

RESEARCH ARTICLE

Deciphering Cis-Regulatory Element Mediated Combinatorial Regulation in Rice under Blast Infected Condition

Arindam Deb¹, Sudip Kundu^{1,2*}

1 Department of Biophysics Molecular Biology and Bioinformatics, University of Calcutta, Kolkata, West Bengal, India, **2** Center of Excellence in Systems Biology and Biomedical Engineering (TEQIP Phase II), University of Calcutta, Kolkata, West Bengal, India

* skmbmg@caluniv.ac.in



Abstract

Combinations of cis-regulatory elements (CREs) present at the promoters facilitate the binding of several transcription factors (TFs), thereby altering the consequent gene expressions. Due to the eminent complexity of the regulatory mechanism, the combinatorics of CRE-mediated transcriptional regulation has been elusive. In this work, we have developed a new methodology that quantifies the co-occurrence tendencies of CREs present in a set of promoter sequences; these co-occurrence scores are filtered in three consecutive steps to test their statistical significance; and the significantly co-occurring CRE pairs are presented as networks. These networks of co-occurring CREs are further transformed to derive higher order of regulatory combinatorics. We have further applied this methodology on the differentially up-regulated gene-sets of rice tissues under fungal (*Magnaporthe*) infected conditions to demonstrate how it helps to understand the CRE-mediated combinatorial gene regulation. Our analysis includes a wide spectrum of biologically important results. The CRE pairs having a strong tendency to co-occur often exhibit very similar joint distribution patterns at the promoters of rice. We couple the network approach with experimental results of plant gene regulation and defense mechanisms and find evidences of auto and cross regulation among TF families, cross-talk among multiple hormone signaling pathways, similarities and dissimilarities in regulatory combinatorics between different tissues, etc. Our analyses have pointed a highly distributed nature of the combinatorial gene regulation facilitating an efficient alteration in response to fungal attack. All together, our proposed methodology could be an important approach in understanding the combinatorial gene regulation. It can be further applied to unravel the tissue and/or condition specific combinatorial gene regulation in other eukaryotic systems with the availability of annotated genomic sequences and suitable experimental data.

OPEN ACCESS

Citation: Deb A, Kundu S (2015) Deciphering Cis-Regulatory Element Mediated Combinatorial Regulation in Rice under Blast Infected Condition. PLoS ONE 10(9): e0137295. doi:10.1371/journal.pone.0137295

Editor: Richard A Wilson, University of Nebraska-Lincoln, UNITED STATES

Received: May 2, 2015

Accepted: August 14, 2015

Published: September 1, 2015

Copyright: © 2015 Deb, Kundu. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: AD is supported by a research fellowship provided by the UGC (RFSMS), Government of India. This work is supported by Center of Excellence in Systems Biology and Biomedical Engineering (TEQIP Phase II), University of Calcutta. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

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

Introduction

Gene regulation is an essential cellular process that controls the expression level of a gene and thereby controls the phenotypic attributes of an organism. Transcriptional control plays a key role in gene regulation [1–3]. The molecular basis of this control is as following: transcription factors (TFs) bind to their corresponding cis-regulatory elements (CREs), located at the promoters (upstream regions) of target genes and regulate the transcriptional phenomenon. Previous studies established that combinatorial control of gene expression involves association of multiple CREs (and their TFs) at the promoters by combinatorial logic among them [2, 4–6]. Such combinatorial regulation can produce a variety of gene expression patterns which are sometimes tissue and/or condition specific [4, 5].

Several computational approaches intended to understand the CRE mediated combinatorial regulation at the promoter regions of genes [7–11]. If two CREs show a higher tendency to co-occur in a set of promoters, there is a strong possibility that their corresponding TFs regulate the expression of those genes in a co-associative fashion [12, 13]. Sudarsanam et al. [8] used cumulative hypergeometric distribution to identify the co-occurring cis-regulatory elements in the yeast genome. Hannenhalli et al. [13] measured the co-occurrence of cis-regulatory elements by the ‘co-localization index’ to estimate synergistically acting TFs. Pati et al. [14] and Chang et al. [15] applied *A priori* algorithm to identify the co-occurrence of cis-regulatory elements. In recent years, Vandenbon et al. [16] proposed a novel measuring technique to estimate unbiased co-occurrence tendencies of cis-regulatory elements based on their frequencies of occurrences at the promoter regions. In addition, previous studies showed that combinatorial transcriptional regulation can be analyzed in the form of co-regulatory network [17]. Aiming to further advancement, we integrated the measurement of CRE co-occurrence and network approach that mimics the combinatorics of transcriptional regulation. Here, we introduced a new mathematical expression to compute the co-occurrence tendencies of CREs using their frequencies of occurrences which allowed us to directly transform these binary relationships into a network.

Over the years, co-associative roles of CREs have been analyzed in model plant systems such as Rice [15, 18]. Though, previous attempts have been undertaken to explore the transcriptional regulatory mechanism of gene expression under abiotic stress conditions [18, 19], little such analysis has been done on biotic stresses. Therefore, as a practical application of our computational approach, we attempted to explore the plausible combinatorics of transcriptional regulation of rice under biotic stress condition. For this analysis, we have chosen a specific biotic stress condition, in which rice plants (leaf and root) were infected by a fungal pathogen *Magnaporthe oryzae*. The fungus *Magnaporthe oryzae* is responsible for the blast disease, causing the biggest disruption of annual rice yield (up to 30%) worldwide [20].

Our aim was to establish a new methodology to estimate the co-occurrence relationships among CREs and transform these relationships into networks; this approach made a systems level analysis of the regulatory mechanism possible. Results of this work allowed us to map the CRE-mediated combinatorial regulation of differentially up-regulated rice genes under blast infected condition. Here, constructing the CRE co-occurrence networks, we pointed the probable regulatory cross-talks among different hormone signaling pathways. Further, a network transformation procedure was used to generate a network (clique-clique network) representing even higher order of combinations. The analysis of network topological features suggested that alteration in the CRE combinations is highly facilitated under blast infected condition. Although applied to the blast infected condition of rice, it is a general methodology that can be applied to other experimental studies as well.

Materials and Methods

In the present work, we introduced a new methodology to identify significantly co-occurring CREs at the promoters of a gene-set. These co-occurrence relationships were further transformed into an edge-weighted network from which higher order combinatorics were estimated. A basic outline of our methodology is presented in [Fig 1](#) and detailed descriptions are given below.

Data Collection

Promoter sequence data. We collected 1 kb upstream promoter sequence data (starting from TSS) [[18](#), [21](#)] of each rice loci from Rice Genome Annotation Project, version 6.1 [[22](#)]. RepeatMasker, version open-4.0.2 [[23](#)] was used to mask interspersed and simple repeats in the sequence data; this screening prevents unwanted discovery of CREs into the stretches of repeated DNAs present in the promoters.

Cis-regulatory element data. An initial dataset of 469 CREs was collected from PLACE (Plant cis-acting regulatory element) database [[24](#)]. This initial dataset was screened to identify the unique entries. The final dataset contained 465 CREs with unique consensus sequences. Signal Scan [[25](#)] tool was used to fetch the occurrences and positions of the CREs in the promoters.

Expression data of blast infected condition. Raw microarray expression data were obtained from Gene Expression Omnibus(GEO) [[26](#)], platform accession number GPL2025, for Rice (*Oryza sativa*) infected with different strains of *Magnaporthe oryzae* (GSE7256, GSE18361) [[27](#), [28](#)]. This platform was designed on Affymetrix GeneChip Rice Genome Array consisting of 57381 probe sets. Raw data were processed and RMA (Robust Multiarray Average) normalization was performed using R and Bioconductor [[29](#)]. Differentially expressed probes were obtained (Log fold change > 2, adjusted P-value by BH < 0.01) for individual time point from each experiment. All probes and their corresponding loci information for rice were extracted from the Affymetrix database [[30](#)]. Control probes were eliminated from the data set prior to further processing. Those probes mapped to multiple loci were considered as ambiguous probes [[31](#), [32](#)] and were removed from mapping. However, multiple probes mapped to single locus were considered as redundant probes [[31](#)] and only the probe having the extreme (maximum/minimum) expression value was mapped to the locus [[33](#), [34](#)]. Finally we got 29766 unique Rice loci mapped to their corresponding probes. We chose our gene-set of interest from these differentially expressed loci. In-house PERL scripts were used for the mapping process.

GO mapping and identifying transcription factors. GO mapping was done using the web-based GO enrichment analysis program available at Rice Oligonucleotide Array Database (<http://www.ricearray.org/>) [[34](#)]. Different biological processes, related to plant defense and stress, were identified containing several differentially up-regulated rice loci. The information about transcription factors of rice (*Oryza sativa subsp. japonica*) was extracted from PlantTFDB database version-3.0 [[35](#)]. A total number of 2408 loci of transcription factors, distributed into 56 families, were obtained. However, only 1321 loci of transcription factors had corresponding probe mapping and were considered in this work.

Measurement of co-occurrence tendency of CREs

Co-occurrence estimation. We introduced a new mathematical expression to quantify the co-occurrence tendency of a pair of CREs in the promoters of a set of genes. The degree of co-occurrence between a CRE pair, (referred as E_1 and E_2), was measured by a parameter termed

