Genome-wide transcriptome analyses of developing seeds from low and normal phytic acid soybean lines

Neelam R. Redekar, Ruslan M. Biyashev, Roderick V. Jensen, Richard F. Helm, Elizabeth A. Grabau and M. A. Saghai Maroof

Abstract

Background: Low phytic acid (lpa) crops are potentially eco-friendly alternative to conventional normal phytic acid (PA) crops, improving mineral bioavailability in monogastric animals as well as decreasing phosphate pollution. The lpa crops developed to date carry mutations that are directly or indirectly associated with PA biosynthesis and accumulation during seed development. These lpa crops typically exhibit altered carbohydrate profiles, increased free phosphate, and lower seedling emergence, the latter of which reduces overall crop yield, hence limiting their large-scale cultivation. Improving lpa crop yield requires an understanding of the downstream effects of the lpa genotype on seed development. Towards that end, we present a comprehensive comparison of gene-expression profiles between lpa and normal PA soybean lines (Glycine max) at five stages of seed development using RNA-Seq approaches. The lpa line used in this study carries single point mutations in a myo-inositol phosphate synthase gene along with two multidrug-resistance protein ABC transporter genes.

Results: RNA sequencing data of lpa and normal PA soybean lines from five seed-developmental stages (total of 30 libraries) were used for differential expression and functional enrichment analyses. A total of 4235 differentially expressed genes, including 512-transcription factor genes were identified. Eighteen biological processes such as apoptosis, glucan metabolism, cellular transport, photosynthesis and 9 transcription factor families including WRKY, CAMTA3 and SNF2 were enriched during seed development. Genes associated with apoptosis, glucan metabolism, and cellular transport showed enhanced expression in early stages of lpa seed development, while those associated with photosynthesis showed decreased expression in late developmental stages. The results suggest that lpa-causing mutations play a role in inducing and suppressing plant defense responses during early and late stages of seed development, respectively.

Conclusions: This study provides a global perspective of transcriptomal changes during soybean seed development in an lpa mutant. The mutants are characterized by earlier expression of genes associated with cell wall biosynthesis and a decrease in photosynthetic genes in late stages. The biological processes and transcription factors identified in this study are signatures of lpa-causing mutations.

Keywords: Phytic acid, myo-inositol phosphate synthase, Multidrug-resistance protein ABC transporter, Seed development, Transcriptomics, Differential gene expression, Functional enrichment, Apoptosis, Photosynthesis
Background

Soybean (Glycine max (L.) Merr.) seed is one of the most important agricultural commodities produced worldwide, generating oils, proteins, and carbohydrates [1]. Final seed composition is influenced by both the genotype and environmental factors [2–5]. Breeding programs endeavor to improve the functional properties, and hence the economic value of soybean by reducing anti-nutritive seed components such as phytic acid. Phytic acid (PA), a major source of phosphorus in seeds, can cause problems such as poor mineral bioavailability and phosphate pollution [6, 7]. Low PA (lpa) crops are therefore highly desirable for reducing anti-nutritional and environmental effects of conventional crops [8–10]. The lpa soybean line ‘V99-5089’ carries a non-lethal, recessive mutation in the myo-inositol phosphate synthase (MIPS) 1 gene, whereas ‘CX-1834’ carries mutations in two multidrug resistant protein (MRP) genes, encoding ATP-binding cassette transporters [11–16]. The MIPS1 gene, expressed during seed development in soybean, is associated with the conversion of glucose-6-phosphate to myo-inositol-3-monophosphate, which is the first step in PA biosynthesis pathway [17, 18]. Loss of function mutation in this gene disrupts this pathway. The MRP genes are also highly expressed in developing embryos, however, the mechanism by which they regulate PA levels in soybean is poorly understood [13].

The PA biosynthesis pathway plays a vital role in maintaining homeostasis. Several pathway intermediates, such as myo-inositol-1,4,5-trisphosphate, act as secondary messengers in signal transduction and are known to regulate growth and developmental processes, including phosphorus and mineral storage, DNA repair, chromatin remodeling, RNA editing and export, ATP generation, regulation of gene expression, regulation of guard cells, auxin metabolism and cell-wall polysaccharide biosynthesis [17, 19–24]. Numerous studies have reported the effects of lpa on plant growth and development. An RNAi-mediated mips1 knockdown in soybean was reported to inhibit seed development along with reduced PA content [25]. Similarly, seed embryo defects were reported for Arabidopsis and common bean (Phaseolus vulgaris L.) mips mutants [26, 27]. The mips mutation regulates raffinose family oligosaccharide pathway, and mutants exhibit impaired pathogen resistance, programmed cell death in leaves, and polar auxin transport causing deformed cotyledon development [26, 28–33]. The mpr mutants are known to exhibit lpa phenotypes in soybean, rice, maize, and Arabidopsis [12, 13, 34, 35]. MRP knockout studies in Arabidopsis exhibit phenotypes such as insensitivity to abscisic acid-mediated germination and unresponsive stomata opening, resulting in reduced transpiration rate and increased drought tolerance; which were rescued by MRP overexpression [36]. Finally, lpa crops are known to show poor seedling emergence, resulting in reduced crop yield, which decreases the agronomic value of lpa crops [37–39].

Despite these diverse physiological responses of different lpa mutations, very little is known about the effect of combining lpa mutations together on seed development and the underlying regulation of gene expression in soybean. Bowen et al. [40] investigated microarray-based gene expression changes in developing embryos of barley lpa mutant. This study identified several differentially expressed genes associated with different cellular processes, such as carbohydrate metabolism, hormonal signaling and transport [40]. Recent developments in sequencing technologies have enhanced the scope of genome-wide gene expression studies to a level far beyond microarray. An advanced generation recombinant inbred line (RIL) 3mlpa with three lpa-causing mutations, derived from a bi-parental cross of V99-5089 and CX-1834, provides a higher reduction in PA content of soybean seed relative to those with single mutations [41]. The 3mlpa triple mutant line and another advanced generation RIL 3MWT without any lpa-causing mutations, derived from the same cross (Table 1) are unique genetic materials for gene expression studies. In this report, we used mRNA-sequencing (RNA-Seq) to study the effect of lpa-causing mutations from MIPS1 and MRP genes on global changes in gene expression profiles of developing soybean seeds. A total of 30 transcriptome datasets derived from five developing seed stages with three biological replicates each of 3mlpa and 3MWT soybean line were sequenced and analyzed. To the best of our knowledge, this is the first extensive report describing the gene regulatory effect of MIPS1 and MRP mutations together. The significantly enriched biological processes and transcription factors identified in this study further our understanding of seed development lpa mutants.

Results and discussion

Differential expression analyses

Whole soybean seeds comprised of cotyledons, endosperm, and seed coat, were sampled in triplicate from 3mlpa and 3MWT lines at five stages of seed development (Fig. 1). These stages were chosen to correspond to seed filling stages post embryo development and before desiccation [2]. More than 961 million sequencing reads were obtained from 30 mRNA libraries, and 86.69 % (more than 833 million) of the reads were mapped to the annotated soybean reference, Williams 82 genome assembly 1.0 sequence (Fig. 2, Additional file 1). Read counts (number of reads mapping to a given gene) were estimated from the sequence mapping data for all the annotated gene models (total of 54,175) of Williams 82 genome (annotation v1.1). Normalized read count data were used for differential expression analyses. Principal component analysis (PCA) and
sample-to-sample distance clustering variance stabilized log₂ transformed normalized read count data for all genes from 30 sample libraries, are shown in Fig. 3. Sample libraries generated from different seed developmental stages were distinctly represented along PC1, in a unidirectional pattern starting from stage 1 to 5 in PCA; whereas those generated from the 3mlpa line were clearly differentiated from their respective 3MWT libraries along PC2 (Fig. 3a). This means that first and second major contributors to variation in the data are seed developmental stages and genotypes, respectively. At the same time, the biological replicates of each stage clustered together, suggesting minimal variance between replicates. Similarities and dissimilarities between individual sample libraries were visualized using heatmap of sample-to-sample distance clustering as shown in Fig. 3b. The sample libraries from early (stages 1–2) and late (stages 4–5) seed development stages were dissimilar to each other, while those from stage 3 are partially similar to both groups.

