Journal of Hazardous Materials 219–220 (2012) 57–68

Contents lists available at SciVerse ScienceDirect

Journal of Hazardous Materials

journal homepage: www.elsevier.com/locate/jhazmat

Molecular dissection of atrazine-responsive transcriptome and gene networks in rice by high-throughput sequencing

Jia Jun Zhang[a,b,c,1], Zhao Sheng Zhou[d,1], Jian Bo Song[e], Zhao Pu Liu[d], Hong Yang[a,b,*]

[a] Jiangsu Key Laboratory of Pesticide Science, College of Science, Nanjing Agricultural University, Nanjing 210095, China
[b] Key Laboratory of Monitoring and Management of Crop Diseases and Pest Insects, Ministry of Agriculture, Nanjing Agricultural University, Nanjing 210095, China
[c] Department of Pharmacy, Tongren Polytechnic College, Tongren 554300, China
[d] Jiangsu Province Key Laboratory of Marine Biology, College of Resources and Environmental Sciences, Nanjing Agricultural University, Nanjing 210095, China
[e] College of Life Science, Nanjing Agricultural University, Nanjing 210095, China

ARTICLE INFO

Article history:
Received 20 November 2011
Received in revised form 10 March 2012
Accepted 13 March 2012
Available online 22 March 2012

Keywords:
Oryza sativa
Atrazine
Transcriptome
Deep sequencing
Toxicology

ABSTRACT

The residue of atrazine (a herbicide) has become hazards in environments due to its intensive use. However, its molecular toxicity to on plants and human beings is not fully understood. In this study, we performed high-throughput sequencing of atrazine-exposed rice (*Oryza sativa*) to analyze global expression and complexity of genes in the crop. Four libraries were constructed from shoots and roots with or without atrazine exposure. We sequenced 5,751,861, 5,790,013, 5,375,999 and 6,039,618 clean tags that corresponded to 220,806, 111,301, 248,802 and 114,338 distinct tags for Root-Atr (root control, atrazine-free), Shoot-Atr (shoot control, atrazine-free), Root + Atr (root treated with atrazine) and Shoot + Atr (shoot treated with atrazine) libraries, respectively. Mapping the clean tags to gene databases generated 18,833–21,007 annotated genes for each library. Most of annotated genes were differentially expressed among the libraries. The most 40 differentially expressed genes were associated with resistance to environmental stress, degradation of xenobiotics and molecular metabolism. Validation of gene expression by quantitative RT-PCR confirmed the deep-sequencing results. The transcriptome sequences were further subjected to Gene Orthology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis and showed modified biological functions and metabolic pathways. Our results not only highlight the transcriptional complexity in rice with atrazine but also represent a major improvement for analyzing transcriptional changes on a large scale in xenobiotics-responsive toxicology.

1. Introduction

The development of human civilization is at the price of environmental contamination. An intensive agriculture over the last decades has led to dramatically elevated input of herbicides (or pesticides) into arable soils [1]. Herbicides are indispensable for modern agriculture in most of countries. They are designed to kill weeds through specific mechanisms, but not specific to their targets. Moreover, when practically used, herbicides are also accumulated in soils where crops are growing. The accumulated herbicides in crops not only exert detrimental effects on crop itself, but also are very harmful to its ecosystems. Due to their emission into environments, toxic herbicides have become global environmental problems. Recent studies have shown that herbicides applied to soils are frequently detected as pollutants in lakes, rivers and even coastal marine waters [1–7]. The mobility of herbicides into groundwater via soil media has become one of the primary approaches leading to the widespread contamination to ecosystems [8,9]. The massive hazardous materials not only affect the quality of crops which directly accumulate herbicides, but also serve as a food chain threatening human health. As they are readily uptaken by crops [10–13], overload of herbicides into crops is most likely to disrupt many biological processes. Thus, it is of great importance to dissect the toxicological and adaptive response to herbicides.

Atrazine (2-chloro-4-ethylamine-6-isopropylamino-S-triazine) is a herbicide widely used for killing broadleaf weeds or selective grasses. Currently, commercial usage of atrazine has been cancelled in the European Union and United States due to its persistent nature in ecosystems and potential genotoxicity to human beings [14]. But, it is still utilized in developing countries. Compared to other species, the toxicity of atrazine may not be so high but its persistence in ecosystems may result in high incidence of cancer [15–17]. In higher plants, atrazine interacts with proteins of Photosystem II at the plastoquinone B binding site and blocks photosynthetic electron transport to generate reactive oxygen species, which consequently damage the photosynthetic

* Corresponding author at: Department of Applied Chemistry, College of Science, Nanjing Agricultural University, Nanjing, 210095, China. Tel.: +86 25 84395207.
E-mail address: hongyang@njau.edu.cn (H. Yang).
[1] The authors contributed equally to the work.

0304-3894/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.jhazmat.2012.03.041

machinery [18]. The gramineous crop is the major target, such as rice that may take up a great deal of atrazine from paddy soils [19,20]. Plants have developed sophisticated strategies to minimize the adverse impact from organic xenobiotics. Several mechanisms for catabolism and detoxicification of herbicides have been proposed [21]. For instance, cytochrome P450 monooxygenases (P450s) and glutathione S-transferases (GSTs)-mediated catabolic processes have been implicated as a pathway to degrade absorbed pesticides [22]. Recently, a glucosylation-based pesticide degradation pathway has been characterized in plants [23]. Herbicide-triggered gene expression and altered metabolic pathways appear to be a basic model used for studying responses of plants to xenobiotics.

Understanding of gene expression and networks that coordinate adaptive responses to toxic contaminants is the first step to dissect the genetic and molecular basis for plant tolerance to pesticides. Recent studies demonstrate that genome-wide analysis of gene expression has resulted in identification of many biological processes of pesticide accumulation, translocation, and degradation in living things [24,25]. For example, the microarray based analysis of transcriptome provides insights into gene expression and miRNA regulatory pathways under pesticide stress [26]. In addition, more than one hundred genes have been profiled in *Arabidopsis* exposed to five herbicides [27]. Also, using microarray many transcripts were shown to be involved in pathways leading to stress, detoxification, innate immunity, and lipid transport in *C. elegans* with chlorpyrifos and diazinon [28]. These results indicate that genome-wide analysis can help to understand the entire herbicide-responsive molecular events. However, the gene chip technique has become more challenging when dealing with the more inventories of gene species, as such that the genes represented by unspecific probe sets and with low expression levels can not be readily detected. Recently, the deep-sequencing technology has become a powerful tool to permit the concomitant sequencing of millions of signatures and identify specific and enriched genes expressed in a single tissue [29,30]. This approach also highlights the benefits of providing more thorough qualitative and quantitative description of gene expression.

Rice (*Oryza sativa*) is one of the most importantly economical crops because it provides the major portion of calories for human diet in Asia and other parts of the world. It is a model species with completed genomic sequences and subjected to frequent exposure to various herbicides. Also, it is an excellent species for studying the molecular and genetic complexity of pesticide tolerance [20,31]. Like other food crops, rice is easy to accumulate pesticides, and the contamination may in turn pose detrimental effects on human beings [32,33]. With exception of examining phytotoxicity of atrazine, very little is known about the molecular mechanisms involved in regulation of plant response to the herbicide. Toward this end, we present the first genome-wide analysis of transcriptome in rice exposed to atrazine using recently developed high-throughput sequencing technology. The goal of this study is to: (1) establish a platform to characterize gene expression on a larger scale that can facilitate detailed characterization on genes regulating crop toxicological responses to atrazine; (2) quantify the transcript abundance in the species under atrazine exposure; and (3) uncover the networks of genes enriched for regulating rice resistance to the herbicide.

2. Materials and methods

2.1. Plant treatments

Uniform seeds of rice (*O. sativa* L. Japonica) were sterilized with 5% H_2O_2, rinsed thoroughly with distilled water and germinated in darkness on a floating plastic net. After germination, seedlings were hydroponically grown in a growth chamber under the conditions of 24/20 °C (day/night), 300 μmol m^{-2} s^{-1} artificial illumination (14 h photoperiod) and 70% humidity [34]. After growing for 4 d, they were treated with 0 (control) and 0.4 mg L^{-1} atrazine for 2, 4 and 6 d, respectively. When harvested, the roots and shoots were separately collected for following analysis.

2.2. RNA isolation

Total RNA was extracted from shoots and roots using Trizol (Invitrogen, Carlsbad, CA) following the manufacturer's instructions. Four libraries were constructed, including two libraries of Shoot-Atr (shoot control, atrazine-free) and Root-Atr (root control, atrazine-free), which were generated from the RNA pool derived from tissues treated with 0 mg L^{-1} atrazine for 2, 4 and 6 d, and the other two libraries of Shoot + Atr (shoot treated with atrazine) and Root + Atr (root treated with atrazine), which were generated from the RNA pool derived from samples treated with 0.4 mg l^{-1} atrazine for 2, 4 and 6 d. Ten μg RNA samples were incubated for 15 min at room temperature with 1 unit of DNaseI (Takara) to remove residual genomic DNA and incubated with 1 μL of 20 mM EDTA for DNaseI inactivation at 65 °C for 10 min. One % agarose gel (Biowest) buffered by Tris-acetate–EDTA was run to indicate the integrity of RNA. All RNA samples were quantified and examined for protein contamination ($A_{260\,nm}/A_{280\,nm}$ ratios) and reagent contamination ($A_{260\,nm}/A_{230\,nm}$ ratios).