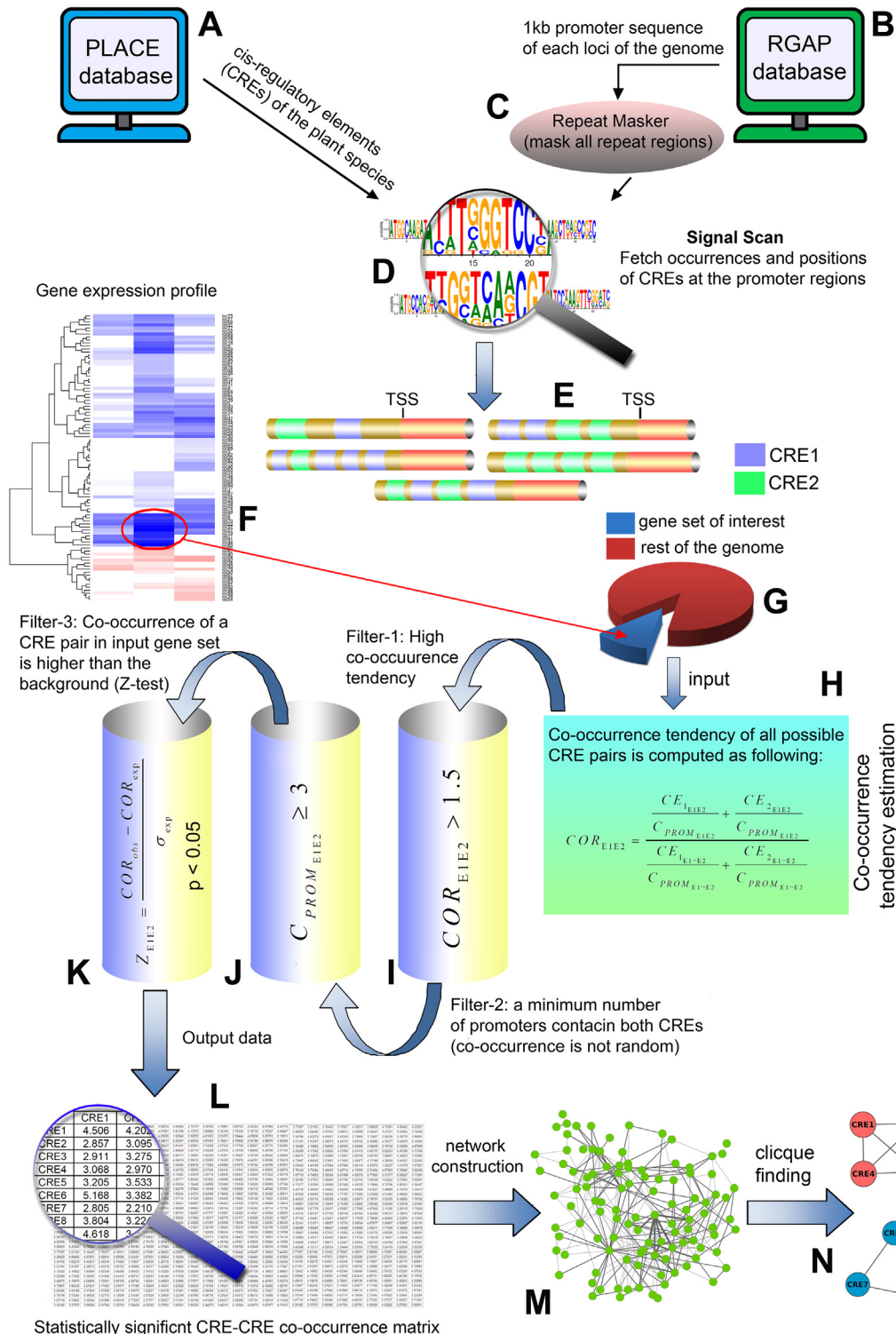


Fig 1. A diagrammatic representation of our methodology. (A) Cis-regulatory element information is collected from PLACE database. (B) Promoter sequence data is collected from Rice Genome Annotation Project database. (C) RepeatMasker masks all the repeated regions in the promoter sequences. (D) Signal Scan tool scans the promoter sequences to estimate the occurrences and positions of the CREs. (E) A cartoon diagram shows the location of individual CREs in the promoter regions. (F, G) Selection of input gene-sets (differentially up-regulated genes) and background (rest of the genome). (H) *COR* values are estimated for all possible CRE pairs. (I) First filtering step: *COR* > 1.5 values are chosen. (J) Second filtering step: only the CRE pairs present in ≥ 3 promoters are selected for further analyses. (K) Third filtering step: significance of a *COR* value in the input gene-set is compared against the

background by a Z-statistics ($p < 0.05$ is chosen). **(L)** Matrix representation of CRE pairs with statistically significant co-occurrence. **(M)** This matrix is transformed into an edge-weighted network (nodes represent individual CREs, edge weights represent the *COR* values). **(N)** Network analysis reveals the unique cliques of CREs.

doi:10.1371/journal.pone.0137295.g001

as the Co-Occurrence Ratio (*COR*).

$$COR_{E1E2} = \frac{\frac{CE1_{E1E2}}{Cprom_{E1E2}} + \frac{CE2_{E1E2}}{Cprom_{E1E2}}}{\frac{CE1_{E1-E2}}{Cprom_{E1-E2}} + \frac{CE2_{E2-E1}}{Cprom_{E2-E1}}} \quad (1)$$

Where, $CE1_{E1E2}$ and $CE2_{E1E2}$ are the count of first (*E1*) and second (*E2*) CREs respectively into those promoters comprising both the CREs. $Cprom_{E1E2}$ is the count of such promoters where both the CREs are present. $CE1_{E1-E2}$ is the count of first CRE (*E1*) in those promoters lacking second CRE (*E2*). $CE2_{E2-E1}$ is the count of second CRE (*E2*) in the promoters lacking first CRE (*E1*). $Cprom_{E1-E2}$ is the number of promoters where first CRE (*E1*) occurs exclusive to the second CRE (*E2*). $Cprom_{E2-E1}$ is the number of promoters where second CRE (*E2*) occurs exclusive to the first CRE (*E1*).

The numerator of the above equation (Eq 1) represents the frequency of joint occurrence of a CRE pair (*E1* and *E2*) and the denominator represents the frequencies of exclusive occurrences of these two CREs. Therefore, *COR* value > 1 indicates higher tendency of co-occurrence of two CREs, whereas, *COR* value < 1 means there is a poor co-occurrence tendency. We did not count the occurrences at overlapping sites of the two CREs to restrict the co-occurrence bias emerging from the overlap (i.e. $CE1_{E1E2}$ and $CE2_{E1E2}$ were not incremented when their occurrences found at overlapping sites) [16]. We present Fig 2 as an example of the *COR* value calculation. For further details see S1 Text.

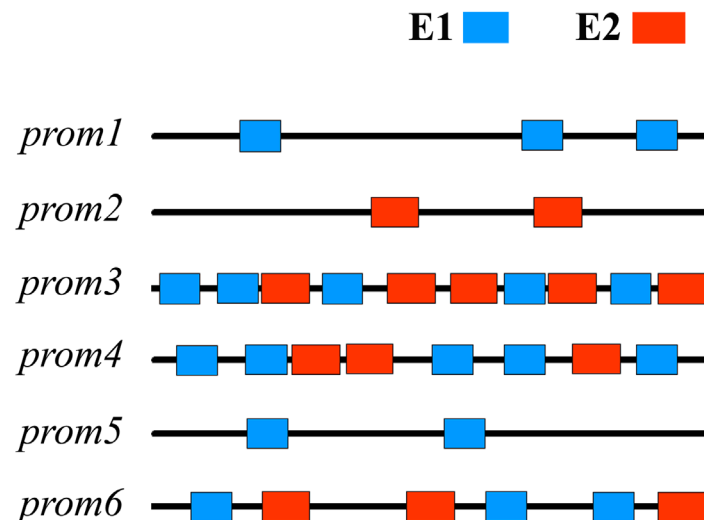


Fig 2. Diagrammatic representation of a pair of CREs (E1, E2) in a set of promoters (*prom1*, *prom2*, ... *prom6*). Out of these 6 promoters, 3 contain both the CREs (so, $Cprom_{E1E2} = 3$), 2 contain the first CRE, *E1* exclusive to *E2* ($Cprom_{E1-E2} = 2$) and the remaining one contains the second CRE, *E2* exclusive to *E1* ($Cprom_{E2-E1} = 1$). Similarly, in the first category of promoters *E1* has occurred 13 times and *E2* has occurred 11 times (hence, $CE1_{E1E2} = 13$ and $CE2_{E1E2} = 11$). The number of occurrences of *E1* exclusive to *E2* is 5 ($CE1_{E1-E2} = 5$) and the number of occurrences of *E2* exclusive to *E1* is 2 ($CE1_{E2-E1} = 2$). In this case the estimated *COR* value is 1.77.

doi:10.1371/journal.pone.0137295.g002

Setting a stringent cutoff of the COR value. We calculated *COR* values for all possible pairs ($^{465}C_2$) of CREs using the rice promoters. A probability distribution plot of the *COR* values is provided in Fig 3. The upper-bound confidence interval of the data was computed at 10^{-6} significance level by performing a permutation Z-statistics. In this step, calculation started with the highest *COR* value (observed set) and permuted the remaining values (background) to check whether the difference met the desired significance level. Next, gradually lower *COR* values were added with the observed set and the former step was repeated until it reached 10^{-6} significance level. This enabled us to identify the minimum value of the upper-bound set, which was 1.498. Next, a non-parametric permutation Mann-Whitney U-test (10000 permutation steps) was performed to check whether this upper-bound *COR* values significantly differ from the rest of the dataset. This U-test showed that data above the *COR* value 1.498 was significantly higher than rest of the dataset with a very strong statistical significance ($p < 10^{-170}$). Therefore, a *COR* value > 1.5 was considered as the stringent cutoff (S1 Text).

Filtering CRE pairs those occur at least in three promoters. A high *COR* value can also appear from rich occurrence of a CRE pair exclusively in a few (1-2) promoters. Therefore, CRE pairs that do not occur in a considerable number of promoters should be eliminated to minimize the false-positives. The CRE pairs (with $COR > 1.5$) were considered for further analysis only if they were present in at least three promoters (i.e. $C_{PROM_{E1E2}} \geq 3$).

Estimation of statistical significance of co-occurrence in a gene-set. When analyzing the co-occurrences of CREs within a specific gene-set of interest, it is important to highlight the *COR* values those are statistically significant with respect to the background data. The statistical significance level of co-occurrence of a pair of CREs was determined by another Z-statistics. We checked how well the *COR* value of a CRE pair in an observed promoter set (gene-set of interest) matched the *COR* value that would be expected for random promoter sets of same

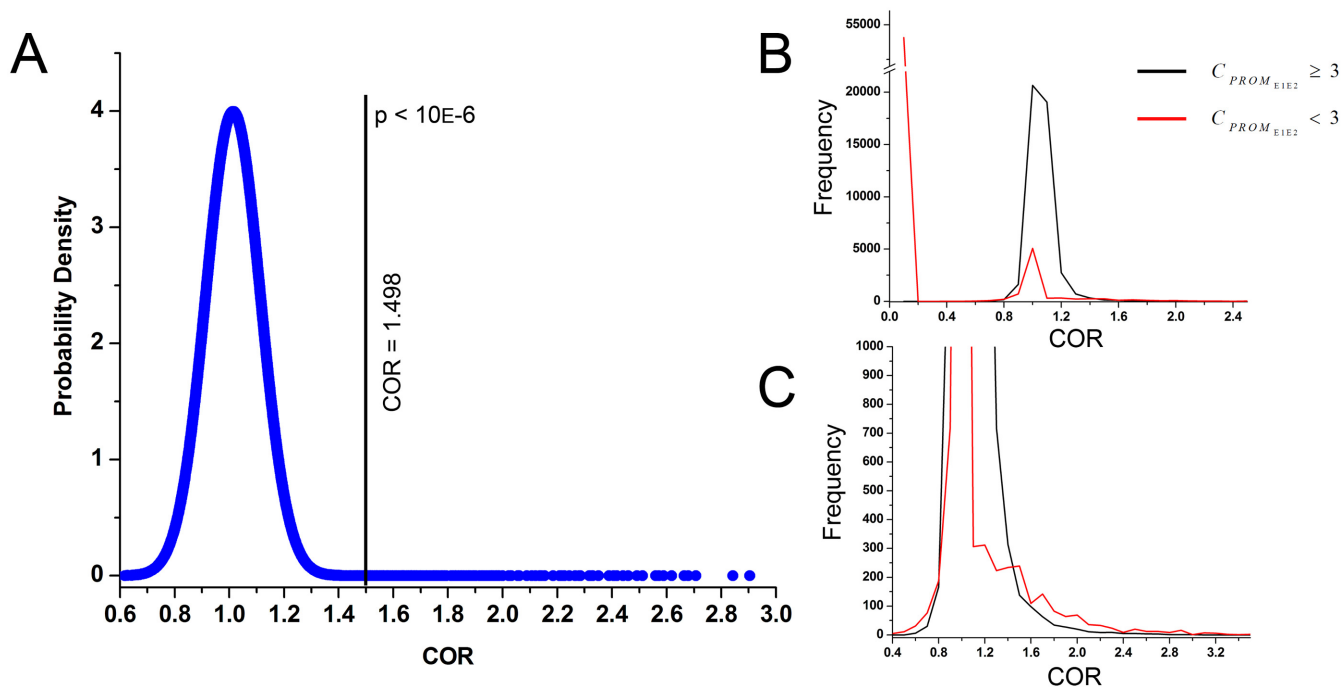


Fig 3. COR value distribution pattern. (A) probability density plot of COR values (B) frequency distribution plot of COR values. (C) magnified view of y axis of the frequency plot of COR values. The red line stands for the COR values of the CRE pairs occurring in < 3 promoters and the black line stands for the COR values of the CRE pairs occurring in ≥ 3 promoters.

doi:10.1371/journal.pone.0137295.g003

size. Based on the Z score (Eq 2), one-tailed p value was estimated for each pair of CREs.

$$Z_{E1E2} = \frac{COR_{obs} - COR_{exp}}{\sigma_{exp}} \tag{2}$$

[COR_{obs} → COR value found in the actual set of promoters, COR_{exp} → expected COR value which is the mean of the COR values calculated from 1000 sets of randomly sampled promoters of the same size of actual set, σ_{exp} → standard deviation calculated on 1000 expected COR values.]

