Computational Molecular Biology 2021, Vol.11, No.1, 1-14
http://bioscipublisher.com/index.php/cmb

Review Open Access

Advances in Transcriptome Analyses Using RNA Sequencing Technology in Soybean Plants [Glycine max]

Xiang Jia Min, Mykaela Wagner, Theoni Kasamias
Department of Biological Sciences, Youngstown State University, OH, 44555, P.R. China
Corresponding author email: xmin@ysue.edu

Computational Molecular Biology, 2021, Vol.11, No.1 doi: 10.5376/cmb.2021.11.0001
Received: 21 Apr., 2021
Accepted: 23 Jan., 2021
Published: 25 Feb., 2021

Copyright © 2021 Min et al., This is an open access article published under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Preferred citation for this article:
Min X.J., Wagner M., and Kasamias T., 2021, Advances in transcriptome analyses using RNA sequencing technology in soybean plants [Glycine max], Computational Molecular Biology, 11(1): 1-14 (doi: 10.5376/cmb.2021.11.0001)

Abstract Soybean [Glycine max] is an important oil and food plant for both humans and animals. Recent development in RNA sequencing (RNA-seq) technology provides a cost effective approach to analyzing transcriptomes of plants at different developmental stages and in responses to different biotic and abiotic challenges. Currently there are over 5 000 RNA-seq datasets in soybean plants publicly available at SRA database in the National Center for Biotechnology Information (NCBI). Such a large number of RNA-seq datasets provide soybean researchers an opportunity as well as a challenge for fully exploring the data to understand soybean biology. A number of research articles have been published on applications of RNA-seq in transcriptome analysis of soybean plants, covering a wide range of topics including growth and development, plant mineral nutrients, responses to environmental stresses, pathogens and pests. In this work we compile and review recent advances of RNA-seq transcriptome analyses including profiling of differential gene expression, gene alternative splicing, and gene regulatory networks in soybean plants, with key findings excerpted from each individual published article.

Keywords Soybean; Glycine max; Transcriptome; RNA-sequencing; Gene expression; Alternative splicing

Introduction

Soybean [Glycine max] is an important food and oil crop, playing an important role in solving world hunger problem (Thoenes, 2004). The first soybean genome with 1.1- gigabase pairs was completely sequenced and 46,430 protein coding genes were predicted (current Genebank assembly accession number: GCA_000004515.4) (Schmutz et al., 2010). Recently, three new soybean genomes were completely sequenced (Accession number: AGCA_003349995.1, GCA_002905335.2, GCA_001269945.2) (Shimomura et al., 2015; Shen et al., 2018). The availability of genome sequences from multiple cultivated soybean varieties and undomesticated soybean plants enhances our understanding of the domestication process of soybean plants and soybean breeding (Kim et al., 2011).

RNA sequencing, commonly known as RNA-seq, is a technology using next-generation sequencing (NGS) to identify and quantify RNA transcripts of a transcriptome in a biological sample (Chu and Corey, 2012). Specifically, RNA-seq has a resolution at one base-pair level, thus, it can be used to determine exon/intron boundaries, to identify alternative gene spliced transcripts, and gene expression levels over time or in different treatments. In addition, it can be used to identify the whole population and their abundance of all types of transcripts including mRNAs and non-coding RNAs, such as small RNA, miRNA, etc. (Wang et al., 2009). Therefore, RNA-seq is widely applied for gene expression profiling, identification of differentially expressed genes (DEGs), and identification of alternative splicing (AS). Sequence Read Archive (SRA) database at the National Center for Biotechnology Information (NCBI, https://www.ncbi.nlm.nih.gov) stores RNA-seq data (https://www.ncbi.nlm.nih.gov/sra). We searched the SRA database and found, by the end of January 2021, there were 5 035 RNA-seq datasets, including 2 400 paired and 2 635 single library layouts. In this work, we compile and briefly review the recent advances in transcriptome analysis by using RNA-seq technology in soybean plants, the relevant publications are summarized in Table 1. It should be noted that transcriptome analysis carried out using microarray technology in soybean plants as well as transcriptome analysis using RNA-seq in wild soybean plants are not included in this work.

Table 1 Summary of recent publications employing RNA-seq technology in soybean research

Research Subject	References
Growth and development	
Different tissue	Severin et al. (2010); Wang et al. (2014); Shamimuzzaman and Vodkin (2013); Wong et al. (2013); Wu et al. (2019); Locke et al. (2018); Kim et al. (2016); Li et al. (2019)
Seed development	Jones and Vodkin (2013); Qi et al. (2018); Shi et al. (2014); Schmidt et al. (2011); Collakova et al. (2013); Aghamirzaie et al. (2013); Lambirth et al. (2015); Goettel et al. (2014); Yang et al. (2019); Yu et al. (2019); Redekar et al. (2015); Du et al. (2017); Liu et al. (2015)
Seed germination	Bellieny-Rabelo et al. (2016); Wang et al. (2016)
Application in breeding	Guo et al. (2018); Zhang et al. (2017a); Li et al. (2015); Song et al. (2016a)
Response to abiotic stress	
Water deficit, waterlogging, or heat	Prince et al. (2015); Song et al. (2016b); Wang et al. (2018); Nakayama et al. (2017); Chen et al. (2016b); Rodrigues et al. (2015)
Shade	Gong et al. (2014)
Nutrients, toxic metals, and other chemicals	Xue et al. (2018); Zeng et al. (2018b); Mo et al. (2019); Lv et al. (2020); Zeng et al. (2018a); Zhang et al. (2017b); Zeng et al. (2019); Whaley et al. (2015); Yang et al. (2018); Wulff et al. (2019); Zhao et al. (2020)
Responses to nodulation or biotic stresses	
Nodule development	Yuan et al. (2016); Yuan et al. (2017)
Virus	Chen et al. (2016a); Zhang et al. (2019); Jo et al. (2017)
Pests: aphid, nematode and pyralid	Neupane et al. (2019a); Neupane et al. (2019b); Song et al. (2019); Zeng et al. (2017); Zeng et al. (2019)
Fungal disease	Dong et al. (2018); Lanubile et al. (2015); McCabe et al. (2018); Calla et al. (2014)
Bacterial diseases	Kim et al. (2011)

1 Growth and Developmental Regulation

1.1 Different tissues during growth and development

The report by Severin et al. (2010) represents an earlier application in soybean transcriptome analysis using RNA-seq. Hierarchical clustering of transcriptional profiles in fourteen diverse tissues reveals three clades with similar profiles: aerial, underground and seed tissues and more than 177 genes that are involved in the economically important seed filling process were identified (http://www.soybase.org/) (Severin et al., 2010). Wang et al. (2014) collected data from 11 meristematic tissues of soybean and found 54 132 annotated genes expressed with a total of 12 810 AS events generated from 7 084 genes, accounting for ~15.9% of 52 406 genes with two or more exons.

Transcriptional reprogramming occurs during the functional transition of cotyledons from non-photosynthetic storage tissue to metabolically active photosynthetic tissue (Shamimuzzaman and Vodkin, 2013). RNA-seq data revealed highly differentially expressed candidate genes regulated by the NAC transcription factor include lipoxygenase, pectin methyl esterase inhibitor, DEAD/DEAH box helicase and homeobox associated proteins, and YABBY-regulated genes include AP2 transcription factor, fatty acid desaturase and WRKY transcription factor.