2.3. Library sequencing and data processing

The 3′-tag digital gene expression libraries were prepared using the Illumina Gene Expression Sample Prep Kit [30]. Briefly, 6 μg total RNA was used for mRNA capture with magnetic oligo (dT) beads. The first and second strand cDNA was synthesized, and the stranded bead-bound cDNA was subsequently digested with NlaIII. The 3′-cDNA fragments attached to the oligo(dT) beads were ligated to the Illumina GEX NlaIII Adapter 1, which contained a recognition site for Endonuclease MmeI for cutting 17 bp downstream of the recognition site (CATG) to produce tags with adapter 1. After removing 3′ fragments with magnetic bead precipitation, an Illumina GEX adapter 2 was introduced to the site of MmeI cleavage. The resulting adapter-ligated cDNA tags were amplified using PCR-primers that were annealed to the adaptor ends for 15 cycles. The 85 base fragments were purified and recovered by 6% polyacrylamide Tris-borate–EDTA gel. The final quality of tagged sequences was checked by Agilent 2100 Bioanalyzer. The four constructed tag libraries underwent Illumina proprietary sequencing chip (flowcell) for cluster generation through situ amplification and was deep-sequenced using Illumina Genome Analyzer. Image analysis, base calling and quality calibration were performed using the Solexa Automated Pipeline, after which the raw data (tag sequences and counts) were produced.

Both raw sequence reads and low quality tags (e.g. <21 nt short tags and those sequenced only once) were removed based on the Illumina pipeline. The remaining high quality sequences were mapped to rice cDNA (ftp://plantbiology.msu.edu/pub/data/Eukaryotic_Projects/o_sativa/annotation_dbs/pseudomolecules/version_6.1/all.dir/all.cDNA) by SOAP. Both sense and antisense sequences were included in the data collection for analyzing the mapping events. All expressed tags were mapped onto a preprocessed database of 17 bases-long sequences of rice cDNA located next to NlaIII restriction site, and only one mismatch was allowed. Tags mapped to more than one transcript were excluded from our analysis. When multiple types of tags were aligned to different positions of the same gene, the gene expression levels were represented by the summation of all.

Table 1
Categorization and abundance of tags. Clean tags are tags after filtering dirty tags (low quality tags) from raw data. Distinct tags are different kinds of tags. Unique tags are the remainder clean tags after removing tags mapped to reference sequences from multiple genes.

Summary		Root-Atr	Shoot-Atr	Root + Atr	Shoot + Atr
Raw tag	Total	5,947,872	5,919,794	5,760,941	6,178,993
	Distinct tag	416,479	240,381	633,553	253,194
Clean tag	Total number	5,751,861	5,790,013	5,375,999	6,039,618
	Distinct tag number	220,806	111,301	248,802	114,338
All tag mapping to gene	Total number	3,974,387	5,218,627	2,747,026	5,170,922
	Total % of clean tag	69.10%	90.13%	51.10%	85.62%
	Distinct tag number	91,020	75,008	62,770	73,862
	Distinct tag % of clean tag	41.22%	67.39%	25.23%	64.60%
Unique tag mapping to gene	Total number	3,826,628	5,044,865	2,649,727	4,985,203
	Total % of clean tag	66.53%	87.13%	49.29%	82.54%
	Distinct tag number	87,230	71,826	60,116	70,818
	Distinct tag % of clean tag	39.51%	64.53%	24.16%	61.94%
All tag-mapped genes	Number	24,377	22,753	22,227	22,603
	% of ref genes	42.92%	40.06%	39.13%	39.80%
Unique tag-mapped genes	Number	21,007	18,827	19,358	18,833
	% of ref genes	36.99%	33.15%	34.08%	33.16%
Mapping to genome	Total number	737,878	302,087	405,603	399,700
	Total % of clean tag	12.83%	5.22%	7.54%	6.62%
	Distinct tag number	32,506	20,773	20,759	21,605
	Distinct tag % of clean tag	14.72%	18.66%	8.34%	18.90%
Unknown tag	Total number	1,039,596	269,299	2,223,370	468,996
	Total % of clean tag	18.07%	4.65%	41.36%	7.77%
	Distinct tag number	97,280	15,520	165,273	18,871
	Distinct tag % of clean tag	44.06%	13.94%	66.43%	16.50%

2.4. Quantitative RT-PCR analysis

Total RNA was extracted from tissues using column plant RNAout Kit (Tiandz). The reverse transcription of the total RNA was carried out at 42 °C and the 25 μL reaction mixture containing μg RNA, 0.5 μg oligo (dT) primers, 12.5 nmol dNTPs, 12.5 units of RNase inhibitor and 5 units of M-MLV reverse transcriptase (Takara). The quantitative RT-PCR was performed on a MyiQ Single Color Real-time PCR system (Bio-Rad) in a final volume of 20 μL containing 2 μL of a 1/10 dilution of cDNA in water, 10 μL of the 2 × SYBR Premix Ex Taq (TaKara) and 200 nM of forward and reverse primers (Supplemental Table S3). The thermal cycling conditions were 1 cycle of 95 °C for 30 s for denaturation and 40 cycles of 95 °C for 5 s and 60 °C for 34 s for annealing and extension. All reactions were run in triplicate by monitoring the dissociation curve to the control dimers. PCR efficiency was determined by a series of 2-fold dilutions of cDNA [30]. Ubiquitin (Os03g13170.1) gene was used as a normalizer and relative expression levels of genes were presented by $2^{-\Delta CT}$ (ΔCT is the differences of CT between the control ubiquitin products and the target gene products).

2.5. Statistical evaluation

Each result in this study was the mean of three replicated treatments and each treatment contained at least 15 seedlings. Statistical analysis was performed to identify differentially expressed genes between the libraries using a rigorous algorithm described previously [35]. The statistical *t*-test was used to identify genes expressed between libraries. We used stringent value false discovery rate (FDR) < 0.001 and the absolute value of $\log_2$ ratio ≥ 1 as the threshold to judge the significant difference of gene expression.

3. Results

3.1. Transcriptome profiling of atrazine-free and -treated libraries

We used the Solexa Genome Analyzer to perform high-throughput tag-sequencing (Tag-seq) analysis on poly (A)-enriched RNAs from four libraries derived from atrazine-free and -treated roots and shoots of rice (*O. sativa*). The total number of tags for Root-Atr (atrazine-free), Shoot-Atr (atrazine-free), Root + Atr (treated with atrazine) and Shoot + Atr (treated with atrazine) was 5.94, 5.92, 5.76 and 6.18 million, respectively (Table 1). The number of tags entitled with distinct tags ranged from 0.24 to 0.63 million. After removal of the low quality tags, a total of 5,751,861, 5,790,013, 5,375,999 and 6,039,618 clean tags, that corresponded to 220,806, 111,301, 248,802 and 114,338 distinct tags for Root-Atr, Shoot-Atr, Root + Atr and Shoot + Atr libraries, respectively, were obtained. Analysis of the total and distinct clean tag counts revealed that their distribution was very similar within the four libraries (Fig. 1). Among the distinct tags, less than 8% had the copy

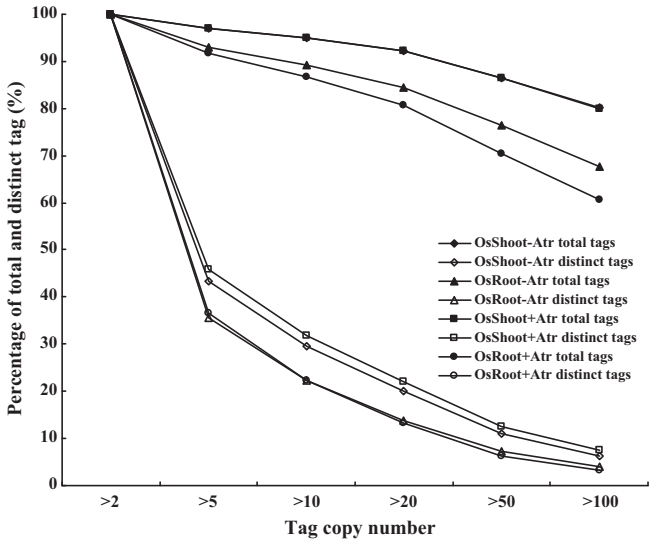

Fig. 1. Distribution of total clean tag (filled) and distinct clean tag (open) counts over different tag dance categories from the four libraries of *Oryza sativa* (Os) with or without atrazine (Atr).

Table 2
The number of sense and antisense transcripts among the four libraries. Sense or antisense gene expression alone or co-expression of both was presented.

	Root-Atr	Shoot-Atr	Root + Atr	Shoot + Atr
Sense	20,145	18,282	18,750	18,334
Sense of ref gene (%)	35.47%	32.19%	33.01%	32.28%
Antisense	14,095	11,285	9928	10,685
Antisense of ref gene (%)	24.82%	19.87%	17.48%	18.81%
Antisense with sense co-expressed	13,233	10,740	9320	10,186
Antisense expressed only	862	545	608	499
Sense with antisense co-expressed (%)	65.69%	58.75%	49.71%	55.56%
Total genes	21,007	18,827	19,358	18,833

number higher than 100 counts and approximate 31.40–38.17% of the tags were present between 5 and 100 copies, while the most abundant distinct tags (54.20–64.52%) were presented between 2 and 5 copies.

Matching the tags to genes is a first step to annotate sequences that illustrates which gene is expressed in response to atrazine. The clean tags were aligned to the reference rice database (ftp://plantbiology.msu.edu/pub/data/Eukaryotic_Projects/o_sativa/annotation_dbs/pseudomolecules/version_6.1/all.dir/all.cDNA). For Root-Atr, Shoot-Atr, Root + Atr and Shoot + Atr libraries, a total of 87,230, 71,826, 60,116 and 70,818 distinct tags can be matched to the reference genes (Table 1). This finally resulted in generation of 21,007, 18,827, 19,358 and 18,833 unique tag-mapped transcripts for Root-Atr, Shoot-Atr, Root + Atr and Shoot + Atr libraries, respectively (Table 1 and Table 2).

