Australian Journal of Crop Science AJCS
AJCS 8(2):276-283 (2014) ISSN:1835-2707

Differentially expressed genes and *in silico* analysis in response to ozone (O_3) stress of soybean leaves

Jun-Cheol Moon[1], Won Cheol Yim[2], Sung Don Lim[1], Kitae Song[2] and Byung-Moo Lee[2,*]

[1]Agriculture and Life Sciences Research Institute, Kangwon National Univ., Chuncheon 200-701 Korea
[2]Plant Molecular Genetic Lab., Dep. of Life Science, Dongguk Univ-Seoul, Seoul 100-715, Korea

***Corresponding author: bmlee@dongguk.edu**

Abstract

This study was conducted to identify specific/predominant genes of soybean leaves fumigated with ozone (O_3) in the fifth trifoliolate stage. Leaves were exposed to O_3 at concentrations of 120 and 200 ppb for four hours on each of five consecutive days. We compared the expression levels of mRNA fragments in untreated versus treated leaves based on, the different banding patterns obtained by agarose gel electrophoresis. These patterns revealed that 408 transcripts were differentially expressed between treated and untreated leaves; 153 transcripts were up-regulated in soybean leaves fumigated with O_3 and 225 transcripts were down-regulated. The putative functions of the expressed sequence tags (ESTs) were identified using BLASTX. Of 408 transcripts, 186 differentially expressed genes (DEGs) in response to O_3 significantly matched previously known genes. To identify the functional modules of DEGs and protein domains, we used the Swiss-Port IPROClass database. The DEGs of soybean leaves in response to O_3 were annotated by cross-species analysis with *Arabidopsis*. The Cytoscape plugin ClueGO was used to identify six network pathways of enrichment over-represented significantly. Most of the O_3 responsive genes identified in this study were related to other biotic and abiotic stresses. The characterization of ESTs enables better understanding of the molecular basis of O_3 responses.

Keywords: Differentially expressed genes, Ozone stress, Soybean, Transcript.
Abbreviations: ACP_annealing control primer; DEGs_Differentially Expressed Genes; FDR_False Discovery Rate; dN_nonsynonymous substittion; dS_synonymous substitution.

Introduction

O_3 is a powerful oxidant that causes damage to plant tissue at high concentrations. Prolonged exposure to O_3 can have many negative effects, including decreased photosynthesis rates, leaf injury, accelerated senescence, and reduced crop yield. These effects are mainly caused by the entry of O_3 gas into the leaf through the stomata and the subsequent generation of various reactive oxygen species (ROS; Laisk et al., 1989). In plants, ROS accumulation has a direct negative effect on the cell causing necrosis/chlorosis and inducing programmed cell death (Mittler et al., 2004; Overmyer et al., 2005). The effects of O_3 are limited by the antioxidative system, which is a complex network of enzymes encoded by multiple genes (Mittler et al., 2004). Apoplastic ascorbate is often considered to be the first line of defense in response to the O_3 absorbed via the stomata (Smirnoff and Pallanca, 1996; Plöchl et al., 2000; Matyssek et al., 2012). In addition to ascorbate, intracellular redox buffering systems include glutathione, flavonoids, α-tocopherol, and ROS enzymes such as catalase (CAT), glutathione *S*-transferases, glutathione reductase (GR), and peroxidases (Mittler et al., 2004). Oxidative stress increases the activity of antioxidant enzymes, including superoxide dismutase (SOD), peroxidase, catalase, and glutathione *S*-transferase (McClung, 1997; Mittler, 2002; Tognolli et al., 2002; Blokhina et al., 2003; Ludwikow et al., 2004). The transcription levels and activities of several ROS change drastically in response to O_3 stress. In addition, many other genes are regulated by O_3 stress at the transcript and protein levels in *Arabidopsis* and rice (Agrawal et al., 2002; Ludwikow et al., 2004). The effects of O_3 resemble the characteristics of systemic acquired resistance (SAR) (Greenberg et al., 1994) and the hypersensitive response (HR) (Tenhaken et al., 1995; Pasqualini et al., 2003). They are also similar to the effects of ethylene, salicylic acid, and jasmonic acid (Rao et al., 2002; Overmyer et al., 2003; Tamaoki et al., 2003a). Soybean is a crucial crop worldwide for the provision of vegetable oil and protein. In addition, the processed products of soybean seeds are economically important for a wide range of products, including food, industry, and agricultural products. Therefore, many studies have assessed the effects of stress on soybeans. The average shoot biomass and seed yield of soybean plants decreases in response to O_3 exposure (Morgan et al., 2003). Chronic O_3 exposure at low concentrations causes serious loss of biomass and seed production, even in the absence of visible symptoms (Fiscus et al., 1997). The physiological characteristics of the soybean were studied in Korean soybean cultivars by Lee et al. (2006).In order to understand the molecular mechanisms of the stress response, DEGs have been studied in stress-induced tissues. However, the expression of DEGs is relatively low in these tissues. In addition, individual transcript detection is not easy in the abundant mRNA in stress-treated tissues. Thus, polymerase chain reaction (PCR) amplification, under highly specific conditions, is required to detect DEGs transcribed at low levels. Novel annealing control primer (ACP)-based differential display (DD)-PCR technology was used by Hwang et al. (2004) to identify DEGs in tissues. This method specifically targets sequence hybridization to the template via a polydeoxyinosine linker (Hwang et al., 2003). In the different progression stages of the stress response, the expression patterns of different gene pairs correlated either negatively or positively, or did not correlate at all. In functional genomics, the increase of data has aided gene expression and protein-protein interactions studies using comparative approaches from basic sequence analysis (Jensen et al., 2006). Bioinformatics approaches are able to identify the function of ozone response

groups of genes. Model species such as *Arabidopsis thaliana* and *Oryza sativa* have large amounts of annotation data. It is now possible to define the gene function by comparing with other model species and crop species. In addition, an orthologous gene enables studies of sequence evolution and gene function. In biological responses, conserved genes and expression levels between species are shown to play a key role (Lu et al., 2009). In this study, our aim was to isolate and characterize differentially expressed genes in soybean leaves treated with O_3. To identify the up-regulated and down-regulated DEGs in the soybean leaves, ACP-based DD-PCR was used. We categorized the putative functions of the DEGs using gene ontology (Ashburner et al., 2000) and the Swiss-Prot IPROClass databases and profiled the functional analysis with Arabidopsis orthologous gene pair.

