

Analysis of the Alfalfa Root Transcriptome in Response to Salinity Stress

Olga A. Postnikova, Jonathan Shao and Lev G. Nemchinov*

USDA/ARS, Plant Sciences Institute, Molecular Plant Pathology Laboratory, 10300 Baltimore Avenue, Beltsville, MD 20705, USA

*Corresponding author: E-mail, Lev.Nemchinov@ars.usda.gov

(Received November 6, 2012; Accepted April 4, 2013)

Salinity is one of the major abiotic factors affecting alfalfa productivity. Identifying genes that control this complex trait will provide critical insights for alfalfa breeding programs. To date, no studies have been published on a deep sequencing-based profiling of the alfalfa transcriptome in response to salinity stress. Observations gathered through research on reference genomes may not always be applicable to alfalfa. In this work, Illumina RNA-sequencing was performed in two alfalfa genotypes contrasting in salt tolerance, in order to estimate a broad spectrum of genes affected by salt stress. A total of 367,619,586 short reads were generated from cDNA libraries originated from roots of both lines. More than 60,000 tentative consensus sequences (TCs) were obtained and, among them, 74.5% had a significant similarity to proteins in the NCBI database. Mining of simple sequence repeats (SSRs) from all TCs revealed 6,496 SSRs belonging to 3,183 annotated unigenes. Bioinformatics analysis showed that the expression of 1,165 genes, including 86 transcription factors (TFs), was significantly altered under salt stress. About 40% of differentially expressed genes were assigned to known gene ontology (GO) categories using Arabidopsis GO. A random check of differentially expressed genes by quantitative real-time PCR confirmed the bioinformatic analysis of the RNA-seq data. A number of salt-responsive genes in both tested genotypes were identified and assigned to functional classes, and gene candidates with roles in the adaptation to salinity were proposed. Alfalfa-specific data on salt-responsive genes obtained in this work will be useful in understanding the molecular mechanisms of salinity tolerance in alfalfa.

Keywords: *Medicago sativa* • Next-generation sequencing • Root transcriptome • Salt stress.

Abbreviations: bHLH, basic helix–loop–helix; CDS, coding sequences; DEG, differentially expressed gene; DFR, dihydroflavonol reductase; FDR, false discovery rate; GO, gene ontology; MS, Murashige and Skoog; NSG, next-generation sequencing; ORF, open reading frame; qRT-PCR, quantitative

real-time PCR; SSR, simple sequence repeat; TF, transcription factor; TCs, tentative consensus sequences.

Introduction

More than one-third of all irrigated lands contain high levels of salts. Salinization of non-irrigated agricultural soils is also a serious issue that leads to lower yields and reduces the ability of crops to take up water. Plant adaptation to salinity is thought to be of three different types (Munns and Tester 2008): tolerance to osmotic shock mediated by calcium signaling; exclusion of salt ions that inhibit metabolic processes inside the cell by transporter proteins and tolerance to already accumulated sodium ions by their internal distribution; and compartmentation away from the cytosol through the operation of a vacuolar Na⁺/H⁺ antiport (Apse et al. 1999). While the physiological reactions underlying these three tolerance mechanisms are well known, understanding of molecular responses to high salinity is often limited to individual components of the model system *Arabidopsis thaliana*, a salt-sensitive species (Ren et al. 2010).

Alfalfa (*Medicago sativa*) is the most extensively cultivated forage legume in the world and the fourth most widely grown crop in the USA, being planted on >23 million acres in all 50 states. Although alfalfa is considered a moderately salt-tolerant species as compared to other legumes (Munns and Tester 2008), salinity stress is among the most problematic environmental factors limiting alfalfa production. Thus, increased salt tolerance in alfalfa has great economic potential (Ottman 1999, Peel et al. 2004).

Salinity research in alfalfa had benefited from a variety of approaches, such as selection of germplasm resources for increased salt tolerance at different growth stages, identification and selection for salt tolerance via physiological traits and molecular markers, regeneration from salt-tolerant cell lines, in vitro selection and electrophysiological studies of traits associated with salt tolerance and genetic engineering through transgene expression (Rumbaugh and Pendery 1990, Winicov and Bastola 1999, Winicov 2000, Djilianov et al. 2003, Smethurst et al. 2008, Monirifar and Barghi 2009, Soltani et al. 2012).

A number of individual genes whose expression is induced by salt stress in alfalfa have recently been isolated and

DDBJ accession number: PRJDB749.

Plant Cell Physiol. 54(7): 1041–1055 (2013) doi:10.1093/pcp/pct056, available online at www.pcp.oxfordjournals.org
Published by Oxford University Press on behalf of Japanese Society of Plant Physiologists.

characterized, and several attempts have been made to obtain a profile of gene expression in alfalfa under salinity conditions (Yang *et al.* 2008, Jin *et al.* 2010, Soto *et al.* 2011, Chen *et al.* 2012, Hua *et al.* 2012). Meanwhile, a comprehensive transcriptomic analysis of alfalfa gene expression in response to salt stress, integrating fragmentary data into signaling pathways and regulatory networks, is still missing from the literature. Without such an analysis, mechanisms underlying phenotypic selection will remain elusive, and breeding for salt tolerance in alfalfa will continue to be challenging.

Genome-wide analyses of plant genes involved in response to salt stress has been conducted in alfalfa's close relative *Medicago truncatula* (Li *et al.* 2009, Zahaf *et al.* 2012), and its recently completed sequencing project has provided an excellent opportunity to use this reference genome for uncovering agronomically important genes in alfalfa (Young *et al.* 2011). However, it is well known that gene expression diversity exists in closely related species. Even though *M. truncatula* and *M. sativa* share a high degree of sequence similarity and a conserved genome structure (Yang *et al.* 2008), observations gathered through research on reference genomes may not always be applicable to alfalfa, a perennial and an outcrossing tetraploid that has a number of unique characteristics reflected in its genome structure and content.

In this work, we performed transcriptional profiling of NaCl-stressed alfalfa roots by next-generation sequencing (NGS) for the assessment of gene expression changes in two distinct alfalfa genotypes: AZ-88NDC, a salt-susceptible accession and AZ-GERM SALT-II, a germplasm salt tolerant at the germination stage (Dobrenz *et al.* 1989, Smith and Fairbanks 1989, Smith 1991). AZ-88NDC was released by the Arizona Agricultural Experiment Station in September 1988 and is the Syn 1 generation of a composite of equal amounts of certified seed from 13 elite non-dormant alfalfa cultivars: Pierce, UC Cibola, Armona, Mecca, Maricopa, WL 515, WL 605, Pioneer 5929, DK 187, CUF 101, Sapphire, Lew and NPI 8391 (Smith and Fairbanks 1989). AZ-GERM SALT-II alfalfa was released by the Arizona Agriculture Experiment Station in July 1988. This germplasm represents the ninth cycle of recurrent mass selection for alfalfa genotypes that germinate at high levels of NaCl. AZ-GERM SALT-II was derived from AZ-GERM SALT-1, which represented the fifth cycle of mass recurrent selection for germination salt tolerance of the parent 'Mesa-Sirsa' (Dobrenz *et al.* 1989).

NGS is currently a method of choice to unravel a diversity of stress responses on a transcriptome-wide scale in non-model plant species, where the complete genome sequence and annotation are not yet available. Using two different genotypes contrasting in their salt tolerance allowed coverage of a broad spectrum of genes influenced by salt stress, including those involved in a general stress response network, in susceptibility to NaCl and in salt adaptation. Using Arabidopsis gene ontology (GO), differentially expressed genes were annotated and assigned to known functional groups, biological processes and regulatory networks. Enrichment analysis revealed over-represented functional groups which in turn helped to predict the

importance of specific biological processes in response to salinity.

The generated reads and *de novo* assemblies were deposited with the DDBJ Sequence Read Archive (http://trace.ddbj.nig.ac.jp/dra/index_e.shtml) with accession number PRJDB749.

Results

Effect of salt on accumulation of ions

AZ-88NDC is a salt-susceptible accession and AZ-GERM SALT-II is a germplasm salt tolerant at the germination stage (Dobrenz *et al.* 1989, Smith and Fairbanks 1989, Smith 1991). Since root samples were collected at the seedling stage and also considering that response to salinity may be substantially different at various growth stages, susceptibility of these two genotypes to salt was additionally confirmed by measuring accumulation of salt ions in roots and shoots.

Measurements of ions concentrations showed accumulation of sodium and chloride ions in roots and shoots of both genotypes in response to salt treatment (Fig. 1A, C). Characteristically, the AZ-GERM SALT-II roots accrued less sodium as compared with AZ-88NDC, which is most probably related to the specific sodium exclusion mechanisms that are important for salt tolerance at the germination, emergence and early developmental stages. Shoots of the salt-tolerant germplasm AZ-GERM SALT-II were capable of accumulating more Na⁺ than the susceptible germplasm AZ-88NDC (Fig. 1A) without the reduction in shoot dry weight biomass observed in AZ-88NDC (data not shown). Accumulation of Cl⁻ was similarly enhanced in the roots of both lines, while in the shoots the increase was more evident in AZ-GERM SALT-II. The potassium concentration decreased in the roots and shoots of both lines (Fig. 1B), although the tolerant line had better K⁺ retention ability in both root and shoot tissues. Tolerant alfalfa varieties have been shown previously to have a higher potassium content under salt stress (Smethurst *et al.* 2008). However, the K⁺/Na⁺ ratio, which shows a positive relationship with salt tolerance (Babu *et al.* 2012), was increased only in the AZ-GERM SALT-II line (Fig. 1D). Overall, these data support known characteristics and/or response of both genotypes to salinity stress (Dobrenz *et al.* 1989, Smith and Fairbanks 1989, Smith 1991).

