Supplementary Information (SI Appendix)

Regulatory network analysis reveals novel regulators of seed desiccation tolerance in *Arabidopsis thaliana*

Sandra Isabel González-Morales, Ricardo A. Chávez Montes, Corina Hayano-Kanashiro, Gerardo Alejo-Jacuinde, Thelma Yerenny Rico-Cambron, Stefan de Folter, and Luis Herrera Estrella.

Supplementary Methodology

Plant materials and growth conditions

Arabidopsis thaliana ecotype Col-0, Ler and Ws-2 were used as wild type in this study. The following mutant lines were used: *lec1-1*, *lec2-1*(1), in Ws-2 background; *abi3-1* and *abi3-5* (2) in Ler background; *fus3-3*(3) in Col-0 background. Plants were grown in a sterile mix of vermiculite and soil in a growth chamber at 22 °C with a 16-h photoperiod at 200 mmol m^2 s.

For RNAseq and carbohydrate profiles, flowers were marked and at specific times after flowering (15, 17 and 21 DAF), siliques were harvested from 24 plants and seeds collected and immediately frozen in liquid nitrogen, and stored at -80°C. All samples were collected in triplicates

Arabidopsis T-DNA insertion mutant lines were obtained from the Nottingham Arabidopsis Stock Centre. Homozygous lines were identified by PCR genotyping to corroborate site of insertion and identify homozygous plants (SI Appendix, Table S10). Plants were grown in the same conditions as described above for the

production of seed. To determine the phenotypes of the T-DNA insertion lines, experiments were performed on three biological replicates of 100 seeds, harvested from a pool of eight plants. The seed storage conditions were 25°C with 10% of relative humidity (RH) for 1, 2, 4 and 8 weeks. Viability of the dry seeds was determined by germination on 0.1X MS media after imbibition of seeds for 2d at 4°C to remove dormancy. To determine ABA sensitivity, seeds were germinated in MS medium in the presence of different concentrations of ABA (mixed isomers; Sigma-Aldrich) 0, 3, 5, 7 and 10 µM.

RNAseq library preparation and sequencing analysis

Total RNA was isolated using Concert Plant RNA Purification reagent (Invitrogen) and then re-purified with the TRIZOL reagent (Invitrogen). To ensure high quality RNA samples, RNeasy MinElute Cleanup kit (Qiagen) was used following the manufacturer's instructions. First and second strand cDNA synthesis was performed using 3 µg of the total RNA mixture using Message Amp-II kit (Ambion) following the manufacturers instruction. 10-12 ng of the synthesized cDNA was transcribed by *in vitro* transcription and the resulting 70-90 µg of antisense RNA (aRNA) were purified using RNAeasy columns (Qiagen). A second round of cDNA synthesis was performed using 20 µg of mRNA as template. cDNA synthesis was performed as described above except that random primers (mostly hexamers) were used for first strand synthesis. This procedure yielded approximately 10 µg of cDNA that was purified using the DNA Clear Kit for cDNA purification (Ambion). Samples were barcoded for multiplexing using the SOLiD Barcoding Kit. Libraries were sequenced for 50 bps on the SOLID 4 platform (Applied Biosystems). The

raw data have been deposited in the Gene Expression Omnibus (GEO) database accession number GSE76015.

Transcriptome data analysis

The SOLiD BioScope Whole Transcriptome Analysis (WTA) version 1.2.1 pipeline for single reads was used to align the reads to the TAIR10 Col-0 reference genome available at The Arabidopsis Information Resource. Two perfect matches per location were allowed, aligned reads per exon were counted, and per base coverage calculated. The WTA pipeline CountTags module provides normalize RPKM (reads per kilobase of exon sequence, per million reads) values along with read counts per exon. Custom bash and MySQL scripts were then used to calculate the number of counts per gene model.

Statistical analyses

Gene counts were normalized using edgeR's (v2.9.16)(4) TMM [trimmed mean M (¼ log fold-change gene expression)] algorithm. For the analysis of differentially expressed genes only genes with at least 5 reads across all samples were included. For the tolerance designation, we tested for differential expression of genes using the multifactor generalized linear models (*glms*) approach in edgeR. We fit negative binomial *glms* with Cox-Reid tagwise dispersion estimates to models that included intolerant and tolerant lines and mutants as factors (SI Appendix, Table. S3). To determine differential expression, we performed likelihood ratio tests by dropping one coefficient from the design matrix (i.e., the "null" model) and comparing it to the full model (4). Differentially expressed genes

(DEGs) between samples were considered based on a 5% false discovery rate (FDR) cutoff.

Categorization and functional analysis

GO annotation analysis on gene clusters was performed using the BiNGO 2.3 plugin tool in Cytoscape version 2.6 with GO_full and GO_slim categories, as described in (5). Over-represented GO_Full categories were identified using a hypergeometric test with a significance threshold of 0.05 after a Benjamini and Hochberg FDR correction(6).

Network analysis

A list of all microarray experiments using the Affymetrix GeneChip ATH1 was downloaded from the EBI ArrayExpress database. The CEL files were manipulated as previously described (7). To avoid possible perturbations of the underlying gene regulatory network, all CEL files corresponding to transgenic samples (mutants, overexpressions, promoter constructions) were excluded. This resulted in 168 CEL files that were normalized using gcRMA under R. CEL files were transformed as previously described (7). The resulting normalized data was used as input for the ARACNe algorithm using the previously reported ARACNe pipeline and list of transcription factors (7).The visualization and manipulation of networks and subnetworks was performed with Cystoscope v2.8.3.

Carbohydrate analysis

50 µg of seeds from 15, 17 and 21 DAF were ground in liquid nitrogen, and soluble sugars extracted twice for 15 min at 75°C with metanol 80% (v/v); then samples

were centrifuged at 5000 rpm for 10 min. The supernatant was dried under vacuum and dissolved in acetonitrile:water (1:1). Carbohydrates were quantified on an LC-MS system composed of the ACQUITY ultra-performance liquid chromatography (UPLC) system (Waters Corporation) fitted to a Q-TOF Premier mass spectrometer (LCT XE Premier, Waters). Conditions for chromatography were as follows: column, ACQUITY HILIC BEH column (2.1×100mm i.d., 1.7 µm, Waters); solvent A, ACN:H_2O (90:10) + 0.1% NaOH ; solvent B, ACN:H_2O (30:70) + NaOH t, 0 to 55 min, initial, 100% A and 0% B; 20 min, 40% A and 60% B; 55 min 100% A and 0% B; flow rate, 0.20 mL/min; injection volume 5 µL and column temperature, 35 °C. Conditions for mass spectrometry under negative mode were as follows, capillary voltage, 2.8 kV; cone voltage, 40 V; source temperature, 350°C; desolvation temperature, 100°C; cone gas flow, 30 L/h; desolvation gas flow, 750 L/h; nebulizer and curtain gas, N_2.The amount of carbohydrates were determined by spectrometer software (MassLynxTM v. 4.1, Micromass) using calibration curves prepared with carbohydrates standards (D-glucose, D-fructose, sucrose, raffinose and stachyose) purchased from Sigma-Aldrich.

Construction of overexpression lines

For pCaV35S::cDNA constructs, the corresponding coding sequence was amplified with the primers indicated for each gene in SI Appendix, Table. S10. PCR products were cloned in pDONR221 and transferred into T-DNA binary vector pFastG02(8). The resultant vectors were used for Agrobacterium-mediated transformation of Arabidopsis Col-0 and *abi3-5* plants using the floral dip transformation method as previously reported (9). Harvested seeds were spread on MS medium containing

20 µM of Phosphinothricin (PPT) for selection of transgenic plants. After selfing, plants with 3:1 segregation rates corresponding to single insertions were selected to produce homozygous lines.

Desiccation tolerance of *abi3-5* TF-overexpressing lines was performed using three biological replicates of 100 seeds, harvested from a pool of 8 plants. Seeds at 25 DAF were harvested and stored at 25°C and 20% RH for 0, 2, 3 and 4 weeks. Viability of the dry seeds was determined by germination on 0.1x MS media after imbibition of seeds for 2d at 4°C to remove dormancy.

Gene expression analysis

Gene-specific primer pairs (SI Appendix, Table S10) designed using the NCBI/ Primer-BLAST tool (10) were used for real-time PCR. A total of 10 µg of RNA was reverse transcribed using SuperScript® III Reverse Transcriptase (Life Technologies) according to the manufacturer's instructions. Reactions were performed with the SYBR Green PCR Master Mix in an ABI 7500 Fast Real-time system. UBQ10 and TIP4L were used as standards for cDNA content normalization. The thermal cycling program was set to 95 °C for 5 min, 40 cycles of 95 °C for 30 s, 60 °C for 30 s, and 72 °C for 1 min. Results were analyzed using the ABI 7500 on-board software, version 2.0.5 (Applied Biosystems). The real-time PCR was conducted with at least three experimental replicates for each biological sample.

SUPPLEMENTARY TEXT.

Global transcriptional analysis of upregulated genes in desiccation tolerant and intolerant Arabidopsis seeds.

The Venn diagrams presented in Fig. S2 show the relationship between upregulated DEGs in *lec1-1*, *abi3-5, fus3-3* and *lec2-1* with respect to their corresponding wild type. Comparison between tolerant and intolerant mutant pairs such as *lec1-1* and *lec2-1* or *abi3-5* and *abi3-1* with their corresponding Wt controls produced a common set of DEGs that are relevant to the pleiotropic phenotype of the mutants. For example, as *lec2* has a similar phenotype to *lec1* but it is still desiccation tolerant, genes that are differentially expressed in both mutants when compared to the WT should have no direct relevance to the tolerance process.

Comparison of differentially expressed genes of all DI mutants yielded a set of genes that differentially expressed in all DI mtants respect to the DT mutants and corresponding WT. At 15 DAF, 89 upregulated genes were common to all intolerant mutants (*lec1, abi3-5* and *fus3-3)* (SI Appendix, Fig. S2). As the level of water loss increased, the number of DEGs also increased; at 17 and 21 DAF a total of 129 and 310 upregulated genes were common to intolerant mutants (SI Appendix, Fig. S2). The expression of these genes is altered in all desiccation intolerant mutants respect to tolerant mutants with related phenotypes or their wild type counterparts (Dataset S8). The highest fold change values of DEGs were reached at 21 DAF, when water loss in the seed is complete.

A subset of DEGs was identified as specific for each mutant. For instance, *fus3-3* showed the highest number of specific DEGs in the three time-points analyzed (827, 1153 and 926 upregulated and 1187, 779 and 587 downregulated genes at 15, 17 and 21 DAF, respectively) (SI Appendix, Fig.S2). The mutant specific DEGs most probably encode genes that are independently regulated by each of these TFs and are related to the specific phenotypes of the corresponding mutant. For example, genes unique to *fus3-3* were mostly related to cell cycle, cell division and ethylene responses, in agreement with previous reports reporting that the *fus3-3* mutation induces ectopic cell divisions in the embryo (11) together with an increased expression of a subset of ethylene responsive genes (12) (SI Appendix, Table S8). Since the reduced expression of these genes is not detected in the other two DT mutants it suggests that these processes are independent of seed DT.

Regulatory networks inference

Datasets of gene expression were obtained as described in the Materials and Methods section. The first dataset, that we denominated TFs-only dataset, is a 168 columns by 2088 rows matrix that correspond to the 2088 TF (SI Appendix, Table S5 and Dataset S9) probe sets present in the ATH1 chip. The second dataset, we named the complete dataset, is 168 by 22810 matrix that contain all 228810 probe sets present in the ATH1 chip. We used both dataset as input to construct regulatory networks using the ARACNe software (13). The ARACNe output is a list of interaction probeset pairs ranked through a Mutual Information value (MI) and its associated p-value (14). Detail for a theoretical background and practical use of

ARACNe have been previously published (7, 14). In a biological context, an interaction between gene A and gene B would indicate that gene A and gene B participate in the same physiological process and, if gene A is a TF and gene B is not TF, the interaction would suggest that the gene A is a transcriptional regulator of gene B (7).

Network construction was concentrated on the 2088 TF probeset and obtained at three data processing inequality (DPI) values, 0.0, 0.1, and 0.2 (SI Appendix, Table S6). DPI is a known information property explained in the (14). A DPI 0.0 means when a three-node triangle is present, the interaction with lowest MI will be removed,as this interaction is considered to represent an indirect interaction. At DP1 1.0 interactions are removed (14) .

After transforming the ARACNe output adjacency files into Cytoscape compatible tables, we obtained the TFs-only (TFsSeedNet, Dataset S9) and complete (FullSeedNet, Dataset S10) databases. SI Appendix, Table S6 shows that the number of edges in the network increases dramatically from DPI 0.0 to DPI 0.1 to DPI 0.2. For this work, a graphical representation of TFSeedNet was obtained at DPI 0.0 while FullSeedNet was obtained at DPI 0.1. The DPI value of FullSeedNet was used to preserve interaction that are removed in DPI 0.0 in order to have a wider view of the genes that could interact with different TFs.

Table S1. Phenotypes of different lines used in the paper.

Phenotype	*Lines*					
	wildtype	*fus3-3*	*lec1-1*	*lec2-1*	*abi3-5*	*abi3-1*
Chlorophyll accumulation in dry seed	NO	NO	YES	YES	YES	NO
Anthocyanin in cotyledons	NO	YES	YES	YES	NO	NO
Storage protein expression	Normal	Reduced	Reduced	Reduced	Reduced	Normal
Ectopic trichomes on Cotyledons	NO	YES	YES	YES	NO	NO
Seed ABA sensitivity	Normal	Normal	Normal	Normal	Reduced	Reduced
Desiccation-tolerant seeds	YES	NO	NO	YES	NO	YES
Genetic Backgroud	-	Col-0	Ws	Ws	Ler	Ler

Table S2. Overview of sequencing and mapping statistics on Arabidopsis genomes.

library	Total reads	Reads mapped (%)	Reads uniquely (%)	Genes detected
***lec1-1*_15daf**	8,907,428	8,907,428 (73.6)	5,909,235 (92.4)	22,011
***lec1-1*_17daf**	6,664,483	5,000,413 (75.0)	4,413,624 (92.0)	22,011
***lec1-1*_21daf**	6,317,389	4,560,674 (72.2)	3,977,475 (91.4)	22,011
***lec2-1*_15daf**	6,469,212	4,983,250 (77.0)	4,601,405 (93.3)	22,011
***lec2-1*_17daf**	6,895,879	5,119,705 (74.2)	4,703,369 (93.3)	22,011
***lec2-1*_21daf**	8,499,125	6,305,852 (74.2)	5,668,640 (93.4)	22,011
Ws_15daf	6,441,380	4,784,939 (74.3)	4,468,431 (93.7)	22,011
Ws_17daf	5,562,550	4,242,351 (76.3)	3,934,393 (93.0)	22,011
Ws_21daf	7,275,259	5,402,772 (74.3)	5,002,734 (93.5)	22,011
***abi3-5*_15daf**	5,112,140	3,941,394 (77.1)	3,572,798 (91.6)	22,011
***abi3-5*_17daf**	5,180,272	3,989,412 (77.0)	3,652,498 (92.4)	22,011
***abi3-5*_21daf**	5,678,982	4,019,785 (70.8)	3,630,208 (91.5)	22,011
***abi3-1*_15daf**	5,963,995	4,355,175, (73.0)	3,993,489 (92.2)	22,011
***abi3-1*_17daf**	4,621,130	3,629,414 (78.5)	3,368,136 (93.0)	22,011
***abi3-1*_21daf**	3,632,793	2,806,398 (77.3)	2,568,817 (92.3)	22,011
Ler_15daf	8,408,818	6,327,513, (75.2)	5,860,618 (92.9)	22,011
Ler_17daf	5,538,070	4,125,875 (74.5)	3,777,419 (92.5)	22,011
Ler_21daf	5,241,362	4,017,069 (76.6)	3,695,767 (92.5)	22,011
***fus3-3*_15daf**	11,058,194	8,442,157 (76.3)	7,717,931 (93.2)	22,011
***fus3-3*_17daf**	13,368,320	10,544,256 (78.9)	9,452,736 (92.3)	22,011
***fus3-3*_21daf**	7,491,466	5,859,888 (78.2)	5,103,317 (91.9)	22,011
Col-0_15daf	6,365,373	4,862,900 (76.4)	4,495,715 (92.8)	22,011
Col-0_17daf	3,510,513	2,693,072 (76.7)	2,464,782 (92.4)	22,011
Col-0_21daf	8,101,628	6,195,345 (76.5)	5,796,084 (93.8)	22,011
Total reads	**162,305,761**	**125,117,037**	**111,829,621**	

Table S3. Generalized linear model fitted to desiccation tolerance seeds data.

Comparison Model	Interpretation	Developmental stage sampled	Differentially Expressed Genes
[(*lec1-1* - Ws) - (*fus3-3* - Col-0) - (*abi3-5* - Ler] **vs** [(*lec2-1* - Ws) - (*abi3-1*- Ler)]	Desiccation tolerance-specific mutants differences intolerant lines **vs** tolerant lines	15 DAF	Upregulated 434 Downregulated 548
[(*lec1-1* - Ws) - (*fus3-3* - Col-0) - (*abi3-5* - Ler] **vs** [(*lec2-1* - Ws) - (*abi3-1*- Ler)]		17 DAF	Upregulated 893 Downregulated 432
[(*lec1-1* - Ws) - (*fus3-3* - Col-0) - (*abi3-5* - Ler] **vs** [(*lec2-1* - Ws) - (*abi3-1*- Ler)]		21 DAF	Upregulate 993 Downregulated 481

Table S4. DEGs from pairwise comparison.