We selected the CRE pairs for which the p value of Z-statistics was < 0.05; these pairs were considered for further analysis.

Thus, we applied a set of relevant statistical screening to fetch only the strong co-occurrence signals out of the background data. This statistically significant co-occurring CREs were used for further analysis. The pipeline described here was implemented using in-house PERL scripts.

Construction of CRE co-occurrence network and clique-clique network

For a set of promoters (gene-set of interest), significantly co-occurring CRE pairs were represented as undirected edge-weighted networks, termed as ‘CRE co-occurrence networks’. In these networks, each CRE represented a node and COR values between the CRE pairs represented the edge weights. The CFinder [36] was used to identify the cliques from each network. Clique(k) was defined as a complete subgraph with k number of nodes. Thus, a clique represents here the combinations of co-occurring CREs in a set of differentially up-regulated genes where their co-occurrences (each CRE co-occurred with every other CRE present into the clique) are statistically enriched. Further, we transformed each CRE co-occurrence network into a clique-clique network where each node represented a clique and an edge weight represented the number of overlapping/sharing CREs between two cliques (Fig 4). Different views of networks, represented here, were constructed using Cytoscape 3.1.1 [37].

Network topology analysis

We looked into the inherent topology of the CRE co-occurrence networks as well as clique-clique networks. The coefficient of assortativity (r) measures the tendency of degree correlation [38]. It is the Pearson correlation coefficient of the degrees at either end of an edge and was defined as:

$$r = \frac{M^{-1} \sum_i j_i k_i - [M^{-1} \sum_i 0.5(j_i + k_i)]^2}{M^{-1} \sum_i 0.5(j_i^2 + k_i^2) - [M^{-1} \sum_i 0.5(j_i + k_i)]^2} \tag{3}$$

Here, j_i and k_i are the degrees of the nodes at the either ends of the i^{th} edge, where, $i = 1, 2, \dots, M$. M is the total number of edges. The networks showing positive and negative r values are assortative and disassortative in nature respectively.

The clustering coefficient (C) is another topological parameter of the network nodes; clustering coefficient (C_i) is a ratio N_i/M_i , where N_i is the number of edges between the neighbors of a i^{th} node, and M_i is the maximum possible number of edges between its neighbors. The network clustering coefficient (\bar{C}) of a network is computed as the average of all the nodes (n) and it represents the cohesiveness of the entire network.

$$\bar{C} = \frac{1}{n} \sum_{i=1}^n C_i \tag{4}$$

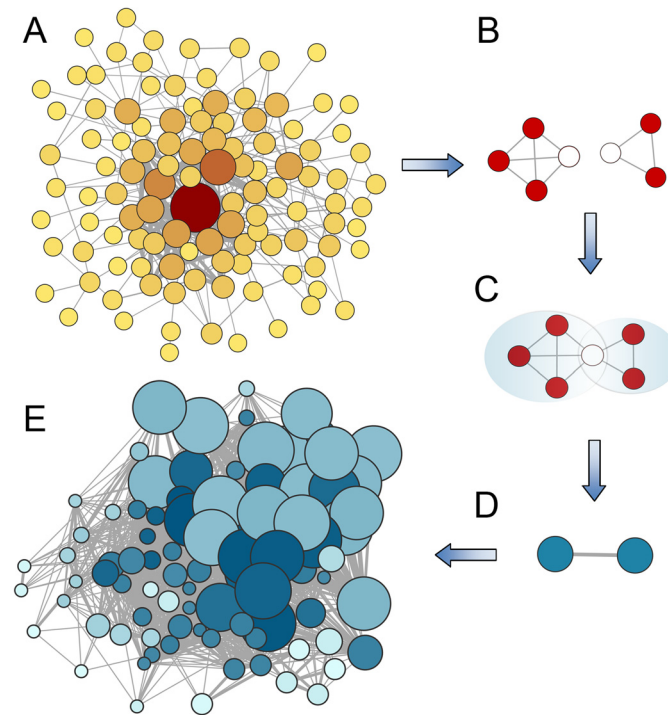


Fig 4. The network transformation. (A) A CRE co-occurrences network. In this network, a node represents a CRE and an edge represents co-occurrence relation (*COR*) in between a pair of nodes (Wider line = higher *COR* value). Node-colour and size represent the degree of a node (darker and bigger = higher degree) (B) Clique identification in the CRE co-occurrences network. (C) Search for common/sharing CRE(s) in between two cliques. (D) Represent each clique as single node and connect two cliques(nodes) if common CREs are found in between them. (E) A clique-clique network. Here a node represents a clique of CREs, while an edge indicates the relation in between two cliques in the terms of sharing/overlapping of CREs. Wider line of an edge indicates higher number of overlapping CREs in between two cliques. The node-size indicates the number of promoters in which the clique has occurred (bigger = more number of promoters) and the node-colour indicates the degree of a node (darker = higher degree).

doi:10.1371/journal.pone.0137295.g004

Measurement of positional distribution patterns of CREs in promoters

We investigated the position-wise distribution patterns of the CRE occurrences. Distribution patterns of their joint as well as exclusive occurrences in rice promoters were studied. Starting from the transcription start site (TSS) we sectioned the 1 kb upstream sequence into 10 intervals, each of 100 nucleotides length. At each interval, the frequency of joint occurrences of the two CREs (when the two CREs E1 and E2 were simultaneously present) as well as their exclusive occurrences (either one of E1 and E2 was present) were computed. At any *i*-th interval the joint occurrence frequency of E1, when it co-occurred with E2 was $F_i(E1E2)$ and the joint occurrence frequency of E2, when it co-occurred with E1 was $F_i(E2E1)$. The Pearson correlation coefficient (*r*) was estimated between $F_i(E1E2)$ and $F_i(E2E1)$ sets (where, $i = 1, 2, \dots, 10$), which measures the similarity of their joint distribution patterns.

Results and Discussion

Differentially up-regulated genes in blast infected tissues of rice

Under any biotic/abiotic stress conditions, gene regulation is reprogrammed; for instance, under pathogen attack, the plant defense response is triggered [39]. Defense response involves

a change in the entire cellular physiology (induced by remodeling of gene expression by TFs), those try to resist the pathogen in diverse ways.

Cellular functions are activated by elevated rates of expression of the related genes and we have used this elevation as the criterion to select the data sets. Activation of genes related to responsive/signaling pathways should be associated with the elevated rates of expression of the TFs those regulate them [40]. The mode of infection of blast fungus defers in leaf and root, which causes tissue-specific defense response mechanism in rice [28]. In the collected gene expression data, a total number of 302 and 1201 loci were differentially up-regulated at 3 DPI (Days Post Inoculation) and 4 DPI in blast infected rice leaf, respectively. On the other hand 677, 55 and 154 loci were differentially up-regulated at 2 DPI, 4 DPI and 6 DPI in the blast infected rice root, respectively. In leaf, the maximum number of differential gene expression was observed at 4 DPI time point, the stage just before of lesion formation (5 DPI), which indicated a large-scale defense response to hemibiotrophic mode of fungal infection [27]. Whereas, in root the maximum number of host gene response was found at 2 DPI time point indicating a substantial and early defense response to biotrophic mode of infection [28]. In addition, Marcel et al. [28] showed that the infection profile of 2 DPI root sample was similar to the two leaf samples (3 DPI and 4 DPI). Though, the 4 DPI time point was common in between leaf and root tissues, in 4 DPI root, the total number of differentially up-regulated genes were only 55. This included only 2 TF genes and few genes related to other GO processes, but excluding those related to defense response. Thus, the 4 DPI root data was found to be insufficient for analysis. Considering all these factors, we selected the input gene-sets from the two expression data sets: 4 DPI of leaf and 2 DPI of root.

In leaf (4 DPI), 112 loci of TFs, belonging to 19 TF families, were up-regulated and in root (2 DPI), 79 loci of TFs, belonging to 12 TF families, were up-regulated out of which 10 TF families (bHLH, bZIP, C2H2, ERF, GRAS, HSF, MYB, NAC, Trihelix and WRKY) were found as common in both tissues. The up-regulated genes from the four TF families (WRKY, ERF, MYB and NAC) were initially chosen for analysis based on the following criteria: (i) the TF families those have well established roles in plants defense response [41–47]; (ii) those were present in higher numbers in both the infected tissues (S1A Fig). In addition, we performed GO functional analysis of the up-regulated rice genes. This identified several GO biological processes, from which we selected four processes based on these two criteria: (i) processes those are known to be involved in plant defense response [27, 48–51] and (ii) those were identified in both the infected tissues. These included “chitin catabolic process” (GO:0006032), “response to stress” (GO:0006950), “defense responses” (GO:0006952) and “G-protein coupled receptor protein signaling pathway” (GO:0007186) (S1B and S1C Fig). We applied our methodology on each of these gene-sets (up-regulated gene-set of 4 TF families and 4 GO processes in infected leaf and root; so total 16 input gene-sets) to fetch the respective CRE co-occurrence networks. Two out of the 16 gene-sets (NAC TF family and “chitin catabolic process” of 2DPI in root) could not be analyzed because the number of loci of each of these sets were insufficient for statistical analyses.

Individual statistics of the co-occurrence networks of each gene-set are provided in the Table 1. The respective lists of input gene-sets, the number of CREs initially found and the number of CRE pairs persisting during the three filtering steps are provided in S1 Table.

Co-occurrence tendencies and positional distribution of CREs in Rice promoters

We computed the *COR* value of CRE pairs in the promoters of the whole rice genome. A total number of 41335 (38%) *COR* reclined close to 1 (0.9–1.1) whereas, 428 (0.39%) CRE pairs were found having $COR \geq 1.5$ (Fig 3).

Table 1. Statistical parameters of CRE co-occurrence networks of the input gene-sets. In the CRE co-occurrence network statistics, *Nodes* refer to the number of CREs present in the network and *Edges* correspond to the number of significantly co-occurring interactions. *Cliques* refers to the number of cliques obtained from each network. *Clique size(k)* refers to the size of cliques (number of nodes in each clique). All the CREs do not form cliques; here *CREs in cliques* refers to the number of CREs those appear in cliques.

Input gene-set	Tissue	No. of promoters	CRE co-occurrence network statistics				
			Nodes	Edges	Cliques	Clique size (k)	CREs in cliques
WRKY	Leaf	20	110	260	38	3, 4	38
	Root	20	113	257	69	3	60
MYB	Leaf	18	112	247	47	3, 4	54
	Root	14	101	212	38	3, 4	47
ERF	Leaf	19	121	347	76	3, 4	97
	Root	23	118	261	44	3	55
NAC	Leaf	15	95	162	17	3, 4	31
	Root	6	—	—	—	—	—
Chitin catabolic process	Leaf	11	93	208	53	3, 4, 5	61
	Root	3	—	—	—	—	—
Response to stress	Leaf	23	122	351	97	3	73
	Root	12	90	210	45	3, 4	61
Defense response	Leaf	19	111	287	88	3, 4, 5	63
	Root	12	83	177	60	3, 4	47
GPCR signaling pathway	Leaf	26	122	291	48	3, 4	57
	Root	11	85	219	59	3, 4, 5, 6	54

doi:10.1371/journal.pone.0137295.t001

The positional distribution of CREs at promoter region remained one of the important factors to estimate their complex association [52]. Highly similar distribution patterns between the joint occurrences of CREs (e.g., DOFCOREZM and GT1CONSENSUS; $r = 0.95$) were associated with their elevated *COR* values ($COR = 1.92$) (Fig 5). Consequently, dissimilar joint distribution patterns of pairs of CREs (e.g., DOFCOREZM and GCCCORE; $r = -0.94$) were associated with low *COR* values ($COR = 0.8$). In our study, a pair of CREs (GT1CONSENSUS and GATABOX) simultaneously exhibited strong co-occurrence tendency ($COR = 1.51$) and highly similar positional distribution pattern ($r = 0.94$) of their joint occurrences. Earlier, experimental results confirmed a cooperative relationship existing between GT1CONSENSUS and GATABOX [53, 54]. In cellular system, combinatorial control may occur either by the combinatorial logic of CREs [6] or by cooperative interactions of TFs [55]. The *COR* value and the distribution patterns of joint occurrences of two CREs at the promoter regions are likely being the indicators of both the co-regulatory phenomena.

