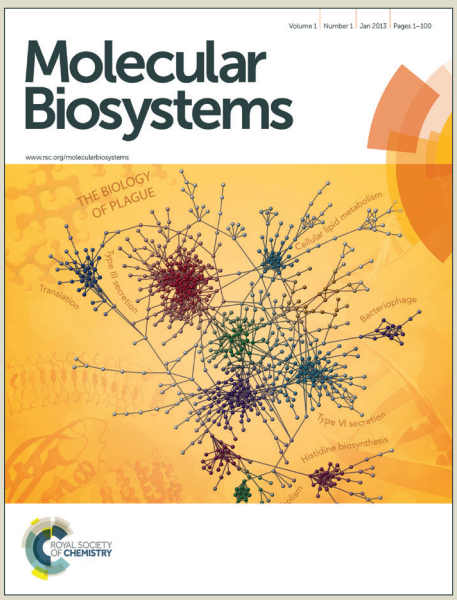
Molecular BioSystems

Accepted Manuscript



ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/molecularbiosystems

ARTICLE TYPE www. rsc.org/xxxxxx | XXXXXXXX

A network analysis of miRNA mediated gene regulation of rice: crosstalk among biological processes

Chittabrata Mal,[a] Arindam Deb,[a] Md. Aftabuddin[b] and Sudip Kundu*[ac]

Received Xth XXXXXXXXXX 20XX, Accepted Xth XXXXXXXXX 20XX
First published on the web Xth XXXXXXXXXX 200X
DOI: 10.1039/b000000x

To understand the network architecture of miRNA mediated regulations at the genomic and functional level of rice, we made an unambiguous annotation of the experimentally verified miRNAs, predicted their targets and possible biological functions they can affect. Some functions, namely translational and protein modifications, photosynthesis are targeted by higher percentage of miRNA. Using transformation procedures, we constructed genome scale miRNA-miRNA functional synergistic network (MFSN). The analysis of MFSN modules help to identify miRNAs co-regulating target genes having several interrelated biological processes. Some of these target genes are also co-expressed in particular condition. For example, the genes co-expressed in drought condition as well as targeted by miRNAs present in a MFSN module have interdependent biological processes namely, photosynthesis, cell-wall biogenesis, root development and xylan synthesis. The stress-induced miRNAs and their distributions; the presence of transcription factors in the target set of MFSN modules were also analyzed.

Introduction

More than half of the world's human population depends on rice as major staple food[1]. Predictably rice-biotechnologists aim to increase the production of rice by generating high yielding, stress tolerant rice cultivars. This calls for understanding the cellular physiology of rice, including the activities of the genes and their regulations, which occurs at the level of transcription, RNA processing, mRNA lifetime and translation etc. The miRNAs are a large family of about 21-22 nucleotide endogenous non-protein-coding regulatory RNA sequences that are the key players in post-transcriptional gene regulation[2]. miRNAs play critical role in cleavage, degradation or translational inhibition of their target mRNAs with a resultant repression of gene expression in animals, plants, and fungi[2–4]. Plant development, response to environmental stress, pathogen invasion and regulation of their own biogenesis etc. are known to be finely tuned by miRNAs[3].

The miRNAs are identified either by computational techniques[5,6] or by direct cloning of small RNAs[7]. The computationally predicted miRNAs and their targets have been widely used in different analyses for various species, like vertebrates[8], insects[9], Arabidopsis and rice[5,6]. Recently, different works have identified mature miRNAs at different developmental phases[10,11] or at different stress conditions[7,12] using high-throughput technologies. It is believed that many miRNA genes and their targets are still to be discovered; and such discoveries may help us to understand the critical role of miRNA mediated regulation in more details.

Genome-scale miRNA regulations have been studied mostly in case of human[13,14] and also in few plants[15]. Rice is one of the most important crops and the miRNA regulation has also been studied in this species. Example of these include miRNA mediated gene regulatory network in roots[16], vegetative and reproductive stage-specific regulation[17], the genome wide analysis of miRNA and their target genes in leaf at senescence[18], etc. While these studies have shed light on condition and tissue specific regulatory role of miRNAs; to the best of our knowledge, a comprehensive analysis of genome scale miRNA-mediated co-target and co-functional network (as studied in human[13]) in rice has not been done so far.

The present work has constructed a genome scale miRNA-miRNA co-targeting and co-functional network of rice and has subsequently identified some crucial modules where the participating nodes can co-target and co-regulate different but related cellular processes. To decipher the broader scheme, we have additionally analyzed miRNAs targeting transcription factors and stress induced miRNAs within the co-functional modules. Further, targets of the miRNAs of the co-functional modules which are co-expressed in drought, salt and both of the stresses are identified and the potential functions of the miRNAs have been assigned.

† Electronic Supplementary Information (ESI) available: [Supplementary files [1-6]. See DOI: 10.1039/b000000x/

[a] *Department of Biophysics, Molecular Biology & Bioinformatics, University of Calcutta, 92, A.P.C. Road, Kolkata 700009, India. *Corresponding Author. E-mail: skbmbg@caluniv.ac.in*
[b] *West Bengal University of Technology, BF-142, Salt Lake, Sector I, Kolkata 700064, India.*
[c] *Center of Excellence in Systems Biology and Biomedical Engineering (TEQIP Phase II), University of Calcutta, India.*

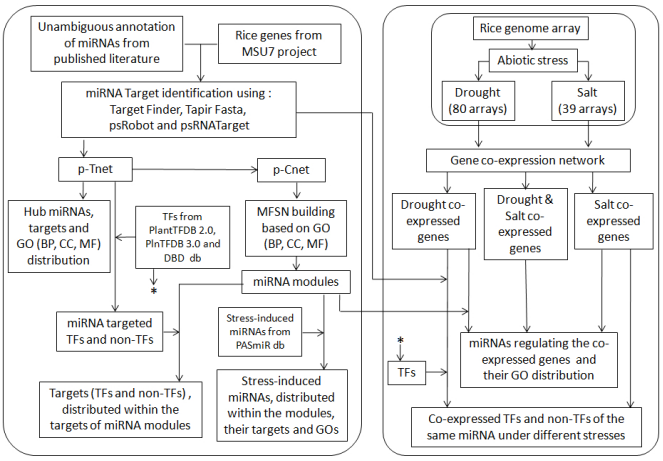


Fig. 1 Schematic of the present work in brief.