Wong et al. (2013) investigated the molecular basis of floral initiation process using RNA-seq. A total of 2 951 genes in shoot apical meristem and 13 609 genes in leaf with significant profile changes were identified. Most changes in mRNA level occurred after 1short-day treatment. Wu et al. (2019) used two common cultivars and four near-isogenic lines of maturity loci of soybean plants in elucidating global gene expression patterns under different photoperiod regimes. Their work provides novel insights in the complex molecular mechanisms of photoperiodic flowering control in soybean. Locke et al. (2018) examined the diurnal patterns of genes involved in photosynthesis-related processes in field-grown soybean transcriptomes and reported 21% of genes were differentially expressed over the course of the day.

To understand the molecular events in abscission process in soybean, Kim et al. (2016) performed RNA-seq analysis in abscission zones and after treating stem/petiole explants with ethylene. Their work revealed that 188 abscission-specific TFs include several TFs containing domains for homeobox, MYB, Zinc finger, bHLH, AP2, NAC, WRKY, YABBY, and auxin-related motifs. Li et al. (2019) identified DEGs in soybean somatic embryogenesis and regeneration processes. Exogenous hormone 2,4-D was used to treat immature embryos. There were 2 666 differentially expressed genes obtained including three genes in hormone pathways. These identified important genes in the soybean regeneration process provided a basis for accelerating the application of biotechnology to soybean for improving its breeding efficiency.

1.2 Seed development

Soybean seeds are an important source of proteins, oil, and carbohydrates used for food, feed, chemical, and biofuel production, thus studying the gene expression profiles in seed development is important for understanding this process. Jones and Vodkin (2013) examined gene expression in seven different stages of seed development using RNA-seq. Over one hundred gene models were identified with high expression exclusively in young seed stages, starting at four days after fertilization. Genes encoding storage proteins such as glycinin and beta-conglycinin reached to their highest expression levels at the stages of largest fresh weight, confirming these storage products are being rapidly accumulated before the seed begins the desiccation process. Qi et al. (2018) profiled the time-course transcriptome patterns of soybean seeds at early maturity, middle maturity, and dry seed stages. Their analysis revealed 7 482 differentially expressed genes and 45 expression patterns clusters.

Soybean cotyledon is the nutrient storage area in seeds and is critical for seed quality and yield. Cotyledon mutants are important for the genetic dissection of embryo patterning and seed development. Shi et al. (2014) characterized a soybean curled cotyledon (cco) mutant. RNA-seq identified 1 093 differentially expressed genes (DEGs) between WT and the cco mutant, and many DEGs were mapped to the hormone biosynthesis and signal transduction pathways. Schmidt et al. (2011) analyzed the storage protein knockdown (SP-) seeds and the wild type seeds. RNA-seq analysis identified 997 were down-regulated and 151 were up-regulated genes with more than 3-fold changes in SP- compared with wildtype seeds (Schmidt et al., 2011). Collakova et al. (2013) assessed detailed temporal transcriptional and metabolic changes in developing soybean embryos, and further using the same dataset, Aghamirzaie et al. (2013) identified different classes of alternatively spliced isoforms and corresponding changes in their levels. Conserved domain analyses revealed that AS resulted in global changes in the number, types, and extent of truncation of functional domains in protein variants (Aghamirzaie et al., 2015). Lambirth et al. (2015) used RNA-seq to survey gene expression in three transgenic soybean lines expressing recombinant protein in seed tissues. There was no correlation between recombinant protein expression level and the quantity of differentially expressed genes detected.

Oil content in soybean seeds is an important breeding parameter. Goettel et al. (2014) sequenced the transcriptomes of soybean seeds from nine lines varying in oil composition and/or total oil content. A total of 8 037 transcript expression polymorphisms and 50,485 transcript sequence polymorphisms (48 792 SNPs and 1 693 small Indels) were identified among the lines. Similarly, Yang et al. (2019) performed transcriptome analysis on oil contents in developing seeds. Through weighted correlation network analysis, six co-expression modules associated with soybean seed oil content and 124 candidate genes potentially affecting soybean seed oil content were identified (Yang et al. 2019). Yu et al. (2019) selected four lines with extreme phenotypes, as well as high or low protein and oil content, from the chromosome segment substitution lines. Five specific genes were identified to be differentially expressed during seed development and subsequently analyzed for their regulatory relationship with miRNAs. The research provides a theoretical basis for the regulation of soybean quality traits. Redekar et al. (2015) compared gene expression profiles during seed development between low phytic acid lines (lpa) and a normal phytic acid line (PA). A total of 4 235 DEGs, including 512-transcription factor genes were identified. Their study provides a global perspective of transcriptional changes during soybean seed development in an lpa mutant.

To understand the gene expression networks controlling soybean seed set and size, Du et al. (2017) performed transcriptome analyses in three early seed developmental stages, using two genotypes with contrasting seed size. They identified modules of coexpressed genes and hub genes for each module and discovered specific modules for the large seed size variety and for seed developmental stages. This work provided new insights into the molecular networks underlying soybean seed size regulation. Seed number per pod is an important component of yield traits in soybean. Liu et al. (2015) identified a natural mutant with an increased number of four-seed pods from a cultivated soybean variety. To identify and understand the function of genes associated with this mutant trait, transcriptome sequencing was performed using three types of tissues: axillary buds, unfertilized ovaries, and young pods at three different growth stages, respectively. They reported there were 55 582 expressed genes and 4 183 DEGs. These results show that the development of four-seed pods is associated with a complex network involving multiple physiological and metabolic pathways. However, the exact mutations remain to be identified.

1.3 Seed germination

Seed germination is a developmental process that produces a new plant from a seed. Bellieny-Rabelo et al. (2016) examined the transcriptional patterns of embryonic axes during germination using RNA-seq reads generated from soybean embryonic axes. Their results support glycolysis, Krebs cycle and cell wall re-modelling are critical processes. Effects of calcium on yield and nutritional qualities of soybean sprout were investigated (Wang et al., 2016). A total of 1 912 genes and 460 proteins involved in the growth, storage material decomposition, and bioactive substance synthesis in soybean sprouts after treated with Ca^{2+} were identified. This study represents the first report of a comprehensive transcriptomic and proteomic analysis of soybean sprout in response to supplemental Ca^{2+}.

1.4 Application in soybean breeding

Guo et al. (2018) reported that RNA-seq data generated from four soybean accessions. A total of 75 209 SNPs were identified, 89.1% of which were located in expressed regions and 27.0% resulted in amino acid changes. The validated SNPs were assessed by genotyping 393 wild and cultivated soybean accessions. These SNPs identified by RNA-seq provide a useful resource for genetic and genomic studies of soybean.