Regulators being regulated: Auto and cross regulation of TF families

Present study pointed that WRKY and MYB were subjected to auto-regulation [41, 42, 56], since their own CREs were present in their promoters. Many WRKY genes themselves were enriched for W box elements (WRKY71OS, WBOXNTERF3, WBOXPCWRKY1, WBOX-ATNPR1 and WBOXNTCHN48) in their promoters, suggesting the possibility of intra-family auto-regulation by themselves. Clique analysis in WRKY promoters revealed that WRKY binding sites, present in their promoters, co-occurred with CREs of other TF families. For example, WRKY71OS, the binding site of rice WRKY71 TF co-occurred with GATABOX (binding site of GATA TF) and MYCCONSENSUSAT (MYC recognition site, bHLH family TF).

Similarly, WBOXNTERF3, another WRKY binding site, formed a combination with ASF1-MOTIFCAMV (binding site of ASF-1, a bZIP family TF) and MYBCORE (MYB binding site).

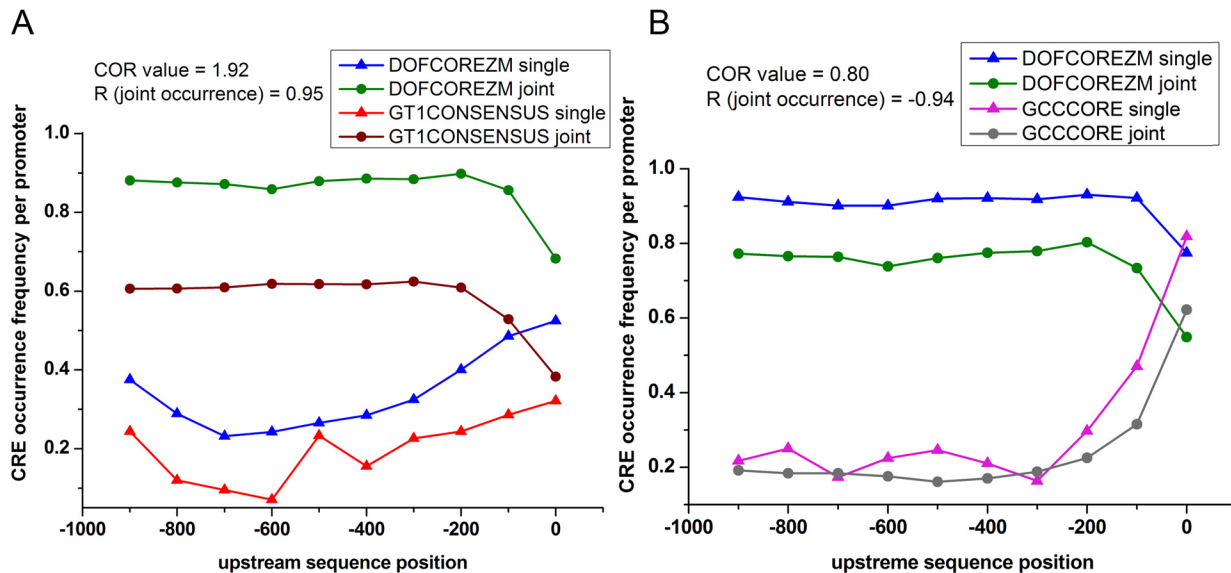


Fig 5. Positional distribution (exclusive/single vs joint occurrence) of CRE pairs. (A) The frequencies of occurrence (single and joint) of two CREs (DOFCOREZM and GT1CONSENSUS) with significantly high COR value (1.92) is presented. The positional distribution patterns of the two CREs are very similar ($R = 0.95$) where they occur together. Moreover, there is a distinct difference in between frequencies of single and joint occurrences of CREs. The joint occurrence frequencies are about 3 times higher than their corresponding single occurrence frequencies, which indicates the tendency of co-occurrence. (B) Positional distribution of DOFCOREZM and GCCCORE exhibit completely different scenario, in which the distribution pattern of single and joint occurrences of each CRE have no difference rather the frequencies of single occurrence are more favored over joint occurrences. As a result, this CRE pair exhibits weak COR value (0.80) and highly dissimilar positional distribution pattern ($R = -0.94$) of joint occurrences.

doi:10.1371/journal.pone.0137295.g005

This indicated the co-regulation of WRKY genes by a number of TF-families (e.g., WRKY, MYC, GATA, bZIP, MYB etc). Thus, the auto-regulatory loops of WRKY TFs are again associated with complex regulatory cross-talks among different TF-families, some of which play crucial role in plant defense response.

The promoters of MYB family had several binding sites (MYBST1, MYB2CONSENSUSAT, MYBCOREATCYCB1, MYBPLANT, MYB1AT, MYBCORE, MYBPZM and TATCCAO-SAMY) of different MYB TFs which indicated the self regulation of this family. A number of cliques of CREs were identified in the MYB promoters pointing the involvement of multiple TF-families in MYB gene regulation. The GT1CONSENSUS (binding site of GT-1, trihelix TF-family) co-occurred with MYCONSENSUSAT (MYC binding site, bHLH TF-family) and DOFCOREZM (core binding site of DOF TFs). In another clique, ARR1AT (recognized by ARR1 TF) co-occurred with RAV1AAT (recognition site of RAV1 TF, containing AP2-like and B3 domain) and CURECORECR (core sequence 'GTAC' of copper responsive element).

The cliques, identified in the promoters of ERF-family genes, contained WRKY binding sites (WRKY71OS, WBOXNTERF3, WBOXATNPR1), bZIP binding sites (ASF1MOTIF-CAMV, DPBFCORED3), MYB binding sites (MYBPZM, TATCCAO-SAMY, MYB1AT, MYB2CONSENSUSAT, MYBCOREATCYCB1, MYBGAHV, MYBST1 and EECRCRH1) and MYC binding sites (MYCONSENSUSAT, MYCATERD1 and MYCATRD22) indicating a probable cross-regulation by these TF families. In the promoters of NAC family, it was observed that a number of GT-1 binding sites (GT1CONSENSUS, GT1GMSCAM4, GT1CORE) co-occurred with MYC, WRKY, MYB, RAV1, ABRE and GATA binding sites.

All together, our result depicted the complexity of transcriptional regulation by revealing the auto/cross regulations among different TF families. The auto/cross regulation probably performs fine tuning in the expression level of the TF genes and thus may help in the plant defense.

Involvement of CRE combinations in plant defense suggests cross-talk among multiple hormone signaling pathways

Salicylic acid (SA) and SA-mediated signaling pathways play pivotal roles in systemic acquired resistance (SAR) against fungal infection in plants [57]. Jasmonic acid (JA) mediated signaling pathways are involved with wounding against necrotrophic pathogens [58, 59]. The crosstalk between SA and JA mediated defense pathways involves both synergistic and antagonistic interactions [60, 61]. Here we investigated whether the CRE combinatorics can reflect the essence of these cross-talks among different hormone signaling pathways including SA and JA.

The following clique: [WRKY71OS, MYCCONSENSUSAT, GT1CONSENSUS, DOFCOREZM] was identified in the co-occurrence network of “Defense Responses” genes (S1 Table) differentially up-regulated in leaf. WRKY71OS is the binding site of rice WRKY71 TF, involved in multiple defense response mechanisms such as wounding and pathogen infection, being induced by defense signaling hormones like SA and JA [62]. Moreover, OsWRKY71 was found to be induced by ABA and was repressed in response to Gibberellic acid (GA) in rice aleurone cells [63]. MYCCONSENSUSAT (MYC recognition site) is the binding site of AtMYC2 which is known to be induced by Abscisic acid (ABA) [64] and is involved in Jasmonic acid (JA) mediated defense responses to pathogen attack and wounding [58, 59]. GT1CONSENSUS is the binding site of GT-1 TF (trihelix family) which influences the SA-inducible PR gene expression [65]. DOFCOREZM is the core binding site of DOF, a class of plant specific TF which acts as transcriptional regulator [66]. In this combination the three CREs (WRKY71OS, MYCCONSENSUSAT and GT1CONSENSUS) showed SA, JA and ABA mediated defense activity. Another example is the following clique: [WBOXATNPR1, ARR1AT, MYCCONSENSUSAT, POLLEN1LELAT52], which was identified in the same network. WBOXATNPR1 (another WRKY binding sequence) was identified in the promoters of NPR1 genes of Arabidopsis, has regulatory role in SA-induced defense gene expression and disease resistance. [67]. ARR1AT is the binding site of ARR1 TF, which is regulated by CK [68] and interact with SA to activate plant defense response [69]. POLLEN1LELAT52 was reported as one of the two co-dependent elements involved in pollen specific activation of tomato lat52 gene [70]. While, the two CREs (WBOXATNPR1, ARR1AT) perform SA dependent functions, MYCCONSENSUSAT takes part in JA mediated defense gene regulation [58, 59].

In the co-occurrence networks of “G-protein coupled receptor protein signaling pathway” genes GCCCORE co-occurred with BIHD1OS (in a clique of leaf) and GT1CONSENSUS (in a clique of root). The GCCCORE (GCC box), reported in the promoters of several PR genes, shows JA dependent defense responses interacting with Ethylene response factors(ERF) [71]. The BIHD1OS acts as the binding site of OsBIHD1, a rice BELL homeodomain TF which is known to be activated in rice blast resistance response [72]. While the co-occurrence of GCCCORE and BIHD1OS indicated the JA mediated defense response in rice, the co-occurrence of GCCCORE and GT1CONSENSUS reflected a possible synergism among the JA and SA dependent defense responses.

These CRE combinations indicated a possible scenario in which SA and JA mediated defense signaling may be operated synergistically or antagonistically depending on the requirements of plant immuno-system. The differentially up-regulated genes of these hormone (SA, JA, CK, ABA and ET) biosynthetic processes in infected tissues also suggested their involvement in plant defense response under fungal attack. Finally, this study illustrated the SA and JA mediated transcriptional regulation among defense response genes by multiple TF families and thus, depicting the cross-talk of different hormonal signaling pathways in defense response mechanism.

Similarities and dissimilarities in CRE combinatorics in blast infected leaf and root

Previously, Marcel et al. [28] established that the biological mode of blast infection is different in leaf and root, and this causes tissue-specific defense response mechanisms of rice. They showed that the gene expression pattern of 2 DPI root sample was similar to the two leaf samples (3 DPI and 4 DPI) and a closer similarity of infection profile was found with 4 DPI leaf. Using these two data sets (4 DPI leaf and 2 DPI root) we investigated whether we can find any similarity/dissimilarity in combinatorial regulation.

We found a number of cliques those exhibit partial differences in between the two data sets. For example, a clique [GATABOX, WRKY71OS, MYCCONSENSUSAT] occurred in the WRKY promoters in leaf and another clique [WRKY71OS, MYCCONSENSUSAT, CURECORECR] was identified in the WRKY promoters in root (Fig 6). While these two cliques share the common CRE pair (WRKY71OS, MYCCONSENSUSAT), the clique in leaf possessed GATABOX (involved in light-response) and the clique in root had CURECORECR (involved in copper and oxygen response) as the non-overlapping nodes. Another example is the following two cliques: [ASF1MOTIFCAMV, MYBCORE, WBOXNTERF3] and [ASF1MOTIFCAMV, MYBCORE, CCAATBOX1]. The first one was identified in leaf promoters and the second one was found in root promoters of WRKY family genes. The biological roles of ASF1MOTIFCAMV, MYBCORE and WBOXNTERF3 were discussed in the previous sections. CCAATBOX1 (binding site of NF-Y) was found in the promoters of HSP (Heat shock protein) genes of tobacco and functioning cooperatively with HSEs (Heat shock elements) to induce the heat-shock activity [73].

