Network-based Transcriptome Analysis of Two Maize Genotypes Identified Pathways Associated With Differences in Salt Tolerance

Taher Mohasseli
University of Mohaghegh Ardabili
Razgar Seyed Rahmani
Ghent University: Universiteit Gent https://orcid.org/0000-0002-6821-4537
Reza Darvishzadeh
Urmia University
Sara Dezhsetan
University of Mohaghegh Ardabili
Kathleen Marchal (kathleen.marchal@ugent.be)
Department of Biochemistry, Genetics and Microbiology, University of Pretoria, Pretoria, South Africa.

Research article

Keywords: Salt stress, Zea mays L, Genotype selection, Expression analysis, Network analysis

Posted Date: November 4th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-100135/v1

Network-based transcriptome analysis of two maize genotypes identified pathways associated with differences in salt tolerance

Taher Mohasseli[1‡], Razgar Seyed Rahmani[2‡], Reza Darvishzadeh[3*], Sara Dezhsetan[1], Kathleen Marchal[2,4,5*]

[1]Department of Plant Breeding, Faculty of Agriculture, University of Mohaghegh Ardabili, Ardabil, Iran.

[2]Department of Plant Biotechnology and Bioinformatics, Ghent University, Ghent, Belgium.

[3]Department of Plant Production and Genetics, Faculty of Agriculture and Natural Resources, Urmia University, Urmia, Iran.

[4]Department of Information Technology, IDLab, imec, Ghent University, Ghent, Belgium

[5]Department of Biochemistry, Genetics and Microbiology, University of Pretoria, Pretoria, South Africa.

[**]Corresponding authors: E-mails: r.darvishzadeh@urmia.ac.ir; kathleen.marchal@ugent.be

‡These authors contributed equally to this work.

Abstract

Background

A better understanding of the molecular effects of salinity stress is key to improving salt tolerance in *Zea mays*. In this study, we combined phenotyping with transcript profiling and network analysis to study genotype-specific differences in salt tolerance in *Zea mays*.

Result

An extensive phenotypic screening identified two genotypes with an extreme phenotypic difference in tolerance towards salt stress. RNA-seq analysis of the selected salt-tolerant (R9) and salt-sensitive (S46) genotype was performed to unveil the molecular mechanism underlying the difference in salt tolerance. GO enrichment and network analysis on the results of the expression analysis identified phosphorylation-dependent signaling processes, ion transportation, oxidation-reduction, glutathione and tryptophan metabolism as the main processes different between the selected tolerant and sensitive genotypes. Genes belonging to the subnetwork enriched for phosphorylation and kinase activity shared a common regulatory element in their promoter region, which matched the binding site of an Arabidopsis TF with known role in salt-stress response.

Conclusion

Network-based transcriptome analysis of two maize genotypes identified pathways associated with differences in genotype-specific salt tolerance and identified a link between transcriptional and posttranslational regulation of salt tolerance.

Keywords

Salt stress, *Zea mays* L, Genotype selection, Expression analysis, Network analysis

Background

Improving response to environmental stresses in crop cultivars is a major challenge in plant breeding. Salt stress is one of the abiotic stresses that negatively affects crop growth and productivity. More than six percent of the world's total land area is affected by excess salt [1]. Maize (*Zea mays* L.) is the third most important world's cereal crop after wheat and rice and is used for food, feed and biofuel [2, 3]. Like most crop species, the majority of maize genotypes are moderately sensitive to salinity and their growth

and production are adversely affected by salt stress [2, 4]. Excessive salt concentrations in the plant root rhizosphere results in osmotic stress, ion imbalance, and oxidative stress [5-7]. Osmotic stress caused by salt stress has similar effects as stress induced by water shortage and leads to water deficiency [8, 9]. Ion imbalance results in cytotoxicity due to the accumulation of ions such as sodium (Na^+) and chloride (Cl^-) in plant cells [2, 5, 10]. An excess Na^+ and Cl^- ions disrupts the uptake of other ions particularly, Ca^{2+} and K^+ which are essential for the catalytic activity of most enzymes [11, 12]. Excess salt can also result in oxidative damage due to the production of reactive oxygen species (ROS), the effect of which depends on the intensity and duration of the stress and the growth stage of the plant [12-15]. To counter the negative impact of salt stress, plants have developed avoidance and tolerance mechanisms. Avoidance is a rapid reaction to prevent or delay the negative impact of salt stress [16]. Tolerance is achieved by a rapid decrease in stomatal conductance, compartmentation of toxic ions into vacuoles, and accumulation of compatible solutes such as proline, glycinebetaine, sugars, proteins, and polyols that result in ionic and osmotic homeostasis [12-15]. Tolerance comes at the expense of a decreased photosynthetic rate and metabolic capacity. Those aforementioned responses are regulated in plants by initiating fast and efficient signaling reactions such as the abscisic acid (ABA)-dependent and independent signaling of which the regulation can be lineage-specific [17-20]. Hence, because of the complexity of the salt stress phenotype, developing genotypes with increased salt-tolerance requires a deeper understanding of the molecular basis underlying the tolerance phenotype [1, 13, 21]. The well-known model plant, Arabidopsis, has been instructive in furthering our understanding of salt tolerance mechanisms. However, as regulation mechanisms vary among plant species, it is difficult to extrapolate results between species [22, 23]. Thanks to its cross-pollinated behavior, maize is highly polymorphic and provides genotypes with different tolerance to salt stress. Exploiting this genotypic and phenotypic variation offers the opportunity to study salt tolerance mechanisms in maize.

Transcriptome analysis reveals genes of which the expression levels are significantly altered between conditions. Although useful to identify the biological processes that are activated or inhibited under these conditions, the mechanisms that led to these changes often remained unexplained [24, 25]. In addition, transcriptome analysis can result in both false positive and negatives: on the one hand spuriously differentially expressed genes can be identified that do not contribute to the process of interest, whereas on the other hand genes that are not being regulated at expression level or that exhibit too subtle

expression changes remain unidentified [26]. Integrating complementary information with the expression data can leverage the information contained in expression data. Network analysis provides an intuitive way to combine expression data with prior information on known molecular interactions or already available functional data [27-29]. Network analysis maps candidate genes identified through expression analysis on an integrated network and searches for subnetworks that connect as many candidate genes as possible [27, 30-32]. By leveraging candidate genes identified through expression analysis with known interaction information, spuriously identified candidate genes can be removed as they will not be part of the subnetworks. In addition, genes relevant to the process of interest that are themselves not regulated at the level of expression are indirectly identified by being part of a connected component/subnetwork to which also many of the candidate genes belong. Such an integrated analysis allows gaining a more comprehensive view on the process of interest [27-29]. Here we applied such an integrated network-based strategy to unveil the molecular mechanisms underlying the differences in salt tolerance between two genotypes of *Zea mays*.

Result

Morphological and physiological response to salinity stress of two selected genotypes

Ninety-three maize inbred lines obtained from different Iranian research centers were evaluated in a randomized complete block design with three replications. Agro-biological and physiochemical traits were assessed under normal and 8 deci-siemens per metre (ds/m) salinity stress during two successive years in pot conditions. Two lines exhibiting statistically significant differences in response to salinity stress were selected, here referred to as R9 and S46 with R9 being the most salt-tolerant and S46 the most salt-sensitive inbred line. Relevant traits were measured in plants of both genotypes that were subjected to salt stress or grown under normal conditions. Samples were taken at 7 and 12 days post the application of the salinity stress. The salt treatment of the growth environment drastically affected the relative water content (RWC), plant weight, height, and stem diameter of the sensitive genotype (S46) after 7 and 12 days while it had a considerably smaller impact on the R9 phenotype (Fig 1). In addition, leaf growth (width and length) was severely affected in plants of the sensitive genotype while being normal in plants of the tolerant line (Figure 1). Limited leaf growth, caused by reducing the number of elongating cells and the rate of cell elongation, is indeed known to be one of the main morphological symptoms of salinity stress [33].

Na^+ and K^+ contents were measured 12 days after salt treatment in the stressed and control plants. Under control conditions, the K^+ content of the two genotypes was comparable but sharply decreased in plants of the sensitive genotype under salt treatment (Fig 1). Unlike the K^+, the Na^+ content did not show significant difference between two genotypes. This indicates that under salt stress the tolerant genotype was effectively taking up K^+ to keep the K^+/Na^+ ratio in balance while this was not the case for the sensitive genotype. Morphologically, interfering with K^+ uptake is known to lead to disturbances in stomatal modulations and causes water loss and necrosis [34]. In our study, neither the resistant nor the sensitive genotype showed excess accumulation of Na^+ in the leaves. Hence, an inefficiency of K^+ uptake led to imbalance in K^+/Na^+ in plants of the sensitive genotype (Figure 1). This result suggests that plants of the tolerant genotype might possess an efficient regulation of K^+ up-take that allows maintaining the K^+ and Na^+ homeostasis under salt stress.

Transcriptome analysis

To explore the molecular differences between the salt-tolerant (R9) and sensitive (S46) genotypes that underlie their difference in salt response, expression profiles of plants belonging to both genotypes were compared under normal and salinity stress using two biological replicates (each replicate represents pooled samples collected at two different time points post the application of the salinity stress). Data were preprocessed as explained in materials and methods. We were interested in identifying genes that showed a significantly different response to salt stress between the R9 and S46 genotypes. This requires identifying all genes that display an altered expression under salt stress versus normal conditions and for which the response is different between two genotypes. Genes that display an altered expression under salt stress versus normal conditions, but that show the same response between the two genotypes were removed as for those genes it is difficult to distinguish whether their altered expression under salt conditions is caused by sub-optimal growth under salt stress or by the genotype-specific triggering of salt-tolerance mechanisms (Fig. 2). In total, our candidate gene list contained 113 genes potentially involved in the genotype-specific response to salt stress (fold change > 1.5, FDR < 0.05) (Additional file 1 and 2). This list contains genes that display an altered expression under salt stress versus normal conditions and for which the alteration in expression is different between two genotypes.

GO analysis of the gene list showed enrichment in 'response to endoplasmic reticulum (ER) stress', 'systematic acquired resistance', 'tryptophan metabolism', 'dicarboxylic acid metabolism' and 'protein retention in Golgi apparatus' (Fig 3). Some of these identified pathways can be related to salinity stress. However only 12 genes from our candidate list belonged to the enriched functions although 111 genes could be mapped to at least one GO biological process term. This indicates that a substantial part of the salt tolerance mechanism remained unexplained by the enrichment analysis.

Network-based analysis

To better explain the mechanisms of salt tolerance reflected by the candidate genes, network-based analysis was performed. We generated first a physical maize interaction network using known interaction data including protein-DNA (regulatory), protein-protein, and metabolic interactions (KEGG). The network scaffold was complemented with additional functional edges derived from the large body of publicly available expression data in maize. Expression derived edges were derived from a co-expression network (Materials and methods). To avoid adding spurious interactions, only highly confident co-expression edges from the co-expression network were added. To assess the relevance of the selected edges, we clustered the high confident co-expression network containing these high confident edges and performed GO enrichment of the obtained clusters. The majority of the clusters showed high confidence enrichment to at least one biological pathway (out of 67 overlapped clusters, 41 clusters showed high enrichment for at least one biological process with p-value < 10e-5), supporting the relevance of the added edges.

Prior to performing network-based analysis, the obtained integrated network was converted into a weighted network using a topology based weighting scheme (Materials and Methods). This weighting scheme aims at weighting the edges based on their global connectivity in the network, and reduces the impact of hubs on the network-analysis: hubs with many neighboring genes risk connecting candidate genes in connected components that do not have direct biological links. To perform network analysis, the candidate genes were mapped on the interaction network (Fig 2). Subsequently, subnetworks were identified that connect as many as possible candidate genes using the least number of edges [27, 31]. Those sub-networks are proxies of pathways that contribute to genotype-specific expression differences.

Six subnetworks of different size connecting the candidate genes were identified (Fig 4). The first subnetwork (Subnetwork1) contains most genes and displays the highest connectivity among its nodes. Subnetwork 1 and 2 consist mostly of protein interactions whereas subnetworks 3-5 contain metabolic interactions only. This indicates that each molecular layer in the network contributes different information and little redundancy exists between interaction networks at different molecular layers, as is observed in many organisms [35-37]. Relatively few co-expression edges have been used to connect the candidate genes on the interaction network, indicating that adding them did not result in over-connecting the network. No regulatory edges were used to connect the genes from the candidate list, indicating that regulatory edges are quite sparse and understudied in the used interaction network.

