

# Regulatory Circuits of MicroRNAs and Transcription Factors in Plant

Xuefeng Zhou<sup>1</sup>, Weixiong Zhang<sup>1,2,\*</sup>

1. Department of Computer Science, Washington University, Saint Louis, USA

2. Department of Genetics, Washington University, Saint Louis, USA

\*E-mail: {xzhou,zhang}@cse.wustl.edu

## Motivation

The functions of miRNAs in the context of gene networks are important, but remain mostly undetermined. In this paper, we studied the cooperation of miRNAs and transcription factors for gene regulation in Arabidopsis.

## Significant motifs characteristic of development and stress response

We selected differentially expressed genes induced by cold, drought and salinity in the root and shoot of Arabidopsis seedlings, and up-regulated genes at different developmental stages with a method called Rank Products [2] (Figure 1). The microarray data were from AtGenExpress [4]. Using our genome-wide WordSpy motif algorithm [7, 8], we discovered many significant motifs. The expression relevance of a motif was measured by a  $p$ -value from a cumulative hypergeometric test [1], and by the average pairwise cosine similarity (gscore) of genes containing the motif. Figure 1E shows the distribution of the motifs from the cold-induced genes in shoot, and 258 motifs obtained with the  $p$ -value cutoff of 0.01 and gscore cutoff of 0.1. Similarly, significant motifs were obtained for the drought and salt stresses, and different development stages.

## miRNA target prediction

We developed a plant miRNA target prediction algorithm, which improves upon three existing methods [3, 6, 9]. The algorithm considered all possible arrangements of mismatches and bulges, and limited the maximal number of mismatches, G-U wobbles and bulges along the alignments of miRNAs and their target sites. Signal-noise ratios were tested by random permutations of miRNAs with zero-order HMMs. With this method, we predicted targets for miRNAs, including miR777, miR779, miR829 and miR830, which had no targets identified before.

## Coordinated gene regulation by miRNAs and TFs and their regulatory circuits

One way to identifying cooperation between a miRNA and a TF is to find statistically significant co-occurrences of TF binding sites and miRNA target sites. To test the significance of a co-occurrence, we computed two statistical scores for each motif-miRNA pair. One is a hypergeometric  $p$ -value [1] based on the number of genes containing the motif, the number of target genes of the miRNA, and the number of miRNA's target genes containing the motif. Another is probability of random observing the co-occurrence based on the local nucleotide compositions of the promoter and the cDNA.

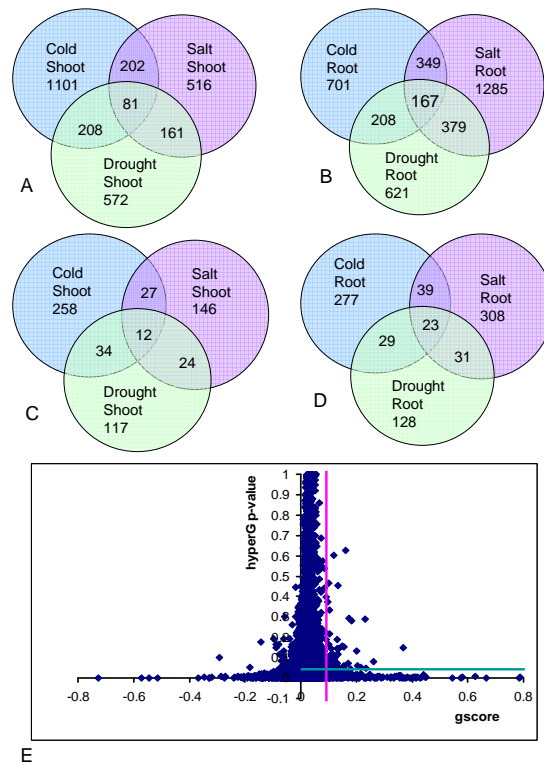


Figure 1: A,B: Number of cold, drought and salt up-regulated genes in Arabidopsis root and shoot, respectively. C,D: Number of significant motifs relevant to expression of genes in A and B, respectively. A total of 2,106 genes are up-regulated, and 1244 significant motifs are discovered. E: Distribution of statistical scores of all motifs from cold up-regulated genes in shoot. Pink and green bars show the cutoff (0.01) of hypergeometric  $p$ -value and the cutoff (0.1) of gscore for expression coherence, respectively. The Y-axis is the  $p$ -value from a hypergeometric test, which quantifies the enrichment of up-regulated genes in all genes containing the corresponding motif. The X-axis is gscore, measuring expression coherence of the genes containing the corresponding motif.

