BMC Plant Biology

RESEARCH Open Access

Water deficit response in nodulated soybean roots: a comprehensive transcriptome and translatome network analysis

María Martha Sainz[1*], Carla V. Filippi[1], Guillermo Eastman[2,3], Mariana Sotelo-Silveira[1], Sofía Zardo[1], Mauro Martínez-Moré[1], José Sotelo-Silveira[2,4*] and Omar Borsani[1*]

Abstract

Background Soybean establishes a mutualistic interaction with nitrogen-fixing rhizobacteria, acquiring most of its nitrogen requirements through symbiotic nitrogen fixation. This crop is susceptible to water deficit; evidence suggests that its nodulation status—whether it is nodulated or not—can influence how it responds to water deficit. The translational control step of gene expression has proven relevant in plants subjected to water deficit.

Results Here, we analyzed soybean roots' differential responses to water deficit at transcriptional, translational, and mixed (transcriptional + translational) levels. Thus, the transcriptome and translatome of four combined-treated soybean roots were analyzed. We found hormone metabolism-related genes among the differentially expressed genes (DEGs) at the translatome level in nodulated and water-restricted plants. Also, weighted gene co-expression network analysis followed by differential expression analysis identified gene modules associated with nodulation and water deficit conditions. Protein-protein interaction network analysis was performed for subsets of mixed DEGs of the modules associated with the plant responses to nodulation, water deficit, or their combination.

Conclusions Our research reveals that the stand-out processes and pathways in the before-mentioned plant responses partially differ; terms related to glutathione metabolism and hormone signal transduction (2 C protein phosphatases) were associated with the response to water deficit, terms related to transmembrane transport, response to abscisic acid, pigment metabolic process were associated with the response to nodulation plus water deficit. Still, two processes were common: galactose metabolism and branched-chain amino acid catabolism. A comprehensive analysis of these processes could lead to identifying new sources of tolerance to drought in soybean.

Keywords Soybean stress response, Symbiosis, Root, Transcriptome, Translatome, Gene networks

*Correspondence:
María Martha Sainz
msainz@fagro.edu.uy
José Sotelo-Silveira
jsotelosilveira@iibce.edu.uy
Omar Borsani
oborsani@fagro.edu.uy

[1]Laboratorio de Bioquímica, Departamento de Biología Vegetal, Facultad de Agronomía, Universidad de la República, Avenida Garzón 780, Montevideo CP 12900, Uruguay
[2]Departamento de Genómica, Instituto de Investigaciones Biológicas Clemente Estable, MEC, Av. Italia 3318, Montevideo CP 11600, Uruguay
[3]Department of Biology, University of Virginia, 485 McCormick Rd, Charlottesville, VA 22904, USA
[4]Departamento de Biología Celular y Molecular, Facultad de Ciencias, Universidad de la República, Iguá, Montevideo 4225, CP 11400, Uruguay

Background

Leguminous plants such as soybean (Glycine max (L.) Merr.) can establish a root nodule endosymbiosis with nitrogen-fixing soil rhizobacteria. This mutualistic interaction induces significant metabolic and nutritional changes in the plant [1, 2]. The symbiotic nodule, the newly developed root organ due to the symbiotic pathway, is where nitrogen fixation and assimilation occur—a process known as symbiotic nitrogen fixation (SNF). The main products of SNF, ureid or amide compounds in determinate or indeterminate nodules, respectively, are exported to the rest of the plant, which in turn provides the rhizobia with photoassimilates [3]. Since developing and maintaining nodules is resource-consuming, the plant exerts tight control over the nodulation process under favorable and sub-optimal growing conditions [4].

Evidence suggests that the nodulation condition of a legume (i.e., nodulated or non-nodulated) can affect its response strategies to water deficit [5–9]; however, it is uncertain which molecular mechanisms are responsible for the differential response. The regulation of gene expression, which can be achieved at the transcriptional and/or post-transcriptional levels (including translational and post-translational events), could explain the plant differential response strategies previously mentioned. In particular, translational control has proven relevant in plants subjected to stressful conditions such as nutrient scarcity [10] or different situations of biotic and abiotic stresses like water deficit [11, 12]. Plants benefit from this step of gene expression regulation, which does not require de novo messenger RNA (mRNA) synthesis but rather refers to the efficiency with which mRNAs already present in cells are translated since it allows them to respond rapidly, thus conferring flexibility and adaptability [13, 14]. At present, it is well established that the poor or variable levels of correlation between the levels of transcripts and proteins found in different organisms is explained by the post-transcriptional steps of gene expression, mainly translation [12, 15–17]. This way, the direct analysis of the subset of mRNAs that are being translated (the translatome) enables more accurate and complete measurement of cell gene expression than the one obtained when only the transcriptome (steady-state mRNA levels) is analyzed [18]. Still, the information gathered from a transcriptomic and a translatomic analysis is complementary, and their comparison allows, for instance, to distinguish between the different levels of regulation of each mRNA.

Despite how advantageous translational control is for plants subjected to stressful conditions, studies that involve a translatomic analysis of nodulated and water-restricted plants are scarce. As an interesting example, we can mention the work recently published by our group, in which we reported that some members of the thioredoxin and glutaredoxin systems were regulated mainly at the translational level in the roots of nodulated soybean plants subjected to water-deficit stress [19], thus resignifying the relevance of these enzymes for having a unique role in nodulated plants subjected to water restriction.

The plant roots are the organs first exposed to water deficit and are thus responsible for sensing water shortage, a main constraint for crop production, and transmitting stress signals to the rest of the plant. This causes developmental, physiological, and metabolic changes in adaptation to the water deficit, leading to the acquisition of resistance at the whole-plant level [20, 21]. The responses of plant roots, primarily those of crop plants, have therefore been the subject of research over the past few decades, and the selection of plants with root traits that improve productivity under drought is of great relevance for geneticists and breeders [22].

This study aimed to analyze the differential responses of roots of nodulated soybean plants to water deficit by evaluating gene expression regulation at the transcriptional, translational, or mixed (transcriptional + translational) levels. Thus, the transcriptome (total RNA fraction) and translatome (polysome-associated RNA fraction) of four combined-treated soybean roots (including the nodulation and water deficit conditions) were analyzed. Genes encoding various enzymes involved in hormone metabolism were found among the differentially expressed genes (DEGs) mainly regulated at the translatome level in nodulated and water-restricted plants. Further, we identified gene modules associated with the nodulation and/or water deficit conditions of soybean plant roots through a weighted gene co-expression network analysis (WGCNA) followed by a differential expression analysis [23]. Gene modules more representative of the DEGs were subjected to enrichment analysis, and selected modules were further analyzed. Protein-protein interaction network analysis was performed for subsets of DEGs of the modules more associated with the plant responses to nodulation, water deficit, or the combination of nodulation + water deficit, highlighting the standout biological processes and metabolic pathways in the before-mentioned plant responses.

Results

Scope of the experimental design and data quality

Our experimental design comprised soybean plants subjected to four treatments as a result of combining two nodulation conditions, i.e., nodulated (N) and non-nodulated (NN) plants, with two hydric conditions, i.e., water-restricted (WR) and well-watered (WW) plants. While this study aimed to analyze the responses of nodulated and water-restricted plants, this comprehensive experimental design allowed us to perform several comparisons

between the before-mentioned treatments to investigate the distinctive responses of plants to the different nodulation and hydric conditions.

In this study, we analyzed the following four comparisons: *i*) N+WR vs. N+WW; *ii*) N+WR vs. NN+WR; *iii*) NN+WR vs. NN+WW; *iv*) N+WW vs. NN+WW (Fig. 1). The N+WR vs. N+WW (*i*) comparison exhibits the response of nodulated plants to water deficit; meanwhile, the *ii* comparison (N+WR vs. NN+WR) shows the particular response to water deficit of nodulated plants with respect to non-nodulated plants. The NN+WR vs. NN+WW (*iii*) comparison evidence the response of non-nodulated plants to water restriction, and the *iv* comparison (N+WW vs. NN+WW) shows the plant responses due to the nodulation process without involving water restriction.

The four comparisons were studied at the transcriptome (total RNA fraction; TOTAL) and the Polysome-Associated RNA level (PAR fraction) to analyze the plant responses to the different nodulation and hydric conditions in a comparative manner and to highlight changes at the loading of mRNA in the translational apparatus versus total RNA levels. The latter allows to assess regulation at the transcriptome (TOTAL) and translatome (PAR) levels as well as mixed responses (TOTAL+PAR) involving those compartments upon the conditions compared.

Exploration of the RNA-seq data using distance matrix analysis (heatmap) and principal component analysis (PCA) showed that all the biological replicates clustered together (Figure S1, A and B). Also, the PCA evidenced that component 1 (PC1) explained the greatest proportion of the variance (36%) separating the samples by hydric condition (pink vs. green). Component 2 (PC2) separated the samples by the nodulation condition (circles vs. triangles). Thus, four well-defined groups were obtained, corresponding to the four combined treatments. As expected, each sample's TOTAL and PAR RNA fractions were found to be very close to each other (Figure S1, B).

The metabolism of several hormones is regulated at the translational level in nodulated and water-restricted plants

The results of the exploratory analysis of the RNA-seq data suggested that the data obtained were suitable for further downstream analysis. In this line, to learn how the nodulation condition and the water-restriction condition affect the plant's transcriptome and translatome, an initial differentially expressed gene (DEG) analysis was performed contrasting nodulated samples with respect to non-nodulated samples (including WR and WW plants) on one side (Fig. 2, A and B), and water-restricted samples with respect to well-watered samples (including N and NN plants) on the other (Fig. 2, C and D). Overall, water restriction significantly impacted gene expression at the transcriptome and translatome level more than nodulation. In the N vs. NN comparison, most DEGs were upregulated, while in the WR vs. WW comparison, more DEGs were downregulated at both TOTAL and PAR levels (Fig. 2, A, B, C, D).

To further identify DEGs with translational regulation in each of the four comparisons between combined treatments analyzed, scatter plots showing the fold change (FC) in the corresponding TOTAL and PAR samples were done (Fig. 2, E, F, G, H). Interestingly, the set of DEGs with the highest FC regulated mainly at the PAR level in nodulated and water-restricted plants (comparisons *i* and *ii*; Fig. 2, E, F; Table S2) includes many genes related to the metabolism of the hormones abscisic acid (ABA), ethylene, auxin and cytokinin such as those encoding the zeaxanthin epoxidase, non-race-specific disease resistance/hairpin-induced (NDR1/HIN1)-like protein 6, 1-aminocyclopropoane-1-carboxylate (ACC) synthase 3, indole-3-acetic acid (IAA)-amido synthetase GH3.10, indole-3-acetate O-methyltransferase 1 (IAMT1), and adenylate isopentenyltransferase 5 (IPT5). Given the key role of these hormones in the plant response mechanisms to the processes under investigation (nodulation and water restriction), the fact that these genes were differentially expressed at the translational level in N+WR plants is highly pertinent. Genes encoding amino acids and sucrose transporters, cell wall proteins and enzymes related to its modification, plant receptor-like kinases (RLKs), protein phosphatases 2 C, heat shock proteins (HSPs), and transcription factors (WRKY 72 A, MADS-box 23, bHLH18, WUSCHEL-related homeobox 4) were also included among those DEGs mainly regulated at the PAR level in N+WR plants (Fig. 2, E, F; Table S2). Examples of DEGs solely regulated at the translational level in comparison *iii* (the one not including nodulated plants) comprised those encoding late embryogenesis abundant (LEA) proteins, heat shock proteins (HSP), inositol phosphate synthase—a rate-limiting enzyme in myo-inositol biosynthesis—transcription factors, and ABA and auxin metabolic enzymes (Fig. 2, G; Table S2). Comparison *iv* (which only involved the nodulation condition) had among its DEGs mainly regulated at the PAR level those coding for enzymes related to cell wall reconstruction or modification, cytokinins and phenylpropanoid biosynthesis, and amino acids transporters, among others (Fig. 2, H; Table S2).

