GM Crops & Food, 10:20–34, 2019

ISSN: 2164-5698 print / 2164-5701 online
DOI: 10.1080/21645698.2019.1598215

RESEARCH PAPER

Unintended effects of transgenic rice revealed by transcriptome and metabolism

Wei Fu,[a]* Chenguang Wang,[a,b,c]* Wenjie Xu,[a,b,c] Pengyu Zhu,[a] Yun Lu,[a] Shuang Wei,[d] Xiyang Wu,[e] Yuping Wu,[a] Yiqiang Zhao,[b] and Shuifang Zhu[a,c]

[a]Chinese Academy of Inspection and Quarantine, Beijing, China; [b]Beijing Advanced Innovation Center for Food Nutrition and Human Health, College of Biological Sciences, China Agricultural University, Beijing, China; [c]College of Plant Protection, China Agricultural University, Beijing, China; [d]Guangdong Entry-Exit Inspection and Quarantine Bureau, Guangzhou, China; [e]Department of Food Science and Engineering, Jinan University, Guangzhou, China

ABSTRACT. Genetically modified (GM) organisms have been developed for decades. However, unintended effects are the main concerns of safety assessment that needs to be carefully investigated. Here, eight varieties of GM rice that were developed in China were selected to assess the unintended effects through transcriptome and metabolism. There are 2892–8758 differentially expressed genes (DEGs) and 7–50 metabolites at significant level between GM varieties and their isogenic counterparts, which were far fewer than that between traditional rice varieties. The function enrichment analysis showed altered transcription in stress-related pathway and starch and sucrose metabolism. DEGs shared among eight GM samples constitute less than 1% of the genes in the genome, and none of them is reported more than four times. The insertion effect on the nearby gene expression and the associated metabolism is only restricted to 50 genes. All the results provide a comprehensive analysis of unintended effects and indication of difference in Chinese transgenic rice based on their backgrounds, transformation, and insertion elements.

KEYWORDS. differentially expressed genes; function; genetically modified rice; metabolism; *Oryza sativa* L.; transcriptome

Correspondence to: Yiqiang Zhao E-mail: yiqiangz@cau.edu.cn; Shuifang Zhu E-mail: zhusf@caiq.gov.cn

*These authors contributed equally to this work.
Received November 11, 2018; Revised March 15, 2019; Accepted 19 March 2019.

INTRODUCTION

Genetically modified (GM) crops have been developed rapidly since the mid-eighties, possessing 185.1 million hectares in. 2016.[12] The top ten countries grow more than 1 million hectares each.[1] GM crops are manipulated by inserting target DNA fragments into the host genome to improve traits, which has been developing into insect-resistant, herbicide-resistant, quality improvement, and stacked traits. For this preference, GM crops have become the possible alternative to improve grain yield and reduce labor cost.[3]

The public and regulators' concerns on GM crops mainly focus on whether GM crops are subjected to strict scrutiny for their safety.[4] It has been agreed internationally through countries and commissions to ensure an substantial equivalence between GM crops and traditional varieties, which is approved internationally.[5–8] The USA has published its research on effects of GM crops on human health, environment, agronomy, and economy; they found no adverse effects attributed to genetic engineering in human population.[9] However, the majority of commercialized GM crops are created by *Agrobacterium*-mediated gene transfer[10–12] and particle bombardment[13–15] that result in random insertion[4] and new proteins, new metabolites or altered levels of existing metabolites that could compromise safety.[8] It needs to be a comprehensive analyzing method to compare the GM crops or products to their counterparts from parental or near isogenic lines.[16,17]

The role of emerging "omics" technologies in the assessment of unintended effects was recently proposed and has been applied for the reconstruction of genome-scale networks for model organisms in the commercialized GM crops,[16,18,19] such as maize, soybean, canola, potato, and stacked GM maize variety.[20–26] Unfortunately, current studies are mostly based on individual events, and it is difficult to evaluate whether there is uniform change in genome-scale.

Rice (*Oryza sativa* L.) is one of the most important food crops, providing food for over half the world's population.[27] Researches on transgenic rice have been extensively carried out, but few are focused on the safety management.[28–30] Although seven transgenic rice cultivars are reported to be approved for cultivation or food and feed use in 2017,[1] most of them are not commercialized across the world. Rice is one of the major crops in China, and transgenic insect-resistant rice TT51-1 (synonym BT63) that harbors a hybrid Cry1Ab/Ac gene was granted a safety certificate for commercial use in 2009 as the first transgenic food crop.[31–33] However, limited published results guided us for better understanding the unintended effects of GM rice.[30,34,35]

In this study, eight genetically engineered rice lines in China and their corresponding wild types were collected. Changes of GM plants at whole genome level and the insertion site for transcriptions and metabolites compared to their counterparts were explored. We evaluated the changes in gene expression and metabolites between GM rice and different traditional maize varieties and between traditional maize varieties. This work provided a comprehensive analysis to evaluate the unintended effects of GM rice and a safety analyzing result of GM rice developed in China.

METHODS

Plant Materials

The experimental transgenic materials and their corresponding counterparts are described in Table 1, which is proved to be genetically stable by providers. GM1 and wild type 1 (WT1) were provided by Prof. Daichang Yang (Wuhan University). GM3, GM4, and WT3 were provided by Prof. Huachen Yu Chinese Academy of Sciences. GM5/WT4 and GM6/WT4 were provided by Prof. Qifa Zhang (Huazhong Agriculture University). GM8 were provided by Prof. Wensheng Huang (Chinese Academy of Inspection and Quarantine). GM2, GM7, WT2, and WT5 were stored at our laboratory. GM1 and GM2 were special because they both possess unique gene inserted to unique parental line, which is completely different from other materials.

Samples Preparation

Seeds of rice were surface sterilized with 10% H_2O_2 (v/v) for 10 min, rinsed thoroughly with

TABLE 1. Sample information included in this study.

