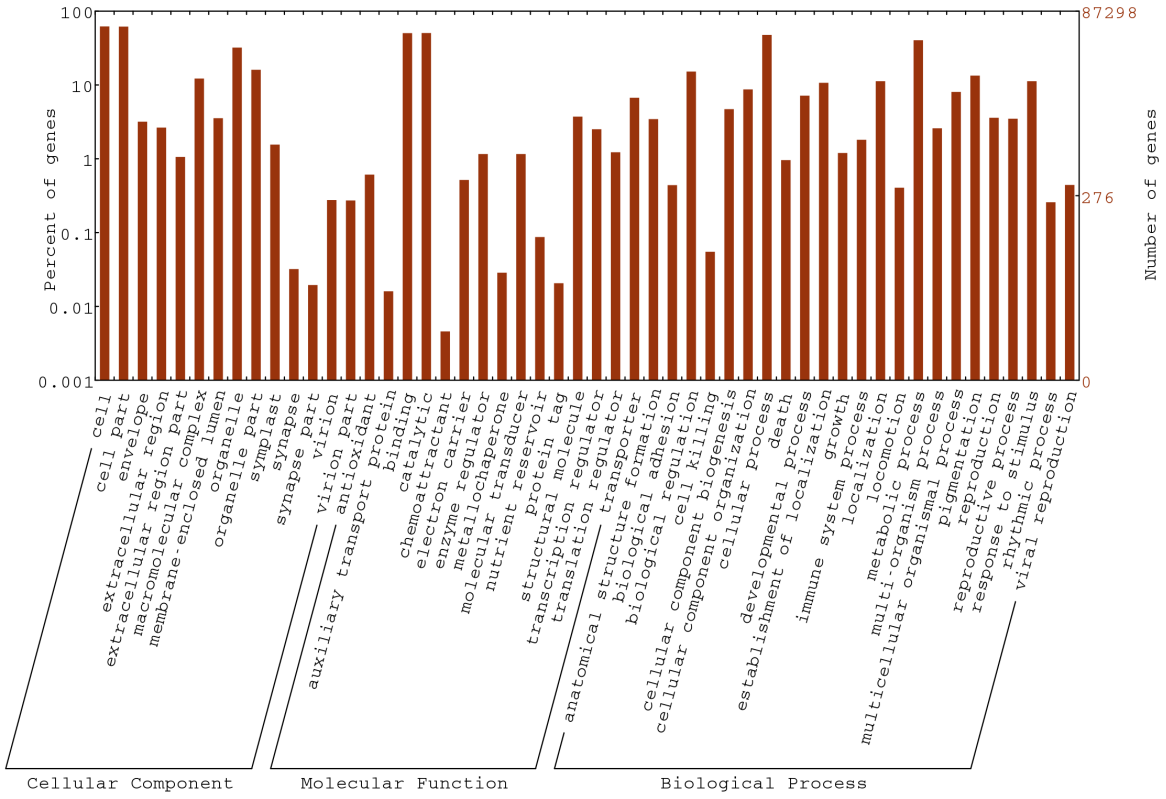

Percent of genes
100
10
1
0.1
0.01
0.001
87298
276
0
Number of genes
cell
cell part
envelope
extracellular region
extracellular region part
macromolecular complex
membrane-enclosed lumen
organelle
organelle part
symplast
synapse
synapse part
virion
virion part
antioxidant
auxiliary transport protein
binding
catalytic
chemoattractant
electron carrier
enzyme regulator
metallochaperone
molecular transducer
nutrient reservoir
protein tag
structural molecule
transcription regulator
translation regulator
transporter
anatomical structure formation
biological adhesion
biological regulation
cell killing
cellular component biogenesis
cellular component organization
cellular process
death
developmental process
establishment of localization
growth
immune system process
localization
locomotion
metabolic process
multi-organism process
multicellular organismal process
pigmentation
process
reproduction
reproductive process
response to stimulus
rhythmic process
viral reproduction
Cellular Component
Molecular Function
Biological Process

