

# Genes and Co-Expression Modules Common to Drought and Bacterial Stress Responses in *Arabidopsis* and Rice

Rafi Shaik, Wusirika Ramakrishna\*

Department of Biological Sciences, Michigan Technological University, Houghton, Michigan, United States of America

## Abstract

Plants are simultaneously exposed to multiple stresses resulting in enormous changes in the molecular landscape within the cell. Identification and characterization of the synergistic and antagonistic components of stress response mechanisms contributing to the cross talk between stresses is of high priority to explore and enhance multiple stress responses. To this end, we performed meta-analysis of drought (abiotic), bacterial (biotic) stress response in rice and *Arabidopsis* by analyzing a total of 386 microarray samples belonging to 20 microarray studies and identified approximately 3100 and 900 DEGs in rice and *Arabidopsis*, respectively. About 38.5% (1214) and 28.7% (272) DEGs were common to drought and bacterial stresses in rice and *Arabidopsis*, respectively. A majority of these common DEGs showed conserved expression status in both stresses. Gene ontology enrichment analysis clearly demarcated the response and regulation of various plant hormones and related biological processes. Fatty acid metabolism and biosynthesis of alkaloids were upregulated and, nitrogen metabolism and photosynthesis was downregulated in both stress conditions. WRKY transcription family genes were highly enriched in all upregulated gene sets while 'CO-like' TF family showed inverse relationship of expression between drought and bacterial stresses. Weighted gene co-expression network analysis divided DEG sets into multiple modules that show high co-expression and identified stress specific hub genes with high connectivity. Detection of consensus modules based on DEGs common to drought and bacterial stress revealed 9 and 4 modules in rice and *Arabidopsis*, respectively, with conserved and reversed co-expression patterns.

**Citation:** Shaik R, Ramakrishna W (2013) Genes and Co-Expression Modules Common to Drought and Bacterial Stress Responses in *Arabidopsis* and Rice. PLoS ONE 8(10): e77261. doi:10.1371/journal.pone.0077261

**Editor:** Shyamal D Peddada, National Institute of Environmental and Health Sciences, United States of America

**Received:** April 28, 2013; **Accepted:** August 30, 2013; **Published:** October 10, 2013

**Copyright:** © 2013 Shaik et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** The authors have no support or funding to report.

**Competing interests:** The authors have declared that no competing interests exist.

\* E-mail: wusirika@mtu.edu

## Introduction

Crop productivity and survival is tightly linked to its environment which is being altered due to climate change, biodiversity loss and degradation of land and freshwater [1] threatening the food security of the world while the food demand is estimated to increase by 70% in 2050 [2,3,4]. According to latest World Agricultural Supply and Demand Estimates (WASDE) report by United States Department of Agriculture (USDA), about 80% of agricultural land is experiencing drought and over 2,000 U.S. counties had been designated as disaster areas [5]. Reflecting the declining environmental conditions, more often than not plants today are exposed simultaneously to multiple stresses resulting in enormous changes in the molecular landscape within the cell. Comprehensive understanding of the regulatory networks that modulate the dynamic adaptive changes in a plant responding to stress is critical to meet future energy needs. Rice and *Arabidopsis* are both model plant organisms representing monocots and dicots respectively. Both the plants have

extensive biological knowledgebase and resources including complete genome sequence and highest number of microarray studies in the plant kingdom. Thus, analysis of stress responsive genes within and between rice and *Arabidopsis* for different kinds of stresses would reveal a number of pivotal attributes spanning across the major plant division, angiosperms.

Advancements in high throughput technologies have resulted in deluge of various kinds of -omic data addressing different aspects of temporal and spatial response in variety of stresses in plants. Microarray technology revolutionized the identification of global transcriptomic changes and today multiple transcriptomic studies exist for the same or related stress conditions. Thus meta-analysis of related microarray studies is increasingly becoming popular to enhance the sensitivity of the hypothesis addressed and validate conclusions [6]. So far, very few meta-analysis studies are available in plant systems [7,8,9,10,11]. Meta-analysis of microarray data from *Arabidopsis* infected with eight different viruses revealed hub genes that are highly connected, organized in modules and are

**Table 1.** Number of microarray studies for each stress condition.

Stress condition	GEO Series	GEO Samples	No. of controls	No. of treatments
RD	5	78	35	43
RB	7	227	61	166
AD	7	114	61	53
AB	5	106	40	66
<b>Total</b>	<b>24</b>	<b>525</b>	<b>197</b>	<b>328</b>

RD: Rice Drought; RB: Rice Bacteria; AD: Arabidopsis Drought; Arabidopsis Bacteria

doi: 10.1371/journal.pone.0077261.t001

central to plant defense response [12]. It is reported that in plants responding to multiple stresses, there exists extensive cross-talk between different stress responses via hormonal signaling pathways [13]. Thus, it is imperative to compare and analyze different kinds of stress responses to find the genes, proteins and metabolites that are common and specific to different kinds of abiotic and biotic stress conditions. Meta-analysis of microarray studies involving samples from a wide range of tissues, developmental stages and different levels of stresses but specific to one stress condition would unravel the universal principles and features related to the stress response. Comparative analysis of such universal molecular profiles from different stresses would allow the identification of unique and shared features. Further, comparison of the stress responsive profiles across diverse plant species would reveal the conserved stress specific mechanisms and uncover orthologous genes that are most critical to the stress response.

Recently, there has been an upsurge in the number of studies reporting global co-expression networks of plants based on genome wide transcriptome data [14,15,16]. A number of tools namely ATTED-II [17], CressExpress [18], RiceArrayNet [19], OryzaExpress [20] and RiceFRIEND [21] based on co-expression networks are available that can be explored to identify novel genes, predict gene functions and characterize gene regulatory networks. A network based analysis in rice identified drought responsive gene modules and found a module with 134 genes specifically associated with both drought tolerant and drought resistant rice varieties [22]. Weighted Gene Co-expression Network Analysis (WGCNA) is one of the latest and popular methodologies to decipher correlation patterns across microarray samples [23]. Implemented in R as a package, WGCNA provides a vast array of functions to detect, analyze and export individual and consensus modules from diverse but related microarray studies. WGCNA has been utilized to detect coexpression modules in *Arabidopsis*, rice, maize, soybean and poplar [14,24,25] and also across species [26].

In this study, we performed large scale comparative transcriptomic analysis via meta-analysis of microarray data on drought and bacterial stress in rice and *Arabidopsis*. To elucidate the cross talk between different stress conditions, knowledge of the expression status of genes involved in stress response is critical. Our analysis revealed the genes that are unique to each stress and those that are shared with other stress conditions. Further, within common genes, we also found genes that were up or downregulated in both stresses

and also genes which showed reversed expression status. Extensive analysis of various gene sets based on Gene Ontologies (GO), KEGG Orthologies (KO) and metabolic pathway analysis unraveled the underlying biological mechanisms related to different stresses. We then performed co-expression network analysis which divided the stress responsive genes into tightly co-expressed modules revealing organization of stress transcriptome.

## Methods

### Selection of stress related microarray studies

Gene Expression Omnibus (GEO) is the central repository for microarray and other forms of high-throughput data [27]. Experiments conducted on the Affymetrix platforms, Rice Genome Array (GPL2025) and *Arabidopsis* ATH1 Genome Array (GPL198) were chosen for this study as they provide extensive gene coverage and are widely used. GEO currently holds 1920 and 9106 samples and 114 and 709 series records (group of related samples) belonging to GPL2025 and GPL198 platforms, respectively. In total, we analyzed 305 and 220 samples of rice and *Arabidopsis*, respectively, belonging to 28 series records. Different series involving samples from various tissues, developmental stages and stress levels were included to identify differentially expressed genes in a wide range of conditions with higher sensitivity and uncover an overarching drought and bacterial stress response profile. The number of selected series, sample records and number of controls and treatments for each stress condition is given in Table 1. Complete list of selected series and sample records including their GEO IDs and brief description is given in Table S1 in File S1.

### Identification of differentially expressed genes

The raw intensity CEL files of the selected samples were downloaded from GEO and intensity values were extracted from the CEL files using the bioconductor package Affy in R [28], quality checked using the package, ArrayQualityMetrics [29] and the samples failing two or more of its quality tests were removed. The samples of each stress were normalized together using Robust Multichip Average (RMA) method [30]. The probes were then matched to their loci based on annotation provided by array element mapping facility at TAIR portal for *Arabidopsis* (<http://www.Arabidopsis.org/portals/expression/microarray/microarrayElementsV2.jsp>) and at

ricechip.org (<http://www.ricechip.org>) for rice. Probes with no match or ambiguously matching multiple loci were discarded. The retained probes and their normalized intensity values were then loaded into oneChannelGUI environment to perform non-specific filtering of probes with relatively small signal distribution using Inter Quartile Range (IQR) filter at most stringent setting (0.5) and probes with very low intensity values (probes below threshold  $\log_2(50)=5.64$  in  $\geq 90\%$  of arrays). An example of resultant distribution of retained probes after filtering is shown in Figure S1.

Differentially expressed genes (DEGs) were identified using Rank Product method [31]. Rank Product is a non-parametric method returning up and down regulated genes, their fold change (FC), p-values and percentage of false predictions (PFP). It was shown to perform better than other methods including significance analysis of microarrays (SAM), Fisher's Inverse  $\chi^2$  test and t-based hierarchical modeling [32] and is widely used for meta-analysis studies combining data sets from different origins of the sample pool to increase the power of identification [6]. We used the function RPadvance of the bioconductor package RankProd [33] which is specifically designed for meta-analysis. The number of permutation tests was set to 250. The function topGene with a PFP cut-off value of  $\leq 0.01$  was used to output differentially expressed genes. Among multiple probes matching the same locus, the probe ID with highest fold change was retained.

The orthologs between rice and *Arabidopsis* were obtained by parsing the gene families reported in GreenPhyIDB [34] which were identified based on analysis of complete proteomes of 16 plant species, cross referencing a number of resources (UniProtKB, Pubmed, InterPro, MEME motifs, KEGG pathways).

### Functional enrichment analysis

Gene ontology analysis was carried out using the Singular Enrichment Analysis (SEA) tool offered by agriGO [35] at default settings of Fisher t-test ( $p < 0.05$ ), False Discovery Rate (FDR) correction by Hochberg method and five minimum number of mapping entries against species specific pre-computed background reference. KEGG orthology (KO) terms associated with a gene correspond to KEGG pathway nodes and BRITE hierarchy nodes [36]. To identify enzymes and proteins encoded by differently expressed genes and their associated metabolic and signaling pathways in each stress condition, we performed enrichment analysis of KO terms and determined the significance based on hypergeometric distribution p-values with  $< 0.05$  cut off value. Further analysis of biological pathways was carried out using the tool Database for Annotation, Visualization and Integrated Discovery (DAVID) v6.7 [37]. Information on transcription factors (TFs) genes in rice and *Arabidopsis* was obtained from the database PInTFDB [38] and analyzed for enrichment of TF families in stress responsive genes.