The differences were found not only in the cliques but also in the frequencies of the CRE-involvement in the cliques. In the CRE co-occurrence network of “Response to stress”, several MYC recognition sites (MYCCONSENSUSAT, MYCATERD1 and MYCATRD22) were found in different cliques of leaf whereas they were absent in the cliques of root. Significant differences in MYB binding sites and their frequencies of involvement in between the cliques of leaf and root were also observed. We found 9 MYB binding sites (MYBCOREATCYCB1, MYBPLANT, MYB2AT, TATCCAOSAMY, BOXLCOREDPCAL, MYB2CONSENSUSAT and MYBGAHV) in the cliques of leaf, while only 2 of them (MYBST1, MYBPZM) were present in the cliques of root with a significantly lower frequency. In addition, 4 WRKY binding sites (WBOXNTCHN48, WBOXATNPR1, WBOXHVISO1 and WRKY71OS) were involved in different cliques of leaf but only one (WBOXATNPR1) appeared in a single clique of root. This observation showed a clear difference of the co-occurrence tendencies of some TF binding site families (e.g., MYC, MYB and WRKY binding sites) between these two data sets.

Our analysis revealed that in spite of their presence (as nodes) in the co-occurrence networks of both leaf and root, several CREs formed cliques in any one of the tissues. For example, in both (leaf and root) co-occurrence networks of “Defense response” genes, 5 CREs (CCAATBOX1, MYCCONSENSUSAT, WBOXATNPR1, WBOXHVISO1 and ARR1AT) were present as common nodes. But the existing edges (co-occurrence relationship) among these CREs were not common in both the networks. In leaf network, 3 CREs (CCAATBOX1, WBOXATNPR1 and ARR1AT) formed a clique whereas, in root network 4 CREs (CCAATBOX1, WBOXATNPR1, MYCCONSENSUSAT and WBOXHVISO1) were combined in a clique (Fig 6). The clique of 3 CREs (CCAATBOX1, WBOXATNPR1 and MYCCONSENSUSAT) was found as the common clique in both cases.

The CRE co-occurrence networks of up-regulated gene-sets of leaf and root shared some common as well as some unique CREs. These results clearly showed that there exists some similarities as well as some dissimilarities in regulatory combinatorics in between the leaf and

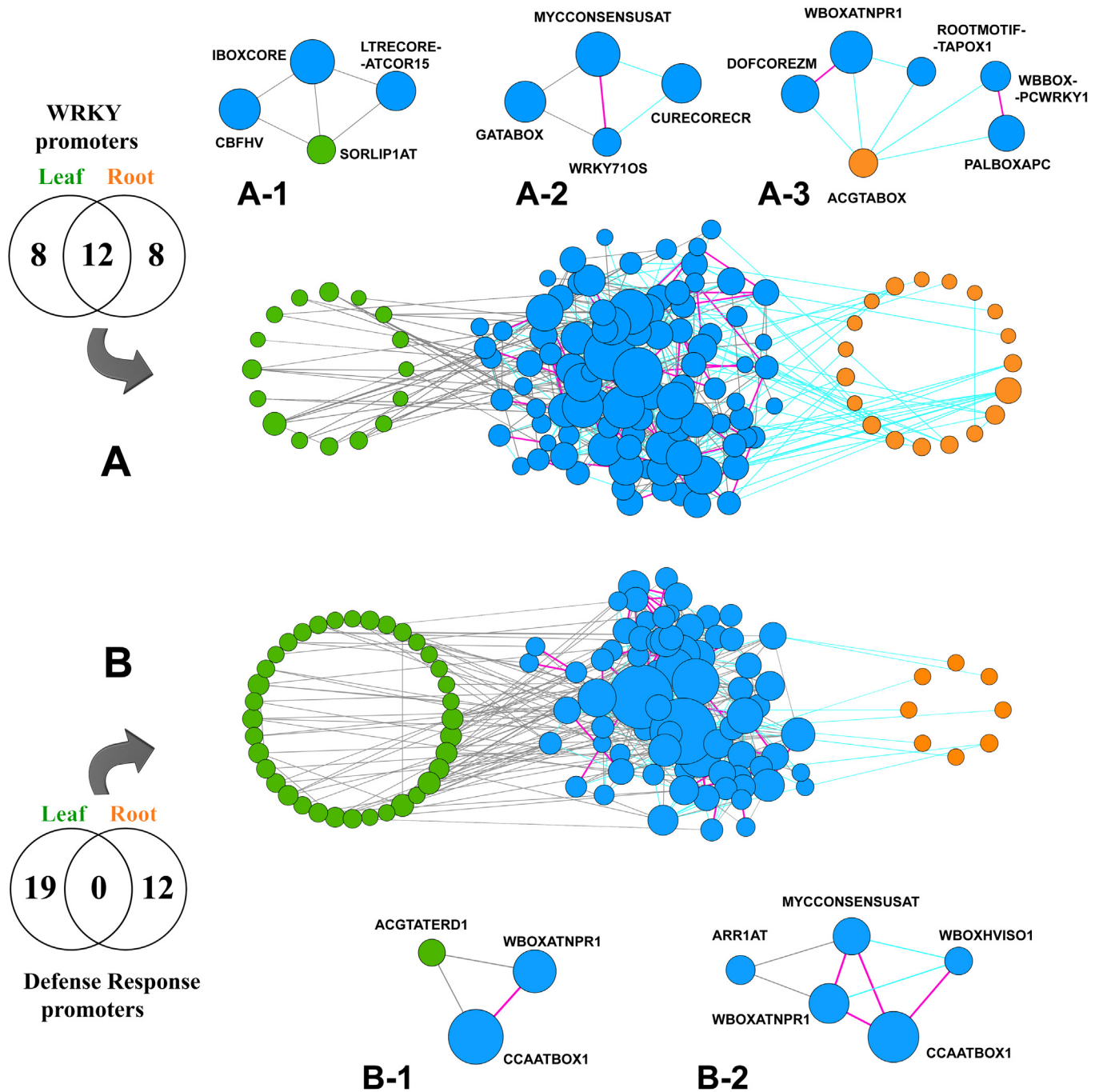


Fig 6. Co-occurrence networks of two pairs of input gene-sets (derived from 4 DPI leaf and 2 DPI root data) and their respective cliques are presented. (A) The CRE co-occurrence networks computed for the up-regulated WRKY genes in rice leaf and root tissues are merged into a single one. Nodes (individual CREs) explicitly found in the leaf are highlighted (green) in the left, while those found in the root are highlighted (orange) in the right hand side of the network. Nodes in the middle (blue) are found in both leaf and root tissues. Gray and cyan colour of edges imply that a significant co-occurrence is found between the two connected CREs explicitly in leaf and root respectively. A pink edge implies that the two connected CREs significantly co-occur in both the tissues. **(A-1)** An example of two cliques [SORLIP1AT, CBFHV, IBOXCORE] and [SORLIP1AT, IBOXCORE, LTRECOREATCOR15], explicitly occurred in the network of leaf. **(A-2)** An example of two cliques, found in leaf and root respectively, where they share a common edge (WRKY71OS—MYCCONSENSUSAT). **(A-3)** An example of three cliques [ACGTABOX, DOFCOREZM, WBOXATNPR1], [ACGTABOX, WBOXATNPR1, ROOTMOTIFTAPOX1] and [ACGTABOX, PALBOXAPC, WBOXPCWRKY1], out of which one clique explicitly occurred in root. **(B)** The merged CRE co-occurrence networks of rice leaf and root, computed for the “Defense Response” gene-sets. **(B-1)** An example of a clique, occurred in the network of leaf. **(B-2)** An example of four cliques, out of which one occurred in the network of leaf, two occurred in the network of root and one is common [CCAATBOX1, WBOXATNPR1, MYCCONSENSUSAT] in both the networks.

doi:10.1371/journal.pone.0137295.g006

root data sets. These dissimilarities might be indicative of the following possibilities: (i) the different biological modes of infection in leaf and root, which induce different defense response patterns, or (ii) the difference of time points (4 DPI in leaf and 2 DPI in root) between the two data sets, or (iii) both.

Network topology indicates quick transcriptional reprogramming under blast infected condition is highly facilitated

Co-occurrence of CREs at the promoters of a gene implies that TFs are likely acting in a co-associative fashion, involving different TF families. Therefore, a CRE co-occurrence network represents transcriptional co-regulation of a set of genes. On the other hand, a clique (co-occurrence combination) of CREs represents one of the possible higher order combinations of TFs those can regulate the expression of that gene. Thus, a clique-clique network illustrates the different possible higher order regulatory combinatorics ($k \geq 3$) and the associations among them. We computed two topological parameters of the CRE co-occurrence networks and the clique-clique networks to obtain their inherent properties.

Assortative mixing is a particular network topology which quantifies the tendency of preferential association of the nodes in a network [38, 74] (Eq 3). A network is said to show assortative mixing, if the nodes of similar degree tend to be connected, and disassortative, when nodes of dissimilar degree remain connected. Similarly, the network clustering coefficient (\bar{C}) is a global topological parameter which reflects the overall cohesiveness of a network (Eq 4). For each CRE co-occurrence network, we computed the coefficients of assortativity (r) and the network clustering coefficients (\bar{C}) (Table 2). CRE co-occurrence networks exhibited disassortative mixing and smaller clustering coefficient values. We transformed each CRE co-occurrence network into a clique-clique network (undirected), in which the nodes represent the cliques of

Table 2. Network property of CRE co-occurrence networks and clique-clique networks of the input gene-sets. The coefficient of assortativity (r) reflects the assortative mixing topology of a network. Negative values of r of the CRE co-occurrence networks indicate their disassortative mixing topology whereas, positive values of r observed in clique-clique networks imply their assortative mixing topology. Further, the network clustering coefficient (\bar{C}) values of the clique-clique networks are significantly higher than their respective CRE co-occurrence networks and this global attribute is an indicator of cohesiveness of the clique-clique networks. Significance is tested by permutation Mann-Whitney U test (No. of permutation = 10000), $p < 0.001$.