We identified a total of 6988 (4235 unique) genes with significant differential expression between the 3mlpa and 3MWT lines at five seed developmental stages (Table 2). Some of these DEGs may be due to non-isogenic background of the experimental lines. Of these differentially expressed genes (DEGs), 3321 (47.5 %) and 3493 (48.5 %) genes were identified as up- and down-regulated in the 3mlpa line. About 174 (2.5 %) and 102 (1.5 %) genes were expressed only in either 3mlpa line or 3MWT, respectively (Table 2). Several genes were differentially expressed in more than one stage, and 192 genes were differentially expressed between 3mlpa versus 3MWT in all five stages of seed development (Fig. 4).

### Functional enrichment analyses

Functional enrichment of gene ontology (GO) and transcription factor (TF) families employing DEGs for each stage was performed using a statistical hyper-geometric test employing the Benjamini-Hochberg method for multiple testing to obtain \( \text{adjusted-}P \)-values. The enriched results were then filtered using \( \text{adjusted-}P \)-value \( < 0.01 \) (or 1 % FDR) to obtain highly significant enriched terms. Table 3 shows highly significant enriched GO terms identified at five seed developmental stages. Although some terms were found in more than one stage or were stage-specific, none were common in all five stages. In order to simplify the interpretations, we grouped the stages into early (stages 1–2) and late phase (stages 4–5) of seed development. Enriched GO terms from stage 3, showed partial overlap with both early and late phases. Most of the DEGs associated with significantly enriched biological processes in early phase of seed development

### Table 1 Characteristics of experimental lines and their parents

| Soy Lines | Genotype       | Phytate | Emergence | Stachyose | Sucrose |
|-----------|----------------|---------|-----------|-----------|---------|
| V99-5089  | mips1/MPR-L/MPR-N | Low     | Low       | Low       | High    |
| CX-1834   | MIPS1/mrp-l/mrp-n | Low     | Low       | Normal    | Normal  |
| 3mlpa     | mips1/mrp-l/mrp-n | Low     | Low       | Normal    | Normal  |
| 3MWT      | MIPS1/MPR-L/MPR-N | Normal  | Normal    | Normal    | Normal  |

All experimental lines represented here are homozygous. Italicized text indicates mutations or mutant line. The soybean lines 3mlpa and 3MWT are two recombinant inbred lines derived from an advanced generation of a cross of V99-5089 with CX-1834

---

Redekar et al. BMC Genomics (2015) 16:1074
were up-regulated, whereas those in late phase were down-regulated in 3mlpa line.

Hierarchical clustering of mean normalized gene expression levels of DEGs associated with the enriched biological processes of cellular glucan metabolism, apoptosis, cellular transport, photosynthesis and glycolysis is shown in Fig. 5. The biological processes of cellular glucan metabolism, apoptosis and cellular transport were up-regulated in early stages (Fig. 5a-d), while those associated with photosynthesis and glycolysis were down-regulated in late stages (Fig. 5e-f) of 3mlpa line. The relative gene expression data from RNA-Seq analysis for two randomly selected DEGs (CESA: Glyma12g36570, and SugT1: Glyma08g06420) were evaluated using qPCR (Fig. 6). The genes CESA and SugT1 are associated with cellular glucan metabolism and cellular transport respectively. The RNA-Seq analysis showed up-regulation of CESA and SugT1 genes in 3mlpa in early stages of seed development. Both real time quantification and RNA-Seq analysis showed similar expression pattern for the CESA and SugT1 genes. There was no significant difference in the gene expression data obtained from qPCR and RNA-Seq analyses at the 0.01 significance level. We also tested relative expression of MIPS1, MRPL-L and MRPL-N genes, which were not differentially expressed in any of our seed stages. Both qPCR and RNA-Seq analyses estimated similar fold change expression with no significant difference between them (Additional file 2: Figure S1).

Soybean genome assembly 1 is annotated to contain 5683 TF genes from 63 TF families [42]. Total of 512 differentially expressed TF genes were identified in this study. These genes belonged to 32, 31, 33, 20, and 34 different TF families in seed developmental stages 1 to 5 respectively. TF family enrichment analysis resulted in identification of 2, 2, 4, 2, and 2 TF families significantly enriched in seed developmental stages 1 to 5, respectively (Table 4). We observed TF families GRAS (Gibberellin-Insensitive, Repressor of ga1-3, Scarecrow), WRKY, ZF-HD (Zinc Finger-Homeodomain), and ZIM (Zinc-finger protein expressed in Inflorescence Meristem) were enriched in early stages (stages 1–2), whereas, families CAMTA (CAMModulin-binding Transcription Activator), GRF (Growth-Regulating Factor1), MBF1 (Multiprotein Bridging Factor 1), SNF2, and TCP (Teosinte branched 1, Cycloidea, PCF) were enriched in later stages (stages 3–5) of seed development. Among these enriched TF families, TCP family was represented in stages 3 and 4, while CAMTA was represented in stages 3–5. All together, these TF families represented 53 unique DEGs. The genes belonging to enriched TF families, such as WRKY, GRAS, ZIM, CAMTA, GRF, and SNF2, were up-regulated, whereas, those belonging to ZF-HD, MBF1B, and TCP were down-regulated in the 3mlpa mutant.

**Regulation of defense response in early seed development of 3mlpa mutant**

‘Apoptosis’ (GO:0006915) and ‘Innate immune response’ (GO:0045087) were enriched in early phases of seed development, particularly in stages 1–3 and stage 3 respectively, with different sets of DEGs identified for these stages (Table 3). While several of these genes are shared between these two categories, a total of 61 DEGs were associated with both. Forty-seven of these genes (77 %) were up-regulated and 14 (23 %) genes were down-regulated in the 3mlpa mutant line. The FC and \( \log_{2} \) FC ratio of apoptosis-related differentially expressed genes are reported in Additional file 3. These defense-related DEGs encoded (a) LRR (leucine rich repeat) domain-containing disease resistance proteins, (b) NB-ARC (nucleotide-binding adaptor shared by APAF-1, R proteins, and CED-4) domain-containing disease resistance proteins, (c) Bcl-2-associated athanogene 1, (d) ADR1-L1 (Activated Disease Resistant 1-like 1), (e) cysteine proteinases, (f) protein kinase and (g) NTP hydrolases. Apoptosis is a common phenotype observed in mips1 mutants of Arabidopsis thaliana [26, 30, 43].