RNA-seq has also been used for studying gene expression in F1 hybrids and their parents in soybean as heterosis has been widely used for increasing crop yield in plant breeding. Zhang et al. (2017a) identified a total of 681 and 899 genes as being differentially expressed between two F1 hybrid lines and their parents, respectively; and 26 common DEGs that showed transgressive up-regulation, representing potential candidate genes for heterosis in soybean F1 hybrids (Zhang et al., 2017a). Cytoplasmic male sterility (CMS) plays an important role in the production of hybrid seeds. However, the molecular mechanism of CMS in soybean remains unclear. Li et al. (2015) reported the comparative transcriptome analysis between cytoplasmic male sterile line NJCMS1A and its near-isogenic maintainer NJCMS1B in soybean. A total of 56 044 genes were mapped to the genome with 339 down-regulated and 26 up-regulated in CMS line compared to the maintainer line. In addition, RNA profiles were compared in a soybean near-isogenic line (cgy-2-NIL) containing the cgy-2 allele, which did not encode the allergenic α-subunit of β-conglycinin, identified 2 193 genes down regulated and 1 350 genes up regulated, in two genotypes, including genes involved in arginine and proline metabolism, which is useful for hypoallergenic soybean breeding programs (Song et al., 2016a).

2 Responses to Abiotic Stress

2.1 Water deficit, waterlogging, or heat

Environmental factors such as water deficit or waterlogging, heat, cold, and nutrient deficits are common for causing stress in field grown crops. Prince et al. (2015) analyzed the leaf tissue transcriptomes of two soybean lines: drought-susceptible (DS) Pana and drought-tolerant (DT) PI 567690. They reported 1 914 in DS and 670 genes in DT with a greater than two-fold change in expression under drought conditions. Song et al. (2016b) reported 6 609 genes in the roots showed differential expression patterns in response to different water-deficit stress levels in soybean Williams 82 genotype including tissue-specific or water-deficit levels specific regulation

of TFs (Song et al., 2016b). Wang et al. (2018) determined the transcriptional responses of soybean to heat, drought and combined stress using RNA-seq. They reported that 1 468 and 1 220 up-regulated and 1 146 and 686 down-regulated genes at 8 h and 24 h after treatment. Comparing the heat, drought and the combined stress-responsive genes identified potential new targets for enhancing stress tolerance of soybean. Their analysis suggests that a synergistic response to drought and heat may exist in soybean. Song et al. (2020) performed a genome-wide analysis of AS events in soybean roots grown under various drought conditions and identified 385, 989, 1 429, and 465 AS events that were significantly differentially spliced under very mild drought stress, mild drought stress, severe drought stress, and recovery after severe drought conditions, respectively.

Nakayama et al. (2017) analyzed root transcriptional responses between flood-tolerant (Embrapa 45) and flood-sensitive (BR4) soybean cultivars under hypoxic stress conditions with RNA-seq. Between cultivars, Embrapa 45 showed less up- and more down-regulated genes. Over 6 600 to 8 300 genes were up-regulated and 8 100 to 11 000 genes were down-regulated in roots sampled at different time over a 24-hour period of submergence treatment. All genes encoding the key enzymes in the glycolysis/gluconeogenesis pathway were up-regulated. Among the down-regulated genes, eight of them encoded peroxidases, which catalyze the conversion of coumaroyl alcohol to hydroxy-phenyl lignin in the final step of lignin biosynthesis. The study provides a foundation for further genomic examination of submergence or flooding tolerance in soybean plants.

Chen et al. (2016b) identified a total of 2 724 and 3 498 DEGs in leaf tissues were identified under drought and flooding treatments, respectively. This study has revealed the involvement of TFs, transporters, and photosynthetic genes. Rodrigues et al. (2015) reported soybean leaf transcriptome fluctuations under water deficit stress during a 24-h time course. There were 4 866 DEGs in soybean plants in response to water deficit. Of these genes, 3 715 were differentially expressed during the light period, from which approximately 9.55% were observed in both light and darkness. There were 887 genes either up- or down-regulated in different periods of the day. Their work identified putative cycling genes that are expressed in soybean leaves under normal developmental conditions and genes whose expression oscillates under conditions of water deficit.

2.2 Shade

Soybean is often planted in intercropping, such as with maize, for fully utilizing the land and light resources, thus studying shading effects is needed in such a system. Gong et al. (2014) analyzed the transcriptome of shaded soybean leaves via RNA-seq technology. They reported that transcription of 1,085 genes in mature leaves and 1 847 genes in young leaves were significantly affected by shade.

2.3 Nutrients and toxic metals

Phosphorus (P) is an essential plant macronutrient. Low phosphate (Pi) availability in soils often limits plant growth and yield. Xue et al. (2018) conducted RNA-seq analysis in soybean nodules grown under P-sufficient and P-deficient conditions. Nodule transcript profiling revealed that a total of 2 055 genes exhibited differential expression patterns between Pi sufficient and deficient conditions. Their results suggest that a complex transcriptional regulatory network participates in soybean nodule adaption to Pi starvation. Zeng et al. (2018b) performed a genome-wide transcriptomic analysis in soybean leaves treated with a short-term Pi-deprivation (24 h). A total of 533 genes were found to be differentially expressed in response to Pi deprivation, including 303 up regulated and 230 repressed by Pi deprivation. Mo et al. (2019) reported, between two Pi availability treatments, a total of 155 metabolites differentially accumulated in soybean roots, and a total of 1 644 DEGs were identified in soybean roots, including 1 199 up-regulated and 445 down-regulated genes. Lv et al. (2020) examined the response mechanisms of circular RNAs (circRNAs) to low Pi (LP) stress using the roots of two contrasting soybean genotypes, Bogao (a LP-sensitive genotype) and Nannong 94156 (a LP-tolerant genotype). In total, 371 novel circRNA candidates, including 120 significantly differentially expressed (DE) circRNAs, were identified across different P levels and genotypes. Integrating all transcriptome data generated under P deficiency treatments may find consistent patterns of gene regulation across these multiple experiments, thus, increase our understanding of the regulation mechanism of plant responses.

Zinc (Zn) is an essential micronutrient in plants. Zeng et al. (2018a) analyzed the transcriptomes of soybean leaves and roots in response to Zn deficiency through RNA-seq. A total of 614 and 1 011 genes were differentially expressed in leaves and roots, respectively, and 88 genes were common in both leaves and roots. Among the DEGs, 20 DEGs are potentially involved in Zn homeostasis and nine DEGs were found to contain zinc-deficiency-response element in their promoter regions. Their results could provide comprehensive insights into the soybean response to Zn deficiency. Cadmium (Cd) is a common pollutant heavy metal. Zhang et al. (2017b) investigated the molecular mechanism of soybean root responding to Cd stress. Soybean seedlings (7-day old) were stressed by Cd (75 Mmol·L-1) for 0, 4, 8, 12 and 48 h. Comparative transcriptome analysis showed 244, 1545, 442 and 1401 of genes responded to the four Cd treatments, respectively, and total 2 670 DEGs were identified. Aluminum (Al) toxicity is a major problem on acid soils. Zhao et al. (2020) carried out comparative transcriptome analysis of two contrasting soybean varieties in response to Al stress. They identified a total of 354 Al-tolerance related genes, showed up-regulated expression by Al in the Al-tolerant soybean variety.

Salt stress causes stunting growth, biomass and yield reduction in soybean. Zeng et al. (2019) used salt tolerant and sensitive soybean lines treated with 250 mM NaCl. A total of 2 374, 998, 1 746, and 630 DEGs between salt-tolerant line and salt-sensitive line, were found at 0, 6, 12, and 24 h after treatment, respectively. The expression patterns of 154 common DEGs among all the time points were examined.

