Significance

Glucosinolates in Brassicales constitute an important group of natural metabolites important for plant defense and human health. Its biosynthetic pathways and transcriptional regulation have been well-studied. Using *Arabidopsis* mutants of important genes in glucosinolate biosynthesis, quantitative proteomics and metabolomics led to identification of many proteins and metabolites that are potentially related to glucosinolate metabolism. This study provides a comprehensive insight into the molecular networks of glucosinolate metabolism, and will facilitate efforts toward engineering and breeding of glucosinolate profiles for enhanced crop defense, and nutritional value.

New nodes and edges in the glucosinolate molecular network revealed by proteomics and metabolomics of *Arabidopsis myb28/29* and *cyp79B2/B3* glucosinolate mutants

Islam Mostafa^{a,b,c}, Ning Zhu^{a,b}, Mi-Jeong Yoo^{a,b}, Kelly M. Balmant^{a,b,d}, Biswapriya B. Misra^{a,b}, Craig Dufresne^e, Maged Abou-Hashem^c, Sixue Chen^{a,b,d,f,*}, Maher El-Domiaty^c

^a Department of Biology, University of Florida, Gainesville, FL 32610, USA

^b Genetics Institute, University of Florida, Gainesville, FL 32610, USA

^c Department of Pharmacognosy, Faculty of Pharmacy, Zagazig University, Zagazig 44519, Egypt

^d Plant Molecular and Cellular Biology Program, University of Florida, Gainesville, FL 32610, USA

^e Thermo Fisher Scientific, West Palm Beach, FL 33407, USA

^f Interdisciplinary Center for Biotechnology Research, University of Florida, Gainesville, FL 32610, USA

List of abbreviations: GO, Gene Ontology; IAOx, Indole-3-acetaldoxime; PP Neg, polar phase separation/negative ion mode electrospray ionization (ESI); RP Neg, reversed phase separation/negative ion mode ESI; RP Pos, reversed phase separation/positive ion mode ESI; T-DNA, transfer-DNA; WT, *Arabidopsis thaliana* wild type.

* To whom correspondence should be addressed: Sixue Chen, Cancer and Genetics Research Complex, University of Florida, Gainesville, FL 32610, USA. Tel.: +1-352-2738330; Fax: +1-352-2738284; Email: schen@ufl.edu

Key words: *Arabidopsis*, genetic perturbation, iTRAQ, stress and defense, aliphatic and indolic glucosinolates

ABSTRACT

Glucosinolates present in Brassicales are important for human health and plant defense against insects and pathogens. Here we investigate the proteomes and metabolomes of *Arabidopsis myb28/29* and *cyp79B2/B3* mutants deficient in aliphatic glucosinolates and indolic glucosinolates, respectively. Quantitative proteomics of the *myb28/29* and *cyp79B2/B3* mutants led to the identification of 2785 proteins, of which 142 proteins showed significant changes in the two mutants compared to wild type (WT). By mapping the differential proteins using STRING, we detected 59 new edges in the glucosinolate metabolic network. These connections can be classified as primary with direct roles in glucosinolate metabolism, secondary related to plant stress responses, and tertiary involved in other biological processes. Gene Ontology analysis of the differential proteins showed high level of enrichment in the nodes belonging to metabolic process including glucosinolate biosynthesis and response to stimulus. Using metabolomics, we quantified 292 metabolites covering a broad spectrum of metabolic pathways, and 89 exhibited differential accumulation patterns between the mutants and WT. The changing metabolites (e.g., γ -glutamyl amino acids, auxins and glucosinolate hydrolysis products) complement our proteomics findings. This study contributes toward engineering and breeding of glucosinolate profiles in plants in efforts to improve human health, crop quality and productivity.

1. Introduction

Glucosinolates present primarily in members of Brassicales [1] are a group of plant secondary metabolites known for their strong anti-carcinogenic activities in animals and humans [2,3]. In plants, they play key roles in defense and interaction with the environment [4,5]. Based on the amino acid precursors of chain-elongated methionines, tryptophan and phenylalanine, glucosinolates are chemically categorized as aliphatic, indolic, and aromatic, respectively [5,6]. Although glucosinolate biosynthetic pathway is well studied in the reference plant *Arabidopsis thaliana* [6-8], our understanding of the molecular networks that regulate glucosinolate metabolism is far from complete [9]. Several studies have been conducted in recent years to explore connections between the glucosinolate biosynthetic pathways and other signaling and metabolic pathways. For examples, reduction in aliphatic glucosinolate production was shown to increase auxin levels [10]. In addition, indolic glucosinolates and auxin share the common precursor indole-3-acetaldoxime (IAOx). Thus, over-production of IAOx resulted in increased levels of indolic glucosinolates and auxin, while blockage of indolic glucosinolate biosynthesis downstream of IAOx was found to channel IAOx to increase auxin levels [11,12]. A direct proportional relationship between aliphatic glucosinolate biosynthesis and jasmonates was also reported [13]. Another group of studies focused on the relationship between glucosinolates, their hydrolysis products, and stress responses [14,15].

MYB28, *MYB29* and *MYB76* encode major transcriptional factors regulating aliphatic glucosinolate biosynthesis [5,16]. For example, they regulate the transcription of *MAM1*, *MAM2* and *MAM3*, encoding the enzymes for converting α -keto acids to chain-elongated methionines. The chain-elongated methionines are further metabolized by CYP79F1 and CYP79F2 to produce aldoximes, which are then converted to *aci*-nitro compounds by CYP83A1 [2], which in turn

converted to thiohydroximate by glutathione S transferase U20 and rearranged to desulfoglucosinolate [13]. SOT17 and SOT18 catalyze the sulfation step of desulfoglucosinolates to form glucosinolates [6]. Similarly, *MYB34*, *MYB51* and *MYB122* encode the major transcriptional factors that regulate the biosynthesis of indolic glucosinolates, and they regulate *TSBI* (important for conversion of indole to tryptophan), *CYP79B2* and *CYP79B3* (aldoxime formation from tryptophan), *CYP83B1* (*aci*-nitro compounds formation from aldoximes), *GSTF10* (thiohydroximate synthesis) and *SOT16* (responsible for the sulfation step in indolic glucosinolate biosynthesis) [2,6,13,16].

It is known that CYP79F1 is involved in the biosynthesis of both short- and long-chain aliphatic glucosinolates, while CYP79F2 exclusively catalyzes the production of long-chain aliphatic glucosinolates [17]. Using *CYP79F1* RNAi lines, and *CYP79F1* and *CYP79F2* knockout mutants, limited production of aliphatic glucosinolates and increased levels of indolic glucosinolates were observed [15, 17]. In addition, a *myb28/29* double mutant was generated to contain trace levels of aliphatic glucosinolates and enhanced levels of indolic glucosinolates [18]. Complementation experiments with either *MYB28* or *MYB29* showed recovery of short-chain aliphatic glucosinolates but not long-chain glucosinolates [18]. Interestingly, *MYB76* appeared to require *MYB28* and *MYB29* expression to produce long-chain glucosinolates [18]. On the other hand, the *cyp79B2/B3* double mutant showed very low concentrations of indolic glucosinolates and auxin, which share the same precursor [19], and also low production of aliphatic glucosinolates compared to WT [20]. Chen et al. [15] used proteomics and metabolomics of *CYP79F1* RNAi lines to discover glucosinolate connections with amino acids, carbohydrates and oxidative stress metabolites. Despite of the information from the above studies, our knowledge of the overall molecular network of glucosinolate metabolism is still limited.

In this study, we aimed to discover and characterize additional nodes and edges in the glucosinolate molecular network. As mentioned above, *MYB28*, *MYB29* and *MYB76* are the major controlling factors of aliphatic glucosinolate biosynthesis, and *MYB76* function is under the control of *MYB28* and *MYB29*. As to indolic glucosinolate biosynthesis, *MYB34*, *MYB51* and *MYB122* are the major controlling factors, but they also affect the tryptophan biosynthesis. To reduce pleiotropic effects, we chose to use the *Arabidopsis* double mutants *myb28/29* and *cyp79B2/B3* perturbed in the production of aliphatic and indolic glucosinolates, respectively. Proteomics and metabolomics together with bioinformatics allowed discovery of new nodes and edges in the glucosinolate molecular network. Specifically, we used iTRAQ approach for the proteomics work, LC-MS and GC-MS for the metabolomics analysis, STRING and AgriGO for molecular networking and gene ontology analysis, respectively.

2. Material and methods

2.1. Plant growth, genotyping and sample collection

The seeds of *Arabidopsis thaliana* (L.) Heynh ecotype Columbia (Col-0) were obtained from the *Arabidopsis* Biological Resource Center (Columbus, OH, USA), and the double mutants of *myb28/29* and *cyp79B2/B3* were kindly provided by Dr. Masami Hirai (RIKEN Plant Science Center, Yokohama, Japan) and Dr. John Celenza (Boston University, Boston, MA, USA), respectively. The seeds were sterilized using 70% ethanol (0.5 min) and 50% sodium hypochlorite (8 min), followed by washing with sterilized water. The sterilized seeds were kept at 4°C for stratification. After four days, the seeds were germinated on Murashige and Skoog medium with 1% sucrose, 0.1% Gamborg's vitamin and 0.8% agar in a controlled growth chamber under a 8 h light/16 h dark cycle at 22°C/20°C, respectively. After seven days, plant

seedlings were transferred to moistened soil and grown under the same conditions for four weeks. DNA extraction was carried out from seven day-old seedlings according to Kasajima et al. [21] and the genotypes of the double mutants were confirmed through genomic PCR using the primers listed in Supplementary Table 1 [18,19], Biolabs Taq 2x Master Mix (Ipswich, MA, USA), and the PCR program: 95°C for 2 min, 35 cycles of denaturation at 95°C for 30 sec, annealing at 50°C for 30 sec, and extension at 68°C for 1 min, followed by a final extension at 68°C for 5 min and holding at 10°C. The amplified segments were separated using 2% agarose gel electrophoresis. The insertion sites were confirmed by DNA sequencing of the PCR fragments generated by LB primer and the gene specific primers. Leaves of five week-old WT, *myb28/29* and *cyp79B2/B3* double mutants were collected and frozen immediately in liquid nitrogen and stored at -80°C. To minimize individual variation, we pooled four plants per each genotype, and six, four, and three biological replicates were prepared for proteomics, metabolomics, and glucosinolate analysis, respectively. To test the effect of glucosinolate metabolic perturbation on plant growth, we measured fresh weight and dry weight of leaves from four plants of each genotype.

2.2. Glucosinolate content analysis

Before glucosinolate extraction, 40 µl of 5 mM benzyl glucosinolate (internal standard) was added to each sample. The glucosinolates were extracted thrice from 20 mg of seeds or 200 mg of fresh leaves and desulfonated as described by Yan et al. [13]. The eluted desulfoglucosinolates were lyophilized and reconstituted in 100 µl and 60 µl of water for the seed and leaf samples, respectively. A total of 25 µl of extract from each sample were injected onto a HPLC system (Shimadzu, Kyoto, Japan), and glucosinolates separation was performed on a C-18 column (250 mm x 4.6 mm, 5µm, ODS-HYPERSIL, Thermo Scientific, USA) and eluted

by running the mobile phase at flow rate 1 ml/min with the same gradient as described by Yan et al. [13]. Peaks were monitored at 229 nm. Peaks identities were confirmed by matching retention times to an in house glucosinolate database. For each glucosinolate, the data were shown as average of the three biological replicates and standard error.

2.3. Protein extraction, iTRAQ labeling and peptide desalting

Proteins were extracted from 500 mg of leaves of WT (four biological replicates), *myb28/29* (six biological replicates) and *cyp79B2/B3* (six biological replicates) using the phenol method [22] for two independent iTRAQ experiments. Each sample was ground in liquid nitrogen and extracted in 1.5 ml of extraction buffer (0.1 M Tris-HCl pH 8.8, 10mM EDTA, 0.9 M sucrose, 0.4% β - mercaptoethanol, and 1% protease inhibitor) and 1.5 ml of phenol saturated with Tris-HCl pH 8.8. The mixture was agitated at room temperature for 2 h. Centrifugation of the mixture was carried out at 5000 xg for 10 min, the phenol phase that containing the proteins was transferred to new tubes and the residue was re-extracted with 1.5 ml of phenol saturated with Tris-HCl pH 8.8 by agitation for 30 min. Proteins were precipitated with 5 volumes of ice cold 0.1 M ammonium acetate in 100% methanol over night at -20 °C. The precipitated protein was pelleted by centrifugation at 20000 xg for 20 min at 4 °C, and washed twice with ice cold 0.1 M ammonium acetate in 100% methanol, and twice with 100% acetone. The washed protein pellets were dissolved in a solubilization buffer (7 M urea, 2 M thiourea, 4% CHAPS). EZQ assay kit was used for measurement of the protein amount [23] according to manufacturer instructions (Invitrogen Inc., Eugene, OR, USA). A total of 100 μ g of protein was precipitated from each sample using ice cold 90% acetone overnight at -20 °C followed by centrifugation at 20000 xg for 15 min at 4 °C. The precipitated proteins were washed with ice cold 80% acetone then with ice cold 100% acetone. The recovered protein pellets were processed according to the

iTRAQ 8-plex kit manual (AB Sciex, Inc., Foster City, CA, USA). In brief, each sample was digested with 4 µg modified trypsin (Promega, Madison, WI, USA) at 37°C for 16 h. The iTRAQ labeling reagents were dissolved in 55 µl of isopropanol, the content of each solubilized iTRAQ tag was transferred to the corresponding sample, and WT replicates were labeled with 113 and 114 tags, *myb28/29* replicates with 115, 116 and 117 tags and *cyp79B2/B3* replicates with 118, 119 and 121 tags at room temperature for 4 h. After 30 min of water quenching, the labeled samples were combined together. The labeled peptides were lyophilized, solubilized in 3% acetonitrile in 0.1% formic acid and desalted on a Macrospin C-18 reverse phase mini-column with a capacity of 30-300 µg (The Nestgroup Inc., Southborough, MA, USA). Before sample loading the column was washed with acetonitrile and 0.1% formic acid. After sample loading it was washed again with 0.1% formic acid and the peptides were eluted with 80% acetonitrile and 0.1% formic acid, followed by lyophilization to dryness.

2.4. Strong cation exchange fractionation and LC-MS/MS

The iTRAQ labeled peptides were fractionated as described by Zhu et al. [24] with minor modification in the gradient. The gradient program included running solvent A (25% (v/v) acetonitrile, 10 mM ammonium formate, and 0.1% (v/v) formic acid, pH 2.8) for 10 min, a linear gradient of 0- 20% solvent B (25% (v/v) acetonitrile and 500 mM ammonium formate, pH 6.8) over 80 min, and then ramping up to 100% solvent B in 5 min and holding for 10 min. Peptide absorbance at 280 nm was monitored, and based on the resulted chromatograms the collected fractions were combined into 12 fractions. The fractions were lyophilized and resuspended in LC solvent A (0.1% formic acid in water, 3% acetonitrile). Each fraction was analyzed on an Easy-nLC 1000 system coupled to a Q-Exactive Orbitrap Plus MS (Thermo Fisher Scientific, Bremen, Germany). The peptides were concentrated on an Acclaim Pepmap 100 pre-column

(20 mm × 75 µm; 3 µm-C18), then separated on a PepMap RSLC analytical column (250 mm × 75 µm; 2 µm-C18) at a flow rate of 350 nl/min using solvent A (0.1% formic acid) and B (0.1% formic acid, 99.9% acetonitrile), the solvent gradient was 2-30 % solvent B in 100 min, 30-98 % solvent B in 10 min then isocratic of 98% solvent B for 10 min. The separated labeled peptides were analyzed on the Q-Exactive Plus MS (Thermo Fisher Scientific, Bremen, Germany), which was run in positive ion mode and data dependent scanning with high collision dissociation (HCD) as previously described by Jones et al. [25] with minor modifications. The chromatography peak width was 4 sec and the default charge state was 3. The full MS resolution was 70000 with a scan range of 400- 2000 m/z , an AGC target of 1e6 and maximum IT of 100 ms. The MS² resolution was 17500 with a scan range of 200- 2000 m/z , an AGC target of 1e6, and maximum IT of 60 ms, a loop count of 10, an isolation window of 2 m/z , the fixed first mass of 105 m/z and the NCE of 28. The underfill ratio was 1%, intensity threshold was 1.7e5, charge exculsion was 1, 5-8, and >8. A lock mass of polysiloxane ion (445.12003 m/z) was used for real-time mass calibration. The spray voltage was set to be 1700, the capillary temperature was 260 °C, and the maximum spray current was 50.

2.5. Protein identification and quantification

Proteome Discoverer 1.4 (Thermo Scientific, Bremen, Germany) was used for protein identification based on searching the MS raw data files against the Arabidopsis *Tair10* database (downloaded in June, 2014) with 35,386 entries. The input criteria were set as following: The minimum precursor mass was set to 300 Da and the maximum to 5000 Da as spectrum property filters. Trypsin digestion (at K and R) with a maximum of two missed cleavage sites, static modification of MMTS on C, and dynamic modifications of iTRAQ8plex on any N terminus, phosphorylation on S, T and Y, oxidation of M, iTRAQ8plex on K. The precursor mass

tolerance was set to 10 ppm and the fragment mass tolerance to 0.01 Da. The maximum retention time difference for spectrum grouping was set to be 1.1. The strict target false discovery rate (FDR) was set to 0.01, while the relaxed target was set to 0.05. The peak integration tolerance was set to 20 ppm, mass precision to 2 ppm and signal/noise threshold to 3. High confidence level of peptides identification was used as a filter for the results, which aided in avoiding uncertainty in protein identification. We adopted the following cutoff values of Xcorr that are commonly used for the SEQUEST algorithm: 2.31 for 2+, 2.41 for 3+, and 2.6 for 4+ and 5+ peptides. Labeled unique peptides were selected for the relative protein quantification. To avoid effect of technical and systematic errors on quantification results, two steps data normalization were performed, in the first, peptides were normalized to the total summation of their intensities based on the 113 tagged peptides. Protein grouping was performed by summation of repeated peptides followed by median calculation for the summed peptides belonging to the same protein. After ratio calculation between each mutant protein value and the corresponding WT value, a second normalization step for all the ratios to a median equal to one was carried out. Log transformation of the normalized ratios and t-testing for significant differences between the WT and mutants protein expression values were conducted with the null hypothesis of $\log_2$ (averaged ratio) = 0. To exclude random errors arising in any replicate fold change filtering at a cutoff >1.2 and <0.8 with p -value < 0.05 was used to determine significant changes in the protein levels [15]. The raw proteomics data were deposited to ProteomeXchange via PRIDE repository [26] with accession number PXD002284.