### Co-expression network analysis

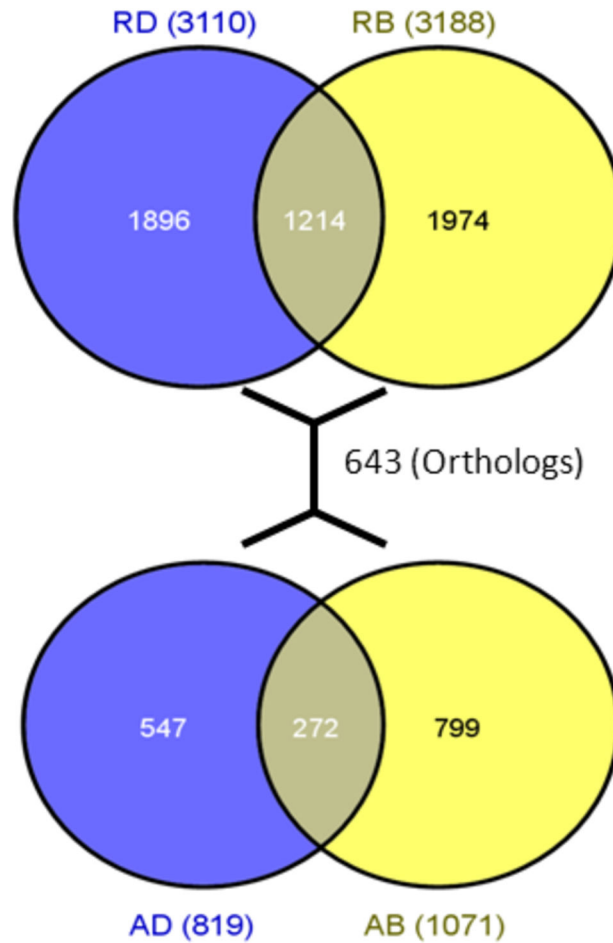
To identify co-expression modules within SRGs, we extracted the normalized, log transformed gene expression values of each stress condition from the microarray

experiments used in meta-analysis and performed Weighted Gene Co-expression Network Analysis (WGCNA) [23]. Briefly, WGCNA procedure calculates Pearson's correlation matrix for all genes, transforms the correlation matrix by raising all values to a power  $\beta$  (soft thresholding as biological networks are small world and scale free [39]), calculates a topological overlap matrix (TOM) from the transformed correlation matrix, converts the topological overlap matrix into a dissimilarity matrix, creates a hierarchical cluster tree based on the dissimilarity matrix, and identifies gene co-expression modules from the hierarchical cluster tree using a dynamic tree cut procedure. The blockwiseModules function of WGCNA package in R was used to generate the modules with powers 8, 6, 14 and 5 for RD, RB, AD and AB, respectively, which best approximate a scale-free topology (model fit  $> 0.8$ ) of the resultant network (Figure S2). For this analysis, module size was 20-30, deep split was set at level 4 and tree merge cut height was 0.15-0.25. Heatmaps were constructed to depict the eigengenes from each identified module. Eigengenes represent a centroid measure of the expression levels of all genes in a cluster. The SRGs common to drought and bacterial stress were analyzed to find consensus modules showing co-expression patterns across stresses using the function blockwiseConsensusModules with the following settings: powers 7 and 10, minimum module size 30 and 15 for rice and *Arabidopsis*, respectively, with the merge cut height set at 0.15.

## Results and Discussion

### Highly conserved expression status of genes common to drought and bacterial stresses

We identified a total of 5084 and 1618 DEGs referred herein as stress responsive genes (SRGs) in rice and *Arabidopsis*, respectively, combining the DEGs in drought and bacterial stresses together that were below PFP  $\leq 0.01$  (Figure 1). Greater than 60% of genes were unique to individual stresses in all cases and AB (*Arabidopsis* Bacteria) had highest percent (~75%) of unique SRGs (799 genes). The number of up and downregulated SRGs are shown in Figure S3A and complete list of genes along with their fold change values is given in Table S2 in File S1. Among the 1214 SRGs common to the stresses studied in rice, majority of the genes were expressed in same direction (72% or 874) with 565 up and 309 downregulated in both drought and bacterial stresses. Similarly, higher number of SRGs (170 out of 272 or 62.5%) common to both stresses studied in *Arabidopsis* was expressed in same direction with 93 and 77 genes up and downregulated, respectively. This finding elucidates that these set of genes and their associated biological processes are altered similarly as part of stress response in a wide range of tissues, developmental stages, stress levels and ecotypes (Table 1A). Among the genes with non-conserved expression pattern, the proportion of genes showing downregulation in drought and upregulation in bacterial stress (255 or 21% of 1214) was higher in rice while upregulation in drought and downregulation in bacterial stress (66 or 24% of 272) was higher in *Arabidopsis* (Figure S3D). Further comparison of PFP values of common SRGs revealed that about 30% (371 out of



**Figure 1. Number of unique and common differentially expressed genes (DEGs) found in rice and *Arabidopsis*.** The number of orthologous genes found between rice and *Arabidopsis* DEGs are also shown. RD: Rice Drought, RB: Rice Bacteria, AD: *Arabidopsis* Drought, AB: *Arabidopsis* Bacteria.

doi: 10.1371/journal.pone.0077261.g001

1214) and 15% (33 out of 272) of rice and *Arabidopsis* genes, respectively, showed very high significance in both the stresses (PFP <0.0001). Especially among common SRGs of rice, ~40% (121/309) of genes that showed conserved upregulation were with PFP <0.0001 in both the stresses.

To further validate the reliability of the identified genes, we compared the DEGs against lists of genes that showed stable expression in multiple studies. None of the 1618 *Arabidopsis* DEGs were part of 39 genes proposed as a superior set of reference genes for the normalization of gene expression by Czechowski et. al [40]. Two of the 5084 rice DEGs, LOC\_Os08g23180 coding for arabinogalactan protein found only in RD DEGs and LOC\_Os02g38920 coding for GAPDH found only in RB DEGs are part of 26 genes reported by Narsai et. al. showing stable expression [41], which probably is due to the limited number of tissues, developmental stages and stress treatments used in their analysis.

The average fold change observed for SRGs was about 1.52, 0.93, 1.28 and 0.99 for RD (Rice Drought), RB (Rice Bacteria), AD (*Arabidopsis* Drought) and AB (*Arabidopsis*

Bacteria) stresses respectively. The number of SRGs with fold change (FC) value  $\geq 1.5$  was higher in drought stress (51% and 26% in RD and AD respectively) and lower in bacterial stress (4% and 3% in RB and AB respectively), majority of which were part of downregulated genes. Especially three genes showed >11 fold downregulation in RD, with LOC\_Os05g47540 annotated as 'CPuORF26 - conserved peptide uORF-containing transcript, expressed' under expressed 20.86 folds. Upstream open reading frames (uORFs) are small open reading frames found in the 5' UTR of mature mRNA which regulate translation of major ORFs (mORFs) that code for transcription factors, signal transduction factors and developmental signal proteins [42]. Multiple studies have reported the involvement of uORFs in translation repression of target genes in response to stress conditions [43]. We found this gene to be downregulated also in RB (FC 1.55). LOC\_Os10g36500 annotated as 'invertase/pectin methylesterase inhibitor family protein' is the second top DEG which was downregulated in both stress conditions (FC 11.34 and 1.20 in RD and RB, respectively). Pectin methylesterase

inhibitors (PMEI) are invertase inhibitor-related defense proteins that play key roles in developmental transitions, wounding, senescence and abiotic stresses [44]. Another gene that was highly downregulated in RD (FC 11.08) but upregulated in RB is LOC\_Os04g39320 annotated as 'expressed protein'. In all the four stresses about 15-20% of SRGs were annotated as just 'expressed protein' or 'protein\_coding' or 'unknown protein' (500, 556, 220, 157 DEGs in RD, RB, AD and AB, respectively) suggesting there are still hundreds of stress responsive genes with little or no functional information. We also found ~1% SRGs (27 and 34 genes in RD and RB, respectively) were annotated as retrotransposon related genes in rice. In *Arabidopsis*, 21 genes showed >4 fold downregulation under drought stress with AT1G22690 annotated as 'gibberellin-responsive protein' and AT5G03350, a legume lectin family protein showing 8.8 and 7.9 FC, respectively.

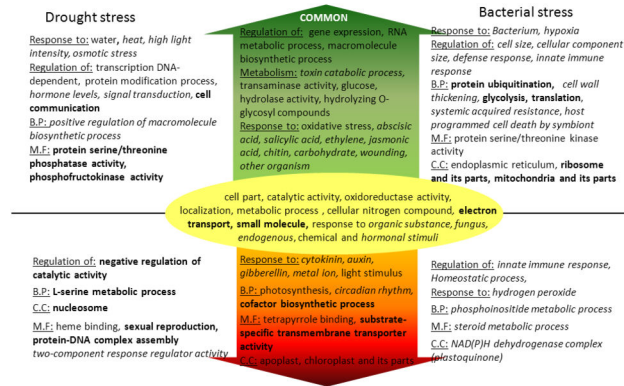
We found 642 orthologous genes between rice and *Arabidopsis* that are involved in stress response (Table S3 in File S1). There were 255 orthologous genes differentially expressed in drought out of which 167 or 65% had their expression status conserved (73 and 94 were up and downregulated, respectively, in both rice and *Arabidopsis* genomes). Similarly, there were 280 orthologous genes differentially expressed in bacterial stress, out of which 211 or 75% had their expression status conserved. Majority of these were upregulated in both the genomes (134 or 63% SRGs). We also analyzed orthologs between AD and RB, and found 72 SRGs with conserved upregulation. On the other hand, there were 102 SRGs with conserved downregulation between AB and RD (Figure S4). There were 9 up and 8 downregulated orthologous genes found in all four stresses. One of these genes is a MYB TF that was highly downregulated, especially in drought (AT2G21650 (FC 3.5) and LOC\_Os01G44390 (FC 5.5)). ARR6 and 7 (two-component response regulators) and their orthologous gene OsRR10 involved in cytokinin response system [45,46] were also downregulated in all stresses. The upregulated genes in all four stress conditions include a NAC TF (AT1G69490 and LOC\_Os03G21060), HAI-1 or highly ABA-induced PP2C gene 1 (AT5G59220 and LOC\_Os05G38290) and a heavy metal-associated domain containing protein (AT5G52760 and LOC\_Os10G38870). Expression of HAI-1 gene was shown to be induced by wound in *Arabidopsis* [47].

