Identification of Potential Cytokinin Responsive Key Genes in Rice Treated With Trans-Zeatin Through Systems Biology Approach

Dwijesh Chandra Mishra[1†], Devender Arora[1,2†], Neeraj Budhlakoti[1], Amolkumar U. Solanke[3], S. V. Amitha CR Mithra[3], Anuj Kumar[1], P. S. Pandey[4], Sudhir Srivastava[1], Sanjeev Kumar[1], M. S. Farooqi[1], S. B. Lal[1], Anil Rai[1] and K. K. Chaturvedi[1*†]

[1]Centre for Agricultural Bioinformatics, ICAR-Indian Agricultural Statistics Research Institute, New Delhi, India, [2]National Institute of Animal Science, Rural Development Administration, Jeonju, South Korea, [3]ICAR-National Institute for Plant Biotechnology, New Delhi, India, [4]Agricultural Education Division, Indian Council of Agricultural Research, New Delhi, India

Rice is an important staple food grain consumed by most of the population around the world. With climate and environmental changes, rice has undergone a tremendous stress state which has impacted crop production and productivity. Plant growth hormones are essential component that controls the overall outcome of the growth and development of the plant. Cytokinin is a hormone that plays an important role in plant immunity and defense systems. *Trans*-zeatin is an active form of cytokinin that can affect plant growth which is mediated by a multi-step two-component phosphorelay system that has different roles in various developmental stages. Systems biology is an approach for pathway analysis to *trans*-zeatin treated rice that could provide a deep understanding of different molecules associated with them. In this study, we have used a weighted gene co-expression network analysis method to identify the functional modules and hub genes involved in the cytokinin pathway. We have identified nine functional modules comprising of different hub genes which contribute to the cytokinin signaling route. The biological significance of these identified hub genes has been tested by applying well-proven statistical techniques to establish the association with the experimentally validated QTLs and annotated by the DAVID server. The establishment of key genes in different pathways has been confirmed. These results will be useful to design new stress-resistant cultivars which can provide sustainable yield in stress-specific conditions.

Keywords: WGCNA, systems biology, co-expression, cytokinin, hub genes, QTLs

INTRODUCTION

Rice (*Oryza sativa*) is an important food grain crop that is consumed worldwide (Kubo and Purevdorj, 2004). The human population is estimated to reach approximately 10.7 billion by 2050 (Speidel, 2000) and accordingly, the demand for consumption of rice will also increase. On the other hand, the productivity of rice is not increasing at the same pace due to various reasons such as poor water management, soil depletion, abiotic stresses (e.g., drought, flooding, and salinity), biotic stresses (e.g., damage caused by bacteria, fungi, insects), etc. (Kumar et al., 2017). Therefore, it becomes necessary to understand the underlying mechanisms in crops to sustain production and fulfill the demand of the growing population (Ma and Michailides, 2007). Plants must be able to react rapidly with various abiotic and biotic stress signals and develop efficient defense responses to cope with adverse conditions arising in the field (Kumar et al., 2015). Understanding the crosstalk mechanisms in a stress response could help in the characterization of synergistic and antagonistic mechanisms (Kim et al., 2021). Plant hormones are the key components of these defense and adaptation mechanisms. To facilitate the adaptations to the stresses, various hormonal pathways are upregulated or downregulated. Change in hormonal influence usually affects the degree of resistance or susceptibility to different stresses (O'Brien and Benkova, 2013). Cytokinins (a class of phytohormones) are central regulators of plant growth and development. Cytokinins regulate numerous developmental and physiological processes such as cytokinesis of shoots and roots, reproductive behavior, leaf senescence, and responses to environmental cues, particularly to light and nutrients (Haberer and Kieber, 2002; Werner et al., 2003). Trans-Zeatin (tz) is an active form of cytokinin involved in managing environmental stress. Cytokinin pathway has been widely studied and a huge amount of gene expression data are available in public repositories (Edgar et al., 2002; Leinonen et al., 2011). These data can be better utilized for constructing gene regulatory networks and identifying key genes which will further help in developing improved rice varieties having the ability to produce high yield and resistance to such abiotic stress and adverse conditions (Stuart et al., 2003; Lelandais et al., 2011).

Key genes regulating plant growth and cytokinin responsive genes involved in development process will help in developing better stress tolerant varieties (Zhang et al., 2011, 2012; Li et al., 2019). Co-expression analysis is one of the important ways to construct such a network and identify the most relevant module (Heyer et al., 1999; Hudson et al., 2012; Gaiteri et al., 2014). A statistical approach, Weighted Gene Co-expression Network Analysis (WGCNA) is an effective way to perform such analysis (Tang et al., 2016; Che et al., 2018; Wu et al., 2018) and it has been successfully used previously in identifying important modules and key hub genes related to rice in different conditions (Tan et al., 2017; Zhang et al., 2018). This approach provides an analytical framework for the analysis of microarray and transcriptomic data and helps in relating the gene expression to external traits. Based on various co-expression analysis studies, we found WGCNA as a well-proven approach

for the identification of functional modules and co-expressed genes from large biological datasets (Kost et al., 2017; Zhang et al., 2018; Mishra et al., 2021). WGCNA is available as an R package and works on the principle of "guilt by association" (Langfelder and Horvath, 2008), that is, a group of genes are more likely to be associated with each other when they have similar expressions (Gillis and Pavlidis, 2012; van Dam et al., 2018). It uses an unsupervised learning method where large-scale data is clustered by building a tree from bottom to top by connecting the two nearest branches or genes. Modules are formed using the hierarchical clustering method and comprised of genes with similar functions. The modules can be further utilized to identify the important key genes. These key or hub genes are centrally connected genes in different modules which are expected to have an important function and play a critical role in maintaining overall harmony within the cell and development of tissue (Sircar and Parekh, 2015).

In this paper, we have performed co-expression analysis on publicly available microarray data retrieved from the NCBI GEO database (Edgar et al., 2002; Barrett et al., 2013) for cytokinin-responsive genes. We have identified novel key genes in each module using sound statistical approaches of co-expression analysis. Furthermore, we have done the biological validation of the novel key genes using well-established and experimentally validated QTLs (quantitative trait locus) of rice.

MATERIALS AND METHODS

Data Summary

Microarray data related to cytokinin effect on root and leaves of rice were retrieved from the NCBI GEO database with reference series GSE6719 (Hirose et al., 2007). In this experiment, roots and leaves of rice seedlings were treated with *trans*-zeatin (an active form of type-A cytokinin hormone) for 30 and 120 min. Three independent replicate treatments were performed for roots and leaves at each time point. To identify cytokinin responsive genes, dimethyl sulfoxide (DMSO) treated roots and leaves for each time point (30 and 120 min) were used as a control. Three independent replicates of DMSO treatment were performed for each organ per time point. Then, the microarray data were generated by using the Affymetrix GeneChip® rice genome array which contains probes to query approximately 48,564 *japonica* and 1,269 *indica* transcripts. There are 24 samples having accession series (GSE6719). The number of transcripts (features) in the expression data was 57,381, of which annotation data was available for only 23,850 transcripts. We have used only the annotated transcripts for WGCNA-based co-expression analysis.

Weighted Gene Co-Expression Analysis

To perform the analysis, we first studied the distribution pattern of grown samples using transcript count data and performed principal component analysis. Following, uniform distribution pattern we have used the R package "WGCNA" for constructing the network (Langfelder and Horvath, 2008). The elements in the

Sample clustering to detect outliers

FIGURE 1 | Clustering dendrogram of samples based on their Euclidean distance.

FIGURE 2 | Analysis of network topology for various soft-thresholding powers. The left panel shows the scale-free fit index (*y*-axis) as a function of the soft-thresholding power (*x*-axis). The right panel displays the mean connectivity (degree, *y*-axis) as a function of the soft-thresholding power (*x*-axis).

co-expression matrix are defined as the weighted value of correlation coefficients. Gene co-expression of a pair of genes i and j were calculated using **Eq. 1** of an unsigned network:

$$A_{ij} = \left| cor\left(E_i, E_j \right) \right|^\beta \qquad (1)$$

where E_i and E_j consist of expression vector profiles of genes i and j across multiple samples and A_{ij} is the adjacency of the unsigned network. Pearson's correlation coefficient was used to identify the similarity of genes. The absolute value of correlation is raised to a power β to create the adjacency matrix. WGCNA emphasizes high correlation by raising the absolute value of correlation to a

FIGURE 3 | Clustering dendrogram of genes, with dissimilarity based on topological overlap, together with assigned merged module colors and the original module colors.

FIGURE 4 | Visualization of the eigengenes network representing the relationships among the modules. The top panel shows a hierarchical clustering dendrogram of the modules based on the dissimilarity score of E_i and E_j given by $1-cor\ (E_i, E_j)$. The heatmap in the bottom panel shows the eigengenes adjacency $A_{ij} = (1 + cor\ (E_i, E_j))/2$.