Materials and methods

Annotation of miRNAs

Mature miRNA sequences, expressed at different conditions, have been collected from literature (Additional file 8). The miRNA sequences have been identified by different experimental techniques including MPSS, pyrosequencing, microarray, deep sequencing, cloning etc. We observe that some of the miRNAs have identical sequences, but different identifiers have been used in different experiments. To identify the unique miRNAs, unambiguous annotation is needed. Such annotation is achieved by assigning new identifier (starts with "om") to 889 unique miRNA sequences. Fig. 1 represents the schematic of our whole work in brief.

miRNA - target genes, GO terms and transcription factors

Large number of tools are available to computationally predict the target genes of miRNAs. Srivastava and co-authors have recently compared their performances in plant miRNA target prediction[19]. They have concluded that Targetfinder alone or in combination with psRNATarget or Tapirhybrid for the miRNA target predictions can provide the most satisfactory results. Firstly, the target genes (TGs) of miRNAs are predicted using each of the four different tools: Target Finder 1.6, Tapir Fasta, psRobot 1.2, and psRNAtarget, with optimal cut-off values (6, 4, 6 and 4, respectively) as suggested by Srivastava et. al. Then, we have considered only the common target genes for our further analysis. Gene Ontology (GO) annotations (biological processes (BP), molecular function (MF) and cellular component (CC)) are downloaded from Biomart[20] (Release 24). Lists of transcription factors are taken from PlantTFDB 2.0, PlnTFDB 3.0 and DBD: Transcription factor prediction database (Release 2). Rice genes are collected from *Oryza sativa* MSU Rice Genome Annotation (Release 7).

Construction of post-transcriptional miRNA - miRNA co-targeting network (p-CNet)

miRNA target data sets are preprocessed as described in the previous section. The p-CNet is constructed following the method of Balaji et. al.[21]. In the p-Cnet, initially an edge is given between a pair of miRNAs if they target a common gene (Fig. 2a); thus 56845 pair-wise interactions among 706 miRNAs have been identified. The edge-weight is denoted by the number of TGs co-targeted by the pair. Then this value is normalized by the ratio of the observed number of shared TGs by a pair of miRNA to the expected number of genes shared by the same pair. The normalized value is termed as co-targeting coefficient (CC) value for each pair. The expected number is calculated by taking the average of shared TGs for a miRNA pair in 60,000 randomly generated post-transcriptional networks where degree distribution remains similar to that of the original network. We have filtered only those miRNAs which have $CC > 1$ (i.e. observed co-targeting association is higher than the random one). Further, considering O_{min} (minimum number of co-targeted TGs)= 3, a p-Cnet with 677 edges between 232 miRNAs is obtained (Fig. 2b). The same threshold of CC value has been used by Balaji et. al.[21].

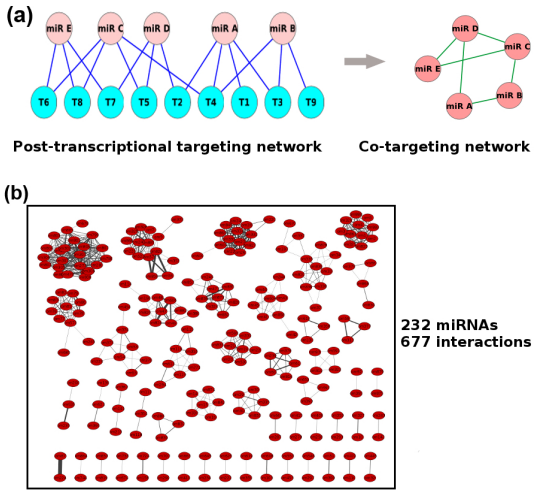


Fig. 2 Post-transcriptional miRNA - miRNA co-targeting network (p-CNet). (a) Procedure to determine co-targeting network. (b) Sparse co-targeting network in rice.

In the p-Cnet, the weight of the edge between miRNAs is a measure of the extent of co-targeting association between pairs of miRNAs over what is expected by chance. Bias aris-

ing due to chance sharing of TGs, especially seen in high-throughput datasets, is normalized by this procedure[21]. We have used this p-Cnet consisting of significant co-targeting interactions of miRNAs in rice for further genome-scale analysis.

Construction of miRNA - miRNA functional synergistic network (MFSN)

For a given miRNA pair (M1 and M2), the co-targeting set (M1 ∩ M2) has been identified. We have taken only those set of miRNA pairs which contain at least three common target genes[13] and also have CC-values > 1. These common target genes have been used for further analysis. By functional enrichment analysis (for different gene ontology domains e.g., biological processes (BP), molecular function (MF) and cellular component (CC) separately) the processes (GO categories) where this target set is enriched are identified. Considering the hypergeometric distribution[13,22], the probability PG_i for M1 ∩ M2 in the GO term '*i*' is calculated according to the following relation,

$$PG_i = 1 - F(n \mid S, T_i, L) = 1 - \sum_{t=0}^{n} \frac{\binom{T_i}{t}\binom{S-T_i}{L-t}}{\binom{S}{L}}$$

where S is the number of all targets (default background distribution), T_i is the total number of genes that are annotated in the GO term '*i*' and targeted by miRNAs, L is the size of M1 ∩ M2, n is the number of targets in M1 ∩ M2 that are also annotated to term '*i*'. Here, $i = 1,2,.....,I$ and I is the total number of GO terms we considered. If $PG_i < 0.05$ and FDR (Benjamini Hochberg)<0.05, we consider that the GO term '*i*' is functionally enriched by the targets of the miRNA pair under study. If at least one GO category is significantly co-regulated by a pair of miRNAs, we define them as synergistic. After assembling all significant miRNA pairs identified above, we have constructed the miRNA - miRNA functional synergistic network (MFSN). A node represents a miRNA, and two nodes are connected if the corresponding miRNA pair has a synergistic action, otherwise no edge, is deemed to exist.