### Comparison of significant gene ontology terms of SRG sets distinguished the roles of different hormones and related processes

We found 623 unique GO terms enriched by SRGs in one or more stress conditions (Table S4 in File S1). We analyzed gene sets that are up or downregulated separately for each stress as shown in Figure S3B. Although the number of SRGs in *Arabidopsis* was only 1/3<sup>rd</sup> compared to those found in rice, total number of significant GO terms in *Arabidopsis* is close to rice reflecting the lack of annotation for a number of rice genes. Four way Venn diagram analysis revealed the number of GO terms common and exclusive to same stress (28 terms between RBU and ABU vs. 4 terms between RBU and ADU)

and same species (68 between ADU and ABU vs. 10 between ADU and RDU) were higher than vice versa (Figure S5). The top most significant GO term found in upregulated gene sets were response to water (FDR 5E-11), ribosome (2.9E-37), response to organic substance (3.4E-31) and response to biotic stimulus (2.4E-30) in RDU, RBU, ADU and ABU, respectively and in downregulated sets were catalytic activity (1.7E-24), photosynthesis (1.7E-16), thylakoid (2.6E-18) and response to chemical stimulus (1.3E-15) in RDD, RBD, ADD and ABD, respectively. The terms, 'polysaccharide catabolic process', 'hydrolase activity, hydrolyzing O-glycosyl compounds', 'aromatic amino acid family metabolic process', 'regulation of gene expression', 'transcription factor activity' were significantly enriched in upregulated gene sets, while 'photosynthesis', 'circadian rhythm', 'cofactor biosynthetic process', 'substrate-specific transmembrane transporter activity' were significantly enriched in downregulated gene sets (Figure 2).

Terms related to hormones and their related functions showed clear distinction between the processes that are up or downregulated in a stress response especially in *Arabidopsis*. While terms related to the hormones auxins, cytokinins and gibberellins were downregulated, abscisic acid, salicylic acid, ethylene and jasmonic acid was upregulated both in drought and bacterial stresses. Abscisic acid (ABA) is known to play a central role in abiotic stress response by inducing stomatal closure resulting in reduction of transpiration [48], regulating root growth, ion channels and gene expression [49]. Further, it was found that ABA can have both positive and negative effect on biotic stress signaling [50]. For example, ABA-induced stomatal closure prevented invasion of microbes through open stomata. Thus, recent findings increasingly suggest ABA as a key player in fine-tuning of cross talk between abiotic and biotic stress responses and therefore ABA production can be the crucial factor determining how well a plant responds to multiple stresses [51]. While the terms 'response to ethylene stimulus' and 'response to salicylic acid stimulus' were found both in ADU and ABU, the terms 'ethylene mediated signaling pathway' and 'salicylic acid mediated signaling pathway' were significant only in ABU, which is in agreement with their known functional roles in defense against pathogens and senescence [52,53]. Further, the terms 'host programmed cell death induced by symbiont' and 'systemic acquired resistance (SAR)' mechanisms that are induced by salicylic acid were also found only in ABU. On the other hand, jasmonic acid biosynthetic process was significant only in ADU although jasmonic acid mediated signaling pathway was significant both in ADU and ABU. Jasmonic acid (JA) plays a key role in defense response especially against necrotrophic pathogens and wounding acting antagonistically to salicylic acid which is majorly involved in resistance to biotrophic pathogens [54]. JA also has a role in the formation of antioxidants that regulate ascorbate and glutathione metabolism [55] explaining our observation of its increased synthesis in drought stress. The downregulation of all of the major plant growth and development promoting hormones such as auxins, cytokinins and gibberellins across diverse stress conditions indicates various processes including cell differentiation, chloroplast biogenesis, flowering and



**Figure 2. Summary of significant GO terms found in different stresses.** Terms in green arrow indicate those that are commonly upregulated in drought and bacterial stress, and the terms besides the green arrow indicate those that are specifically upregulated in one stress. Similarly the terms in red arrow and those besides indicate the terms that are downregulated in both stresses and specific to one stress respectively. Terms in yellow oval were found both in up and downregulated gene sets. The terms in bold and those in italics are highly significantly found in rice and *Arabidopsis* respectively. B.P: Biological Process, M.F: Molecular Function, C.C: Cellular Component.

doi: 10.1371/journal.pone.0077261.g002

reproduction [56,57], controlled by them are pushed to backseat while processes related to reprogramming of metabolism, gene expression, balancing of homeostasis and modulation of defense and immunity are given higher priority. The above observations are further supported by a number of terms related to photosynthesis and biosynthesis of its components including 'chloroplast', 'photosystem', 'photosynthetic membrane', 'photosynthesis, light reaction', 'photosynthetic electron transport chain' that were highly enriched in all four downregulated gene sets but none of the upregulated gene sets.

GO terms related to various metabolic processes including carbohydrates, amino acids, proteins, ribosomes, translation and nucleobases were significantly enriched in RBU. Translation is a highly energy expensive process and its regulation via protein phosphorylation, initiation factor isoforms, RNA sequence element interactions, and small RNAs enable cells to rapidly and reversibly control gene expression in response to environmental changes [58]. Upregulation of a number of translation related GO terms in rice under bacterial stress suggests cellular adjustments at translational level upon bacterial infection. The term 'response to water' was highly enriched in RDU (FDR 5E-11) and ADU (FDR 4.6E-19) and the term 'response to water deprivation' was highly enriched in ADU (1.5E-18). A number of terms related to regulation of gene expression and metabolic processes including 'transcription factor activity', 'nucleic acid metabolism', and 'chitin catabolic process' were enriched in three or all of the upregulated gene sets. Both positive and negative regulation of response to stimulus was found in ABU. The term 'negative regulation of defense response' was also significantly enriched in ABU (FDR 8.5E-05). The SRGs associated with the above GO term, EDS1 (Enhanced disease susceptibility 1) and PAD4 (Phytoalexin deficient 4) directly interact and induce salicylic acid biosynthesis in response to biotrophic pathogens [59]. A mutant of EDS1 was found to be disease resistant [60].

### KEGG pathway analysis of SRGs identified upregulation of fatty acid, aromatic amino acid, glucose metabolism and biosynthesis pathways of secondary metabolites

The enriched KEGG orthology (KO) terms in different SRG sets revealed many similar patterns as that of GO analysis that can be seen by the top KO terms and their associated pathways in Table 2. Enrichment of 'jasmonate ZIM domain-containing' proteins (JAZs) and 'auxin responsive GH3 gene family' proteins in the upregulated SRGs of *Arabidopsis* substantiate recent findings that these proteins negatively regulate downstream processes of hormonal activity especially those related to plant growth and development [61,62,63]. On the other hand, KO terms, 'two-component response regulator ARR-A family' involved in negative regulation of cytokinin signaling via phospho relay [64] and 'SAUR family proteins' which are primary auxin-inducible genes involved in auxin transport and organ elongation [65] were enriched in downregulated gene sets of both the stresses. Reactive oxygen species (ROS) have been proposed as a central component of plant adaptation to both biotic and abiotic stresses [66]. Glutathione S-transferase (GST) plays a key role in scavenging ROS and detoxification and is differentially activated by stress-induced plant growth regulators [67]. GST was upregulated in both the stresses and was also part of ADD.

A number of terms related to enzymes involved in biosynthetic pathways of amino acids including 'peroxidase', 'tyrosine aminotransferase' and 'serine O-acetyltransferase' were part of downregulated gene sets (Table S5 in File S1). The KO term 'Cellulose synthase A (CesA)' was highly enriched in RDD. Several studies reported disruption of genes involved in biosynthesis of cellulose enhanced stress tolerance [68,69,70]. As also revealed by GO analysis, the term 'small subunit ribosomal protein S4e' was enriched in RBU and

**Table 2.** Top KO terms and their associated pathways.

KO ID	Definition	Pathway	Ref No.	ADD	ADU	ABD	ABU	
K00430	Peroxidase	Phenylalanine and methane metabolism; Phenylpropanoid biosynthesis	144	11				
K00511	Squalene monooxygenase	Sesquiterpenoid and triterpenoid biosynthesis	6			3		
K00799	Glutathione S-transferase	Glutathione metabolism	46	4	6		7	
K09580	Protein disulfide-isomerase A1	-	7				3	
K13464	Jasmonate ZIM domain-containing protein	Plant hormone signal transduction; Plant-pathogen interaction	20		7			
K14487	Auxin responsive GH3 gene family	Plant hormone signal transduction	18				4	
K14488	SAUR family protein	Plant hormone signal transduction	59	9		5		
K14492	Two-component response regulator ARR-A family	Plant hormone signal transduction	10	4		3		
K14497	Protein phosphatase 2C	Plant hormone signal transduction	9		4			
					<b>RDD</b>	<b>RDU</b>	<b>RBD</b>	<b>RBU</b>
K00815	Tyrosine aminotransferase	Phenylalanine, tyrosine and tryptophan biosynthesis	4		3	3		
K02639	Ferredoxin	Photosynthesis	5			4		
K02987	Small subunit ribosomal protein S4e	Ribosome	3				3	
K03283	Heat shock 70kDa protein 1/8	Spliceosome; MAPK signaling pathway; Protein processing in ER	6		3		4	
K05953	Nicotianamine synthase	-	3	3				
K09874	Aquaporin NIP	-	6	4				
K10999	Cellulose synthase A	Starch and sucrose metabolism	9	6				
K00640	Serine O-acetyltransferase	Cysteine and methionine metabolism	5			3		

doi: 10.1371/journal.pone.0077261.t002

'ferredoxin' involved in photosynthesis was enriched in RBD. Heat shock protein 70 (Hsp70) is one of the most abundant heat shock proteins in eukaryotic cells which bind to hydrophobic patches of partially unfolded proteins preventing protein aggregation [71]. Hsp70 was enriched in both the upregulated gene sets of rice.

The KEGG pathways found significant by the tool DAVID with p-value <0.05 in SRG sets are shown in Figure S6. The pathway 'fatty acid metabolism' was enriched both in RDU and RBU. Plants acclimating to stress modulate membrane fluidity and levels of oleic acid and linolenic acid using lipases facilitating proper functioning of critical integral proteins during stress [72].  $\alpha$ -linolenic acid released under stress from chloroplast membranes is a major parent compound for an array of messenger compounds derived via oxidative modification by ROS [73] including jasmonic acid [74,75]. The pathway ' $\alpha$ -linolenic acid metabolism' was highly significant in ADU and RBU. A number of pathways related to biosynthesis of secondary metabolites were enriched in upregulated sets including biosynthesis of alkaloids from shikimates, purines, histidine, terpenoids and polyketides. Phenylpropanoids, derived from a very limited set of core structures of shikimate pathway are modified by oxygenases, ligases, oxidoreductases and transferases to generate an enormous number of secondary metabolites (>200,000) including lignins, suberins and tannins which contribute substantially to the robustness of plants facing stress [76] and are also implicated in providing nutritional and medicinal benefits for animals and humans due to their potent antioxidant activity [77]. Phenylpropanoid biosynthesis was enriched in drought especially in rice but was found both in up and downregulated gene sets suggesting differential regulation of the enzymes resulting in synthesis of different end-products. The biosynthetic pathway of flavonoids from phenylpropanoid derivatives was enriched in ABD.