Genetically modified (GM) materials	Event name	Wild type (WT)	Insert gene	Promoter	Terminator	Insertion site	References
GM1	114-7-2	WT1 (Taipei 309)	*Oryza sativa* recombinant human serum albumin (OsrHSA)	Rice Gt13a promoter	NOS[2]	Os04g0604600	[36]
GM2	KMD	WT2 (Xianyou 10)	Bt CrylAb(Hpt)[1]	CaMV 35S	NOS	OS02G0620533-OS 02G0620600	[37]
GM3	Huahui No.1	WT3 (Minghui 63-Beijing)	Bt CrylAb/crylAc	Actin	NOS	OS10G0173800-OS 10G0174548	[38]
GM4	Bt 63 (TT51-1)	WT3 (Minghui 63-Beijing)	Bt CrylAb/crylAc	Actin	NOS	OS10G0173800-OS 10G0174548	[33]
GM5	T2A-1	WT4 (Minghui 63-Wuhan)	Bt Cry2A	Maize ubiquitin promoter	NOS	Os12g0631600	[39]
GM6	T1C-19	WT4 (Minghui 63-Wuhan)	Bt CrylC	Maize ubiquitin promoter	NOS	OS11G0124300-OS 11G0124500	[40]
GM7	Kefeng-6	WT5 (Minghui 86)	Bt CrylAc/CptII(Hpt)[3]	Maize ubiquitin promoter	NOS	Os04g0648750	[41]
GM8	Kefeng-8	WT5 (Minghui 86)	Bt CrylAc/CptII[4]	Maize ubiquitin promoter	NOS	Os11g0430000	[41]

[1]*Bt* gene was from *Bacillus thuringiensis* to confer resistance to insects, such as Bt Cry1Ab, Bt Cry1Ac, Bt Cry2A, Bt Cry1Ab/1Ac, etc.
[2]NOS is the terminator of nopaline synthase gene.
[3]Hpt gene is hygromycin phosphotransferase gene from *Streptomyces hygroscopicus*.
[4]CptII is a cowpea trypsin inhibitor.

distilled water, and germinated on moist filter paper for 3 days in an incubator at 25°C. After germination, the seeds were then transferred to a net floating on 0.5 mM $CaCl_2$ in a plastic container. After 6 days, the seedlings were transferred to a 10-L plastic pot in a growth chamber at 400 μM photons/m^2 s and 28°C/18°C under a 14/10-h light/dark regime. The nutrient solution for rice was half-strength Kimura B solution containing the macronutrients, including 0.18 mM $(NH_4)_2SO_4$, 0.27 mM $MgSO_4{\cdot}7H_2O$, 0.09 mM KNO_3, 0.18 mM $Ca(NO_3)_2{\cdot}4H_2O$, and 0.09 mM KH_2PO_4 and the micronutrients, including 20 μM NaEDTAFe·$3H_2O$, 6.7 μM $MnCl_2{\cdot}4H_2O$, 9.4 μM H_3BO_3, 0.015 μM $(NH_4)_6Mo_7O_{24}{\cdot}4H_2O$, 0.15 μM $ZnSO_4{\cdot}7H_2O$, and 0.16 mM $CuSO_4{\cdot}5H_2O$. The nutrient solution was prepared with distilled water, aerated daily, and renewed every 3 days. The leaves were harvested at the trefoil stage, frozen with liquid nitrogen for RNA extraction. Each sample was prepared in triplicate for sampling pooling and sequencing.

RNA-Seq and Bioinformatic Analysis

Total RNA was extracted using a protocol based on the Trizol reagent (Invitrogen Life Technologies, Carlsbad, CA, USA) and purified with the Qiagen RNeasy MiniElute Cleanup kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. Extracted RNA was then quantified by a NanoDrop N2000 spectrophotometer (Thermo Fisher Scientific Inc., Wilmington, DE, USA). OD 260/280 nm absorption ratios (mean = 2.02 ± 0.02) were used to confirm the integrity and purity of the RNA samples. The library for sequencing was prepared using the Illumina mRNA-Seq 8-Sample Prep Kit following the manufacturer's instructions (Illumina, USA). The 125 base paired-end sequencing was performed on the Illumina Hiseq-2000 platform (Illumina, USA). The sequenced reads were further quality-filtered using FastQC 0.11.5 tools to remove poor-quality sequences. Clean reads were mapped to the *Oryza Sativa* reference genome IRGSP-1.0.22 (ftp://ftp.ensemblgenomes.org/pub/release-38/plants/fasta/oryza_sativa/) using TopHat 2.1.0 with the default parameters. The normalized expression level (fragments per kilobase million) for each gene was then estimated using Cufflinks 2.2.0 based on the alignment from the last step.[42] We used analysis of variance to detect the DEGs for each pair of transgenic rice line and their isogenic wild type line through in-house SAS scripts. DEGs among isogenic wild type were also calculated through this method. Principle component analysis (PCA), hierarchical clustering, and *K*-means analysis to compare the difference between GM and WT lines were performed through in-house SAS scripts.

Metabolite Profiling

The metabolomic profiling analysis was performed by Metabolon (Durham, NC, USA).[43] The metabolomic platforms consisted of three independent platforms: ultrahigh performance liquid chromatography/tandem-mass spectrometry (UPLC/MS/MS^2) optimized for basic species, UPLC/MS/MS^2 optimized for acidic species, and gas chromatography/mass spectrometry (GC/MS) for volatile species. Three biological replicates were provided for each sample.

Pathway Mapping and Analysis

Genes with Ensembl ID were integrated and mapped on gene ontology (GO) terms (http://systemsbiology.cau.edu.cn/agriGOv2/download/871_slimGO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways (http://structuralbiology.cau.edu.cn/PlantGSEA/database/Osa_KEGG) through in-house SAS scripts. The Pathview package was used to perform the KEGG pathway analysis of metabolites.[44] Metabolites of each sample at significant level were integrated and linked with related protein and genes through STITCH website 5.0 (http://stitch.embl.de/cgi/).

Insertion Analysis

Flanking sequence of each GM line was first obtained from publications and Chinese applications (Table 1). Their locations were determined by blasting analysis of flanking sequence to *Oryza*

Sativa reference genome. The coding genes surrounding the insertion site were retrieved and grouped according to the distance to the insertion site. Each cluster contained several numbers of genes. The average *P* value of genes in one cluster was calculated. These five *P* values were then used to score the inversion number to demonstrate the insertion influence on nearby genes. Such analysis was performed five times, as numbers of coding genes contained in one cluster were 1, 3, 5, 10, 20, 30, and 50, respectively. A low inversion number, which should be derived from the consecutive increase of average *P* value in the further adjacent clusters, would indicate the existence of hierarchical continuous effect due to gene insertion. Negative control was calculated based on the corresponding insertion site and the nearby transcriptome of non-GM materials.

qRT-PCR

Total RNAs were isolated with a Trizol reagent (Invitrogen, CA, USA). First-strand cDNA was synthesized using a ReverTra Ace® qPCR RT Kit (TOYOBO, Japan), and PCR reaction was performed with a SYBR Green® Realtime PCR Master Mix (TOYOBO, Japan) on a LightCycler 480 II (Roche, CH). Each reaction mix contains 2 μL cDNA, 0.3 μM both primers, 10 μL Master Mix, and ddH_2O was added to form a total volume of 20 μL. The PCR reaction condition was as follows: 95°C for 1 min, following 95°C for 10 s, 60°C for 60 s for 40 cycles. A fluorescence signal was detected at the end of each cycle. Three replicates were arranged for every data point. Actin was used as an internal control.