Module identification

A module in the co-functional network has been defined as a *k*-clique, i.e. a highly-dense sub-graph with '*k*' number of miRNAs where all miRNAs having co-functional association with other miRNAs in the sub-graph. Modules are identified using Cytoscape[23] AllegroMCODE plug-in with *k*-core value of 2. Each module has a unique composition of miRNAs and the same miRNA or the same pair do not occur in more than one module.

Stress induced coexpressed miRNA targets

Co-expression network from microarray data We have collected drought and salt stressed microarray data (Additional file 9) from Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo/). We have used No3CoGP tool[24] (GCRMA as data normalization method) to identify co-expressed gene sets in salt, drought and in both the stresses.

Identification of coexpressed genes regulated by miRNAs in MFSN modules The co-expressed genes which are present in drought, salt and both of the stresses have been identified using the No3CoGP tool. Next we extracted genes which are regulated by the miRNAs from MFSN modules and which functions those miRNAs regulate. Further, the coexpressed TFs and non-TFs targeted by the same miRNA under different stresses have been identified.

Results and Discussions

General Properties of *Oryza sativa* genome-scale post-Transcriptional regulatory network

Considering the miRNAs and their target genes (TGs) as nodes; and interactions of miRNAs with their target genes as edges; we have constructed a genome-scale post-transcriptional targeting network (p-Tnet) of *Oryza sativa*. The network has 6230 nodes (706 miRNAs and 5524 TGs) and 7928 edges. Each miRNA has an average of 11 TGs, while each TG is targeted by an average of ∼1.43 miRNAs. A list of unique miRNA sequences with the new and old identifiers is presented in Additional file 1. The distributions of miRNAs and TGs show that while a large number of miRNAs (70%) target lesser number of TG (i.e., below average, 1 to 11), a few miRNAs (30%) target higher number of TG (i.e., above average, 12 to 100).

Transcription factors are targeted by higher number of miRNAs

1566 transcription factors (TFs) have been taken from PlantTFDB 2.0, PlnTFDB 3.0 and DBD: Transcription factor prediction database (Release 2). Dataset in *Oryza sativa* MSU Rice Genome Annotation release 7 contains 66,153 genes. The results show that 117 TFs and 5407 non-TFs are targeted by miRNAs through 459 and 7469 interactions, respectively. One TF, on average, is targeted by ~ 4 (more precisely 3.94) miRNAs and this number is higher than the average number (1.38) of miRNAs targeting one non-TFs. This result is similar to the previous study (but with different dataset) by Archak and Nagaraju[25].

miRNAs acting as mater regulators

It is always interesting to find the hubs of a network as they play significant roles in the network architecture. A hub miRNA can target and thus regulate a large number of TGs. We define the top 10% of miRNAs in the p-Tnet with high out-going connectivities (i.e., >(average + standard deviation) of the entire population) as hubs. The predicted hubs have >26 TGs. One sample t-test ($p<0.0001$) showed that the degrees >26 are significantly higher than the rest of the population. Interestingly, these 10% of total miRNAs, acting as hubs, have more than half of the interactions (edges) of the network. Some of these hubs are om510 (miR3979-5p), om255 (osa-miR5490), om197 (miR5819) etc. and each of which targets >70 TGs. These hub miRNAs may have the potential to act as master regulators and their list is presented in Additional file 2.

The average incoming connectivity (>1) of TGs and the degree distribution of p-Tnet clearly indicate that like transcriptional regulatory network of genes[21]; this post-transcriptional regulatory network of genes mediated by miRNAs also has co-targeting regulatory architecture.

Characteristics of Co-targeting association

To understand the co-targeting association we have constructed the post- transcriptional co-targeting network (p-Cnet) using a network transformation procedure described in Methods. Here, only those pairs having coregulation coefficient (CC) (for definition, see methods) value > 1 have been considered. The resultant network has 232 miRNA nodes with 677 edges among them (miRNA-miRNA connectivities). Each of the miRNA pairs in p-Cnet can target more than one target gene. 677 miRNA pairs of p-Cnet can target 5524 TGs through 5266 interactions. Thus, each miRNA-pair targets an average of ∼8 (to be precise, 7.7) TGs. However, the network is sparse (i.e. nodes of the network are not fully connected) in nature (Fig. 2b) and it has many clusters of miRNAs.

miRNA functional synergistic network and modules

For each miRNA pair, we have identified their co-targeting genes as a target subset and then, have identified candidate functional GO category by performing functional enrichment analysis (described in method section). We have performed functional enrichment for different gene ontology domains e.g., biological processes (BP), molecular function (MF) and cellular component (CC) separately. However, here we have described results of the biological processes only. Assembling all synergistic miRNA pairs (3518 pairwise interactions among 204 miRNAs) we have constructed miRNA functional synergistic network (MFSN) (Fig. 3).

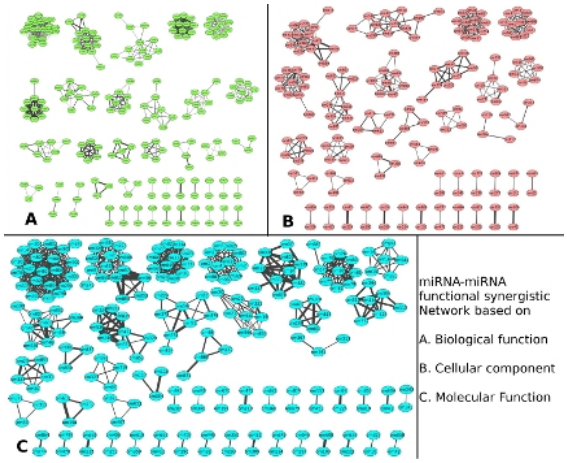


Fig. 3 MFSN networks based on different GO categories namely BP (biological process), CC (cellular component) and MF (molecular function)

