

## Invited Review Article

## Soybean Omics and Biotechnology in China

Yong Guo, Xiao-Bo Wang, Wei He, Guo-An Zhou, Bing-Fu Guo, Le Zhang, Zhang-Xiong Liu, Zhong-Qin Luo, Li-Hui Wang, Li-Juan Qiu\*

The National Key Facility for Crop Gene Resources and Genetic Improvement (NFCRI) / Key Lab of Germplasm Utilization (MOA), Institute of Crop Science, Chinese Academy of Agricultural Sciences, Beijing 100081, P. R. China

\*Corresponding author: qiu\_ljuan@263.net

## Abstract

Key scientific objectives of soybean biology research are to identify and define the functions of soybean genes, and uncover when and how genes work together to produce phenotypic traits. Novel omics and biotechnology tools offer enormous potential to investigate the regulatory networks controlling agronomic traits in soybean. Rapid progress in soybean genome sequencing has facilitated research in soybean omics and biotechnology in China. An overview of soybean omics and biotechnology research in China is presented in this article, which focuses on recent applications of functional genomics, transcriptomics, proteomics, and metabolomics research. Soybean research in China is mainly based on various platforms such as molecular marker development, germplasm resources, mutation library construction, and cDNA library construction. Among the different agronomic traits of soybean, quality, stresses, and development are three major traits studied by Chinese soybean scientists. The future prospects of soybean research in China also are discussed.

**Keywords:** *Glycine max*; functional genomics; transcriptomics; proteomics; metabolomics; transgenic technology; China.

**Abbreviations:** ABA- abscisic acid; AOS- allene oxide synthase; DGAT- diacylglycerol acyltransferase; DREB- dehydration-responsive element-binding protein; DTA- 6-diethyl aminoethyl hexanoate; ER- endoplasmic reticulum; EST- expressed sequence tag; GST- glutathione S-transferase; MAS- marker assisted selection; MeJA- methyl jasmonate; MEP- methylerythritol phosphate; NBS- nucleotide-binding site; PAL- phenylalanine ammonia-lyase; PEG- polyethylene glycol; PGR- plant growth regulator; PSY- phytoene synthase; PTA- *Pinellia ternata* agglutinins; QTL- quantitative trait locus; RCA- rubisco activase; RLK- receptor-like kinase; SAM- shoot apical meristem; SAMDC- S-adenosylmethionine decarboxylase; SAR- system acquired resistance; SCN- soybean cyst nematode; SOD<sub>M</sub>- SOD simulation material; SPS- sucrose phosphate synthase; SMV- soybean mosaic virus; SSH- subtractive hybridization.

## Introduction

Soybean [*Glycine max* (L.) Merr.] is one of the most important agricultural crops for oil and protein and has a natural capacity to fix nitrogen from the atmosphere (Singh and Hymowitz 1999). With a compact genome spanning approximately 1.1 giga-base pairs (Gb), completion of the soybean genome sequence opens a new and exciting chapter in our quest to functionally characterize all annotated genes in soybean (Schmutz, et al. 2010). A systematic approach to characterizing these genes will allow us to dissect and understand the regulatory networks and evolutionary processes controlling such complex traits as yield, oil content, protein composition, responses to abiotic and biotic stresses, and plant development. Soybean omics research, including functional genomics, transcriptomics, proteomics, and metabolomics, provides new approaches and strategies to study important traits of soybean. Therefore, development of omics research provides useful tools for discovering important genes and their functions, and can drive the development of soybean biotechnology such as transgenics. China is one of the most important countries for soybean production and consumption; modern soybean research in China originated in the early twentieth century (Xiong, et al. 2008). Along with advancements in genetics and molecular biology, soybean research in China has entered an

era of rapid growth. For example, numerous genes have been cloned and characterized in soybean (Table 1); soybean transcriptomics, proteomics, and metabolomics research have been initiated, and soybean biotechnology has been adopted and utilized widely. The major players for the soybean research in China include the Chinese Academy of Agricultural Sciences (CAAS), The Chinese Academy of Sciences (CAS), several universities such as Nanjing Agricultural University and Northeast Agricultural University, and provincial academies of agricultural sciences. Here, we report the major accomplishments and current status of soybean omics and biotechnology research in China.

