J. Biochem. 2020;167(1):67–78 doi:10.1093/jb/mvz081

JB THE JOURNAL OF BIOCHEMISTRY

iTRAQ-based quantitative proteomic analysis of two transgenic soybean lines and the corresponding non-genetically modified isogenic variety

Received August 5, 2019; accepted September 27, 2019; published online October 9, 2019

Weixiao Liu[1,*], Zhe Zhang[1], Xuri Liu[2] and Wujun Jin[1,†]

[1]Biotechnology Research Institute, Chinese Agricultural and Academic Sciences, No. 12 Zhongguancun South St., Haidian District, Beijing 100081, P.R. China and [2]Department of Food and Biological Engineering, Handan Polytechnic College, No.141 Zhuhe Road, Hanshan District, Handan 056001, P.R. China

*Weixiao Liu, Biotechnology Research Institute, Chinese Agriculture Academy of Sciences, No. 12 Zhongguancun South St., Haidian District, Beijing 100081, P.R. China. Tel: +86-10-8210-9852, Fax: +86-10-8210-7657, email: liuweixiao@caas.cn

†Wujun Jin, Biotechnology Research Institute, Chinese Agriculture Academy of Sciences, No. 12 Zhongguancun South St., Haidian District, Beijing 100081, P.R. China. Tel: +86-10-8210-9852, Fax: +86-10-8210-7657, email: jinwujun@caas.cn

To investigate the unintended effects of genetically modified (GM) crops, an isobaric tags for relative and absolute quantitation (iTRAQ)-based comparative proteomic analysis was performed with seed cotyledons of two GM soybean lines, MON87705 and MON87701×MON89788, and the corresponding non-transgenic isogenic variety A3525. Thirty-five differentially abundant proteins (DAPs) were identified in MON87705/A3525, 27 of which were upregulated and 8 downregulated. Thirty-eight DAPs were identified from the MON87701×MON89788/A3525 sample, including 29 upregulated proteins and 9 downregulated proteins. Pathway analysis showed that most of these DAPs participate in protein processing in endoplasmic reticulum and in metabolic pathways. Protein–protein interaction analysis of these DAPs demonstrated that the main interacting proteins are associated with post-translational modification, protein turnover, chaperones and signal transduction mechanisms. Nevertheless, these DAPs were not identified as new unintended toxins or allergens and only showed changes in abundance. All these results suggest that the seed cotyledon proteomic profiles of the two GM soybean lines studied were not dramatically altered compared with that of their natural isogenic control.

Keywords: genetic modification; iTRAQ; qRT-PCR; quantitative proteomics; soybean seed cotyledons.

Abbreviations: 2-DE, two-dimensional electrophoresis; co-DAPs, common differentially abundant proteins; COG, clusters of orthologous groups; DAPs, differentially abundant proteins; DEPs, differentially expressed proteins; ELISAs, enzyme-linked immunosorbent assays; GM, genetically modified; GO, gene ontology; iTRAQ, isobaric tags for relative and absolute quantitation; KEGG, Kyoto encyclopedia of genes and genomes; qPCR, quantitative real-time PCR; qRT-PCR, quantitative real-time polymerase chain reaction.

Soybean (*Glycine max*) is an important economic crop in many countries. The planting area of transgenic soybeans is 94.1 million hectares, accounting for approximately half of the total planting area of genetically modified (GM) crops. GM crops are modified by insertion of exogenous genes that synthesize new products to improve the nutrient composition or enhance the insect or herbicide resistance of crops (*1*). The rapid development of GM crops has led to economic benefits for farmers by both reducing the use of chemical pesticides and increasing crop yields; however, insertion of exogenous DNA fragments may lead to the deletion, insertion or rearrangement of some genes, thereby affecting some biological pathways or resulting in the formation of new allergens or toxins (*2*, *3*). Thus, the food and environmental safety of GM crops must be carefully evaluated (*4*, *5*).

The basic evaluation methods for GM crops, namely, PCR-based detection of specific DNA sequences, ELISAs and immune test strip-based detection of specific gene products, detect predicted intended effects (*3*, *6*, *7*). Omics profiling techniques enable comprehensive measurement and comparison of the transcripts, proteins and metabolites of crops and provide detailed information regarding global changes in GM crops (*6*, *8–11*). Among these profiling techniques, proteomics, a method for evaluating unintended effects directly at the protein level, has been widely used to evaluate the unintended effects of GM crops (*12–19*).

Proteomics has developed continuously over the past 20 years. 2-DE coupled with MS, as a first-generation proteomic technique, has been frequently used in proteomic research for many years (*2*, *6*, *20*, *21*). Subsequently, isobaric tags for relative and absolute quantitation (iTRAQ) technology was rapidly developed as a quantitative proteomic analysis method with high accuracy, sensitivity and repeatability. Indeed, iTRAQ has been widely used in recent years. Several studies have reported the efficiency of this method for detecting differentially abundant proteins (DAPs) in potato (*22*), rice (*23*), maize (*20*) and soybean (*24–27*). In this paper, the DAPs of soybean seed cotyledons from three different soybean lines, namely, MON87705, MON87701×MON89788 (two GM lines) and A3525 (their corresponding non-GM control), were comprehensively analysed. MON87705 has been modified by insertion of the exogenous DNA fragments *fatb1-A* and *fad2-1A* to improve nutrient contents and *cp4 epsps* to enhance tolerance to herbicides. MON87701×MON89788 is modified with herbicide (*cp4 epsps*) and insect (*cry1Ac*) resistance genes. MON87705 and MON87701×MON89788 have been

approved for commercial cultivation or consumption (ISAAA, 2018). These two varieties have been modified with different foreign genes (*fatb1-A*, *fad2-1A in* MON87705 and *Cry 1Ac* in MON87701×MON89788) and the same foreign genes *(cp4 epsps* in both varieties) with the same non-GM control. Therefore, these soybean varieties were chosen for study. This study would simultaneously compare the effects of the same and different foreign genes on the same receptor, and further expand the depth and breadth of our knowledge of GM crops in a case-by-case manner.

Materials and Methods

Reagents and materials

Trypsin, reducing reagent, cysteine-blocking reagent and the dissolution buffer in the iTRAQ Kit and iTRAQ 8plex Kit were purchased from AB Sciex Corporation (Washington, DC, USA). Thiourea, ammonia, formic acid, methyl alcohol and bovine serum albumin were purchased from Sigma-Aldrich Corporation (MO, USA). Acetonitrile was purchased from Merck (NJ, USA). High performance liquid chromatography (HPLC)-quality water was obtained from a Cascada TM IX water purification system (Pall Co., NY, USA). Methanol (HPLC-grade) was purchased from Thermo Fisher Scientific (MA, USA). Urea and CHAPS (3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate) were purchased from Bio-Rad Laboratories, Inc. (CA, USA). The Durashell-C18 column was purchased from Agela (DE, USA).

Plant materials

Seed cotyledons of GM soybean lines, MON87705 and MON87701×MON89788, and the corresponding non-GM isogenic variety A3525 were studied in this paper. Their genetic information is provided in Supplementary Table S1 and Reference (*28*). Cotyledons of soybean seeds with full grains and uniform sizes were selected as experimental materials.

DNA extraction and event-specific PCR of transgenic soybeans

Genomic DNA was extracted from soybean seed cotyledons using the EasyPure Plant Genomic DNA Kit (Tiangen, Beijing, China) according to the manufacturer's instructions. The genomic DNA concentrations were quantified using a NanoDrop 2000 (Thermo Scientific, USA). The DNA was stored at −20°C until further analysis. Event-specific PCR was performed according to the Chinese National Standards MOA-2122-4-2014 (*28*), MOA-2259-7-2015 (*29*) and MOA-1485-6-2010 (*30*). The sequences of the primers used and the sizes of the amplified DNA fragments are listed in Supplementary Table S5.

Protein preparation

Three biological replicates of the studied lines of soybean seed cotyledons were used for protein profiling in this study. A total of 20 grain seed cotyledons of each soybean line were ground in liquid N_2 separately, and the total proteins were extracted with 1 ml of lysis buffer containing 7 M urea, 2 M thiourea, 0.1% CHAPS and protease inhibitor. After centrifugation at 15,000×g for 20 min at 4°C, the supernatant was collected and transferred to a fresh tube. The concentration of the extracted protein was measured using the Bradford protein assay (*31*, *32*).

