



Transcriptomic data-driven discovery of global regulatory features of rice seeds developing under heat stress



Mohammad Mazharul Islam^a, Jaspreet Sandhu^b, Harkamal Walia^b, Rajib Saha^{a,*}

^a Department of Chemical and Biomolecular Engineering, University of Nebraska-Lincoln, Lincoln, NE 68588, USA

^b Department of Agronomy and Horticulture, University of Nebraska-Lincoln, Lincoln, NE 68583, USA

ARTICLE INFO

Article history:

Received 21 July 2020

Received in revised form 10 September 2020

Accepted 11 September 2020

Available online 18 September 2020

Keywords:

Minimal regulatory network

Heat stress

Rice seed development

Global regulation

Optimization

Transcriptomics

ABSTRACT

Plants respond to abiotic stressors through a suite of strategies including differential regulation of stress-responsive genes. Hence, characterizing the influences of the relevant global regulators or on stress-related transcription factors is critical to understand plant stress response. Rice seed development is highly sensitive to elevated temperatures. To elucidate the extent and directional hierarchy of gene regulation in rice seeds under heat stress, we developed and implemented a robust multi-level optimization-based algorithm called Minimal Regulatory Network identifier (MiReN). MiReN could predict the minimal regulatory relationship between a gene and its potential regulators from our temporal transcriptomic dataset. MiReN predictions for global regulators including stress-responsive gene *Slender Rice 1 (SLR1)* and disease resistance gene *XA21* were validated with published literature. It also predicted novel regulatory influences of other major regulators such as Kinesin-like proteins KIN12C and STD1, and WD repeat-containing protein WD40. Out of the 228 stress-responsive transcription factors identified, we predicted *de novo* regulatory influences on three major groups (MADS-box M-type, MYB, and bZIP) and investigated their physiological impacts during stress. Overall, MiReN results can facilitate new experimental studies to enhance our understanding of global regulatory mechanisms triggered during heat stress, which can potentially accelerate the development of stress-tolerant cultivars.

© 2020 The Author(s). Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Rice (*Oryza sativa*) is the primary food crop for more than half of the human population. Rice has a highly mature genomics and genetic toolbox with multiple high-quality finished genomes, 3000 rice resequenced genomes, and efficient genetic engineering and gene editing protocols. Rice yield, nutrient content, and cooking quality are highly sensitive to abiotic stresses. This is particularly true for high temperatures during reproductive development when the grains are transitioning through rapid development and grain weight accumulation [1–5]. This poses a serious threat to global food security with the increasing atmospheric temperature. Therefore, there is a clear need for a comprehensive effort to elucidate the phenotypic differences and/or plasticity of rice under these stressful conditions. Such information can be useful for developing biomarkers for breeding stress tolerant rice germplasm. Since rice shares extensive synteny and collinearity with other grasses and monocot crops [6–8],

fundamental discoveries in rice can also be leveraged for improving other major monocot crops of high economic importance such as maize, wheat and sorghum, as well as candidate energy crops such as switchgrass [9]. However, due to the complex nature of protein translation, folding, and degradation, complex formation, posttranslational modification, allosteric regulation, and substrate availability, the expression patterns at the transcript as well as protein levels and metabolite abundances do not always correlate linearly [10]. This becomes even more complicated since the time frame for imposition of abiotic stresses can vary widely, thus creating stress-specific response thresholds at cellular and whole-plant level. Heat stress typically can occur more rapidly relative to some of the other stressors such as drought and nutrient deficiency, and therefore, mitigates some of the temporal challenges.

A number of studies focused on identifying heat stress-responsive genes and understanding the molecular mechanisms underlying these responses as well as the cellular consequence of these stress responses [11–13]. With the fast progress in molecular biology, significant research effort was geared towards uncovering the adaptive and/or defensive mechanisms at the molecular level in plants in stressed conditions [14,15]. Moreover, the elucidation

* Corresponding author.

E-mail address: rsaha2@unl.edu (R. Saha).

of the topology of genetic regulatory networks using high-throughput omics data is a pressing challenge driving numerous research efforts [16]. For example, Cooper *et al.* focused on developing a network of genes associated with developmental and stress responses by identifying and localizing genes to stress-tolerance trait genetic loci [17]. Dasika *et al.* developed an optimization-based framework to infer time delays in gene regulatory networks from expression data [18]. Chen *et al.* developed BNArray [19], an algorithm implemented in R environment, which statistically evaluates potential high-scoring Bayesian network sets and directed, dense coherent significant sub-networks from microarray data.

A small number of global regulatory genes/proteins usually play a central role in integrating the regulatory architecture of both prokaryotes and eukaryotes and modulates the transcriptional responses to environmental stressors [20–23]. Although previous studies have explored the co-expression patterns of stress-responsive genes, they do not address the hierarchy in regulatory relationships and the role of the major regulatory players. Also, the elucidation of directional hierarchy and the parsimonious nature of the regulatory relationships have not been explicitly addressed yet. Extracting these features requires a careful systems-level analysis of the temporal gene expression profiles. Some recent studies attempted to address this knowledge gap. Zhang *et al.* proposed Context Based Dependency Network (CBDN), a method for reconstructing directed gene regulatory network [24] by evaluating the magnitude of changes in expression dependencies between genes. To identify major transcriptional regulators in rice involved in the metabolic adjustments necessary for adaptation to drought, Mohanty *et al.* used correlative analysis of the patterns of differential spatio-temporal expression profile and cis-element enrichment in leaf, root, and young panicle of rice plants subjected to drought stress during tillering, booting, and panicle elongation stages [25]. In a subsequent work, they identified that crosstalk between a number of key transcription factors and different phytohormones is responsible for survival of an *alcohol dehydrogenase 1 (ADH1)*-deficient rice mutant under complete submergence [26]. Mueller *et al.* developed an optimization-based algorithm to identify the regulatory influence networks for cyanobacterial strains, namely *Synechocystis* PCC 6803 and *Cyanosphaera* ATCC 51142 [27]. While these studies focused on identifying the regulatory effect of known global regulators, a data-driven systematic approach to decipher the molecular mechanisms that drive the expression pattern of these major regulatory players is yet to be developed. Moreover, the overwhelming abundance of multi-level omics data presents a challenge to interpret and incorporate the information into any system-level study of metabolism in an efficient fashion [28,29]. A reasonably robust and minimal model that incorporates the necessary and sufficient set of regulatory influences grounded on experimental data will be critical. Therefore, a clustering-based method to identify global regulators and to define a minimal network involving the global regulators has the potential to capture the emergent biological shifts and stress response mechanisms while being computationally tractable and efficient.

In this work, we attempt to elucidate the heat stress response mechanism in developing rice seeds. To this end, we used differential gene expression analysis on our temporal transcriptomic dataset collected from developing rice seeds under moderate heat stress. This analysis yielded 415 stress-responsive genes as well as 6720 genes differentially expressed across the early developmental stage of the seed. Clustering analysis was then used to develop a minimal gene coexpression network and to identify the highly connected “hub” genes. These genes included many functionally important genes producing putative ribosomal proteins, DNA binding domain containing proteins, histone domain containing proteins, complex carbohydrate synthases, proteinase

inhibitors, Defensin-like (DEFL) family proteins, elongation factors, lipase/acylhydrolases and glycosyl hydrolases, kinesin motor domain containing proteins, protease inhibitors, membrane ATPase/synthases, protein kinases, and several groups of transcription factors related to growth, immune system and biotic and abiotic stress response. Once the candidate regulatory genes were identified, Minimal Regulatory Network identifier (MiReN), a Mixed Integer Linear Programming (MILP) optimization-based tool developed in this work, was implemented to decipher the minimal regulatory relationships. MiReN provides the necessary flexibility in predicting the influence of a master regulator on stress-responsive genes and of a group of genes/regulators on important transcription factor(s). MiReN predictions were validated against published gene regulatory information for multiple global regulators in rice, including the stress-responsive gene *Slender Rice 1 (SLR1)* and the disease resistance gene *XA21*. MiReN was also applied to predict the regulatory influence of several known regulators of seed size and quality, and cell division like the kinesin motor-domain containing proteins KIN12C and STD1, and the WD-repeat containing protein OsWD40. Of a total of 228 stress-responsive rice transcription factors identified, we additionally present MiReN-predicted *de novo* regulatory influences on the three major groups of heat stress responsive transcription factors namely MADS-box M-type, MYB, and bZIP and explain how they potentially impact the physiology of the plant under stress. The novel predictions from MiReN will drive new hypotheses development and experimental designs to further our understanding of plant stress response and subsequently accelerate the identification of genetic intervention strategies for developing stress tolerant crop cultivars.

2. Results

2.1. Differential gene expression under heat stress

Differential gene expression analysis was used to identify 6755 genes from the transcriptomic data, among which 415 were stress responsive (listed in Supplemental information 1). The average read count for each of the genes across all the time points in control and stress samples were then filtered to discard genes with low read count to avoid noise in the transcriptomic data. This yielded a total of 3441 genes of interest that were differentially expressed across the 48-hour period under heat stress. Parametric Gene Set Enrichment Analysis identified 39 Gene Ontology (GO) terms related to cellular processes that are significantly differentially enriched (shown in Fig. 1). Genes involved in photosynthesis, precursor metabolite synthesis, energy metabolism, regulation of cellular components, and response to stress and abiotic stimulus were consistently upregulated under heat stress across all time points. On the other hand, genes related to protein translation, macromolecule synthesis, signaling processes, flowering, reproduction, growth and organismal development, transport, cell cycle, and metabolic processes involving proteins and nucleic acids showed significant downregulation in stressed condition. This set of differentially expressed genes defines the genes of interest for *de novo* identification of their global regulators using the MiReN algorithm.