Subnetwork1 is highly enriched for **"phosphorylation" related processes** (FDR: 6.6e-6) with 13 genes annotated to those processes. As expected, some of the genes in our selected subnetwork annotated as phosphorylation-related were not in our candidate gene list (*Zm00001d049727*, *Zm00001d048054*, *Zm00001d010234*, *Zm00001d020355*, *Zm00001d043562*, *Zm00001d005135*, *Zm00001d017525*). These genes were not significantly differentially expressed themselves but are recovered as 'connector' genes because of their high connectivity to the differentially expressed candidate genes (Fig 4, subnetwork1, gray nodes). *Zm00001d020355* and *Zm00001d017525* are involved in the "stress-activated protein kinase signaling" pathway (GO:0031098). Given these enrichments and the fact that this subnetwork 1 is biased towards protein-protein interaction, differences in salt-tolerance mechanism between B9 and S46 seem mostly related to protein mediated post-translational protein modifications and signaling (Fig 5).

The second-largest subnetwork (subnetwork 2) is highly enriched for '**oxidation-reduction'** related processes **(**GO:0055114**)** (Fig 5), indicating that the candidate genes in this subnetwork are relevant for ROS homeostasis and regulation during salt stress which is in line with the literature [38-40]. The candidate genes in this sub-network (red nodes) are not directly connected to each other, but are connected through a few connector genes (Fig 4, subnetwork 2). The smaller sub-networks (4, 5, 6) are enriched for glutathione metabolism, tryptophan biosynthesis and lignin biosynthesis, respectively (Fig 4 and Fig 5), processes that have been documented in the literature to relate to salt stress [2, 13, 41-45].

Transporters with known roles in ion homeostasis were identified in subnetwork 1 and 3. Among those were genes with GTPase activity (*Zm00001d039091*, *Zm00001d011474*, and *Zm00001d039090*).

Genes with GTPase activity have shown to be essential during ion homeostasis, particularly for the maintenance of Na^+ and K^+ homeostasis under salt stress [46]. For example, the overexpression of one of the GTPase activity gene, PtRabE1b, conferred salt tolerance in poplar [47]. GTPase and ATPase activity genes are also known to interact with genes involved in salt tolerance [48, 49]. In addition, GTPase activity was shown to be required for the reorganization of microtubules, a key response mechanism during salt stress in plants [20].

Motif detection

We identified 5 sub-networks containing candidate genes. These subnetworks are likely to reflect processes involved in salt tolerance. Given that the genes in these subnetworks are triggered by salt stress at the level of transcription, we assumed that they might be under control of the same transcriptional program that acts under salt stress. Given the large underrepresentation of transcriptional interactions in our network, we were not able to unveil this regulatory program through network analysis. Here we use as alternative a *de novo* approach based on motif analysis to recover the missing regulatory program. We hereby focused on the largest subnetwork 1, which contains genes enriched for phosphorylation and kinase activity. Assuming that the genes of this subnetwork were regulated by the same TF, we searched for a statistically overrepresented motif in their promoter region. As *de novo* motif detection is very sensitive to the presence of false positives, we excluded the genes that are only marginally connected in the subnetwork (genes with at most one edge connectivity to any other gene in the network).

The 1kb upstream of the transcription start site (TSS) was selected to search for regulatory elements. When the gene was located on the negative strand, the reverse complement of the sequence was considered. We found one motif overrepresented in the selected genes of subnetwork 1. This motif occurred in 17 out of 21 sequences and showed a positional bias towards the TSS in most of the sequences (Fig 6). The identified motif turned out to be highly similar to a known Arabidopsis motif from the DAP-2016 dataset [50], representing the binding site of FUF1 (AT1G71450). FUF1 is a member of DREB subfamily A-4 of ERF/AP2 transcription factors and was shown to exhibit a salt specific expression response in Arabidopsis root tissue [51]. FUF1 has a unique homolog in maize named *Zm00001d048991* and is alike in Arabidopsis, expressed in a tissue-specific way [52].

Discussion

In this study, we selected two genotypes with an extremely different phenotypic behavior towards salt stress and subjected them to expression analysis. We identified genotype-specific differences in salt induced gene expression. The genes that were salt induced in a genotype-specific way were functionally characterized by complementing a GO enrichment with network analysis [27, 53]. Hereto we used an integrated network as a backbone to search for subnetworks that most optimally connect the differentially expressed genes of our candidate gene list. These subnetworks reflect the pathways of relevance triggered by the studied process. The advantage of this approach is that the backbone network used to drive the analysis contains next to well-annotated connections between genes also less well documented edges, derived from the large body of publicly available expression data on *Zea Mays*. As such also the less characterized differentially expressed genes (i.e. genes without GO annotation) can be assigned to a common process, here reflected by a subnetwork. Indeed we observed that trough network analysis considerably more genes of the candidate gene list could be explained than by performing a GO enrichment analysis (of the total 113 prioritized candidate genes, GO enrichment and network analysis could explain 12 and 38 genes, respectively). In addition to these 38 recovered candidate genes, network analysis also identified 28 connector genes. These genes were not identified as differentially expressed under salt stress, but belong to the same network neighborhood as the differentially expressed candidate genes, indicating that they also might play a role in the salt tolerance mechanism.

Cellular potassium homeostasis is known to be one of the main contributors to salt tolerance [54]. Salt stress activates the salt-overly-sensitive (SOS) system to maintain ion homeostasis [55]. The SOS system perceives salt stress through the rise in free cytosolic calcium [41]. Those calcium signals activate kinases and subsequent phosphorylation cascades which result in regulating gene expression in response to salt stress. Subnetwork 1 is highly enriched for phosphorylation related processes. This process was not identified by performing GO analysis directly on the candidate gene list: phosphorylation related processes are likely to be regulated by post-transcriptionally rather than by transcriptional mechanisms. Hence, they cannot be identified through differential expression analysis only. However, by combining expression with network analysis genes belonging to these processes can be recovered as connector genes. This was indeed the case: subnetwork 1 contains next to genes that are differentially expressed

under salt stress (*Zm00001d051915*, *Zm00001d018799*, *Zm00001d029258*, *Zm00001d011573*, *Zm00001d019411*, *Zm00001d022166*) and that were annotated as phosphorylation-related also several connector genes associated to phosphorylation-related processes (*Zm00001d049727*, *Zm00001d048054*, *Zm00001d010234*, *Zm00001d020355*, *Zm00001d043562*, *Zm00001d005135*, *Zm00001d017525*) (Fig 4). This result indicates that the activated phosphorylation cascades contribute to the maintenance of potassium homeostasis in the tolerant genotype (R9) which is in line with the phenotypic results (Fig 1).

Next to the subnetwork enriched for phosphorylation, network analysis identified several sub-networks enriched for pathways and processes with known relevance to salt stress (ion transportation, oxidation-reduction, glutathione and tryptophan metabolism) [5, 13, 21]. The enrichment of processes related to 'oxidation reduction' is in line with previous findings that showed how salt-tolerant genotypes exhibit strong responses against ROS molecules by the upregulation of antioxidant defense mechanisms [56, 57]. This up-regulation of ROS detoxification in tolerant genotypes also helps to minimize the oxidative damage to proteins, lipids and carbohydrates in order to maintain growth [2, 6, 13]. Besides their negative effects, ROS can also be employed as signaling molecules to trigger a cascade of signaling pathways that result in the perception and activation of salt responses [58]. Therefore, the extra produced ROS in the tolerant genotype might help to sense the salt stress in an earlier stage and more easily adapt to the salt stress. [38-40]. ROS detoxification mechanisms are activated to counteract the negative effects of ROS. Glutathione is known to provide protection against salt stress-induced oxidative damage [42, 44]. Tryptophan is the major precursor of indole acetic acid (IAA), the most common plant hormone of the auxin class, and precursor of several compatible solutes involved in ionic and osmotic homeostasis under salinity [2, 13, 43]. Regulating lignin biosynthesis is known to be the main contributor to enhance salt tolerance in plants [45, 59].

Transporters with GTPase activity (*Zm00001d039091*, *Zm00001d011474*, and *Zm00001d039090*) which were enriched in subnetwork 1 have shown to be involved in signaling and ion homeostasis (particularly of Na^+ and K^+) under salt stress [20, 47] and might indeed relate to the aberrations in ion homeostasis we observed at the phenotypic level. *De novo* motif detection allowed identifying a statistically overrepresented binding site in the upstream region of the genes belong to subnetwork 1 which was highly similar to the motif of a TF in Arabidopsis of which the role in salt-stress

response is well documented. Genomic variation in this TF could act as a transfactor and explain the differences in stress response between the salt-tolerant and sensitive genotypes.

Conclusion

Network-based transcriptome analysis of two mays genotypes identified pathways associated with differences in salt tolerance and identified a novel link between transcriptional and posttranslational regulation of salt tolerance.

Methods

Plant Material, Salt stress condition and RNA preparation:

Two maize genotypes with different sensitivity to salt stress were selected after screening ninety-three maize inbred lines for agro-biological and physiochemical traits under normal and 8 ds/m salinity stress during two successive years in pot conditions. Seeds were kindly provided by Seed and Plant Improvement Institute Karaj Iran; Razi University of Kermanshah; and Ferdowsi University of Mashhad. From the 93 inbred lines, S46 was identified as the line most sensitive and R9 as the line most tolerant to salinity stress. The R9 line was generated by the Razi University of Kermanshah and was kindly provided in the context of this study. The S46 line is a commercially available inbred line (Mo17) that has been widely used in many genetic and molecular studies. Seeds were soaked overnight in distilled water and then cultured in 25×20 cm pots containing perlite and peat moss at a ratio of 2:3. Plants were grown under standard conditions (22 to 26°C, 16 h light/8 h dark photoperiod and 70% relative humidity) at Urmia University. Irrigation was performed with 250 ml of Hoagland solution each other day. When the plants reached the 8-leaf stage, they were divided in two groups: irrigation of one group was continued with a Hoagland solution mimicking normal conditions (electrical conductivity (EC) of 2.7 ds.m-1) whereas the other group received Hoagland solution inducing salinity stress (electrical conductivity (EC) of 8 ds.m-1). Salinity stress of 8 ds.m-1 was obtained by dissolving NaCl in Hoagland solution. The seventh and eighth leaves of respectively the control plants and the plants subjected to salinity stress were sampled 7 and 12 days post the application of salinity stress. To extract RNA, the leaf tissue of three

plants were pooled for each biological replicate. Total RNA extraction from each sample was performed using RNX-Plus TM extraction solution (Sinaclon, Iran), according to the manufacturer's recommendation and was followed by DNase digestion. After quality control procedures using gel electrophoresis, nanodrop and Agilent 2100 bioassays, mRNA was purified. Equal quantities of total RNA obtained for each time point (7 and 12 days after salinity stress) were pooled for RNA sequencing.

RNA sequencing, quality check, differential expression analysis and GO enrichment

Illumina sequencing was performed by Berry Genomics Company, China. The cDNA libraries were sequenced on a Novaseq 6000 generating 150 bp paired-end reads. The raw data have been deposited in the Sequence Read Archive (SRA) accession number SRP273987 (https://www.ncbi.nlm.nih.gov/Traces/study/?acc=SRP273987&o=acc_s%3Aa). Sequencing quality was checked on the paired-end reads using FASTQC [60]. Low quality reads were filtered, adaptor contaminations and low quality bases were trimmed using Trimmomatic [61]. Hisat2 was used to map the reads on the maize B73 reference genome (Zea_mays.B73_RefGen_v4.dna.toplevel.fa.gz) (73% overall unique alignment rate) [62]. For quantification of expression levels, htseq-count was used along with the genome annotation from Ensembl release 43 (Zea_mays.B73_RefGen_v4.43.gtf.gz) [63]. Differential expression analysis was performed using edgeR and differentially expressed (DE) genes were identified using fold change > 1.5 and FDR < 0.05 [64].

We are interested in identifying genotype-specific expression changes that can explain the difference in phenotypic behavior between the sensitive and tolerant lines. These genes are referred to as candidate genes. To select these genes an ANCOVA model was used: gene expression is modeled as the response variable, ‘salt treatment’ and ‘genotype’ as the explanatory variables. The genes of interest are modeled as the ones having an interaction effect. By only considering the interaction term in the model, genes that are affected to the same degree (fold) by salinity stress in both lines are removed because they cannot be associated with genotype specific differences in response to salinity stress.