2.6. Glucosinolate networking and Gene Ontology analysis

To test the connection between the proteins with significant changes in this study with those known to be involved in glucosinolate metabolism [9,13], we generated molecular network

maps in the confidence view using STRING [27]. The strength of interaction in the confidence view is reflected by the thickness of the edges. Node interaction prediction was based on neighborhood, gene fusion, co-occurrence, co-expression, published results, and text-mining information. To determine the level of functional enrichment between the significantly changed proteins, a complete GO analysis using agriGO database was carried out [28]. We used the Singular Enrichment Analysis (SEA) tool and Arabidopsis genome locus TAIR10 as a background. We selected Fisher test for comparing the input list to the background, Yekutieli as multi-test adjustment method, 0.05 as a cutoff significance level and adjusted the minimum number of mapping entries to 5. The results from the two mutants were compared together using the SEA cross comparison tool.

2.7. Metabolomics analysis

Four biological replicates from each genotype (35 mg of dry weight) were used for the metabolomics study. The analysis was conducted by Metabolon Inc. (Durham, NC, USA). Automated MicroLab STAR® system from Hamilton Company (Reno, NV, USA) was used for sample preparation, and metabolites were extracted in methanol by vigorous shaking for 2 min using a Geno/Grinder 2000 homogenizer (Glen Mills Inc., NJ, USA) followed by centrifugation. The resulting extracts were subjected to GC-MS analysis and UPLC-MS/MS analyses using RP Pos, RP Neg and PP Neg. Samples analyzed by GC-MS were dried under vacuum for 18 h, then derivatized using bistrimethyl-silyltrifluoroacetamide under dried nitrogen. Derivatized metabolites were separated on a 5% diphenyl / 95% dimethyl polysiloxane fused silica column (20 m x 0.18 mm ID; 0.18 um film thickness) as described by Yan et al. [13]. The LC-MS samples were dried and solubilized in acidic or basic LC-compatible solvents: 0.1% formic acid in water for positive mode and 6.5 mM ammonium bicarbonate in water at pH 8 for negative

mode. For LC-MS, a Waters UPLC BEH C-18 (2.1 x 100 mm, 1.7 μ m) column was used in an Acquity UPLC system (Waters, Millford, MA, USA). The samples dissolved in 0.1 % formic acid were eluted at a 350 μ l/min flow rate using 0.1% formic acid in water as solvent A and 0.1% formic acid in methanol as solvent B, the gradient was 0.5% B to 70% B in 4 min, 70 to 98% B in 0.5 min, isocratic 98% B for 0.9 min, back to 0.5% B in 0.2 min and isocratic 0.5% B for 5.4 min. The 6.5 mM ammonium bicarbonate solubilized samples were eluted using the same gradient and flow rate with solvent A consisting of 6.5 mM ammonium bicarbonate in water at pH 8 and solvent B consisting of 6.5 mM ammonium bicarbonate in 95% methanol in water. For polar phase separation, samples were eluted from a HILIC column (Waters UPLC BEH Amide 2.1 x 150 mm, 1.7 μ m) using a gradient consisting of water (A) and acetonitrile (B) with 10 mM ammonium formate at 500 μ l/min flow rate. The gradient was ramped from 5% B to 50% B in 3.5 min then to 95% B in 2 min followed by isocratic 95% B for 1 min then fall to 5% B in 0.2 min and fixed at that concentration for 4.3 min. The mass analyses were conducted by interfacing the UPLC to a Q-Exactive high resolution/accurate mass spectrometer (Thermo Scientific Inc., San Jose, CA, USA) connected with a heated electrospray ionization (HESI-II) source and Orbitrap mass analyzer operated at 35,000 mass resolution. For RP Pos, RP Neg and PP Neg, the sheath gas flow was 80, 75 and 60 au, the auxiliary gas flow was 12, 15 and 20 au, the spray voltage was 4, 2.75 and 3 kv, the source heat temp was 400, 400 and 380 $^{\circ}$ C and the ion transfer tube temp was 300, 300 and 400 $^{\circ}$ C, respectively. The scan range was from 80-1000 m/z . Alternation between MS and MS² scans was in 9 s using a dynamic exclusion time of 3 s. The MS max fill-time was 60 ms, the MS² max time was 120 ms and the MS isolation window was 3 m/z [29]. Metabolite identification was relied on comparing retention index, accurate mass +/- 0.005 amu and MS/MS spectral data with the authentic compound information in the Metabolon library. Peaks were quantified using area under the curve. Metabolites appearing in at

least six samples out of the 12 analyzed samples were considered as true positives for our study. To evaluate the real difference between each mutant's metabolites relative to WT, we carried out statistical analysis relying on two sample independent t-test using the SPSS statistics 22 software (<http://www-01.ibm.com/software/analytics/spss/>). Differences were considered significant when the p -value < 0.05 . KEGG (<http://www.genome.jp/kegg/pathway.html>) was used for plotting the metabolite changes onto the metabolic pathways [13].

3. Results

3.1. Molecular characterization, Morphological and chemical phenotypes of the *myb28/29* and *cyp79B2/B3* mutants

Sequencing results of the PCR amplified DNA fragments of the different mutants showed that T-DNA insertion sites in the *myb28/29* are at 182 bp before the starting codon in the promoter region of *MYB28* and at 10 bp before the starting codon in the 5' UTR of *MYB29* (Fig. 1A). In *cyp79B2/B3*, the T-DNA insertions are located at 1509 bp after the start codon in the second exon of *CYP79B2* and at 1365 after the start codon in the intron of *CYP79B3* (Fig. 1A). Gel electrophoresis of the amplified DNA fragments from the mutants in comparison with the WT confirmed the presence of homozygous double mutants of *myb29/29* and *cyp79B2/B3* (Fig. 1B). In agreement with Zhao et al. [19], we found a tandem T-DNA insertion with LBb1.3 flanking sequences at the two sides of T-DNA insertion in *CYP79B3*.

Morphological examination of the tested genotypes revealed no obvious differences between *myb28/29* and WT, while the *cyp79B2/B3* plants appeared chlorotic at leaf tips and the leaves tended to roll toward the lower epidermis (Fig. 1C). The growth of *myb28/29* double

mutant was found to be retarded in the first few days after germination, but it could catch up with the WT later [30]. By comparing the average fresh and dry weights of leaves, it became obvious that WT growth was slightly better than the mutants (Supplementary Table 2), consistent with previous findings [31].

HPLC analysis of the *myb28/29* mutant seeds showed very low levels of 8-methylsulfinyloctyl (8MSOO), 7-methylthioheptyl (7MTH) and 8-methylthiooctyl (8MTO) glucosinolates and no detection of other aliphatic glucosinolates. The indolic glucosinolates, 4-hydroxy-indolyl-3-methyl (4OHI3M), 4-methoxy-3-indolylmethyl (4MTI3M) and 1-methoxy-3-indolylmethyl (1MTI3M) showed decreases in the *myb28/29* seeds compared to WT, while there is no significant change in the levels of indol-3-ylmethyl (I3M) glucosinolate (Supplementary Table 3). In the *myb28/29* leaves, the aliphatic glucosinolate 5-methylsulfinylpentyl (5MSOP) was increased, while 6-methylsulfinylhexyl (6MSOH) and 4-methylthiobutyl (4MTB) glucosinolates were decreased. There were no significant changes in the indolic glucosinolates (Supplementary Table 3). Overall, the total aliphatic glucosinolate levels decreased to less than 0.29% and 14.9% of WT levels in the *myb28/29* seeds and leaves, respectively, whereas the total indolic glucosinolate levels were not significantly changed.

In the *cyp79B2/B3* seeds, the levels of aliphatic glucosinolates (except 6-methylthiohexyl (6MTH)) remained unchanged compared to WT, while all the indolic glucosinolates were significantly decreased (Supplementary Table 3). In *cyp79B2/B3* leaves, except for significantly decreased 8-methylsulfinyloctyl (8MSOO), other aliphatic glucosinolates were not changed. In contrast, all the indolic glucosinolates were significantly decreased (Supplementary Table 3). The total aliphatic glucosinolate levels decreased to about 88.4% and 78.2% of WT in the seeds and leaves, respectively. The total levels of indolic glucosinolates were decreased to 10.6% and

4.7% of WT in the seeds and leaves, respectively. These data showed predicted perturbation of aliphatic glucosinolates in the *myb28/29*, and indolic glucosinolates in the *cyp79B2/B3* [18,19].

3.2. Overview of the *myb28/29* and *cyp79B2/B3* proteomes

Analysis of the iTRAQ labeled peptides of WT, *myb28/29* and *cyp79B2/B3* resulted in identification of 2785 proteins in two independent experiments based on high confidence peptide filtering (Supplementary Table 4). A total of 1833 proteins were shared between the two experiments, while 480 and 472 were unique to experiment 1 and 2, respectively (Fig. 2A). A total of 2100 proteins were quantified based on the labeled unique peptides. Quantification of the proteins from the mutants relative to WT based on a *p*-value < 0.05 and a fold change cutoff (>1.2 and < 0.8) revealed 56 increased (Fig. 2B) and 86 decreased proteins (Fig. 2C). We employed STRING analysis to test the connections among the proteins affected by the glucosinolate perturbation and those known to be involved in the glucosinolate metabolism [9,13]. As a result, we identified new edges and nodes in the glucosinolate molecular networks (Figs. 3 and 4) and their implications are discussed in detail in the following sections. We classified these nodes as primary with direct roles in glucosinolate metabolism, secondary related to plant stress responses, and tertiary involved in other biological processes.

3.2.1 Shared proteome changes between the *myb28/29* and *cyp79B2/B3* mutants

Among the proteins showing significant differences, there were 36 shared proteins between the two mutants with 8 increased and 28 decreased in levels (Table 1). Based on the STRING generated molecular networks (Figs. 3 and 4) and the levels of shared increased and decreased proteins (Table 1), seven new nodes were found to form biologically relevant edges to

the glucosinolate metabolism, including glucan endo-1,3-beta-glucosidase (acidic isoform) (BGL2), glutathione S-transferase F7 (GSTF7), glutathione S-transferase F2 (GSTF2), DNA-damage-repair/tolerance protein (DRT102), mediator of RNA polymerase II transcription subunit 37c (MED37D), Aha1 domain-containing protein (At3g12050) (T21B14.13), and L-ascorbate peroxidase 1 (APX1). These proteins are important in stress and defense responses [32-34,39,40,49-52] and can be considered as secondary nodes. As a secondary effect of glucosinolate metabolic perturbation, translation initiation factor isoform 4E (EIF4E), 30S ribosomal protein S8 (RPS8), 40S ribosomal protein S8-1 (RPS8A) and Myb domain protein 79 (AtMYB79) formed new tertiary nodes in the glucosinolate molecular network. Several proteins formed new nodes in either aliphatic or indolic glucosinolate molecular network. For example, phosphoglycerate mutase (PGM) in the shared protein category formed a new tertiary node with the aliphatic glucosinolate molecular network. Similarly, three proteins, aldose 1-epimerase family protein (F14M19.180), partner of Y14-mago (PYM), and NADH dehydrogenase flavoprotein 2 (At4g02580) formed new tertiary nodes in the indolic glucosinolate molecular network. The cysteine protease component of protease-inhibitor complex (At5g43060) known to be related to stress response [48] formed a secondary node with the indolic glucosinolate molecular network.

3.2.2. Proteome changes specific to the *myb28/29* mutant

As a result of the proteome study of the *myb28/29* mutant, we found 11 and 24 unique proteins to be significantly increased or decreased compared to WT, respectively (Table 2). The STRING analysis of the differential proteins revealed 11 novel nodes connected to the glucosinolate metabolic network (Fig. 3). Two of these nodes have primary connections to methionine and consequently to aliphatic glucosinolate metabolism, and they are peptide

methionine sulfoxide reductase A2 (PMSR2) [81] and 5-methyltetrahydropteroyltriglutamate-homocysteine methyltransferase 2 (ATMS2) [83]. As a membrane oxygenase electron carrier [78], cytochrome B5 isoform C (CYTB5-C) node may be considered biologically relevant to glucosinolate metabolism. Its secondary connection to glucosinolate metabolism may be through its role in the oxidation/reduction reactions that are associated with biotic stress like glutathione-S-transferases [79]. In addition, a lectin-like protein (At3g16530) may be another biologically relevant secondary node due to its immune response activity to bacterial pathogens [67]. Other nodes such as enolase 1 (ENO1), NADP-dependent malic enzyme 2 (NADP-ME2), glyceraldehyde-3-phosphate dehydrogenase (GAPC1), L-3-cyanoalanine synthase/cysteine synthase D1 (CYSD1), putative dehydrogenase (At3g17810) (PYD1A), 60S ribosomal protein L7a-1 (RPL7A) and pectinesterase/pectinesterase inhibitor 18 (PME18) could be considered to be tertiary as there is no evidence for them to be directly involved in the glucosinolate metabolism or related to glucosinolate physiological functions.

3.2.3. Proteome changes specific to the cyp79B2/B3 mutant

Thirty-seven increased proteins and 34 decreased proteins exhibited specific changes in the *cyp79B2/B3* mutant (Table 3). Mapping of these proteins using STRING found 30 newly formed nodes (Fig. 4). The role of anthranilate phosphoribosyltransferase (PAT1) [94,95], phosphoserine aminotransferase (PSAT) [96] and indole-3-glycerol phosphate synthase (IGPS) [109] in tryptophan biosynthesis classified these nodes as primary to the glucosinolate molecular network. The stress and defense response-related proteins such as CaLB domain family protein (At4g34150) [90], lectin-like protein (At5g03350) [67], peroxidase 34 (PRXCB) [97,98], hypersensitive-induced response protein 2 (HIR2) [107], protein SGT1 homolog B (SGT1B)

[114] and thioredoxin H5 (TRX5) [115] constituted biologically relevant secondary nodes in the glucosinolate molecular network. In addition, cytochrome c oxidase subunit 5b-2 (COX5B-2) as a mitochondrial electron transporter may be considered as a tertiary node due to its involvement in the oxidation/reduction reactions similar to the cytochrome B5 isoform C in *myb28/29*. The nodes of DNA replication licensing factor (MCM4), Hyaluronan/mRNA binding family protein (At5g47210), ATP-dependent zinc metalloprotease (FTSH4), nudix hydrolase 7 (NUDT7), protein disulfide-isomerase 2-3 (PDIL2-3), protein disulfide isomerase-like 1-1 (PDIL1-1), chaperonin 60 subunit beta 3 (CPN60B3), flotillin-like protein 1 (FLOT1), heat shock 70 kDa protein 10 (HSP70-10), 3-hydroxyisobutyryl-CoA hydrolase-like protein 3 (At4g13360), DNA-directed RNA polymerase subunit alpha (RPOA), 26S proteasome non-ATPase regulatory subunit 4 (RPN10), 40S ribosomal protein S6-2 (RPS6B), 26S proteasome non-ATPase regulatory subunit 1 homolog A (RPN2A), eukaryotic translation initiation factor 3 subunit C (EIF3C1), branched-chain amino acid aminotransferase-like protein 3 (At5g57850), ankyrin repeat family protein (ANK), LisH/CRA/RING-U-box domains-containing protein (At4g37880), glutamine synthetase cytosolic isozyme 1-1 (GLN1-1) and PP2A regulatory subunit TAP46 (TAP46) were categorized as tertiary connections. Surprisingly, we detected a side network formed of 13 uniquely changed proteins in the *cyp79B2/B3* mutant, plus fatty-acid-binding protein 1 (FAP1), triphosphate tunnel metalloenzyme 1 (At1g73980), DnaJ/Hsp40 cysteine-rich protein (F1P2.200) and plastid-lipid-associated protein 5 (PAP5) from the shared proteins affected by the indolic glucosinolate perturbation. Some of the nodes may be biologically relevant, such as the stress related PsbP-like protein 1 (PPL1) [111], nucleus-encoded thylakoid protein (CEST) [133] and wound-responsive protein-related (At3g07230, T1B9.10). The other nodes include light-harvesting-like protein (LIL3.1), photosystem II 10 kDa polypeptide (PSBR), ribulose biphosphate carboxylase small chain 1B (RBCS-1B), BolA like protein (At1g55805),

ferredoxin-NADP reductase leaf isozyme 1 (FNR1), 2-methyl-6-phytyl-1,4-hydroquinone methyltransferase (APG1), pentatricopeptide repeat-containing protein (CRR2), Unknown protein (At2g21960), putative uncharacterized protein (At1g33810) and metallo-beta-lactamase domain-containing protein (F15D2.25).

3.2.4. GO enrichment of significantly changed proteins

The biological processes (BP), cellular components (CC) and molecular functions (MF) of the proteins with significant level changes relative to WT were analyzed using agriGO. A complete GO was used to annotate 71 changed proteins in the *myb28/29*. The SEA results reflected that 121 GO terms were enriched for BP (Supplementary Figs. 1), 33 for CC (Supplementary Figs. 2), and two for MF (Supplementary Figs. 3). Similarly, in the *cyp79B2/B3* mutant, 107 changed proteins were annotated. The SEA resulted in 232 GO terms enrichment for BP (Supplementary Figs. 4), 47 for CC (Supplementary Figs. 5) and 6 for MF (Supplementary Figs. 6). A cross comparison of the SEA results from the two mutants using SEACOMPARE showed that two and four GO terms specifically enriched in *myb28/29* BP and CC, respectively, while 113, 18, and four GO terms were enriched specifically in *cyp79B2/B3* BP, CC, and MF, respectively (Supplementary Table 5). The highly enriched GO terms in BP can be categorized into two main processes, i.e., metabolic process and response to stimulus. Secondary metabolic process (including nitrogenous and sulfur compound metabolism, carbohydrate metabolism, indole, organic acid and ketone metabolism), and response to stress (including biotic and abiotic stimulus and defense response) represented the highly enriched nodes between the two mutants (Fig. 5, Supplementary Table 5). The high enrichment of the glucosinolate metabolic process,

response to biotic stimulus and response to stress in the two mutants (Fig. 5, Supplementary Figs. 1 and 4) was anticipated and supports the utility of the tools used in this study.