## Platform of Basic and Applied Soybean Research in China

Molecular marker technology is the basis of gene function and other omics research. So far, a number of important soybean trait loci/QTLs have been identified and mapped in China. These traits included pest resistance [such as soybean cyst nematode (SCN), soybean mosaic virus (SMV), and *Phytophthora* root rot], abiotic stress tolerance (drought, salt, etc.), seed quality attributes (protein and oil content, lipoxygenase-free), yield-related traits, and morphological

characteristics (root mass, narrow leaves, curved stem, short petiole, and fertility). Many important agronomic traits, including yield, seed quality, and stress tolerance, are quantitative traits controlled by multiple genes and influenced by environmental conditions. Therefore, molecular markers identified are often environment and/or population/genetic background specific. However, less repeatable molecular markers identified using different materials in different years at different locations are being characterized in China. Soybean originated in China (Qiu and Chang 2010) and soybean germplasm is the basis of soybean research. Many germplasm accessions contain new or useful genes that can significantly enhance the quality of agronomic traits of soybean. Therefore, collection of germplasms will be very useful for indentifying new genes related to soybean traits of importance (Qiu, et al. 2010). China has the largest collection of soybean germplasms in the world, with more than 30,000 in the National Germplasm Bank. Based on the strategy of germplasm core collection, in which a limited set of accessions with a minimum amount of repetitiveness were chosen to represent a maximum genetic diversity of entire germplasm resources (Frankel and Brown 1984), a large number of soybean germplasm collections have been evaluated properly and soybean core collections were established with both simple sequence repeat (SSR) marker data and agronomic traits (Qiu, et al. 2009; Wang, et al. 2006). These soybean core collections are valuable materials for new trait discovery and novel gene characterization. For example, a determinate stem (*Dt1*) locus was identified from these soybean core collections (Tian, et al. 2010). One of the best strategies for reverse genetics is based on mutagenesis. Soybean mutant libraries in China were constructed by treating seeds of “Nannong 94-16” with EMS or  $\text{NaN}_3$ - $^{60}\text{Co}$   $\gamma$ -rays combined irradiation, and seeds of “Nannong 86-4” with  $^{60}\text{Co}$   $\gamma$ -rays irradiation and EMS separately (Chen and Yu 2009; Han, et al. 2008; Han, et al. 2007). Approximately 400 mutants were identified from these libraries, relating to plant development, stress tolerance, and seed quality attributes. Full-length cDNA clones are valuable resources for functional analyses of genes, not only at the transcriptional level but also at the translational level (Nishiyama, et al. 2003). Soybean scientists in China have established full-length cDNA libraries for a vast collection of expressed sequences. To date, full-length soybean cDNA libraries have been constructed from leaf (Dong, et al. 2006), seeds at different developmental stages (Li, et al. 2010), plants grown under drought or low temperature stresses (Yang, et al. 2009), and even from wild soybean seed (Wang, et al. 2005). All these full-length clones are important resources for further functional studies and could be utilized widely in soybean biological studies.

### Soybean Quality Traits

Traits related to soybean quality are among the most important agronomic traits. Understanding the genetic basis of seed quality traits is critical for improving the quality of soybean. Soybean quality traits include protein and oil content, fatty acid composition, isoflavone content and other chemical, physical, and nutritional characteristics. Chinese researchers have made notable progress in research on the biology of quality traits, particularly in mapping a large number of loci/QTLs associated with protein content (Liu, et al. 2009; Shan, et al. 2009; Wu, et al. 2001; Yang, et al. 2004; Zhang, et al. 2004), oil content (Wang, et al. 2008; Wei, et al. 2009; Xu, et al. 2007; Zheng, et al. 2006), and lipoxygenase deficiency (Sun, et al. 2004), providing useful information for dissecting respective gene functions. The pathways for lipid biosynthesis are important processes in soybean and identifying genes involved in such