Input gene-set	Tissue	CRE co-occurrence network		clique-clique network	
		r	\bar{C}	r	\bar{C}
WRKY	Leaf	-0.176	0.052	0.238	0.690
WRKY	Root	-0.176	0.127	0.504	0.684
MYB	Leaf	-0.131	0.118	0.536	0.725
MYB	Root	-0.161	0.134	0.094	0.775
ERF	Leaf	-0.138	0.165	0.433	0.675
ERF	Root	-0.104	0.127	0.439	0.622
NAC	Leaf	-0.007	0.086	0.627	0.435
NAC	Root	—	—	—	—
Defense response	Leaf	-0.116	0.211	0.273	0.828
Defense response	Root	-0.132	0.209	0.112	0.871
Response to stress	Leaf	-0.161	0.133	0.538	0.672
Response to stress	Root	-0.103	0.135	0.325	0.679
Chitin catabolic process	Leaf	-0.033	0.231	0.459	0.723
Chitin catabolic process	Root	—	—	—	—
GPCR signaling pathway	Leaf	-0.07	0.115	0.365	0.781
GPCR signaling pathway	Root	-0.006	0.237	0.313	0.721

doi:10.1371/journal.pone.0137295.t002

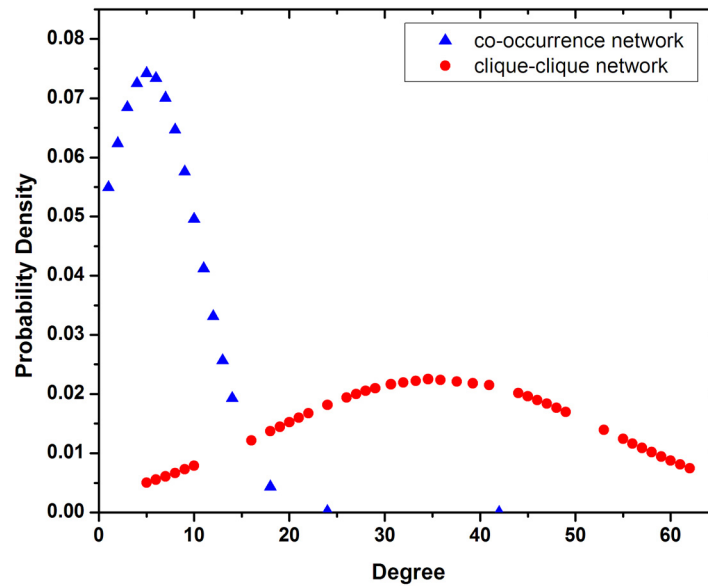


Fig 7. Probability density of degree distribution of CRE co-occurrence network and clique-clique network. The degree distribution pattern shows that there is a better distributed architecture in the clique-clique network compared to the CRE co-occurrence network.

doi:10.1371/journal.pone.0137295.g007

CREs and two cliques are connected if they have ≥ 1 sharing CREs. Surprisingly, this transformation entirely changed the inherent topology of the networks. In contrast to the disassortative nature of CRE co-occurrence networks, the clique-clique networks exhibited assortative mixing topology with significantly elevated network clustering coefficient values. For example, the CRE co-occurrence network of MYB TF-family in rice leaf exhibited disassortative topology ($r = -0.131$), while its clique-clique network was assortative ($r = 0.536$). In addition, its network clustering coefficient also elevated from 0.118 to 0.725 (Table 2). Balaji et al. [75] transformed the Yeast transcriptional regulatory network into a co-regulatory network (looking into higher order of combinations); this network transformation revealed the hidden distributed architecture. They observed that at co-regulatory level, a distributed architecture emerges, in which there exists a large number of coordinating partners for any transcription factor. Here, transforming the CRE co-occurrence network into a clique-clique network (represents even higher order of co-regulation) we showed that the later one reflects stronger distributed architecture (Fig 7), supported by high associativity and cohesiveness.

In cellular system, an efficient reprogramming of transcriptional regulation is necessary for quick response against external perturbation like pathogen attack. A clique-clique network represents the higher order combinatorics of transcription regulation of a set of genes, the ability of the transcriptional reprogramming should be reflected in its topology. Here, we observed that they simultaneously exhibit strong assortative mixing of nodes as well as higher clustering coefficients. These features illustrate that the CRE combinations, central to the regulation, tend to associate with several other CRE combinations those are also central to the regulation. This might be an indication of the ability of efficient alteration in the combinatorial logic of CREs so that a quick change of transcriptional regulation may occur at the time of blast infection.

Summary

Here, we proposed a methodology that first quantified the co-occurrence tendencies of CREs in a gene-set of interest, filtered these scores to obtain significantly co-occurring CRE pairs,

transformed these pairs into networks and analyzed these networks to fetch higher order CRE combinatorics. This allowed us to perform system level analysis of regulatory phenomena. To demonstrate the potential applicability of our methodology, we applied it to the blast infected condition of rice tissues.

We observed that strong co-occurrence of some CRE pairs are associated with a specific positional distribution pattern of their joint occurrences at the promoter regions. In a few cases, such highly similar positional distribution and simultaneous strong co-occurrence tendency of CREs were associated with the respective TF cooperativity.

Transcriptional control of gene expression involves auto/cross regulation among various TF families. The co-occurring CREs found at the promoters of TF genes indicated auto-regulation and cross regulation. In this study, preparing CRE co-occurrence networks, we showed that the essence of cross-talk among multiple hormone signaling pathways can be captured in the form of cliques of CREs. For example, during blast infection, we found that CREs associated with SA and JA mediated signaling pathways made up several cliques in an intermingled fashion. CRE co-occurrence networks of leaf and root data sets revealed a number of differences as well as similarities of CRE combinatorics in between them.

Social networks exhibit assortative mixing topology which is generally lacking in the non-social networks (e.g., biological networks) [76, 77], with a few exceptions like, amino acid contact networks [78] and co-evolutionary networks within macromolecular complexes [79]. Here, the disassortative CRE co-occurrence networks were transformed into assortative clique-clique networks. This observation clearly unraveled the hidden property of the higher order of combinatorial regulation with an indication that biological networks might exhibit a social network-like behavior.

Transcription regulation and its involvement in altering cellular physiology under various stressed conditions has been proven difficult to resolve due to the inherent complexity of the process. As a result, even after decades of research, a systematic description of *in vivo* transcription regulation remains elusive. In this study, we conveyed a potential methodology to compute co-occurrence tendencies among cis-regulatory elements and its application to reveal the regulatory cross-talk among multiple defense signaling pathways. One can apply this computational approach to explore the combinatorial gene regulation in other eukaryotes, under any cellular condition.

Supporting Information

S1 Text. This file describes the significance of the new metric COR value to estimate the co-occurrence tendencies of CREs. The calculation of COR value under different possible cases are illustrated. In addition, the significance of COR value cutoff 1.5 is shown here. (PDF)

S1 Fig. Differentially up-regulated rice genes under blast infected conditions. (A) Differentially up-regulated TFs in infected leaf and root of rice. The X axis represents individual TF-family and the Y axis represents the numbers of differentially up-regulated TF genes from each family. (B), (C) The pie charts representing the GO biological processes which mapped to ≥ 10 differentially up-regulated loci of infected rice leaf and root, respectively. In leaf 882 (out of 1201) loci and in root 502 (out of 677) loci are shown here. (D) The intersect region of each Venn diagram indicates the number of common up-regulated loci of respective gene-sets in between leaf (4 DPI) and root (2 DPI). (PNG)

S1 Table. Detailed information about the input gene-sets and the results obtained by using the proposed methodology.

(XLS)

Acknowledgments

The authors acknowledge Saurav Mallik, University of Calcutta, for his critical assessments on the manuscript, assistance in network analysis and discussions on the results. AD acknowledges UGC(RFSMS), Government of India, for his fellowship.

Author Contributions

Conceived and designed the experiments: AD SK. Performed the experiments: AD. Analyzed the data: AD. Contributed reagents/materials/analysis tools: AD. Wrote the paper: AD SK.

References

1. Mitchell PJ, Tjian R. Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. *Science*. 1989; 245(4916):371–378. doi: [10.1126/science.2667136](https://doi.org/10.1126/science.2667136) PMID: [2667136](https://pubmed.ncbi.nlm.nih.gov/2667136/)
2. Levine M, Tjian R. Transcription regulation and animal diversity. *Nature*. 2003; 424(6945):147–151. doi: [10.1038/nature01763](https://doi.org/10.1038/nature01763) PMID: [12853946](https://pubmed.ncbi.nlm.nih.gov/12853946/)
3. Brem RB, Yvert G, Clinton R, Kruglyak L. Genetic dissection of transcriptional regulation in budding yeast. *Science*. 2002; 296(5568):752–755. doi: [10.1126/science.1069516](https://doi.org/10.1126/science.1069516) PMID: [11923494](https://pubmed.ncbi.nlm.nih.gov/11923494/)
4. Benfey PN, Chua NH. The cauliflower mosaic virus 35S promoter: combinatorial regulation of transcription in plants. *Science*. 1990; 250(4983):959–966. doi: [10.1126/science.250.4983.959](https://doi.org/10.1126/science.250.4983.959) PMID: [17746920](https://pubmed.ncbi.nlm.nih.gov/17746920/)
5. Singh KB. Transcriptional regulation in plants: the importance of combinatorial control. *Plant Physiology*. 1998; 118(4):1111–1120. doi: [10.1104/pp.118.4.1111](https://doi.org/10.1104/pp.118.4.1111) PMID: [9847085](https://pubmed.ncbi.nlm.nih.gov/9847085/)
6. Reményi A, Scholer H, Wilmanns M. Combinatorial control of gene expression. *Nature structural & molecular biology*. 2004; 11(9):812–815. doi: [10.1038/nsmb820](https://doi.org/10.1038/nsmb820)
7. Pilpel Y, Sudarsanam P, Church GM. Identifying regulatory networks by combinatorial analysis of promoter elements. *Nature genetics*. 2001; 29(2):153–159. doi: [10.1038/ng724](https://doi.org/10.1038/ng724) PMID: [11547334](https://pubmed.ncbi.nlm.nih.gov/11547334/)
8. Sudarsanam P, Pilpel Y, Church GM. Genome-wide co-occurrence of promoter elements reveals a cis-regulatory cassette of rRNA transcription motifs in *Saccharomyces cerevisiae*. *Genome research*. 2002; 12(11):1723–1731. doi: [10.1101/gr.301202](https://doi.org/10.1101/gr.301202) PMID: [12421759](https://pubmed.ncbi.nlm.nih.gov/12421759/)
9. Kato M, Hata N, Banerjee N, Fitcher B, Zhang MQ. Identifying combinatorial regulation of transcription factors and binding motifs. *Genome biology*. 2004; 5(8):R56. doi: [10.1186/gb-2004-5-8-r56](https://doi.org/10.1186/gb-2004-5-8-r56) PMID: [15287978](https://pubmed.ncbi.nlm.nih.gov/15287978/)
10. Yu X, Lin J, Masuda T, Esumi N, Zack DJ, Qian J. Genome-wide prediction and characterization of interactions between transcription factors in *Saccharomyces cerevisiae*. *Nucleic acids research*. 2006; 34(3):917–927. doi: [10.1093/nar/gkj487](https://doi.org/10.1093/nar/gkj487) PMID: [16464824](https://pubmed.ncbi.nlm.nih.gov/16464824/)
11. Murakami K, Imanishi T, Gojbori T, Nakai K. Two different classes of co-occurring motif pairs found by a novel visualization method in human promoter regions. *BMC genomics*. 2008; 9(1):112. doi: [10.1186/1471-2164-9-112](https://doi.org/10.1186/1471-2164-9-112) PMID: [18312685](https://pubmed.ncbi.nlm.nih.gov/18312685/)
12. Berman BP, Nibu Y, Pfeiffer BD, Tomancak P, Celniker SE, Levine M, et al. Exploiting transcription factor binding site clustering to identify cis-regulatory modules involved in pattern formation in the *Drosophila* genome. *Proceedings of the National Academy of Sciences*. 2002; 99(2):757–762. doi: [10.1073/pnas.231608898](https://doi.org/10.1073/pnas.231608898)
13. Hannenhalli S, Levy S. Predicting transcription factor synergism. *Nucleic acids research*. 2002; 30(19):4278–4284. doi: [10.1093/nar/gkf535](https://doi.org/10.1093/nar/gkf535) PMID: [12364607](https://pubmed.ncbi.nlm.nih.gov/12364607/)
14. Pati A, Vasquez-Robinet C, Heath L, Grene R, Murali T. XcisClique: analysis of regulatory bicliques. *BMC bioinformatics*. 2006; 7(1):218. doi: [10.1186/1471-2105-7-218](https://doi.org/10.1186/1471-2105-7-218) PMID: [16630346](https://pubmed.ncbi.nlm.nih.gov/16630346/)
15. Chang WC, Lee TY, Huang HD, Huang HY, Pan RL. PlantPAN: Plant promoter analysis navigator, for identifying combinatorial cis-regulatory elements with distance constraint in plant gene groups. *Bmc Genomics*. 2008; 9(1):561. doi: [10.1186/1471-2164-9-561](https://doi.org/10.1186/1471-2164-9-561) PMID: [19036138](https://pubmed.ncbi.nlm.nih.gov/19036138/)