2.2. Co-expression network and clustering behavior

We next sought to examine the co-expression patterns of the 3441 differentially expressed genes to develop gene coexpression networks, with the goal to identify the highly coexpressed “hub” genes which can be considered as putative global regulators for many stress-responsive genes. We used a threshold of >0.9 on

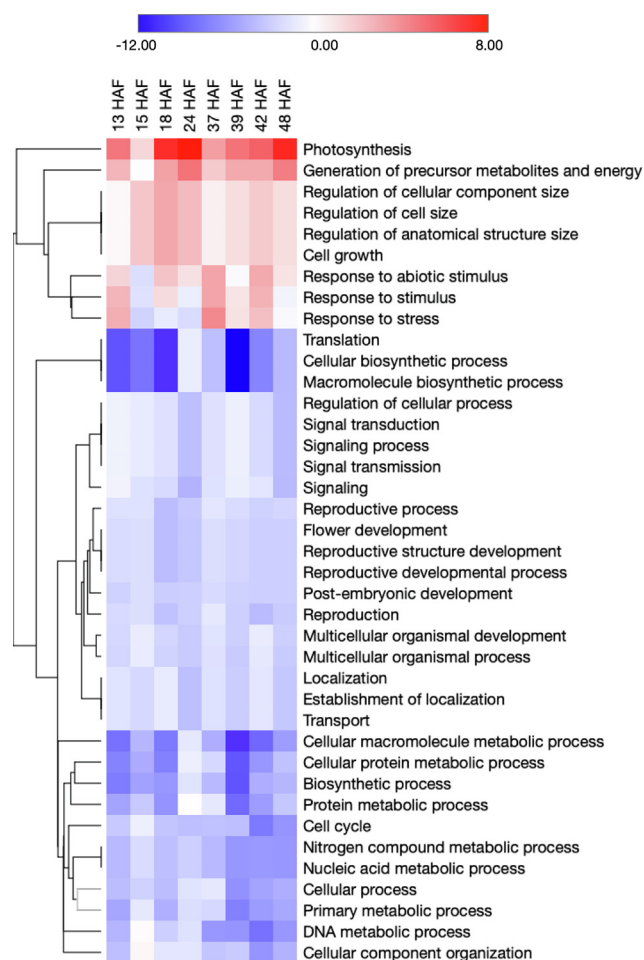


Fig. 1. Enrichment of gene ontology terms related to cellular processes differentially enriched under stress condition. The color scale represents the z-score of differential enrichment in stress vs. control condition.

the Pearson's correlation coefficient with an adjusted p-value of <0.05 for development of the coexpression networks. [Figures S1–S3 \(Supplemental information 2\)](#) shows the coexpression networks in control condition, stress condition, and rearrangements of regulatory relations in the differential coexpression network in stress with respect to control conditions, respectively. [Supplemental information 3](#) contains the full coexpression analysis results.

Based on the cluster size (see methods), a total of 267 genes in control condition and 409 genes in stress condition were identified as “hub” genes. These highly connected genes are used in the MiReN framework as the potential global regulatory players for the target genes of interest identified in the previous section. These “hub” genes included previously identified MADS-box genes [30] as well as a large number of genes with putative functions ([Supplemental information 3](#)). The most dominant groups of hub genes were responsible for the production of putative ribosomal proteins, DNA binding domain containing proteins, histone domain containing proteins, complex carbohydrate synthases, proteinase inhibitors, DEFL family proteins, elongation factors, lipase/acylhydrolases and glycosyl hydrolases, kinesin motor domain containing proteins, protease inhibitors, membrane ATPase/synthases, protein kinases, and several groups of transcription factors related to growth, immune system and biotic and abiotic stress response. While there is significant overlap of the nodes and edges in control and stress conditions, there is also noticeable separation of the clusters between the two conditions ([supplemental Figures](#)

[S1–S3](#)). Overall, the major cluster uniquely identified in stress condition includes genes for brassinosteroid signaling, calmodulin-dependent protein kinases, histone domain containing proteins, cyclin-dependent kinases, endonuclease, elongation factors, kinesin motor domain containing proteins, F-box domain containing proteins, ribosomal proteins L4, L5, L7/L12, L29 and S13p/ S18e, and WD repeat-containing proteins.

2.3. MiReN-predicted regulatory influences of known regulators in rice

MiReN, the multi-objective Mixed-Integer Linear Programming (MILP) optimization tool developed in this work (see Methods), was applied to predict the regulatory networks for several of the highly connected hub genes from the coexpression analysis. These hub genes include many known regulators related to stress- and immune-response, seed size and quality, and cell division [31–41]. MiReN-predicted regulatory networks for some of the known regulators (e.g., *Slender Rice 1*, *SLR1* and the disease resistance gene *XA21*) were validated to test the performance of the optimization framework in predicting regulations correctly.

MiReN predicted negative interaction coefficients for each of the genes in the GA-responsive gene family ([Fig. 2A](#)), as identified by Jan and Komatsu 2006 [42], which implies a repressive action. The *Slender Rice 1* (*SLR1*) gene and the corresponding Rice DELLA protein *SLR1* are responsible for the repression of several Gibberellic acid response genes in rice [34,42]. These genes play crucial roles in controlling plant height [43,44], regulates stomatal development and patterning [45], modules the elongation of pedicels and/or secondary branches [46,47], and sometimes enhances the resistance to insects [48] in rice. The *slr1-1* mutant is a constitutive gibberellin (GA) response phenotype that elongates as if saturated with GAs [49]. On the other hand, rice stress interactome shows that *XA21*, an immune receptor that confers resistance to the bacterial blight disease [50,51], was intimately connected to *WRKY* and *SnRK* regulators [31]. *OsWRKY62* and *OsWRKY76* act as negative regulators of *XA21*-mediated immunity [32], while *OsWRKY71* is a positive regulator [33]. Other studies [34] reported that the regulation is ill-defined, without any convincing proof of positive or negative regulations. MiReN predicts that *XA21* directly positively regulates *WRKY71* and *WRKY76* ([Fig. 2B](#)), instead of through *WRKY62*. Also, *XA21* slightly inhibits another loci of *SnRK1* (*LOC_Os03g63450*).

Two major kinesin-motor domain containing proteins that were identified as putative global regulators in our coexpression analysis are *KIN12C* and *STD1*. The influences of these two regulators on several important stress-responsive genes are shown in [Fig. 2C](#). Both of these proteins regulate important biological processes via RNA polymerases, ethylene-responsive transcription factor, and bZIP transcription factors. MiReN predicted that both of these regulators positively regulate ethylene-responsive transcription factors *LOC_Os09g20350* and *LOC_Os02g51670*, which are members of the AP2-EREBP family and are key regulators of seed developmental processes, seed shattering and seed size; and *LOC_Os09g35920*, a mediator of RNA polymerase II transcription subunit 10 [36,41].

In [Fig. 2D](#), the genes positively regulated by another putative global regulator, cell division cycle gene *OsWD40* (*LOC_Os02g47180*), are shown. *OsWD40* performs diversified biological functions in rice, including seed growth and development [38]. The genes that are the most strongly regulated by *WD40* are ethylene-responsive transcription factors *LOC_Os09g20350* and *LOC_Os02g51670*; and MYB family transcription factor *LOC_Os11g10130*. Other positively regulated genes are related to histone-mediated post-translational modifications, and reproductive pathways.

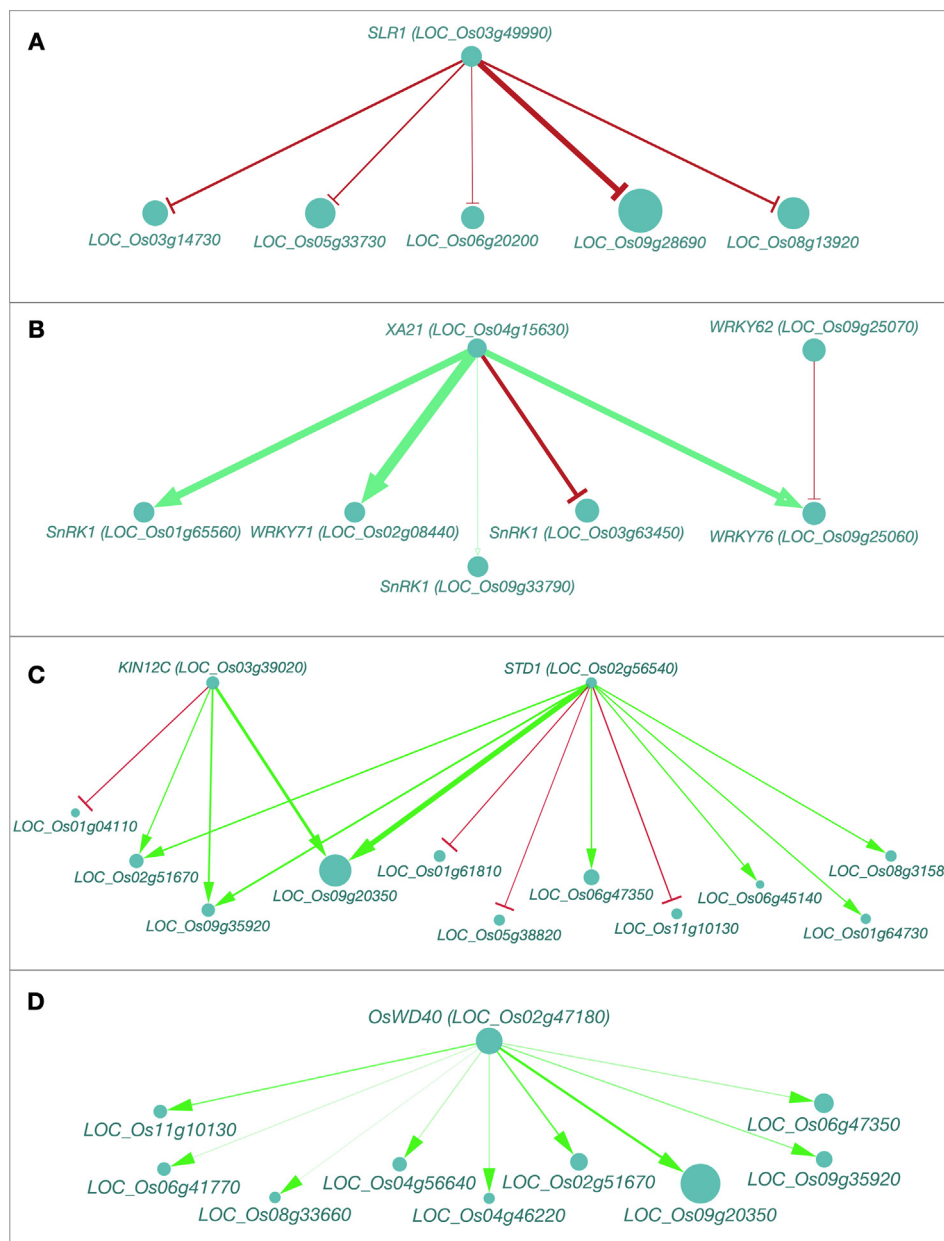


Fig. 2. MiReN-predicted and experimentally identified regulatory networks for (A) *Slender Rice 1*, (B) Immune receptor XA21 (C) Kinesin motor-domain containing proteins KIN12C and STD1, and (D) WD-repeat containing protein OsWD40. The size of the nodes represents the average expression value of the genes. The green and red colors of the edges represent positive (activation) and negative (repression) regulation, respectively. The thickness of the edge lines represents the magnitude of the regulation. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

2.4. MiReN-predicted de novo regulatory influences on stress-responsive rice transcription factors

In addition to predicting the regulatory influences of known or putative global regulators on other genes, MiReN can predict novel regulatory influences on any gene of interest. From the transcriptomic analysis, a total of 228 heat-responsive transcription factors were identified, which were then grouped according to their functional role in rice seed metabolism. The major groups were transcription elongation factors, ethylene response genes, histone-like proteins, calmodulin-binding proteins, RNA polymerase, and bZIP, MYB and MADS-Box transcription factors. MiReN-predicted regulatory networks for three major groups of heat stress responsive transcription factors (MADS-box M-type, MYB, and bZIP) are presented below and shown in Fig. 3, with the complete MiReN results presented in Supplemental information 4.