Comparison	Upregulated DEGs	Downregulated DEGs
***lec1-1* vs Ws 15 DAF**	1535	992
***lec1-1* vs Ws 17 DAF**	2164	1441
***lec1-1* vs Ws 21 DAF**	2397	1057
***lec2-1* vs Ws 15 DAF**	418	98
***lec2-1* vs Ws 17 DAF**	450	23
***lec2-1* vs Ws 21 DAF**	883	50
***abi3-5* vs Ler 15 DAF**	476	538
***abi3-5* vs Ler 17 DAF**	946	1128
***abi3-5* vs Ler 21 DAF**	2104	1660
***abi3-1* vs Ler 15 DAF**	2	0
***abi3-1* vs Ler 17 DAF**	425	397
***abi3-1* vs Ler 21 DAF**	0	0
***fus3-3* vs Col-0 15 DAF**	1610	1807
***fus3-3* vs Col-0 17 DAF**	2522	1429
***fus3-3* vs Col-0 21 DAF**	2066	1766

Table S5 . CEL. Files used in this study

Accession	Description of experiment	CEL_file
E-ATMX-1	Transcription profiling of Arabidopsis wild type seeds grown under sulfur-deficient conditions	CEL_C1.CEL
E-ATMX-1	Transcription profiling of Arabidopsis wild type seeds grown under sulfur-deficient conditions	CEL_C2.CEL
E-ATMX-1	Transcription profiling of Arabidopsis wild type seeds grown under sulfur-deficient conditions	CEL_DS1.CEL
E-ATMX-1	Transcription profiling of Arabidopsis wild type seeds grown under sulfur-deficient conditions	CEL_DS2.CEL
E-GEOD-1051	Transcription profiling by array of Arabidopsis seeds mutant for lec1	GSM10445.CEL
E-GEOD-1051	Transcription profiling by array of Arabidopsis seeds mutant for lec1	GSM10448.CEL
E-GEOD-1051	Transcription profiling by array of Arabidopsis seeds mutant for lec1	GSM10449.CEL
E-GEOD-1051	Transcription profiling by array of Arabidopsis seeds mutant for lec1	GSM10451.CEL
E-GEOD-1051	Transcription profiling by array of Arabidopsis seeds mutant for lec1	GSM10453.CEL
E-GEOD-1051	Transcription profiling by array of Arabidopsis seeds mutant for lec1	GSM10477.CEL
E-GEOD-1051	Transcription profiling by array of Arabidopsis seeds mutant for lec1	GSM10481.CEL
E-GEOD-11262	Transcription profiling by array of Arabidopsis seed compartments at the globular embryo stage	GSM284384.CEL
E-GEOD-11262	Transcription profiling by array of Arabidopsis seed compartments at the globular embryo stage	GSM284385.CEL
E-GEOD-11262	Transcription profiling by array of Arabidopsis seed compartments at the globular embryo stage	GSM284386.CEL
E-GEOD-11262	Transcription profiling by array of Arabidopsis seed compartments at the globular embryo stage	GSM284387.CEL
E-GEOD-11262	Transcription profiling by array of Arabidopsis seed compartments at the globular embryo stage	GSM284389.CEL
E-GEOD-	Transcription profiling by array of Arabidopsis seed	GSM284391.CEL

11262	compartments at the globular embryo stage	
E-GEOD-11262	Transcription profiling by array of Arabidopsis seed compartments at the globular embryo stage	GSM284392.CEL
E-GEOD-11262	Transcription profiling by array of Arabidopsis seed compartments at the globular embryo stage	GSM284394.CEL
E-GEOD-11262	Transcription profiling by array of Arabidopsis seed compartments at the globular embryo stage	GSM284395.CEL
E-GEOD-11262	Transcription profiling by array of Arabidopsis seed compartments at the globular embryo stage	GSM284396.CEL
E-GEOD-11262	Transcription profiling by array of Arabidopsis seed compartments at the globular embryo stage	GSM284397.CEL
E-GEOD-11262	Transcription profiling by array of Arabidopsis seed compartments at the globular embryo stage	GSM284398.CEL
E-GEOD-11852	Transcription profiling by array of Arabidopsis mutant for pickle after treatment with uniconazole	GSM299322.CEL
E-GEOD-11852	Transcription profiling by array of Arabidopsis mutant for pickle after treatment with uniconazole	GSM299323.CEL
E-GEOD-11852	Transcription profiling by array of Arabidopsis mutant for pickle after treatment with uniconazole	GSM299324.CEL
E-GEOD-11852	Transcription profiling by array of Arabidopsis mutant for pickle after treatment with uniconazole	GSM299325.CEL
E-GEOD-11852	Transcription profiling by array of Arabidopsis mutant for pickle after treatment with uniconazole	GSM299326.CEL
E-GEOD-11852	Transcription profiling by array of Arabidopsis mutant for pickle after treatment with uniconazole	GSM299333.CEL
E-GEOD-11852	Transcription profiling by array of Arabidopsis mutant for pickle after treatment with uniconazole	GSM299335.CEL
E-GEOD-11852	Transcription profiling by array of Arabidopsis mutant for pickle after treatment with uniconazole	GSM299336.CEL
E-GEOD-11852	Transcription profiling by array of Arabidopsis mutant for pickle after treatment with uniconazole	GSM299337.CEL
E-GEOD-12402	Transcription profiling by array of Arabidopsis seed compartments at the pre-globular stage	GSM311273.CEL
E-GEOD-12402	Transcription profiling by array of Arabidopsis seed compartments at the pre-globular stage	GSM311274.CEL

E-GEOD-12402	Transcription profiling by array of Arabidopsis seed compartments at the pre-globular stage	GSM311275.CEL
E-GEOD-12402	Transcription profiling by array of Arabidopsis seed compartments at the pre-globular stage	GSM311277.CEL
E-GEOD-12402	Transcription profiling by array of Arabidopsis seed compartments at the pre-globular stage	GSM311278.CEL
E-GEOD-12402	Transcription profiling by array of Arabidopsis seed compartments at the pre-globular stage	GSM311279.CEL
E-GEOD-12402	Transcription profiling by array of Arabidopsis seed compartments at the pre-globular stage	GSM311280.CEL
E-GEOD-12402	Transcription profiling by array of Arabidopsis seed compartments at the pre-globular stage	GSM311281.CEL
E-GEOD-12402	Transcription profiling by array of Arabidopsis seed compartments at the pre-globular stage	GSM311282.CEL
E-GEOD-12402	Transcription profiling by array of Arabidopsis seed compartments at the pre-globular stage	GSM311283.CEL
E-GEOD-12402	Transcription profiling by array of Arabidopsis seed compartments at the pre-globular stage	GSM311284.CEL
E-GEOD-12402	Transcription profiling by array of Arabidopsis seed compartments at the pre-globular stage	GSM311285.CEL
E-GEOD-12402	Transcription profiling by array of Arabidopsis seed compartments at the pre-globular stage	GSM311286.CEL
E-GEOD-12403	Transcription profiling by array of Arabidopsis seed compartments at the linear-cotyledon stage	GSM311288.CEL
E-GEOD-12403	Transcription profiling by array of Arabidopsis seed compartments at the linear-cotyledon stage	GSM311289.CEL
E-GEOD-12403	Transcription profiling by array of Arabidopsis seed compartments at the linear-cotyledon stage	GSM311290.CEL
E-GEOD-12403	Transcription profiling by array of Arabidopsis seed compartments at the linear-cotyledon stage	GSM311291.CEL
E-GEOD-12403	Transcription profiling by array of Arabidopsis seed compartments at the linear-cotyledon stage	GSM311292.CEL
E-GEOD-12403	Transcription profiling by array of Arabidopsis seed compartments at the linear-cotyledon stage	GSM311293.CEL
E-GEOD-12403	Transcription profiling by array of Arabidopsis seed compartments at the linear-cotyledon stage	GSM311294.CEL

E-GEOD-12403	Transcription profiling by array of Arabidopsis seed compartments at the linear-cotyledon stage	GSM311295.CEL
E-GEOD-12403	Transcription profiling by array of Arabidopsis seed compartments at the linear-cotyledon stage	GSM311296.CEL
E-GEOD-14229	Transcription profiling by array of Arabidopsis mutant for myb5	GSM356564.CEL
E-GEOD-14229	Transcription profiling by array of Arabidopsis mutant for myb5	GSM356565.CEL
E-GEOD-14229	Transcription profiling by array of Arabidopsis mutant for myb5	GSM356566.CEL
E-GEOD-14229	Transcription profiling by array of Arabidopsis mutant for myb5	GSM356567.CEL
E-GEOD-15160	Transcription profiling by array of Arabidopsis seed compartments at the heart stage	GSM378647.CEL
E-GEOD-15160	Transcription profiling by array of Arabidopsis seed compartments at the heart stage	GSM378649.CEL
E-GEOD-15160	Transcription profiling by array of Arabidopsis seed compartments at the heart stage	GSM378650.CEL
E-GEOD-15160	Transcription profiling by array of Arabidopsis seed compartments at the heart stage	GSM378651.CEL
E-GEOD-15160	Transcription profiling by array of Arabidopsis seed compartments at the heart stage	GSM378653.CEL
E-GEOD-15160	Transcription profiling by array of Arabidopsis seed compartments at the heart stage	GSM378654.CEL
E-GEOD-15160	Transcription profiling by array of Arabidopsis seed compartments at the heart stage	GSM378655.CEL
E-GEOD-15160	Transcription profiling by array of Arabidopsis seed compartments at the heart stage	GSM378656.CEL
E-GEOD-15160	Transcription profiling by array of Arabidopsis seed compartments at the heart stage	GSM378657.CEL
E-GEOD-15160	Transcription profiling by array of Arabidopsis seed compartments at the heart stage	GSM378658.CEL
E-GEOD-15160	Transcription profiling by array of Arabidopsis seed compartments at the heart stage	GSM378659.CEL
E-GEOD-15160	Transcription profiling by array of Arabidopsis seed compartments at the heart stage	GSM378660.CEL

E-GEOD-15165	Transcription profiling by array of Arabidopsis seed compartments at the mature green stage	GSM378733.CEL
E-GEOD-15165	Transcription profiling by array of Arabidopsis seed compartments at the mature green stage	GSM378734.CEL
E-GEOD-15165	Transcription profiling by array of Arabidopsis seed compartments at the mature green stage	GSM378735.CEL
E-GEOD-15165	Transcription profiling by array of Arabidopsis seed compartments at the mature green stage	GSM378736.CEL
E-GEOD-15165	Transcription profiling by array of Arabidopsis seed compartments at the mature green stage	GSM378737.CEL
E-GEOD-15165	Transcription profiling by array of Arabidopsis seed compartments at the mature green stage	GSM378738.CEL
E-GEOD-15165	Transcription profiling by array of Arabidopsis seed compartments at the mature green stage	GSM378739.CEL
E-GEOD-15165	Transcription profiling by array of Arabidopsis seed compartments at the mature green stage	GSM378740.CEL
E-GEOD-15165	Transcription profiling by array of Arabidopsis seed compartments at the mature green stage	GSM378741.CEL
E-GEOD-15165	Transcription profiling by array of Arabidopsis seed compartments at the mature green stage	GSM378742.CEL
E-GEOD-15165	Transcription profiling by array of Arabidopsis seed compartments at the mature green stage	GSM378743.CEL
E-GEOD-15165	Transcription profiling by array of Arabidopsis seed compartments at the mature green stage	GSM378744.CEL
E-GEOD-15165	Transcription profiling by array of Arabidopsis seed compartments at the mature green stage	GSM378745.CEL
E-GEOD-15165	Transcription profiling by array of Arabidopsis seed compartments at the mature green stage	GSM378746.CEL
E-GEOD-20039	Expression data from Arabidopsis seed compartments at the bending coyledon stage	GSM501157.CEL
E-GEOD-20039	Expression data from Arabidopsis seed compartments at the bending coyledon stage	GSM501158.CEL
E-GEOD-20039	Expression data from Arabidopsis seed compartments at the bending coyledon stage	GSM501159.CEL
E-GEOD-20039	Expression data from Arabidopsis seed compartments at the bending coyledon stage	GSM501160.CEL

E-GEOD-20039	Expression data from Arabidopsis seed compartments at the bending coyledon stage	GSM501161.CEL
E-GEOD-20039	Expression data from Arabidopsis seed compartments at the bending coyledon stage	GSM501162.CEL
E-GEOD-20039	Expression data from Arabidopsis seed compartments at the bending coyledon stage	GSM501163.CEL
E-GEOD-20039	Expression data from Arabidopsis seed compartments at the bending coyledon stage	GSM501164.CEL
E-GEOD-20039	Expression data from Arabidopsis seed compartments at the bending coyledon stage	GSM501165.CEL
E-GEOD-20039	Expression data from Arabidopsis seed compartments at the bending coyledon stage	GSM501166.CEL
E-GEOD-20039	Expression data from Arabidopsis seed compartments at the bending coyledon stage	GSM501167.CEL
E-GEOD-20039	Expression data from Arabidopsis seed compartments at the bending coyledon stage	GSM501168.CEL
E-GEOD-24887	Transcription profiling by array of Arabidopsis wild-type and dcl1-15 torpedo-staged embryos	GSM612043.CEL
E-GEOD-24887	Transcription profiling by array of Arabidopsis wild-type and dcl1-15 torpedo-staged embryos	GSM612044.CEL
E-GEOD-24887	Transcription profiling by array of Arabidopsis wild-type and dcl1-15 torpedo-staged embryos	GSM612045.CEL
E-GEOD-28747	Transcription profiling by array of Arabidopsis ft-1 mutants	GSM712386.CEL
E-GEOD-28747	Transcription profiling by array of Arabidopsis ft-1 mutants	GSM712387.CEL
E-GEOD-28747	Transcription profiling by array of Arabidopsis ft-1 mutants	GSM712388.CEL
E-GEOD-28747	Transcription profiling by array of Arabidopsis ft-1 mutants	GSM712392.CEL
E-GEOD-28747	Transcription profiling by array of Arabidopsis ft-1 mutants	GSM712393.CEL
E-GEOD-28747	Transcription profiling by array of Arabidopsis ft-1 mutants	GSM712394.CEL
E-GEOD-28747	Transcription profiling by array of Arabidopsis ft-1 mutants	GSM712395.CEL

E-GEOD-28747	Transcription profiling by array of Arabidopsis ft-1 mutants	GSM712396.CEL
E-GEOD-28747	Transcription profiling by array of Arabidopsis ft-1 mutants	GSM712397.CEL
E-GEOD-30223	Transcription profiling by array of Arabidopsis seeds during the germination ripening, stratification and germination	GSM748466_Har_rep_1_Arab.CEL
E-GEOD-30223	Transcription profiling by array of Arabidopsis seeds during the germination ripening, stratification and germination	GSM748467_Har_rep_2_Arab.CEL
E-GEOD-30223	Transcription profiling by array of Arabidopsis seeds during the germination ripening, stratification and germination	GSM748468_Har_rep_3_Arab.CEL
E-GEOD-30223	Transcription profiling by array of Arabidopsis seeds during the germination ripening, stratification and germination	GSM748469_0_h_dry_rep_1_Arab.CEL
E-GEOD-30223	Transcription profiling by array of Arabidopsis seeds during the germination ripening, stratification and germination	GSM748470_0_h_dry_rep_2_Arab.CEL
E-GEOD-30223	Transcription profiling by array of Arabidopsis seeds during the germination ripening, stratification and germination	GSM748471_0_h_dry_rep_3_Arab.CEL
E-GEOD-5634	Transcription profiling by array of Arabidopsis seeds and siliques at different developmental stages	GSM131694.CEL
E-GEOD-5634	Transcription profiling by array of Arabidopsis seeds and siliques at different developmental stages	GSM131695.CEL
E-GEOD-5634	Transcription profiling by array of Arabidopsis seeds and siliques at different developmental stages	GSM131696.CEL
E-GEOD-5634	Transcription profiling by array of Arabidopsis seeds and siliques at different developmental stages	GSM131697.CEL
E-GEOD-5634	Transcription profiling by array of Arabidopsis seeds and siliques at different developmental stages	GSM131698.CEL
E-GEOD-5634	Transcription profiling by array of Arabidopsis seeds and siliques at different developmental stages	GSM131699.CEL
E-GEOD-5634	Transcription profiling by array of Arabidopsis seeds and siliques at different developmental stages	GSM131700.CEL
E-GEOD-	Transcription profiling by array of Arabidopsis seeds	GSM131701.CEL

5634	and siliques at different developmental stages	
E-GEOD-5634	Transcription profiling by array of Arabidopsis seeds and siliques at different developmental stages	GSM131702.CEL
E-GEOD-5634	Transcription profiling by array of Arabidopsis seeds and siliques at different developmental stages	GSM131703.CEL
E-GEOD-5634	Transcription profiling by array of Arabidopsis seeds and siliques at different developmental stages	GSM131705.CEL
E-GEOD-5634	Transcription profiling by array of Arabidopsis seeds and siliques at different developmental stages	GSM131706.CEL
E-GEOD-5634	Transcription profiling by array of Arabidopsis seeds and siliques at different developmental stages	GSM131707.CEL
E-GEOD-5634	Transcription profiling by array of Arabidopsis seeds and siliques at different developmental stages	GSM131708.CEL
E-GEOD-5641	Transcription profiling by array of Arabidopsis mutant for mdh	GSM131848.CEL
E-GEOD-5641	Transcription profiling by array of Arabidopsis mutant for mdh	GSM131849.CEL
E-GEOD-5687	Transcription profiling by array of Arabidopsis seeds after incubation at different temperatures	GSM133117.CEL
E-GEOD-5687	Transcription profiling by array of Arabidopsis seeds after incubation at different temperatures	GSM133118.CEL
E-GEOD-5687	Transcription profiling by array of Arabidopsis seeds after incubation at different temperatures	GSM133119.CEL
E-GEOD-5687	Transcription profiling by array of Arabidopsis seeds after incubation at different temperatures	GSM133120.CEL
E-GEOD-5700	Transcription profiling by array of Arabidopsis after treatment with abscisic acid	GSM133304.CEL
E-GEOD-5700	Transcription profiling by array of Arabidopsis after treatment with abscisic acid	GSM133310.CEL
E-GEOD-5701	Transcription profiling by array of Arabidopsis after treatment with gibberellin	GSM133312.CEL
E-GEOD-5701	Transcription profiling by array of Arabidopsis after treatment with gibberellin	GSM133314.CEL
E-GEOD-5701	Transcription profiling by array of Arabidopsis after treatment with gibberellin	GSM133315.CEL

E-GEOD-5701	Transcription profiling by array of Arabidopsis after treatment with gibberellin	GSM133316.CEL
E-GEOD-5701	Transcription profiling by array of Arabidopsis after treatment with gibberellin	GSM133318.CEL
E-GEOD-5701	Transcription profiling by array of Arabidopsis after treatment with gibberellin	GSM133320.CEL
E-GEOD-5701	Transcription profiling by array of Arabidopsis after treatment with gibberellin	GSM133321.CEL
E-GEOD-5701	Transcription profiling by array of Arabidopsis after treatment with gibberellin	GSM133322.CEL
E-GEOD-5730	Transcription profiling by array of Arabidopsis embryo tissues	GSM133758.CEL
E-GEOD-5730	Transcription profiling by array of Arabidopsis embryo tissues	GSM133759.CEL
E-GEOD-5730	Transcription profiling by array of Arabidopsis embryo tissues	GSM133761.CEL
E-GEOD-5730	Transcription profiling by array of Arabidopsis embryo tissues	GSM133762.CEL
E-GEOD-5730	Transcription profiling by array of Arabidopsis embryo tissues	GSM133763.CEL
E-GEOD-5730	Transcription profiling by array of Arabidopsis embryo tissues	GSM133767.CEL
E-GEOD-5730	Transcription profiling by array of Arabidopsis embryo tissues	GSM133769.CEL
E-GEOD-5730	Transcription profiling by array of Arabidopsis embryo tissues	GSM133770.CEL
E-GEOD-5730	Transcription profiling by array of Arabidopsis embryo tissues	GSM133771.CEL
E-GEOD-5730	Transcription profiling by array of Arabidopsis embryo tissues	GSM133772.CEL
E-GEOD-5730	Transcription profiling by array of Arabidopsis embryo tissues	GSM133775.CEL
E-GEOD-5730	Transcription profiling by array of Arabidopsis embryo tissues	GSM133777.CEL
E-GEOD-5730	Transcription profiling by array of Arabidopsis embryo tissues	GSM133779.CEL

E-GEOD-5730	Transcription profiling by array of Arabidopsis embryo tissues	GSM133780.CEL
E-GEOD-5730	Transcription profiling by array of Arabidopsis embryo tissues	GSM133781.CEL
E-MEXP-1607	Transcription profiling by array of Arabidosis seeds after treatment with EL6 and GR24 to investigate the effect on germination	DMSO_rep1.CEL
E-MEXP-849	Transcription profiling of Arabidopsis seed and flowers to identify DELLA regulated transcripts	PJR-Dongni-seed-la-er-210404.CEL
E-MEXP-849	Transcription profiling of Arabidopsis seed and flowers to identify DELLA regulated transcripts	PJR-Dongni-seed-la-er-230304.CEL
E-TABM-1007	Transcription profiling by array of Arabidopsis mutant for fis2	caquinof_20080327_wt_2-v4.CEL
E-TABM-1007	Transcription profiling by array of Arabidopsis mutant for fis2	caquinof_20080327_wt_3-v4.CEL
E-TABM-17	Transcription profiling by array of organism parts from different strains of Arabidopsis	ATGE_79_A.CEL
E-TABM-17	Transcription profiling by array of organism parts from different strains of Arabidopsis	ATGE_79_B.CEL

Table S6. Number of nodes and edges in the TFsSeedNet and FullSeedNet obtained at DPI 0.0, 0.1 and 0.2

DPI	Nodes	Edges
TFsSeedNet		
0.0	2068	3745
0.1	2068	16608
0.2	2068	59079
FullSeedNet		
0.0	22553	89045
0.1	22553	372424
0.2	22553	2570778

Nodes and edges present at DPI 0.0 are part of the network obtained at DPI 0.1 and both are part of the network obtained at DPI 0.2.