processes are very important for soybean quality improvement. Diacylglycerol acyltransferase (DGAT), a key enzyme in triacylglycerol synthesis, catalyzes the final acylation of the Kennedy pathway. The *GmDGAT* gene was cloned from *Glycine max* by searching for homologs of characterized sequences from *Arabidopsis*. Alleles were isolated from 13 soybean accessions and divergences among deduced amino acid sequences were compared. (Wang, et al. 2006). The endoplasmic reticulum (ER)-associated oleate desaturase (FAD2) is the key enzyme responsible for the production of linoleic acid in non-photosynthetic tissues. Two soybean *FAD2* isoforms, *FAD2-1B* and *FAD2-3*, were cloned and characterized, indicating that these two genes were expressed in developing soybean seeds. Yeast cells transformed with *FAD2-3* coding region accumulated a considerable amount of linoleic acid (18:2), normally not present in wild-type yeast cells (Li, et al. 2007; Li, et al. 2008). Two Dof-like genes, *GmDof4* and *GmDof11*, were characterized to be involved in the regulation of the lipid content in soybean seeds by activating, respectively, the acetyl CoA carboxylase and long-chain-acyl CoA synthetase genes (Zhang, et al. 2007). The function of *GmDof4* was confirmed through a transgenic soybean cultivar “Suinong 14” (Du, et al. 2010). By using RNAi technology, the expression of lipoxygenase genes in soybean seed could be effectively inhibited and the amount of oil content increased, which improves the quality of soybeans (Ma, et al. 2009). In addition, a number of differentially expression cDNAs related to oil content were obtained by suppression subtractive hybridization (SSH) (Cheng, et al. 2007). Protein composition is another important quality trait for soybean. The terminal step of soybean cysteine synthesis is catalyzed by O-acetylserine(thiol)lyase. Two soybean O-acetylserine(thiol)lyase genes, *GmOASTLA* and *GsOAS-TLI*, were cloned and characterized from cultivated and wild soybean (Ning, et al. 2010; Zhang, et al. 2008). Increased accumulation of transcripts and higher cysteine content in transgenic tobacco plants that over-expressed *GmOASTLA* showed that *GmOASTLA* was involved in soybean cysteine synthesis (Ning, et al. 2010). Soybean seed ferritin is a model for studying the structure and function of plant ferritins. A novel ferritin gene, *SferH-5*, has been cloned from 7-day-old soybean seedlings (Dong, et al. 2007). This research revealed that recombinant SferH-5 was able to assemble, together with co-expressed SferH-2, as a functional soybean seed ferritin-like complex, H-5/H-2 (Dong, et al. 2007). Soybean is not only rich in high-quality protein and fat, but also contains a variety of isoflavones with biological activities. Isoflavones are legume-specific flavonoids best known for their potential for cancer prevention and their phytoestrogenic properties. Presence of genistin (genistein 7-O-glucoside) as the major isoflavone metabolite could be detected by LC-MS/MS analysis of tomato plants transformed with soybean isoflavone synthase (*GmIFS2*) cDNA under the control of the CaMV 35S promoter (Shih, et al. 2008). DXS catalyses the first committed step of the methylerythritol phosphate (MEP) pathway, which is an alternative route for isoprenoid biosynthetic. Research on the function of *GmDXS1* in transgenic tobacco implied that *GmDXS1* catalyzed the first potential regulatory step in photosynthetic pigment biosynthesis via the MEP pathway (Zhang, et al. 2009). In addition, studies on metabolic mechanisms of isoflavones have been reported. Exogenous methyl jasmonate (MeJA) could promote the synthesis of soybean isoflavones. Exogenous MeJA increased the activity of phenylalanine ammonia-lyase (PAL) in leaves and enhanced the expression level and activity of *PAL* gene in grains, indicating that MeJA could be involved in the metabolism of isoflavones (Ma, et al. 2007). In China, genetic improvement of

