**Journal of Genetics** (2020)99:35
https://doi.org/10.1007/s12041-020-01195-w



**RESEARCH ARTICLE**

# Identification of drought stress-responsive genes in rice (*Oryza sativa*) by meta-analysis of microarray data

PREETI SIROHI, BIRENDRA S. YADAV, SHADMA AFZAL, ASHUTOSH MANI and NAND K. SINGH*

*Department of Biotechnology, Motilal Nehru National Institute of Technology Allahabad, Prayagraj 211 004, India*
*For correspondence. E-mail: singhnand@gmail.com, nksingh@mnnit.ac.in.

Received 19 August 2019; revised 8 January 2020; accepted 12 February 2020

**Abstract.** Meta-analysis provides a systematic access to the previously studied microarray datasets that can recognize several common signatures of stresses. Three different datasets of abiotic stresses on rice were used for meta-analysis. These microarray datasets were normalized to regulate data for technical variation, as opposed to biological differences between the samples. A *t*-test was performed to recognize the differentially-expressed genes (DEGs) between stressed and normal samples. Gene ontology enrichment analysis revealed the functional distribution of DEGs in different stressed conditions. Further analysis was carried out using software RICE NET DB and divided into three different categories: biological process (homoiothermy and protein amino acid phosphorylation), cellular component (nucleus and membrane), and molecular function (zinc ion binding ad DNA binding). The study revealed that 5686 genes were constantly expressed differentially in *Oryza sativa* (2089 upregulated and 3597 downregulated). The lowest *P* value ($P = 0.003756$) among upregulated DEGs was observed for naringenin, 2-oxoglutrate 3-dioxygenase protein. The lowest *P* value ($P = 0.002866816$) among the downregulated DEGs was also recorded for retrotransposon protein. The network constructed from 48 genes revealed 10 hub genes that are connected with topological genes. These hub genes are stress responsive genes that may also be regarded as the marker genes for drought stress response. Our study reported a new set of hub genes (reference genes) that have potentially significant role in development of stress tolerant rice.

**Keywords.** meta-analysis; drought; abiotic stress; downregulated; upregulated; differentially expressed genes; hub genes.

## Introduction

The natural environment of the plant is composed of several rate-limiting constituents, these may regulate the plant growth in several ways. Abiotic stresses are one of the major factors altering the metabolic pathway of plant system (Cramer *et al.* 2011). They follow a stringent pathway, depends on the type of stimulant (drought, water logging, cold heat and adverse condition). These stimulants or stresses may alter the normal plant growth as a result of differential gene expression.

Abiotic stress such as drought, water-logging/flooding, cold, heat, alkalinity/acidity of soil and metal toxicity adversely affect the growth, development, yield and seed quality of plants/crops which can lead to changes/alteration in gene expression. Under stress conditions, plants have developed complex mechanisms to distinguish external signals which allow them to acclimatize in changing environmental conditions for their survival. The emergence of high throughput technologies such as microarray and next-generation sequencing is used for the study of the unusual genomic alteration in biotic and abiotic stressed plants. Using these techniques, researchers can detect the genomic alteration in the plant cells caused by specific stimulants.

Plant stress usually reflects sudden changes in the plant system that ultimately lends a signature effect on the plant system. These variations are gradually adopted for the survival of the plant. These variations are not a result of single variable, it is a multi-gene system that regulates and maintain the physiological homeostasis of the plant system (Velazquez *et al.* 2011). Drought is found to be one of the major abiotic stress affecting the plant growth and yield. It reduces the rate of photosynthesis in the cells by exerting several undesired stress on the cells, such as hardening of the cell wall, accumulation of reactive oxygen species (ROS) and secondary metabolites (You and Chan 2015). During a study, it came to light that the

Conceived and designed the experiment, NKS and AM; performed the experiments, PS and BSY; analysed the data; PS and SA; wrote the paper, PS.

Published online: 22 April 2020

drought affected the total DNA methylation pattern which counted as an average total of 12.1% methylation differences when calculated across different tissue, genotype and developmental stage (Saraswat *et al.* 2017). In plant system, the cell wall is composed of cellulose and connected to hemicellulose. Besides these major components, it also has a trace amount of phenolics, esterase, pectin, expansin and several other proteins. Drought and other stress cause the production of ROS leading to crosslinking of phenolics with glycoproteins resulting in hardening of the cell wall (Tenhaken 2014). This hardening of cell wall ultimately reduces the leaf surface area and photosynthetic rate per unit area (Basu *et al.* 2016).

Plants have diverse mechanisms to sustain drought at cellular and physiological levels. At the physiological level, they reduce water loss by a decreased rate of diffusion through stomata and other parts, increases water absorption with the enhanced root system, smaller and succulent leaves to diminish water deficiency. Drought stress in plant can be managed by executing few tactics such as marker-assisted selection, breeding and mass screening and an extracellular spray of hormones and osmoprotectants (Fang and Xiong 2015).

The plants experiencing stresses, i.e. biotic or abiotic, independently and facing positive and negative influence have shown crosstalk between them (Abuqamar *et al.* 2009). These studies have offered opportunities to improve plants for fighting different individual stress tolerance (Mao *et al.* 2010). Currently, the gene expression data from separate experiments of biotic and abiotic stress have been exploited to find out shared stress-responsive genes (Shaik and Ramakrishna 2013, 2014). Plants in nature are generally tested by various biotic and abiotic stresses. Plants capable of tolerating two or more independent stresses are not really capable of tolerating these stress altogether (Atkinson and Urwin 2012; Ramegowda and Kumar 2015). The reports showed that plants are capable of coping up with coinciding biotic and abiotic stresses through demonstration of relevant responses which cannot be comprehended by specifically concluding the outcomes from specific stress studies where each stress is applied separately (Bostock *et al.* 2014). There is a requirement of comprehending coinciding biotic and abiotic stress tolerance of plants because sufficient work was not done for this purpose. The answer to this problem is combining the different stressed conditions data and find out specific markers in response to stress through meta-analysis.

Meta-analysis of available data has the potential to explore the transcriptomic studies (Feichtinger *et al.* 2012). It is based on the statistical analysis of multiple studies on the similar experimental conditions, based on this one can easily identify the variation and have a reason for this alteration. By comparing the studies on the statistical background, it increases the reliability of the outcome with a given set of data that may be in the form of genes called as differentially-expressed genes (DEGs) (Ramasamy *et al.* 2008).

Keeping the above situation in mind, this work was designed to study abiotic stress on plants together by comparing the microarray datasets of different stresses on rice (*Oryza sativa*). The objective of this study was that a meta-analysis of freely accessible microarray datasets of different biotic and abiotic stresses on rice can recognize a common mark of stresses and apply a meta-analysis method on various microarray datasets which then approves the signature in individual datasets (Daves *et al.* 2011).