16. Vandenberg A, Kumagai Y, Akira S, Standley DM. A novel unbiased measure for motif co-occurrence predicts combinatorial regulation of transcription. *BMC genomics*. 2012; 13(Suppl 7):S11. doi: [10.1186/1471-2164-13-S7-S11](https://doi.org/10.1186/1471-2164-13-S7-S11) PMID: [23282148](https://pubmed.ncbi.nlm.nih.gov/23282148/)
17. Balaji S, Babu MM, Iyer LM, Luscombe NM, Aravind L. Comprehensive analysis of combinatorial regulation using the transcriptional regulatory network of yeast. *Journal of molecular biology*. 2006; 360(1):213–227. doi: [10.1016/j.jmb.2006.04.029](https://doi.org/10.1016/j.jmb.2006.04.029) PMID: [16762362](https://pubmed.ncbi.nlm.nih.gov/16762362/)
18. Lindlöf A, Bräutigam M, Chawade A, Olsson O, Olsson B. In silico analysis of promoter regions from cold-induced genes in rice (*Oryza sativa* L.) and *Arabidopsis thaliana* reveals the importance of combinatorial control. *Bioinformatics*. 2009; 25(11):1345–1348. doi: [10.1093/bioinformatics/btp172](https://doi.org/10.1093/bioinformatics/btp172) PMID: [19321735](https://pubmed.ncbi.nlm.nih.gov/19321735/)
19. Mohanty B, Herath V, Wijaya E, Yeo HC, de los Reyes BG, Lee DY. Patterns of cis-element enrichment reveal potential regulatory modules involved in the transcriptional regulation of anoxia response of japonica rice. *Gene*. 2012; 511(2):235–242. doi: [10.1016/j.gene.2012.09.048](https://doi.org/10.1016/j.gene.2012.09.048) PMID: [23010196](https://pubmed.ncbi.nlm.nih.gov/23010196/)
20. Talbot NJ. On the trail of a cereal killer: exploring the biology of *Magnaporthe grisea*. *Annual Reviews in Microbiology*. 2003; 57(1):177–202. doi: [10.1146/annurev.micro.57.030502.090957](https://doi.org/10.1146/annurev.micro.57.030502.090957)
21. Lenka SK, Lohia B, Kumar A, Chinnusamy V, Bansal KC. Genome-wide targeted prediction of ABA responsive genes in rice based on over-represented cis-motif in co-expressed genes. *Plant molecular biology*. 2009; 69(3):261–271. doi: [10.1007/s11103-008-9423-4](https://doi.org/10.1007/s11103-008-9423-4) PMID: [18998058](https://pubmed.ncbi.nlm.nih.gov/18998058/)
22. Ouyang S, Zhu W, Hamilton J, Lin H, Campbell M, Childs K, et al. The TIGR rice genome annotation resource: improvements and new features. *Nucleic acids research*. 2007; 35(suppl 1):D883–D887. doi: [10.1093/nar/gkl976](https://doi.org/10.1093/nar/gkl976) PMID: [17145706](https://pubmed.ncbi.nlm.nih.gov/17145706/)
23. Smit AF, Hubley R, Green P. RepeatMasker Open-3.0; 1996.
24. Higo K, Ugawa Y, Iwamoto M, Korenaga T. Plant cis-acting regulatory DNA elements (PLACE) database: 1999. *Nucleic acids research*. 1999; 27(1):297–300. doi: [10.1093/nar/27.1.297](https://doi.org/10.1093/nar/27.1.297) PMID: [9847208](https://pubmed.ncbi.nlm.nih.gov/9847208/)
25. Prestridge DS. SIGNAL SCAN: a computer program that scans DNA sequences for eukaryotic transcriptional elements. *Computer applications in the biosciences: CABIOS*. 1991; 7(2):203–206. PMID: [2059845](https://pubmed.ncbi.nlm.nih.gov/2059845/)
26. Edgar R, Domrachev M, Lash AE. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic acids research*. 2002; 30(1):207–210. doi: [10.1093/nar/30.1.207](https://doi.org/10.1093/nar/30.1.207) PMID: [11752295](https://pubmed.ncbi.nlm.nih.gov/11752295/)
27. Ribot C, Hirsch J, Balzergue S, Tharreau D, Nottéghem JL, Lebrun MH, et al. Susceptibility of rice to the blast fungus, *Magnaporthe grisea*. *Journal of plant physiology*. 2008; 165(1):114–124. doi: [10.1016/j.jplph.2007.06.013](https://doi.org/10.1016/j.jplph.2007.06.013) PMID: [17905473](https://pubmed.ncbi.nlm.nih.gov/17905473/)
28. Marcel S, Sawers R, Oakeley E, Angliker H, Paszkowski U. Tissue-adapted invasion strategies of the rice blast fungus *Magnaporthe oryzae*. *The Plant Cell*. 2010; 22(9):3177–3187. doi: [10.1105/tpc.110.078048](https://doi.org/10.1105/tpc.110.078048) PMID: [20858844](https://pubmed.ncbi.nlm.nih.gov/20858844/)
29. Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, et al. Bioconductor: open software development for computational biology and bioinformatics. *Genome biology*. 2004; 5(10):R80. doi: [10.1186/gb-2004-5-10-r80](https://doi.org/10.1186/gb-2004-5-10-r80) PMID: [15461798](https://pubmed.ncbi.nlm.nih.gov/15461798/)
30. Liu G, Loraine AE, Shigeta R, Cline M, Cheng J, Valmeekam V, et al. NetAffx: Affymetrix probesets and annotations. *Nucleic acids research*. 2003; 31(1):82–86. doi: [10.1093/nar/gkg121](https://doi.org/10.1093/nar/gkg121) PMID: [12519953](https://pubmed.ncbi.nlm.nih.gov/12519953/)
31. Ficklin SP, Luo F, Feltus FA. The association of multiple interacting genes with specific phenotypes in rice using gene coexpression networks. *Plant physiology*. 2010; 154(1):13–24. doi: [10.1104/pp.110.159459](https://doi.org/10.1104/pp.110.159459) PMID: [20668062](https://pubmed.ncbi.nlm.nih.gov/20668062/)
32. Risueño A, Fontanillo C, Dinger M, De Las Rivas J. GATEExplorer: genomic and transcriptomic explorer; mapping expression probes to gene loci, transcripts, exons and ncRNAs. *BMC bioinformatics*. 2010; 11(1):221. doi: [10.1186/1471-2105-11-221](https://doi.org/10.1186/1471-2105-11-221) PMID: [20429936](https://pubmed.ncbi.nlm.nih.gov/20429936/)
33. Dozmorov M, Giles C, Wren J. Predicting gene ontology from a global meta-analysis of 1-color microarray experiments. *BMC bioinformatics*. 2011; 12(Suppl 10):S14. doi: [10.1186/1471-2105-12-S10-S14](https://doi.org/10.1186/1471-2105-12-S10-S14) PMID: [22166114](https://pubmed.ncbi.nlm.nih.gov/22166114/)
34. Cao P, Jung KH, Choi D, Hwang D, Zhu J, Ronald PC. The rice oligonucleotide array database: an atlas of rice gene expression. *Rice*. 2012; 5(1):1–9. doi: [10.1186/1939-8433-5-17](https://doi.org/10.1186/1939-8433-5-17)
35. Jin J, Zhang H, Kong L, Gao G, Luo J. PlantTFDB 3.0: a portal for the functional and evolutionary study of plant transcription factors. *Nucleic acids research*. 2014; 42(D1):D1182–D1187. doi: [10.1093/nar/gkt1016](https://doi.org/10.1093/nar/gkt1016) PMID: [24174544](https://pubmed.ncbi.nlm.nih.gov/24174544/)
36. Adamcsek B, Palla G, Farkas IJ, Derényi I, Vicsek T. CFinder: locating cliques and overlapping modules in biological networks. *Bioinformatics*. 2006; 22(8):1021–1023. doi: [10.1093/bioinformatics/btl039](https://doi.org/10.1093/bioinformatics/btl039) PMID: [16473872](https://pubmed.ncbi.nlm.nih.gov/16473872/)

37. Smoot ME, Ono K, Ruscheinski J, Wang PL, Ideker T. Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics*. 2011; 27(3):431–432. doi: [10.1093/bioinformatics/btq675](https://doi.org/10.1093/bioinformatics/btq675) PMID: [21149340](https://pubmed.ncbi.nlm.nih.gov/21149340/)
38. Newman ME. Assortative mixing in networks. *Physical review letters*. 2002; 89(20):208701. doi: [10.1103/PhysRevLett.89.208701](https://doi.org/10.1103/PhysRevLett.89.208701) PMID: [12443515](https://pubmed.ncbi.nlm.nih.gov/12443515/)
39. Wei T, Ou B, Li J, Zhao Y, Guo D, Zhu Y, et al. Transcriptional Profiling of Rice Early Response to Magnaporthe oryzae Identified OsWRKYs as Important Regulators in Rice Blast Resistance. *PLoS one*. 2013; 8(3):e59720. doi: [10.1371/journal.pone.0059720](https://doi.org/10.1371/journal.pone.0059720) PMID: [23544090](https://pubmed.ncbi.nlm.nih.gov/23544090/)
40. Whitmarsh AJ. Regulation of gene transcription by mitogen-activated protein kinase signaling pathways. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*. 2007; 1773(8):1285–1298. doi: [10.1016/j.bbamcr.2006.11.011](https://doi.org/10.1016/j.bbamcr.2006.11.011)
41. Pandey SP, Somssich IE. The role of WRKY transcription factors in plant immunity. *Plant physiology*. 2009; 150(4):1648–1655. doi: [10.1104/pp.109.138990](https://doi.org/10.1104/pp.109.138990) PMID: [19420325](https://pubmed.ncbi.nlm.nih.gov/19420325/)
42. Van Verk MC, Gatz C, Linthorst HJ. Transcriptional regulation of plant defense responses. *Advances in Botanical Research*. 2009; 51:397–438. doi: [10.1016/S0065-2296\(09\)51010-5](https://doi.org/10.1016/S0065-2296(09)51010-5)
43. Raffaele S, Rivas S. Regulate and be regulated: integration of defense and other signals by the AtMYB30 transcription factor. *Frontiers in plant science*. 2013; 4. doi: [10.3389/fpls.2013.00098](https://doi.org/10.3389/fpls.2013.00098) PMID: [23596456](https://pubmed.ncbi.nlm.nih.gov/23596456/)
44. Grennan AK. Ethylene response factors in jasmonate signaling and defense response. *Plant physiology*. 2008; 146(4):1457–1458. doi: [10.1104/pp.104.900254](https://doi.org/10.1104/pp.104.900254) PMID: [18390488](https://pubmed.ncbi.nlm.nih.gov/18390488/)
45. Pré M, Atallah M, Champion A, De Vos M, Pieterse CM, Memelink J. The AP2/ERF domain transcription factor ORA59 integrates jasmonic acid and ethylene signals in plant defense. *Plant Physiology*. 2008; 147(3):1347–1357. doi: [10.1104/pp.108.117523](https://doi.org/10.1104/pp.108.117523) PMID: [18467450](https://pubmed.ncbi.nlm.nih.gov/18467450/)
46. Lin R, Zhao W, Meng X, Wang M, Peng Y. Rice gene OsNAC19 encodes a novel NAC-domain transcription factor and responds to infection by Magnaporthe grisea. *Plant science*. 2007; 172(1):120–130. doi: [10.1016/j.plantsci.2006.07.019](https://doi.org/10.1016/j.plantsci.2006.07.019)
47. Sun L, Zhang H, Li D, Huang L, Hong Y, Ding XS, et al. Functions of rice NAC transcriptional factors, ONAC122 and ONAC131, in defense responses against Magnaporthe grisea. *Plant molecular biology*. 2013; 81(1-2):41–56. doi: [10.1007/s11103-012-9981-3](https://doi.org/10.1007/s11103-012-9981-3) PMID: [23103994](https://pubmed.ncbi.nlm.nih.gov/23103994/)
48. Consortium GO, et al. The Gene Ontology (GO) database and informatics resource. *Nucleic acids research*. 2004; 32(suppl 1):D258–D261. doi: [10.1093/nar/gkh036](https://doi.org/10.1093/nar/gkh036)
49. Nishizawa Y, Nishio Z, Nakazono K, Soma M, Nakajima E, Ugaki M, et al. Enhanced resistance to blast (Magnaporthe grisea) in transgenic Japonica rice by constitutive expression of rice chitinase. *Theoretical and applied genetics*. 1999; 99(3-4):383–390. doi: [10.1007/s001220051248](https://doi.org/10.1007/s001220051248) PMID: [22665169](https://pubmed.ncbi.nlm.nih.gov/22665169/)
50. Mentlak TA, Kombrink A, Shinya T, Ryder LS, Otomo I, Saitoh H, et al. Effector-mediated suppression of chitin-triggered immunity by Magnaporthe oryzae is necessary for rice blast disease. *The Plant Cell*. 2012; 24(1):322–335. doi: [10.1105/tpc.111.092957](https://doi.org/10.1105/tpc.111.092957) PMID: [22267486](https://pubmed.ncbi.nlm.nih.gov/22267486/)
51. Assmann SM. G protein regulation of disease resistance during infection of rice with rice blast fungus. *Science Signaling*. 2005; 2005(310):cm13–cm13. doi: [10.1126/stke.3102005cm13](https://doi.org/10.1126/stke.3102005cm13)
52. Vardhanabhuti S, Wang J, Hannenhalli S. Position and distance specificity are important determinants of cis-regulatory motifs in addition to evolutionary conservation. *Nucleic acids research*. 2007; 35(10):3203–3213. doi: [10.1093/nar/gkm201](https://doi.org/10.1093/nar/gkm201) PMID: [17452354](https://pubmed.ncbi.nlm.nih.gov/17452354/)
53. Puente P, Wei N, Deng XW. Combinatorial interplay of promoter elements constitutes the minimal determinants for light and developmental control of gene expression in Arabidopsis. *The EMBO journal*. 1996; 15(14):3732. PMID: [8670877](https://pubmed.ncbi.nlm.nih.gov/8670877/)
54. Zhou DX. Regulatory mechanism of plant gene transcription by GT-elements and GT-factors. *Trends in plant science*. 1999; 4(6):210–214. doi: [10.1016/S1360-1385\(99\)01418-1](https://doi.org/10.1016/S1360-1385(99)01418-1) PMID: [10366876](https://pubmed.ncbi.nlm.nih.gov/10366876/)
55. Banerjee N, Zhang MQ. Identifying cooperativity among transcription factors controlling the cell cycle in yeast. *Nucleic acids research*. 2003; 31(23):7024–7031. doi: [10.1093/nar/gkg894](https://doi.org/10.1093/nar/gkg894) PMID: [14627835](https://pubmed.ncbi.nlm.nih.gov/14627835/)
56. Baudry A, Caboche M, Lepiniec L. TT8 controls its own expression in a feedback regulation involving TTG1 and homologous MYB and bHLH factors, allowing a strong and cell-specific accumulation of flavonoids in Arabidopsis thaliana. *The Plant Journal*. 2006; 46(5):768–779. doi: [10.1111/j.1365-313X.2006.02733.x](https://doi.org/10.1111/j.1365-313X.2006.02733.x) PMID: [16709193](https://pubmed.ncbi.nlm.nih.gov/16709193/)
57. Métraux J, Signer H, Ryals J, Ward E, Wyss-Benz M, Gaudin J, et al. Increase in salicylic acid at the onset of systemic acquired resistance in cucumber. *Science*. 1990; 250(4983):1004–1006. doi: [10.1126/science.250.4983.1004](https://doi.org/10.1126/science.250.4983.1004) PMID: [17746926](https://pubmed.ncbi.nlm.nih.gov/17746926/)
58. Boter M, Ruiz-Rivero O, Abdeen A, Prat S. Conserved MYC transcription factors play a key role in jasmonate signaling both in tomato and Arabidopsis. *Genes & Development*. 2004; 18(13):1577–1591. doi: [10.1101/gad.297704](https://doi.org/10.1101/gad.297704)