FIGURE 5 | Visualizing the gene network using a heatmap plot. The heatmap depicts the Topological Overlap Matrix (TOM). The light red color represents low overlap and progressively darker shades of red represent higher overlap. Blocks of dark colors along the diagonal depict modules. The gene dendrogram and module assignment are also shown along the left side and the top of the Figure.

power $\beta \geq 1$. The correlation network among the genes is built based on the adjacency matrix. Each gene is considered as a node and an edge between two genes is built if it passes a set threshold of co-expression strength. A soft threshold value of 6 was used to determine the significant edges for connecting pairs of genes. Furthermore, modules within the networks were identified by calling the R function "cutreeDynamic" available in the WGCNA package, which helps in identifying the minimum number of large modules with a strong association of genes. The function was applied after identifying the soft threshold of the adjacency matrix, and cut them respectively to get modules related to different functions. These identified modules were further used to identify the key or hub genes responsible for governing specific biological functions in cell development.

In order to identify the modules that were highly related, module similarity was quantified by eigengenes correlation (Shi et al., 2015). The eigengenes of a module are defined as the eigenvector associated with the first principal component of the expression matrix

(Langfelder and Horvath, 2007). Highly related modules were identified using module significance. Module significance is defined as the average gene significance of all the genes in the modules to access the association of a module to the phenotype.

Visualization of Network and Hub Genes Identification

Various methods have been used for the identification of hub genes from a large dataset (Bader and Hogue, 2003; Chin et al., 2014). These methods mainly focused on hub gene identification, based only on gene connection degrees in the gene co-expression network. A heat map was generated to compare the expression pattern across the samples and assigned different colors to genes with similar values (Zhao et al., 2014). Genes are interconnected in each module and possess specific functions. After identifying the most significant module, identification of key or hub genes was
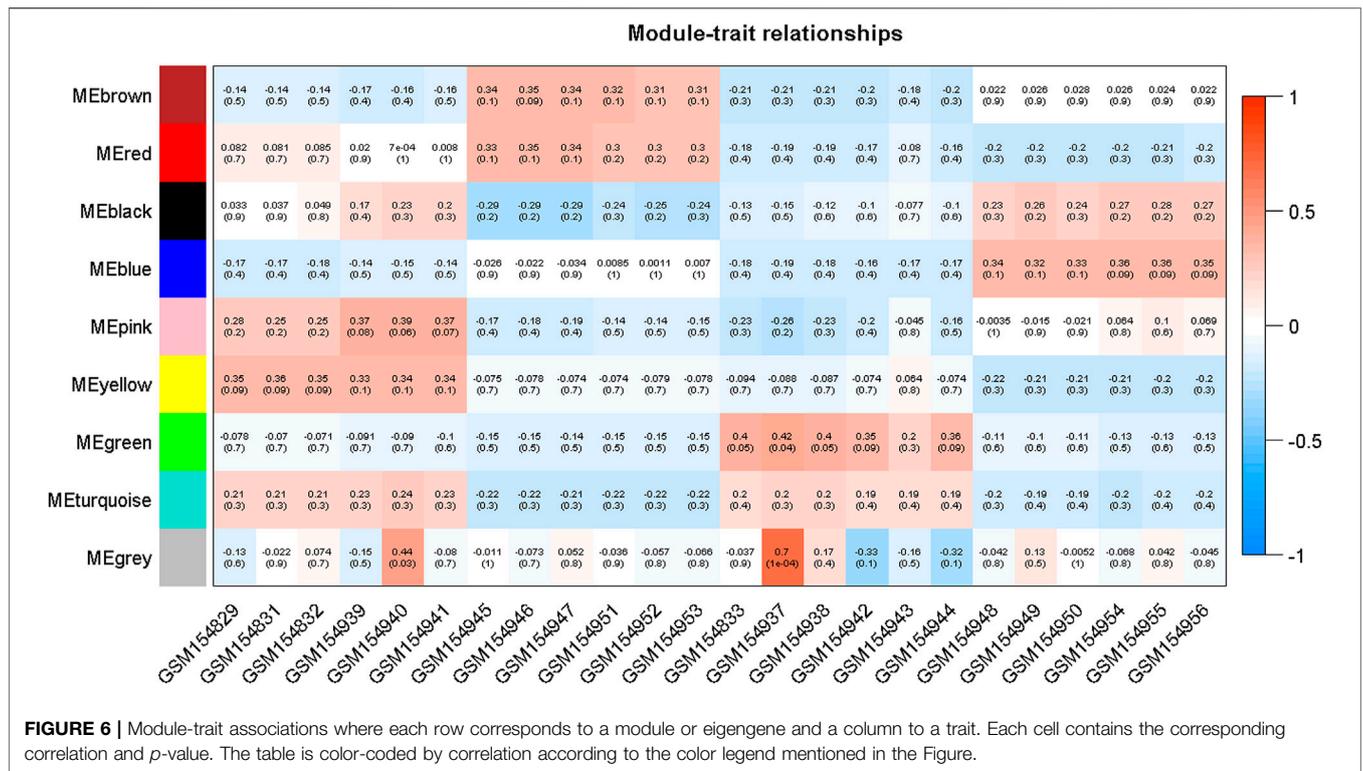
FIGURE 6 | Module-trait associations where each row corresponds to a module or eigengene and a column to a trait. Each cell contains the corresponding correlation and *p*-value. The table is color-coded by correlation according to the color legend mentioned in the Figure.

TABLE 1 | Identified hub genes with respect to their level of connectivity in the turquoise module.

S.No	Hub gene	Function/Annotation	Locus location	Acc. No
1	Os04g0295400	Horcolin/Jasmonate-induced protein, putative, expressed	LOC4335407/Chr 4, NC_029259.1 (12989808.12993215)	AK067477
2	Os03g0425800	uncharacterized LOC4333157	LOC4333157/Chr 3, NC_029258.1 (17780383.17784105)	AK100427
3	Os02g0731200	MADS-box transcription factor 57	LOC4330621/Chr 2, NC_029257.1 (30456659.30462758)	AY177702
4	Os01g0558800	uncharacterized LOC4324676	LOC4324676/Chr 1, NC_029256.1 (21151568.21155639, complement)	AK068120

TABLE 2 | Identified hub genes with respect to their level of connectivity in the blue module.

S.No	Hub gene	Function	Locus location	Acc. No
1	Os07g0424400	probable cellulose synthase A catalytic subunit 3 [UDP-forming]	LOC4343049/Chr 7, NC_029262.1 (13741571.13747205, complement)	AK120236
2	Os02g0511800	uncharacterized LOC4329457	LOC4329457/Chr 2, NC_029257.1 (18343563.18347153, complement)	AK069817
3	Os08g0505200	uncharacterized LOC4345975	LOC4345975/Chr 8, NC_029263.1 (24986078.24990484)	AK067190
4	Os01g0276800	Glucosidase 2 subunit beta	LOC4324264/Chr 1, NC_029256.1 (9702711.9709700, complement)	AK108476

carried out using an R package, "dhga" (Differential Hub Gene Analysis) (Das et al., 2017).

Gene Annotation and Gene Ontology Analysis

The annotation of expressed genes was performed using the Institute for Genomic Research (TIGR) database which

includes information regarding biological processes (BP), molecular function (MF), and cellular components (CC) (Ouyang et al., 2007). Furthermore, the identified genes from identified modules were submitted to the OryzaExpress database (Hamada et al., 2011) and converted the ids into probeID (Affymatrix). These ids were subsequently converted into geneid using DAVID (the database for

TABLE 3 | Identified hub genes with respect to their level of connectivity in the brown module.

S.No	Hub gene	Function	Locus location	Acc. No
1	Os01g0743600	ATP-dependent protease La domain containing protein, expressed	LOC4326165/Chr 1, NC_029256.1 (31078612.31088072, complement)	AK102317
2	Os06g0319800	Os01g0512300	Os01g0512300/Chr 1, NC_008394.3 (18047906.18050788, complement)	AK107048
3	Os07g0636600	dirigent protein 5	LOC4344033/Chr 7, NC_029262.1 (26445635.26446517, complement)	AK106022
4	Os05g0128100	uncharacterized LOC4337691	LOC4337691/Chr 5, NC_029260.1 (1652011.1653860)	AK108556

TABLE 4 | Identified hub genes with respect to their level of connectivity in the green module.

S.No	Hub gene	Function	Locus location	Acc. No
1	Os07g0171300	uncharacterized LOC4342515	LOC4342515/Chr 7, NC_029262.1 (3773297.3779772)	AK100663
2	Os03g0610800	serpin-ZXB	LOC4333434/Chr 3, NC_029258.1 (23054420.23068346)	AK107194
3	Os01g0827600	exocyst complex component EXO70B1	LOC4327525/Chr 1, NC_029256.1 (35412698.35416714)	AK122173
4	Os08g0200400	KH domain-containing protein At4g18375	LOC4344900/Chr 8, NC_029263.1 (5810049.5816102, complement)	AK067859

TABLE 5 | Identified hub genes with respect to their level of connectivity in the yellow module.