Trypsin digestion and iTRAQ labelling

The extracted protein solution containing 200 μg of protein was digested with 4 μg of trypsin overnight at 37°C. Protein reduction, blocking of cysteine residues and digestion were performed according to the manufacturer's protocol included with the iTRAQ kit. The digested peptides were transferred to vials containing individual iTRAQ reagents according to the standard iTRAQ protocol for the 8-plex kit. The tags used were 114 Da for MON87705, 115 Da for MON87701×MON89788 and 117 Da for A3525 (*32*). The labelled samples were pooled in equal amounts, centrifuged under a vacuum and freeze-dried.

LC and MS/MS analyses

The peptide mixture was re-dissolved in solution A (98% ddH_2O and 2% acetonitrile, pH 10.0) and then fractionated by high pH (10.0) separation using the RIGOL L-3000RP-HPLC system (Beijing Puyuan Power Technology Co., Ltd.), and 100 μg of the mixture was desalted and fractionated using a Durashell-C18 reverse phase column. Next, solution B (98% acetonitrile and 2% ddH_2O) was added, and the pH was adjusted to 10.0. After separation, the fractions were resuspended in 20 μl of solution C (0.1% formic acid and 2% methanol in water), separated using a nano-liquid chromatography (LC) system (Thermo Fisher Scientific Corp., MA, USA) and analysed on-line using electrospray tandem MS.

Nano-LC-mass spectrometry/mass spectrometry (MS/MS) experiments were performed using an EASY-nLC 1000 coupled with a Q-Exactive system. Peptides were loaded on a nanocolumn (EASY-Spray column, C18) balanced with solvent D (0.1% formic acid acetonitrile solution). The Q-Exactive mass spectrometer was operated in data-dependent mode to switch automatically between MS and MS/MS acquisitions. Surveys of the full-scan MS spectra (m/z 350-1800) were acquired with a mass resolution of 70,000 FWHM, followed by 15 sequential high-energy collisional dissociation MS/MS scans with a resolution of 17,500 FWHM.

Analysis of proteomic data and bioinformatics

The original files generated by the Q-Exactive system were analysed using Proteome Discoverer 1.4 software (Thermo Fisher Scientific, Waltham, MA, USA), and protein identification was performed using the Mascot search engine (Matrix Science, London, UK; version 2.3.02) against the UniProt *Glycine max* (soybean) database. All identified proteins were matched with at least one peptide at a confidence 95% and FDR ≥ 1% (*25*, *33*). A3525 labelled with iTRAQ tag 117 was used as the reference to determine the fold change value. Proteins with a fold change ≥ 1.5 or ≤ 0.67 and *P*-values ≤ 0.05 (*17*, *26*, *34*) in the samples of MON87705 (iTRAQ tag 114)/A3525 and MON87701×MON89788 (iTRAQ tag 115)/A3525 were considered DAPs. A heatmap of the identified proteins was generated by the online ClustVis tool (https://biit.cs.ut.ee/clustvis/). The clusters of orthologous groups of proteins (COG, http://www.ncbi.nlm.nih.gov/COG/) system was employed for functional annotation of the identified proteins. Functional classification of the DAPs was performed by gene ontology (GO) annotation using the online DAVID tool (http://david.abcc.ncifcrf.gov). Pathway analysis of the DAPs was carried out using the Kyoto Encyclopedia of Genes and Genomes (KEGG) database (http://www.genome.jp/kegg/). The protein–protein interaction network was mapped using the online STRING 10.5 tool (http://string-db.org) and visualized by Cytoscape (3.2.0).

Western blotting

Soybean cotyledon proteins of A3525, MON87705 and MON87701×MON89788 were ground with liquid N_2. Proteins were extracted with lysate buffer (100 mM Tris-HCl (pH 8.0), 100 mM NaCl, 500 mM sucrose, 10 mM EDTA, 10 mM DTT, 2% Triton X-100, 1 mM protein inhibitor and 2 mM PMSF) and analysed by Western blotting with the antibodies CP4 EPSPS mAb (Shanghai Youlong Biotech Co., Ltd.) and actin pAb (Beijing Bioeasy Technology Co., Ltd.).

Quantitative reverse transcription PCR

A total of 1.0 g of soybean seed cotyledons was used for total RNA extraction with Ambion Pure Link plant RNA reagent (Life Technologies, Invitrogen, Carlsbad, CA, USA). RNA integrity was analysed by agarose gel electrophoresis. RNA (1.0 μg) was reverse transcribed with M-MLV reverse transcriptase (Life Technologies, Invitrogen) according to the manufacturer's instructions. The primers used for quantitative real-time PCR (qPCR), which are listed in Tables III and IV, were designed using Primer Premier 5.0. Quantitative reverse transcription PCR (qRT-PCR) was performed using the SYBR Green qRT-PCR Kit (Bio-Rad, Hercules CA, USA) according to the manufacturer's instructions with three biological replicates. The expression of Lectin (Lectin F: 5′-CCAGCTTCGC CGCTTCCTTC-3′; Lectin R: 5′-GAAGGCAAGCCCATCTGCAA GCC-3′) was measured as an internal control. All reactions were conducted on a CFX96 real-time system (Bio-Rad). The RT-qPCR data were analysed using the relative quantification $2^{-\Delta\Delta CT}$ method (*35*).

Results

Analysis of GM soybean lines

Soybean seed cotyledons were used to study proteomic differences between two GM lines and their corresponding non-GM isogenic control. The genetic relationships among the studied soybean lines are shown in Supplementary Table S1 and in our previously published report (*32*). Event-specific PCR (*36*, *37*) was used to detect specific events in transgenic soybean lines, and the target DNA fragment was obtained from GM soybean lines (*32*). Western blotting was used to determine the expression abundance of CP4 EPSPS. High expression levels of CP4 EPSPS were observed in the GM soybean lines MON87705 and MON87701×MON89788 (Supplementary Fig. S1).

Protein expression pattern analysis

Cluster analysis of the identified proteins was conducted using the ClustVis tool (https://biit.cs.ut.ee/clustvis/). The proteins were grouped based on expression level. The protein expression patterns of MON87705 and MON87701×MON89788 share higher similarity with those of MON87705 and A3525 and those of MON87701×MON89788 and A3525. The similarity of MON87701×MON89788 with A3525 is higher than that between MON87705 and A3525 (Fig. 1A).

COG functional analysis for all identified proteins

All the quantitative identified proteins were clustered into 24 COG categories (Fig. 1B). Among these categories, translation, ribosomal structure and biogenesis and post-translational modification, protein turnover and chaperones, accounting for 29% of the detected proteins, represented two of the largest groups associated with protein biogenesis and modification in soybean seeds. These groups were followed by 14% of the proteins in general function prediction only, 10% in energy production and conversion, 10% in carbohydrate transport and metabolism, 9% in amino acid transport and metabolism, 4% in lipid transport and metabolism, 3% in cell wall/membrane/envelope biogenesis and 3% in secondary metabolite biosynthesis, transport and catabolism. These proteins are involved in energy conversion and metabolism in soybean seeds. Finally, the low-abundance categories, such as signal transduction mechanisms; replication, recombination and repair; cell cycle control, cell division, chromosome partitioning, chromatin structure and dynamics; and RNA processing and modification, accounted for a low proportion (1–2%) of the identified proteins, which was consistent with the regulatory functions of these proteins in soybean seeds.

Analysis of the DAPs

A total of 2,403 non-redundant proteins were detected by iTRAQ (*32*). Among them, 2,369 proteins (listed in Supplementary Table S2) were quantitatively identified in the MON87705/A3525 and MON87701×MON89788/A3525 samples. Further analysis showed that 35 DAPs were identified by comparison of MON87705 with A3525, 27 of which were upregulated and 8 of which were downregulated. Thirty-eight DAPs were identified in the MON87701×MON89788/A3525 sample, including 29 upregulated proteins and 9 downregulated proteins (Table I, Supplementary Tables S3 and S4).