Therefore, for each gene (g), its expression (y_{ig}) was modeled by negative binomial (NB) distribution using salt treatment (S) and genotype (G) as explaining variables as follows.

$$y_{ig} \sim NB(\mu_{ig}, \phi_g)$$
$$E[y_{ig}] = \mu$$
$$\log(\mu_{ig}) = \eta_{ig}$$
$$\eta_{ig} = \beta_0 + \beta_{S.g} X_{iS} + \beta_{G.g} X_{iG} + \beta_{S.g:G.g} X_{iS} X_{iG}$$

X_{iS} is an indicator variable where $X_{iS} = 1$ if the sample is salt treated and $X_{iS} = 0$ otherwise. Likewise, the indicator variable $X_{iG} = 1$ if sample was taken from the sensitive genotype and $X_{iG} = 0$ if the sample was taken from the tolerant genotype. The tolerant genotype grown under normal conditions was considered the reference. For each gene, the mean expression in each treatment can be derived from the defined model as:

$$\text{Tolerant\_Control} = \beta_0$$

$$\text{Tolerant Salinity} = \beta_0 + \beta_S X_S$$

$$\text{Sensitive\_Control} = \beta_0 + \beta_G X_G$$

$$\text{Sensitive\_Salinity} = \beta_0 + \beta_G X_G + \beta_S X_S + \beta_{S:G} X_S X_G$$

Finally, for each gene the difference in response to salinity between the tolerant and sensitive genotypes after correcting for their differences in expression under control conditions can be derived as:

$$(\text{Sensitive\_salinity} - \text{Sensitive\_control}) - (\text{Tolerant\_Salinity} - \text{Tolerant\_control}) = \beta_{S:G}$$

Where $\beta_{S:G}$ corresponds to the interaction effect of the explaining variables in the design matrix. The design and contrasting matrices in edgeR were set based on the formula mentioned above.

Gene ontology was downloaded from maizegdb.org. GO enrichment was performed using ClusterProfiler [65] and GO enrichment graphs were created using R and Plaza [66].

Constructing the interaction network

A high confident protein interaction network for maize was downloaded from the PPIM database [67]. Since the PPIM database uses the old version of gene ids (B73-v3 reference genome), the ids were mapped to the B73-v4 reference genome using the v3 to v4 mapping ids provided by maizegdb.org. This resulted in a protein interaction network consisting of 10868 unique genes as nodes and 157934 interactions as edges. The Maize metabolic interactions were downloaded from KEGG [68] and converted to a network using KEGGgraph [69]. Corresponding v4 gene ids were retrieved from NCBI

using the Entrez search of Biopython [70], resulting in a metabolic network with 54807 edges between 4052 genes. A maize specific regulatory network was downloaded from PlantRegMap (*http://plantregmap.cbi.pku.edu.cn/*) [71]. To avoid adding too many unreliable regulatory links to the network, only interactions supported by functional transcription binding site (FTBS, PlantRegMap) were considered, resulting in 12759 edges between 5251 genes. The three obtained networks were merged. Duplicated edges were removed: if the duplicated edge was present in KEGG (as a most reliable source), this one was retained as metabolic edge and edges redundant with the KEGG edge derived from other sources would be removed. In the absence of a KEGG edge, the regulatory edge was retained at the expense of removing the edges derived from the protein interaction network.

Refinement of the interaction network using RNAseq expression data:

An improved version of the B73 genome, assembled from long-reads, with updated gene annotations (AGPv4) is available for maize [72]. However, most interactions for *Zea mays* are in each of their respective databases still annotated with the older annotation (v3). When performing the mapping between both versions, we noticed that a considerable portion of the v3 genome could not be mapped to v4 annotation. Simply converting loci between the deprecated versions of the genome and the v4 version seemed not sufficient as the v4 version contained several newly identified genes not yet present in the deprecated versions. To remove likely false positive edges resulting from the mapping errors between the older and the more recent genome versions, we exploited the large availability of expression data in maize in which expression has been profiled across a large number of samples, in multiple tissues and different conditions [73]. We integrated the 8 RNAseq samples from our study with the aforementioned dataset. Our final expression compendium consisted of 282 samples spanning the majority of developmental stages and tissues in *Zea mays*. To reduce the number of spurious and condition irrelevant edges in our network, we removed all genes and their interactions from the network if the gene was not expressed in at least 5 samples of our compendium (FPKM >1). On the other hand, for the genes in the genome release (v4-B73) that had no counterpart in v3 and hence also not in the interaction network, we added functional interactions derived from the co-expression network. To build the co-expression network the ‘rank of correlation coefficient’ (RCC) was used to determine the degree of pairwise co-

expression [74]. The RCC values were transformed into Mutual Rank scores (MR scores) based on the following formula:

$$MR(AB) = \sqrt{RCC_{(A\rightarrow B)} * RCC_{(B\rightarrow A)}}$$

Where rank $_{(A\rightarrow B)}$ is the RCC of gene B with gene A.

Smaller MR scores correspond to a higher degree of pairwise correlation between two genes and can be converted to a network edge weight using the following formula taken from Obayashi and Kinoshita, 2009:

$$weight_{(A\rightarrow B)} = e^{-(MR_{(A\rightarrow B)}-1)/5}$$

guaranteeing that the range of edge weights in the co-expression network scales between 0 and 1 [74]. To further filter spurious links, only the strongest coexpression links were retained using a threshold of 0.9 on the coefficient of Pearson correlation. To calculate the Pearson correlation between the expression vectors of two genes log transformed FPKM values were used. A small value (one) was added to the FPKM values prior to the log transformation in order to avoid having undefined values for zero values. In addition, genes that were not expressed in at least 10% of the samples were excluded from the dataset prior to calculating the correlation in order to avoid adding spurious relations for lowly expressed genes. To assess the quality of the coexpression network, we searched for clusters of connected genes using ClusterOne with default parameters. ClusterOne allows for overlap between clusters and only returns highly connected components [75]. GO enrichment was performed on the clusters obtained by ClusterOne. From the coexpression network, 66383 functional interactions were extracted and added to the aforementioned interaction network. This resulted in a final interaction network consisting of 269731 unique edges between 21236 genes (Additional file 3).

Weighting the network and performing network analysis

In a probabilistic network analysis, weights on the edges between nodes are used to drive the search for subnetworks. To design a weighting scheme we performed a Random walk with Restart (RWR) (restart parameter = 0.05) [76, 77]. The RWR uses as input the degree normalized adjacency matrix of the

interaction network. The RWR performs a topology based weighting reducing the impact of hubs. However, RWR produces relatively small values for most genes whereas genes that are marginally connected in the network get a relatively high weight. As we do not want to bias the network search too much towards the marginally connected genes, we rescaled the RWR obtained values using the following heuristic formula.

$$rescaled\ weight(s\_t) = \frac{e^{\text{weight(s\_t)}}}{e}$$

Where "s" is the source node and "t" the target node. This formula transforms the minimal weigh to 0.3 ensuring that all edges are considered during network analysis but that edges between highly connected genes (higher RWR value) remain higher in weight.

Extracting subnetworks from the weighted interaction network was performed using 'Phenetic', a probabilistic pathfinding approach [27, 31] with following parameters: mode: downstream; min cost: 0.1; max cost: 5; step size: log scale between max and min cost with 28 steps; Path-length=4; k-best paths: 20. For each edge cost, the highest scoring subnetwork was selected and a stability score was computed for all the subnetworks with this edge cost. For each cost, the subnetwork is rejected if either it has a low stability score (minimal stability score is 0.5) or is too large (max size is 80). The final subnetwork was extracted as a combination of all these "best networks" that passed the tests.

De novo motif detection was performed using MEME [78]. The promoter regions were defined as 1kb upstream of TSS and first-order model of sequences was used as background. The promoter sequences were searched for motifs with a width ranging from 6 to 15 using the sequence of the lead (+) strand only.

Figures

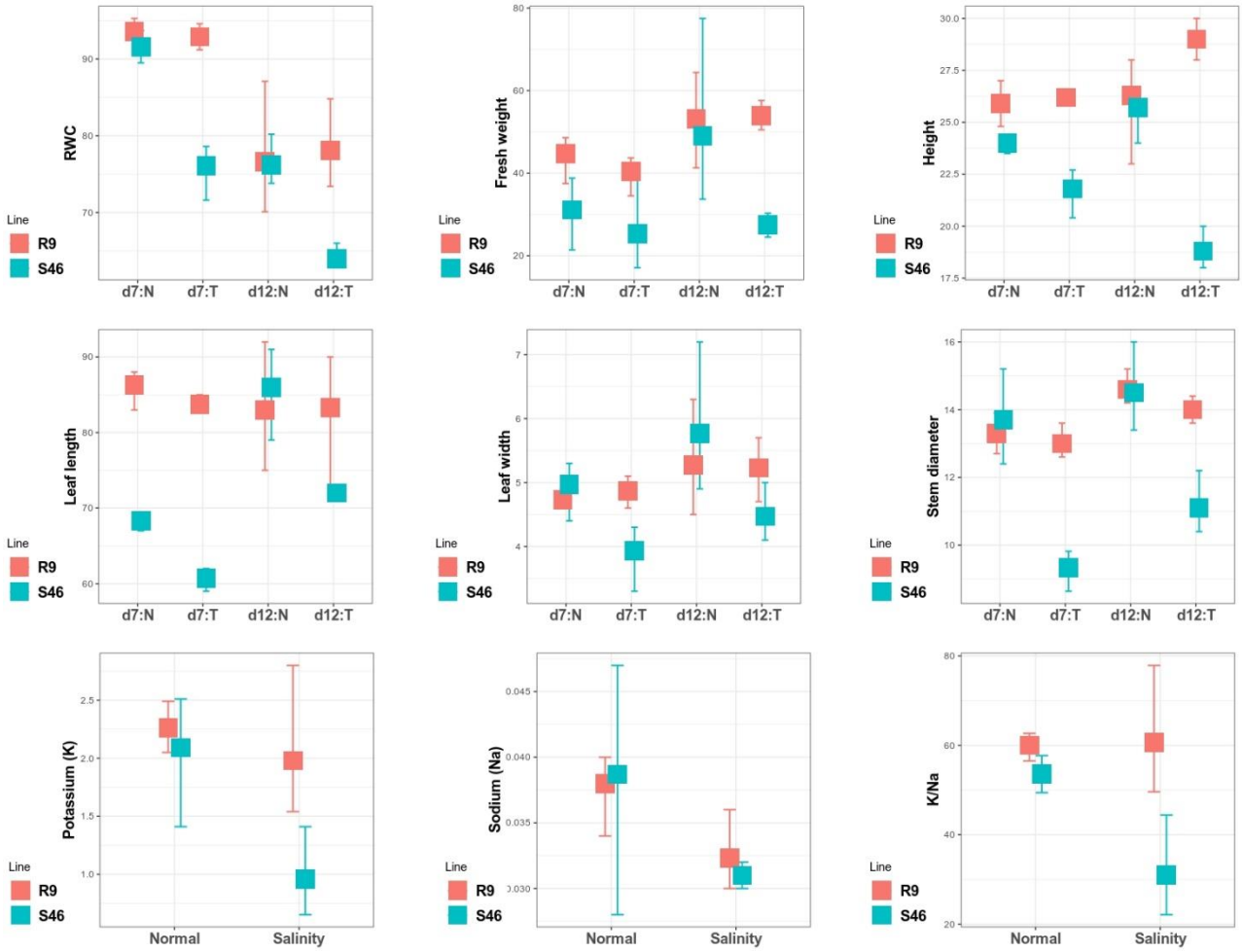


Fig. 1: Comparing morphological and physiological traits between R9 and S46 under normal and salinity imposed conditions (values obtained from samples taken at 7 and 12 days' post-salt treatment). The boxes show the mean based on three biological repeats and bars represent the 95 percent confidence interval for the means. Potassium and sodium content measured 12 days after salt treatment. Abbreviations: d7: N indicates day 7 under control condition; d7:T indicates day 7 after salt treatment; d12:N indicates day 12 under control, and d12:T indicates day 12 under salt treatment, RWC: relative water content.