The defense-related genes identified in this study lie upstream of the apoptotic processes. For example, the LRR and NB-ARC domain-containing disease resistance genes are involved in initiating the defense response via induction of salicylic acid (SA) and formation of reactive oxygen species (ROS), both of which ultimately leading to cell death. Arabidopsis thaliana mips1 mutants were shown to accumulate SA and ceramide, both associated with formation of a cell death lesion area, with the phenotype rescued by treating with either myo-inositol
or galactinol [26, 30, 43]. Apoptosis in mips1 mutants is also regulated by light intensity and ROS such as peroxisomal hydrogen peroxide and the superoxide anion [26, 30, 43]. Myo-inositol mitigates SA-dependent apoptosis and defense responses [43]. Although apoptosis is associated with mips1 mutation and not with mrp mutation, induction of the defense-related genes in 3mlpa may not be solely due to mips1 mutation. The ADR1-L1 genes (Glyma14g08700 and Glyma17g36420), which belong to a subgroup of CNL-A clade of coiled-coil NBS-LRR gene family, showed two-fold higher expression in 3mlpa mutant [44]. The ADR1 genes are positive regulators of SA-mediated cell death and defense response [45]. Enhanced expression of ADR1 gene was previously shown to establish drought tolerance in presence of SA [46]. And interestingly, the mrp5 mutants of Arabidopsis are also known to be drought tolerant [36]. Role of MRP genes in regulating defense responses is still unknown,

**Fig. 3** Biological sample variability. a Principle component analysis plot explains the variance in gene expression data from biological sample libraries along PC1 or X-axis and PC2 or Y-axis. b Sample clustering heatmap representing sample-to-sample distance. Blue color suggests similarity between sample libraries. Samples A and B correspond to 3mlpa mutant, and 3MWT, respectively, e.g. A3-2, means 3mlpa-stage3-replicate2 (See Additional file 3 for more information)
and further studies are required to confirm the association of the drought tolerant phenotype of the *mrp* mutants and up regulation of ADR1 genes as in 3mlpa. In addition, several WRKY transcription factor family genes that regulate expression of other defense-related genes were also up-regulated in early stages of seed development in 3mlpa mutant. In summary, the *lpa*-causing mutations in 3mlpa mutant may play a role in initiating defense responses during seed development.

### Table 2 Differential gene expression between 3mlpa and 3MWT

| Stages | 1 | 2 | 3 | 4 | 5 |
|--------|---|---|---|---|---|
| DEGs up-regulated in 3mlpa | 831 | 1114 | 788 | 269 | 493 |
| DEGs down-regulated in 3mlpa | 695 | 677 | 560 | 415 | 1146 |
| DEGs only expressed in 3mlpa | 29 | 41 | 39 | 32 | 33 |
| DEGs only expressed in 3MWT | 26 | 20 | 17 | 21 | 18 |

Out of total 6988 DEGs identified in five seed developmental stages, 4235 were unique, meaning only counted once. Remaining genes were repeatedly identified in more than one stage.

**Regulation of photosynthesis in late stages of seed development in 3mlpa mutant**

‘Photosynthesis’ (GO:0015979) was enriched in late phase (stages 3–5) of seed development, with most of the photosynthesis-related DEGs from these stages were down-regulated in 3mlpa mutant (Table 3, Fig. 5). Fifty-six DEGs were associated with photosynthesis and of these only one gene (Glyma15g40131) was up-regulated in 3mlpa mutant. These DEGs genes encode different subunits in photosystem (PS) I (psaD, psaE, psaF, psaG, psaH, psaK, psaL, and psaN), and PS II (psbA, psbE, psbP, psbQ, psbW, psbX, and psbY), light-harvesting chlorophyll complex proteins from PS I (LHCA1 and LHCA2) and PS II (LHCB1, LHCB4, LHCB5), and a single gene encoding for magnesium protoporphyrin IX methyl transferase (CHLM) (Additional file 4). A previous study on transcriptome profiles of soybean cultivar ‘Williams’ reported expression of photosynthesis-related genes early in development when seed size was 25–50 mg [47]. This would correspond to stages 2–3 in our investigation. Bowen et al. [40] previously reported two photosynthesis related probesets that were differentially expressed in

---

**Fig. 4** Overlap of differentially expressed genes between developmental stages. Each oval in this Venn diagram corresponds to a seed developmental stage. Numbers within each oval represents significant differentially expressed genes between 3mlpa and 3MWT overlapping between- or unique to-seed developmental stages.
M955 maize line with lpa at 7 days post anthesis. One representing chloroplast nucleoid DNA binding protein was down-regulated, while another representing early light inducing protein, was up-regulated in M955 line [40]. Although these proteins were not identified in our study, regulation of photosynthesis-related genes in lpa mutants is a common observation.

The PSI and PSII protein complexes are involved in capturing light energy during photosynthesis. Differential expression of these genes may translate into different numbers of PSI and PSII protein complexes between 3mlpa and 3MWT. The methyl transferase CHLM is involved in chlorophyll metabolism where it catalyzes the formation of Mg-protoporphyrin IX monomethyl ester [48]. Bruggeman et al. [49] have reviewed the association of chloroplast activity with the formation of cell death lesions, where excess light energy can trigger cell death. Down regulation of photosynthesis-related genes in 3mlpa may prevent apoptosis by limiting the light energy during late stages of seed development. Cell death lesion formation in mips1 Arabidopsis mutants is dependent on light intensity and ROS accumulation [26, 30, 43], and abolishing PSII assembly was shown to reduce cell death lesions [50]. Tetrapyrrole biosynthesis is also regulated to prevent accumulation of ROS, which can also lead to cell death [49]. Moreover, we also identified enhanced expression of CAMTA3/SIGNAL RESPONSIVE 1 (SR1) TF genes in 3mlpa during late stages of seed development. CAMTA3/SR1 is a negative regulator of SA-mediated plant immunity and it inhibits cell death [51]. Therefore, these observations strongly suggest that the up-regulation of cell death inducing genes in early stages of seed development in 3mlpa mutant is suppressed in late stages. Myo-inositol promotes photosynthesis in Mesembryanthemum crystallinum so another explanation for down-regulation of photosynthesis-related genes in 3mlpa mutant may be the mips1 mutation, which reduces myo-inositol levels [52].

**Table 3** Enriched gene ontology terms associated with biological processes

| Stage | GO term | Biological processes | DEG | P-value | FDR |
|-------|---------|----------------------|-----|---------|-----|
| 1     | GO:0006915 | Apoptosis | 31  | 4.0E-04 | 3.4E-03 |
|       | GO:0006073 | Cellular Glucan Metabolic Process | 17  | 6.8E-05 | 7.4E-04 |
|       | GO:0006334 | Nucleosome Assembly | 15  | 6.9E-06 | 1.2E-04 |
|       | GO:0006412 | Translation | 71  | 2.9E-10 | 6.4E-08 |
| 2     | GO:0006915 | Apoptosis | 39  | 1.1E-05 | 7.9E-04 |
|       | GO:0006073 | Cellular Glucan Metabolic Process | 18  | 1.0E-04 | 3.4E-03 |
|       | GO:0006857 | Oligopeptide Transport | 15  | 1.6E-04 | 3.5E-03 |
|       | GO:0055114 | Oxidation Reduction | 140 | 3.6E-06 | 7.9E-04 |
|       | GO:0055085 | Transmembrane Transport | 68  | 3.4E-04 | 6.3E-03 |
| 3     | GO:0006915 | Apoptosis | 42  | 1.9E-10 | 1.7E-08 |
|       | GO:0006096 | Glycolysis | 10  | 1.2E-03 | 9.8E-03 |
|       | GO:0045087 | Innate Immune Response | 20  | 2.2E-06 | 5.1E-05 |
|       | GO:0015979 | Photosynthesis | 20  | 6.0E-07 | 2.2E-05 |
| 4     | GO:0006091 | Generation of precursor metabolites and energy (Glycolysis and Photosynthesis) | 14  | 7.0E-06 | 4.9E-04 |
|       | GO:0015979 | Photosynthesis | 16  | 4.1E-09 | 5.6E-07 |
| 5     | GO:0008652 | Cellular Amino acid Biosynthetic Process | 18  | 1.0E-03 | 4.0E-03 |
|       | GO:0006096 | Glycolysis | 12  | 7.3E-04 | 3.8E-03 |
|       | GO:0006184 | GTP Catabolic Process | 7  | 6.6E-04 | 3.6E-03 |
|       | GO:0008610 | Lipid Biosynthetic Process | 30  | 9.2E-04 | 3.8E-03 |
|       | GO:0006108 | Malate Metabolic Process | 6  | 8.3E-05 | 5.7E-04 |
|       | GO:0006334 | Nucleosome Assembly | 20  | 9.2E-09 | 2.1E-07 |
|       | GO:0015979 | Photosynthesis | 55  | 4.4E-30 | 1.1E-27 |
|       | GO:0009765 | Photosynthesis, Light Harvesting | 14  | 1.3E-09 | 6.6E-08 |
|       | GO:0006814 | Sodium ion Transport | 5  | 8.2E-04 | 3.8E-03 |
|       | GO:0006414 | Translational Elongation | 9  | 4.6E-04 | 2.7E-03 |
|       | GO:0006412 | Translation | 63  | 1.7E-06 | 1.9E-05 |

