American Journal of Molecular Biology, 2011, 1, 156-173 **AJMB**
doi:10.4236/ajmb.2011.13016 Published Online October 2011 (http://www.SciRP.org/journal/ajmb/).

Gene expression profiling in soybean under aluminum stress: genes differentially expressed between Al-tolerant and Al-sensitive genotypes

Dechassa Duressa[1,2], Khairy M. Soliman[1], Robert W. Taylor[1], Dongquan Chen[3]

[1]Department of Natural Resources and Environmental Sciences, Alabama A&M University, Normal, Alabama, USA;
[2]USDA-ARS, Salinas, California, USA;
[3]Biostatistics and Bioinformatics Unit, Comprehensive Cancer Center, Division of Preventive Medicine, Department of Medicine, University of Alabama at Birmingham, Birmingham, USA.

Email: [1]khairy.soliman@aamu.edu

Received 20 July 2011; revised 29 July 2011; accepted 16 August 2011.

ABSTRACT

It is well documented that aluminum (Al) toxicity is the most important constraint to crop production on acid soils and soybean is one of the most Al sensitive plant species. To advance our understanding of the molecular and genetic mechanisms of Al-tolerance in soybean we compared root tip (1 cm long) transcriptome profiles of an Al-tolerant (PI 416937) and Al-sensitive (Young) soybean genotypes using a combination of DNA microarrays and quantitative real-time PCR gene expression profiling technologies, in a time-course experiment (2, 12, 48, 72 h post Al treatment). We observed many genes differentially expressed between the two genotypes in constitutive and non-constitutive manner. The most likely candidate Al-tolerance genes expressed at high level in PI 416937 relative to Young include the previously reported transcription factors, auxin down regulated-like protein (*ADR*6-like) and, basic leucine zipper (*bZIP* 94), sulfur transmembrane transport protein and lipid transfer protein; and several novel genes that include rare cold inducible protein (*RCI*2*B*), *GPI*-transamidase, malonyl-*COA:* Isoflavone 7-O-glucoside-6″-O-malontransferase, a cell proliferation protein (*WPP*2), oleosin protein, pectinestrease inhibitor, and impaired sucrose induction1; whereas genes negatively correlated with Al-tolerance, namely cellulose synthase and calcium transporters were down regulated in Al-tolerant PI 416937 compared to the Al-sensitive Young. The possible mechanisms of how these genes contribute to Al-tolerance trait are discussed. In conclusion, transcriptome profile comparisons of Al-tolerant and Al-sensitive soybean genotypes revealed novel putative Al-tolerance genes. These genes deserve further functional characterization for eventual utilization in developing soybean germplasm adapted to high aluminum soils.

Keywords: Soybean, Al Tolerance, Gene Expression, Microarray

1. INTRODUCTION

Aluminum (Al) toxicity is a major constraint to crop production on acid soils. In view of the fact that 40% of the world's arable land is acidic [1], Al toxicity remains as a major hurdle for increasing world food production, especially in developing tropical and subtropical regions where increase in food production is much needed. Aluminum reduces crop yield through root growth inhibition and impairment in nutrient and water uptake. Plants tolerate aluminum via several mechanisms that include, 1) exclusion mechanism that involve chelation of aluminum with root secreted organic ligands-mainly citrate, malate and oxalate in the rhizosphere, 2) by possessing low cell wall polysaccharides and, 3) by internal detoxification that involves complexation of Al in the symplast with organic ligands and subsequent sequestration in the vacuole [2-6]. The level of Al tolerance varies from species to species and among genotypes within species. Although physiologically a simple trait, the molecular mechanism of Al tolerance and toxicity largely remains elusive. Two Al tolerance genes Aluminum Induced Malate Transporter 1 (*ALMT*1) in Arabidopsis [7] and Aluminum Induced Multidrug Exporter (*ALMATE*) in sorghum [8] have so far been cloned.

Soybean is one of the most Al sensitive plant species with root growth inhibition of up to 50% in sensitive cultivars with Al^{3+} activity of only 1.5 μM in culture

solution [9]. Genetic [10] and physiological analyses [11] show Al tolerance in soybean is a complex trait. In a population derived from a cross between the Al tolerant genotype PI 416937 and the Al sensitive cultivar Young quantitative trait loci mapping revealed five DNA markers associated with Al tolerance [10]. Using differential display gene expression profiling [12] identified three genes associated with soybean Al tolerance namely phosphenopyruvate, carboxylase (*PEPC*), homologous of translationally controlled tumor proteins (*TCTP*) and inosine 5'-monophosphate dehydrogenases (*IMPDH*). Using subtractive hybridization approach [13] identified two putative soybean Al tolerance genes, namely soybean aluminum induced 3-2 (*Sali*3-2*)* and soybean aluminum induced 5-4a (*Sali*5-4*a).* The two research groups [12,13] used a pair of Al tolerant and sensitive genotypes to identify genes specifically regulated by Al in the tolerant types, however, the methods used were not sensitive enough to detected cascade of genes and pathways involved in soybean Al tolerance and toxicity. The aim of the present work was to identify candidate soybean Al-tolerance genes using the Al-tolerant soybean PI 416937 and the Al-sensitive Young employing the technique of DNA microarrays to understand specific gene function related to Al tolerance trait. Such approach was recently used in wheat [14,15], maize [16], *Medicago truncatula* [17,18] and *Arabidopsis* [19] to discern the molecular basis of Al tolerance in the respective species.

2. MATERIALS AND METHODS

2.1. Plant Genotype and Growth Conditions

Soybean plant introduction (PI 416937) is well characterized for its Al-tolerance, including exclusion of Al from entering root tip [11,47]. Al -sensitive soybean genotype Young served as a control. Seeds were surface sterilized with 20% clorox (v/v) for 12 minutes, rinsed with distilled-deionized water several times, and were germinated in deionized water moistened germination paper at 25˚C in an incubator for 72 hours. Seedlings with uniform tap root length from the germinated seeds were transferred to black painted pots filled with approximately 4 L 800 µM $CaCl_2$ solution with 0 or 10 µM Al in Conviron growth chamber (16/8 hour light/ dark cycle, temp. 28˚C /20˚C, photosynthetic photon density of 100 µmol m^{-2} s^{-1}). The pH of the culture solution was adjusted to 4.3. After 2, 12, 48 or 72 h of exposure to Al treatment, 0-1 cm section of the primary root tips of approximately 15 plants per pot were harvested and immediately flash-frozen in liquid nitrogen, and stored at –70˚C for RNA extraction. Three independent replicates were used per treatment.

2.2. RNA Extraction, Microarray Procedure and Data Analysis

Total RNA was extracted from 100 mg root tissue samples using a Qiagen RNeasy plant RNA isolation kit following manufacturer's protocol (Qiagen, Inc.). The Affymetrix GeneChip Soybean Genome Array with over 68,000 probe sets *Glycine Max L*. and wild soybean combined was used for microarray analysis of the soybean genome for Al tolerance. Three chips were used per treatment. Detailed procedures for RNA labeling and array analysis are described in the Affymetrix GeneChip Expression Technical Manual. Briefly, the quality of total RNA was determined using RNA 6000 Nano chip on the Agilent BioAnalyzer 2100 prior to double–stranded cDNA synthesis. Total RNA in the amount of 2 µg was used for double-stranded cDNA generation by linear amplification using oligo dT-T7 primer and reverse transcriptase (RT). Subsequently, biotin-labeled cRNA was synthesized by *in vitro* transcription (IVT) using ENZO High Yield IVT kit (ENZO). Quality and quantity of cRNA was assessed using RNA 6000 Nano chip on Agilent BioAnalyzer 2100. Fifteen microgram cRNA was used for hybridization. Arrays were hybridized overnight at 45˚C for 16 hours in a GeneArray Hybridization Oven 640 (Affymetrix). The next day arrays were washed and stained in the Fluidics Station 450 (Affymetrix) and scanned by the High Resolution Gene Chip Scanner 3000 (Affymetrix).

Gene expression levels were determined using Gene Chip Operating Software (GCOS 1.1, Affymetrix). The expression levels were subjected to data query and data mining in Data Mining Tool (DMT). Statistical analysis of the data was conducted using the software packages Array Assist Enterprise together with Pathway Architect (Stratagene/Agilent, Santa Clara, CA). Briefly, the raw Gene Chip files (Cel and CHP) from Gene Chip Operating Software (GCOS, Affymetrix, CA) were uploaded, background-subtracted, variance stabilized, and normalized with the GC-RMA method [48]. Gene expression level of the control treatment was used as a baseline to calculate the intensity ratio/fold changes (FC). The ratio was log2-transformed before further statistical analysis. The p-values were obtained by an unpaired t-test assuming unequal variance. Significantly up-regulated and down-regulated genes were annotated using non-redundant protein databases accessed by BLASTX at National Center for Biotechnology Information (NCBI).