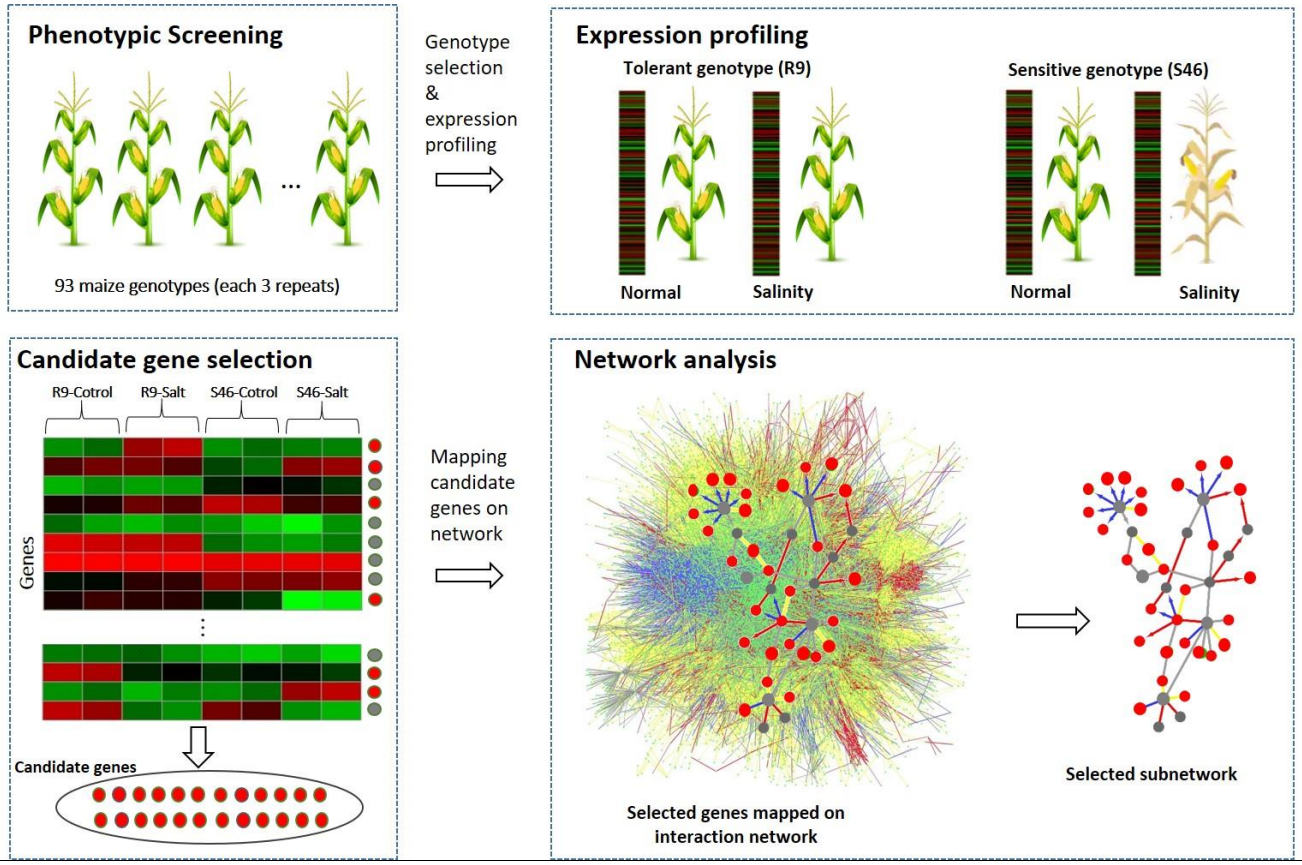
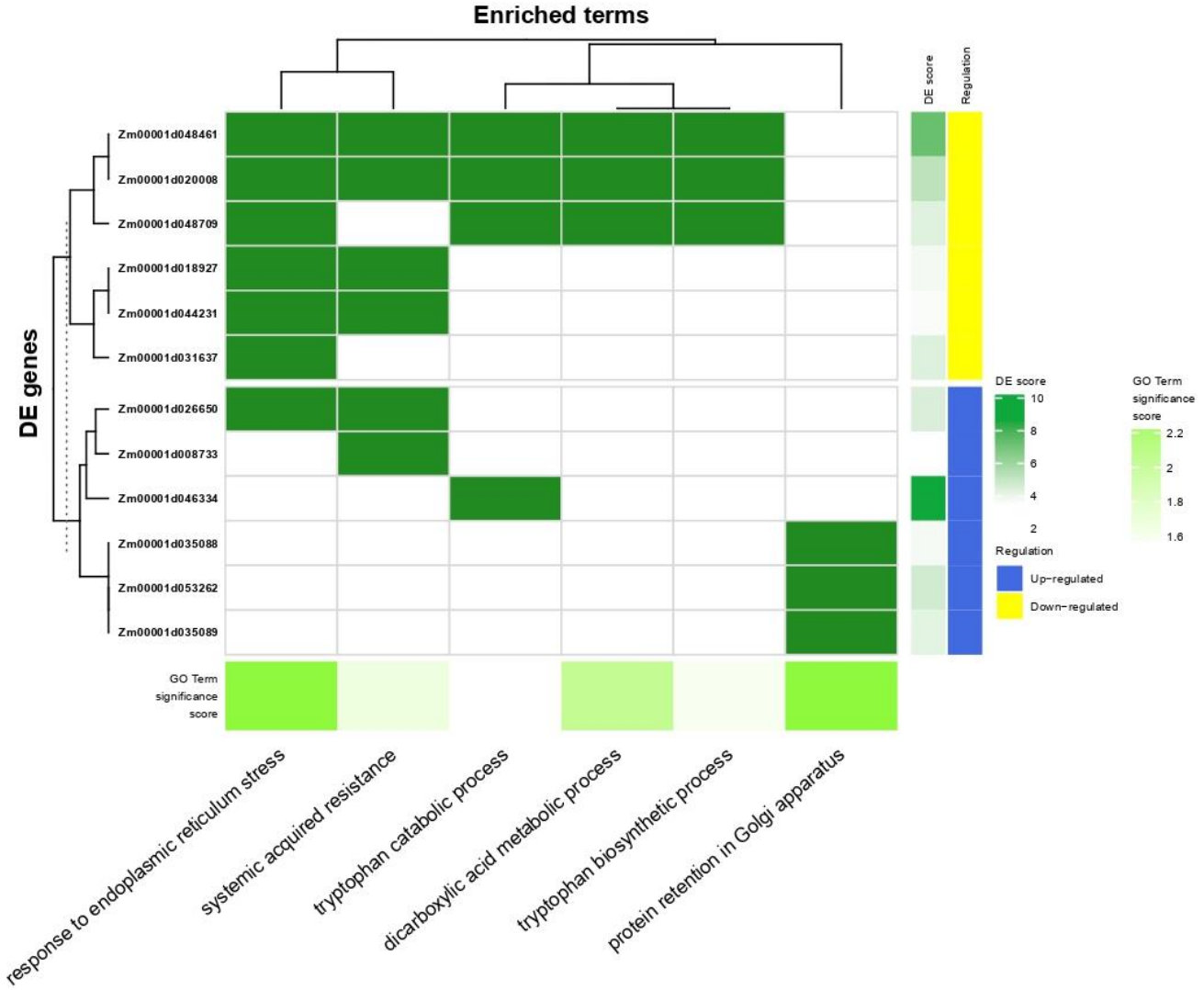


Fig. 2: The flow of mapping candidate genes resulting from the expression analysis on the interaction network. Phenotype screening: from 93 maize genotypes, two genotypes with an extremely different phenotypic behavior towards salt stress (R9:Tolerant, S46:Sensitive) were selected. Expression profiling: gene expression of the R9 and S46 genotypes profiled under normal and salinity imposed conditions. Candidate gene selection: the genes that responded differently to salinity stress between two genotypes are selected as candidate genes (red genes on the right hand of heatmap). Network analysis: an integrated interaction network was compiled from different sources (regulatory, protein-protein, metabolic and coexpression networks) and was subsequently converted to a probabilistic interaction network. Candidate genes were mapped on this probabilistic interaction network and subnetworks extracted using Phenetic. Red nodes are candidate genes (identified as differentially expressed in a genotype specific way) and grey nodes indicate connector genes recovered by the network analysis.



464

465

466

467 **Fig. 3**: GO enrichment for the candidate genes. The overrepresented GO term and the candidate genes are shown
468 on respectively the x-axis and y-axis. Green indicates that the corresponding gene is present in the GO class, the
469 DE score reflects the degree to which the candidate gene is differentially expressed (log fold change), Yellow and
470 blue indicate whether the gene was up versus down regulated under salt stress comparing S46 to S9 after correcting
471 for differences in expression under normal conditions.

472

473

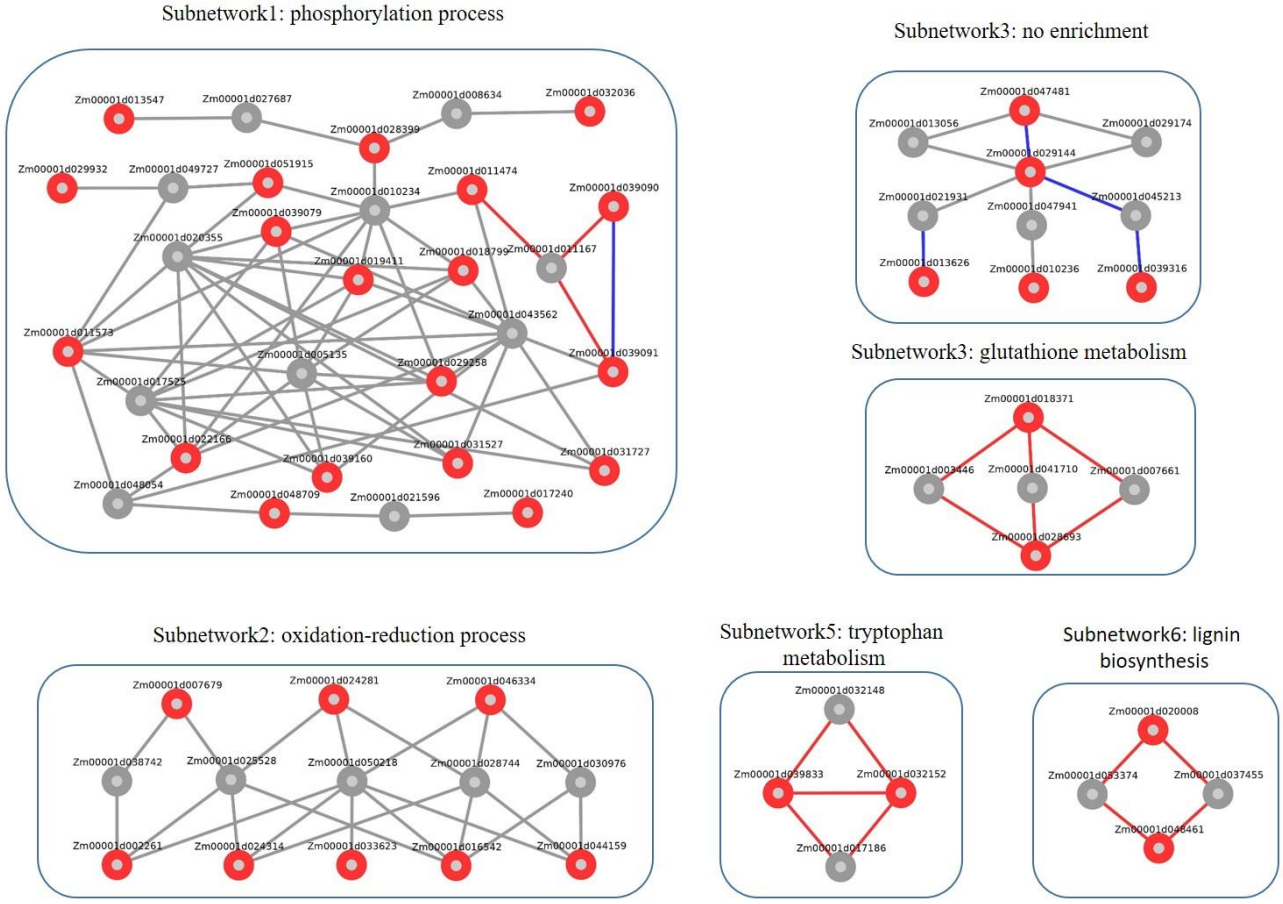


Fig. 4: Results of the network analysis. Red nodes represent candidate genes obtained from the expression analysis, grey nodes are not identified as differentially expressed, but were identified by the network analysis as connector genes (genes needed to connect the candidate genes). Edge color: red indicates a metabolic interaction, grey a protein-protein interaction, and blue a co-expression derived interaction. No regulatory interactions were recovered in any of the sub-networks