The GO terms identified in this study are more specialized (or, child) terms in the GO hierarchy.
Fig. 5 Mean normalized gene expression profiles of DEGs associated with biological processes (a) Cellular glucan metabolic process, (b) Apoptosis, (c) Oligopeptide transporters, (d) Transmembrane transporters, (e) Photosynthesis, and (f) Glycolysis. Hierarchical clustering of mean normalized gene expression values based on euclidean distance between seed developmental stages of 3mlpa and 3MWT. Rows represent genes, while columns represent samples. Green color indicates higher gene expression values.
Fig. 6 Relative gene expression of DEGs for RNA-Seq data validation. Fold change between 3mlpa and 3MWT at respective seed developmental stages for genes encoding (a) Cellulose synthase (CESA: Glyma12g36570) and (b) Sugar transporter (SugT1: Glyma08g06420). Green and orange bars indicate mean fold change values from RNA-Seq and qPCR experiments, respectively. There was no significant difference in the gene expression profiles estimated using qPCR and RNA-Seq analyses at significance level of 0.01 (See Additional file 8 for more information)

Table 4 Transcription factor families significantly enriched in developing seed stages

| Stages | TF family | FDR     | Gene symbols     | DEGs associated with the TF family | Functions                                                                 |
|--------|-----------|---------|------------------|-----------------------------------|---------------------------------------------------------------------------|
| 1      | WRKY      | 4.92E-04| WRKY13, WRKY40, WRKY29, WRKY6, WRKY28, WRKY23, WRKY15, WRKY11 | Glyma1g29720, Glyma08g23380, Glyma08g02160, Glyma13g44730, Glyma09g00820, Glyma12g10350, Glyma08g08720, Glyma13g38630, Glyma15g11680, Glyma03g37940, Glyma17g18480, Glyma05g20710, Glyma06g08120, Glyma04g08060 | Associated with plant defense, senescence, and abiotic stress [73] |
|        | GRAS      | 3.19E-04| SCL1, SCL3, SCL5, SCL14, PAT1, SGR7 | Glyma04g28490, Glyma18g09030, Glyma08g43780, Glyma11g14670, Glyma14g27290, Glyma13g09220, Glyma15g04160, Glyma14g01960, Glyma14g01020, Glyma17g17400, Glyma03g37940 | Involved in gibberellin signaling, phytochrome A signal transduction, controls radial patterning [74–77] |
| 2      | ZIM       | 8.60E-03| JAZ6, JAZ12      | Glyma17g04850, Glyma16g01220, Glyma07g04630 | Repressor of jasmonate responses [78] |
|        | ZF-HD     | 7.44E-03| HB22, HB24, HB33 | Glyma06g09970, Glyma04g09910, Glyma06g06120, Glyma11g07360, Glyma14g35770 | Regulator of ABA-response [79] |
| 3      | TCP       | 3.23E-05| TCP3, TCP20      | Glyma09g42140, Glyma16g05840, Glyma19g26560, Glyma12g35720, Glyma06g34330, Glyma13g34690, Glyma03g02090, Glyma09g42120 | Controls cell expansion and morphogenesis, negatively regulates auxin response, and promotes flavonoid synthesis [80, 81] |
|        | SNF2      | 1.63E-03| EDA16, RGD3, PIE1, SYD, CHR5 | Glyma07g31180, Glyma09g36910, Glyma02g29380, Glyma17g02540, Glyma13g25310, Glyma02g45000 | Embryo sac development, repressor of flowering, chromatin remodeling, gravitropism [82] |
| 4      | GRF       | 8.68E-03| Glyma19g28010   | Lygma10g31190, Glyma08g14370, Glyma17g04310 | Cotyledon growth [83] |
|        | CAMTA     | 4.18E-04| CAMTA3, SRI     | Glyma05g31190, Glyma08g14370, Glyma17g04310 | Negative regulator of plant immunity [51, 84] |
| 5      | TCP       | 9.25E-03| TCP3            | Glyma13g34690, Glyma09g42120, Glyma09g42140 | Negatively regulates auxin response, and promotes flavonoid synthesis [80] |
|        | MBF1      | 5.27E-03| MFB1B           | Glyma06g42890 | Negative regulator of ABA-dependent inhibition of germination [85] |
|        | CAMTA     | 9.57E-03| CAMTA3, SRI    | Glyma08g11080, Glyma08g14370 | Negative regulator of plant immunity [51, 84] |
Regulation of glucan metabolism in early seed development of 3mlpa mutant

Cellular glucan metabolic process (GO:0006073) was found enriched in both stage 1 and 2 (early phase of seed development), however the DEGs associated with this process were stage-dependent (Table 3, the FC and log2 FC ratios for these genes are reported in Additional file 5). Twenty-seven genes were up-regulated, and two genes (Glyma02g47076 and Glyma04g43470) were down-regulated in 3mlpa mutant. Six genes were differentially expressed in both stages and all were up-regulated in 3mlpa mutant. The FC of up-regulated genes ranged between 1.6–60 and 1.3–18 for stages 1 and 2, respectively. In case of down-regulated genes, the FC was between 0.2–0.6 for stages 1 and 2, respectively. Two genes (Glyma09g07070 and Glyma15g18360) showed no expression or zero read count in 3MWT resulting in infinite (∞) value for FC and log2 FC ratio. The differentially expressed genes associated with this process encoded two enzymes: (a) cellulose synthase (also, known as glucan synthase, or CESA, EC 2.4.1.12); and (b) xylolucan endotransglycosylase/hydrolase (also known as xylolucan: xylolucosyl transferase or XET, EC 2.4.1.207). Functional enrichment also showed over-representation of genes associated with XET enzymatic activity (GO:0016762). The CESA and XET enzymes are involved in synthesis of building units of cell wall, i.e. cellulose, and xylolucan (hemicellulose) chains, respectively. Cellulose synthase activity is an integral component of cell wall synthesis. A previous study on transcriptome profiles of soybean line 'A81–356022' reported higher cellulose synthase activity in developing seeds at 25–28 days post flowering [53]. Differential expression of the CESA and XET genes in early seed development suggests regulation of cell wall synthesis in 3mlpa mutant line. Role of myo-inositol in cell wall synthesis via the myo-inositol oxidation pathway, resulting in formation precursors of pectin and hemicellulose, is well established and regulation of cell wall biosynthesis in mips mutants is therefore expected [23, 54]. The contribution of either the mips or mrp mutations towards this response lies out of scope of this experiment, however, it is possible that 3mlpa mutants initiate formation of cell wall polysaccharides at earlier stages of development relative to the non-mutants.