The sequencing saturation for the four libraries was analyzed to estimate whether the sequencing depth was sufficient for the transcriptome coverage. It was shown that genes mapped by unambiguous tags increased with the total number of tags. When sequencing counts were 3 M or more than 3 M, the number of detected genes was saturated (Supplemental Fig.S1). We further analyzed distribution of the distinct tag number in each pair of libraries and found that more than 94% distinct tags had a ratio within 5 folds (Supplemental Fig.S2). In this study, we sequenced a large proportion of antisense genes, with 17.5–24.8% being matched by the reference genes for the four libraries (Table 2; Supplemental Table S1). These antisense genes may be expressed coordinately with sense genes.

3.2. Variation of numerous genes expressed in different libraries

We matched the unique tags from each library to the reference gene sequences. Analysis of the four libraries showed that Root-Atr, Shoot-Atr, Root + Atr and Shoot + Atr had 20,145, 18,282, 18,750 and 18,334 sense transcripts, respectively (Table 2). There were 21,007 and 18,827 transcripts detected from control roots (Root-Atr) and shoots (Shoot-Atr). Between the two libraries, 16,679 existed in the both libraries, and 3466 were present only in Root-Atr and 1603 in Shoot-Atr library (Fig. 2A), indicating that under normal condition more transcripts were expressed in roots than in shoots. It is shown that fewer transcripts were found in atrazine-treated roots than the control (atrazine-free roots) (Fig. 2B). By contrast, the shoots had more transcripts under atrazine exposure than those of non-atrazine exposure (Fig. 2C). We also found the change in the number of transcripts between Root + Atr and Shoot + Atr (Fig. 2D). The varied transcripts within the libraries indicate that expression of the genes was altered under atrazine stress.

3.3. Regulation of gene expression by atrazine

Variations of gene expression between two libraries give insights into the molecular events involved in plant tolerance to atrazine toxicity. To understand each gene expression within the libraries, we analyzed the transcript abundance of tag-mapped genes in our datasets by counting the number of transcripts per million (TPM) clean tags. We normalized the read density measurement, and then used false discovery rates (FDRs) < 0.001 and

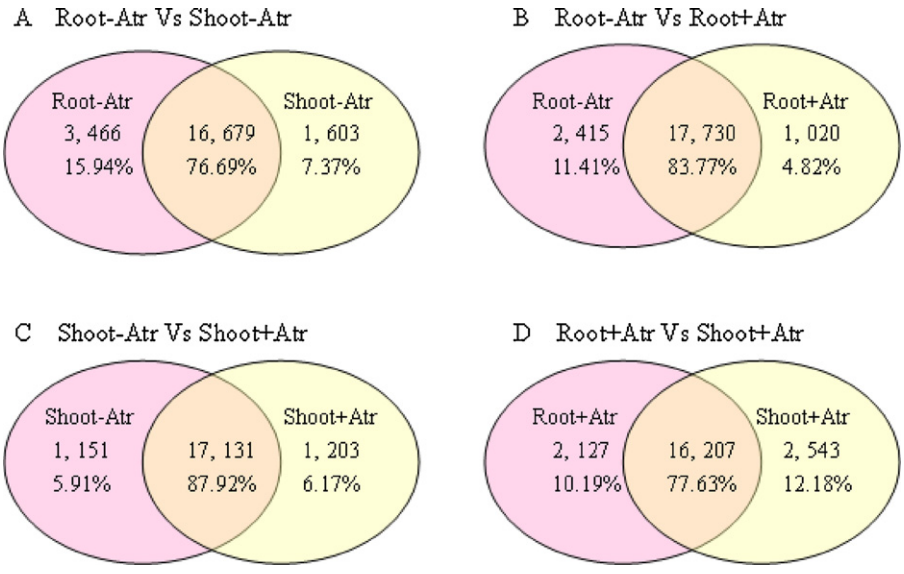

Fig. 2. Analysis of tag-mapped transcripts within the four libraries. Venn diagram of the quantitative genes classification into those specifically expressed in one library or those expressed in both libraries (Atr: atrazine).

Table 3
Differentially expressed genes between libraries. All unique tag-mapped genes were examined for their expression differences between the different libraries. Numbers of differentially expressed genes represent across sense transcripts, using threshold values FDR < 0.001 and $|\log_2 \text{ratio}| \geq 1$ for controlling false discovery rates.

	Root-Atr vs shoot-Atr	Root-Atr vs root + Atr	Shoot-Atr vs shoot + Atr	Root + Atr vs shoot + Atr
Total	9744	6714	3596	8172
Up-regulated	3523	1319	2265	5226
Down-regulated	6221	5395	1331	2946
Expressed only in former	1453	379	44	885
Expressed only in latter	595	128	77	730

the absolute value of $\log_2$ ratio ≥ 1 as a threshold to estimate the statistical significance of gene expression [30]. A number of genes were found to be expressed differentially between two libraries (Supplemental Fig.S3).

There are 6714 genes differently expressed between Root-Atr and Root + Atr libraries (Table 3). Among these 1319 genes were up-regulated and 5395 were down-regulated. Under atrazine stress, 128 genes were expressed only in the Root + Atr library and 379 were only expressed in Root-Atr library, suggesting that treatment with atrazine led to null expression of many genes in roots. In contrast, 77 transcripts were detected only in Shoot + Atr library, whereas 44 transcripts were only found in Shoot-Atr plants. Also, 1331 genes were down-regulated and 2265 were up-regulated in the Shoot-Atr/Shoot + Atr dataset, indicating that more genes were depressed by atrazine exposure.

Because transcripts varied from one library to another, we presented the 20 most abundantly expressed genes in each library. As shown in Supplemental Table S2, most of the genes were differentially expressed in the four libraries. For instance, gene (*Os10g11500*) coding for a SCP-like extracellular protein was expressed substantially in roots but can be hardly expressed in shoots. Transcripts such as *Os01g24710* encoding jacalin-like lectin domain containing protein and *Os04g56430* encoding cysteine-rich receptor-like protein kinase in roots were found to be substantially induced by atrazine exposure. By contrast, gene (*Os08g44680*) coding for a photosystem I reaction center subunit II in roots was depressed by atrazine treatments. However, gene *Os12g43600* coding for a RNA recognition motif-containing protein was found to be present in all datasets and abundantly expressed in all libraries. Several other genes such as *Os05g15770* and *Os11g44810* coding for glycosyl hydrolase and auxin-repressed protein, respectively, were shown in two or three libraries.

We analyzed further 40 genes that were most differentially regulated by atrazine in roots or shoots. The relative abundance was expressed as a TPM ratio of +Atrazine vs −Atrazine responsive transcripts. The top three differentially expressed genes in +Atrazine/−Atrazine roots encode glycosyl hydrolases (*Os05g15880* and *Os11g47530*) and B12D protein (*Os07g41340*) (Table 4). B12D protein is an unidentified protein. In addition, several other enzymes such as aldehyde oxidase (*Os03g58380*) and peroxidase (*Os07g44590*) were strongly in response to atrazine exposure. In shoots *Os07g34520* and *Os04g40990*, that were the top differentially expressed genes, encode isocitrate lyase and malate synthase, respectively. Interestingly, both isocitrate lyase and malate synthase are enzymes that work together in the glyoxylate cycle. Isocitrate lyase catalyzes the cleavage of isocitrate to succinate and glyoxylate, while malate synthase catalyzes the second step of the glyoxylate bypass, the condensation of acetyl coenzyme A and glyoxylate to form malate. In this dataset, we also found several other interesting genes coding for glycine-rich cell wall structural protein (*Os02g37480*), cytochrome P450 (*Os08g39660* and *Os08g39730*) and glutathione S-transferase (*Os10g22310*), all of which appear to involve resistance to xenobiotics. There are some genes for some transcription factors (e.g. bZIP domain-containing protein and HBP-1b). Additionally, several other transcripts for transporters (MDR-like ABC transporter and POT domain containing peptide transporter) were identified to be negatively regulated by atrazine.

3.4. Validation of tag-mapped genes by qRT-PCR

To confirm the tag-mapped genes in rice exposed to atrazine, eight genes were selected at random and subjected to transcriptional validation. All genes could be successfully amplified by quantitative RT-PCR with gene primers listed in Supplemental Table S3. It is shown that expression pattern of all genes by qRT-PCR was in good agreement with that by deep-sequencing based digital gene expression (DGE) (Fig. 3).

3.5. Analysis on GO functional categories and KEGG pathways

To understand whether enrichment of atrazine-responsive genes was associated with their biological significance, we carried out a functional clustering analysis [36]. Most of the annotated genes could be grouped together in Gene Ontology (GO: http://www.geneontology.org) terms. For instance, there were 5612, 13,491, and 6978 genes in Root-Atr/Root + Atr pair, that could be grouped into biological process, cellular component, and molecular function categories (Table 5). Genes that were subjected to GO term analysis in other library pairs were summarized in (Supplemental Tables S4, S5). For Root-Atr/Root + Atr pair, we observed that the top 10 percent subgroups in "Biological Process" were involved in cellular protein metabolic process, gene expression, biological regulation, etc. (Table 5). The most remarkable subgroup was that in response to stimuli. This group of genes in response to biotic or abiotic stresses also can be induced by atrazine. Besides, gene groups related to transport and carbohydrate metabolic process were considerably enriched among the transcripts. By contrast, genes encoding carbohydrate boisynthestic enzymes or proteins were least enriched.