GO annotation for DAPs

Annotation was performed using the online DAVID tool (http://david.abcc.ncifcrf.gov) to reveal the functions of these identified DAPs. Depending on their functional annotation, these DAPs were classified into three large groups: molecular functions, biological processes and cellular components. The results showed that the DAPs of MON87705/A3525 can be annotated into 29 functional groups, including 17 biological processes, 7 cellular components and 5 molecular functions (Fig. 2A). In biological processes, the DAPs are mainly involved in cellular processes, metabolic processes and responses to stimuli. In cellular components, the DAPs are mainly involved in the cell, cell part and organelle. In molecular functions, the DAPs are mainly involved in catalytic activity, binding and nutrient reservoir activity. The DAPs of MON87701×MON89788/A3525 were annotated into 31 functional groups, including 16 biological processes, 8 cellular components and 7 molecular functions (Fig. 2B). In biological processes, the DAPs are mainly involved in metabolic processes, cellular processes and responses to stimuli. In cellular components, the DAPs are mainly involved in the cell, cell part and organelle. In molecular functions, the DAPs are mainly involved in catalytic activity, binding and nutrient reservoir activity. The results also showed that the downregulated proteins of MON87701×MON89788/A3525 participated in almost all biological processes, but the downregulated proteins of MON87705/A3525 were only involved in some biological processes, such as metabolic processes, cellular processes and responses to stimuli.

KEGG analysis of all the DAPs

KEGG pathway analysis was carried out against the KEGG pathway database (http://www.genome.jp/kegg/). The results showed that the DAPs of MON87705/A3525 were annotated into 19 KEGG pathways (Fig. 3A). Protein processing in the endoplasmic reticulum (ko04141, four upregulated proteins, Supplementary Fig. S2) and metabolic pathways (ko01100, two upregulated proteins and two downregulated proteins) were the primary enriched pathways, followed by RNA transport (ko03013, two upregulated proteins). The DAPs of MON87701×MON89788/A3525, including 15 upregulated and 6 downregulated proteins, were annotated into 26 KEGG pathways (Fig. 3B). Metabolic pathways (ko01100, seven upregulated proteins and four downregulated proteins) were the primary enriched pathways, followed by protein processing in the endoplasmic reticulum (ko04141, three upregulated proteins, Supplementary Fig. S2) and biosynthesis of secondary metabolites (ko01110, three downregulated proteins).

Protein–protein interaction analysis

To identify protein–protein interaction networks, the identified DAPs were analysed by the online STRING 10.5 tool (http://string-db.org). A protein interaction

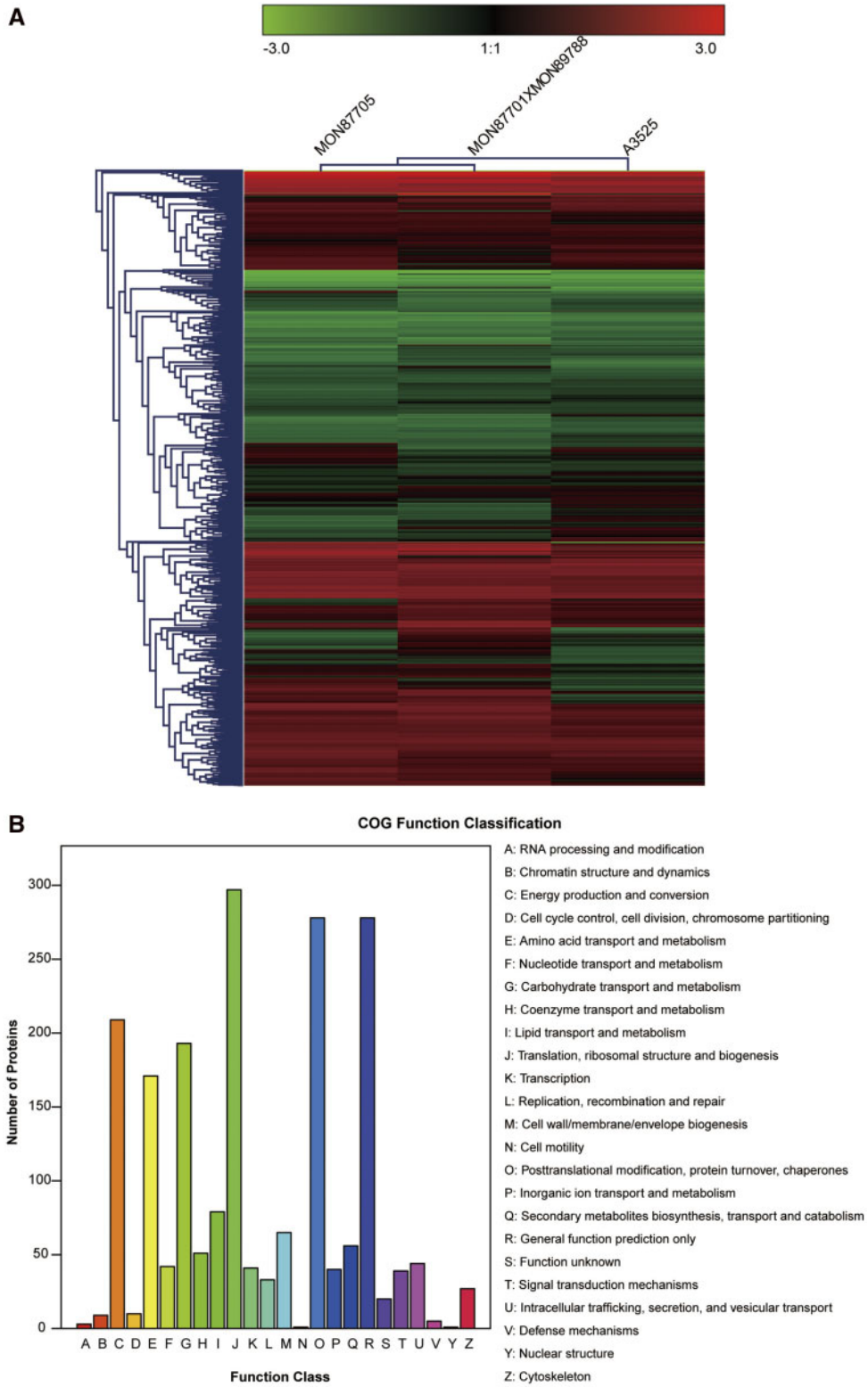

Fig. 1 Protein expression pattern analysis and COG function analysis. (**A**) Cluster map comparing the protein expression patterns of MON87705, MON87701×MON89788 and A3525. Red indicates relatively high expression, green indicates relatively low expression and black indicates the same expression levels in the two lines. (**B**) COG functional classes of all the identified proteins. The letters under the X axis represent the COG categories listed on the right of the column, and the Y axis represents the number of proteins.

network was constructed and visualized with Cytoscape (3.2.0) (Supplementary Table S3). Fourteen DAPs of MON87705/A3525, including 10 upregulated and 4 downregulated DAPs, were included in 5 main clusters (Fig. 4A). Among these DAPs, five interacting proteins were mainly related to post-translational modification, protein turnover and chaperones, while three interacting proteins were

Table I. Summary of the protein identification data

Database	No. of repeats		Protein groups	
			MON87705/A3525	MON87701×MON89788/A3525
Glycine max	1st		1,584	1,585
	2nd		1,619	1,621
	3rd		1,617	1,617
	Total no. of proteins		2,369	2,369
	DAPs	Up	27	29
		Down	8	8
		Total	35	38

related to signal transduction mechanisms. Nine DAPs of MON87701×MON89788/A3525, including five upregulated and four downregulated DAPs, were included in four main clusters (Fig. 4B). Among these DAPs, two interacting proteins were related to signal transduction mechanisms, two interacting proteins were mainly related to post-translational modification, protein turnover, and chaperones, and two interacting proteins were related to amino acid transport and metabolism.

Eleven co-DAPs were identified in two GM soybean lines

Among the DAPs, 11 DAPs were simultaneously identified in two GM soybean lines. O22378/metallothionein-II protein and I1KQW4/uncharacterized protein were downregulated in both. A0A0B2NW50/lipoxygenase was downregulated in the MON87705/A3525 sample, but upregulated in the MON87701×MON89788/A3525 sample. Other co-DAPs were upregulated in two GM soybean lines (Fig. 5). P45458/Malate synthase, A0A0B2NW50/Lipoxygenase and Q71LY8/CP4 EPSPS are involved in KEGG metabolic pathways, and I1KQW4/uncharacterized protein participates in plant-pathogen interaction. A0A0B2QAS8/glycine-rich RNA-binding protein 2 is involved in ribosome and K7K8E5/uncharacterized protein in RNA transport. Other co-DAPs were not annotated in KEGG (Table II).

qRT-PCR analysis of DAPs

To explore changes in the identified DAPs at the transcriptional level, 14 and 17 representative DAPs of the MON87705/A3525 and MON87701×MON89788/A3525 samples, respectively, were selected for qRT-PCR to assess gene expression. The transcriptional patterns of these genes are shown in Fig. 6A and B. Compared with the patterns of the tested genes at the transcriptional and translational levels, six DAPs exhibited similar trends at both the transcriptional and translational levels in the MON87705/A3525 sample. Four DAPs exhibited no significant change at the transcriptional level, and the transcriptional expression trends of four DAPs were the opposite of the corresponding translational expression trends (Fig. 6A, Table III). Twelve DAPs exhibited similar trends at both the transcriptional and translational levels in the MON87701×MON89788/A3525 sample. The transcriptional expression trend of one DAP showed no significant change. The transcriptional expression trends of four DAPs were the opposite of the corresponding translational expression trends (Fig. 6B, Table IV).