Figure 3

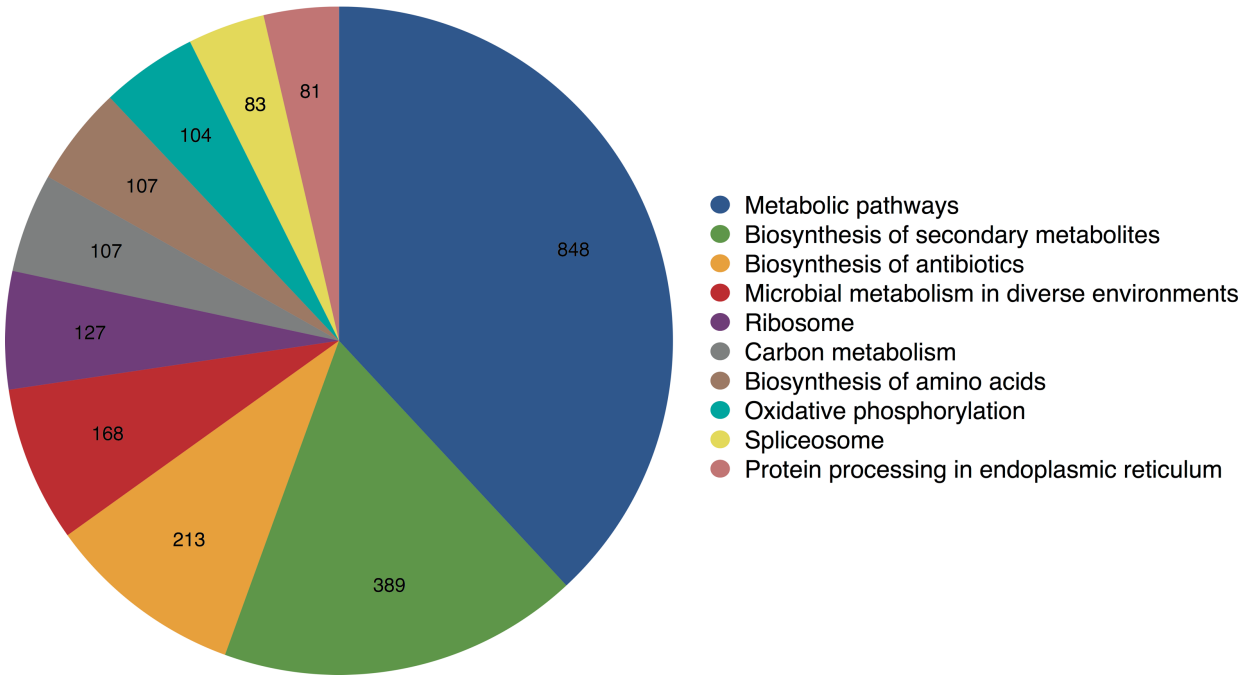
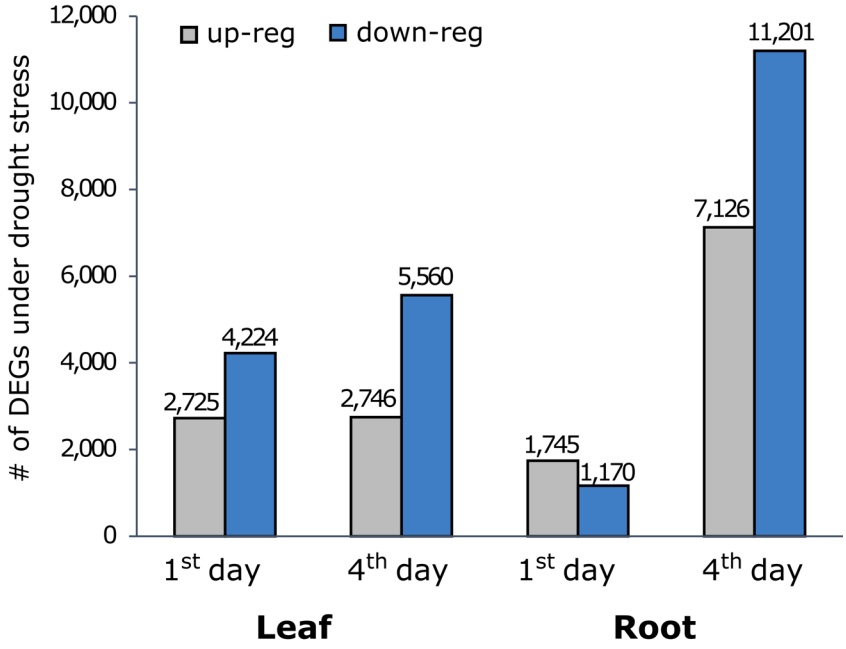

848
389
213
168
127
107
107
104
83
81
Metabolic pathways
Biosynthesis of secondary metabolites
Biosynthesis of antibiotics
Microbial metabolism in diverse environments
Ribosome
Carbon metabolism
Biosynthesis of amino acids
Oxidative phosphorylation
Spliceosome
Protein processing in endoplasmic reticulum