3.3. Metabolomic overview of the *myb28/29* and *cyp79B2/B3* mutants

Metabolomics of WT, *myb28/29* and *cyp79B2/B3* resulted in the identification of 292 metabolites. The detected metabolites belong to different metabolic pathways of amino acids (30.48%), nucleotides (15.07%), carbohydrates (14.04%), lipids (14.04%), cofactors, prosthetic groups and electron carriers (8.56%), peptides (8.56%), secondary metabolism (7.19%), hormone metabolism (1.71%), and xenobiotics (0.34%) (Supplementary Table 6). Out of the 292 metabolites, 89 showed significant level changes or were completely absent compared to the WT (Fig. 6). For example, four metabolites (sulforaphane, sulforaphane-glutathione, sulforaphane-N-acetyl-cysteine, and sulforaphane-cysteine-glycine) were not detected in the *myb28/29* and sulforaphane was the only metabolite with a significant decrease in the *cyp79B2/B3*. Eight metabolites (indolepropionate, indolelactate, indole-3-carboxylic acid, methyl indole-3-acetate, indole-3-acetamide, indoleacetate, glycerophosphorylcholine, and glycerophosphoethanolamine) were not detected in the *cyp79B2/B3*, and the first four exhibited significant increases in the *myb28/29*. Six metabolites (phenylalanine, 2-hydroxyadipate, pipecolate, 3-hydroxy-3-methylglutarate, glutarate, and gentisic acid-5-glucoside) showed opposite changes in the two mutants, i.e., decreased in the *myb28/29* and increased in the *cyp79B2/B3* (Fig. 6B). As to the remaining seventy-one metabolites, we found that 10 increased specifically in the *myb28/29*, 28 increased specifically in the *cyp79B2/B3*, two increased in both mutants, 16 decreased specifically in the *myb28/29*, eight decreased specifically in the *cyp79B2/B3* and seven decreased in both mutants (Fig. 6C).

462

463 *3.3.1. Significant changes of the myb28/29 and cyp79B2/B3 shared Metabolome*

464 A glucosinolate hydrolysis product 3-methylsulfinylpropyl isothiocyanate exhibited 0.1%
465 of WT level in *myb28/29* and about 30% of WT level in *cyp79B2/B3* (Fig. 7, Supplementary
466 Table 6). In both mutants, glycine and methionine showed decreased levels, while phenylalanine,
467 2-hydroxyadipate, and pipecolate exhibited decreased levels in the *myb28/29* and increased
468 levels in the *cyp79B2/B3* (Fig. 8). Gentisic acid-5-glucoside, an aromatic secondary metabolite,
469 showed increased levels in the *cyp79B2/B3* in agreement with the observed increased activity of
470 aromatic amino acid metabolism. Interestingly, however, its level in the *myb28/29* was decreased
471 (Fig. 8). Elevated levels of some γ -glutamyl amino acids were observed, and the increase of γ -
472 glutamylthreonine was shared in both mutants (Fig. 9). As to the nucleotide metabolic pathway,
473 the purine metabolism related metabolite N6-succinyladenosine is the only metabolite that
474 increased in both mutants. The purine metabolism related metabolites adenosine-2',3'-cyclic
475 monophosphate, and guanosine-2',3'-cyclic monophosphate and the pyrimidine metabolism
476 related metabolite uridine-2',3'-cyclic monophosphate were all decreased in both mutants
477 (Supplementary Table 6). Other changes include decreased citramalate from the C5 branched
478 dibasic acid metabolic pathway in both mutants, and decreased 3-hydroxy-3-methylglutarate and
479 glutarate from the dicarboxylic fatty acid metabolic pathway in the *myb28/29*, but they were
480 increased in the *cyp79B2/B3* mutant (Supplementary Table 6).

481

482 *3.3.2. Unique metabolome changes in the myb28/29 mutant*

As a result of *MYB28/29* mutation, there was a direct perturbation of the glucosinolate metabolism, which is reflected by the absence of sulforaphane and its conjugation products sulforaphane-glutathione, sulforaphane-N-acetyl-cysteine and sulforaphane-cysteine-glycine (Fig. 7). In addition, we observed 2-3 fold elevated levels of the tryptophan metabolites indolepropionate, indolelactate and the auxin metabolites indole-3-carboxylic acid, methyl indole-3-acetate, and 2-oxindole-3-acetate, but no significant changes in indolic glucosinolate levels (Fig. 7, Supplementary Table 6). The glutathione γ -glutamyl amino acid cycle showed increases in the levels of some of its metabolites such as the glutathione reduced form, γ -glutamylglutamine and γ -glutamylleucine (Fig. 9). Pantothenate and its precursor alanine showed increases in levels together with nicotinamide adenine dinucleotide (Fig. 8). These increases reflect enrichment in the cofactors, prosthetic groups and electron carriers. At the same time trigonelline was the only metabolite decreased significantly in the cofactor, prosthetic group and electron carrier group. The dipeptide cysteinylglycine exhibited an increased level, while the nucleotide metabolism exhibited different patterns, e.g., increase in the levels of the purine metabolism related urate and 1-methyladenine and decrease in the pyrimidine metabolism related 2'-deoxycytidine in the *myb28/29* (Supplementary Table 6). In general, carbohydrate and dicarboxylic acid metabolism can be noted as decreased, as represented by the low levels of xylose, erythronate, glucuronate, myo-inositol, 3-deoxyoctulosonate, fumarate, malate and 2-hydroxyglutarate. Decreases of some amino acid related metabolites including N-acetylvaline, homoserine, methionine sulfoxide, N-acetylputrescine, tyramine and N-acetylleucine were also observed (Fig. 8, Supplementary Table 6).

3.3.3. Unique metabolome changes in the *cyp79B2/B3* mutant

Mutation of the *CYP79B2/B3* directly affected the indolic compound production, which is obvious from the absence of indolepropionate, indolelactate, indole-3-acetamide, indole-3-carboxylic acid, methyl indole-3-acetate, and indoleacetate in the mutant (Fig. 7, Supplementary Table 6). The increased levels of serine, O-sulfo-L-tyrosine, tyrosine, N-acetylphenylalanine, tryptophan, salicylate and 4-hydroxybenzoate reflected an expected metabolite diversion and activation of the aromatic amino acid derived pathways (Fig. 8, Supplementary Table 6). The elevated levels of amino acids, dipeptides, γ -glutamyl amino acids, purine metabolites and fatty acids may reflect the stressful status of the plant cells [13,15,139-141]. This notion is strongly supported by the high levels of the stress related metabolites in the *cyp79B2/B3* mutant, e.g., the aspartate family metabolites threonine, 2-aminoadipate and 6-oxopiperidine-2-carboxylic acid, the glutamate family N-acetylproline, the dipeptides valylleucine, phenylalanylglycine and leucylglycine, the γ -glutamyl amino acid γ -glutamylphenylalanine, the purines xanthine, adenosine, 2'-deoxyguanosine, 7-methylguanosine, guanosine, N6-carbamoylthreonyladenosine, N6-methyladenosine and allantoic acid, and the fatty acids malonate, and 1-palmitoylglycerophosphoinositol (Figs. 8 and 9, Supplementary Table 6). The chlorotic and curling leaf phenotype of the *cyp79B2/B3* mutant is also indicative of intracellular metabolic stress (Fig. 1). In agreement with Kliebenstein et al. [20] and our glucosinolate analysis results, which showed decrease in the total aliphatic glucosinolate level, sulforaphane also exhibited a lower level than WT (Fig. 7). Increased production of nicotinate ribonucleoside, choline and stachydrine alkaloid were observed. N-acetylalanine, kynurenate, 2-isopropylmalate, shikimate, 2-pyrrolidinone, norophthalmate, erythritol and benzamide showed lower levels than WT. Unexplained absence of glycerophosphorylcholine and glycerophosphoethanolamine in the *cyp79B2/B3* mutant was observed (Supplementary Table 6).

4. Discussion

Based on the results from STRING networking and agriGO functional analyses of the differential proteins, as well as the metabolomics data, we focus on discussing the effects of glucosinolate perturbation on cellular metabolic pathways, stress responses, and plant growth and defense trade-off.

4.1. Impact of genetic perturbation of glucosinolate biosynthesis on related metabolic pathways

Due to the mutation of *MYB28/29* known to be involved in controlling aliphatic glucosinolate biosynthesis [16,18], the levels of some enzymes that participate in the aliphatic glucosinolates biosynthesis were decreased, e.g., methionine aminotransferase BCAT4 involved in converting methionine into 4-methylthio-2-oxobutanoate in the methionine chain-elongation process [88], chloroplastic methylthioalkylmalate synthase 1 in condensing 4-methylthio-2-oxobutanoate with acetyl-CoA [58,59], and 3-isopropylmalate isomerase small subunit 1 working with LeuD1 or LeuD2 in methionine chain-elongation of aliphatic glucosinolate biosynthesis [6,82] (Fig. 7). Cytochrome P450 83A1 and glutathione S-transferase U20 producing thiohydroximates from chain-elongated methionine derived aldoximes [13,87], and cytosolic sulfotransferases 17 and 18 required for desulfoglucosinolate sulfation [69-71] were also decreased (Fig. 7). The decreased activities of aliphatic glucosinolate biosynthesis clearly feedback inhibits the precursor methionine biosynthesis, e.g., via decreasing the expression of 5-methyltetrahydropteroyltriglutamate-homocysteine methyltransferase 2, which catalyzes the conversion of homocysteine to methionine [83], and peptide methionine sulfoxide reductase A2 involved in the conversion of methionine sulfoxide to methionine [81] (Fig. 8). As a result, the levels of methionine and its precursors methionine sulfoxide and homoserine [13,81] were also low in the mutant compared to WT (Fig. 8).

In the *cyp79B2/B3* mutant, cellular activities for compensating the decreased levels of indolic glucosinolates resulted from the mutation of *CYP79B2/B3* were observed. Several enzymes closely related to indolic glucosinolate biosynthesis were found at higher levels than the WT, e.g., enzymes related to the biosynthesis of tryptophan (the key precursor for indolic glucosinolate biosynthesis [15]) including chloroplastic tryptophan synthase beta chain 1 required for the formation of tryptophan from (1S,2R)-1-C-(indol-3-yl) glycerol 3-phosphate and serine, chloroplastic anthranilate phosphoribosyltransferase catalyzing the biosynthesis of N-(5-phosphoribosyl)anthranilate from anthranilate [94,95], chloroplastic indole-3-glycerol phosphate synthase converting 1-(O-carboxyphenylalanine)-1-deoxyribulose 5-phosphate to (1S,2R)-1-C-(indol-3-yl) glycerol 3-phosphate [109], and chloroplastic phosphoserine aminotransferase enzyme for serine biosynthesis by forming 3-phospho-L-serine from 3-phospho-hydroxypyruvate [96] (Fig. 8). Cytochrome P450 83B1 that known to convert indole-3-acetaldoxime to aci-Nitro compound [13,99-103] was also increased (Fig. 7). Other enzymes showing higher levels in the *cyp79B2/B3* than WT include S-alkyl-thiohydroximate lyase SUR1 an enzyme catalyzing the synthesis of thiohydroximate from S-(alkylacetohydroximoyl)-cysteine [12], UDP-glycosyltransferase 74B1 for the glucosylation of the thiohydroximate group to form desulfoglucosinolates [89], and sulfotransferase 16 catalyzing the sulfate conjugation of desulfoglucosinolates [15,69-71] (Fig. 7). In addition, methionine S-methyltransferase showed lower levels in the *cyp79B2/B3* than WT (Fig. 8), which is important in S-methylmethionine formation and plays a role in the transport of phloem sulfur and the synthesis of selenium methylmethionine [130,131].

Flavin-containing monooxygenase 3 is known to play a key role in glucosinolate metabolism by converting methylthioalkyl glucosinolates into methylsulfinylalkyl glucosinolates [55], its amount was decreased in both mutants. In *myb28/29* this reduction may be attributed to

577 the absence of the precursor 4-methylthiobutyl glucosinolate (4MTB) in the seeds and very low
 578 concentrations in the leaves, and potentiated by the absence of the product 4-methylsulfinylbutyl
 579 glucosinolate (4MSOB) (Fig. 7, Supplementary Table 3). In the *cyp79B2/B3* mutant, low levels
 580 of methylthioalkylmalate synthase 1 (in methionine chain-elongation) and methionine were
 581 observed (Fig. 7) and they may contribute to the observed decrease in the total aliphatic
 582 glucosinolates (Supplementary Table 3). As a major representative for this group 4MTB
 583 exhibited decreasing trend in the seeds, although its change is not statistically significant ($p =$
 584 0.0557), this reduction may explain the decrease of FMO GS-OX3 level relative to WT. The
 585 decrease in the total aliphatic glucosinolate concentration in *cyp79B2/B3* is consistent with the
 586 previous finding [20]. Myrosinase 2 known by its glucosinolate hydrolysis activity [74] was
 587 decreased significantly in *myb28/29* (Fig. 7, Table 2) which may be due to the absence or very
 588 low concentrations of its substrates [15] (Fig. 7, Supplementary Table 3). In the *cyp79B2/B3*
 589 mutant, myrosinase 2 exhibited lower levels than WT (albeit not significant) (Supplementary
 590 Table 4). This is in agreement with the metabolomics data showing low levels of 3-
 591 methylsulfinylpropyl isothiocyanate and sulforaphane (the product of 4-methylsulfinylbutyl
 592 glucosinolate) in the mutant (Fig. 7, Supplementary Table 6). Overall, there appeared to be two
 593 types of cross-talk: the first is between the aliphatic and indolic glucosinolates, i.e., the decreased
 594 production of the aliphatic glucosinolates enhances the total indolic glucosinolates level by
 595 activation of the indole biosynthetic pathway as shown in the metabolomics data, while the
 596 decrease in indolic glucosinolates production decreases the aliphatic glucosinolate production by
 597 decreasing their biosynthetic precursor methionine and its converting enzymes, e.g.,
 598 methylthioalkylmalate synthase 1 and methionine S-methyltransferase. This finding is in
 599 agreement with data from *CYP79F1* RNAi lines [15], *myb28/29* double mutant [18], *CYP79F1*
 600 and *CYP79F2* knockout mutants [17] and *cyp79B2/B3* glucosinolate profile [20]. The other

cross-talk is between the glucosinolate levels and their hydrolyzing enzymes, i.e., when glucosinolate levels are low, the myrosinase levels are also low. However, the mechanisms underlying such cross-talks are not known.

4.2. Decrease of glucosinolate production results in activation of stress-related pathways

Several studies reported the relationships between glucosinolates and stress conditions such as salinity, drought, temperature, light, nutrient deficiency [14] and oxidative stress [15]. For example, Martinez-Ballesta et al [142] described that aliphatic glucosinolate deficiency in *myb28/29* affected plant hydraulic conductivity under salt stress. Our results added further evidence supporting the relationship between glucosinolates and cellular stress responses. The decrease in the glucosinolate levels resulted in stress status stimulation as indicated by changes of proteins and metabolites related to stress response. On the proteome level, it is interesting to note the elevated levels in GSTF2 and GSTF7 in both mutants (Table 1). Both enzymes were shown to play important roles in detoxification under stress conditions, e.g., by catalyzing the conjugation of reduced glutathione with xenobiotics [33,34,39,40]. In addition, cold-responsive proteins such as adenine nucleotide alpha hydrolases-like superfamily protein in *myb28/29* (Table 2) and CaLB domain family protein in the *cyp79B2/B3* mutant [90] (Table 3) were increased. Peroxidase 34 was increased in *cyp79B2/B3*, and it is known to be activated under stress conditions to remove hydrogen peroxide and other toxic compounds [97,98]. One of the lactoylglutathione lyase / glyoxalase I family protein (At5g57040) (involved in methylglyoxal detoxification [105]) and a NADP-dependent alkenal double bond reductase P1 (catalyzing phenylpropanol 7-8 double bond reduction and activated during anti-oxidant defense [106]) showed higher levels in the *cyp79B2/B3* mutant (Table 3). In addition, chloroplastic PsbP-like

protein 1 involved in photo-damage repair [111], thioredoxin H5 in stress response [115] and a wound-responsive protein (At3g07230) proposed to respond to wounding were all increased in the *cyp79B2/B3* (Table 3). As a part of the glutathione biosynthetic pathway, some of the γ -glutamyl amino acids exhibited elevated levels relative to WT (Fig. 9), and they are known to be involved in oxidative stress response [139]. Another indicator of the oxidative stress condition is the elevated levels of dipeptides, which reflect increased activities of proteolysis [15]. In this study, three dipeptides showed higher levels in the *cyp79B2/B3* together with the protein degradation products N-acetylphenylalanine and N-acetylproline, and one dipeptide showed increased level in the *myb28/29* (Supplementary Table 6). Other lines of evidence for the cellular stressful state of the *cyp79B2/B3* include the elevated levels of some amino acids and purines relative to WT [140,141] (Fig. 8, Supplementary Table 6) and the chlorotic phenotype (Fig. 1).