The MiReN-predicted regulatory network for three MADS-box M-type genes of interest (*OsMADS64*, locus ID *LOC_Os04g31804*; *OsMADS82*, locus ID *LOC_Os04g24800*; and *OsMADS77*, locus ID *LOC_Os09g02780*) is shown in Fig. 3A. Most influential positive regulators of MADS-box M-type transcription factors (have a positive interaction coefficient) include F-box domain and kelch repeat containing proteins, protein and peptide transporters, RNA recognition motifs, transcriptional repressors, retrotransposons, cyclin-dependent kinases, energy molecule producers, galactosyltransferase, and calcium/calmodulin-binding proteins. Major repressors (having a negative interaction coefficient) of MADS-box M-type transcription factors include NOP2 Nucleolar Proteins, sucrose transporters, ATP/GTP-binding proteins, mitogen-activated protein kinase MKK1, salt and drought induced proteins, and regulators of chromosome condensation domain containing proteins. The MiReN-predicted regulatory network for five MYB transcription

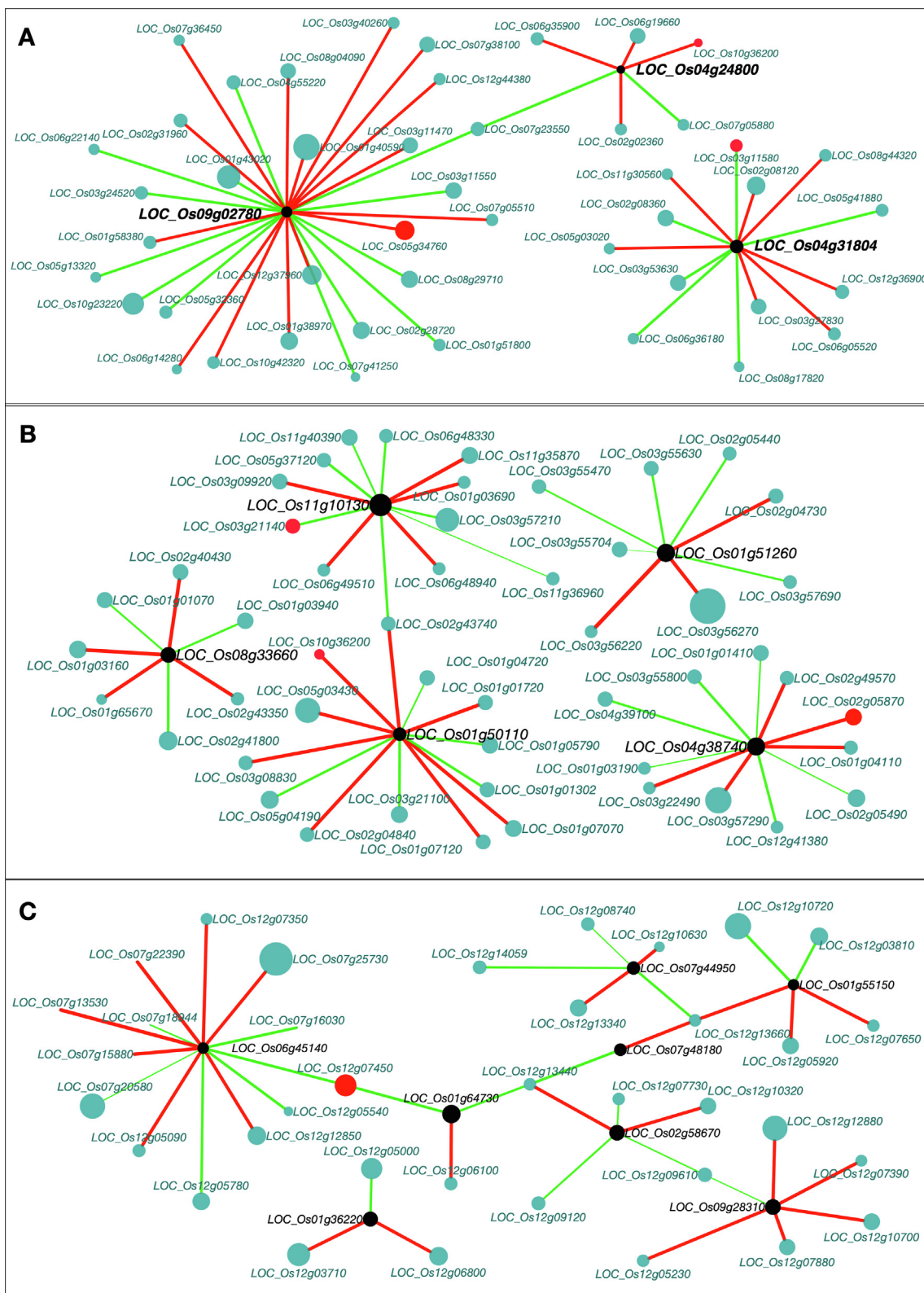


Fig. 3. Minimal regulatory networks involving the MADS-Box M-type (A), MYB (B), and bZIP (C) genes in control and stress conditions. The black nodes are the genes of interest (MADS-Box M-type genes, MYB, or bZIP genes). The size of the node circles corresponds to the average expression value of the genes. The edge colors are green for positive regulation and red for negative regulation; the width of the edges correspond to the magnitude of the regulation coefficient. The “hub” genes identified in this work are highlighted as red nodes. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

factors (locus IDs *LOC_Os01g51260*, *LOC_Os11g10130*, *LOC_Os08g33660*, *LOC_Os04g38740*, and *LOC_Os01g50110*) is shown in Fig. 3B. Most influential positive regulators of MYB transcription

factors include DnaJ and DnaK domain containing proteins, C-type lectin domain-containing proteins, abscisic aldehyde oxidase, and other signaling proteins. On the other hand, several stress induced

proteins, glycosyl transferase family 14 proteins, peroxisomal membrane anchor protein (Pex14p) domain containing protein, OsFtsH7 (FtsH protease), RWD domain containing protein, and HEAT repeat family proteins were predicted to be the most influential negative regulators of MYB transcription factors. Transcription factors in the bZIP group are mostly regulated by proteins involved in nuclear localization, transcription, DNA repair, RNA processing, epigenetic regulation (*LOC_Os02g05660*, *LOC_Os07g43316*, *LOC_Os02g05470*), phosphatase and decarboxylase proteins (see Fig. 3C).

3. Discussion

Temperature is one of the primary drivers of growth and development for crops [52]. However, crop yield and grain quality at maturation is adversely affected by heat stress during early developmental stage. Our results indicate a genome-wide modulation of gene expression when rice seeds are subjected to moderate heat stress. The complexity of the differential expression levels across the whole genome put a challenge in understanding stress response in a systematic fashion, which led to our efforts in developing MiReN. It is a mixed integer linear programming optimization-based framework to decipher the minimal regulatory network and the hierarchy in regulatory relationships among stress-responsive genes. MiReN attempts to minimize the number of regulatory influence (activation/repression) on a gene/transcription factor while minimizing the deviation from the transcriptomic experiment data. It quantifies the regulatory influence (interaction coefficient in the mathematical formulation) of a global regulator/gene on a gene or transcription factor, thereby allowing the ranking of regulatory influences of global regulators on genes and choose the most important ones. The biggest strength of a framework of this nature is the bidirectional discovery of regulatory information: either to infer the influence of a global regulator on downstream genes or to find the most important regulators of genes of interest. In this work, MIREN predicted the regulatory influences of known (*Slender Rice 1 (SLR1)* and the disease resistance gene *XA21*) and putative (Kinesin-like proteins *KIN12C* and *STD1* and WD repeat-containing protein *WD40*) global regulators. The results for the known ones were further validated with published literature. In addition, with MIREN we predicted *de novo* regulatory influences on three major groups of transcription factors (i.e., MADS-box M-type, MYB, and bZIP) and investigated their physiological impacts on the plant under stress.

The differential gene expression analysis and the coexpression analysis were used as preprocessing tools leading to the MiReN framework. The genes with high clustering coefficients in the coexpression network were used as the candidate global regulators in MiReN, while the genes identified from the differential expression analysis are the target genes for which we predict the global regulatory influences. We also verified the expression patterns of two known rice seed regulators, *OsFIE1 (LOC_Os08g04290)* and *OsMADS87 (LOC_Os03g38610)* by real-time quantitative PCR for control samples (see Supplemental information 5). While the regulators involved in seed and spikelet development are of importance, they didn't appear to be significantly differentially expressed at the early developmental stage in our samples (see Supplemental information 1). From the differential gene expression analysis, we observed that biological processes including photosynthesis, precursor metabolite synthesis, energy metabolism, regulation of cellular components, and response to stress and abiotic stimulus were consistently upregulated under heat stress. These upregulations support prior observation that moderate heat stress during early seed development improved seed germination rate and advanced seedling establishment, suggesting that

short-term episodes of elevated temperature play a positive role in enhancing seed performance [53,54]. The early heat stress response in rice seed also triggers protection mechanisms against cell or tissue damage from exposure to high temperatures [55,56]. For example, Heat Shock Factor (HSF) genes were observed to be differentially overexpressed in our study, similar to the observations in other plant species in prior studies [57–61]. This facilitates the acquisition of an autonomous thermo-tolerance caused by a moderate/sub-lethal heat stress to overcome subsequent exposure to higher temperatures [62–64].