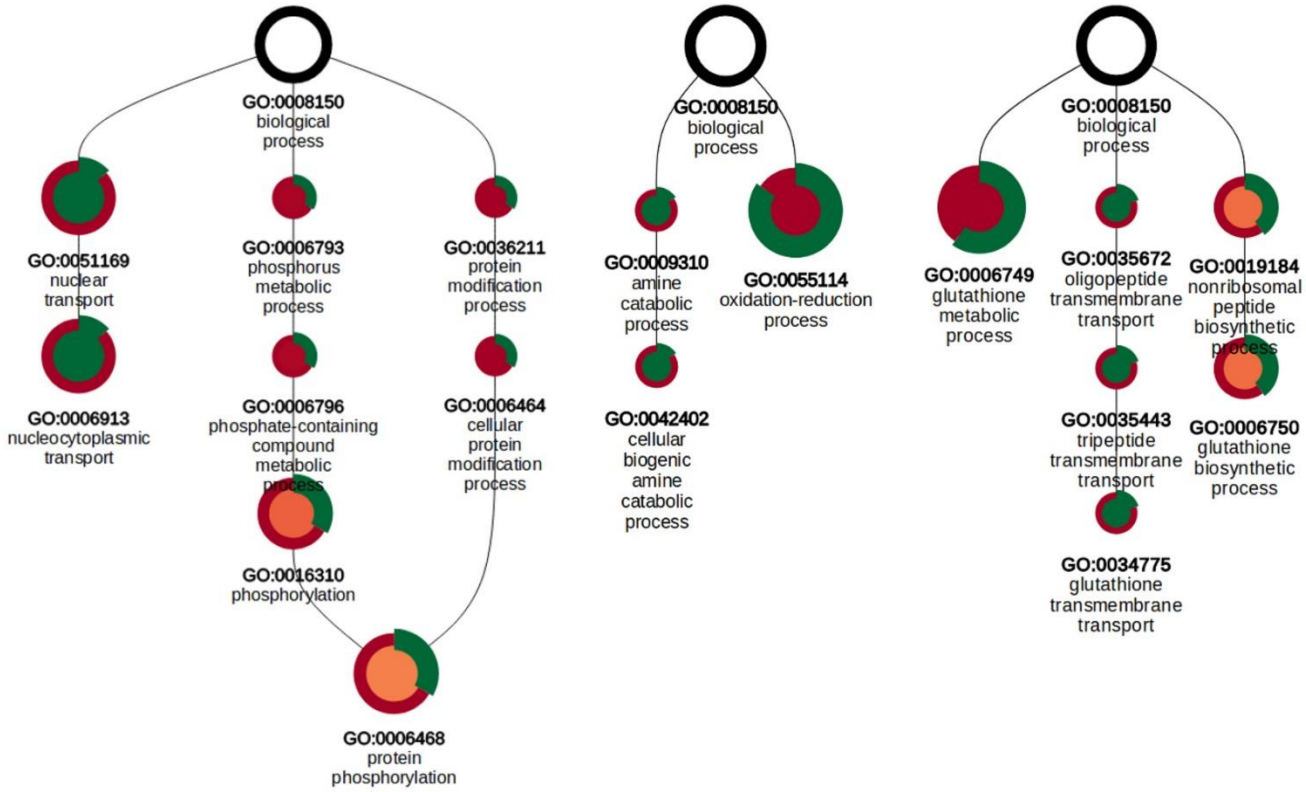


Fig. 5: The hierarchy of the GO enrichment result for sub-network 1 (left), sub-network 2 (middle) and sub-network 4 (right). Sub-network3 was not enriched for any specific biological process and the GO results for sub-network 5 and 6 are not shown as they consist of four genes only, all of which are annotated in KEGG metabolic pathways. Node size is scaled by the Bonferroni corrected p-value for enrichment; node color is determined by the enrichment fold such that green shows the highest and red shows the lowest fold enrichment; the node outer band reflects the 'Percentage Present' i.e. the percentage of genes that are annotated with the enriched ontology term (indicated by the green part of the ring).

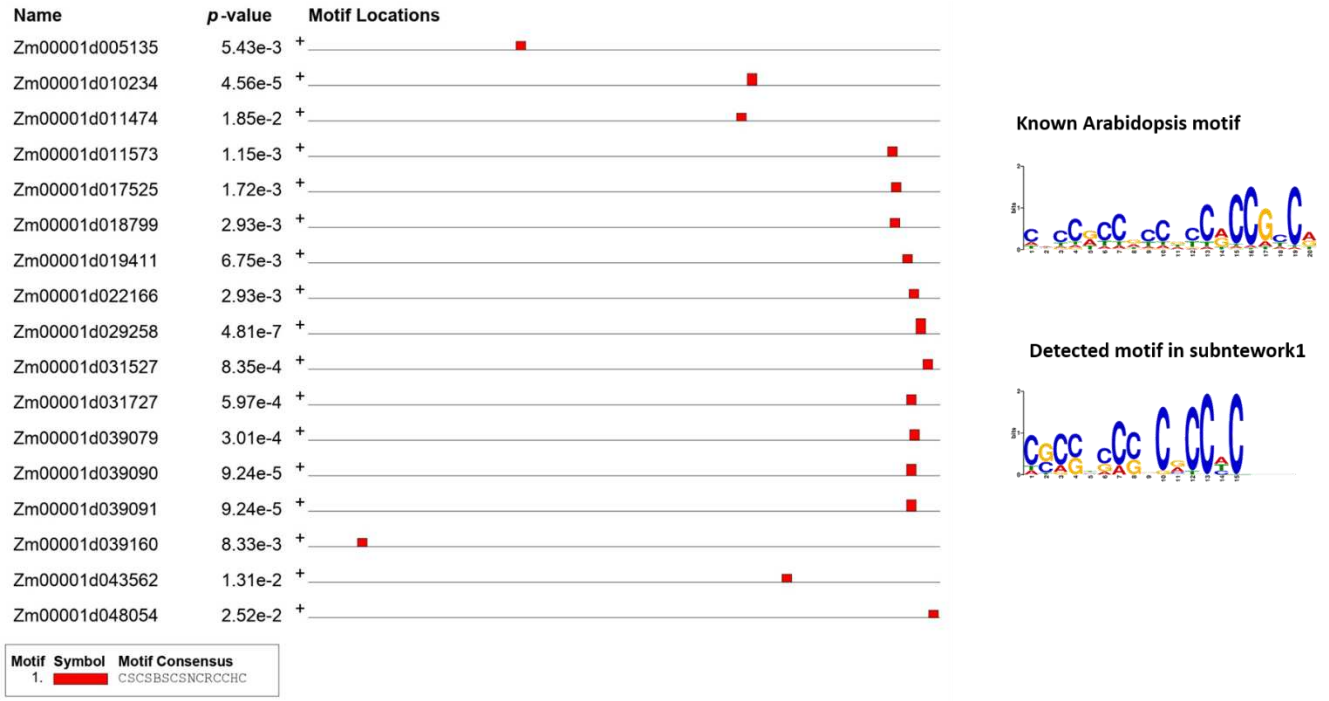


Fig. 6: Regulatory element overrepresented in the promoter sequences of the genes in sub-network 1. The identified motif shows a strong bias towards the TSS and high similarity with a known Arabidopsis motif representative for the FuF1 binding site (q-value 2.32e-04). The first logo is the Arabidopsis cis-regulatory element obtained from a DAP experiment and the second one represents the de novo detected motif in subnetwork1.

Availability of data and material

The RNA-seq data has been deposited to the National Centre for Biotechnology Information (NCBI) Sequence Read Archive (SRA) under the submission name: SRP273987 (link to the data).

Abbreviations

R9: Salt tolerant genotype

S46: Salt-sensitive genotype

GO: Gene Ontology

TF: Transcription Factor

ROS: Reactive oxygen species

ABA: Abscisic acid

RWC: Relative water content

FDR: False discovery rate

FUF1: Fyf Up-Regulating 321 Factor 1

DREB: Dehydration-responsive element-binding protein

ERF/AP2 : Apetala2/Ethylene Responsive Factor

ANCOVA: Analysis of covariance

SOS: Salt-overly-sensitive

IAA: Indole acetic acid

EC: Electrical conductivity

S: Salt treatment

G: Genotype

NB: negative binomial distribution

FPKM: Fragments per kilobase of transcript per million mapped reads

RCC: Rank of correlation coefficient

MR: Mutual Rank

RWR: Random walk with restart

TSS: Transcription start site

Acknowledgements

Seeds studied in this research were kindly provided by Seed and Plant Improvement Institute Karaj Iran; Razi University of Kermanshah; and Ferdowsi University of Mashhad.

Funding

The research was funded by the Office of Vice Chancellor for Research, Urmia University, Urmia, Iran (Project No. 94/101/T.T) and by grants of the Fonds Wetenschappelijk Onderzoek-Vlaanderen (FWO, Belgium) [3G046318, G.0371.06, 3G045620]. Razgar Seyed Rahmani holds a personal grant of the Ministry of Science Research and Technology, Iran.

Authors' contributions

Conceptualization: R.D, KM, R.S.R; Material preparation and data collection: T. M; Formal analysis and writing the original draft: R.S.R, T.M; Writing - review and editing: R.D, K.M; reviewing the paper: S.D; Supervision: R.D, S.D, K.M. All authors read and approved the final manuscript.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Munns R, Tester M. Mechanisms of salinity tolerance. Annu Rev Plant Biol. 2008;59:651-81.
2. Farooq M, Hussain M, Wakeel A, Siddique KH. Salt stress in maize: effects, resistance mechanisms, and management. A review. Agronomy for Sustainable Development. 2015;35(2):461-81.
3. Shiferaw B, Prasanna BM, Hellin J, Bänziger M. Crops that feed the world 6. Past successes and future challenges to the role played by maize in global food security. Food Security. 2011;3(3):307.
4. Chinnusamy V, Jagendorf A, Zhu J-K. Understanding and improving salt tolerance in plants. Crop science. 2005;45(2):437-48.
5. Serrano R, Rodriguez-Navarro A. Ion homeostasis during salt stress in plants. Current opinion in cell biology. 2001;13(4):399-404.
6. Van Zelm E, Zhang Y, Testerink C. Salt tolerance mechanisms of plants. Annual Review of Plant Biology. 2020;71:403-33.
7. Zhu J-K, Hasegawa PM, Bressan RA, Bohnert HJ. Molecular aspects of osmotic stress in plants. Critical Reviews in Plant Sciences. 1997;16(3):253-77.
8. Liang W, Ma X, Wan P, Liu L. Plant salt-tolerance mechanism: a review. Biochemical and biophysical research communications. 2018;495(1):286-91.
9. Zhu J-K. Cell signaling under salt, water and cold stresses. Current opinion in plant biology. 2001;4(5):401-6.
10. Mansour M, Salama K, Ali F, Abou Hadid A. Cell and plant responses to NaCl in Zea mays L. cultivars differing in salt tolerance. General and Applied Plant Physiology. 2005;31(1-2):29-41.
11. Fu H-H, Luan S. AtKUP1: a dual-affinity K+ transporter from Arabidopsis. The Plant Cell. 1998;10(1):63-73.
12. Isayenkov S. Physiological and molecular aspects of salt stress in plants. Cytology and Genetics. 2012;46(5):302-18.
13. Isayenkov SV, Maathuis FJ. Plant salinity stress: Many unanswered questions remain. Frontiers in plant science. 2019;10.
14. Ranathunge K, Schreiber L. Water and solute permeabilities of Arabidopsis roots in relation to the amount and composition of aliphatic suberin. Journal of experimental botany. 2011;62(6):1961-74.
15. Tsugane K, Kobayashi K, Niwa Y, Ohba Y, Wada K, Kobayashi H. A recessive Arabidopsis mutant that grows photoautotrophically under salt stress shows enhanced active oxygen detoxification. The Plant Cell. 1999;11(7):1195-206.
16. Rajendran K, Tester M, Roy SJ. Quantifying the three main components of salinity tolerance in cereals. Plant, cell & environment. 2009;32(3):237-49.
17. Boudsocq M, Droillard M-J, Barbier-Brygoo H, Laurière C. Different phosphorylation mechanisms are involved in the activation of sucrose non-fermenting 1 related protein kinases 2 by osmotic stresses and abscisic acid. Plant molecular biology. 2007;63(4):491-503.
18. Gu L, Liu Y, Zong X, Liu L, Li D-P, Li D-Q. Overexpression of maize mitogen-activated protein kinase gene, ZmSIMK1 in Arabidopsis increases tolerance to salt stress. Molecular biology reports. 2010;37(8):4067-73.