Moreover, the presence of some genes in nodulated and water restricted plants (particularly highlighted in comparison *ii*) with a notable transcriptional upregulation (Up-TOTAL; the ones with the higher FC at the TOTAL level) is another intriguing aspect depicted in Fig. 2 (Fig. 2, F; Table S2). Strikingly, despite their transcriptional upregulation, these transcripts seem to be

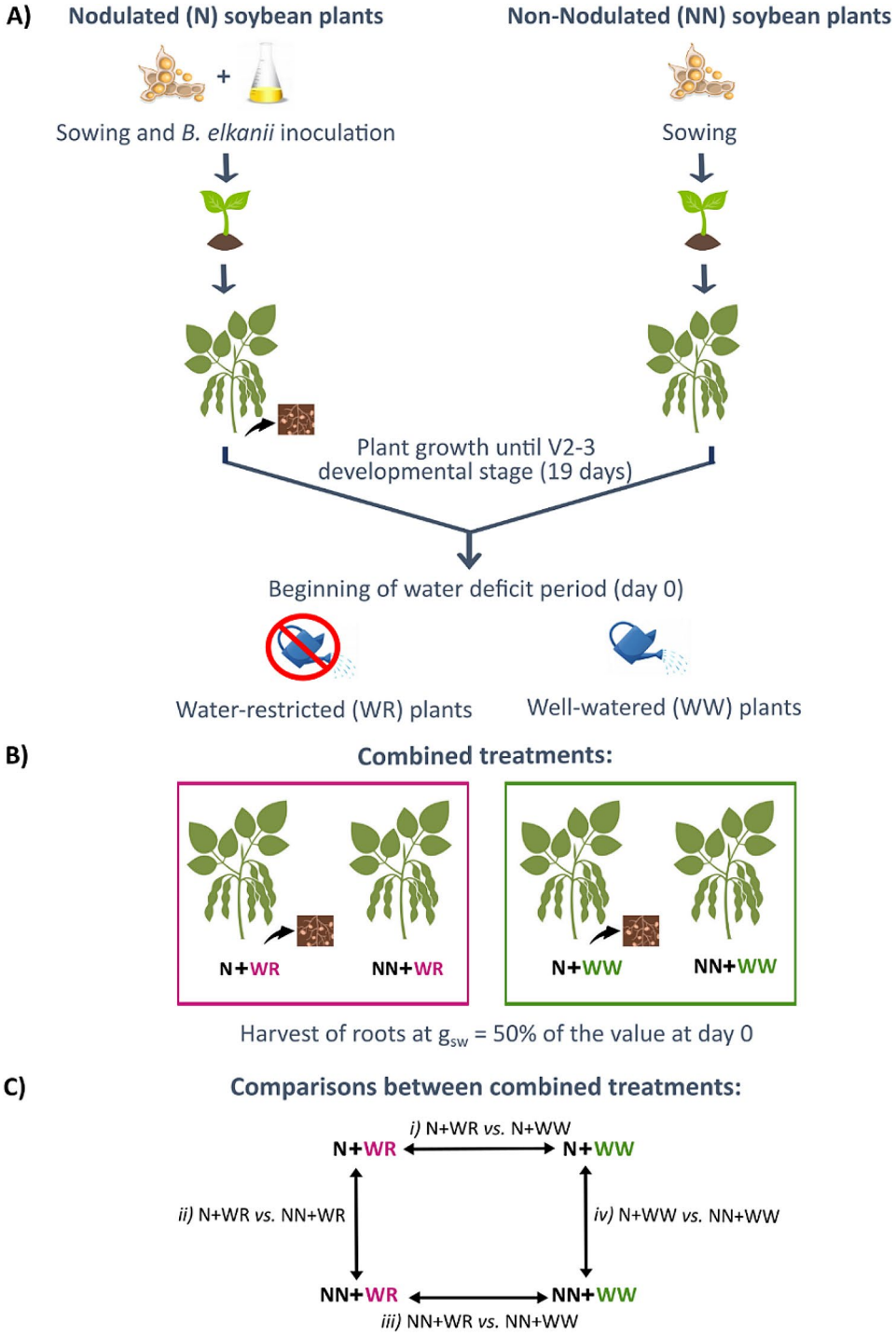

Fig. 1 Schematic overview of the methodology followed for the obtention of the combined treated soybean roots. The comparisons among the combined treatments are also shown. **(A)** The scheme gives an overview of the steps followed from plant growth, inoculation (only for nodulated plants), and the establishment of the water deficit stress assay (only for water-restricted plants). **(B)** As a result, soybean plants subjected to four combined treatments were obtained: nodulated and water-restricted plants (N + WR), non-nodulated and water-restricted plants (NN + WR), nodulated and well-watered plants (N + WW), non-nodulated and well-watered plants (NN + WW). **(C)** Four comparisons between combined treatments were analyzed: *i*) N + WR vs. N + WW; *ii*) N + WR vs. NN + WR; *iii*) NN + WR vs. NN + WW; *iv*) N + WW vs. NN + WW
N: nodulated; NN: non-nodulated; WR: water-restricted; WW: well-watered; g_{sw}: stomatal conductance

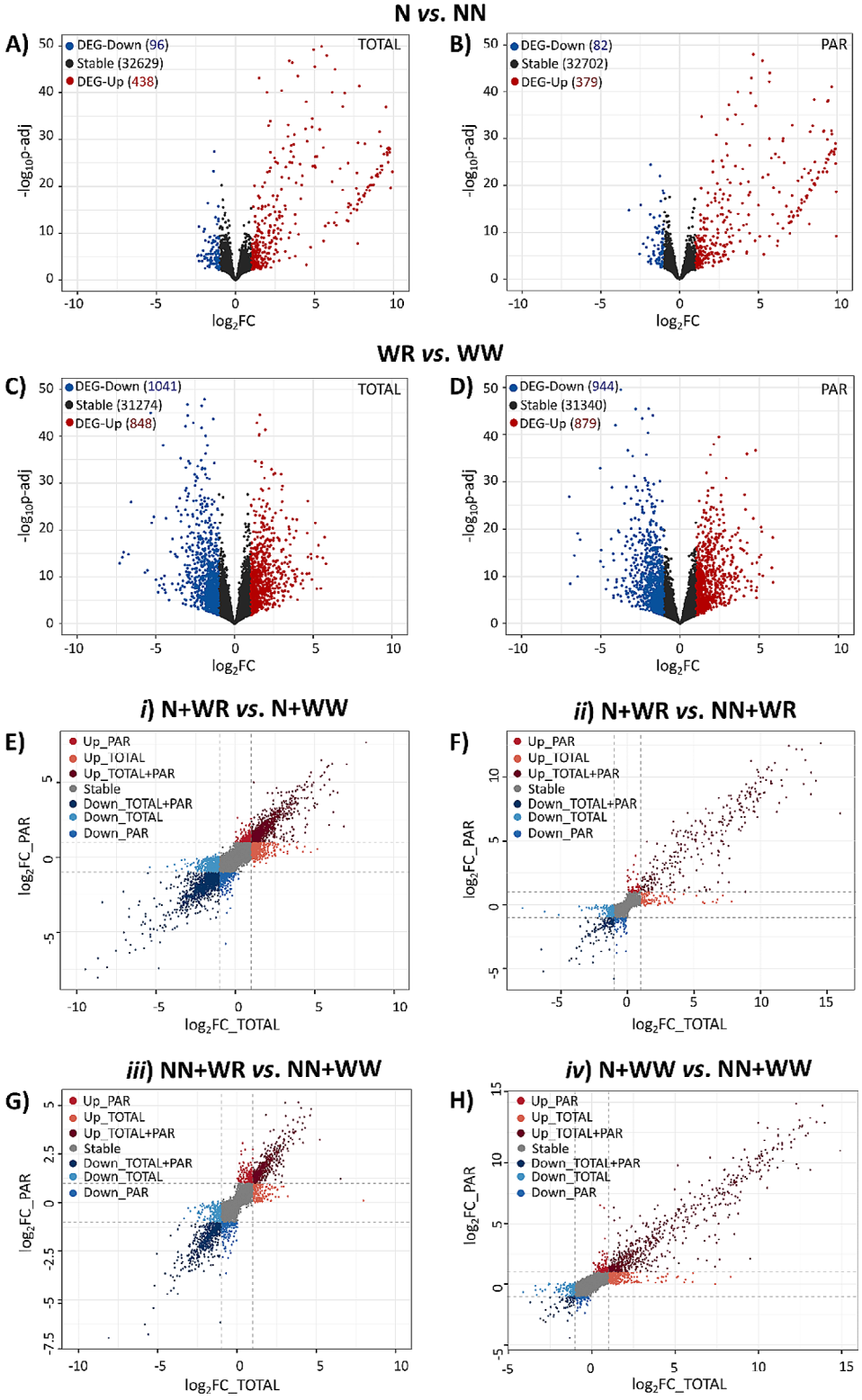

Fig. 2 Transcriptional and translational responses of soybean roots to nodulation and water restriction. Volcano plots showing stable genes, down- and up-regulated DEGs in the comparison between nodulated vs. non-nodulated plants at the TOTAL **(A)** and PAR **(B)** levels. Volcano plots showing stable genes, down- and up-regulated DEGs in the comparison between water-restricted vs. well-watered plants at the TOTAL **(C)** and PAR **(D)** levels. Scatterplots showing the FC in the TOTAL and PAR samples in the comparisons between the combined treatments analyzed: *(i)* N + WR vs. N + WW **(E)**; *(ii)* N + WR vs. NN + WR **(F)**; *(iii)* NN + WR vs. NN + WW **(G)**; *(iv)* N + WW vs. NN + WW **(H)**. Each dot represents the $\log_2$ of the FC in TOTAL and PAR samples. DEGs are shown as colored dots

hindered from associating with polysomes and undergoing active translation. This fact can suggest a potential sequestration of these transcripts within biomolecular condensates, such as P bodies or stress granules [24], offering a convincing avenue for additional research.

Weighted gene co-expression network analysis identified 25 gene modules with coordinated changes in the combined-treated soybean roots

WGCNA was performed using the 24 samples comprising the four combined-treated roots (in triplicates) and each sample's TOTAL and PAR fractions. The analysis identified 25 co-expression modules, as shown in Fig. 3A, using 26,793 genes, which were the ones left after genes with low coefficients of variation and/or low counts were removed (Table S3). These 25 distinct expression modules contained between 27 and 6276 genes and were further merged into seven groups based on their eigengene expression patterns (Fig. 3A). As can be seen, the distinct groups were differentially detected in the different samples and RNA fractions. It is worth mentioning that the first group (ME5, ME1, ME14, ME13, ME16, ME22) was mainly detected in polysome-associated RNA fraction across all treatments, with the highest detection in non-nodulated plants without water restriction (NN+WW treatment). The second group (ME11, ME9, ME12) exhibited detection in both RNA fractions (TOTAL and PAR) of the N+WW treatment, specifically associated with nodulation. Group three (ME18, ME7, ME24, ME20) included genes preferentially detected in well-watered conditions (NN+WW and N+WW treatments). Notably, the highest and lowest transcript detection were observed in NN+WW and N+WR treatments, respectively, highlighting the specificity of these genes to the treatment combination involving the most contrasting conditions. Group six (ME21, ME6, ME3, ME10, ME23) exhibited higher detection in N+WR and NN+WR treatments compared to the other two well-watered combined treatments. This is graphically presented in Fig. 3A.

The topological representation of the expression modules of the WGCNA network was projected in different colors using the R package network [25] (Fig. 3B). All modules, except for M17, showed edge adjacency values higher than the threshold (0.15) and thus are present on the network. This way, the whole network contained 8,944 nodes and 1,378,580 edges, disposed of as a central core, and several smaller sub-networks disconnected from the central core (Fig. 3B).