On the other hand, genes related to protein translation, macromolecule synthesis, signaling processes, flowering, reproduction, growth and organismal development, transport, cell cycle, and metabolic processes involving proteins and nucleic acids showed significant down-regulation in the stressed condition. Many of the genes in the pathways from starch metabolism (especially those encoding the subunits of starch synthase) were observed to be downregulated at high temperature, which can potentially result in poor grain quality and reduced size and weight, as suggested by previous studies [65–68]. A number of genes involved in the biosynthesis of Gibberellic acid (for example *OsGA2OX3*, *OsCPS2*, and *OsCPS4*) showed reduced expression level under heat stress, which is consistent with previous reports [54]. The repression of Gibberellic acid production can have drastic effects on development of the seed and the plant overall. Our results indicate a consistent repression of the genes producing 14-3-3 proteins under heat stress, which function as signal regulators and modulate nitrate reductase under limiting carbon and nitrogen conditions and play a crucial role in the modulation of ATPase/synthase complex and plasma membrane expansion [69,70] in stressed conditions. The repression of 14-3-3 proteins in turn relaxes the regulation on ATPase production, as evident from the upregulation of membrane ATPase/synthase observed in heat stress from our results. Many of the genes that were found to be significantly heat-responsive in our study were also associated with responses to other abiotic stresses such as drought, salt, cold, and submergence, as found in other studies [11,71]. For example, genes for salt stress responsive proteins (Salt stress root protein *RS1*, *LOC_Os01g13210*; DREPP plasma membrane polypeptide family protein, *LOC_Os02g18410*), abscisic acid-responsive proteins (*LOC_Os11g06720*), wound/stress protein (*LOC_Os02g51710*) and zinc finger A20 and AN1 domain-containing stress-associated proteins (*LOC_Os06g41010*) were identified as heat-sensitive in this study. In addition, six universal stress protein-producing genes were differentially expressed in our study, which indicates that there is significant overlap of responses under multiple different abiotic stresses.

The co-expression and clustering analysis provided important insights about the interactions within the groups of differentially expressed genes during the early developmental stage of rice seed under heat stress. For example, one major cluster uniquely detected in the coexpression network in stress condition involved the “hub” genes that are responsible for brassinosteroid insensitive 1-associated kinase 1 (*Osl-BAK1*), which is associated with grain filling and seed development in rice [72]. The brassinosteroid class of steroid hormones regulates plant development and physiology, including controlling division, elongation and differentiation of various cell types [73]. BR-mediated signaling has been demonstrated to play a role in the adaptation to biotic stresses such as insect and microbial attack, and abiotic stresses such as drought, temperature changes, and salinity via crosstalk with other hormones such as ABA, GA, auxin, cytokinin, jasmonic acid (JA), salicylic acid (SA), and ethylene [74–77]. The clustering of *Osl-BAK1* with *WRKY* and *MYB* transcription factors and zinc-finger proteins in our differential coexpression network supports these findings, since at elevated temperatures, repression of brassinosteroid

insensitive 1-associated kinases affects the biochemical activity and induced protein misfolding and degradation [78–80], which can primarily impede seed development. These highly connected genes in the coexpression network are hypothesized as potential global regulators, whose regulatory influence on stress responsive genes in developing rice seed could be further explored with MiReN.

The ability of MiReN to identify the hierarchy of the regulatory influences among the complex gene network allows for parsimoniously pinpointing the source of the response of a group of genes when the plant is subjected to stress. Therefore, engineering interventions designed for address stress-induced effects in plants can be focused toward a lower number of highly influential global regulators instead of a vast number of genes. While a coexpression network can only describe the patterns in differential expressions, MiReN is a way forward in deciphering the hierarchy in gene regulation. In this study, MiReN could successfully predict known small regulatory interaction networks for multiple global regulators in rice as identified from our coexpression analysis, e.g., the stress-responsive gene Slender Rice 1 (*SLR1*) and the disease resistance gene *XA21*. It should be noted that the *SLR1* gene and the *SLR1*-repressed GA pathway have circadian regulation, which is also demonstrated by the expression pattern of 39 genes over the course of the day (starting from sunrise to sunset) and transcription factors regulated by *SLR1* in our dataset. For example, the light-induced Rice1 regulator (*LIR1*) that regulates the attachment of leaf-type ferredoxin-NADP⁺ oxidoreductase to the thylakoid membrane in rice [81] was found to be increasing in expression throughout the day, with the maximum expression at the end of the day (24, 48 HAF). In addition, the circadian rhythm observed in the regulation of chromosomal condensation is important to drive oscillations in gene expression and determine the circadian transcriptional output [82]. Circadian rhythm of gene expression is also (positively or negatively) regulated by multiple transcription factors including MYB, bZIP, and ethylene-responsive transcription factors, which also showed oscillatory behavior in our experiment [83–85].

MiReN was also applied to predict the regulatory networks for several known regulators of seed size and quality, and cell division like the kinesin motor-domain containing proteins KIN12C and STD1; and the WD-repeat containing protein OsWD40. Kinesin-like protein KIN12C controls microtubule-based movements in the cytoskeleton and mediates the developmental processes of male gametophyte, embryo, seedling, and seeds. It also participates in critical cellular events during cell division and lengthening [35,39,40]. On the other hand, *STD1* (*stemless dwarf 1*) encodes a phragmoplast-associated kinesin-related protein that is highly expressed in the actively dividing tissues like developing seeds and plays important roles in the regulation of cell division and plant development [37]. Both the kinesin-like proteins and the WD-repeat containing proteins were observed to be downregulated under heat stress in our transcriptomic experiment while the genes involved in the regulation of cell growth and size were overexpressed (see Fig. 1). While these regulators have been known to regulate important physiological processes involved in cell division, cell cycling, seed size and maturity, the regulatory influence of these regulators in specific genes or transcription factors could only be hypothesized or predicted from our current study. The positive regulation of these two groups of regulators on genes related to developmental processes and seed size explains why seed development is highly likely to be hampered under elevated temperatures.

In addition, we predicted the minimal regulatory networks for several stress responsive transcription factors (including MADS-box M-type, MYB, and bZIP) that control various biophysical functions important for growth, development and immunity. The

minimal regulatory networks for these stress-responsive transcription factors (Supplementary Information 4) provide useful insights into the regulatory influences affecting the physiology of developing rice seed under heat stress.

MADS-box transcription factors are responsible for the regulation of developmental process while acting in conjunction with other unrelated transcription factors. Many MADS-box transcription factors including the M-type and MIFCc-types (i.e., MADS 2, 3, 6, 14, 21, 27, 56, 64, 77, 82) were found to show significant differential expression under heat stress in our study (see Fig. 1 which shows the downregulation of reproductive, embryonic, and organismal development processes under stress). Similar observation was confirmed by previous experimental findings [17,86], which reflects on their importance in the growth and development of rice seed. The positive and negative regulators of MADS-box transcription factors identified by MiReN (see results section) have numerous crucial roles in signal transduction, grain size and quality, regulation of cell division and cell cycle, gene silencing, reproductive growth and morphology, food storage, evolution, and stress tolerance [30,87–92].

MYB group of transcription factors render tolerance to elevated temperatures and modulates amino acid metabolism in plants [93]. Consistent with their metabolic and regulatory functions related to stimulus response, many MYB genes showed differential overexpression in our experiment (see Fig. 1). The positive and negative regulators of MYB transcription factors predicted by MiReN play important roles in heat shock, cell–cell communication, development, abiotic stress response, seed maturation, and defense mechanisms. For example, DnaJ and DnaK chaperon proteins are well known proteins that were previously identified to have a heat shock element in their promoter regions [94]. Lectin Receptor-Like Kinases are fundamental to plant life and have important roles in cell-to-cell communication; development of cell wall structure, cell growth, and defense strategies [95]. These are also involved in abiotic stress responses, including salinity and osmotic stress [96], not only in rice but also in other monocot plants like wheat [97]. Phytohormone abscisic acid (ABA) regulates numerous growth and developmental processes including seed maturation, germination and response to abiotic environmental cues [98]. Abscisic aldehyde oxidase is an important step in ABA biosynthesis and degradation that ensures ABA homeostasis during drought stress and seed development [99]. Glycosyltransferase proteins regulate programmed cell death [100] and participate in important biological processes including hormone homeostasis, flower and fruit pigmentation, cell wall development, and defense responses [101,102], which are sensitive to abiotic stresses and therefore negatively affect the heat-sensitive MYB transcription factors. Peroxisomal membrane anchor protein like Pex14p are significantly induced by exogenous H₂O₂ and are important for producing fertile offspring [103–106]. FtsH Protease Is Involved in Development, Oxidative and light stress Response and Heat Shock Control in plants [107].

bZIP transcription factors regulate a number of plant processes such as seed development, light signaling, floral induction and flower development, biotic and abiotic stresses, ABA signaling, and hormonal response [108–111]. These are regulated by a number of proteins involved in DNA and RNA metabolism. For example, one of the most influential positive regulators, the DEAD-box RNA helicase family proteins, control RNA metabolism and a range of cellular functions including response to abiotic stresses [112]. SAP domain-containing proteins regulate stress signaling by modulating the expression of endogenous stress-related genes [113]. Therefore, it is expected that while subjected to elevated temperatures, these regulators would activate the stress-responsive pathways in seed, which is also evident from the upregulation of stress-responsive biological process (see Fig. 1). On the other hand,

small GTPases were predicted to be a repressor of bZIP genes, which are key regulator do cellular transformation and are involved in controlling cell growth and shape and [114]. Although a positive or negative regulatory mechanism is not identified in literature, MiReN results show strong repression of the bZIP family by these proteins.