Discussion

With the rapid development of GM crops, the unintended effects resulting from genetic modification are the most controversial issues associated with transgenic crops. Genes associated with herbicide resistance, insect resistance and nutritional improvement have been used to facilitate the generation of GM crops. In this paper, the iTRAQ protein profiling approach was applied to investigate differences in seed cotyledon proteins between the GM soybean lines MON87705 (modified with herbicide resistance and nutritional improvement genes) and MON87701×MON89788 (modified with herbicide and insect resistance genes) and the corresponding non-GM isogenic variety A3525.

The iTRAQ protein profiles detected 35 and 38 DAPs in the MON87705/A3525 and MON87701×MON89788/A3525 samples, respectively, but these numbers were lower than 2% of the total number (2,369) of detected quantitative proteins. Most are involved in translation, ribosomal structure and biogenesis and post-translational modification, protein turnover and chaperones and general function prediction. Furthermore, these DAPs were not identified as new unintended toxins and/or allergens and only showed changes in abundance, indicating that the proteomic profiles of the twoGM soybean lines studied were not dramatically altered compared with that of their natural isogenic control A3525. These results are consistent with the results previously reported for some other GM crops (*2*, *6*, *38–40*). Furthermore, the unintended effects are not unique to GM crops but are also observed during conventional plant breeding (*6*, *25*) and hybridization (*2*, *41*, *42*).

An iTRAQ-based proteomic analysis of soybean seed cotyledons from four GM lines and three natural genotypic soybean lines conducted to identify common differentially expressed proteins (cDEPs) among different GM soybean lines and natural genotypic soybean lines was reported by our research group in 2018 (*32*). To analyse seven soybean lines at the same time, we analysed the results of iTRAQ experiments individually to screen more DAPs and effectively assess the common unintended effects of the studied GM soybean lines. A total of 67 and 61 DAPs were identified from 3 iTRAQ replicates of the MON87705/A3525 and MON87701×MON89788/A3525 samples, respectively. These DAPs were classified and analysed based on whether they were detected 2 or 3 times. However, this study focused on the particular genetic modification event; 35 and 38 DAPs were selected from the MON87705/A3525 and MON87701×MON89788/A3525 samples, respectively, according to a comprehensive analysis of the results of three iTRAQ experiments. These DAPs included almost all of the DAPs

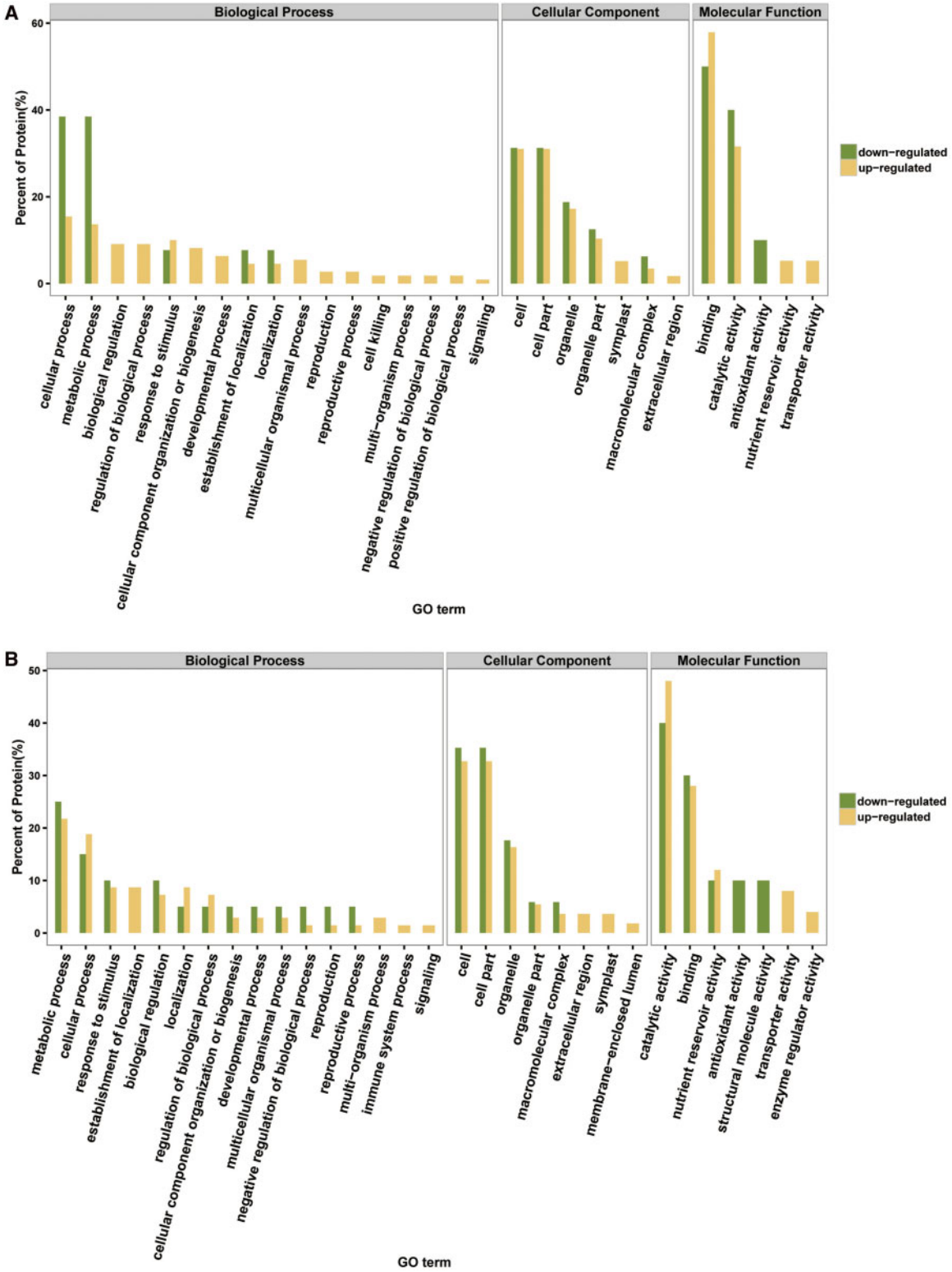

Fig. 2 GO annotations of the identified DAPs. DAPs were annotated into three main categories, including biological processes, cellular components and molecular functions, to determine the functions of the identified DAPs between the GM soybean lines and their parents: (**A**) MON87705/A3525; (**B**) MON87701×MON89788/A3525. Yellow indicates upregulated proteins and green indicates downregulated proteins.

detected three times and some DAPs detected two times from previously published data (*32*).