2.4 Other chemicals

Whaley et al. (2015) reported RNA-seq analysis of two soybean cultivars including an ozone intolerant cultivar (Mandarin-Ottawa) and an ozone resistant cultivar (Fiskeby III) following exposure to ozone. Acid rain influences soybean growth and productivity. Yang et al. (2018) conducted RNA-seq to examine changes in gene expression when soybean was exposed to simulated acid rain (SAR). In total, 54 175 expression genes were found, including 2 016 DEGs. A total of 416 genes were considered to be closely related to the response to SAR. These genes related to the regulation of sulfur and nitrogen metabolism, carbohydrate metabolism, photosynthesis, and reactive oxygen species were among those DEGs in response to SAR.

Neonicotinoids are widely used systemic insecticides that have been associated with spider mite outbreaks on diverse plants. Wulff et al. (2019) used RNA-seq to explore how neonicotinoids modify gene expression in soybean by exposing soybean to two neonicotinoid insecticides, thiamethoxam applied to seeds and imidacloprid applied as a soil drench and exposing a subset of these plants to spider mites (*Tetranychus cinnabarinus*). Applications of both insecticides downregulated genes involved in plant-pathogen interactions, phytohormone pathways, phenylpropanoid pathway, and cell wall biosynthesis. These effects were especially evident in plants exposed to thiamethoxam.

3 Responses to Nodulation or Biotic Stress

3.1 Nodule development

Soybean plants establish a symbiosis with nitrogen fixing bacteria by developing nodules. This process plays a significant agronomic and ecological role in soybean production. Yuan et al. (2016) reported differential gene expression responding to different rhizobium strains in soybean roots. Two symbiotic systems (*Bradyrhizobium japonicum* strain 113-2-soybean and *Sinorhizobium fredii* USDA205-soybean) with notable different nodulation phenotypes were studied. The DEGs uncovered in this study provides molecular candidates for better understanding the mechanisms of symbiotic host-specificity and explaining the different symbiotic effects between soybean roots inoculated with different strains. Yuan et al. (2017) further investigated differential gene expression of root nodules at five developmental stages inoculated with *B. japonicum* strain 113-2. Genes primarily encoding symbiotic nitrogen fixation-related proteins, cysteine proteases, cystatins and cysteine-rich proteins, as well as proteins involving plant-pathogen interactions were identified. Their study provides molecular material for further investigations into the mechanisms of nitrogen fixation at different soybean developmental stages.

3.2 Virus

Plant virus may affect plant growth using a regulation mechanism via microRNAs and associated host target genes. Chen et al. (2016a) performed small RNA (sRNA)-seq, degradome-seq and as well as a genome-wide transcriptome analysis to profile the global gene and miRNA expression in soybean following infections by three different Soybean mosaic virus (SMV) isolates. Their analyses revealed a total of 253 soybean miRNAs and 125 transcripts as potential miRNA targets, as well as total 2 679 DEGs in response to SMV infection. Zhang et al. (2019) compared the transcriptomes of SMV infected plants and reported a total of 3 548 and 4 319 DEGs in soybean plants infected with SMV under normal light and in the shade, respectively.

Jo et al. (2017) used publicly available soybean transcriptome data to identify viruses infecting soybean. Of the screened transcriptomes, a soybean transcriptome for soybean seed development analysis was found containing several virus-associated sequences. Five viruses, including SMV, infecting soybean were identified by de novo transcriptome assembly. A near complete consensus genome sequence of SMV China was assembled using transcriptome data. Based on phylogenetic analysis, the consensus genome sequence of SMV China was closely related to SMV isolates from South Korea. Their work demonstrated the application of soybean transcriptome data for virus genome assembly.

3.3 Pests: aphid, nematode and pyralid

Neupane et al. (2019b) investigated transcriptome responses of soybean plants colonized by soybean aphid (*Aphis glycines* Matsumura; SBA) including avirulent and virulent biotypes after 11 days of feeding. Ten RNA datasets are reported with 266 million sequence reads (55.2 GB) obtained from pooled samples derived from the leaves collected at day 1 and day 11 post SBA infestation. However, the data have not been thoroughly analyzed. Neupane et al. (2019a) generated RNA-seq data from soybean cyst nematode (*Heterodera glycines*; SCN) -resistant and SCN-susceptible soybean cultivars with a combination of soybean aphid infection treatment to evaluate the three-way interactions. Song et al. (2019) used soybean cultivar Lee, which exhibits resistant to SCNT (a new type of SCN) and susceptible to HG 1.2.3.4.7 (SCNs) and identified 3 746 and 602 DEGs in the resistant interaction and in the susceptible interaction, respectively.

Bean pyralid is one of the major soybean leaf-feeding insects in China. Zeng et al. (2017) identified 1 744 DEGs in the leaves of two different varieties fed by bean pyralid for 48 h, compared to 0 h. Zeng et al. (2019) performed small RNA and transcriptome sequencing in bean pyralid larvae resistant and susceptible soybean varieties, and differentially expressed miRNAs and their target genes were identified in both varieties in comparing controls with larvae feeding treatments as well as between varieties.

3.4 Fungal disease

Soybean downy mildew (SDM) caused by *Peronospora manshurica* (Pm) is a common disease of soybean. To examine the defense mechanism of soybean response to Pm infection, differential expression of WRKY TFs in SDM-high resistant (HR) and SDM-high susceptible (HS) genotypes were analyzed (Dong et al., 2018). Total 22 WRKY TFs were differentially expressed in HR and HS genotype, while 16 WRKY TFs were found to be specific in response to fungal inoculation. Their result indicates that the GmWRKY31 might regulate the GmSAGT1 gene expression and involve in SA-mediated immune responses in soybean.

Fusarium oxysporum is one of the most common fungal pathogens in soybean. Lanubile et al. (2015) investigated the molecular aspects of the interactions of a partially resistant soybean genotype with non-pathogenic/pathogenic isolates of *F. oxysporum* at 72 and 96 h post inoculation (hpi). A peak of highly differentially expressed genes (HDEGs) was triggered at 72 hpi in soybean roots and the number of HDEGs was about eight times higher in response to the pathogenic isolate compared to the non-pathogenic one (1,659 vs. 203 HDEGs, respectively). Their data may be useful in the developing new methods of broadening resistance of soybean to *F. oxysporum*. Brown stem rot, a fungal disease caused by *Phialophora gregata*, reduces soybean yield. McCabe et al. (2018) conducted RNA-seq of infected and mock-infected leaf, stem, and root tissues of a resistant (PI 437970, Rbs3)

and susceptible (Corsoy 79) genotype. A cluster of genes encoding receptor-like proteins were identified as candidates for the Rbs3 resistance gene. Their research demonstrates the utility of combining contrasting genotypes, gene expression, and classical genetic studies to characterize complex disease resistance loci.

Sclerotinia sclerotiorum is a serious fungal pathogen. The major virulence factor of this pathogen is oxalic acid (OA). Plants that express enzymes that degrade OA, such as oxalate oxidase (OxO), are very resistant to *S. sclerotiorum*. Calla et al. (2014) compared the gene expression levels between leaves of a transgenic soybean carrying an OxO gene (OxO) and its parent AC Colibri (AC) using both microarray and RNA-seq technology. A total of 936 DEGs were identified, and among them, 630 showed up-regulation in both AC and OxO lines infiltrated with OA, relative to the samples infiltrated with water, at pH 2.4. Ferritin, a gene that encodes for an iron storage protein, was one of the strongest genes induced by OA. These results suggest that *S. sclerotiorum* benefits from the ability of OA to free iron from plant proteins by inducing host cell death (Calla et al., 2014).