We propose MiReN as a powerful and highly predictive hypothesis-generation tool to further our understanding of the global regulatory mechanism in play. While the current study has focused on developing rice seeds under moderate heat stress, MiReN, as a toolbox, can be applied to any plant species and any kind of differential growth conditions. The prediction capability of MiReN relies on the quality and temporal/spatial resolution of the transcriptomic data. Since the inner objective in the MiReN formulation is to minimize the discrepancy of optimal regulatory influence with expression data for the genes and regulators, MiReN is benefitted when high-resolution expression data is available. It should also be noted that utilizing MiReN framework for a large set of genes is challenging due to the polynomial scaling of the optimization problem to be solved. One approach used to overcome this issue in the current study is high degree of parallelization during computation. Furthermore, the inner level and outer level objective functions in MiReN are sensitive to the magnitude of gene expression values as well as the slack variables used to model the discrepancy with experimental data. As a result, MiReN results should be interpreted not in terms of the absolute values of the interaction coefficients, but rather as a ranking from the highest to the lowest regulatory impact on the gene(s) of interest. For example, MiReN-predicted regulatory network for Immune receptor XA21 should be interpreted as *WRKY71* being the most influenced by *XA21*, but not necessarily changing its expression 20.4 times as of *XA21*. One particular advantage of the MiReN framework is that it can be customized to suit the analysis of regulatory influences on any molecular regulation level (i.e., gene expression, protein abundance etc.). MiReN also is benefitted from greater number of data points during transcriptomic or proteomic experiments because the optimization algorithm behaves more robustly when more data is available. Therefore, availability of multi-level and high-resolution experimental omics data will facilitate the generation of high-quality predictions which can be easily integrated into any systems-level study of stress response.

4. Conclusions

We currently lack a genome-scale understanding of how increased temperature shapes the transcriptome of a rapidly growing young seed. Given the impact of heat stress on final seed size and grain quality, a better understanding of how the critical transition from the coenocytic to the cellularized state of endosperm is affected by heat stress at the molecular level is essential for identifying targets for genetic edits that can accelerate the development of more stress-tolerant plant cultivars. In addition, while the current study focused on the heat stress response in developing rice seed, it should be highlighted that globally rice is affected by other abiotic stresses including but not limited to, drought, salinity, flooding/submergence, cold, low/high nitrogen and phosphorus etc. Due to the complex nature of the physiological changes in multiple molecular levels, the signal transduction genes and stress-tolerance gene often demonstrate overlapping functionalities when plants are subjected to abiotic stressors [115–118]. For example, the ethylene-response transcription factors and the GA pathway regulators that was found to be stress-responsive in the current study were also reported to be playing important role in regulating gene expression during submergence [119]. It is important to highlight the MiReN, being a generalized algorithm, can be

useful for any level of omics' data generated from experiments. Being a flexible tool, MiReN can facilitate both the discovery of global regulatory players during stress or the unknown metabolic modulations caused by known global regulators. It can also serve as a key tool for computational mathematical frameworks to incorporate information from multi-level 'omics' datasets into genome-scale models of plant metabolism to elucidate limiting reaction step(s) or pathway(s) related to multiple stress tolerance and eventually identify and propose a suit of strategies (i.e., gene up- or down-regulation or mutation) to develop rice varieties that are more resilient during this transient but highly sensitive window of reproductive development. These genetic intervention strategies can also be applicable to other major monocot species such as wheat and maize.

5. Materials and methods

5.1. Data collection

Rice (*Oryza sativa*) cv. 'Kitaake' plants were grown in 4-inch square pots under optimum greenhouse conditions (16-h-light/8-h-dark cycle at 28 °C/25 °C) until flowering. At flowering, uniformly looking plants were moved to in percival growth chambers maintained at control conditions (28/25 °C, 16/8h light/dark). The spikelets were marked at time of fertilization on two consecutive days and plants were maintained at control growth chamber conditions. To investigate the transcriptional responses of rice early seed development under heat stress, we exposed the plants to moderate heat stress (35 °C). Previous studies show that exposure to short-term moderate heat stress (35 °C) during early seed development reduces the seed size at maturity [30]. Therefore, we selected 35 °C for simulating the moderate heat stress.

When the marked seeds/spikelets reached to 36 h after fertilization (HAF) and 12 HAF, half of plants were moved to stress chamber (35 °C). The developing seeds without husk were collected from control and stress plants at time points indicated in Fig. 4. The time points were decided with an idea to capture early (i.e. within few hours after imposing the stress) heat stress responses. Two biological replicates were collected for each of control and stress time-point. For each sample, developing seeds were pooled from 2 to 3 plants. Total RNA isolated using RNAeasy mini-elute kit (Qiagen) was sequenced using illumina-sequencing single end 100 bp reads. RNA-seq analysis including trimming low-quality reads, read alignment and read counting was performed as described in Chen et al., 2016 [30]. Trimmomatic was used to remove low quality reads and trimmed reads were aligned to rice MSU (v 7.0) genome using TopHat [120,121]. To get number of reads per gene, HTSeq-count was used in "union" resolution mode [122].

For RT-qPCR, a 20 µl reverse transcription reaction was performed to synthesize cDNA (iScript cDNA Synthesis kit, Invitrogen). RT-PCR was performed using *OsFIE1* and *OsMADS87* specific primers (see Supplemental information 5) (IQ SYBR Green Supermix, Bio-Rad). Relative expression was calculated using Delta-Delta-Cq method [123].

5.2. Differential expression analysis

Differential gene expression analysis was used to identify the stress responsive genes from the transcriptomic data as the first step of the data analysis (see Fig. 5). The Reads Per Kilobase of transcript, per Million mapped reads (RPKM) values were filtered with a cutoff of 20 reads (maximum among all conditions and time points). In addition, the genes with a standard deviation in gene expression levels (normalized read count values) less than 2%

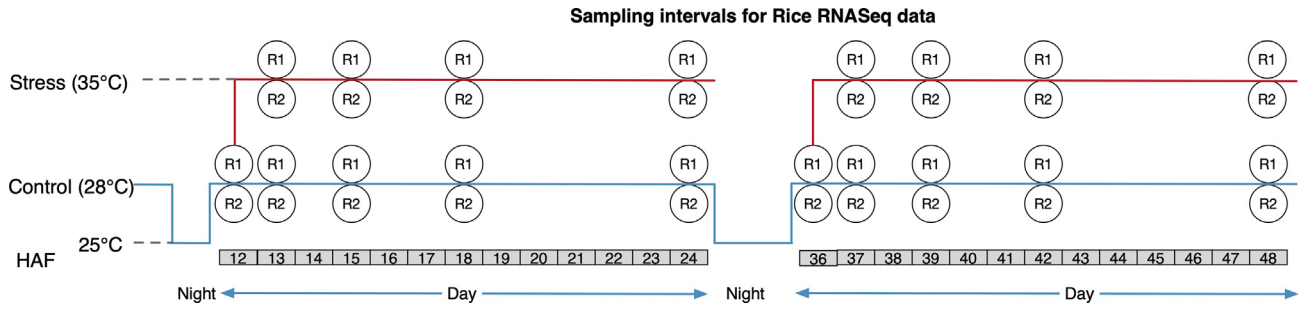


Fig. 4. Transcriptomic experimental design of control and stressed samples of developing rice seed. The collected time points in Hours After Fertilization (HAF) are indicated at the bottom. The temperature profile for the control and stress samples are indicated by blue and red lines, respectively. For each control and stress samples at every time points, two biological replicates (indicated by “R”) were collected. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

across all data points were discarded from the analysis. The DESeq algorithm in R software package “Bioconductor” was used for differential gene expression analysis [124]. DESeq employs negative binomial distribution and a shrinkage estimator for the distribution’s variance methods to test for differential expression [124]. The average of the raw read counts from two replicates were used to calculate the fold change and the $\log_2(\text{foldchange})$ of the genes. Genes with a $\log_2(\text{foldchange})$ value of 1 or higher were considered overexpressed and genes with a $\log_2(\text{foldchange})$ value of -1 or lower were considered underexpressed, while satisfying an adjusted p-value of <0.05 [125]. GO functional enrichment analysis was performed to identify significantly enriched biological processes and molecular functions in stress condition. The GO analysis was carried out by AgriGO software with $\text{FDR} \leq 0.05$ based on biological process (BP), molecular function (MF) and cellular component (CC) GO terms [126,127]. Heatmap of gene set enrichment was developed in Morpheus (<https://software.broadinstitute.org/morpheus/>) using the z-score of differential enrichment for each GO term in the stressed sample compared to the control sample at every time point.

Co-expression networks were created based on the correlation coefficients calculated in Matlab version 2018b (Mathworks Inc.) using the Pearson Correlation method with an adjusted p-value of <0.05 for testing the hypothesis of no correlation against the alternative hypothesis of a nonzero correlation. Genes with a conservative absolute value of Pearson’s correlation coefficient >0.9 were considered to be present in the coexpression network to obtain a set of coexpressed genes with very high correlation, similar to usual practice in recent studies [128–131]. Genes with a cluster size of >10 in the coexpression networks were considered as potential “hub” genes for minimal regulatory network analysis (as illustrated in Fig. 5).

5.3. Application of MiReN optimization framework to identify minimal regulatory networks

MiReN (Minimal Regulatory Network identifier) is an optimization-based algorithm to predict a minimal regulatory relationship between a particular genetic entity and the potential regulatory entities form a temporal transcriptomic data. The Mixed Integer Linear Programming (MILP) mathematical formulation for MiReN is given below.

$$\text{minimize } D_{i,t}^+, D_{i,t}^-, A_{i,i'}, y_{i,i} \sum_i \sum_{i' (i \neq i')} y_{i,i} + \sum_i \sum_t (D_{i,t}^+ + D_{i,t}^-) \quad (1)$$

subject to,

$$\sum_i y_{i,i} \leq \text{cutoff} \quad \forall i \in I \quad (2)$$

$$D_{i,t}^+ - D_{i,t}^- + \Delta t \sum_i A_{i,i'} X_{i',t} = X_{i,t+1} - X_{i,t} \quad \forall i \in I, \forall t \in T - 1 \quad (3)$$

$$LB_{i,i'} y_{i,i} \leq A_{i,i'} \leq UB_{i,i'} y_{i,i} \quad \forall i \in I, \forall i' \in I \quad (4)$$

$$D_{i,t}^+, D_{i,t}^- \geq 0 \quad \forall i \in I, \forall t \in T \quad (5)$$

$$A_{i,i'} \in \mathfrak{R} \quad \forall i \in I, \forall i' \in I \quad (6)$$

$$y_{i,i} \in \{0, 1\} \quad \forall i \in I, \forall i' \in I \quad (7)$$

The set I is the set of all genes/proteins in the regulatory network, for which expression data is available from experiment. This formulation is initially blind to which of these genes/proteins are acting as regulators of other genes/proteins. Therefore, the set of genes are a superset of regulators, but they are not identified beforehand. The formulation assumes interaction between each gene/protein pair i and i' and defines their interaction coefficient as $A_{i,i'}$. The sign and magnitude of $A_{i,i'}$ describes the type (activation/inhibition) and extent of the regulation of gene/protein i by gene/protein i' , respectively. The variable $y_{i,i}$ is defined as the interaction identifier, which is a binary variable, assuming a value of 0 if there is not interaction between i and i' . $D_{i,t}^+$ and $D_{i,t}^-$ are the slack variables that describes the deviation of the model predicted expression value of gene i at time t and the experimental values ($X_{i,t}$). $LB_{i,i'}$ and $UB_{i,i'}$ are arbitrary lower and upper bounds for the interaction coefficient. The optimization algorithm was solved using the General Algebraic Modeling System (GAMS) version 24.7.4 with IBM CPLEX solver on the High-performance cluster computing system at the University of Nebraska-Lincoln.