Results and discussions

Plant materials and DEGs sequence analysis

In soybean cultivars of Korean, Jinpumkong was reported sensitive to response of ozone with physiological characters by Lee et al. (2006). In 200 ppb concentration, it was appeared intense ozone injury than 120 ppb. Also the recovery of leaves was delayed by increasing of ozone fumigation periods. It was shown similarly with previous study in the phenotypic character. To identify the genes specifically or predominantly expressed in ozone treated soybean leaves, this study compared the differential expression levels of the mRNA fragments of treated leaves and untreated leaves using differential banding patterns in agarose gels. The differentially expressed gene (DEG) bands were detected in the untreated leaves (control) and ozone treated leaves using a combination of 120 arbitrary primers and two anchored oligo (dT) primers from GeneFishing^{TM} DEG kits (Seegene, Korea). A total of 408 DEGs were expressed in the ozone treated soybean leaves (cv. Jinpumkong). There were 153 up-regulated DEGs and 225 down-regulated DEGs. The PCR was performed twice in positive case. (Fig. 1). For sequence analysis, the EST sequence was decoded to FASTA format by the Phred program and filtered to remove contaminants such as the vector and poor quality sequence. Directly, the contig and singlet were assembled using the ESTs cluster by the CAP3 program (Huang and Madan, 1999). The sequence homologies were determined by BLASTN of the NCBI nr database and the non-significant sequences were removed at a cut-off e value of 10^{-4}. The nucleotide sequences of 382 DEGs in the ozone fumigated treatments were analyzed. We constructed the total 213 contigs including 65 up-regulated contigs and 148 down-regulated contigs and made a list of the total 186 best matched sequences that included 60 up-regulated ESTs and 126 down-regulated ESTs .

Functional enrichment analysis, pathway analysis and domain search

To identify functional modules of DEGs and protein domains that were significant, we used the GO and Swiss-Prot IPROClass databases (Table 1). The terms of twenty-three proteins (10 up-regulation and 13 down-regulation) were specifically associated with *P*-value thresholds of 0.01 and included the protein domains of thirty-two candidates. In up-regulation, the number of domains ranged from 1 to 14 and the percentages of associated domain ranged from 15.5 to 91.9. The number of domains and the percentage of associated domains for down-regulation ranged from 1 to 11 and 17.2 to 49.9, respectively. These analyses indicated main domain term that shown 11 in up regulation and 6 in down regulation, respectively. The Pfam domains, PF00005 (ATP-binding domain of ABC transporters), PF00025 (ADP-ribosylation factor family), PF00067 (Cytochrome P450), PF00071 (Ras family), PF00081 (Iron/manganese superoxide dismutases, alpha-hairpin domain), PF00106 (short chain dehydrogenase), PF00125 (Core histone H2A/H2B/H3/H4), PF02347 (Glycine cleavage system P-protein), PF031712 (OG-Fe(II) oxygenase superfamily), PF07992 (Pyridine nucleotide-disulphide oxidoreductase), PF12338 (Ribulose-1,5-bisphosphate carboxylase small subunit) were observed amongst the up-regulated genes. The other way, the down-regulated protein domains were indicated as PF0007 (6RNA recognition motif), PF00203 (Ribosomal protein S19), PF00234 (Protease inhibitor/seed storage/LTP family), PF00504 (Chlorophyll A-B binding protein), PF01471 (putative peptidoglycan binding domain), PF04535 (domain of unknown function; DUF588). Previous studies have reported many genes induced by the O_3 response, such as those encoding glutathione *S*-transferase (Glyma15g40190), copper-binding protein (Glyma10g06250), ribulose-1,5-bisphosphate carboxylase (Glyma19g06370), lipoxygenase (Glyma07g00890 and Glyma15g03050), superoxide dismutase (Glyma10g33710) and chlorophyll a/b-binding protein (*cab*; Glyma14g01130) (Maccarrone et al., 1992; Conklin and Last, 1995; Glick et al., 1995; Kubo et al., 1995; Kliebenstein et al., 1998; Miller et al., 1999; Tamaoki et al., 2003a, b), and our results support these observations (Suppl. Table 1). Our results also agree with a study by Bicke et al. (2001) showing glutathione accumulation in response to O_3 treatment. Miller et al. (1999) suggested that O_3 treatment accelerates the normal rate of foliar senescence in *Arabidopsis* plants, which might be explained by the reduction in chlorophyll content, photosynthetic capacity, and efficiency of photosynthesis in leaves exposed to O_3 (Reich, 1983; Held et al., 1991; Nie et al., 1993). Moreover, the loss of Rubisco protein is associated with O_3-induced senescence (Pell and Pearson, 1983; Nie et al., 1993; Pell et al., 1997). O_3 exposure leads to reduced transcript levels of *cab*, small subunit of Rubisco (*rbc*S), and large subunit of Rubisco (*rbc*L) in potato, and reductions in *cab* and *rbc*S transcript levels in *Arabidopsis* and tobacco (Bahl and Kahl, 1995; Conklin and Last, 1995; Glick et al., 1995). These previous findings showed that O_3 treatment can stimulate natural senescence in many plant species. This study supports the findings that O_3 treatment reduces the levels of these transcripts. Kawasaki et al. (2001) reported the up-regulation of many genes encoding glycine-rich proteins, ABA and stress-induced proteins, metallothionein-like proteins, glutathione *S*-transferase, and ascorbate peroxidase in salt-tolerant rice. Seki et al. (2002) identified cold-, drought-, and salinity-induced target genes as well as stress-related transcription factor family members such as helix-loop-helix using *Arabidopsis* cDNAs. These results suggest cross-talk between the salt and drought stress signaling processes. Helix-loop-helix (Basic helix-loop-helix protein; bhlh; Glyma03g29710) was down-regulated in response to ozone treatment, inferring cross-talk between ozone stress and other stress signaling processes. As shown supplemental Table 1, Plastidic (AldP; Glyma14g01470), ribulose-1,5-bisphosphate carboxylase small subunit (Glyma19g06370) and chlorophyll a/b binding protein (Glyma18g03220) were regulated by the ozone response. Plastidic catalyzes the cleavage of plastidic fructose-bisphosphate (FBP; Glyma14g01470) into D-glyceraldehyde-3-phosphate (GAP) and dihydroxyacetone phosphate (DHAP). Two forms of are present in higher plants: cytoplasmic and plastidic. The cytoplasmic (AldC) gene is inducible under anaerobic conditions and is considered to have an important role in producing ATP by stimulating glycolysis under such conditions (Andrews et al., 1994; Umeda and Uchimiya, 1994). AldP is involved in the photosynthetic carbon reduction cycle and catalyzes the synthesis of FBP from GAP and DHAP. AldP gene expression is regulated by light

Table 1. Functional annotation results of protein domains showing significant regulation upon ozone treatment.