S.No	Hub gene	Function	Locus location	Acc. No
1	Os09g0442400	protein GAMETE EXPRESSED 1	LOC4347181/Chr 9, NC_029264.1 (16443693.16448158)	AK106970
2	Os08g0522400	putative L-ascorbate peroxidase 6	LOC4346078/Chr 8, NC_029263.1 (2,5971775.25974968, complement)	AK065893
3	Os03g0654600	chlorophyll (ide) b reductase NOL, chloroplastic	LOC4333604/Chr 3, NC_029258.1 (25520290.25525342, complement)	CB669633
4	Os07g0550600	benzyl alcohol O-benzoyltransferase	LOC4343545/Chr 7, NC_029262.1 (21854513.21856895)	AK109553

TABLE 6 | Identified hub genes with respect to their level of connectivity in the red module.

S.No	Hub gene	Function	Locus location	Acc. No
1	Os12g0566000	boron transporter 1	LOC4352546/Chr 12, NC_029267.1 (23248819.23253256)	AK100510
2	Os04g0658300	ribulose bisphosphate carboxylase/oxygenase activase, chloroplastic	LOC4337267/Chr 4, NC_029259.1 (33575149.33579656, complement)	AK067399
3	Os05g0358200	DNA primase small subunit	LOC9267485/Chr 5, NC_029260.1 (17006498.17011733, complement)	AK073973
4	Os11g0707000	ribulose bisphosphate carboxylase/oxygenase activase, chloroplastic	LOC4351224/Chr 11, NC_029266.1 (28932976.28936094, complement)	CB673145

TABLE 7 | Identified hub genes with respect to their level of connectivity in the black module.

S.No	Hub gene	Function	Locus location	Acc. No
1	Os11g0491400	uncharacterized LOC4350546	LOC4350546/Chr 11, NC_029266.1 (17379762.17382920)	AK068341
2	Os07g0271500	bisdemethoxycurcumin synthase	LOC4342896/Chr 7, NC_029262.1 (10018732.10020733)	AK109558
3	Os09g0482740	uncharacterized LOC9271634	LOC9271634/Chr 9, NC_029264.1 (18576964.18581257)	AK061852
4	LOC4338611	lichenase-2	LOC4338611/Chr 5, NC_029260.1 (18106236.18110996, complement)	CB628871

TABLE 8 | Identified hub genes with respect to their level of connectivity in the pink module.

S.No	Hub gene	Function	Locus location	Acc. No
1	Os07g0187700	SEC12-like protein 1	LOC4342601/Chr 7, NC_029262.1 (4682485.4687647)	AK111777
2	Os05g0568800	bradykinin-potentiating and C-type natriuretic peptides	LOC4339650/Chr 5, NC_029260.1 (28310927.28311938)	AK059883
3	Os04g0334700	aspartic proteinase-like protein 2	LOC4335504/Chr 4, NC_029259.1 (15623634.15653865, complement)	AK120870
4	Os03g0359000	uncharacterized LOC4332880	LOC4332880/Chr 3, NC_029258.1 (13927427.13935150, complement)	AK064512

TABLE 9 | Identified hub genes with respect to their level of connectivity in the grey module.

S.No	Hub gene	Function	Locus location	Acc. No
1	Os07g0131600	hexose carrier protein HEX6	LOC4342334/Chr 7, NC_029262.1 (1669273.1671384)	AK068296
2	Os03g0704100	probable plastid-lipid-associated protein 4, chloroplastic	LOC4333849/Chr 3, NC_029258.1 (28304914.28307925)	AK070474
3	Os06g0130400	probable aminotransferase ACS12	LOC4340002/Chr 6, NC_029261.1 (1629690.1633597)	AK065212
4	Os03g0322500	14 kDa zinc-binding protein	LOC4332685/Chr 3, NC_029258.1 (11672780.11677395)	AK121029
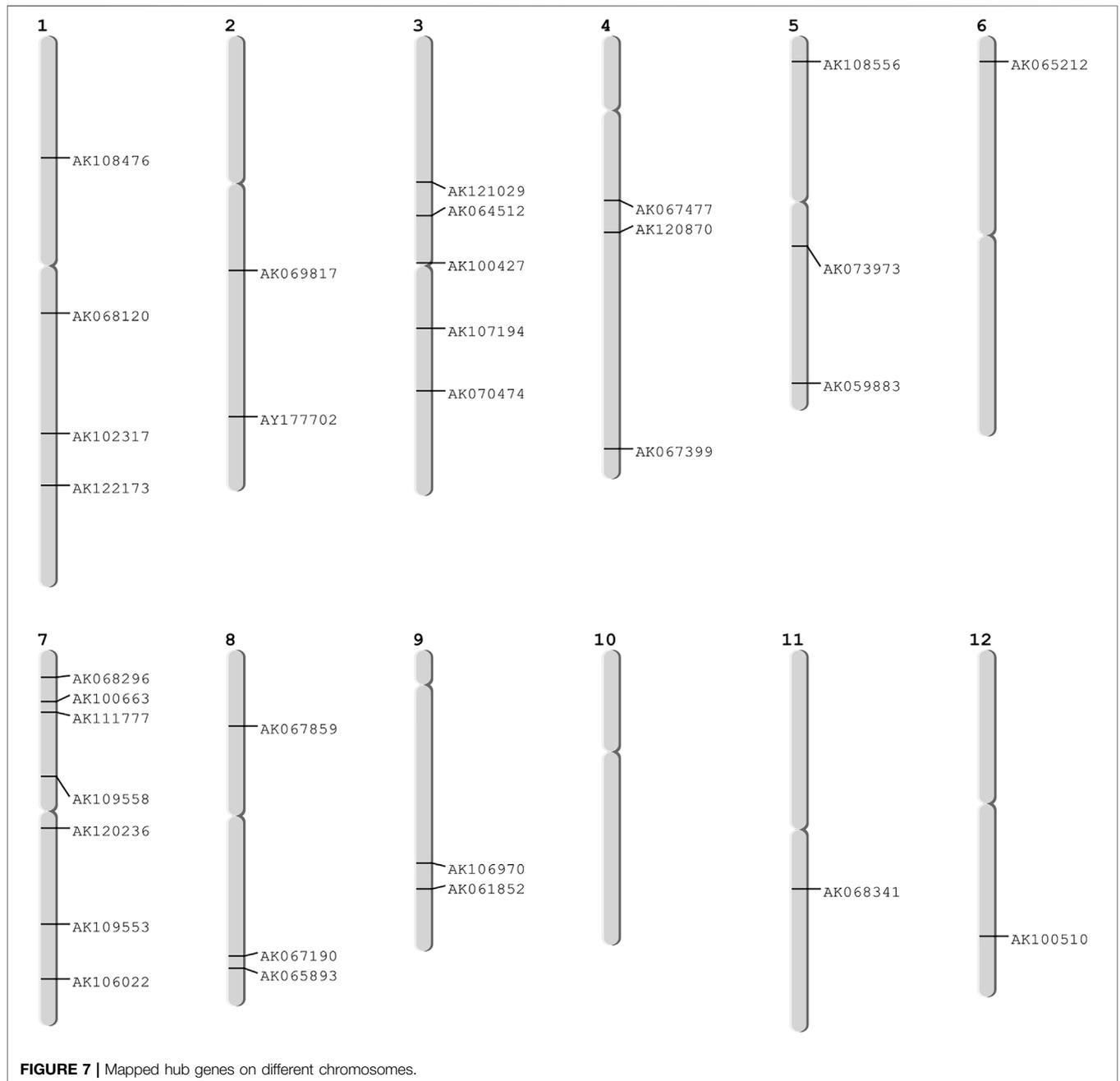
FIGURE 7 | Mapped hub genes on different chromosomes.