59. Lorenzo O, Chico JM, Sánchez-Serrano JJ, Solano R. JASMONATE-INSENSITIVE1 encodes a MYC transcription factor essential to discriminate between different jasmonate-regulated defense responses in Arabidopsis. *The Plant Cell Online*. 2004; 16(7):1938–1950. doi: [10.1105/tpc.022319](https://doi.org/10.1105/tpc.022319)
60. Kunkel BN, Brooks DM. Cross talk between signaling pathways in pathogen defense. *Current opinion in plant biology*. 2002; 5(4):325–331. doi: [10.1016/S1369-5266\(02\)00275-3](https://doi.org/10.1016/S1369-5266(02)00275-3) PMID: [12179966](https://pubmed.ncbi.nlm.nih.gov/12179966/)
61. Koornneef A, Leon-Reyes A, Ritsema T, Verhage A, Den Otter FC, Van Loon L, et al. Kinetics of salicylate-mediated suppression of jasmonate signaling reveal a role for redox modulation. *Plant Physiology*. 2008; 147(3):1358–1368. doi: [10.1104/pp.108.121392](https://doi.org/10.1104/pp.108.121392) PMID: [18539774](https://pubmed.ncbi.nlm.nih.gov/18539774/)
62. Liu X, Bai X, Wang X, Chu C. OsWRKY71, a rice transcription factor, is involved in rice defense response. *Journal of plant physiology*. 2007; 164(8):969–979. doi: [10.1016/j.jplph.2006.07.006](https://doi.org/10.1016/j.jplph.2006.07.006) PMID: [16919842](https://pubmed.ncbi.nlm.nih.gov/16919842/)
63. Xie Z, Zhang ZL, Zou X, Yang G, Komatsu S, Shen QJ. Interactions of two abscisic-acid induced WRKY genes in repressing gibberellin signaling in aleurone cells. *The Plant Journal*. 2006; 46(2):231–242. doi: [10.1111/j.1365-313X.2006.02694.x](https://doi.org/10.1111/j.1365-313X.2006.02694.x) PMID: [16623886](https://pubmed.ncbi.nlm.nih.gov/16623886/)
64. Abe H, Urao T, Ito T, Seki M, Shinozaki K, Yamaguchi-Shinozaki K. Arabidopsis AtMYC2 (bHLH) and AtMYB2 (MYB) function as transcriptional activators in abscisic acid signaling. *The Plant Cell Online*. 2003; 15(1):63–78. doi: [10.1105/tpc.006130](https://doi.org/10.1105/tpc.006130)
65. Buchel AS, Brederode FT, Bol JF, Linthorst HJ. Mutation of GT-1 binding sites in the Pr-1A promoter influences the level of inducible gene expression in vivo. *Plant molecular biology*. 1999; 40(3):387–396. doi: [10.1023/A:1006144505121](https://doi.org/10.1023/A:1006144505121) PMID: [10437823](https://pubmed.ncbi.nlm.nih.gov/10437823/)
66. Yanagisawa S. Dof domain proteins: plant-specific transcription factors associated with diverse phenomena unique to plants. *Plant and Cell Physiology*. 2004; 45(4):386–391. doi: [10.1093/pcp/pch055](https://doi.org/10.1093/pcp/pch055) PMID: [15111712](https://pubmed.ncbi.nlm.nih.gov/15111712/)
67. Yu D, Chen C, Chen Z. Evidence for an important role of WRKY DNA binding proteins in the regulation of NPR1 gene expression. *The Plant Cell Online*. 2001; 13(7):1527–1540. doi: [10.2307/3871384](https://doi.org/10.2307/3871384)
68. Ross EJ, Stone JM, Elowsky CG, Arredondo-Peter R, Klucas RV, Sarath G. Activation of the *Oryza sativa* non-symbiotic haemoglobin-2 promoter by the cytokinin-regulated transcription factor, ARR1. *Journal of experimental botany*. 2004; 55(403):1721–1731. doi: [10.1093/jxb/erh211](https://doi.org/10.1093/jxb/erh211) PMID: [15258171](https://pubmed.ncbi.nlm.nih.gov/15258171/)
69. Choi J, Choi D, Lee S, Ryu CM, Hwang I. Cytokinins and plant immunity: old foes or new friends? *Trends in plant science*. 2011; 16(7):388–394. doi: [10.1016/j.tplants.2011.03.003](https://doi.org/10.1016/j.tplants.2011.03.003) PMID: [21470894](https://pubmed.ncbi.nlm.nih.gov/21470894/)
70. Bate N, Twell D. Functional architecture of a late pollen promoter: pollen-specific transcription is developmentally regulated by multiple stage-specific and co-dependent activator elements. *Plant molecular biology*. 1998; 37(5):859–869. doi: [10.1023/A:1006095023050](https://doi.org/10.1023/A:1006095023050) PMID: [9678581](https://pubmed.ncbi.nlm.nih.gov/9678581/)
71. Brown RL, Kazan K, McGrath KC, Maclean DJ, Manners JM. A role for the GCC-box in jasmonate-mediated activation of the PDF1.2 gene of Arabidopsis. *Plant Physiology*. 2003; 132(2):1020–1032. doi: [10.1104/pp.102.017814](https://doi.org/10.1104/pp.102.017814) PMID: [12805630](https://pubmed.ncbi.nlm.nih.gov/12805630/)
72. Luo H, Song F, Goodman R, Zheng Z. Up-Regulation of OsBIHD1, a Rice Gene Encoding BELL Homeodomain Transcriptional Factor, in Disease Resistance Responses. *Plant Biology*. 2005; 7(5):459–468. doi: [10.1055/s-2005-865851](https://doi.org/10.1055/s-2005-865851) PMID: [16163610](https://pubmed.ncbi.nlm.nih.gov/16163610/)
73. Rieping M, Schöffl F. Synergistic effect of upstream sequences, CCAAT box elements, and HSE sequences for enhanced expression of chimaeric heat shock genes in transgenic tobacco. *Molecular and General Genetics MGG*. 1992; 231(2):226–232. PMID: [1736093](https://pubmed.ncbi.nlm.nih.gov/1736093/)
74. Newman ME. Mixing patterns in networks. *Physical Review E*. 2003; 67(2):026126. doi: [10.1103/PhysRevE.67.026126](https://doi.org/10.1103/PhysRevE.67.026126)
75. Balaji S, Iyer LM, Aravind L, Babu MM. Uncovering a hidden distributed architecture behind scale-free transcriptional regulatory networks. *Journal of molecular biology*. 2006; 360(1):204–212. doi: [10.1016/j.jmb.2006.04.026](https://doi.org/10.1016/j.jmb.2006.04.026) PMID: [16730024](https://pubmed.ncbi.nlm.nih.gov/16730024/)
76. Newman ME, Park J. Why social networks are different from other types of networks. *Physical Review E*. 2003; 68(3):036122. doi: [10.1103/PhysRevE.68.036122](https://doi.org/10.1103/PhysRevE.68.036122)
77. Redner S. Networks: teasing out the missing links. *Nature*. 2008; 453(7191):47–48. doi: [10.1038/453047a](https://doi.org/10.1038/453047a) PMID: [18451851](https://pubmed.ncbi.nlm.nih.gov/18451851/)
78. Aftabuddin M, Kundu S. Hydrophobic, hydrophilic, and charged amino acid networks within protein. *Biophysical journal*. 2007; 93(1):225–231. doi: [10.1529/biophysj.106.098004](https://doi.org/10.1529/biophysj.106.098004) PMID: [17172302](https://pubmed.ncbi.nlm.nih.gov/17172302/)
79. Mallik S, Akashi H, Kundu S. Assembly constraints drive co-evolution among ribosomal constituents. *Nucleic Acids Research*. 2015; 43(11):5352–5363. doi: [10.1093/nar/gkv448](https://doi.org/10.1093/nar/gkv448) PMID: [25956649](https://pubmed.ncbi.nlm.nih.gov/25956649/)