We used Kyoto Encyclopedia of Genes and Genomes (KEGG: http://www.genome.jp/kegg) ontology assignments to classify functional annotations of the identified genes [37]. The KEGG pathway database records networks of molecular interactions in cells, as well as their variants specific to particular organisms, that can help to understand the biological functions of genes. We identified 3399, 1861, 4844, and 4151 differentially expressed genes (DEGs) for the four library pairs (Root-Atr/Root + Atr, Shoot-Atr/Shoot + Atr, Root-Atr/Shoot-Atr, and Root-Atr/Shoot-Atr), and each of them could be further assigned to 28, 18, 35, and 42 pathways, respectively (Table 6; Supplemental Table S6). It is interesting to find that among the assignments, DEGs belonging to "Metabolic pathway" were most abundantly presented in the four library pairs, and each of them comprised 804 (23.65%), 461 (24.77%), 1211 (25.00%) and 1049 (25.27%) genes for Root-Atr/Root + Atr, Shoot-Atr/Shoot + Atr, Root-Atr/Shoot-Atr, and Root + Atr/Shoot + Atr libraries, respectively. The pathway of "Biosynthesis of secondary metabolites" was second dominant in the number of differentially expressed genes (279–774). The other pathways contain only fewer than 5 percent number of DEGs. It is worthy to note that a proportion of genes were grouped into secondary metabolic pathways, such as phenylpropaniod and phenylalanine biosynthesis or metabolism;

Table 4
Top 40 annotated genes most differentially expressed in roots and shoots in the presence of atrazine (+Atr) and absence of atrazine (−Atr). This calculation is based on the expressed tag frequency. $\log_2$ ratio > 0 indicates genes stimulated and $\log_2$ ratio < 0 indicates genes depressed by atrazine stress.

Gene ID	Root + Atr/Root-Atr			
	Root + Atr	Root-Atr	$\log_2$ ratio	Description
LOC_Os05g15880	109.56	0.01	13.42	Glycosyl hydrolase
LOC_Os07g41340	49.48	0.01	12.27	B12D protein
LOC_Os11g47530	43.90	0.01	12.10	Glycosyl hydrolase
LOC_Os10g31420	34.78	0.01	11.76	Retrotransposon protein
LOC_Os07g34520	30.69	0.01	11.58	Isocitrate lyase
LOC_Os06g01590	29.58	0.01	11.53	Lactate/malate dehydrogenase
LOC_Os10g20350	23.81	0.01	11.22	MATE efflux family protein
LOC_Os03g48770	23.62	0.01	11.21	Cupin domain containing protein
LOC_Os03g58380	20.83	0.01	11.02	Aldehyde oxidase 2
LOC_Os08g04210	17.30	0.01	10.76	Cysteine-rich repeat secretory protein 55 precursor
LOC_Os12g13890	16.37	0.01	10.68	Retrotransposon protein
LOC_Os02g29370	15.63	0.01	10.61	Transposon protein
LOC_Os09g37540	15.44	0.01	10.59	Uncharacterized protein PA4923
LOC_Os09g27930	14.14	0.01	10.47	Ubiquitin fusion protein
LOC_Os05g18090	11.53	0.01	10.17	SHR5-receptor-like kinase
LOC_Os06g03970	10.60	0.01	10.05	Receptor-like protein kinase 5 precursor
LOC_Os01g19529	10.04	0.01	9.97	Pentatricopeptide
LOC_Os08g23870	9.86	0.01	9.95	Late embryogenesis abundant group 1
LOC_Os09g29690	8.18	0.01	9.68	Beta-expansin precursor
LOC_Os02g51440	7.63	0.01	9.58	Laccase precursor protein
LOC_Os12g16690	0.01	8.87	−9.79	Zinc finger
LOC_Os06g51190	0.01	8.87	−9.79	Lysine ketoglutarate reductase trans-splicing related 1
LOC_Os10g40470	0.01	9.04	−9.82	LTPL142 – Protease inhibitor/seed storage/LTP family protein precursor
LOC_Os04g54940	0.01	9.21	−9.85	ATEB1A-like microtubule associated protein
LOC_Os02g55410	0.01	9.74	−9.93	MCM5 – Putative minichromosome maintenance MCM complex subunit 5
LOC_Os12g39980	0.01	9.91	−9.95	Kinesin motor domain containing protein
LOC_Os09g29560	0.01	10.43	−10.03	OsWAK83 – OsWAK pseudogene
LOC_Os08g38270	0.01	10.95	−10.10	fasciclin domain containing protein
LOC_Os07g42420	0.01	11.13	−10.12	3-oxoacyl-synthase
LOC_Os03g11210	0.01	12.17	−10.25	DUF538 domain containing protein
LOC_Os11g32510	0.01	14.26	−10.48	OsGH3-13 – Auxin-responsive GH3 gene family member
LOC_Os07g44590	0.01	16.86	−10.72	Peroxidase precursor
LOC_Os08g40720	0.01	18.08	−10.82	FAD-binding and arabino-lactone oxidase domains containing protein
LOC_Os03g16760	0.01	19.30	−10.91	Protein phosphatase 2C
LOC_Os01g50030	0.01	20.34	−10.99	CPuORF25 – conserved peptide uORF-containing transcript
LOC_Os01g65110	0.01	20.69	−11.01	POT family protein
LOC_Os07g15460	0.01	25.56	−11.32	Metal transporter Nramp6
LOC_Os10g31530	0.01	26.77	−11.39	Glycine-rich cell wall structural protein
LOC_Os03g09980	0.01	34.25	−11.74	Sulfate transporter
LOC_Os05g47540	0.01	45.38	−12.15	CPuORF26 – conserved peptide uORF-containing transcript

Gene ID	Shoot + Atr/Shoot-Atr			
	Shoot + Atr	Shoot-Atr	$\log_2$ ratio	Description
LOC_Os07g34520	377.34	0.01	15.20	Isocitrate lyase
LOC_Os04g40990	275.51	0.01	14.75	Malate synthase
LOC_Os01g62260	92.72	0.01	13.18	Thaumatin
LOC_Os01g42860	49.01	0.01	12.26	Inhibitor I family protein
LOC_Os01g03360	24.34	0.01	11.25	BBTI5 – Bowman-Birk type bran trypsin inhibitor precursor
LOC_Os01g55940	19.21	0.01	10.91	OsGH3.2 – Probable indole-3-acetic acid-amido synthetase
LOC_Os01g06560	18.54	0.01	10.86	Transcription factor HBP-1b
LOC_Os07g08240	17.88	0.01	10.80	ZmEBE-1 protein
LOC_Os02g37480	13.91	0.01	10.44	Glycine-rich cell wall structural protein
LOC_Os09g25060	13.41	0.01	10.39	OsWRKY76 – Superfamily of TFs having WRKY and zinc finger domains
LOC_Os12g39400	13.08	0.01	10.35	ZOS12-09 – C2H2 zinc finger protein
LOC_Os07g26110	9.44	0.01	9.88	Membrane associated DUF588 domain containing protein
LOC_Os04g41640	8.78	0.01	9.78	HEV2 – Hevein family protein precursor
LOC_Os11g47560	6.95	0.01	9.44	Glycosyl hydrolase
LOC_Os03g33090	6.29	0.01	9.30	DUF260 domain containing protein
LOC_Os09g34920	6.13	0.01	9.26	Glycosyl hydrolase family 29
LOC_Os08g39660	5.8	0.01	9.18	Cytochrome P450
LOC_Os08g39730	5.8	0.01	9.18	Cytochrome P450
LOC_Os04g53210	5.8	0.01	9.18	Hydroxyacid oxidase 1
LOC_Os11g16940	5.3	0.01	9.05	Transposon protein
LOC_Os03g60190	0.01	2.42	−7.92	Oxidoreductase
LOC_Os11g37060	0.01	2.42	−7.92	OsFBDUF52 – F-box and DUF domain containing protein
LOC_Os05g07000	0.01	2.42	−7.92	Splicing factor
LOC_Os03g08940	0.01	2.42	−7.92	Conserved hypothetical protein
LOC_Os02g33944	0.01	2.59	−8.02	Transposon protein
LOC_Os01g47050	0.01	2.59	−8.02	OsFBK1 – F-box domain and kelch repeat containing protein
LOC_Os01g09580	0.01	2.59	−8.02	CAMK_CAMK_like_Aur_like.1 – CAMK includes calcium/calmodulin depedent protein kinases
LOC_Os01g52160	0.01	2.59	−8.02	Heavy metal-associated domain containing protein

Table 4 (*Continued*)

Gene ID	Shoot + Atr/Shoot-Atr			
	Shoot + Atr	Shoot-Atr	$\log_2$ ratio	Description
LOC_Os01g03180	0.01	2.76	−8.11	Retrotransposon protein
LOC_Os12g06200	0.01	2.94	−8.20	E2F family transcription factor protein
LOC_Os01g12920	0.01	3.11	−8.28	Thioesterase family protein
LOC_Os08g33740	0.01	3.11	−8.28	CSLA11 – cellulose synthase-like family A
LOC_Os01g50080	0.01	3.45	−8.43	MDR-like ABC transporter
LOC_Os02g34580	0.01	3.45	−8.43	Ammonium transporter protein
LOC_Os01g71310	0.01	3.63	−8.50	Cytokinin dehydrogenase precursor
LOC_Os10g22310	0.01	3.97	−8.63	Glutathione S-transferase GST 26
LOC_Os12g24320	0.01	4.15	−8.70	ATPase 3
LOC_Os05g08370	0.01	4.66	−8.86	CESA1 – cellulose synthase
LOC_Os10g05780	0.01	5.18	−9.02	POT domain containing peptide transporter
LOC_Os01g58760	0.01	5.87	−9.20	bZIP transcription factor domain containing protein