3.5 Bacterial diseases

Xanthomonas axonopodis pv. glycines (Xag) causes bacterial leaf pustule (BLP) disease in soybean. Kim et al. (2011) investigated the plant defense mechanisms against Xag using BLP-susceptible and BLP-resistant soybean near isogenic lines (NIL) and reported 1 978 and 783 genes were up- or down-regulated, respectively, in the BLP-resistant NIL relative to the BLP-susceptible NIL. Pathogen-associated molecular pattern (PAMP) and damage-associated molecular pattern (DAMP) receptors and the genes induced by these receptors were highly expressed in the BLP-resistant NIL. Their results indicate that the defense mechanisms involved in the recognition of PAMPs or DAMPs and a high level of accumulation of defense-related gene products may contribute to BLP resistance in soybean.

4 Applications of RNA-seq Data

Yu et al. (2014) constructed the Soybean Functional Genomics Database (SFGD) using Generic Genome Browser (Gbrowse) as the core platform by integrating microarray expression profiling with 255 samples from 14 groups' experiments and mRNA-seq data with 30 samples from four groups' experiments, including spatial and temporal transcriptome data for different soybean development stages and environmental stresses. The SFGD includes a gene co-expression regulatory network containing 23 267 genes and 1 873 miRNA-target pairs, and a group of acyl-lipid pathways containing 221 enzymes and more than 1 550 genes. The SFGD is publically accessible (http://bioinformatics.cau.edu.cn/SFGD/). Using the gene expression values of RNA-seq data generated by the Soybean Gene Atlas project, Wang et al. (2019a) constructed a database of gene regulatory network which was named as SoyCSN (Soybean context-specific network, http://soykb.org/SoyCSN). The database allows users to search and download co-expressed genes and predicted protein-protein interactomes across all tissues. It is a useful resource for systematically exploring gene regulatory mechanisms and gene relationships for soybean researchers and molecular breeders. The homeodomain leucine zipper (HD-Zip) transcription factor family is one of the largest plant specific super families including genes with roles in modulation of plant growth and response to environmental stresses. Belamkar et al. (2014) identified 101 members of the HD-Zip gene family in soybean cv. 'Williams 82', and characterized their expression under dehydration and salt stress. Recently, Machado et al. (2020) analyzed 1 298 RNA-Seq samples and constructed a comprehensive soybean (Glycine max) expression atlas. They found that 94% of the annotated genes (52 737/56 044) had detectable expression in at least one sample. The dataset can be downloaded or accessed through a user-friendly web interface (http://venanciogroup.uenf.br/resources/). This transcriptome atlas will facilitate research on soybean genetics and genomics. Almeida-Silva et al. (2020) investigated global gene co-expression networks constructed with data from 1 284 RNA-Seq experiments from 15 distinct tissues and predicted functions for 93 of 106 hubs without functional description in soybean.

Publicly available RNA-seq datasets are rich resources for datamining of genes with specific functionalities. Nawaz et al. (2017) performed genome-wide digital expression analysis of cellulose synthase (CS) gene superfamily, consisting of 78 putative CS genes. Liew et al. (2013) analyzed the RNA-seq data, which were

generated by Wong et al. (2013), for identification of histone modifiers and RNA silencing genes in soybean. Their analysis reveals that the extensive activation of genes that are usually involved in the epigenetic programming and RNAi gene silencing in the soybean shoot apical meristem are reprogrammed for floral development following an exposure to inductive conditions. Zhu et al. (2013) developed an automated bioinformatics method to predict gene regulatory networks from the quantitative expression values of DEGs based on RNA-seq transcriptome data of a cell in different stages and conditions, integrating transcriptional, genomic and gene function data. They applied the method to the RNA-seq transcriptome data generated for soybean root hair cells in three different development stages of nodulation after rhizobium infection (Libault et al., 2010).

Mostly likely we have missed out some publications related RNA-seq applications in soybean research, however, we noticed that most of transcriptome analyses, as reviewed above, performed identification of DEGs, GO analysis and pathway analysis. Only a few of published articles carried out AS analysis (Wang et al., 2014; Aghamirzaie et al., 2015; Wang et al., 2019b; Song et al. 2020). Given the importance of AS in regulation of gene functions in growth and development and in response to various biotic and abiotic stresses (Reddy et al., 2013; Staiger and Brown, 2013), AS analysis needs to be included as an essential part of RNA-seq data analysis in plants. Thus, such a large amount of RNA-seq datasets publicly available in soybean plants provide an unprecedented opportunity for identifying alternatively spliced genes and novel transcripts isoforms generated by AS in this plant. Further integration of all available RNA-seq data for advancing our understanding of gene expression regulatory networks remains to be a challenge, however, will be warranted to greatly benefit soybean plant researchers.

Authors' contributions

XM conceived the study and prepared the manuscript. MW and TK collected the literature, collected RNA-seq data and performed data analysis. All authors read and approved the final manuscript.

Acknowledgements

The work is supported by the Youngstown State University Research Professorship award to XM. The Ohio Supercomputer Center is acknowledged for providing computational resources for the project.

References

Aghamirzaie D., Batra D., Heath L.S., Schneider A., Grene R., and Collakova E., 2015, Transcriptome-wide functional characterization reveals novel relationships among differentially expressed transcripts in developing soybean embryos, BMC Genomics, 16(928): 1-23
https://doi.org/10.1186/s12864-015-2108-x
PMid:26572793 PMCid:PMC4647491

Aghamirzaie D., Nabiyouni M., Fang Y., Klumas C., Heath L.S., Grene R., and Collakova E., 2013, Changes in RNA splicing in developing soybean (Glycine max) embryos, Biology, 2(4): 1311-1337
https://doi.org/10.3390/biology2041311
PMid:24833227 PMCid:PMC4009788

Almeida-Silva F., Moharana K.C., Machado F.B., et al., 2020, Exploring the complexity of soybean (*Glycine max*) transcriptional regulation using global gene co-expression networks, Planta, 252(6): 104
https://doi.org/10.1007/s00425-020-03499-8
PMid:33196909

Belamkar V., Weeks N.T., Bharti A.K., et al., 2014, Comprehensive characterization and RNA-seq profiling of the HD-Zip transcription factor family in soybean (*Glycine max*) during dehydration and salt stress, BMC Genomics, 15(1): 950
https://doi.org/10.1186/1471-2164-15-950
PMid:25362847 PMCid:PMC4226900

Bellieny-Rabelo D., De Oliveira E.A., da Silva Ribeiro E., et al., 2016, Transcriptome analysis uncovers key regulatory and metabolic aspects of soybean embryonic axes during germination, Scientific Reports, 6: 36009
https://doi.org/10.1038/srep36009
PMid:27824062 PMCid:PMC5099898

Calla B., Blahut-Beatty L., Koziol L., et al., 2014, Transcriptome analyses suggest a disturbance of iron homeostasis in soybean leaves during white mould disease establishment, Molecular Plant Pathology, 15(6): 576-588
https://doi.org/10.1111/mpp.12113
PMid:24330102 PMCid:PMC6638882