2.3. Quantitative Real-Time PCR

Quantitative real-time PCR analysis was employed to validate the relative change in expression of genes for selected gene panels from microarray experiments using the Roche Diagontics light Cycler® 480 System with

SYBR green detection format (Roche Diagnostic, Corp). RNA extraction and quality test was as described above. Prior to cDNA synthesis, RNA samples were treated with Applied Biosystems Turbo DNAse-free™ DNase (Ambion, Inc.,) to remove DNA contamination. Briefly, 2 µl 10x DNase I buffer and 1 µl rDNase I were added to 20 µl RNA sample and the mix was incubated at 37˚C for 30 min in water bath. Afterwards, 2 µl re-suspended DNase inactivation reagent was added and the samples mixed well and incubated at room temperature for 3 min. Samples were then centrifuged at 10,000 x g for 1.5 min (Eppendorf centrifuge 5415 D) in 1.6 ml centrifuge tubes and the supernatants were transferred to fresh tubes.

cDNA was synthesized from DNase treated RNA samples using a Roche Diagnostics Transcriptor First Strand cDNA Synthesis Kit (Roche Diagnostics, Corp) was as follows. Prior to the procedure, reagents and samples were thawed on ice and reagents were briefly centrifuged at 10,000 rpm. A cDNA synthesis reaction was set-up in 0.5 ml PCR tubes with 1 µg total RNA, 1 µl of Oligo (dT) primer, 4 µl reaction buffer, 0.5 µl RNase inhibitor, 2 µl deoxyribonucleotides (1 mM each), and 0.5 µl reverse transcriptase added in that order. Samples were brought to 20 µl total reaction volume with PCR grade water and centrifuged for 1 min @ 340 rpm (Ep-pendorf centrifuge 5804R), incubated at 55˚C for 30 min. Reverse transcriptase was inactivated by an additional incubation for 5 min at 85°C, all in a DNA Engine (PTC-200) thermocycler (MJ-research). The cDNA concentration and quality was determined using Nano-Drop Spectrophotometer brand ND-1000 (Nano-Drop Technologies, Inc. A total reaction volume of 11 µl comprising 2 µl cDNA sample, 2 µl each of the reverse and forward primers at 0.2 µM concentration and 5 µl SYBR mix was prepared in 96-well plates (Roche Diagnostics) in two biological and three technical replicates for each gene. The real-time PCR profile was 95˚C for 5 min, 45 cycles of amplification at 95˚C for 10 second, 55˚C for 20 second and 72˚C for 20 second, and melting 95˚C for 1 min, 65˚C for 1 min and 95˚C continuous, and cooling at 40˚C for 30 second. Negative controls, in which cDNA sample was replaced with PCR grade water for each primer pairs, were included in each run. Sample wells were individually assessed for data quality and PCR product specificity was verified by melting curve analysis. Expression level of target genes was normalized using in-run beta tubulin gene as internal control and transcript concentration ratios were calculated using the ΔΔCT –Method [49]. Fold change was calculated as treatment to control ratio and correlated with results from microarray.

3. RESULTS AND DISCUSSION

3.1. Principal Component Analysis (PCA), (A)

Principal component analysis was conducted to evaluate whether or not the two genotypes differ in gene expression pattern. The results clearly demonstrated that the two genotypes had a distinct transcriptome profile (**Figure 1**). An assessment of genes differentially expressed between the genotypes revealed that the majority of the genes are constitutively expressed at higher or lower levels in PI 416937 compared to Young (**Figure 4**), a reflection of genetic background difference. By fitting an ANOVA model to gene expression data in a maize Al study [16] observed large genotypic effect and small treatment effect, consistent with our finding.

3.2. Overall Assessment of Gene Expression Changes (B)

The number of genes differentially expressed between PI 416937 and Young increased with treatment time peaking at 72 h post treatment (**Figures 2**, **3**). The ratio of up-regulated to down-regulated genes was 52/20, 71/41, 61/56 and 137/786 for 2, 12, 48, and 72 h, respectively. These results also show that more genes were up-regulated in Young as exposure time to Al treatment increased indicating that the sensitive genotypes responds in a nonspecific manner the longer the treatment duration, likely, as a syndrome of Al toxicity rather than a mechanism of tolerance. Similar results were reported by [16] in maize and [18] in *Medicago truncatula.*

To determine the type of induction of genes differentially expressed, the two genotypes were further compared with or without aluminum treatment. The results demonstrated that the majority of the genes are either constitutively up- or down-regulated in PI 416937 in comparison to Young (**Figure 4**), suggesting that perhaps the constitutive nature of the Al tolerance mechanism is genetically inherited, that is, an Al tolerant genotype has genes that are expressed at high level compared to the sensitive genotype in the absence or presence of Al. This conclusion would not have been possible without making control to control comparison of tolerant vs. sensitive genotype pairs as was done here. These results are corroborated by findings of [7] in wheat, [8] in sorghum, and [20] in barley; that the malate and citrate transporters Al tolerance genes cloned so far are constitutively expressed at high level in tolerant compared to sensitive genotypes. Classification of the differentially expressed genes based on cellular function shows that 12% is related to stress response, 9% to transport, 5% to signaling, 5% to cell structure, and 3% to transcription factors. The remaining 42% is comprised of genes of unknown function (**Figure 5**).

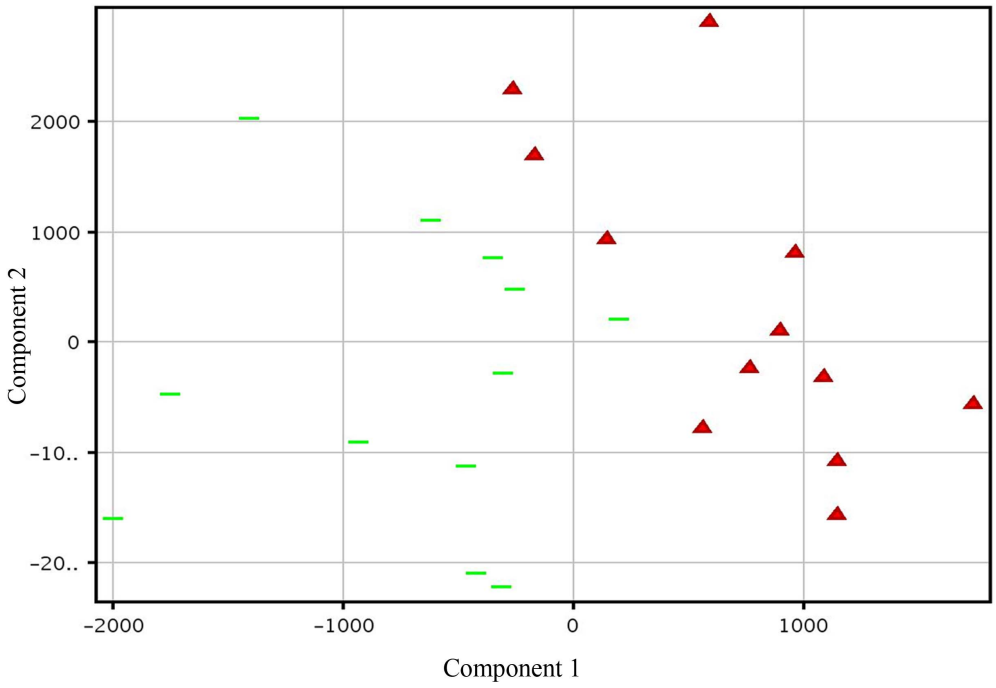


Figure 1. Principal Component Analysis (PCA) of genotypic gene expression profiles based on 24 chips with or without aluminum combined. Red triangles (PI 416937) green bars (Young). A clear separation of the gene expression profiles of the genotypes is shown.

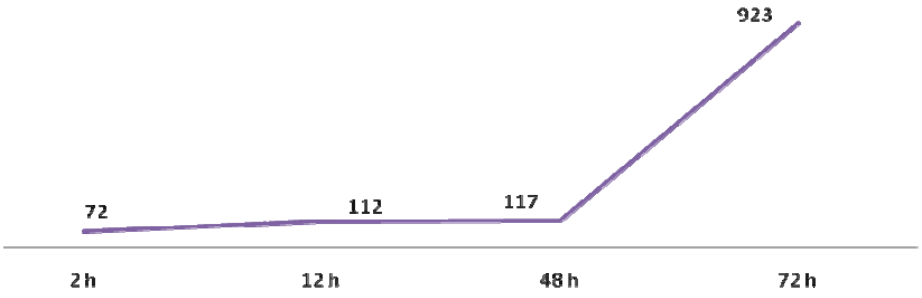
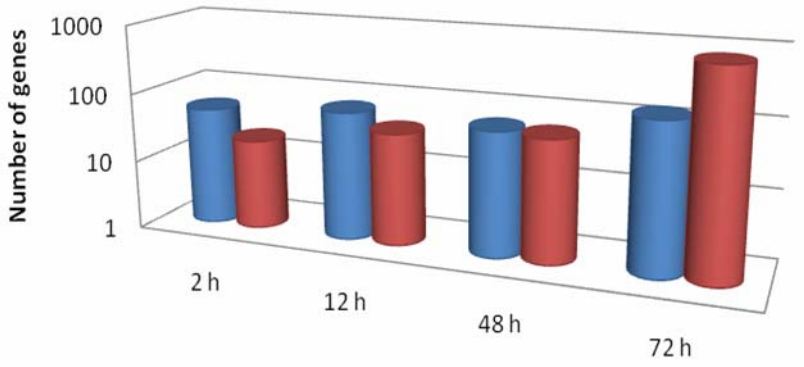


Figure 2. Number of differentially expressed genes as a function of aluminum treatment time (h) in PI 416937 vs. Young comparisons.

3.3. Differentially Expressed Genes by Functional Category (D)

3.3.1. Genes Related to Transcription Factors

Classification of genes differentially expressed between PI 416937 and Young based on cellular function showed that the number of transcription factors detected was small, constituting only 3% of the total expressed genes (**Figure 5**). At 2 h post treatment the homeobox transcription factor (Gma. 24251) was up-regulated by ≈ 13 fold (**Table 1**). In a similar study [23] identified the homeobox domain transcription factor as the regulator of cold tolerance in *Arabidopsis* in addition to its regulatory effect on plant growth and development. At 12 h post Al treatment two transcription factors *WRKY70* (Gma.4281) and basic-helix-loop helix (Gma.16666) were down-regulated (Additional file 1). The *ADR6*-like transcription factor (Gma.28057) previously reported by [13] was up-regulated 5.18 fold at 48 h post treatment (Additional file 2). A basic leucine zipper (*bZip94*) transcription factor (Gma.17306) was up-regulated 5.18 and 6.45 fold at 48 h and 72 h, respectively, in a constitutive manner. In the legume family, *bZip* transcription factors are reported to regulate drought stress response and seed development in *Phaseolus* species and pathogen defense response and plant development in soybean [24].



	2 h	12 h	48 h	72 h
up	52	71	61	137
down	20	41	56	786

Figure 3. Number of up-regulated and down-regulated genes at each time point in PI 416937 vs. Young comparison (P < 0.01; FC>3).