It is well known that glucosinolates play a key role in plant defense [4]. The mutants *myb28/29* and *cyp79B2/B3* lacking aliphatic and indolic glucosinolates, respectively were found to be more susceptible to fungal infection than WT [143]. In addition, the susceptibility of the *myb28/29* plants to *Pseudomonas* was reported and attributed to the lack of sulforaphane (the hydrolysis product of aliphatic glucosinolates) [144]. In this study, it is interesting to note that some enzymes related to defense mechanism were affected by glucosinolate metabolism perturbation. For example, glucan endo-1,3-beta-glucosidase (acidic isoform) involved in plant response to pathogens [32] and GSTF2 involved in the transport and binding of defense-related bioactive products [39,40] were increased in both mutants (Table 1). In the *myb28/29*, a lectin-like protein At3g16530 was increased (Table 2), and it is known to play a role in Arabidopsis immunological response against *Pseudomonas syringae* [67]. Hypersensitive-induced response protein 2 that potentiates disease defense [107] and SGT1 homolog B protein in the immunity system [114] showed higher levels in the *cyp79B2/B3* than WT (Table 3). At the metabolome

level, accumulation of aromatic compounds such as salicylate, 4-hydroxybenzoate and gentisic acid-5-glucoside was observed in the *cyp79B2/B3* (Fig. 8). These compounds are known for their antimicrobial and antioxidant activities [145-148], and they may compensate the low levels of the indolic glucosinolates in plant defense. The low concentration of shikimate in the *cyp79B2/B3* may be attributed to its turnover to produce the above aromatic compounds (Fig. 8).

4.3. Glucosinolate biosynthesis may be an exception to the theory of growth and defense balance

It is known that plant defense and growth are irreversibly correlated, i.e., plants utilize metabolites and energy during the defense/stress response and consequently compromise their growth [13,149]. Therefore, low levels of glucosinolates in the mutants might allow the metabolic intermediates/energy to be allocated to enhance plant growth. However, this is obviously not the case. Our growth measurement (Fig. 1, Supplementary Table 2), proteomics data (Tables 1, 2 and 3) and metabolomics data (Supplementary Table 6) support previous findings concerning the decreased growth of the *myb28/29* and *cyp79B2/B3* mutants [19,30,31]. Therefore, it appears that perturbation in glucosinolate metabolism may be an exception to the growth and defense balance theory. This may be attributed to the generation of stressful status in the mutant cells as aforementioned.

The metabolomics data shed more light on the observed plant growth. For example, indolic carboxylic acids derived from tryptophan were not detected in the *cyp79B2/B3* as a result of *CYP79B2/B3* mutation, while they showed higher levels in the *myb28/29* (Fig. 7, Supplementary Table 6). It is well-known that this group of auxin compounds is important in plant growth [150]. Several studies have explored the connection between glucosinolates and plant growth regulatory hormones [10-12,15,89]. Based on their findings and our results, the

increase of indolic carboxylic acids in the *myb28/29* may be a secondary effect of increased production of indolic glucosinolates [10] and concurrent degradation of indole acetic acids to produce indolic carboxylic acids [151]. The absence of indolic carboxylic acids in the *cyp79B2/B3* was clearly a result of blockage of the indole-3-acetaldoxime production [11,12]. In addition, the increased levels of some amino acids in the *cyp79B2/B3* may indirectly affect plant growth, as the accumulation of amino acids were shown to decrease nitrate uptake and as a consequence affect plant growth [152]. These findings might explain the growth phenotype of the *cyp79B2/B3*. Purine metabolic cycle is highly correlated to plant growth [141]. The metabolomic analysis revealed increased activity in the purine metabolism, as indicated by the elevated levels of nine purines in the *cyp79B2/B3* and three in the *myb28/29* mutant (Supplementary Table 6).

In this study changes in growth related proteins can be classified into two groups; the first group contains proteins showing decreased levels, and the second includes proteins showing elevated levels in the mutants. The first group included eukaryotic initiation factor 4B in protein synthesis [44] and DnaJ/Hsp40 cysteine-rich protein important in protein binding. Both were decreased in the two mutants (Table 1). In the *myb28/29*, pectinesterase/pectinesterase inhibitor 18, a cell wall modifier enzyme [86] was representative for the first type (Table 2). The *cyp79B2/B3* mutant contained several of growth and protein synthesis-related enzymes related to the first type, which are 40S ribosomal protein S6-2 involved in cytokinin regulation and consequently plant growth [118,119], eukaryotic translation initiation factor 3 subunit C with a role in protein synthesis [121], ankyrin repeat family protein involved in protein interactions [127], Ran-binding protein 1 homolog c involved in primary root growth [128], Ran-binding protein 1 homolog a with protein transport activity, PP2A regulatory subunit TAP46 involved in plant growth [137], and prolyl 4-hydroxylase 7 involved in cell wall formation (Table 3). The

second group is represented by L-isoaspartate O-methyltransferase 1 involved in L-isoaspartyl residue methyl esterification and damaged protein repair/degradation [35-37], nudC domain-containing protein necessary for Arabidopsis embryo apical domain partitioning [41] and translation initiation factor isoform 4E in protein synthesis [42,43]. These proteins all exhibited increased levels in both mutants (Table 1). In addition, the following *cyp79B2/B3* protein levels were high relative to the WT: protein disulfide isomerase 2-3 involved in protein folding through disulfide bond rearrangement, protein disulfide isomerase-like 1-1 involved in programmed cell death and disulfide bond rearrangement [104], chloroplastic chaperonin 60 subunit beta 3 involved in protein folding [108] and mitochondrial heat shock 70 kDa protein 10 preventing protein aggregation and involved in protein folding [112] (Table 3). The increases of these proteins provided another line of evidence for the intracellular stress imposed by the indole glucosinolate perturbation.

5. Conclusion

Proteomics and metabolomics have clearly shown utility in unraveling cellular molecular changes in response to perturbations of aliphatic glucosinolate biosynthesis (using *myb28/29* mutant) and indolic glucosinolate biosynthesis (using *cyp79B2/B3* mutant). In this study, 2785 proteins were identified, of which 142 proteins showed significant changes in the mutants compared to WT. In addition, 292 metabolites were identified and 89 exhibited different patterns between the mutants and WT. Using STRING, many new nodes and edges in the glucosinolate metabolic network could be derived. The processes of glucosinolate biosynthesis, degradation, cellular stress responses, and growth and defense related biological processes were highly enriched in the data set. Interestingly, we observed cross-talks between the syntheses of indolic

glucosinolates and aliphatic glucosinolates, as well as between glucosinolates and their hydrolyzing enzymes. It is also interesting to note that reduction of the specialized defense metabolites (glucosinolates) did not lead to growth enhancement. This result may be attributed to the cellular stress conditions revealed in the mutants. Improved understanding of glucosinolate molecular networks will facilitate engineering and breeding of glucosinolate production in crops for improving quality, yield, defense, and human health.

Acknowledgements

We would like to thank Dr. Danny Alexander, Metabolon Inc, Durham, North Carolina, USA for advice in metabolomics data analysis and interpretation, Dr. Jin Koh, Interdisciplinary Center for Biotechnology Research, University of Florida, Gainesville, Florida, USA for technical advice, Tong Zhang, Sisi Geng, Lihong Liu, Joe Collins and all the Chen laboratory members for helpful discussions and co-operation. Juan M. Gonzales is thanked for critical reading of the manuscript. This study was funded by the US National Science Foundation CAREER Award 0845162 to S.C., and a visiting scholarship from the Egyptian Government represented by the Egyptian Cultural and Educational Bureau at Washington DC, USA to I.M.

References

- 739 [1] J.W. Fahey, A.T. Zalcman, P. Talalay, The chemical diversity and distribution of
740 glucosinolates and isothiocyanates among plants, *Phytochemistry* 56 (2001) 5–51.
- 741 [2] C.D. Grubb, S. Abel, Glucosinolate metabolism and its control, *Trends Plant Sci.* 11
742 (2006) 89–100.
- 743 [3] T.A. Shapiro, J.W. Fahey, K.L. Wade, K.K. Stephenson, P. Talalay, Chemoprotective
744 glucosinolates and isothiocyanates of broccoli sprouts: Metabolism and excretion in
745 humans, *Cancer Epidemiol. Biomarkers Prev.* 10 (2001) 501–508.
- 746 [4] B.A. Halkier, J. Gershenzon, Biology and biochemistry of glucosinolates, *Annu. Rev.*
747 *Plant Biol.* 57 (2006) 303–333.
- 748 [5] X. Yan, S. Chen, Regulation of plant glucosinolate metabolism, *Planta* 226 (2007) 1343–
749 1352.
- 750 [6] I.E. Sønderby, F. Geu-Flores, B.A. Halkier, Biosynthesis of glucosinolates - gene
751 discovery and beyond, *Trends Plant Sci.* 15 (2010) 283–290.
- 752 [7] M.Y. Hirai, K. Sugiyama, Y. Sawada, T. Tohge, T. Obayashi, A. Suzuki, R. Araki, N.
753 Sakurai, H. Suzuki, K. Aoki, H. Goda, O.I. Nishizawa, D. Shibata, K. Saito, Omics-based
754 identification of *Arabidopsis* Myb transcription factors regulating aliphatic glucosinolate
755 biosynthesis, *Proc. Natl. Acad. Sci. U. S. A.* 104 (2007) 6478–6483.
- 756 [8] T. Gigolashvili, B. Berger, H.P. Mock, C. Müller, B. Weisshaar, U.I. Flügge, The
757 transcription factor HIG1/MYB51 regulates indolic glucosinolate biosynthesis in
758 *Arabidopsis thaliana*, *Plant J.* 50 (2007) 886–901.
- 759 [9] Y. Chen, X. Yan, S. Chen, Bioinformatic analysis of molecular network of glucosinolate
760 biosynthesis, *Comput. Biol. Chem.* 35 (2011) 10–18.

- 761 [10] B. Reintanz, M. Lehen, M. Reichelt, J. Gershenzon, M. Kowalczyk, G. Sandberg, M.
762 Godde, R. Uhl, K. Palme, *bus*, a Bushy Arabidopsis *CYP79F1* knockout mutant with
763 abolished synthesis of short-chain aliphatic glucosinolates, *Plant Cell* 13 (2001) 351–367.
- 764 [11] M.D. Mikkelsen, C.H. Hansen, U. Wittstock, B.A. Halkier, Cytochrome P450 CYP79B2
765 from Arabidopsis catalyzes the conversion of tryptophan to indole-3-acetaldoxime, a
766 precursor of indole glucosinolates and indole-3-acetic acid, *J. Biol. Chem.* 275 (2000)
767 33712–33717.
- 768 [12] M.D. Mikkelsen, P. Naur, B.A. Halkier, Arabidopsis mutants in the C-S lyase of
769 glucosinolate biosynthesis establish a critical role for indole-3-acetaldoxime in auxin
770 homeostasis, *Plant J.* 37 (2004) 770–777.
- 771 [13] H. Yan, M.-J. Yoo, J. Koh, L. Liu, Y. Chen, D. Acikgoz, Q. Wang, S. Chen, Molecular
772 reprogramming of Arabidopsis in response to perturbation of jasmonate signaling, *J.*
773 *Proteome Res.* 13 (2014) 5751–5766.
- 774 [14] M. D.-C. Martínez-Ballesta, D.A. Moreno, M. Carvajal, The physiological importance of
775 glucosinolates on plant response to abiotic stress in *Brassica*, *Int. J. Mol. Sci.* 14 (2013)
776 11607–11625.
- 777 [15] Y.-Z. Chen, Q.-Y. Pang, Y. He, N. Zhu, I. Branstrom, X.-F. Yan, S. Chen, Proteomics and
778 metabolomics of Arabidopsis responses to perturbation of glucosinolate biosynthesis, *Mol.*
779 *Plant* 5 (2012) 1138–1150.
- 780 [16] H. Frerigmann, C. Böttcher, D. Baatout, T. Gigolashvili, Glucosinolates are produced in
781 trichomes of *Arabidopsis thaliana*, *Front. Plant Sci.* 3 (2012) 1–13.
- 782 [17] S. Chen, E. Glawischnig, K. Jørgensen, P. Naur, B. Jørgensen, C.E. Olsen, C.H. Hansen, H.
783 Rasmussen, J.A. Pickett, B.A. Halkier, CYP79F1 and CYP79F2 have distinct functions in
784 the biosynthesis of aliphatic glucosinolates in Arabidopsis, *Plant J.* 33 (2003) 923–937.

785 [18] Y. Li, Y. Sawada, A. Hirai, M. Sato, A. Kuwahara, X. Yan, M.Y. Hirai, Novel insights
786 into the function of Arabidopsis R2R3-MYB transcription factors regulating aliphatic
787 glucosinolate biosynthesis, *Plant Cell Physiol.* 54 (2013) 1335–1344.

788 [19] Y. Zhao, A.K. Hull, N.R. Gupta, K.A. Goss, J. Alonso, J.R. Ecker, J. Normanly, J. Chory,
789 J.L. Celenza, Trp-dependent auxin biosynthesis in Arabidopsis: involvement of
790 cytochrome P450s *CYP79B2* and *CYP79B3*, *Genes Dev.* 16 (2002) 3100– 3112.

791 [20] D.J. Kliebenstein, H.C. Rowe, K.J. Denby, Secondary metabolites influence
792 Arabidopsis/Botrytis interactions: variation in host production and pathogen sensitivity,
793 *Plant J.* 44 (2005) 25–36.

794 [21] I. Kasajima, Y. Ide, N. Ohkama-Ohtsu, H. Hayashi, T. Yoneyama, T. Fujiwara, A
795 protocol for rapid DNA extraction from *Arabidopsis thaliana* for PCR analysis, *Plant Mol.*
796 *Biol. Report.* 22 (2004) 49–52.

797 [22] D. Kambiranda, R. Katam, S.M. Basha, S. Siebert, iTRAQ-based quantitative proteomics
798 of developing and ripening muscadine grape berry, *J. Proteome Res.* 13 (2014) 555–569.

799 [23] S. Weist, M. Eravci, O. Broedel, S. Fuxius, S. Eravci, A. Baumgartner, Results and
800 reliability of protein quantification for two-dimensional gel electrophoresis strongly
801 depend on the type of protein sample and the method employed, *Proteomics* 8 (2008)
802 3389–3396.

803 [24] M. Zhu, B. Simons, N. Zhu, D.G. Oppenheimer, S. Chen, Analysis of abscisic acid
804 responsive proteins in *Brassica napus* guard cells by multiplexed isobaric tagging, *J.*
805 *Proteomics* 73 (2010) 790–805.

806 [25] K.A. Jones, P.D. Kim, B.B. Patel, S.G.Kelsen, A. Braverman, D.J. Swinton, P.R. Gafken,
807 L.A. Jones, W.S. Lane, J.M. Neveu, H.C. Leung, S.A. Shaffer, J.D. Leszyk, B.A. Stanley,

- T.E. Fox, A. Stanley, M.J. Hall, H. Hample, C.D South, A. De la Chapelle, R.W. Burt, D.A. Jones, L. Kopelovich, A.T. Yeung, Immunodepletion plasma proteomics by tripleTOF 5600 and Orbitrap elite/LTQ-Orbitrap Velos/Q exactive mass spectrometers, *J Proteome Res* 12 (2013) 4351–4365.
- [26] J.A. Vizcaíno, R. Côté, F. Reisinger, J.M. Foster, M. Mueller, J. Rameseder, H. Hermjakob, L. Martens, A guide to the Proteomics Identifications Database proteomics data repository, *Proteomics* 9 (2009) 4276–4283.
- [27] J. Baldrianová, M. Černý, J. Novák, P. L. Jedelský, E. Divíšková, B. Brzobohatý, Arabidopsis proteome responses to the smoke-derived growth regulator karrikin, *J. proteomics* 120 (2015) 7–20.
- [28] C. Silva-Sanchez, S. Chen, N. Zhu, Q.B. Li, P.S. Chourey, Proteomic comparison of basal endosperm in maize miniature1 mutant and its wild-type Mn1, *Front. Plant Sci.* 4:211 (2013) 10.3389/fpls.2013.00211.
- [29] A.M. Evans, B.R. Bridgewater, Q. Liu, M.W. Mitchell, R.J. Robinson, H. Dai, S.J. Stewart, C.D. DeHaven, L.A.D. Miller, High resolution mass spectrometry improves data quantity and quality as compared to unit mass resolution mass spectrometry in high-throughput profiling metabolomics, *Metabolomics* 4 (2014) 132. doi:10.4172/2153-0769.1000132.
- [30] J. Beekwilder, W. Van Leeuwen, N.M. Van Dam, M. Bertossi, V. Grandi, L. Mizzi, M. Soloviev, L. Szabados, J.W. Molthoff, B. Schipper, H. Verbocht, R.C.H. de Vos, P. Morandini, M.G.M. Aarts, A. Bovy, The impact of the absence of aliphatic glucosinolates on insect herbivory in Arabidopsis, *PLoS One* 3 (2008) e2068.

- 830 [31] T. Züst, B. Joseph, K.K. Shimizu, D.J. Kliebenstein, L.A. Turnbull, Using knockout
831 mutants to reveal the growth costs of defensive traits, *Proc. R. Soc. B: Biological Sciences*
832 (2011) rspb20102475.
- 833 [32] R. Zavaliev, A. Levy, A. Gera, B.L. Epel, Subcellular dynamics and role of Arabidopsis
834 β -1,3-glucanases in cell-to-cell movement of tobamoviruses, *Mol. Plant. Microbe. Interact.*
835 26 (2013) 1016–1030.
- 836 [33] D. Lieberherr, U. Wagner, P.H. Dubuis, J.P. Métraux, F. Mauch, The rapid induction of
837 glutathione S-transferases *AtGSTF2* and *AtGSTF6* by avirulent *Pseudomonas syringae* is
838 the result of combined salicylic acid and ethylene signaling, *Plant Cell Physiol.* 44 (2003)
839 750–757.
- 840 [34] A.P. Smith, B.P. DeRidder, W.J. Guo, E.H. Seeley, F.E. Regnier, P.B. Goldsbrough,
841 Proteomic analysis of Arabidopsis glutathione S-transferases from benoxacor- and copper-
842 treated seedlings, *J. Biol. Chem.* 279 (2004) 26098–26104.
- 843 [35] Q. Xu, M.P. Belcastro, S.T. Villa, R.D. Dinkins, S.G. Clarke, A.B. Downie, A second
844 protein L-isoaspartyl methyltransferase gene in Arabidopsis produces two transcripts
845 whose products are sequestered in the nucleus, *Plant Physiol.* 136 (2004) 2652–2664.
- 846 [36] S.T. Villa, Q. Xu, A.B. Downie, S.G. Clarke, *Arabidopsis* protein repair L-isoaspartyl
847 methyltransferases: Predominant activities at lethal temperatures, *Physiol. Plant.* 128
848 (2006) 581–592.
- 849 [37] L. Ogé, G. Bourdais, J. Bove, B. Collet, B. Godin, F. Granier, J.-P. Boutin, D. Job, M.
850 Jullien, P. Grappin, Protein repair L-isoaspartyl methyltransferase 1 is involved in both
851 seed longevity and germination vigor in Arabidopsis, *Plant Cell* 20 (2008) 3022–3037.