Regulation of cellular transport in early seed development of 3mlpa mutant

Processes involving oligopeptide transporters (GO:0006857) and transmembrane (GO:0055085) transporters were enriched in stage 2 (Table 3). Although several transporter genes were differentially expressed in other stages, the GO term associated with the transporter activity was not significantly enriched in those stages. Together, oligopeptide and transmembrane transport activity were associated with a total of 79 unique differentially expressed genes and 86 % of these genes were found up-regulated in the mutant line. Twenty-nine (44 %) of these up-regulated genes encode for multidrug transporters, including 15 from major facilitator superfamily (MFS), six from multidrug and toxic compound extrusion (MATE) efflux carrier superfamily, five from multidrug resistance superfamily, two genes for P-glycoprotein (PGP) and 1 gene for ATP binding cassette (ABC) subfamily (B4) (Additional file 6). Most of the multidrug transporters are annotated as being involved in removal of toxic compounds from the cell [55]. Recently, a MFS transporter, Zinc-Induced Facilitator-Like 1 from Arabidopsis was reported to be associated with polar auxin transport in roots, as well as regulation of stomata for drought stress tolerance [56]. The ABC B4 transporters are also involved in auxin-gradient dependent polar auxin transport in root [57], and PGP transporters are involved in cellular and long distance transport of auxin [58]. Defects in mips mutant embryos were previously associated with impaired endomembrane system and lack of polar auxin transport [32]. The genes encoding monosaccharide transporters such as Sugar Transport Protein, Inositol Transporter and Polyol/Monosaccharide transporter were also differentially expressed in our dataset. These genes are involved in transport of sugars such as glucose, fructose, galactose, mannose, xylose, sorbitol, mannitol, xylitol, and epimers and derivatives of myo-inositol [59]. Other genes encoding transporters/carriers of cationic amino acids, oligopeptides, potassium, sulfate, nitrates, zinc, chloride and dicarboxylate ions were also identified in 3mlpa mutant line. Although substrates of MRP-L and MRP-N transporters are still unknown, enrichment of the transporter genes may be compensating for the loss of MRP-L and MRP-N.

Regulation of raffinose family oligosaccharide biosynthesis in developing seeds of 3mlpa mutant

The raffinose family oligosaccharides such as raffinose and stachyose are considered to be anti-nutritive components in seeds [60]. The biosynthesis of raffinose family oligosaccharide (RFO) involves three steps: (a) galactinol synthase (GS) catalyzes the formation of galactinol from myo-inositol and UDP-galactose, (b) raffinose synthase (RS) catalyzes the formation of raffinose by adding sucrose to galactinol and releasing a myo-inositol molecule, and (c) stachyose synthase (SS) catalyzes the formation of stachyose by adding galactinol to raffinose and releasing myo-inositol (Fig. 7). The mips1 mutation is associated with low stachyose phenotype in the parental line V99-5089 [15]. However, the 3mlpa mutant shows normal stachyose levels in mature seeds despite having mips1 mutation [41]. The genes encoding RFO biosynthesis pathway enzymes are expressed in seed filling and desiccation stages of soybean seed development [61]. None
of these genes showed significant differential expression at 1 % FDR, which may be one possible explanation for normal stachyose phenotype in 3mlpa mutant. Figure 7 represents RFOs biosynthesis pathway, and the relative gene expression values for selected genes in the pathway.

There are six genes (Glyma03g38080, Glyma03g38910, Glyma10g28610, Glyma19g40680, Glyma19g41550, and Glyma20g22700) annotated as GS in first assembly of reference soybean genome (Glyma.Wm82.a1.0). The Glyma19g41550 gene showed highest-FC in RNA-Seq experiment, so we estimated relative expression for this GS gene using qPCR. The relative transcript levels of GS gene showed a gradual increase in 3mlpa during seed development (Fig. 7a). RNA-Seq data showed similar gradual increase in expression in 3mlpa during seed development, but expression dropped at stage 4 (Fig. 7a). Glyma06g18890 encodes RS (also known as RS2 or Rsm1) is expressed during seed development in Williams 82 soybean line, and is known to control the raffinose and stachyose content in soybean seeds [60, 62]. The relative transcript levels of the RS gene showed more than 2-fold change in 3mlpa in late phase (stages 3–5) of seed development, with no difference between the stages 3–5 (Fig. 7b). The fold change of RS remained below 2 in 3mlpa during early stages. This validates the RNA-Seq data except for stage 1, where fold change of RS was close to 3 in 3mlpa (Fig. 7b). Glyma19g40550 is the only SS encoding gene in Glyma.Wm82.a1.0. At stage 1 and 5, SS gene showed significant difference in the fold change obtained from qPCR and RNA-Seq analyses. The relative transcript levels of the SS gene showed almost no change between 3mlpa and 3MWT with fold change below 2 (Fig. 7c). It is possible that mechanisms other than differential expression are involved in regulation of RFO biosynthesis pathway, which does not allow us to elaborate further on these observations.

Fig. 7 Raffinose family oligosaccharide biosynthesis pathway. Bar graphs indicate fold change (3mlpa over 3MWT) ratio of normalized expression values for genes (a) galactinol synthase, (b) raffinose synthase and (c) stachyose synthase at 5 stages of seed development. Green and orange bars indicate mean fold change values from RNA-Seq and qPCR experiments, respectively. Gene expression profiles showing significant difference between qPCR and RNA-Seq analyses are indicated by asterisk (*) (Refer Additional file 8 for more information).
**Conclusions**

The PA biosynthesis pathway intermediates are involved in many growth and developmental processes. Several genes encoding key enzymes regulating this pathway are mutated to obtain lpa crops. We used a transcriptomics approach to compare the gene expression profiles of a lpa soybean line (3mlpa) carrying three (mips1, mrp-l, and mrp-n) mutations and a non-mutant line, 3MWT, at 5 stages of soybean seed development. The differential expression and functional enrichment analyses indicated regulation of biological processes such as glucan synthesis, cell death and photosynthesis. We also identified regulated transcription factor families, including WRKY, CAMTA, GRAS and ZIM. These results delineate the metabolic events associated with regulation of the PA biosynthetic pathway in the presence of lpa-causing mutations. We also quantified transcript levels of genes involved in the raffinose family oligosaccharides pathway in 3mlpa mutant. Overall, these results contribute to an understanding of regulation of metabolism in 3mlpa mutant during seed development.

**Methods**

**Genetic material and background**

The experimental lines of soybean (Glycine max (L.) Merr.) used in this study were: (a) a triple homozygous mutant line, designated as ‘3mlpa’ (mips1/mrp-l/mrp-n) with lpa, and (b) a sibling line with homozygous non-mutant alleles, designated as ‘3MWTL’ (MIPS1/MRP-L/MRP-N) with normal PA (Table 1). These lines were developed from a cross of V99-5089 with CX-1834 [16]. V99-5089 soybean experimental line carries a point mutation in MIPS1 gene (chromosome 11) that results in lpa, low stachyose, and high sucrose phenotype [15]. Similarly, CX-1834 soybean line carries point mutations in two MRP genes, namely, MRP-L and MRP-N, located on chromosomes 19 and 3 respectively. Both mrp mutations are required to obtain the lpa phenotype, and have no effect on seed stachyose and sucrose contents [12, 16, 63]. The MIPS1 gene does not interact with MRP-L and MRP-N genes. Although the 3mlpa line, carrying mutations in MIPS1 and two MRP genes, shows reduced PA content, the stachyose and sucrose phenotype associated with the mips1 mutation was rescued (Table 1) [41]. The low phytic acid soybean exhibits reduced seed emergence, however, the molecular basis of this relationship is still unknown.