Biosynthesis and metabolic pathways of aromatic amino acids, phenylalanine, tyrosine and tryptophan and degradation pathways of lysine, valine, leucine and isoleucine were enriched in upregulated gene sets (ABU, ADU, RBU and RDU). The aromatic amino acids are also synthesized via the shikimate pathway playing crucial roles in plant growth, development, reproduction, defense, and environmental responses [78,79]. Recent reports indicate reduction in starch biosynthesis and accumulation, and increased consumption of storage substances under drought [80,81] resulting in elevated levels of hexose sugars (glucose and fructose) [82]. Our analysis revealed upregulation of starch and sucrose metabolism, glycolysis/gluconeogenesis and pentose phosphate pathway in both drought and bacterial stresses. As observed in GO analysis, a number of pathways related to photosynthesis were enriched in downregulated gene sets including porphyrin and chlorophyll metabolism, carbon fixation in photosynthetic organisms and carotenoid biosynthesis. Similar to the observation of enrichment of GO term 'nucleobase, nucleoside, nucleotide and nucleic acid biosynthetic process' in RDD, 'amino sugar and nucleotide sugar metabolism' pathway was also enriched in RDD.

#### Differential enrichment 'CO-like' TF family members under drought and bacterial stresses

Out of the 82 and 56 known TF families/regulators, 34 (41%) and 38 (67%) were found in one or more gene sets of *Arabidopsis* and rice, respectively (Figure S3C). A comparative list of the number of TFs belonging to each TF family found in different stresses is given in Table S6 in File S1. Among the large TF families, higher numbers of NAC, ERF, AP2-EREBP and C2H2 family members were found in upregulated gene sets while higher numbers of bHLH and MYB-related family

members were found in downregulated gene sets. WRKY TFs were the highest in the upregulated set of bacterial stress both in rice and *Arabidopsis*. WRKY TFs are considered to be at the heart of global regulation of plant immunity by modulating its immediate downstream target genes which include MAP kinases and other TFs [83]. 'CO-like' TF family members were the highest in RBU (17 TFs) and RDD (11) but low in RBD (1) and RDU (3) indicating an inverse expression relationship between drought and bacterial stress. CO (*CONSTANS*) gene and other members of CO-like TF family play an important role in regulation of flowering and act between the circadian clock and genes controlling meristem identity [84]. A high number of HSF (heat shock transcription factor) family members were found in upregulated gene sets of rice. Seven HD-ZIP (heomodomain leucine zipper motif) members were found in RDU only. Out of 16 Tify family members in *Arabidopsis*, seven were found in ADU. Tify is a novel TF family with JAZ motifs, is implicated to play a critical role in jasmonate signaling pathway [85]. Members of this family were reported to be strongly induced under drought conferring improved tolerance to drought and high salinity [86].

### Gene network analysis revealed tightly co-expressed modules of SRG sets

Gene coexpression networks, built using a set of microarray samples as input, can help elucidate tightly coexpressed modules that are a mixture of genes with known and unknown functions, identify hub genes, and candidate genes which can be used as biomarkers [15,87]. Using Weighted Gene Co-expression Network Analysis (WGCNA), we divided SRGs into 11, 10, 5, 8 modules of RD, RB, AD and AB, respectively, excluding a grey color module listing genes that did not significantly co-express with any other group of genes (Table S7 in File S1). The module of each SRG indicated by module color,  $k_{IM}$  (intramodular connectivity), a measure of how well connected or co-expressed a given gene is, with respect to other genes in its module, MM (Module Membership), a measure of module membership correlating its gene expression profile with the Module Eigengene (ME, which is the first principal component of a given module also considered as a representative of the gene expression profile of the module) [23] and p-values are given in Table S8 in File S1.

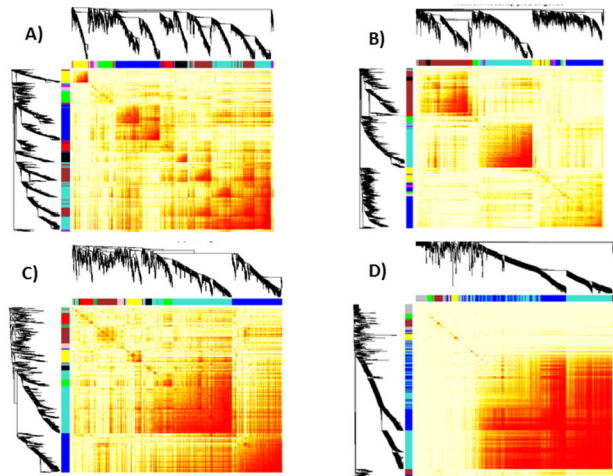
The long length of the dendrogram branches and corresponding intense red color in the heat maps of co-expression modules illustrate high co-expression of SRGs within modules and less co-expression outside the module (Figure 3). We used unsigned correlations so that positively and negatively correlated genes could be grouped into the same module. Yet, a number of modules showed high enrichment of either up or down regulated genes (Table S7 in File S1). For example, the largest module (turquoise) found in RD with 846 SRGs was made up of 663 (78%) downregulated genes while the second largest module (blue) with 763 SRGs was made up of 618 (79%) upregulated genes. We compared the 11 RD modules detected by us against 15 drought-responsive modules of rice found by another recent study using Markov Cluster (MCL) algorithm [22]. Out of those 15 modules, 14 modules were made up of 28-75% of RD SRGs, most of

which significantly overlapped with one of the RD modules (Table S9 in File S1). For example, module 2 found by Zhang et al [22] was made up of 213 genes, out of which 146 (68.5%) were part of SRGs and 116 (90%) of those overlapped with RD turquoise module. The module eigengene (ME) of the RD turquoise module has low values in all drought arrays compared to control indicating that most of the genes are downregulated (green color in the heatmap) (Figure 4A). The top functional terms enriched in this module were predominantly related to photosynthesis. In the blue module, ME has higher values in all drought arrays compared to control indicating that most of the genes are upregulated (red color in the heatmap) (Figure 4B). The top functional term of blue module was 'response to water' followed by protein domains 'dehydrin' and 'LEA'. Late embryogenesis abundant (LEA) proteins are extremely hydrophilic proteins implicated in desiccation tolerance and stabilization of proteins and membranes during drying [88]. The blue module had a very high number of TFs than turquoise (64 compared to 38 TFs) (Table S7 in File S1) although it was made up of less number of genes than turquoise. Majority of blue module TFs were from ERF and NAC families while turquoise had higher number of bZIP and CO-like TFs.

Functional enrichment analysis of each of the co-expression modules revealed a number of significant terms with FDR <0.05 (Table S10 in File S1) especially in *Arabidopsis* which were proportional to their module size. However, in rice, there was large variation in number of significant functional terms compared to module size (Table S7 in File S1). For instance, the RD module brown (size 732) had 83 significant terms but blue (size 763) had only 8 terms with FDR <0.05. Further analysis of these modules revealed higher number of genes annotated as 'expressed protein', 'DUF – Domain of unknown function' in blue module (129, 26 and 260) compared to brown module (96, 9 and 211). There were 51 and 278 genes in blue and brown modules, respectively, with high intramodular connectivity ( $k_{IM}$  value >100), out of which 11 and 31 genes were annotated as 'expressed protein' in blue and brown modules, respectively. These genes would be important candidates for further investigation as they might be playing significant role in stress response.

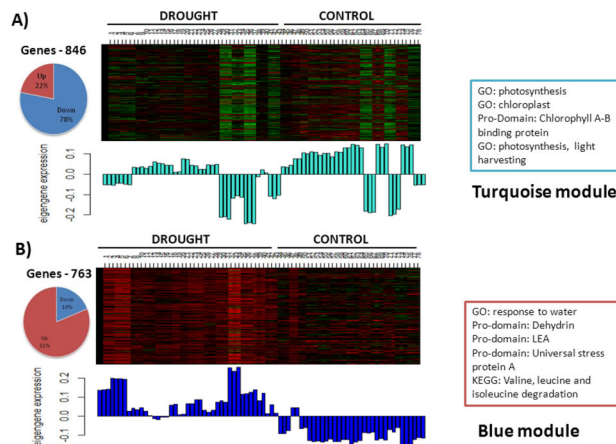
Under same stress, a number of modules in rice and *Arabidopsis* showed relatedness in functionality, indicating conservation of co-expression of functionally related genes across species. The module AD turquoise was related to RD brown with shared terms, response to oxidative stress (GO: 0006979, AD FDR=5.18E-07, RD FDR=5.32E-07) and calcium ion binding. The module AD blue was related to RD turquoise with terms photosynthesis (GO:0015979, AD FDR=9.24E-07, RD FDR=2.0E-19) and other similar terms. AB yellow was related to RB magenta with shared terms, 'aromatic compound biosynthetic process' and 'cellular amino acid biosynthetic process'. RB red with 203 upregulated genes out of 206, had 26 TFs which is double the percent of TFs found in other modules. Most of the TFs in this module belong to WRKY and MYB families with the top gene being a MYB TF, LOC\_Os04g43680. The only downregulated genes in this module are LOC\_Os05g37820 (major facilitator family





**Figure 3. Dendrograms and heatmaps of SRGs divided into tightly co-expressed modules by the R statistical package WGCNA.** A) RB, B) RD, C) AB, and D) AD. The DEGs were clustered based on co-expression patterns as represented by the dendrogram and correlation heat map. Clusters of like-regulated genes are referred to as modules and are indicated by different colors. Grey color represents the genes that could not be assigned to a module. Intensity of red coloring in the heat map indicates strength of correlation between pairs of genes on a linear scale.

doi: 10.1371/journal.pone.0077261.g003



**Figure 4. Heatmaps of turquoise and blue modules in rice under drought stress.** The x-axis represents microarray samples grouped into drought treated and control samples and y-axis represents genes found in the module. Below the heatmap the corresponding module eigengene expression values are shown. The most significant functional terms found in the module are also shown. The number of genes found in each module and the percentage of up and downregulated genes in each module are shown as a pie chart.

doi: 10.1371/journal.pone.0077261.g004

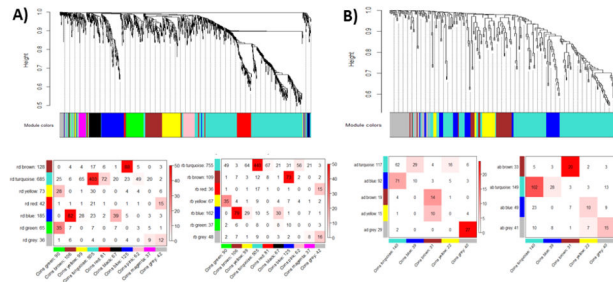
transporter), LOC\_Os09g35010 (dehydration-responsive element-binding protein) and LOC\_Os02g51910 (cytokinin-O-glucosyltransferase 2).