Total size and composition of the alfalfa root transcriptome as revealed by RNA-seq

AZ-88NDC. A total of 184,596,972 short reads (mock and experiment) were generated from cDNA libraries originated from the whole roots of AZ-88NDC. They were assembled into tentative consensus sequences (TCs) [The Gene Index Databases, Dana Farber Cancer Institute (URL: <http://www.danafarber.org/>)]. TCs are created by assembling expressed sequence tags (ESTs) into virtual transcripts and may contain partial cDNA sequences or, in many cases, represent full-length transcripts. TCs are actual assemblies, rather than clusters of overlapping

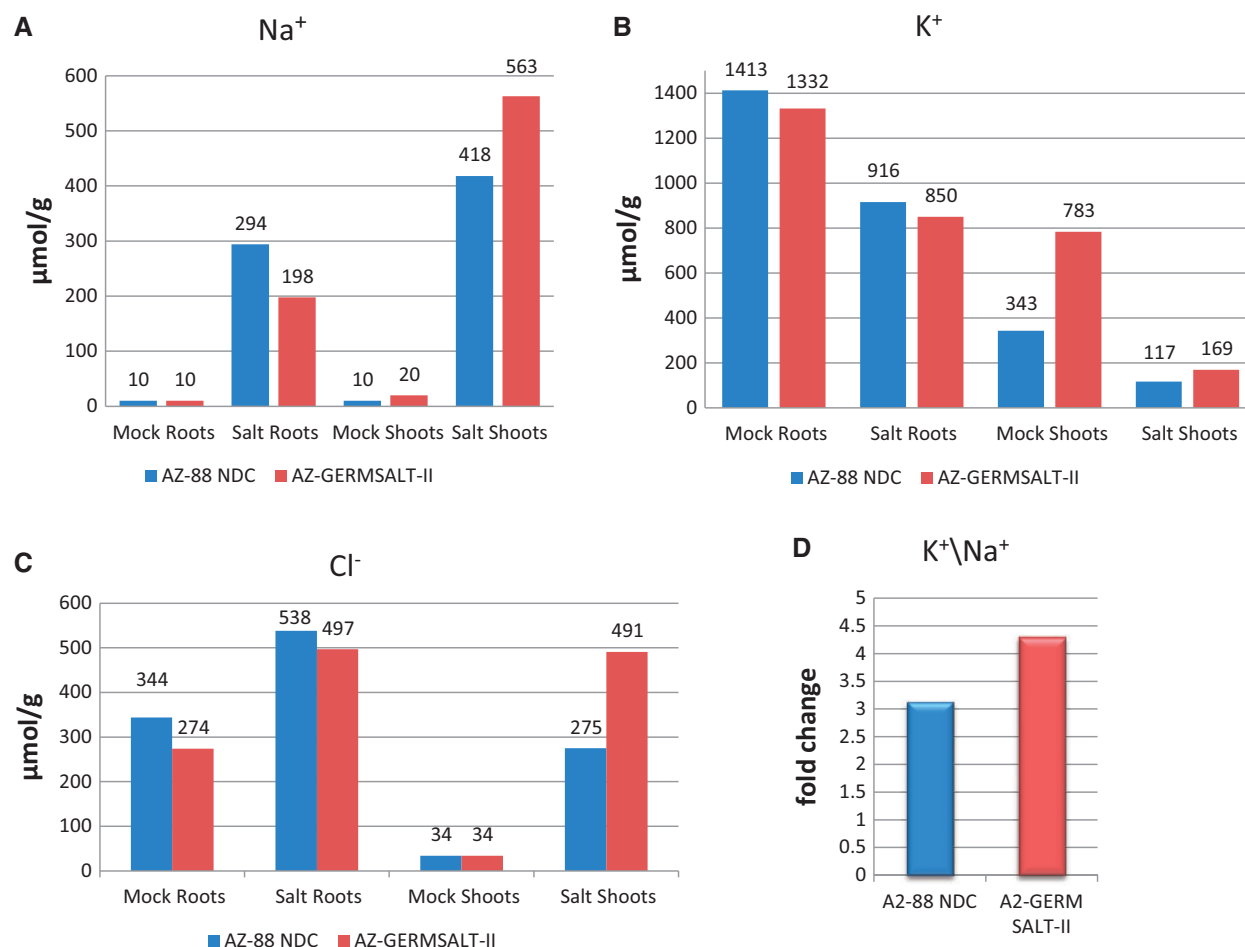


Fig. 1 Ion concentrations in two genotypes under salinity stress. (A–C) Na⁺, K⁺ and Cl⁻ ion concentrations in roots and shoots of salt-treated and non-treated plants. (D) Ratio between K⁺ and Na⁺ concentrations in roots under salt stress.

sequences. A total of 60,290 TCs were obtained and, among them, 44,922 (74.5%) had a significant similarity (blastx 1e-10) to genes in the protein database (**Table 1**). Unannotated sequences fell between 100 and 2,500 bp in length, with the majority of TCs being 150 bp. Only 1.5% of them contained open reading frames (ORFs) at least 100 amino acids long, while others were often interrupted by stop codons. Out of 44,922 assembled TCs, about 90% had similarities or homologs to the *M. truncatula* proteins (Mt3.5v3; blastx cut-off 1e-10). Among 44,922 TCs, 19,625 were non-redundant and 12,620 (64%) had significant hits in *M. truncatula* (blastx 1e-10). The assembly contained 11,449 complete ORFs or 58% of all the non-redundant genes (19,626).

Mining of simple sequence repeats (SSRs; also referred to as microsatellites) from the 60,290 assembled TCs (redundant blast hits often targeted different regions of the same genes) revealed 4,542 total SSRs of six classes (mono-, di-, tri-, tetra-, penta- and hexanucleotide repeats). Among them were 3,240 SSRs belonging to 2,156 annotated unigenes (**Supplementary Table S1**).

AZ-GERM SALT-II. A total of 183,022,614 short reads were generated from cDNA libraries of AZ-GERM SALT-II germplasm

Table 1 De novo assembly statistics

	AZ-88NDC	AZ-GERM SALT-II
Short reads	184,596,972	183,022,614
TCs	60,290	61,874
Non-redundant blastx 1E-10	19,625	19,942
<i>M. truncatula</i> blastx 1E-10	12,620	13,613

(salt tolerant at the germination stage). They were assembled into 61,874 TCs; 45,705 TCs (73.9%) had hits (1e-10) in the nr (NCBI) protein database (**Table 1**). The amount and the size distribution of unannotated sequences were similar to those of AZ-88NDC. Among annotated TCs 19,942 were non-redundant and 13,613 (68%) had significant hits in the *M. truncatula* protein database (**Table 1**). The percentage of the complete ORFs found in the assembly was similar to that of AZ-88NDC.

A total of 4,911 SSRs were identified in all TCs, including 3,526 SSRs in 2,323 annotated unigenes (**Supplementary Table S1**). On average, 9% of TCs contained more than one SSR.

Comparisons of the root transcriptome between the two genotypes. A composition analysis of the root transcriptome in each of the genotypes showed that data obtained by RNA-seq are sufficient for gene expression profiling (Table 1). About 71% of all assembled TCs had 70–100% identity with the protein database (Supplementary Tables S2A, S2B).

The number of common genes found between both lines was 16,504; other unigenes were genotype-specific or differentially expressed low abundant transcripts. This number (16,504) included 1,296 common unigenes containing SSRs. We have also found 274 polymorphic SSRs that were unique for each genotype.

Approximately 39% of the assembled non-redundant TCs in both genotypes had functions assignable by Arabidopsis GO classifications. The GO project aims at standardizing the representation of gene and gene products across species and databases. As seen in Supplementary Table S3, the distribution of TCs among GO functional categories did not differ between the two genotypes.

To examine whether the distribution of our TCs among GO categories is similar to that of *M. truncatula*, we compared the percentage of each GO category with that found in the *M. truncatula* coding sequences (Mt3.0 CDS; Supplementary Table S3). The comparison resulted in similar levels of representation for most of the functional classes, even though all the sequences in the de novo assembly were derived from the root tissues of *M. sativa* vs. all tissues in Mt3.0 CDS. This suggests that TCs identified in this work can serve as a reference sequence database for genomic analysis in alfalfa. TCs in each line were further functionally assigned using the MapMan software tool (Thimm et al. 2004). The composition of the alfalfa root transcriptome was next compared with the available microarray platform, Affymetrix Medicago GenChip Array. It was important to determine how many array probes would correctly bind to alfalfa de novo-assembled TCs identified in this study. More specifically, probe ID should precisely match the respective gene in alfalfa. Following this logic, we performed the analysis with the Affymetrix Medicago GenChip array and found that only 8,727 unique genes (36.5%) represented by probe sets on the array would correctly hybridize with alfalfa root samples. However, the number of matching probes could potentially increase, since some of the sequences on the Affymetrix array might not be present in the alfalfa root transcriptome.

Identification of the salt-responsive transcripts and their functional categorization

Differential expression [>2 fold, false discovery rate (FDR) <0.025] of alfalfa genes was assessed by mapping reads to the de novo-assembled TCs or to the reference genome of *M. truncatula* (see the Materials and Methods) (Fig. 2).

AZ-88NDC. Using de novo assembled TCs, we found 876 differentially expressed TCs in the AZ-88NDC genotype. Among them were 488 differentially expressed genes (DEGs) with

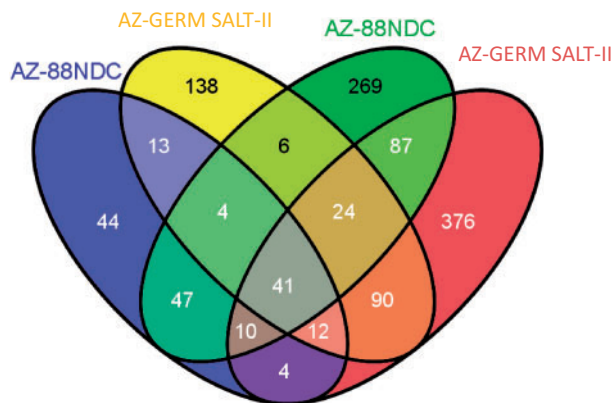


Fig. 2 Comparison between the amounts of DEGs found in the two genotypes by mapping reads to the *de novo* assembled TCs or to the reference genome of *Medicago truncatula*. A Venn diagram depicting the number of statistically significant (>2 -fold) DEGs when the *de novo* transcriptome (green and red ovals, FDR <0.025) or the *M. truncatula* 3.5 data release (blue and yellow ovals, FDR $<1E-04$) were used as a reference for row reads mapping.

non-redundant IDs (blastx $1e-10$). We have also uncovered 161 potentially novel differentially expressed TCs [weak BLAST hits against the nr database ($>1e-10$)]. When *M. truncatula* (Mt3.5 data release) was used as a reference potentially to enrich the resulting transcriptome (Ji et al. 2011), 438 genes significantly (FDR <0.025) changed their expression level in response to salt stress. Nevertheless, when the accuracy of both methods was additionally confirmed by a quantitative real-time PCR (qRT-PCR) screen of the arbitrarily selected 41 DEGs identified by the *de novo* and the reference-based methods, the former showed a slightly higher fidelity ($r = 0.83$ and $r = 0.60$, respectively).