Chen H., Arsovski A.A., Yu K., and Wang A., 2016a, Genome-wide investigation using sRNA-seq, degradome-seq and transcriptome-seq reveals regulatory networks of microRNAs and their target genes in soybean during Soybean mosaic virus infection, PLoS One, 11(3): e0150582
https://doi.org/10.1371/journal.pone.0150582
PMid:26963095 PMCid:PMC4786119

Chen W., Yao Q., Patil G.B., et al., 2016b, Identification and comparative analysis of differential gene expression in soybean leaf tissue under drought and flooding stress revealed by RNA-seq, Frontiers in Plant Science, 7: 1044
https://doi.org/10.3389/fpls.2016.01044

Chu Y., and Corey D.R., 2012, RNA sequencing: platform selection, experimental design, and data interpretation. Nucleic Acid Therapeutics, 22(4): 271-274
https://doi.org/10.1089/nat.2012.0367
PMid:22830413 PMCid:PMC3426205

Collakova E., Aghamirzaie D., Fang Y., et al., 2013, Metabolic and transcriptional reprogramming in developing soybean (Glycine max) embryos, Metabolites, 3: 347-372 347-372
https://doi.org/10.3390/metabo3020347
PMid:24957996 PMCid:PMC3901275

Dong H., Tan J., Li M., et al., 2018, Transcriptome analysis of soybean WRKY TFs in response to *Peronospora manshurica* infection, Genomics, 111(6): 1412-1422
https://doi.org/10.1016/j.ygeno.2018.09.014
PMid:30267765

Du J., Wang S., He C., et al., 2017, Identification of regulatory networks and hub genes controlling soybean seed set and size using RNA sequencing analysis., J. Experimental Botany, 68(8): 1955-1972
https://doi.org/10.1093/jxb/erw460
PMid:28087653 PMCid:PMC5429000

Goettel W., Xia E., Upchurch R., et al., 2014, Identification and characterization of transcript polymorphisms in soybean lines varying in oil composition and content, BMC Genomics, 15(1): 299
https://doi.org/10.1186/1471-2164-15-299
PMid:24755115 PMCid:PMC4023607

Gong W., Qi P., Du J., et al., 2014, Transcriptome analysis of shade-induced inhibition on leaf size in relay intercropped soybean, PLoS One, 9(6): e98465
https://doi.org/10.1371/journal.pone.0098465
PMid:24886785 PMCid:PMC4041726

Guo Y., Su B., Tang J., et al., 2018, Gene-based SNP identification and validation in soybean using next-generation transcriptome sequencing, Molecular Genetics and Genomics, 293(3): 623-633
https://doi.org/10.1007/s00438-017-1410-5
PMid:29280001

Jo Y., Choi H., Bae M., et al., 2017, De novo genome assembly and single nucleotide variations for soybean mosaic virus using soybean seed transcriptome data, The Plant Pathology J., 33(5): 478
https://doi.org/10.5423/PPJ.OA.03.2017.0060
PMid:29018311 PMCid:PMC5624490

Jones S.I., and Vodkin L.O., 2013, Using RNA-seq to profile soybean seed development from fertilization to maturity, PloS one, 8(3): e59270
https://doi.org/10.1371/journal.pone.0059270
PMid:23555009 PMCid:PMC3598657

Kim J., Yang J., Yang R., et al., 2016, Transcriptome analysis of soybean leaf abscission identifies transcriptional regulators of organ polarity and cell fate, Frontiers in Plant Science, 7: 125
https://doi.org/10.3389/fpls.2016.00125

Kim K.H., Kang Y.J., Kim D.H., et al., 2011, RNA-seq analysis of a soybean near-isogenic line carrying bacterial leaf pustule-resistant and-susceptible alleles, DNA Res., 18(6): 483-497
https://doi.org/10.1093/dnares/dsr033
PMid:21987089 PMCid:PMC3223079

Lambirth K.C., Whaley A.M., Blakley I.C., et al., 2015, A comparison of transgenic and wild type soybean seeds: analysis of transcriptome profiles using RNA-Seq, BMC Biotechnology, 15(1): 89
https://doi.org/10.1186/s12896-015-0207-z
PMid:26427366 PMCid:PMC4591623

Lanubile A., Muppirala U.K., Severin A.J., et al., 2015, Transcriptome profiling of soybean (Glycine max) roots challenged with pathogenic and non-pathogenic isolates of *Fusarium oxysporum*, BMC Genomics, 16(1): 1089
https://doi.org/10.1186/s12864-015-2318-2
PMid:26689712 PMCid:PMC4687377

Li J., Han S., Ding X., et al., 2015, Comparative transcriptome analysis between the cytoplasmic male sterile line NJCMS1A and its maintainer NJCMS1B in soybean (Glycine max (L.) Merr.), PLoS One, 10(5): e0126771
https://doi.org/10.1371/journal.pone.0126771
PMid:25985300 PMCid:PMC4436259

Li S.N., Cheng P., Bai Y.Q., et al., 2019, Analysis of soybean somatic embryogenesis using chromosome segment substitution lines and transcriptome sequencing, Genes, 10(11): 943
https://doi.org/10.3390/genes10110943
PMid:31752416 PMCid:PMC6896167

Libault M., Farmer A., Joshi T., et al., 2010, An integrated transcriptome atlas of the crop model *Glycine max*, and its use in comparative analyses in plants, Plant J., 63(1): 86-99
https://doi.org/10.1111/j.1365-313X.2010.04222.x
PMid:20408999

Liew L.C., Singh M.B. and Bhalla P.L., 2013, An RNA-seq transcriptome analysis of histone modifiers and RNA silencing genes in soybean during floral initiation process, PLoS One, 8(10): e77502
https://doi.org/10.1371/journal.pone.0077502
PMid:24147010 PMCid:PMC3797736

Liu Z.Z., Yao D., Zhang J., et al., 2015, Identification of genes associated with the increased number of four-seed pods in soybean (*Glycine max* L.) using transcriptome analysis, Genetics Molec. Res., 14(4): 18895-18912
https://doi.org/10.4238/2015.December.28.39
PMid:26782540

Locke A.M., Slattery R.A., and Ort D.R., 2018, Field-grown soybean transcriptome shows diurnal patterns in photosynthesis-related processes, Plant Direct, 2(12): e00099
https://doi.org/10.1002/pld3.99
PMid:31245700 PMCid:PMC6508813

Lv L., Yu K., Lü H., et al., 2020, Transcriptome-wide identification of novel circular RNAs in soybean in response to low-phosphorus stress, PloS one, 15(1): e0227243
https://doi.org/10.1371/journal.pone.0227243
PMid:31961887 PMCid:PMC6974154

Machado F.B., Moharana K.C., Almeida-Silva F., et al., 2020, Systematic analysis of 1298 RNA-Seq samples and construction of a comprehensive soybean (Glycine max) expression atlas, Plant J., 103(5): 1894-1909
https://doi.org/10.1111/tpj.14850
PMid:32445587

McCabe C.E., Cianzio S.R., O'Rourke J.A., and Graham M.A., 2018, Leveraging RNA-seq to characterize resistance to Brown stem rot and the Rbs3 locus in soybean, Molecular Plant-Microbe Interactions, 31(10): 1083-1094
https://doi.org/10.1094/MPMI-01-18-0009-R
PMid:30004290