Table S7. Motifs enrichment in DT subnetworks.

Genes of subnetworks	Motif identified	No. Genes with this element	No. of elements	P-value
snetFullDT1 at 15DAF	ABRE-like binding site motif	28	42	< 10 - 3
	AtMYC2 BS in RD22	42	59	< 10 - 4
	GADOWNAT	17	21	< 10 - 3
	RY-repeat promoter motif	9	20	< 10 - 3
	ACGTABREMOTIFA2OSEM	24	33	< 10 - 4
	CACGTGMOTIF	25	66	< 10 - 4
	MYCATERD1	42	59	< 10 - 4
snetFullDT1 at 17DAF	ABFs binding site motif	12	15	< 10 - 5
	ABRE-like binding site motif	44	85	< 10 - 10
	ACGTABREMOTIFA2OSEM	36	70	< 10 - 10
	DRE core motif	30	41	< 10 - 4
	GBF1/2/3 BS in ADH1	7	14	< 10 - 4
	RY-repeat promoter motif	11	24	< 10 - 4
	ABRE binding site motif	19	23	< 10 - 9
	ABREATRD22	10	12	< 10 - 5
	CACGTGMOTIF	36	108	< 10 - 10
	GADOWNAT	28	42	< 10 - 10
	GBOXLERBCS	12	15	< 10 - 6
	Z-box promoter motif	8	8	< 10 - 3
snetFullDT1 at 21 DAF	ABFs binding site motif	17	22	< 10 - 8
	ABRE-like binding site motif	59	113	< 10 - 10
	ACGTABREMOTIFA2OSEM	52	89	< 10 - 10
	CACGTGMOTIF	54	148	< 10 - 10
	GADOWNAT	35	46	< 10 - 10

	GBOXLERBCS	16	21	< 10 - 8
	RAV1-B binding site motif	21	23	< 10 - 3
	ABRE binding site motif	25	31	< 10 - 10
	ABREATRD22	12	14	< 10 - 5
	AtMYC2 BS in RD22	51	69	< 10 - 4
	DRE core motif	33	41	< 10 - 3
	GBF1/2/3 BS in ADH1	10	20	< 10 - 5
	MYCATERD1	51	69	< 10 - 4
	RY-repeat promoter motif	10	22	< 10 - 3
snetFullDT2 at 17DAF	ABFs binding site motif	20	24	< 10 - 10
	ABRE-like binding site motif	74	158	< 10 - 10
	ACGTABREMOTIFA2OSEM	67	129	< 10 - 10
	CBF1 BS in cor15a	5	5	< 10 - 3
	GADOWNAT	49	74	< 10 - 10
	GBOXLERBCS	19	23	< 10 - 10
	RY-repeat promoter motif	12	26	< 10 - 3
	ABRE binding site motif	32	40	< 10 - 10
	ABREATRD22	18	23	< 10 - 10
	CACGTGMOTIF	56	178	< 10 - 10
	DRE core motif	43	59	< 10 - 4
	GBF1/2/3 BS in ADH1	10	20	< 10 - 5
	Hexamer promoter motif	22	25	< 10 - 3
	Z-box promoter motif	18	18	< 10 - 10
snetFullDT2 at 21DAF	ABFs binding site motif	24	31	< 10 - 10
	ABRE-like binding site motif	78	169	< 10 - 9
	ACGTABREMOTIFA2OSEM	72	136	< 10 - 10
	CACGTGMOTIF	64	192	< 10 - 10
	GADOWNAT	51	72	< 10 - 9
	GBOXLERBCS	22	29	< 10 - 9
	Z-box promoter motif	15	15	< 10 - 6

	ABRE binding site motif	36	47	< 10 - 9
	ABREATRD22	19	24	< 10 - 9
	AtMYC2 BS in RD22	63	87	< 10 - 4
	DRE core motif	41	53	< 10 - 3
	GBF1/2/3 BS in ADH1	12	24	< 10 - 6
	MYCATERD1	63	87	< 10 - 4

Survey was performed in 1000 bp maximum upstream range cutting off at adjacent genes. Data obtained using the Athena Web tools (http://www.bioinformatics2.wsu.edu/cgibin/Athena/cgi/home.pl).

Table S8. Expression pattern of genes related with different process

Locus	Description	Fold-change *abi3-5*/Ler			Fold-change *fus3-3*/Col-0			Fold-change *lec1-1*/Ws			Fold-change *lec2-1*/Ws			Fold-change *abi3-1*/Ler		
		15DAF	17DAF	21DAF	15DAF	17DAF	21DAF	15DAF	17DAF	21DAF	15DAF	17DAF	21DAF	15DAF	17DAF	21DAF
Anthocyanins genes related																
AT5G13930	Chalcone and stilbene synthase family protein	0.966	-1.422	5.494	0.712	5.698	2.705	1.255	6.341	5.668	1.7	5.695	4.111	0.591	-4.687	1.845
AT5G07990	Cytochrome P450 superfamily protein	-1.12	-3.867	1.946	-0.57	1.786	-0.497	-0.765	0.413	1.655	-0.708	-0.182	1.028	1.044	-2.177	0.082
AT5G42800	dihydroflavonol 4-reductase	-7.402	-2.274	0.76	5.958	5.608	0.561	0.669	2.321	6.845	3.837	4.562	9.328	-2.699	-1.683	-4.872
AT3G51240	flavanone 3-hydroxylase	-1.294	-2.573	1.404	-0.78	1.624	2.593	-0.719	1.978	1.213	0.544	1.301	0.97	0.432	-1.73	0.819
AT5G17220	glutathione S-transferase phi 12	-2.315	-0.73	7.564	8.071	10.475	5.747	8.583	9.782	3.074	10.62	10.62	5.819	-0.816	-4.565	0
AT3G29590	HXXXD-type acyl-transferase family protein	0	0	7.073	7.763	9.183	-0.158	6.475	8.398	7.004	8.006	8.006	7.299	0	0	0
AT3G55970	jasmonate-regulated gene 21	3.203	-4.565	4.265	3.417	4.729	0.395	3.827	2.9	0	6.667	6.667	6.616	0	-4.565	4.179
AT4G22880	leucoanthocyanidin dioxygenase	-3.109	3.835	6.704	5.465	9.275	1.572	3.502	8.963	1.292	5.726	9.922	3.575	-6.312	0	0
AT1G66390	myb domain protein 90	3.203	3.835	3.776	3.798	1.158	1.1	6.324	1.868	5.583	8.72	3.492	9.092	3.756	0	-3.92
AT5G54060	UDP-glucose:flavonoid 3-o-glucosyltransferase	0	0	7.469	7.4	8.895	3.669	8	10.17	7.214	10.101	10.101	8.625	3.756	0	0
AT4G14090	UDP-Glycosyltransferase superfamily protein	0.936	0.744	3.444	1.458	4.488	1.957	1.71	4.596	1.499	3.167	4.393	2.055	-1.234	-2.093	0.588
AT5G49690	UDP-Glycosyltransferase superfamily protein	1.784	6.209	-1.252	4.745	-2.571	-0.314	1.295	-0.686	-1.445	4.524	0.526	-0.495	0.147	7.365	0.44
Chlorophyll synthesis genes related																
AT1G08520	ALBINA 1	2.582	3.87	3.116	2.512	1.463	0.512	9.278	5.347	2.842	9.261	4.995	1.667	0.477	0.431	-1.04
AT1G69740	Aldolase superfamily protein	1.48	0.909	1.642	1.674	1.472	2.122	1.92	1.926	0.684	1.083	1.375	0.069	0.347	-0.638	0.339
AT1G03475	Coproporphyrinogen III oxidase	0.666	0.922	1.653	1.897	1.673	1.056	1.534	1.948	0.493	0.897	1.423	-0.251	-0.486	0.826	0.089
AT3G56940	dicarboxylate diiron protein, putative (Crd1)	1.927	2.314	7.571	1.559	6.11	2.684	4.667	6.078	5.899	5.307	5.513	5.97	0.658	-4.339	0.684
AT3G59400	enzyme binding;tetrapyrrole binding	1.093	1.598	3.152	0.181	6.445	2.991	7.57	8.425	6.462	8.623	8.623	7.686	-0.947	-5.534	1.218
AT2G30390	ferrochelatase 2	1.581	1.468	3.169	1.121	1.884	0.087	0.941	1.177	0.882	1.013	1.201	0.841	1.229	0.587	0.271
AT1G48520	GLU-ADT subunit B	1.865	3.007	2.294	2.729	2.45	0.034	1.714	-0.357	0.854	-0.22	-0.614	0.918	0.892	1.64	0.082
AT3G48730	glutamate-1-semialdehyde 2,1-aminomutase 2	1.871	2.131	2.776	3.442	2.443	2.221	3.658	4.595	1.614	1.955	2.821	0.39	-0.183	0.246	0.376
AT5G63570	glutamate-1-semialdehyde-2,1-aminomutase	0.99	1.086	1.573	3.187	2.275	0.68	2.541	2.35	1.356	1.41	1.069	0.503	-0.091	0.334	0.254
AT1G58290	Glutamyl-tRNA reductase family protein	1.437	3.427	3.756	1.25	8.432	1.542	5.787	4.925	4.986	5.458	4.426	5.345	-0.686	-1.109	-4.872
AT2G26550	heme oxygenase 2	0.982	5.757	1.717	2.398	0.727	-0.162	1.73	7.787	2.685	-1.168	5.546	2.283	0.517	5.965	-5.847
AT5G08280	hydroxymethylbilane synthase	1.707	1.794	2.497	2.658	1.035	3.219	2.56	4.324	1.373	1.378	2.711	0.282	-0.744	0.681	0.422
AT5G45930	magnesium chelatase i2	2.087	4.783	8	2.626	7.563	6.156	7.393	7.654	6.348	5.546	5.546	5.915	-4.549	0	0

AT5G13630	magnesium-chelatase subunit chlH, chloroplast, putative / Mg-protoporphyrin IX chelatase, putative (CHLH)	0.869	2.169	10.135	0.663	7.767	0.532	4.011	9.884	4.951	5.553	9.749	6.385	0.597	-6.518	0
AT4G25080	magnesium-protoporphyrin IX methyltransferase	1.714	2.51	2.781	1.22	1.977	2.273	3.528	4.141	1.606	2.595	3.295	0.939	0.637	-3.436	0.274
AT5G18660	NAD(P)-binding Rossmann-fold superfamily protein	1.892	3.03	9.36	1.623	9.191	3.342	3.585	5.846	8.857	2.202	4.39	7.568	-0.161	-4.565	0
AT4G18480	P-loop containing nucleoside triphosphate hydrolases superfamily protein	1.44	1.604	10.468	1.898	4.672	3.485	9.685	9.923	3.056	9.102	9.102	3.083	-0.423	-1.928	0
AT5G54190	protochlorophyllide oxidoreductase A	-0.427	0.913	5.792	4.909	4.509	8.82	3.292	6.173	3.849	4.27	4.27	4.806	-4.549	-5.534	-4.872
AT4G27440	protochlorophyllide oxidoreductase B	2.405	4.195	8.658	2.528	6.791	1.84	4.83	6.372	3.438	5.545	5.732	4.203	-2.23	-1.127	2.515
AT1G03630	protochlorophyllide oxidoreductase C	3.141	6.209	9.523	3.499	8.378	4.128	8.229	3.696	3.513	8.688	3.46	3.995	1.093	4.425	4.179
AT5G63290	Radical SAM superfamily protein	0.28	5.924	-0.218	1.342	-1.803	1.797	6.674	5.721	-0.908	5.804	5.804	-0.739	0.561	6.953	-0.308
AT1G45110	Tetrapyrrole (Corrin/Porphyrin) Methylases	2.042	1.359	1.878	0.721	1.296	3.064	5.862	2.144	0.34	3.343	-0.924	-0.093	0.147	-4.565	2.197
AT3G14110	Tetratricopeptide repeat (TPR)-like superfamily protein	1.341	1.313	0.998	1.813	2.153	2.763	2.142	3.212	2.05	1.152	1.431	0.928	-0.237	-0.435	0.518
AT3G51820	UbiA prenyltransferase family protein	1.843	1.889	4.334	2.623	2.47	1.87	3.153	4.349	2.71	3.688	4.596	2.65	0.847	-1.354	0.052
AT2G40490	Uroporphyrinogen decarboxylase	0.925	0.226	2.773	2.039	0.764	1.483	1.112	3.005	1.531	0.623	1.658	0.601	-0.573	-2.288	0.398
AT3G14930	Uroporphyrinogen decarboxylase	0.514	0.359	2.934	1.504	1.665	1.052	1.369	2.435	1.388	0.915	1.452	0.976	0.293	-0.386	1.195

Leaf development genes related

AT1G63650	basic helix-loop-helix (bHLH) DNA-binding superfamily protein	0.541	6.743	6.863	0.394	1.965	-0.158	1.257	2.059	6.224	-0.113	0.794	0	-5.518	0	0
AT5G41315	basic helix-loop-helix (bHLH) DNA-binding superfamily protein	0	0	5.226	2.106	0.512	-1.667	4.522	0	5.583	3.343	3.343	0	0	0	0
AT1G79840	HD-ZIP IV family of homeobox-leucine zipper protein with lipid-binding START domain	1.386	0.34	1.703	0.893	1.697	3.361	3.539	5.366	2.42	2.373	3.373	1.252	0.012	-1.579	0.362
AT3G61150	homeodomain GLABROUS 1	-0.034	-1.052	1.126	1.524	0.607	0.54	-0.064	7.869	1.404	-2.276	5.546	0.493	0.158	-1.94	1.269
AT1G17920	homeodomain GLABROUS 12	-1.231	3.835	5.948	2.242	5.408	1.572	5.625	-0.481	3.305	0	-5.228	-0.48	-1.61	0	4.179
AT1G05230	homeodomain GLABROUS 2	-1.03	0.12	1.5	0.585	0.548	0.866	7.466	0.901	0.892	7.11	0.102	0.926	-0.162	0.215	0.969
AT2G46410	Homeodomain-like superfamily protein	4.122	4.385	3.717	3.417	5.868	5.581	1.295	6.173	2.426	-0.854	3.343	-0.48	3.756	4.425	5.138
AT5G53200	Homeodomain-like superfamily protein	4.679	4.385	2.822	-0.32	5.295	-1.043	1.168	1.266	5.366	0.39	1.28	0	3.756	0	0
AT5G11060	KNOTTED1-like homeobox gene 4	-0.427	-0.73	5.443	4.348	0.327	-2.103	5.342	1.771	2.931	7.203	2.936	3.495	0.719	0.826	0
AT4G08150	KNOTTED-like from Arabidopsis thaliana	3.141	7.595	4.01	6.294	0.014	-0.161	5.862	1.068	-0.237	5.546	0.318	-0.997	1.093	4.425	2.197
AT3G27920	myb domain protein 0	0	0	2.822	2.898	2.883	0	4.522	0	4.4	4.27	4.27	0	0	0	0
AT2G26580	plant-specific transcription factor YABBY family protein	1.299	-1.321	6.786	-2.365	7.237	-0.809	-0.212	1.08	6.666	-1.423	-0.504	6.258	0.081	-2.272	0
AT2G45190	Plant-specific transcription factor YABBY family protein	-0.094	-0.357	3.253	-1.838	2.389	0.112	-0.766	1.04	4.376	-1.039	0.132	3.687	0.183	-1.589	-5.44

Antioxidants genes related

AT1G67600	Acid phosphatase/vanadium-dependent haloperoxidase-related protein	-0.86	-2.261	-4.882	-2.899	-2.986	-2.333	-1.817	-3.045	-4.31	-0.296	-1.047	-1.214	-0.738	0.852	-0.189
AT1G08570	atypical CYS HIS rich thioredoxin 4	-1.269	-2.709	-1.996	-2.309	-2.894	-1.258	-2.593	-2.05	-2.004	-0.263	-0.864	-1.189	-0.46	-0.946	0.396
AT1G20630	catalase 1	-0.943	-1.576	-1.875	-0.87	-0.406	-2.861	-2.153	-2.638	-1.182	-0.583	-0.705	-0.283	-0.056	-0.126	-0.081

AT4G35090	catalase 2	2.202	2.12	0.111	1.997	-1.167	1.314	0.54	0.331	-1.24	1.249	0.387	-0.815	0.195	0.6	0.547
AT1G08830	copper/zinc superoxide dismutase 1	-0.5	-0.151	-0.277	-4.661	-1.978	-3.23	-1.353	-2.297	-0.587	-0.543	-0.588	-0.616	-1.028	0.296	-0.512
AT2G28190	copper/zinc superoxide dismutase 2	-1.529	-0.886	-1.636	-7.577	-6.414	-4.189	-3.327	-3.465	-2.001	-0.201	-0.67	-0.761	-1.435	-0.092	-0.108
AT5G18100	copper/zinc superoxide dismutase 3	0.658	0.993	-2.974	-1.387	-1.088	0.6	0.322	-0.539	-0.784	0.517	-0.42	-0.828	-0.517	1.137	-1.262
AT3G63080	glutathione peroxidase 5	0.309	0.12	-0.295	-1.679	-0.463	-0.624	-0.403	0.006	-0.866	-0.087	0.604	0.026	0.389	0.01	-0.19
AT4G31870	glutathione peroxidase 7	-4.549	-6.109	-3.342	-2.8	-4.409	-1.667	-1.864	-2.849	-3.053	1.224	-0.819	-1.988	-0.793	1.065	0.943
AT2G18950	homogentisate phytyltransferase 1	-0.684	-0.77	2.659	-1.373	1.554	-2.036	-0.197	-0.805	1.636	0.398	0.155	2.317	0.159	-0.991	-1.262
AT3G56350	Iron/manganese superoxide dismutase family protein	-8.603	-7.674	-8.801	-3.198	-3.207	-1.255	-2.686	-2.647	-2.818	-0.155	-0.29	-1.221	-0.549	-0.413	-0.572
AT1G54870	NAD(P)-binding Rossmann-fold superfamily protein	-5.222	-6.019	-8.25	-2.15	-3.945	-2.077	-3.71	-3.928	-2.984	0.036	-0.956	-1.236	-0.381	0.969	0.277
AT1G68850	Peroxidase superfamily protein	-1.147	-1.758	-1.155	0.442	5.295	-7.184	-1.153	-1.871	0	-1.005	-1.079	0	0.245	0.338	-4.872
AT1G71695	Peroxidase superfamily protein	-0.073	-1.647	-0.108	-2.793	1.65	-1.417	-1.445	-0.651	1.903	-0.985	0.186	1.138	0.021	-2.858	-0.198
AT1G77100	Peroxidase superfamily protein	1.688	2.071	-0.82	0.682	1.697	-3.322	-0.293	-0.433	0.997	0.989	0.097	2.2	-0.272	3.465	-1.493
AT2G38380	Peroxidase superfamily protein	-1.569	-3.705	-2.879	-0.247	2.191	-7.233	-1.203	-1.926	-5.6	-1.176	-1.365	0.658	0.054	-0.897	-0.443
AT3G01420	Peroxidase superfamily protein	-1.991	-2.541	-2.819	-0.207	1.324	-6.64	-1.665	-2.748	1.292	-0.431	-0.877	2.543	-0.384	-0.126	-1.878
AT4G21960	Peroxidase superfamily protein	0.645	-0.009	3.452	-2.245	3.62	-0.947	-1.525	-0.096	4.365	-0.684	0.39	3.998	-0.261	-1.858	-1.237
AT4G25980	Peroxidase superfamily protein	-3.003	-4.361	4.265	-4.026	5.868	-6.262	-0.924	1.137	-1.181	-2.871	1.081	-0.533	-0.037	-3.616	0
AT4G37520	Peroxidase superfamily protein	-1.912	-2.385	-3.025	-2.475	0.038	-6.297	-1.733	-2.06	0	-0.918	-0.491	2.677	-0.266	0.158	-5.847
AT4G37530	Peroxidase superfamily protein	-3.609	-6.109	0	-3.794	5.295	-8.304	-2.951	-2.307	0	-2.895	-2.648	0	-3.609	-6.109	0
AT5G42180	Peroxidase superfamily protein	-1.725	-0.73	-3.92	-0.024	4.893	-7.048	-0.971	-1.15	2.946	-3.176	-1.939	0	-0.597	2.388	-3.92
AT2G23240	Plant EC metallothionein-like protein, family 15	-0.447	-2.275	-7.786	0.911	-3.732	2.328	-2.867	-3.17	-3.118	0.611	-1.223	-1.53	0.421	2.288	-0.303
AT2G42000	Plant EC metallothionein-like protein, family 15	-0.802	-2.61	-6.861	-2.642	-4.001	-0.87	-1.367	-2.424	-1.678	0.302	-1.076	-0.683	-0.566	0.951	-0.633
AT1G45145	thioredoxin H-type 5	0.384	-0.587	-1.606	-1.153	-1.387	1.749	0.226	-0.856	-0.9	0.487	-0.494	-0.678	0.159	1.3	0.514
AT4G32770	tocopherol cyclase, chloroplast / vitamin E deficient 1 (VTE1) / sucrose export defective 1 (SXD1)	-1.269	-2.247	-1.078	-1.735	0.024	-2.699	-1.343	-1.635	-0.595	-0.367	-0.398	-0.171	0.128	-1.351	-0.386