## Materials and methods

### *Microarray datasets selection*

The publicly available microarray studies were searched using several keywords and their combinations such as '*O. sativa*, abiotic stress, gene expression, microarray and genome' by using Gene Expression Omnibus (GEO) database for candidate genes in *O. sativa* gene expression datasets. The data collected further screened for eligibility and the duplicate data was excluded (figure 1).

The available experimental datasets and their corresponding experimental conditions were only attributed towards the gene expression profiling in *O. sativa* between controlled and stressed conditions. The data of stressed samples were collected after administration of stress conditions. In the present study, a meta-analysis was performed as per the guidelines of PRISMA statement. Data were accessed from original studies having GEO accession number, analysis platform, the number of cases and controls, gene expression data and related references (Yang *et al.* 2014). Among the data, two of them were collected from root tissues and one from the seedling of the rice. The data with GEO ID GSE36661 administered with drought stress, similarly GSE62308 and GSE64576 administered with ABA-regulated drought stress (table 1).

### *Data curation*

In data curation, normalization of available data and related parameters is a very crucial step in comparing microarray datasets. It became very difficult in a direct comparison between altered datasets from various sources; these differences arise mainly due to the use of the different platform, gene nomenclature and tissue used as a control. Variation in the normalization may lead to the probability of distorting comparative outcome; it reduces the authentic computation of candidate gene expression changes. As a consequence, there may be a need to consider a globally accepted normalization pathway for minimal inconsistency. Z score transformation method is very reliable and sensitive tool to compute the expression potencies of each probe in gene expression profiles and computed by the given formula.

$$\text{Z score} = \frac{(\text{xi} - \bar{\text{x}})}{\delta},$$

where xi denotes raw data for each gene; $\bar{\text{x}}$ denotes average gene intensity within a single experiment and $\delta$ denotes

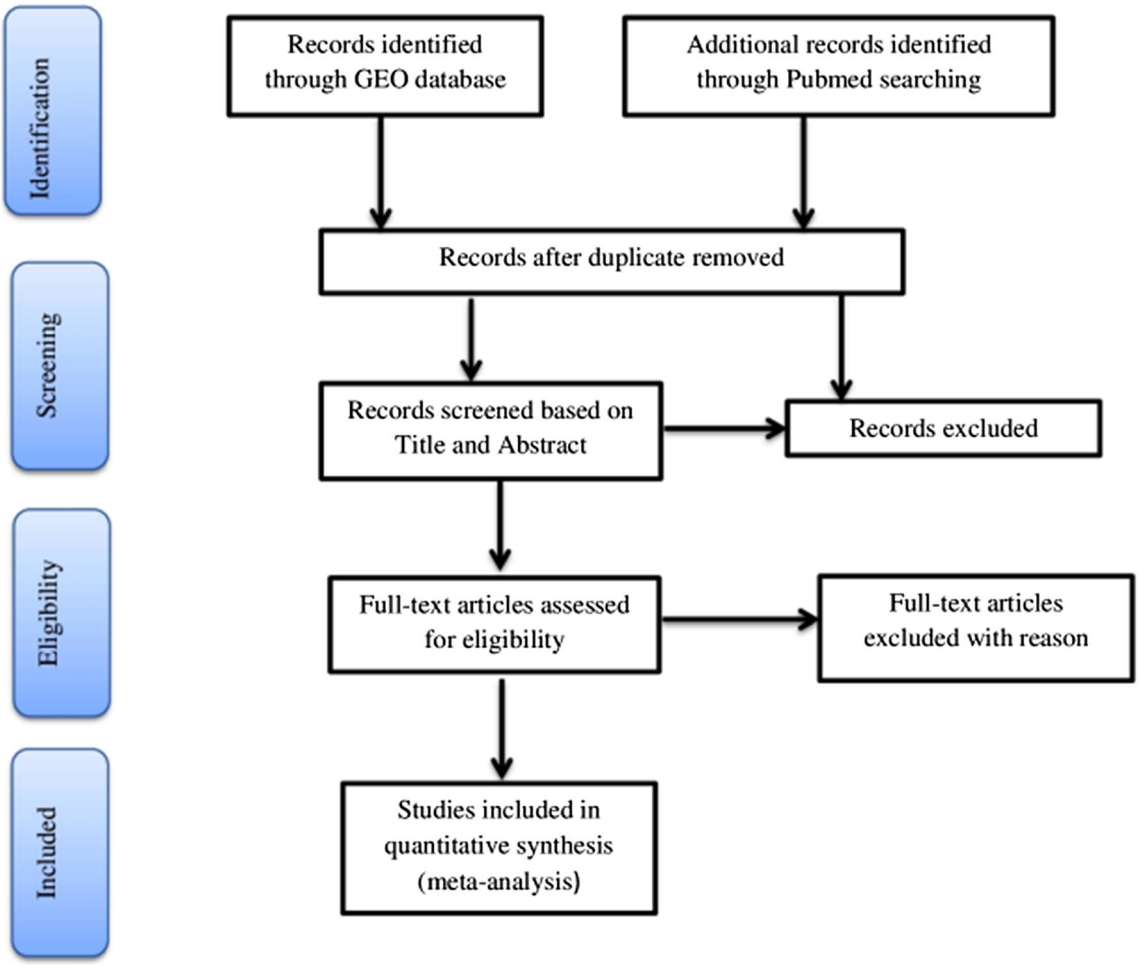


**Figure 1.** Flowchart of the selected process of microarray datasets for the meta-analysis.

**Table 1.** Characteristics of the individual studies included in study.

| GEO ID | Sample count (case : control) | Platform | Tissue | Type of stress |
|---|---|---|---|---|
| GSE36661 | 3:3 | Affymetrix rice genome array | Root | Disease and drought |
| GSE62308 | 2:2 | Affymetrix rice genome array | Seedling | ABA-regulated drought stress |
| GSE64576 | 4:4 | Affymetrix rice genome array | Root | ABA-regulated drought stress |

standard deviation (SD) of all measured potencies (Yang *et al.* 2014).

### *Statistical analysis*

To identify the DEGs between stressed and control samples statistical significance analysis of microarray (SAM) was used. Microsoft Excel 2010 was used to analyze the data. The specific *t*-test (one-tailed, paired *t*-test) was performed to validate the significance of DEGs with a 'comparative difference' score for screened candidate genes (Yang *et al.* 2014). The average expression change from various expression forms the standard deviation of values for that gene is termed as *P* value (Hou *et al.* 2014). DEGs were selected from those genes which showed at least two-fold changes equivalent to a false discovery rate (FDR) $< 0.01$ (Tusher *et al.* 2001).

### *Venn diagram*

Venn diagrams were prepared using online tool Venny v2.1. The individual microarray data was normalized and fold change values were calculated. Fold change value $> 2$ were selected as upregulated DEGs and fold change value $< 2$ were selected as downregulated DEGs. The list of selected reference IDs were pasted in the online tool which will provide the common upregulated and common downregulated DEGs among all three microarray data. Venn diagrams

were also constructed for tissues vs conditions as shown in figure 2.