MiReN can also be formulated with L1 regularization as an additional penalty term in the objective function. The Mixed Integer Linear Programming (MILP) mathematical formulation for MiReN with L1 regularization is presented in Supplemental information 2. While it possible that multiple different combinations of highly impactful regulators on a set of genes can yield different networks, the MiReN framework with L0 regularization imposes a strong penalty and avoids the issue. From paired sample statistical analysis of a sample case (results in Supplemental information 6) with low (0.001), medium (1.0), and high (1000) values of the penalization in L1 regularized formulation of MiReN, a significant change in the values of the regulation coefficients was not observed compared to the L0 formulation.

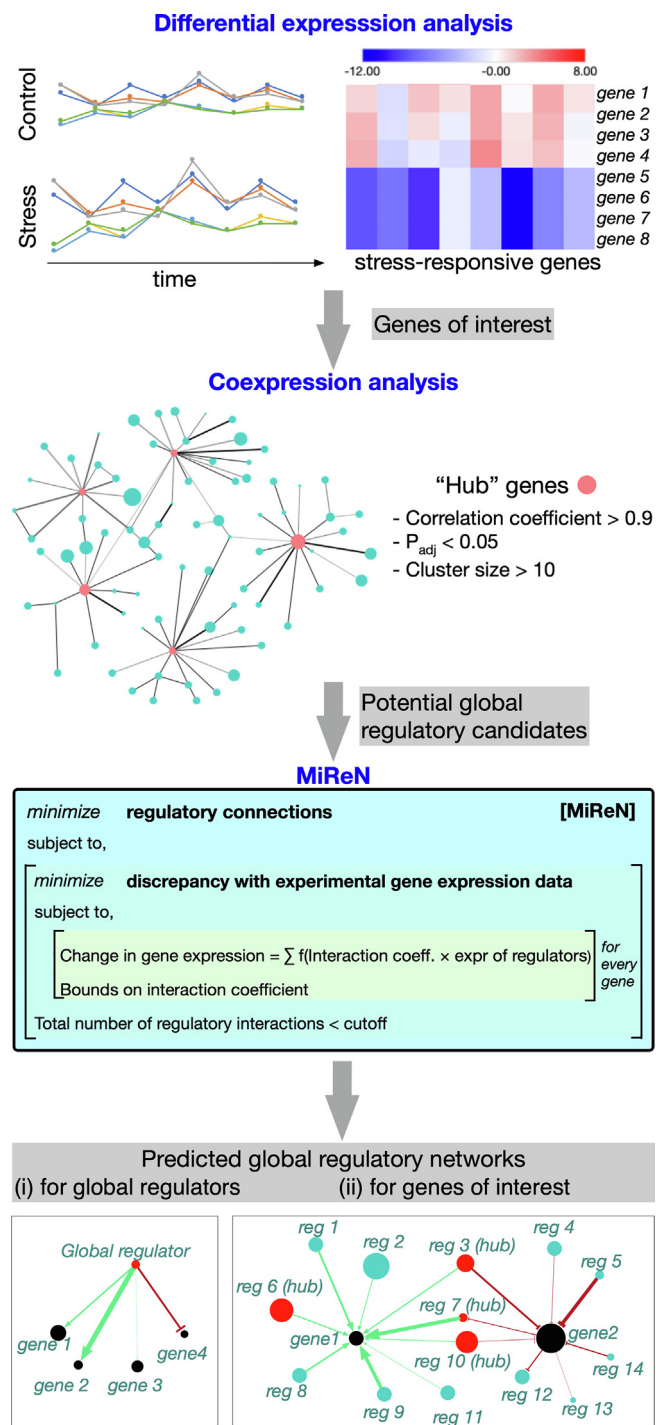


Fig. 5. Step-by-step workflow to identify stress responsive genes of interest and the potential regulator candidates to be analyzed using the MiReN framework.

5.4. Regulatory network creation and visualization

Network visualization was performed in Cytoscape [132] version 3.5.x on a linux-based high-performance cluster computing system using Prefuse Force Directed Layout (<http://prefuse.org/>) and yFiles Organic Layout (<http://www.yworks.com/>) with automatic edge bundling and manual repositioning. DyNet toolbox in Cytoscape was used to show the changes in node presence, connections and node and edge attributes (e.g., expression value or regulation coefficient). The “merge” tool in Cytoscape was used to make

new networks from the union, intersection, and subtraction of the original networks.

Data availability

All datasets generated or analyzed for this study are included in the manuscript and the supplementary files. All computer codes in General Algebraic Modeling System (GAMS) and Matlab used to analyze data during this study are available at a public github repository at the DOI: <https://doi.org/10.5281/zenodo.3832242> under GNU General Public License v3.0.

Funding

This work was supported by University of Nebraska-Lincoln Faculty Startup Grant [21–1106–4308] to RS, University of Nebraska-Lincoln Research Council Interdisciplinary Grant [26–1106–9001–008] to both RS and HW, and NSF Award [Grant numbers 1736192, 1564621] to HW. If the funding sources had no involvement in the design of the study, analysis of data, or the preparation of the manuscript.

CRedit authorship contribution statement

Mohammad Mazharul Islam: Methodology, Software, Formal analysis, Investigation, Writing - original draft, Visualization. **Jaspreet Sandhu:** Methodology, Data curation, Investigation, Writing - original draft. **Harkamal Walia:** Conceptualization, Resources, Writing - review & editing, Supervision, Funding acquisition. **Rajib Saha:** Conceptualization, Resources, Writing - review & editing, Supervision, Project administration, Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work has been completed utilizing the Holland Computing Center of the University of Nebraska, which receives support from the Nebraska Research Initiative.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.csbj.2020.09.022>.

References

- [1] Aghamolki MTK et al. Response of yield and morphological characteristic of rice cultivars to heat stress at different growth stages. *Int J Biol, Biomolecular, Agric, Food Biotechnol Eng* 2014;86:98–100.
- [2] Peng S et al. Rice yields decline with higher night temperature from global warming. *Proc Natl Acad Sci U S A* 2004;101(27):9971–5.
- [3] Cheng LR et al. Genetic Analysis of Cold Tolerance at Seedling Stage and Heat Tolerance at Anthesis in Rice (*Oryza sativa* L.). *J Integrative Agric* 2012;11(3):359–67.
- [4] Jagadish SVK et al. Genetic Analysis of Heat Tolerance at Anthesis in Rice. *Crop Sci* 2010;50(5):1633–41.
- [5] Long SP, Ort DR. More than taking the heat: crops and global change. *Curr Opin Plant Biol* 2010;13(3):241–8.
- [6] Shimamoto K, Kyoizuka J. Rice as a model for comparative genomics of plants. *Annu Rev Plant Biol* 2002;53:399–419.
- [7] Flavell R. Role of model plant species. *Methods Mol Biol* 2009;513:1–18.
- [8] Devos KM, Gale MD. Genome relationships: the grass model in current research. *Plant Cell* 2000;12(5):637–46.