DEG analysis identified eight co-expressed modules as the more relevant ones in the plant responses to nodulation and/or water deficit

To explore further the plant responses to the different experimental conditions, the DEGs and their status (i.e., up- or down-regulated) found in the four comparisons between the combined treatments (Fig. 1C) were highlighted on the WGCNA network (Fig. 4). Moreover, the number of genes present in the three regulation levels analyzed (TOTAL, PAR, and TOTAL+PAR) was further discriminated in Table S4. Comparison *i*, which exhibited the response of nodulated plants to water deficit, had the greatest number of DEGs that localized within specific modules. M3, M5, M6, M7, M9, M10, M11, M12, M18, and M19 were the most representative modules in this case and included a similar number of up- and down-regulated DEGs (Fig. 4A; Table S4). The DEGs obtained in comparison *ii*, which showed how plants change their response to water deficit when involved in a symbiotic interaction with rhizobia, belonged mainly to modules 3, 7, and 13 (Fig. 4B). In this case, most DEGs (84%) were down-regulated (blue dots; Fig. 4B; Table S4). Comparison *iii* (NN+WR vs. NN+WW) evidenced the response of soybean plants to water restriction; here, the DEGs were mainly localized within M5, M7, M12, and M23 and comprised a similar number of up- and down-regulated DEGs (Fig. 4C; Table S4). The comparison that showed the plant responses due to the nodulation process without involving any water restriction, comparison *iv*, presented its DEGs principally in modules 9, 10, 11, and 12 (Fig. 4D), with 72% of them being up-regulated (red dots; Fig. 4D; Table S4). Notably, the most contrasting situation regarding the status of the DEGs was observed between comparison *i* and *iv* (Fig. 4A and D, respectively), where DEGs within modules 9, 10, 11, and 12 exhibited opposite regulation; e.g., DEGs in M11 were up-regulated in comparison *iv* whereas down-regulated in comparison *i*. Hence, many DEGs up-regulated due to the nodulation process inverted their expression when the water deficit was imposed on the nodulated plants.

Based on the number of DEGs localized at each of the WGCNA modules (Table S4), the eight most representative ones (M3, M5, M6, M7, M9, M10, M11, and M12) were selected for further analysis (Table S5). Table S5 complements the data presented in Fig. 4 by specifying the regulation at TOTAL, PAR, or TOTAL+PAR level for each comparison in which the DEG condition was achieved, at each status, and for each of the eight selected modules. Moreover, in Table S5, the number of the modules' hub genes (10% most connected) is shown in parentheses since we wondered whether any of the DEGs corresponded to any of the hub genes in each modules, as highly connected genes are often more important for the functionality of networks that other nodes [26]. In fact, in three out of the four comparisons analyzed, some DEGs were also hub genes, with the N+WR vs. NN+WR comparison (*ii*) having the most hub genes regarding its DEGs (12%: 25 out of 214). Comparison *iii* (NN+WR vs. NN+WW) contained no hub genes among their DEGs.

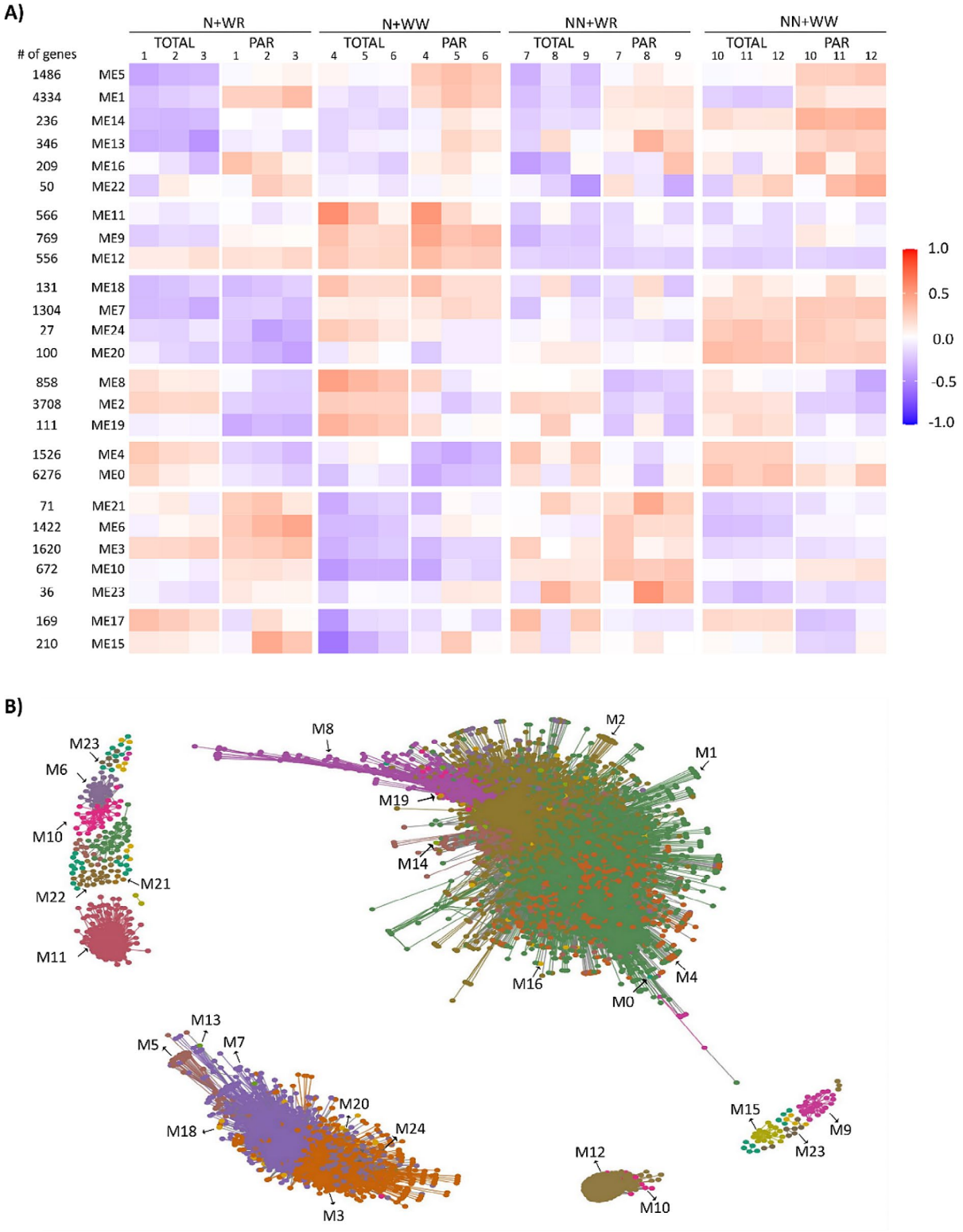

Fig. 3 Gene co-expression network of soybean plants subjected to four combined treatments comprising nodulation and water restriction conditions and total RNA and polysome-associated mRNA samples. **(A)** Heatmap of module eigengene (ME) expression profiles in the indicated combined treatments and RNA fraction (TOTAL, PAR). The three biological replicates of each treatment and the number of genes in each module are also shown. **(B)** Projection of expression modules (M) of the weighted gene co-expression network where 24 out of 25 modules are visualized; M17 is not visualized due to a low edge adjacency value. The network contained 8,944 nodes and 1,378,580 edges disposed as a central core and several smaller sub-networks disconnected from the central core

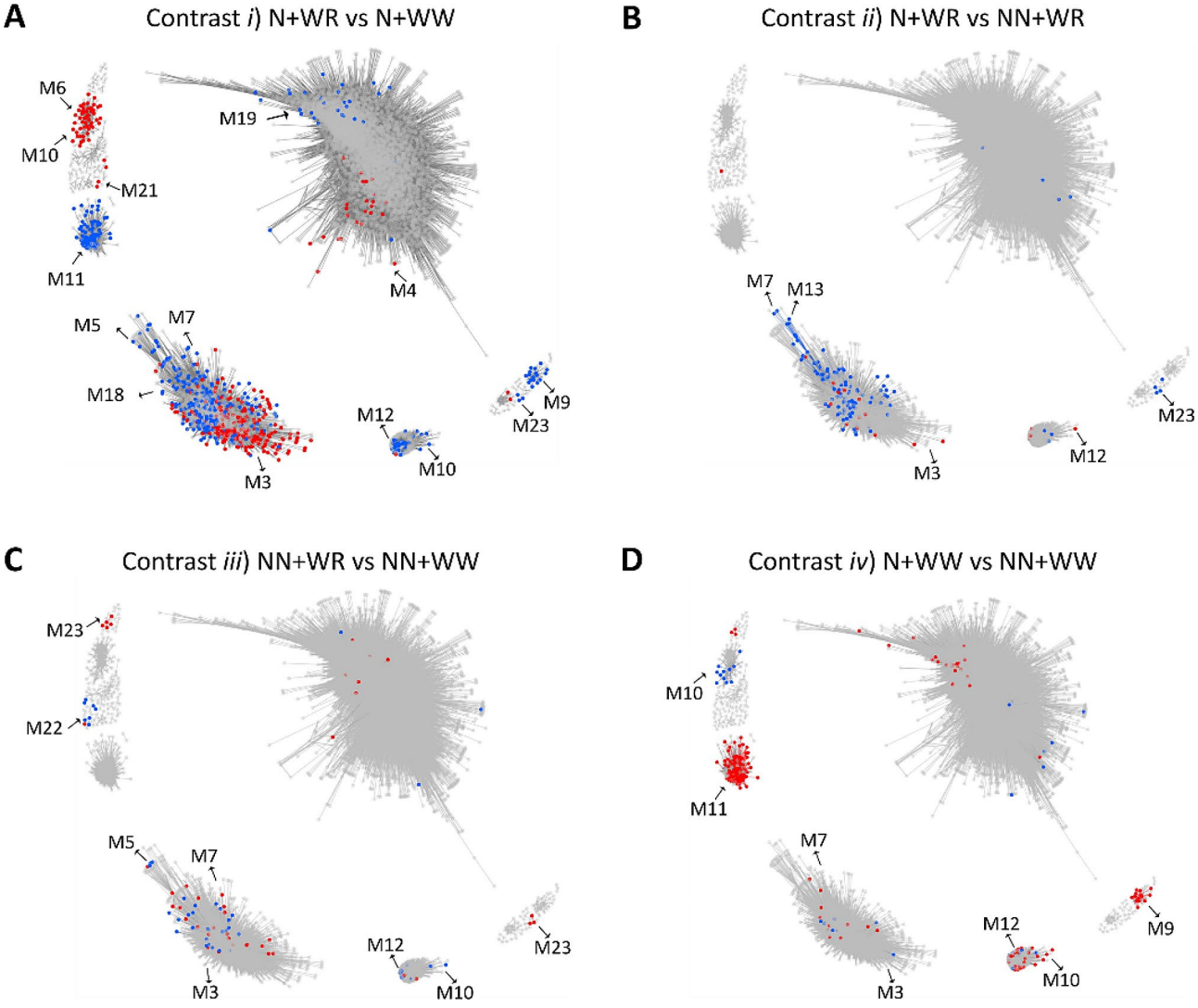

Fig. 4 Identification of differentially expressed genes (DEGs) at TOTAL, PAR, and TOTAL + PAR regulation levels in the comparisons between combined treatments within the WGCNA network. **(A)** Comparison *i*) N + WR vs. N + WW. **(B)** Comparison *ii*) N + WR vs. NN + WR. **(C)** Comparison *iii*) NN + WR vs. NN + WW. **(D)** Comparison *iv*) N + WW vs. NN + WW. Red dots represent up-regulated DEGs. Blue dots represent down-regulated DEGs. The four combined treatments were nodulated and water-restricted plants (N + WR), nodulated and well-watered plants (N + WW), non-nodulated and water-restricted plants (NN + WR), non-nodulated and well-watered plants (NN + WW). M: module

Furthermore, compared to the percentage found for the just mentioned comparisons *ii* (12%) and *iii* (0%), comparisons *i* (N+WR vs. N+WW) and *iv* (N+WW vs. NN+WW) showed an intermediate percentage of hub genes among their DEGs (4%: 70 out of 1636, and 5%: 21 out of 410, respectively). Our findings suggest that, whereas the nodulation condition affected the differential expression of some hub genes, the interaction between nodulation and water deficit (relative to non-nodulation) affected the differential expression of a higher proportion of hub genes (Table S5).