Next, we have analyzed the modular structure of the miRNA functional synergistic network (MFSN) and identified 19 modules (see methods). The modules having higher number of miRNAs are less in number. The largest module consists of 20 miRNAs and the smallest one with only 3 miRNAs. The significant miRNA pairs, their targeted GO biological processes and respective TGs in each module are presented in Additional file 3. Next, we have analyzed biological functions of each module and have identified that some modules are involved in multiple biological processes (e.g., Module 5 in 57 different processes like glucuronoxylan metabolic process (GO:0010413), phosphorylation (GO:0016310), callose deposition in cell wall (GO:0052543), cell wall macromolecule metabolic process (GO:0044036), regulation of photosynthesis, light reaction (GO:0042548) etc.), while other modules are involved in less number of biological process (e.g., Module 17 in 3 processes like regulation of transcription, DNA-templated (GO:0006355), transcription, DNA-templated (GO:0006351), and root development(GO:0048364).

We have also analyzed the over-representation of any biological function within the modules. The biological function "regulation of transcription, DNA-dependent (GO:0006355)" is present in 12 different modules (Module 2 to 12 and 19); protein phosphorylation (GO:0006468) is present in 10 different modules (Module 1,2, 3, 5, 7, 8, 12, 13, 14 and 17); "transport (GO:0006810)" is present in 4 different modules (Module 7, 8, 11, 12) etc.

Next, we have analysed how the transcription factors are distributed within the targets of miRNA cofunctional modules. As expected, the TFs of same family are mostly present in one module. For example,

members of TF family AP2-EREBP (LOC_Os07g13170.1, LOC_Os04g55560.2 etc.) are co-targeted by miRNAs of Module 7; and miR families 172, 18, 84 are present in Module 7. Aukerman and Sakai postulated that downregulation of AP2-like target genes by miR172 resulting in the promotion of flowering[26,27]. MADS transcription factor family genes (LOC_Os02g36924.1, LOC_Os02g49840.1 etc.) are the potential targets of miRNA family 444 present in Module 5. It has been already reported that MADS genes are collectively regulated by miR444[26]. ARF family TFs (LOC_Os02g06910.1, LOC_Os04g57610.1, LOC_Os06g46410.1) are present in Module 11. Thus, we have identified the set of miRNAs regulating the co-functional transcription factors. Moreover, sometimes a combination of different miRNA families can co-regulate cofunctional TFs (e.g., HB family TFs are targeted by miRNAs of Module 2 and 5). The TFs, the miRNA pairs targeting them and the GO biological processes within each module are presented in Additional file 4.

From the database PASmiR[28], we have collected miRNAs induced in different stresses. Out of 87 miRNA families present in PASmiR, there are only 47 families which have targets in our dataset. In the MFSN modules, we have found that stress induced miRNAs are distributed in 14 modules (out of 19) and the details are present in Additional File 5.

For example, miR169 is present in Module 8. The experimental evidence[28] suggests that this miR169 is induced differently under varying conditions. For example, it is up-regulated in excess aluminium, drought, ozone, high salt, phosphorus deficiency and down-regulated in abscisic acid, low temperature and nitrogen deficiency. We found that protein phosphorylation (GO:0006468), translation (GO:0006412), transcription, DNA-templated (GO:0006351), ER to Golgi vesicle-mediated transport (GO:0006888) etc. are regulated by miRNAs of Module 8. It is expected that the varying expressions of miRNA169 under different stresses might change the amount of its targeted gene due to post-transcriptional regulation. Thus, the miRNA pairs induced during abiotic stresses have enormous potential to regulate different biological processes.

Biological functions of miRNA targets at different levels

p-Tnet The analysis of p-Tnet helps us to identify hub miRNAs (e.g., om510 (miR3979-5p, oryza_test-m0135-5p), om255 (osa-miR5490, miR11, oryza_test-m0141-3p, oryza _test-m0142-3p), om197 (miR5819), om141 (miR2055.2')) which can control the post-transcriptional regulations of a large number of genes.

As expected, these hubs that act as master regulators in cell, are few in numbers. Further, the GO (BP, CC, MF) functional classification of the genes targeted by the miRNAs indicate

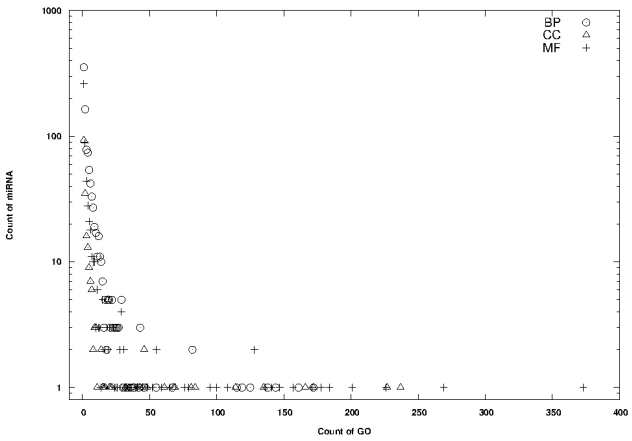


Fig. 4 Distributions of miRNA and their regulating different GO categories namely BP (biological process), CC (cellular component) and MF (molecular function).

that miRNA can regulate different important functions of a cell. The distributions of miRNA and their regulating GO categories also show that while a large number of miRNA regulate a small number of biological functions, a few miRNAs regulate a large number of biological functions (Fig. 4). TGs of 319 (out of 889) miRNAs could not be classified in the known GO categories. In particular, 31.8% miRNAs regulate more than 11 (above average) biological processes. For example, om765, om776, om141, om347 are involved in 71, 69, 57, 57 biological processes, respectively. The GO distribution of 570 miRNAs are described in Additional file 6.