Figure 4

a


12,000
10,000
8,000
6,000
4,000
2,000
0
# of DEGs under drought stress
up-reg
down-reg
2,725
4,224
2,746
5,560
1,745
1,170
7,126
11,201
1st day
4th day
1st day
4th day
Leaf
Root

b

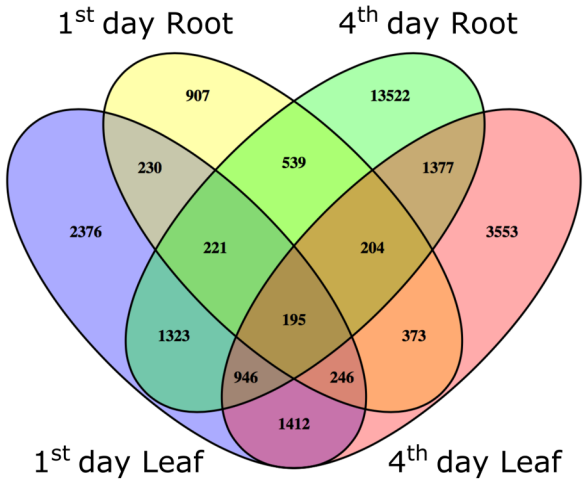

1st day Root
4th day Root
907
13522
230
539
1377
2376
221
204
3553
195
1323
373
946
246
1412
1st day Leaf
4th day Leaf