Another interesting focus point regarding our WGCNA+DEG analysis is the common DEGs among two comparisons that are discriminated according to their expression modules. The common DEGs found between comparisons *i* and *iii* are genes related to the plant responses to the water deficit, independently of the nodulation condition (Fig. 2C and D). Instead, the ones found in common between comparisons *ii* and *iv* are genes related to nodulation, regardless of the hydric condition (Fig. 2A and B). In both cases, there were a significant number of DEGs at TOTAL+PAR regulation level in some modules (Table S6). Interestingly, these DEGs were specific to certain co-expression modules (M3, M5, M6, M7, M12) and were hub genes in some cases (Table S7). Moreover, the common DEGs between comparisons *i* and *iii* were mainly present in M3, M5, M6, and M7, whereas the common DEGs between comparisons *ii* and

iv were almost entirely found in M12 and were almost entirely upregulated (Table S7). It is noteworthy that 27% (30 out of 113) of these upregulated DEGs were also hub genes of the module. In conclusion, most genes related to the plant responses to the water deficit were localized within four specific WGCNA expression modules (M3, M5, M6, M7), but, among these, mainly within M3 and M7. Also, almost all the genes related to nodulation were specifically localized in M12. Our RNA-seq data were validated by analysing the expression data of another soybean genotype—Génesis—which was assayed in parallel with the Don Mario genotype (not shown). We analysed expression data of the genes shown in Table S7 and found high positive Pearson correlation coefficients among the common genes between both genotypes in the different comparisons analysed in this study.

M3 comprised 67 hub genes within the network; 60 were DEGs (Table S3). Among them, 50 were common between comparisons *i* and *iii* at TOTAL+PAR level, and most [46] were up-regulated (Table S7). M5 and M6 each had 21 hub genes; five DEGs in M5 and 13 in M6 (Table S3). In M7, 75 out of 78 hub genes in M7 were DEGs, with 63 common in comparisons *i* and *iii* at TOTAL+PAR level (Table S7). M9 contained three hub genes coding histone H3.2 (Table S3). All five hub genes in M10 were up-regulated in N+WR plants (regarding N+WW plants, comparison *i*) at TOTAL+PAR level. Also, three were down-regulated in comparison *ii* (Table S3; Table S5). Regarding M11, the 18 hub genes were DEGs (Table S3), with down-regulation in comparison *i* at PAR and TOTAL+PAR levels and up-regulation in comparison *iv* at TOTAL and TOTAL+PAR levels (Table S5). Notably, 33% (6 out of 18) of M11 hub genes were transcription factors (TFs) (Table S3). All 30 M12' hub genes were up-regulated DEGs common between comparisons *ii* and *iv* at TOTAL+PAR level (Table S7), and 20 were down-regulated DEGs in comparison *i* at TOTAL, PAR, and TOTAL+PAR levels (Table S5).

Characterization of biological processes and pathways in the selected WGCNA modules

Subsequently, Gene Ontology (GO) functional enrichment analysis (biological process; BP) was used to identify and rank overrepresented functional categories [26] of the previously selected modules (M3, M5, M6, M7, M9, M10, M11, and M12) (Figure S2). M3, M7, M11, and M12 were chosen for further interpretation based on the DEGs characteristics (shown in Table S5 and Table S7) and on our interest in the BP and pathways in which they are enriched (Fig. 5). In addition to the GO_PB shown in Figure S2, this figure depicts the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis and the modules network with the hub genes highlighted in the central part. The different topologies of the networks, together with the centrality of the hub genes (big dots), can be appreciated. Notably, the nodes (small dots) in M12' network displayed a sharp peripheral distribution, while the networks of the other modules had a more uniform node distribution (Fig. 5).

Module 3, closely related to the response of nodulated plants to water deficit, was strongly associated with dephosphorylation (GO:0016311), cation transmembrane transport (GO: 0098655), and fatty acid metabolism (GO:0006631) processes (Fig. 5A). The KEGG pathways with higher gene counts included galactose and glycerophospholipid metabolisms, degradation of branched-chain amino acids (BCAAs), and glutathione (GSH) metabolism (Fig. 5A). M7 was enriched in protein phosphorylation (GO:0006351), cellular oxidant detoxification (GO:0098869), and hydrogen peroxide catabolism (GO:0042744) processes, the first of which had over 200 gene counts (Fig. 5B). Moreover, KEGG enrichment highlighted phenylpropanoid biosynthesis and the MAPK signaling pathway (Fig. 5B). The two GO-BPs with the greatest gene counts in M11 were regulation of transcription (GO:0006351) and protein ubiquitination (GO:0016567) (Fig. 5C). KEGG analysis, in this case, showed enrichment in the MAPK signaling pathway, plant-pathogen interaction, and phosphatidylinositol signaling system (Fig. 5C). M12 was strongly associated with nodulation, peptidyl-serine modification, and protein phosphorylation processes, with circadian rhythm as the only identified KEGG-enriched pathway (Fig. 5D). The absence of an enriched pathway related to nodulation in M12 may be due to the lack of information in the KEGG database.

Stand-out processes and pathways in the plant responses to nodulation, water deficit or the combination of both conditions

Combining information from different data layers may lead to new biologically interpretable associations [26]. Here, we further investigate the DEGs' more relevant WGCNA modules for protein-protein interaction (PPI)-enriched networks, from which novel biological knowledge can be derived. The online STRING tool, which integrates known and predicted associations between proteins, including both physical interactions and functional associations [27], was used to construct different PPI networks.

On the one hand, two particular subsets of DEGs obtained in comparison *i* (N+WR vs. N+WW) and *ii* (N+WR vs. NN+WR) were analyzed to seek potential interactions between them. These particular subsets of DEGs are highlighted in bold in Table S5, and the PPI networks are depicted in Fig. 6.

In comparison *i*, the subset of DEGs selected comprised M3 up-regulated genes at TOTAL+PAR level

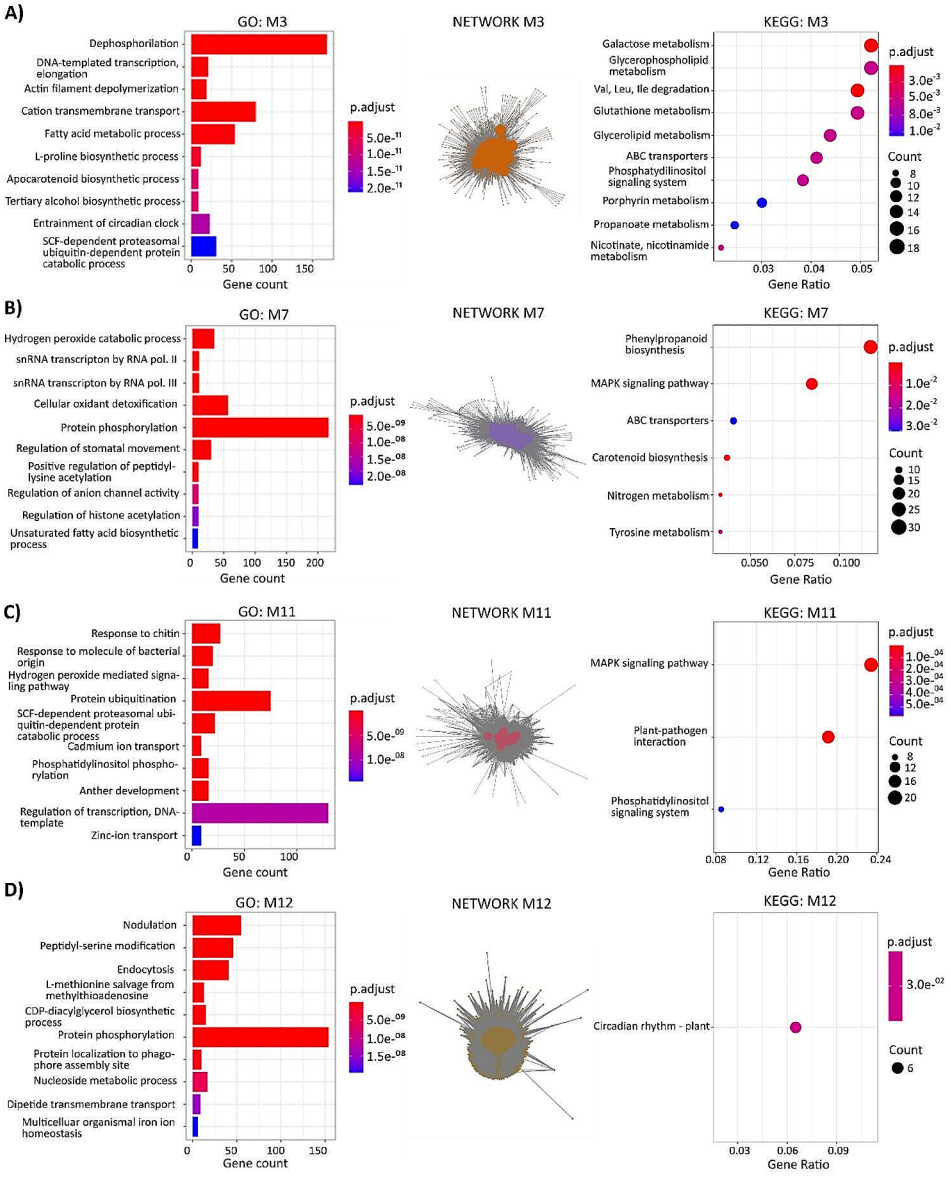

Fig. 5 Functional and pathway enrichment analysis and network representation of co-expression modules M3 (**A**), M7 (**B**), M11 (**C**) and M12 (**D**). The top ten Gene Ontology Biological Process (GO-BP) terms are shown for each module (y-axis, left chart). The nodes of the networks are highlighted with the colour of the corresponding module (according to Fig. 3B), and hub genes are the ones with the bigger dots. The modules' KEGG-enriched pathways are depicted in the right chart. The hub genes were defined as the 10% of genes with the highest connectivity

(Fig. 6A). The over-represented functional terms obtained in this PPI network depict the main processes that suffered changes due to up-regulation in nodulated plants subjected to water deficit, pinpointing them as candidates for further investigations. These processes can be mainly described by the following GO_BP terms and KEGG pathway: response to abiotic stress, protein folding, response to abscisic acid (ABA), transmembrane transport, pigment metabolic process, cellular amino acid catabolic process, branched-chain amino acids (BCAA) catabolic process, and galactose metabolism (Fig. 6A). The first two BPs mentioned above shared many nodes (proteins), mainly heat shock proteins (HSPs) from different families. Among the proteins associated with the term transmembrane transport were ABC (ATP-binding cassette) transporters, calcium, chloride, and sodium channels, and MATE (Multidrug And Toxic Compound Extrusion) family transporters. The BCAA catabolism BP was entirely comprised within the cellular amino acid catabolism cluster; however, its identification highlights the importance of these small hydrocarbon-branched amino acids, which serve as precursors to secondary metabolites [28] and whose breakdown process is relevant in plants subject to different environmental