Among the modules found in AD, brown (size 64) had 63 upregulated genes and 22 TF genes (34%), and showed enrichment of 44 functional terms including response to various hormones and endogenous stimuli like water deprivation, salt, cold, temperature and chitin. There were 6 TFs including WRKY33 and WRKY40 in the top 10 genes in this module based on  $k_{IM}$  values. Among AB modules, yellow module with

90% (72 out of 80) downregulated genes and 18.75% of TFs showed enrichment of a number of terms related to secondary metabolic process including biosynthesis of aromatic compounds, flavonoids and phenylpropanoids.

### Consensus co-expression modules of drought and bacterial stresses

The expression profiles of the SRGs common to drought and bacterial stresses was utilized to detect consensus modules



**Figure 5. Clustering dendrogram of genes and consensus modules found in A) rice and B) Arabidopsis.** The correspondence between consensus modules and modules found individually in drought and bacterial stress based on the expression values of the common genes are also shown as a table. Each row of the table corresponds to individual stress specific module (labeled by color as well as text along with the number of genes in the module), and each column corresponds to one consensus module. Numbers in the table indicate gene counts in the intersection of the corresponding modules. Coloring of the table encodes  $-\log(p)$ , with  $p$  being the Fisher's exact test  $p$ -value for the overlap of the two modules. The stronger red color indicates more significant overlap.

doi: 10.1371/journal.pone.0077261.g005

**Table 3.** List of consensus co-expression modules found in each stress gene set.

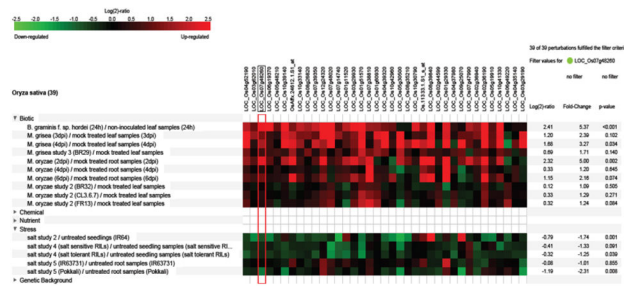
Rice Modules	Module Size	No. of TFs	RD Down/Up	RB Down/Up	Conservation of gene expression (%)
<b>Brown</b>	106	10	8/98	4/102	94.34
<b>Red</b>	81	2	72/9	75/6	93.83
<b>Yellow</b>	99	8	32/67	39/60	90.91
<b>Turquoise</b>	505	19	396/109	363/142	82.38
<b>Blue</b>	125	5	114/11	88/37	72.80
<b>Green</b>	90	5	17/73	33/57	48.89
<b>Pink</b>	62	2	57/5	28/34	46.77
<b>Black</b>	67	7	62/5	2/65	7.46
<b>Magenta</b>	37	3	37/0	1/36	2.70
Arabidopsis Modules			AD Down/Up	AB Down/Up	
<b>Turquoise</b>	140	14	63/77	98/42	53.57
<b>Blue</b>	39	6	22/17	20/19	64.10
<b>Brown</b>	31	7	2/29	1/30	90.32
<b>Yellow</b>	22	0	7/15	0/22	68.18

doi: 10.1371/journal.pone.0077261.t003

that would reveal sets of genes with similar co-expression patterns in both the stresses. We found 9 and 4 consensus modules (excluding grey module for genes that did not co-express with others) based on 1214 and 272 SRGs differentially expressed both in drought and bacterial stress in rice and *Arabidopsis*, respectively (Figure 5 and Table 3). The color coded tables below the dendrograms in Figure 5 show the correspondence between consensus modules and modules found individually in drought and bacterial stress revealing several of the modules with preserved module structure. Consensus modules brown, turquoise and blue in rice and turquoise and brown in *Arabidopsis* showed significant overlap with their counterparts indicating the module structure in

drought and bacterial stress to be very similar. A complete list of SRGs with their consensus modules and  $k_{ME}$  values which is a measure of module membership by correlating its gene expression profile with its module eigengene is given in Table S11 in File S1.

Among the 9 consensus modules found in rice, three modules showed conservation of differential expression in >90% of genes. Of these, module red contains majorly downregulated genes while brown contains upregulated genes. Red module was enriched with terms 'ribonucleoprotein' and 'rotamase' and brown was enriched with terms 'valine, leucine and isoleucine degradation' and 'NAM protein'. Interestingly, two modules (magenta and black) showed >92% of genes with



**Figure 6. Gene expression profile of rice consensus module magenta under fungal and salt stresses using the tool Genevestigator.** The heatmap shows color coded values based on Log<sub>2</sub>-ratio of test/control samples in different studies. A brief description of test/control samples including tissue, treatment and cultivar is also given. The top TF gene WRKY47 (LOC\_Os07g48260) is highlighted with a red box and the corresponding log<sub>2</sub>-ratio, fold change and p-values across different microarray studies are shown.

doi: 10.1371/journal.pone.0077261.g006

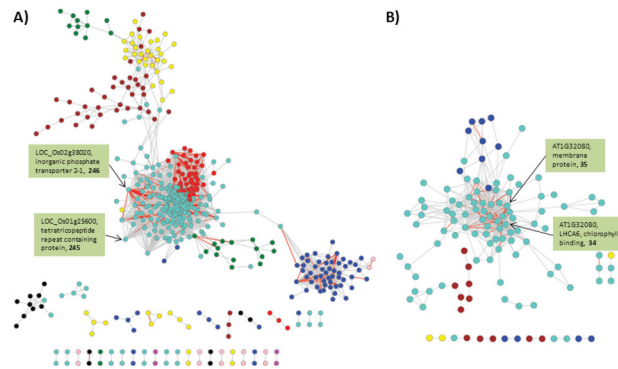
reversed expression status suggesting that these set of genes possibly play a co-ordinated role specific to the stress condition. Most of the genes in these modules were downregulated under drought but upregulated under bacterial stress elucidating the differences in abiotic and biotic stress responses. We further investigated if this trend can be observed in other stresses, using the tool Genevestigator [89]. Analysis of the expression profile of genes in magenta module under salt (3 microarray studies) and fungal (3 studies and 2 pathogens, *B. graminis* and *M. oryzae*) stress conditions identified most of the genes to be highly up and downregulated under fungal and salt stresses, respectively (Figure 6). Magenta color module showed enrichment of GO terms 'electron transport' and 'oxidoreductase activity' and black was significantly enriched in the following protein domains: 'Glycoside hydrolase, chitinase active site', 'DNA-binding WRKY', 'Bet v I allergen' and 'VQ'. Bet v 1 belongs to plant pathogenesis-related proteins (PR-10) family that is involved in plant development and defense systems via interactions with plant hormones [90]. VQ is a small motif found only in plants. A recent study has shown that VQ motif containing proteins act as co-activators of WRKY33 in *Arabidopsis* as part of plant defense response [91,92]. The gene LOC\_Os01g61080 (WRKY24) which is an ortholog of WRKY33 of *Arabidopsis* was also part of black module. Occurrence of VQ motif containing genes (LOC\_Os05g44270 and LOC\_Os03g20440) and WRKY24 in the same module and upregulation of all three under bacterial stress and downregulation under drought stress suggests these genes play a similar role in rice.

Among the *Arabidopsis* consensus modules, brown and yellow were made up of mostly upregulated genes. In brown module, 28 (90%) out of 31 SRGs were upregulated in both the stresses. It contained four WRKY TF genes including WRKY33 (AT2G38470) which was also found in rice consensus module black. The top three SRGs of brown module based on  $k_{IM}$  values are AT3G23250 (MYB15), AT2G22880 (VQ motif-containing protein) and AT3G25780 (allene oxide cyclase 3), which is one of the enzymes involved in jasmonic acid biosynthesis. The top three SRGs found in yellow module are AT5G67340 (armadillo/beta-catenin repeat family protein)

which functions in ubiquitin-protein ligase activity, AT4G01700 (chitinase family protein) and AT5G50200 (wound-responsive gene 3), which encodes a high-affinity nitrate transporter.

We further analyzed the consensus co-expression modules by constructing a network based on co-expressed genes with high absolute Pearson correlation coefficient ( $r > 0.8$ ) in both drought and bacterial stresses. There were 16,576 edges between 585 co-expressed genes in rice (Figure 7A and Table S12A in File S1). One of the top edges was between LOC\_Os02g43790, an ethylene-responsive TF and LOC\_Os02g41510, a MYB TF with  $r > 0.98$  in both stresses. Color coding of nodes in network with their consensus module color showed clear grouping of genes from the same module with high number of intra-modular edges. For example, majority of blue consensus module genes had edges within the group and was largely isolated from all other modules. This indicates that these set of genes are co-regulated and exhibit stress specific co-functionality. Gene ontology analysis revealed enrichment of terms 'cytoplasmic membrane-bounded vesicle' (FDR=0.003) and 'endopeptidase activity' (FDR=0.02). Interestingly, these blue module genes were connected to the largest module (turquoise) via only one gene, LOC\_Os05g09724 a HAD (haloacid dehalogenase) superfamily phosphatase which are involved in diverse housekeeping and secondary metabolism activities [93]. Red module showed the highest percent of genes (71 out of 81 or 87.6%) with a number of edges having  $r > 0.8$  in both the stresses. Black module had 15 genes (out of 67 or 22%) including 5 TFs with edges showing  $r > 0.8$ , all of which showing non-conserved expression status between drought and bacterial stresses. In *Arabidopsis*, there were 509 edges between 119 genes showing  $r > 0.8$  in both stresses. Color coding the nodes with consensus module colors revealed that most of the edges were between genes of turquoise module (Figure 7B and Table S12B in File S1). The top most co-expressed genes were AT3G51420 (Strictosidine synthase-like 4) and AT1G70760 (Chlororespiratory reduction 23) with  $r > 0.98$  in both the stresses.