19. Liu W, Li R-J, Han T-T, Cai W, Fu Z-W, Lu Y-T. Salt stress reduces root meristem size by nitric oxide-mediated modulation of auxin accumulation and signaling in Arabidopsis. Plant physiology. 2015;168(1):343-56.
20. Yang Y, Guo Y. Elucidating the molecular mechanisms mediating plant salt-stress responses. New Phytologist. 2018;217(2):523-39.
21. Roy SJ, Negrão S, Tester M. Salt resistant crop plants. Current opinion in Biotechnology. 2014;26:115-24.
22. Gao Y, Lu Y, Wu M, Liang E, Li Y, Zhang D, et al. Ability to remove Na+ and retain K+ correlates with salt tolerance in two maize inbred lines seedlings. Frontiers in plant science. 2016;7:1716.
23. Sun W, Xu X, Zhu H, Liu A, Liu L, Li J, et al. Comparative transcriptomic profiling of a salt-tolerant wild tomato species and a salt-sensitive tomato cultivar. Plant and Cell Physiology. 2010;51(6):997-1006.
24. Emmert-Streib F, Glazko GV. Pathway analysis of expression data: deciphering functional building blocks of complex diseases. PLoS computational biology. 2011;7(5).
25. Li X, Li W, Zeng M, Zheng R, Li M. Network-based methods for predicting essential genes or proteins: a survey. Briefings in bioinformatics. 2020;21(2):566-83.
26. García-Ortega LF, Martínez O. How many genes are expressed in a transcriptome? Estimation and results for RNA-Seq. PLoS One. 2015;10(6).
27. De Maeyer D, Weytjens B, Renkens J, De Raedt L, Marchal K. PheNetic: network-based interpretation of molecular profiling data. Nucleic acids research. 2015;43(W1):W244-W50.
28. Kosová K, Vítámvás P, Urban MO, Klíma M, Roy A, Prášil IT. Biological networks underlying abiotic stress tolerance in temperate crops—a proteomic perspective. International journal of molecular sciences. 2015;16(9):20913-42.
29. Suravajhala P, Kogelman LJ, Kadarmideen HN. Multi-omic data integration and analysis using systems genomics approaches: methods and applications in animal production, health and welfare. Genetics Selection Evolution. 2016;48(1):38.
30. Tuncbag N, McCallum S, Huang S-sC, Fraenkel E. SteinerNet: a web server for integrating 'omic'data to discover hidden components of response pathways. Nucleic acids research. 2012;40(W1):W505-W9.
31. De Maeyer D, Renkens J, Cloots L, De Raedt L, Marchal K. PheNetic: network-based interpretation of unstructured gene lists in E. coli. Molecular BioSystems. 2013;9(7):1594-603.
32. Jiang Y, Liang Y, Wang D, Xu D, Joshi T. A dynamic programing approach to integrate gene expression data and network information for pathway model generation. Bioinformatics. 2020;36(1):169-76.
33. Szalai G, Janda T. Effect of salt stress on the salicylic acid synthesis in young maize (Zea mays L.) plants. Journal of agronomy and crop science. 2009;195(3):165-71.
34. Shahzad M, Witzel K, Zörb C, Mühling K. Growth-related changes in subcellular ion patterns in maize leaves (Zea mays L.) under salt stress. Journal of Agronomy and Crop Science. 2012;198(1):46-56.
35. Fukushima A, Kusano M, Redestig H, Arita M, Saito K. Integrated omics approaches in plant systems biology. Current opinion in chemical biology. 2009;13(5-6):532-8.
36. Pinu FR, Beale DJ, Paten AM, Kouremenos K, Swarup S, Schirra HJ, et al. Systems biology and multi-omics integration: Viewpoints from the metabolomics research community. Metabolites. 2019;9(4):76.

640 37. Yuan JS, Galbraith DW, Dai SY, Griffin P, Stewart Jr CN. Plant systems biology comes of age.
641 Trends in plant science. 2008;13(4):165-71.
642 38. de Azevedo Neto AD, Prisco JT, Enéas-Filho J, Medeiros J-VR, Gomes-Filho E. Hydrogen
643 peroxide pre-treatment induces salt-stress acclimation in maize plants. Journal of Plant Physiology.
644 2005;162(10):1114-22.
645 39. Hichem H, Mounir D. Differential responses of two maize (Zea mays L.) varieties to salt stress:
646 changes on polyphenols composition of foliage and oxidative damages. Industrial Crops and Products.
647 2009;30(1):144-51.
648 40. Singh PK, Singh R, Singh S. Cinnamic acid induced changes in reactive oxygen species
649 scavenging enzymes and protein profile in maize (Zea mays L.) plants grown under salt stress.
650 Physiology and Molecular Biology of Plants. 2013;19(1):53-9.
651 41. Knight H, Trewavas AJ, Knight MR. Calcium signalling in Arabidopsis thaliana responding to
652 drought and salinity. The Plant Journal. 1997;12(5):1067-78.
653 42. Ruiz J, Blumwald E. Salinity-induced glutathione synthesis in Brassica napus. Planta.
654 2002;214(6):965-9.
655 43. Zahir Z, Shah MK, Naveed M, Akhter MJ. Substrate-dependent auxin production by Rhizobium
656 phaseoli improves the growth and yield of Vigna radiata L. under salt stress conditions. J Microbiol
657 Biotechnol. 2010;20(9):1288-94.
658 44. Nazar R, Umar S, Khan NA. Exogenous salicylic acid improves photosynthesis and growth
659 through increase in ascorbate-glutathione metabolism and S assimilation in mustard under salt stress.
660 Plant signaling & behavior. 2015;10(3):e1003751.
661 45. Shafi A, Chauhan R, Gill T, Swarnkar MK, Sreenivasulu Y, Kumar S, et al. Expression of SOD
662 and APX genes positively regulates secondary cell wall biosynthesis and promotes plant growth and yield
663 in Arabidopsis under salt stress. Plant Molecular Biology. 2015;87(6):615-31.
664 46. Ngara R, Ndimba R, Borch-Jensen J, Jensen ON, Ndimba B. Identification and profiling of
665 salinity stress-responsive proteins in Sorghum bicolor seedlings. Journal of Proteomics.
666 2012;75(13):4139-50.
667 47. Zhang J, Li Y, Liu B, Wang L, Zhang L, Hu J, et al. Characterization of the Populus Rab family
668 genes and the function of PtRabE1b in salt tolerance. BMC plant biology. 2018;18(1):124.
669 48. Bolte S, Schiene K, Dietz K-J. Characterization of a small GTP-binding protein of the rab 5 family
670 in Mesembryanthemum crystallinum with increased level of expression during early salt stress. Plant
671 molecular biology. 2000;42(6):923-35.
672 49. Mirzaei M, Pascovici D, Atwell BJ, Haynes PA. Differential regulation of aquaporins, small GTP
673 ases and V-ATP ases proteins in rice leaves subjected to drought stress and recovery. Proteomics.
674 2012;12(6):864-77.
675 50. O'Malley RC, Huang S-sC, Song L, Lewsey MG, Bartlett A, Nery JR, et al. Cistrome and
676 epicistrome features shape the regulatory DNA landscape. Cell. 2016;165(5):1280-92.
677 51. Ma S, Gong Q, Bohnert HJ. Dissecting salt stress pathways. Journal of experimental botany.
678 2006;57(5):1097-107.
679 52. Walley JW, Sartor RC, Shen Z, Schmitz RJ, Wu KJ, Urich MA, et al. Integration of omic networks
680 in a developmental atlas of maize. Science. 2016;353(6301):814-8.
681 53. Li H-D, Menon R, Omenn GS, Guan Y. The emerging era of genomic data integration for
682 analyzing splice isoform function. Trends in Genetics. 2014;30(8):340-7.
683 54. Munns R, Tester M. Mechanisms of salinity tolerance. Annu Rev Plant Biol. 2008;59:651-81.

55. Ji H, Pardo JM, Batelli G, Van Oosten MJ, Bressan RA, Li X. The Salt Overly Sensitive (SOS) pathway: established and emerging roles. Molecular plant. 2013;6(2):275-86.
56. Sairam R, Tyagi A. Physiology and molecular biology of salinity stress tolerance in plants. Current science. 2004;86(3):407-21.
57. Zhu J-K. Plant salt tolerance. Trends in plant science. 2001;6(2):66-71.
58. Choudhury S, Panda P, Sahoo L, Panda SK. Reactive oxygen species signaling in plants under abiotic stress. Plant signaling & behavior. 2013;8(4):e23681.
59. Liu Q, Luo L, Zheng L. Lignins: biosynthesis and biological functions in plants. International journal of molecular sciences. 2018;19(2):335.
60. Andrews S. FastQC: a quality control tool for high throughput sequence data. Babraham Bioinformatics, Babraham Institute, Cambridge, United Kingdom; 2010.
61. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30(15):2114-20.
62. Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nature biotechnology. 2019;37(8):907-15.
63. Anders S, Pyl PT, Huber W. HTSeq—a Python framework to work with high-throughput sequencing data. Bioinformatics. 2015;31(2):166-9.
64. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2010;26(1):139-40.
65. Yu G, Wang L-G, Han Y, He Q-Y. clusterProfiler: an R package for comparing biological themes among gene clusters. Omics: a journal of integrative biology. 2012;16(5):284-7.
66. Proost S, Van Bel M, Sterck L, Billiau K, Van Parys T, Van de Peer Y, et al. PLAZA: a comparative genomics resource to study gene and genome evolution in plants. The Plant Cell. 2009;21(12):3718-31.
67. Zhu G, Wu A, Xu X-J, Xiao P-P, Lu L, Liu J, et al. PPIM: a protein-protein interaction database for maize. Plant physiology. 2016;170(2):618-26.
68. Kanehisa M, Goto S. KEGG: Kyoto Encyclopedia of Genes and Genomes Nucleic Acids Res. January. 2000;1(28):1.
69. Zhang JD, Wiemann S. KEGGgraph: a graph approach to KEGG PATHWAY in R and bioconductor. Bioinformatics. 2009;25(11):1470-1.
70. Cock PJ, Antao T, Chang JT, Chapman BA, Cox CJ, Dalke A, et al. Biopython: freely available Python tools for computational molecular biology and bioinformatics. Bioinformatics. 2009;25(11):1422-3.
71. Tian F, Yang D-C, Meng Y-Q, Jin J, Gao G. PlantRegMap: charting functional regulatory maps in plants. Nucleic acids research. 2020;48(D1):D1104-D13.
72. Jiao Y, Peluso P, Shi J, Liang T, Stitzer MC, Wang B, et al. Improved maize reference genome with single-molecule technologies. Nature. 2017;546(7659):524-7.
73. Hoopes GM, Hamilton JP, Wood JC, Esteban E, Pasha A, Vaillancourt B, et al. An updated gene atlas for maize reveals organ-specific and stress-induced genes. The Plant Journal. 2019;97(6):1154-67.
74. Obayashi T, Kinoshita K. Rank of correlation coefficient as a comparable measure for biological significance of gene coexpression. DNA research. 2009;16(5):249-60.
75. Nepusz T, Yu H, Paccanaro A. Detecting overlapping protein complexes in protein-protein interaction networks. Nature methods. 2012;9(5):471.
76. Valdeolivas A, Tichit L, Navarro C, Perrin S, Odelin G, Levy N, et al. Random walk with restart on multiplex and heterogeneous biological networks. Bioinformatics. 2019;35(3):497-505.

77. Verbeke LP, Van den Eynden J, Fierro AC, Demeester P, Fostier J, Marchal K. Pathway relevance ranking for tumor samples through network-based data integration. PloS one. 2015;10(7).
78. Bailey TL, Johnson J, Grant CE, Noble WS. The MEME suite. Nucleic acids research. 2015;43(W1):W39-W49.

Supplementary information

Additional file 1: List of differentially expressed genes between sensitive and tolerant genotypes under salt stress after correcting for differences under control conditions (log2 fold change and p-values).

Additional file 2: Heatmap showing the expression values (log2 FPKM) of the candidate genes.

Additional file 3: Integrated interaction network form KEGG, protein-protein, regulatory and coexpression edges.