### *Functional classification of DEGs*

Functional distribution and biological significance of screened DEGs were further analysed on RICE NET DB (Narsai *et al.* 2013). These online tools perform the gene ontology (GO) enrichment analysis and find the biological significance of candidate gene in a wide variety of datasets (Madrid *et al.* 2012).

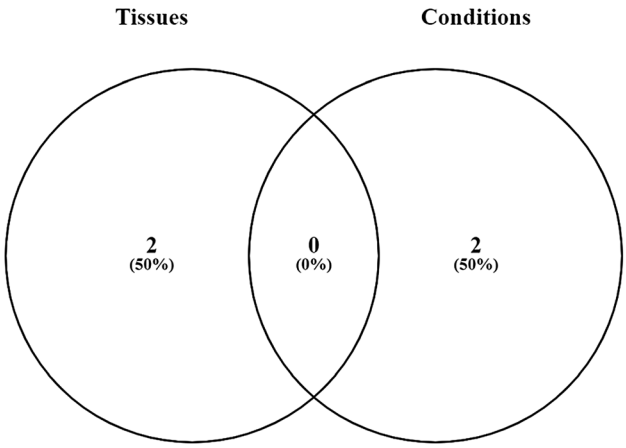


**Figure 2.** Venn diagrams of tissues vs conditions of microarray data.

### *Biological network construction*

Coexpression networks are useful for associate genes that are involved in the same biological pathway or that are of protein complexes (Moyano *et al.* 2015). The coexpression network was constructed for the genes greater than fold change ± 1.5 and *P* value less than 0.05 by constructing the coefficient correlation matrix. Those genes that have Pearson's coefficient correlation above the cut-off value, i.e. ± 0.95 are used for construction of biological network using expression correlation and network analyzer by network analyst of Cytoscape (Shannon *et al.* 2003). In biological network, nodes represent the genes and nodes represent the connectivity between the genes. Hub genes (most connected genes among the biological network) were also screened using cytoHubba plugin of cytoscape (Chin *et al.* 2014).

## Results and discussion

Drought is among the major abiotic stresses which limits the crop production worldwide. Drought can be described as the deficiency of water for a period of time which results in

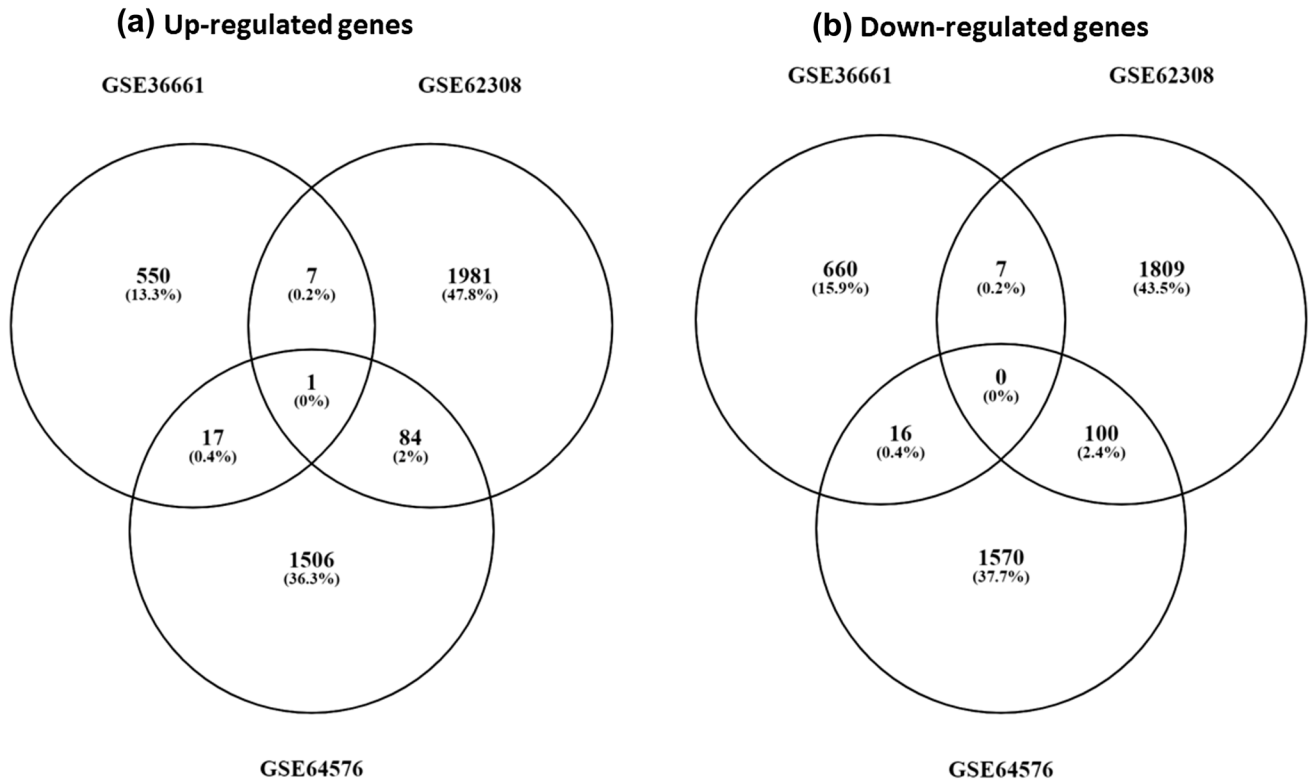


**Figure 3.** Venn diagrams summarizing microarray data analysis: (a) Venn diagram showing number of upregulated and (b) downregulated genes in rice after drought stress. Only genes with log2 fold change above 2 were considered for this analysis.

water shortage, depleting soil moisture and cause adverse effects on plants. Drought drastically devastates plant growth and development with considerable reductions in crop growth rate and biomass accumulation. Meta-analysis approach integrates DEGs from microarray datasets which were expressed consistently with statistical significance and performed GO enrichment analysis. High-throughput transcriptomic data enable meta-analysis of multiple datasets which lead to discovering robust candidate gene for stress (Fang *et al.* 2015). In this study, we identified DEGs by comparing transcriptomic responses in stress and normal rice. Most important candidate genes were identified in pathways involved in abiotic stress. The gene expression of rice during abiotic stress indicated that the expression levels of these DEGs were changed with induction of stress in rice (Yu *et al.* 2014).

The Venn diagram was prepared by the induction of stress in rice to identify the exclusively upregulated and downregulated genes. Five hundred and fifty genes were exclusively upregulated in GSE36661, 1981 genes in GSE62308 and 1506 genes in GSE64576 (figure 3a). Only one gene was found to upregulated in all three data conditions. Six hundred and sixty genes were exclusively downregulated in GSE36661, 1809 in GSE62308 and 1570 in GSE64576 (figure 3b). There was no common gene found to be downregulated among them.