RESULTS

GM Rice Lines Show Less Differences to Wild Counterparts than WT Comparisons

We obtained more than 903 million 125 bp pair-end reads that passed the quality filters in these 13 samples; each of which is analyzed by triplicate sequencing data. The metabolism data obtained from GC–MS and LC–MS are also satisfied for further analysis (5: Figure S1 and 6: Figure S2).

A total of 218 metabolites and 35,678 genes were quantified across 13 lines, and the DEGs have been selectively verified through q-PCR experiment (Table S4). There are 2892–8758 DEGs ($P < 0.05$) between GM variety and their isogenic counterparts (Table 2), which were far fewer than that between traditional rice varieties (*t*-test, DF = 16, $t = 3.31$, $P = 0.004$). The same trend was also found at the level of metabolites (Table 2), which means that only 7–50 metabolites at significant level were found in WG (differential analysis between genetically modified rices and their corresponding counterparts) comparisons (*t*-test, DF = 16, $t = 2.77$, $P = 0.014$). After subtracting those overlapped genes that existed in both WW (differential analysis between two isogenic counterparts in this research) and WG comparisons, the number of DEGs (unique mRNAs in Table 2) was decreased to 164–667. These results indicate that transcriptional turbulences caused by transgenic constructs are much less than that by genetic and environmental influence. This is consistent with the previous findings that environment and/or naturally species variations affect gene expression more strongly than the genetic modification.[21,45]

PCA was further used to show the differences between the WG and the WW comparisons (Fig. 1). About 45.6% of the total variance was explained by the first three PCs (PC1, 26%; PC2, 13%; and PC3, 5%), suggesting that these wild types are moderately separated by gene expression patterns (Fig. 1A). GM samples tend to be grouped with the respective non-GM counterparts in plots (Fig. 1B). Analysis on metabolites shows similar trends as the expression data (Fig. 1C,D). Hierarchical clustering showed that each GM sample and its corresponding counterpart belong to the same class (Fig. 1). This is the first comprehensive analysis on major Chinese GM rice events that GM events seem to introduce less variation on the transcripts and metabolites than that between non-GM counterparts.

TABLE 2. Significant gene expressions and metabolites summary of all GM rice varieties and their corresponding counterparts.

WW comparisons[1]	Metabolites	mRNAs	Ratio of metabolites	Ratio of mRNAs	WG comparisons[2]	Metabolites	mRNAs	Ratio of metabolites	Ratio of mRNAs	Unique mRNAs
W12	74	9538	33.94%	26.73%	WT1/GM1(WG1)	33	2566	15.14%	7.19%	216
W13	88	11454	40.37%	32.10%	WT2/GM2(WG2)	50	5174	22.94%	14.50%	298
W14	54	12403	24.77%	34.76%	WT3/GM3(WG3)	15	2892	7.34%	8.11%	185
W15	58	16390	26.61%	45.94%	WT3/GM4(WG4)	10	4977	4.59%	13.95%	667
W23	56	3954	25.69%	11.08%	WT4/GM5(WG5)	7	8758	3.21%	24.55%	164
W24	25	8413	11.47%	23.58%	WT4/GM6(WG6)	8	3681	3.67%	10.32%	167
W25	19	9880	8.72%	27.69%	WT5/GM7(WG7)	23	7301	10.55%	20.46%	250
W34	30	4072	13.76%	11.41%	WT5/GM8(WG8)	19	3027	8.72%	8.48%	396
W35	52	9651	23.85%	27.05%						
W45	18	12456	8.26%	34.91%						

[1]WW comparisons mean differential analysis between two isogenic counterparts in this research. For example, W12 means comparison between WT1 and WT2, and so forth.
[2]WG comparisons mean differential analysis between genetically modified rice varieties and their corresponding counterparts. For example, WG1 means comparison between WT1 and GM1, and so forth.