Figures

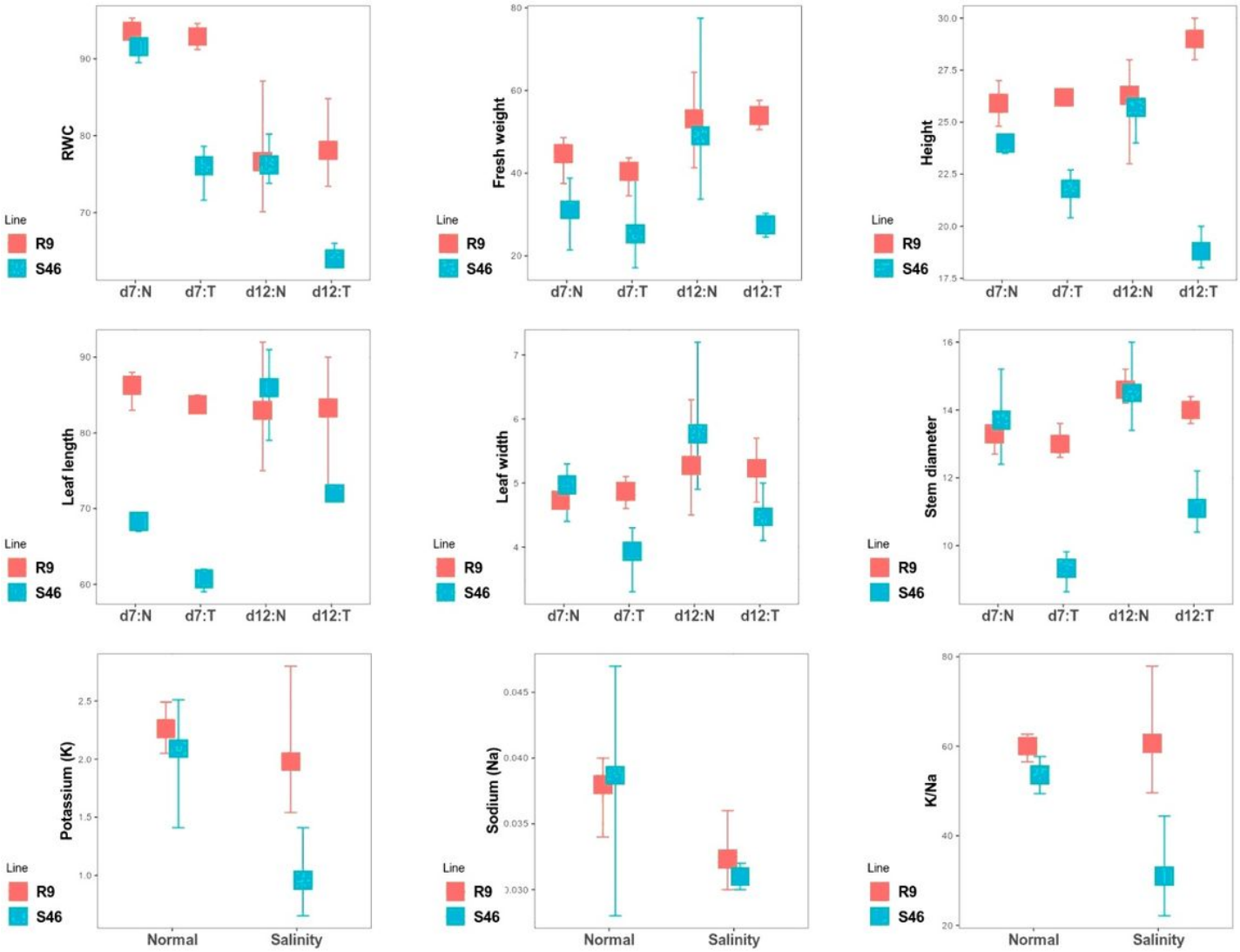


Figure 1

Comparing morphological and physiological traits between R9 and S46 under normal and salinity imposed conditions (values obtained from samples taken at 7 and 12 days' post-salt treatment). The boxes show the mean based on three biological repeats and bars represent the 95 percent confidence interval for the means. Potassium and sodium content measured 12 days after salt treatment. Abbreviations: d7: N indicates day 7 under control condition; d7:T indicates day 7 after salt treatment; d12:N indicates day 12 under control, and d12:T indicates day 12 under salt treatment, RWC: relative water content.

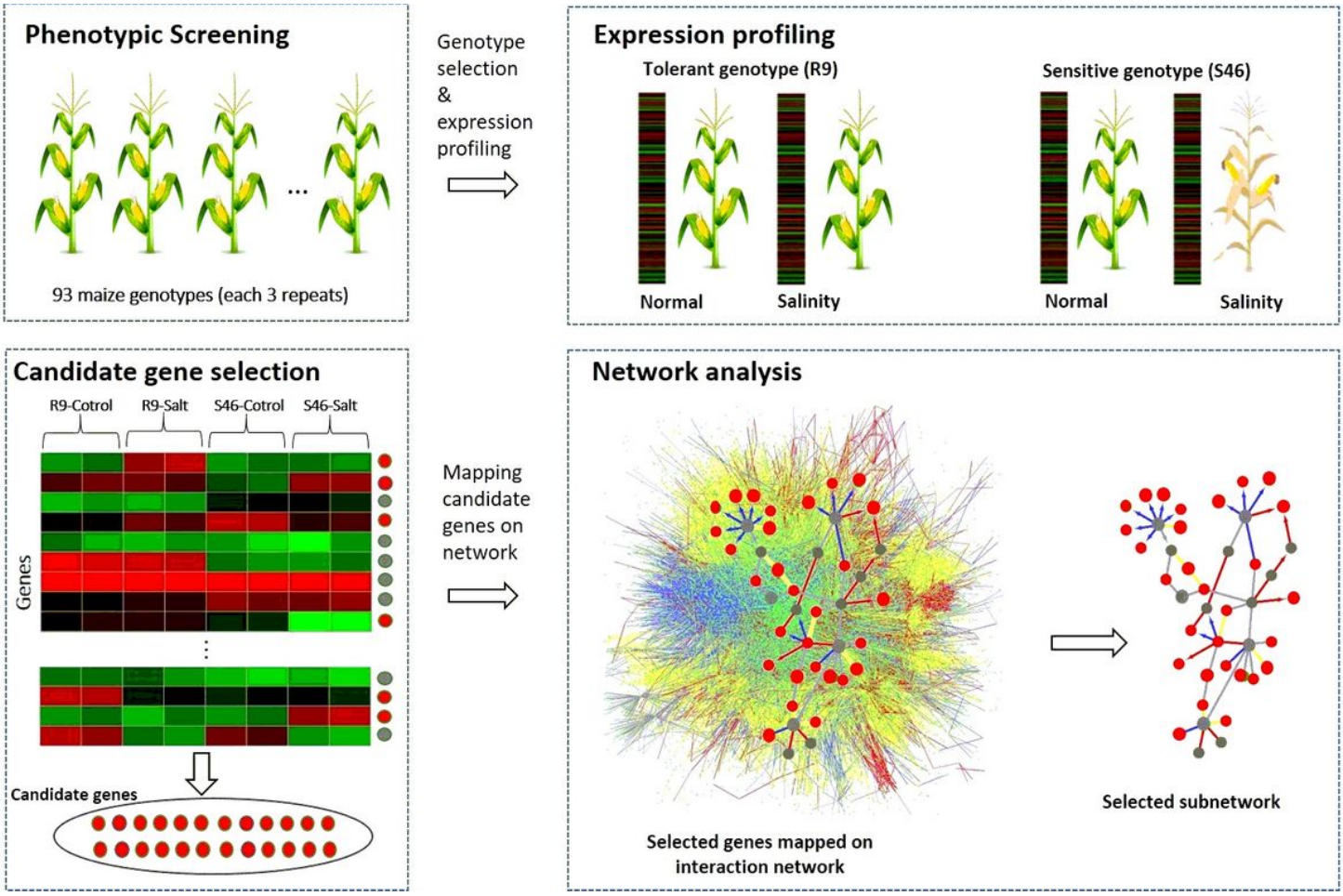


Figure 2

The flow of mapping candidate genes resulting from the expression analysis on the interaction network. Phenotype screening: from 93 maize genotypes, two genotypes with an extremely different phenotypic behavior towards salt stress (R9:Tolerant, S46:Sensitive) were selected. Expression profiling: gene expression of the R9 and S46 genotypes profiled under normal and salinity imposed conditions. Candidate gene selection: the genes that responded differently to salinity stress between two genotypes are selected as candidate genes (red genes on the right hand of heatmap). Network analysis: an integrated interaction network was compiled from different sources (regulatory, protein-protein, metabolic and coexpression networks) and was subsequently converted to a probabilistic interaction network. Candidate genes were mapped on this probabilistic interaction network and subnetworks extracted using Phenetic. Red nodes are candidate genes (identified as differentially expressed in a genotype specific way) and grey nodes indicate connector genes recovered by the network analysis.

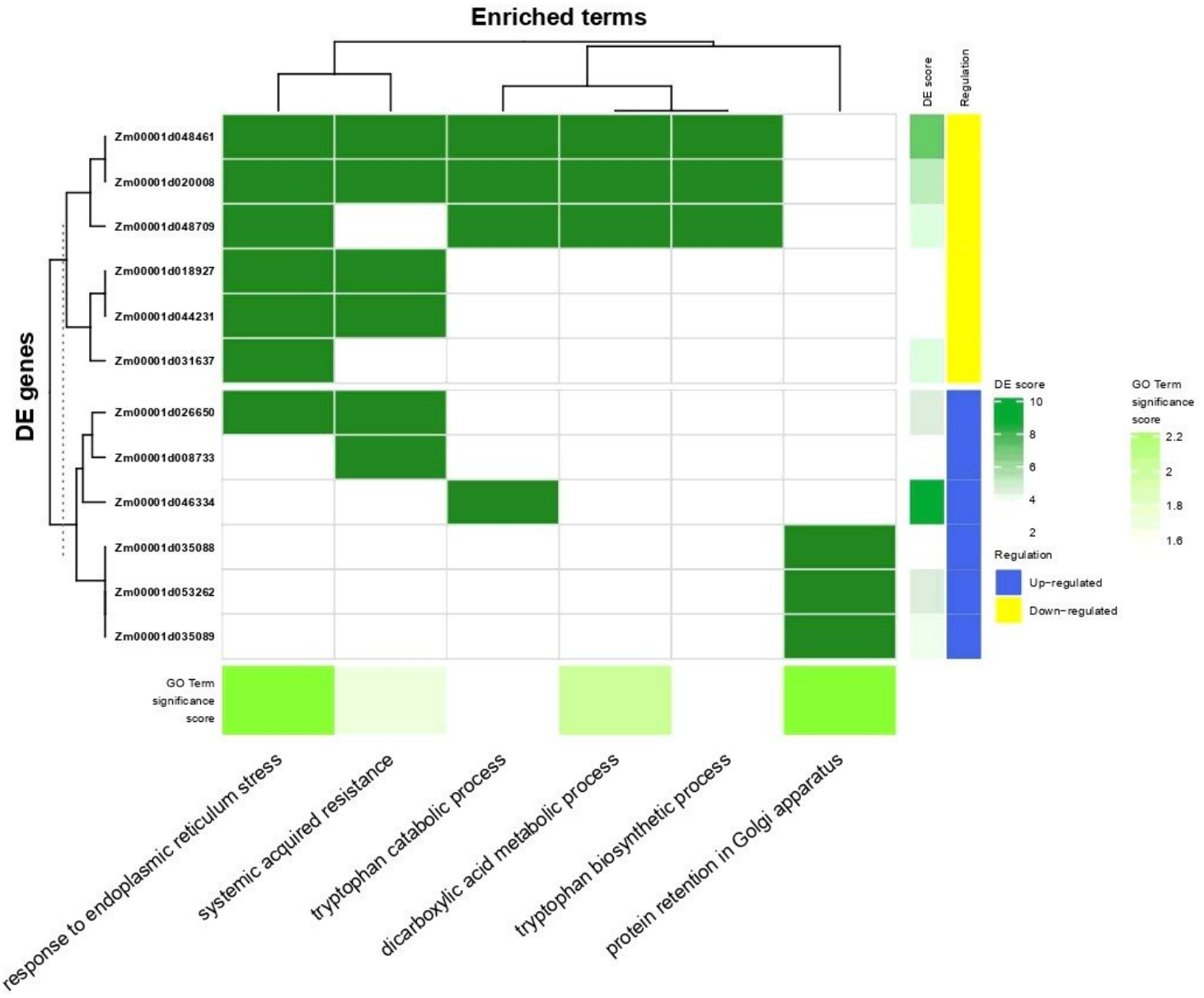


Figure 3

GO enrichment for the candidate genes. The overrepresented GO term and the candidate genes are shown on respectively the x-axis and y-axis. Green indicates that the corresponding gene is present in the GO class, the DE score reflects the degree to which the candidate gene is differentially expressed (log fold change), Yellow and blue indicate whether the gene was up versus down regulated under salt stress comparing S46 to S9 after correcting for differences in expression under normal conditions.