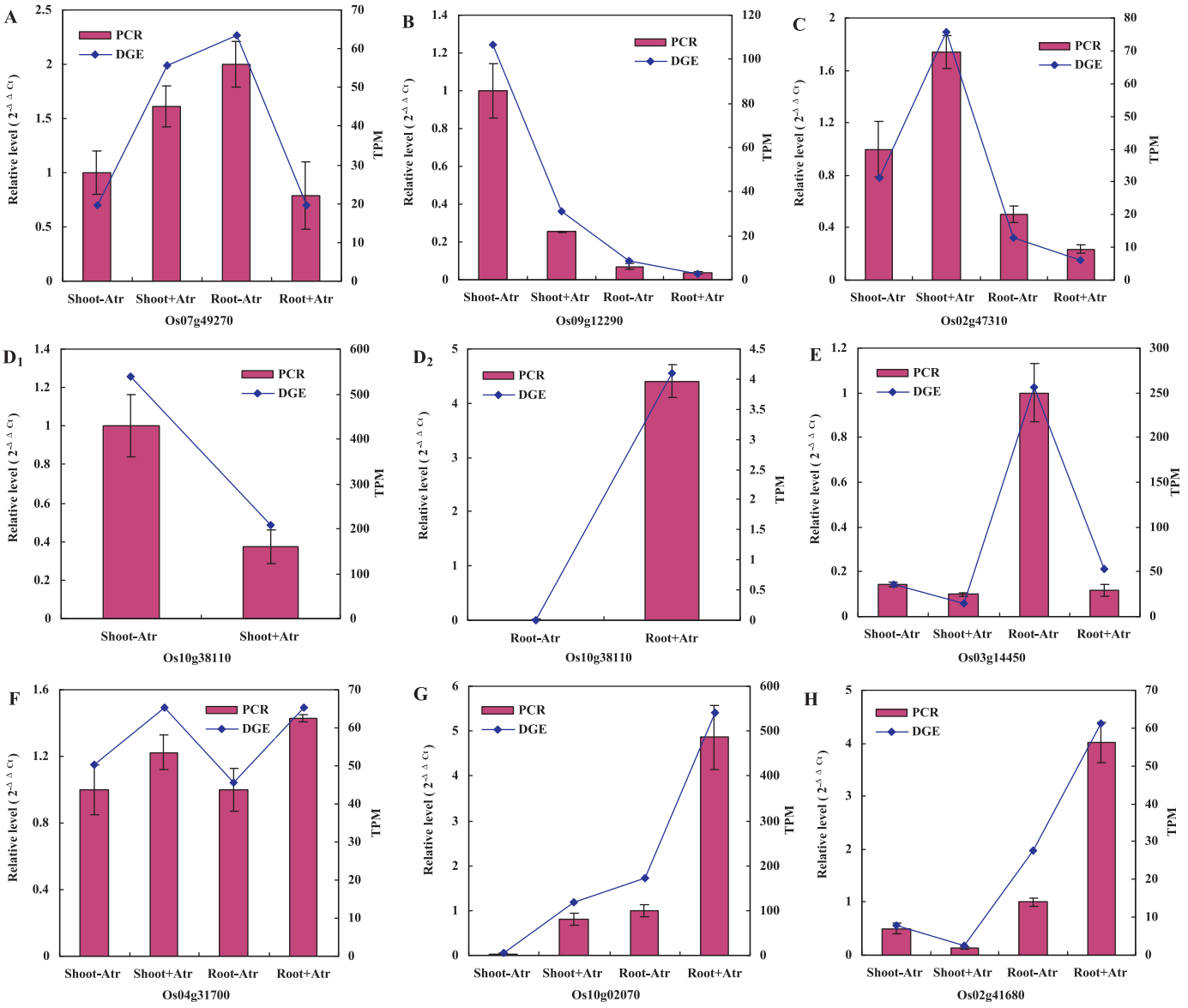

Fig. 3. Quantitative real time RT-PCR validation of selected tag-mapped genes or digital gene expression (DGE) from rice shoots and roots. These genes encode proteins as presented as follows. (A), Os07g49270: AMP deaminase; (B), Os09g12290: bifunctional aspartokinase/homoserine dehydrogenase; (C), Os02g47310: cyclopropane-fatty-acyl-phospholipid synthase; (D), Os10g38110: cytochrome P450; (E), Os03g14450: enolase; F, Os04g31700: methylisocitrate lyase 2; (G), Os10g02070: peroxidase; (H), Os02g41680: phenylalanine ammonia-lyase. TPM: transcripts per million clean tags. Atr: atrazine. Vertical bars represent standard deviation of the mean ($n = 3$) between the treatment and control.

Table 5
Significantly enriched gene ontology (GO) terms in shoot/root-Atr and shoot/root + Atr libraries of rice (*Oryza sativa* Japonica Group). GO terms including three ontologies: biological process, cellular component and molecular function with corrected *P*-value < 0.01, indicate significantly enriched in differentially expressed genes (DEGs).

GO term	Cluster frequency	Genome frequency of use	Corrected *P*-value
	Root-Atr and Root + Atr		
BP: Biological Process	DEGs (936)	Genes (5612)	
Cellular protein metabolic process	22.4% (210,33↑,177↓)	16.0% (899)	6.48E−06
Gene expression	20.0% (187,24↑,163↓)	14.4% (810)	1.30E−04
Biological regulation	19.9% (186,29↑,157↓)	13.9% (782)	1.39E−05
Regulation of biological process	19.1% (179,28↑,151↓)	13.6% (764)	8.97E−05
Regulation of cellular process	18.3% (171,26↑,145↓)	12.7% (715)	4.02E−05
Localization	15.3% (143,22↑,121↓)	10.4% (586)	1.70E−04
Transport	15.2% (142,22↑,120↓)	10.4% (583)	2.10E−04
Establishment of localization	15.2% (142,22↑,120↓)	10.4% (583)	2.10E−04
Carbohydrate metabolic process	12.1% (113,32↑,81↓)	6.4% (359)	3.81E−10
Rsponse to stimulus	11.4% (107,28↑,79↓)	6.9% (390)	1.28E−05
Cellular carbohydrate metabolic process	6.9% (65,17↑,48↓)	3.4% (192)	1.72E−06
Organic acid metabolic process	6.9% (65,20↑,45↓)	4.2% (236)	8.07E−03
Carboxylic acid metabolic process	6.9% (65,20↑,45↓)	4.2% (236)	8.07E−03
Cellular ketone metabolic process	6.9% (65,20↑,45↓)	4.2% (236)	8.07E−03
Oxoacid metabolic process	6.9% (65,20↑,45↓)	4.2% (236)	8.07E−03
Cellular amino acid and derivative metabolic process	5.4% (51,14↑,37↓)	2.9% (165)	2.15E−03
Amine metabolic process	5.3% (50,18↑,32↓)	2.9% (160)	1.89E−03
Cellular component biogenesis	3.5% (33,4↑,29↓)	1.7% (93)	4.92E−03
Anatomical structure formation	3.4% (32,4↑,28↓)	1.6% (89)	4.92E−03
Alcohol metabolic process	3.3% (31,8↑,23↓)	1.5% (82)	2.03E−03
Secondary metabolic process	3.3% (31,16↑,15↓)	1.5% (85)	4.87E−03
Monosaccharide metabolic process	3.0% (28,8↑,20↓)	1.3% (71)	2.33E−03
Cellular aromatic compound metabolic process	3.0% (28,8↑,20↓)	1.3% (75)	8.05E−03
Regulation of biological quality	2.9% (27,6↑,21↓)	1.1% (63)	5.00E−04
Homeostatic process	2.4% (22,5↑,17↓)	0.8% (47)	8.70E−04
Cellular amino acid derivative metabolic process	2.2% (21,6↑,15↓)	0.9% (48)	5.76E−03
Carbohydrate biosynthetic process	2.1% (20,5↑,15↓)	0.8% (45)	7.06E−03
CC: Cellular Component	DEGs (2243)	Genes (13,491)	
Membrane-bounded vesicle	37.2% (834,205↑,629↓)	30.8% (4152)	1.09E−10
Cytoplasmic vesicle	37.2% (834,205↑,629↓)	30.8% (4160)	1.79E−10
Membrane	22.7% (509,87↑,422↓)	16.6% (2243)	2.82E−14
Macromolecular complex	6.6% (147,17↑,130↓)	4.4% (598)	3.60E−05
Nucleus	6.4% (144,24↑,120↓)	3.8% (515)	5.00E−09
Membrane part	4.6% (104,10↑,94↓)	2.9% (387)	2.57E−05
Intracellular non−membrane−bounded organelle	4.3% (97,11↑,86↓)	2.6% (354)	2.49E−05
Intracellular organelle part	4.2% (95,12↑,83↓)	2.7% (359)	1.80E−04
Organelle part	4.2% (95,12↑,83↓)	2.7% (362)	2.70E−04
Integral to membrane	3.5% (79,7↑,72↓)	2.0% (274)	3.75E−05
Intrinsic to membrane	3.5% (79,7↑,72↓)	2.0% (275)	4.42E−05
Extracellular region	1.6% (35,14↑,21↓)	0.5% (71)	2.78E−08
Cell periphery	1.4% (31,4↑,27↓)	0.6% (75)	5.05E−05
Plasma membrane	1.2% (27,4↑,23↓)	0.5% (61)	5.52E−05
Cytoskeleton	0.8% (18,2↑,16↓)	0.3% (41)	5.59E−03
Apoplast	0.6% (13,9↑,4↓)	0.2% (24)	4.40E−03
MF: Molecular Function	DEGs (1191)	Genes (6978)	
Nucleotide binding	18.3% (218,39↑,179↓)	11.9% (830)	2.48E−10
Protein binding	16.5% (196,18↑,178↓)	10.0% (698)	4.37E−12
Purine nucleotide binding	16.4% (195,35↑,160↓)	10.8% (755)	2.83E−08
Ribonucleotide binding	15.9% (189,34↑,155↓)	10.5% (736)	1.00E−07
Purine ribonucleotide binding	15.9% (189,34↑,155↓)	10.5% (736)	1.00E−07
Purine nucleoside binding	13.9% (166,31↑,135↓)	9.6% (670)	2.36E−05
Adenyl nucleotide binding	13.9% (166,31↑,135↓)	9.6% (670)	2.36E−05
Nucleoside binding	13.9% (166,31↑,135↓)	9.7% (675)	4.07E−05
ATP binding	13.4% (160,30↑,130↓)	9.3% (651)	7.51E−05
Adenyl ribonucleotide binding	13.4% (160,30↑,130↓)	9.3% (651)	7.51E−05
Cation binding	12.5% (149,38↑,111↓)	8.0% (555)	4.10E−07
Ion binding	12.5% (149,38↑,111↓)	8.0% (559)	7.09E−07
Metal ion binding	12.0% (143,36↑,107↓)	7.7% (540)	2.52E−06
Oxidoreductase activity	10.0% (119,31↑,88↓)	6.7% (466)	4.20E−04
Transition metal ion binding	7.7% (92,26↑,66↓)	4.5% (313)	8.84E−06
Transferase activity, transferring alkyl or aryl (other than methyl) groups	2.4% (28,18↑,10↓)	0.9% (64)	1.80E−04
Glutathione transferase activity	1.8% (22,17↑,5↓)	0.6% (39)	1.10E−05
UDP-glucosyltransferase activity	0.9% (11,1↑,10↓)	0.2% (15)	9.20E−04
Manganese ion binding	0.9% (11,9↑,2↓)	0.2% (17)	6.01E−03
Glucosyltransferase activity	0.9% (11,1↑,10↓)	0.2% (17)	6.01E−03