When we further tested unique DEGs mapped to the *M. truncatula* transcriptome only, the correlation coefficient had an even lower value ($r = 0.18$). Some of the target genes were not amplified at all when *M. truncatula* sequences were used to construct the primers, which indicated that the FDR <0.025 cut-off is not sufficient for the reference-based reads mapping (see the Discussion for further explanation). Increasing the threshold to FDR $<1e-4$ led to identification of 175 DEGs. A total of 102 shared DEGs were found by both methods (*de novo* and reference based) and the linear correlation coefficient between these DEGs was $r = 0.9$ (Fig. 2). Finally, we combined the two data sets (*de novo* and reference based) to increase the accuracy of the resulting transcriptome. The final list contained 561 unigenes and, among them, 41 transcription factors (TFs) (Supplementary Table S4). A MapMan representation of the combined data set for AZ-88NDC is shown in Supplementary Fig. S1A and B.

Presumably these genes reflect general transcriptomic responses to salt and include all common genes affected by salt stress in roots of the susceptible check. Among those 561 genes 288 were induced by NaCl and 273 were repressed by at least 2.0-fold (Table 2). Relative to the 19,625 non-redundant TCs of

Table 2 The number of up- and down-regulated DEGs in each line

	AZ-88NDC	AZ-GERM SALT-II
Up-regulated	288	468
Down-regulated	273	337

M. sativa assembled in this study, it means that 2.9% of the accumulated transcripts were responsive to salt treatment.

To find out what biological processes are implicated in response to salt, we assigned them to known GO categories. First, a background was created by taking a portion of each GO category in the set of DEGs and the TCs in the de novo-assembled root transcriptome. Next, the involvement of each GO category in salt stress was projected by calculating a ratio between the portion of the category in the set of DEGs and in the background (de novo assembly). An increase in the category portion would indicate a greater probability of its involvement in response to salinity stress (Fig. 3).

Processes represented by the GO terms 'carbohydrate metabolic process', 'response to stress (biotic and abiotic)', 'lipid metabolic process', 'cell communication' and 'signal transduction' were mobilized in both lines. GO categories that participated in response to salt stress in the AZ-88NDC line only were 'transport' and 'reproduction' (Fig. 3). The proportions of the GO terms 'response to biotic stimuli' and 'response to abiotic stimuli' in the domain 'biological process' were increased in AZ-88NDC. Genes localized in the GO 'extracellular region', 'cell wall' and 'external encapsulating' terms were also involved in response to stress. In the domain 'molecular function', a portion of the GO 'carbohydrate binding', 'receptor', 'hydrolase' and 'transporter' activities were increased in both lines among the genes that responded differentially to salt stress. The GO terms 'nuclease activity', 'DNA binding' and 'transcription factor activity' were found only in AZ-88NDC.

One of the important ways to gain insights into the biological meaning of the large-scale genomic experiment is to describe how the list of genes that share specific common characteristics, such as differential expression, relates to the set of all genes (population set) revealed by the high-throughput trial (Grossmann et al. 2007). An over-represented category of genes revealed in the salt-susceptible check AZ-88NDC using the SEA tool agriGO (Du et al. 2010) was the GO term 'lipid transport'.

AZ-GERM SALT-II. Analysis of the de novo transcriptomes showed that in the AZ-GERM SALT-II (salt-tolerant line at germination stage) 1,303 TCs significantly (>2-fold, FDR <0.025) changed their expression level in response to 150 mM salt. Among them were 644 with non-redundant IDs (blastx e-value cut-off 1e-10) (Fig. 2). As many as 277 differentially expressed TCs were found in AZ-GERM SALT-II that did not result in any strong BLAST hits against the nr database with an e-value of $\leq 1e-10$.

When reads were mapped to a reference genome of *M. truncatula*, 619 (FDR <0.025) and 328 (FDR <1E-04) genes significantly changed their expression level in response to salt stress (Fig. 2). A combination of these two data sets led to the identification of 805 DEGs, including 167 shared DEGs found by both methods. A combined list of DEGs contained 61 TFs (Supplementary Table S4).

A MapMan representation of the data set for AZ-GERM SALT-II is shown in Supplementary Fig. S2A and B. Among those 805 genes, 468 were induced by NaCl and 337 were repressed by at least 2.0-fold (Table 2). Relative to the 19,942 assembled non-redundant TCs, the percentage of the transcripts affected by salt stress was 4%.

Assignment of the DEGs to GO terms resulted in a somewhat different distribution (Fig. 3). For example, fractions of the categories 'response to stress' and 'transport' were decreased, but the fraction of the responsive genes located in the 'extracellular region' and 'cell wall' categories increased as compared with the AZ-88NDC genotype. The most noticeable increase in gene fractions in the GO domain 'molecular function' was in the two categories 'hydrolase activity' and 'enzyme regulator activity'. Some of the over-represented categories included 'hydrolase activity' (GO:0016798), 'oxidoreductase activity' (GO:0016491) and 'extracellular region' (GO:0005576).

Comparison of the differentially expressed genes between the two genotypes in response to salt stress. To estimate the differential effect of salt stress on gene expression in the two genotypes, we compared sets of DEGs between AZ-88NDC and AZ-GERM SALT-II. The resultant DEG sets obtained by the two different methods for each of the genotypes as well as the combined DEGs are presented in Fig. 2. The combined list of DEGs contained 201 common genes and 964 distinct genes differentially expressed in both genotypes in response to salt treatment (Supplementary Table S5). Presumably, common DEGs represent general transcriptomic responses to salt stress in roots of *M. sativa* and may be similar in other accessions.

The distribution of the common DEGs among functional GO categories in response to salt is shown in Supplementary Table S5. The prevalent categories include 'binding', 'metabolic process', 'membrane' and 'transport'. There were 105 common up-regulated and 88 down-regulated genes between the two genotypes. Only eight genes were differentially expressed in response to salt in the two lines.

In both accessions, there were more up-regulated genes in the GO category 'catabolic process', whereas in the categories 'signal transduction' and 'receptor activity' there were more down-regulated genes (Fig. 4). Characteristically, genes in 'response to stress' and 'kinase activity' categories were more repressed in the AZ-88NDC line, whereas in the AZ-GERM SALT-II the situation was reversed (more up-regulated genes found in the same categories). More up-regulated genes were present in the category 'extracellular region' in the AZ-GERM SALT-II line than in AZ-88NDC.

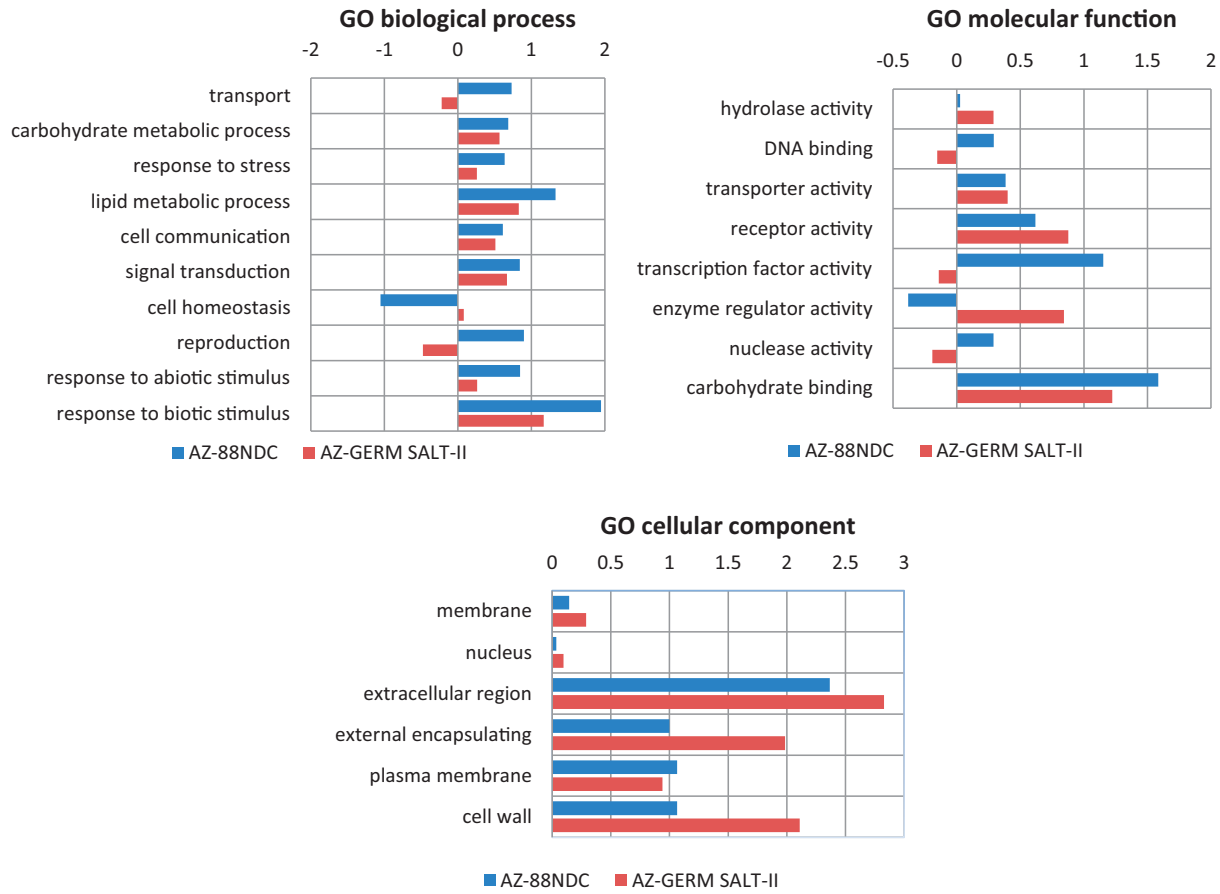


Fig. 3 Distribution of GO categories involved in the salt stress response among three major GO domains. The bar represents the log₂-transformed ratio between category portions in DEGs and the background root transcriptome. A ratio greater than zero indicates a gain in the portion of the GO category in response to salt.