### *Detecting genes associated with drought stress*

During this study, a collection of total three expression profiling studies were used as per the inclusion criteria, among these profiles, nine treated samples of *O. sativa* and nine control samples were included. In this study, two roots samples and one seedling samples were utilized. Detecting genes linked with *O. sativa* to find out the genetic markers which are engaged in the development of stress condition during growth of *O. sativa*, the probe ID should be similar for all the data and microarray platform, which denote a named gene. The probe ID should be converted to gene ID or locus ID. The expression value was logarithmically transformed (base 2) which gives a total of 57,381 genes for this study. For each gene, expression values were changed to the Z-score for the objective of global normalization. The changes in gene expression between stressed/treated and control/normal *O. sativa* were investigated using the assembled expression compilation. After that, SAM method was used to identify the DEGs between stressed and control samples. A total of 5686 genes with an FDR of minimum 2 and –2 in upregulated and downregulated genes, after applying a minimum *P* value of 0.01, were found with changed expression in samples of stressed *O. sativa* in comparison to control. Of the total 5686 DEGs, the number of upregulated genes is 2089 and the number of downregulated genes is 3597 at the level of 1% significance. The DEG with lowest *P* value ($P = 0.003756$) among

**Table 2.** Details of downregulated genes obtained during meta-analysis.

| Affymetrix ID | Locus ID | Gene expression | Fold change | *P* value |
|---|---|---|---|---|
| OsAffx.30383.1.S1_at | LOC_Os10g10720 | Retrotransposon protein | – 4.286387491 | 0.002866816 |
| Os.16233.1.S1_a_at | LOC_Os04g09900 | Diterpene phytoalexin precursor biosynthetic process pathway and ent-kaurene synthase, chloroplast precursor | – 8.718961239 | 0.006254085 |
| OsAffx.19894.1.S1_at | LOC_Os12g30150 | CAMK_CAMK_like.47 - CAMK includes calcium/calmodulin dependent protein kinases | – 2.317901962 | 0.008920827 |
| OsAffx.4679.1.S1_at | LOC_Os05g50730 | Expressed protein | – 5.779275181 | 0.009377673 |
| OsAffx.13951.2.S1_s_at | LOC_Os04g24310 | Retrotransposon protein and jasmonate- induced protein | – 3.883709153 | 0.009379355 |
| | LOC_Os04g24319 | | | |
| | LOC_Os04g24478 | | | |
| Os.46470.2.S1_x_at | LOC_Os10g09110 | Cytochrome P450 | – 3.167526416 | 0.012599032 |
| | LOC_Os05g32730 | and tetratricopeptide -like helical | | |
| OsAffx.26087.1.S1_at | LOC_Os06g32310 | Hypothetical protein | – 3.991934519 | 0.013500274 |
| OsAffx.21309.1.S1_at | LOC_Os07g43180 | SKP1-like protein 1B and ubiquitin-dependent protein catabolic process | – 5.591710319 | 0.014619313 |
| OsAffx.27515.1.S1_x_at | LOC_Os06g08690 | Leucine-rich repeat receptor protein kinase EXS precursor, | – 3.387832634 | 0.015094976 |
| | LOC_Os06g08710 | receptor-like protein kinase precursor | | |
| Os.4784.1.S1_at | LOC_Os09g31040 | EF hand family protein | – 3.829875411 | 0.015256946 |

**Table 3.** Details of upregulated genes obtained during meta-analysis.

| Affymetrix ID | Locus ID | Fold change | Gene expression | *P* value |
|---|---|---|---|---|
| Os.7095.1.S1_at | LOC_Os04g49210 | 2.200212 | Naringenin,2-oxoglutarate 3-dioxygenase | 0.003756 |
| Os.16317.1.S1_at | LOC_Os04g45480 | 3.152641 | Heat shock protein STI | 0.004203 |
| Os.2292.3.S1_x_at | LOC_Os03g53340 | 3.090799 | HSF-type DNA-binding domain containing protein | 0.005923 |
| Os.10942.1.S1_a_at | LOC_Os10g28340 | 6.058194 | Heat stress transcription factor | 0.006883 |
| Os.49648.1.S1_s_at | LOC_Os02g54140 | 3.414359 | hsp20/alpha crystallin family protein | 0.007126 |
| Os.55306.1.S1_at | LOC_Os09g27330 | 3.232658 | Oxidoreductase / transition metal ion binding protein | 0.007448 |
| Os.11376.1.S1_at | LOC_Os06g06490 | 2.300383 | U-box domain containing heat shock protein | 0.008226 |
| Os.5574.1.S1_s_at | LOC_Os02g08490 | 4.631049 | Hypothetical protein | 0.009367 |
| Os.8032.1.S1_at | LOC_Os04g57440 | 2.292794 | Oryzain beta chain precursor | 0.009751 |
| Os.5817.1.S1_at | LOC_Os06g09560 | 4.495812 | Heat shock protein DnaJ | 0.010119 |

upregulated DEGs express naringenin, 2-oxoglutarate 3-dioxygenase protein. The predicted subcellular location of this protein is cytosol chloroplast. This protein is involved in flavonoid biosynthesis pathway, which is a part of secondary metabolite biosynthesis. The DEG with lowest *P* value ($P = 0.002866816$) among the downregulated DEGs express retrotransposon protein which is a member of cytochrome P450 family protein. Retrotransposons are regulators of gene expression mediated through RNA intermediate (Elbarbary *et al.* 2016). Therefore during the drought condition, plants become incapable of synthesis of proteins helpful to overcome the stress condition (tables 2 & 3).

Top 10 significantly downregulated and upregulated DEGs are listed in tables 2&3. Among the most upregulated DEGs, most of the genes are responsible for the proteins like heat shock protein STI, heat shock protein DnaJ, U-box domain containing heat shock protein, hsp20/alpha crystallin family protein. In response to drought stress, plants develop protective strategies to cope up with it. Heat shock proteins synthesis is one of the protective tools of the plants to provide a defense against drought and heat stress (Virdi *et al.* 2015).