Regulation	Terms[†]	*P*-Value[‡]	Pfam ID	No. of Domains	% Associated Domains	Benjamini[§]
Up regulated	nucleotide-binding	4.40E-283	PF00009, PF00091, PF00560, PF00005, PF00004, PF00069, PF07714, PF00702, PF02359, PF00025, PF00012, PF03106, PF00071, PF00403	14	42.7	3.69E-281
	atp-binding	1.85E-210	PF00069, PF00702, PF07714, PF02359, PF00005, PF00560, PF00004, PF03106, PF00012, PF00403	10	37.0	3.89E-209
	oxidoreductase	4.14E-49	PF12338, PF03171, PF02347, PF00067, PF01477, PF07992, PF00081, PF00106	8	16.2	4.34E-48
	nucleus	2.81E-09	PF00125, PF00249, PF00808, PF02309, PF03106	5	15.5	1.31E-08
	nucleotide-binding	0	PF00069, PF00005, PF00025, PF00071	4	91.9	0
	serine/threonine-protein kinase	1.32E-229	PF00069, PF07714, PF00560	3	20.9	3.71E-228
	leucine-rich repeat	6.59E-249	PF00069, PF07714, PF00560	3	15.9	2.77E-247
	atp-binding	0	PF00069, PF00005	2	78.3	0
	serine/threonine-protein kinase	0	PF00069	1	67.4	0
	kinase	1.47E-122	PF00069	1	16.1	2.64E-121
Down regulated	nucleotide-binding	7.38E-292	PF00069, PF00702, PF07714, PF00009, PF00091, PF02359, PF00560, PF00004, PF03106, PF00012, PF00403	11	49.9	2.29E-290
	atp-binding	2.21E-287	PF00069, PF00702, PF07714, PF02359, PF00560, PF00004, PF03106, PF00012, PF00403	9	46.7	4.56E-286
	nucleus	7.88E-34	PF00249, PF00808, PF02309, PF03106	4	20.2	1.22E-32
	metal-binding	1.55E-05	PF00076, PF01471, PF01477	3	31.8	5.41E-05
	serine/threonine-protein kinase	0	PF00069, PF07714, PF00560	3	29.9	0
	leucine-rich repeat	0	PF00069, PF07714, PF00560	3	22.7	0
	zinc	3.57E-05	PF00076, PF01471	2	21.0	1.13E-04
	chloroplast	1.48E-23	PF00203, PF00504	2	18.5	1.30E-22
	rna-binding	6.69E-32	PF00203, PF00076	2	17.2	1.17E-30
	membrane	1.19E-06	PF04535	1	20.4	4.61E-06
	transmembrane	4.90E-05	PF04535	1	20.4	1.43E-04
	lipid-binding	1.77E-47	PF00234	1	18.5	6.20E-46
	transport	3.42E-09	PF00234	1	18.5	1.50E-08

[†]A group of terms having similar biological meaning due to sharing similar domain members in Swiss-Prot IPROClass database. [‡]The *p*-value associated each annotation terms inside, a modified one-tail Fisger Exact Probailty value for gene-enrichment analysis. Since *p*-value <= 0.01, our gene list is specifically associated in term. [§]A Benjamini statistics is used to adjustment for *p*-value correction.

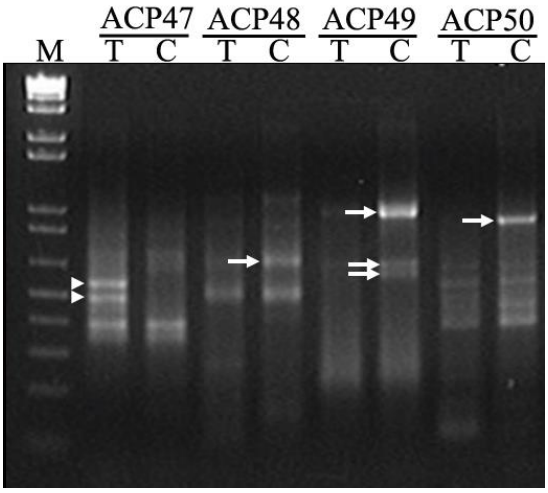

Fig 1. The differential expressed genes (DEGs) using the annealing control primer (ACP) after the O3 treatment. ACP (47; 48; 49; 50): ACP series number, M: 1kb plus DNA ladder, T: O3 treated leaves of soybean, C: control, arrow: down regulated DEGs, arrow head: up regulated DEGs.

(Kagaya et al., 1995). The isolation of two AldP genes from a salt-tolerant *Nicotiana* species, *N. paniculata*, and their expression regulation among three *Nicotiana* species was compared, showing different salt-tolerance abilities under salt stress (Yamada et al., 2000).