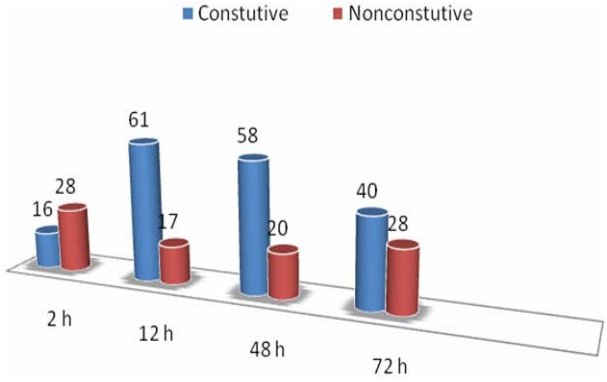


Figure 4. Number of constitutive and non constitutive genes differentially expressed between PI 416937 and Young in time-course experiment.

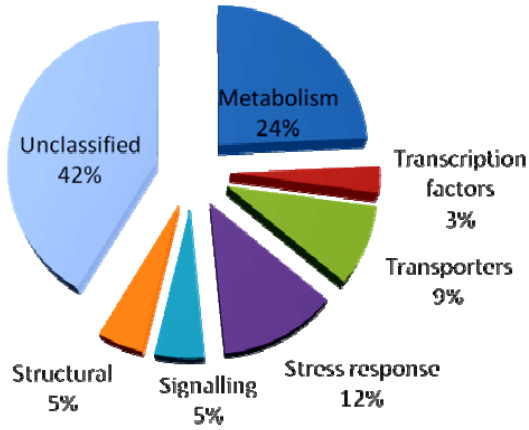


Figure 5. Functional category of genes differentially expressed between PI 416937 and Young.

3.3.2. Genes Related to Transporters

Regulated influx and efflux of substances across biological membranes is a vital component of cellular stress responses. Several classes of transporters, symporters and antiporters, lipid transfer proteins, carbohydrate transporters, sulfate trans-membrane transporters, and (*ATPase*) inward rectifier potassium channel were differentially expressed between PI 416937 and Young. At 2 h, k+–*ATPase* inward rectifier of potassium channel was up-regulated 8.54 fold (**Table 1**). Potassium plays important roles in biological systems ranging from maintenance of membrane potential, electrical neutralization of ionic groups, osmoregulation, and control of cell membrane polarization, ion homeostasis, enzyme activation, signal transduction and many other physiological functions [25]. These results are consistent with the finding that potassium channel has been shown to be up-regulated [25] during early stages of plant exposure to cold, drought, salt stress and abscisic acid. Sec 14 cytosolic factor family protein-a lipid transport protein (Gma.34414) was consistently up-regulated at 12, 48, and 72 h with a fold change of over 11 times in a constitutive manner (Additional files 1, 2 & 3). Furthermore, another lipid transfer protein (Gma.3880) was up-regulated by 9.15, 6.79, and 13.89 fold at 12, 48 and 72 h respectively. Lipid transport proteins facilitate transport of lipids to cell walls for biosynthesis of cutin layers and surface waxes as a defense mechanism in response to pathogen attack [26]. They are also induced by abiotic stresses including aluminum stress in wheat roots [14]. A recent report by [27] suggests that lipid transport proteins loosen cell wall in a nonhydrolytic mode and enhance cell elongation, the role traditionally attributed to expansin. Aluminum stress inhibits root growth by restricting cell wall expansion. The higher expression level

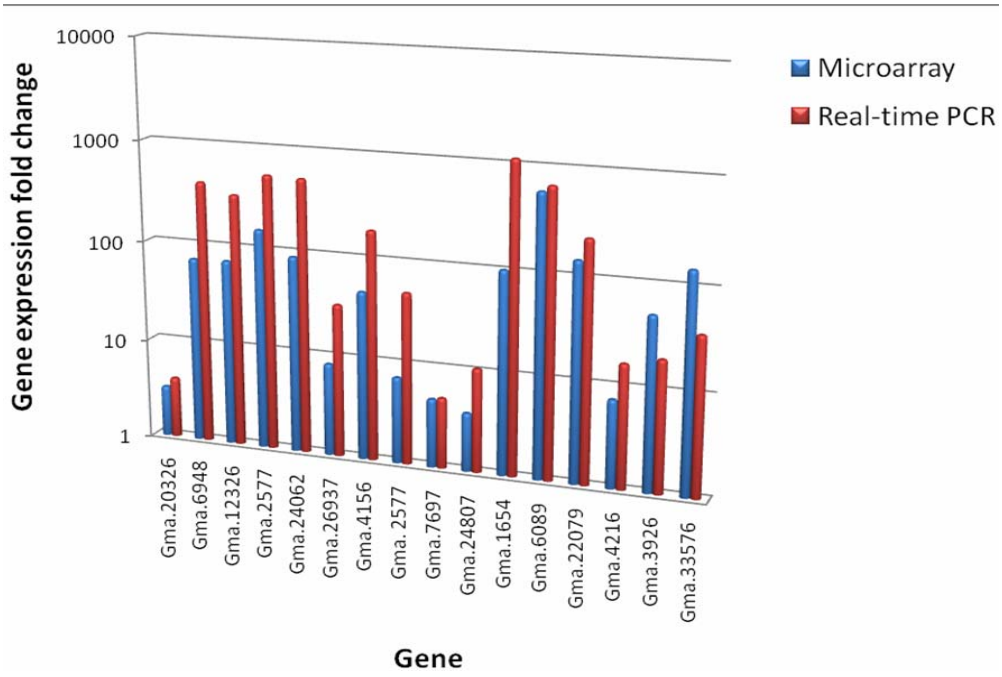


Figure 6. Comparison of microarray and quantitative real-time gene expression in fold change for selected gene panels.

of lipid transfer proteins in Al tolerant soybean suggests a role for these proteins in soybean Al tolerance mechanisms. Sulfate trans-membrane proteins Gma.37580 (12 h, Additional file 1) and Gma. 172 8097 (72 h, Additional file 3) were constitutively expressed more than 4 fold in PI 416937 compared to Young. Sulfate is a negatively charged ion and could possibly be exuded by roots as Al chelator similar to organic acid anions, phenolics and phosphate [1]. Two Calcium antiporters genes Gma 42414 (12 h, Additional file 1) and Gma.468 (48 h, Additional file 2) were constitutively expressed at approximately 4 fold lower in PI 416937 than in Young. Disruption of cellular calcium homeostasis by calcium influx to cytoplasm has been suggested as the primary trigger of the Al toxicity syndrome [28]. Aluminum increases cystosolic calcium activity that leads to disruption in physiological and biochemical processes and ultimately reduced growth. The expression of calcium antiporters/secondary calcium transporters at low level in Al tolerant PI 416937 suggests maintenance of calcium homeostasis under Al stress in this genotype. Plant sugar transporters are induced by pathogen attack to meet the energy demand imposed on cells under stress [29]. Similarly, Al stress has been shown to cause mitochondrial dysfunction and *ATP* depletion [30]. Several investtigators [19] identified multiple Al induced sugar transporters in Arabidopsis. In this study, one sugar: hydrogen antiporter gene Gma.17205 was constitutively expressed at 4 fold higher in PI 416937, 12 and 48 h post treatment (Additional files 1 & 2).

3.3.3. Genes Related to Stress Response

It has been known that plants have developed defense responses to biotic and abiotic stresses over evolutionary time as a survival mechanism. Regulation of gene expression in response to a stress factor is a key component of such response at the molecular level. Aluminum toxicity triggers perturbation in cellular processes and in response induces change in gene expression in both tolerance related genes-mechanism of Al tolerance; and stress associated genes-a manifestation of Al toxicity. As one criterion for distinguishing between genes expressed as manifestation of Al toxicity and genes associated with Al tolerance, genes differentially expressed between an Al tolerant and sensitive genotypes were considered in discussion of this paper. Several stress related genes showed differences in expression level in Al tolerant PI 416937 compared to Al sensitive Young. Several pathogenesis related proteins, oxidative stress related proteins, chaperonin proteins, heat shock proteins, desiccation related proteins, and cold stress related proteins were differentially expressed between the two genotypes. A rare cold inducible gene-RCI2B (Gma.27795) was constitutively expressed at 223.73, 257.05, and 253.35 fold higher in PI 416937 at 2, 48, and 72 h, respectively (**Table 1**, Additional files 2&3). This gene was not previously reported in plant Al literature and is probably a key novel Al tolerance gene in soybean, and thus would be an interesting candidate gene for further characterization. Its molecular mechanism of action is not yet known. In

Table 1. Genes significantly (p<0.01, FC >4) expressed at higher or lower level in PI 416937 compared to Young 2 h post aluminum treatment.