Similarly, a few downregulated DEGs are observed in response to drought stress which regulates the proteins like

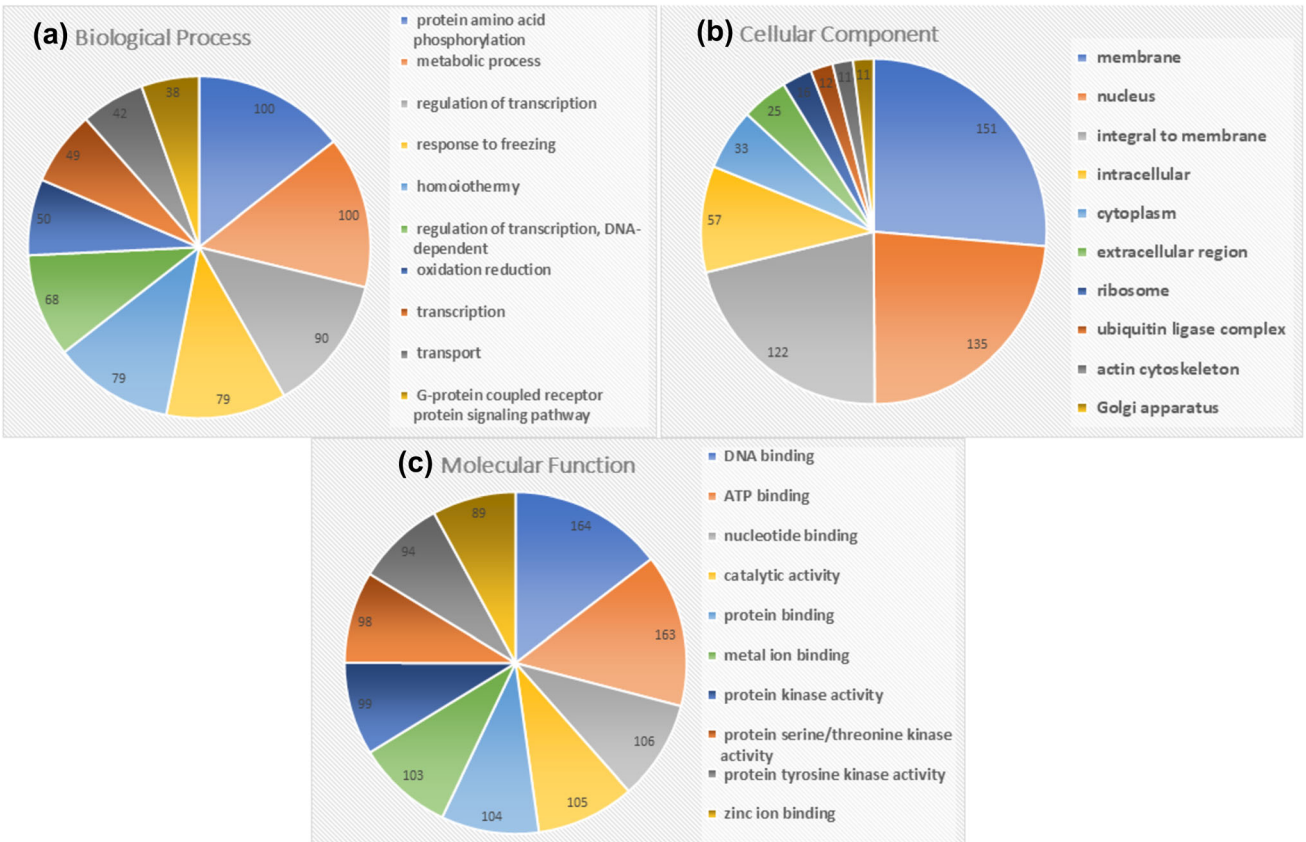


**Figure 4.** The top 10 enriched GO terms of upregulated DEGs. (a) Biological process for DEGs; (b) cellular component for DEGs; (c) molecular function for DEGs.

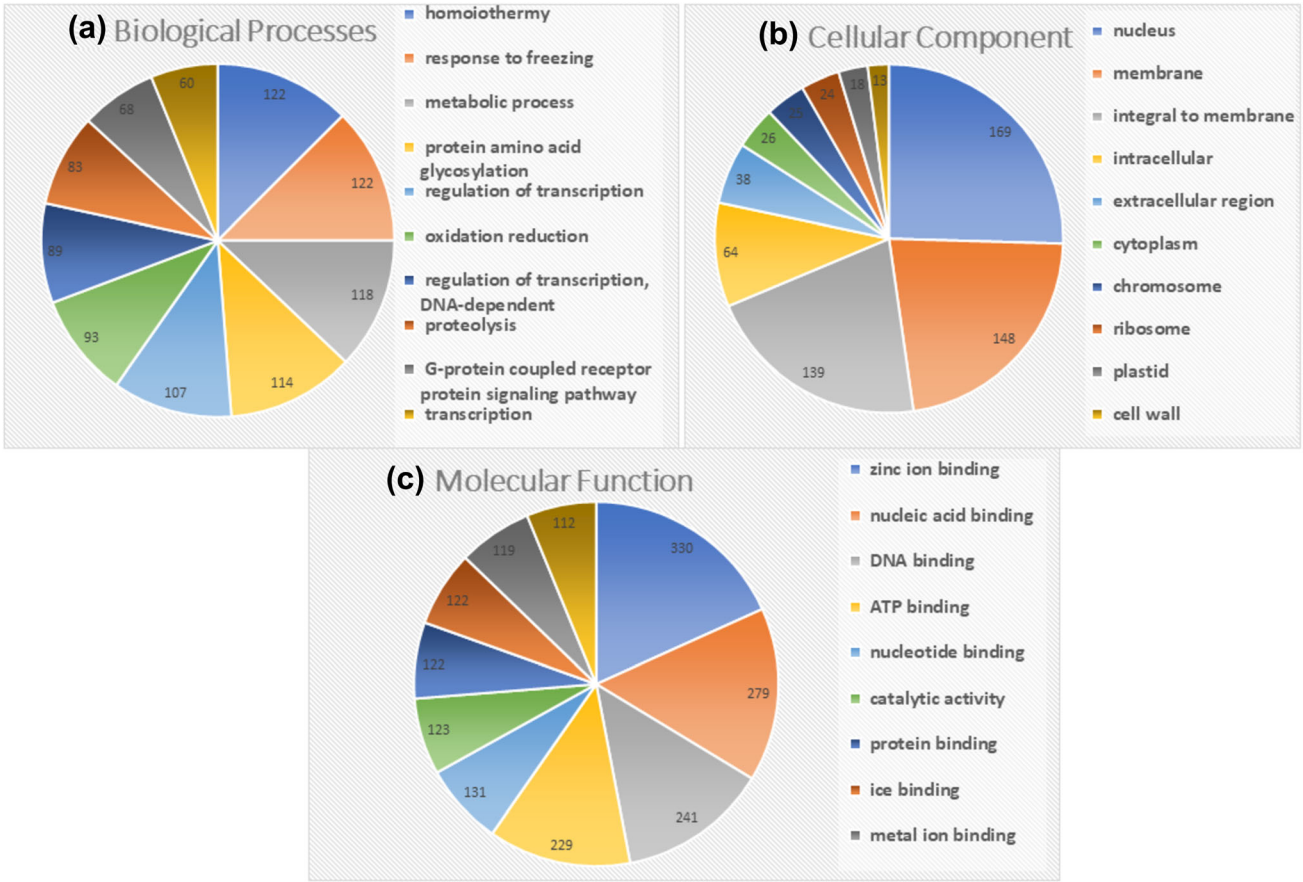


**Figure 5.** The top 10 enriched GO terms of downregulated DEGs. (a) Biological process for DEGs; (b) cellular component for DEGs; (c) molecular function for DEGs.