Gene functional analysis based on cross-species

For cross-species analysis, the orthologous genes among the regulated DEGs by ozone treatment were compared with Arabidopsis (Table 2). As shown in table 2, a total of 186 DEGs including 60 up-regulated DEGs and 126 down-regulated DEGs were identified. Compared to Arabidopsis, we found 78 orthologous gene pairs that included 24 up-regulated and 54 down-regulated genes. The orthologous gene pairs are listed in table 3. Global dN/dS values were calculated for all pairwise combinations of gene family members in Arabidopsis and soybean using the method of PAML (Yang, 1997). We calculated the dN/dS ratios between the regulated DEGs by ozone treatment in soybean and Arabidopsis. As calculated by pairwise comparisons within each regulation, up-regulated DEGs revealed dN/dS ratios of less than 1.000 (from 0.019 to 0.375) and down-regulated DEGs revealed dN/dS ratios of less than 1.000 (from 0.025 to 0.928). In order to classify orthologous pairs of DEGs into pathway categories, we assigned each gene to major GO categories representing its primary function, based on GO terms restriction level of over 6. We also mapped each GO category to a KEGG pathway to identify the biosynthetic pathways related to protein and mRNA metabolism as well as several alkaloid related terms, suggesting that the concordance on ozone stress condition and cross species was at a functional level. The pathway analysis of orthologous DEGs was performed to identify significantly overrepresented pathways. Our analysis determined glycolysis/gluconeogenesis, biosynthesis of terpenoids and steroids, carbon fixation in photosynthetic organisms and biosynthesis of alkaloids as the most significant pathways involved. Our approach of pathway analysis indicated statistically significant overrepresentation of genes associated with response to stress as radiation, light stimulus and photosystem. The GO analysis was performed on the identified Arabidopsis orthologous DEGs using GO database via Kappa statistics. GO consists of hierarchically structured vocabularies that gene description in terms of their associated biological processes, molecular functions and cellular components. The most of genes were belong to multiple nodes in each hierarchy. To identify this given DEGs, overrepresentation of GO categories in subgaph of a biological network was need. The Cytoscape plugin ClueGO was applied to identify six significantly overrepresented enrichment network pathways. This network result for a set of DEGs as visualize in Cytoscape. Due to the interdependency of functional categories in the GO hierarchy that included the tetraterpenoid metabolic process and gibberellin metabolic process (A), glucose catabolic process (B), transcription from plastid promoter (C), protein import into nucleus, docking (D), glycine decarboxylation via glycin cleavage system (E) and ATP biosynthetic process (F) in GO biological process (Fig. 2). In DEG gene groups, genes in one of the major corresponding pathways were mostly distributed in the ATP biosynthetic process. One out of six represented pathways was related with nitrogen compound metabolic process (Fig. 2C) and five represented pathways were related with cellular process in DEG gene groups. In addition, we found that tetraterpenoid metabolic process and gibbrellin metabolic process shared three biological processes (Fig. 2A). These were the carotenoid biosynthetic process, tetraterpenoid biosynthetic process and terpenoid biosynthetic process.

Materials and methods

Plant materials and ozone treatment

Soybean (*Glycine max*; cv. Jinpumkong and Chengjakong) plants were grown in soil until the fifth trifoliolat stage in a controlled environment greenhouse (28/25°C) under a 16 h photoperiod. Plants were placed on pot each 25 cm in diameter. For ozone fumigations, a loaf design greenhouse, 182 X 304 X 121 cm high along the center beam, approximately 6 m^3, was constructed and covered with 0.008-in polyvinylchloride glazing. Ozone was generated by passing high-grade oxygen through the OZ portable ozone generator (Fischer Industries, Germany). For RNA extraction, plants were fumigated with 120 ppb and 200 ppb ozone for 4 h, on each of 5 consecutive days (4 h with added ozone and 20 h without added ozone per day) in an airtight greenhouse, respectively. Control plants were exposed to ambient air in a similar condition. After exposure, the 3rd, 4th and 5th leaves were harvested, immediately frozen in liquid nitrogen and stored at -80°C till further use.

Total RNA extraction and first-strand cDNA preparation

The extraction of total RNA was carried out using the Trizol kit (Invitrogen). RNA quantification and integrity were double checked by electrophoresis in a 1.2% agarose gel and spectrophotometer. DNA contamination was removed from RNA using DNase (Qiagen, Valencia, CA, USA) treatment. Total RNA was used in the synthesis of first-strand cDNA by reverse transcriptase, as described by Hwang et al. (2003). First-strand cDNA synthesis was performed using the poly(dI) linker between the 3′ and 5′ end sequences of a primer, which generates a region with lower melting temperature (Tm) via the formation of a bubble-like structure at specific temperatures. The total RNA (3 µg) was incubated with 2 µl of dT-ACP1 ($CTGTGAATGCTGCGACTACGATIIIII(T)_{18}$;10 µM) at 80°C for 3 min. Subsequently, reverse transcription was performed for 1.5 h at 42°C in 20 µl of final reaction buffer containing purified total RNA, 4 µl of 5x first strand buffer (Invitrogen), 5 µl dNTPs (each 2 mM), 0.5 µl of RNasin Plus RNase Inhibitor (40 U/µl, Promega), and 1 µl of SuperScript™ II RNase H-Reverse Transcriptase (200 U/µl, Invitrogen). The first-strand cDNA samples were diluted by the addition of 80 µl of DNase-free water.

Table 2. Summary of Arabidopsis orthologous gene pairs.