Mo X., Zhang M., Liang C., et al., 2019, Integration of metabolome and transcriptome analyses highlights soybean roots responding to phosphorus deficiency by modulating phosphorylated metabolite processes, Plant Physiol. Biochem., 139: 697-706
https://doi.org/10.1016/j.plaphy.2019.04.033
PMid:31054472

Nakayama T.J., Rodrigues F.A., Neumaier N., et al., 2017, Insights into soybean transcriptome reconfiguration under hypoxic stress: Functional, regulatory, structural, and compositional characterization, PloS one, 12(11): e0187920
https://doi.org/10.1371/journal.pone.0187920
PMid:29145496 PMCid:PMC5690659

Nawaz M.A., Rehman H.M., Baloch F.S., et al., 2017, Genome and transcriptome-wide analyses of cellulose synthase gene superfamily in soybean, J. Plant Physiol., 215: 163-175
https://doi.org/10.1016/j.jplph.2017.04.009
PMid:28704793

Neupane S., Mathew F.M., Varenhorst A.J., and Nepal M.P., 2019a, Transcriptome profiling of interaction effects of soybean cyst nematodes and soybean aphids on soybean, Scientific Data, 6(1): 133
https://doi.org/10.1038/s41597-019-0140-4
PMid:31341170 PMCid:PMC6656750

Neupane S., Varenhorst A.J., and Nepal M.P., 2019b, Transcriptome profiling of induced susceptibility effects on soybean-soybean aphid (Hemiptera: Aphididae) interaction, BMC Res. Notes, 12(1): 325
https://doi.org/10.1186/s13104-019-4372-3
PMid:31182145 PMCid:PMC6558899

Prince S.J., Joshi T., Mutava R.N., et al., 2015, Comparative analysis of the drought-responsive transcriptome in soybean lines contrasting for canopy wilting, Plant Science, 240: 65-78
https://doi.org/10.1016/j.plantsci.2015.08.017
PMid:26475188

Qi Z., Zhang Z., Wang Z., et al., 2018, Meta-analysis and transcriptome profiling reveal hub genes for soybean seed storage composition during seed development, Plant, Cell & Environment, 41(9): 2109-2127
https://doi.org/10.1111/pce.13175
PMid:29486529

Reddy A.S., Marquez Y., Kalyna M., et al., 2013, Complexity of the alternative splicing landscape in plants, Plant Cell, 25: 3657-3683
https://doi.org/10.1105/tpc.113.117523
PMid:24179125 PMCid:PMC3877793

Redekar N.R., Biyashev R.M., Jensen R.V., et al., 2015, Genome-wide transcriptome analyses of developing seeds from low and normal phytic acid soybean lines, BMC Genomics, 16(1): 1074
https://doi.org/10.1186/s12864-015-2283-9
PMid:26678836 PMCid:PMC4683714

Rodrigues F.A., Fuganti-Pagliarini R., Marcolino-Gomes J., et al., 2015, Daytime soybean transcriptome fluctuations during water deficit stress, BMC Genomics, 16(1): 505
https://doi.org/10.1186/s12864-015-1731-x
PMid:26149272 PMCid:PMC4491896

Schmidt M.A., Barbazuk W.B., Sandford M., et al., 2011, Silencing of soybean seed storage proteins results in a rebalanced protein composition preserving seed protein content without major collateral changes in the metabolome and transcriptome, Plant Physiol., 156(1): 330-345
https://doi.org/10.1104/pp.111.173807
PMid:21398260 PMCid:PMC3091051

Schmutz J., Cannon S.B., Schlueter J., et al., 2010, Genome sequence of the palaeopolyploid soybean, Nature, 463(7278): 178-183

Severin A.J., Woody J.L., Bolon Y.T., et al., 2010, RNA-seq Atlas of Glycine max: a guide to the soybean transcriptome, BMC Plant Biol., 10(1): 160
https://doi.org/10.1186/1471-2229-10-160
PMid:20687943 PMCid:PMC3017786

Shamimuzzaman M., and Vodkin L., 2013, Genome-wide identification of binding sites for NAC and YABBY transcription factors and co-regulated genes during soybean seedling development by ChIP-Seq and RNA-seq, BMC Genomics, 14(1): 477
https://doi.org/10.1186/1471-2164-14-477
PMid:23865409 PMCid:PMC3720225

Shen Y., Liu J., Geng H., et al., 2018, De novo assembly of a Chinese soybean genome, Science China Life Sciences, 61(8): 871-884
https://doi.org/10.1007/s11427-018-9360-0
PMid:30062469

Shi G., Huang F., Gong Y., et al., 2014, RNA-seq analysis reveals that multiple phytohormone biosynthesis and signal transduction pathways are reprogrammed in curled-cotyledons mutant of soybean [*Glycine max* (L.) Merr.], BMC Genomics, 15(1): 510
https://doi.org/10.1186/1471-2164-15-510
PMid:24952381 PMCid:PMC4078243

Shimomura M., Kanamori H., Komatsu S., et al., 2015, The *Glycine max* cv. Enrei genome for improvement of Japanese soybean cultivars, International J. Genomics, 2015: 358127
https://doi.org/10.1155/2015/358127
PMid:26199933 PMCid:PMC4493290

Song B., An L., Han Y., et al., 2016a, Transcriptome profile of near-isogenic soybean lines for β-conglycinin α-subunit deficiency during seed maturation, PloS one, 11(8): e0159723
https://doi.org/10.1371/journal.pone.0159723
PMid:27532666 PMCid:PMC4988716

Song L., Pan Z., Chen L., et al., 2020, Analysis of whole transcriptome RNA-seq data reveals many alternative splicing events in soybean roots under drought stress conditions, Genes (Basel), 11(12): 1520
https://doi.org/10.3390/genes11121520
PMid:33352659 PMCid:PMC7765832

Song L., Prince S., Valliyodan B., et al., 2016b, Genome-wide transcriptome analysis of soybean primary root under varying water-deficit conditions, BMC Genomics, 17(1): 57
https://doi.org/10.1186/s12864-016-2378-y
PMid:26769043 PMCid:PMC4714440

Song W., Qi N., Liang C., et al., 2019, Soybean root transcriptome profiling reveals a nonhost resistant response during *Heterodera glycines* infection, PloS one, 14(5): e0217130
https://doi.org/10.1371/journal.pone.0217130
PMid:31125369 PMCid:PMC6534303

Staiger D., and Brown J.W., 2013, Alternative splicing at the intersection of biological timing, development, and stress responses, Plant Cell, 25: 3640-3656
https://doi.org/10.1105/tpc.113.113803
PMid:24179132 PMCid:PMC3877812

Thoenes P., 2004, The role of soybeans in fighting world hunger. In Paper presented by the FAO at the VIIth World Soybean Research Conference held in Foz do Iguassu, Brazil, pp.1-5

Wang J., Hossain M.S., Lyu Z., et al., 2019a, SoyCSN: Soybean context-specific network analysis and prediction based on tissue-specific transcriptome data, Plant Direct, 3(9): e00167
https://doi.org/10.1002/pld3.167

Wang L., Cao C., Ma Q., et al., 2014, RNA-seq analyses of multiple meristems of soybean: novel and alternative transcripts, evolutionary and functional implications, BMC Plant Biol., 14(1): 169
https://doi.org/10.1186/1471-2229-14-169
PMid:24939556 PMCid:PMC4070088