retrotransposon protein, diterpene phytoalexin precursor biosynthetic process pathway and ent-kaurene synthase, chloroplast precursor, CAMK_CAMK_like.47 - CAMK includes calcium/calmodulin-dependent protein kinases. Calmodulin (CaM) acts as an integrator of different stress signalling pathways, which allows plants to maintain homeostasis between different cellular processes (Virdi *et al.* 2015). In case of drought stress, plants encounter osmotic stress which however, further induces a chain of various responses at the molecular and cellular levels. Due to osmotic stress, the concentration of the cytosolic $Ca^{2+}$ increases which transduces $Ca^{2+}$ signals. $Ca^{2+}$-ion signalling induces appropriate cellular responses to overcome the damage caused by the drought stress (Zeng *et al.* 2015).

### *Functional annotation*

The biological significance of the DEGs could be understood by performing GO enrichment analysis from *O. sativa*. A typical descriptive model and functional annotation and categorization to study the gene set information were provided by gene ontology. GO groups are arranged into three categories, namely, biological process, cellular component and molecular function. Genes with the minimum significance level of ($P < 0.01$) 1% were selected and were tested against the background set of all genes with GO annotations. The biological process, molecular functions, and cellular components were investigated separately by web-based software RICE NET DB. The GO terms found for biological process are significantly enriched in protein amino acid phosphorylation (GO:0006468, hyper $P = 0.0000$) (figure 4a) and homoiothermy (GO:0042309, hyper $P = 0.0000$) (figure 5a) while for cellular component, the enriched GO terms were membrane (GO:0016020, hyper $P = 0.0000$) (figure 4b) and nucleus (GO:0005634, hyper $P = 0.0000$) (figure 5b), and for molecular functions, the enriched GO terms were DNA binding (GO:0003677, hyper $P = 0.0000$) (figure 4c) and zinc ion binding (GO:0008270, hyper $P = 0.0000$) (figure 5c).

Based on the GO analysis, the upregulated genes mostly performed the DNA binding, i.e. transcription factor localized in the membrane and participate in protein amino acid phosphorylation activity, while downregulated genes mostly performed the zinc ion binding factor localized in the nucleus and participate in homoiothermy, which indicated

that the temperature regulation of plant was affected due to downregualtion of biological process homiothermy. DNA-binding transcription factor such as CaM7 regulates plant response to light signals to reduce the probable damage caused by drought stress. Post-translational modifications (PTMs) involve protein phosphorylation which changes protein function, protein–protein interaction and cellular localization. Phosphorylated drought-responsive proteins play major role in signalling, transcription and photosynthesis, as well as in protein synthesis. The investigation of physiological, molecular and proteomic studies related to drought-responsive traits gives insights for further understanding of plant drought tolerance (Wang *et al.* 2016).

### *Biological network analysis*

A set of 52 genes having fold change $> \pm 1.5$ and $P$-value $< 0.05$ was taken for construction of coefficient correlation matrix. Of the 52 genes, 48 cleared the cut-off and biological network was constructed (figure 6). Different parameters of network obtained from microarray expression data of rice during drought stress is shown in table 4. Clustering coefficient is low which represents the property of biological network.

The top 10 hub genes with their degree which were identified from network is shown in table 5. The maximum of connectivity, i.e. degree of gene was 12 during drought stress. These genes are differentially expressed during different conditions and are most connected genes, having key role in different biological process and molecular function (figure 7).

The hub genes identified during network analysis were alpha amylase isozyme 3D which belongs to gylcosyl hydrolase 13 family; myb family transcription factor APL required for the phloem identity and regulates the expression of transcription factor NAC045 (direct the sieve element enucleation and cytosol degradation), they may also activate the transcription of specific genes involved in phosphate uptake or assimilation. Heat stress transcription factors (transcriptional regulators) were identified as hub genes that specifically binds DNA of heat shock promotor elements (HSE). OsWRKY71 (transcription factor) identified hub gene might function as a transcriptional regulator in rice defense signalling pathways. WRKY proteins are a large family of transcription factors that mainly participate in plant biotic stress responses, therefore they are responsible for the development of drought stress tolerance in rice (Liu *et al.* 2006).

Drought creates water deficiency in plants/rice which affect the physiological functions of the rice as rice require a large amount of water for its physiological functions. Hence, it

**Table 4.** Different parameters of biological network obtained from microarray expression data of rice on the basis of Pearson's coefficient correlation during exposure of drought stress using cytoscape software.

| | |
|---|---|
| Clustering coefficient | 0.3409 |
| Shortest path | 337 |
| Characteristic path length | 1.7032 |
| Average no. of neighbours | 7.958 |
| Network density | 0 |
| No. of nodes | 48 |

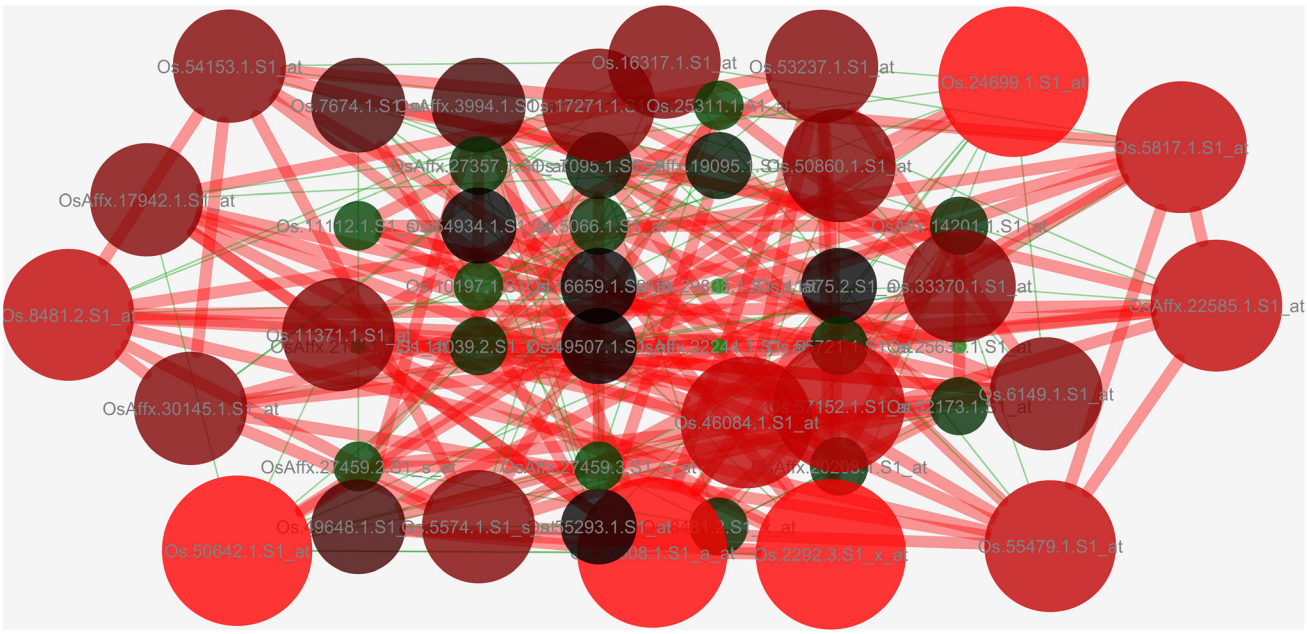


**Figure 6.** Biological network constructed on the basis of Pearson's coefficient correlation using the expression data of rice under drought stress.