DNA repair genes related

AT3G50880	DNA glycosylase superfamily protein	-0.32	-0.897	-1.825	-1.429	-1.191	-0.834	-1.116	-1.236	-0.666	-0.554	-0.904	-0.299	0.318	0.143	0.439
AT5G16630	DNA repair protein Rad4 family	-0.085	1.027	-1.403	0.44	-4.209	-1.667	0.021	-1.423	-1.763	1.394	0.103	-1.198	0.348	2.301	-0.521
AT1G48620	high mobility group A5	-0.488	-0.236	-2.058	0.419	-0.678	-1.001	-0.903	-0.711	-0.874	0.477	0.203	-0.182	-0.223	0.99	0.213
AT1G08170	Histone superfamily protein	-6.957	-6.929	-7.996	-2.374	-6.475	-1.465	-4.463	-4.945	-4.802	1.762	-0.614	-2.164	-0.297	0.278	0.036
AT3G27360	Histone superfamily protein	-1.764	-0.768	-2.97	-1.749	-1.912	-0.077	0.041	-0.373	-1.104	-0.505	-0.497	-0.421	-1.741	2.433	0.137
AT5G63190	MA3 domain-containing protein	0.071	0.121	-1.628	1.716	-1.129	-0.753	-0.272	-0.784	-0.6	0.226	-0.396	-0.98	-0.336	0.867	0.301
AT4G29170	Mnd1 family protein	1.677	-0.185	-3.471	1.957	-1.19	1.705	6.243	0.406	-1.754	5.232	-0.977	-0.926	0.719	1.065	0.274
AT1G58060	RNA helicase family protein	-1.711	1.508	-0.938	-0.946	4.332	-3.379	-1.146	-1.15	-0.046	-0.113	-0.186	-1.322	-0.565	-0.14	0.271

Table S9. Genes and T-DNA insertion seeds (SALK Line) examined from the snetFullDT1 and snetFullDT2 for seed germination phenotypes

Locus	Aliases	Description	SALK lines
AT1G21000	PLATZ1	PLATZ transcription factor family protein	SALK_005374
AT1G76590	PLATZ2	PLATZ transcription factor family protein	SALK_016183
AT4G26050	PIRL8	plant intracellular ras group-related LRR 8	SALK_001827
AT1G01720	ATAF1	NAC (No Apical Meristem) domain transcriptional regulator superfamily protein	SALK_067648
AT2G47180	GOLS1	galactinol synthase 1	SALK_013103 SALK_121059
AT1G56600	GOLS2	galactinol synthase 2	SALK_075769 SALK_101144
AT1G77950	AGL67	AGAMOUS-like 67	SALK_050367 SALK_013790
AT5G01670	NRS1	NAD(P)-linked oxidoreductase superfamily protein	SALK_150592 SALK_111609
AT4G01970	STS	stachyose synthase	SALK_021300
AT1G01250	ERF23	Integrase-type DNA-binding superfamily protein	SALK_044761
AT1G75490	DREB2D	Integrase-type DNA-binding superfamily protein	SALK_023056
AT3G54510	ERD4	Early-responsive to dehydration stress protein (ERD4)	SALK_055548
AT4G04840	MSRB6	methionine sulfoxide reductase B6	SALK_039712
AT2G19320		unknown protein	SALK_092667
AT5G54070	HSF9	heat shock transcription factor A9	SALK_063950
AT1G05510	OBAP1A	OBAP1A, oil body-associated protein 1A	SALK_017397 SALK_011689

Table S10. Primers used in this study

Primers	5´-3’
	Genotyping SALK lines
SALK_005374	Fw TGATTGTGGTTTTCGATAGCC Rv CACGTTAAGTAACGTGGCAGC
SALK_016183	Fw ATCATGAGCATTTCAAATGCC Rv GAATGAAGCGTGATCCAAGTC
SALK_001827	Fw ACCATTTGATTCGGCCTATTC Rv TTTGAGTTGACGGAGAGCTTC
SALK_067648	Fw TCCCAGGGACAGAAAATATCC Rv AAATATTAAATTGATTGCGGCAC
SALK_013103	Fw TTCGAAACAAAAATTGAACCG Rv TTGTTTGACCTACCAGATGG
SALK_121059	Fw ATGAATGTATGTTTCGCAGGC Rv TTGTTTGACCTACCAGATGG
SALK_075769	Fw TCATAGACTTGATTGGTTTCCG Rv AACAACACAGCCTTGATCCAC
SALK_101144	Fw TCATAGACTTGATTGGTTTCCG Rv AACAACACAGCCTTGATCCAC
SALK_050367	Fw AAAGGTATCGGATGTTTTCGG Rv TTGGTAACAGCTTATGGACCC
SALK_013790	Fw TACAGTCCAGAAGCCACCAAC Rv ATTTCTCGTGGCAACGTTATG
LB3.1	ATTTTGCCGATTTCGGAAC
	cDNA (CDS)
PLATZ1	Fw GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGATAAGAACAGAGGAAGAAGAAGA Rv GGGGACCACTTTGTACAAGAAAGCTGGGTCTTAGAACGGAGCACGGTGAG
PLATZ2	Fw GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGGGACCGATGATGATGAGAGC Rv GGGGACCACTTTGTACAAGAAAGCTGGGTCTTAGAAAGGAGCACGGTGTGGGA
AGL67	Fw GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGGGTCGGGTTAAATTGGAGT Rv GGGGACCACTTTGTACAAGAAAGCTGGGTCTCATTTTGGATTTTGGTTCGAA
DREB2G	Fw GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGGAAGAAGAGCAACCTCCG Rv GGGGACCACTTTGTACAAGAAAGCTGGGTCTCAGAACCAATTCCATGGATGTTGA
DREB2D	Fw GGGGACAAGTTTGTACAAAAAAGCAGGCTTCACATGTCATCCATAGAGCCAAA Rv GGGGACCACTTTGTACAAGAAAGCTGGGTCGTGACACCTCAAAGTGGGGA
	Real Time PCR
ABI5	Fw AAACATGCATTGGCGGAGTT Rv TCAATGTCCGCAATCTCCCG
ATAF1	Fw TCAGGCTGGATGATTGGGTT Rv GGCGGAGGCATAACCATCTC
UGE3	Fw GGAGCTCACGAGAGTGGAAG Rv CTAACCGCGCTACCATCCAT
TTP5	Fw GCCAAAGAGCTTATGGAACACC Rv CCTTTGTTCACACCCTGTGG
LEA4	Fw GAGAAACGCGTCAGCACAAC Rv ATCTGATGTGTCCCAGTGCC
ChlADR	Fw CCAATTGAAGTGGCACCATCC Rv CCACAGCTCCTCCATTAGGG
PLATZ2	Fw TCGGCAAAGGAGTCACCAAT Rv TCCTCCAAGTTTGCAACCGA
ABI3	Fw GAACACCGGCGATTTTGTGA

	Rv TTTTGTCCGCTCGGTTGTCT
GOLS1	Fw GTGTACAATCTCGTCCTTGCG Rv AACCCGCTGCACAGTAGTGA
PLATZ1	Fw TGCAAGCTTGGAGGAATGAGG Rv AGCTTCATCTGATTCCGACCC
UBQ10	Fw GCGTCTTCGTGGTGGTTTCTA Rv GTCGAGTCACTTTGCAGGC
TIP4L	Fw AGTCATGCCAAGCTCATGGT Rv TCAACTGGATACCCTTTCGCA

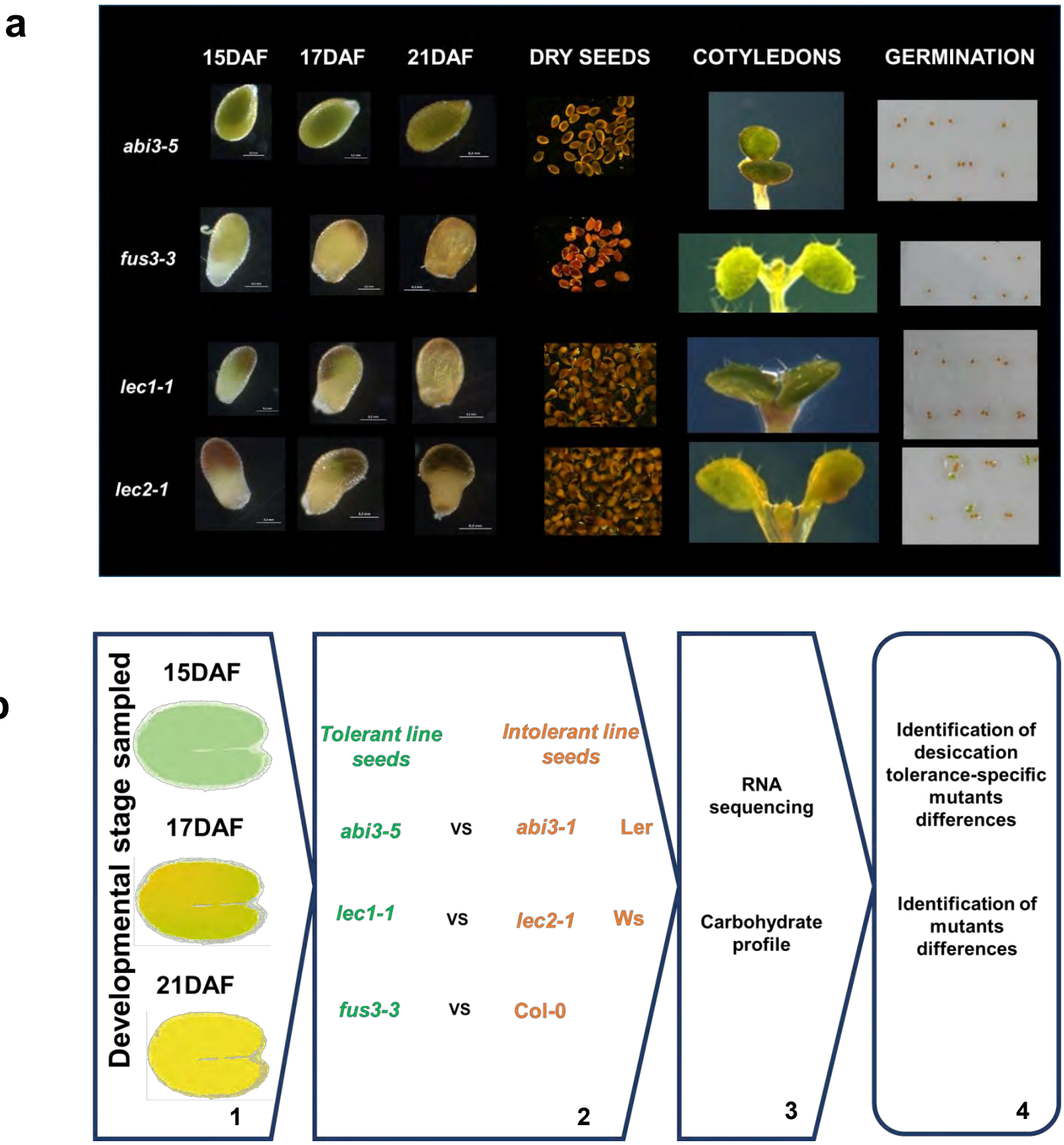


Figure S1. Phenotype of the different lines and general strategy used in the work. (a) Seed phenotypes of each mutant at 15,17 ,21 DAF, dry seeds, trichome cotyledons and germination. (b) General approach used to identify desiccation tolerance differences.1) Seeds were sampled at 15, 17 and 21DAF 2) Different lines: desiccation intolerant mutants, tolerant mutants and wildtypes were collected. Contrasts were performed tolerant line seeds vs intolerant line seeds at 15, 17 and 21 DAF. 3) RNA sequencing and carbohydrate profiles were performed from each lines sampled. 4) To RNA sequencing, the DT-specific differences were identified by generalized linear model and to carbohydrates analysis, the DT–specific differences were identified by pair comparision (mutant *vs* wildtype)

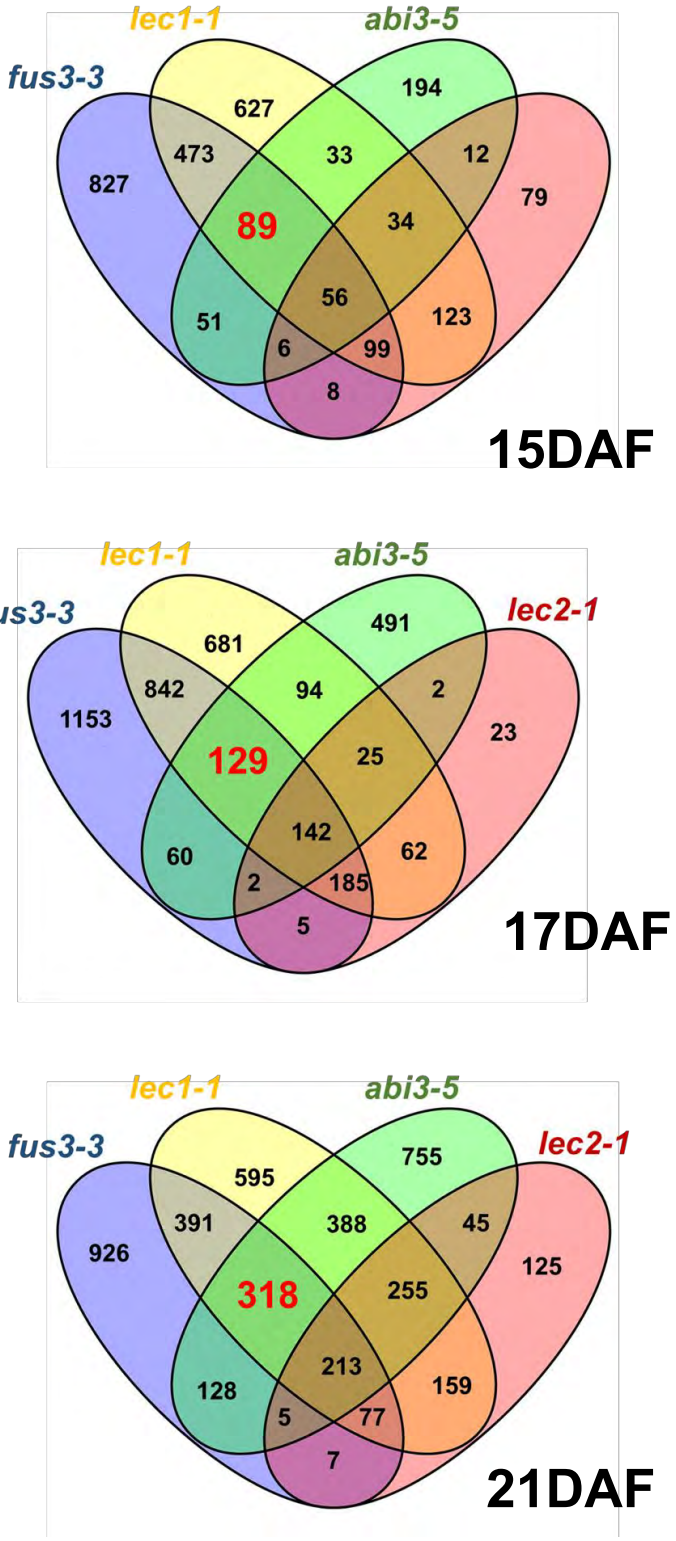


Figure S2. Venn diagrams of upregulated genes. Venn diagrams showing the number and distribution of differentially upregulate genes across the tested mutants lines at 15, 17 and 21 DAF, respectively. The number of shared genes is indicated in red.