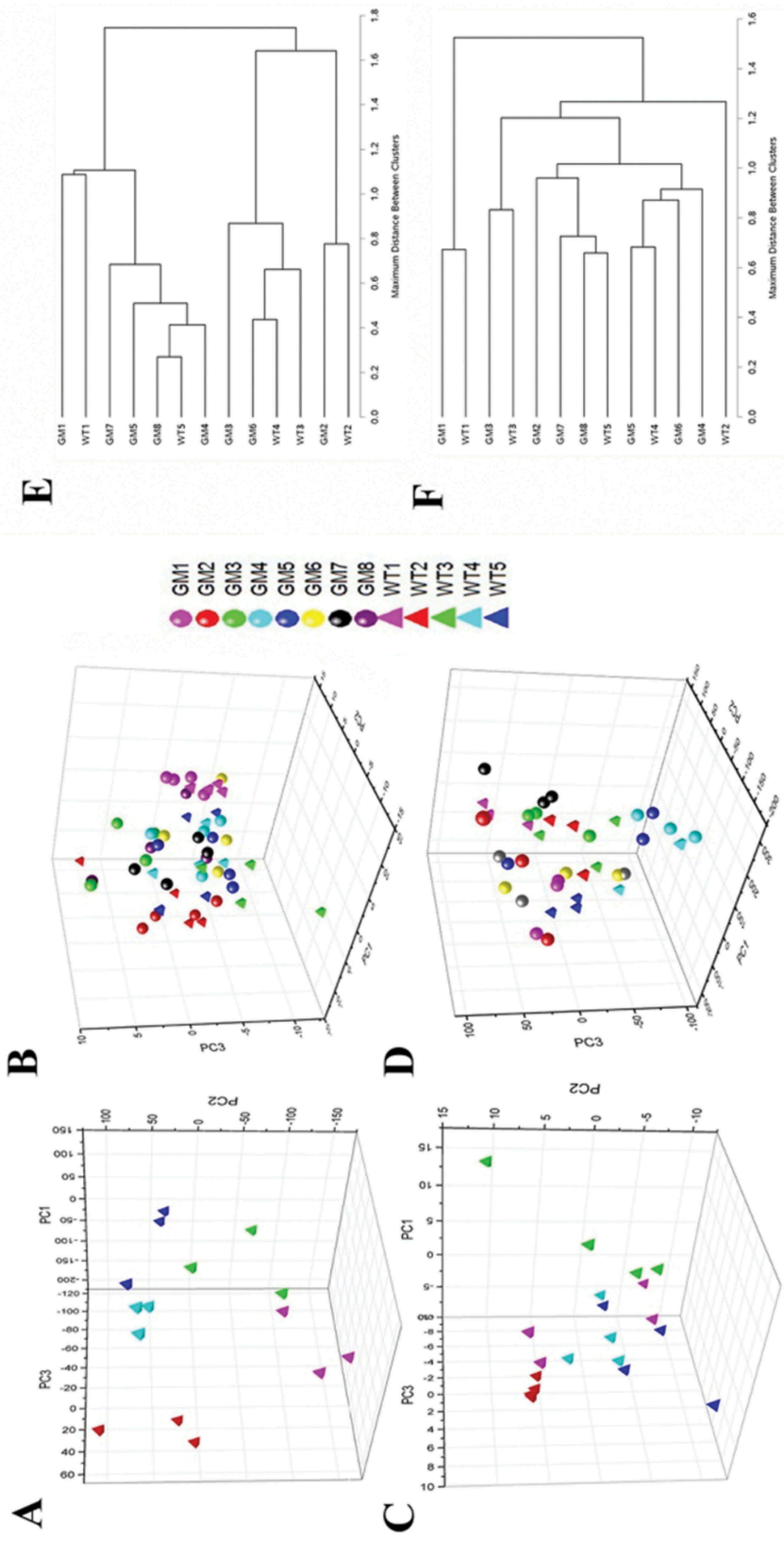

FIGURE 1. PCA and clustering analysis of transcripts (A, B, and E) and metabolites (C, D, and F) of GM and WT samples, respectively.

Few Overlapped DEGs, Metabolites, and Enriched Functions were Found Between GM Lines

For unique DEGs reported by each WG comparison, less than 15% were reported two times. Sixty six and eight genes were reported three or four times, respectively (Table S1). The function of these genes mainly attributed to response to external stimulus or stress and transformation is obviously one of these stresses (Table 3). The metabolites analysis offered the similar conclusions, which means that nine overlapped metabolites that reported three times at significant concentration only influenced amino acid metabolism between WG comparisons (Table S1).

The *K*-means analysis further revealed that WG groups are close to each other based on their parental genotype and the insertion element (Table 4), such as WG7 and WG8 that belong to the same class because of their similar parent origin. However, GM5 and GM6 are also based on the same parent background but belong to two different groups, indicating that effects caused by transgene are not solely parent dependent. Moreover, GM5 and GM6 are based on the same isogenic wild type counterpart (Minghui 63) and shared 18 DEGs and no metabolites at significant level, which is mainly related to RNA splicing and biosynthesis of secondary metabolites (Table S2). GM3 and GM4 were developed using the same construct and transformation method (pFHBT1) and shared four DEGs and no metabolites. Difference between GM rice and non-GM counterparts is related to insertion element, background, and other factors and should be evaluated case-by-case.

The GO analysis mainly contributed to the process of macromolecule metabolism, DNA and RNA biosynthesis, and plant stress/defense signaling. The representative GO and KEGG enrichment with a cutoff $P < 0.05$ is shown in Table 5 and Table S3. Most GO overlaps contributed to the biological and energy metabolism, which appeared five times. GM3 was distantly placed with WT3 in PCA analysis

TABLE 4. *K*-means analysis of WG comparisons.

	C2	C3	C4	C5*
WG1	1	1	1	1
WG2	2	3	4	3
WG3	2	3	2	2
WG4	2	3	4	5
WG5	2	3	4	5
WG6	2	2	3	4
WG7	2	3	4	5
WG8	2	3	4	5

*C2–C5 means clusters are classified into 2, 3, 4, and 5, respectively.

TABLE 3. List of gene IDs that reported four times and KEGG compound IDs that reported three times.

Gene IDs	Times	Function	Compound IDs	Times	Metabolites
OS05G0144900	4	Sucrose-phosphatase	C00022	3	Pyruvate
OS06G0222100	4	Trehalose-phosphate phosphatase 8	C00025	3	Glutamate
OS07G0463600	4	Pentadecatrienyl resorcinol *O*-methyltransferase	C00072	3	Ascorbate (vitamin C)
OS08G0459600	4	12-Oxophytodienoate reductase 7	C00073	3	Methionine
OS11G0135000	4	Zinc induced facilitator	C00847	3	Pyridoxate
OS10G0188100	4	Transcription initiation factor IIF subunit alpha	C00864	3	Pantothenate
OS02G0620400	4	Uncharacterized	C01035	3	4-Guanidinobutanoate
OS02G0690000	4	Uncharacterized	C01235	3	Galactinol
			C10646	3	Lariciresinol

TABLE 5. GO and KEGG enrichment of significantly induced and repressed genes for all WG comparisons.

Description	Times	WG comparisons
GO overlap		
Regulation of transcription, DNA dependent	5	WG2/WG4/WG5/WG7/WG8
Regulation of primary metabolic process	5	WG2/WG4/WG5/WG7/WG8
Regulation of nucleobase-containing compound metabolic process	5	WG2/WG4/WG5/WG7/WG8
Regulation of nitrogen compound metabolic process	5	WG2/WG4/WG5/WG7/WG8
Regulation of macromolecule metabolic process	5	WG2/WG4/WG5/WG7/WG8
Regulation of cellular metabolic process	5	WG2/WG4/WG5/WG7/WG8
Regulation of biosynthetic process	5	WG2/WG4/WG5/WG7/WG8
RNA biosynthetic process	3	WG4/WG6/WG8
Defense response	3	WG1/WG2/WG3
Regulation of biological process	2	WG2/WG8
Programmed cell death	2	WG2/WG3
Photosynthesis	2	WG1/WG7
Lipid transport	2	WG2/WG7
Death	2	WG2/WG3
Cell death	2	WG2/WG3
Apoptotic process	2	WG2/WG3
KEGG overlap		
Metabolic pathways	8	WG1–WG8
Biosynthesis of secondary metabolites	6	WG1–WG6
Starch and sucrose metabolism	3	WG2, WG5, WG8
Ubiquitin mediated proteolysis	2	WG2, WG7
Plant hormone signal transduction	2	WG2, WG7
Photosynthesis	2	WG1, WG3

(Fig. 1); however, the GO analysis of WG3 was mainly enriched with defense response. Of all the six KEGG pathways, metabolic pathways exist in all eight GM events. Three GM events (WG2, WG5, and WG8) showed transcriptional modification in starch and sucrose metabolism that is related to plant defense response to the biotic or environment stress.[46–48] These results tend to indicate that the acquisition of desired traits is accompanied by transcript levels of regulation of transcriptional process that is associated with environmental stress responses.[49–52]

Metabolites are Not Linked with DEGs in Response to Transgene

In order to investigate the effect of transgene on the entire metabolic pathways that functioned from transcript to metabolites, we linked the metabolites in each WG comparison to the corresponding coding genes and evaluated their expression. Each WG comparison has no more than two DEGs that are directly linked to metabolism profiling; most of which were related to citrate cycle and amino acid metabolism (Table 6). Additionally, WG1, WG3, and WG6 all showed variation on the defense response, mainly focused on the transmembrane transportation, metal tolerance, and transcription activation.