annotation, visualization, and integrated discovery) (Huang et al., 2007) for KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway analysis web server to decipher the role of genes varies from BP, CC, MF, and KEGG pathway analysis respectively against *oryza sativa* database and observed in REVIGO database (Supek et al., 2011). We found the
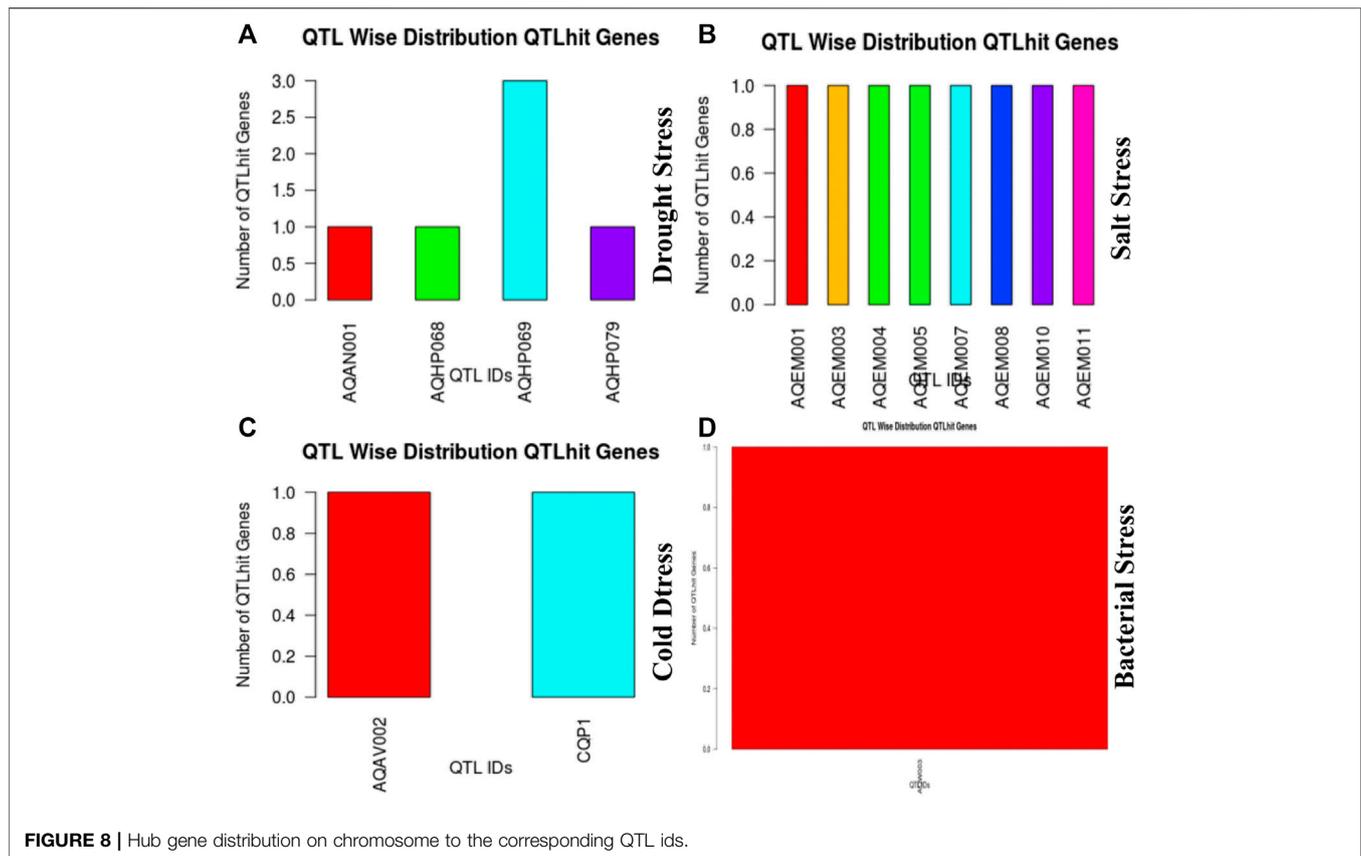
FIGURE 8 | Hub gene distribution on chromosome to the corresponding QTL ids.

involvement of these hub genes by observing the changes in the expression level of treated and control (normal) conditions [S1_file].

Validation of Identified Hub Genes

We validated the identified hub genes using *in-silico* approach. In the *in-silico* approach, we identified the chromosome-wise distribution of the hub genes and mapped these genes on chromosomes using web server oryzabase (https://shigen.nig.ac.jp/rice/oryzabase/) (Kurata and Yamazaki, 2006). Further, we used a statistical framework approach to test the association of these hub genes with well-known experimentally validated and genetically rich QTLs data reported for biotic and abiotic stress conditions (Jaiswal et al., 2002). We have used an R package "GSAQ" (https://cran.r-project.org/web/packages/GSAQ) (Das et al., 2018) for mapping the identified hub genes with QTLs on respective chromosomes. GSAQ provides a platform to associate the selected hub genes to the corresponding overlapped QTL-IDs with their genomic positions. Furthermore, the identified hub genes were also validated through pathway analysis.

RESULTS

Weighted Co-expression Network Construction and Module Identification

Clustering of the samples (**Figure 1**) suggests that there is no outlier present in the data. Power β was obtained through two types of graphs given in **Figure 2**: 1) soft threshold values of β (*x*-axis) *vs* scale-free topology model fit scaled R^2 (*y*-axis) and 2) soft threshold values of β (*x*-axis) vs mean connectivity scores (*y*-axis). The optimal value of β obtained using these graphs is 6 with R^2 value 0.8. This value of β was further used to produce hierarchical clustering (**Figure 3**).

A dynamic hierarchical tree algorithm was used to divide the clustering tree constructed from the differentially expressed genes, resulting in 24 different co-expression modules in the data which were named as blue (3,972 genes), brown (3,064 genes), green (1801 genes), yellow (2,387 genes), and red (1,272 genes) and turquoise (5,670 genes), Black (1,115 genes), Cyan (275 genes), dark green (85 genes), dark red (113), dark turquoise (83 genes), green yellow (338 genes), grey (35 genes), grey60 (162 genes), light cyan (246 genes), light green (144 genes), light yellow (141

FIGURE 9 | The classification of identified hub genes represented by heat map using DAVID software. It is noted that eight genes (*y*-axis) are classified into four classes (*x*-axis).

genes), magenta (644 genes), midnight blue (252 genes), pink (835 genes), purple (442 genes), royal blue (122 genes), salmon (296 genes), and tan (326 genes) (**Figure 4**). After clustering, the genes were grouped into modules (subnetworks) depicted in different colors for easy identification (**Figure 5**).

Moreover, in order to explore the relationship between identified modules and the experimental samples (traits), we calculated and tested the correlation coefficients. These correlation coefficients along with their *p*-values for module-trait relationship were depicted (**Figure 6**). In **Figure 6**, the red color shows a strong positive correlation and the blue color displays a strong negative correlation. Furthermore, the centralized hub genes were identified from these modules through statistical analysis with the help of R-package "dhga" using a weighted gene score. A total of 36 hub genes were identified in 9 modules and their detailed description (function/annotation, location, and accession number) are mentioned in **Table 1**; **Table 2**; **Table 3**; **Table 4**; **Table 5**; **Table 6**: **Table 7**; **Table 8**; **Table 9**.

Validation of Identified Hub Genes

These 36 identified hub genes were mapped on rice chromosomes and 32 genes were found to be located at various rice chromosomes (**Figure 7**). In the process of finding the association of these genes with the well-known QTLs related to salt, cold, drought, bacterial stress, we found that 17 out of the 36 identified hub genes were associated with these QTLs mapped on various chromosomes (**Figure 8**). We also performed pathway analysis and found that the expression of these genes either increased or decreased during the period of treatment in five different pathways. Heat map analysis was conducted through DAVID to produce a matrix of enriched GO terms with the identified genes. The green and black color shading on the heat map matrix indicate a positive and negative correlation between the enriched GO term and identified DMSO and tz-treated hub genes, respectively (**Figure 9**).

DISCUSSION

The productivity of rice is severely affected due to various biotic and abiotic stresses. Therefore, to develop a variety that is resistant to these stresses, there is an urgent need to identify important hub genes governing the whole production process. Rice genome consists of around 58,000 genes (Cao et al., 2012)