Figure 5

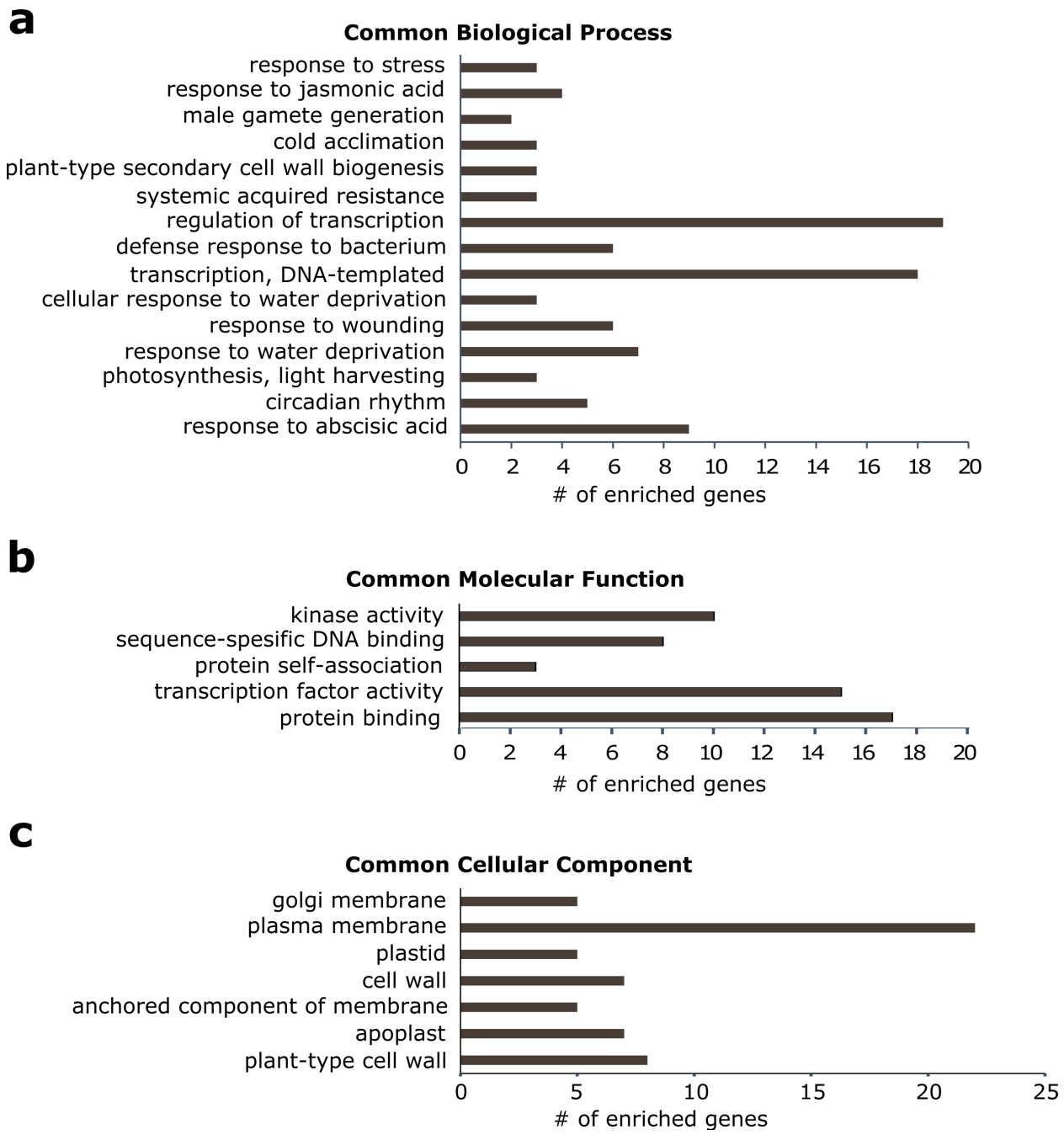

a
Common Biological Process
response to stress
response to jasmonic acid
male gamete generation
cold acclimation
plant-type secondary cell wall biogenesis
systemic acquired resistance
regulation of transcription
defense response to bacterium
transcription, DNA-templated
cellular response to water deprivation
response to wounding
response to water deprivation
photosynthesis, light harvesting
circadian rhythm
response to abscisic acid
0 2 4 6 8 10 12 14 16 18 20
# of enriched genes
b
Common Molecular Function
kinase activity
sequence-spesific DNA binding
protein self-association
transcription factor activity
protein binding
0 2 4 6 8 10 12 14 16 18 20
# of enriched genes
c
Common Cellular Component
golgi membrane
plasma membrane
plastid
cell wall
anchored component of membrane
apoplast
plant-type cell wall
0 5 10 15 20 25
# of enriched genes