Insertion Analysis Suggest only Short Range Affections

By blasting analysis of the recorded flanking sequences from each transgenic event to *Oryza Sativa* reference genome, we identified the insertion sites of each transgenic event in the gnome. Results showed that the insertion sites of WG1, WG5, WG7, and WG8 were located at the coding sequence, while all others are located on the intergenic region. The insertion site was also analyzed to determine if it locates in the high-

TABLE 6. Differentially expressed genes that associated with metabolites and their functions.

WG comparisons	Gene ID	Description	Category
WG1	OS10G0493600 OS07G0461900	Alpha-galactosidase Acetylornithine aminotransferase	Metabolism/Defense response Metabolism
WG2	OS06G0256500	Glucose-6-phosphate	Metabolism
WG3	OS12G0221700 OS07G0689600	Leucine-rich repeat receptor Nicotianamine synthase	Signal reception and transduction Defense response
WG4	OS03G0718000	Anthranilate synthase beta subunit	Metabolism
WG5	OS03G0136900	Aconitate hydratase	Metabolism
WG6	OS09G0567900 OS01G0316100	Uridine nucleosidase Sarcosine oxidase	Metabolism Stress/Defense response
WG7	OS03G0136900	Aconitate hydratase	Metabolism
WG8	–	–	–

expression region or gene enrichment region, as these regions seemed much significant on the gene expression and the corresponding metabolic regulation. By randomly selecting sampling gene expression for 5000 times, the insertion site of GM5 was determined to be located on the relatively high-expression region.

To test whether insertion results in a server disturb on gene expression for the nearby genes than the remote genes, we calculated the inversion number (the cardinality of inversion set of a set of numbers) for each sample (step size in Table 7). For each test in the WG comparison, a higher *P* value indicates less changes at transcriptional level. To describe the trends of ripple effect, we calculated the inversion number based on *P* values for the adjacent clusters. We calculated the inversion number of each step size. As shown in Table 6, we did not observe zero for any step size, and we did not observe a clear trend of increasing inversion number with increasing step size either. It suggests that the ripple effect might not exist. The insertion event would only bring short range effect. By using random WW comparison as negative control, we recalculated the inversion number and found the similar result with that from the canonical WG comparison, which confirms the sporadic effect on the nearby genes. We next test the short-range effect by checking if the *P* value of the first cluster (P1) is the smallest in that of all the five clusters. By using the results from 100 permutations as negative control, we found GM3 and GM4, but not GM5-exhibited influence on up to three nearest genes, while all other insertion events failed to pass the threshold from the null distribution of inversion number. Altogether, the insertion events did not impose long-range effects on the gene expression; even if it located on the high expression region. Even in few cases, short-range effects are observed, and it seems not to be a common, but case-specific, pattern.

TABLE 7. Inversion number calculation of each stepsize between GM materials and negative control.

GM materials							Negative control					
Stepsize	**3**	**5**	**10**	**20**	**30**	**50**	**3**	**5**	**10**	**20**	**30**	**50**
WG1	**6***	7	6	**6**	**7**	7	4	7	7	8	8	9
WG2	**3**	7	6	7	4	6	7	9	9	6	6	**6**
WG3	7	**4**	**5**	3	**6**	6	11	11	10	11	11	9
WG4	3	5	**2**	**5**	**6**	**6**	11	12	10	10	11	9
WG5	**5**	**6**	10	10	11	7	10	9	5	**7**	**7**	**8**
WG6	7	**6**	6	9	9	8	9	12	9	11	10	7
WG7	**2**	**7**	7	11	11	7	8	9	11	7	6	6
WG8	12	10	6	10	9	8	5	7	8	7	**6**	**7**

*Number in bold indicates that this stepsize satisfied the condition that *P* value of the first cluster is the biggest in that of all the five clusters.

DISCUSSION

Unintended effects of GM crops are the main concern of administrators and citizens after their intended effects have been thoroughly investigated. Many efforts have been made to test if

GM crops are substantially equivalent as safe non-GM comparators, following a large body of high-quality compositional data (around 50–150 analytes) that need to be determined according to principles outlined in the Organization for Economic Cooperation and Development.[7] This concern promoted the use of modern technologies, especially an unbiased high-throughput "omics" technology to validate the substantial equivalence.[16] In these transgenic rice varieties, *Bt* genes and the expressed proteins are not native to plants and exert no known metabolic activity in plant, and so is the OsrHSA recombinant protein. However, we still found eight DEGs that reported four times in WG comparisons, especially in insect-resistant *Bt* rice (Table 3). Of these unintended DEGs, two are related to starch and sucrose metabolism, representing sucrose-phosphatase (OS05G0144900) and trehalose-phosphate phosphatase (OS06G0222100), respectively. These enzymes are mainly involved in the sugar metabolism and photosynthesis.[53,54] Under drought stress conditions, the critical roles of sucrose, glycine-betaine, and trehalose in the starch and sucrose metabolism (ko00500) have been actively researched to understand the tolerance of plants to biotics, dehydration, and insect invasion.[46–48] Other genes are related to plant response to abiotic response, such as *OPR* (OS08G0459600) that is functional in the metabolism of jasmonic acid biosynthesis,[55] *TFIIF* (OS10G0188100) that is involved in the process of transcription,[56] and *ZIF* (OS11G0135000) that involved in the signaling of plant immunity.[57] Additionally, DEGs that were linked directly to metabolism profiling were mainly involved in amino acid metabolism, which could partly because of the increased demand of inserted protein biosynthesis.[58] These DEGs expressed in GM rice should be further studied for their edible and environmental safety.