We have also observed that there are few TGs which can be post-transcriptionally controlled by a large number of miRNAs. For example, LOC_Os01g20720.1, LOC_Os10g26820.1, LOC_Os05g34220.1 are potential targets of 21, 19 and 16 miRNAs respectively. Among them, it is already reported that the defense responsive NBS-LRR gene family are targeted by miNRAs[29,30]. This indicate several possibilities: (i) there may be functional redundancies among miRNAs targeting a gene, (ii) tissue (condition) specific expression of different miRNAs can regulate the same gene in different tissues (conditions), (iii) there may exist combinatorial regulation of a gene targeted by different miRNAs. The regulation by miR164 family in Arabidopsis[31] is one of such example of functional redundancies. The authors experimentally demonstrated that the disruption of shoot development by a single gene mutation vary drastically from the loss of entire miR164 family. Next example is for tissue specific expression of the genes of a miRNA family. Yanjie et. al. showed that while miR169a expressed abundantly in cotyledon, hypocotyl, stoma and root vascular tissues; miR169c uniquely expressed

in the shoot apex, root tip and the lateral root primordia[32]. Analyzing the expression of miR169, they further suggested for a cooperative regulatory mechanism of miRNA's action.

Next we observe that some groups of biological processes are targeted by higher percentage of miRNAs. For example, 18% miRNAs are involved in translational and protein modification, 16.3% miRNAs are involved in photosynthesis. Some groups of biological processes are targeted by lower percentage of miRNAs. For example, 6% miRNAs are involved in cellular growth, 7% miRNAs are involved in homeostasis, 8.5% miRNAs are regulating stress and signaling etc. A pie-chart of percentage of miRNAs regulating different groups of biological processes is described in the Fig. 5.

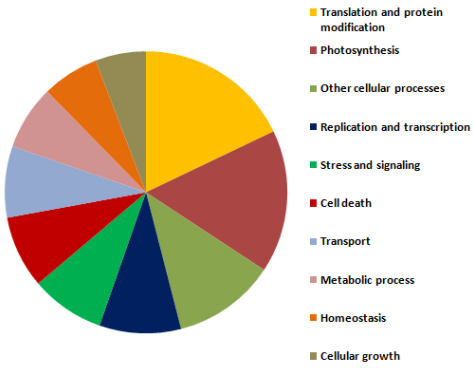


Fig. 5 Pie-chart of miRNAs regulating different groups of biological processes.

p-Cnet In general, the higher the CC-value of a miRNA pair indicates the higher association of the target genes of that miRNA pair. We have identified that many miRNA pairs have a relatively low CC value, but a small number of miRNA pairs have very high co-targeting coefficients.

Interestingly, we observe that co-targeting network is enable to identify the miRNA pairs regulating related biological processes. For example, at CC value $>$ 10.0 and $<$ 30.0, om334:om332, om869:om26 are regulating determination of bilateral symmetry (GO:0009855), embryonic pattern specification (GO:0009880), adaxial/abaxial axis specification (GO:0009943), polarity specification of adaxial/abaxial axis (GO:0009944), adaxial/abaxial pattern specification (GO:0009955) etc. All these functions related to growth and development[33,34]. At CC value $>$ 5.0 and $<$ 10.0, om54:om374, om74:0m375, om53:om374 are regulating sulfate assimilation (GO:0000103), carbohydrate metabolic process (GO:0005975), protein phosphorylation (GO:0006468), metabolic process (GO:0008152), sulfate transport (GO:0008272) etc. It is expected as well as reported that sulfate transport has role in plant metabolic process[35] and is also linked with phosphorylation/dephosphorylation process[36]. It may be hypothesized that these processes might be co-regulated by these miRNA pairs.

miRNA target gene coexpression analysis identifies sets of miRNAs associated with related functions

It has been hypothesised that TGs of miRNAs are likely to be coexpressed[37,38]. Using No3CoGP tool[24] we compared the drought and salt coexpressed genes and identified the co-expressed genes in drought, salt and both of the stresses. From the MFSN modules we already obtained the co-targeted co-functional set of miRNAs. Next, we identified the gene sets co-targeted by miRNAs of a module as well as co-expressed in a particular stress condition. Some evidences in support of this are given below.

Drought co-expressed genes and MFSN Module 5: We identified 7 TGs (LOC_Os02g36924, LOC_Os04g23910, LOC_Os04g51350, LOC _Os05g47560, LOC_Os06g23980, LOC_Os07g04690, LOC_Os08g33488) co-expressed during drought stress, among the TGs targeted by miRNAs of Module 5. The miRNA pairs om80:om571, om76:om571, om78:om77, om79:om571, om80:om77, om82:om80 etc. co-target MADS TF family (LOC_Os02g36924), which is involved in callose deposition in cell wall (GO:0052543), cell wall macromolecule metabolic process (GO:0044036), lateral root development (GO:0048527), xylan biosynthetic process (GO:0045492), cell wall biogenesis (GO:0042546), etc. On the other hand, the miRNA pairs om79:om78, om80:om78 and om80:om79 co-target serine-threonine kinase protein (LOC_Os05g47560), having a role in regulation of photosynthesis (GO:0042548), circadian rhythm (GO:0007623), phosphorylation (GO:0016310) etc. Interestingly, all these miRNA pairs belong to Module 5 (Fig. 6) of MFSN network and literature survey suggests that the functions of their targets are interlinked with one another. It is already well-established that the MADS box promotes the flowering by enhancing chloroplast development and photosynthesis[39,40]. The photosynthesis is tightly coupled with cell wall biogenesis[41,42], root development[43], nutrient response[44] and xylan biosynthetic process[45] in the developing plant. Again root development depends on cell wall biogenesis[46] which has a link to xylan synthesis[47,48]. On the other hand, Stomatal regulation is connected to nutrient response[49] and photosynthesis[50].

Salt co-expressed genes and MFSN Module 13: Among the co-targeted genes of miRNAs of MFSN Module 13 (Fig. 7) we observed that 7 TGs (LOC_Os01g13190, LOC_Os01g44330, LOC_Os01g63180, LOC_Os01g63200, LOC_Os05g38390, LOC _Os09g26820, LOC_Os11g48060) are co-expressed during salt stress.