- 852 [38] M.N. Ngaki, G.V. Louie, R.N. Philippe, G. Manning, F. Pojer, M.E. Bowman, L. Li, E.
853 Larsen, E.S. Wurtele, J.P. Noel, Evolution of the chalcone-isomerase fold from fatty-acid
854 binding to stereospecific catalysis, *Nature* 485 (2012) 530–533.
- 855 [39] D.P. Dixon, J.D. Sellars, R. Edwards, The Arabidopsis phi class glutathione transferase
856 AtGSTF2: binding and regulation by biologically active heterocyclic ligands, *Biochem. J.*
857 438 (2011) 63–70.
- 858 [40] U. Wagner, R. Edwards, D.P. Dixon, F. Mauch, Probing the diversity of the Arabidopsis
859 glutathione S-transferase gene family, *Plant Mol. Biol.* 49 (2002) 515–532.
- 860 [41] D.E. Perez, J.S. Hoyer, A.I. Johnson, Z.R. Moody, J. Lopez, N.J. Kaplinsky, BOBBER1 is
861 a noncanonical Arabidopsis small heat shock protein required for both development and
862 thermotolerance, *Plant Physiol.* 151 (2009) 241–252.
- 863 [42] S. Wittmann, H. Chatel, M.G. Fortin, J.F. Laliberté, Interaction of the viral protein
864 genome linked of turnip mosaic potyvirus with the translational eukaryotic initiation
865 factor (iso) 4E of *Arabidopsis thaliana* using the yeast two-hybrid system, *Virology* 234
866 (1997) 84–92.
- 867 [43] A.D. Lellis, K.D. Kasschau, S.A. Whitham, J.C. Carrington, Loss-of-susceptibility
868 mutants of *Arabidopsis thaliana* reveal an essential role for eIF(iso)4E during potyvirus
869 infection, *Curr. Biol.* 12 (2002) 1046–1051.
- 870 [44] A.M. Metz, K.C. Wong, S.A. Malmström, K.S. Browning, Eukaryotic initiation factor 4B
871 from wheat and *Arabidopsis thaliana* is a member of a multigene family, *Biochem.*
872 *Biophys. Res. Commun.* 266 (1999) 314–321.
- 873 [45] D.N. Richardson, M.F. Rogers, A. Labadorf, A. Ben-Hur, H. Guo, A.H. Paterson, A.S.N.
874 Reddy, Comparative analysis of serine/arginine-rich proteins across 27 eukaryotes:

875 Insights into sub-family classification and extent of alternative splicing, PLoS One 6
876 (2011) e24542.

877 [46] N.I. Park, D.G. Muench, Biochemical and cellular characterization of the plant ortholog of
878 PYM, a protein that interacts with the exon junction complex core proteins Mago and Y14,
879 Planta 225 (2007) 625–639.

880 [47] N. Sauer, K. Friedländer, U. Gräml-Wicke, Primary structure, genomic organization and
881 heterologous expression of a glucose transporter from *Arabidopsis thaliana*, EMBO J. 9
882 (1990) 3045–3050.

883 [48] B. Zhang, D. Tremousaygue, N. Denancé, H.P. van Esse, A.C. Hörger, P. Dabos, D.
884 Goffner, B.P.H.J. Thomma, R.A.L. van der Hoorn, H. and Tuominen, PIRIN2 stabilizes
885 cysteine protease XCP2 and increases susceptibility to the vascular pathogen *Ralstonia*
886 *solanacearum* in Arabidopsis, Plant J. 79 (2014) 1009–1019.

887 [49] M.S. Bae, E.J. Cho, E.Y. Choi, O.K. Park, Analysis of the Arabidopsis nuclear proteome
888 and its response to cold stress, Plant J. 36 (2003) 652–663.

889 [50] L.D. Noël, G. Cagna, J. Stuttmann, L. Wirthmüller, S. Betsuyaku, C.-P. Witte, R. Bhat, N.
890 Pochon, T. Colby, J.E. Parker, Interaction between SGT1 and cytosolic/nuclear HSC70
891 chaperones regulates Arabidopsis immune responses, Plant Cell 19 (2007) 4061–4076.

892 [51] A. Finka, R.U. Mattoo, P. Goloubinoff, Meta-analysis of heat-and chemically upregulated
893 chaperone genes in plant and human cells, Cell Stress Chaperone 16 (2011) 15–31.

894 [52] V. Shulaev, K. Schlauch, R. Mittler, Cytosolic ascorbate peroxidase 1 is a central
895 component of the reactive oxygen gene network of Arabidopsis, Plant Cell 17 (2005) 268–
896 281.

- 897 [53] W. Kammerloher, U. Fischer, G.P. Piechotka, A.R. Schäffner, Water channels in the
898 plant plasma membrane cloned by immunoselection from a mammalian expression system,
899 Plant J. 6 (1994) 187–199.
- 900 [54] R. Stracke, M. Werber, B. Weisshaar, The *R2R3-MYB* gene family in *Arabidopsis*
901 *thaliana*, Curr. Opin. Plant Biol. 4 (2001) 447–56.
- 902 [55] J. Li, B.G. Hansen, J.A. Ober, D.J. Kliebenstein, B.A. Halkier, Subclade of flavin-
903 monooxygenases involved in aliphatic glucosinolate biosynthesis, Plant Physiol. 148
904 (2008) 1721–1733.
- 905 [56] M. Burow, A. Losansky, R. Müller, A. Plock, D.J. Kliebenstein, U. Wittstock, The genetic
906 basis of constitutive and herbivore-induced ESP-independent nitrile formation in
907 *Arabidopsis*, Plant Physiol. 149 (2009) 561–574.
- 908 [57] R. Kissen, A.M. Bones, Nitrile-specifier proteins involved in glucosinolate hydrolysis in
909 *Arabidopsis thaliana*, J. Biol. Chem. 284 (2009) 12057–12070.
- 910 [58] M. Benderoth, S. Textor, A.J. Windsor, T. Mitchell-Olds, J. Gershenzon, J. Kroymann,
911 Positive selection driving diversification in plant secondary metabolism, Proc. Natl. Acad.
912 Sci. U. S. A. 103 (2006) 9118–9123.
- 913 [59] S. Textor, S. Bartram, J. Kroymann, K.L. Falk, A. Hick, J.A. Pickett, J. Gershenzon,
914 Biosynthesis of methionine-derived glucosinolates in *Arabidopsis thaliana*: Recombinant
915 expression and characterization of methylthioalkylmalate synthase, the condensing
916 enzyme of the chain-elongation cycle, Planta 218 (2004) 1026–1035.
- 917 [60] Y. Yang, R. Sulpice, A. Himmelbach, M. Meinhard, A. Christmann, E. Grill, Fibrillin
918 expression is regulated by abscisic acid response regulators and is involved in abscisic
919 acid-mediated photoprotection, Proc. Natl. Acad. Sci. U. S. A. 103 (2006) 6061–6066.

920 [61] A. Youssef, Y. Laizet, M.A. Block, E. Maréchal, J.P. Alcaraz, T.R. Larson, D. Pontier, J.
921 Gaffé, M. Kuntz, Plant lipid-associated fibrillin proteins condition jasmonate production
922 under photosynthetic stress, *Plant J.* 61 (2010) 436–445.

923 [62] J.C. Wang, Cellular roles of DNA topoisomerases: a molecular perspective, *Nat. Rev. Mol.*
924 *Cell Biol.* 3 (2002) 430–440.

925 [63] V. Prabhakar, T. Löttgert, T. Gigolashvili, K. Bell, U.I. Flügge, R.E. Häusler, Molecular
926 and functional characterization of the plastid-localized Phosphoenolpyruvate enolase
927 (ENO1) from *Arabidopsis thaliana*, *FEBS Lett.* 583 (2009) 983–991.

928 [64] T. Kiyosue, J.K. Beetham, F. Pinot, B.D. Hammock, K. Yamaguchi-Shinozaki, K.
929 Shinozaki, Characterization of an *Arabidopsis* cDNA for a soluble epoxide hydrolase gene
930 that is inducible by auxin and water stress, *Plant J.* 6 (1994) 259–269.

931 [65] L.M. Voll, M.B. Zell, T. Engelsdorf, A. Saur, M.G. Wheeler, M.F. Drincovich, A.P.M.
932 Weber, V.G. Maurino, Loss of cytosolic NADP-malic enzyme 2 in *Arabidopsis thaliana*
933 is associated with enhanced susceptibility to *Colletotrichum higginsianum*, *New Phytol.*
934 195 (2012) 189–202.

935 [66] M.C.G. Wheeler, F.M. Drincovich, C.S. Andreo, U. Flugge, M.A. Tronconi, V.G.
936 Maurino, A Comprehensive Analysis of the NADP-Malic Enzyme Gene Family of
937 *Arabidopsis*, *Plant Physiol.* 139 (2005) 39–51.

938 [67] G. Armijo, P. Salinas, M.I. Monteoliva, A. Seguel, C. García, E. Villarroel-Candia, W.
939 Song, A.R. van der Krol, M.E. Álvarez, L. Holuigue, A salicylic acid-induced lectin-like
940 protein plays a positive role in the effector-triggered immunity response of *Arabidopsis*
941 *thaliana* to *Pseudomonas syringae* Avr-Rpm1, *Mol. Plant. Microbe. Interact.* 26 (2013)
942 1395–406.

- 943 [68] S.P. Rius, P. Casati, A.A. Iglesias, D.F. Gomez-Casati, Characterization of Arabidopsis
944 lines deficient in GAPC-1, a cytosolic NAD-dependent glyceraldehyde-3-phosphate
945 dehydrogenase, *Plant Physiol.* 148 (2008) 1655–1667.
- 946 [69] M. Piotrowski, A. Schemenewitz, A. Lopukhina, A. Müller, T. Janowitz, E.W. Weiler, C.
947 Oecking, Desulfoglucosinolate sulfotransferases from *Arabidopsis thaliana* catalyze the
948 final step in the biosynthesis of the glucosinolate core structure, *J. Biol. Chem.* 279 (2004)
949 50717–50725.
- 950 [70] M. Klein, M. Reichelt, J. Gershenzon, J. Papenbrock, The three desulfoglucosinolate
951 sulfotransferase proteins in Arabidopsis have different substrate specificities and are
952 differentially expressed, *FEBS J.* 273 (2006) 122–136.
- 953 [71] M. Klein, J. Papenbrock, Kinetics and substrate specificities of desulfo-glucosinolate
954 sulfotransferases in *Arabidopsis thaliana*, *Physiol. Plant.* 135 (2009) 140–149.
- 955 [72] Y. Hatzfeld, A. Maruyama, A. Schmidt, M. Noji, K. Ishizawa, K. Saito, β -Cyanoalanine
956 synthase is a mitochondrial cysteine synthase-like protein in spinach and Arabidopsis,
957 *Plant Physiol.* 123 (2000) 1163–1171.
- 958 [73] R. Zrenner, H. Riegler, C.R. Marquard, P.R. Lange, C. Geserick, C.E. Bartosz, C.T. Chen,
959 R.D. Slocum, A functional analysis of the pyrimidine catabolic pathway in Arabidopsis,
960 *New Phytol.* 183 (2009) 117–132.
- 961 [74] C. Barth, G. Jander, Arabidopsis myrosinases TGG1 and TGG2 have redundant function
962 in glucosinolate breakdown and insect defense, *Plant J.* 46 (2006) 549–562.
- 963 [75] W.C.S. Tai, D.K. Banfield, AtBS14a and AtBS14b, two Bet1/Sft1-like SNAREs from
964 *Arabidopsis thaliana* that complement mutations in the yeast SFT1 gene, *FEBS Lett.* 500
965 (2001) 177–182.

- 966 [76] T. Ogawa, K. Yoshimura, H. Miyake, K. Ishikawa, D. Ito, N. Tanabe, S. Shigeoka,
967 Molecular characterization of organelle-type Nudix hydrolases in *Arabidopsis*, *Plant*
968 *Physiol.* 148 (2008) 1412–1424.
- 969 [77] F. Chen, L. Jiang, J. Zheng, R. Huang, Z. Hong, Y. Huang, H. Wang, Identification of
970 a co-segregative protein associated with the tillering trait in rice (*Oryza sativa* L.), *POJ.* 6
971 (2013) 36–45.
- 972 [78] R. Kumar, L.S.P. Tran, A.K. Neelakandan, H.T. Nguyen, Higher plant cytochrome b5
973 polypeptides modulate fatty acid desaturation, *PLoS One* 7 (2012) e31370.
- 974 [79] P. Reymond, N. Bodenhausen, R.M. Van Poecke, V. Krishnamurthy, M. Dicke, E.E.
975 Farmer, A conserved transcript pattern in response to a specialist and a generalist
976 herbivore, *Plant Cell* 16 (2004) 3132–3147.
- 977 [80] S. Baud, S. Wuillème, B. Dubreucq, A. de Almeida, C. Vuagnat, L. Lepiniec, M. Miquel,
978 C. Rochat, Function of plastidial pyruvate kinases in seeds of *Arabidopsis thaliana*, *Plant*
979 *J.* 52 (2007) 405–419.
- 980 [81] U. Bechtold, D.J. Murphy, P.M. Mullineaux, *Arabidopsis* peptide methionine sulfoxide
981 reductase2 prevents cellular oxidative damage in long nights, *Plant Cell* 16 (2004) 908–
982 919.
- 983 [82] Y. He, B. Chen, Q. Pang, J.M. Strul, S. Chen, Functional specification of *Arabidopsis*
984 isopropylmalate isomerases in glucosinolate and leucine biosynthesis, *Plant Cell Physiol.*
985 51 (2010) 1480–1487.
- 986 [83] S. Ravanel, M.A. Block, P. Rippert, S. Jabrin, G. Curien, F. Rébeillé, R. Douce,
987 Methionine metabolism in plants: Chloroplasts are autonomous for *de novo* methionine

988 synthesis and can import S-adenosylmethionine from the cytosol, J. Biol. Chem. 279
989 (2004) 22548–22557.

990 [84] I.-C. Chen, I.-C. Huang, M.-J. Liu, Z.-G. Wang, S.-S. Chung, H.-L. Hsieh, Glutathione S-
991 transferase interacting with far-red insensitive 219 is involved in phytochrome A-mediated
992 signaling in Arabidopsis, Plant Physiol. 143 (2007) 1189–1202.

993 [85] S.A. Saracco, M. Hansson, M. Scalf, J.M. Walker, L.M. Smith, R.D. Vierstra, Tandem
994 affinity purification and mass spectrometric analysis of ubiquitylated proteins in
995 Arabidopsis, Plant J. 59 (2009) 344–358.

996 [86] C. De-la-Peña, D.V. Badri, J.M. Vivanco, Novel role for pectin methylesterase in
997 Arabidopsis: A new function showing ribosome-inactivating protein (RIP) activity,
998 Biochim. Biophys. Acta - Gen. Subj. 1780 (2008) 773–783.

999 [87] C. Weis, U. Hildebrandt, T. Hoffmann, C. Hemetsberger, S. Pfeilmeier, C. König, W.
1000 Schwab, R. Eichmann, R. Hüchelhoven, CYP83A1 is required for metabolic compatibility
1001 of Arabidopsis with the adapted powdery mildew fungus *Erysiphe cruciferarum*, New
1002 Phytol. 202 (2014) 1310–1319.

1003 [88] J. Schuster, T. Knill, M. Reichelt, J. Gershenzon, S. Binder, Branched-chain
1004 aminotransferase4 is part of the chain elongation pathway in the biosynthesis of
1005 methionine-derived glucosinolates in Arabidopsis, Plant Cell 18 (2006) 2664–2679.

1006 [89] C.D. Grubb, B.J. Zipp, J. Ludwig-Müller, M.N. Masuno, T.F. Molinski, S. Abel,
1007 Arabidopsis glucosyltransferase UGT74B1 functions in glucosinolate biosynthesis and
1008 auxin homeostasis, Plant J. 40 (2004) 893–908.

1009 [90] Y. Kawamura, M. Uemura, Mass spectrometric approach for identifying putative plasma
1010 membrane proteins of Arabidopsis leaves associated with cold acclimation, Plant J. 36
1011 (2003) 141–154.

- 1012 [91] K. Marta, G. Marta, U. Adam, J. Hanna, The significance of Arabidopsis AAA proteases
1013 for activity and assembly/stability of mitochondrial OXPHOS complexes, *Physiol. Plant.*
1014 129 (2007) 135–142.
- 1015 [92] J.T. Trost, D.A. Chisholm, D.B. Jordan, B.A. Diner, The D1 C-terminal processing
1016 protease of photosystem II from *Scenedesmus obliquus*: Protein purification and gene
1017 characterization in wild type and processing mutants, *J. Biol. Chem.* 272 (1997) 20348–
1018 20356.
- 1019 [93] T. Ogawa, Y. Ueda, K. Yoshimura, S. Shigeoka, Comprehensive analysis of cytosolic
1020 nudix hydrolases in *Arabidopsis thaliana*, *J. Biol. Chem.* 280 (2005) 25277–25283.
- 1021 [94] A.B. Rose, R.L. Last, Introns act post-transcriptionally to increase expression of the
1022 *Arabidopsis thaliana* tryptophan pathway gene PAT1, *Plant J.* 11 (1997) 455–464.
- 1023 [95] A.B. Rose, J. Li, R.L. Last, An allelic series of blue fluorescent trp1 mutants of
1024 *Arabidopsis thaliana*, *Genetics* 145 (1997) 197–205.
- 1025 [96] R.M. Benstein, K. Ludewig, S. Wulfert, S. Wittek, T. Gigolashvili, H. Frerigmann, M.
1026 Gierth, U.-I Flügge, S. Krueger, Arabidopsis phosphoglycerate dehydrogenase1 of the
1027 phosphoserine pathway is essential for development and required for ammonium
1028 assimilation and tryptophan biosynthesis, *Plant Cell* 25 (2013) 1–20.
- 1029 [97] J.A. O'Brien, A. Daudi, P. Finch, V.S. Butt, J.P. Whitelegge, P. Souda, F.M. Ausubel, G.P.
1030 Bolwell, A peroxidase-dependent apoplastic oxidative burst in cultured arabidopsis cells
1031 functions in MAMP-elicited defense, *Plant Physiol.* 158 (2012) 2013–2027.
- 1032 [98] L. Østergaard, A.G. Pedersen, H.M. Jespersen, S. Brunak, K.G. Welinder, Computational
1033 analyses and annotations of the Arabidopsis peroxidase gene family, *FEBS Lett.* 433
1034 (1998) 98–102.