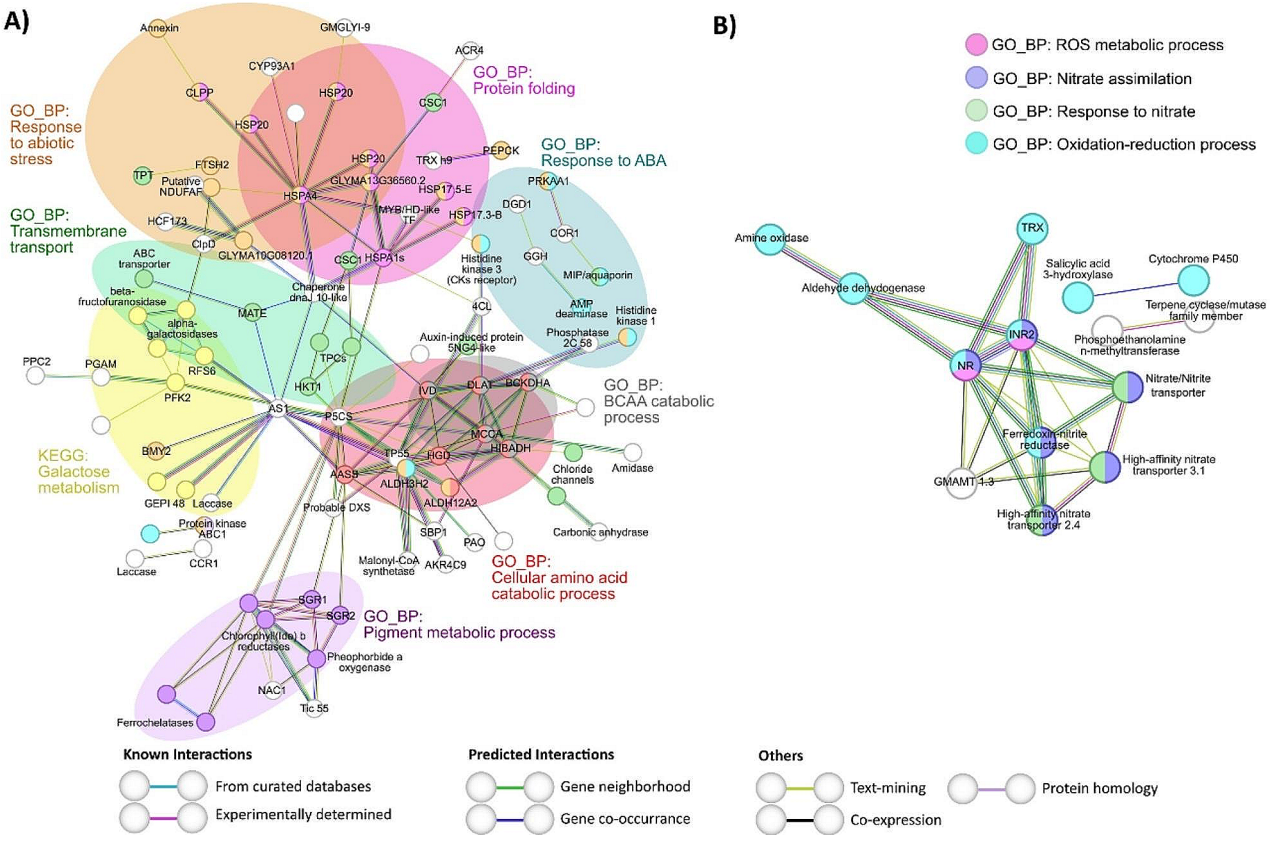

Fig. 6 String analysis derived protein-protein interaction networks obtained from two subsets of differentially expressed genes discriminated by WGCNA modules at TOTAL + PAR level of comparisons *i* **(A)** and *ii* **(B). (A)** Network obtained from WGCNA M3 up-regulated genes in nodulated and water-restricted plants with respect to nodulated and well-watered plants (*i*) N + WR vs. N + WW comparison) (highlighted in bold in Table S5). **(B)** Network obtained from WGCNA M7 down-regulated genes in nodulated and water-restricted plants with respect to non-nodulated and water-restricted plants (*ii*) N + WR vs. NN + WR comparison) (highlighted in bold in Table S5). The network nodes represent proteins. Connecting lines denote protein-protein associations whose colors represent the type of interaction evidence as defined in the legend. The GO_BP and KEGG functional terms are also specified in the legends according to their colors. Disconnected nodes, i.e., proteins not showing any interaction in the network, were deleted

conditions such as long-term darkness [29] and water deficit [30]. Nodes that exclusively belonged to the cellular amino acid catabolic process comprised proteins involved in the degradation of Phe, Tyr, Pro, and Lys. The proteins that were included in the galactose metabolism KEGG functional term are related to cell wall metabolism. The pigment metabolic process term mainly comprises proteins related to chlorophyll breakdown, such as stay green proteins, chlorophyll(Ide) b reductases, and pheophorbide a oxygenase (Fig. 6A).

In comparison *ii*, the selected subset of DEGs corresponded to M7 down-regulated genes also at TOTAL+PAR level (Fig. 6B). The PPI network of this particular subset of DEGs presented enriched terms related to nitrate metabolism (nitrate assimilation and response to nitrate), ROS metabolic process, and oxidation-reduction process (Fig. 6B). In fact, in nodulated and water-restricted plants, the down-regulation of proteins related to nitrate uptake and reduction makes sense when compared to non-nodulated (nitrate-fed) plants.

On the other, the two other subsets of DEGs selected to seek potential interactions corresponded to comparisons *i* and *ii* common DEGs (Fig. 7A) and to comparisons *ii* and *iv* common DEGs (Fig. 7B). These two particular subsets of DEGs are highlighted in bold in Table S7. For comparisons *i* and *iii* common DEGs (related to water deficit responses, independently of the nodulation condition), the subset analyzed included M3 up-regulated genes at TOTAL+PAR level. For comparisons *ii* and *iv* common DEGs (related to nodulation, regardless of the hydric condition), the subset comprised M12 up-regulated genes, also at TOTAL+PAR level. The PPI network for common DEGs in comparisons *i* and *iii* revealed enrichment in four KEGG pathway functional terms: "glutathione (GSH) metabolism", "plant hormone signal transduction", "BCAA degradation", and "galactose metabolism" (Fig. 7A). Gray and yellow-outlined terms, BCAA degradation and galactose metabolism, were also present in the network depicting processes related to nodulated plants' response to water deficit (Fig. 6A).

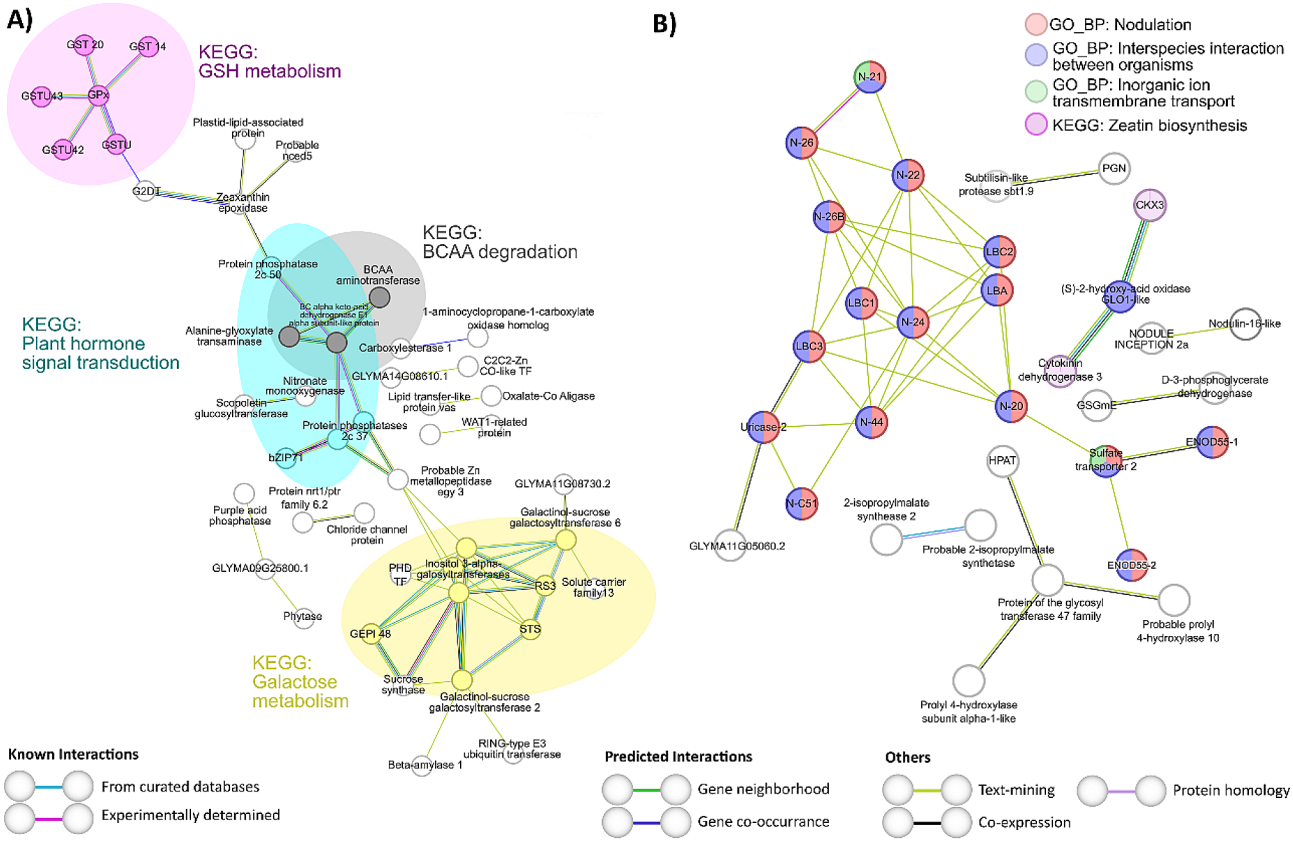

Fig. 7 String analysis derived protein-protein interaction networks obtained from the differentially expressed genes common to comparisons *i*) and *iii*) **(A)** and common to comparisons *ii*) and *iv*) **(B)**, discriminated by WGCNA modules at TOTAL + PAR level. **(A)** Network obtained from WGCNA M3 up-regulated genes (highlighted in bold in Table S7). **(B)** Network obtained from WGCNA M12 up-regulated genes (highlighted in bold in Table S7). The network nodes represent proteins. Connecting lines denote protein-protein associations whose colors represent the type of interaction evidence as specified in the legend. The GO_BP and KEGG functional terms are also specified in the legends according to their colors. Disconnected nodes, i.e., proteins not showing any interaction in the network, were deleted

However, proteins associated with these terms varied between the two figures. The "GSH metabolism" term included glutathione peroxidase (GPx) and glutathione S-transferase (GSTs) enzymes, while "plant hormone signal transduction" involved protein phosphatases of the 2 C family and an ABA-responsive element binding factor (bZIP71) (Fig. 7A). The PPI network for DEGs related to the nodulation process (those common in comparisons *ii* and *iv*) showed enrichment in various processes, including "nodulation," "interspecies interaction between organisms," "inorganic ion transmembrane transport," and "zeatin biosynthesis" (Fig. 7B).

Discussion

Studies on how soybean plants respond to water-deficit stress have been mainly conducted on non-nodulated plants [31–33], which in this work has been analyzed in comparison *iii* (NN+WR vs. NN+WW) (Fig. 1C). However, all the major producer countries of soybean rely on the high symbiotic nitrogen fixation capacity of this plant [34]. Hence, evaluating the responses of non-nodulated legume plants to abiotic constraints is not the most relevant approximation if we aim to explore the crops' response in the field. Moreover, evidence suggests that nodulated and non-nodulated plants respond differentially when subjected to water deficit [6–9], denoting even more, the relevance of dissecting how nodulated legume plants respond to water constraints. In addition to being different, the response of nodulated plants to water deficit is more complex in terms of the number of DEGs found (see Table S4, Table S5 and Fig. 4A and C for comparisons *i* and *iii*). Actually, comparison *iii* (the one not including nodulation) had the lowest number of DEGs of the four comparisons analyzed (1881, 308, 263, and 523 DEGs for comparisons *i*, *ii*, *iii*, and *iv*, respectively; Table S4).