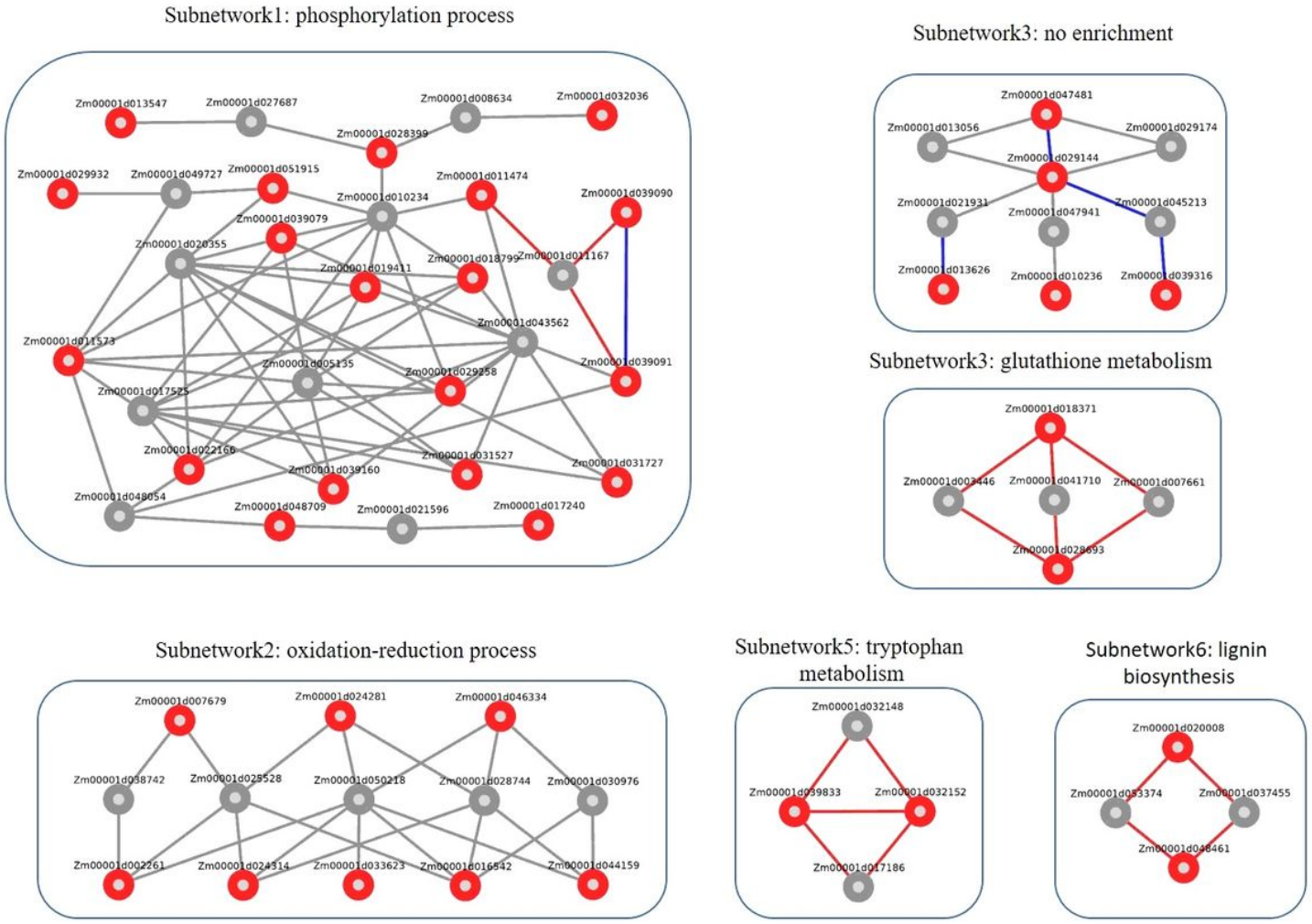


Figure 4

Results of the network analysis. Red nodes represent candidate genes obtained from the expression analysis, grey nodes are not identified as differentially expressed, but were identified by the network analysis as connector genes (genes needed to connect the candidate genes). Edge color: red indicates a metabolic interaction, grey a protein-protein interaction, and blue a co-expression derived interaction. No regulatory interactions were recovered in any of the sub-networks

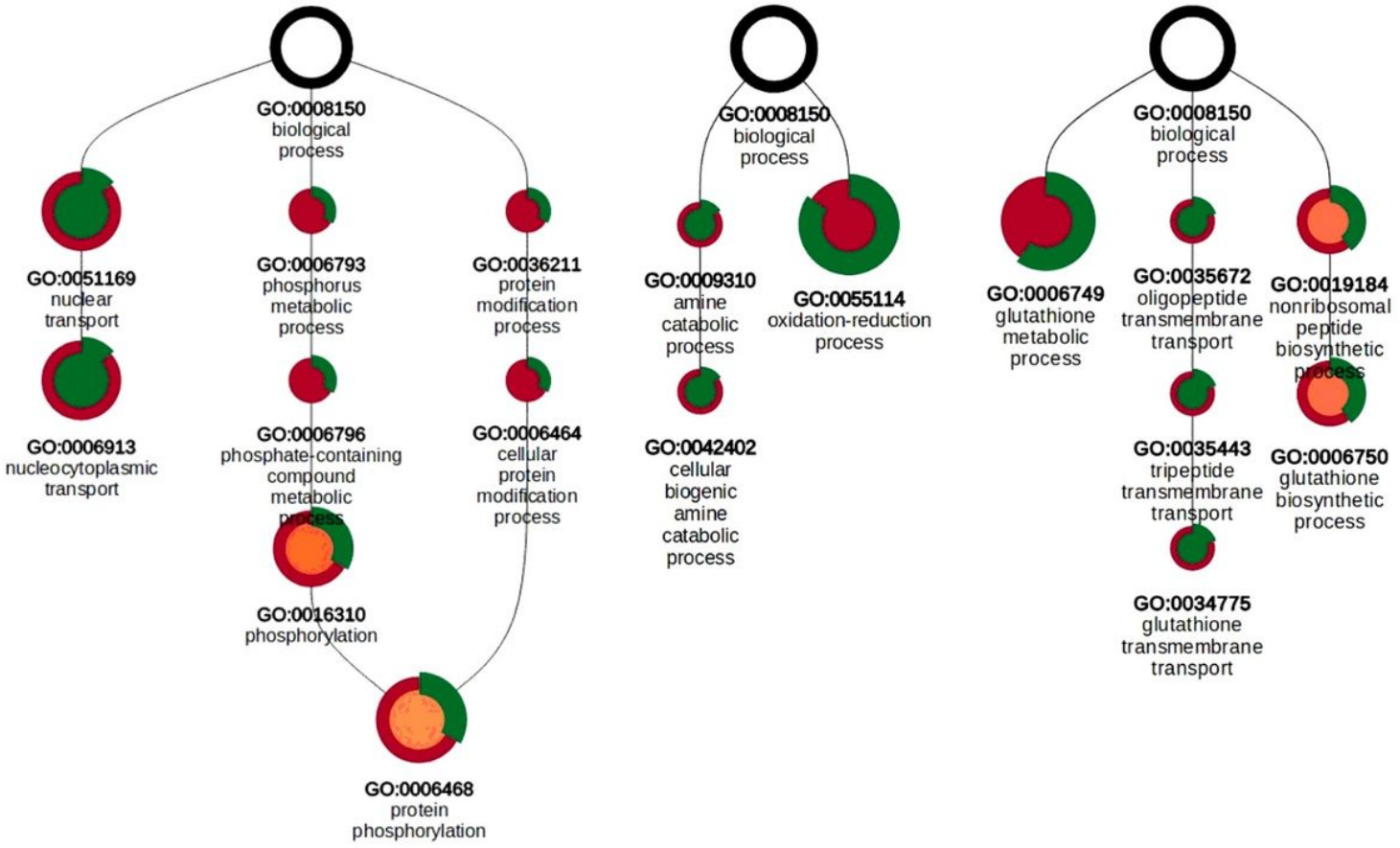


Figure 5

The hierarchy of the GO enrichment result for sub-network 1 (left), sub-network 2 (middle) and sub-network 4 (right). Sub-network3 was not enriched for any specific biological process and the GO results for sub-network 5 and 6 are not shown as they consist of four genes only, all of which are annotated in KEGG metabolic pathways. Node size is scaled by the Bonferroni corrected p-value for enrichment; node color is determined by the enrichment fold such that green shows the highest and red shows the lowest fold enrichment; the node outer band reflects the 'Percentage Present' i.e. the percentage of genes that are annotated with the enriched ontology term (indicated by the green part of the ring).

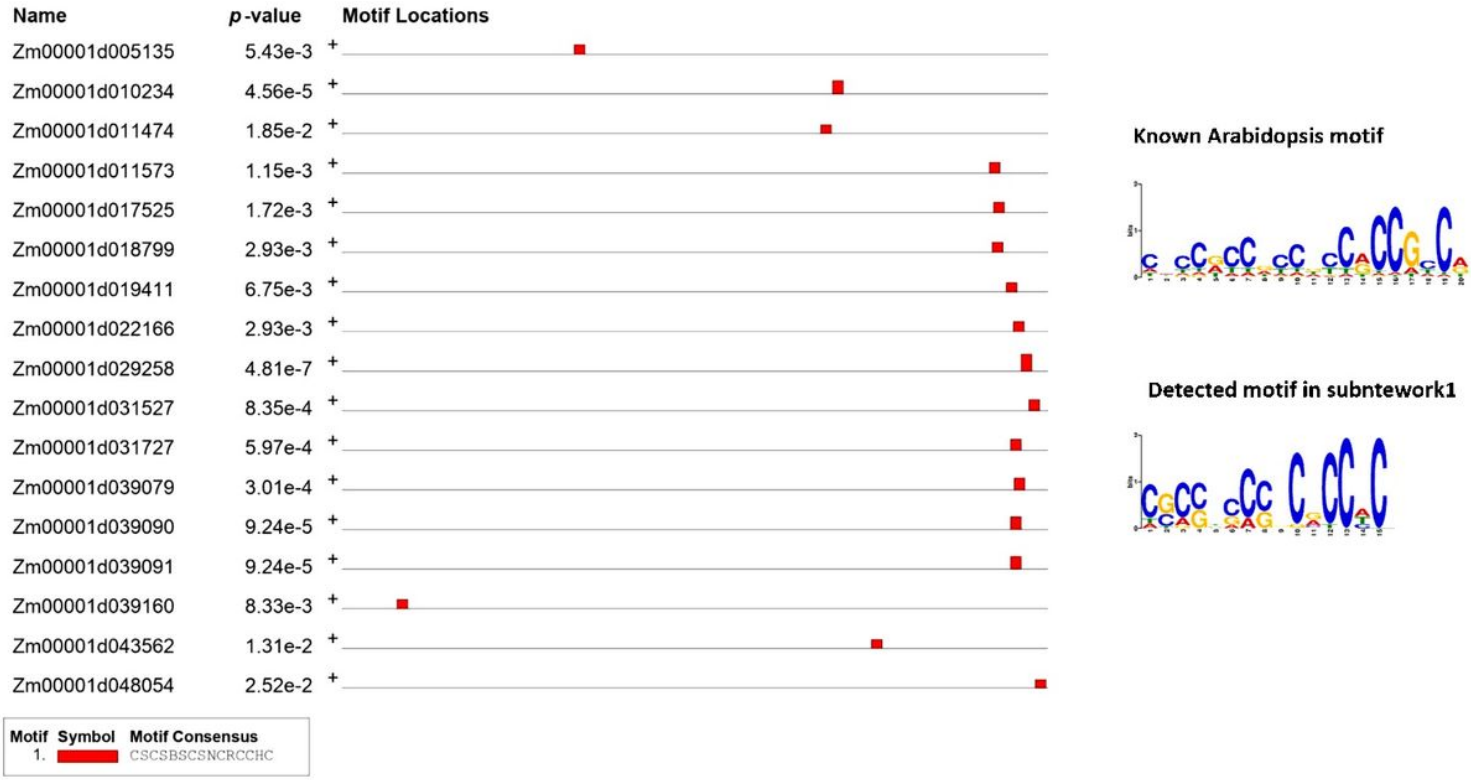


Figure 6

Regulatory element overrepresented in the promoter sequences of the genes in sub-network 1. The identified motif shows a strong bias towards the TSS and high similarity with a known Arabidopsis motif representative for the FuF1 binding site (q-value 2.32e-04). The first logo is the Arabidopsis cis-regulatory element obtained from a DAP experiment and the second one represents the de novo detected motif in subnetwork1.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Additionalfile1.xlsx
- Additionalfile2.jpeg
- Additionalfile3.txt