Regulation	Soybean†	Arabidopsis†	e-value‡	dN§	dS§	dN/dS	Description
Up regulation	Glyma03g39200	AT1G10350	0	0.73	3.77	0.20	dnaj heat-shock protein
	Glyma04g01020	AT2G21330	0	0.16	1.35	0.12	fructose bisphosphate aldolase
	Glyma05g00780	AT1G69410	0	0.08	1.31	0.06	translation initiation factor eif-5a
	Glyma05g36770	AT4G10925	1.00E-176	0.18	1.89	0.10	f-box protein skip8
	Glyma06g01310	AT2G21580	0	0.05	1.48	0.03	40s ribosomal protein s25-2
	Glyma06g07400	AT5G20010	0	0.04	2.05	0.02	gtp-binding nuclear protein ran-3-like
	Glyma06g11970	AT1G77450	1.00E-79	0.21	2.47	0.09	nac domain protein nac2
	Glyma06g16600	AT5G08720	2.00E-150	0.30	2.56	0.12	uncharacterized protein
	Glyma08g08770	AT1G29910	0	0.09	1.22	0.07	photosystem ii type i chlorophyll a b binding protein
	Glyma08g11480	AT3G23810	0	0.07	1.30	0.05	s-adenosyl-l-homocysteine hydrolase
	Glyma08g22600	AT1G52570	0	0.11	2.61	0.04	phospholipase d
	Glyma11g01520	AT3G03150	0	0.44	N/A	N/A	seed specific protein bn15d1b
	Glyma11g14250	AT1G07400	0	0.60	1.67	0.36	small heat shock protein
	Glyma13g35950	AT4G27280	1.00E-133	0.29	N/A	N/A	protein
	Glyma13g44080	AT1G80180	5.00E-145	0.39	N/A	N/A	uncharacterized protein
	Glyma14g14360	AT4G32272	4.00E-177	0.18	1.29	0.14	udp-sugar transporter-like protein
	Glyma14g39950	AT1G76070	0	0.49	1.31	0.38	uncharacterized protein
	Glyma15g40190	AT1G78320	0	0.29	N/A	N/A	glutathione transferase
	Glyma17g11130	AT1G69410	0	0.07	1.19	0.06	translation initiation factor eif-5a
	Glyma18g01860	AT4G40030	0	0.03	1.35	0.02	histone h3
	Glyma18g02610	AT3G10985	0	0.27	1.05	0.26	senescence associated protein 20
	Glyma18g04080	AT2G39730	0	0.14	1.04	0.14	ribulose bisphosphate carboxylase oxygenase activase
	Glyma19g37100	AT2G36750	0	0.37	2.23	0.17	glucosyl transferase
	Glyma19g40690	AT1G02430	7.00E-160	0.41	1.48	0.28	adp-ribosylation factor 1
Down regulation	Glyma01g06030	AT1G69530	0	0.11	N/A	N/A	expansin a1
	Glyma01g35910	AT4G19170	1E-47	0.18	4.19	0.04	carotenoid cleavage dioxygenase 4
	Glyma02g16000	AT2G30570	0	0.26	1.63	0.16	photosystem ii reaction center w chloroplast
	Glyma02g46970	AT5G47640	0	0.27	N/A	N/A	nuclear transcription factor y subunit b-3
	Glyma03g29710	AT2G42300	0	0.39	2.63	0.15	transcription factor bhlh60
	Glyma03g31580	AT2G30570	2E-48	0.35	1.26	0.28	photosystem ii reaction center w chloroplast
	Glyma03g33990	AT3G09840	0	0.05	2.14	0.03	cell division cycle protein 48 homolog
	Glyma03g34310	AT2G36830	0	0.10	1.04	0.09	aquaporin tip1-1
	Glyma04g01020	AT2G21330	0	0.16	1.35	0.12	fructose bisphosphate aldolase
	Glyma04g09550	AT5G65670	0	0.32	1.54	0.21	indole-3-acetic acid inducible 9
	Glyma05g04210	AT1G62750	0	0.10	2.10	0.05	elongation factor chloroplastic-like
	Glyma05g25810	AT1G29910	0	0.08	1.25	0.06	photosystem ii type i chlorophyll a b binding protein
	Glyma05g31450	AT4G39800	0	0.08	2.18	0.03	inositol-3-phosphate synthase
	Glyma06g01310	AT2G21580	0	0.05	1.48	0.03	40s ribosomal protein s25-2
	Glyma06g07200	AT2G25490	0	0.32	1.97	0.16	f-box protein
	Glyma06g08030	AT3G15070	0	1.64	1.77	0.93	c3hc4-type ring finger protein
	Glyma06g12250	AT1G21670	0	0.38	1.78	0.22	uncharacterized protein
	Glyma06g12720	AT1G43850	6E-179	0.22	1.87	0.12	seuss transcriptional co-regulator
	Glyma06g16600	AT5G08720	3E-140	0.30	2.56	0.12	uncharacterized protein
	Glyma07g36870	AT3G17040	5E-115	0.30	2.73	0.11	protein high chlorophyll fluorescent 107
	Glyma07g38090	AT4G15620	0	0.50	1.61	0.31	uncharacterized protein
	Glyma08g14280	AT4G39970	0	0.27	2.56	0.11	protein chromosomal-like
	Glyma08g18800	AT1G78390	2E-125	0.43	2.96	0.14	9-cis-epoxycarotenoid dioxygenase
	Glyma08g40850	AT1G14410	0	0.27	2.32	0.12	ssdna-binding transcriptional regulator
	Glyma09g00820	AT1G62300	0	0.34	N/A	N/A	wrky dna-binding protein 31
	Glyma09g07590	AT5G57030	0	0.18	1.97	0.09	lycopene epsilon cyclase
	Glyma10g06250	AT2G36950	0	0.42	N/A	N/A	heavy metal transport detoxification domain-containing protein
	Glyma10g41300	AT2G01100	0	0.54	2.06	0.26	uncharacterized protein
	Glyma11g01520	AT3G03150	0	0.44	N/A	N/A	seed specific protein bn15d1b
	Glyma11g04900	AT3G49750	0	0.26	2.27	0.11	receptor like protein 44
	Glyma11g07160	AT3G50685	0	0.24	N/A	N/A	uncharacterized protein
	Glyma11g34230	AT2G39730	0	0.14	1.06	0.13	ribulose bisphosphate carboxylase oxygenase activase
	Glyma11g35500	AT2G30130	0	0.66	N/A	N/A	lob domain-containing protein 12-like
	Glyma12g06920	AT2G28800	0	0.20	1.92	0.10	inner membrane protein albino3
	Glyma12g07160	AT3G55280	0	0.11	2.65	0.04	60s ribosomal protein l23a
	Glyma12g08060	AT2G28400	0	0.46	N/A	N/A	uncharacterized protein

Glyma12g16560	AT5G54680	1E-170	0.20	1.50	0.13	transcription factor ilr3
Glyma13g20830	AT2G37220	0	0.32	1.55	0.20	rna-binding protein
Glyma13g37010	AT2G20570	0	0.42	N/A	N/A	gbf s pro-rich region-interacting factor 1
Glyma14g01130	AT2G05070	0	0.10	1.58	0.06	chlorophyll a b binding protein
Glyma14g01470	AT4G26520	0	0.15	2.79	0.05	fructose-bisphosphate aldolase
Glyma14g36140	AT4G24480	0	0.32	1.48	0.22	serine threonine-protein kinase ctr1-like
Glyma15g11490	AT4G04640	0	0.10	3.29	0.03	atp synthase gamma chain
Glyma15g41540	AT1G79550	0	0.12	1.65	0.07	phosphoglycerate kinase
Glyma17g01800	AT1G11650	0	0.25	2.36	0.11	rna-binding post-transcriptional regulator csx1-like
Glyma17g04210	AT1G48030	2E-85	0.11	2.05	0.05	dihydrolipoyl dehydrogenase 1
Glyma17g05500	AT3G14490	0	0.89	N/A	N/A	terpene c1 domain-containing protein
Glyma17g34690	AT2G26080	0	0.11	1.67	0.06	glycine dehydrogenase
Glyma18g04510	AT5G16710	0	0.24	1.68	0.14	dehydroascorbate reductase
Glyma18g44510	AT5G01810	0	0.53	N/A	N/A	serine-threonine kinase
Glyma19g35560	AT3G09440	0	0.06	1.30	0.05	protein heat shock protein 70-3
Glyma19g36130	AT2G37240	0	0.26	1.72	0.15	thioredoxin-like protein
Glyma19g44200	AT3G61610	0	0.17	1.88	0.09	aldose 1-
Glyma20g37940	AT3G05420	0	0.19	1.28	0.15	acyl- binding protein 4

†Soybean and Arabidopsis sequence ID from Phytozome. ‡The BLASTX is used to find the best alignment Soybean genes versus Arabidopsis sequence from Phytozome.
§The methods for estimating dN;nonsynonymous substittion and dS; synonymous substitution were used for finding orthologous.