- [9] Paterson AH, Freeling M, Sasaki T. Grains of knowledge: genomics of model cereals. *Genome Res* 2005;15(12):1643–50.
- [10] Hoppe A. What mRNA Abundances Can Tell us about Metabolism. *Metabolites* 2012;2(3):614.
- [11] Jung KH et al. Genome-wide identification and analysis of early heat stress responsive genes in rice. *J Plant Biol* 2012;55(6):458–68.
- [12] Jung KH et al. Genome-wide expression analysis of HSP70 family genes in rice and identification of a cytosolic HSP70 gene highly induced under heat stress. *Funct Integr Genomics* 2013;13(3):391–402.
- [13] Kumar M et al. Genome-wide identification and analysis of genes, conserved between japonica and indica rice cultivars, that respond to low-temperature stress at the vegetative growth stage. *Front Plant Sci* 2017;8.
- [14] Yamaguchi-Shinozaki K, Shinozaki K. Organization of cis-acting regulatory elements in osmotic- and cold-stress-responsive promoters. *Trends Plant Sci* 2005;10(2):88–94.
- [15] Yamaguchi-Shinozaki K, Shinozaki K. Transcriptional regulatory networks in cellular responses and tolerance to dehydration and cold stresses. *Annu Rev Plant Biol* 2006;57:781–803.
- [16] Huynh-Thu VA et al. Inferring regulatory networks from expression data using tree-based methods. *PLoS ONE* 2010;5(9).
- [17] Cooper B et al. A network of rice genes associated with stress response and seed development. *Proc Natl Acad Sci U S A* 2003;100(8):4945–50.
- [18] Dasika MS, Gupta A, Maranas CD. A mixed integer linear programming (MILP) framework for inferring time delay in gene regulatory networks. *Pac Symp Biocomput* 2004:474–85.
- [19] Chen X, Chen M, Ning K. BNArray: an R package for constructing gene regulatory networks from microarray data by using Bayesian network. *Bioinformatics* 2006;22(23):2952–4.
- [20] Paul L et al. Integration of regulatory signals through involvement of multiple global regulators: control of the *Escherichia coli* *gltBDF* operon by Lrp, IHF, Crp, and ArgR. *BMC Microbiol* 2007;7:2.
- [21] Martinez-Antonio A, Collado-Vides J. Identifying global regulators in transcriptional regulatory networks in bacteria. *Curr Opin Microbiol* 2003;6(5):482–9.
- [22] Sun C et al. Understanding the genetic and epigenetic architecture in complex network of rice flowering pathways. *Protein Cell* 2014;5(12):889–98.
- [23] Kaur C et al. Analysis of global gene expression profile of rice in response to methylglyoxal indicates its possible role as a stress signal molecule. *Front Plant Sci* 2015;6:682.
- [24] Zhang L et al. Reconstructing directed gene regulatory network by only gene expression data. *BMC Genomics* 2016;17(Suppl 4):430.
- [25] Mohanty B et al. Identification of candidate network hubs involved in metabolic adjustments of rice under drought stress by integrating transcriptome data and genome-scale metabolic network. *Plant Sci* 2016;242:224–39.
- [26] Mohanty B et al. Transcriptional regulatory mechanism of alcohol dehydrogenase 1-deficient mutant of rice for cell survival under complete submergence. *Rice (N Y)* 2016;9(1):51.
- [27] Mueller TJ et al. Identifying Regulatory Changes to Facilitate Nitrogen Fixation in the Nondiazotroph *Synechocystis* sp PCC 6803. *ACS Synth Biol* 2016;5(3):250–8.
- [28] Zhou G, Xia J. Using OmicsNet for Network Integration and 3D Visualization. *Curr Protocols Bioinformatics* 2019;65(1):e69.
- [29] Raja K et al. A Review of Recent Advancement in Integrating Omics Data with Lattice Mining towards Biomedical Discoveries. *Int J Genomics* 2017;2017:6213474.
- [30] Chen C et al. Heat stress yields a unique MADS box transcription factor in determining seed size and thermal sensitivity. *Plant Physiol* 2016;171(1):606–22.
- [31] Seo YS et al. Towards establishment of a rice stress response interactome. *PLoS Genet* 2011;7(4):e1002020.
- [32] Peng Y et al. OsWRKY62 is a negative regulator of basal and Xa21-mediated defense against *Xanthomonas oryzae* pv. *oryzae* in rice. *Mol Plant* 2008;1(3):446–58.
- [33] Liu X et al. OsWRKY71, a rice transcription factor, is involved in rice defense response. *J Plant Physiol* 2007;164(8):969–79.
- [34] Sharma R et al. Recent advances in dissecting stress-regulatory crosstalk in rice. *Mol Plant* 2013;6(2):250–60.
- [35] Muller S, Livanos P. Plant Kinesin-12: Localization Heterogeneity and Functional Implications. *Int J Mol Sci* 2019;20(17).
- [36] Jiang L et al. The APETALA2-Like Transcription Factor SUPERNUMERARY BRACT Controls Rice Seed Shattering and Seed Size. *Plant Cell* 2019;31(1):17–36.
- [37] Fang J et al. Reduction of ATPase activity in the rice kinesin protein Stemless Dwarf 1 inhibits cell division and organ development. *Plant J* 2018;96(3):620–34.
- [38] Ouyang Y et al. Genomic survey, expression profile and co-expression network analysis of OsWD40 family in rice. *BMC Genomics* 2012;13:100.
- [39] Li J, Xu Y, Chong K. The novel functions of kinesin motor proteins in plants. *Protoplasma* 2012;249(Suppl 2):S95–S100.
- [40] Kitagawa K et al. A novel kinesin 13 protein regulating rice seed length. *Plant Cell Physiol* 2010;51(8):1315–29.
- [41] Shigyo M, Hasebe M, Ito M. Molecular evolution of the AP2 subfamily. *Gene* 2006;366(2):256–65.
- [42] Jan A, Komatsu S. Functional Characterization of Gibberellin-Regulated Genes in Rice Using Microarray System. *Genomics, Proteomics & Bioinformatics* 2006;4(3):137–44.
- [43] Zhang Y et al. Gibberellin homeostasis and plant height control by EUI and a role for gibberellin in root gravity responses in rice. *Cell Res* 2008;18(3):412–21.
- [44] Ueguchi-Tanaka M et al. GIBBERELLIN INSENSITIVE DWARF1 encodes a soluble receptor for gibberellin. *Nature* 2005;437(7059):693–8.
- [45] Du H et al. GID1 modulates stomatal response and submergence tolerance involving abscisic acid and gibberellic acid signaling in rice. *J Integr Plant Biol* 2015;57(11):954–68.
- [46] Jan A et al. Characterization of a xyloglucan endotransglucosylase gene that is up-regulated by gibberellin in rice. *Plant Physiol* 2004;136(3):3670–81.
- [47] Jiang G et al. Regulation of inflorescence branch development in rice through a novel pathway involving the pentatricopeptide repeat protein *sped1-D*. *Genetics* 2014;197(4):1395–407.
- [48] Chen L et al. Overexpression of OsGID1 Enhances the Resistance of Rice to the Brown Planthopper *Nilaparvata lugens*. *Int J Mol Sci* 2018;19(9).
- [49] Ikeda Y et al. slender rice, a constitutive gibberellin response mutant, is caused by a null mutation of the SLR1 gene, an ortholog of the height-regulating gene *GAI/RGA/RHT/D8*. *Plant Cell* 2001;13(5):999–1010.
- [50] Song WY et al. A receptor kinase-like protein encoded by the rice disease resistance gene, *Xa21*. *Science* 1995;270(5243):1804–6.
- [51] Ronald PC, Beutler B. Plant and animal sensors of conserved microbial signatures. *Science* 2010;330(6007):1061–4.
- [52] Shah F et al. Impact of high-temperature stress on rice plant and its traits related to tolerance. *J Agric Sci* 2011;149(5):545–56.
- [53] Reksiri W et al. Expression and promoter analysis of six heat stress-inducible genes in rice. *ScientificWorldJ* 2013;2013:397401.
- [54] Begcy K, Sandhu J, Walia H. Transient Heat Stress During Early Seed Development Primes Germination and Seedling Establishment in Rice. *Front Plant Sci* 2018;9:1768.
- [55] Maestri E et al. Molecular genetics of heat tolerance and heat shock proteins in cereals. *Plant Mol Biol* 2002;48(5–6):667–81.
- [56] Weerakoon WMW, Maruyama A, Ohba K. Impact of Humidity on Temperature-Induced Grain Sterility in Rice (*Oryza sativa* L.). *J Agron Crop Sci* 2008;194(2):135–40.
- [57] Nover L et al. Arabidopsis and the heat stress transcription factor world: how many heat stress transcription factors do we need?. *Cell Stress Chaperones* 2001;6(3):177–89.
- [58] Scharf KD et al. The plant heat stress transcription factor (Hsf) family: structure, function and evolution. *Biochim Biophys Acta* 2012;1819(2):104–19.
- [59] Lavania D, Dhingra A, Grover A. Analysis of transactivation potential of rice (*Oryza sativa* L.) heat shock factors. *Planta* 2018;247(6):1267–76.
- [60] Chauhan H et al. Heat shock factors in rice (*Oryza sativa* L.): genome-wide expression analysis during reproductive development and abiotic stress. *Mol Genet Genomics* 2011;286(2):171–87.
- [61] Chauhan H et al. A seed preferential heat shock transcription factor from wheat provides abiotic stress tolerance and yield enhancement in transgenic Arabidopsis under heat stress environment. *PLoS ONE* 2013;8(11):e79577.
- [62] Burke JJ, O'Mahony PJ, Oliver MJ. Isolation of Arabidopsis mutants lacking components of acquired thermotolerance. *Plant Physiol* 2000;123(2):575–88.
- [63] Larkindale J, Vierling E. Core genome responses involved in acclimation to high temperature. *Plant Physiol* 2008;146(2):748–61.
- [64] Chang P-F-L et al. Induction of a cDNA clone from rice encoding a class II small heat shock protein by heat stress, mechanical injury, and salicylic acid. *Plant Sci* 2007;172(1):64–75.
- [65] Xu Q et al. Regulation of Sucrose Transporters and Phloem Loading in Response to Environmental Cues. *Plant Physiol* 2018;176(1):930–45.
- [66] Cheng F et al. Temperature induced changes in the starch components and biosynthetic enzymes of two rice varieties. *Plant Growth Regul* 2005;46(1):87–95.
- [67] Yamakawa H et al. Comprehensive expression profiling of rice grain filling-related genes under high temperature using DNA microarray. *Plant Physiol* 2007;144(1):258–77.
- [68] Tanamachi K et al. Differential responses to high temperature during maturation in heat-stress-tolerant cultivars of Japonica rice. *Plant Prod Sci* 2016;19(2):300–8.
- [69] Baunsgaard L et al. The 14-3-3 proteins associate with the plant plasma membrane H(+)-ATPase to generate a fusicoccin binding complex and a fusicoccin responsive system. *Plant J* 1998;13(5):661–71.
- [70] Bachmann M et al. 14-3-3 proteins associate with the regulatory phosphorylation site of spinach leaf nitrate reductase in an isoform-specific manner and reduce dephosphorylation of Ser-543 by endogenous protein phosphatases. *FEBS Lett* 1996;398(1):26–30.
- [71] Priya P, Jain M. RiceSRTFDB: a database of rice transcription factors containing comprehensive expression, cis-regulatory element and mutant information to facilitate gene function analysis. *Database (Oxford)* 2013;2013:p. bat027.
- [72] Khew CY et al. Brassinosteroid insensitive 1-associated kinase 1 (OsI-BAK1) is associated with grain filling and leaf development in rice. *J Plant Physiol* 2015;182:23–32.