EPSPS, 3-phosphoshikimate 1-carboxyvinyltransferase, is a key enzyme in the shikimic acid pathway (*43–45*). Upregulated expression of CP4 EPSPS was observed in MON87705 and MON87701×MON89788 by both iTRAQ and Western blotting, while other enzymes of the shikimic acid pathway were not identified as DAPs. Q71LY8/CP4 EPSPS in MON87705, but not in MON87701×MON89788, was observed to be

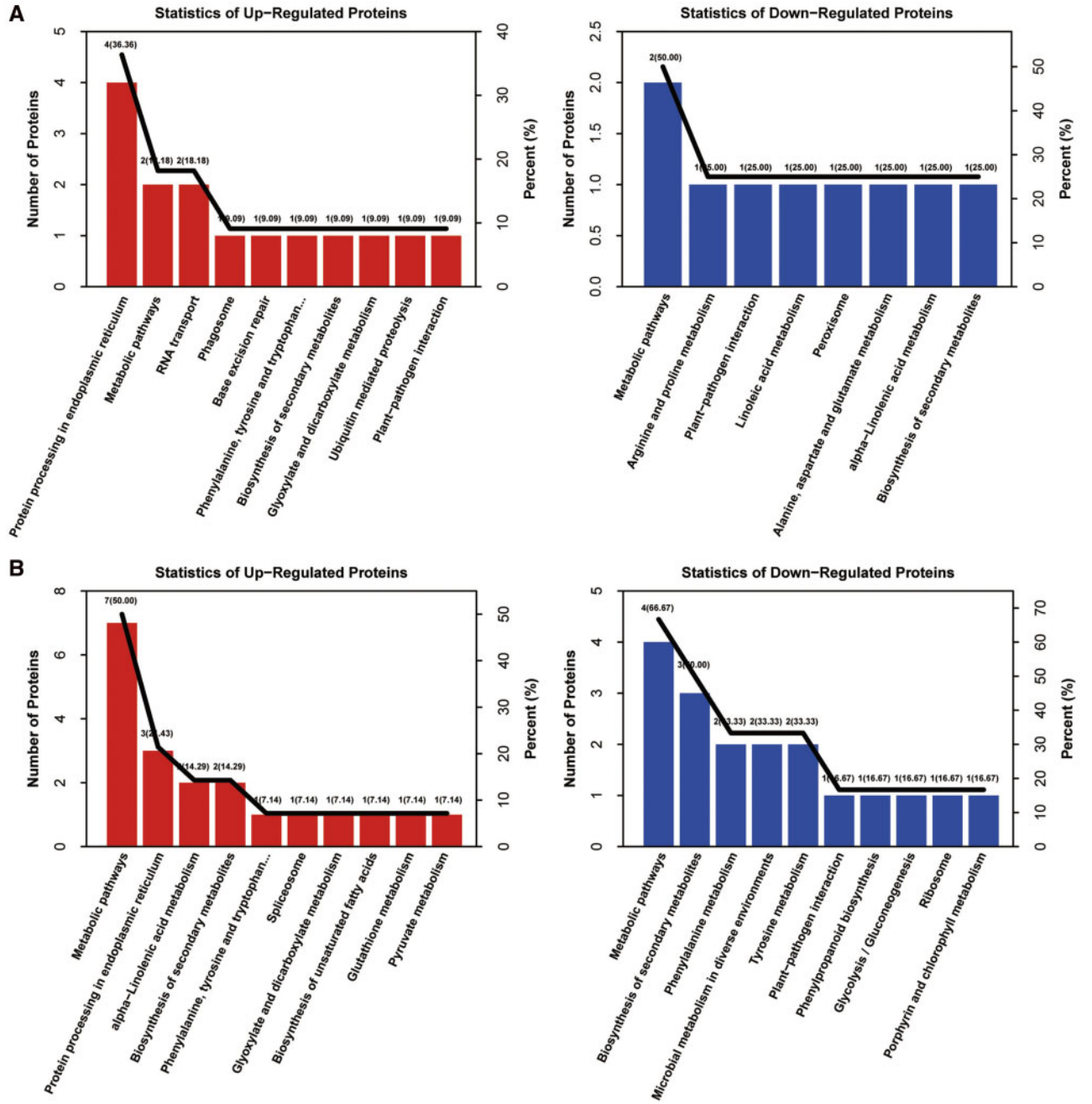

Fig. 3 KEGG pathway analysis of upregulated and downregulated proteins of MON87705/A3525 (A) and MON87701×MON89788/A3525 (B).

involved in protein–protein interactions (Fig. 4). Among other identified co-DAPs, I1KQW4 in both GM soybeans is involved in protein–protein interactions, whereas K7K8E5 was only observed in MON87705 and Q852U4/Glycinin G3 only in MON87701×MON89788.

KEGG analysis of the identified DAPs showed that several in the two GM soybean lines were enriched in protein processing in the endoplasmic reticulum pathway. Calnexin, UDP-glucose: glycoprotein glucosyltransferase, protein disulfide-isomerase and ubiquitin-conjugating enzyme E2 (Ubc6/7) were screened from the MON87705/A3525 sample (Supplementary Fig. S2). Translocation protein SEC62 (Sec 62/63), PDI and ubiquitin domain-containing protein DSK2 were screened from the MON87701×MON89788/A3525 sample (Supplementary Fig. S3). The ubiquitin-conjugating enzyme E2 is the key enzyme in protein ubiquitination. The interaction between ubiquitin-conjugating enzyme E2 and ubiquitin ligase E3 determines the ubiquitin chain linkage, which further determines the fate of the ubiquitinated protein (*46*, *47*). The ubiquitin domain-containing protein DSK2 modulates proteasome-mediated protein degradation (*48*), which may be the mechanism underlying soybean self-protection stimulated by the expression of an exogenous protein that can be degraded and removed by protein processing via the endoplasmic reticulum pathway. The identified co-DAPs in two GM soybean lines participate in metabolic, plant-pathogen interaction, ribosome

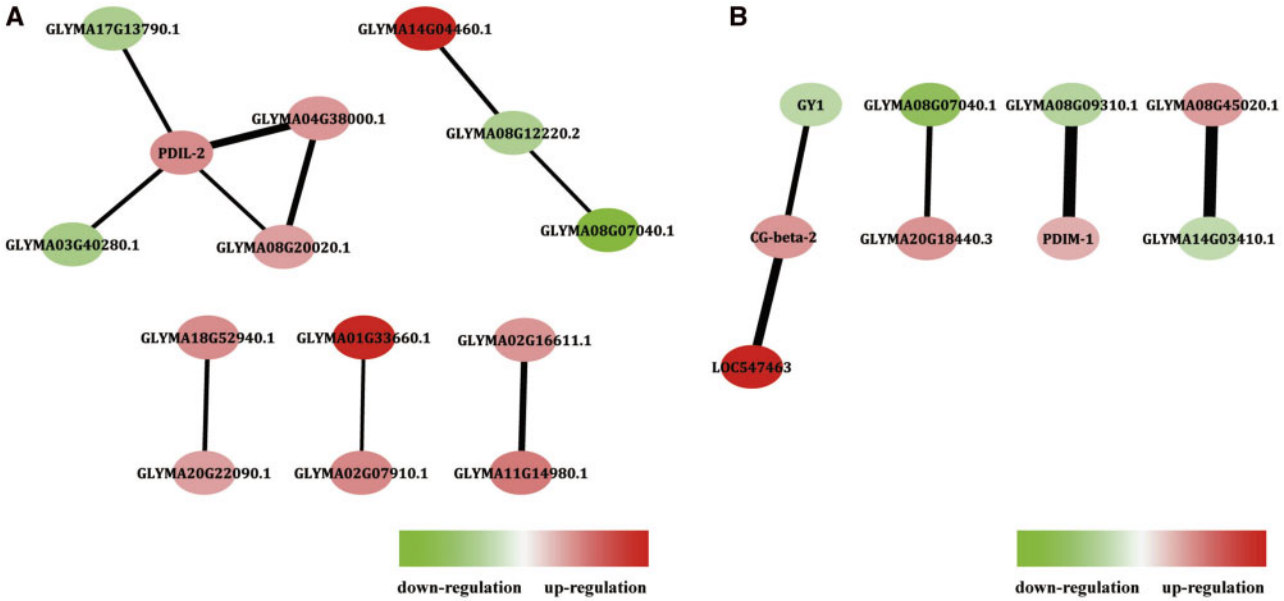

Fig. 4 Protein–protein interaction analysis of the identified DAPs in MON87705/A3525 (A) and MON87701×MON89788/A3525 (B). Upregulated proteins are marked in red, and downregulated proteins are marked in green. The edge is used to represent the interaction between two proteins, and edge thickness indicates the confidence score for the interaction.