The miRNA pairs om816:om60, om378:om239, om60:om378, om816:om378, om816 :om239 co-target

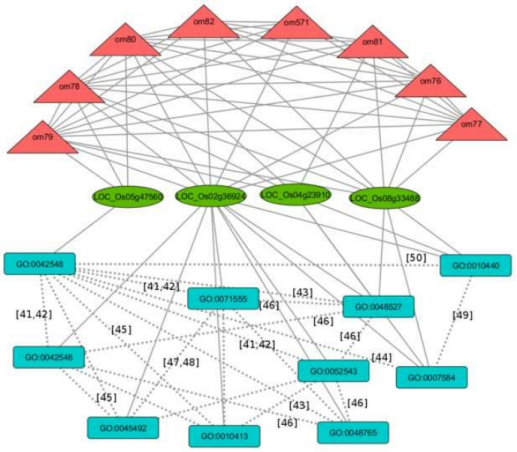


Fig. 6 MFSN Module 5: miRNAs (first layer), drought co-expressed target genes (second layer) and GO processes (third layer). Literature support (reference number) for the cross-talk between two GO biological processes are given in the [ ] beside the link line. TG description: MADS TF family (LOC Os02g36924), serine-threonine kinase protien (LOC Os05g47560), OsMADS25 - MADS-box family gene (LOC Os04g23910), OsMADS23 - MADS-box family gene (LOC Os08g33488); GO descriptions: callose deposition in cell wall (GO:0052543), lateral root development (GO:0048527), xylan biosynthetic process (GO:0045492), cell wall biogenesis (GO:0042546), regulation of photosynthesis (GO:0042548), phosphorylation (GO: 0016310), glucuronoxylan metabolic process (GO:0010413), cell wall organization (GO:0071555), stomatal lineage progression (GO:0010440), response to nutrient (GO:0007584), root hair cell differentiation (GO:0048765).

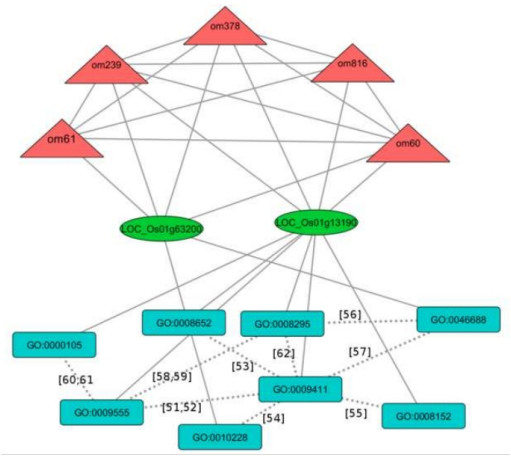


Fig. 7 MFSN Module 13: miRNAs (first layer), salt co-expressed target genes (second layer) and GO processes (third layer). Literature support (reference number) for the cross-talk between two GO biological processes are given in the [ ] beside the link line. TG description: laccase precursor protein (LOC Os01g63200), histidinol dehydrogenase, chloroplast precursor (LOC Os01g13190); GO description: cellular amino acid biosynthetic process (GO:0008652), histidine biosynthetic process (GO:0000105), pollen development (GO:0009555), response to UV (GO:0009411), spermidine biosynthetic process (GO:0008295), response to copper ion (GO:0046688), vegetative to reproductive phase transition of meristem (GO:0010228), metabolic process (GO:0008152).

LOC_Os01g13190 (histidinol dehydrogenase, chloroplast precursor) which is involved in cellular amino acid biosynthetic process (GO:0008652), histidine biosynthetic process (GO:0000105), pollen development (GO:0009555), response to UV (GO:0009411), spermidine biosynthetic process (GO:0008295). The miRNA pairs om61:om60, om61 :om239, om60:om378 co-target LOC_Os05 g38390, LOC_Os11g48060 (laccase precursor protein) which are involved in lignin catabolic process (GO:0046274), response to copper ion (GO:0046688), vegetative to reproductive phase transition of meristem (GO:0010228). UV stress plays role in pollen development[51,52], amino acid biosynthesis[53], vegetative growth[54] and metabolism[55]. Spermidine biosynthesis is interlinked with UV, water or copper stress[56]. It is reported that UV and copper shows similar stress response[57]. Pollen development is dependent on spermidine biosynthesis[58,59], and histidine[60,61]. UV stress also regulates spermidine biosynthesis[62]. The above examples support that the module of the MFSN network can cluster the miRNAs whose targets can act together to perform interconnected biological functions within the cell.

Functional classification of miRNA regulated condition specific TFs and non-TFs

By investigating the transcriptional regulation of microRNA target genes, Wang et al. showed that different groups of target genes of the same miRNA are co-expressed under different conditions[63]. Here, we have identified 462 and 464 miRNAs which have targets in the drought and salt co-expressed gene set, respectively; comparing the co-expression of genes in the drought stress with the salt stress. Among these miRNAs, 80 and 84 miRNAs have both the coexpressed TFs and non-TFs in the drought and salt stress respectively (Additional file 7). For example, om82 (miR444) has 3 TFs and 7 non-TFs which are coexpressed in both the drought and salt stress. Further, we have classified the miRNAs which have co-expressed targets in only drought, only salt and in both of the stresses.

Future direction

Our analysis on co-functional modules identified sets of miRNAs having the potential to post-transcriptionally co-regulate known inter-related cellular activities. Moreover, based on

the co-functional modular analysis (Additional File 3), one can generate hypothesis to test the possible dependencies of several cellular processes, which can be verified by doing related experiments. Further, we want to mention that while the cross-talk among a subset of GO biological processes within a module are supported by existing literature; one can exploit the potential relationships among other biological processes present within the same co-functional module but we could not find any evidence of their direct relationship through literature search.