**Table 5.** Hub genes obtained from biological network of drought stress data of rice (*O. sativa*).

| | Hub genes | Degree | Gene name |
|---|---|---|---|
| 1 | Os.24699.1.S1_at | 12 | myb family transcription factor APL (LOC4330431) |
| 2 | Os.10908.1.S1_a_at | 12 | Alpha-amylase isozyme 3D (LOC4345814) |
| 3 | Os.50642.1.S1_at | 12 | Heat stress transcription factor B-2a (LOC4336701) |
| 4 | Os.2292.3.S1_x_at | 12 | Heat stress transcription factor A-2a (LOC4334080) |
| 5 | Os.57152.1.S1_at | 11 | Uncharacterized (LOC4329890) |
| 6 | Os.46084.1.S1_at | 11 | Uncharacterized |
| 7 | Os.55479.1.S1_at | 11 | EID1-like F-box protein 3 (LOC9268106) |
| 8 | OsAffx.22585.1.S1_at | 11 | Probable WRKY transcription factor 30 (LOC4340188) |
| 9 | Os.8481.2.S1_at | 11 | Protein OPI10 homology (LOC4332189) |
| 10 | Os.5817.1.S1_at | 11 | dnaJ homology subfamily B member 6 (LOC4340388) |

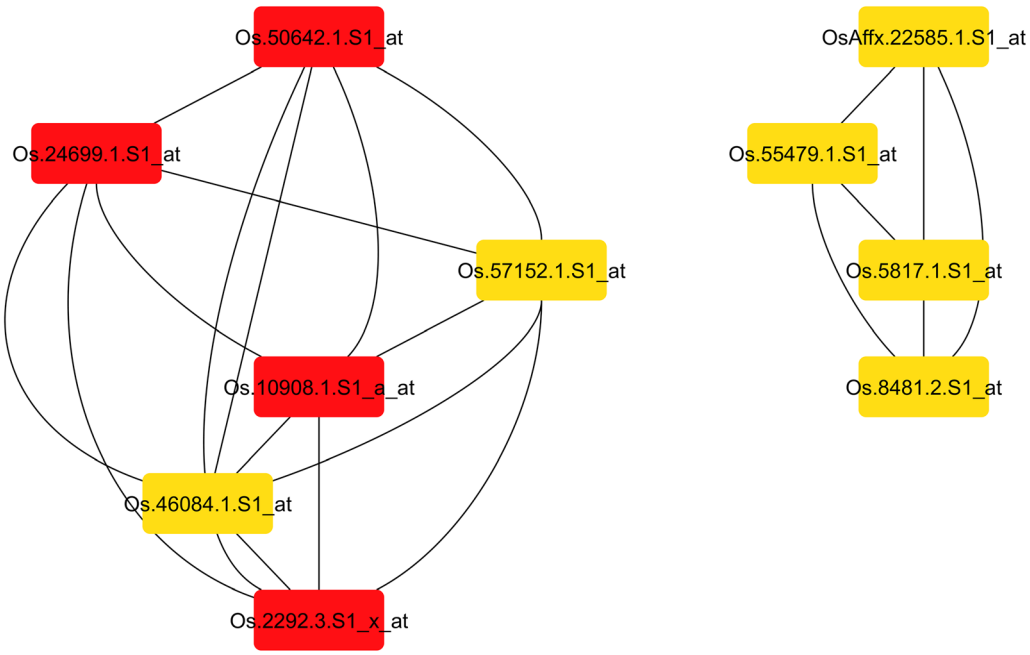


**Figure 7.** Biological networks constructed from expression data of rice under drought stress with hub genes, i.e. most connected genes.

adversely affects the yield of the rice crop. To overcome the problem of drought and develop drought resistant/tolerant rice varieties, the knowledge of the morphological, biochemical and molecular mechanisms involved in rice against drought is very important for rice breeders (Nahar *et al.* 2016).

In conclusion, 5686 genes were consistently expressed differentially in *O. sativa*, among which 2089 genes were upregulated and 3597 genes were found downregulated. The meta-analysis based on gene expression data of stressed rice have shown the fundamental differences between normal and stressed rice which includes DEGs along with their biological function and it may contribute to identify potential candidate genes of abiotic stress. Against drought stress, the proteins like heat shock protein STI, heat shock protein DnaJ, U-box domain containing heat shock protein, hsp20/alpha crystallin family protein were found upregulated to increase the defense mechanism of plants and on the other hand, proteins like retrotransposon protein, diterpene phytoalexin precursor biosynthetic process pathway and ent-kaurene synthase, chloroplast precursor, CAMK_CAMK_like.47 - CAMK includes calcium/calmodulin dependent protein kinases were found downregulated to enhance the plants adaptation during stress. In response to drought stress, intracellular $Ca^{2+}$ levels changes and induce signalling pathways which help plants to cope with the changing environmental conditions. CaM is one of the important proteins that decodes $Ca^{2+}$ signals and regulates activities of diverse proteins. The heat stress transcription factors play a pivotal role in regulating the drought stress condition by regulating heat shock elements/promoters and help the plants to overcome this situation. This study gives a broad view for researchers with respect to the available different microarray dataset which can be used to find out how plants overcome different stresses/diseases. The identified hub genes also provide a platform to develop a drought tolerant rice varieties.

**Acknowledgments**

Authors are thankful to the Director, Motilal Nehru National Institute of Technology Allahabad, for providing research lab and financial support for performing this experiment. The author is thankful to the Ministry of Human Resource and Development, Govt. of India, New Delhi, India for providing scholarship during this tenure. Authors also sincerely thank Mr Manish P. Singh, MNNIT Allahabad for his help in manuscript correction.

## References

Abuqamar S., Luo H., Laluk K., Mickelbart M. V. and Mengiste T. 2009 Crosstalk between biotic and abiotic stress responses in tomato is mediated by the AIM1 transcription factor. *Plant J.* **58**, 347–360.