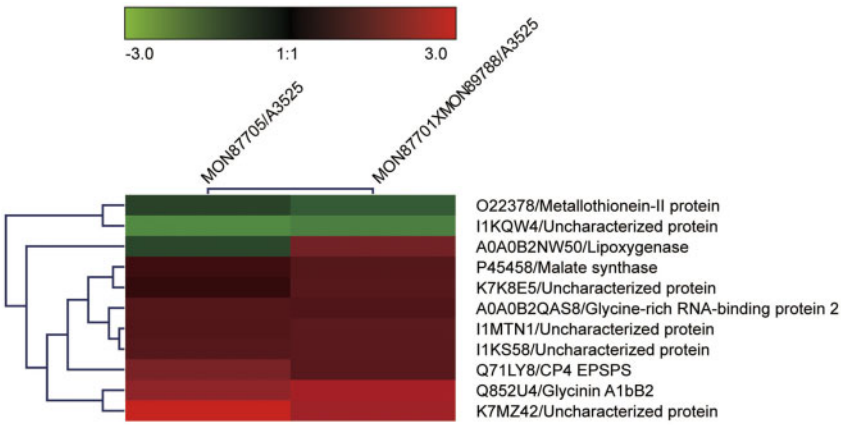

Fig. 5 The expression pattern of co-DAPs identified in two GM soybean lines. Red indicates relatively high expression, green indicates relatively low expression and black indicates the same expression levels in the two lines.

Table II. Summary of the identified co-DAPs

Accession/Name	KEGG function	Subcellular location
Q852U4/Glycinin A1bB2-784		Nuclear
P45458/Malate synthase	Metabolic pathways; glyoxylate and dicarboxylate metabolism; pyruvate metabolism; microbial metabolism in diverse environments	
A0A0B2NW50/Lipoxygenase	Metabolic pathways; linoleic acid metabolism; alpha-linolenic acid metabolism	Cytoplasmic
O22378/Metallothionein-II protein		Extracellular
I1MTN1/Uncharacterized protein		Plasma membrane
I1KQW4/Uncharacterized protein	Plant-pathogen interaction	Plasma membrane
A0A0B2QAS8/Glycine-rich RNA-binding protein 2	Ribosome	
K7K8E5/Uncharacterized protein	RNA transport	Cytoplasmic
I1KS58/Uncharacterized protein		Extracellular
Q71LY8/CP4 EPSPS	Metabolic pathways; phenylalanine, tyrosine and tryptophan biosynthesis; biosynthesis of secondary metabolites	Chloroplast
K7MZ42/Uncharacterized protein		Plasma membrane

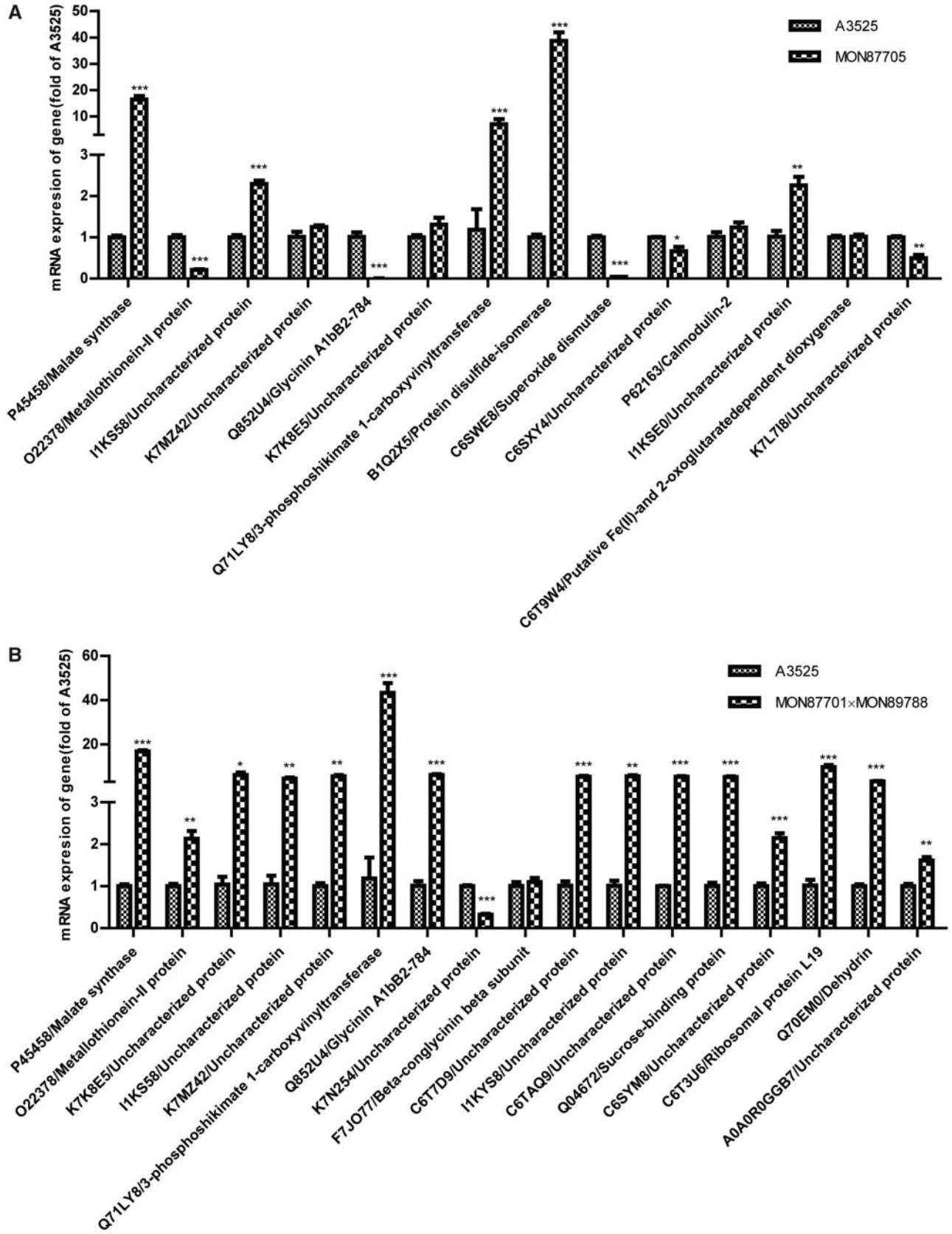

Fig. 6 qRT-PCR analysis of the gene expression patterns of selected DAPs of MON87705/A3525 (A) and MON87701×MON89788/A3525 (B). Error bars represent the standard deviation (SD) among the three replicates. The asterisks represent significant difference compared with A3525, as indicated by the t-test (*$P<0.05$; **$P<0.01$ and ***$P<0.001$).

Table III. qRT-PCR verification of MON87705/A3525

ID	Primer-F	Primer-R	Change	
			Proteomics	RT-qPCR
Lectin	CCAGCTTCGCCGCTTCCTTC	GAAGGCAAGCCCATCTGCAAGCC		
P45458	GCTTTGAAATTCGTTGCTGACTTG	CACCCACTCTTGCTCTCTGATG	Up	Up
O22378	TGACAACTGGAGGCGGTGATC	TCTTGGGACATGAGCATGGATTG	Down	Down
I1KS58	GGAGAGGTGTGACAAGGTGATG	CCTTCCTGTTCCTGAGTAGATAGC	Up	Up
K7MZ42	ATGGTTGCTCAGTGGCTATATGTG	CAGTCAATTCTCTCCCTCTCTTCG	Up	No
Q852U4	ACTCCTGTTGTTGCCGTTTCTC	TACCTCCTTGCTGCTTCTGTGG	Up	Down
K7K8E5	GCTTCTGTGGCTGCCTGTATATC	ACCTCCTTCCAGTGCCTTGC	Up	No
Q71LY8	CCTCCGCACAGGTGAAGTCC	CCGTCCGCATCCGTCTCG	Up	Up
B1Q2X5	CTCCTCTTCTCCTCCCTCTTTCTC	TGGCGTCATCGTCGTGGTAG	Up	Up
C6SWE8	AATGTCGGTGATGATGGTACTGTC	CGCCCTTCCTATGATGTTGTTTGG	Down	Down
C6SXY4	GCTGGACACCTGTTCATACGG	GACTTCCTTACACATCGGCTAACC	Up	Down
P62163	GACCAGATCGCCGAGTTCAAG	CGCATCACAGTCCCAAGTTCC	Up	No
I1KSE0	GTGGTGGGTGCTCGTAGGG	CCTCTTCACACTCACATCGTTCC	Down	Up
C6T9W4	TGCCCTGACCCTGAAGTTGTAG	ACTATCACCATCACTGCCTCTTAC	Down	No
K7L7I8	GGGAGATGGACGGCTATCGG	TCGGAACTTCTTCAGGTCAACAAC	Up	Down