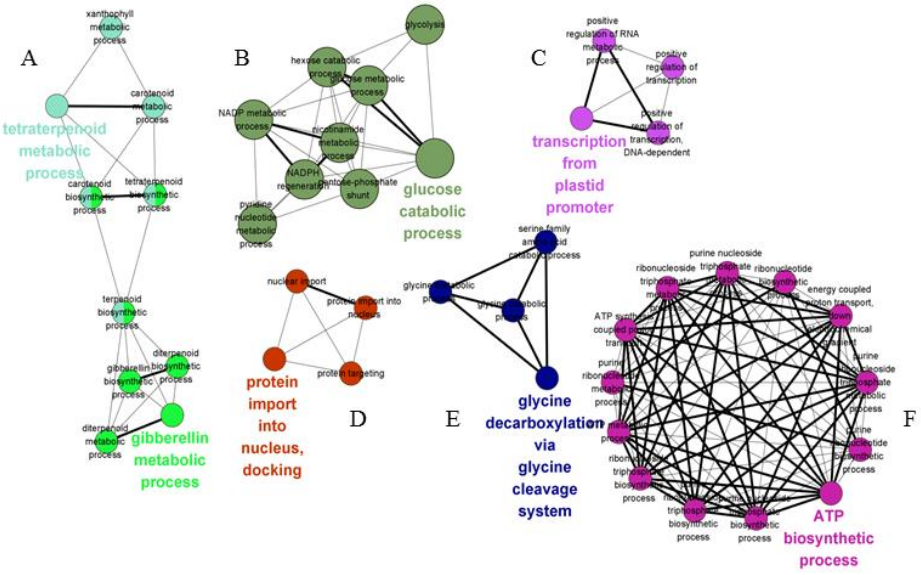

Fig 2. The over-representation analysis by using the Gene Ontology biological process database (The size of node represented the integration of genes and thickness of edge, which showed a significant Kappa value).

ACP-based PCR

The diluted first-strand cDNAs were then subjected to second-strand cDNA synthesis by random PCR amplification using dT-ACP2 (reverse primer; CTGTGAATGCTGCGACTA-CGATIIIII(T)$_{15}$) and one of arbitrary 120 ACPs (forward primer; GeneFishing^{TM} DEG kits, Seegene, Korea) as primers. The second-strand cDNA synthesis was performed in a final reaction volume of 20 µl containing diluted first-strand cDNA (50 ng), 2 µl of 5 µM arbitrary ACP primer, 1 µl of 10 µM dT-ACP2, and 10 µl of 2x SeeAmp^{TM}ACP^{TM} Master Mix (GeneFishing^{TM}DEG kits, Seegene, Korea). The PCR protocol for second-strand synthesis included one cycle at 94°C for 1 min, followed by 50°C for 3 min, and 72°C for 1 min. After completing the second-strand DNA synthesis, 40 cycles were performed. Each cycle involved denaturation at 95°C for 20 sec, annealing at 65°C for 40 sec, extension at 72°C for 40 sec, and a final extension step of 5 min at 72°C to complete the reaction. The amplified products were separated on 2% agarose gels, and then stained with ethidium bromide.

Cloning and soybean DEGs sequence analysis

The differentially expressed transcripts were eluted and cloned into a T&A cloning vector (Real Biotech Corp., Taiwan). The ligation mixture was transformed into competent *Escherichia coli* DH5a cells; plated onto LB agar containing 100 ug mL^{-1} ampicillin, 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) and 80 ug mL^{-1} X-gal; and incubated at 37°C overnight. White colonies that contained the insert were then grown overnight in LB liquid media supplemented with 100 ug mL^{-1} ampicillin. Plasmid DNA was isolated from randomly selected, white colonies by a commercial kit (Invitrogen, USA). To verify the identity of insert DNA, isolated plasmids were sequenced automatically. DNA sequencing was performed by dye-terminator cycle sequencing using the BigDye^{TM} Terminator v3.0 Ready Reaction Cycle Sequencing Kit (Applied Biosystems, USA), followed by ethanol precipitation of the extension products. Both procedures were performed according to the manufacturer's instructions. The M13 primers were used to generate sequences for all isolated cDNAs. The trace files were processed using Phred (score>20), followed by phrap

Table 3. Profile of the orthologous gene pairs.

	No. of genes	Orthologue gene pair vs. Arabidopsis
Up-regulated	60	24
Down-regulated	126	54
Total	186	78

assembly into contigs by clustering a minimum continuous stretch of 100 bp (Ewing and Green, 1998). Assembled sequences were visualized and manually edited using Consed (Gordon et al., 1998). The contigs of the assembled sequences were determine by BLASTX (Altschul et al., 1990) analysis (a cut-off < e-25) following download of Soybean amino acid sequence from Phytozome (Goodstein et al., 2011). The best match results were used as query for BLASTP vs. Arabidopsis. All the EST data are publicly available through the NCBI, with GenBank dbEST accession nos. EG702072 - EG702322. RT-PCR amplifications were conducted in a total volume of 25 uL, containing 1 uL of 10x diluted cDNA, 10x reaction buffer containing 20 mM $MgCl_2$, 2.5 mM of each dNTPs, and 10 pmoles of each gene-specific primer, and 1 U Taq polymerase (RBC). The PCR program was run as follows: 5 min at 94°C, followed by cycles of 30s of denaturation at 94°C, 30 s of annealing and 60s for polymerization at 72°C, with a final extension step at 72°C.