Table 5 (*Continued*)

	Shoot-Atr and Shoot + Atr		
BP: Biological Process	DEGs (526)	Genes (5612)	
Biological regulation	21.7% (114, 67↑, 47↓)	13.9% (782)	1.00E−04
Regulation of biological process	20.7% (109, 64↑, 45↓)	13.6% (764)	6.40E−04
Regulation of metabolic process	14.8% (78, 43↑, 35↓)	8.6% (481)	1.90E−04
Carbohydrate metabolic process	14.6% (77, 40↑, 37↓)	6.4% (359)	2.88E−10
Regulation of primary metabolic process	14.6% (77, 42↑, 35↓)	8.2% (461)	6.55E−05
Regulation of macromolecule metabolic process	14.6% (77, 42↑, 35↓)	8.5% (477)	2.70E−04
Response to stimulus	13.1% (69, 44↑, 25↓)	6.9% (390)	3.23E−05
Regulation of cellular metabolic process	12.9% (68, 38↑, 30↓)	7.9% (442)	7.49E−03
Regulation of macromolecule biosynthetic process	12.7% (67, 37↑, 30↓)	7.8% (436)	9.05E−03
Regulation of biosynthetic process	12.7% (67, 37↑, 30↓)	7.8% (437)	9.77E−03
Regulation of cellular biosynthetic process	12.7% (67, 37↑, 30↓)	7.8% (437)	9.77E−03
Organic acid metabolic process	9.7% (51, 22↑, 29↓)	4.2% (236)	2.51E−06
Carboxylic acid metabolic process	9.7% (51, 22↑, 29↓)	4.2% (236)	2.51E−06
Cellular ketone metabolic process	9.7% (51, 22↑, 29↓)	4.2% (236)	2.51E−06
Oxoacid metabolic process	9.7% (51, 22↑, 29↓)	4.2% (236)	2.51E−06
Cellular carbohydrate metabolic process	8.7% (46, 22↑, 24↓)	3.4% (192)	4.49E−07
Amine metabolic process	7.8% (41, 21↑, 20↓)	2.9% (160)	4.44E−07
Cellular amino acid and derivative metabolic process	7.0% (37, 15↑, 22↓)	2.9% (165)	1.30E−04
Cellular amino acid metabolic process	5.7% (30, 12↑, 18↓)	2.3% (129)	7.90E−04
Cellular amine metabolic process	5.7% (30, 12↑, 18↓)	2.3% (129)	7.90E−04
Alcohol metabolic process	4.8% (25, 13↑, 12↓)	1.5% (82)	2.31E−05
Monosaccharide metabolic process	3.6% (19, 10↑, 9↓)	1.3% (71)	7.89E−03
Carbohydrate catabolic process	3.4% (18, 10↑, 8↓)	1.2% (66)	9.89E−03
Protein folding	2.7% (14, 8↑, 6↓)	0.7% (42)	6.78E−03
Sulfur metabolic process	2.3% (12, 2↑, 10↓)	0.5% (27)	8.40E−04
Sulfur compound biosynthetic process	1.9% (10, 2↑, 8↓)	0.3% (19)	9.10E−04
Sulfur, amino acid metabolic process	1.5% (8, 1↑, 7↓)	0.3% (15)	8.99E−03
Photosynthesis, light harvesting	1.3% (7, 0↑, 7↓)	0.1% (8)	2.00E−04
Sulfur, amino acid biosynthetic process	1.3% (7, 1↑, 6↓)	0.2% (11)	6.46E−03
CC: Cellular Component	DEGs (1182)	Genes (1,3491)	
Membrane	22.3% (264, 135↑, 129↓)	16.6% (2243)	7.13E−06
MF: Molecular Function	DEGs (667)	Genes (6978)	
Protein binding	16.2% (108, 73↑, 35↓)	10.0% (698)	3.16E−05
Cation binding	13.9% (93, 53↑, 40↓)	8.0% (555)	6.02E−06
Ion binding	13.9% (93, 53↑, 40↓)	8.0% (559)	8.67E−06
Metal ion binding	13.5% (90, 50↑, 40↓)	7.7% (540)	1.37E−05
Oxidoreductase activity	12.3% (82, 44↑, 38↓)	6.7% (466)	4.92E−06
Ligase activity	4.3% (29, 14↑, 15↓)	1.6% (112)	1.10E−04
Isomerase activity	2.8% (19, 7↑, 12↓)	1.0% (69)	4.66E−03

additionally, transcripts involved in glutathione, ascorbate, and aldarate metabolism were enriched in Root-Atr/Root + Atr and Shoot-Atr/Shoot + Atr datasets.

4. Discussion

Although genome-wide gene expression in rice has been identified for growth and development [38,39], molecular information on toxicology with organic xenobiotics is lacking. A global analysis of transcriptome would facilitate our understanding of systemic gene expression and regulatory mechanisms for plant adaptive responses to xenobiotic stress. In this study, we identified a large number of genes differently expressed in rice tissues with atrazine using the recently developed high-throughput sequencing technology [30,40]. Creating four libraries from rice tissues exposed to atrazine and performing a direct DEG analysis allowed us to identify 18,827–21,007 essential and dynamic genes for each library. These genes are responsible for enzymes or proteins involved in synthesis of diverse essential substances (e.g. carbonhydrates, proteins, and lipid) and secondary metabolites associated with currently known or unknown biological pathways. Thus, the present study represented so far the deepest analysis on rice transcriptome in response to toxic organic compounds.

Response of plants to atrazine undergoes several biological processes, including perception of stress signal, regulation of gene expression, metabolic reinforcement, and physiological acclimation. Understanding of altered gene expression gives insights into molecular and physiological responses to atrazine. Our analysis revealed at least 6714 and 3596 enriched genes in roots and shoots with atrazine, respectively (Table 3). These genes were differentially expressed in the presence of atrazine, and a certain proportion of them responded specifically to atrazine exposure. For instance, in the root 1319 transcripts were up-regulated and 5395 down-regulated by atrazine treatment (Table 3). Notably, 128 genes were expressed only in atrazine-treated roots, whereas 379 genes were preferentially expressed in untreated roots. In shoots, there were 2265 and 1331 genes that were up- and down-regulated in response to atrazine. However, compared with roots, more genes in shoots were induced by atrazine. These results suggest that expression of genes were reprogrammed in the presence of atrazine.

We demonstrated forty genes that were most differentially expressed in Root + Atr vs Root-Atr and Shoot + Atr vs Shoot-Atr libraries (Table 4). These genes are encoding proteins or enzymes regulating diverse biological processes. Glycosyl hydrolases in plants belong a big family and are believed as secretory proteins for modification of polysaccharides in cell wall [41]. This gene, along with several other transcripts such as expansion, peroxidase, and glycine-rich structural protein genes, are involved in cell wall build-up or reinforcement. An example comes from peroxidase which catalyzes the conversion of H_2O_2 into H_2O by oxidizing

Table 6
Significant pathway enrichment analysis in KEGG in each pair of *Oryza sativa* libraries. Pathways with *P*-value < 0.05 and *Q* value < 0.05 indicate significantly enriched in differentially expressed genes (DEGs).