**Plant growth, and sampling**

For each of the two experimental lines, 48 plants were grown in 12 pots (four plants per pot) containing Metro-Mix® 360 (Sun Gro) soilless media, over-layered with topsoil GardenPro ULTRA®. All plants were grown in the same growth chamber unit, with controlled conditions, as follows: 14/10 h (day/night) photoperiod, 24/16 °C (day/night), light intensities in the range of 300 and 400 μE and 50–60 % relative humidity. About 41–47 plants from each experimental line were used for sampling developing seeds. Seed length was the criterion for sampling different developmental stages. Pods were randomly selected, opened and, seed length was measured. Five developmental stages were defined by seed length as: (Stage 1) between 2–4 mm; (Stage 2) between 4–6 mm; (Stage 3) between 6–8 mm; (Stage 4) between 8–10 mm; and, (Stage 5) between 10–12 mm (Fig. 1). Three biological replicates for each stage were taken, where each replicate sample was represented by a minimum of 10–15 seeds (stages 1–2), and at least 3 seeds (stages 3–5), collected from different pods on separate plants. Sampled seeds were immediately frozen in liquid nitrogen and stored at −70 °C.

**RNA extraction, library preparation, and mRNA sequencing**

Frozen seeds were ground to a fine powder and total RNAs were extracted using RNasey Plant Mini Kit, with on column DNase digestion (QIAGEN). RNA concentrations were determined by UV spectrophotometry (260 nm, NanoDrop 1000, Thermo Fischer Scientific). RNA concentrations were normalized to 200 ng/μl and the RNA integrity number (RIN) was measured using a Bioanalyzer (Agilent Technologies). Total RNA samples with RIN values ranging between 9.0–10.0, an indication of high quality, were kept. High quality total RNA samples (50 μl each) were used for library preparation and mRNA sequencing at the Génome Québec Innovation Centre, Canada. A total of 30 cDNA libraries were generated using the TruSeq RNA sample preparation kit (Illumina) and sequenced on five lanes of a HiSeq2000 sequencing system (Illumina), to obtain single-end 100-bp long RNA-seq reads. Six libraries representing three biological replicates of single sampling stage from both 3mlpa and 3MWT were multiplexed together in single lane.

**Transcriptomics data processing and analysis**

Sequencing data quality control was performed prior to data analysis. Sequencing reads were then mapped/aligned to the well-annotated “Williams 82” soybean reference genome (assembly Glyma.Wm82.a1.0, annotation v1.1) using the splice-aware mapping tool, TopHat, v2.0.8 [42, 64, 65]. Sequence mapping data was used to estimate expression values for annotated genes using HTSeq-count ‘Union’ mode [66]. Differential gene expression analyses were performed using the statistical tool, DESeq, v1.12.1 [67]. Fold-change (FC) was calculated by dividing mean normalized gene expression value in 3mlpa over that in 3MWT. Most significant genes were identified at 1 % false discovery rate (FDR) calculated using P-value adjusted for multiple testing using Benjamini-Hochberg method [52].
Functional enrichment analysis was performed to identify ontology terms and pathways represented by these significant genes using online AgriGO tool [68]. The enriched results were then filtered using 1 % FDR to obtain highly significant enriched terms. R-script for statistical hyper-geometric test was used to identify significantly enriched transcription factor families. Most significant GO terms and transcription factor families were identified at 1 % FDR using Benjamini-Hochberg method.

**Quantitative real time PCR**

RNA-Seq output was validated using quantitative PCR (qPCR). First strand cDNA was synthesized from two μg high quality total RNA (see above) using High Capacity RNA-to-cDNA kit (Applied Biosystems) following manufacturer's instructions, from a total of 30 samples comprising three biological replicates for each time point. Two μl of stock cDNA from each of the 30 samples was diluted ten fold in water. Quantitative real time PCR was performed in 20 μl total reaction volume, constituted of four μl of diluted cDNA, 10 μl of 2x SYBR Green PCR Master Mix (Applied Biosystems), 0.4 μl of 10 mM each gene-specific primers, and 3.2 μl of UltraPure™ distilled water. PCR conditions were: 50 °C for 2 min, 95 °C for 10 min, followed by 40 cycles at 95 °C for 20 s and primer annealing temperature for 1 min, using 7500 Real Time PCR system (Applied Biosystems). Melting curve analyses were performed to test primer specificity. Target gene specific primers were designed according to the soybean reference sequence using Primer3.0 for six randomly selected most significant differentially expressed genes [69, 70]. Primer information is provided in Additional file 7. Ubiquitin 10 (UBQ10) gene was used as a reference gene to normalize target gene transcript level amongst all samples [71]. For estimating target and reference gene efficiency, equal volumes of diluted cDNA from all samples were pooled together. Standard curve of Ct values was generated using five-fold serial dilution of this pooled cDNA sample. PCR efficiency was estimated using slope of this standard curve as: $E = 10^{\left(-1/\text{slope}\right)}$. Relative quantification of target gene transcript was estimated using Pfaffl (or, efficiency correction) method [72]. The fold change gene expression values of stage-specific samples estimated using qPCR and RNA-Seq analyses were first tested for normality and to compare variance using the Shapiro-Wilk test and F-test, respectively. The Student's t and Mann-Whitney-Wilcoxon tests were performed to check the null hypothesis that there is no significant difference in qPCR and RNA-Seq data at the significance level of 0.01 (Additional file 8).

**Availability of supporting data**

The data sets supporting the results of this article are available in the NCBI-Gene Expression Omnibus (GEO, http://www.ncbi.nlm.nih.gov/projects/geo/) accession no. GSE75575.

### Additional files

**Additional file 1:** Sequencing data from developing seed tissue of soybean. (XLSX 57 kb)

**Additional file 2:** Figure S1. Relative gene expression of MIPS1, MRP-L, MRP-N. Fold change between 3mpla and 3mWT at respective seed developmental stages for genes encoding (a) MIPS1, (b) MRP-L and (c) MRP-N. Green and orange bars indicate mean fold change values from RNA-Seq and qPCR experiments, respectively. There was no significant difference in the gene expression profiles estimated using qPCR and RNA-Seq analyses at significance level of 0.01 (See Additional file 7 for more information) (PDF 27 kb)

**Additional file 3:** Defense related genes differentially expressed in early stages of seed development. Column ‘S’ reports the stages of development in which the gene is differentially expressed. Genes down-regulated in 3mpla mutant are indicated by negative values of log2 ratio. Adjusted P-values are only reported for the stages in which the gene is differentially expressed. (XLSX 60 kb)

**Additional file 4:** Photosynthesis related genes differentially expressed in late stages of seed development. (XLSX 49 kb)

**Additional file 5:** Cellular glucan metabolism related genes differentially expressed in early stages of seed development. The genes up-regulated in 3mpla mutant in both stages are indicated by asterisk (*) symbol. The genes differentially expressed only in stage 1 and stage 2 are indicated by section sign (§) and double dagger (‡) symbols, respectively. The non-significant adjusted P-values are indicated by dash (-). The genes down-regulated in 3mpla mutant are indicated by negative values of log2 ratio. Zero read count (no gene expression) in 3mWT, resulting in infinite value for FC and log2 ratio indicated by infinite sign (∞). (XLSX 59 kb)