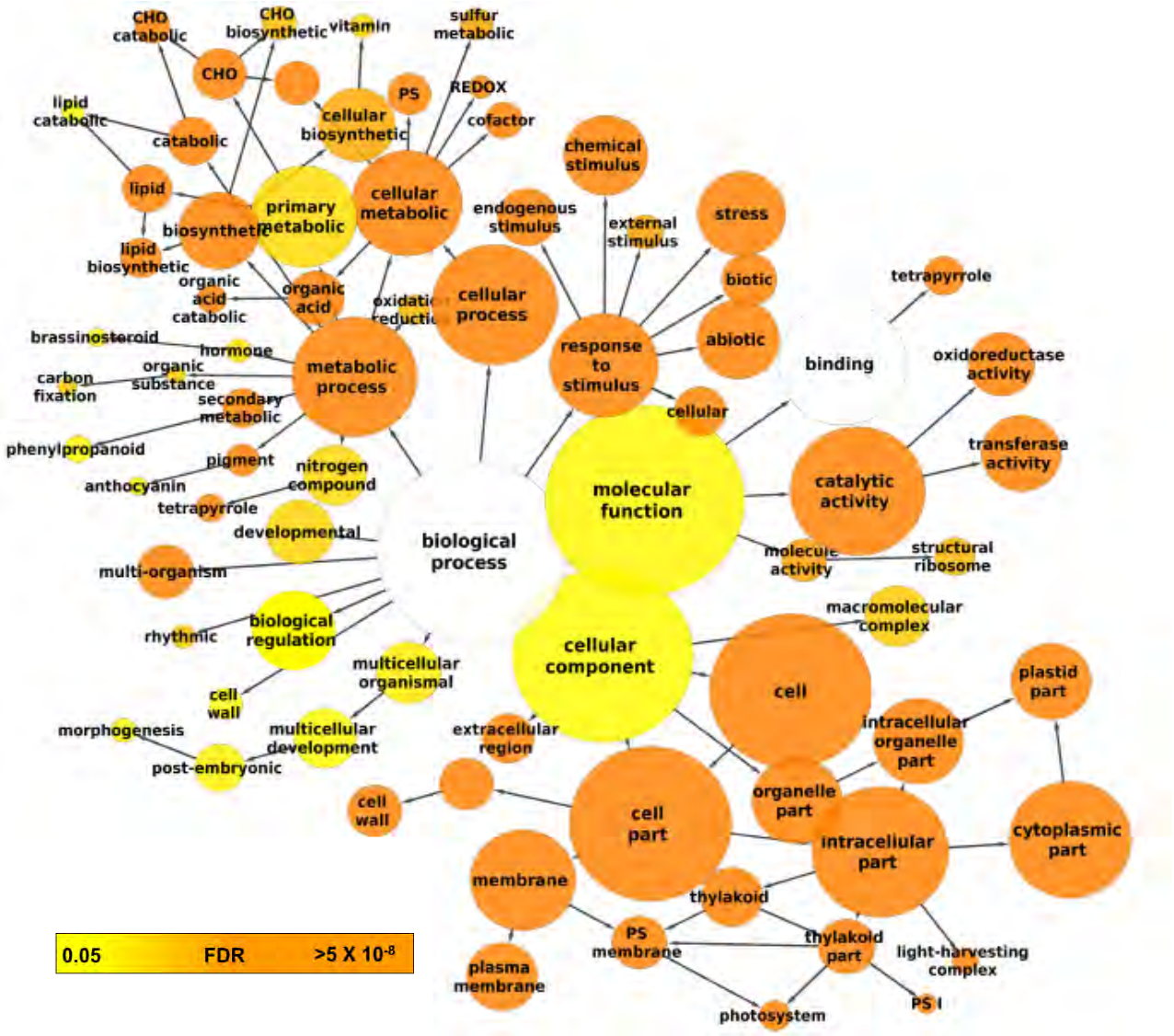


Figure S3. Enriched GO Terms of upregulated genes in *lec1, lec2, fus3 and abi3-5*. The graph is a BiNGO visualization of overrepresented GO terms for upregulated (2,218 genes) present in Supplementary Tables 5,6 and 7. Uncolored nodes are not overrepresented, but they may be the parents of overrepresented terms. Colored nodes represent GO terms that are significantly overrepresented (Benjamini and Hochberg corrected P value < 0.05), the node color indicates significance as shown in the color bar. A more detailed analysis of the GO categories is shown in Supplemental Table 12. PS, Photosynthesis; CHO carbohydrates.

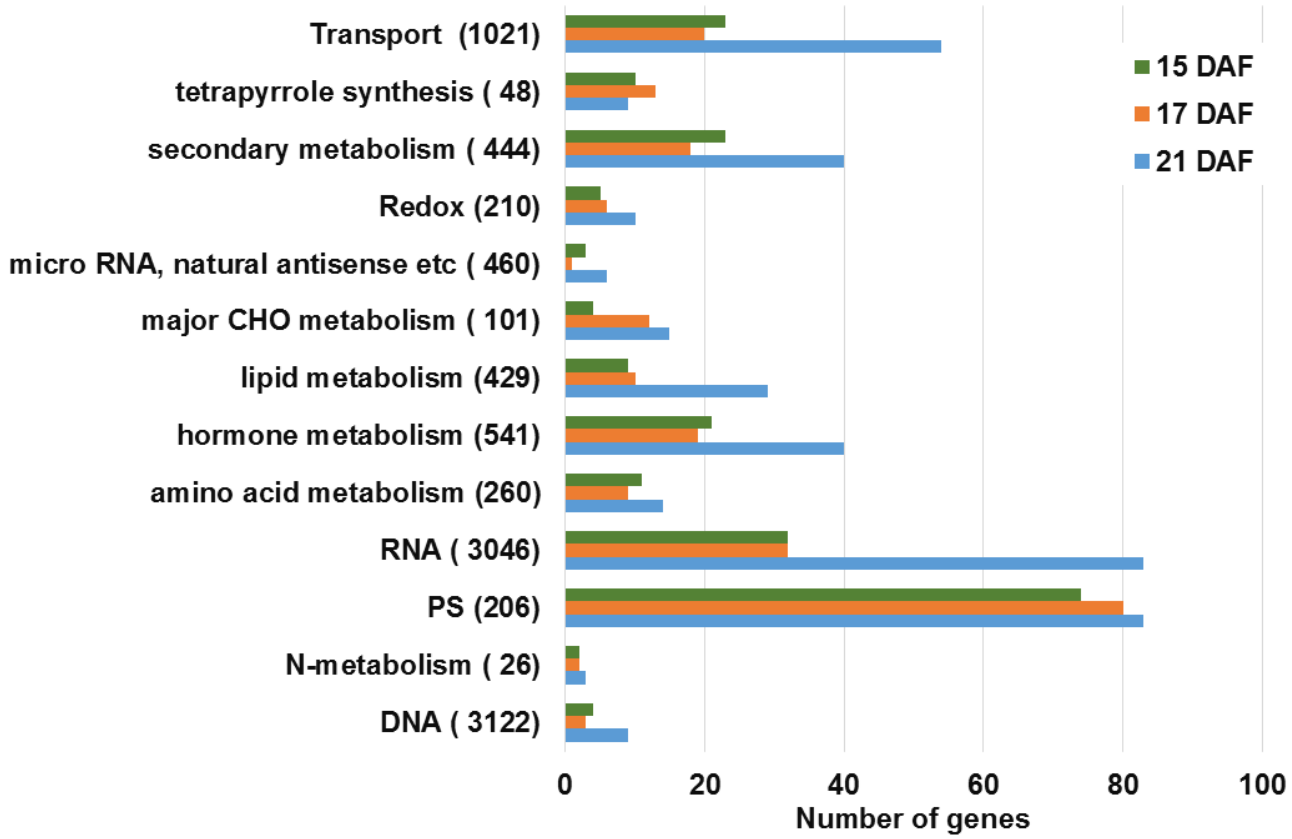


Figure S4. **MAPMAN-based functional classification of upregulated DEGs in *lec1, lec2, fus3 and abi3-5* during seed desiccation tolerance.** (a) MAPMAN classification of the transcripts using the web-tool Classification Superviewer (http://bar.utoronto.ca). All categories are significant (P value <0.05). CHO, carbohydrate; PS, photosynthesis.

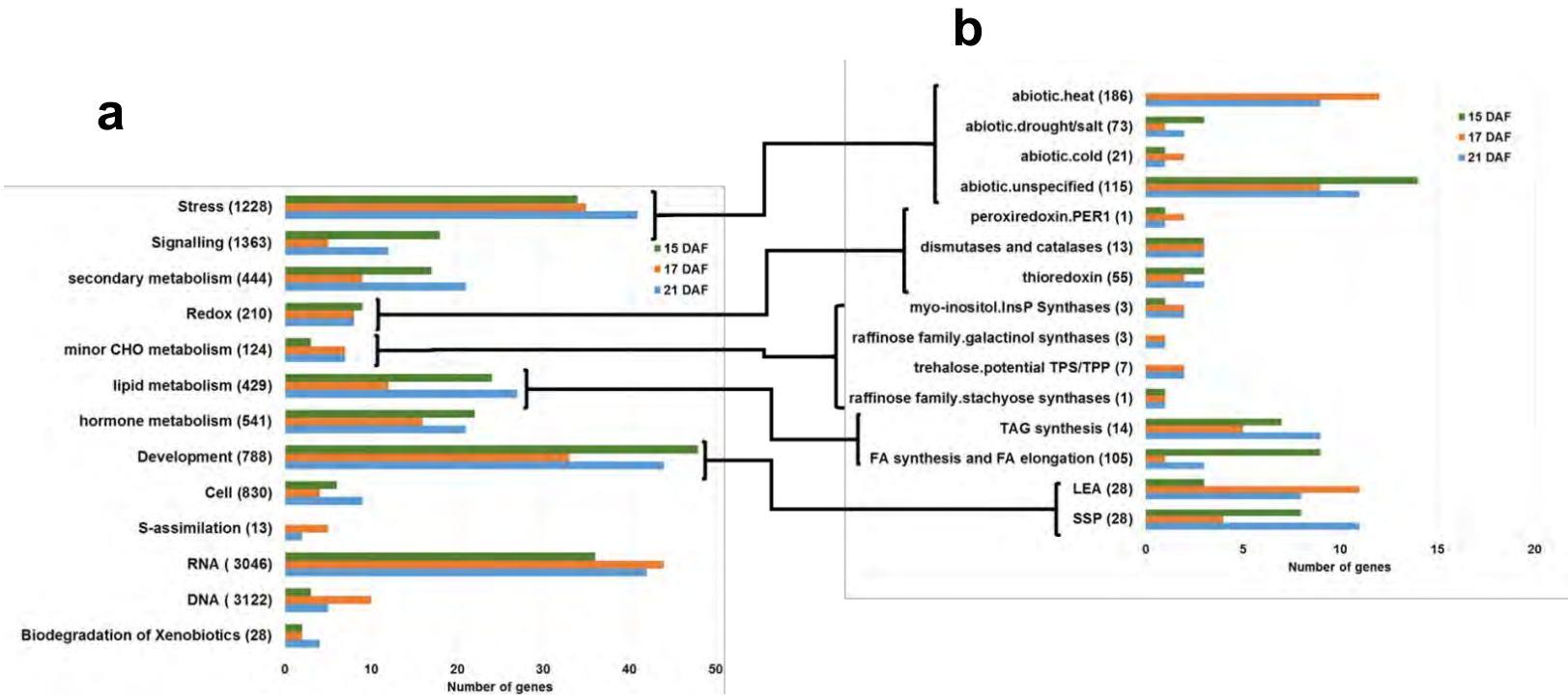


Figure S5. **MAPMAN-based functional classification of downregulated DEGs in desiccation intolerant lines during seed desiccation tolerance.** MAPMAN (a) categories and (b) subcategories classification of the downregulated transcripts using the web-tool Classification Superviewer (http://bar.utoronto.ca). All categories are significant (P value <0.05). CHO, carbohydrate; TPS, Trehalose-6-Phosphate Synthase; TPP, Trehalose-6-Phosphate Phosphatase; TAG, triacylglycerol; LEA, late embryogenesis abundant; SSP, storage seed proteins.

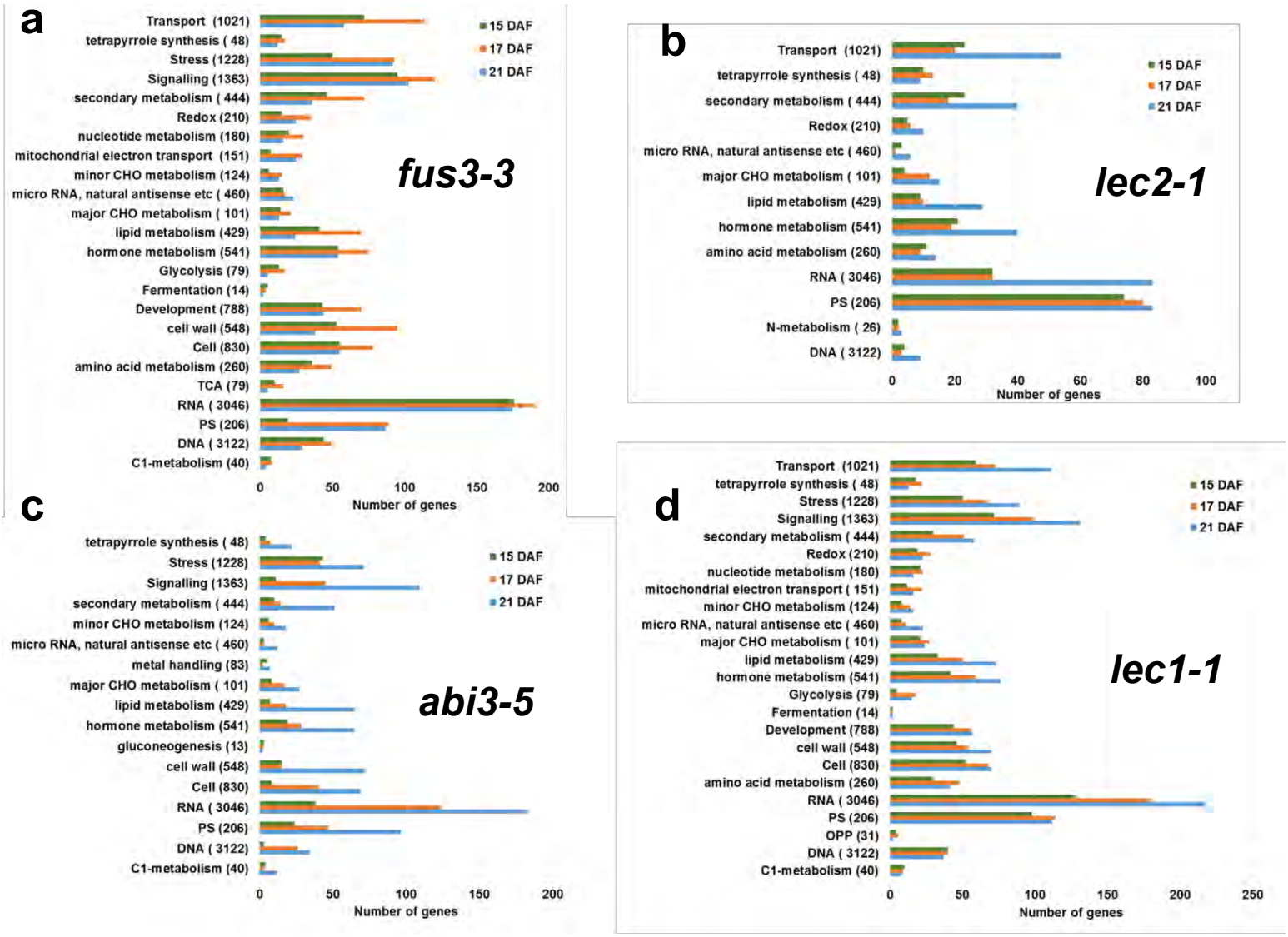
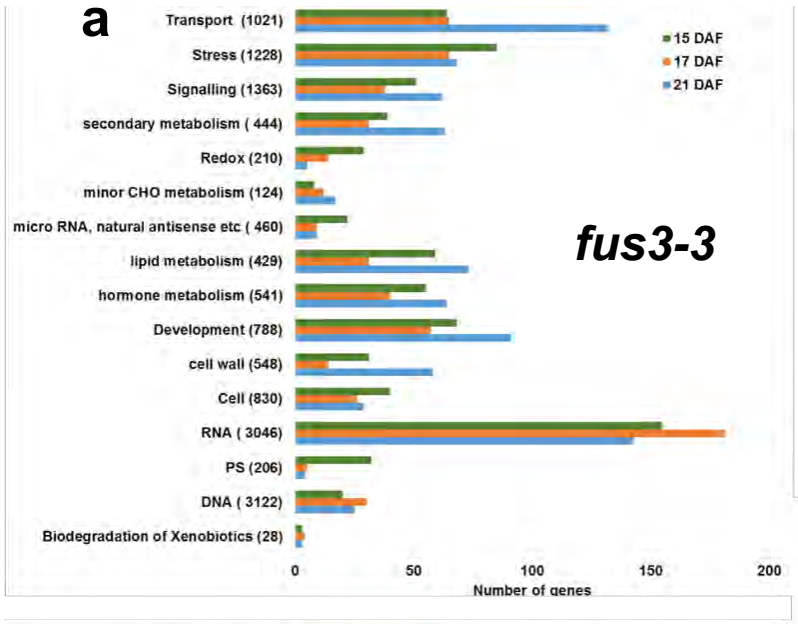
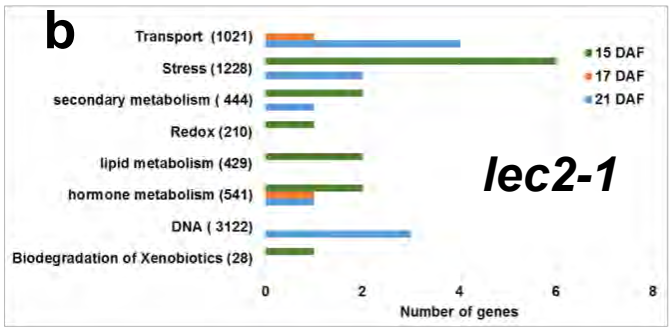
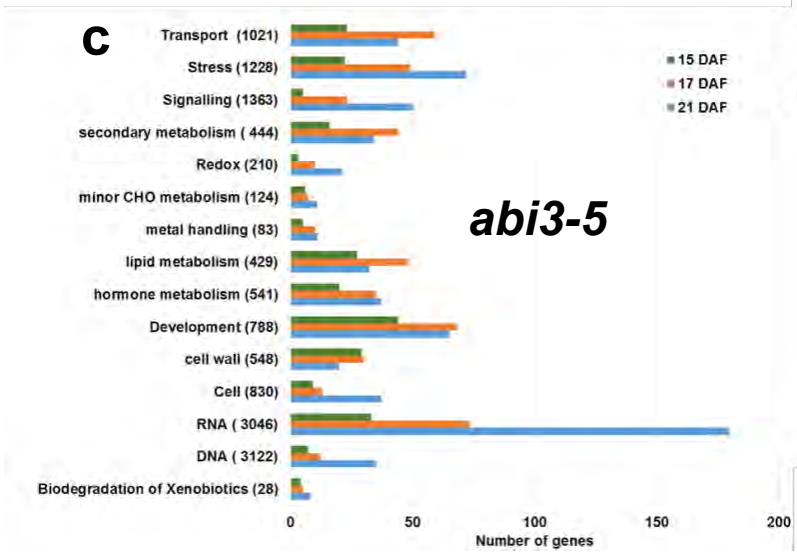


Figure S6. MAPMAN-based functional classification of upregulated DEGs in each mutant line during seed desiccation tolerance. (a) *fus3-3* vs Col-0 (b) *lec2-1* vs Ws (c) *abi3-5* vs Ler and (d) *lec1-1* vs Ws contrasts. MAPMAN classification of the transcripts using the web-tool Classification Superviewer (http://bar.utoronto.ca). All categories are significant (P value <0.05). CHO, carbohydrate; OPP, oxidative pentose phosphate; PS, photosynthesis; TCA, tricarboxylic.







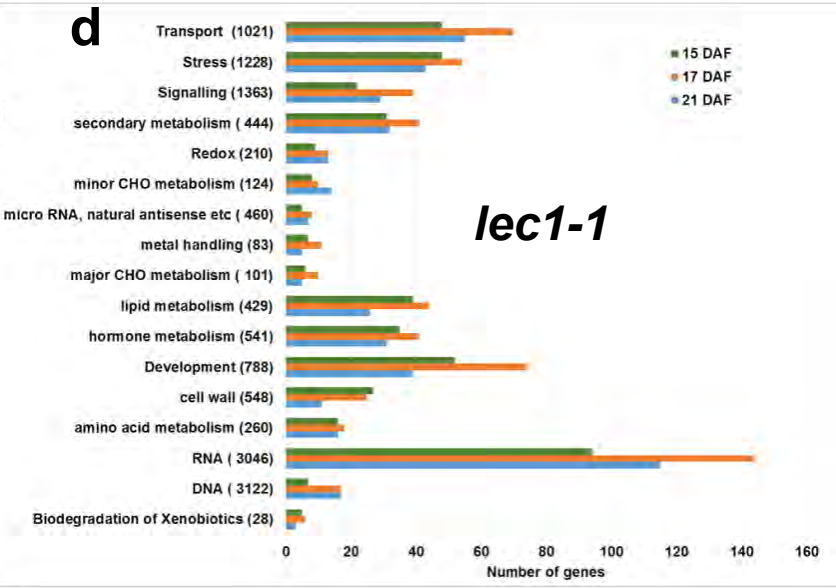


Figure S7. MAPMAN-based functional classification of downregulated DEGs in each mutant line during seed desiccation tolerance. (a) *fus3-3* vs Col-0 (b) *lec2-1* vs Ws (c) *abi3-5* vs Ler and (d) *lec1-1* vs Ws contrasts. MAPMAN classification of the transcripts using the web-tool Classification Superviewer (http://bar.utoronto.ca). All categories are significant (P-value <0.05). CHO, carbohydrate.

