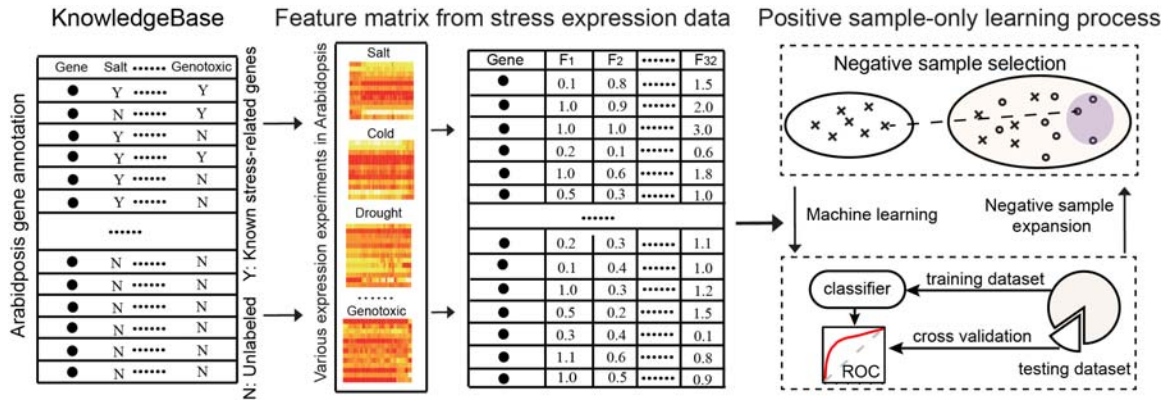
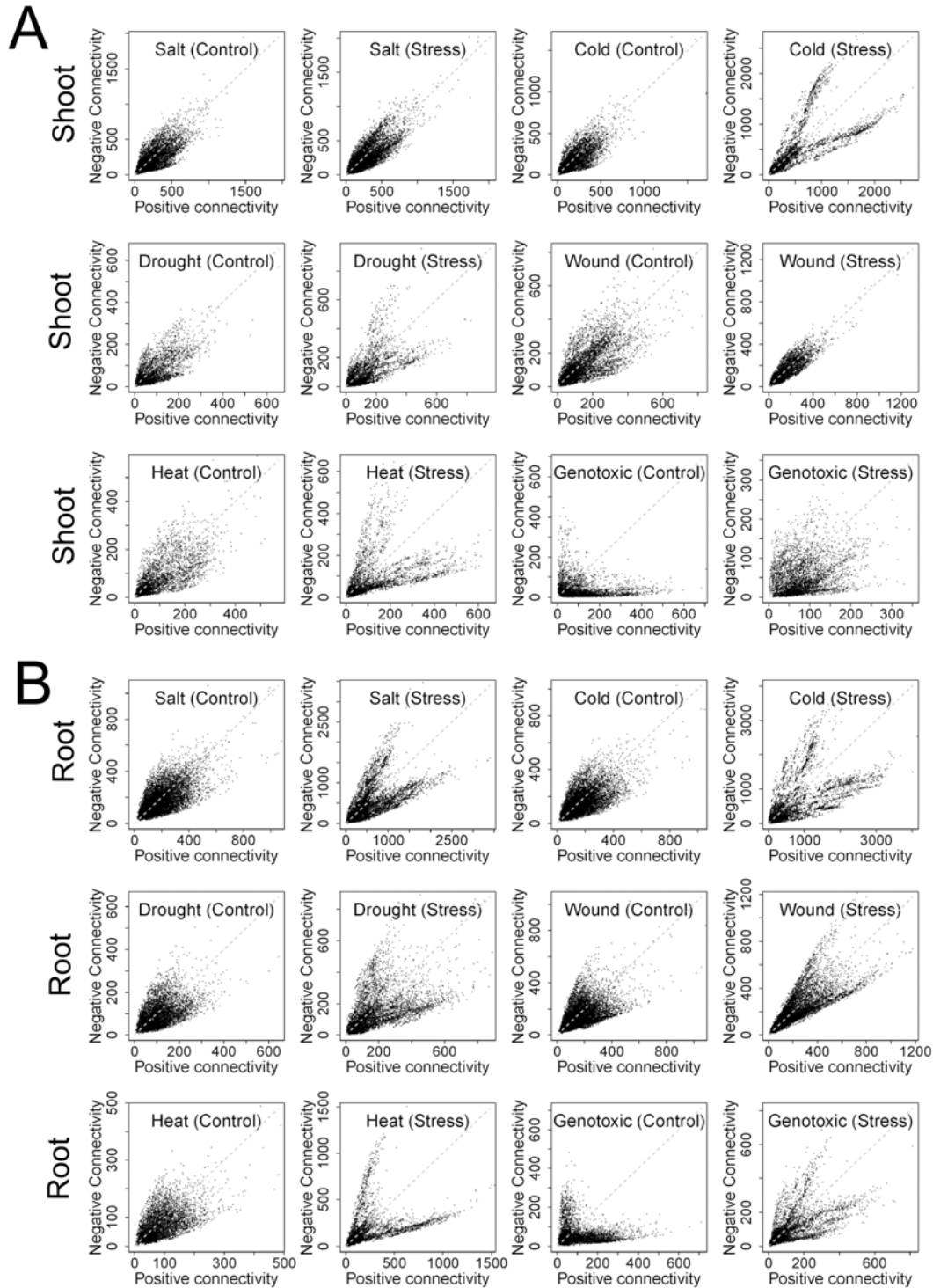