The differential response strategies—depending on the nodulation condition—of soybean plants to water deficit could be explained by gene expression regulation changes. In particular, translational control of gene expression allows cells to respond to a stimulus quickly, providing flexibility and adaptability to organisms [18,

35, 36]. Therefore, here, we analyzed the transcriptome and the translatome of the plants' roots to identify genes subjected to translational control. ABA and ethylene are two crucial hormones that regulate plant responses to various stresses [37]. Additionally, cytokinin and auxin play essential roles in controlling cell proliferation, differentiation, and processes like root nodule development in plants [38]. Hence, the fact that we found DEGs mainly regulated at the translatome level in N+WR plants (with respect to N+WW and/or NN+WR plants) encoding various enzymes involved in ABA biosynthesis and signaling, such as zeaxanthin epoxidase and NDR1/HIN1-like protein 6, ethylene biosynthesis, like ACC synthase 3, which catalyzes the formation of ACC (a direct precursor of its biosynthesis), auxins—specifically IAA—inactivation, like IAA-amido synthetase GH3.10 and IAMT1, and cytokinin biosynthesis, such as IPT5, is very interesting, have not been addressed previously in our experimental conditions and is a focus point for future research regarding the fine-tuning of the molecular mechanisms governing the response of nodulated plants to water restriction (Fig. 2, E, F; Table S2). Genes with mixed (transcriptional+translational) regulation are the most prevalent, so we performed the last part of our analysis (PPI network analysis; Figs. 6 and 7) with the DEGs having the before-mentioned mixed regulation.

The analysis of our transcriptomic data through a pipeline that combines WGCNA+DEGs [23] has proven effective for the identification of co-expressed gene modules (Fig. 3) enriched with the DEGs obtained in the different comparisons. Interestingly, these DEGs co-localized specifically in some modules rather than being uniformly distributed throughout the WGCNA network (Fig. 4; Table S5). This allows the association between module/s and processes underlying the plant responses being evidenced in each of the comparisons. An even more interesting aspect of this analysis that arises when the regulation status (up- or down-regulated) of the DEGs is analyzed is what is depicted in Fig. 4 for comparisons *i* and *iv* (Fig. 4A and D). It can be seen that many DEGs induced by nodulation and localized within M9, M10, M11, and M12 inverted their expression when water deficit was imposed on the nodulated plants. Notably, the biological processes represented in these modules were translation, DNA replication, regulation of transcription, protein ubiquitination, nodulation, and protein phosphorylation (Figure S2). This evidences once again the association between the DEGs' enriched modules and the plant response mechanisms under evaluation and also that performing the differential expression analysis in the functional context of WGCNA facilitates the interpretation of complex scenarios as in the current study [39, 40]. Likewise, this strategy was proper when we analyzed the DEGs common to comparisons *i* and *iii* and to *ii* and *iv*, i.e., the DEGs common to the plants' responses to water deficit and nodulation, respectively (Table S6, Table S7). On the one hand, M3 and M7 can be assigned as the principal modules associated with the mechanisms that underlie the plants' responses to water deficit independently of the nodulation condition (Table S7). Instead, M12 can be categorically assigned as the one module associated with the nodulation process, regardless of the hydric condition (Table S7).

To further investigate which molecular mechanisms were involved in the plants' responses either to water deficit or nodulation, but mainly their combination, we performed GO_BP terms and KEGG pathways enrichment analysis (Fig. 5) and PPI network analysis (Figs. 6 and 7) for the most interesting modules and subsets of DEGs, respectively. M3 was identified as closely related to both the response of nodulated plants to water deficit—since most of the DEGs (91%) that localized on it correspond to comparison *i* DEGs (Table S5)—and, together with M7, to the plants' responses to water deficit independently of the nodulation condition, considering that 85% of the common DEGs between comparisons *i* and *iii* localized within these two modules (Table S7). However, the stand-out processes and pathways in the before-mentioned plant responses differ both as to quantity and type of process or pathway, as can be seen in the PPI networks depicted in Figs. 6A and 7A. As previously mentioned, terms related to "response to abiotic stress", "protein folding", "transmembrane transport", "response to ABA", "cellular amino acid catabolic process", and "pigment catabolic process" were associated with the plant responses to nodulation+water deficit (Fig. 6A). Instead, fewer terms were exclusively found to be associated with the plants' response to water deficit: "GSH metabolism" and "hormone signal transduction" (2 C protein phosphatases) (Fig. 7A), which is in agreement with other studies [31, 41, 42]. These findings support what has already been suggested (when considering the number of DEGs) about the response of nodulated plants to water deficit being unique and more complex if compare to that of non-nodulated plants.

Although two processes ("galactose metabolism" and "BCAA catabolism") were common to both responses, the proteins associated with those terms in each case were different (Figs. 6A and 7A). Proteins related to cell wall metabolism (included in the term "galactose metabolism") in the nodulated plant´s responses to water deficit were UDP-glucose epimerases, which are involved in channeling UDP-D-galactose into cell wall polymers [43], and alpha-galactosidase 3, which might have a role in cell wall loosening and expansion [44] (Fig. 6A). Rather, proteins related to the term "galactose metabolism" in the plant's responses to water deficit were inositol 3-alpha-galactosyltransferases (galactinol

synthases), stachyose synthase (involved in the pathway that converts raffinose into stachyose), and galactinol-sucrose galactosyl transferases (also known as raffinose synthases). These enzymes have been shown to enhance drought tolerance in maize and Arabidopsis by raffinose synthesis or galactinol hydrolysis [45] (Fig. 7A). The term "BCAA catabolism" in the response of nodulated plants to water deficit (Fig. 6A) comprised mainly proteins associated with leucine, valine, and isovaleric acid catabolism, but also the branched-chain keto acid dehydrogenase E1 subunit alpha (BCKDHA) which is a component of the complex that catalyzes the second major step in the BCAAs catabolism. Interestingly, the DLAT (Dihydrolipoyl transacetylase) protein, a component of the pyruvate dehydrogenase complex, is also included in the "BCAA catabolic process", confirming the connection that has been suggested between the degradation of these amino acids and the energetic status of the cells during drought conditions, considering they serve as alternative carbon sources as their degradation products are substrates of the tricarboxylic acid cycle [30, 46]. Further, DLAT presented experimentally determined interactions with two ABA and abiotic stress-responsive histidine kinases: histidine kinase 1, for which a role as an osmosensor in water deficit conditions has been demonstrated in Arabidopsis [47], and the cytokinin receptor histidine kinase 3. BCAAs have been suggested as important components of the ABA-mediated drought response [46].

The relevance of identifying the transcriptomic and translatomic changes and related molecular mechanisms contributing to nodulated soybean plants' responses to water deficit is clear. The experimental setup performed in this work allowed us to approach the above mentioned mechanisms through the up-regulated biological processes and metabolic pathways exclusively associated with nodulated and water-restricted plants. Moreover, the identification of the biological processes that, as a result of the nodulation process, were up-regulated but subsequently became down-regulated when the nodulated plants were subjected to water deficit shed light on the understanding of how nodulated plants respond to water deficit (Fig. 8). Further comprehensive analysis of these processes could lead to identifying new sources of tolerance to drought in soybean.

Conclusion

Our study presents the transcriptional and translational response of the root of nodulated soybean plants to water restriction, an approach that is not commonly documented in the literature. We found that the response of these plants´ roots to water restriction is distinct and more complex in terms of DEGs, both at the transcriptional and translational level, than that of the roots of non-nodulated plants. Our approach allowed to identify specifically the stand-out processes in the roots of water-restricted nodulated plants, which include transmembrane transport, response to ABA, and pigment metabolism. Moreover, the inclusion of the translatome analysis brought to light that in these plants´ roots, the metabolism of the phytohormones ABA, ethylene, auxin, and cytokinin is regulated mainly at the translational level, which has not been addressed previously in the literature, underscoring the significance of including this layer of analysis. The experimental setup and the data analysis pipeline followed allowed us to approach the main molecular mechanisms contributing to nodulated soybean plant responses to water deficit, whose comprehensive analysis may reveal novel sources of tolerance to drought in soybean.

Materials and methods

Plant growth and drought assay

The assay was carried out with the commercial Don Mario 6.8i (DM; gently provided by Sergio Ceretta from Instituto Nacional de Investigaciones Agropecuarias (INIA) - La Estanzuela, Uruguay) soybean [*Glycine max* (L.) Merr.] genotype as specified in [19]. Briefly, plants were grown in a 0.5 L plastic bottle filled with a mix of sand: vermiculite (1:1) in a growth chamber under controlled temperature, photoperiod, and humidity conditions. The U1302 *Bradirhizobium elkanii* strain was used for the inoculated plants (Fig. 1A). During the first 19 days after sowing (V2-V3 developmental stage), soybean seedlings were grown without water restriction keeping the substrate at field capacity with B&D-medium [48] supplemented with KNO_3 (0.5 mM and 5 mM final concentration for nodulated and non-nodulated plants, respectively). From day 20 after sowing, which corresponds to day 0 of the water deficit period, watering was withdrawn to the water-restricted (WR) plants. On the contrary, the well-watered (WW) plants were maintained without water restriction throughout the assay (Fig. 1A). The substrate water content was measured daily by gravimetry (water gravimetric content) during the growth and water deficit period (Figure S3). The imposition of the water deficit was monitored through stomatal conductance (g_{sw}) measurements (Porometer Model SC-1, Decagon Device). Daily measurements were performed for all plants from day 0 until the end of the water deficit period, determined for each WR plant when the g_{SW} value was approximately 50% of the one obtained on day 0 (Figure S3). At the end of the water deficit period of each WR plant, the roots were harvested and kept at -80 °C until polysomal fraction purification. The roots of the WW plants were harvested together with the WR plants and kept at -80 °C until polysomal fraction purification. At the moment of harvesting the nodulated plants, nodules were detached from the roots and both