Gene functional analysis based on cross-species

In order to evaluate cross-species analysis of protein-coding regions between orthologous genes dN (nonsynonymous substittion), dS (synonymous substitution), and dN/dS ratio were measured between Soybean vs. Arabidopsis BLASTX match. The dN/dS ratios were calculated by using the PAML package (Yang, 2007) with the Nei-Gojobori method (Nei and Gojobori, 1986). For functional analysis, the application Blast2Go (Conesa et al., 2005) was used to complement existing homologous Arabidopsis gene annotation. Next, gene annotation was further supplemented with BLASTX, conserved InterProScan and Pfam (Finn et al., 2008), as well as function detection of homologous sequences and literature searches. Keywords derived from published reports and the National Center for Biotechnology Information (NCBI) entry data were added to each gene. Using this combined information, a functionally driven classification scheme (based in part on GO terms) was analyzed. We mapped the external classification system to GO. To estimate the enrichment of GO terms, ClueGO was used (Bindea et al., 2009) which evaluates the enrichment of the principle GO categories, including cellular components, biological processes, and molecular functions with hypergeometric distribution and False Discovery Rate (FDR). In order to determine significantly over-represented GO terms, the terms with a *P*-value less than 0.05 were considered Kappa significant value. Genes classified as significantly over-represented were validated by the Bejamini test.

Conclusion

The results showed various functions of the DEGs by the ozone response, and inferred that the DEGs would be affected by numerous factors. This DEG profiling by ozone treatment highlights the possible signaling pathways and molecular biological process for future research. The regulatory pathways that offer such acclimation responses to ozone, as well as the relationships between ozone and other types of abiotic stress, would be expected.

Acknowledgments

This work was carried out with the support of "Cooperative Research Program for Agriculture Science & Technology Development (Project No. PJ907047)", Rural Development Administration, Republic of Korea.

References

Agrawal GK, Rakwal R, Yonekura M, Kubo A, Saji H (2002) Proteome analysis of differentially displayed proteins as a tool for investigating ozone stress in rice (Oryza sativa L.) seedlings. Proteomics. 2: 947-959.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol. 215: 403-410.

Andrews DL, MacAlpine DM, Johnson JR, Kelley PM, Cobb BG, Drew MC (1994) Differential induction of mRNAs for the glycolytic and ethanolic fermentative pathways by hypoxia and anoxia in maize seedlings. Plant Physiol. 106: 1575-1582.

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, et al. (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet. 25: 25-29.

Bahl A, Kahl G (1995) Air pollutant stress changes the steady-state transcript levels of three photosynthesis genes. Environ Pollut. 88: 57–65.

Bicke JA, Setterdahl AT, Knaff DB, Chen Y, Pitcher LH, Zilinskas BA, Leustek T (2001) Regulation of the plant-type 5′-adenylyl sulfate reductase by oxidative stress. Biochemistry. 40:9040-9048.

Bindea G, Mlecnik B, Hackl H, Charoentong P, Tosolini M, Kirilovsky A, Fridman WH, Pages F, Trajanoski Z, Galon J (2009) ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. Bioinformatics. 25: 1091-1093.

Blokhina O, Virolainen E, Fagerstedt KV (2003) Antioxidants, oxidative damage and oxygen deprivation stress: a review. Ann Bot. 91: 179-194.

Conesa A, Gotz S, Garcia-Gomez JM, Terol J, Talon M, Robles M (2005) Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics. 21: 3674-3676.

Conklin PL, Last RL (1995) Differential accumulation of antioxidant mRNAs in *Arabidopsis thaliana* exposed to ozone. Plant Physiol. 109: 203-212.

Ewing B, Green P (1998) Base-calling of automated sequencer traces using phred. II. Error probabilities. Genome Res. 8: 186-194.

Finn RD, Tate J, Mistry J, Coggill PC, Sammut SJ, Hotz HR, Ceric G, Forslund K, Eddy SR, Sonnhammer EL, et al. (2008) The Pfam protein families database. Nucleic Acids Res. 36: D281-288.

Fiscus EL, Reid CD, Miller JE, Heagle AS (1997) Elevated CO_2 reduces O_3 flux and O_3-induced yield losses in soybeans: possible implications for elevated CO_2 studies. J Exp Bot. 48: 307-313.

Glick RE, Schlagnhaufer CD, Arteca RN, Pell EJ (1995) Ozone-induced ethylene emission accelerates the loss of ribulose-1,5-bisphosphate carboxylase/oxygenase and nuclear-encoded mRNAs in senescing potato leaves. Plant Physiol. 109: 891-898.

Goodstein DM, Shu S, Howson R, Neupane R, Hayes RD, Fazo J, Mitros T, Dirks W, Hellsten U, Putnam N, Rokhsar DS (2011) Phytozome: a comparative platform for green plant genomics. Nucleic Acids Res. 40:D1178-D1186.

Gordon D, Abajian C, Green P (1998) Consed: a graphical tool for sequence finishing. Genome Res. 8: 195-202.

Greenberg JT, Guo A, Klessig DF, Ausubel FM (1994) Programmed cell death in plants: a pathogen-triggered response activated coordinately with multiple defense functions. Cell. 77: 551-563.