We are aware that there are several limitations in computation based miRNA research. The problems lie in (i) false target prediction and (ii) though there are many genes having putative binding sites for multiple miRNAs[64,65], some putative sites may be non-functional in vivo, etc. Experimental verification of any hypothesis is needed. However, it is not very easy to perform the experiment to study the whole genome level miRNA gene regulatory network. Recently, it has been demonstrated that the manipulation of miRNAs can change the miRNA mediated gene regulatory network and the authors speculate that this may be used to design the desired effect in Arabidopsis[66]. Thus, our findings might help the researchers in designing efficient rice cultivars.

Conclusion

In conclusion, ambiguities regarding the repeated annotations of some miRNA sequences have been removed from the available experimental data in the present work. p-Tnet analysis identifies a number of master miRNA regulators (hubs) which can target large number of genes. The functional processes which may be regulated by large set of miRNAs have also been reported here. Potential roles of miRNAs (whose target genes are involved in different gene ontology domains e.g., biological processes, molecular function and cellular component) along with the interacted miRNAs can be detected by means of MFSN construction. We also identified different groups of target genes (TFs and non-TFs) of the same miRNA which are co-expressed under different stresses. The further study on the miRNA functional modules and TFs connection will provide a new approach towards the mechanism study of the alterations of miRNAs in different stresses.

Acknowledgements

CM acknowledges UGC (RFSMS) and DST-PURSE, Government of India and AD acknowledges UGC (RFSMS), Government of India for financial support. SK acknowledges Center of Excellence in Systems Biology and Biomedical Engineering (TEQIP Phase II), University of Calcutta, India for partial support. All authors acknowledge DIC, Calcutta University for computational facilities.

Author's contributions

CM and SK conceived of the study and prepared the manuscript. MA, AD and CM performed the analysis using in-house PERL code. All authors read and approved the final manuscript.

References

1 S. S. Gnanamanickam, *Biological Control of Rice Diseases*, Springer, 2009, pp. 1–11.
2 B. Bartel and D. P. Bartel, *Plant Physiol*, 2003, **132**, 709–17.
3 D. P. Bartel, *Cell*, 2004, **116**, 281–97.
4 J. C. Carrington and V. Ambros, *Science*, 2003, **301**, 336–8.
5 E. Bonnet, J. Wuyts, P. Rouze and Y. Van de Peer, *Proc Natl Acad Sci USA*, 2004, **101**, 11511–6.
6 M. W. Jones-Rhoades and D. P. Bartel, *Mol Cell*, 2004, **14**, 787–99.
7 R. Sunkar and J. K. Zhu, *Plant Cell*, 2004, **16**, 2001–19.
8 B. P. Lewis, I. H. Shih, M. W. Jones-Rhoades, D. P. Bartel and C. B. Burge, *Cell*, 2003, **115**, 787–8.
9 N. Rajewsky and N. D. Socci, *Dev Biol*, 2004, **267**, 529–35.
10 Q. H. Zhu, A. Spriggs, L. Matthew *et al.*, *Genome Res.*, 2008, **18**, 1456–1465.
11 L. J. Xue, J. J. Zhang and H. W. Xue, *Nucl. Acids Res.*, 2009, **37**, 916–930.
12 M. Gu, K. Xu, A. Chen, Y. Zhu, G. Tang and G. Xu, *Physiol Plant*, 2010, **138**, 226–237.
13 J. Xu, *Nucl. Acids Res.*, 2011, **39**, 825–836.
14 Y. Friedman, O. Balaga and M. Linial, *Adv Exp Med Biol.*, 2013, **774**, 317–37.
15 J. Je-Gun and F. Zhangjun, *Bioinformatics*, 2009, **25**, 387–393.
16 X. Ma, C. Shao, H. Wang, Y. Jin and Y. Meng, *BMC genomics*, 2013, **14**, 510.
17 Y. Meng, C. Shao, H. Wang, X. Ma and M. Chen, *New Phytologist*, 2013, **197**, 441–453.
18 X. Xu, H. Bai, C. Liu, E. Chen, Q. Chen, J. Zhuang and B. Shen, *PloS ONE*, 2014, **9**, e114313.
19 P. K. Srivastava, T. R. Moturu, P. Pandey, I. T. Baldwin and S. P. Pandey, *BMC Genomics*, 2014, **15**, 348.
20 D. Smedley, S. Haider, B. Ballester, R. Holland, D. London, G. Thorisson and A. Kasprzyk, *BMC Genomics*, 2009, **10**, 22.
21 S. Balaji, B. M. Madan, L. M. Iyer, N. M. Luscombe and L. Aravind, *Journal of Molecular Biology*, 2006, **360**, 213–227.
22 M. Aftabuddin, C. Mal, A. Deb and S. Kundu, *Genomics, proteomics & bioinformatics*, 2014, **12**, 133–136.
23 P. Shannon, A. Markiel, O. Ozier, N. S. Baliga, J. T. Wang, D. Ramage, N. Amin, B. Schwikowski and T. Ideker, *Genome Res.*, 2003, **13**, 2498–504.
24 C. Mal, M. Aftabudddin and S. Kundu, *BMC research notes*, 2014, **7**, 886.
25 S. Archak and J. Nagaraju, *Genomics Proteomics Bioinformatics*, 2007, **5**, 196–206.
26 Y. Li, Y. Zheng, C. Addo-Quaye, L. Zhang, A. Saini, G. Jagadeeswaran, M. J. Axtell, W. Zhang and R. Sunkar, *The Plant Journal*, 2010, **62**, 742–759.
27 M. J. Aukerman and H. Sakai, *Plant Cell*, 2003, **15**, 2730–2741.
28 S. Zhang, Y. Yue, L. Sheng, Y. Wu, G. Fan, A. Li, X. Hu, M. Shangguan and C. Wei, *BMC Plant Biol.*, 2013, **13**, 33.
29 V. M. González, S. Müller, D. Baulcombe and P. Puigdomènech, *Molecular Plant*, 2014.
30 P. V. Shivaprasad, H.-M. Chen, K. Patel, D. M. Bond, B. A. Santos and D. C. Baulcombe, *The Plant Cell Online*, 2012, **24**, 859–874.