DEGs that are different in the two lines correspond to genotype-specific responses to salt. Among 964 distinct DEGs, 604 were specific to the salt-tolerant line AZ-GERM SALT-II and 360 to the salt-sensitive line AZ-88NDC (**Supplementary Table S5**). Based on the MapMan annotation, 50 DEGs (among them 35 unique for this line) in AZ-88NDC are known to be involved in stress responses and 23 of them respond to abiotic stress. In AZ-GERM SALT-II germplasm, 63 DEGs (48 unique for this line) were classified as responsive to stress and, among them, 26 DEGs are known to respond to abiotic factors (**Supplementary Table S5**).

We further assessed the distribution of TFs in both lines using MapMan functional classes. Five and eight TFs in AZ-88NDC and AZ-GERM SALT-II, respectively, contained an AP2/ERF domain specific for the AP2/ERBP supergene family that plays a variety of roles throughout the plant life cycle (Riechmann and Meyerowitz 1998). A group of basic helix–loop–helix (bHLH)-type TFs, which are involved in root growth and development, were found in both genotypes (three in AZ-88NDC and four in AZ-GERM SALT-II), among the genes responsive to salinity stress in the alfalfa root transcriptome (**Supplementary Table S4**). There were several TF

families, such as, for example, three up-regulated GRAS TFs (Pysh *et al.* 1999), that responded to salt treatment in AZ-GERM SALT-II but not in AZ-88NDC.

As seen in **Fig. 1**, the salt-tolerant line AZ-GERM SALT-II accumulates less sodium under salinity stress and has a better K⁺/Na⁺ ratio in roots, as compared with the salt-sensitive line AZ-88NDC. The AZ-GERM SALT-II line has better K⁺ uptake and less accumulation of Cl⁻ ion in roots under normal conditions (**Fig. 1C, D**). To identify genes that are possibly employed by AZ-GERM SALT-II for different ion distribution and salt tolerance, we assessed the level of basal gene expression between both genotypes under normal control conditions.

A total of 544 DEGs were found in the two lines under control conditions (**Supplementary Table S6**). Some of the genes such as albumin (Medtr3g067430) and HVA22 had increased basal levels in the AZ-GERM SALT-II line. Albumin can potentially play an osmoprotectant role and HVA22 regulates vesicular traffic in stressed cells and is induced by environmental stresses, such as dehydration, salinity and extreme temperatures (Brands *et al.* 2002).

Comparison between DEGs found under control conditions and common DEGs found in both lines under salt treatment

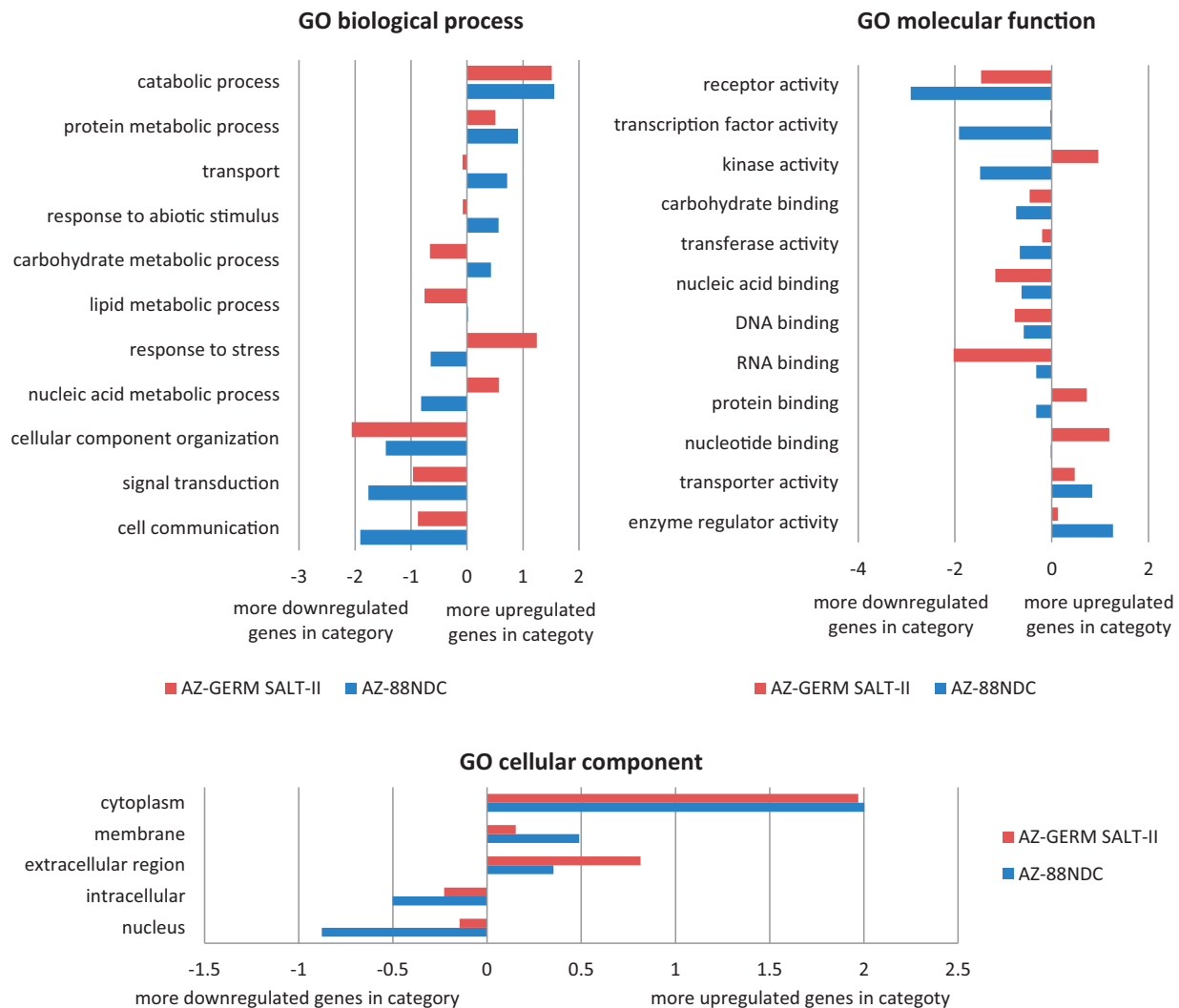


Fig. 4 Distribution of the sets of up- and down-regulated genes among GO categories in response to salt in the two tested genotypes. The bar represents the log₂-transformed ratio between category portions in up-regulated and down-regulated gene sets. A ratio greater than zero indicates that there were more up-regulated genes in the category.

resulted in only 17 common genes (FDR <0.025), thus indicating that different sets of genes are active in the two lines under salinity stress as compared with normal conditions.

To confirm this, we compared the 544 genes (see above) with all 1,165 DEGs identified in both lines under salt treatment. This resulted in 56 and 100 common genes in AZ-88NDC and AZ-GERM SALT-II, respectively, that had a different basal expression level between the two genotypes and were also involved in response to salt stress. Several DEGs involved in secondary metabolism of phenylpropanoids and isoflavones, which possess antioxidant activities, were identified among these common genes (Supplementary Table S6). Other genes associated with the antioxidant system and ion transport functions were also identified, such as the K⁺/H⁺ antiporter and vacuolar H⁺-pumping ATPase (Supplementary Table S6).

Using the total size and composition of the alfalfa root transcriptome as well as the response of each genotype to salt stress,

we have attempted to estimate the number of genes, whose expression levels could be determined using available *M. truncatula* microarray platforms. As a result, we found that only about 50% of all DEGs identified in this study are represented on the Affymetrix chip of *M. truncatula* and would correctly bind to the microarray probes.

QRT-PCR with selected gene candidates from both lines

To validate NGS results further, we performed qRT-PCR with 45 selected genes differentially expressed in both genotypes in response to salt stress. The RNA-seq-generated data showed a strong correlation with the qRT-PCR results (Pearson correlation coefficients $r = 0.84$; Supplementary Table S7). Table 3 and Fig. 5 show the most interesting genes differentially expressed in both lines. Among up-regulated genes are dihydroflavonol reductase (DFR; Medtr3g031650.1), a key enzyme

Table 3 Validation of NGS data by qRT-PCR

Gene ID	AZ-88NDC		Gene ID	AZ-GERM SALT-II	
	Log ₂ -fold change			Log ₂ -fold change	
	NGS	qRT-PCR		NGS	qRT-PCR
Medtr1g075030.1	4.79	4.08	Medtr2g060880.1	6.54	3.15
Medtr3g031650.1	3.69	2.43	Medtr5g043550.1	4.06	2.30
Medtr4g098570.1	3.04	1.80	Medtr2g010590.1	3.42	1.57
Medtr5g043550.1	2.83	1.89	Medtr3g114530.1	3.13	2.66
Medtr3g114530.1	2.71	2.00	Medtr1g073170.1	2.90	0.93
Medtr1g073170.1	2.43	3.06	Medtr3g005420.1	2.81	3.25
Medtr8g020630.1	1.25	3.21	Medtr3g031650.1	2.23	1.16
Medtr4g098850.1	-1.59	-1.76	Medtr3g103960.1	-1.95	-1.76
Medtr3g103960.1	-1.63	-2.51	Medtr3g071740.1	-2.17	-1.23
Medtr5g024020.1	-1.95	-1.08	Medtr2g020710.1	-2.92	-1.47
Medtr1g012710.1	-2.06	-1.90	Medtr1g093600.1	-3.03	-1.17
Medtr3g071740.1	-2.76	-2.34	Medtr5g024020.1	-3.06	-0.99
Medtr1g093600.1	-3.26	-1.02	Medtr1g012710.1	-3.36	-0.52
Medtr2g020710.1	-3.81	-2.86	Medtr4g098850.1	-3.83	-1.62

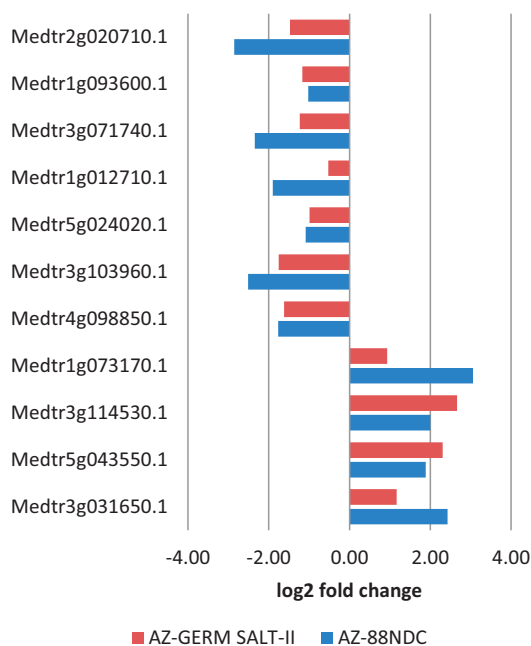


Fig. 5 QRT-PCR screening of the selected DEGs shared between the two genotypes.

involved in anthocyanin biosynthesis, C2H2 zinc finger protein (Medtr2g060880.1), a transcription activator whose expression levels increased >6-fold, class III acidic chitinase (Medtr5g043550.1) of plant chitinases that are known to play a role in defense, growth and development (Grover, 2012), sieve element-occluding protein (Medtr1g075030.1), a putative component of the forisome (Pélissier et al. 2008), and other genes (Table 3).