FIGURE 10 | Distribution of samples based on different stress at different time intervals and obtained uniform distribution of data in all four coordinates (Szklarczyk et al., 2017).
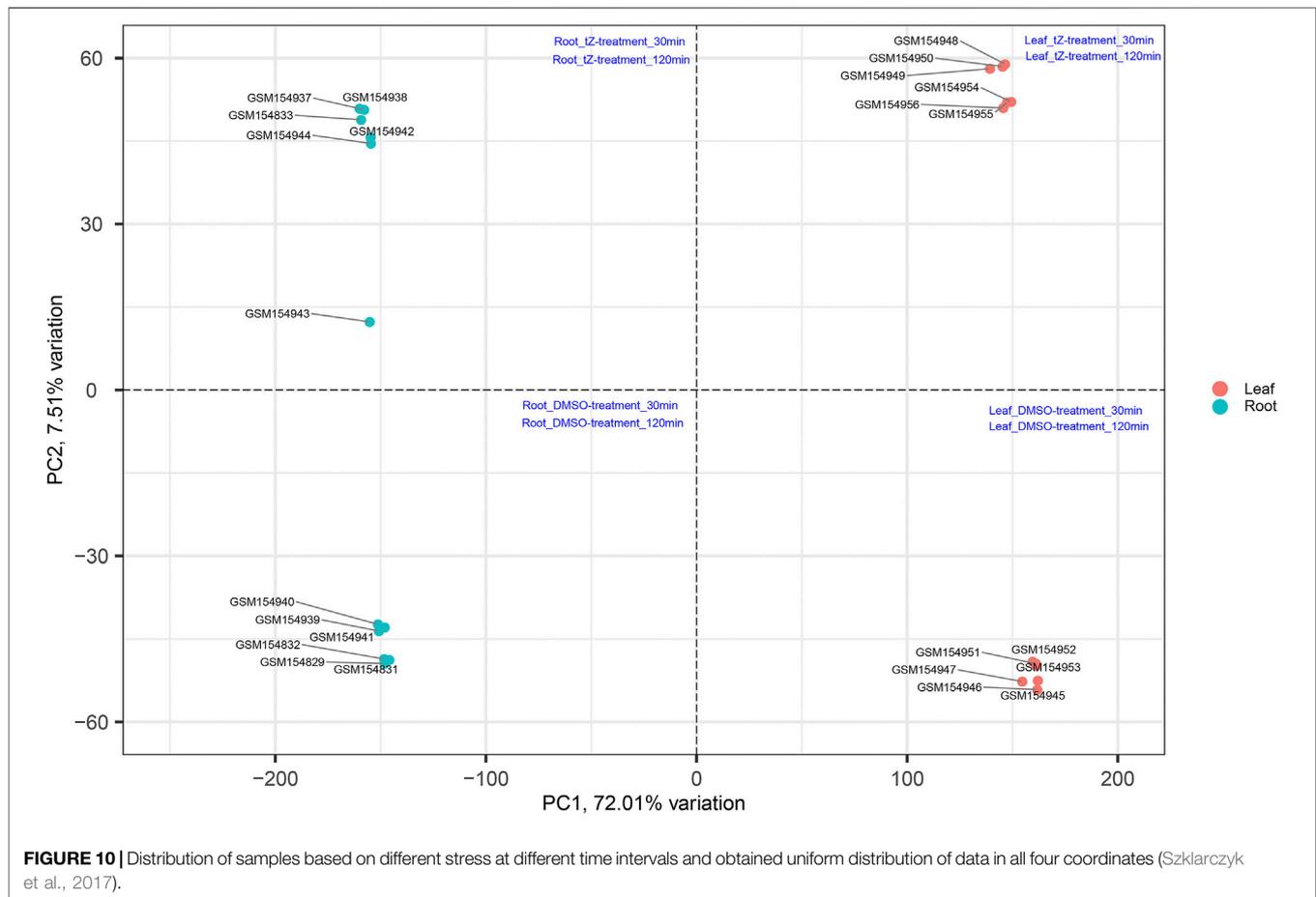
and conventional approaches can identify several hundred genes related to these stresses. However, by using conventional approaches, it becomes difficult to identify the few centrally important genes that play important roles in cellular functions to cope with these stresses. This problem can be solved by applying systems biology approaches (Arora et al., 2019). In this study, we have performed a comprehensive analysis on existing data retrieved from NCBI to understand the potential genes and mechanisms involved in such processes by first performing PCA to validate the distribution and uniformity of data evenly (**Figure 10**) and subsequently obtained significant modules associated with the biological functions regulating the growth and development of the plant. Moreover, hub genes in these modules were also identified that play an important role during cytokinin signaling and are crucial in plant growth and development.

Amongst the identified modules, the hub genes identified in these modules were observed to be involved in various processes. For example, in the turquoise module, the top hub gene (Os04g0295400) is located in chr4 and its important function is to encode Jasmonate-induced protein. Though little is known about this function in rice, Jasmonate-induced proteins are already reported for immunity and development in other plants (Wasternack and Hause, 2013; Campos et al., 2014). In-

depth characterization of this gene is further required as this family of genes are reported to play important roles such as in the defense systems against lethal disease and bacterial blight (Yamada et al., 2012; Taniguchi et al., 2014) and it may also be involved in stress management as both these stresses (abiotic and biotic stress) are interlinked with each other (Cao et al., 2017). Similarly, in the blue module, we identified the Os07g0424400 hub gene in chr7 that played an important role in cellulose synthase A catalytic subunit 3 [UDP-forming] that governs a major mechanism of cell wall formation (Taylor et al., 2000) and ultimately helps in supporting the plant growth against abiotic stress. In the brown module, the top potential hub gene, Os01g0743600, is located on chr1 with the reported function of ATP-dependent protease La domain containing protein (Koodathingal et al., 2009) and it is one of the key components in providing protection against the harmful effects of unfolded proteins. It is activated by stress conditions in the endoplasmic reticulum (ER) and it supports plant defense as well as response to abiotic stresses (Bao and Howell, 2017).

The top potential key gene in the green module, Os07g0171300, is still not fully characterized, but annotation results suggest it has a key influence in the protein kinase-like domain superfamily which is believed to be a conserved protein domain mainly involved in most signaling and regulatory
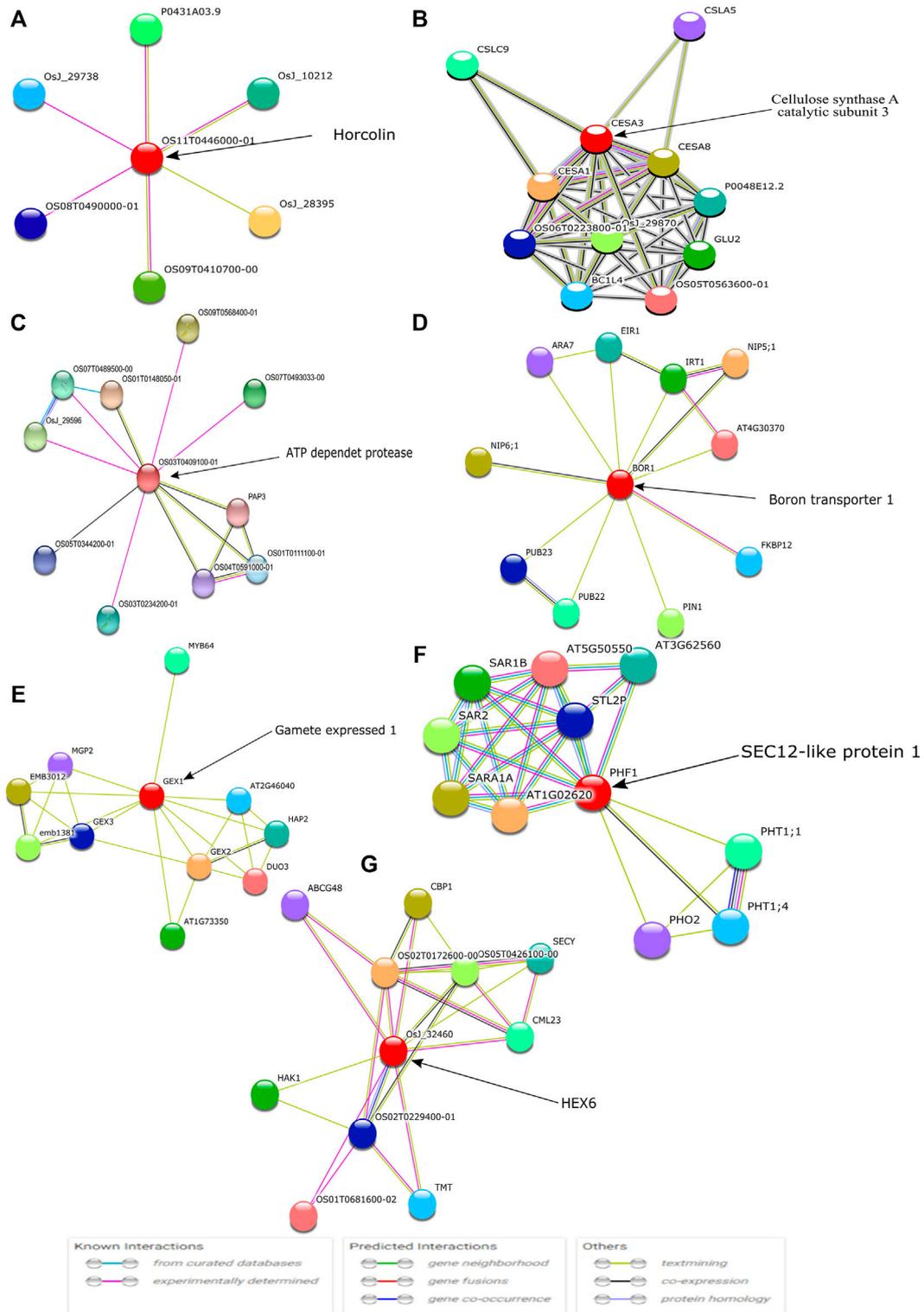
FIGURE 11 | The identified top hub genes are denoted with circles and are known as the first shell of interactors. Each color symbol signifies a specific interaction either known or predicted, as mentioned in previous studies.