1035 [99] I. Barlier, M. Kowalczyk, A. Marchant, K. Ljung, R. Bhalerao, M. Bennett, G. Sandberg,
1036 C. Bellini, The *SUR2* gene of *Arabidopsis thaliana* encodes the cytochrome P450
1037 CYP83B1, a modulator of auxin homeostasis, Proc. Natl. Acad. Sci. U. S. A. 97 (2000)
1038 14819–14824.

1039 [100] C.H. Hansen, L. Du, P. Naur, C.E. Olsen, K.B. Axelsen, A.J. Hick, J.A. Pickett, B.A.
1040 Halkier, CYP83B1 Is the oxime-metabolizing enzyme in the glucosinolate pathway in
1041 Arabidopsis, J. Biol. Chem. 276 (2001) 24790–24796.

1042 [101] S. Bak, F.E. Tax, K.A. Feldmann, D.W. Galbraith, R. Feyereisen, CYP83B1, a
1043 cytochrome P450 at the metabolic branch point in auxin and indole glucosinolate
1044 biosynthesis in Arabidopsis, Plant Cell 13 (2001) 101–111.

1045 [102] S. Bak, R. Feyereisen, The involvement of two p450 enzymes, CYP83B1 and CYP83A1,
1046 in auxin homeostasis and glucosinolate biosynthesis, Plant Physiol. 127 (2001) 108–118.

1047 [103] G. Smolen, J. Bender, Arabidopsis cytochrome P450 *cyp83B1* mutations activate the
1048 tryptophan biosynthetic pathway, Genetics 160 (2002) 323–332.

1049 [104] C. Andème Ondzighi, D.A. Christopher, E.J. Cho, S.-C. Chang, L.A. Staehelin,
1050 Arabidopsis protein disulfide isomerase-5 inhibits cysteine proteases during trafficking to
1051 vacuoles before programmed cell death of the endothelium in developing seeds, Plant Cell
1052 20 (2008) 2205–2220.

1053 [105] A. Mustafiz, A.K. Singh, A. Pareek, S.K. Sopory, S.L. Singla-Pareek, Genome-wide
1054 analysis of rice and Arabidopsis identifies two glyoxalase genes that are highly expressed
1055 in abiotic stresses, Funct. Integr. Genomics 11 (2011) 293–305.

1056 [106] B. Youn, S.J. Kim, S.G.A. Moinuddin, C. Lee, D.L. Bedgar, A.R. Harper, L.B. Davin,
1057 N.G. Lewis, C. Kang, Mechanistic and structural studies of apoform, binary, and ternary

1058 complexes of the Arabidopsis alkenal double bond reductase At5g16970, J. Biol. Chem.
1059 281 (2006) 40076–40088.

1060 [107] Y. Qi, K. Tsuda, L.V. Nguyen, X. Wang, J. Lin, A.S. Murphy, J. Glazebrook, H. Thordal-
1061 Christensen, F. Katagiri, Physical association of Arabidopsis hypersensitive induced
1062 reaction proteins (HIRs) with the immune receptor RPS2, J. Biol. Chem. 286 (2011)
1063 31297–31307.

1064 [108] A.L. Bonshtien, A. Parnas, R. Sharkia, A. Niv, I. Mizrahi, A. Azem, C. Weiss,
1065 Differential effects of co-chaperonin homologs on cpn60 oligomers, Cell Stress
1066 Chaperones 14 (2009) 509–519.

1067 [109] J. Ouyang, X. Shao, J. Li, Indole-3-glycerol phosphate, a branchpoint of indole-3-acetic
1068 acid biosynthesis from the tryptophan biosynthetic pathway in *Arabidopsis thaliana*, Plant
1069 J. 24 (2000) 327–333.

1070 [110] Y. Olvera-Carrillo, J.L. Reyes, A.A. Covarrubias, Late embryogenesis abundant proteins
1071 Versatile players in the plant adaptation to water limiting environments, Plant Signal Behav.
1072 6 (2011) 586–589.

1073 [111] S. Ishihara, A. Takabayashi, K. Ido, T. Endo, K. Ifuku, F. Sato, Distinct functions for the
1074 two PsbP-like proteins PPL1 and PPL2 in the chloroplast thylakoid lumen of Arabidopsis,
1075 Plant Physiol. 145 (2007) 668–679.

1076 [112] P.-H. Su, H.-M. Li, Arabidopsis stromal 70-kD heat shock proteins are essential for plant
1077 development and important for thermotolerance of germinating seeds, Plant Physiol. 146
1078 (2008) 1231–1241.

1079 [113] D.B. Jordan, K.O. Bacot, T.J. Carlson, M. Kessel, P.V. Viitanen, Plant riboflavin
1080 biosynthesis. Cloning, chloroplast localization, expression, purification, and partial
1081 characterization of spinach lumazine synthase, J. Biol. Chem. 274 (1999) 22114–22121.

1082 [114] M. Tör, P. Gordon, A. Cuzick, T. Eulgem, E. Sinapidou, F. Mert-Türk, C. Can, J.L. Dangl,
1083 E.B. Holub, Arabidopsis SGT1b is required for defense signaling conferred by several
1084 downy mildew resistance genes, Plant Cell 14 (2002) 993–1003.

1085 [115] D. Yamazaki, K. Motohashi, T. Kasama, Y. Hara, T. Hisabori, Target proteins of the
1086 cytosolic thioredoxins in *Arabidopsis thaliana*, Plant Cell Physiol. 45 (2004) 18–27.

1087 [116] S. Van Nocker, Q. Deveraux, M. Rechsteiner, R.D. Vierstra, Arabidopsis *MBP1* gene
1088 encodes a conserved ubiquitin recognition component of the 26S proteasome, Proc. Natl.
1089 Acad. Sci. U. S. A. 93 (1996) 856–860.

1090 [117] Y.-L. Lin, H. Fu, In vivo relevance of substrate recognition function of major Arabidopsis
1091 ubiquitin receptors, Plant Signal. Behav. 7 (2012) 722–727.

1092 [118] M. Černý, A. Kuklová, W. Hoehenwarter, L. Fragner, O. Novák, G. Rotková, P.L.
1093 Jedelský, K. Žáková, M. Šmehilová, M. Strnad, W. Weckwerth, B. Brzobohatý, Proteome
1094 and metabolome profiling of cytokinin action in Arabidopsis identifying both distinct and
1095 similar responses to cytokinin down- and up-regulation, J. Exp. Bot. 64 (2013) 4193–4206.

1096 [119] I. Ruvinsky, O. Meyuhas, Ribosomal protein S6 phosphorylation: from protein synthesis
1097 to cell size, Trends Biochem. Sci. 31 (2006) 342–348.

1098 [120] A.J. Book, N.P. Gladman, S.S. Lee, M. Scalf, L.M. Smith, R.D. Vierstra, Affinity
1099 purification of the Arabidopsis 26 S proteasome reveals a diverse array of plant
1100 proteolytic complexes, J. Biol. Chem. 285 (2010) 25554–25569.

1101 [121] A. Yahalom, T.H. Kim, E. Winter, B. Karniol, A.G. von Arnim, D.A. Chamovitz,
1102 Arabidopsis eIF3e (INT-6) associates with both eIF3c and the COP9 signalosome subunit
1103 CSN7, J. Biol. Chem. 276 (2001) 334–340.

1104 [122] S. Kitamura, N. Shikazono, A. Tanaka, *TRANSPARENT TESTA 19* is involved in the
 1105 accumulation of both anthocyanins and proanthocyanidins in *Arabidopsis*, *Plant J.* 37
 1106 (2004) 104–114.

1107 [123] X. Li, P. Gao, D. Cui, L. Wu, I. Parkin, R. Saberianfar, R. Menassa, H. Pan, N. Westcott,
 1108 M.Y. Gruber, The *Arabidopsis tt19-4* mutant differentially accumulates proanthocyanidin
 1109 and anthocyanin through a 3' amino acid substitution in glutathione S-transferase, *Plant*,
 1110 *Cell Environ.* 34 (2011) 374–388.

1111 [124] M. Suorsa, S. Sirpiö, Y. Allahverdiyeva, V. Paakkarinen, F. Mamedov, S. Styring, E.M.
 1112 Aro, PsbR, a missing link in the assembly of the oxygen-evolving complex of plant
 1113 photosystem II, *J. Biol. Chem.* 281 (2006) 145–150.

1114 [125] E. Krebbers, J. Seurinck, L. Herdies, A.R. Cashmore, M.P. Timko, Four genes in two
 1115 diverged subfamilies encode the ribulose-1,5-bisphosphate carboxylase small subunit
 1116 polypeptides of *Arabidopsis thaliana*, *Plant Mol. Biol.* 11 (1988) 745–759.

1117 [126] J. Couturier, H.C. Wu, T. Dhalleine, H. Pégeot, D. Sudre, J.M. Gualberto, J.P. Jacquot, F.
 1118 Gaymard, F. Vignols, N. Rouhier, Monothiol glutaredoxin-BolA interactions: Redox
 1119 control of *Arabidopsis thaliana* BolA2 and SufE1, *Mol. Plant* 7 (2014) 187–205.

1120 [127] C. Becerra, T. Jahrmann, P. Puigdomènech, C.M. Vicent, Ankyrin repeat-containing
 1121 proteins in *Arabidopsis*: characterization of a novel and abundant group of genes coding
 1122 ankyrin-transmembrane proteins, *Gene* 340 (2004) 111–121.

1123 [128] S.H. Kim, D. Arnold, A. Lloyd, S.J. Roux, Antisense expression of an *Arabidopsis* ran
 1124 binding protein renders transgenic roots hypersensitive to auxin and alters auxin-induced
 1125 root growth and development by arresting mitotic progress, *Plant Cell* 13 (2001) 2619–
 1126 2630.

- 1127 [129] G.T. Hanke, S. Okutani, Y. Satomi, T. Takao, A. Suzuki, T. Hase, Multiple iso-proteins of
 1128 FNR in Arabidopsis: evidence for different contributions to chloroplast function and
 1129 nitrogen assimilation, *Plant, Cell Environ.* 28 (2005) 1146–1157.
- 1130 [130] F. Bourgis, S. Roje, M.L. Nuccio, D.B. Fisher, M.C. Tarczynski, C. Li, C. Herschbach, H.
 1131 Rennenberg, M.J. Pimenta, T.L. Shen, D.A. Gage, A.D. Hanson, S-methylmethionine
 1132 plays a major role in phloem sulfur transport and is synthesized by a novel type of
 1133 methyltransferase, *Plant Cell* 11 (1999) 1485–1498.
- 1134 [131] A. Tagmount, A. Berken, N. Terry, An essential role of S-adenosyl-L-methionine:L-
 1135 methionine S-methyltransferase in selenium volatilization by plants. Methylation of
 1136 selenomethionine to selenium-methyl-L-selenium-methionine, the precursor of volatile
 1137 selenium, *Plant Physiol.* 130 (2002) 847–856.
- 1138 [132] Z. Cheng, S. Sattler, H. Maeda, Y. Sakuragi, D.A. Bryant, D. DellaPenna, Highly
 1139 divergent methyltransferases catalyze a conserved reaction in tocopherol and
 1140 plastoquinone synthesis in cyanobacteria and photosynthetic eukaryotes, *Plant Cell* 15
 1141 (2003) 2343–2356.
- 1142 [133] N. Yokotani, M. Higuchi, Y. Kondou, T. Ichikawa, M. Iwabuchi, H. Hirochika, M. Matsui,
 1143 K. Oda, A novel chloroplast protein, CEST induces tolerance to multiple environmental
 1144 stresses and reduces photooxidative damage in transgenic Arabidopsis, *J. Exp. Bot.* 62
 1145 (2011) 557–569.
- 1146 [134] S.D.G. Marshall, J.J. Putterill, K.M. Plummer, R.D. Newcomb, The carboxylesterase gene
 1147 family from *Arabidopsis thaliana*, *J. Mol. Evol.* 57 (2003) 487–500.
- 1148 [135] K. Ishiyama, E. Inoue, A. Watanabe-Takahashi, M. Obara, T. Yamaya, H. Takahashi,
 1149 Kinetic properties and ammonium-dependent regulation of cytosolic isoenzymes of
 1150 glutamine synthetase in Arabidopsis, *J. Biol. Chem.* 279 (2004) 16598–16605.

1151 [136] Y. Mudgil, S.-H. Shiu, S.L. Stone, J.N. Salt, D.R. Goring, A large complement of the
 1152 predicted Arabidopsis ARM repeat proteins are members of the U-box E3 ubiquitin ligase
 1153 family, *Plant Physiol.* 134 (2004) 59–66.

1154 [137] C.S. Ahn, J.-A. Han, H.-S. Lee, S. Lee, H.-S. Pai, The PP2A regulatory subunit Tap46, a
 1155 component of the TOR signaling pathway, modulates growth and metabolism in plants,
 1156 *Plant Cell* 23 (2011) 185–209.

1157 [138] M. Hashimoto, T. Endo, G. Peltier, M. Tasaka, T. Shikanai, A nucleus-encoded factor,
 1158 CRR2, is essential for the expression of chloroplast *ndhB* in Arabidopsis, *Plant J.* 36 (2003)
 1159 541–549.

1160 [139] M.J. Oliver, L. Guo, D.C. Alexander, J.A. Ryals, B.W.M. Wone, J.C. Cushman, A sister
 1161 group contrast using untargeted global metabolomic analysis delineates the biochemical
 1162 regulation underlying desiccation tolerance in *Sporobolus stapfianus*, *Plant Cell* 23 (2011)
 1163 1231–1248.

1164 [140] V.K. Rai, Role of amino acids in plant responses to stresses, *Biol. Plantarum* 45 (2002)
 1165 481–487.

1166 [141] C. Stasolla, R. Katahira, T.A. Thorpe, H. Ashihara, Purine and pyrimidine nucleotide
 1167 metabolism in higher plants, *J. Plant Physiol.* 160 (2003) 1271–1295.

1168 [142] M. Martínez-Ballesta, D.A. Moreno-Fernández, D. Castejón, C. Ochando, P.A. Morandini,
 1169 M. Carvajal M, The impact of the absence of aliphatic glucosinolates on water transport
 1170 under salt stress in *Arabidopsis thaliana*, *Front Plant Sci.* 6 (2015) PMC4502342

1171 [143] K. Buxdorf, H. Yaffe, O. Barda, M. Levy, The effects of glucosinolates and their
 1172 breakdown products on necrotrophic fungi, *PLoS One.* 8 (2013) e70771.

1173 [144] J. Fan, C. Crooks, G. Creissen, L. Hill, S. Fairhurst, P. Doerner, C. Lamb, *Pseudomonas*
1174 *sax* genes overcome aliphatic isothiocyanate-mediated non-host resistance in *Arabidopsis*,
1175 *Science*. 331 (2011) 1185–1188.

1176 [145] J. Smith-Becker, E. Marois, E.J. Huguet, S.L. Midland, J.J. Sims, N.T. Keen,
1177 Accumulation of salicylic acid and 4-hydroxybenzoic acid in phloem fluids of cucumber
1178 during systemic acquired resistance is preceded by a transient increase in phenylalanine
1179 ammonia-lyase activity in petioles and stems, *Plant Physiol*. 116 (1998) 231–238.

1180 [146] C. Nawrath, J.P. Métraux, Salicylic acid induction-deficient mutants of *Arabidopsis*
1181 express *PR-2* and *PR-5* and accumulate high levels of camalexin after pathogen
1182 inoculation, *Plant Cell* 11 (1999) 1393–1404.

1183 [147] H. Vanacker, H. Lu, D.N. Rate, J.T. Greenberg, A role for salicylic acid and NPR1 in
1184 regulating cell growth in *Arabidopsis*, *Plant J*. 28 (2001) 209–216.

1185 [148] R. Merkl, I. Hrádková, V. Filip, J. Šmidrkal, Antimicrobial and antioxidant properties of
1186 phenolic acids alkyl esters, *Czech J Food Sci* 28 (2010) 275–279.

1187 [149] P.A.V. Zandt, Plant defense, growth, and habitat: a comparative assessment of
1188 constitutive and induced resistance, *Ecology* 88 (2007) 1984–1993.

1189 [150] S. Sugawara, S. Hishiyama, Y. Jikumaru, A. Hanada, T. Nishimura, T. Koshiba, Y. Zhao,
1190 Y. Kamiya, H. Kasahara, Biochemical analyses of indole-3-acetaldoxime-dependent auxin
1191 biosynthesis in *Arabidopsis*, *Proc. Natl. Acad. Sci. U. S. A.* 106 (2009) 5430–5435.

1192 [151] O. Sagee, J. Riov, R. Goren, Ethylene-enhanced catabolism of [¹⁴C] indole-3-acetic acid
1193 to indole-3-carboxylic acid in citrus leaf tissues, *Plant physiol*. 92 (1990) 54–60.

1194 [152] X. Liu, K. Lee, Effect of Mixed Amino Acids on Crop Growth, *Agricultural Science*
1195 (2012) 119–158. Dr. Godwin Aflakpui (Ed.), ISBN: 978-953-51-0567-1, InTech.

Figure captions

Fig. 1. Genotyping and phenotyping of *Arabidopsis thaliana myb28/29* and *cyp79B2/B3* mutants.