Among the down-regulated genes were sugar/inositol transporter ERD6-like 16 (Medtr2g020710.1), which is an abiotic stress-responsive transporter expressed specifically in roots (Legume IP database, The Samuel Roberts Noble Foundation), inositol-145-trisphosphate 5-phosphatase (Medtr4g098850.1) (Munnik and Vermeer, 2010), ABA receptor PYL6 (Medtr3g071740.1) (Qin et al. 2011), and other genes.

Discussion

The primary goal of this work was to identify a broad spectrum of salt-responsive genes in alfalfa in order to estimate a range of physiological, metabolic and cellular processes influenced by salt stress. This knowledge would allow characterization of the most vulnerable pathways associated with susceptibility to salt in alfalfa and thus identification of salt-sensitive germplasm. This information would also be used to identify genes/pathways that are involved in salt stress resistance.

The root system is the first to be exposed to and to cope with soil salinity, and evaluation of plant responses to salt is often based on the assessment of the root environment. In this study, we analyzed the root transcriptome of two different alfalfa accessions, AZ-88NDC and AZ-GERM SALT-II, in response to salinity stress. The former genotype is described as susceptible check (Smith and Fairbanks 1989, Smith 1991), while the latter accession is usually characterized as salt tolerant at the germination stage (Dobrenz et al. 1989).

Defining salt tolerance in alfalfa, especially at the germination phase, is difficult due to the substantial differences in response to salinity at various growth stages (Scasta et al. 2012). Behind these differences lie specific gene expression programs regulating tolerance to salt at each particular growth

period. Germplasm can be tolerant at germination, based on percentage survival, but sensitive to salt during emergence, early vegetative development and later developmental stages, based on growth reductions (Lauchli and Grattan 2007). Root samples for NGS were collected at the seedling stage, whereas AZ-GERM SALT-II is regarded as a salt-tolerant line during seed germination. Hence, there was a chance that tolerance could be changed at the time of sample collection.

Measurement of the concentration of salt ions in both genotypes has shown considerable accumulation of sodium and chloride ions in roots and shoots of both genotypes in response to salt treatment. Importantly, however, that sodium accumulation was lower in the roots of AZ-GERM SALT-II, suggesting a mechanism of Na⁺ exclusion from uptake in the roots of this germplasm. Shoots of AZ-GERM SALT-II were also capable of accumulating and storing more Na⁺ than the susceptible germplasm AZ-88NDC. A similar tendency of ion distribution was previously demonstrated for salt-tolerant AZ-GERM SALT-II germplasm as well as for the salt-adapted genotype of *M. truncatula* (Ashraf and O'Leary 1994). Sodium uptake and compartmentalization at the cellular and intracellular level in the leaf tissue is one of the known mechanisms of salinity tolerance (Munns and Tester 2008, Smethurst et al. 2008).

The study was conducted in the following order: first, sequence reads derived from each of the genotypes separately as well as from both of them combined were assembled into the respective contigs and transcripts. Next, the total size and composition of the resultant root transcriptomes were analyzed. Lastly, a detailed comparison of the transcriptome data between the two lines and assignment of salt-responsive genes to known functional groups, biological processes, using Arabidopsis GO was performed.

Several important findings emerged from our study. A global root transcriptome profile of *M. sativa* (both mock and experiment) derived from a large number (megabases of nucleotide sequences) of transcripts (sequence reads) contained a significant portion of unannotated genes or genes without any homology in *M. truncatula*. When *M. truncatula* sequences were used as a reference to identify DEGs, the accuracy of the obtained data, according to qRT-PCR screening, was low. This indicated that the two species may have a significant difference on the DNA level. The pool of common TCs shared between the two genotypes was significant (16,504) and included a majority of the DEGs found in both lines. Many common DEGs, including TFs, were found in both genotypes. The distribution of all TCs among functional categories did not differ between the genotypes; however, the distribution of the DEGs differed. A number of SSRs have been identified; after further analysis and evaluation, they may be used to search for quantitative trait loci (QTLs) associated with salt tolerance.

Unannotated tentative consensus sequences

Predominant self-incompatibility and obligate outcrossing may have supported higher levels of heterozygosity in alfalfa

populations as compared with its diploid self-fertile relative *M. truncatula*. This would undoubtedly contribute to a greater genetic diversity of *M. sativa*, its variability and adaptation to the environment.

Unlike microarray technology that is limited to studying expression of previously annotated genes, Illumina RNA-seq method generates whole transcriptome profiles and is capable of identifying new genes, transcript variants and genetic polymorphisms, as well as non-coding regulatory sequences (Mizuno et al. 2010). In addition, RNA-seq permits estimation of gene expression levels more precisely due to some known limitations of microarray technology such as, for example, probe length. When 50 bp reads from alfalfa root mRNA were assembled into TCs and submitted to blastx against the nr protein database, about 75% of them had BLAST hits and roughly 25% had no BLAST hits at all (unannotated transcripts). When an ORF search was performed in the unannotated transcripts derived from both lines, only 466 ORFs were found. However, no conserved domains (CD-search tool E-value <1e-10) were identified in these ORFs. Among approximately 75% annotated transcripts, about 90% had hits with the *M. truncatula* genome and 10% did not. When a BLAST search was performed on those genes that were not identified in *M. truncatula*, we found (based on the first-best hit) homologous genes in 24 plant species other than *M. truncatula*. About 74% of them belonged to *Glycine max*, 5% to *Vitis vinifera*, 3% to *Populus trichocarpa*, 1% to *Pisum sativum* and 17% to other species. By presumption, these could be alfalfa-specific genes, that are absent in its close relative *M. truncatula*, which might shape the difference between these two species.

Differences at the DNA level between *M. sativa* and *M. truncatula*

To estimate possible differences on the DNA level between *M. sativa* and *M. truncatula*, a blastn search of our experimentally found TCs was performed against the *M. truncatula* CDS. Results showed that 24% of all TCs (based on the combined length of TCs) did not have significant (1e-10) hits in *M. truncatula* CDS. Considering a high overall percentage of identity based on the best hit (77–100%), part of those alfalfa-specific TCs have probably mapped to the wrong sites in the *M. truncatula* CDS. Sequence identity distribution of the TCs (>300 bp) on the *M. truncatula* CDS are shown in Fig. 6.

Differentially expressed genes

Differences in gene expression changes between the two alfalfa lines in response to salt stress are outlined in Table 2 and Supplementary Tables S5 and S6.

A total of 1,165 unique DEGs were identified in both genotypes in response to salt stress. Among 604 DEGs specific to the salt-tolerant line AZ-GERM SALT-II only, we found 12 different ABC transporters (including 10 up-regulated) implicated in transporting stress-related secondary metabolites such as alkaloids, terpenoids, polyphenols and quinines (Zhang et al. 2012),

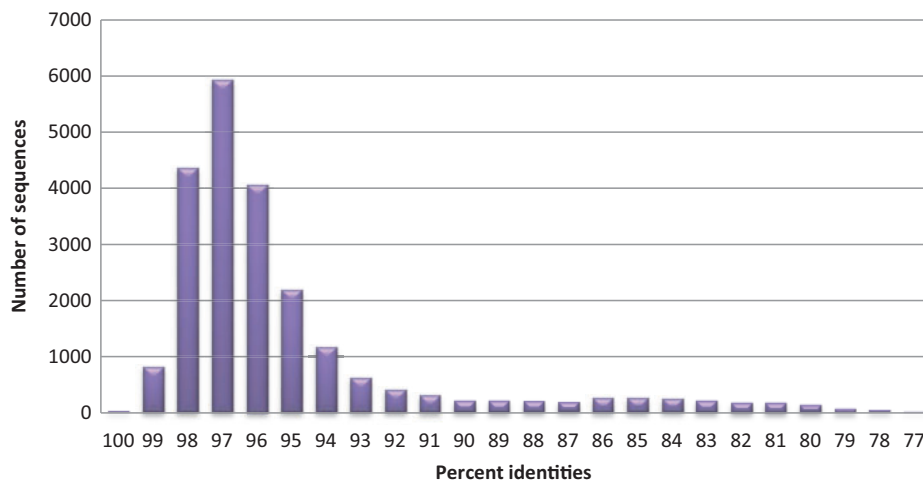


Fig. 6 Sequence identity distribution of TCs (>300 bp) to *M. truncatula* CDS.

and three nitrate transporters, one of which was up-regulated and the other two down-regulated (**Supplementary Table S6**). It is known that moderate NaCl stress affects nitrate assimilation (Dluzniewska *et al.* 2007). A gene encoding remorin (Bariola *et al.* 2004), a plant-specific plasma membrane protein (Medtr5g010590.1), was salt induced in AZ-GERM SALT-II. Remorin may contribute to the stabilization of the damaged plasma membrane under salt stress (Zhang *et al.* 2012).