FIGURE 12 | Gene ontology analysis viz biological process, cellular component, and molecular function terms associated with different modules performed using REVIGO.
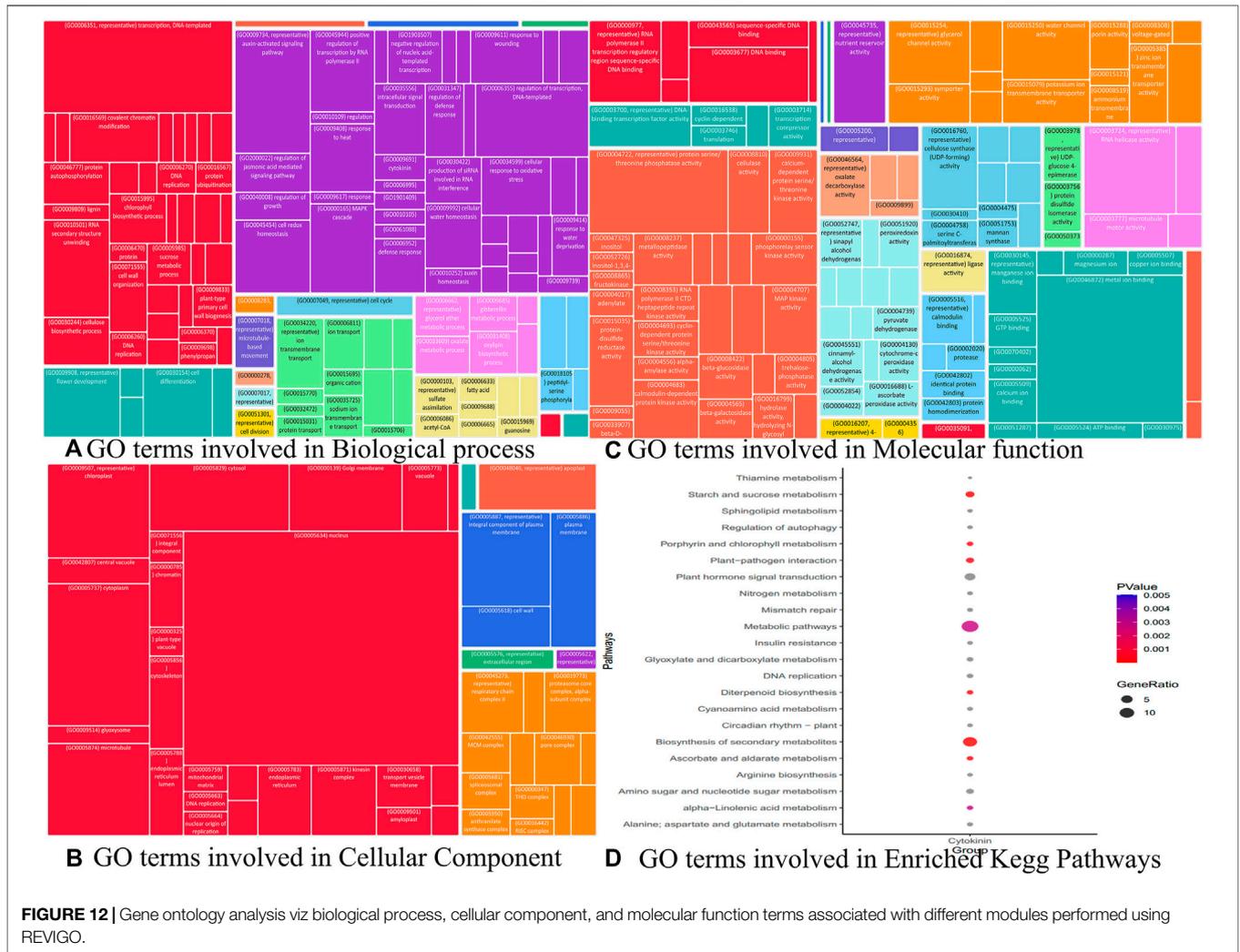
processes in the eukaryotic cell, and as a switch in controlling biological processes such as metabolism, transcription, cell moment, apoptosis, etc.

In the yellow module, the identified key gene Os12g0566000 codes for boron transporter 1, and is located at LOC4352546 of Chr12 and mapped with QTL identification number AK100510. Boron is essential for maintaining the integrity of plants cell walls. It exhibits an important structural role in shaping the cell by providing mechanical strength via cross-linking of cell wall rhamnogalacturonan II (RG-II) to form a stable three-dimensional pectic network which contributes to the mechanical properties of cell wall structure (Funakawa and Miwa, 2015). It is also reported that boron expression deficiency inhibits plant photosynthetic capacity (Zhao and Oosterhuis, 2002; Kastori et al., 2008) and directly impacts the total yield of the crop.

In the red module, the identified key gene Os09g0442400 codes for protein Gamete expressed 1, and is located at LOC4347181 of Chr9 and mapped with QTL identification number AK106970. Gamete expressed 1 protein is mainly responsible for fertilization. It has a dual function during

gametophyte development and early embryogenesis and it is required for correct pollen maturation (Engel et al., 2005). In the pink module, the identified key gene Os07g0187700 codes for protein SEC12-like protein1, and is located at LOC4342601 of Chr7 and mapped with QTL identification number AK111777. Phosphate Transporter 1 (PHT1) is a plant-specific SEC12 gene that encodes phosphate transporter involved in phosphate uptake by facilitating the trafficking of PHT1-1/PHT1; 1 from the ER to the plasma membrane that enables the ER exit of a high-affinity phosphate transporter (González et al., 2005). The top key hub gene identified in the grey module, Os07g0131600, codes for HEX6 protein which is one of the hexose carrier proteins, and is located at LOC4342334 of Chr7 mapped with QTL identification number AK068296. HEX6 protein has an active uptake of hexose with an important role in glucose/hydrogen symport (Boles and Hollenberg, 1997). These different hub genes directly or indirectly govern the main function of the positive build-up of overall cell growth but there are leftover top key genes Os07g0171300 in the green module and Os11g0491400 in the black modules that need to be further characterized to understand their role at different stages of crop development.

Further, the identified key hub genes were visualized using the STRING database for protein-protein interaction (Szklarczyk et al., 2017). The STRING database helped in the identification of direct (physical) interactions and indirect (functional) interactions as long as the interactions were specific and biologically meaningful. Out of the 9 obtained modules, seven genes were found to be associated with neighboring proteins (**Figure 11**).

Similarly, we identified other important hub genes in each module (**Table 1** to **Table 9**) which are not yet fully explored with respect to cytokinin signaling to maintain the harmony of cell and rice growth mechanisms. Although we identified 36 key genes, we were interested in understanding the role of genes in various processes such as BP, CC, and MF to further delineate the role of these genes in playing different roles in the development of cytokinin-related responses. The Gene Ontology terms in these processes include transcription (GO:0006351), auxin-activated signaling pathway (GO:0009734), MAPK cascade (GO:0000165), and regulation of transcription in BP (GO:0006355) (**Figure 12A**), whereas the GO terms mostly constituted in CC include nucleus (GO:0005634), cytoplasm (GO:0005737), chloroplast (GO:0009507), and microtubule (GO:0005874) (**Figure 12B**) and composed of molecular functions metal ion binding (GO:0046872), protein serine/threonine phosphatase activity (GO:0004722), ATP-dependent RNA helicase activity (GO:0004004), and RNA polymerase II regulatory region sequence-specific DNA binding (GO:0000977) are some (**Figure 12C**). Likewise, the Kegg pathway also revealed the enrichment of GO terms such as plant hormone signal transduction (osa04075), starch and sucrose metabolism (osa00500), and diterpenoid biosynthesis (osa00904) (**Figure 12D**). In the absence of wet lab experiments, these identified hub genes were validated with the help of well-known QTLs and pathways using an *in-silico* approach. The analysis indicates the involvement of identified hub genes in these stress conditions as these genes are found to be associated with biotic and abiotic stress-related QTLs and pathways. These results indicate the role of identified hub genes in the regulation of plant growth and development. However, these hub genes need further attention at the molecular level through wet lab experiments to improve the traits which will be useful in enhancing the productivity of the crop.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

DM, DA, NB, AS, SM, AK, and SS conducted the experiments. DA, NB, and SS analyzed the data, DM and DA drafted the manuscript. MF, SL, AR, PP, and KC proofread the manuscript. AR and KC revised the manuscript. All authors agreed to the final manuscript.