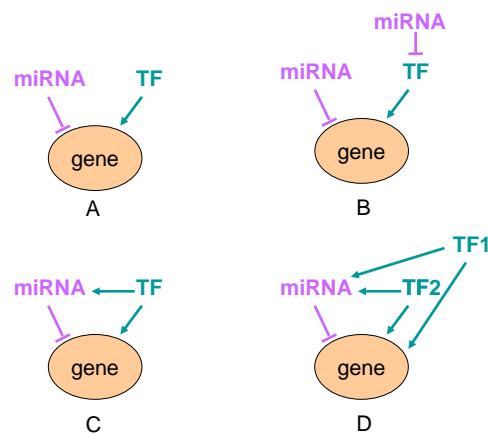


Figure 2: Four regulatory circuits mediated by miRNAs and TFs in plant. A: Cooperated TF-miRNA pair; B: Indirectly cooperated miRNA-miRNA pair; C: Feedforward loop; D: Composite feedforward loop.

We analyzed the co-occurrence of miRNA target sites and 325 known regulatory motifs. Among 4541 motif-miRNA pairs, 1239 (27.3%) were significant. Similarly, we analyzed the co-occurrences of miRNA target sites and the motifs extracted from stress-induced genes. We identified 1120 motif-miRNA pairs for the new motifs in stress-induced genes. Among them, 197 (17.3%) were significant. For new motifs from up-regulated genes at different development stages, 5272 motif-miRNA pairs were obtained and 1510 (28.6%) were significant.

We further tested whether these regulatory circuits, such as feedback and feedforward loops [5], existed in the motif-miRNA pairs discussed above. As shown in Figure 2, three more circuits were discovered. Furthermore, we found that many of the motif-miRNA pairs and other three regulatory circuits were conserved in rice. Genes containing these motif-miRNA pairs were expressed more coherently than others.

## References

- [1] D.G. Altman. *Practical Statistics for Medical Research*. Chapman & Hall/CRC, 1991.
- [2] R. Breitling, P. Armengaud, A. Amtmann, and P. Herzyk. Rank products: a simple, yet powerful, new method to detect differentially regulated genes in replicated microarray experiments. *FEBS Lett*, 573(1-3):83–92, Aug 2004.
- [3] M.W. Jones-Rhoades and D.P. Bartel. Computational identification of plant microRNAs and their targets, including a stress-induced miRNA. *Mol Cell*, 14(6):787–99, Jun 2004.
- [4] J. Kilian, D. Whitehead, J. Horak, D. Wanke, S. Weinl, O. Batistic, C. D’Angelo, E. Bornberg-Bauer, J. Kudla, and K. Harter. The AtGenExpress global stress expression data set: Protocols, evaluation and model data analysis of UV-B light, drought and cold stress responses. *Plant J.*, 50(2):347–63, Apr. 2007.
- [5] S.S. Shen-Orr, R. Milo, S. Mangan, and U. Alon. Network motifs in the transcriptional regulation network of escherichia coli. *Nat Genet*, 31(1):64–8, May 2002.
- [6] R. Sunkar and J.K. Zhu. Novel and stress-regulated microRNAs and other small RNAs from Arabidopsis. *Plant Cell*, 16(8):2001–19, Aug 2004.
- [7] G. Wang, T. Yu, and W. Zhang. WordSpy: identifying transcription factor binding motifs by building a dictionary and learning a grammar. *Nucleic Acids Res*, 33(Web Server issue):W412–6, Jul 2005.
- [8] G. Wang and W. Zhang. A steganalysis-based approach to comprehensive identification and characterization of functional regulatory elements. *Genome Biol*, 7(6):R49, 2006.
- [9] Y. Zhang. miru: an automated plant mirna target prediction server. *Nucleic Acids Res*, 33(Web Server issue):W701–4, Jul 2005.