Between the two lines, only four genes (out of the 50 and 63 DEGs in AZ-88NDC and AZ-GERM SALT-II, respectively) were previously demonstrated to respond to salt or drought stress, based on the MapMan annotations. These are Medtr2g033550.1 (desiccation-related protein), Medtr2g081610.1, Medtr3g109490.1 (RD22, nutrient reservoir, responsive to dehydration; mediated by ABA) and Medtr5g042560.1 (DUF221 early-responsive to dehydration). The rest of the DEGs, including 45 unique genes in the salt-tolerant line AZ-GERM SALT-II, were not shown to be involved in salt stress prior to this work (**Supplementary Table S5**). Interestingly enough, both lines contained a substantial number (16–20) of pathogenesis-related (PR) and other disease-responsive genes. It is known that abiotic stresses can elicit defense-related protein induction that can serve a role in protection of cellular structures against the stress (Van Loon *et al.* 2006).

Some of the salt-responsive genes, especially those involved in the antioxidant system and transport functions (ion translocations across cell membrane), may provide a clue to understanding the mechanisms responsible for salt tolerance in AZ-GERM SALT-II. For example, K^+/H^+ antiporter (Medtr7g099800) and vacuolar H^+ -pumping ATPase (Medtr8g076150) were up-regulated in the salt-tolerant line on the basal level, i.e. under normal conditions (**Supplementary Table S6**). It is known that higher activity of antiporters is associated with salt tolerance (Munns and Tester 2008). The increased activity of vacuolar H^+ -ATPases is indicative of the Na^+ sequestration and osmotic adjustment under salt stress (Zhang *et al.* 2012).

Among up-regulated genes are DFR (Medtr3g031650.1), a key enzyme involved in anthocyanin biosynthesis, and C2H2

zinc finger protein, a transcription activator whose expression levels increased >6-fold. It was previously reported that transgenic rice overexpressing DFR increased tolerance to multiple environmental stresses (Uchimiya *et al.* 2002). The Cys2/His2-type (C2H2) zinc finger proteins have been implicated in plant development and stress responses, including salt tolerance (Sun *et al.* 2010).

There has also been a considerable activity of genes encoding secondary metabolites such as phenylpropanoids that are known to take part in lignin biosynthesis of the cell wall and antioxidant activity, and play important roles in interactions with the biotic and abiotic environment (Dixon *et al.* 2002, Bonawitz *et al.* 2012). More than 10 genes involved in the phenylpropanoid pathway were up-regulated in the salt-tolerant line AZ-GERM SALT-II (**Supplementary Table S6**). Class III acidic chitinase of plant chitinases was also up-regulated. Chitinase genes are expressed in diverse conditions. They are strongly induced when host plant cells are challenged by pathogen stress and are also involved in plant abiotic stress responses such as osmotic, salt, cold, wounding and heavy metal stresses (Grover 2012).

Underexpression of inositol-145-trisphosphate 5-phosphatase, registered in both genotypes under salt conditions, can cause accumulation of the inositol-145-trisphosphate (IP_3) gene product in the cells. IP_3 is a soluble secondary messenger molecule that is involved in mobilization of intracellular Ca^{2+} . IP_3 -induced increases in cytoplasmic Ca^{2+} are responsible for a large number of receptor-initiated signaling pathways, including activation of stress responses via Ca^{2+} -regulated TFs and Ca^{2+} -responsive promoter elements (Taylor and Thorn 2001, Burnette *et al.* 2003, Stevenson *et al.* 2003).

Down-regulation of the ABA receptor PYL6 can make alfalfa less sensitive to ABA signaling in response to salt stress. ABA is a stress hormone that rapidly accumulates in response to stress. Once stress is relieved, ABA should be degraded so that normal plant growth and functions can resume. Some previous studies on a complex ABA signaling in plants indicated that low levels

of ABA are more adaptive under salt stress and a higher concentration of ABA can be inhibitory to germination and leaf expansion (Zhu 2000).

Ninety-nine DEGs had hits in 10 plant species other than *M. truncatula*, with the range of response to salt treatment from about -3-fold (down-regulation) to +6-fold (up-regulation, log₂-transformed). Apparently, these DEGs represent salt-responsive genes specific for *M. sativa*. As many as 201 common DEGs were found in both genotypes among annotated and unannotated genes, and this number included 15 TFs. The activity (induced or repressed) of 193 DEGs was similar in both lines, showing the roles of these genes in response to salt.

Although the distribution of all TCs among functional categories did not vary between the genotypes, the distribution of the DEGs differed, thus reflecting a genotype-specific response to salt stress (Fig. 3). Among GO categories with the most contrasting gene distribution were 'transport', 'hydrolase activity', 'enzyme regulator activity' and 'cell wall'. On the other hand, amongst the most similar GO categories (only positive values) were 'carbohydrate metabolic process', 'lipid metabolic process', 'signal transduction', 'response to stress', 'extracellular region', 'transporter activity' and 'receptor activity'. In the susceptible line AZ-88NDC, there were more repressed genes in the 'response to stress' and 'kinase activity' categories. In contrast, in AZ-GERM SALT-II, more up-regulated genes were found in the same categories and also in the category 'extracellular region' (Fig. 4). The GO term 'extracellular region' includes the child term 'apoplast', that is external to the plasma membrane and includes cell walls, intercellular spaces and the lumen of dead structures such as xylem vessels. Water and solutes pass freely through it. More importantly, different ion channels and transporters crucial for ion uptake and regulation of xylem loading are located in the apoplast.

However, only about 40% of all genes were assigned to a GO category. Therefore, the current picture may change as more information about gene functions in legumes is elucidated. In addition, functions of similar proteins may be different in *Arabidopsis* and legume species such as *M. sativa*.

Based on this information, we assembled lists of genes falling under the contrasting or similar categories (Supplementary Table S8). This partitioning provides a simplified view of the involvement of different biological pathways in genotype-specific and species-specific responses to salinity.

TFs found in the total transcriptome and differentially expressed TFs

The assembled transcriptome contained ~726 unique TFs that had homology with *M. truncatula*. They were classified according to the Legume TF database (Li et al. 2012). Among them were 86 TFs responsive to salt treatment that represented 12% of all TFs. They include major TF families linked to plant stress responses, such as MYB, AP2/ERBP, zinc finger family, bHLH and bZIP. The composition of these differentially expressed TFs in both genotypes differs from that of the rest of

the TFs identified in the transcriptomes. The most represented TF family in both lines was AP2/ERBP, based on the >2-fold increase in the portion of this TF class vs. the full list of TFs found in this study. This family of TFs, which has not been characterized in alfalfa, is known to play an important role in responses to abiotic and biotic factors in model plants with annotated genomes such as *A. thaliana* but has not been characterized in alfalfa. A group of bHLH-type TFs, involved in root growth and development, were found in both genotypes (three in AZ-88NDC and four in AZ-GERM SALT-II; Supplementary Table S4). It has been previously demonstrated that bHLH-type TFs are linked to the adaptation of *M. truncatula* to saline soil environments (Zahaf et al. 2012). GRAS TFs that responded to salt treatment in AZ-GERM SALT-II but not in AZ-88NDC are associated with root development programs (Pysh et al. 1999).

MYB TFs (similar to Medtr1g073170.1, regulation of transcription) and HB Homeobox TFs (similar to Medtr8g026960) were among the most up-regulated, while AP2/ERBP and APETALA2/ethylene-responsive element-binding protein family (similar to Medtr1g093600.1 and Medtr5g053920.1, respectively) were among the most down-regulated TFs under salt treatment in both lines (Supplementary Table S4).

Induction or repression of MYB TFs, representatives of the largest superfamily of plant TFs involved in a variety of processes (Cominelli and Tonelli 2009, Dubos et al. 2010), indicates that metabolic, developmental, signal transduction and defense-related pathways are all affected by salinity. For instance, Medtr1g073170.1, a member of the R2R3-MYB subfamily that has been characterized for its roles in phenylpropanoid metabolism, determination of specialized cell morphology and responses to biotic and abiotic stresses (Cominelli and Tonelli 2009), is up-regulated >5-fold in AZ-88NDC and >7-fold in AZ-GERM SALT-II.

There were several TF families and/or different representatives of the same families that responded to salt stress only in AZ-GERM SALT-II, such as GRAS, ARR (response regulators involved in cytokinin signal transduction; Sakai et al. 2001), JUMONJI (histone lysine demethylase; Agger et al. 2008) and MYB (Table 4).

Conclusions

We have demonstrated by high-throughput sequencing technology that salinity stress in alfalfa affects a variety of genes belonging to different biological processes, signaling pathways and regulatory networks. Different sets of genes were active in the two lines under salinity stress as compared with normal conditions. Among the most influenced, based on expression levels in both lines, were many unknown or predicted proteins and proteins with known function such as DFR, MYB59, sugar transporter ERD6-like 16 and inositol-145-trisphosphate 5-phosphatase. In the salt-tolerant line, we have identified 604 genes that changed their expression in response to salinity. They included genes implicated in transporting stress-related secondary metabolites, nitrogen uptake and stabilization of the

Table 4 Transcription factors differentially expressed in the two genotypes under salinity stress

Transcription factor family	AZ-88NDC	AZ-GERM SALT-II
MYB	3	10
AP2/EREBP	5	8
bZIP	5	6
HB, Homeobox	2	4
bHLH	3	4
GRAS	0	3
CCAAT	3	3
C3H zinc finger	2	3
AS2	3	3
C2H2 zinc finger	2	3
ARR	0	2
MADS	2	1
JUMONJI	0	1
G2-like, GARP	2	0
WRKY	1	0
TCP	1	0
NAC	1	0
Other	6	10
Total	41	61

plasma membrane, as well as several families of salt-responsive TFs. The most affected biological functions were assigned to GO categories using Arabidopsis GO. Genes responsive to stress were identified in alfalfa that are absent in its close relative, *M. truncatula*.

One of the important conclusions of this work is that only about 36.5% of the unique genes on the *Medicago* microarray would hybridize with alfalfa TCs identified in this study, even though the number of matching probes could be more, since some of the probes on the Affymetrix array might not be present in the alfalfa root transcriptome. Among salt-responsive genes, about 50% of all discovered DEGs could be found on and/or correctly bind to *M. truncatula* microarrays.

Overall, species-specific data on global transcription profiling of salt-responsive genes obtained in this study will be useful in understanding the molecular mechanisms of salinity tolerance in alfalfa needed to maximize yield and economic return.