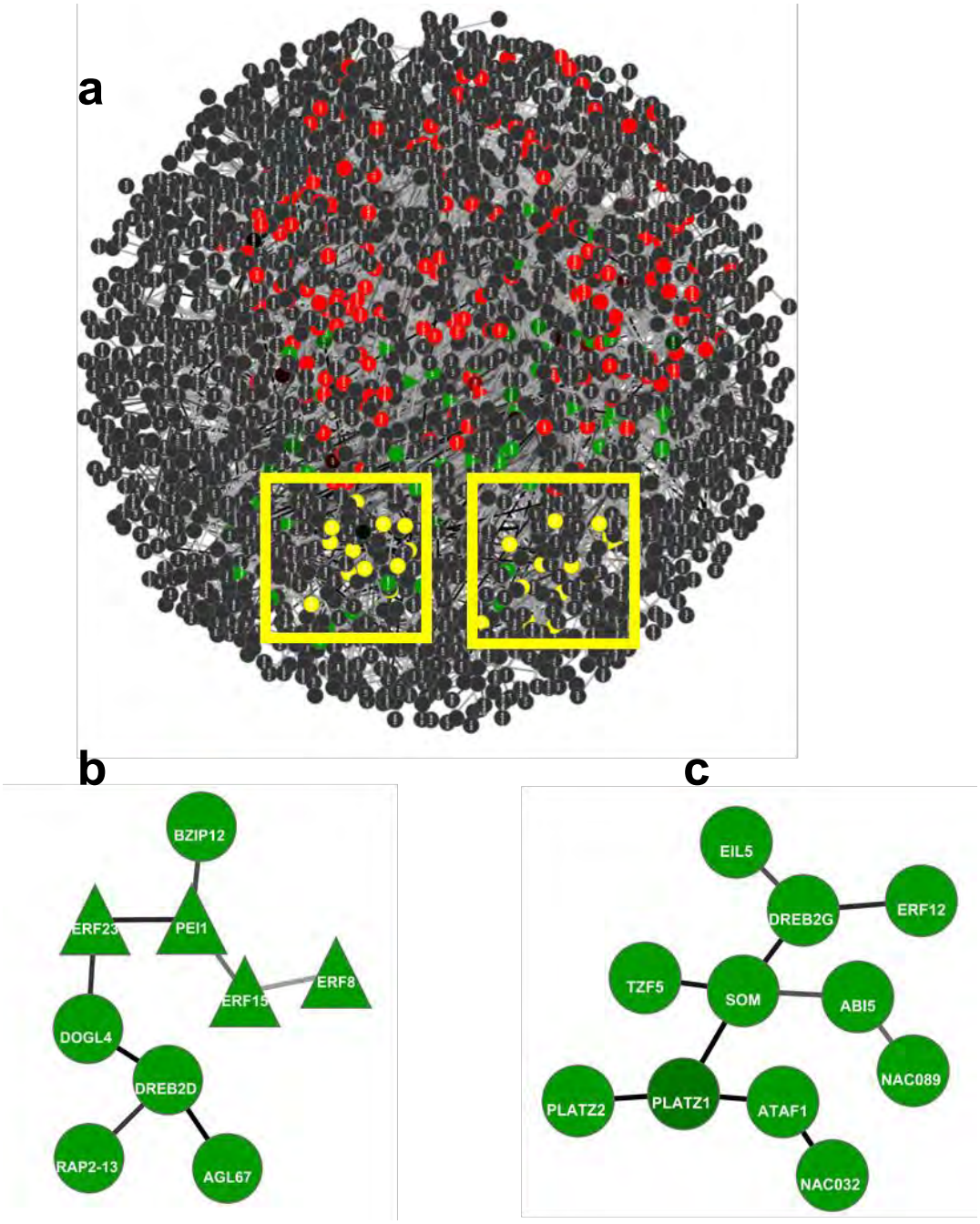


Figure S8. The whole gene regulatory network TFsSeedNet. (a) Overview of the TFsSeedNet obtained at DPI0.0. (b) TFsSeed-sNetDT1 and (c) TFsSeed-sNetDT2 subnetworks of the most representative downregulated in desiccation intolerant lines. Genes are represented as nodes and inferred interactions as edges (lines). Nodes are colored grey, except upregulated TFs in red and downregulated TFs in green from DT analysis. Edge width and color intensity is proportional to the Mutual Information (MI) value of the interaction, with higher MI values corresponding to thicker and darker edges. Triangular nodes represent genes expressed exclusively at 15DAF and circular nodes represent genes expressed at more than one time point.

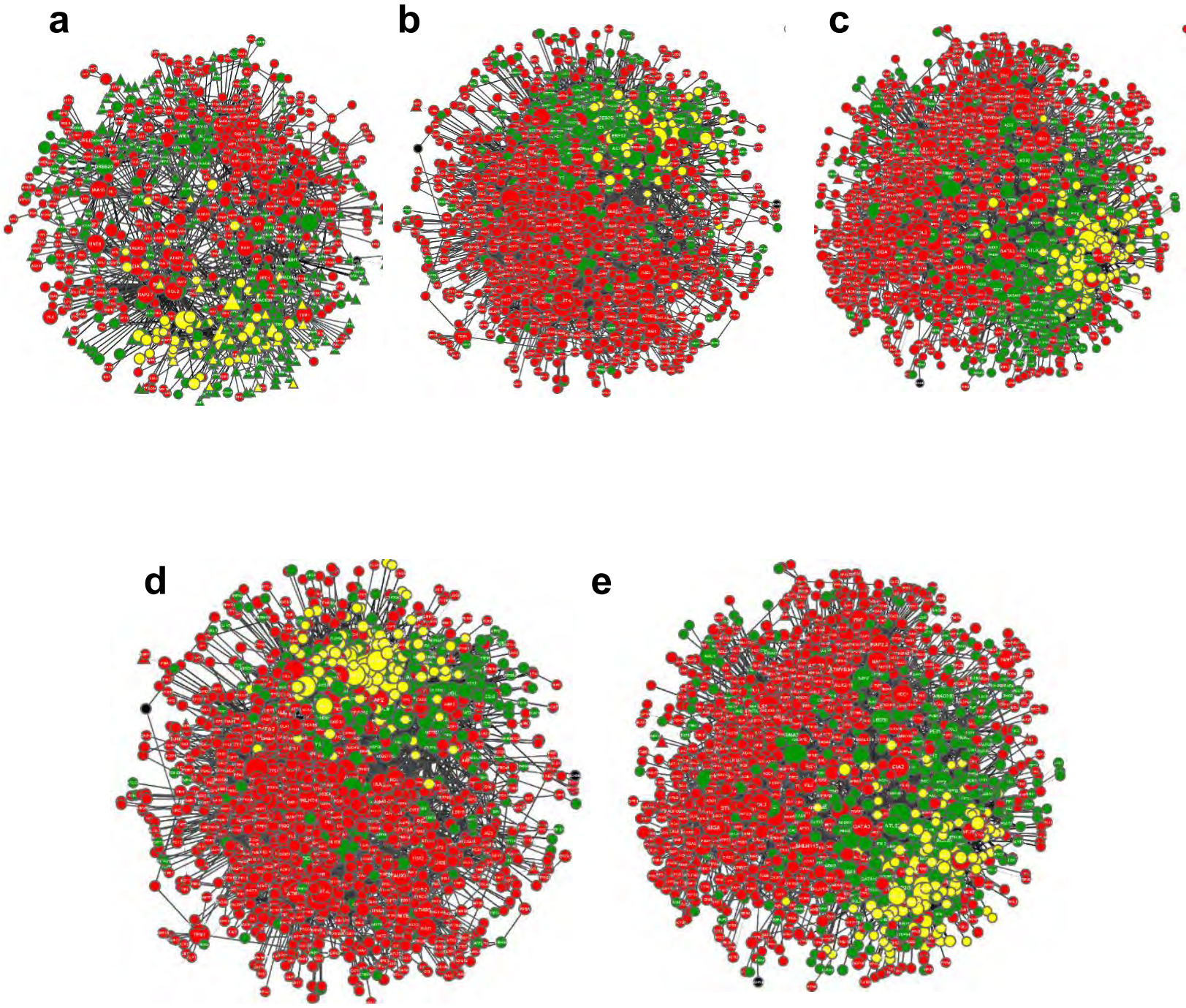


Figure S9. **The whole temporal gene regulatory network FullSeedNet**. (a), (b) and (c) represent the FullSeed-sNetDT1 related to nutrient storage. (d) and (e) represent the FullSeed-sNetDT2 related to cellular protection mechanism. Overview of the FullSeedNet obtained at DPI 0.1 at (a) 15, (b) and (d) 17 and (c) and (e) 21 DAF. Genes are represented as nodes and inferred interactions as edges. Nodes are colored grey, except upregulated genes in red and downregulated genes in green from DT analysis. Edges width and color intensity is proportional to the Mutual Information (MI) value of the interaction, with higher MI values corresponding to thicker and darker edges. Yellow nodes in represent the FullSeed-sNetDT1 and FullSeed-sNetDT2 at 15, 17 and 21 DAF..

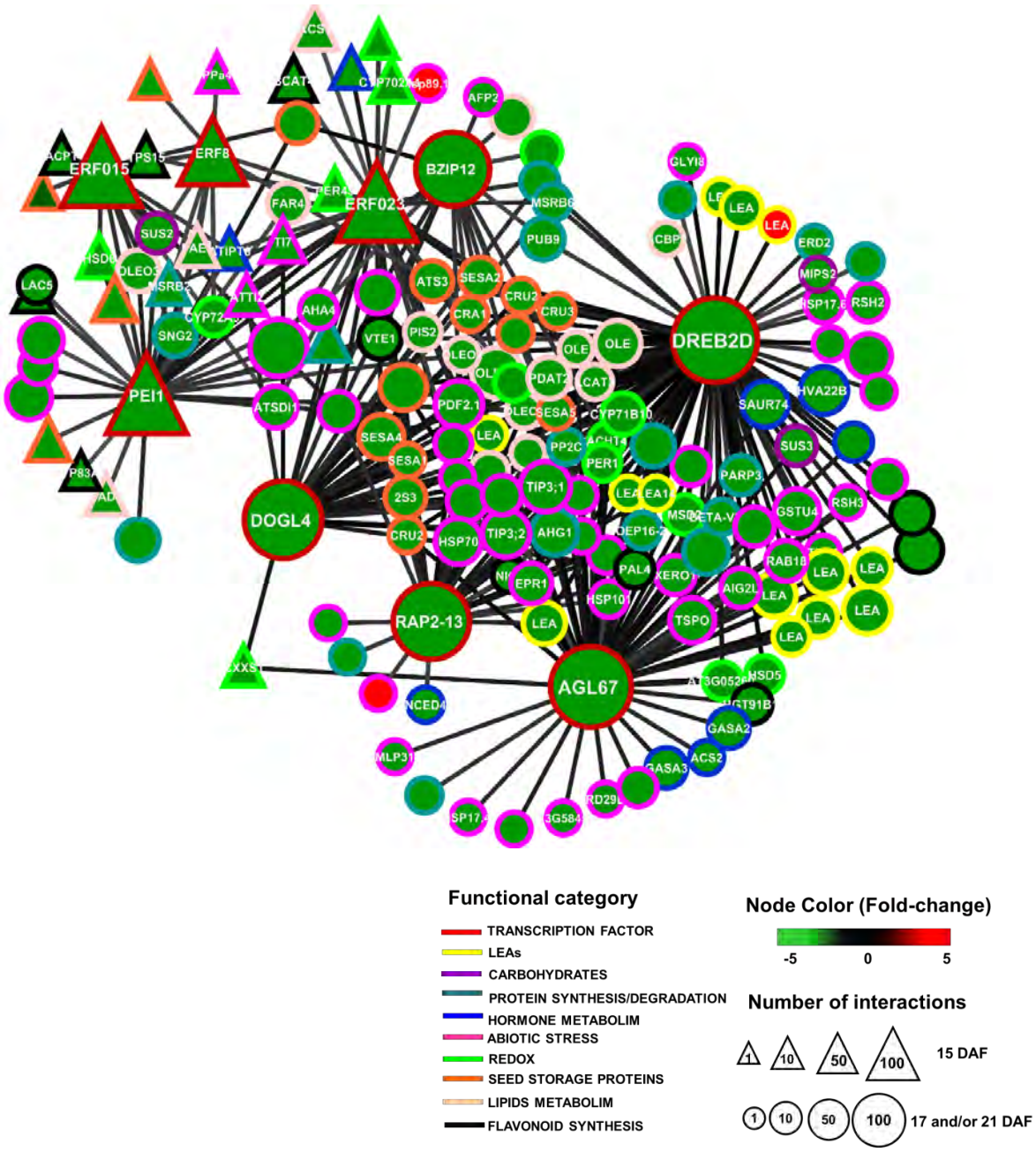


Figure S10. **FullSeed-sNetDT1 related to stress and nutrient storage**. Subnetworks were obtained from FullSeedNet at DPI 0.1. Genes are represented as nodes and inferred interactions as edges. Information box shows attributes in the networks. Triangular nodes represent genes expressed exclusively at 15DAF and circular nodes represent genes expressed at more than one time point. The border node color represent the functional category see Supplementary Table 18 and 19. Edge width and color intensity is proportional to the Mutual Information (MI) value of the interaction, with higher MI values corresponding to thicker and darker edges.

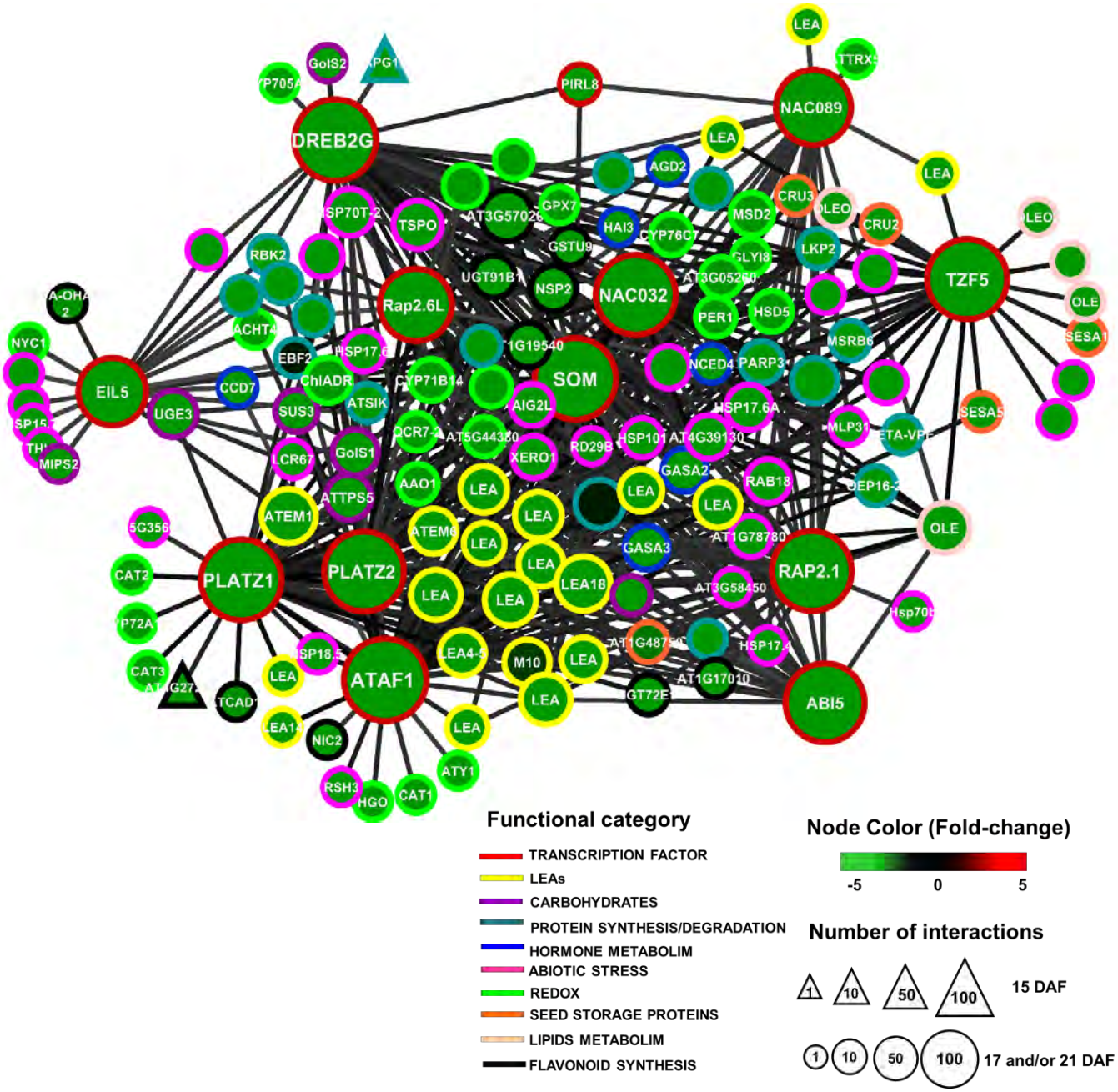


Figure S11. FullSeed-sNetDT1 related to cellular protection mechanisms. Subnetworks were obtained from FullSeedNet at DPI 0.1. Genes are represented as nodes and inferred interactions as edges. Information box shows attributes in the networks. Triangular nodes represent genes expressed exclusively at 15DAF and circular nodes represent genes expressed at more than one time point. The border node color represents the functional category (see Supplementary Table 20 and 21). Edge width and color intensity is proportional to the Mutual Information (MI) value of the interaction, with higher MI values corresponding to thicker and darker edges.