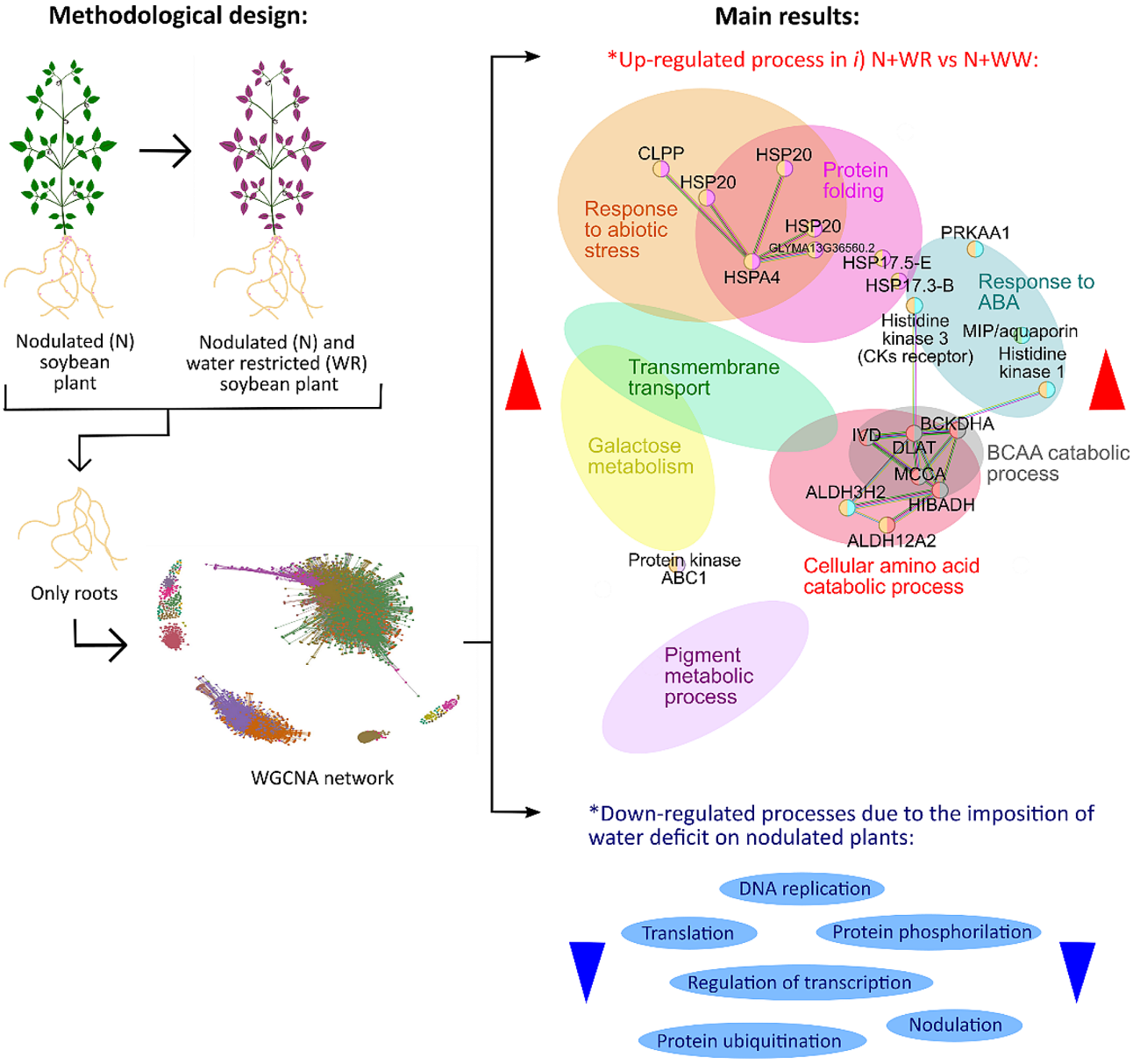

Fig. 8 Schematic overview of the methodological design and main results of the work. Total and polysome-associated mRNA were extracted from the roots of N and WR soybean plants, along with their respective controls (NN and WW plants, not shown). The extracted mRNA underwent RNA-seq analysis, and subsequent weighted gene co-expression network analysis (WGCNA) was conducted. The figure depicts over-represented up-regulated functional processes linked to the response of N plants to WR (comparison *i*), as well as processes that were down-regulated due to the imposition of WR on N plants. N: nodulated; NN: non-nodulated; WR: water-restricted; WW: well-watered

organs were kept at -80 °C separately. Only the roots were analysed in this work.

The drought assay's experimental design was a completely randomized one consisting of four combined treatments with five biological replicates each one, thus comprising a total of 20 pots. The experimental unit was one pot with one plant. The four combined treatments were nodulated (N) water-restricted (WR) plants (N+WR), nodulated well-watered (WW) plants (N+WW), non-nodulated (NN) water-restricted plants (NN+WR) and non-nodulated well-watered plants (NN+WW) (Fig. 1B).

Polysomal fraction purification

The polysomal fraction purification by sucrose cushion centrifugation was performed according to [49–51] with some modifications. Two mL of packed volume of frozen pulverized roots (850 mg) were homogenized in 4 mL of polysome extraction buffer ([49] for buffer composition) using mortar and pestle. Homogenates were maintained on ice with gentle shaking until all samples were

processed and clarified by centrifugation at 16,000 g for 15 min. Next, the homogenate was filtered with cheesecloth and the centrifugation step was repeated. 500 µL of the supernatant was reserved for isolation of total RNA (TOTAL). Two mL of the remaining supernatant was loaded on 12% and 33.5% sucrose layers cushion in 13.2 mL tubes (Ultra-Clear, Beckman Coulter, United States, 344,059) and centrifuge in a Beckman L-100 K class S ultracentrifuge (W40 Ti swinging bucket rotor) at 4 °C for 2 h at 35,000 rpm. Refer to [49] for sucrose cushion stock solutions and layers preparation. After centrifugation, the polysomal fraction was recovered as a pellet and resuspended in 200 µL of polysome resuspension buffer ([49] for buffer composition) pipetting up and down several times. The resuspended polysomal pellet was maintained for 30 min at 4 °C and then regular RNA purification was performed to obtain the polysome-associated mRNA (PAR) fraction.

The confirmation of the presence of polysomes in the pellet obtained through sucrose cushion centrifugation was assessed through a continuous sucrose density gradient. The gradients were made from 15% to 50% sucrose solutions using a linear gradient maker and a peristaltic pump. Once cold, the centrifuge tubes containing the gradients were loaded with 200 µL of the resuspended polysomal pellet and centrifuged in the same conditions as the sucrose cushion centrifugation. The gradient fractions were analyzed using a fixed-wavelength (254 nm) detector (Cole-Parmer, United States, EW-42664-35).

RNA extraction and transcriptome sequencing

TOTAL and PAR RNA fraction extraction was performed with TRizol LS reagent (Invitrogen, United States, 10296-028). The resuspended polysomal pellet and the extract reserved for isolation of total RNA were homogenized in 750 µL of TRizol, and the procedure shown in [49] was followed. After the resuspension of the RNA pellets, RNA concentration and integrity were measured using an Agilent 2100 bioanalyzer (Agilent Technologies, Inc., United States). Samples with a RIN (RNA integrity number) > 7.0 and > 1.0 µg were sent to Macrogen Inc. (South Korea) for library preparation and sequencing. TruSeq Stranded mRNA paired-end cDNA libraries were made and sequenced by the Illumina high-throughput sequencing platform. TOTAL and PAR samples from three biological replicates per combined treatment were sent for analysis.

Data analysis

Sequencing read processing

Per sample sequencing quality was visually inspected using FastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Trimmomatic was used for adapter remotion and low-sequencing-quality bases trimming. Only those trimmed sequences longer than 80 bp, with overall quality higher than 30, were retained for further analysis. Gene expression, on a transcript level, was quantified using salmon quasi-mapping mode (v0.12.0 [52]. Salmon default parameters were used, except the GC bias correction, which was enabled. The index for mapping was built from the most recent version of the reference Glycine max transcriptome, as retrieved from NCBI (GCF_000004515.6). Transcript read counts were then aggregated to gene level using the R package Tximport (v 1.2.0 [53]). Descriptive data, as initial read counts, percentage of data retained after quality control, mapping rate, beside others, is presented in Table S1. The sequencing data is available in the NCBI Sequence Read Archive (SRA) under the accession number PRJNA868178.

Differentially expressed genes (DEG) analysis

Initial descriptive analysis, as principal component analysis (PCA) of the samples and heatmap, were conducted using R base functions, and the R packages ggplot2 [54] and pheatmap [55]. Statistical analysis of differential expression analysis was performed with DESeq2 (version 1.16.1 [56], using the automated DESeq2 function. Differentially expressed genes (DEG) were defined as those with $|log2FC| > 1$ and Benjamini–Hochberg adjusted p-value (padj) < 0.05. Cluster analysis was performed using VennDiagramm [57], in order to identify DEG specifics for each comparison.

DEGs that co-localized with WGCNA modules of interest were subsequently used for Gene Ontology enrichment analysis, using topGO [58] and the weight01 method. Fisher's exact tests were applied to the biological process (BP) category. GO terms with a FDR < 0.05 were considered significantly enriched. BP_GO enrichment plots, depicting the top-10 significant GO terms, were generated using the 'enrichment_barplot' function of topGO. The gseKEGG function, as implemented in the clusterProfiler package [59] was used for KEGG pathway gene set enrichment analysis. For PPI analysis, stringDB ([27]; 11.5 version; https://string-db.org/) was used with default parameters, using protein sequences in FASTA format as input. For visualization purposes, unconnected nodes were hidden, while the most significant GO terms and/or KEGG pathways were colored.

WGCNA analysis

A co-expression network analysis was performed using the WGCNA package (v1.71 [60]. For this analysis, DEseq2 (version 1.16.1 [56]; see below) normalized data was initially filtered, to remove genes with low coefficient of variation and/or low counts (< than 50 in more than 50% of the samples), as low-expressed features tend to reflect noise. Then, the pickSoftThreshold function in WGCNA was used to select the soft threshold power for

the scale-free topology network. This way, the 1-step network construction and module detection function was applied, with power set to 17, maxBlockSize defined as the total number of genes, mergeCutHeight set to 0.20 and signed TOMType (as signed networks preserve the natural continuity of the correlation).

To determine whether the generated modules are associated with treatment groups, an Eigengene (hypothetical central gene) was calculated for each module, using the 'moduleEigengenes' function of WGCNA. A heatmap, depicting the relationship between module Eigengene and treatment groups, was generated using pheatmap [55].

Network was generated using the R package network [25], with the edge adjacency set to 0.15. For visualization purposes, modules were projected in different colors. Alternatively, DEG genes were colored in the network, based on their regulation status (i.e. up/down regulated). Gene connectivity was estimated using the 'intramodularConnectivity' function of WGCNA. Hub genes were defined as the top 10% genes of highest connectivity, following the recommendations of [61]. Finally, transcription factor annotation was performed following [62].

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12870-024-05280-5.

Supplementary Material 1

Supplementary Material 2

Supplementary Material 3

Supplementary Material 4

Supplementary Material 5

Supplementary Material 6

Acknowledgements
We thank Joaquina Farías (Centro Universitario Regional NorEste-Universidad de la República) for helpful comments on the manuscript. We also thank the Sistema Nacional de Investigadores (ANII) (María Martha Sainz, Carla Valeria Filippi, Guillermo Eastman, Mariana Sotelo-Silveira, José Sotelo-Silveira, Omar Borsani).

Author contributions
M.M.S., J.S.-S. and O.B. conceived the study; M.M.S., C.V.F., G.E., S. Z. and M.M.-M. performed the experiments; M.M.S., C.V.F., G.E., M.S.-S. and J.S.-S. analyzed the data; M.M.S. wrote the original draft except for the data analysis section of Materials and Methods; C.V.F. wrote that part. All authors have reviewed and edited the manuscript.

Funding
This work was supported by CSIC I + D 2020 Grant No. 282, FVF 2017 Grant No. 210 (María Martha Sainz), Programa de Desarrollo de las Ciencias Básicas (PEDECIBA) (María Martha Sainz, Carla Valeria Filippi, Guillermo Eastman, Mariana Sotelo-Silveira, Omar Borsani, José Sotelo-Silveira), and Red Nacional de Biotecnología Agrícola: RTS_1_2014_1-ANII (Omar Borsani).

Data availability
All datasets supporting the results of this study are included within the article and its supplementary information. The sequencing data is available in the NCBI Sequence Read Archive (SRA) under the accession number PRJNA868178.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 16 April 2024 / Accepted: 10 June 2024
Published online: 21 June 2024