A. *MYB28/29* and *CYP79B2/B3* gene structure with T-DNA insertion sites. Primers used for genotyping are indicated and the sequences can be found in Table S1. B. 2% agarose gel electrophoretogram showing homozygosity of the mutants. C. Morphology of WT and the two mutants at five week-old rosette stage. The curling and slight chlorotic leaf tips of *cyp79B2/B3* are indicated by arrows.

Fig. 2. Venn diagrams showing the number of identified proteins and their significant changes. A.

The number of proteins identified in the two independent iTRAQ experiments using high peptide confidence level (1% FDR). B. The number of proteins showing significant increases in the two mutants relative to WT (fold change >1.2 with $p < 0.05$). C. The number of proteins showing significant decreases in the two mutants relative to WT (fold change <0.8 with $p < 0.05$).

Fig. 3. Mapping of the significantly changed proteins in *myb28/29* to known glucosinolate

biosynthetic pathway using STRING. Red balls indicate proteins known to be involved in the glucosinolate biosynthesis, grey balls indicate newly identified proteins, italic labelling highlights proteins showing changes in both mutants, and non-italic labelling indicates proteins showing unique changes in *myb28/29*. Green, brown and violet labels show proteins with primary (glucosinolate biosynthesis related), secondary (stress and defense related) and tertiary (other biological processes) connections, respectively, and aqua label indicates unknown function. The thickness of the edges reflects the strength of connection between the proteins based on neighborhood, gene fusion, co-occurrence, co-expression, previous experiments, and text-mining information at medium confidence score. Asterisk (*) indicates manual connection

derived from literature. Double asterisk (**) means nodes suggested to be involved in glucosinolate metabolism by TAIR10 and AgriGO analysis. The full name of the proteins can be found in the abbreviation and protein name columns in Tables 1 and 2.

Fig. 4. Mapping of the significantly changed proteins in *cyp79B2/B3* to known glucosinolate biosynthetic pathway using STRING in the confidence view. Red balls indicate proteins known to be involved in the glucosinolate biosynthesis, grey balls indicate newly identified proteins, Italic labelling highlights proteins showing changes in both mutants, and non-italic labelling indicates proteins showing unique changes in *cyp79B2/B3*. Green, brown and violet labels show proteins with primary (glucosinolate biosynthesis related), secondary (stress and defense related) and tertiary (other biological processes) connections, respectively, and aqua label indicates unknown function. The thickness of the edges reflects the strength of connection between the proteins based on neighborhood, gene fusion, co-occurrence, co-expression, previous experiments, and text-mining information at medium confidence score. Asterisk (*) indicates manual connection derived from literature. Double asterisk (**) means nodes suggested to be involved in glucosinolate metabolism by TAIR10 and AgriGO analysis. The full name of the proteins can be found in the abbreviation and protein name columns in Tables 1 and 3.

Fig. 5. Snapshots of the biological process-enriched GO terms using agriGO SEA for significantly changed proteins in *myb28/29* and *cyp79B2/B3*. The color scales represent the significance level of the functional enrichment from low (yellow) to high (red), except the dark gray denoting no enrichment.

Fig. 6. Diagrams showing the numbers of identified metabolites and their significant changes in the *myb28/29* and *cyp79B2/B3* mutants relative to WT. A. The number of identified metabolites by each analytical platform (RP, reverse phase, PP, polar phase, Pos, positive mode, Neg,

negative mode). B. Venn diagram showing the number of the metabolites with specific patterns of changes between the mutants. Up and down-arrows indicate significantly increased and decreased metabolites, respectively. The number in the shared region reflects metabolites detected in both mutants, while others are unique to either *myb89/29* or *cyp79B2/B3*. C. Venn diagrams showing metabolites in the two mutants with significant increases (top panel) or significant decreases (bottom panel) in levels ($p < 0.05$).

Fig. 7. Changes in the aliphatic and indolic glucosinolate biosynthetic pathways in *myb28/29* and *cyp79B2/B3*. Enzymes exhibiting significant changes in the proteomics study are shown in brown color. Asterisk (*) indicates metabolites analyzed by HPLC. Abbreviations: 1MTI3M, 1-Methoxy-3-indolylmethyl glucosinolate; 3 MSOP ITC, 3-methylsulfinylpropyl isothiocyanate; 3 MSOP, 3-methylsulfinylpropyl glucosinolate; 4 MSOB, 4-methylsulfinylbutyl glucosinolate; 4 MTB, 4-methylthiobutyl glucosinolate; 4MTI3M, 4-Methoxy-3-indolylmethyl glucosinolate; 4OHI3M, 4-hydroxy-indolyl-3-methyl glucosinolate; BCAT4, Branched-chain aminotransferase 4; CYP71A12, Cytochrome P450 71A12; CYP71A13, Indoleacetaldoxime dehydratase; CYP79B2, Cytochrome P450 79B2; CYP79B3, Cytochrome P450 79B3; CYP83A1, Cytochrome P450 83A1; CYP83B1(SUR2), Cytochrome P450 83B1; FMO3, Flavin-containing monooxygenase-3; GST, glutathione S-transferase; GSTF10, glutathione S-transferase F 10; GSTU20, Glutathione S-transferase U20; I3M, indol-3-ylmethyl glucosinolate; IPMD, 3-isopropylmalate dehydratase small subunit 1; MAM1, methylthioalkylmalate synthase 1; NSP1, Nitrile-specifier protein 1; SOT16, sulfotransferase 16; SOT17, sulfotransferase 17; SOT18, sulfotransferase 18; SUR1, S-alkyl-thiohydroximate lyase SUR1; TGG2, myrosinase 2. UGT74B1, UDP-glycosyltransferase 74B1.

1264 **Fig. 8.** Metabolite and protein changes in glycolysis and TCA cycle in *myb28/29* and
1265 *cyp79B2/B3*. Metabolites that were detected by metabolomics approaches are shown in blue
1266 color. Enzymes that exhibited significant changes in the proteomics study are shown in brown
1267 color. Abbreviations: ATMS2, 5-methyltetrahydropteroyltriglutamate-homocysteine
1268 methyltransferase 2; ENO1, enolase 1; GLN1-1, glutamine synthetase cytosolic isozyme 1-1;
1269 IGPS, indole-3-glycerol phosphate synthase; IPMD, 3-isopropylmalate dehydratase small
1270 subunit 1; MMT1, methionine S-methyltransferase; NADP-ME2, NADP-dependent malic
1271 enzyme 2; P4H7, prolyl 4-hydroxylase 7; PAT1, anthranilate phosphoribosyltransferase; PGM,
1272 phospho glycerate mutase family protein; PIMT1, L-isoaspartate O-methyltransferase 1; PMSR2,
1273 peptide methionine sulfoxide reductase A2; PSAT, phosphoserine aminotransferase; PyK,
1274 pyruvate kinase; RBCS-1B, ribulose biphosphate carboxylase; TSB1, tryptophan synthase beta
1275 chain 1.

1276 **Fig. 9.** Metabolite and protein changes in the glutathione γ -glutamyl amino acid cycle.
1277 Metabolites detected in the metabolomics study are shown in blue color. Enzymes that exhibited
1278 significant changes in the proteomics study are shown in brown color.

1279 **Table 1**
1280 List of shared proteins showing significant changes in *myb28/29* and *cyp79B2/B3* mutants relative to WT and their biological functions.

Accession number	Locus	Protein name	Abbreviation*	FC <i>myb</i> ^a	<i>p</i> -value <i>myb</i> ^b	FC <i>cyp</i> ^a	<i>p</i> -value <i>cyp</i> ^b	Function	Ref
P33157	At3g57260	Glucan endo-1,3-beta-glucosidase	BGL2	2.080	0.014	3.911	0.002	Response to pathogens	[32]
Q9SRY5	At1g02920	Glutathione S-transferase F7	GSTF7	1.351	0.001	1.389	0.000	Stress stimuli detoxification effect	[33,34]
Q42539	At3g48330	L-isoaspartate O-methyltransferase 1	PIMT1	1.341	0.000	1.288	0.012	Damaged protein repair or degradation	[35-37]
Q940G5	At4g25900	Aldose 1-epimerase	F14M19.180	1.227	0.010	1.209	0.003	Isomerization and binding of carbohydrate	UniProt
Q9M1X2	At3g63170	Fatty-acid-binding protein 1 (C18:3)	FAP1	1.222	0.001	1.269	0.011	Metabolism of fatty acid	[38]
P46422	At4g02520	Glutathione S-transferase F2	GSTF2	1.219	0.001	1.409	0.002	Conjugation of defense related bioactive compounds	[39,40]
Q8VXX3	At5g58740	Nudc domain-containing protein	AT5G58740	1.205	0.019	1.242	0.009	Arabidopsis embryo apical domain partitioning and patterning	[41]
O04663	At5g35620	Translation initiation factor isoform 4E	EIF4E	1.204	0.013	1.220	0.000	Protein synthesis	[42,43]
Q9LIN5	At3g26400	Eukaryotic initiation factor 4B	EIF4B1	0.784	0.022	0.783	0.031	Protein synthesis	[44]
Q9C9B9	At1g73980	Triphosphate tunnel metalloenzyme 1	F2P9.15	0.783	0.014	0.785	0.003	Correlated to kinase and adenylate cyclase activity	NCBI
Q9SJA6	At2g24590	Serine/arginine-rich splicing factor	RSZ22A	0.776	0.006	0.766	0.009	Identification of intron and formation of spliceosome	[45]
Q9LPZ4	At1g11400	Partner of Y14-mago	PYM	0.767	0.011	0.768	0.001	Controlling gene expression post transcription	[46]
P23586	At1g11260	Sugar transport protein 1	STP1	0.765	0.028	0.678	0.001	Monosaccharide transport	[47]
Q9FMH8	At5g43060	Cysteine protease component of protease-inhibitor complex	AT5G43060	0.742	0.008	0.729	0.019	Pathogen stress response	[48]
Q05212	At3g04880	DNA-damage-repair/toleration protein	DRT102	0.731	0.048	0.729	0.000	DNA binding and cold stress	[49]
P56801	ATCG00770	30S ribosomal protein S8	RPS8	0.716	0.005	0.689	0.003	30S ribosomal subunit formation	UniProt
P22954	At5g02490	Probable mediator of RNA polymerase II subunit 37c	MED37D	0.715	0.016	0.735	0.010	Controlling RNA polymerase II-dependent genes transcription and heat and pathogen response	UniProt, [50]
Q93VG5	At5g20290	40S ribosomal protein S8-1	RPS8A	0.712	0.008	0.782	0.006	Ribosome formation	UniProt
Q8VXZ7	At3g56310	Melibiose family protein	AT3G56310	0.712	0.001	0.786	0.000	Carbohydrate metabolism	UniProt
Q8L475	At1g74640	Alpha/beta-hydrolases	AT1G74640	0.712	0.010	0.703	0.010	Hydrolase action	UniProt
Q9LHL7	At3g12050	Aha1 domain-containing protein	T21B14.13	0.698	0.006	0.788	0.033	ATPase and chaperone binding, stress response	TAIR,[51]
Q05431	At1g07890	L-ascorbate peroxidase 1	APX1	0.693	0.033	0.764	0.004	Elimination of hydrogen peroxide	[52]
F4IAX0	At1g31690	Putative copper amine oxidase	AT1G31690	0.671	0.003	0.799	0.037	Amine oxidase activity, copper ion and quinone binding	UniProt
Q5EAH9	At3g55760	Unknown protein	AT3G55760	0.669	0.014	0.654	0.000		
O22769	At4g02580	NADH dehydrogenase [ubiquinone]	AT4G02580	0.643	0.000	0.636	0.000	Mitochondrial membrane respiratory chain complex	UniProt
Q8RWG1	At4g31390	aarf domain- protein kinase	AT4G31390	0.642	0.030	0.619	0.011	Protein kinase activity	UniProt, EMBL
Q9M214	At3g60450	Phosphoglycerate mutase	PGM	0.615	0.001	0.711	0.011	Glycolysis	UniProt
P43286	At3g53420	Aquaporin PIP2-1	PIP2-1	0.582	0.005	0.546	0.000	Water transport	[53]
Q9T0G9	At4g13480	Myb domain protein 79	AtMYB79	0.572	0.020	0.670	0.046	Binding of DNA	[54]

Q9SXE1	At1g62560	Flavin-containing monooxygenase-3	FMO3	0.511	0.005	0.521	0.008	Glucosinolate metabolism	[55]
Q9SDM9	At3g16400	Nitrile-specifier protein 1	NSP1	0.483	0.004	0.460	0.000	Glucosinolate hydrolysis and nitrile formation	[56,57]
Q93VA8	At1g76010	Alba DNA/RNA-binding protein	AT1G76010	0.409	0.004	0.412	0.001	Binding of nucleic acid	UniProt, TAIR
Q9FG67	At5g23010	Methylthioalkylmalate synthase 1	MAM1	0.350	0.000	0.587	0.008	Methionine side chain elongation in glucosinolate biosynthesis	[58,59]
Q9SN73	At3g47650	Dnaj/Hsp40 cysteine-rich protein	F1P2.200	0.348	0.000	0.330	0.002	Binding of protein	TAIR
Q6DBN2	At3g26080	Plastid-lipid-associated protein 5	PAP5	0.343	0.003	0.343	0.001	Light stress protection through jasmonate biosynthesis	[60,61]
O49393	At4g31980	Putative protein	F10N7.210	0.335	0.000	0.334	0.000		

1281

1282 Note: *Abbreviations for the shared proteins in Figs 3 and 4. ^a Fold change at cut-off point > 1.2 or < 0.8. ^b *p*-value < 0.05.

1283

1284

1285

1286 Table 2

1287 List of unique proteins in *myb28/29* showing significant level changes relative to WT and their biological functions

1288

Accession number	Locus tag	Protein name	Abbreviation*	FC ^a	<i>p</i> -value ^b	Function	Ref
Q9FJ79	At5g55310	DNA topoisomerase 1 beta	TOP1BETA	2.625	0.006	DNA topological transformation catalysis	[62]
F4IZC0	At3g50190	Uncharacterized protein	AT3G50190	1.796	0.021		
Q9ZUX4	At2g27730	Mitochondrial uncharacterized protein	AT2G27730	1.375	0.003	Copper binding and photorespiration	UniProt, EMBL
Q8LDV3	At4g13200	Uncharacterized protein	AT4G13200	1.306	0.011		
Q9C9C4	At1g74030	Chloroplastic enolase 1	ENO1	1.289	0.015	Conversion of glucose 6-phosphate to phosphoenolpyruvate	[13,15,63]
P25855	At2g35370	Glycine cleavage system H protein 1	GDH1	1.263	0.028	Glycine degradation	UniProt
Q42566	At2g26740	Epoxide hydrolase	AtSEH	1.246	0.042	Convert epoxides to diols	[64]
Q9LYG3	At5g11670	NADP-dependent malic enzyme 2	NADP-ME2	1.231	0.019	Oxidative decarboxylation of L-malate into pyruvate	[65,66]
Q9LK72	At3g16530	Lectin-like protein	AT3G16530	1.210	0.004	Immunological response against <i>P. syringae</i>	[67]
Q9M328	At3g53990	Adenine nucleotide alpha hydrolase-like	F5K20_290	1.207	0.026	Cold response and hydrolase action	UniProt, EMBL
P25858	At3g04120	Glyceraldehyde-3-phosphate dehydrogenase	GAPC1	1.205	0.001	Glycolysis process	[68]
Q9FZ80	At1g18590	Cytosolic sulfotransferase 17	SOT17	0.792	0.040	Glucosinolate biosynthesis	[69-71]

Q84MD7	At1g72880	Survival protein phosphatase/nucleotidase	sure-like	AT1G72880	0.783	0.022	Nucleotidase and hydrolase action	UniProt, EMBL
Q9LHA6	At3g28220	TRAF-like family protein		AT3G28220	0.779	0.023	Protease action and response to salt stress	UniProt, EMBL
Q9S6Z7	At3g04940	L-3-cyanoalanine synthase/cysteine synthase D1		CYSD1	0.767	0.026	Synthesis of cysteine	[72]
Q9LVI9	At3g17810	Putative dehydrogenase		PYD1A	0.763	0.009	Metabolism of pyrimidine	[73]
Q9C5C2	At5g25980	Myrosinase 2		TGG2	0.761	0.007	Glucosinolate hydrolysis	[74]
Q93WN0	At4g14040	Selenium-binding protein 2		SBP2	0.758	0.000	Binding of selenium	UniProt
Q9M2J9	At3g58170	Bet1-like SNARE 1-1		BET11	0.756	0.043	Intra-cellular membrane fusion and vesicular transport	[75]
Q8L831	At1g79690	Nudix hydrolase 3		NUDT3	0.749	0.004	Nucleotide sugar hydrolysis	[76]
Q9C9C9	At1g74090	Cytosolic sulfotransferase 18		SOT18	0.732	0.000	Glucosinolate biosynthesis	[69-71]
Q501F7	At1g21440	Plastid localized mutase		AT1G21440	0.732	0.007	C-P bond formation	[77]
Q9ZNV4	At2g46650	Cytochrome B5 isoform C		CYTB5-C	0.685	0.003	Membrane oxygenase electron carrier, stress response	[78,79]
Q9FNN1	At5g08570	Pyruvate kinase		PyK	0.680	0.033	Pyruvate and ATP irreversible synthesis	[80]
Q9LY15	At5g07460	Peptide methionine sulfoxide reductase A2		PMSR2	0.668	0.004	Protecting cells from oxidative stress by reducing methionine sulfoxide to methionine	[81]
Q9ZW84	At2g43100	3-isopropylmalate dehydratase small subunit 1		AtLeuC	0.647	0.011	Methionine chain elongation of aliphatic glucosinolate biosynthesis	[6,82]
Q9SRV5	At3g03780	5-methyltetrahydropteroyltriglutamate-homocysteine methyltransferase 2		ATMS2	0.633	0.000	Metabolism of methionine	[83]
Q8L7C9	At1g78370	Glutathione S-transferase U20		GSTU20	0.616	0.000	Glucosinolate formation and controls far red light on growth	[9,13,84]
P49692	At2g47610	60S ribosomal protein l7a-1		RPL7A	0.606	0.006	Ribosome formation and translation	[85]
Q1JPL7	At1g11580	Pectinesterase/pectinesterase inhibitor 18		PME18	0.589	0.023	Cell wall modifier	[86]
Q8LDQ8	At5g24165	Unknown protein		AT5G24165	0.579	0.027		
Q9LDA7	At3g15260	Probable protein phosphatase 2C 39		AT3G15260	0.556	0.016	Binding of metal ion	UniProt
P48421	At4g13770	Cytochrome P450 83A1		CYP83A1	0.483	0.004	Glucosinolate biosynthesis	[87]
Q9LZ83	At5g04420	Galactose oxidase/kelch repeat superfamily		T32M21_20	0.435	0.007	Galactose oxidation	UniProt
Q9LE06	At3g19710	Methionine aminotransferase BCAT4		BCAT4	0.406	0.000	Glucosinolate biosynthesis	[88]

1289 Note: *Abbreviations for unique proteins in Fig. 3. ^a Fold change at cut-off point > 1.2 or < 0.8. ^b *p*-value < 0.05.