The environment was even shown to play a stronger effect in the protein, gene expression, and metabolite levels of the GM samples than gene modification by previous publications, which showed similar results for transcriptomics in maize, wheat, soybean, and potato.[20–22,25,45,59] Our research on GM rice developed in China indicated less variation in gene expression and metabolites than non-GM varieties (Table 2). Only half of GM events shared DEGs without function enrichment so that gene insertion did not pose metabolic change on the rice physiology. WG2 and WG3 are associated with some stress-related mechanisms such as programmed cell death and apoptotic process.[60–62] However, *Bt* rice is resistant to insects through making them lethal after taking *Bt* protein, which inevitably stimulate some defense pathways according to the previous report.[63] High-level expression of *Bt* genes as well as maker genes and artificial sequence of gene regulators in this rice may act as a physiological sensor to disturb stress-related pathways. This observation has been proved in the assessment of GM2 and GM3 previously.[58,63,64] These could possibly be common effects caused by *Bt* transgene. However, different DEGs and metabolites in significance existed in WG comparisons suggest that this effect may also be related to construct, background, and transformation method. Moreover, gene insertion-induced pathological variation occurred in WG3, but not in WG4 that was cultivated with GM3 and Zhenshan 97, providing an implication that crossover between GM and dominant rice that compromise the unintended effects to stabilize the genetics of inserted gene. Gene insertion is indicative of the regulation of stress-related genes and pathways that may act a role to protect plant from environmental changes and should fully be evaluated for plant genetic transformation technologies. This observation was also validated through linkage analysis between DEGs and metabolites at significant level, where nine candidate DEGs that linked metabolites at significant level attributes to metabolic process and stress-related pathways. This result also suggested that GM rice by inserting foreign genes was much like the process of exposure to abiotic stress, which induced the expression of defense genes for adaptation.[46,47]

This study provides an unbiased, comprehensive profiling of unintended effects in GM rice. Considering their breeding development,

the extensive vetting is involved in the generation and selection of one or a few "elite" events, which artificially prevents the "bad" GM crops that might produce "bad" metabolites or unintended effects for food and feed use.[26] On the other hand, it is also revealed in this study that difference of metabolite levels is less than that of transcripts level, no matter in GM or WT lines (Figure S3). This is also corroborated by previous studies on variation in human chromatin and primate transcript state.[65,66] GM rice is genetically stable through molecular breeding and undergoing assessment to be approved for commercialization, and their unintended effects based on omics technologies should be evaluated further.

CONCLUSIONS

Transcriptomic and metabolic analysis of Chinese transgenic rice that are potential for further commercialization and their isogenic counterparts were conducted in this study. Changes of transgene insertion in rice exposed less difference than other rice varieties. Difference between GM rice and non-GM counterparts is limited to defense-related pathways and is related to insertion element, background, and transformation process. Gene insertion in the genome does not impose long-range effects on the gene expression, even if it located on the high expression region.

ACKNOWLEDGMENTS

We thank Prof. Wentao Xu, Prof. Jianbing Yan, and Prof. Daoxin Xie for providing useful suggestions.

FUNDING

This work is supported by National Science and Technology Major Project: [Grant Number 2018ZX08012-001].

DISCLOSURE OF POTENTIAL CONFLICTS OF INTEREST

No potential conflicts of interest were disclosed.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

Not applicable.

AVAILABILITY OF DATA AND MATERIALS

RNA-seq read data have been deposited on the NCBI database under the accession number SRA230931-SRA230941.

COMPETING INTERESTS

The authors declare that they have no competing interests.

AUTHOR CONTRIBUTIONS

FW and WCG performed the experiment and analyzed the gene and metabolites mapping data. ZPY and XWJ performed the GO and KEGG analysis. ZYQ and ZSF conceived and designed the experiment. FW and ZSF provided the project resources. FW, WCG, and XWJ wrote the manuscript.

SUPPLEMENTARY MATERIAL

Supplemental data for this article can be accessed on the publisher's website.

ABBREVIATIONS

ANOVA	analysis of variance
DEGs	differentially expressed genes
FPKM	fragments per kilobase million

GM	genetically modified
GO	gene ontology
KEGG	Kyoto Encyclopedia of Genes and Genomes
OPR	Oxophytodienoate reductase
PCA	principle component analysis
qPCR	quantitative polymerase chain reaction
TFIIF	transcription initiation factor IIF
WT	wild type
ZIF	zinc-induced facilitator

REFERENCES

1. James C (2018) Global Status of Commercialized Biotech/GM Crops: 2017 ISAAA Brief
2. Qaim M, Zilberman D. Yield effects of genetically modified crops in developing countries. Science. 2003;299:900–02. doi:10.1126/science.1080609.
3. Aldemita RR, Reano IM, Solis RO, Hautea RA. Trends in global approvals of biotech crops (1992-2014). GM Crops Food. 2015;6:150–66. doi:10.1080/21645698.2015.1056972.
4. Baker JM, Hawkins ND, Ward JL, Lovegrove A, Napier JA, Shewry PR, Beale MH. A metabolomic study of substantial equivalence of field-grown genetically modified wheat. Plant Biotechnol J. 2006;4:381–92. doi:10.1111/j.1467-7652.2006.00197.x.
5. FAO/WHO (2000) Safety Aspects of Genetically Modified Foods of Plant Origin Report of a Joint FAO/WHO Expert Consultation on Foods Derived from Biotechnology
6. Kok EJ, Kuiper HA. Comparative safety assessment for biotech crops. Trends Biotechnol. 2003;21:439–44. doi:10.1016/j.tibtech.2003.08.003.
7. OECD (2006) An introduction to the food/feed safety consensus documents of the task force. Series on the Safety of Novel Foods and Feeds, No.14.:7–9
8. Kuiper HA, Kleter GA, Noteborn HP, Kok EJ. Assessment of the food safety issues related to genetically modified foods. Plant J. 2001;27:503–28.
9.. National Academies of Sciences E, Medicine. 2017. Genetically engineered crops: experiences and prospects. Washington (DC): National Academies Press.
10. Kim HA, Utomo SD, Kwon SY, Min SR, Kim JS, Yoo HS, Choi PS. The development of herbicide-resistant maize: stable Agrobacterium-mediated transformation of maize using explants of type II embryogenic calli. Plant Biotechnol Rep. 2009;3:277–83. doi:10.1007/s11816-009-0099-2.
11. Zhao Z-Y, Gu W, Cai T, Tagliani L, Hondred D, Bond D, Schroeder S, Rudert M, Pierce D. High throughput genetic transformation mediated by Agrobacterium tumefaciens in maize. Mol Breed. 2002;8:323–33. doi:10.1023/A:1015243600325.
12. Gould J, Devey M, Hasegawa O, Ulian EC, Peterson G, Smith RH. Transformation of Zea mays L. using Agrobacterium tumefaciens and the shoot apex. Plant Physiol. 1991;95:426–34.
13. Vain P, McMullen MD, Finer JJ. Osmotic treatment enhances particle bombardment-mediated transient and stable transformation of maize. Plant Cell Rep. 1993;12:84–88. doi:10.1007/BF00241940.
14. Gordon-Kamm WJ, Spencer TM, Mangano ML, Adams TR, Daines RJ, Start WG, O'Brien JV, Chambers SA, Adams WR Jr., Willetts NG, et al. Transformation of maize cells and regeneration of fertile transgenic. Plants Plant Cell. 1990;2:603–18. doi:10.1105/tpc.2.7.603.
15. Klein TM, Kornstein L, Sanford JC, Fromm ME. Genetic transformation of maize cells by particle bombardment. Plant Physiol. 1989;91:440–44.
16. Ricroch AE. Assessment of GE food safety using '-omics' techniques and long-term animal feeding studies. N Biotechnol. 2013;30:349–54. doi:10.1016/j.nbt.2012.12.001.
17. Strauss SH, Sax JK. Ending event-based regulation of GMO crops. Nat Biotechnol. 2016;34:474–77. doi:10.1038/nbt.3541.
18. Bassel GW, Gaudinier A, Brady SM, Hennig L, Rhee SY, De Smet I. Systems analysis of plant functional, transcriptional, physical interaction, and metabolic networks. Plant Cell. 2012;24:3859–75. doi:10.1105/tpc.112.100776.
19. Pauwels K, De Keersmaecker SC, De Schrijver A, Du Jardin P, Roosens NH, Herman P. Next-generation sequencing as a tool for the molecular characterisation and risk assessment of genetically modified plants: added value or not? Trends Food Sci Technol. 2015;45:319–26. doi:10.1016/j.tifs.2015.07.009.
20. Cheng K, Beaulieu J, Iquira E, Belzile F, Fortin M, Strömvik M. Effect of transgenes on global gene expression in soybean is within the natural range of variation of conventional cultivars. J Agric Food Chem. 2008;56:3057–67. doi:10.1021/jf073505i.
21. Barros E, Lezar S, Anttonen MJ, Van Dijk JP, Röhlig RM, Kok EJ, Engel KH. Comparison of two GM maize varieties with a near-isogenic non-GM variety using transcriptomics, proteomics and metabolomics. Plant Biotechnol J. 2010;8:436–51. doi:10.1111/j.1467-7652.2009.00487.x.
22. Batista R, Saibo N, Lourenço T, Oliveira MM. Microarray analyses reveal that plant mutagenesis may induce more transcriptomic changes than transgene insertion. Proc Natl Acad Sci USA. 2008;105:3640–45. doi:10.1073/pnas.0707881105.
23. Wang X, Zhang X, Yang J, Wang Z. Effect on transcriptome and metabolome of stacked transgenic