Pathway	Number of DEGs with pathway annotation	Number of genes with pathway annotation	*P* value	*Q* value	Pathway ID
Root-Atr vs Root + Atr	(3399)	(16,810)			
Metabolic pathways	804 (23.65%, 183↑, 621↓)	3281 (19.52%)	1.29E−11	1.52E−09	ko01100
Biosynthesis of secondary metabolites	508 (14.95%, 163↑, 345↓)	2129 (12.67%)	6.12E−06	1.81E−04	ko01110
Phenylpropanoid biosynthesis	160 (4.71%, 55↑, 105↓)	631 (3.75%)	8.26E−04	6.50E−03	ko00940
Ribosome	124 (3.65%, 13↑, 111↓)	460 (2.74%)	2.55E−04	2.64E−03	ko03010
Starch and sucrose metabolism	109 (3.21%, 20↑, 89↓)	429 (2.55%)	4.83E−03	2.31E−02	ko00500
Phenylalanine metabolism	93 (2.74%, 29↑, 64↓)	270 (1.61%)	2.78E−08	1.56E−06	ko00360
Oxidative phosphorylation	68 (2.00%, 9↑, 59↓)	222 (1.32%)	1.43E−04	2.11E−03	ko00190
Purine metabolism	68 (2.00%, 5↑, 63↓)	259 (1.54%)	1.07E−02	4.53E−02	ko00230
Glycolysis/Gluconeogenesis	60 (1.77%, 17↑, 43↓)	185 (1.1%)	5.88E−05	1.16E−03	ko00010
Amino sugar and nucleotide sugar metabolism	58 (1.71%, 20↑, 38↓)	189 (1.12%)	4.02E−04	3.65E−03	ko00520
Glutathione metabolism	57 (1.68%, 31↑, 26↓)	142 (0.84%)	3.95E−08	1.56E−06	ko00480
Endocytosis	49 (1.44%, 8↑, 41↓)	173 (1.03%)	6.47E−03	2.93E−02	ko04144
Nitrogen metabolism	44 (1.29%, 17↑, 27↓)	118 (0.7%)	1.31E−05	3.08E−04	ko00910
Phagosome	42 (1.24%, 4↑, 38↓)	120 (0.71%)	1.11E−04	1.88E−03	ko04145
Pyruvate metabolism	38 (1.12%, 10↑, 28↓)	108 (0.64%)	2.04E−04	2.41E−03	ko00620
Ascorbate and aldarate metabolism	37 (1.09%, 11↑, 26↓)	104 (0.62%)	1.90E−04	2.41E−03	ko00053
Fructose and mannose metabolism	34 (1.00%, 5↑, 29↓)	107 (0.64%)	3.16E−03	1.86E−02	ko00051
Aminoacyl-tRNA biosynthesis	31 (0.91%, 5↑, 26↓)	95 (0.57%)	2.99E−03	1.85E−02	ko00970
Glycine, serine and threonine metabolism	29 (0.85%, 9↑, 20↓)	79 (0.47%)	5.00E−04	4.22E−03	ko00260
Tyrosine metabolism	28 (0.82%, 12↑, 16↓)	85 (0.51%)	3.99E−03	2.24E−02	ko00350
Butanoate metabolism	27 (0.79%, 4↑, 23↓)	79 (0.47%)	2.60E−03	1.80E−02	ko00650
Citrate cycle (TCA cycle)	26 (0.76%, 5↑, 21↓)	66 (0.39%)	2.68E−04	2.64E−03	ko00020
Inositol phosphate metabolism	26 (0.76%, 4↑,22↓)	78 (0.46%)	4.53E−03	2.31E−02	ko00562
Pentose phosphate pathway	23 (0.68%, 3↑, 20↓)	67 (0.4%)	4.89E−03	2.31E−02	ko00030
Sulfur metabolism	22 (0.65%, 6↑, 16↓)	63 (0.37%)	4.63E−03	2.31E−02	ko00920
Valine, leucine and isoleucine degradation	22 (0.65%, 9↑, 13↓)	67 (0.4%)	1.04E−02	4.53E−02	ko00280
Histidine metabolism	18 (0.53%, 5↑, 13↓)	44 (0.26%)	1.37E−03	1.01E−02	ko00340
Non-homologous end-joining	10 (0.29%, 0↑, 10↓)	20 (0.12%)	2.81E−03	1.84E−02	ko03450
Shoot-Atr vs Shoot + Atr	(1861)	(16,810)			
Metabolic pathways	461 (24.77%, 254↑, 207↓)	3281 (19.52%)	2.06E−09	2.37E−07	ko01100
Biosynthesis of secondary metabolites	279 (14.99%, 154↑, 125↓)	2129 (12.67%)	9.54E−04	9.02E−03	ko01110
Spliceosome	89 (4.78%, 64↑, 25↓)	511 (3.04%)	8.86E−06	2.55E−04	ko03040
Glycolysis/Gluconeogenesis	38 (2.04%, 18↑, 20↓)	185 (1.10%)	1.18E−04	1.70E−03	ko00010
Glycerolipid metabolism	29 (1.56%, 21↑, 8↓)	118 (0.70%)	2.60E−05	5.97E−04	ko00561
Pyruvate metabolism	27 (1.45%, 17↑, 10↓)	108 (0.64%)	3.52E−05	6.74E−04	ko00620
Fructose and mannose metabolism	25 (1.34%, 10↑, 15↓)	107 (0.64%)	2.16E−04	2.76E−03	ko00051
Fatty acid metabolism	25 (1.34%, 20↑, 5↓)	115 (0.68%)	6.87E−04	7.19E−03	ko00071
Nitrogen metabolism	25 (1.34%, 12↑, 13↓)	118 (0.70%)	1.02E−03	9.02E−03	ko00910
Sulfur metabolism	22 (1.18%, 5↑, 17↓)	63 (0.37%)	4.82E−07	1.85E−05	ko00920
Carbon fixation in photosynthetic organisms	21 (1.13%, 8↑, 13↓)	103 (0.61%)	4.02E−03	2.89E−02	ko00710
Alanine, aspartate and glutamate metabolism	20 (1.07%, 11↑, 9↓)	98 (0.58%)	4.86E−03	3.29E−02	ko00250
Valine, leucine and isoleucine degradation	19 (1.02%, 15↑, 4↓)	67 (0.40%)	8.11E−05	1.33E−03	ko00280
Ubiquinone and other terpenoid-quinone biosynthesis	17 (0.91%, 11↑, 6↓)	73 (0.43%)	2.18E−03	1.69E−02	ko00130
Butanoate metabolism	17 (0.91%, 9↑, 8↓)	79 (0.47%)	5.19E−03	3.32E−02	ko00650
Selenocompound metabolism	16 (0.86%, 5↑, 11↓)	59 (0.35%)	5.04E−04	5.80E−03	ko00450
Pentose phosphate pathway	16 (0.86%, 6↑, 10↓)	67 (0.40%)	2.20E−03	1.69E−02	ko00030
Photosynthesis–antenna proteins	12 (0.64%, 0↑, 12↓)	17 (0.10%)	1.19E−08	6.84E−07	ko00196

various secondary metabolites such as phenolic compounds. In plants, the group III peroxidase is the major type, which is composed of numerous isoenzymes and mediates biosynthesis of secondary metabolites and cross-linking of matrix polysaccharides in cell wall [42,43]. In this study, one of peroxidase genes (*Os10g02070*) was up-regulated by atrazine treatment using both deep-sequencing and qRT-PCR approaches (Fig. 3). Induction of group III peroxidase activity by pesticides has been found in many food crops such as wheat [44,45], rapeseed [46], and green algae [47]. This suggests that activation of peroxidase gene expression may facilitate the cross-linking of secondary metabolites or cell wall solidification, and confer resistance to atrazine toxicity. It is worthy to note that two genes coding for laccase and cytochrome P450 were up-regulated by atrazine in rice (Table 4). In higher plants, both laccase and cytochrome P450 have been implicated in detoxification and catabolism of herbicides [1,48]. The process for herbicide degradation undergoes three steps: conversion (e.g. hydrophobic chemicals are converted into less hydropholic metabolites), conjugation (with the aid of glutathione S-transferases), and compartmentation (into vacuoles or other organelles) [2].

One major advantage of the high throughput sequencing is the direct estimation of gene expression [36]. Moreover, millions of generated rice sequences allowed us to analyze enrichment of various genes. The identified genes within the libraries displayed varied abundance from one another. According to the GO and KEGG distribution, the majority of genes were enriched to involve metabolic processes (e.g. genes for metabolism of carbohydrates, organic acids, sulphate, amino acids, secondary metabolites, etc.) and regulation of cellular process (Tables 5 and 6). These results indicate that plant exposure to atrazine altered many basic biological pathways associated with metabolism of carbohydrates, proteins, lipids, and nucleic acids. We also retrieved many important genes grouped into the three categories including response to intracellular and environmental stimulus, glutathione transferase activity, and oxidoreductase activity. Using the database categories, we identified 107 genes for stimulus-responsive protein, 22 genes coding for

glutathione transferases, and 119 genes for oxidoreductases, all of which were differentially regulated by atrazine.

In conclusion, we presented an extensive survey of atrazine-responsive genes, showing differential expression in rice tissues. These genes are involved in many metabolic processes in responses to atrazine. Our results not only demonstrate transcriptional complexity in rice with atrazine, but also open up a possibility for identification of the genes in regulating plant tolerance to xenobiotic stress. Furthermore, the high-throughput sequencing platform described here serves as a powerful approach for developing biomarkers monitoring molecular response to the organic hazards.

Acknowledgment

The authors acknowledge the financial support of the National Natural Science Foundation of China (No. 21077055) and Special Fund for Agro-scientific Research in the Public Interest (No. 201203022) from the Ministry of Agriculture of China for this study.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jhazmat.2012.03.041.