Out of the 642 orthologous SRGs between rice and *Arabidopsis*, 92 were part of 585 genes showing high



**Figure 7. Coexpression network of SRGs common to drought and bacterial stresses.** (A) rice (B) *Arabidopsis*. Nodes are color coded based on consensus modules found by WGCNA. Edges are constructed between genes with correlation coefficient ( $r$ )  $> 0.8$ . The edges with  $r > 0.8$  are shown in red. The gene IDs, description and number of edges of top 2 nodes in each of the networks are shown in green colored boxes.

doi: 10.1371/journal.pone.0077261.g007

coexpression in rice and 49 were part of 119 genes showing high coexpression in *Arabidopsis*. Out of these, 15 orthologs were part of genes showing high coexpression in both rice and *Arabidopsis* (Table 4) including four genes (PEX11B, ACA1, CYCP2 and CRR23) that showed conserved downregulation in both stresses in both the species. The genes peroxisomal biogenesis factor 11 (PEX11) and cyclin P2 had high number of co-expression edges in both the stresses. Four genes including two TFs (MYB and NAC) were upregulated in three or more stress genesets. The orthologous genes LOC\_Os01G72530 and AT1G76650 coding for calcium sensor proteins showed conserved upregulation in both the stresses. AT1G76650 was previously shown to be involved in abiotic and biotic stresses such as wound and metal stress response [94,95].

## Conclusion

In this study, we performed meta-analysis of microarray studies and identified differentially expressed genes in rice and

*Arabidopsis* from a wide variety of samples under drought and bacterial stresses. This type of approach enhances sensitivity in the identification of important stress response genes that could be missed by studies that are limited to specific tissue or developmental stage or level of stress. Comparative analysis of the DEGs identified common stress responsive genes between stresses and across species. Functional enrichment analysis revealed the biological processes, cellular pathways and transcription factor families that are commonly and exclusively altered under different stresses. The knowledge gained in this study on various molecular mechanisms like biosynthesis of secondary metabolites and stress specific roles of plant hormones vastly adds on to our understanding of stress response and its regulation. Weighted gene co-expression network analysis divided genes into individual and consensus modules and revealed sets of genes with conserved and reversed expression status. A number of genes with high connectivity, conserved expression but with poor annotation were identified. We propose these genes as potential candidates for stress response engineering.

**Table 4.** Orthologous SRGs with high co-expression edges and their expression status in drought and bacterial stresses.

TAIR ID (AT)	TAIR annotation	MSU ID (LOC_Os)	MSU annotation	No. of edges (rice)	No. of edges (Arabidopsis)	AD	AB	RD	RB
3G47430	PEX11B	04G45210	Peroxisomal biogenesis factor 11	220	28	Down	Down	Down	Down
4G15440	HPL1 (Hydroperoxide lyase 1); electron carrier/ heme binding / iron ion binding / monooxygenase	02G02000	Cytochrome P450	217	3	Up	Down	Down	Down
3G52720	ACA1 (Alpha carbonic anhydrase 1); carbonate dehydratase/ zinc ion binding	02G33030	Bifunctional monodehydroascorbate reductase and carbonic anhydrasenelectarin-3	197	4	Down	Down	Down	Down
3G21870	CYCP2;1 (cyclin p2;1); cyclin- dependent protein kinase	04G53680	Cyclin	184	14	Down	Down	Down	Down
1G42550	PMI1 (Plastid movement impaired1)	09G38090	Expressed protein	130	15	Up	Down	Down	Down
1G19150	LHCA6; chlorophyll binding	09G26810	Chlorophyll A-B binding protein	116	34	Up	Down	Down	Down
4G16146	Protein_coding	03G19220	Expressed protein	28	2	Up	Down	Down	Down
5G22580	Protein_coding	11G05290	Stress responsive A/B Barrel domain containing protein	24	1	Down	Up	Down	Down
5G13930	TT4 (Transparent testa 4); naringenin-chalcone synthase	11G32650	Chalcone synthase	14	1	Up	Down	Down	Down
3G23250	MYB15 (MYB domain protein 15); DNA binding / transcription factor	02G41510	MYB family transcription factor	7	2	Up	Up	Down	Up
1G76650	Calcium-binding EF hand family protein	01G72530	OsCML31 - Calmodulin-related calcium sensor	6	1	Up	Up	Up	Up
3G04070	Anac047 ( <i>Arabidopsis</i> NAC domain containing protein 47); transcription factor	03G21060	No apical meristem protein	4	2	Down	Up	Up	Up
5G22630	ADT5 (arogenate dehydratase 5)	04G33390	Prephenate dehydratase domain containing protein	3	2	Up	Up	Down	Up
1G70760	CRR23 (chlororespiratory reduction 23)	05G28090	Expressed protein	1	33	Down	Down	Down	Down
4G23290	Protein kinase family protein	07G35370	TKL_IRAK_DUF26-1c.15 - DUF26 kinase	1	21	Down	Down	Down	Up

doi: 10.1371/journal.pone.0077261.t004

## Supporting Information

**Figure S1. Histogram plot of intensities from probe sets of rice under bacterial stress studies before and after inter-quartile range filtering and intensity filtering.**  
(TIF)

**Figure S2. Analysis of network topology for various soft-thresholding powers.** Y-axis indicates scale-free fit index as a function of the soft-thresholding power (x-axis). The red line indicates soft power at which the scale-free fit index cut-off value 0.8 is reached.  
(TIF)

**Figure S3. Distribution of up and downregulated genes and GO terms in different SRG sets.** A) Number of up and downregulated genes found in each stress B) Number of significant GO terms and C) Number of transcription factor (TFs) genes found in up and downregulated genes of each stress. D) Number of common SRGs showing conserved gene expression status between drought and bacterial stress. Both Down and Both Up indicate genes with conserved expression status and, D (Drought) Down - B (Bacteria) Up and D Up - B Down indicate genes with non-conserved expression pattern.  
(TIF)

**Figure S4. Conservation of expression status of orthologous SRGs between rice and *Arabidopsis*.**  
(TIF)

**Figure S5. Four way Venn diagram comparing significant GO terms found in A) up and B) downregulated SRG sets.**  
(TIF)

**Figure S6. Significant KEGG pathways identified by the functional enrichment analysis tool DAVID in different SRG sets.** X-axis shows fold enrichment of the pathway by

comparing number of genes of a pathway found in SRG set to total number of genes in pathway found in the genome. Pathways found significant ( $p$ -value  $< 0.05$ ) in A) upregulated SRG sets and B) downregulated SRG sets. RDU: Rice Drought Up, RBU: Rice Bacteria Up, ADU: Arabidopsis Drought Up, ABU: Arabidopsis Drought Up, RDD: Rice Drought Down, RBD: Rice Bacteria Down, ADD: Arabidopsis Drought Down, ABD: Arabidopsis Drought Down.  
(TIF)

**File S1.** Description GEO series and samples. Table S2: List of differentially expressed genes. Table S3: Orthologous DEGs between rice and Arabidopsis. Table S4: List of significant GO terms. Table S5: List of significant KEGG orthology (K.O) terms found in various up and downregulated gene sets below hypergeometric  $p$ -value  $\leq 0.05$ . Table S6: Number of TFs belonging to different TF families in in up and downregulated stress gene sets. Table S7: List of co-expression modules found in each stress gene set. Table S8: Modules of each SRG set along with their kIM (intramodular connectivity), MM (Module Membership) and  $p$ -values. Table S9: Comparison of RD modules against those detected by Zhang et.al. (2012) by MCL. Table S10: List of significant functional terms associated with genes in different co-expression modules. Table S11: List of DEGs in consensus modules along with their module membership (kME) values, z-score and associated  $p$ -values. Table S12: Common genes between drought and bacterial stresses showing correlation coefficient  $r > 0.8$ .  
(XLSX)

## Author Contributions

Conceived and designed the experiments: WR RS. Performed the experiments: RS. Analyzed the data: RS WR. Contributed reagents/materials/analysis tools: WR. Wrote the manuscript: RS WR.

## References

- Foley JA, Ramankutty N, Brauman KA, Cassidy ES, Gerber JS et al. (2011) Solutions for a cultivated planet. *Nature* 478: 337-342. doi: 10.1038/nature10452. PubMed: 21993620.
- Godfray HC (2011) Food for thought. *Proc Natl Acad Sci U S A* 108: 19845-19846. doi:10.1073/pnas.1118568109. PubMed: 22123955.
- Tester M, Langridge P (2010) Breeding technologies to increase crop production in a changing world. *Science* 327: 818-822. doi:10.1126/science.1183700. PubMed: 20150489.
- Reynolds MP, Hellin J, Govaerts B, Kosina P, Sonder K et al. (2012) Global crop improvement networks to bridge technology gaps. *J Exp Bot* 63: 1-12. doi:10.1093/jxb/err241. PubMed: 21926090.
- WASDE (2012) US Drought 2012: Farm and Food Impacts. Available: <http://www.ers.usda.gov/topics/in-the-news/us-drought-2012-farm-and-food-impacts.aspx>.
- Tseng GC, Ghosh D, Feingold E (2012) Comprehensive literature review and statistical considerations for microarray meta-analysis. *Nucleic Acids Res*, 40: 3785-99. PubMed: 22262733.
- Ghanekar R, Srinivasasainagendra V, Page GP (2008) Cross-chip probe matching tool: A web-based tool for linking microarray probes within and across plant species. *Int J Plants Genomics*, 2008: 451327. PubMed: 18949054
- Meier S, Bastian R, Donaldson L, Murray S, Bajic V et al. (2008) Co-expression and promoter content analyses assign a role in biotic and abiotic stress responses to plant natriuretic peptides. *Bmc. Plant Biol* 8: 24.
- Adie BA, Pérez-Pérez J, Pérez-Pérez MM, Godoy M, Sánchez-Serrano JJ et al. (2007) ABA is an essential signal for plant resistance to pathogens affecting JA biosynthesis and the activation of defenses in Arabidopsis. *Plant Cell* 19: 1665-1681. doi:10.1105/tpc.106.048041. PubMed: 17513501.
- Finka A, Mattoo RU, Goloubinoff P (2011) Meta-analysis of heat- and chemically upregulated chaperone genes in plant and human cells. *Cell Stress Chaperones* 16: 15-31. doi:10.1007/s12192-010-0216-8. PubMed: 20694844.
- Cohen D, Bogeat-Triboulot MB, Tisserant E, Balzergue S, Martin-Magniette ML et al. (2010) Comparative transcriptomics of drought responses in *Populus*: a meta-analysis of genome-wide expression profiling in mature leaves and root apices across two genotypes. *Bmc Genomics* 11: 630. doi:10.1186/1471-2164-11-630. PubMed: 21073700.
- Rodrigo G, Carrera J, Ruiz-Ferrer V, del Toro FJ, Llave C et al. (2012) A meta-analysis reveals the commonalities and differences in Arabidopsis thaliana response to different viral pathogens. *PLOS ONE* 7: e40526. doi:10.1371/journal.pone.0040526. PubMed: 22808182.
- Seki M, Narusaka M, Ishida J, Nanjo T, Fujita M et al. (2002) Monitoring the expression profiles of 7000 Arabidopsis genes under drought, cold and high-salinity stresses using a full-length cDNA