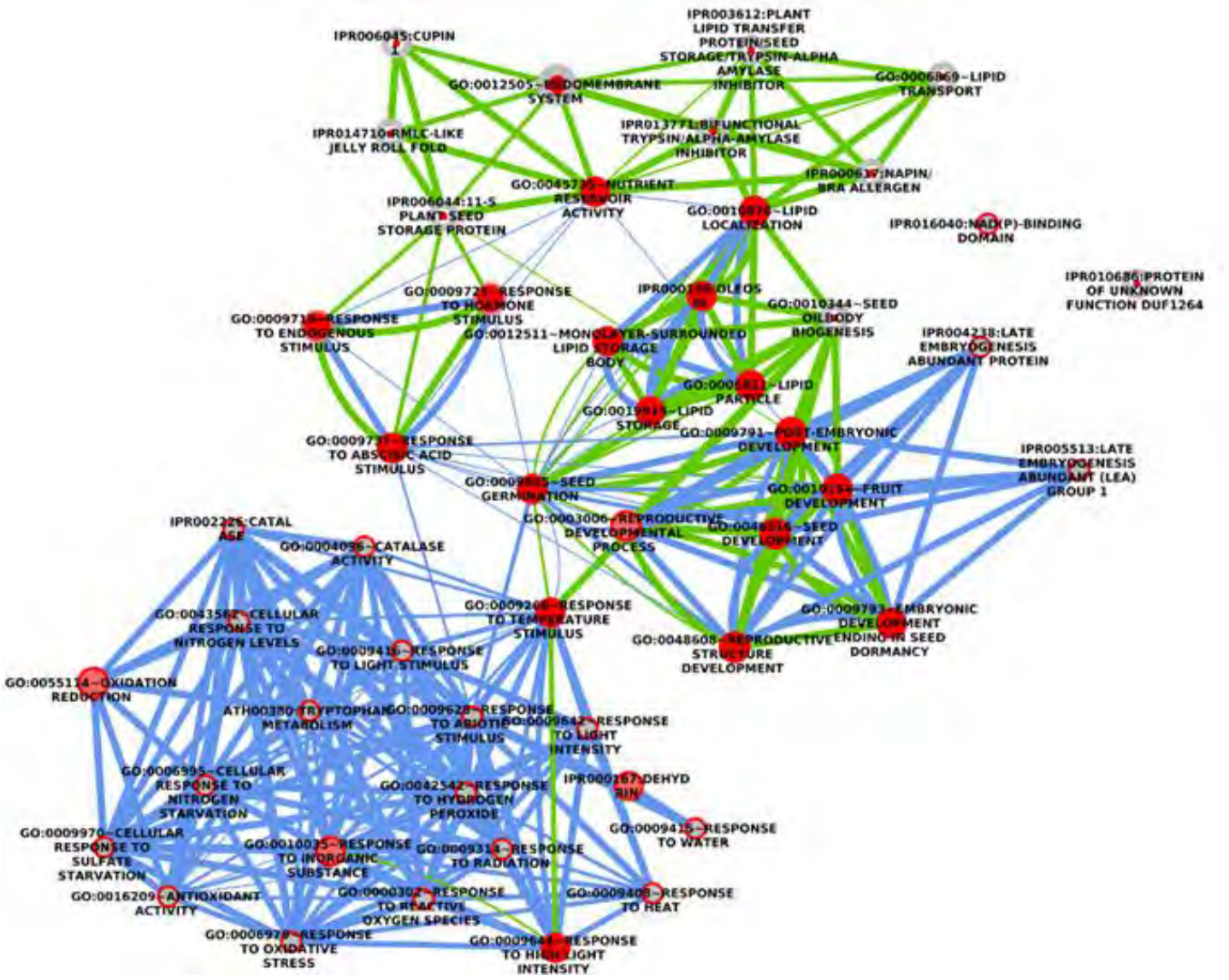


Figure S12. Comparative analysis of the functional enriched category between FullSeed-sNetDT1 and FullSeed-sNetDT2 . . Enrichment Map graph of category enrichments calculated with David for genes in snetFullDT1 and snetFullDT2. Node (inner circle) size corresponds to the number of genes in snetFullDT1 within the geneset. Node border (outer circle) size corresponds to the number of genes in snetFullDT2 within the geneset Nodes in red belong to genes categories that are only enriched in snetFullDT1**.** Nodes border red represent gene categories that are enriched snetFullDT2. Edge size corresponds to the number of genes that overlap between the two connected genesets. Green edges corresponds to snetFullDT1 and blue corresponds to snetFullDT2.

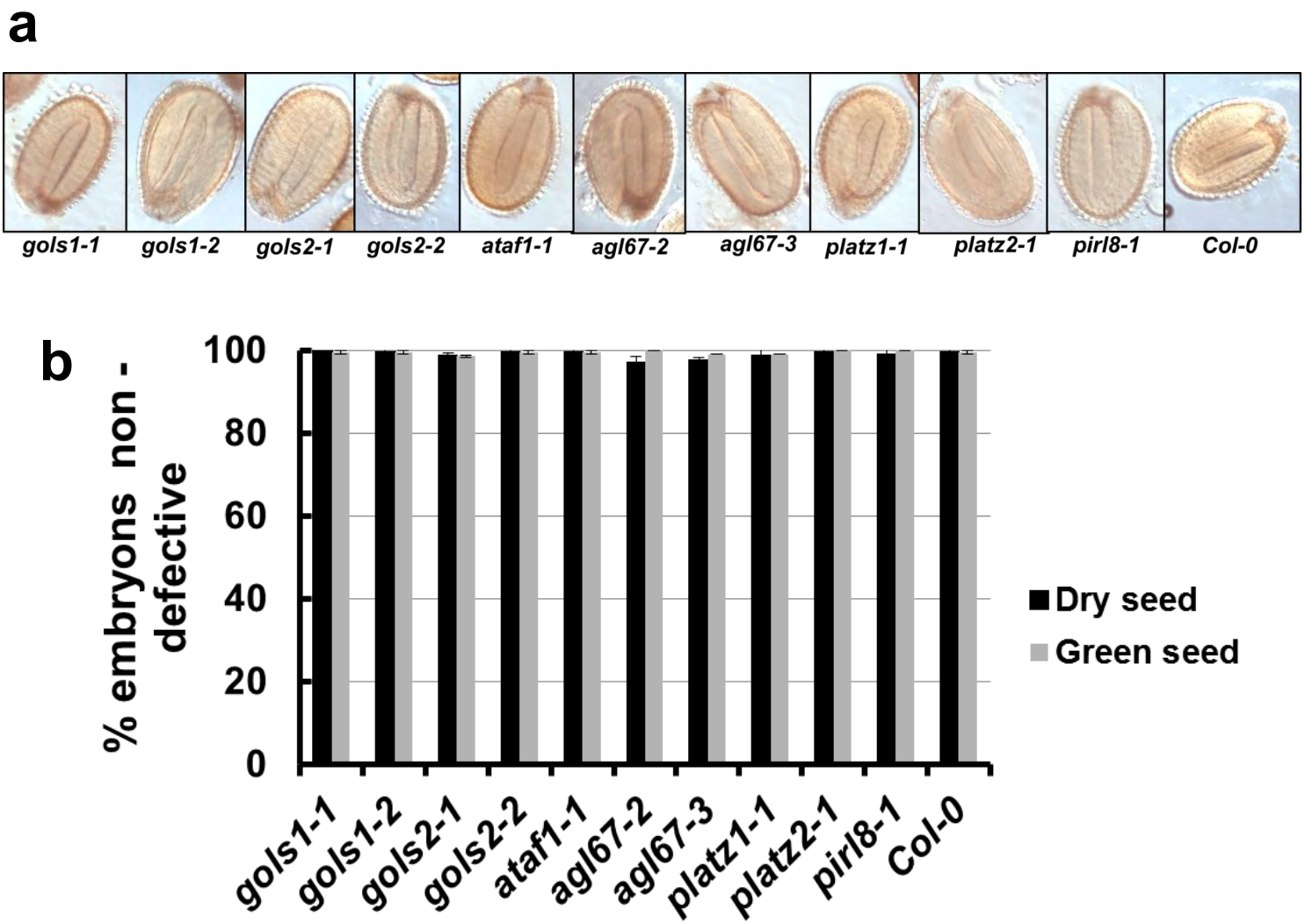


Figure S13. Embryos defective in T-DNA mutants seeds. (a) The visualization and (b) counting defective seeds were performed clearing dry and mature green seeds with Hoyer's solution overnight and visualized with microscope equipped with Nomarski optics. Values are means ± SD of two biological replicates for 100 seeds.

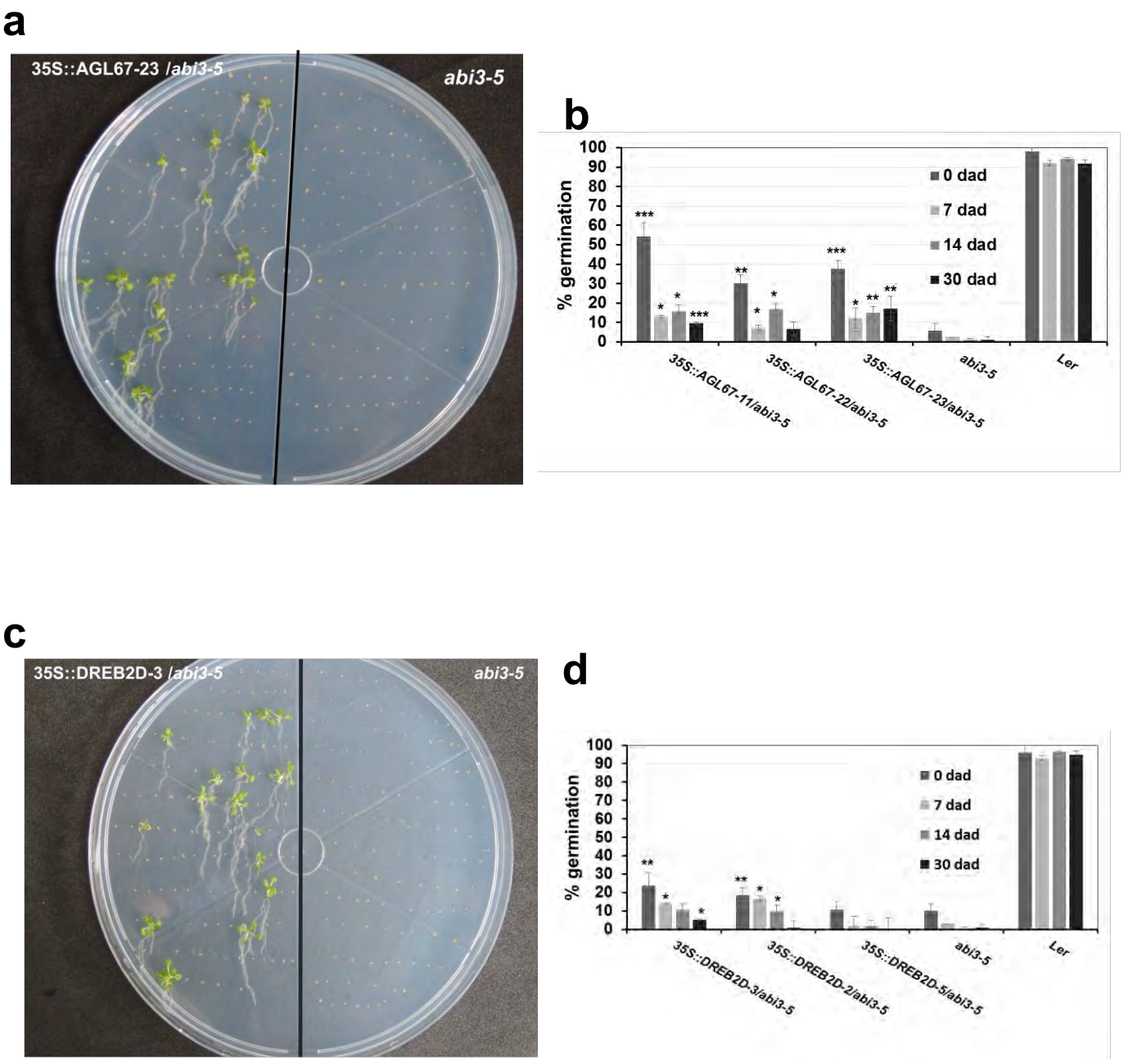


Figure S14. Overexpression of AGL67 and DREB2A partially rescues *abi3-5* intolerance desiccation phenotype. (a) Germination efficiency of 35S::AGL67-23/*abi3-5*; (b) Germination percentage of Ler, *abi3-5* and 35S::AGL67/*abi3-5* lines; (c) Germination efficiency of 35S::DREB2D-3/*abi3-5* (b) Germination percentage of Ler, *abi3-5* and 35S::DREB2D-3/*abi3-5* lines. For all experiments seeds at 21 DAF were stored for 0, 7, 14 and 30 days after desiccation (dad) to assess DT. Values are means ± SD of three biological replicates for 100 seeds. Statistical analysis of the data was done using a student t-test. Bars with asterisks are significantly different from the control (* = P<0.05, ** = P < 0.01, *** P<0.0001).

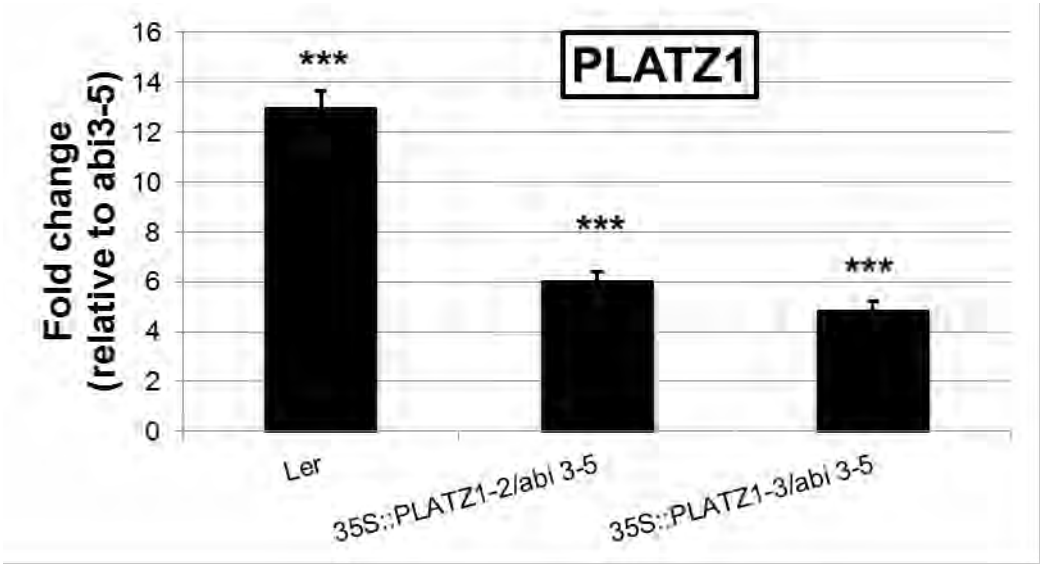


Fig. S15. Real Time PCRs of PLATZ1 in 35S::PLATZ1/abi3-5 lines seeds at 21DAF. *abi3-5* was used as the reference sample. UBQ10 and TIP4L served as internal controls. Data are the means ± SD of two biological replicates and three technical replicates. Values are means and ±SD of two biological replicates .Bars with asterisks are significantly different from the control (Student's t-test, * = P<0.05, ** = P < 0.01, *** P<0.0001).

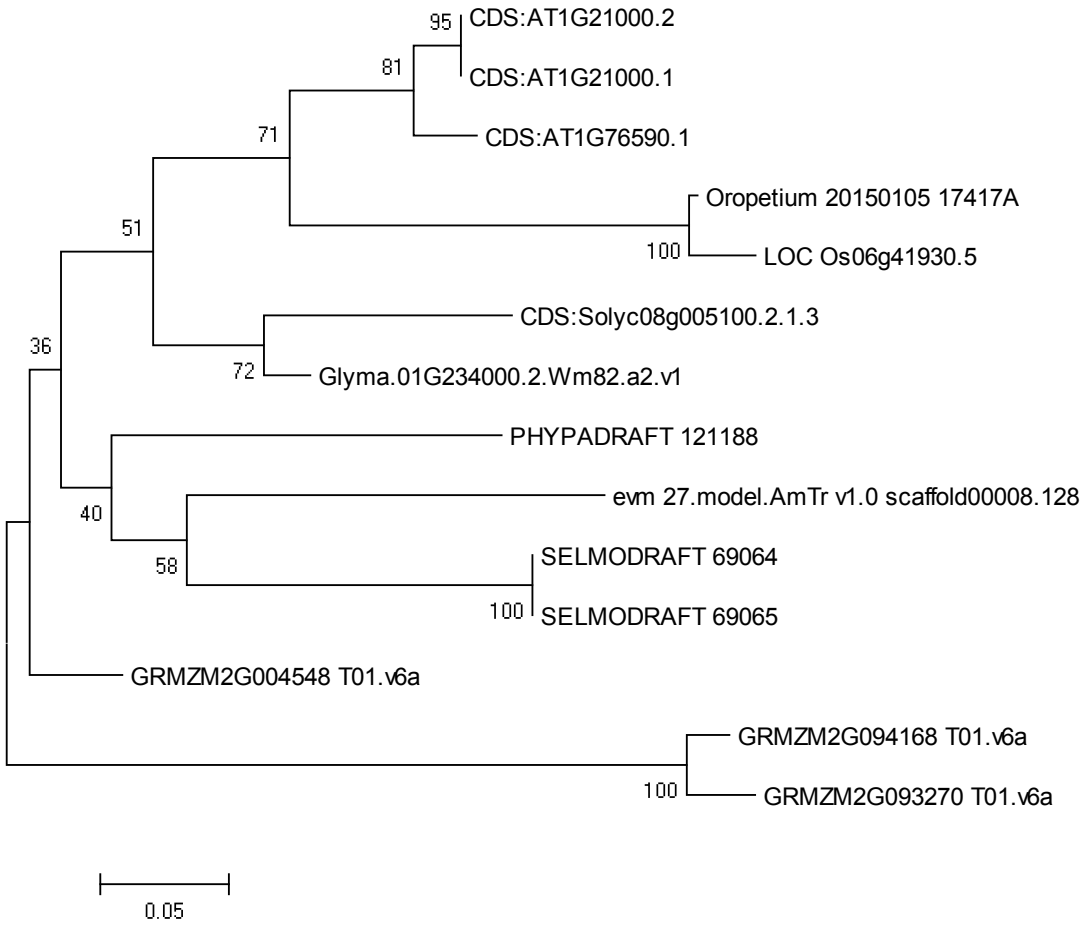


Figure S16. **Phylogenetic analysis of PLATZ1 and PLATZ2**.The NJ tree was constructed from the amino acid sequences of the PLATZ1 and PLATZ2 domain using the MEGA 6 program based on the JTT model. A consensus tree (after 1000 bootstrap samplings) is shown, and support values are indicated on the sides of important nodes. The BLAST was performed with CoGe Blast tool (https://genomevolution.org/coge/CoGeBlast.pl). PHYPADRAFT 121188 (*Physcomitella pattens* genome v.1 from NCBI database), SELMODRAFT 69064 and SELMODRAFT 69065 (*Selaginella mollendorffi* genome v.1 from NCBI database), evm 27.model.AmTr v1.0 scaffold00008.128 (*Amborella trichopoda* genome v.2 from Amborella Genome Consortium), LOC Os06g41930.5 (*Oryza sativa japonica* from MSU: Rice Genome Annotation v.7), Oropetium 20150105 17417A (*Oropetium thomaeum* genome v.1 from PacBio), Solyc08g005100.2.1.3 (*Solanum Lycopersicum* genome v.2 from SGN), Glyma.01G234000.2.Wm82.a2.v1 (*Glycine Max* genome v.1 from JGI), AT1G21000 and AT1G76590 (Arabidopsis thaliana genome TAIR10 from Ensembl Plants) GRMZM2G004548 T01.v6a, GRMZM2G094168 T01.v6a and GRMZM2G093270 T01.v6a (*Zea mays* genome v.6 from Ensembl Plants).