maize containing insecticidal cry and glyphosate tolerance epsps genes. Plant J. 2018. doi:10.1111/tpj.13825.
24. Llorente B, Alonso GD, BravooAlmonacid F, Rodríguez V, López MG, Carrari F, Torres HN, Flawiá MM. Safety assessment of nonbrowning potatoes: opening the discussion about the relevance of substantial equivalence on next generation biotech crops. Plant Biotechnol J. 2011;9:136–50. doi:10.1111/j.1467-7652.2010.00534.x.
25. Baudo MM, Lyons R, Powers S, Pastori GM, Edwards KJ, Holdsworth MJ, Shewry PR. Transgenesis has less impact on the transcriptome of wheat grain than conventional breeding. Plant Biotechnol J. 2006;4:369–80. doi:10.1111/pbi.2006.4.issue-4.
26. Ladics GS, Bartholomaeus A, Bregitzer P, Doerrer NG, Gray A, Holzhauser T, Jordan M, Keese P, Kok E, Macdonald P. Genetic basis and detection of unintended effects in genetically modified crop plants. Transgenic Res. 2015;24:587–603. doi:10.1007/s11248-015-9867-7.
27. Muthayya S, Sugimoto JD, Montgomery S, Maberly GF. An overview of global rice production, supply, trade, and consumption. Ann N Y Acad Sci. 2014;1324:7–14. doi:10.1111/nyas.12540.
28. Toriyama K, Arimoto Y, Uchimiya H, Hinata K. Transgenic rice plants after direct gene transfer into protoplasts. Nat Biotechnol. 1988;6:1072–74. doi:10.1038/nbt0988-1072.
29. Bajaj S, Mohanty A. Recent advances in rice biotechnology—towards genetically superior transgenic rice. Plant Biotechnol J. 2005;3:275–307. doi:10.1111/j.1467-7652.2005.00130.x.
30. Montero M, Coll A, Nadal A, Messeguer J, Pla M. Only half the transcriptomic differences between resistant genetically modified and conventional rice are associated with the transgene. Plant Biotechnol J. 2011;9:693–702. doi:10.1111/j.1467-7652.2010.00572.x.
31. Lu C. The first approved transgenic rice in China. GM Crops. 2010;1:113–15. doi:10.4161/gmcr.1.3.12377.
32. Lu B-R. Challenges of transgenic crop commercialization in China. Nat Plants. 2016;2:16077. doi:10.1038/nplants.2016.77.
33. Tu J, Zhang G, Datta K, Xu C, He Y, Zhang Q, Khush GS, Datta SK. Field performance of transgenic elite commercial hybrid rice expressing Bacillus thuringiensis δ-endotoxin. Nat Biotechnol. 2000;18:1101–04. doi:10.1038/80310. https://www.nature.com/articles/nbt1000_1101
34. Harrigan GG, Lundry D, Drury S, Berman K, Riordan SG, Nemeth MA, Ridley WP, Glenn KC. Natural variation in crop composition and the impact of transgenesis. Nat Biotechnol. 2010;28:402–04. doi:10.1038/nbt0510-402.
35. Wang Y, Lan Q, Zhao X, Xu W, Li F, Wang Q, Chen R. Comparative profiling of microRNA expression in soybean seeds from genetically modified plants and their near-isogenic parental lines. PLoS One. 2016;11:e0155896. doi:10.1371/journal.pone.0155896.
36. Zhang L, Huang X, Zhu S. An event-specific real-time PCR detection system for the transgenic rice line 114-7-2 of producing functional human serum albumin. Eur Food Res Technol. 2014;239:403–08. doi:10.1007/s00217-014-2234-8.
37. Babekova R, Funk T, Pecoraro S, Engel KH, Busch U. Development of an event-specific Real-time PCR detection method for the transgenic Bt rice line KMD1. Eur Food Res Technol. 2009;228:707–16. doi:10.1007/s00217-008-0981-0.
38. Gang W, Wu Y, Nie S, Li Z, Ling X, Cao Y, Lu C. Real-time PCR method for detection of the transgenic rice event TT51-1. Food Chem. 2010;119:417–22. doi:10.1016/j.foodchem.2009.08.031.
39. Chen H, Tang W, Xu C, Li X, Lin Y, Zhang Q. Transgenic indica rice plants harboring a synthetic cry2A* gene of Bacillus thuringiensis exhibit enhanced resistance against lepidopteran rice pests. Theor Appl Genet. 2005;111:1330–37. doi:10.1007/s00122-005-0062-8.
40. Tang W, Chen H, Xu C, Li X, Lin Y, Zhang Q. Development of insect-resistant transgenic indica rice with a synthetic cry1C* gene. Mol Breed. 2006;18:1–10. doi:10.1007/s11032-006-9002-9.
41. Wang WX, Zhu TH, Lai FX, Fu Q. Event-specific qualitative and quantitative detection of transgenic rice Kefeng-8 by characterization of the transgene flanking sequence. Eur Food Res Technol. 2012;234:477–484. https://link.springer.com/article/10.1007/s00217-011-1654-y.
42. Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL, Pachter L. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc. 2012;7:562.
43. Krumsiek J, Suhre K, Evans AM, Mitchell MW, Mohney RP, Milburn MV, Wägele B, Römisch-Margl W, Illig T, Adamski J. Mining the unknown: a systems approach to metabolite identification combining genetic and metabolic information. PLoS Genet. 2012;8:e1003005. doi:10.1371/journal.pgen.1003005.
44. Luo W, Brouwer C. Pathview: an R/Bioconductor package for pathway-based data integration and visualization. Bioinformatics. 2013; 29:1830–31. doi:10.1093/bioinformatics/btt285.
45. Kogel K-H, Voll LM, Schäfer P, Jansen C, Wu Y, Langen G, Imani J, Hofmann J, Schmiedl A, Sonnewald S. Transcriptome and metabolome profiling of field-grown transgenic barley lack induced