- microarray. *The Plant journal : for cell and molecular biology* 31: 279-292.
14. Downs GS, Bi YM, Colasanti J, Wu W, Chen X et al. (2013) A developmental transcriptional network for *Zea mays* defines coexpression modules. *Plant Physiol*.
  15. Ficklin SP, Luo F, Feltus FA (2010) The association of multiple interacting genes with specific phenotypes in rice using gene coexpression networks. *Plant Physiol* 154: 13-24. doi:10.1104/pp.110.159459. PubMed: 20668062.
  16. Mochida K, Uehara-Yamaguchi Y, Yoshida T, Sakurai T, Shinozaki K (2011) Global landscape of a co-expressed gene network in barley and its application to gene discovery in Triticeae crops. *Plant Cell Physiol* 52: 785-803. doi:10.1093/pcp/pcr035. PubMed: 21441235.
  17. Obayashi T, Hayashi S, Saeki M, Ohta H, Kinoshita K (2009) ATTED-II provides coexpressed gene networks for Arabidopsis. *Nucleic Acids Res* 37: D987-D991. doi:10.1093/nar/gkn807. PubMed: 18953027.
  18. Srinivasasainagendra V, Page GP, Mehta T, Coulibaly I, Loraine AE (2008) CressExpress: a tool for large-scale mining of expression data from Arabidopsis. *Plant Physiol* 147: 1004-1016. doi:10.1104/pp.107.115535. PubMed: 18467456.
  19. Lee TH, Kim YK, Pham TT, Song SI, Kim JK et al. (2009) RiceArrayNet: a database for correlating gene expression from transcriptome profiling, and its application to the analysis of coexpressed genes in rice. *Plant Physiol* 151: 16-33. doi:10.1104/pp.109.139030. PubMed: 19605550.
  20. Hamada K, Hongo K, Suwabe K, Shimizu A, Nagayama T et al. (2011) OryzaExpress: an integrated database of gene expression networks and omics annotations in rice. *Plant Cell Physiol* 52: 220-229. doi:10.1093/pcp/pcq195. PubMed: 21186175.
  21. Sato Y, Namiki N, Takehisa H, Kamatsuki K, Minami H et al. (2013) RiceFRIEND: a platform for retrieving coexpressed gene networks in rice. *Nucleic Acids Res* 41: D1214-D1221. doi:10.1093/nar/gks1122. PubMed: 23180784.
  22. Zhang L, Yu S, Zuo K, Luo L, Tang K (2012) Identification of gene modules associated with drought response in rice by network-based analysis. *PLOS ONE* 7: e33748. doi:10.1371/journal.pone.0033748. PubMed: 22662107.
  23. Langfelder P, Horvath S (2008) WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 9: 559. doi:10.1186/1471-2105-9-559. PubMed: 19114008.
  24. Weston DJ, Karve AA, Gunter LE, Jawdy SS, Yang X et al. (2011) Comparative physiology and transcriptional networks underlying the heat shock response in *Populus trichocarpa*, *Arabidopsis thaliana* and *Glycine max*. *Plant Cell Environ* 34: 1488-1506. doi:10.1111/j.1365-3040.2011.02347.x. PubMed: 21554326.
  25. Childs KL, Davidson RM, Buell CR (2011) Gene Coexpression Network Analysis as a Source of Functional Annotation for Rice Genes. *PLOS ONE* 6: e22196. PubMed: 21799793.
  26. Ficklin SP, Feltus FA (2011) Gene coexpression network alignment and conservation of gene modules between two grass species: maize and rice. *Plant Physiol* 156: 1244-1256. doi:10.1104/pp.111.173047. PubMed: 21606319.
  27. Barrett T, Troup DB, Wilhite SE, Ledoux P, Rudnev D et al. (2007) NCBI GEO: mining tens of millions of expression profiles—database and tools update. *Nucleic Acids Res* 35: D760-D765. doi:10.1093/nar/gkl887. PubMed: 17099226.
  28. Gautier L, Cope L, Bolstad BM, Irizarry RA (2004) affy—analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics* 20: 307-315. doi:10.1093/bioinformatics/btg405. PubMed: 14960456.
  29. Kauffmann A, Gentleman R, Huber W (2009) arrayQualityMetrics—a bioconductor package for quality assessment of microarray data. *Bioinformatics* 25: 415-416. doi:10.1093/bioinformatics/btn647. PubMed: 19106121.
  30. Irizarry RA, Bolstad BM, Collin F, Cope LM, Hobbs B et al. (2003) Summaries of Affymetrix GeneChip probe level data. *Nucleic Acids Res* 31: e15. doi:10.1093/nar/gng015. PubMed: 12582260.
  31. Breitling R, Armengaud P, Amtmann A, Herzyk P (2004) Rank products: a simple, yet powerful, new method to detect differentially regulated genes in replicated microarray experiments. *FEBS Lett* 573: 83-92. doi:10.1016/j.febslet.2004.07.055. PubMed: 15327980.
  32. Hong F, Breitling R (2008) A comparison of meta-analysis methods for detecting differentially expressed genes in microarray experiments. *Bioinformatics* 24: 374-382. doi:10.1093/bioinformatics/btm620. PubMed: 18204063.
  33. Hong F, Breitling R, McEntee CW, Wittner BS, Nemhauser JL et al. (2006) RankProd: a bioconductor package for detecting differentially expressed genes in meta-analysis. *Bioinformatics* 22: 2825-2827. doi:10.1093/bioinformatics/btl476. PubMed: 16982708.
  34. Rouard M, Guignon V, Aluome C, Laporte MA, Droc G et al. (2011) GreenPhylDB v2.0: comparative and functional genomics in plants. *Nucleic Acids Res* 39: D1095-D1102. doi:10.1093/nar/gkq847. PubMed: 20864446.
  35. Du Z, Zhou X, Ling Y, Zhang Z, Su Z (2010) agriGO: a GO analysis toolkit for the agricultural community. *Nucleic Acids Res* 38: W64-W70. doi:10.1093/nar/gkq310. PubMed: 20435677.
  36. Mao XZ, Cai T, Olyarchuk JG, Wei LP (2005) Automated genome annotation and pathway identification using the KEGG Orthology (KO) as a controlled vocabulary. *Bioinformatics* 21: 3787-3793. doi:10.1093/bioinformatics/bti430. PubMed: 15817693.
  37. Huang DW, Sherman BT, Lempicki RA (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 4: 44-57. PubMed: 19131956.
  38. Pérez-Rodríguez P, Riaño-Pachón DM, Corrêa LG, Rensing SA, Kersten B et al. (2010) PlnTFDB: updated content and new features of the plant transcription factor database. *Nucleic Acids Res* 38: D822-D827. doi:10.1093/nar/gkp805. PubMed: 19858103.
  39. Albert R (2007) Network inference, analysis, and modeling in systems biology. *Plant Cell* 19: 3327-3338. doi:10.1105/tpc.107.054700. PubMed: 18055607.
  40. Czechowski T, Stitt M, Altmann T, Udvardi MK, Scheible WR (2005) Genome-wide identification and testing of superior reference genes for transcript normalization in Arabidopsis. *Plant Physiol* 139: 5-17. doi:10.1104/pp.105.063743. PubMed: 16166256.
  41. Narsai R, Ivanova A, Ng S, Whelan J (2010) Defining reference genes in *Oryza sativa* using organ, development, biotic and abiotic transcriptome datasets. *Bmc. Plant Biol* 10: 56.
  42. Hayden CA, Jorgensen RA (2007) Identification of novel conserved peptide uORF homology groups in Arabidopsis and rice reveals ancient eukaryotic origin of select groups and preferential association with transcription factor-encoding genes. *BMC Biol* 5: 32. doi:10.1186/1741-7007-5-32. PubMed: 17663791.
  43. Jorgensen RA, Dorantes-Acosta AE (2012) Conserved Peptide Upstream Open Reading Frames are Associated with Regulatory Genes in Angiosperms. *Front Plants Science* 3: 191.
  44. An SH, Sohn KH, Choi HW, Hwang IS, Lee SC et al. (2008) Pepper pectin methylesterase inhibitor protein CaPMEI1 is required for antifungal activity, basal disease resistance and abiotic stress tolerance. *Planta* 228: 61-78. doi:10.1007/s00425-008-0719-z. PubMed: 18327607.
  45. Tsai YC, Weir NR, Hill K, Zhang WJ, Kim HJ et al. (2012) Characterization of Genes Involved in Cytokinin Signaling and Metabolism from Rice. *Plant Physiol* 158: 1666-1684. doi:10.1104/pp.111.192765. PubMed: 22383541.
  46. Hwang I, Sheen J (2001) Two-component circuitry in Arabidopsis cytokinin signal transduction. *Nature* 413: 383-389. doi:10.1038/35096500. PubMed: 11574878.
  47. Zhang J, Li X, He Z, Zhao X, Wang Q et al. (2012) Molecular character of a phosphatase 2C (PP2C) gene relation to stress tolerance in Arabidopsis thaliana. *Mol Biol Rep* 40: 2633-44. PubMed: 23268310.
  48. Pantin F, Monnet F, Jannaud D, Costa JM, Renaud J et al. (2013) The dual effect of abscisic acid on stomata. *New Phytol* 197: 65-72. doi:10.1111/nph.12013. PubMed: 23106390.
  49. Duan L, Dietrich D, Ng CH, Chan PM, Bhalerao R et al. (2013) Endodermal ABA Signaling Promotes Lateral Root Quiescence during Salt Stress in Arabidopsis Seedlings. *Plant Cell*.
  50. Melotto M, Underwood W, Koczan J, Nomura K, He SY (2006) Plant stomata function in innate immunity against bacterial invasion. *Cell* 126: 969-980. doi:10.1016/j.cell.2006.06.054. PubMed: 16959575.
  51. Atkinson NJ, Urwin PE (2012) The interaction of plant biotic and abiotic stresses: from genes to the field. *J Exp Bot* 63: 3523-3543. doi:10.1093/jxb/ers100. PubMed: 22467407.
  52. Vlot AC, Dempsey DA, Klessig DF (2009) Salicylic Acid, a multifaceted hormone to combat disease. *Annu Rev Phytopathol* 47: 177-206. doi:10.1146/annurev.phyto.050908.135202. PubMed: 19400653.
  53. Wilkinson S, Kudoyarova GR, Veselov DS, Arkhipova TN, Davies WJ (2012) Plant hormone interactions: innovative targets for crop breeding and management. *J Exp Bot* 63: 3499-3509. doi:10.1093/jxb/ers148. PubMed: 22641615.
  54. Thaler JS, Humphrey PT, Whiteman NK (2012) Evolution of jasmonate and salicylate signal crosstalk. *Trends Plant Sci* 17: 260-270. doi:10.1016/j.tplants.2012.02.010. PubMed: 22498450.
  55. Brossa R, Lopez-Carbonell M, Jubany-Mari T, Alegre L (2011) Interplay Between Abscisic Acid and Jasmonic Acid and its Role in Water-oxidative Stress in Wild-type, ABA-deficient, JA-deficient, and Ascorbate-deficient Arabidopsis Plants. *J Plant Growth Regul* 30: 322-333. doi:10.1007/s00344-011-9194-z.