Atkinson N. J. and Urwin P. E. 2012 The interaction of plant biotic and abiotic stresses: from genes to the field. *J. Exp. Bot.* **63**, 3523–3543.

Basu S., Ramegowda V., Kumar A. and Pereria A. 2016 Plant adaptation to drought stress. F1000Res, **5** (F1000 Faculty Rev) 1554.

Bostock R. M., Pye M. F. and Roubtsova T. V. 2014 Predisposition in plant disease: exploiting the nexus in abiotic and biotic stress perception and response. *Ann. Rev. Phytopathol.* **52**, 517–549.

Chin C., Chen S., Wu H. H., Ho C. W., Ko M. T. and Lin C. Y. 2014 cytoHubba: identifying hub objects and sub-networks from complex interactome. *BMC Syst. Biol.* **8**, S11.

Cramer G. R., Urano K., Delrot S., Pezzotti M. and Shinozaki K. 2011 Effects of abiotic stress on plants: a systems biology perspective. *BMC Plant Biol.* **11**, 163.

Daves M. H., Hilsenbeck S. G., Lau C. C. and Man T. K. 2011 Meta-analysis of multiple microarray datasets reveals a common gene signature of metastasis in solid tumors. *BMC Med. Genomics* **4**, 56–67.

Elbarbary R. A., Lucas B. A. and Maquat L. E. 2016 Retrotransposons as regulators of gene expression. *Science* **351**, aac7247.

Fang F., Pan J., Xu L., Li G. and Wang J. 2015 Identification of potential transcriptomic markers in developing ankylosing spondylitis: a meta-analysis of gene expression profiles. *BioMed Res. Int.* **2015**, 1–6.

Fang Y. and Xiong L. 2015 General mechanisms of drought response and their application in drought resistance improvement in plants. *Cell. Mol. Life Sci.* **72**, 673–689.

Feichtinger J., Thallinger G., McFarlane R. and Larcombe L. 2012 Microarray meta-analysis: from data to expression to biological relationships. In *Computational medicine* (ed. Z. Trajanoski), pp. 59–77. Springer, Vienna.

Hou D. L., Chen L., Liu B., Song L. N. and Fang T. 2014 Identification of common gene networks responsive to radiotherapy in human cancer cells. *Cancer Gene Ther.* **21**, 542–548.

Liu X., Bai X., Wang X. and Chu C. 2006 OsWRKY71, a rice transcription factor, is involved in rice defense response. *J. Plant Physiol.* **164**, 969–979.

Madrid D. T., Cadenas R. N. and Montano A. P. 2012 GeneCodis3: a non-redundant and modular enrichment analysis tool for functional genomics. *Nucleic Acids Res.* **40**, W478–W483.

Mao X., Zhang H., Tian S., Chang X. and Jing R. 2010 TaSnRK2.4, an SNF1-type serine/threonine protein kinase of wheat (*Triticum aestivum* L.), confers enhanced multistress tolerance in Arabidopsis. *J. Exp. Bot.* **61**, 683–696.

Moyano T. C., Vidal E. A., Contreras-López O. and Gutiérrez R. A. 2015 Constructing simple biological networks for understanding complex high-throughput data in plants. In *Plant functional genomics. methods in molecular biology* (ed. J. Alonso and A. Stepanova), vol. 1284. Humana Press, New York.

Nahar S., Kalita J., Sahoo L. and Tanti B. 2016 Morphophysiological and molecular effects of drought stress in rice. *Ann. Plant Sci.* 1409–1416.

Narsai R., Devenish J., Castleden I., Narsai K., Xu L., Shou H. *et al.* 2013 Rice DB: an *Oryza* information portal linking annotation, subcellular location, function, expression, regulation, and evolutionary information for rice and Arabidopsis. *Plant J.* **76**, 1057–1073.

Ramasamy A., Mondry A., Holmes C. C. and Altman D. G. 2008 Key issues in conducting a meta-analysis of gene expression microarray datasets. *PLoS Med.* **5**, e184.

Ramegowda V. and Kumar M. S. 2015 The interactive effects of simultaneous biotic and abiotic stresses on plants: mechanistic understanding from drought and pathogen combination. *J. Plant Physiol.* **176**, 47–54.

Saraswat S., Yadav A. K., Sirohi P. and Singh N. K. 2017 Role of epigenetics in crop improvement: water and heat stress. *J. Plant Biol.* **60**, 231–240.

Shaik R. and Ramakrishna W. 2013 Genes and co-expression modules common to drought and bacterial stress responses in *Arabidopsis* and rice. *PLoS One* **8**, e77261.

Shaik R. and Ramakrishna W. 2014 Machine learning approaches distinguish multiple stress conditions using stress-responsive genes and identify candidate genes for broad resistance in rice. *Plant Physiol.* **164**, 481–495.

Shannon P., Markiel A., Ozier O., Baliga N. S., Wang J. T., Ramage D. *et al*. 2003 Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504.

Tenhaken R. 2014 Cell wall remodeling under abiotic stress. *Front. Plant Sci.* **5**, 771.

Tusher V. G., Tibshirani and Chu G. 2001 Significance analysis of microarrays applied to the ionizing radiation response. *Proc. Natl. Acad. Sci. USA* **98**, 5116–5121.

Velazquez S. F., Guerra R. R. and Calderon L. S. 2011 Abiotic and biotic stress response crosstalk in plants. In *Abiotic stress response in plants – physiological, biochemical and genetic perspectives* (ed. S. Arun Kumar and B. Venkateswarlu), pp. 3–26. InTech Open Access Publisher.

Virdi A. S., Singh S. and Singh P. 2015 Abiotic stress responses in plants: roles of calmodulin-regulated proteins. *Front. Plant Sci.* **6**, 809.

Wang X., Cai X., Xu C., Wang Q. and Dai S. 2016 Drought-responsive mechanisms in plant leaves revealed by proteomics. *Int. J. Mol. Sci.* **17**, 1706–1735.

Yang Z., Chen Y., Fu Y., Yang Y., Zhang Y., Chen Y. and Li D. 2014 Meta-analysis of differentially expressed genes in osteosarcoma based on gene expression data. *BMC Med. Genet.* **15**, 80–87.

You J. and Chan Z. 2015 ROS regulation during abiotic stress responses in crop plants. *Front. Plant Sci.* **6**, 1092.

Yu C., Chen H. H., Tian F., Leach J. E. and He C. 2014 Differentially-expressed genes in rice infected by *Xanthomonas oryzae* pv. *oryzae* relative to a flagellin-deficient mutant reveal potential functions of flagellin in host–pathogen interactions. *Rice* **7**, 20–30.

Zeng H., Xu L., Singh A., Wang H., Du L. and Poovaiah B. W. 2015 Involvement of calmodulin and calmodulin-like proteins in plant responses to abiotic stresses. *Front. Plant Sci.* **6**, 600.

Corresponding editor: H. A. Ranganath