Unigene ID	Fold Change	Functional Category	Annotationb
Gma.27795c	223.73(up)	stress response	RCI2B/response to cold
Gma.17961	29.73(up)	structural	glycine rich protein/oleosin protein, lipid sequestration 1e-20
Gma.15007	23.55(up)	metabolism	ferredoxin2 electron carrier/ Photosynthetic electron transfer chain
Gma.17744c	21.99(down)	stress response	gluthatione-s-transferase29
Gma.6487	18.31(up)	unclassified	unknown
Gma.7535c	17.96(up)	stress response	malonyl-COA: isoflavone 7-O-glucoside-6"-O-malontransferase
Gma.21391[c]	12.91(up)	unclassified	unknown
Gma.5672	13.97(up)	metabolism	asparagines synthase 2
Gma.24251	12.89(up)	transcription factor	homebox7 transcription factor
Gma.23189	12.60(up)	metabolism	hydrolase/metabolic process
Gma.9092	10.82(up)	unclassified	hypothetical protein
Gma.10987	10.62(up)	metabolism	Rubsico activase
Gma.5716[c]	10.34(up)	unclassified	unknown
Gma.2098	10.19(up)	unclassified	unknown
Gma.22527	9.70(up)	unclassified	unknown
Gma.6245c	9.53(up)	stress response	pathogenesis related thaumatin family protein
Gma.5568c	9.53(up)	stress response	germin like protein/metal ion binding in apoplast
Gma.17291	9.42(up)	unclassified	unknown
Gma.1134	8.54 (up)	transport	K+ ATPase/ inward rectifier potassium channel
Gma.4434[c]	7.91(down)	unclassified	plastocynanin like domain protien/copper ion binding in memberane
Gma.13640	7.89(down)	metabolism	cellulose synthase like-3/cellulose synthase activity
Gma.2166	7.88(up)	cell signaling	protein kinase family protein/Phosphorylation function
Gma.19712	7.8(down)	unclassified	unknown
Gma.5672	7.79(up)	unclassified	unknown
BI316842	7.72(up)	unclassified	unknown
Gma.9683[c]	7.64(up)	unclassified	unknown
Gma.16838	7.25(up)	metabolism	photosystem I subunit H-2
Gma.395[c]	6.42(up)	metabolism	WPP2 domain protein/lateral root development/mitosis
Gma.17162[c]	6.30(up)	unclassified	hypothetical integral membrane protein
Gma.33691[c]	5.80(up)	stress response	quinine reductase family protein/oxidoreductase
Gma.3314[c]	5.53(up)	stress response	cysteine proteinase inhibitor/cysteine type endo-peptidase inhibitor activity
Gma.36991	5.03(up)	metabolism	beta-xylosidase/hydrolase-hydrolysis O-glycosyl cpds
Gma.32786	4.98(up)	metabolism	caleosin-related protein/lipase activity
Gma.27015[c]	4.68(up)	cell signaling	octicosapeptide PB1 domain protein
Gma.5609c	4.66(down)	metabolism	pyrophosphorylase4/pyrophosphatase-metabolic process
Gma.11191[c]	4.66(down)	unclassified	hypothetical protein
Gma.15538	4.39(up)	stress response	glutaredoxin protein/arsenate reductase
Gma.1783	4.28(up)	metabolism	2-oxoglutarate dependent dioxgenase/hydrolase
Gma.10731	4.11(up)	metabolism	photo system II subunit O-2/oxygen evolving
Gma.15	4.02(up)	stress response	LEAS protein/desiccation protective
Gma.681	4.02(up)	unclassified	unknown
BI787321	4.00(up)	unclassified	hypothetical protein
Gma.4613	4.00(up)	structural	rubber elongation factor (REF) protein binds to rubber particles
Gma.47043.99 (up)		unclassified	unknown

[a]significance thresholds (fold change ≥ 4, p ≤ 0.01), [b] e-value (< e-10) = the probability that the match has no biological basis, [c]constitutively expressed, [d]up = up-regulated in PI 416937 relative to Young, down = down-regulated in PI 416937 relative to Young. Note: constitutively expressed means the gene is expressed at higher or lower level in PI 416937 relative to Young with or without aluminum treatment.

Arabidopsis [31] demonstrated that the expression of this gene is negatively regulated by light but induced in etiolated seedlings and roots of adult plants. Also, transgenic plants over-expressing this gene had enhanced tolerance to salt and dehydration. Its higher expression level in a drought and aluminum tolerant soybean line here suggests it might confer multi-stress tolerance. Oxidative stress related proteins including gluthatione-s-trnasferase, germin-like protein, quinine reductase previously reported as Al responsive [15,18,19] showed differential expression between PI 416937 and Young at 2 h (**Table 1**). Gluthathione-s-transferase (*GST*) was down-regulated, whereas, the other two were up-regulated. *GST* is a key oxidative stress responsive gene and its low expression level in PI 416937 indicates that this genotype might not undergo server oxidative stress as the sensitive genotype. This finding is in agreement with the observations of [9,10] that less Al enters PI 416937 roots compared to Young. Over-expression of GST in Arabidopsis has been shown to conferred Al tolerance [32]. However, the current consensus is that oxidative stress gene expression is more of a manifestation of Al toxicity rather than a tolerance mechanism [16, 33]. A gene for isoflavone biosynthesis malonyl-COA: isoflavone 7-O-glucoside-6″-O- malontransefrase (Gma.7535) was up-regulated 17.96 and 54.65 fold at 2, and 72 h respectively. Flavonoids can neutralize Al toxicity by chelating Al ion in vivo or in vitro [1] and/or by acting as antioxidant [18]. Several pathogenesis related proteins, thaumatin protein, cysteine protease inhibitor, xyloglycan specific fungus endoglycananse inhibitor, pathogenesis related protein, disease resistance protein, and chitinase were expressed at higher level in PI 416937 than in Young (**Table 1**, Additional files 1, 2&3). The role of pathogenesis related proteins in Al tolerance is equivocal. In one study [32] over-expressed peroxidase and proteinase inhibitor genes in *Arabidopsis* and found that transgenic plants did not show better Al tolerance level than controls. On the other hand [34], over-expressed pepper basic pathogenesis related protein 1 gene in tobacco and found enhanced tolerance to the heavy metal cadmium and pathogen infection. Desiccation related protein, heat shock protein, and chaperonin protein were also differentially expressed between PI 416937 and Young. Late embroyogenesis abundant (*LEAS*) desiccation tolerance gene (Gma.15) was expressed 4.02 fold higher in P I 416937 than in Young at 2 h (**Table 1**). This is consistent with the fact that PI 416937 is both drought and Al tolerant [10]. *DNAJ* heat shock protein previously reported to be Al induced at protein level in Al tolerant soybean genotype [35] was up-regulated 5.32 and 6.45 fold at 48 h and 72 h , respectively (Additional files 2, 3). A chaperonin protein (Gma.26538) was expressed at 6.67 fold lower in PI 426937 12 h post treatment (Additional file 1).

3.3.4. Genes Related to Metabolism

Plants undergo change in cellular metabolism upon exposure to Al either as a manifestation of Al toxicity or as a mechanism of Al tolerance [15,18]. Aluminum interferes with nutrient uptake, transportation and utilization [36]. Phosphorus deficiency due to Al-phosphate precipitation in rhizosphere, cell wall or in symplast [37] combined with Al induced mitochondrial dysfunction depletes *ATP* and other nucleoside phosphates [30]. Probably as a mechanism for increasing cellular orthophosphate supply under Al stress, a phosphorylase 4 (Gma.5609) gene was expressed at 4.91,6.61, and 5.23 fold higher in PI 416937 than in Young at 246 12, 48 and 72 h respectively (Additional files 1, 2&3). The expression of this gene was 4.66 fold lower in PI 416937 than in Young at 2 h (**Table 1**). A cellulose synthase like-3 gene (Gma.13640) involved in cellulose synthesis was expressed at 7.89 and 6.15 fold lower in PI 416937 than in Young at 2 h and 12 h, respectively (**Table 1** and additional file 1). Cell wall polysaccharides level is inversely related to plant Al tolerance [33,38]. The lower cellulose synthase activity in PI 416937 could be an indication of Al tolerance mechanism. Cell wall metabolism enzymes have been suggested to alleviate Al induced root growth inhibition by remodeling cell wall architecture [3, 18]. Pursuant with this, a gene for beta-xylosidase (Gma. 36991) involved in hydrolysis of O-glycoysl compounds was expressed 5.03 fold higher in PI 416937 at 2 h (**Table 1**). Furthermore, another gene coding for glycoysl hydrolase family3 protein (AW75634) was expressed constitutively at 4.60 fold higher in PI 416937 (Additional file 1). A pectinestrase gene (Gma.22124) was up-regulated at 12 h (Additional file 1). Pectinestrease activity has been shown to correlate with increase in Al sensitivity due to demethylation of pectin by the enzyme and increase in its Al adsorption capacity [3]. It is worth noting that a gene for pectin methylesetrase inhibitor family protein (Gma.31645) was up- regulated at 72 h (Additional file 3) probably deactivating pectinestrease, thus conferring Al tolerance via the exclusion mechanism.

Many eukaryotic proteins are anchored to plasma membrane by glycosyl phosphatidylinositol (*GPI*) [39]. *GPI* transamidase is an endoplasmic reticulum localized protein that transfers performed *GPI* to proteins with *GPI* attachment signal in the carboxyl terminal. Aluminum tolerance genes such as organic acid ion and ATP binding cassette (*ABC*) transporters are membrane proteins. *GPI* transamidase was strongly induced in the present work. It was constitutively expressed at 35.17,

41.82 and 51.79 fold higher at 12 h, 48 and 72 h respectively in PI 416937(Additional files 1, 2&3). Two other enzymes involved in posttranslational modification of proteins, signal peptidases, showed difference in expression level between the two genotypes. One of them Gma.2057 was down regulate at 12 h (Additional file 1), whereas, Gma.22290 was constitutively expressed in PI 416937 at higher level at 48 h (Additional file 2). Some authors [19] observed up-regulation of peptidases in Arabidopsis under Al stress. Impaired sucrose induction1 gene (Gma.4033) was expressed at 4.95 and 4.27 fold higher at 12 and 48 h respectively, in PI 416937 (Additional files 1&2). Impaired sucrose induction1 encodes a conserved plant-specific protein that couples carbohydrate availability to gene expression and plant growth [40]. Mutants of this gene do not utilize carbohydrate resources efficiently. Since aluminum causes energy shortage the up- regulation of these genes in Al tolerant PI 416937 suggests an efficient utilization of the available carbohydrate under Al stress. A gene for polyamine-sperm dine biosynthesis (Gma.21460) was up-regulated 4.26 fold at 12 h (Additional file 1). Polyamines are essential for cell differentiation and growth and in plant Al stress tolerance [41]. Exogenous application of polyamines to Al containing solution culture by [41] showed enhanced Al tolerance in *Crocus sativus L*. They ascribed the amelioration of Al toxicity to exclusion of Al from entering plant roots. On the other hand [42] observed root growth inhibition in rice due to extreme accumulation of putrescine type polyamine in roots under Al stress. These studies suggest that the exact role of polyamines in Al tolerance may be species specific and that exact role remains largely unknown. Genes encoding *UDP*-glycosyl transferases Gma.6457 and Gma.30046 at 12 h (Additional file 1), Gma.6457 and Gma.2213 at 48 h (Additional file 2), and Gma.30046 and Gma.2213 at 72 h 290 (Additional file 3) were expressed at high level in PI 416937. *UDP*-glycolsyl transferases that glucoyslate plant secondary metabolites and hormones have been shown to be triggered by wounding, pathogen infection, and oxidative stress [43]. The glycosylation reaction converts secondary metabolites and hormones to inactive form for storage in vacuole or plastids that become available as needed. Glycosylation of low molecular weight molecules like harmful metabolites or environmental compounds also allow the solubilization of these compounds in water for detoxification and modulation of their biological activity [43]. Biosynthesis and homeostasis of secondary metabolites with antioxidant and detoxification property such polyphenols, benzoic acid and terpeniods could relate these enzymes to Al tolerance.