Materials and Methods

Plant material and growth conditions

Two lines were used in the experiments: a susceptible check AZ-88NDC (Smith and Fairbanks 1989, Smith 1991) and AZ-GERM SALT-II, germplasm salt tolerant at the germination stage (Dobrenz et al. 1989). Prior to the experiments, seeds were tested for their germination ability/salt sensitivity at increased NaCl concentration (250 mM). Germinated and hard seeds were counted after 7 d. Germination of control and salt-treated seeds

was computed by dividing the number of seeds germinated under salt conditions by the corrected total to obtain the fraction of germinated seeds (Rumbaugh, 1991). To optimize salt concentrations, seedlings were grown on Murashige and Skoog (MS) medium (Murashige and Skoog 1962) containing 100–250 mM NaCl. Since the latter concentration caused severe inhibition of growth, moderate concentrations of 150 mM were used in the experiments with germinated seedlings.

Salt stress was imitated by incubation in a growth chamber on half-strength MS medium containing 3% sucrose and an optimized salt concentration (Wang et al. 2009). For germination on the medium, seeds were scarified with H₂SO₄, surface-sterilized with 70% ethanol for 1 min and with 0.5% sodium hypochloride solution for 5 min, rinsed with distilled water and placed on half-strength MS medium, pH 5.7. One-week-old seedlings were transferred to salt or control (no salt added) medium. Root samples were collected (4–5 bioreplicates for treatment) after 1 week under salt stress. The whole roots were cut off the seedlings, frozen in liquid nitrogen and ground to a fine powder for RNA extraction.

Ion measurements

Ion concentrations were determined by VHG Labs (Manchester, NH, USA) on a fee-for-service basis. Forty roots and 40 shoots per genotype were analyzed (20 for mock and 20 for salt-treated plants). Samples were pooled in order to reduce the effects of biological variation and to have a sufficient amount of tissue for the accurate measurements. For the Na⁺ and K⁺ determinations, the samples were completely digested in HNO₃. They were run on a PE Optima ICP-OES spectrometer against NIST traceable standards and controls. For Cl⁻ determination, samples were placed in deionized water, sonicated for 30 min and left for 24 h to extract the Cl⁻. They were run on a Metrohm 881 Compact Ion Chromatograph against NIST traceable standards and controls.

RNA extraction, first-strand synthesis and quantitative real-time PCR

Total RNA was extracted by maceration of the tissues in liquid nitrogen followed by isolation of total RNA with a FastRNATM Pro Green Kit (MP Biomedicals), and subsequent purification using RNA easy columns (RNeasy Mini Kit, Qiagen Inc.). cDNA was synthesized using SuperScript First-Strand cDNA Synthesis System according to the manufacturer's directions (Invitrogen Corp.). Real-time PCR was performed with an iQ SYBR Green Supermix kit (Bio-Rad Laboratories, Inc.) on the MiniOpticon Real-Time PCR system (Bio-Rad) using the following parameters: 94°C for 1 min (one cycle); 94°C for 30 s, 60°C for 30 s, 72°C for 30 s (30 cycles). Amplification was performed in four biological and two technical replicas. The *M. sativa* actin-depolymerizing factor (ADF) gene, a putative housekeeping gene, and NP_001237047, an unknown gene from this study with little variation in expression levels, were used as reference in all real-time PCR experiments (Yang et al. 2011). NP_001237047

showed much less variation between salt and mock samples, and qRT-PCR results were more consistent with processed RNA-seq data than when ADF was used.

RNA-seq, transcriptome assembly and analysis

NGS was performed by the Expression Analysis DNA/RNA Services (Durham, NC, USA) for a fee. Reads were obtained from the Illumina HiSeq 2000 sequencing system. There were four replicates for each experimental condition. For each of the two lines, eight samples were sequenced (four control and four salt). A total of 184,596,972 and 183,022,614 short reads (50 bp) were generated from cDNA libraries originated from AZ-88NDC (susceptible check) and AZ-GERM SALT-II (salt tolerant at the germination stage) germplasms, respectively.

Two approaches were used for detection of DEGs. In the first approach, *M. truncatula* was used as a matrix for mapping of reads. This was accomplished with the help of Bowtie-0.12.7, an ultrafast, memory-efficient short read aligner (<http://bowtie-bio.sourceforge.net/index.shtml>; Langmead et al. 2009). Obtained count data were used to estimate the expression level of transcripts applying the DEseq method (Anders and Huber 2010). De novo assembly was performed directly from a set of sequence reads using Velvet 1.2.03 and Oases 0.2.06 genome assemblers (Martin and Wang, 2011). The resultant transcriptome served as a template for gene expression profiling. A paired-end assembly was conducted for each of the studied alfalfa lines. This analysis was done applying various *k*-mers (25, 31, 33, 35, 45 and 47) and default parameters. The average insert size was 280 nucleotides. The mean value and N50 number from each data set were examined and found to be the highest and best while using a *k*-mer of 45. Using a *k*-mer of 45, the mean value for each contig was 348 bases and the N50 number was 591 bases. We examined the quality of the transcripts by blasting them against the *M. truncatula* genome and examining the percentage identity. A subset of transcripts were chosen at random and examined individually and manually for an ORF and for the 3'- and 5'-untranslated regions. Overall, a *k*-mer of 45 gave the best results. Assemblies that resulted in complete intact transcripts were chosen in order to estimate expression levels correctly. At lower *k*-mer values, the transcripts were less intact and often were fragmented. At *k*-mer 47, the N50 value decreased.

After transcriptome assembly, genes were annotated based on the blastx hits and assigned to known functional groups, biological processes, using GO tools, such as the Gene Ontology database and MapMan (Thimm et al. 2004). Functional categorization by GO terms was performed based on the best BLASTX hits against the nr database using BLAST2GO software (Conesa et al. 2005). Venn diagrams were produced using the Venny tool (Oliveros 2007).

Supplementary data

Supplementary data are available at PCP online.

Funding

This work was supported by the United States Department of Agriculture, Agricultural Research Service.

Acknowledgments

We thank Dr R.A. Owens for his invaluable contribution to the project and Dr. R.E. Davis for critical reading of the manuscript.

References

- Agger, K., Christensen, J., Cloos, P.A. and Helin, K. (2008) The emerging functions of histone demethylases. *Curr. Opin. Genet. Dev.* 2: 159–168.
- Anders, S. and Huber, W. (2010) Differential expression analysis for sequence count data. *Genome Biol.* 11: R106.
- Apse, M.P., Aharon, G.S., Snedden, W.A. and Blumwald, E. (1999) Salt tolerance conferred by overexpression of a vacuolar Na⁺/H⁺ antiport in *Arabidopsis*. *Science* 285: 1256–1258.
- Ashraf, M. and O'Leary, J.W. (1994) Ion distribution in leaves of varying ages in salt-tolerant lines of alfalfa under salt stress. *J. Plant Nutr.* 17: 1463–1476.
- Babu, M.A., Singh, D. and Gothandam, K.M. (2012) The effect of salinity on growth, hormones and mineral elements in leaf and fruit of tomato cultivar PKM1. *J. Anim. Plant Sci.* 22: 159–164.
- Bariola, P.A., Retelska, D., Stasiak, A., Kammerer, R.A., Fleming, A., Hijri, M. et al. (2004) Remorins form a novel family of coiled coil-forming oligomeric and filamentous proteins associated with apical, vascular and embryonic tissues in plants. *Plant Mol. Biol.* 55: 579–594.
- Bonawitz, N.D., Soltau, W.L., Blatchley, M.R., Powers, B.L., Hurlock, A.K., Seals, L.A. et al. (2012) REF4 and RFR1, subunits of the transcriptional coregulatory complex mediator, are required for phenylpropanoid homeostasis in *Arabidopsis*. *J. Biol. Chem.* 287: 5434–5445.
- Brands, A. and Ho, T.H. (2002) Function of a plant stress-induced gene, HVA22. Synthetic enhancement screen with its yeast homolog reveals its role in vesicular traffic. *Plant Physiol.* 130: 1121–1131.
- Burnette, R.N., Gunesequera, B.M. and Gillaspay, G.E. (2003) An *Arabidopsis* inositol 5-phosphatase gain-of-function alters abscisic acid signaling. *Plant Physiol.* 132: 1011–1019.
- Chen, T., Yang, Q., Zhang, X., Ding, W. and Gruber, M. (2012) An alfalfa (*Medicago sativa* L.) ethylene response factor gene, MsERF11, enhances salt tolerance in transgenic *Arabidopsis*. *Plant Cell Rep.* 31: 1737–1746.
- Cominelli, E. and Tonelli, C. (2009) A new role for plant R2R3-MYB transcription factors in cell cycle regulation. *Cell Res.* 19: 1231–1232.
- Conesa, A., Götz, S., García-Gómez, J.M., Terol, J., Talón, M. and Robles, M. (2005) Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 18: 3674–3676.
- Dixon, R.A., Achnine, L., Kota, P., Liu, C.J., Reddy, M.S. and Wang, L. (2002) The phenylpropanoid pathway and plant defence—a genomics perspective. *Mol. Plant Pathol.* 3: 371–390.
- Djilianov, D., Prinsen, E., Oden, S., Onckelen, H. and Müller, J. (2003) Nodulation under salt stress of alfalfa lines obtained after in vitro selection for osmotic tolerance. *Plant Sci.* 165: 887–894.