**Additional file 6:** Enriched multidrug transporter genes up-regulated in 3mpla (XLSX 51 kb)

**Additional file 7:** Primers used for quantitative real-time PCR. Gene ID for MIPS1 was not available for the first reference soybean assembly; therefore it is indicated as ‘NA’. Published primer sequences were obtained for MIPS1 [18], RS2 [63] and housekeeping gene, UBQ10 [71] (XLSX 45 kb)

**Additional file 8:** Statistical test results for qPCR and RNA-Seq data comparison. The stage-specific data for each gene was tested for normality and sample variances using Shapiro-Wilk’s test and F-test, respectively. This was followed by Student’s t-test and Mann-Whitney U test to check the null hypothesis that there is no significant difference in the gene expression data estimated using qPCR and RNA-Seq analyses, at the significance level of 0.01. Samples that showed significant difference in the expression between qPCR and RNA-Seq data are indicated by asterisk (*) symbol. The sample for which qPCR and RNA-Seq data showed unequal variance is indicated by hash (#) symbol (XLSX 51 kb)

### Abbreviations

PA: Phytic acid; Lpa: Low phytic acid; MIPS1: myo-inositol phosphate synthase 1; MRP: Multidrug resistant protein; RNA-Seq: mRNA sequencing; PCO: Principal component analysis; DEG: Differentially expressed gene; GO: Gene ontology; TF: Transcription factor; CES: Cellulose synthase; XET: Xyloglucan endotransglucosylase/hydrolase; SugT1: Sugar transporter; GRAS: Gibberellin-insensitive, repressor of ga1 endotransglucosylase/hydrolase; SA: Salicylic acid; PSI: Photosystem I; PSII: Photosystem II; LHC: Light harvesting complex; CHLM: Magnesium protoporphyrin IX methyl transferase; MFS: Major facilitator superfamily; MATE: Multidrug and toxic compound extrusion; ABC: ATP binding cassette; PGP: P-glycoprotein; FKO: Raffinose family
Acknowledgements

Funding for this work was through the Bio-design and Bioprocessing Research Center (BBRC) and the John Lee Pratt Fellowship Program, both at Virginia Tech. Additional funding was provided by the Virginia Tech Agricultural Experiment Station Hatch Program. We would like to thank support team of the Advanced Research Computing (ARC) server and Translational Plant Sciences’ MAGYK server at Virginia Tech. We also would like to thank Dr. Victor Raboy, from United States Department of Agriculture’s Agricultural Research Service in Aberdeen, Idaho, Dr. Song Li, Dr. John McDowell, Dr. Guillaume Pilot, and Dr. Glenda Gillaspy from Virginia Tech, for providing helpful comments to improve this manuscript.

Author details

1Department of Crop and Soil Environmental Sciences, Virginia Tech, 185 AgQuad Lane, 24061 Blacksburg, VA, USA. 2Department of Biological Sciences, Virginia Tech, Life Science I building, 24061 Blacksburg, VA, USA. 3Department of Biochemistry, Virginia Tech, Life Science I building, 24061 Blacksburg, VA, USA. 4Department of Plant Pathology, Physiology, and Weed Science, Virginia Tech, Price Hall, 24061 Blacksburg, VA, USA.

Received: 30 August 2015 Accepted: 10 December 2015

Published online: 18 December 2015

References

1. Wilson JH. Seed composition. In: Boerma HR, Specht JE, editors. Soybeans: improvement, production, and uses. 3rd ed. Madison: American Society of Agronomy, Crop Science Society of America, Soil Science Society of America; 2004. p. 621–68.

2. Meinke DW, Chen J, Bechy RN. Expression of storage-protein genes during soybean seed development. Planta. 1981;153(2):130–9. doi:10.1007/BF00384094.

3. Fehr WR, Hrocz AK, Johnson LS, Murphy PA, Nott JD, Padilla GI, et al. Genotype and Environment Influence on Protein Components of Soybean. Crop Sci. 2003;43(2):511. doi:10.2135/cropsci2003.0511.

4. Breene WM, Lin S, Hardman L, Of J, Protein and oil content of soybeans from different geographic locations. J Am Oil Chem Soc. 1988;65(12):1927–31. doi:10.1007/bf02546009.

5. Bennett JO, Krishnan HB. Long-term study of weather effects on soybean seed composition. Korean J Crop Sci. 2005;50(1):32–8.

6. Raboy V. Accumulation and storage of phosphate and minerals. Cellular and Molecular Biology of Plant Seed Development. Dordrecht: Kluwer Academic Publishers; 1997. p. 447–71.

7. Cosgrove DJ, Irving GCJ. Inositol phosphates: their chemistry, biochemistry, and physiology. Amsterdam: Elsevier Scientific Pub. Co; 1980.

8. Zhou JR, Fordyce EJ, Raboy V, Dickinson DB, Burns RA, Wu H, et al. Reduction of phytic acid in soybean products improves zinc bioavailability in rats. J Nutr. 1992;122(12):2466–73.

9. Hill BE, Sutton AL, Richert BT. Effects of low-phytic acid com, low-phytic acid soybean meal, and phytase on nutrient digestibility and excretion in growing pigs. J Anim Sci. 2009;87(4):1518–27. doi:10.2527/jas.2008-1219.

10. Htto JK, Sauer WC, Zhang Y, Cervantes M, Liao SF, Araiza BA, et al. The effect of feeding low-phytate barley-soybean meal diets differing in protein content to growing pigs on the excretion of phosphorus and calcium. J Anim Sci. 2007;85(5):1700–5. doi:10.2527/jas.2006-249.

11. Wilcox JR, Premachandra GS, Young KA, Raboy V. Isolation of High Seed Inorganic P, Low-Phytate Soybean Mutants. Crop Sci. 2000;40(6):1601. doi:10.2135/cropsci2000.4061601x.
33. Obendorf RL, Zimmerman AD, Zhang Q, Castillo A, Kosina SM, Bryant EG, et al. Accumulation of Soluble Carbohydrates during Seed Development and Maturation of Low-Raffinose, Low-Stachyose Soybean. Crop Sci. 2009; 49(1):329–41. doi:10.2135/cropsres.2008.06370.

34. Nagy R, Grob H, Weder B, Green P, Klein M, Frelet-Barrant A, et al. The Arabidopsis ATP-binding cassette protein AtMRPS/AtABCBS is a high affinity inositol hexakisphosphate transporter involved in guard cell signaling and phytate storage. J Biol Chem. 2009;284(48):33614–22. doi:10.1074/jbc.M109.030247.

35. Xu XH, Zhao HJ, Liu QL, Frank T, Engel KH, An G, et al. Mutations of the multi-drug resistance-associated protein ABC transporter gene 5 result in reduction of phytic acid in rice seeds. TAG Theor Appl Genet. 2009;119(1):75–83. doi:10.1007/s00122-009-0918-1.

36. Klein M, Perful-Barbecho L, Frelet A, Gaedeke N, Reinhardt D, Mueller-Roeber B, et al. The plant multidrug resistance ABC transporter AtMRPS is involved in guard cell hormonal signaling and water use. Plant J. 2003;33(3):119–29. doi:10.1046/j.1365-313X.2003.016012.x.

37. Gao Y, Biayath RM, Marof MaS, Glover NM, Tucker DM, Buss GR. Validation of Low-Phytate QTLs and Evaluation of Seedling Emergence of Low-Phytate Soybeans. Crop Sci. 2008;48(4):1335. doi:10.2135/cropsci2007.06.01133.