3.3.5. Cell Signaling, Structural and Cell Cycle Genes

Like other environmental or endogenous signals, perception and transmission of Al by the cell constitutes the initial steps of plant response to Al toxicity. Protein kinases and phosphatases which activate proteins by phosphorylation and dephosphorylation, respectively, are the key players in cell signaling. Cell wall associated receptor kinase (*WAK1*) is the first of such genes identified to be involved in Al signaling [44]. In our study, a protein kinase family gene Gma.2166 was up-regulated 7.88 fold at 2 h (**Table 1**). A second protein kinase family gene BI317550 was down-regulated 8 fold at 12 h (Additional file 1) and up-regulated 7.18 fold at 48 h (Additional file 2). Protien kinases activate proteins by phosphorylating serine/theorine residues of target proteins. Protein phosphatase 2c that activates target signaling proteins by opposite action of dephosphorylation was also up-regulated 4.54 and 5.43 fold at 12 and 72 h respectively (Additional file 1&2). Two other kinases, wall associated kinase Gma.27299 (Additional file 1) and transmembrane receptor kinases Gma.39148 (Additional file 2) were down-regulated. In Arabidobsis [19] identified several aluminum responsive protein kinases and phosphatases similar to the ones reported in our study. It worths noting that a novel cell cycle gene, encoding *WPP2* domain protein found to be involved in mitosis and tap root elongation and lateral root proliferation [45] was constitutively expressed at 6.42, 9.25, 12.25 and 10.94 fold higher in PI 416139 at 2, 12, 48 and 72 h respectively (Table 1, Additional files 1, 2&3). The main mechanism of Al induced root growth inhibition occurs by hindrance of cell division and cell elongation. We postulate that the higher expression level of cell division stimulating and lateral root proliferation gene in PI 416937 may partly contribute to its Al and drought tolerance characteristics. Histone H1-3 DNA binding protein (Gma.5833) was constitutively and consistently up-regulated 10.27, 8.37, and 12.56 fold at 12, 48 and 72 h respectively (Additional files 1, 2&3). Histone proteins are required for nucleosome assembly and their induction under Al stress was previously reported in Arabidopsis [18]. A lipid sequestration oleosin protein (Gma.17961) and octicosapeptide *PB1* domain protein with signaling function were up-regulated at 2 h. Oleosin protein has been recently shown by mutation analysis to be involved in cold tolerance in oilseeds of Arabidopsis [46].

4. CONCLUSIONS

We compared the transcriptome profiles of Al-tolerant and Al-sensitive soybean genotypes to identify potential genetic factors underlying Al tolerance trait. Our results

uncovered several novel putative genes which might potentially have influence on soybean Al tolerance. Among these are, rare cold inducible protein (*RCI2B)*, a cell proliferation protein (*WPP2*), pectinestrease inhibitor, *ADR6*-like transcription factor, oleosin protein and malonyl-COA: Isoflavone 7-O-glucoside-6″-O-malontransferase. The transcription factor, *ADR6* is an auxin down regulated gene. Al suppresses auxin biosynthesis and transport in root system which might be one possible mechanism of Al induced root growth inhibition. Conversely, *ADR6* is triggered under Al stress probably acting in parallel pathway to auxin to restore root growth under Al stress. Root cell wall rigidification by Al binding is one principal mechanism of Al toxicity. Cell wall metabolism enzymes and proteins are induced under Al stress and may counteract Al effects on root cell walls. It is increasingly evident that these proteins as well as cell wall pectin and hemicelluloses contents are important determinants of Al tolerance. Evidence from this study implies that cell wall remodeling enzymes and proteins may play role in soybean Al tolerance. In conclusion, as judged by levels and patterns of expression difference between tolerant and sensitive genotypes, cellular function, and mechanisms of aluminum tolerance and toxicity some of the genes described in this study could be the genetic determinants of Al-tolerance trait in soybean. These genes deserve further functional characterization for eventual utilization in developing soybean germplasm adapted to high aluminum soils.

5. ACKNOWLEDGEMENTS

We are grateful to USDA Plant Genetic Resources and Dr. Thomas Carter of North Carolina State University for generous gifts of the seeds of soybean genotypes used in this study. We thank Dr. Ernie Cebert of Alabama A&M University for his kind assistance in seed multiplication. We are indebted to Dr. Michael Crowley, Genomics Core facility at University of Alabama at Birmingham for his assistance and training on microarray analysis in this project. This research was supported in part by NSF/EPSCOR grant no. 05026. The financial support of Alabama Graduate Research Scholars Program is greatly appreciated.

REFERENCES

[1] Kochian, L.V, Hoekenga, O.A. and Piñeros, M.A. (2004) How do crop plants tolerate acid soils? Mechanism of aluminum tolerance and phosphorous efficiency. *Annual Review of Plant Biology*, **55**, 459-493. doi:10.1146/annurev.arplant.55.031903.141655

[2] Ma, J.F., Ryan, P.R. and Delhaize, E. (2001) Aluminum tolerance in plants and the complexing role of organic acids. *Trends in Plant Science*, **6**, 273-278. doi:10.1016/S1360-1385(01)01961-6

[3] Eticha, D., Staß, A. and Horst, W.J. (2005) Localization of aluminum in the maize root apex: Canmorin detect cell wall-bound aluminum. *Journal of Experimental Botany*, **56**, 1351-1357. doi:10.1093/jxb/eri136

[4] Hossain, Z.A.K., Koyama, H. and Hara, T. (2006) Growth and cell wall properties of two wheat cultivars differing in their sensitivity to aluminum stress. *Journal of Plant Physiology*, **163**, 39-47. doi:10.1016/j.jplph.2005.02.008

[5] Mortia, A.,Yanagisawa, O., Takatsu, S., Maeda, S. and Hiradate, S. (2008) Mechanism for the detoxification of aluminum in roots of tea plant (Camellia sinsensis (L) Kuntze). *Phytochemistry*, **69**, 147-153. doi:10.1016/j.phytochem.2007.06.007

[6] Watanabe, T., Osaki, M., Yano, H. and Rao, I.M. (2006) Internal mechanisms of plant adaptation to aluminum toxicity and phosphorus starvation in three tropical forages. *Journal of Plant Nutrition*, **29**, 1243-1255. doi:10.1080/01904160600767484

[7] Sasaki, T., Yamamoto, Y., Ezaki, B., Katsuhara, M., Ahn, S.J., Ryan, P.R., Delhaize, E. and Matsumoto, H.A. (2004) Wheat gene encoding aluminum activated malate transporter. *The Plant Journal*, **37**, 645-653. doi:10.1111/j.1365-313X.2003.01991.x

[8] Magalhaes, J.V., Liu, J., Guimará, C.T., Lana, U.G.P, Alves, V.M.C., Wang, Y.H., Schaffert, R.E., Hoekenga, O.A., Piñehaff, M.A., klein, P.E., Carneiro, N.P., Coelho, C.M., Trick, H.N. and Kochian, L.V.(2007) A gene in the multidrug and toxic compound extrusion (MATE) family confers aluminum tolerance in sorghum. *Nature Genetics*, **39**, 1156-1161. doi:10.1038/ng2074

[9] Silva, I.R., Smyth, T.J., Israel, D.W., Raper, C.D., Ruffy, T.W. (2001) Magnesium is more efficient than calcium in alleviating aluminum rhizotoxicity in soybean and its ameliorative effect is not explained by the gouy-chapman-stern model. *Plant Cell Physiology*, **42**, 538-545. doi:10.1093/pcp/pce066

[10] Bianchi-Hall, C.M., Carter, T.E., Bailey, M.A., Mian, M.A.R., Ashley, D.A., Boerma, H,E., Arellano C, Hussey, R.S. and Parrott, W.A. (2000) Aluminum tolerance associated with quantitative trait loci derived from soybean PI 416937 in hydroponics. *Crop Science*, **40**, 538-545. doi:10.2135/cropsci2000.402538x

[11] Nian, H., Yang, Z., Huang, H. and Yan, X. (2004) Citrate secretion induced by aluminum stress may not be a key mechanism responsible for differential aluminum tolerance of some soybean genotypes. *Journal of Plant Nutrition*, **27**, 2047-2066. doi:10.1081/PLN-200030112

[12] Ermolayev, V., Weschke, W. and Manteuffel, R. (2003) Comparison of Al-induced gene expression in sensitive and tolerant soybean cultivars. *Journal of Experimental Botany*, **54**, 2745-2756. doi:10.1093/jxb/erg302

[13] Ragland, M. and Soliman, K.M. (1997) Two genes induced by Al in soybean roots. *Plant Physiology*, **114**, 395.