References

1. Oldroyd GED, Murray JD, Poole PS, Downie JA. The rules of engagement in the legume-rhizobial symbiosis. Annu Rev Genet. 2011;45:119–44.
2. Concha C, Doerner P. The impact of the rhizobia – legume symbiosis on host root system architecture. J Exp Bot. 2020;71(13):3902–21.
3. Van Heerden PDR, De Beer M, Mellet DJ, Maphike HS, Foit W. Growth media effects on shoot physiology, nodule numbers and symbiotic nitrogen fixation in soybean. South African J Bot [Internet]. 2007 Nov [cited 2014 Oct 21];73(4):600–5. http://www.sciencedirect.com/science/article/pii/S025462990700316X.
4. Ferguson BJ, Mens C, Su H, Jones CH, Zhang M, Hastwell AH, et al. Legume nodulation: the host controls the party. Plant Cell Environ. 2019;42:41–51.
5. Antolín MC, Yoller J, Sánchez-Díaz M. Effects of temporary drought on nitrate-fed and nitrogen-fixing alfalfa plants. Plant Sci. 1995;107(2):159–65.
6. Borsani O, Díaz P, Monza J. Proline is involved in water stress responses of Lotus corniculatus nitrogen fixing and nitrate fed plants. J Plant Physiol. 1999;155(2):269–73.
7. Lodeiro AR, González P, Hernández A, Balagué LJ, Favelukes G. Comparison of drought tolerance in nitrogen-fixing and inorganic nitrogen-grown common beans. Plant Sci. 2000;154(1):31–41.
8. Staudinger C, Mehmeti-Tershani V, Gil-Quintana E, Gonzalez EM, Hofhansl F, Bachmann G et al. Evidence for a rhizobia-induced drought stress response strategy in Medicago truncatula. J Proteomics [Internet]. 2016;136:202–13. https://doi.org/10.1016/j.jprot.2016.01.006.
9. Liu Y, Guo Z, Shi H. Rhizobium Symbiosis leads to increased Drought Tolerance in Chinese milk vetch (Astragalus sinicus L). Agronomy. 2022;12(725):1–11.
10. Bazin J, Baerenfaller K, Gosai SJ, Gregory BD, Crespi M, Bailey-serres J. Global analysis of ribosome-associated noncoding RNAs unveils new modes of translational regulation. P Natl Acad Sci USA. 2017;E10018–27.
11. Kawaguchi R, Girke T, Bray EA, Bailey-Serres J. Differential mRNA translation contributes to gene regulation under non-stress and dehydration stress conditions in Arabidopsis thaliana. Plant J [Internet]. 2004 Jun [cited 2014 Jul 25];38(5):823–39. http://www.ncbi.nlm.nih.gov/pubmed/15144383.
12. Lei L, Shi J, Chen J, Zhang M, Sun S, Xie S, et al. Ribosome profiling reveals dynamic translational landscape in maize seedlings under drought stress. Plant J. 2015;84(6):1206–8.
13. Lee TA, Bailey-serres J. Integrative Analysis from the Epigenome to Translatome uncovers patterns of Dominant Nuclear regulation during transient stress. Plant Cell. 2019;31(November):2573–95.
14. Urquidi Camacho RA, Lokdarshi A, von Arnim AG. Translational gene regulation in plants: a green new deal. Wiley Interdiscip Rev RNA. 2020;11(6):1–40.
15. Piccirillo CA, Bjur E, Topisirovic I, Sonenberg N, Larsson O. Translational control of immune responses: from transcripts to translatomes. Nat Immunol. 2014;15(6):503–11.
16. Becker K, Bluhm A, Casas-vila N, Dinges N, Dejung M, Sayols S et al. Quantifying post-transcriptional regulation in the development of Drosophila melanogaster. Nat Commun [Internet]. 2018; https://doi.org/10.1038/s41467-018-07455-9.
17. Traubenik S, Reynoso A, Hobecker K, Lancia M, Hummel M, Rosen B, et al. Reprogramming of Root cells during Nitrogen-fixing symbiosis involves

Dynamic Polysome Association of Coding and Noncoding RNAs. Plant Cell. 2020;32(February):352–73.
18. Sablok G, Powell JJ, Kazan K. Emerging roles and Landscape of translating mRNAs in plants. Front Plant Sci. 2017;8(September):1–9.
19. Sainz MM, Filippi CV, Eastman G, Sotelo-Silveira JR, Borsani O, Sotelo-Silveira M. Analysis of Thioredoxins and glutaredoxins in soybean: evidence of translational regulation under Water Restriction. Antioxidants. 2022;11(1622):1–22.
20. Kang J, Peng Y, Xu W. Crop Root responses to Drought stress: Molecular mechanisms, Nutrient regulations, and interactions with microorganisms in the Rhizosphere. Int J Mol Sci. 2022;23(16):1–26.
21. Takahashi F, Kuromori T, Urano K, Yamaguchi-Shinozaki K, Shinozaki K. Drought stress responses and resistance in plants: from Cellular responses to Long-Distance Intercellular Communication. Front Plant Sci. 2020;11(September):1–14.
22. Comas LH, Becker SR, Cruz VMV, Byrne PF, Dierig DA. Root traits contributing to plant productivity under drought. Front Funct Plant Ecol. 2013.
23. Sánchez-Baizán N, Ribas L, Piferrer F. Improved biomarker discovery through a plot twist in transcriptomic data analysis. BMC Biol [Internet]. 2022;20:1–26. https://doi.org/10.1186/s12915-022-01398-w.
24. Parker DM, Winkenbach LP, Osborne Nishimura E. It's just a phase: exploring the relationship between mRNA, Biomolecular condensates, and Translational Control. Front Genet. 2022;13(June):1–16.
25. Butts CT. Network: a package for managing relational data in R. J Stat Softw. 2008;24(2).
26. van Dam S, Võsa U, van der Graaf A, Franke L, de Magalhães JP. Gene co-expression analysis for functional classification and gene-disease predictions. Brief Bioinform. 2018;19(4):575–92.
27. Szklarczyk D, Gable AL, Nastou KC, Lyon D, Kirsch R, Pyysalo S, et al. The STRING database in 2021: customizable protein–protein networks, and functional characterization of user-uploaded gene/measurement sets. Nucleic Acids Res. 2021;49(18):10800.
28. DeKraker JW, Gershenzon J. From amino acid to glucosinolate biosynthesis: protein sequence changes in the evolution of methylthioalkylmalate synthase in Arabidopsis. Plant Cell. 2011;23(1):38–53.
29. Peng C, Uygun S, Shiu SH, Last RL. The impact of the branched-chain keto-acid dehydrogenase complex on amino acid homeostasis in Arabidopsis. Plant Physiol. 2015;169(3):1807–20.
30. Pires MV, Júnior AAP, Medeiros DB, Daloso DM, Pham PA, Barros KA. The influence of alternative pathways of respiration that utilize branched-chain amino acids following water shortage in Arabidopsis. Plant Cell Environ. 2016;39:1304–19.
31. Van Ha C, Watanabe Y, Tran UT, Le DT, Tanaka M, Nguyen KH, et al. Comparative analysis of root transcriptomes from two contrasting drought-responsive Williams 82 and DT2008 soybean cultivars under normal and dehydration conditions. Front Plant Sci. 2015;6(AUG):1–12.
32. Song L, Prince S, Valliyodan B, Joshi T, Maldonado dos Santos JV, Wang J et al. Genome-wide transcriptome analysis of soybean primary root under varying water-deficit conditions. BMC Genomics [Internet]. 2016;17(57):1–17. http://www.biomedcentral.com/1471-2164/17/57.
33. Wang X, Wu Z, Zhou Q, Wang X, Song S, Dong S. Physiological Response of Soybean Plants to Water Deficit. Front Plant Sci. 2022;12(January).
34. Maluk M, Giles M, Wardell GE, Akramin AT, Ferrando-molina F, Murdoch A et al. Biological nitrogen fi xation by soybean (Glycine max [L.] Merr.), a novel, high protein crop in Scotland, requires inoculation with non-native bradyrhizobia. Front Agron. 2023;(5:1196873).
35. Hummel M, Rahmani F, Smeekens S, Hanson J. Sucrose-mediated translational control. Ann Bot. 2009;104:1–7.
36. Merchante C, Stepanova AN, Alonso JM. Translation regulation in plants: an interesting past, an exciting present and a promising future. Plant J. 2017;90:628–53.
37. Müller M. Foes or friends: ABA and ethylene interaction under abiotic stress. Plants. 2021;10(448):1–7.
38. Suzaki T, Ito M, Kawaguchi M. Genetic basis of cytokinin and auxin functions during root nodule development. Front Plant Sci. 2013;4(MAR):1–6.
39. Li W, Wang L, Wu YUE, Yuan Z, Zhou J. Weighted gene co – expression network analysis to identify key modules and hub genes associated with atrial fibrillation. Int J Mol Med. 2020;45:401–16.
40. Sferra G, Fantozzi D, Scippa GS, Trupiano D. Key pathways and genes of Arabidopsis thaliana and Arabidopsis Halleri roots under Cadmium stress responses: differences and similarities. Plants. 2023;12(1793):1–19.
41. Lin L, Wang J, Wang Q, Ji M, Hong S, Shang L, et al. Transcriptome Approach reveals the response mechanism of Heimia myrtifolia (Lythraceae, Myrtales) to Drought stress. Front Plant Sci. 2022;13(July):1–14.
42. Ji W, Yu H, Shangguan Y, Cao J, Chen X, Zhao L et al. Transcriptome profiling of Gossypium Anomalum Seedlings reveals key regulators and metabolic pathways in response to Drought stress. Plants. 2023;12(2).
43. Rösti J, Barton CJ, Albrecht S, Dupree P, Pauly M, Findlay K, et al. UDP-glucose 4-epimerase isoforms UGE2 and UGE4 cooperate in providing UDP-galactose for cell wall biosynthesis and growth of Arabidopsis thaliana. Plant Cell. 2007;19(5):1565–79.
44. Chuankhayan P, Lee RH, Guan HH, Lin CC, Chen NC, Huang YC, et al. Structural insight into the hydrolase and synthase activities of an alkaline α-galactosidase from Arabidopsis from complexes with substrate/product. Acta Crystallogr Sect D Struct Biol. 2023;79:154–67.
45. Li T, Zhang Y, Liu Y, Li X, Hao G, Han Q et al. Raffinose synthase enhances drought tolerance through raffinose synthesis or galactinol hydrolysis in maize and Arabidopsis plants. J Biol Chem [Internet]. 2020;295(23):8064–77. https://doi.org/10.1074/jbc.RA120.013948.
46. Zivanovic B, Milic Komic S, Tosti T, Vidovic M, Prokic L, Veljovic Jovanovic S. Leaf Soluble sugars and free amino acids as important components of abscisic acid — mediated Drought Response in Tomato. Plants. 2020;9(1147):1–17.
47. Wohlbach DJ, Quirino BF, Sussman MR. Analysis of the Arabidopsis histidine kinase ATHK1 reveals a connection between vegetative osmotic stress sensing and seed maturation. Plant Cell. 2008;20(4):1101–17.
48. Broughton WJ, Dilworth MJ. Control of Leghaemoglobin Synthesis in Snake beans. Biochem J. 1971;125(4):1075–80.
49. Sainz MM, Filippi CV, Eastman G, Sotelo-Silveira M, Martinez CM, Borsani O et al. Polysome Purification from Soybean Symbiotic Nodules. J Vis Exp [Internet]. 2022; https://medium.com/@arifwicaksanaa/pengertian-use-case-a7e576e1b6bf.
50. Di Paolo A, Eeastman G, Mesquita-Ribeiro R, Farias J, Macklin A, Kislinger T, et al. PDCD4 regulates axonal growth by translational repression of neurite growth-related genes and is modulated during nerve injury responses. RNA. 2020;26(11):1637–53.
51. Smircich P, Eastman G, Bispo S, Duhagon MA, Guerra-Slompo EP, Garat B, et al. Ribosome profiling reveals translation control as a key mechanism generating differential gene expression in Trypanosoma Cruzi. BMC Genomics. 2015;16(443):1–14.
52. Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. Salmon provides fast and bias-aware quantification of transcript expression. Nat Methods [Internet]. 2017;14(4):417–9. https://doi.org/10.1038/nmeth.4197.
53. Soneson C, Love MI, Robinson MD. Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. F1000Research. 2016;4:1–23.
54. Wickham H. ggplot2: elegant graphics for data analysis. New York: Springer-; 2016.
55. Kolde R. pheatmap: Pretty Heatmaps. 2019.
56. Love MI, Huber W, Anders S. Moderated estimation of Fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(550):1–21.
57. Chen H. VennDiagram: generate high-resolution venn and euler plots. R package version 1.7.3. 2022. pp. 1–35.
58. Alexa A, Rahnenfuhrer J, topGO. Enrichment Analysis for Gene Ontology. 2023.
59. Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z et al. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. Innov [Internet]. 2021;2(3):100141. https://doi.org/10.1016/j.xinn.2021.100141.
60. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics. 2008;9(559):1–13.
61. Seo CH, Kim JR, Kim MS, Cho KH. Hub genes with positive feedbacks function as master switches in developmental gene regulatory networks. Bioinformatics. 2009;25(15):1898–904.
62. Jin J, Tian F, Yang DC, Meng YQ, Kong L, Luo J, et al. PlantTFDB 4.0: toward a central hub for transcription factors and regulatory interactions in plants. Nucleic Acids Res. 2017;45(D1):D1040–5.

Publisher's Note