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.780599/full#supplementary-material

REFERENCES

Arora, D., Chaudhary, R., and Singh, A. (2019). System Biology Approach to Identify Potential Receptor for Targeting Cancer and Biomolecular Interaction Studies of Indole[2,1-a]Isoquinoline Derivative as Anticancerous Drug Candidate against it. *Interdiscip. Sci. Comput. Life Sci.* 11 (1), 125–134. doi:10.1007/s12539-017-0249-0

Bader, G. D., and Hogue, C. W. (2003). An Automated Method for Finding Molecular Complexes in Large Protein Interaction Networks. *BMC Bioinformatics* 4 (4), 2–27. doi:10.1186/1471-2105-4-2

Bao, Y., and Howell, S. H. (2017). The Unfolded Protein Response Supports Plant Development and Defense as Well as Responses to Abiotic Stress. *Front. Plant Sci.* 8, 344. doi:10.3389/FPLS.2017.00344

Barrett, T., Wilhite, S. E., Ledoux, P., Evangelista, C., Kim, I. F., Tomashevsky, M., et al. (2013). NCBI GEO: Archive for Functional Genomics Data Sets-Update. *Nucleic Acids Res.* 41, D991–D995. doi:10.1093/NAR/GKS1193

Boles, E., and Hollenberg, C. P. (1997). The Molecular Genetics of Hexose Transport in Yeasts. *FEMS Microbiol. Rev.* 21, 85–111. doi:10.1111/J.1574-6976.1997.TB00346.X

Campos, M. L., Kang, J.-H., and Howe, G. A. (2014). Jasmonate-triggered Plant Immunity. *J. Chem. Ecol.* 40, 657–675. doi:10.1007/S10886-014-0468-3

Cao, P., Jung, K.-H., Choi, D., Hwang, D., Zhu, J., and Ronald, P. C. (2012). The Rice Oligonucleotide Array Database: an Atlas of rice Gene Expression. *Rice* 5, 17. doi:10.1186/1939-8433-5-17

Cao, X.-Q., Jiang, Z.-H., Yi, Y.-Y., Yang, Y., Ke, L.-P., Pei, Z.-M., et al. (2017). Biotic and Abiotic Stresses Activate Different Ca2+ Permeable Channels in Arabidopsis. *Front. Plant Sci.* 8, 83. doi:10.3389/FPLS.2017.00083

Che, T., Li, D., Jin, L., Fu, Y., Liu, Y., Liu, P., et al. (2018). Long Non-coding RNAs and mRNAs Profiling during Spleen Development in Pig. *PLoS One* 13, e0193552. doi:10.1371/JOURNAL.PONE.0193552

Chin, C.-H., Chen, S.-H., 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, 1–7. doi:10.1186/1752-0509-8-S4-S11

Das, S., Meher, P. K., Rai, A., Bhar, L. M., and Mandal, B. N. (2017). Statistical Approaches for Gene Selection, Hub Gene Identification and Module Interaction in Gene Co-expression Network Analysis: An Application to Aluminum Stress in Soybean (Glycine max L.). *PLoS One* 12, e0169605. doi:10.1371/JOURNAL.PONE.0169605

Das, S., Rai, A., Mishra, D. C., and Rai, S. N. (2018). Statistical Approach for Gene Set Analysis with Trait Specific Quantitative Trait Loci. *Sci. Rep.* 8 (8), 1–12. doi:10.1038/s41598-018-19736-w

Edgar, R., Domrachev, M., and Lash, A. E. (2002). Gene Expression Omnibus: NCBI Gene Expression and Hybridization Array Data Repository. *Nucleic Acids Res.* 30, 207–210. doi:10.1093/NAR/30.1.207

Engel, M. L., Holmes-Davis, R., and McCormick, S. (2005). Green Sperm. Identification of Male Gamete Promoters in Arabidopsis. *Plant Physiol.* 138, 2124–2133. doi:10.1104/PP.104.054213

Funakawa, H., and Miwa, K. (2015). Synthesis of Borate Cross-Linked Rhamnogalacturonan II. *Front. Plant Sci.* 6, 223. doi:10.3389/FPLS.2015.00223

Gaiteri, C., Ding, Y., French, B., Tseng, G. C., and Sibille, E. (2014). Beyond Modules and Hubs: the Potential of Gene Coexpression Networks for Investigating Molecular Mechanisms of Complex Brain Disorders. *Genes, Brain Behav.* 13, 13–24. doi:10.1111/GBB.12106

Gillis, J., and Pavlidis, P. (2012). "Guilt by Association" Is the Exception rather Than the Rule in Gene Networks. *PLOS Comput. Biol.* 8, e1002444. doi:10.1371/JOURNAL.PCBI.1002444

González, E., Solano, R., Rubio, V., Leyva, A., and Paz-Ares, J. (2005). PHOSPHATE TRANSPORTER TRAFFIC FACILITATOR1 Is a Plant-specific SEC12-Related Protein that Enables the Endoplasmic Reticulum Exit of a High-Affinity Phosphate Transporter in Arabidopsis [W]. *Plant Cell* 17, 3500–3512. doi:10.1105/TPC.105.036640

Haberer, G., and Kieber, J. J. (2002). Cytokinins. New Insights into a Classic Phytohormone. *Plant Physiol.* 128, 354–362. doi:10.1104/PP.010773

Hamada, K., Hongo, K., Suwabe, K., Shimizu, A., Nagayama, T., Abe, R., et al. (2011). OryzaExpress: an Integrated Database of Gene Expression Networks and Omics Annotations in rice. *Plant Cel. Physiol.* 52 (2), 220–229. doi:10.1093/pcp/pcq195

Heyer, L. J., Kruglyak, S., and Yooseph, S. (1999). Exploring Expression Data: Identification and Analysis of Coexpressed Genes. *Genome Res.* 9, 1106–1115. doi:10.1101/GR.9.11.1106

Hirose, N., Makita, N., Kojima, M., Kamada-Nobusada, T., and Sakakibara, H. (2007). Overexpression of a Type-A Response Regulator Alters rice Morphology and Cytokinin Metabolism. *Plant Cel Physiol.* 48, 523–539. doi:10.1093/PCP/PCM022

Huang, D., Sherman, B. T., Tan, Q., Collins, J. R., Alvord, W. G., Roayaei, J., et al. (2007). The DAVID Gene Functional Classification Tool: a Novel Biological Module-Centric Algorithm to Functionally Analyze Large Gene Lists. *Genome Biol.* 8 (8), R183. doi:10.1186/GB-2007-8-9-R183

Hudson, N. J., Dalrymple, B. P., and Reverter, A. (2012). Beyond Differential Expression: the Quest for Causal Mutations and Effector Molecules. *BMC Genomics* 13 (13), 356. doi:10.1186/1471-2164-13-356

Jaiswal, P., Ware, D., Ni, J., Chang, K., Zhao, W., Schmidt, S., et al. (2002). Gramene: Development and Integration of Trait and Gene Ontologies for rice. *Comp. Funct. Genomics* 3, 132–136. doi:10.1002/CFG.156

Kastori, R., Plesničar, M., Panković, D., and Sakač, Z. (2008). Photosynthesis, Chlorophyll Fluorescence and Soluble Carbohydrates in sunflower Leaves as Affected by boron Deficiency. *J. Plant Nutr.* 18, 1751. doi:10.1080/01904169509365021

Kim, H., Kim, H., Seong, P., Arora, D., Shin, D., Park, W., et al. (2021). Transcriptomic Response under Heat Stress in Chickens Revealed the Regulation of Genes and Alteration of Metabolism to Maintain Homeostasis. *Animals* 11 (8), 2241. doi:10.3390/ani11082241

Koodathingal, P., Jaffe, N. E., Kraut, D. A., Prakash, S., Fishbain, S., Herman, C., et al. (2009). ATP-dependent Proteases Differ Substantially in Their Ability to Unfold Globular Proteins. *J. Biol. Chem.* 284, 18674–18684. doi:10.1074/JBC.M900783200

Kost, M. A., Perales, H. R., Wijeratne, S., Wijeratne, A. J., Stockinger, E., and Mercer, K. L. (2017). Differentiated Transcriptional Signatures in the maize Landraces of Chiapas, Mexico. *BMC Genomics* 18, 1–14. doi:10.1186/S12864-017-4005-Y

Kubo, M., and Purevdorj, M. (2004). The Future of rice Production and Consumption. *J. Food Distrib. Res.* 35, 128–142. doi:10.22004/AG.ECON.27145

Kumar, A., Kumar, S., Kumar, A., Sharma, N., Sharma, M., Singh, K. P., et al. (2017). Homology Modeling, Molecular Docking and Molecular Dynamics Based Functional Insights into rice Urease Bound to Urea. *Proc. Natl. Acad. Sci. India, Sect. B Biol. Sci.* 88, 1539–1548. doi:10.1007/S40011-017-0898-0