Table IV. qRT-PCR verification of MON87701×MON89788/A3525

ID	Primer-F	Primer-R	Change	
			Proteomic	RT-qPCR
Lectin	CCAGCTTCGCCGCTTCCTTC	GAAGGCAAGCCCATCTGCAAGCC		
P45458	GCTTTGAAATTCGTTGCTGACTTG	CACCCACTCTTGCTCTCTGATG	Up	Up
O22378	TGACAACTGGAGGCGGTGATC	TCTTGGGACATGAGCATGGATTG	Down	Up
K7K8E5	GCTTCTGTGGCTGCCTGTATATC	ACCTCCTTCCAGTGCCTTGC	Up	Up
I1KS58	GGAGAGGTGTGACAAGGTGATG	CCTTCCTGTTCCTGAGTAGATAGC	Up	Up
K7MZ42	ATGGTTGCTCAGTGGCTATATGTG	CAGTCAATTCTCTCCCTCTCTTCG	Up	Up
Q71LY8	CCTCCGCACAGGTGAAGTCC	CCGTCCGCATCCGTCTCG	Up	Up
Q852U4	ACTCCTGTTGTTGCCGTTTCTC	TACCTCCTTGCTGCTTCTGTGG	Up	Up
K7N254	AGGAACAACAGACGCCCAAAG	TTCAAGCCTAGCAACCTCTTCAAG	Up	Down
F7J077	GATTTCCTCCTCTTTGTCCTTAGCG	GTCTCTGTCGTCGTTGTTCACC	Up	No
C6T7D9	TACTGTGTTGAAGGGTGTTGTTACTG	TGAGCCAAAGACGGGTGAGC	Up	Up
I1KYS8	CCACCTCCGACGCTCCTC	ACCAATTCCTTATGATCTCTAACCTTC	UP	Up
C6TAQ9	GAGTGGGATTTTGGCGTGTTTTC	GTTTCCAGCCCTCTCAAGTGC	Up	Up
Q04672	GTCCCCACGCCACTTTGATTC	TTTCACTTTCCCTCACCAACCC	Up	Up
C6SYM8	TTGGTTCTGAGGATGATTGGATGG	AGCACAAACAACGAAACAAAGAGC	Down	Up
C6T3U6	CCAAGATTCACTCCCGCTCAC	GCACACGCATCCTTCTCATCC	Down	Up
Q70EM0	GCTTCGTCGTTCCTCCAGTTC	CCTTCTTCCTCCTCCCACCTTG	Up	Up
A0A0R0GGB7	AGGGCAGGGTGGTGGTATTG	GTCAACTGTTCCATCTCGGTAGG	Up	Up

and RNA transport pathways. Protein–protein interaction analysis also showed that most interacting proteins are involved in post-translational modification, protein turnover and chaperones.

In this study, several co-DAPs and representative DAPs involved in protein processing in the endoplasmic reticulum and/or in protein–protein interactions were selected for qRT-PCR to assess gene expression. Most of the DAPs showed similar patterns at the protein and transcript levels, while others exhibited inconsistent or even opposite trends for transcription and translation. These results are consistent with previously reported results for GM maize (*49*) and cotton (*17*), perhaps because protein expression is regulated during various biological processes, such as transcription, post-transcriptional modification, translation and post-translational modification.

Supplementary Data

Supplementary Data are available at *JB* Online.

Acknowledgements

We thank Dr. Liang Li and Dr. Yusong Wan for critical reading of the manuscript. We are grateful to Dr. Yang Yang (Genecreate Biological Engineering Co., Ltd., Wuhan, China) for technical assistance with the iTRAQ-MS.

Funding

This work was supported by the Fundamental Research Funds for Central Non-profit Scientific Institutions (*1610392018008*) and the Major Project Multipoint identification and comprehensive evaluation of transgenic crops (No. 2017ZX08013001).

Conflict of Interest

None declared.

References

1. http://www.isaaa.org/gmapprovaldatabase/default.asp.
2. Ren, Y., Lv, J., Wang, H., Li, L., Peng, Y., and Qu, L.J. (2009) A comparative proteomics approach to detect unintended effects in transgenic Arabidopsis. *J. Genet. Genomics* **36**, 629–639
3. Cellini, F., Chesson, A., Colquhoun, I., Constable, A., Davies, H.V., Engel, K.H., Gatehouse, A.M.R., Kärenlampi, S., Kok, E.J., Leguay, J.-J., Lehesranta, S., Noteborn, H.P.J.M., Pedersen, J., and Smith, M. (2004) Unintended effects and their detection in genetically modified crops. *Food Chem. Toxicol.* **42**, 1089–1125
4. Li, X., He, X.Y., Luo, Y.B., Xiao, G.Y., Jiang, X.B., and Huang, K.L. (2008) Comparative analysis of nutritional composition between herbicide-tolerant rice with bar gene and its non-transgenic counterpart. *J. Food Compos. Anal.* **21**, 535–539
5. Han, J.H., Yang, Y.X., Chen, S.R., Wang, Z., Yang, X.L., Wang, G.D., and Men, J.H. (2005) Comparison of nutrient composition of parental rice and rice genetically modified with cowpea trypsin inhibitor in China. *J. Food Compos. Anal.* **18**, 297–302
6. Gong, C.Y., Li, Q., Yu, H.T., Wang, Z.Z., and Wang, T. (2012) Proteomics insight into the biological safety of transgenic modification of rice as compared with conventional genetic breeding and spontaneous genotypic variation. *J. Proteome Res.* **11**, 3019–3029
7. Wang, X.J., Zhang, X., Yang, J.T., and Wang, Z.X. (2018) Effect on transcriptome and metabolome of stacked transgenic maize containing insecticidal cry and glyphosate tolerance EPSPS genes. *Plant J.* **93**, 1007–1016
8. Ricroch, A.E., Berge, J.B., and Kuntz, M. (2011) Evaluation of genetically engineered crops using transcriptomic, proteomic, and metabolomic profiling techniques. *Plant Physiol.* **155**, 1752–1761
9. Kuiper, H.A., Kok, E.J., and Engel, K.H. (2003) Exploitation of molecular profiling techniques for GM food safety assessment. *Curr. Opin. Biotech.* **14**, 238–243
10. Baudo, M.M., Lyons, R., Powers, S., Pastori, G.M., Edwards, K.J., Holdsworth, M.J., and Shewry, P.R. (2006) Transgenesis has less impact on the transcriptome of wheat grain than conventional breeding. *Plant Biotechnology J.* **4**, 369–380
11. Barros, E., Lezar, S., Anttonen, M.J., van Dijk, J.P., Rohlig, R.M., Kok, E.J., and Engel, K.H. (2010) Comparison of two GM maize varieties with a near-isogenic non-GM variety using transcriptomics, proteomics and metabolomics. *Plant Biotechnol. J.* **8**, 436–451
12. Coll, A., Nadal, A., Rossignol, M., Puigdomenech, P., and Pla, M. (2011) Proteomic analysis of MON810 and comparable non-GM maize varieties grown in agricultural fields. *Transgenic Res.* **20**, 939–949
13. Scossa, F., Laudencia-Chingcuanco, D., Anderson, O.D., Vensel, W.H., Lafiandra, D., D'Ovidio, R., and Masci, S. (2008) Comparative proteomic and transcriptional profiling of a bread wheat cultivar and its derived transgenic line overexpressing a low molecular weight glutenin subunit gene in the endosperm. *Proteomics.* **8**, 2948–2966
14. Kok, E.J., and Kuiper, H.A. (2003) Comparative safety assessment for biotech crops. *Trends Biotechnol.* **21**, 439–444
15. Chassy, B.N. (2002) Food safety evaluation of crops produced through biotechnology. *J. Am. Coll. Nutr.* **21**, 166s–173s
16. Brandao, A.R., Barbosa, H.S., and Arruda, M.A.Z. (2010) Image analysis of two-dimensional gel electrophoresis for comparative proteomics of transgenic and non-transgenic soybean seeds. *J. Proteomics* **73**, 1433–1440
17. Wang, L.M., Wang, X.C., Jin, X., Jia, R.Z., Huang, Q.X., Tan, Y.H., and Guo, A.P. (2015) Comparative proteomics of Bt-transgenic and non-transgenic cotton leaves. *Proteome Sci.* **13**, 1–15
18. Wang, Y., Xu, W.T., Zhao, W.W., Hao, J.R., Luo, Y.B., Tang, X.G., Zhang, Y., and Huang, K.L. (2012) Comparative analysis of the proteomic and nutritional composition of transgenic rice seeds with Cry1ab/ac genes and their non-transgenic counterparts. *J. Cereal Sci.* **55**, 226–233
19. Khalf, M., Goulet, C., Vorster, J., Brunelle, F., Anguenot, R., Fliss, I., and Michaud, D. (2010) Tubers from potato lines expressing a tomato Kunitz protease inhibitor are substantially equivalent to parental and transgenic controls. *Plant Biotechnol. J.* **8**, 155–169
20. Tan, Y., Yi, X., Wang, L., Peng, C., Sun, Y., Wang, D., Zhang, J., Guo, A., and Wang, X. (2016) Comparative proteomics of leaves from phytase-transgenic maize and its non-transgenic isogenic variety. *Front. Plant Sci.* **7**, 1211
21. Barbosa, H.S., Arruda, S.C., Azevedo, R.A., and Arruda, M.A. (2012) New insights on proteomics of transgenic soybean seeds: evaluation of differential expressions of enzymes and proteins. *Anal. Bioanal. Chem.* **402**, 299–314
22. Lim, S., Borza, T., Peters, R.D., Coffin, R.H., Al-Mughrabi, K.I., Pinto, D.M., and Wang-Pruski, G. (2013) Proteomics analysis suggests broad functional changes in potato leaves triggered by phosphites and a complex indirect mode of action against *Phytophthora infestans*. *J. Proteomics* **93**, 207–223
23. Qian, D.D., Tian, L.H., and Qu, L.Q. (2015) Proteomic analysis of endoplasmic reticulum stress responses in rice seeds. *Sci. Rep.* **5**, 14255
24. Qin, J., Gu, F., Liu, D., Yin, C.C., Zhao, S.J., Chen, H., Zhang, J.A., Yang, C.Y., Zhan, X., and Zhang, M.C. (2013) Proteomic analysis of elite soybean Jidou17 and its parents using iTRAQ-based quantitative approaches. *Proteome Sci.* **11**, 12
25. Qin, J., Zhang, J.N., Wang, F.M., Wang, J.H., Zheng, Z., Yin, C.C., Chen, H., Shi, A.N., Zhang, B., Chen, P.Y., and Zhang, M.C. (2017) iTRAQ protein profile analysis of developmental dynamics in soybean [Glycine max (L.) Merr.] leaves. *Plos One*. **12**, e0181910
26. Zeng, W.Y., Sun, Z.D., Cai, Z.Y., Chen, H.Z., Lai, Z.G., Yang, S.Z., and Tang, X.M. (2017) Proteomic analysis by iTRAQ-MRM of soybean resistance to *Lamprosema indicate*. *BMC Genomics* **18**, 444
27. Ji, W., Cong, R., Li, S., Li, R., Qin, Z.W., Li, Y.J., Zhou, X.L., and Chen, S.X., Li, J. (2016) Comparative proteomic analysis of soybean leaves and roots by iTRAQ provides insights into response mechanisms to short-term salt stress. *Front Plant Sci* **7**, 573
28. Detection of genetically modified plants and derived products-qualitative PCR method for herbicide-tolerant and quality improved soybean MON87705 and its derivates. MOA-2122-4-2014.
29. Detection of genetically modified plants and derived products-qualitative PCR method for insect-resistant soybean MON87701 and its derivates. MOA-2259-7-2015