- [73] Zhu JY, Sae-Seaw J, Wang ZY. Brassinosteroid signalling. *Development* 2013;140(8):1615–20.
- [74] Hao J, Yin Y, Fei SZ. Brassinosteroid signaling network: implications on yield and stress tolerance. *Plant Cell Rep* 2013;32(7):1017–30.
- [75] De Bruyne L, Hofte M, De Vleeschauwer D. Connecting growth and defense: the emerging roles of brassinosteroids and gibberellins in plant innate immunity. *Mol Plant* 2014;7(6):943–59.
- [76] Lozano-Durán R, Zipfel C. Trade-off between growth and immunity: role of brassinosteroids. *Trends Plant Sci* 2015;20(1):12–9.
- [77] Nolan T, Chen J, Yin Y. Cross-talk of Brassinosteroid signaling in controlling growth and stress responses. *Biochem J* 2017;474(16):2641–61.
- [78] Martínez C et al. PIF4-induced BR synthesis is critical to diurnal and thermomorphogenic growth. *EMBO J* 2018;37(23):e99552.
- [79] Zhang X et al. A Temperature-Sensitive Misfolded bri1-301 Receptor Requires Its Kinase Activity to Promote Growth. *Plant Physiol* 2018;178(4):1704–19.
- [80] Planas-Riverola A et al. Brassinosteroid signaling in plant development and adaptation to stress. *Development* 2019;146(5).
- [81] Reimmann C, Dudler R. Circadian rhythmicity in the expression of a novel light-regulated rice gene. *Plant Mol Biol* 1993;22(1):165–70.
- [82] Pacheco-Bernal I, Becerril-Perez F, Aguilar-Arnal L. Circadian rhythms in the three-dimensional genome: implications of chromatin interactions for cyclic transcription. *Clin Epigenetics* 2019;11(1):79.
- [83] Onai K, Ishiura M. PHYTOCLOCK 1 encoding a novel GARP protein essential for the Arabidopsis circadian clock. *Genes Cells* 2005;10(10):963–72.
- [84] Zhou F et al. The bZIP transcription factor HY5 interacts with the promoter of the monoterpene synthase gene QH6 in modulating its rhythmic expression. *Front Plant Sci* 2015;6:304.
- [85] Phukan UJ et al. Regulation of Apetala2/Ethylene Response Factors in Plants. *Front Plant Sci* 2017;8:150.
- [86] Arora R et al. MADS-box gene family in rice: genome-wide identification, organization and expression profiling during reproductive development and stress. *BMC Genomics* 2007;8:242.
- [87] Miyazaki S et al. Rice MEL2, the RNA recognition motif (RRM) protein, binds in vitro to meiosis-expressed genes containing U-rich RNA consensus sequences in the 3'-UTR. *Plant Mol Biol* 2015;89(3):293–307.
- [88] Craig KL, Tyers M. The F-box: a new motif for ubiquitin dependent proteolysis in cell cycle regulation and signal transduction. *Prog Biophys Mol Biol* 1999;72(3):299–328.
- [89] Hong F et al. Overexpression of the rFCA RNA recognition motif affects morphologies modifications in rice (*Oryza sativa* L.). *Biosci Rep* 2007;27(4–5):225–34.
- [90] Hirochika H et al. Retrotransposons of rice involved in mutations induced by tissue culture. *Proc Natl Acad Sci U S A* 1996;93(15):7783–8.
- [91] Vitte C, Panaud O, Quesneville H. LTR retrotransposons in rice (*Oryza sativa*, L.): recent burst amplifications followed by rapid DNA loss. *BMC Genomics* 2007;8:218.
- [92] Wang F, Jing W, Zhang W. The mitogen-activated protein kinase cascade MKK1-MPK4 mediates salt signaling in rice. *Plant Sci* 2014;227:181–9.
- [93] El-Kereamy A et al. The rice R2R3-MYB transcription factor OsMYB55 is involved in the tolerance to high temperature and modulates amino acid metabolism. *PLoS ONE* 2012;7(12):e52030.
- [94] Wilkins O et al. EGRINs (Environmental Gene Regulatory Influence Networks) in Rice That Function in the Response to Water Deficit, High Temperature, and Agricultural Environments. *Plant Cell* 2016;28(10):2365–84.
- [95] Bellande K et al. Plant Lectins and Lectin Receptor-Like Kinases: How Do They Sense the Outside?. *Int J Mol Sci* 2017;18(6).
- [96] Vaid N, Macovei A, Tuteja N. Knights in action: lectin receptor-like kinases in plant development and stress responses. *Mol Plant* 2013;6(5):1405–18.
- [97] Shumayla, et al., Molecular Characterization and Global Expression Analysis of Lectin Receptor Kinases in Bread Wheat (*Triticum aestivum*). *PLoS One*, 2016;11(4):e0153925.
- [98] Silveira RD et al. Expression of drought tolerance genes in tropical upland rice cultivars (*Oryza sativa*). *Genet Mol Res* 2015;14(3):8181–200.
- [99] Kumar KA et al. Identification of genes controlling aba accumulation in rice during drought stress and seed maturation. *Int J Adv Biotechnol Res* 2013;4(4):481–7.
- [100] Ke S et al. Mutation in a putative glycosyltransferase-like gene causes programmed cell death and early leaf senescence in rice. *Rice (N Y)* 2019;12(1):7.
- [101] Moon S et al. Rice glycosyltransferase1 encodes a glycosyltransferase essential for pollen wall formation. *Plant Physiol* 2013;161(2):663–75.
- [102] Yang C et al. MALE STERILITY1 is required for tapetal development and pollen wall biosynthesis. *Plant Cell* 2007;19(11):3530–48.
- [103] Su T et al. Dynamics of Peroxisome Homeostasis and Its Role in Stress Response and Signaling in Plants. *Front Plant Sci* 2019;10:705.
- [104] Monroe-Augustus M et al. Matrix proteins are inefficiently imported into Arabidopsis peroxisomes lacking the receptor-docking peroxin PEX14. *Plant Mol Biol* 2011;77(1–2):1–15.
- [105] Burkhardt SE, Lingard MJ, Bartel B. Genetic dissection of peroxisome-associated matrix protein degradation in Arabidopsis thaliana. *Genetics* 2013;193(1):125–41.
- [106] Kao YT, Gonzalez KL, Bartel B. Peroxisome Function, Biogenesis, and Dynamics in Plants. *Plant Physiol* 2018;176(1):162–77.
- [107] Kato Y, Sakamoto W. FtsH Protease in the Thylakoid Membrane: Physiological Functions and the Regulation of Protease Activity. *Front Plant Sci* 2018;9:855.
- [108] Chern MS, Eiben HG, Bustos MM. The developmentally regulated bZIP factor ROM1 modulates transcription from lectin and storage protein genes in bean embryos. *Plant J* 1996;10(1):135–48.
- [109] Bensmihen S, Giraudat J, Parcy F. Characterization of three homologous basic leucine zipper transcription factors (bZIP) of the AB15 family during Arabidopsis thaliana embryo maturation. *J Exp Bot* 2005;56(412):597–603.
- [110] Jakoby M et al. bZIP transcription factors in Arabidopsis. *Trends Plant Sci* 2002;7(3):106–11.
- [111] Nijhawan A et al. Genomic survey and gene expression analysis of the basic leucine zipper transcription factor family in rice. *Plant Physiol* 2008;146(2):333–50.
- [112] Macovei A et al. A new DEAD-box helicase ATP-binding protein (OsABP) from rice is responsive to abiotic stress. *Plant Signal Behav* 2012;7(9):1138–43.
- [113] Kothari KS et al. Rice Stress Associated Protein 1 (OsSAP1) Interacts with Amino transferase (OsAMTR1) and Pathogenesis-Related 1a Protein (OsSCP) and Regulates Abiotic Stress Responses. *Front Plant Sci* 2016;7:1057.
- [114] McCormick F. Ras-related proteins in signal transduction and growth control. *Mol Reprod Dev* 1995;42(4):500–6.
- [115] Grennan AK. Abiotic stress in rice. An “omic” approach. *Plant Physiol* 2006;140(4):1139–41.
- [116] Narsai R et al. Antagonistic, overlapping and distinct responses to biotic stress in rice (*Oryza sativa*) and interactions with abiotic stress. *BMC Genomics* 2013;14:93.
- [117] Rabbani MA et al. Monitoring expression profiles of rice genes under cold, drought, and high-salinity stresses and abscisic acid application using cDNA microarray and RNA gel-blot analyses. *Plant Physiol* 2003;133(4):1755–67.
- [118] Hasegawa PM et al. Plant Cellular and Molecular Responses to High Salinity. *Annu Rev Plant Physiol Plant Mol Biol* 2000;51:463–99.
- [119] Schmitz AJ et al. SUB1A-mediated submergence tolerance response in rice involves differential regulation of the brassinosteroid pathway. *New Phytol* 2013;198(4):1060–70.
- [120] Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 2014;30(15):2114–20.
- [121] Trapnell C, Pachter L, Salzberg SL. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* 2009;25(9):1105–11.
- [122] Anders S, Pyl PT, Huber W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* 2015;31(2):166–9.
- [123] Haimes J, Kelley M. Demonstration of a $\Delta\Delta Cq$ Calculation Method to Compute Relative Gene Expression from qPCR Data. 2014; Available from: <https://horizondiscovery.com/-/media/Files/Horizon/resources/Technical-manuals/delta-cq-solaris-technote.pdf> [cited 2020 September 5, 2020].
- [124] Anders S, Huber W. Differential expression analysis for sequence count data. *Genome Biol* 2010;11(10):R106.
- [125] Rodriguez-Esteban R, Jiang X. Differential gene expression in disease: a comparison between high-throughput studies and the literature. *BMC Med Genomics* 2017;10(1):59.
- [126] Tian T et al. agriGO v2.0: a GO analysis toolkit for the agricultural community, 2017 update. *Nucleic Acids Res* 2017;45(W1):W122–9.
- [127] Du Z, et al. agriGO: a GO analysis toolkit for the agricultural community. *Nucleic Acids Res*, 2010;38(Web Server issue):W64–70.
- [128] Mukaka MM. Statistics corner: a guide to appropriate use of correlation coefficient in medical research. *Malawi Med J* 2012;24(3):69–71.
- [129] Liesecke F et al. Ranking genome-wide correlation measurements improves microarray and RNA-seq based global and targeted co-expression networks. *Sci Rep* 2018;8(1):10885.
- [130] Batushansky A, Toubiana D, Fait A. Correlation-Based Network Generation, Visualization, and Analysis as a Powerful Tool in Biological Studies: A Case Study in Cancer Cell Metabolism. *Biomed Res Int* 2016;2016:8313272.
- [131] Luo F et al. Constructing gene co-expression networks and predicting functions of unknown genes by random matrix theory. *BMC Bioinf* 2007;8:299.
- [132] Shannon P et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 2003;13(11):2498–504.