- Dluzniewska, P., Gessler, A., Dietrich, H., Schnitzler, J.P., Teuber, M. and Rennenberg, H. (2007) Nitrogen uptake and metabolism in *Populus × canescens* as affected by salinity. *New Phytol.* 173: 279–293.
- Dobrenz, A.K., Robinson, D.L., Smith, S.E. and Poteet, D.C. (1989) Registration of AZ-GERM SALT-II nondormant alfalfa germplasm. *Crop Sci.* 29: 493.
- Du, Z., Zhou, X., Ling, Y., Zhang, Z. and Su, Z. (2010) AgriGO: a GO analysis toolkit for the agricultural community. *Nucleic Acids Res.* 38: W64–W70.
- Dubos, C., Stracke, R., Grotewold, E., Weisshaar, B., Martin, C. and Lepiniec, L. (2010) MYB transcription factors in *Arabidopsis*. *Trends Plant Sci.* 15: 573–581.
- Grossmann, S., Bauer, S., Robinson, P.N. and Vingron, M. (2007) Improved detection of overrepresentation of Gene-Ontology annotations with parent child analysis. *Bioinformatics* 23: 3024–3031.
- Grover, A. (2012) Plant chitinases: genetic diversity and physiological roles. *Crit. Rev. Plant Sci.* 31: 57–73.
- Hua, Y., Zhang, B.X., Cai, H., Li, Y., Bai, X., Ji, W. et al. (2012) Stress-inducible expression of GsSAMS2 enhances salt tolerance in transgenic *Medicago sativa*. *Afr. J. Biotechnol.* 11: 4030–4038.
- Ji, Y., Shi, Y., Ding, G. and Li, Y. (2011) A new strategy for better genome assembly from very short reads. *BMC Bioinformatics* 12: 493.
- Jin, H., Sun, Y., Yang, Q., Chao, Y., Kang, J., Jin, H. et al. (2010) Screening of genes induced by salt stress from alfalfa. *Mol. Biol. Rep.* 37: 745–753.
- Langmead, B., Trapnell, C., Pop, M. and Salzberg, S.L. (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 10: R25.
- Läuchli, A. and Grattan, S.R. (2007) Plant growth and development under salinity stress. In *Advances in Molecular Breeding Toward Drought and Salt Tolerant Crops*. Edited by Jenks, M.A., Hasegawa, P.M. and Jain, S.M. pp. 1–32. Springer, Berlin.
- Li, D., Su, Z., Dong, J. and Wang, T. (2009) An expression database for roots of the model legume *Medicago truncatula* under salt stress. *BMC Genomics* 10: 517.
- Li, J., Dai, X., Liu, T. and Zhao, P.X. (2012) LegumelP: an integrative database for comparative genomics and transcriptomics of model legumes. *Nucleic Acids Res.* 40: D1221–D1229.
- Martin, J.A. and Wang, Z. (2011) Next-generation transcriptome assembly. *Nat. Rev. Genet.* 12: 671–682.
- Mizuno, H., Kawahara, Y., Sakai, H., Kanamori, H., Wakimoto, H., Yamagata, H. et al. (2010) Massive parallel sequencing of mRNA in identification of unannotated salinity stress-inducible transcripts in rice (*Oryza sativa* L.). *BMC Genomics* 11: 683.
- Monirifar, H. and Barghi, M. (2009) Identification and selection for salt tolerance in alfalfa (*Medicago sativa* L.) ecotypes via physiological traits. *Not. Sci. Biol.* 1: 63–66.
- Munnik, T. and Vermeer, J.E. (2010) Osmotic stress-induced phosphoinositide and inositol phosphate signaling in plants. *Plant Cell Environ.* 33: 655–669.
- Munns, R. and Tester, M. (2008) Mechanisms of salinity tolerance. *Annu. Rev. Plant Biol.* 59: 651–681.
- Murashige, T. and Skoog, F. (1962) A revised medium for rapid growth and bioassays with tobacco cultures. *Physiol. Plant.* 15: 473–479.
- Oliveros, J.C. (2007) VENNY. An interactive tool for comparing lists with Venn diagrams. <http://bioinfoqg.cnb.csic.es/tools/venny/index.html>.
- Ottman, M. (1999) Water stress and alfalfa production. In *Handbook of Plant and Crop Stress*, 2nd edn. Edited by Pessaraki, M. pp. 879–889. CRC Press, Boca Raton, FL.
- Peel, M.D., Waldron, B.L., Jensen, K.B., Chatterton, N.J., Horton, H. and Dudley, L.M. (2004) Screening for salinity tolerance in alfalfa: a repeatable method. *Crop Sci.* 44: 2049–2053.
- Pélissier, H.C., Peters, W.S., Collier, R., van Bel, A.J. and Knoblauch, M. (2008) GFP tagging of sieve element occlusion (SEO) proteins results in green fluorescent forisomes. *Plant Cell Physiol.* 49: 1699–1710.
- Pysh, L.D., Wysocka-Diller, J.W., Camilleri, C., Bouchez, D. and Benfey, P.N. (1999) The GRAS gene family in *Arabidopsis*: sequence characterization and basic expression analysis of the SCARECROW-LIKE genes. *Plant J.* 18: 111–119.
- Qin, F., Shinozaki, K. and Yamaguchi-Shinozaki, K. (2011) Achievements and challenges in understanding plant abiotic stress responses and tolerance. *Plant Cell Physiol.* 52: 1569–1582.
- Ren, Z., Zheng, Z., Chinnusamy, V., Zhu, J., Cui, X., Iida, K. et al. (2010) RAS1, a quantitative trait locus for salt tolerance and ABA sensitivity in *Arabidopsis*. *Proc. Natl Acad. Sci. USA* 107: 5669–5674.
- Riechmann, J.L. and Meyerowitz, E.M. (1998) The AP2/EREBP family of plant transcription factors. *Biol. Chem.* 379: 633–646.
- Rumbaugh, M.D. (1991) Salt tolerance of germinating alfalfa seeds. In *North American Alfalfa Improvement Conference. Standard tests to characterize alfalfa cultivars*. <http://www.naaic.org/stdtests/saltseeds.htm>.
- Rumbaugh, M. and Pendery, B. (1990) Germination salt resistance of alfalfa (*Medicago sativa* L.) germplasm in relation to subspecies and centers of diversity. *Plant Soil* 124: 47–51.
- Sakai, H., Honma, T., Aoyama, T., Sato, S., Kato, T., Tabata, S. et al. (2001) ARR1, a transcription factor for genes immediately responsive to cytokinins. *Science* 294: 1519–1521.
- Scasta, J.D., Trostle, C.L. and Foster, M.A. (2012) Evaluating alfalfa (*Medicago sativa* L.) cultivars for salt tolerance using laboratory, greenhouse and field methods. *J. Agric. Sci.* 4: 90–103.
- Smethurst, C.F., Rix, K., Garnett, T., Auricht, G., Bayart, A., Lane, P. et al. (2008) Multiple traits associated with salt tolerance in lucerne: revealing the underlying cellular mechanisms. *Funct. Plant. Biol.* 35: 640–650.
- Smith, S.E. (1991) Forage production under salt stress. Standard tests to characterize alfalfa cultivars. 3rd edn. <http://www.naaic.org/resource/stdtests.php>.
- Smith, S.E. and Fairbanks, D.J. (1989) Registration of AZ-88MS and AZ-88NDC, nondormant male sterile and male fertile maintainer alfalfa germplasms. *Crop Sci.* 29: 1095.
- Soltani, A., Khodarahmpour, Z., Jafari, A.A. and Nakhjavan, S. (2012) Selection of alfalfa (*Medicago sativa* L.) cultivars for salt stress tolerance using germination indices. *Afr. J. Biotechnol.* 11: 7899–7905.
- Soto, G., Stritzler, M., Lisi, C., Alleva, K., Pagano, M.E., Ardila, F. et al. (2011) Acetoacetyl-CoA thiolase regulates the mevalonate pathway during abiotic stress adaptation. *J. Exp. Bot.* 62: 5699–5711.
- Stevenson, J.M., Perera, I.Y., Heilmann, I., Persson, S. and Boss, W.F. (2000) Inositol signaling and plant growth. *Trends Plant Sci.* 5: 252–258.
- Sun, S.J., Guo, S.Q., Yang, X., Bao, Y.M., Tang, H.J., Sun, H. et al. (2010) Functional analysis of a novel Cys2/His2-type zinc finger protein involved in salt tolerance in rice. *J. Exp. Bot.* 61: 2807–2818.
- Taylor, C.W. and Thorn, P. (2001) Calcium signalling: IP3 rises again . . . and again. *Curr. Biol.* 11: R352–R355.

- Thimm, O., Blasing, O., Gibon, Y., Nagel, A., Meyer, S., Krüger, P. et al. (2004) MAPMAN: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. *Plant J.* 37: 914–939.
- Uchimiya, H., Fujii, S., Huang, J., Fushimi, T., Nishioka, M., Kim, K.M. et al. (2002) Transgenic rice plants conferring increased tolerance to rice blast and multiple environmental stresses. *Mol. Breed.* 9: 25–31.
- Van Loon, L.C., Rep, M. and Pieterse, C.M. (2006) Significance of inducible defense-related proteins in infected plants. *Annu. Rev. Phytopathol.* 44: 135–162.
- Wang, W.B., Kim, Y.H., Lee, H.S., Kim, K.Y., Deng, X.P. and Kwak, S.S. (2009) Analysis of antioxidant enzyme activity during germination of alfalfa under salt and drought stresses. *Plant Physiol. Biochem.* 47: 570–577.
- Winicov, I. (2000) Alfin1 transcription factor overexpression enhances plant root growth under normal and saline conditions and improves salt tolerance in alfalfa. *Planta* 210: 416–422.
- Winicov, I. and Bastola, D.R. (1999) Transgenic overexpression of the transcription factor Alfin1 enhances expression of the endogenous MsPRP2 gene in alfalfa and improves salinity tolerance of the plants. *Plant Physiol.* 120: 473–480.
- Yang, S., Gao, M., Xu, C., Gao, J., Deshpande, S., Lin, S. et al. (2008) Alfalfa benefits from *Medicago truncatula*: the RCT1 gene from *M. truncatula* confers broad-spectrum resistance to anthracnose in alfalfa. *Proc. Natl Acad. Sci. USA* 105: 12164–12169.
- Yang, S.S., Tu, Z.J., Cheung, F., Xu, W.W., Lamb, J.F., Jung, H.J. et al. (2011) Using RNA-Seq for gene identification, polymorphism detection and transcript profiling in two alfalfa genotypes with divergent cell wall composition in stems. *BMC Genomics* 12: 199.
- Young, N.D., Debelle, F., Oldroyd, G.E., Geurts, R., Cannon, S.B., Udvardi, M.K. et al. (2011) The *Medicago* genome provides insight into the evolution of rhizobial symbioses. *Nature* 480: 520–524.
- Zahaf, O., Blanchet, S., de Zélicourt, A., Alunni, B., Plet, J., Laffont, C. et al. (2012) Comparative transcriptomic analysis of salt adaptation in roots of contrasting *Medicago truncatula* genotypes. *Mol. Plant* 1–14.
- Zhang, H., Han, B., Wang, T., Chen, S., Li, H., Zhang, Y. et al. (2012) Mechanisms of plant salt response: insights from proteomics. *J. Proteome Res.* 11: 49–67.
- Zhu, J.K. (2000) Genetic analysis of plant salt tolerance using *Arabidopsis*. *Plant Physiol.* 124: 941–948.