1290

1291 **Table 3**

1292 List of unique proteins in *cyp79B2/B3* showing significant level changes relative to WT and their biological functions

1293

Accession number	Locus tag	Protein name	Abbreviation*	FC ^a	p-value ^b	Function	Ref
Q9SJ03	At2g21960	Unknown protein	AT2G21960	2.096	0.001		
Q9C9D0	At1g74100	Cytosolic sulfotransferase 16	SOT16	1.654	0.009	Glucosinolate biosynthesis	[69-71]
O48676	At1g24100	UDP-glycosyltransferase 74B1	UGT74B1	1.503	0.033	Glucosinolate formation	[89]
Q9SSB8	At1g80230	Mitochondrial cytochrome c oxidase subunit 5b-2	COX5B-2	1.486	0.000	Transport of mitochondrial electrons	UniProt
Q945K9	At4g34150	CaLB domain family protein	AT4G34150	1.461	0.007	Cold response	[90]
Q9LZF5	At5g03350	Lectin-like protein	AT5G03350	1.416	0.001	Binding of carbohydrate and salicylic acid stimulus response	[67]
Q0WVF5	At2g16440	DNA replication licensing factor MCM4	MCM4	1.411	0.037	DNA helicase	UniProt
Q9LVT8	At5g47210	Hyaluronan/mRNA binding family protein	AT5G47210	1.393	0.011	Binding of RNA	TAIR
O80983	At2g26140	Mitochondrial ATP-dependent zinc metalloprotease	FTSH4	1.379	0.034	Mitochondrial oxidative phosphorylation	[91]
F4KHG6	At5g46390	Carboxyl-terminal-processing peptidase 1	CTPA1	1.377	0.020	Proteolysis of the chloroplastic D1 protein of photosystem II	[92]
Q9SU14	At4g12720	Nudix hydrolase 7	NUDT7	1.361	0.003	Nucleoside diphosphates hydrolysis	[93]
P14671	At5g54810	Chloroplastic tryptophan synthase beta chain 1	TSB1	1.357	0.000	Tryptophan biosynthesis	[9,79], UniProt
Q02166	At5g17990	Chloroplastic anthranilate phosphoribosyltransferase	PAT1	1.356	0.000	Biosynthesis of tryptophan	[94,95]
O48773	At2g32920	Protein disulfide-isomerase 2-3	PDIL2-3	1.341	0.028	Folding of protein	UniProt
Q96255	At4g35630	Chloroplastic phosphoserine aminotransferase	PSAT	1.312	0.024	Biosynthesis of serine	[96]
Q9SMU8	At3g49120	Peroxidase 34	PRXCB	1.306	0.002	Stress conditions (toxics oxidation and H ₂ O ₂ removal)	[97,98]
O65782	At4g31500	Cytochrome P450 83B1	SUR2	1.305	0.000	Auxin, tryptophan and indolic glucosinolate biosynthesis	[99-103]
Q9XI01	At1g21750	Protein disulfide isomerase-like 1-1	PDIL1-1	1.298	0.003	Programmed cell death and disulfide bond rearrangement	[104]
Q9LTR8	At5g57040	Lactoylglutathione lyase / glyoxalase I family	AT5G57040	1.271	0.002	Detoxification of methylglyoxal	[105]
Q39172	At5g16970	NADP-dependent alkenal double bond reductase P1	AT-AER	1.269	0.009	Phenylpropanal substrate 7-8 double bond reduction and activated during anti-oxidant defense	[106]
Q9CAR7	At1g69840	Hypersensitive-induced response protein 2	HIR2	1.268	0.032	Potentiate defense against diseases	[107]
C0Z361	At5g56500	Chloroplastic chaperonin 60 subunit beta 3	CPN60B3	1.267	0.018	Folding of protein	[108]
Q8L9M8	At1g33810	Putative uncharacterized protein	AT1G33810	1.264	0.003		
P49572	At2g04400	Chloroplastic indole-3-glycerol phosphate synthase	IGPS	1.254	0.000	Tryptophan biosynthesis	[109]
Q8S8R1	At2g23120	Late embryogenesis abundant protein	AT2G23120	1.252	0.024	Seed desiccation and plant adaptation to water deficit	[110]
P82538	At3g55330	Chloroplastic psbp-like protein 1	PPL1	1.251	0.004	Photodamage fixation	[111]

Q9SFV3	At3g07230	Wound-responsive protein-related	T1B9.10	1.238	0.012	Response to wounds	TAIR
Q501E6	At5g25250	Flotillin-like protein 1	FLOT1	1.237	0.031	Caveolae formation	UniProt
Q9LDZ0	At5g09590	Mitochondrial heat shock 70 kDa protein 10	HSP70-10	1.232	0.001	Preventing protein aggregation and participates in folding	[112]
Q9T0K7	At4g13360	3-hydroxyisobutyryl-coa hydrolase-like protein 3	AT4G13360	1.232	0.038	Hydrolase action	UniProt
O80575	At2g44050	6,7-dimethyl-8-ribityllumazine synthase	COS1	1.232	0.001	Biosynthesis of riboflavin	[113]
Q9SIV0	At2g20610	S-alkyl-thiohydroximate lyase SUR1	SUR1	1.232	0.000	Biosynthesis of glucosinolates	[12]
Q9SUT5	At4g11260	Protein SGT1 homolog B	SGT1B	1.228	0.004	Immunity activation	[114]
Q9SL48	At2g17980	SEC1 family transport protein SLY1	SLY1	1.228	0.028	Exocytosis process	UniProt
Q39241	At1g45145	Thioredoxin H5	TRX5	1.226	0.030	Stress response	[115]
P56762	ATCG00740	DNA-directed RNA polymerase subunit alpha	RPOA	1.217	0.006	DNA transcription and binding	UniProt
Q9SYX1	At4g17600	Light-harvesting-like protein	LIL3.1	1.203	0.005	Binding and transcription of DNA	UniProt
Q9FMT8	At5g14120	Major facilitator protein (nodulin like protein)	AT5G14120	0.799	0.010		UniProt
P55034	At4g38630	26S proteasome non-ATPase regulatory subunit 4	RPN10	0.796	0.030	Ubiquitinated substrates turnover	[116,117]
P51430	At5g10360	40S ribosomal protein S6-2	RPS6B	0.796	0.005	Regulating cytokinin, cell and plant growth	[118,119]
Q9SZ83	At4g09670	Uncharacterized oxidoreductase	AT4G09670	0.796	0.037	Oxidoreductase action	UniProt
Q9MAQ0	At1g32900	Probable granule-bound starch synthase 1	WAXY	0.794	0.001	Endospermal amylose synthesis	UniProt
Q9C535	At1g29700	Metallo-beta-lactamase domain-containing protein	F15D2.25	0.793	0.006		
O48844	At2g32730	26S proteasome non-ATPase subunit 1 homolog A	RPN2A	0.790	0.037	Controlling the 26S proteasome subunit in protein degradation	[120]
Q93Z91	At5g42660	At5g42660/mjb21_3	AT5G42660	0.790	0.031	Glycosyltransfer activity	UniProt
O49160	At3g56150	Eukaryotic translation initiation factor 3 subunit C	EIF3C1	0.789	0.011	Protein synthesis	[121]
Q9FE46	At5g17220	Glutathione S-transferase F12	GSTF12	0.785	0.044	Involved in vacuole transport	[122,123]
Q9ZVA2	At1g78830	Curculin-like (mannose-binding) lectin family protein	F9K20.12	0.783	0.041	Binding of carbohydrates	UniProt
P27202	At1g79040	Chloroplastic photosystem II 10 kda polypeptide	PSBR	0.777	0.027	Oxygen release	[124]
Q8L493	At5g57850	Branched-chain amino acid aminotransferase-like 3	AT5G57850	0.770	0.000	Amino group transfer between amino acids	UniProt
P10796	At5g38430	Ribulose biphosphate carboxylase small chain 1B	RBCS-1B	0.765	0.044	Fixation of carbon dioxide	[125]
Q682I1	At1g55805	BolA like protein	AT1G55805	0.760	0.047	Controlling the redox reactions	[126]
Q8GYH5	At5g54610	Ankyrin repeat family protein	ANK	0.758	0.046	Interactions between proteins	[127]
Q9T075	At4g37880	LisH/CRA/RING-U-box domains-containing protein	AT4G37880	0.753	0.018	Binding of zinc	UniProt
P92985	At5g58590	Ran-binding protein 1 homolog c	RANBP1C	0.749	0.008	Primary root growth	[128]
Q9LMK7	At1g07140	Ran-binding protein 1 homolog a	RANBP1A	0.748	0.021	Protein transport activity	UniProt
Q9FKW6	At5g66190	Chloroplastic ferredoxin-NADP reductase	FNR1	0.746	0.004	Oxidation reduction reactions	[129]

Q9LTB2	At5g49810	Methionine S-methyltransferase	MMT1	0.743	0.003	S-methylmethionine /selenium methylmethionine formation	[130,131]
Q9SW43	At4g34920	PLC-like phosphodiesterase	T11I11.160	0.725	0.048	Hydrolase action of phosphoric diester	UniProt
Q9LY74	At3g63410	2-methyl-6-phytyl-1,4-hydroquinone methyltransferase	APG1	0.724	0.007	Biosynthesis of tocoopherol and plastoquinone	[132]
Q9LU01	At5g44650	Nucleus-encoded thylakoid protein	CEST	0.713	0.029	Stress withstanding	[133]
Q93VC9	At1g02305	Cysteine proteinase superfamily protein	AT1G02305	0.678	0.029	Endopeptidase action	UniProt
Q9C8I1	At1g51730	Ubiquitin-conjugating enzyme family protein	F19C24.6	0.663	0.007	Ubiquitin conjugation	TAIR
Q9FX93	At1g49650	Probable carboxylesterase 4	CXE4	0.655	0.024	Fatty acid ester hydrolysis	[134]
Q9LT92	At5g51880	Iron ion binding / oxidoreductase	AT5G51880	0.635	0.000	Binding of iron and oxidoreductase action	UniProt
Q56WN1	At5g37600	Glutamine synthetase cytosolic isozyme 1-1	GLN1-1	0.615	0.004	Glutamine formation	[135]
Q9LEQ9	At5g14790	ARM repeat superfamily protein	T9L3_90	0.567	0.009	Ubiquitin ligation	[136]
F4I3P9	At1g63680	UDP-N-acetylmuamoylanyl-d-glutamate-2,6-diaminopimelate ligase	MURE	0.563	0.005	Amino acid ligation and binding of ATP	UniProt
Q8LDQ4	At5g53000	PP2A regulatory subunit TAP46	TAP46	0.484	0.013	Control of plant growth	[137]
Q8L970	At3g28480	Probable proyl 4-hydroxylase 7	P4H7	0.444	0.003	Cell wall formation	UniProt
Q9STF3	At3g46790	Chloroplastic pentatricopeptide repeat-containing protein	CRR2	0.411	0.013	Modulating RNA between rps7 and ndhb transcripts (a pre-step in ndhb translation)	[138]

1294 Note: *Abbreviations for unique proteins in Fig. 4. ^a Fold change at cut-off point > 1.2 or < 0.8. ^b *p*-value < 0.05.

1295

Figure 1

Figure 2

Figure 3

Figure 4

GO Term	Description	myb28/29	cyp79B2/B3	GO Term	Description	myb28/29	cyp79B2/B3
GO:0019760	glucosinolate metabolic process	12	9	GO:0006952	defense response	16	26
GO:0016137	glycoside metabolic process	13	10	GO:0019318	hexose metabolic process	10	12
GO:0019748	secondary metabolic process	22	21	GO:0046686	response to cadmium ion	9	10
GO:0009683	indoleacetic acid metabolic process	9	9	GO:0006007	glucose catabolic process	9	10
GO:0005975	carbohydrate metabolic process	26	26	GO:0044237	cellular metabolic process	48	71
GO:0006790	sulfur metabolic process	16	13	GO:0046164	alcohol catabolic process	9	10
GO:0044272	sulfur compound biosynthetic process	14	11	GO:0006006	glucose metabolic process	9	11
GO:0019761	glucosinolate biosynthetic process	10	8	GO:0008152	metabolic process	52	83
GO:0009851	auxin biosynthetic process	9	9	GO:0006807	nitrogen compound metabolic process	30	43
GO:0034641	cellular nitrogen compound metabolic process	20	23	GO:0044238	primary metabolic process	47	65
GO:0019344	cysteine biosynthetic process	10	6	GO:0010038	response to metal ion	9	11
GO:0044262	cellular carbohydrate metabolic process	22	23	GO:0034976	response to endoplasmic reticulum stress	7	15
GO:0042434	indole derivative metabolic process	9	9	GO:0005996	monosaccharide metabolic process	10	14
GO:0006534	cysteine metabolic process	10	6	GO:0046483	heterocycle metabolic process	11	15
GO:0016138	glycoside biosynthetic process	10	8	GO:0006970	response to osmotic stress	10	14
GO:0009070	serine family amino acid biosynthetic process	10	7	GO:0009814	defense response, incompatible interaction	8	18
GO:0000097	sulfur amino acid biosynthetic process	11	6	GO:0042221	response to chemical stimulus	23	38
GO:0009607	response to biotic stimulus	21	33	GO:0006091	generation of precursor metabolites and energy	9	14
GO:0034754	cellular hormone metabolic process	9	9	GO:0009617	response to bacterium	8	16
GO:0044271	cellular nitrogen compound biosynthetic process	15	17	GO:0009266	response to temperature stimulus	10	17
GO:0009069	serine family amino acid metabolic process	10	8	GO:0006098	pentose-phosphate shunt	5	6
GO:0006950	response to stress	31	44	GO:0006955	immune response	10	22
GO:0032787	monocarboxylic acid metabolic process	19	27	GO:0006739	NADP metabolic process	5	6
GO:0044283	small molecule biosynthetic process	21	28	GO:0009310	amine catabolic process	5	7
GO:0006725	cellular aromatic compound metabolic process	16	19	GO:0019362	pyridine nucleotide metabolic process	5	6
GO:0044248	cellular catabolic process	22	28	GO:0009987	cellular process	51	79
GO:0043436	oxoacid metabolic process	22	31	GO:0043603	cellular amide metabolic process	5	6
GO:0042180	cellular ketone metabolic process	22	31	GO:0009820	alkaloid metabolic process	5	6
GO:0009056	catabolic process	22	29	GO:0033554	cellular response to stress	12	24
GO:0051707	response to other organism	18	28	GO:0009058	biosynthetic process	34	47
GO:0044275	cellular carbohydrate catabolic process	12	12	GO:0010035	response to inorganic substance	10	14
GO:0016053	organic acid biosynthetic process	16	20	GO:0044249	cellular biosynthetic process	33	45
GO:0016052	carbohydrate catabolic process	12	12	GO:0042742	defense response to bacterium	6	14
GO:0006519	cellular amino acid and derivative metabolic process	17	19	GO:0045087	innate immune response	9	21
GO:0000096	sulfur amino acid metabolic process	11	7	GO:0006066	alcohol metabolic process	10	16
GO:0050896	response to stimulus	37	57	GO:0046700	heterocycle catabolic process	5	7
GO:0009308	amine metabolic process	15	17	GO:0006733	oxidoreduction coenzyme metabolic process	5	7
GO:0042446	hormone biosynthetic process	9	9	GO:0065008	regulation of biological quality	12	16
GO:0044281	small molecule metabolic process	29	45	GO:0009651	response to salt stress	8	13
GO:0008652	cellular amino acid biosynthetic process	11	11	GO:0009627	systemic acquired resistance	6	16
GO:0044282	small molecule catabolic process	14	16	GO:0051716	cellular response to stimulus	14	30
GO:0034637	cellular carbohydrate biosynthetic process	14	14	GO:0006575	cellular amino acid derivative metabolic process	7	12
GO:0009309	amine biosynthetic process	11	11	GO:0010033	response to organic substance	15	24
GO:0044106	cellular amine metabolic process	14	15	GO:0070271	protein complex biogenesis	6	12
GO:0042445	hormone metabolic process	9	9	GO:0016054	organic acid catabolic process	5	8
GO:0006520	cellular amino acid metabolic process	13	14	GO:0006979	response to oxidative stress	6	10
GO:0016051	carbohydrate biosynthetic process	14	14	GO:0009409	response to cold	6	9
GO:0051704	multi-organism process	18	28	GO:0006568	tryptophan metabolic process		7
GO:0009628	response to abiotic stimulus	21	34	GO:0019438	aromatic compound biosynthetic process		13
GO:0044265	cellular macromolecule catabolic process	14	15	GO:0071395	cellular response to jasmonic acid stimulus		7
GO:0010817	regulation of hormone levels	9	9	GO:0009697	salicylic acid biosynthetic process		6
GO:0009057	macromolecule catabolic process	14	16	GO:0000165	MAPKKK cascade		6

Figure 5

Figure 6

Aliphatic glucosinolate biosynthesis

Indolic glucosinolate biosynthesis

Figure 7

Figure 8

Figure 9