[14] Guo, P., Bai, G., Carver, B., and Li, R. (2007) Transcriptional analysis between two wheat near-isogeniclines contrasting in aluminum tolerance under aluminum stress. *Mol Genet Genomics*, **277**, 1-12. doi:10.1007/s00438-006-0169-x

[15] Hounde, M. and Diallo, A.O. (2008) Identification of genes and pathways associated with aluminum stress and

tolerance using transcriptome profiling of wheat near-isogenic lines. *BMC Genomics*, **9**, 400. doi:10.1186/1471-2164-9-400

[16] Maron, L.G., Kirst, M., Mao, C., Milner, M.J., Menossi, M. and Kochian L.V. (2008) Transcriptional profiling of aluminum toxicity and tolerance responses in maize roots. *New Phytologist*, **179**, 116-128. doi:10.1111/j.1469-8137.2008.02440.x

[17] Chandran, D., Sharopova, N., Ivashuta, S., Gantt, J.S., VandenBosch, K.A. and Samac, D.A. (2008a) Transcriptome profiling identified novel genes associated with aluminum toxicity, resistance and tolerance in Medicago truncatula. *Planta*, **228**, 151-166. doi:10.1007/s00425-008-0726-0

[18] Chandran, D., Sharopov, N., VandenBosch, K.A., Garvin, D.V. and Samac, D.A. (2008b) Physiological and molecular characterization of aluminum resistance in Medicago truncatula. *BMC Plant Biology*, **8**, 89. doi:10.1186/1471-2229-8-89

[19] Kumari, M., Taylor, G.J. and Deyholos, M.K. (2008) Transcriptome responses to aluminum stress in roots of Arabidopsis. *Mol Genet Genomics*, **279**, 339-357. doi:10.1007/s00438-007-0316-z

[20] Furukawa, J., Yamaji, N., Wang, H., Mitan, N., Murata, Y., Sato, K., Katsuhara, M., Takeda, K. and Ma, J.F. (2007) An aluminum activated citrate transporter in barley. *Plant cell Physiology*, **48**, 1081-1091. doi:10.1093/pcp/pcm091

[21] Wu, Z., Soliman, K.M., Bolton, J.J., Saha, S. and Jenkins, J.N. (2008) Identification of differentially expressed genes associated with cotton fiber development in a chromosomal substitution line (CS-B22sh). *Funct Integr Genomics*, **8**, 165-174. doi:10.1007/s10142-007-0064-5

[22] Morey, J.S., Ryan, J.C. and Van Dolah, F.M. (2006) Microarray validation factors influencing correlation between oligonucleotide microarrays and real-time PCR. *Biology Proceeding online*, **8**, 175-193.

[23] Zhu, J., Shi, H., Lee, B.H., Damsz, B., Cheng, S., Stirm, V. and Zhu, J.K. (2004) An arabidopsis homeodomain transcription factor gene, HOS9, mediates cold tolerance through a CBF-independent pathway. *Proceedings of the National Academy of Sciences of the United States of America*, **101**, 9873-9878. doi:10.1073/pnas.0403166101

[24] Udvardi, M.K., Kakar, K., Wandrey, M,, Montanari, O., Murray, J., Andriankaja, A., Zhang, J.Y., Benedito, V., Hofer, J.M.I., Chueng, F. and Town, C.D. (2007) Legume transcription factors: Global regulators of plant development and response to the environment. *Plant Physiology*, **144**, 538-549. doi:10.1104/pp.107.098061

[25] Zhang, Y., Wang, Z., Zhang, L., Cao, Y., Huang, D. and Tang, K. (2006) Molecular cloningand stress-dependent regulation of potassium channel gene in Chinese cabbage (Brassica rapa asp. Pekinensis).*Plant Physiology*, **163**, 968-978. doi:10.1016/j.jplph.2005.09.002

[26] Kader, J.C. (1997) Lipid transfer proteins: a puzzling family of plant proteins. *Trends in Plant Science*, **2**, 66-70. doi:10.1016/S1360-1385(97)82565-4

[27] Nieuwald, J., Feron,.R., Huisman, B.A.H., Fasolino, A., Hilbers, C.W., Derksen, J. and Marian, C. (2009) Lipid transfer proteins enhance cell wall extension in tobacco. *The Plant Cell* 2005, **17**, 2009-2019.

[28] Rengel, Z. and Zhang, W.H. (2003) Role of dynamics of intracellular calcium in aluminum toxicity syndrome. *New Physiologist*, **159**, 295. doi:10.1046/j.1469-8137.2003.00821.x

[29] Williams, L.E., Lemoine, R. and Sauer, N. (2000) Sugar transporters in higher plants-a diversity of roles and complex regulation. *Trends in Plant Science*, **5**, 283-290. doi:10.1016/S1360-1385(00)01681-2

[30] Yamamoto, Y., Kobayashi, Y., Devi, S.R., Rikiishi, S. and Matsumoto, H. (2002) Aluminum toxicity is associated with mitochondrial dysfunction and the production of reactive oxygen species in plant cells. *Plant Physiology*, **128**, 63-72. doi:10.1104/pp.010417

[31] Llorenta, F., Cobollo, R.M.L., Catalá, R., Zapater, J.M. and Salinas, J. (2002) A novel cold-inducible gene from Arabidopsis, RCI3, encodes a peroxidase that constitutes a component for stress tolerance. *The Plant Journal*, **32**, 13-24. doi:10.1046/j.1365-313X.2002.01398.x

[32] Ezaki, B., Gardner, R.C., Ezaki, Y. and Matsumoto, H. (2000) Expression of aluminum-induced genes in transgenic arabidopsis plants can ameliorate aluminum stress and/or oxidative stress. *Plant Physiology*, **127**, 918-927. doi:10.1104/pp.010399

[33] Liu, Q., Yang, J.L., He, L.S., Li, Y.Y. and Zheng, S.J (2008) Effect of aluminum on cell wall, plasma 555 membrane, antioxidants and root elongation in triticale. *Biologica Planttarum*, **2**, 87-92. doi:10.1007/s10535-008-0014-7

[34] Sarowar, S., Kim, Y.J., Kim, E.N., Kim, K.D., Hwang, B.K.H., Islam, R. and Shin, J.S. (2005) Over expression of a pepper basic pathogenesis-related protein 1 gene in tobacco plants enhances resistance to 559 heavy metal and pathogen stresses. *Plant Cell Report*, **24**, 216-224. doi:10.1007/s00299-005-0928-x

[35] Zhen, Y., Qi, J.L., Wang, S.S., Su, J., Xu, G.H., Zhang, M.S. and Miao, L.V. (2007) Comparative proteome analysis of differentially expressed proteins induced by Al toxicity in soybean. *Physiologia Plantrarum*, **131**, 542-554. doi:10.1111/j.1399-3054.2007.00979.x

[36] Rout, G.R. and Samantaray, S. (2001) Aluminum toxicity in plants: A review. *Agronomie*, **21**, 3-21. doi:10.1051/agro:2001105

[37] Jemo, M., Abaidoo, R.C., Nolte, C. and Johannes, W. (2006) Aluminum resistance of cowpea as affected by phosphorus-deficiency stress. *Journal of Plant Physiology*, **164**, 442-451. doi:10.1016/j.jplph.2005.12.010

[38] Yang, J.L., Zhang, L.Y., Zhang, Y.J., Wu, S.S., Wu, Y.R. and Zheng, P. (2008) Cell wall polysaccharides are specifically involved in the exclusion of aluminum from the rice root apex. *Plant Physiology*, **146**, 602-611. doi:10.1104/pp.107.111989

[39] Ohishi, K., Nagamune, K., Maeda, Y. and Kinoshita, T. (2003) Two subunits of GPI transamidase GPI8 and P-IG-T form a functionally important intermolecular disulfide bridge. *Journal of Biochemistry*, 16, 13959-13967.

[40] Rook, F., Corke, F., Baier, M., Holman, R., May, A.G. and Bevan, M.W. (2006) Impaired sucrose induction1 encodes a conserved plant-specific protein that couples carbohydrate availability to gene expression and plant growth. *Plant Journal*, **46**, 1045-1058. doi:10.1111/j.1365-313X.2006.02765.x

[41] Chen, W., Xu, C., Zhao, E.B., Wang, E.X. and Wang, E.Y. (2008) Improved Al tolerance of saffron (*Crocus sativus*

L.) by exogenous polyamines. *Acta Physiology Plant*, **30**, 121-127. doi:10.1007/s11738-007-0100-z

[42] Wang, J.W. and Kao, C.H. (2006) Aluminum inhibited root growth of rice seedlings is mediated through putrescine accumulation. *Plant Soil*, **288**, 373-381. doi:10.1007/s11104-006-9127-y

[43] Sepu, G., Jime,.I., Benı, P.R., Porta, H. and Sosa, M.R. (2005) A red beet (Beta vulgaris) UDP-glucosyltranferase gene induced by wounding, bacterial infiltration and oxidative stress. *Journal of Experimental Botany*, **56**, 605-611. doi:10.1093/jxb/eri036

[44] Sivaguru, M., Ezaki, B., He, Z.H.,Tong, H., Osawa, H., Baluska, F.,Volkmann, F. and Matsumoto, H. (2003) Aluminum induced gene expression and protein localization of a cell wall-associated receptor kinase in arabidopsis. *Plant Physiology*, **32**, 2256-2266. doi:10.1104/pp.103.022129

[45] Patel, S., Rose, A., Meulia, T., Dixit, R., Richard, J.C. and Meiera, I. (2004) Arabidopsis WPP-domain proteins are developmentally associated with the nuclear envelope and promote cell division. *The Plant Cell*, **16**, 3260-3273. doi:10.1105/tpc.104.026740

[46] Shimada, T. (2000) Oilseeds in arabidopsis thalina. *The Plant Journal*, **55**, 798-809. doi:10.1111/j.1365-313X.2008.03553.x

[47] Silva, I.R., Smyth, T.J., Moxley, D.F., Carter, T.E., Allen, N.S. and Ruffy, T.W. (2000) Aluminum accumulation at the nuclei of cells in the root tip. fluorescence detection using lumogallion and confocal laser scanning microscopy. *Plant Physiology*, **23**, 543-552. doi:10.1104/pp.123.2.543

[48] Wu, Z., Irizarry, A.R., Gentleman, R., Murillo, F.M. and Spencer, F. (2004) A model-based background adjustment for oligonucleotide expression arrays. *Journal of the American Statistical Association*, **99**, 909. doi:10.1198/016214504000000683

[49] Pfaffi, M.W. (2001) A new mathematical model for relative quantification in real time RT- PCR. *Nucleic Acids Research*, **29**, 2002-2007.