30. Detection of genetically modified plants and derived products-qualitative PCR method for herbicide-tolerant soybean MON89788 and its derivates. MOA-1485-6-2010
31. Bradford, M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**, 248–254
32. Liu, W., Xu, W., Li, L., Dong, M., Wan, Y., He, X., Huang, K., and Jin, W. (2018) iTRAQ-based quantitative tissue proteomic analysis of differentially expressed proteins (DEPs) in non-transgenic and transgenic soybean seeds. *Scientific Reports* **8**, 17681
33. Baldrianova, J., Cerny, M., Novak, J., Jedelsky, P.L., Diviskova, E., and Brzobohaty, B. (2015) Arabidopsis proteome responses to the smoke-derived growth regulator karrikin. *J Proteomics*. **120**, 7–20
34. Liu, Y.B., Zhang, Y.X., Song, S.Q., Li, J.S., Stewart, C.N., Wei, W., Zhao, Y.J., and Wang, W.Q. (2015) A proteomic analysis of seeds from Bt-transgenic *Brassica napus* and hybrids with wild *B. juncea*. *Sci. Rep*. **5**, 15480
35. Livak, K.J., and Schmittgen, T.D. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(T)(-Delta Delta C) method. *Methods* **25**, 402–408
36. Wu, Y.H., Wu, G., Xiao, L., and Lu, C.M. (2007) Event-specific qualitative and quantitative PCR detection methods for transgenic rapeseed hybrids MS1 x RF1 and MS1 x RF2. *J. Agric. Food Chem.* **55**, 8380–8389
37. Rao, J., Yang, L.T., Guo, J.C., Quan, S., Chen, G.H., Zhao, X.X., Zhang, D.B., and Shi, J.X. (2016) Development of event-specific qualitative and quantitative PCR detection methods for the transgenic maize BVLA430101. *Eur. Food Res. Technol.* **242**, 1277–1284
38. Arruda, S.C., Barbosa, H.S., Azevedo, R.A., and Arruda, M.A. (2013) Comparative studies focusing on transgenic through cp4EPSPS gene and non-transgenic soybean plants: an analysis of protein species and enzymes. *J. Proteomics* **93**, 107–116
39. Ruebelt, M.C., Lipp, M., Reynolds, T.L., Schmuke, J.J., Astwood, J.D., DellaPenna, D., Engel, K.H., and Jany, K.D. (2006) Application of two-dimensional gel electrophoresis to interrogate alterations in the proteome of gentically modified crops. 3. Assessing unintended effects. *J. Agric. Food Chem.* **54**, 2169–2177
40. Albo, A.G., Mila, S., Digilio, G., Motto, M., Aime, S., and Corpillo, D. (2007) Proteomic analysis of a genetically modified maize flour carrying Cry1Ab gene and comparison to the corresponding wild-type. *Maydica* **52**, 443–455
41. Jin, X.N., Fu, Z.Y., Ding, D., Li, W.H., Liu, Z.H., Hu, Y.M., and Tang, J.H. (2014) Proteomic analysis of plumules and coleoptiles in maize between hybrids and their corresponding inbred lines. *Acta Physiol. Plant.* **36**, 355–370
42. Liu, Y., Zhang, Y.X., Song, S.Q., Li, J., Neal Stewart, C. Jr., Wei, W., Zhao, Y., and Wang, W.Q. (2015) A proteomic analysis of seeds from Bt-transgenic Brassica napus and hybrids with wild B. *Sci. Rep*. **5**, 15480
43. Zabalza, A., Orcaray, L., Fernández-Escalada, M., Zulet-González, A., and Royuela, M. (2017) The pattern of shikimate pathway and phenylpropanoids after inhibition by glyphosate or quinate feeding in pea roots. *Pestic. Biochem. Physiol.* **141**, 96–102
44. Herrmann, K.M. (1995) The shikimate pathway: early steps in the biosynthesis of aromatic compounds. *Plant Cell* **7**, 907–919
45. Mir, R., Jallu, S., and Singh, T.P. (2015) The shikimate pathway: review of amino acid sequence, function and three-dimensional structures of the enzymes. *Crit. Rev. Microbiol.* **41**, 172–189
46. Liu, W., Shang, Y., and Li, W. (2015) gp78 elongates of polyubiquitin chains from the distal end through the cooperation of its G2BR and CUE domains. *Sci. Rep.* **4**, 7138
47. Liu, W., Shang, Y., Zeng, Y., Liu, C., Li, Y., Zhai, L., Wang, P., Lou, J., Xu, P., Ye, Y., and Li, W. (2014) Dimeric Ube2g2 simultaneously engages donor and acceptor ubiquitins to form Lys48-linked ubiquitin chains. *EMBO J.* **33**, 46–61
48. Liu, C., Liu, W., Ye, Y., and Li, W. (2017) Ufd2p synthesizes branched ubiquitin chains to promote the degradation of substrates modified with atypical chains. *Nat. Commun.* **8**, 14274
49. Tan, Y.H., Tong, Z., Yang, Q., Sun, Y., Jin, X., Peng, C.Z., Guo, A.P., and Wang, X.C. (2017) Proteomic analysis of phytase transgenic and non-transgenic maize seeds. *Sci. Rep.* **7**, 9246