**Supplemental Figure 1. Workflow of the Gene Filtering Process based on the ML-based RandomForest Classification Model using the Positive Sample-only Learning (PSOL) Algorithm.**

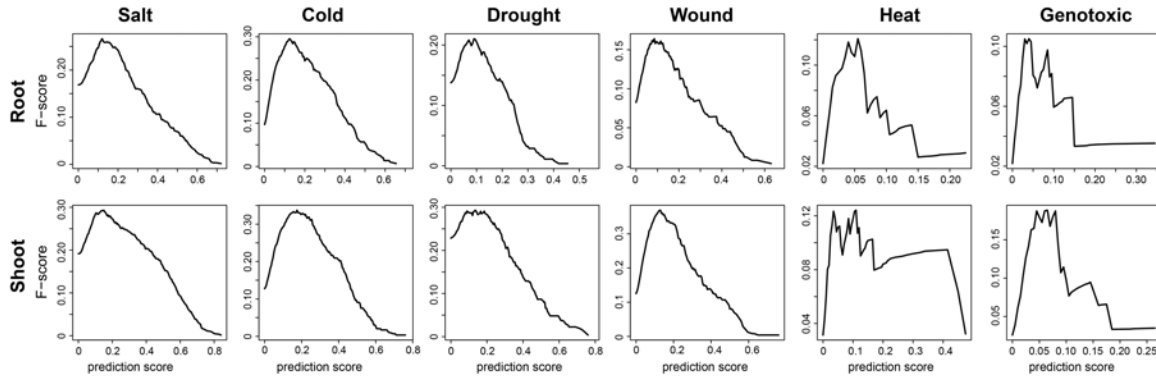


**Supplemental Figure 2. Comparison of the Distributions of Two Network Characteristics – Positive and Negative Connectivity – of Known Stress-related Genes and “Informative” Genes in Control and Stress networks to Illustrate Network Changes in the Two Networks.**



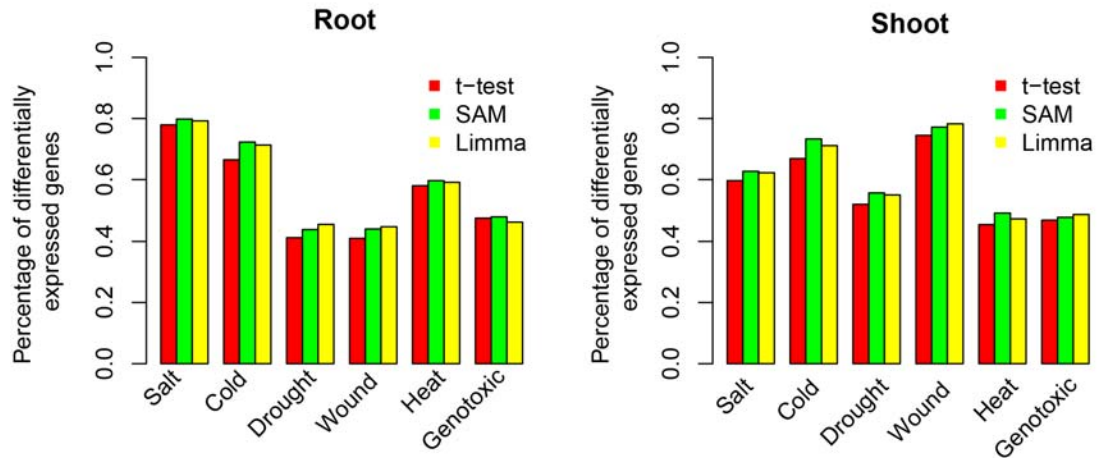
For salt, cold, drought, wound, and heat stresses, connectivity tends to increase after stress treatments, which indicates that intensive stress-responsive regulation occurred. Under genotoxic stress, both the positive and negative connectivity of genes decreased in the stress network compared with the control network, which is consistent with the fact that genotoxic stress could destroy the proper DNA structures and thus interfere with the proper interactions of DNA-binding proteins, such as TFs, and targeted DNA regions, such as gene promoters. In addition, the shape of the distribution of positive and negative connectivity might reflect the extent of stress-induced expression changes. While the “V” shape distribution indicates strong regulation either in the positive or negative direction, the “comet” shape distribution indicates relatively stable gene expression. For example, in the salt (root) experiment, the significantly increased connectivity and clearly separated positive and negative connectivity indicates that the gene expression regulation was more dramatic compared with the salt (shoot) sample. The other example is genotoxic stress in roots: in the control network, the positive connectivity was higher than the negative connectivity, whereas in the stress network, the clear separation of positive and negative connectivity disappeared. This pattern suggests that genotoxic stress usually influences the appropriate binding of TFs and target genes.

**Supplemental Figure 3. Determination of the Prediction Threshold of the RF Classification Model Using the F-score Algorithm.**



The optimal prediction score was selected when the F-score reaches the maximal value.

**Supplemental Figure 4. The Percentages of Differentially Expressed Genes in mDNA-predicted Stress-related Candidate Genes.**



The significance level (*p-value*) of differential expression was calculated with t-test, SAM and Limma using R/Bioconductor package “GeneSelector”. The *p-value* cutoff was 0.05.

**Supplemental Figure 5. Number of Salt Stress-related Genes in the GO Category “Response to Stimulus” and its Three Daughter GO Categories.**