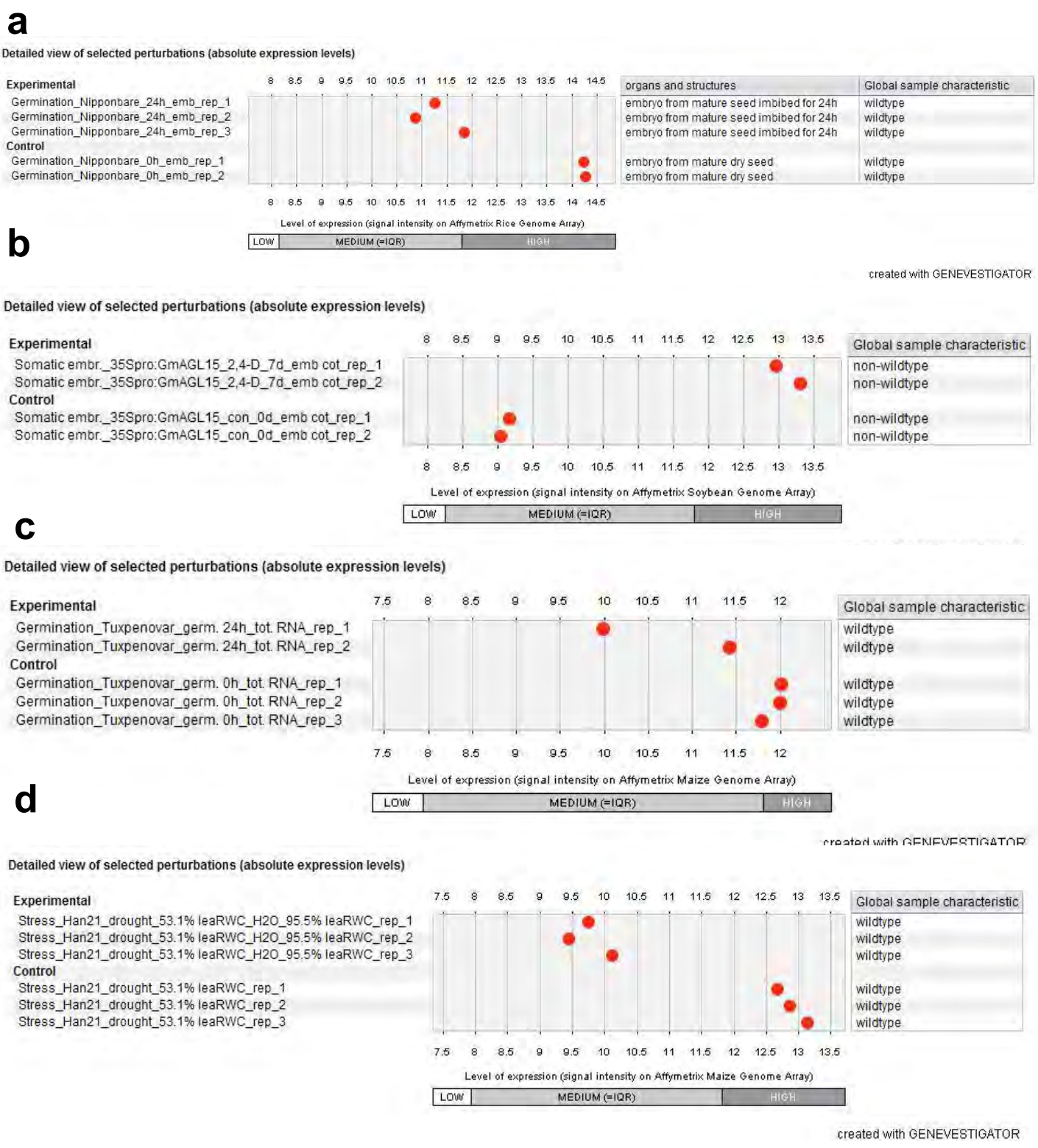


Figure S17. Expression pattern of possible orthologues of PLATZ1. Expression analysis was determined using the Genevestigator database**.** (a) LOC Os06g41930.5 (*Oryza sativa japonica*) (b) Glyma.01G234000.2.Wm82.a2.v1 (*Glycine Max*), (c) and (d) GRMZM2G004548 T01.v6a (*Zea mays).*

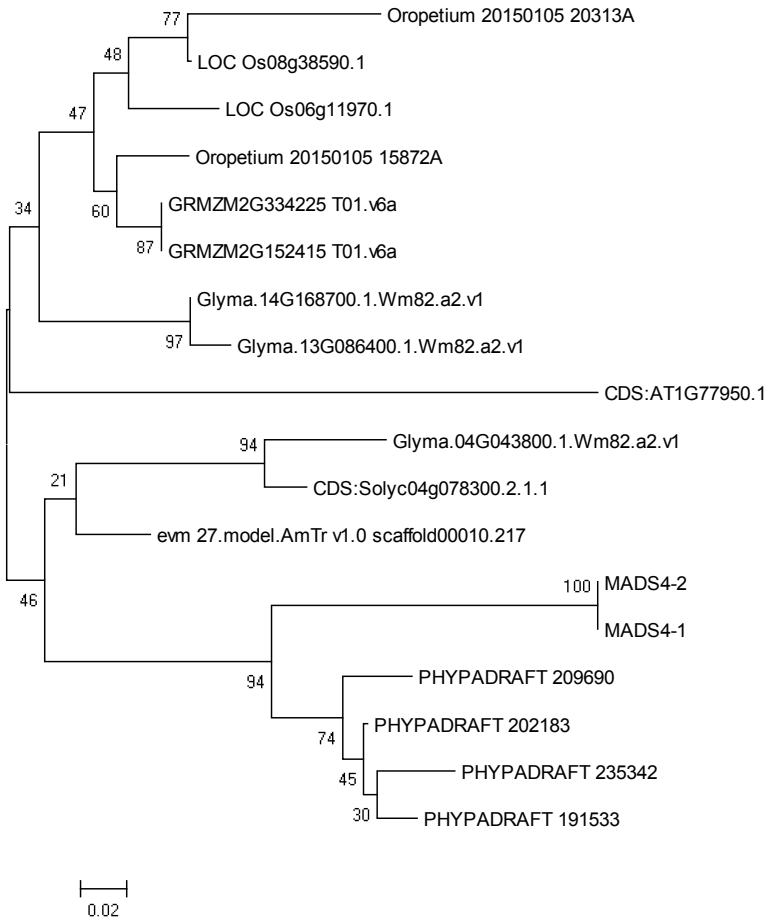


Figure S18. **Phylogenetic of AGL67.** The NJ tree was constructed from the amino acid sequences of the AGL67 domain using the MEGA6 program based on the JTT model. A consensus tree (after 1000 bootstrap samplings) is shown, and support values are indicated on the sides of important nodes. The BLAST was performed with CoGe Blast tool (https://genomevolution.org/coge/CoGeBlast.pl). PHYPADRAFT 202183, PHYPADRAFT 235342, PHYPADRAFT 191533 and PHYPADRAFT 209690 (*Physcomitella pattens* genome v.1 from NCBI database), MADS4-2 and MADS4-1 (*Selaginella mollendorffi* genome v.1 from NCBI database), evm 27.model.AmTr v1.0 scaffold00010.217 (*Amborella trichopoda* genome v.2 from Amborella Genome Consortium), LOC Os08g38590.1 and LOC Os06g11970.1 (*Oryza sativa japonica* from MSU: Rice Genome Annotation v.7), Oropetium 20150105 20313A and Oropetium 20150105 15872Ac (*Oropetium thomaeum* genome v.1 from PacBio), Solyc04g078300.2.1.1 (*Solanum Lycopersicum* genome v.2 from SGN), Glyma.04G043800.1.Wm82.a2.v1, Glyma.14G168700.1.Wm82.a2.v1 and Glyma.13G086400.1.Wm82.a2.v1 (*Glycine Max* genome v.1 from JGI), AT1G77950 (*Arabidopsis thaliana* genome TAIR10 from Ensembl Plants) GRMZM2G334225 T01.v6a and GRMZM2G152415 T01.v6a (*Zea mays* genome v.6 from Ensembl Plants).

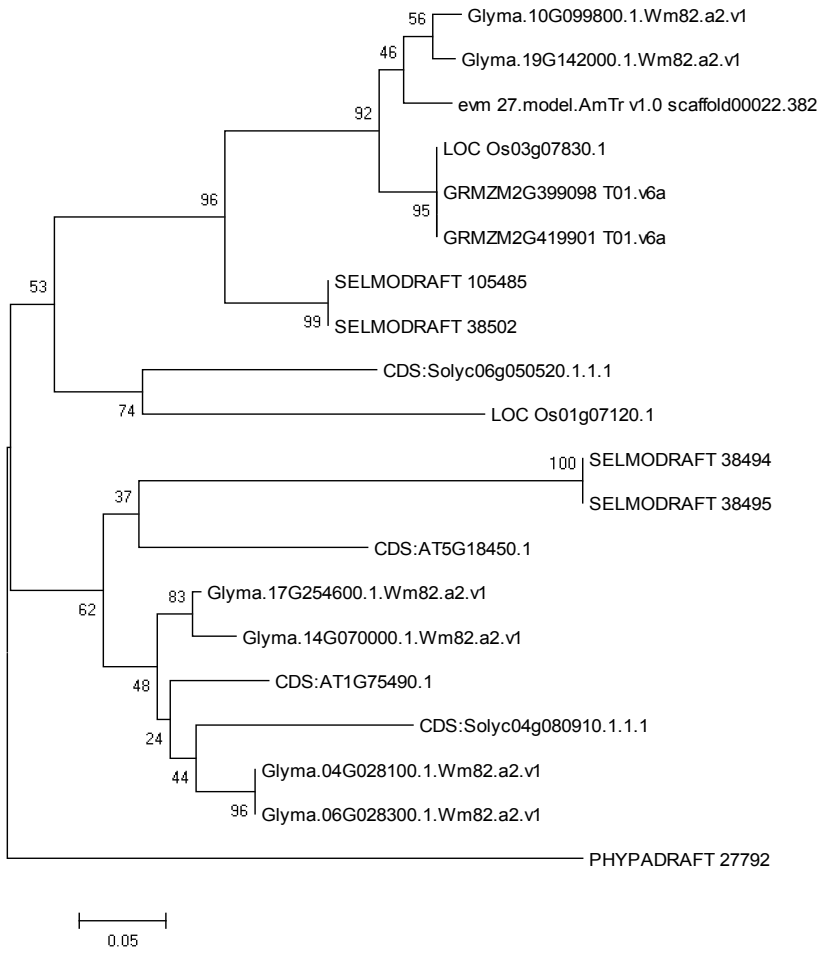


Figure S19. **Phylogeny of DREB2D and DREB2G**. The NJ tree was constructed from the amino acid sequences of the DREB2D and DREB2G domain using the MEGA6 program based on the JTT model. A consensus tree (after 1000 bootstrap samplings) is shown, and support values are indicated on the sides of important nodes. The BLAST was performed with CoGe Blast tool (https://genomevolution.org/coge/CoGeBlast.pl), PHYPADRAFT 27792 (*Physcomitella pattens* genome v.1 from NCBI database), SELMODRAFT 69064 and SELMODRAFT 105485, SELMODRAFT 38502, SELMODRAFT 38494 and SELMODRAFT 38495 (*Selaginella mollendorffi* genome v.1 from NCBI database), evm 27.model.AmTr v1.0 scaffold00022.382 (*Amborella trichopoda* genome v.2 from Amborella Genome Consortium), LOC Os03g07830.1 (*Oryza sativa japonica* from MSU: Rice Genome Annotation v.7) Solyc06g050520.1.1.1 and Solyc04g080910.1.1.1 (*Solanum Lycopersicum* genome v.2 from SGN), Glyma.10G099800.1.Wm82.a2.v1, Glyma.19G142000.1.Wm82.a2.v1, Glyma.17G254600.1.Wm82.a2.v1, Glyma.14G070000.1.Wm82.a2.v1, Glyma.04G028100.1.Wm82.a2.v1 and Glyma.06G028300.1.Wm82.a2.v1 (*Glycine Max genome* v.1 from JGI), AT5G18450.1 and AT1G75490.1 (*Arabidopsis thaliana* genome TAIR10 from Ensembl Plants) GRMZM2G399098 T01.v6a and GRMZM2G419901 T01.v6a (*Zea mays* genome v.6 from Ensembl Plants).

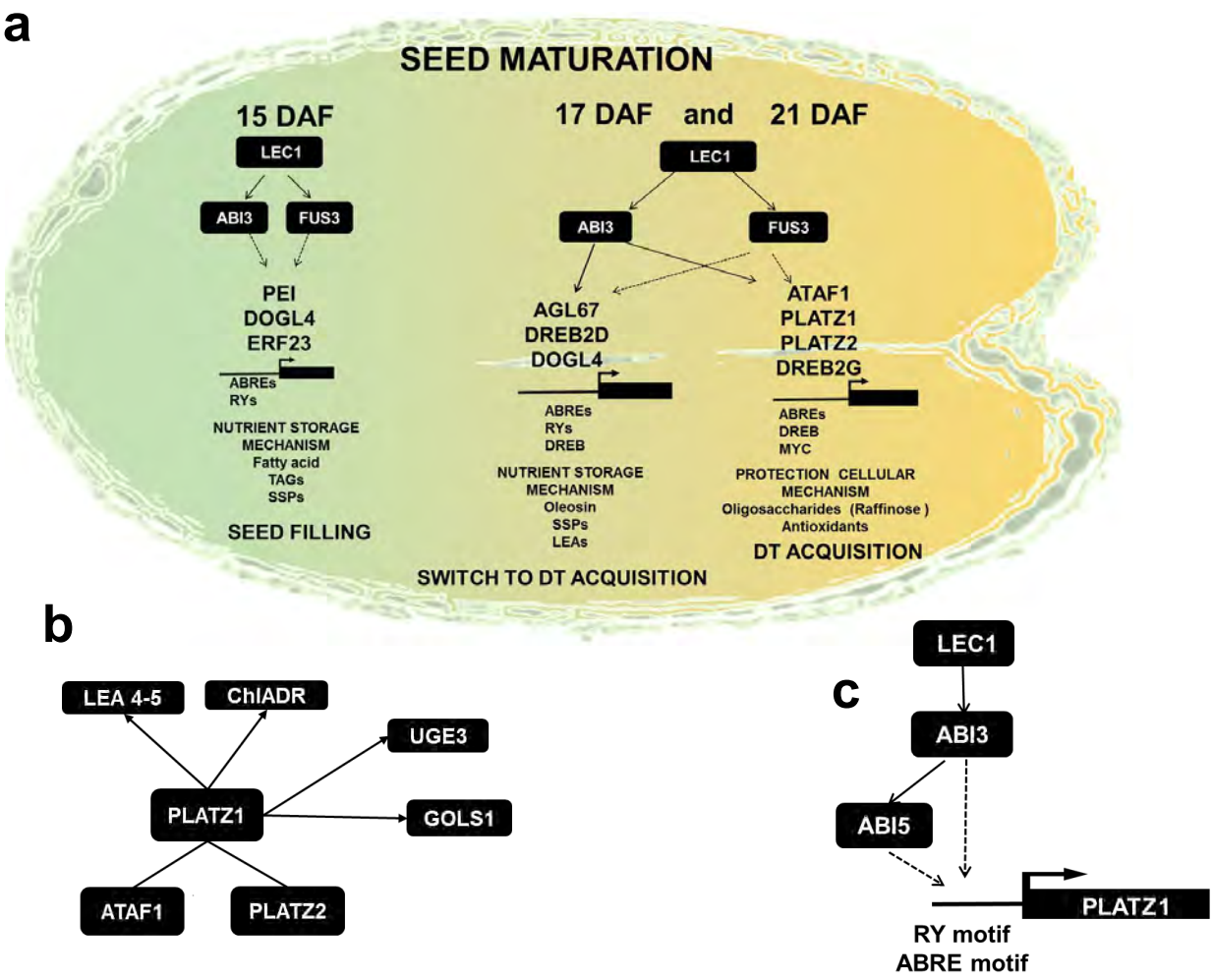


Figure S20. Hypothetical model of desiccation tolerance seed regulation during seed maturation. (a) Downstream of the LEC1, ABI3 and FUS3 master regulators, the expression of a number of TFs is activated at 15, 17 and 21 DAF. At 15 DAF the expression of genes related to storage of reserve compounds (seed filling) is activated. The cognate sites for these TFs are present in the promoter of maturation-specific genes. At 17 and 21DAF the expression of genes related to storage of reserve compounds is maintained, and a the expression of a different set of TFs is activated which in turn regulate the expression of effector genes involved in cell protection mechanism responsible for the acquisition of desiccation tolerance. The cognate binding sites for these second set of TFs is present in dehydration responsive genes. (b) PLATZ1 together with PLATZ2 and ATAF1 activate genes related to stress response and cell protection mechanisms, including gene belonging to LEA protein families, genes involved in the production of antioxidants and protective oligosaccharides. (c) PLATZ1 is a novel regulator of DT, which is transcriptionally activated by ABI3 and/or ABI5 in a network conserved during the evolution of land plants.

References to supplementary text

1. Meinke DW, Franzmann LH, Nickle TC, & Yeung EC (1994) Leafy Cotyledon Mutants of Arabidopsis. *The Plant cell* 6(8):1049-1064.
2. Ooms J, Leon-Kloosterziel KM, Bartels D, Koornneef M, & Karssen CM (1993) Acquisition of Desiccation Tolerance and Longevity in Seeds of Arabidopsis thaliana (A Comparative Study Using Abscisic Acid-Insensitive abi3 Mutants). *Plant physiology* 102(4):1185-1191.
3. Keith K, Kraml M, Dengler NG, & McCourt P (1994) fusca3: A Heterochronic Mutation Affecting Late Embryo Development in Arabidopsis. *The Plant cell* 6(5):589-600.
4. McCarthy DJ, Chen Y, & Smyth GK (2012) Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic acids research* 40(10):4288-4297.
5. Maere S, Heymans K, & Kuiper M (2005) BiNGO: a Cytoscape plugin to assess overrepresentation of Gene Ontology categories in Biological Networks. *Bioinformatics* 21(16):3448-3449.
6. Benjamini Y & Hochberg Y (1995) Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)* 57(1):289-300.
7. Chavez Montes RA*, et al.* (2014) ARACNe-based inference, using curated microarray data, of Arabidopsis thaliana root transcriptional regulatory networks. *BMC plant biology* 14:97.
8. Shimada TL, Shimada T, & Hara-Nishimura I (2010) A rapid and non-destructive screenable marker, FAST, for identifying transformed seeds of Arabidopsis thaliana. *The Plant Journal* 61(3):519-528.
9. Martinez-Trujillo M, Limones-Briones V, Cabrera-Ponce J, & Herrera-Estrella L (2004) Improving transformation efficiency ofArabidopsis thaliana by modifying the floral dip method. *Plant Mol Biol Rep* 22(1):63-70.
10. Ye J*, et al.* (2012) Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. *BMC bioinformatics* 13:134.
11. Gazzarrini S, Tsuchiya Y, Lumba S, Okamoto M, & McCourt P (2004) The transcription factor FUSCA3 controls developmental timing in Arabidopsis through the hormones gibberellin and abscisic acid. *Developmental cell* 7(3):373-385.
12. Lumba S*, et al.* (2012) The embryonic leaf identity gene FUSCA3 regulates vegetative phase transitions by negatively modulating ethylene-regulated gene expression in Arabidopsis. *BMC biology* 10:8.
13. Basso K*, et al.* (2005) Reverse engineering of regulatory networks in human B cells. *Nature genetics* 37(4):382-390.
14. Margolin AA*, et al.* (2006) ARACNE: an algorithm for the reconstruction of gene regulatory networks in a mammalian cellular context. *BMC bioinformatics* 7 Suppl 1:S7.