56. Bari R, Jones JD (2009) Role of plant hormones in plant defence responses. *Plant Mol Biol* 69: 473-488. doi:10.1007/s11103-008-9435-0. PubMed: 19083153.
57. Cui X, Luan S (2012) A new wave of hormone research: crosstalk mechanisms. *Molecular Plants* 5: 959-960. doi:10.1093/mp/sss090.
58. Muench DG, Zhang C, Dahodwala M (2012) Control of cytoplasmic translation in plants. *Wiley Interdiscip Rev RNA* 3: 178-194. doi:10.1002/wrna.1104. PubMed: 22215505.
59. Shah J (2003) The salicylic acid loop in plant defense. *Curr Opin Plant Biol* 6: 365-371. doi:10.1016/S1369-5266(03)00058-X. PubMed: 12873532.
60. Frye CA, Tang DZ, Innes RW (2001) Negative regulation of defense responses in plants by a conserved MAPKK kinase. *Proc Natl Acad Sci U S A* 98: 373-378. doi:10.1073/pnas.98.1.373. PubMed: 11114160.
61. Cheng Z, Sun L, Qi T, Zhang B, Peng W et al. (2011) The bHLH transcription factor MYC3 interacts with the Jasmonate ZIM-domain proteins to mediate jasmonate response in Arabidopsis. *Molecular Plants* 4: 279-288. doi:10.1093/mp/ssp073. PubMed: 21242320.
62. Park JE, Park JY, Kim YS, Staswick PE, Jeon J et al. (2007) GH3-mediated auxin homeostasis links growth regulation with stress adaptation response in Arabidopsis. *J Biol Chem* 282: 10036-10046. doi:10.1074/jbc.M610524200. PubMed: 17276977.
63. Chung HS, Koo AJ, Gao X, Jayanty S, Thines B et al. (2008) Regulation and function of Arabidopsis JASMONATE ZIM-domain genes in response to wounding and herbivory. *Plant Physiol* 146: 952-964. doi:10.1104/pp.107.115691. PubMed: 18223147.
64. To JP, Deruère J, Maxwell BB, Morris VF, Hutchison CE et al. (2007) Cytokinin regulates type-A Arabidopsis Response Regulator activity and protein stability via two-component phosphorelay. *Plant Cell* 19: 3901-3914. doi:10.1105/tpc.107.052662. PubMed: 18065689.
65. Chae K, Isaacs CG, Reeves PH, Maloney GS, Muday GK et al. (2012) Arabidopsis SMALL AUXIN UP RNA63 promotes hypocotyl and stamen filament elongation. *The Plant journal : for cell and molecular biology* 71: 684-697.
66. Dat J, Vandenabeele S, Vranová E, Van Montagu M, Inzé D et al. (2000) Dual action of the active oxygen species during plant stress responses. *Cell Mol Life Sci CMLS* 57: 779-795. doi:10.1007/s000180050041. PubMed: 10892343.
67. Moons A (2005) Regulatory and functional interactions of plant growth regulators and plant glutathione S-transferases (GSTs). *Vitam Horm* 72: 155-202. doi:10.1016/S0083-6729(05)72005-7. PubMed: 16492471.
68. Hernández-Blanco C, Feng DX, Hu J, Sánchez-Vallet A, Deslandes L et al. (2007) Impairment of cellulose synthases required for Arabidopsis secondary cell wall formation enhances disease resistance. *Plant Cell* 19: 890-903. doi:10.1105/tpc.106.048058. PubMed: 17351116.
69. Chen Z, Hong X, Zhang H, Wang Y, Li X et al. (2005) Disruption of the cellulose synthase gene, *AtCesA8/IRX1*, enhances drought and osmotic stress tolerance in Arabidopsis. *The Plant journal : for cell and molecular biology* 43: 273-283.
70. Song XQ, Liu LF, Jiang YJ, Zhang BC, Gao YP et al. (2013) Disruption of Secondary Wall Cellulose Biosynthesis Alters Cadmium Translocation and Tolerance in Rice Plants. *Molecular Plant*.
71. Mayer MP, Bukau B (2005) Hsp70 chaperones: cellular functions and molecular mechanism. *Cell Mol Life Sci CMLS* 62: 670-684. doi:10.1007/s00018-004-4464-6. PubMed: 15770419.
72. Upchurch RG (2008) Fatty acid unsaturation, mobilization, and regulation in the response of plants to stress. *Biotechnol Lett* 30: 967-977. doi:10.1007/s10529-008-9639-z. PubMed: 18227974.
73. Demmig-Adams B, Cohu CM, Amiard V, van Zadelhoff G, Veldink GA et al. (2013) Emerging trade-offs – impact of photoprotectants (PsbS, xanthophylls, and vitamin E) on oxylipins as regulators of development and defense. *New Phytol* 197: 720–729. doi:10.1111/nph.12100. PubMed: 23418633.
74. Gfeller A, Dubugnon L, Liechti R, Farmer EE (2010) Jasmonate Biochemical Pathway. *Sci Signal* 3: cm3-. PubMed: 20159849.
75. Staswick PE (2008) JAZing up jasmonate signaling. *Trends Plant Sci* 13: 66-71. doi:10.5363/tits.13.3\_66. PubMed: 18261950.
76. Vogt T (2010) Phenylpropanoid Biosynthesis. *Molecular Plants* 3: 2-20. doi:10.1093/mp/ssp106. PubMed: 20035037.
77. Tohge T, Watanabe M, Hoefgen R, Fernie AR (2013) The evolution of phenylpropanoid metabolism in the green lineage. *Crit Rev Biochem Mol Biol*.
78. Tzin V, Galili G (2010) New Insights into the Shikimate and Aromatic Amino Acids Biosynthesis Pathways in Plants. *Molecular Plants* 3: 956-972. doi:10.1093/mp/ssp048. PubMed: 20817774.
79. Maeda H, Dudareva N (2012) The Shikimate Pathway and Aromatic Amino Acid Biosynthesis in Plants. *Annu Rev Plant Biol*, Vol 63 63: 73-105. PubMed: 22554242.
80. Harb A, Krishnan A, Ambavaram MM, Pereira A (2010) Molecular and physiological analysis of drought stress in Arabidopsis reveals early responses leading to acclimation in plant growth. *Plant Physiol* 154: 1254-1271. doi:10.1104/pp.110.161752. PubMed: 20807999.
81. Lenka SK, Katiyar A, Chinnusamy V, Bansal KC (2011) Comparative analysis of drought-responsive transcriptome in Indica rice genotypes with contrasting drought tolerance. *Plant Biotechnol J* 9: 315-327. doi:10.1111/j.1467-7652.2010.00560.x. PubMed: 20809928.
82. Shu L, Lou Q, Ma C, Ding W, Zhou J et al. (2011) Genetic, proteomic and metabolic analysis of the regulation of energy storage in rice seedlings in response to drought. *Proteomics* 11: 4122-4138. doi:10.1002/pmic.201000485. PubMed: 21818852.
83. Pandey SP, Somssich IE (2009) The role of WRKY transcription factors in plant immunity. *Plant Physiol* 150: 1648-1655. doi:10.1104/pp.109.138990. PubMed: 19420325.
84. Griffiths S, Dunford RP, Coupland G, Laurie DA (2003) The evolution of CONSTANS-like gene families in barley, rice, and Arabidopsis. *Plant Physiol* 131: 1855-1867. doi:10.1104/pp.102.016188. PubMed: 12692345.
85. Chung HS, Howe GA (2009) A critical role for the TIFY motif in repression of jasmonate signaling by a stabilized splice variant of the JASMONATE ZIM-domain protein JAZ10 in Arabidopsis. *Plant Cell* 21: 131-145. doi:10.1105/tpc.108.064097. PubMed: 19151223.
86. Ye H, Du H, Tang N, Li X, Xiong L (2009) Identification and expression profiling analysis of TIFY family genes involved in stress and phytohormone responses in rice. *Plant Mol Biol* 71: 291-305. doi:10.1007/s11103-009-9524-8. PubMed: 19618278.
87. Allen JD, Xie Y, Chen M, Girard L, Xiao G (2012) Comparing statistical methods for constructing large scale gene networks. *PLOS ONE* 7: e29348. doi:10.1371/journal.pone.0029348. PubMed: 22272232.
88. Hand SC, Menze MA, Toner M, Boswell L, Moore D (2011) LEA Proteins During Water Stress: Not Just for Plants Anymore. *Annu Rev Physiol*, Vol 73 73: 115-134. PubMed: 21034219.
89. Zimmermann P, Hirsch-Hoffmann M, Hennig L, Gruissem W (2004) GENEVESTIGATOR. Arabidopsis microarray database and analysis toolbox. *Plant Physiol* 136: 2621-2632. doi:10.1104/pp.104.046367. PubMed: 15375207.
90. Fernandes H, Michalska K, Sikorski M, Jaskolski M (2013) Structural and functional aspects of PR-10 proteins. *FEBS J* 280: 1169-1199. doi:10.1111/febs.12114. PubMed: 23289796.
91. Cheng Y, Zhou Y, Yang Y, Chi YJ, Zhou J et al. (2012) Structural and functional analysis of VQ motif-containing proteins in Arabidopsis as interacting proteins of WRKY transcription factors. *Plant Physiol* 159: 810-825. doi:10.1104/pp.112.196816. PubMed: 22535423.
92. Lai Z, Li Y, Wang F, Cheng Y, Fan B et al. (2011) Arabidopsis sigma factor binding proteins are activators of the WRKY33 transcription factor in plant defense. *Plant Cell* 23: 3824-3841. doi:10.1105/tpc.111.090571. PubMed: 21990940.
93. Allen KN, Dunaway-Mariano D (2009) Markers of fitness in a successful enzyme superfamily. *Curr Opin Struct Biol* 19: 658-665. doi:10.1016/j.sbi.2009.09.008. PubMed: 19889535.
94. Kim CY, Bove J, Assmann SM (2008) Overexpression of wound-responsive RNA-binding proteins induces leaf senescence and hypersensitive-like cell death. *New Phytol* 180: 57-70. doi:10.1111/j.1469-8137.2008.02557.x. PubMed: 18705666.
95. Abercrombie JM, Halfhill MD, Ranjan P, Rao MR, Saxton AM et al. (2008) Transcriptional responses of Arabidopsis thaliana plants to As (V) stress. *Bmc. Plant Biol* 8: 87.