**Table 1.** Research Updates on Soybean Genes in China

Gene	Putative Function	Functional Verification	References
<b>Quality Trait Related Genes</b>			
<i>GmDGAT</i>	Triacylglycerol synthesis	N.D. <sup>a</sup>	Wang et al., 2006a
<i>FAD2-1B;FAD2-3</i>	Accumulate linoleic acid	Yeast	Li et al., 2007, 2008a
<i>GmDof4;GmDof11</i>	Regulation of lipid content	<i>Arabidopsis</i>	Wang et al., 2007a
<i>GmOASTL4;GmOAS-TL1</i>	Increase cysteine content	Tobacco	Zhang et al., 2008a; Ning et al., 2010
<i>SferH-5</i>	Assemble ferritin-like complex	N.D.	Dong et al., 2007
<i>GmIFS2</i>	Isoflavone biosynthase	Tomato	Shih et al., 2008
<i>GmDXS1</i>	Photosynthetic pigment biosynthase	Tobacco	Zhang et al., 2009b
<b>Stress Resistance Related Genes</b>			
<i>GmPHD2</i>	Salt stress tolerance	<i>Arabidopsis</i>	Wei et al., 2009b
<i>GmGT-2A;GmGT-2B</i>	Salt, freezing and drought tolerance	<i>Arabidopsis</i>	Xie et al., 2009
<i>GmDREBa;GmDREBb;GmDRE Bc</i>	Induced by salt, drought, and cold	N.D.	Li et al., 2005
<i>GmDREB2;GmDREB3</i>	Cold, drought, and high salt tolerance	<i>Arabidopsis</i> tobacco	Chen et al., 2006; Chen et al., 2009a
<i>GmbZIP44;GmbZIP62;GmbZIP 78</i>	ABA signaling; salt and freezing tolerance	<i>Arabidopsis</i>	Liao et al., 2008a; Liao et al., 2008c
<i>GmMYB76;GmMYB92;GmMYB 177</i>	Salt and freezing tolerance	<i>Arabidopsis</i>	Liao et al., 2008b
<i>GmWRKY13;GmWRKY27;GmW RKY54</i>	Increase sensitivity to salt and mannitol stress, decreases sensitivity to ABA	<i>Arabidopsis</i>	Zhou et al., 2008
<i>GmERF3;GmERF4;</i>	Increase salt and drought stresses tolerance	tobacco	Zhang et al., 2009a; Zhang et al., 2010
<i>GmSGR</i>	Reduce ABA sensitivity and enhance salt sensitivity	<i>Arabidopsis</i>	Wang et al., 2008a
<i>GmUBC2</i>	Salt and drought tolerance	<i>Arabidopsis</i>	Zhou et al., 2010
<i>GsGST</i>	Drought and salt tolerance	tobacco	Ji et al., 2010
<i>GmRFP1</i>	ABA signaling and stress responses	N.D.	Du et al., 2010a
<i>GmPAP3</i>	Enhance ROS scavenging and stress tolerance	<i>Arabidopsis</i>	Liao et al., 2003; Li et al., 2008b
<i>GmSAMDC1</i>	Induced by salt, drought and cold	N.D.	Tian et al., 2004c
<i>GmCLC1;GmNHX1</i>	Reduce toxic effects of ions	BY-2 cells	Li et al., 2006a
<i>GmCAX1</i>	Tolerant to elevated Li <sup>+</sup> and Na <sup>+</sup> levels	<i>Arabidopsis</i>	Luo et al., 2005
<i>GmHsfA1</i>	Enhance heat tolerant	Soybean	Chen et al., 2006; Zhu et al., 2006
<i>GmPti1</i>	Induced by salicylic acid and wounding	N.D.	Tian et al., 2004a
<i>SbPRP</i>	Regulated by salt and drought stresses	N.D.	He et al., 2002
<i>GmKR1;GmKR4</i>	Induced by infection of SMV	N.D.	He et al., 2003; Wang et al., 2004
<i>GmHZ1</i>	Response to SMV infection	N.D.	Wang et al., 2005b
<i>GmAOS</i>	Resistance to cotton worm	Tobacco	Wu et al., 2008a
<b>Growth and Development Related Genes</b>			
<i>GmTfl1/GmTFL1b</i>	Control stem growth habit	Soybean <i>Arabidopsis</i>	Liu et al., 2010; Tian et al., 2010
<i>GmCRY1a;GmCRY2a</i>	Major regulator of photoperiodic flowering	<i>Arabidopsis</i>	Zhang et al., 2008b
<i>GmLCL1;GmLCL2;GmTOC1</i>	Components of central clock	N.D.	Liu et al., 2009a
<i>GmNMH7</i>	Related to floral development	N.D.	Wu et al., 2006
<i>GmRCA</i>	Rubisco activase	N.D.	Yin et al., 2010
<i>GmKNT1</i>	Leaf and internode and inflorescence development	<i>Arabidopsis</i>	Liu et al., 2008
<i>GmALDJ3-1;GmZFP1</i>	Flower development	N.D.	Huang et al., 2006a; Huang et al., 2010
<i>GmSARK</i>	Regulate leaf senescence	Soybean	Li et al., 2006b
<i>GmNAC1-GmNAC6</i>	Diverse and vital physiological processes during development	N.D.	Meng et al., 2007