Additional materials

Additional file 1 Table of genes differential expressed between Al-tolerant PI 416937 and Al-sensitive Young 12 h post aluminum treatment

Unigene ID	Fold Change	Functional Category	Annotation[b]
Gma.19126[c]	35.17(up)	metabolism	GPI transamidase PIG-s-related phosphatidylinositol-glycan biosynthesis class S protein
Gma.7535[c]	27.02(up)	stress response	malonyl-COA isoflavone 7-0-glucoside 6"0- malony transferase
Gma.6457[c]	26.07(up)	metabolism	transferase activity/transfer glycosyl group
Gma.9718[c]	17.29(up)	unclassified	unknown
Gma.34414[c]	17.25(up)	transport	Sec14 cytosolic factor family protein
Gma.13976[c]	17.08(up)	unclassified	unknown
Gma.13499[c]	16.14(up)	metabolism	embryo defective 1075 carboxylase
Gma.17427[c]	16.11(up)	unclassified	unknown
Gma.16916[c]	15.90(up)	transport	Sec14 cytosolic factor family protein
Gma.30046[c]	15.71(up)	metabolism	UDP-glucosyl transferase/para-aminobenzoic acid metabolism
Gma.26943[c]	14.59(up)	transport	Sec14 cytosolic factor protein

Gma.8198[c]	14.08(up)	unclassified	hypothetical protein
Gma.25191	12.02(down)	unclassified	unknown
Gma.17162[c]	11.68(up)	unclassified	hypothetical protein
Gma.21642[c]	11.27(down)	unclassified	unknown
Gma.30710[c]	10.40(up)	transport	Sec14 cytosolic factor family protein
Gma.2057	10.14(down)	metabolism	Signal peptidase/proteolysis
Gma.5833[c]	10.27(up)	structural	histone1-3 DNA binding
Gma.27299	10.14(down)	cell signaling	WAK5/wall associated kinase, kinase activity
Gma.4281	9.94(down)	transcription factor	WRKY70 DNA binding protein transcription factor
Gma.2213[c]	9.54(up)	metabolism	UDP-glucoronosyl/transfers glycosyl group Para-amino benzoic acid biosynthesis process
Gma.395	9.25(up)	cell cycle	WPP2 domain protein1/lateral root development
Gma.3880[c]	9.15(up)	transport	lipid transfer protein-related/lipid binding/lipid transfer
Gma.21421[c]	9.0(up)	unclassified	hypothetical protein
Gma.21625[c]	8.92(down)	unclassified	unknown
Gma.22330[c]	8.12(up)	unclassified	unknown
BI317550[c]	8.00(down)	signaling	protein kinase family protein/kinase activity
Gma.15692[c]	7.95(down)	unclassified	unknown
Gma.33997[c]	7.62(down)		Ran GTpase/chromatin binding, zinc ion binding
Gma.21391[c]	7.30(up)	unclassified	unknown
BI788325[c]	7.26(down)	unclassified	unknown
Gma.4819	7.09(up)	unclassified	hypothetical protein
Gma.10737[c]	6.98(down)	unclassified	hypothetical protein
Gma.22290[c]	6.73(up)	metabolism	serine carboxypeptidase/proteolysis
Gma.24523[c]	6.68(down)	metabolism	dehydrodolichyl diphosphate synthase/ dolichol biosynthesis
Gma.26538[c]	6.67(down)	stress response	chaperonin putative
BU551389[c]	6.64(down)	metabolism	reverse transcriptase
Gma.36363	6.59(up)	metabolism	UDP-glucosyl transferase
Gma.33184	6.54(up)	stress response	Xyloglucan specific fungus endoglucanase inhibitor
Gma.9683[c]	6.53(up)	unclassified	unknown
Gma.8292	6.19(up)	stress response	functional resistance protein KR1
Gma.13640[c]	6.15(down)	metabolism	cellulose synthase-like A3/cellulose synthase
Gma.10407c	6.02(up)	metabolism	phospholipid diacylglycerol acyltransferase lipid metabolism
Gma.6707c	6.01(down)	transport	dicarboxylate/tricarboxylate carrier/mitochondrial transport
Gma.2256c	5.96(down)	structural	ribosomal protein P2/ 60S acidic ribosomal protein

BI971552c	5.82(up)	unclassified	hypothetical protein
Gma.22124	5.70(up)	metabolism	pectinestrase family protein/cell wall modification
Gma.10907	5.58(up)	unclassified	hypothetical protein
Gma.10858c	5.33(down)	cell signaling	jasmonate-z-domain protein3/ jasmonic acid mediated signaling
Gma.6245c	5.28(up)	stress response	pathogenesis related thaumatin protein family
Gma.22107c	5.24(up)	metabolism	hydrolase homolo6 ADP-ribose diphosphatase
BI316842 c	5.15(up)	unclassified	unknown
Gma.5622	5.06(down)	unclassified	unknown
Gma.4033c	4.95(up)	metabolism	impaired sucrose induction1/controls growth and Development
Gma.14036c	4.92(down)	unclassified	unknown
Gma.5609c	4.91(up)	metabolism	pyrophosphorylase4/ inorganic diphosphatase activity
Gma.37580c	4.75(up)	transport	sulfate transmembrane transporter
Gma.15804c	4.82(down)	unclassified	hypothetical protein
Gma.16666c	4.78(down)	transcription factor	basic helix-loop-helix family protein/transcription factor
Gma.510616c	4.71(up)	unclassified	unknown
Gma.15007c	4.64(up)	metabolism	iron-sulfur cluster binding protein/electron transfer
AW756534c	4.60(up)	metabolism	glycosyl hydrolase family 3 protein/o-glycosyl cpds
Gma.24361c	4.54(up)	signaling	protein phosphotase 2c/phosphatase activity
Gma.22442c	4.53(down)	unclassified	unknown
Gma.17931c	4.47(down)	metabolism	choloroplast thylakoid processing peptidase
Gma.13006c	4.47(up)	unclassified	hypothetical protein
Gma.2848c	4.35(up)	cell cycle	ribonucleotide-diphosphate reductase /DNA repair cell cycle regul.
Gma.39148	4.34(down)	signaling	receptor kinase/transmembrane receptor protein
Gma.21460c	4.26(up)	metabolism	SPDS1 spermidine synthase activity/spermidine biosynthesis
Gma.1288c	4.18(up)	unclassified	unknown
Gma.24428	4.15(up)	unclassified	unknown
Gma.22248c	4.13(down)	structural	myosin-like protein/actin filament based movement
Gma.4958	4.13(down)	metabolism	acyl-activating enzyme 12/catalytic activity metabolic process
Gma.17205c	4.13(up)	transport	zinc induced facilitator-like1 carbohydrate transmembrane transporter activity
Gma.20904	4.09(up)	unclassified	hypothetical protein
Gma.42414c	4.06(down)	transport	ATCAx6 calcium exchanger/calcium: cation antiporter activity
Gma.17306c	4.05(up)	structural	structural molecule/ cell adhension
Gma.20157	4.03(up)	metabolism	UDP-Gl cNAc dolichol phosphate N-acylglucosamine Transferase

a significance thresholds (fold change ≥ 4, p ≤ 0.01), b e-value (< e-10) = the probability that the match has no biological basis, cconstitutively expressed, dup=up-regulated in PI 416937 relative to Young, down=down-regulated in PI 416937 relative to Young. Note: constitutively expressed means the gene is expressed at higher or lower level in PI 416937 relative to Young with or without aluminum treatment.

Additional file 2 Table of genes differentially expressed between Al-tolerant PI 416937 and Al-sensitive Young 48 h post aluminum treatment.

Unigene ID	Fold Change	Functional Category	Annotationb
Gma.27795[c]	257.05(up)	stress response	RCI2B rare cold inducible 2B /response to cold
Gma.9308[c]	45.92(up)	unclassified	unknown
Gma.6457[c]	43.79(down)	metabolism	glycosyl transferase family protein/transferase activity
Gma.8292	42.19(up)	transcription factor	WRKY19 /transcription factor
Gma.19126[c]	41.82(up)	metabolism	transamidase PIG-S (phosphatidylinositolglycan biosynthesis class S protein
Gma.12911	35.54(down)	metabolism	salicylic acid methyl transferase-like protein
Gma.16558[c]	27.31(down)	Transcription factor	embryro defective/enzyme activator transcription regulator
Gma.7335[c]	26.98(up)	signaling	plastid movement impaired and calcium mediated Signaling
Gma.13976[c]	24.63(up)	unclassified	unknown
Gma.1043[c]	22.26(down)	unclassified	unknown
Gma.13499[c]	18.88(down)	metabolism	embryo defective 1075 carboxylase/amino acid metabolism
Gma 21642c	16.11(down)	unclassified	unknown
Gma.22330c	15.71(up)	unclassified	unknown
Gma. 2057c	15.25(down)	metabolism	signal peptidase/peptidase activity- proteolysis
Gma. 26538c	14.33(down)	stress response	chaperonin
Gma.8198c	13.75(up)	unclassified	hypothetical protein
Gma.29488c	13.44(up)	unclassified	unknown
Gma.17947	13.37(down)	stress response	17.6 kDa class I small heat shock protein
Gma.16916c	12.69(up)	transport	sec14 cytosolic factor family protein
Gma.395c	12.25(up)	cell cycle	WPP2 domain protein/lateral root development/mitosis,
Gma.9718c	12.14(up)	unclassified	unknown
Gma.17162c	11.91(up)	unclassified	hypothetical protein
Gma.30046c	11.63(up)	metabolism	UDP-glucosyl transferase
Gma.34414c	11.47(up)	transport	sec14 cytosolic family protein/transport phosphoglyceride
Gma.681	11.16(up)	unclassified	unknown
Gma.12558c	10.96(down)	transport	non-intrinsic ABC protein (Arabidopsis) ATpase activ-ity/transmembrane transport
Gma.9140c	10.45(down)	unclassified	hypothetical protein
Gma.8966	10.28(down)	unknown	unknown
Gma.2256c	10.00(down)	structural	60 S acidic ribosomal protein/structural component of ribo-some-translational elongation
Gma.21625c	9.89(down)	unclassified	unknown
Gma.2213c	9.77(up)	metabolism	UDP-glucoronosyl transferase
Gma.22107	9.32(up)	metabolism	nudix hydrolase/ADP-ribose diphosphatase activity
Gma.10735c	9.14(up)	stress response	chitinase/cell wall catabolic process
Gma.21391c	9.00(up)	unclassified	unknown
Gma.30710c	8.76(up)	transport	sec14 cytosolic factor family protein/ phosphoglyceride trans-port family protein
Gma.5833c	8.37(up)	structural	histone H1-like protein DNA-binding
Gma.36226c	8.08(up)	unclassified	unknown
Gma.24523c	7.92(up)	metabolism	dehydrodolichyl diphosphate synthase Terpenoid biosynthesis process
Gma.6245c	7.62(up)	stress response	pathogenesis related family protein