38. Raboy V. Seeds for a better future: low phytate grains help to overcome malnutrition and reduce pollution. Trends Plant Sci. 2001;6(8):458–62. doi:10.1016/S1360-1385(01)01889-7.

39. Zhao HJ, Liu QL, Fu H-W, Xu X-H, Wu D-X, Shu Q-Y. Effect of non-lethal low phytic acid mutations on grain yield and seed viability in field. Field Crop Res. 2008;108(3):206–11. doi:10.1016/j.fcr.2008.05.006.

40. Bowen DE, Souza EJ, Guttieri MJ, Raboy V, Fu J. A Low Phytic Acid Barley Mutation Alters Seed Gene Expression. Crop Sci. 2007;47(5):525–149. doi:10.2135/cropsres2006.07.0565tpg.

41. Glover N. The Genetic Basis of Phytate, Oligosaccharide Content, and Emergence in Soybean. Virginia Tech: Blacksburg, VA, 2011.

42. Schmutz J, Cannon SB, Schlueter J, Ma J, MITOS T, Nelson W, et al. Genome sequence of the paleopolyploid soybean. Nature. 2010;463(7278):178–83. doi:10.1038/nature08670.

43. Chaoch S, Nooter G. Myo-inositol abolishes salicylic acid-dependent cell death and pathogen defence responses triggered by peroxisomal hydrogen peroxide. New Phytol. 2010;188(3):711–8. doi:10.1111/j.1469-8137.2010.03453.x.

44. Meyers BC. Genome-Wide Analysis of NBS-LRR-Encoding Genes in Arabidopsis. Plant Cell Online. 2003;15(8):809–34. doi:10.1105/tpc.030908.

45. Bonardi V, Tang S, Stallmait A, Roberts M, Cherki S, Dangl JL. Expanded functions for a family of plant intracellular immune receptors beyond specific recognition of pathogen effectors. Proc Natl Acad Sci U S A 2003;100(18):9663–8. doi:10.1073/pnas.1131725108.

46. Chini A, Grant JJ, Seki M, Shinozaki K, Loake G. Drought tolerance established by enhanced expression of the CC-NBS-LRR gene, ADR1, requires salicylic acid, peroxide. New Phytol. 2010;188(3):711. doi:10.1111/j.1469-8137.2010.03453.x.

47. Jones SJ, Gonzalez DO, Vodkin LO. Flux of transcript patterns during soybean seed development. BMC Genomics. 2010;11:136. doi:10.1186/1471-2164-11-136.

48. Tanaka R, Tanaka A. Tetrapyrrole biosynthesis in higher plants. Annu Rev Plant Biol. 2007;58:321–48. doi:10.1146/annurev.arplant.57.032905.105448.

49. Bruggeman Q, Raynaud C, Benhamed M, Delaure M. To die or not to die? Lessons from lesion mimic mutants. Front Plant Sci. 2015;6:24. doi:10.3389/fpls.2015.00024.

50. Skonecck JA, Marof MaS, Chang C, Buss GR. Identification of Candidate Gene Mutation Associated With Low Stachyose Phytase in Soybean Line PI200508. Crop Sci. 2000;49(1):247. doi:10.2135/cropsci2000.07.0403.

51. Geider R, Murray AS. The ABC of auxin transport: the role of p-glycoproteins in plant development. FEBS Lett. 2006;580(1):1094–102. doi:10.1016/j.febslet.2005.11.054.

52. Stevinski TL. Diverse functional roles of monosaccharide transporters and their homologs in vascular plants: a physiological perspective. Mol Plant. 2011;4(4):641–62. doi:10.1093/mp/sos051.

53. Derking EC, Bilyeu KD. Association of a Soybean Raffinose Synthase Gene with Low Raffinose and Stachyose Seed Phenotype. Plant Genome J. 2008;1(2):135. doi:10.3835/plantgenome2008.06.0321.

54. Raboy V. Seeds for a better future: low phytate grains help to overcome malnutrition and reduce pollution. Trends Plant Sci. 2001;6(8):458–62. doi:10.1016/S1360-1385(01)01889-7.

55. Andes S, Pyl PT, Huber W. HTSeq – A Python framework to work with high-throughput sequencing data. 2014. doi:10.1101/002824.

56. Andes S, Huber W. Differential expression analysis for sequence count data. Genome Biol. 2010;11(10):R106. doi:10.1186/gb-2010-11-10-106.

57. Hu R, Fan C, Li H, Zhao Q, Fu YF. Evaluation of putative reference genes for gene expression normalization in soybean by quantitative real-time RT-PCR. BMC Mol Biol. 2009;10:93. doi:10.1186/1471-2199-10-93.

58. Pysh LD, Wysocka-Diller JW, Camilleri C, Bouchez D, Benfey PN. The GRAS gene family in Arabidopsis: sequence characterization and basic expression analysis of the SCARECROW-LIKE genes. Plant J. 1999;18(1):111–20. doi:10.1046/j.1365-313X.1999.00431.x.

59. Zhang Z-L, Ogawa M, Fleet CM, Zentella R, Hu J, Heo J-O, et al. SCARECROW-LIKE 3 promotes gibberellin signaling by antagonizing master growth repressor SHORT-ROOT in Arabidopsis. Plant J. 2011;67(4):641–50. doi:10.1111/j.1365-313X.2011.04818.x.

60. Chini A, Fonseca S, Fernandez G, Adie B, Chico JM, Lorenzo O, et al. The JAZ family of repressors is the missing link in jasmonate signalling. Nature. 2007;448(7154):666–71.
79. Wang L, Hua D, He J, Duan Y, Chen Z, Hong X, et al. Auxin Response Factor2 (ARF2) and its regulated homeodomain gene HB33 mediate abscisic acid response in Arabidopsis. PLoS Genet. 2011;7(7), e1002172. doi:10.1371/journal.pgen.1002172.

80. Li S, Zachgo S. TCP3 interacts with R2R3-MYB proteins, promotes flavonoid biosynthesis and negatively regulates the auxin response in Arabidopsis thaliana. Plant J. 2013;76(6):901–13. doi:10.1111/tpj.12348.

81. Herve C, Dabos P, Bardet C, Jaunet A, Auria MC, Ramboer A, et al. In vivo interference with ATCP20 function induces severe plant growth alterations and deregulates the expression of many genes important for development. Plant Physiol. 2009;149(3):1462–77. doi:10.1104/pp.108.126136.

82. Shaked H, Avivi-Ragolsky N, Levy AA. Involvement of the Arabidopsis SWI2/SNF2 chromatin remodeling gene family in DNA damage response and recombination. Genetics. 2006;173(2):985–94. doi:10.1534/genetics.105.051664.

83. Kim JH, Choi D, Kende H. The AtGRF family of putative transcription factors is involved in leaf and cotyledon growth in Arabidopsis. Plant J. 2003;36(1):94–104. doi:10.1046/j.1365-313X.2003.01862.x.

84. Galon Y, Nave R, Boyce JM, Nachmiás D, Knight MR, Fromm H. Calmodulin-binding transcription activator (CAMTA) 3 mediates biotic defense responses in Arabidopsis. FEBS Lett. 2008;582(6):943–8. doi:10.1016/j.febslet.2008.02.037.

85. Di Mauro MF, Iglesias MJ, Arce DP, Valle EM, Arnold RB, Tsuda K, et al. MBF1s regulate ABA-dependent germination of Arabidopsis seeds. Plant Signal Behav. 2012;7(2):188–92. doi:10.4161/psb.18843.