References

[1] K.E. Gerhardt, X. Huang, B.R. Glick, B.M. Greenberg, Phytoremediation and rhizoremediation of organic soil contaminants, Plant Sci. 176 (2009) 20–30.
[2] J.W. Denser, Toxicity of mixtures of pesticides in aquatic systems, Pest Manag. Sci. 56 (2000) 516–520.
[3] O.S. Fatoki, O.R. Awofolu, Levels of organochlorine pesticide residues in marine-, surface-, ground-, and drinking waters from the Eastern Cape Province of South Africa, J. Environ. Sci. Health Part B 39 (2004) 101–114.
[4] G.D. Hela, A. Lambropoulou, I.K. Konstantinou, T.A. Albanis, Environmental monitoring and ecological risk assessment for pesticide contamination and effects in Lake Pamvotis, northwestern Greece, Environ. Toxicol. Chem. 24 (2005) 1548–1556.
[5] K.M.A. Holvoet, P. Seuntjens, P.A. Vanrolleghem, Monitoring and modeling pesticide fate in surface waters at the catchment scale, Ecol. Model. 209 (2007) 53–64.
[6] M. Arias-Estévez, E. López-Periago, E. Martínez-Carballo, J. Simal-Gándara, J.C. Mejuto, L. García-Río, The mobility and degradation of pesticides in soils and the pollution of groundwater resources, Agric. Ecosyst. Environ. 123 (2008) 247–260.
[7] V. Belessi, D. Lambropoulou, I. Konstantinou, R. Zbori, J. Tucek, D. Jancik, T. Albanis, D. Petridis, Structure and photocatalytic performance of magnetically separable titania photocatalysts for the degradation of propachlor, Appl. Catal. B: Environ. 87 (2009) 181–189.
[8] K. Majumdar, N. Singh, Effect of soil amendments on sorption and mobility of metribuzin in soils, Chemosphere 66 (2007) 630–637.
[9] L. Jiang, L. Ma, Y. Sui, S.Q. Han, H. Yang, Mobilization and plant accumulation of prometryne in soil by two different sources of organic matter, J. Environ. Monit. 13 (2011) 1935–1943.
[10] Y.H. Su, Y.G. Zhu, Bioconcentration of atrazine and chlorophenols into roots and shoots of rice seedlings, Environ. Pollut. 139 (2006) 32–39.
[11] P.K. Biswas, S.K. Pramanik, S.R. Mitra, A. Bhattacharyya, Persistence of napropamide in/on tea under north-east Indian climatic condition, Bull. Environ. Contam. Toxicol. 79 (2007) 566–569.
[12] R. Zhang, J. Cui, H.M. Zhu, H. Yang, Effect of dissolved organic matters on napropamide availability and ecotoxicity in rapeseed (*Brassica napus*), J. Agric. Food Chem. 58 (2010) 3232–3240.
[13] L.E. Cui, H. Yang, Accumulation and residue of napropamide in alfalfa (*Medicago sativa*) and soil involved in toxic response, J. Hazard. Mater. 190 (2011) 81–86.
[14] D.H. Bennett, W.E. Kastenberg, T.E. McKone, A multimedia, multiple pathway risk assessment of atrazine: the impact of age differentiated exposure including joint uncertainty and variability, Reliab. Eng. Syst. Saf. 63 (1999) 185–198.
[15] A.D. Carter, Herbicide movement in soils: principles, pathways and processes, Weed Res. 40 (2000) 113–122.
[16] J.G. Gu, Y.Z. Fan, J.D. Gu, Biodegradability of atrazine, cyanazine and dicamba under methanogenic condition in three soils of China, Chemosphere 52 (2003) 1515–1521.
[17] N. Sathiakumar, P.A. MacLennan, J. Mandel, E. Delzell, A review of epidemiologic studies of triazine herbicides and cancer, Crit. Rev. Toxicol. 41 (2011) 1–34.
[18] A.W. Rutherford, A. Krieger-Liszkay, Herbicide-induced oxidative stress in photosystem II, Trends Biochem. Sci. 6 (2001) 648–653.
[19] Y.H. Su, Y.G. Zhu, A.J. Lin, X.H. Zhang, Interaction between cadmium and atrazine during uptake by rice seedlings (*Oryza sativa* L.), Chemosphere 60 (2005) 802–809.
[20] M.T. Moore, R. Kröger, Effect of three insecticides and two herbicides on rice (*Oryza sativa*) seedling germination and growth, Arch. Environ. Contam. Toxicol. 59 (2010) 574–581.
[21] H. Kawahigashi, Transgenic plants for phytoremediation of herbicides, Curr. Opin. Biotechnol. 20 (2009) 225–230.
[22] C.A. James, S.E. Strand, Phytoremediation of small organic contaminants using transgenic plants, Curr. Opin. Biotechnol. 20 (2009) 237–241.
[23] M. Brazier-Hicks, W.A. Offen, M.C. Gershater, T.J. Revett, E.K. Lim, D.J. Nowles, G.J. Davies, R. Edwards, Characterization and engineering of the bifunctional N- and O-gulcosyltransferase involved in xenobiotics metabolism in plants, Proc. Natl. Aacd. Sci. U.S.A. 104 (2007) 20238–20243.
[24] S.M. Mense, A. Sengupta, C. Lan, M. Zhou, G. Bentsman, D.J. Volsky, R.M. Whyatt, F.P. Perera, L. Zhang, The Common insecticides yfluthrin and chlorpyrifos alter the expression of a subset of genes with diverse functions in primary human astrocytes, Toxicol. Sci. 93 (2006) 125–135.
[25] F. Ramel, C. Sulmon, F. Cabello-Hurtado, L. Taconnat, M. Martin-Magniette, J. Renou, A.E. Amrani, I. Couée, G. Gouesbet, Genome-wide interacting effects of sucrose and herbicide-mediated stress in *Arabidopsis thaliana*: novel insights into atrazine toxicity and sucrose-induced tolerance, BMC Genomics 8 (2007) 450.
[26] T. Unver, M. Bakar, R.C. Shearman, H. Budak, Genome-wide profiling and analysis of festuca arundinacea miRNAs and transcriptomes in response to foliar glyphosate application, Mol. Genet. Genomics 283 (2010) 397–413.
[27] M. Das, J.R. Reichman, G. Haberer, G. Welzl, F.F. Aceituno, M.T. Mader, L.S. Watrud, T.G. Pfleeger, R.A. Gutiérrez, A.R. Schäffner, D.M. Olszyk, A composite transcriptional signature differentiates responses towards closely related herbicides in *Arabidopsis thaliana* and *Brassica napus*, Plant Mol. Biol. 72 (2010) 545–556.
[28] A.L. Viñuela, B. Snoek, J.A.G. Riksen, J.E. Kammenga, Genome-wide gene expression analysis in response to organophosphorus pesticide chlorpyrifos and diazinon in *C. elegans*, PLoS One 5 (2010) 12145.
[29] S.C. Schuster, Next-generation sequencing transforms today's biology, Nat. Methods 5 (2008) 16–18.
[30] Q.Q. Wang, F. Liu, X.S. Chen, X.J. Ma, H.Q. Zeng, Z.M. Yang, Transcriptome profiling of early developing cotton fibre by deep-sequencing reveals significantly differential expression of genes in a fuzzless/lintless mutant, Genomics 96 (2010) 369–376.
[31] S.Q. Huang, J. Peng, C.X. Qiu, Z.M. Yang, Heavy metal-regulated new microRNAs from rice, J. Inorg. Biochem. 103 (2009) 282–287.
[32] S. Sondhia, Determination of imazosulfuron persistence in rice crop and soil, Environ. Monit. Assess. 137 (2008) 205–211.
[33] K. Hayashi, A. Nakae, Y. Fukushima, K. Sakamoto, T. Furuichi, K. Kitahara, Y. Miyazaki, C. Ikenoue, S. Matumoto, T. Toda, Contamination of rice by etofenprox, diethylphthalate and alkylphenols. Effects on first delivery and sperm count in mice, J. Toxicol. Sci. 35 (2010) 49–55.
[34] G.L. Wu, J. Cui, L. Tao, H. Yang, Fluroxypyr triggers oxidative damage by producing superoxide and hydrogen peroxide in rice (*Oryza sativa*), Ecotoxicology 19 (2010) 124–132.
[35] S. Audic, J.M. Claverie, The significance of digital gene expression profiles, Genome Res. 7 (1997) 986–995.
[36] C. Molina, B. Rotter, R. Horres, S.M. Udupa, B. Besser, L. Bellarmino, M. Baum, H. Matsumura, R. Teruchi, G. Kahl, P. Winter, SuperSAGE: the drought stress-responsive transcriptome of chickpea roots, BMC Genomics 9 (2008) 553.
[37] M. Kanehisa, S. Goto, S. Kawashima, Y. Okuno, M. Hattori, The KEGG resource for deciphering the genome, Nucleic Acids Res. 32 (2004) D277–D280.
[38] H. Yang, L. Hu, T. Hurek, B. Reinhold-Hurek, Global characterization of the root transcriptome of a wild species of rice, *Oryza longistaminata*, by deep sequencing, BMC Genomics 11 (2010) 705.
[39] R. Venu, M. Sreerekha, K. Nobuta, A. Beló, Y. Ning, G. An, B.C. Meyers, G.L. Wang, Deep sequencing reveals the complex and coordinated transcriptional regulation of genes related grain quality in rice cultivars, BMC Genomics 12 (2011) 190.
[40] A.S. Morrissy, R.D. Morin, A. Delaney, T. Zeng, H. McDonald, S. Jones, Y. Zhao, M. Hirst, M.A. Marra, Next-generation tag sequencing for cancer gene expression profiling, Genome Res. 19 (2010) 1825–1835.
[41] E.J. Lee, Y. Matsumura, K. Soga, T. Hoson, N. Koizumi, Glycosyl hydrolases of cell wall are induced by sugar starvation in arabidopsis, Plant Cell Physiol. 48 (2007) 405–413.
[42] C. Dunand, M. Crèvecoeur, C. Penel, Distribution of superoxide and hydrogen peroxide in Arabidopsis root and their influence on root development: possible interaction with peroxidases, New Phytol. 174 (2007) 332–341.
[43] Y.J. Xue, L. Tao, Z.M. Yang, Aluminum-induced cell wall peroxidase activity and lignin synthesis are differentially regulated by jasmonate and nitric oxide, J. Agric. Food Chem. 56 (2008) 9676–9684.
[44] X.L. Yin, L. Jiang, N.H. Song, H. Yang, Toxic reactivity of wheat (*Triticum aestivum*) plants to herbicide isoproturon, J. Agric. Food Chem. 56 (2008) 4825–4831.

[45] N.H. Song, S. Zhang, M. Hong, H. Yang, Impact of dissolved organic matter on bioavailability of chlorotoluronn to wheat, Environ. Pollut. 158 (2010) 906–912.
[46] J. Cui, R. Zhang, G.L. Wu, H.M. Zhu, H. Yang, Salicylic acid reduces napropamide toxicity by preventing its accumulation in rapeseed (*Brassica napus* L.), Arch. Environ. Contam. Toxicol. 59 (2010) 100–108.
[47] S. Zhang, C.B. Qiu, Y. Zhou, Z.P. Jin, H. Yang, Bioaccumulation and degradation of pesticide fluroxypyr are associated with toxic tolerance in green alga *Chlamydomonas reinhardtii*, Ecotoxicology 20 (2011) 337–347.
[48] J. Majeau, S.K. Brar, R.D. Tyagi, Laccases for removal of recalcitrant and emerging pollutants, Bioresour. Technol. 101 (2010) 2331–2350.