differences but show cultivar-specific variances. Proc Natl Acad Sci USA. 2010;107:6198–203. doi:10.1073/pnas.1001945107.
46. Tauzin AS, Giardina T (2014) Sucrose and invertases, a part of the plant defense response to the biotic stresses
47. Singh V, Louis J, Ayre BG, Reese JC, Shah J. TREHALOSE PHOSPHATE SYNT HASE11response totrehalose metabolism promotes Arabidopsis thaliana defense against the phloembiotic stressescMyzus persicae. Plant J. 2011 ;67:94–104. doi:10.1111/j.1365-313X.2011.04583.x.
48. Reddy AR, Chaitanya KV, Vivekanandan M. Drought-induced responses of photosynthesis and antioxidant metabolism in higher plants. J Plant Physiol. 2004;161:1189–202.
49. Shiu S-H, Karlowski WM, Pan R, Tzeng Y-H, Mayer KF, Li W-H. Comparative analysis of the receptor-like kinase family in Arabidopsis and rice. Plant Cell. 2004;16:1220–34. doi:10.1105/tpc.020834.
50. Mahalingam R, Gomez-Buitrago A, Eckardt N, Shah N, Guevara-Garcia A, Day P, Raina R, Fedoroff NV. Characterizing the stress/defense transcriptome of Arabidopsis. Genome Biol. 2003;4: R20. doi:10.1186/gb-2003-4-3-r20.
51. Lee I, Seo Y-S, Coltrane D, Hwang S, Oh T, Marcotte EM, Ronald PC. Genetic dissection of the biotic stress response using a genome-scale gene network for rice. Proc Natl Acad Sci USA. 2011;108:18548–53. doi:10.1073/pnas.1110384108.
52. Xiong L, Zhu JK. Abiotic stress signal transduction in plants: molecular and genetic perspectives. Physiol Plant. 2001;112:152–66.
53. Anh NQ, Sheng L, Wi SG, Hanhong B, Dae-Seok L, Hyeun-Jong B. Pronounced phenotypic changes in transgenic tobacco plants overexpressing sucrose synthase may reveal a novel sugar signaling pathway. Front Plant Sci. 2015;6:1216. doi:10.3389/fpls.2015.01216.
54. Rolland F, Baena-Gonzalez E, Sheen J. Sugar sensing and signaling in plants: conserved and novel mechanisms. Annu Rev Plant Biol. 2006;57:675. doi:10.1146/annurev.arplant.57.032905.105441.
55. Kazan K. Diverse roles of jasmonates and ethylene in abiotic stress tolerance. Trends Plant Sci. 2015;20:219–29. doi:10.1016/j.tplants.2015.02.001.
56. Blombach F, Smollett KL, Grohmann D, Werner F. Molecular mechanisms of transcription initiation—structure, function, and evolution of TFE/TFIIE-like factors and open complex formation. J Mol Biol. 2016;428:2592–606. doi:10.1016/j.jmb.2016.04.016.
57. Laloum T, Martín G, Duque P. Alternative splicing control of abiotic stress responses. Trends Plant Sci. 2017;23:140–150. doi: 10.1016/j.tplants.2017.09.019.
58. Liu Z, Li Y, Zhao J, Chen X, Jian G, Peng Y, Qi F. Differentially expressed genes distributed over chromosomes and implicated in certain biological processes for site insertion genetically modified rice Kemingdao. Int J Biol Sci. 2012a;8:953–63. doi:10.7150/ijbs.4527.
59. Catchpole GS, Beckmann M, Enot DP, Mondhe M, Zywicki B, Taylor J, Hardy N, Smith A, King RD, Kell DB. Hierarchical metabolomics demonstrates substantial compositional similarity between genetically modified and conventional potato crops. Proc Natl Acad Sci USA. 2005;102:14458–62. doi:10.1073/pnas.0503955102.
60. Liu J-X, Howell SH. Endoplasmic reticulum protein quality control and its relationship to environmental stress responses in plants. Plant Cell. 2010;22:2930–42. doi:10.1105/tpc.110.078154.
61. Heil M, Ton J. Long-distance signalling in plant defence. Trends Plant Sci. 2008;13:264–72. doi:10.1016/j.tplants.2008.03.005.
62. Blein J-P, Coutos-Thévenot P, Marion D, Ponchet M. From elicitins to lipid-transfer proteins: a new insight in cell signalling involved in plant defence mechanisms. Trends Plant Sci. 2002;7:293–96.
63. Liu Z, Zhao J, Li Y, Zhang W, Jian G, Peng Y, Qi F. Non-uniform distribution pattern for differentially expressed genes of transgenic rice Huahui 1 at different developmental stages and environments. PLoS One. 2012b;7:e37078. doi:10.1371/journal.pone.0037078.
64. Miki B, Abdeen A, Manabe Y, MacDonald P. Selectable marker genes and unintended changes to the plant transcriptome. Plant Biotechnol J. 2009;7:211–18. doi:10.1111/j.1467-7652.2009.00400.x.
65. Khan Z, Ford MJ, Cusanovich DA, Mitrano A, Pritchard JK, Gilad Y. Primate transcript and protein expression levels evolve under compensatory selection pressures. Science. 2013;342:1100–04. doi:10.1126/science.1242379.
66. Kasowski M, Kyriazopoulou-Panagiotopoulou S, Grubert F, Zaugg JB, Kundaje A, Liu Y, Boyle AP, Zhang QC, Zakharia F, Spacek DV. Extensive variation in chromatin states across humans. Science. 2013;342:750–52. doi:10.1126/science.1242510.