BI317550c	7.18(up)	cell signaling	protein kinase family protein
Gma.8456	7.08(up)	stress response	resistance protein KR2
Gma.9683c	7.02(up)	unclassified	unknown
Gma.7840c	6.86(up)	unclassified	unknown
Gma.3880c	6.79(up)	transport	lipid transfer/lipid binding-lipid transport
Gma.6707c	6.69(up)	transport	dicarboxylate/tricarboxylate carrier, binding oxidative phosphorylation /cell wall component
Gma.1343c	6.72(up)	unclassified	unknown
Gma.5609c	6.61(up)	metabolism	pyrophosphorylase/ inorganic diphosphatase metabolic process
BI316842c	6.25(up)	unclassified	unknown
Gma.6883c	6.16(up)	unclassified	unknown
Gma.38915c	6.00(down)	structural	Ran GTpase binding/zinc ion binding regulator of chromosome condensation
Gma.17427	5.98(up)	unclassified	unknown
Gma.33146	5.88(down)	response to stress	glutathione-s-transferase/toxin catabolic process
Gma.22248	5.83(down)	structural	myosin-like protein XIF/motor activity
BE822969c`	5.74(up)	unclassified	unknown
Gma.28057	5.68(up)	Transcription factor	ADR6 transcription factor/Sal5-4a protein
Gma.15692c	5.62(up)	unclassified	unknown
Gma.24361c	5.43(up)	signaling	protein phosphatase 2c
Gma.34568c	5.40(up)	response to stress	disease resistance/protein binding
Gma.1840c	5.32(up)	Response to stress	DNAJ heat shock protein/N-terminal
Gma.12898c	5.29(up)	unclassified	unknown
Gma.17306c	5.18(up)	transcription factor	bzip (bzip94) transcription factor
Gma.31872c	4.94(up)	metabolism	polyubiquitin
Gma. 413	4.74(up)	stress response	peroxidase/response to oxidative stress (prx4) coumarine and phenylpropanoid biosynthesis
Gma.8417c	4.68(up)	signaling	calmodulin-like domain protein kinase
Gma.42414c	4.68(down)	transport	calcium exchanger 6/calcium antiporter
Gma.9308	4.49(up)	unclassified	unknown
Gma.13006c	4.46(up)	unclassified	hypothetical protein
Gma.33748c	4.46(down)	stress response	glutathione-s-transferase/toxin catabolic activity
Gma.33997c	4.40(down)	unclassified	unknown
Gma.17205	4.33(up)	transport	zinc induced facilitator-like1/sugar: hydrogen symporter activity
Gma.10961	4.30(up)	unclassified	unknown
Gma.3345	4.28(up)	metabolism	class III alcohol dehydrogenase5 metabolic process
Gma.4033c	4.27(up)	metabolism	impaired sucrose induction1/carbohydrate metabolism regulation
BI971205	4.25(down)	unclassified	unknown
Gma.31872	4.22(down)	metabolism	polyubiquitin
Gma.17053	4.21(down)	stress response	glyceraldehydes dehydrogenase/response to stress
Gma.31645	4.08(up)	metabolism	invertase/pectin methyl esterase inhibitor family protein
Gma. 22290*	4.12(up)	metabolism	serine peptidase family like10/caboxy peptidase

[a] significance thresholds (fold change ≥ 4, $p \leq 0.01$), [b] e-value (< e-10) = the probability that the match has no biological basis, [c]constitutively expressed, [d]up=up-regulated in PI 416937 relative to Young, down=down-regulated in PI 416937 relative to Young. Note: constitutively expressed means the gene is expressed at higher or lower level in PI 416937 relative to Young with or without aluminum treatment.

Additional file 3 Table of genes differentially expressed between Al-tolerant PI 416937 and Al-sensitive Young 72 h post aluminum treatment.

Unigene ID	Fold Change	Functional Category	Annotationb
Gma. 30979	8.4	unclassified	Cab3 chlorophyll a/b binding protein
Gma.12839	12.77	unclassified	unknown
Gma.12898c	10.05	unclassified	unknown
Gma.13035	5.23	unclassified	unknown
Gma.13231	3.60	structural	VAP27-2 (VAMP/SYNAPTOBREVIN 27-2 structural molecule
Gma.13976[c]	55.61	unclassified	unknown
Gma.3529	5.31	unclassified	unknown
Gma.5716[c]	9.21	unclassified	unknown
Gma.6184	17.57	unclassified	3-dehydroquinate dehydratase/ NADP binding embryo defective 3004
Gma.681	6.67	unclassified	unknown
Gma.7719[c]	5.43	unclassified	unknown
Gma.7840[c]	7.07	unclassified	unknown
Gma.9308[c]	15.23	unclassified	unknown
Gma.9683[c]	6.06	unclassified	unknown
Gma.9718[c]	10.96	unclassified	unknown
Gma.26808[c]	5.08	unclassified	hypothetical protein
Gma.32966[c]	5.35	metabolism	pseudouridine synthase family protein
Gma.21391[c]	8.09	unclassified	unknown
Gma.1343[c]	10.99	unclassified	unknown
Gma.18784[c]	4.18	unclassified	unknown
Gma.22330[c]	5.67	unclassified	unknown
Gma.16334	6.59	stress response	Uridylytransferase-related protein/response to cold
Gma.5609	5.23	metabolism	Pyrophosphorylase4 inorganic diphosphatase metabolic process
Gma.16838	8.09	metabolism	photosystem I subunit H-2/function photosynthesis
Gma.27795[c]	253.35	stress response	rare cold-inducible 2b (RCI2b)
Gma.15377	5.72	metabolism	PSAG/photosystem I-NADP+ reduction
Gma.8198[c]	13.45	unclassified	unknown
Gma.14306[c]	5.68	unclassified	pentatricopeptide (PPR) repeat-containing protein
Gma.17451	16.48	metabolism	polyubiquitin10 protein/aging
Gma.8097[c]	4.31	transport	sulfate transmembrane transporter protein activity
Gma.2213	6.69	metabolism	UDP-glucoronosyl/UDP-glucosyl transferase
Gma.26162	4.18	metabolism	catalytic activity/nucleoside metabolic process
Gma.2360	5.43	metabolism	light harvesting complex of photo system II chlorophyll binding
Gma.16710	12.28	transport	phospholipid transfer protein homolg1/lipid transport
Gma.17162[c]	15.22	unclassified	integral membrane protein

Gma.16916c	10.18	transport	Sec14 cytosolic factor family protein/ transport function
Gma.19126c	51.79	metabolism	transamidase component PIG-S phosphatidy linositol-glycan biosynthesis class S protein
Gma.30046c	8.11	metabolism	UDP-glycosyl transferase/para-aminobenzoic-acid metabolism
Gma.18216c	5.55	unclassified	hydroxysteriod dehydrogenase5 oxidoreductase
Gma.1840c	6.45	stress response	DNAJ heat schock N-terminal domain containing protein protein folding
Gma.27681	6.55	transport	protease inhibitor protein/lipid binding/lipid transport
Gma.6245c	12.40	stress response	pathogenesis related family protein
Gma.24018c	5.20	metabolism	dienelacetone hydrolase family protein hydrolase activity act on carboxylic esters
Gma.5833c	12.56	structural	histone H1-3/DNA binding
Gma.3880c	13.89	transport	lipid transfer protein
Gma.395c	10.94	cell cycle	WPP2 domain protein2/lateral root development
Gma.31645	6.70	unclassified	invertase/pectin methylestrase inhibitor family protein
Gma.5539c	4.59	metabolism	chloroplast thylakoid lumen protein
Gma.6883c	4.90	signaling	disease resistance protein/ATP binding transmembrane receptor activity
Gma.34414c	11.23	transport	Sec14 cytosolic factor family protein
Gma.8121	4.86	metabolism	SHM7 (Serine hydroxyl methyltransferase)
		metabolism	glycine hydroxymethyltransferase activity
Gma.3345c	4.93	metabolism	alcohol dehydrogenase/metabolism of xenobiotics by cytochrome P450
Gma.17306c	6.45	transcription factor	bzip (bzip94) transcription factor
Gma.10882	4.32	unclassified	small nuclear ribonucleoprotein/nucleic acid binding
Gma.30979	5.04	unclassified	chlorophyll a/b-binding protein
Gma.7535c	54.65	stress response	malonyl-COA: isoflavone 7-O-glucoside 6"-0-malonyl trans-ferase (MT7)
BM091947	4.02	unclassified	unknown
CD390824	5.54	unclassified	unknown
BI943856c	4.28	unclassified	unknown
Gma.37041	5.22	stress response	inositol pentabisphosphate 2-kinase (IPK1)/stress response
Gma.3529	4.03	unclassified	unknown
Gma.12898c	7.15	unclassified	unknown
BI316842c	4.38	unclassified	unknown
Gma.39401	5.17	unclassified	hypothetical protein
Gma.5555	4.60	unclassified	unknown
BM519816	7.02	unclassified	unknown
BF598929	8.23	unclassified	embryodefgective3004-dehydroquinate binding/ catalytic activity
Gma.8134	3.75	metabolism	FMN adenyltransferase activity riboflavin biosynthetic process

[a] significance thresholds (fold change ≥ 4, p ≤ 0.01), [b] e-value (< e-10) = the probability that the match has no biological basis, [c] constitutively expressed, [d] all genes in this table are up-regulated in PI 416937 relative to Young. Note: constitutively expressed means the gene is expressed at higher or lower level in PI 416937 relative to Young with or without aluminum treatment.