Kumar, R. R., Goswami, S., Sharma, S. K., Kala, Y. K., Rai, G. K., Mishra, D. C., et al. (2015). Harnessing Next Generation Sequencing in Climate Change: RNA-Seq Analysis of Heat Stress-Responsive Genes in Wheat (*Triticum aestivum* L.). *OMICS: A J. Integr. Biol.* 19 (10), 632–647. doi:10.1089/omi.2015.0097

Kurata, N., and Yamazaki, Y. (2006). Oryzabase. An Integrated Biological and Genome Information Database for rice. *Plant Physiol.* 140, 12–17. doi:10.1104/PP.105.063008

Langfelder, P., and Horvath, S. (2007). Eigengene Networks for Studying the Relationships between Co-expression Modules. *BMC Syst. Biol.* 1 (1), 1–17. doi:10.1186/1752-0509-1-54

Langfelder, P., and Horvath, S. (2008). WGCNA: an R Package for Weighted Correlation Network Analysis. *BMC Bioinformatics* 9 (9), 1. doi:10.1186/1471-2105-9-559

Leinonen, R., Sugawara, H., and Shumway, M.International Nucleotide Sequence Database Collaboration (2011). The Sequence Read Archive. *Nucleic Acids Res.* 39, D19–D21. doi:10.1093/NAR/GKQ1019

Lelandais, G., Goudot, C., and Devaux, F. (2011). The Evolution of Gene Expression Regulatory Networks in Yeasts. *Comptes Rendus Biologies* 334, 655–661. doi:10.1016/J.CRVI.2011.05.014

Li, X., Liu, Z., Mi, M., Zhang, C., Xiao, Y., Liu, X., et al. (2019). Identification of Hub Genes and Key Pathways Associated with Angioimmunoblastic T-Cell Lymphoma Using Weighted Gene Co-expression Network Analysis. *Cmar* 11, 5209–5220. doi:10.2147/CMAR.S185030

Ma, Z., and Michailides, T. J. (2005). Genetic Structure of Botrytis Cinerea Populations from Different Host Plants in California. *Plant Dis.* 89, 1083–1089. doi:10.1094/PD-89-1083

Mishra, D. C., Arora, D., Kumar, R. R., Goswami, S., Varshney, S., Budhlakoti, N., et al. (2021). Weighted Gene Co-expression Analysis for Identification of Key Genes Regulating Heat Stress in Wheat. *Cereal Res. Commun.* 49 (1), 73–81. doi:10.1007/s42976-020-00072-7

O'Brien, J. A., and Benková, E. (2013). Cytokinin Cross-Talking during Biotic and Abiotic Stress Responses. *Front. Plant Sci.* 4, 451. doi:10.3389/FPLS.2013.00451

Ouyang, S., Zhu, W., Hamilton, J., Lin, H., Campbell, M., Childs, K., et al. (2007). The TIGR Rice Genome Annotation Resource: Improvements and New Features. *Nucleic Acids Res.* 35, D883–D887. doi:10.1093/NAR/GKL976

Shi, K., Bing, Z. T., Cao, G. Q., Guo, L., Cao, Y. N., Jiang, H. O., et al. (2015). Identify the Signature Genes for Diagnose of Uveal Melanoma by Weight Gene Co-expression Network Analysis. *Int. J. Ophthalmol.* 8, 269–274. doi:10.3980/J.ISSN.2222-3959.2015.02.10

Sircar, S., and Parekh, N. (2015). Functional Characterization of Drought-Responsive Modules and Genes in Oryza Sativa: a Network-Based Approach. *Front. Genet.* 6, 256. doi:10.3389/FGENE.2015.00256

Speidel, J. J. (2000). Environment and Health: 1. Population, Consumption and Human Health. *CMAJ* 163(5): 551–556. Available at: /labs/pmc/articles/PMC80465/(Accessed September 3, 2021).

Stuart, J. M., Segal, E., Koller, D., and Kim, S. K. (2003). A Gene-Coexpression Network for Global Discovery of Conserved Genetic Modules. *Science* 302, 249–255. doi:10.1126/SCIENCE.1087447

Supek, F., Bošnjak, M., Škunca, N., and Šmuc, T. (2011). REVIGO Summarizes and Visualizes Long Lists of Gene Ontology Terms. *PLoS One* 6 (7), e21800. doi:10.1371/journal.pone.0021800

Szklarczyk, D., Morris, J. H., Cook, H., Kuhn, M., Wyder, S., Simonovic, M., et al. (2017). The STRING Database in 2017: Quality-Controlled Protein-Protein Association Networks, Made Broadly Accessible. *Nucleic Acids Res.* 45, D362–D368. doi:10.1093/NAR/GKW937

Tan, M., Cheng, D., Yang, Y., Zhang, G., Qin, M., Chen, J., et al. (2017). Co-expression Network Analysis of the Transcriptomes of rice Roots Exposed to Various Cadmium Stresses Reveals Universal Cadmium-Responsive Genes. *BMC Plant Biol.* 17, 1–18. doi:10.1186/S12870-017-1143-Y

Tang, R., Zhu, W., Song, X., Lin, X., Cai, J., Wang, M., et al. (2016). Genome-Wide Identification and Function Analyses of Heat Shock Transcription Factors in Potato. *Front. Plant Sci.* 7, 490. doi:10.3389/FPLS.2016.00490

Taniguchi, S., Miyoshi, S., Tamaoki, D., Yamada, S., Tanaka, K., Uji, Y., et al. (2014). Isolation of Jasmonate-Induced Sesquiterpene Synthase of rice: Product of Which Has an Antifungal Activity against Magnaporthe Oryzae. *J. Plant Physiol.* 171, 625–632. doi:10.1016/J.JPLPH.2014.01.007

Taylor, N. G., Laurie, S., and Turner, S. R. (2000). Multiple Cellulose Synthase Catalytic Subunits Are Required for Cellulose Synthesis in Arabidopsis. *The Plant Cell* 12, 2529. doi:10.1105/TPC.12.12.252910.2307/3871246

van Dam, S., Võsa, U., van der Graaf, A., Franke, L., and de Magalhães, J. P. (2018). Gene Co-expression Analysis for Functional Classification and Gene-Disease Predictions. *Brief. Bioinform.* 19, bbw139. doi:10.1093/BIB/BBW139

Wasternack, C., and Hause, B. (2013). Jasmonates: Biosynthesis, Perception, Signal Transduction and Action in Plant Stress Response, Growth and Development. An Update to the 2007 Review in Annals of Botany. *Ann. Bot.* 111, 1021–1058. doi:10.1093/AOB/MCT067

Werner, T., Motyka, V., Laucou, V., Smets, R., Van Onckelen, H., and Schmülling, T. (2003). Cytokinin-deficient Transgenic Arabidopsis Plants Show Multiple Developmental Alterations Indicating Opposite Functions of Cytokinins in the Regulation of Shoot and Root Meristem Activity. *Plant Cell* 15, 2532–2550. doi:10.1105/TPC.014928

Wu, D., Zhou, M., Li, L., Leng, X., Zhang, Z., Wang, N., et al. (2019). Time Series Analysis of Gene Changes and Processes after Burn with Human Gene Expression Profiles. *Burns* 45, 387–397. doi:10.1016/j.burns.2018.08.022

Yamada, S., Kano, A., Tamaoki, D., Miyamoto, A., Shishido, H., Miyoshi, S., et al. (2012). Involvement of OsJAZ8 in Jasmonate-Induced Resistance to Bacterial Blight in rice. *Plant Cel Physiol* 53, 2060–2072. doi:10.1093/PCP/PCS145

Zhang, J., Yang, Y., Wang, Y., Zhang, J., Wang, Z., Yin, M., et al. (2011). Identification of Hub Genes Related to the Recovery Phase of Irradiation Injury by Microarray and Integrated Gene Network Analysis. *PLoS One* 6, e24680. doi:10.1371/JOURNAL.PONE.0024680

Zhang, J., Zhao, W., Fu, R., Fu, C., Wang, L., Liu, H., et al. (2018). Comparison of Gene Co-networks Reveals the Molecular Mechanisms of the rice (Oryza Sativa L.) Response to Rhizoctonia solani AG1 IA Infection. *Funct. Integr. Genomics* 18, 545–557. doi:10.1007/S10142-018-0607-Y

Zhang, L., Yu, S., Zuo, K., Luo, L., and 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

Zhao, D., and Oosterhuis, D. M. (2002). Cotton Carbon Exchange, Nonstructural Carbohydrates, and boron Distribution in Tissues during Development of boron Deficiency. *Field Crops Res.* 78, 75–87. doi:10.1016/S0378-4290(02)00095-3

Zhao, S., Guo, Y., Sheng, Q., and Shyr, Y. (2014). Advanced Heat Map and Clustering Analysis Using Heatmap3. *Biomed. Res. Int.* 2014, 1–6. doi:10.1155/2014/986048