31 P. Sieber, F. Wellmer, J. Gheyselinck, J. L. Riechmann and E. M. Meyerowitz, *Development*, 2007, **134**, 1051–1060.

32 Y. Li, Y. Fu, L. Ji, C. Wu and C. Zheng, *Plant science*, 2010, **178**, 271–280.

33 T. Sachs, *Positional controls in plant development/edited by PW Barlow and DJ Carr*, 1984.

34 S. Medvedev, *Russian Journal of Plant Physiology*, 2012, **59**, 502–514.

35 H. Takahashi, *International review of cell and molecular biology*, 2010, **281**, 129–159.

36 A. Maruyama-Nakashita, Y. Nakamura, A. Watanabe-Takahashi, T. Yamaya and H. Takahashi, *Plant and cell physiology*, 2004, **45**, 340–345.

37 V. A. Gennarino, G. D'Angelo, G. Dharmalingam, S. Fernandez, G. Russolillo, R. Sanges, M. Mutarelli, V. Belcastro, A. Ballabio, P. Verde *et al.*, *Genome research*, 2012, **22**, 1163–1172.

38 C. Cheng and L. M. Li, *PloS ONE*, 2008, **3**, e1989.

39 G.-Z. Qu, T. Zheng, G. Liu, W. Wang, L. Zang, H. Liu and C. Yang, *PloS ONE*, 2013, **8**, e63398.

40 C. Liu, J. Zhang, N. Zhang, H. Shan, K. Su, J. Zhang, Z. Meng, H. Kong and Z. Chen, *Molecular biology and evolution*, 2010, **27**, 1598–1611.

41 V. V. Lozovaya, O. A. Zabotina and J. M. Widholm, *Plant physiology*, 1996, **111**, 921–929.

42 M. Torres, F. X. Niell and F. L. Figueroa, *Journal of applied phycology*, 1995, **7**, 167–174.

43 H. E. Flores, Y. Dai, J. L. Cuello, I. E. Maldonado-Mendoza and V. M. Loyola-Vargas, *Plant physiology*, 1993, **101**, 363–371.

44 D. J. Longstreth and P. S. Nobel, *Plant Physiology*, 1980, **65**, 541–543.

45 R. Vanholme, V. Storme, B. Vanholme, L. Sundin, J. H. Christensen, G. Goeminne, C. Halpin, A. Rohde, K. Morreel and W. Boerjan, *The Plant Cell Online*, 2012, **24**, 3506–3529.

46 E. Nielsen, *Plant Cell Monographs*, 2009, **12**, 85–102.

47 A. Oikawa, H. J. Joshi, E. A. Rennie, B. Ebert, C. Manisseri, J. L. Heazlewood and H. V. Scheller, *PLoS ONE*, 2010, **5**, e15481.

48 M. Bosch, C.-D. Mayer, A. Cookson and I. S. Donnison, *Journal of experimental botany*, 2011, **62**, 3545–3561.

49 J. A. Nadeau and F. D. Sack, *The Arabidopsis book/American Society of Plant Biologists*, 2002, **1**, year.

50 K. Kusumi, S. Hirotsuka, T. Kumamaru and K. Iba, *Journal of experimental botany*, 2012, ers216.

51 H. Feng, L. An, L. Tan, Z. Hou and X. Wang, *The journal of applied ecology*, 2002, **13**, 814–818.

52 C. Zhang, Y.-P. Yang and Y.-W. Duan, *Scientific reports*, 2014, **4**, year.

53 C. M. Sorrels, P. J. Proteau and W. H. Gerwick, *Applied and environmental microbiology*, 2009, **75**, 4861–4869.

54 E. A. Kravets, L. B. Zelena, E. P. Zabara and Y. B. Blume, *Emirates Journal of Food and Agriculture*, 2012, **24**, year.

55 M. Kusano, T. Tohge, A. Fukushima, M. Kobayashi, N. Hayashi, H. Otsuki, Y. Kondou, H. Goto, M. Kawashima, F. Matsuda *et al.*, *The Plant Journal*, 2011, **67**, 354–369.

56 Q. Fariduddin, P. Varshney, M. Yusuf and A. Ahmad, *Journal of Plant Interactions*, 2013, **8**, 1–16.

57 T. S. Babu, T. A. Akhtar, M. A. Lampi, S. Tripuranthakam, D. G. Dixon and B. M. Greenberg, *Plant and Cell Physiology*, 2003, **44**, 1320–1329.

58 T. D. Quilichini, A. L. Samuels and C. J. Douglas, *The Plant Cell Online*, 2014, **26**, 4483–4498.

59 G. Falasca, M. Franceschetti, N. Bagni, M. Altamura and R. Biasi, *Plant Physiology and Biochemistry*, 2010, **48**, 565–573.

60 L. N. Petersen, S. Marineo, S. Mandalà, F. Davids, B. T. Sewell and R. A. Ingle, *Plant physiology*, 2010, **152**, 1186–1196.

61 R. Muralla, C. Sweeney, A. Stepansky, T. Leustek and D. Meinke, *Plant physiology*, 2007, **144**, 890–903.

62 D. Todorova, Z. Katerova, E. Shopova, A. Nikolova, N. Georgieva, I. Sergiev and S. Mapelli, *Biotechnology & Biotechnological Equipment*, 2013, **27**, 3798–3802.

63 Y. Wang, X. Li and H. Hu, *Genomics*, 2011, **98**, 445–452.

64 A. Krek, D. Grn, M. N. Poy, R. Wolf, L. Rosenberg, E. J. Epstein, P. MacMenamin, I. Piedade, K. C. Gunsalus, M. Stoffel and N. Rajewsky, *Nat. Genet.*, 2005, **37**, 495–500.

65 J. Krtzfeldt, N. Rajewsky, R. Braich, K. G. Rajeev, T. Tuschl, M. Manoharan and M. Stoffel, *Nature*, 2005, **438**, 685–689.

66 S. Ivashuta, I. R. Banks, B. E. Wiggins, Y. Zhang, T. E. Ziegler *et al.*, *PLoS ONE*, 2011, **6**, e21330.