Held AA, Mooney HA, Gorham JN (1991) Acclimation to ozone stress in radish: leaf demography and photosynthesis. New Phytol. 118: 417–423.
Huang X, Madan A (1999) CAP3: A DNA sequence assembly program. Genome Res. 9: 868-877.
Hwang IT, Kim YJ, Kim SH, Kwak CI, Gu YY, Chun JY (2003) Annealing control primer system for improving specificity of PCR amplification. BioTechniques. 35: 1180-1184.
Hwang K-C, Cui X-S, Park S-P, Shin M-R, Park S-Y, Kim E-Y, Kim N-H (2004) Identification of differentially regulated genes in bovine blastocysts using an annealing control primer system. Mol Reprod Dev. 69: 43-51.
Jensen LJ, Jensen TS, de Lichtenberg U, Brunak S, Bork P (2006) Co-evolution of transcriptional and post-translational cell-cycle regulation. Nature. 443: 594-597.
Kagaya Y, Nakamura H, Ejiri S-i, Tsutsumi K-i, Hidaka S (1995) The promoter from the rice nuclear gene encoding chloroplast aldolase confers mesophyll-specific and light-regulated expression in transgenic tobacco. Mol Gen Genet. 248: 668-674.
Kawasaki S, Borchert C, Deyholos M, Wang H, Brazille S, Kawai K, Galbraith D, Bohnert HJ (2001) Gene expression profiles during the initial phase of salt stress in rice. Plant Cell. 13: 889-905.
Kliebenstein DJ, Monde R-A, Last RL (1998) Superoxide dismutase in Arabidopsis: an eclectic enzyme family with disparate regulation and protein localization. Plant Physiol. 118: 637-650.
Kubo A, Saji H, Tanaka K, Kondo N (1995) Expression of arabidopsis cytosolic ascorbate peroxidase gene in response to ozone or sulfur dioxide. Plant Mol Biol. 29: 479-489.
Laisk A, Kull O, Moldau H (1989) Ozone concentration in leaf intercellular air spaces is close to zero. Plant Physiol. 90:1163-1163.
Lee J-E, Koo B-C, Kim Y-G, Park H-K, Kwon Y-U, Lee B-M (2006) Responses of physiological to ozone stress in soybean. Korean J Crop Sci. 51: 310-311.
Lu Y, Huggins P, Bar-Joseph Z (2009) Cross species analysis of microarray expression data. Bioinformatics. 25: 1476-1483.
Ludwikow A, Gallois P, Sadowski J (2004) Ozone-induced oxidative stress response in Arabidopsis: transcription profiling by microarray approach. Cell Mol Biol Lett. 9: 829-842.
Maccarrone M, Veldink GA, Vliegenthart JF (1992) Thermal injury and ozone stress affect soybean lipoxygenases expression. FEBS Lett. 309: 225-230.
Matyssek R. Wieser G. Calfapietra C. de Vries W. Dizengremel P. Ernst D. Joliver Y. Mikkelsen TN. Mohren GMJ. Le Thiec D. Tuovinen J-P. Weatherall A. Paoletti E (2012) Forests under climate change and air pollution: gaps in understanding and future directions for research. Environ Pollut. 160:57-65.
McClung CR (1997) Regulation of catalases in Arabidopsis. Free Radical Bio Med. 23: 489-496.
Miller JD, Arteca RN, Pell EJ (1999) Senescence-associated gene expression during ozone-induced leaf senescence in Arabidopsis. Plant Physiol. 120: 1015-1024.
Mittler R (2002) Oxidative stress, antioxidants and stress tolerance. Trends Plant Sci. 7: 405-410.
Mittler R. Vanderauwera S. Gollery M. Van Breusegem F (2004) Reactive oxygen gene network of plants. Trends Plant Sci. 9:490-498.
Morgan PB, Ainsworth EA, Long SP (2003) How does elevated ozone impact soybean? A meta-analysis of photosynthesis, growth and yield. Plant Cell Environ. 26: 1317-1328.
Nei M, Gojobori T (1986) Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Mol Biol Evol. 3: 418-426.
Nie GY, Tomasevic M, Baker NR (1993) Effects of ozone on the photosynthetic apparatus and leaf proteins during leaf development in wheat. Plant Cell Environ. 16: 643–651.
Overmyer K, Brosché M, Kangasjarvi J (2003) Reactive oxygen species and hormonal control of cell death. Trends Plant Sci. 8: 335-342.
Overmyer K. Brosché M. Pellinen R. Kuittinen T. Tuominen H. Ahlfors R. Keinänen M. Saarma M. Scheel D. Kangasjärvi J (2005) Ozone-induced programmed cell death in the Arabidopsis radical-induced cell death1 mutant. Plant Physiol. 137:1092-1104.
Pasqualini S, Piccioni C, Reale L, Ederli L, Della Torre G, Ferranti F (2003) Ozone-induced cell death in tobacco cultivar Bel W3 plants. The role of programmed cell death in lesion formation. Plant Physiol. 133: 1122-1134.
Pell EJ, Pearson NS (1983) Ozone-induced reduction in quantity of ribulose-1,5-bisphosphate carboxylase in alfalfa foliage. Plant Physiol. 73: 185–187.
Pell EJ, Schlagnhaufer CD, Arteca RN (1997) Ozone-induced oxidative stress: mechanisms of action and reaction. Physiol Plantarum. 100: 264–273.
Plöchl M. Lyons T. Ollerenshaw J. Barnes J (2000) Simulating ozone detoxification in the leaf apoplast through the direct reaction with ascorbate. Planta. 210:454-467.
Rao MV, Lee H-i, Davis KR (2002) Ozone-induced ethylene production is dependent on salicylic acid, and both salicylic acid and ethylene act in concert to regulate ozone-induced cell death. Plant J. 32: 447-456.
Reich PB (1983) Effects of low concentrations of O_3 on net photosynthesis, dark respiration, and chlorophyll contents in aging hybrid poplar leaves. Plant Physiol. 73: 291–296.
Seki M, Narusaka M, Ishida J, Nanjo T, Fujita M, Oono Y, Kamiya A, Nakajima M, Enju A, Sakurai T, et al. (2002) Monitoring the expression profiles of 7000 Arabidopsis genes under drought, cold and high-salinity stresses using a full-length cDNA microarray. Plant J. 31: 279-292.
Smirnoff N. Pallanca JE (1996) Ascorbate metabolism in relation to oxidative stress. Biochem Soc T. 24:472-478.
Tamaoki M, Matsuyama T, Kanna M, Nakajima N, Kubo A, Aono M, Saji H (2003a) Differential ozone sensitivity among Arabidopsis accessions and its relevance to ethylene synthesis. Planta. 216: 552-560.
Tamaoki M, Nakajima N, Kubo A, Aono M, Matsuyama T, Saji H (2003b) Transcriptome analysis of O_3-exposed Arabidopsis reveals that multiple signal pathways act mutually antagonistically to induce gene expression. Plant Mol Biol. 53: 443-456.
Tenhaken R, Levine A, Brisson LF, Dixon RA, Lamb C (1995) Function of the oxidative burst in hypersensitive disease resistance. Proc Natl Acad Sci USA. 92: 4158-4163.
Tognolli M, Penel C, Greppin H, Simon P (2002) Analysis and expression of the class III peroxidase large gene family in *Arabidopsis thaliana*. Gene. 288: 129-138.
Umeda M, Uchimiya H (1994) Differential transcript levels of genes associated with glycolysis and alcohol fermentation in rice plants (*Oryza sativa* L.) under submergence stress. Plant Physiol. 106: 1015-1022.
Yamada S, Komori T, Hashimoto A, Kuwata S, Imaseki H, Kubo T (2000) Differential expression of plastidic aldolase genes in Nicotiana plants under salt stress. Plant Sci. 154: 61-69.
Yang Z (1997) PAML: a program package for phylogenetic analysis by maximum likelihood. Comput Appl Biosci. 13: 555-556.
Yang Z (2007) PAML 4: Phylogenetic analysis by maximum likelihood. Mol Biol Evol. 24: 1586-1591.