Wang L., Liu L., Ma Y., et al., 2018, Transcriptome profilling analysis characterized the gene expression patterns responded to combined drought and heat stresses in soybean, Computational Biol. Chem., 77: 413-429
https://doi.org/10.1016/j.compbiolchem.2018.09.012
PMid:30476702

Wang X., Yang R., Zhou Y., and Gu Z., 2016, A comparative transcriptome and proteomics analysis reveals the positive effect of supplementary Ca2+ on soybean sprout yield and nutritional qualities, J. Proteomics, 143: 161-172
https://doi.org/10.1016/j.jprot.2016.04.020
PMid:27108549

Wang Z., Gerstein M., and Snyder M., 2009, RNA-seq: a revolutionary tool for transcriptomics, Nature Reviews Genetics, 10(1): 57-63
https://doi.org/10.1038/nrg2484
PMid:19015660 PMCid:PMC2949280

Wang Z., Zhang H., and Gong W., 2019b, Genome-wide identification and comparative analysis of alternative splicing across four legume species, Planta, 249(4): 1133-1142
https://doi.org/10.1007/s00425-018-03073-3
PMid:30603789

Whaley A., Sheridan J., Safari S., et al., 2015, RNA-seq analysis reveals genetic response and tolerance mechanisms to ozone exposure in soybean, BMC Genomics, 16(1): 426
https://doi.org/10.1186/s12864-015-1637-7
PMid:26040850 PMCid:PMC4456062

Wong C.E., Singh M.B., and Bhalla P.L., 2013, The dynamics of soybean leaf and shoot apical meristem transcriptome undergoing floral initiation process, PLoS One, 8(6): e65319
https://doi.org/10.1371/journal.pone.0065319
PMid:23762343 PMCid:PMC3675103

Wu F., Kang X.K., Wang M., et al., 2019, Transcriptome-enabled network inference revealed the GmCOL1 feed-forward loop and its roles in photoperiodic flowering of soybean, Frontiers Plant Science, 10: 1221
https://doi.org/10.3389/fpls.2019.01221
PMid:31787988 PMCid:PMC6856076

Wulff J.A., Kiani M., Regan K., et al., 2019, Neonicotinoid insecticides alter the transcriptome of soybean and decrease plant resistance. Intern. J. Molec. Sci., 20(3): 783
https://doi.org/10.3390/ijms20030783
PMid:30759791 PMCid:PMC6387383

Xue Y., Zhuang Q., Zhu S., et al., 2018, Genome wide transcriptome analysis reveals complex regulatory mechanisms underlying phosphate homeostasis in soybean nodules, Intern. J. Molec. Sci., 19(10): 2924
https://doi.org/10.3390/ijms19102924
PMid:30261621 PMCid:PMC6213598

Yang L., Xu Y., Zhang R., et al., 2018, Comprehensive transcriptome profiling of soybean leaves in response to simulated acid rain, Ecotoxicology Environmental Safety, 158: 18-27
https://doi.org/10.1016/j.ecoenv.2018.04.015
PMid:29656160

Yang S., Miao L., He J., et al., 2019, Dynamic transcriptome changes related to oil accumulation in developing soybean seeds, Intern. J. Molec. Sci., 20(9): 2202
https://doi.org/10.3390/ijms20092202
PMid:31060266 PMCid:PMC6539092

Yu J., Zhang Z., Wei J., et al., 2014, SFGD: a comprehensive platform for mining functional information from soybean transcriptome data and its use in identifying acyl-lipid metabolism pathways, BMC Genomics, 15(1): 271
https://doi.org/10.1186/1471-2164-15-271
PMid:24712981 PMCid:PMC4051163

Yu J.Y., Zhang Z.G., Huang S.Y., et al., 2019, Analysis of miRNAs targeted storage regulatory genes during soybean seed development based on transcriptome sequencing, Genes, 10(6): 408
https://doi.org/10.3390/genes10060408
PMid:31142023 PMCid:PMC6628032

Yuan S., Li R., Chen S., et al., 2016, RNA-seq analysis of differential gene expression responding to different rhizobium strains in soybean (*Glycine max*) roots, Frontiers Plant Science, 7: 721
https://doi.org/10.3389/fpls.2016.00721

Yuan S.L., Li R., Chen H.F., et al., 2017, RNA-seq analysis of nodule development at five different developmental stages of soybean (*Glycine max*) inoculated with *Bradyrhizobium japonicum* strain 113-2, Sci. Rep., 7: 42248
https://doi.org/10.1038/srep42248
PMid:28169364 PMCid:PMC5294573

Zeng A., Chen P., Korth K.L., et al., 2019, RNA sequencing analysis of salt tolerance in soybean (*Glycine max*), Genomics, 111(4): 629-635
https://doi.org/10.1016/j.ygeno.2018.03.020
PMid:29626511

Zeng H., Zhang X., Ding M., et al., 2018a, Transcriptome profiles of soybean leaves and roots in response to zinc deficiency, Physiologia Plantarum, 167(3): 330-351
https://doi.org/10.1111/ppl.12894
PMid:30536844

Zeng H., Zhang X., Zhang X., et al., 2018b, Early transcriptomic response to phosphate deprivation in soybean leaves as revealed by RNA-sequencing, Intern. J. Molec. Sci., 19(7): 2145
https://doi.org/10.3390/ijms19072145
PMid:30041471 PMCid:PMC6073939

Zeng W., Sun Z., Cai Z., et al., 2017, Comparative transcriptome analysis of soybean response to bean pyralid larvae, BMC Genomics, 18(1): 871
https://doi.org/10.1186/s12864-017-4256-7
PMid:29132375 PMCid:PMC5683215

Zeng W., Sun Z., Lai Z., et al., 2019, Determination of the miRNAs related to bean pyralid larvae resistance in soybean using small RNA and transcriptome sequencing, Intern. J. Molec. Sci., 20(12): 2966
https://doi.org/10.3390/ijms20122966
PMid:31216642 PMCid:PMC6628378

Zhang C., Lin C., Fu F., et al., 2017a, Comparative transcriptome analysis of flower heterosis in two soybean F1 hybrids by RNA-seq, PloS one, 12(7): e0181061
https://doi.org/10.1371/journal.pone.0181061
PMid:28708857 PMCid:PMC5510844

Zhang L., Shang J., Wang W., et al., 2019, Comparison of transcriptome differences in soybean response to soybean mosaic virus under normal light and in the shade, Viruses, 11(9): 793
https://doi.org/10.3390/v11090793
PMid:31470502 PMCid:PMC6784153

Zhang X.N., Piao C.L., Dong Y.K., and Cui M.L., 2017b, Transcriptome analysis of response to heavy metal Cd stress in soybean root, J. Applied Ecology, 28(5): 1633-1641

Zhao L., Cui J., Cai Y., et al., 2020, Comparative transcriptome analysis of two contrasting soybean varieties in response to Aluminum toxicity, Int. J. Mol. Sci., 21(12): 4316
https://doi.org/10.3390/ijms21124316
PMid:32560405 PMCid:PMC7352676

Zhu M., Dahmen J.L., Stacey G., and Cheng J., 2013, Predicting gene regulatory networks of soybean nodulation from RNA-seq transcriptome data, BMC Bioinformatics, 14(1): 278
https://doi.org/10.1186/1471-2105-14-278
PMid:24053776 PMCid:PMC3854569