Figure 6

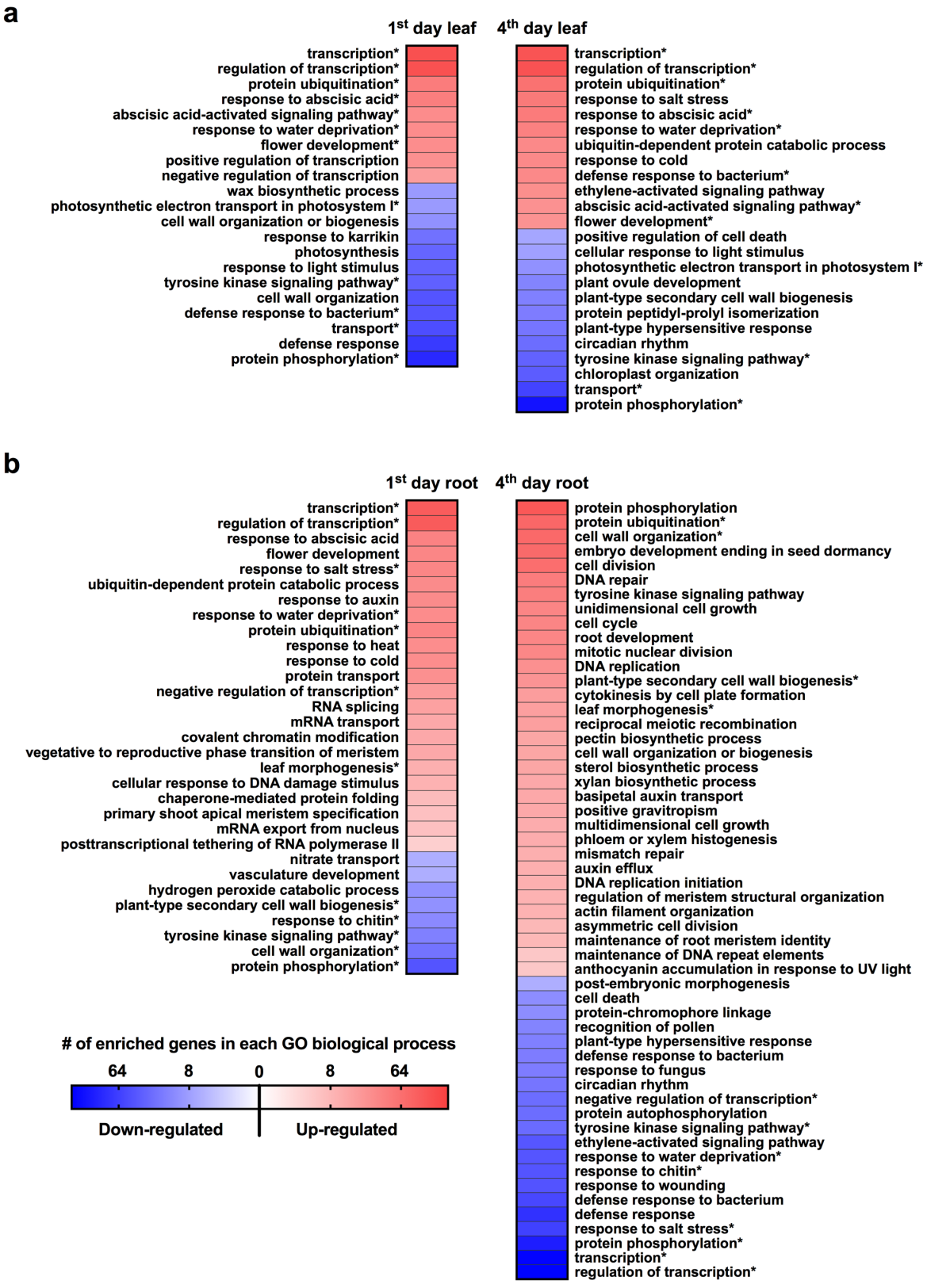

a
1st day leaf
4th day leaf
transcription*
regulation of transcription*
protein ubiquitination*
response to abscisic acid*
abscisic acid-activated signaling pathway*
response to water deprivation*
flower development*
positive regulation of transcription
negative regulation of transcription
wax biosynthetic process
photosynthetic electron transport in photosystem I*
cell wall organization or biogenesis
response to karrikin
photosynthesis
response to light stimulus
tyrosine kinase signaling pathway*
cell wall organization
defense response to bacterium*
transport*
defense response
protein phosphorylation*
transcription*
regulation of transcription*
protein ubiquitination*
response to salt stress
response to abscisic acid*
response to water deprivation*
ubiquitin-dependent protein catabolic process
response to cold
defense response to bacterium*
ethylene-activated signaling pathway
abscisic acid-activated signaling pathway*
flower development*
positive regulation of cell death
cellular response to light stimulus
photosynthetic electron transport in photosystem I*
plant ovule development
plant-type secondary cell wall biogenesis
protein peptidyl-prolyl isomerization
plant-type hypersensitive response
circadian rhythm
tyrosine kinase signaling pathway*
chloroplast organization
transport*
protein phosphorylation*
b
1st day root
4th day root
transcription*
regulation of transcription*
response to abscisic acid
flower development
response to salt stress*
ubiquitin-dependent protein catabolic process
response to auxin
response to water deprivation*
protein ubiquitination*
response to heat
response to cold
protein transport
negative regulation of transcription*
RNA splicing
mRNA transport
covalent chromatin modification
vegetative to reproductive phase transition of meristem
leaf morphogenesis*
cellular response to DNA damage stimulus
chaperone-mediated protein folding
primary shoot apical meristem specification
mRNA export from nucleus
posttranscriptional tethering of RNA polymerase II
nitrate transport
vasculature development
hydrogen peroxide catabolic process
plant-type secondary cell wall biogenesis*
response to chitin*
tyrosine kinase signaling pathway*
cell wall organization*
protein phosphorylation*
# of enriched genes in each GO biological process
64
8
0
8
64
Down-regulated
Up-regulated
protein phosphorylation
protein ubiquitination*
cell wall organization*
embryo development ending in seed dormancy
cell division
DNA repair
tyrosine kinase signaling pathway
unidimensional cell growth
cell cycle
root development
mitotic nuclear division
DNA replication
plant-type secondary cell wall biogenesis*
cytokinesis by cell plate formation
leaf morphogenesis*
reciprocal meiotic recombination
pectin biosynthetic process
cell wall organization or biogenesis
sterol biosynthetic process
xylan biosynthetic process
basipetal auxin transport
positive gravitropism
multidimensional cell growth
phloem or xylem histogenesis
mismatch repair
auxin efflux
DNA replication initiation
regulation of meristem structural organization
actin filament organization
asymmetric cell division
maintenance of root meristem identity
maintenance of DNA repeat elements
anthocyanin accumulation in response to UV light
post-embryonic morphogenesis
cell death
protein-chromophore linkage
recognition of pollen
plant-type hypersensitive response
defense response to bacterium
response to fungus
circadian rhythm
negative regulation of transcription*
protein autophosphorylation
tyrosine kinase signaling pathway*
ethylene-activated signaling pathway
response to water deprivation*
response to chitin*
response to wounding
defense response to bacterium
defense response
response to salt stress*
protein phosphorylation*
transcription*
regulation of transcription*

Figure 7