<sup>a</sup> No data provided

soybean quality via biotechnology is still at the infant stage; however, research progress has been quite fast-paced. The *Mortierella isabellina*  $\Delta^6$ -fatty acid desaturase gene has been transformed into soybean by *Agrobacterium*-mediated genetic transformation (Bo, et al. 2003). The phytoene synthase (PSY) gene was integrated into the soybean genome by particle bombardment or *Agrobacterium*-mediated transformation to improve seed quality, including increased carotenoid content of transgenic seeds (Gong, et al. 2005; Gong, et al. 2006). After the maize sucrose phosphate synthase (SPS) gene was transformed into the soybean genome by *Agrobacterium*-mediated transformation of the explants of embryos tip, sucrose content in transgenic plants increased significantly (Song, et al. 2008).

### Soybean Stress Resistance/Tolerance

Biotic and abiotic stresses are major factors that affect plant growth and development; thus it is important to investigate the mechanisms by which plants respond to environmental stresses. In China, many loci/QTLs that control soybean responses to various stresses have been discovered, including salt tolerance (Guo, et al. 2000), drought resistance (Liu, et al. 2005), lodging-resistance (Huang, et al. 2008), low phosphate stress (Cui, et al. 2007), and flood tolerance (Wang, et al. 2008). To understand the effect of abiotic stresses on gene expression, a full-length cDNA library of *Glycine soja* was constructed (Ji, et al. 2006). A total of 2,003 cleaned ESTs sequences were assembled to generate 375 contigs and 696 singlets; most expressed sequences from *Glycine soja* exhibited similarity with orthologs from *Glycine max*. In addition, cDNA libraries for soybean leaves, drought and cold stresses, and wild soybean seed were constructed (Dong, et al. 2006; Wang, et al. 2005; Yang, et al. 2009). In response to abiotic stresses, many loci/QTLs have been identified for SMV resistance (Zheng, et al. 2001), SCN resistance (Lu, et al. 2008), cotton warm resistance (Liu, et al. 2005), bean pyralid resistance (Li, et al. 2009) and *Phytophthora* root rot resistance (Han, et al. 2006). These genetic studies have provided in-depth insights into soybean functional genes and genetic/biochemical mechanisms related to plant responses to biotic and abiotic stresses.

#### (1) Abiotic stresses

Several transcription factor families have been found to be induced by abiotic stresses, including DREB, ERF/AP2, WRKY, MYB, bZIP, PHD, and GT-2 families. Six *GmPHD* genes encoding Alfin1-type PHD finger proteins and two *GmGT-2* transcription factor genes were identified and their expressions responded differentially to drought, salt, cold and ABA treatments (Wei, et al. 2009; Xie, et al. 2009). Transgenic *Arabidopsis* plants over-expressing the *GmPHD2* gene achieved greater salt tolerance by diminishing oxidative stress through regulation of downstream genes compared with the wild type plants (Wei, et al. 2009). Over-expression of the *GmGT-2A* and *GmGT-2B* improved plant tolerance to salt, freezing and drought stresses in transgenic *Arabidopsis* plants. These findings indicated that *GmPHD* and *GmGT-2* transcription factors confer stress tolerances through regulation of a common set of specific genes (Wei, et al. 2009). DREB (dehydration-responsive element-binding protein) transcription factors have important roles in stress-related regulation networks in plants. Five DREB homologue genes, *GmDREBa*, *GmDREBb*, *GmDREBc*, *GmDREB2* and *GmDREB3*, have been isolated from soybean (Chen, et al. 2007; Chen, et al. 2009; Li, et al. 2005). Several of these genes were induced by salt, drought, and cold stresses in leaves of soybean seedlings (Li, et

al. 2005). Over-expression of *GmDREB2* and *GmDREB3* in *Arabidopsis* and tobacco could enhance tolerance to cold, drought, and high salt stresses, and accumulation of higher levels of free proline compared to wild type plants under drought conditions (Chen, et al. 2007; Chen, et al. 2009). Plant basic-leucine zipper (bZIP) transcription factors play important roles in many biological processes. Functional studies of four bZIP transcription factor genes (*GmbZIP44*, *GmbZIP62*, *GmbZIP78* and *GmbZIP132*) in transgenic *Arabidopsis* revealed that these genes could be negative regulators of ABA signaling and function in salt and freezing tolerance (Liao, et al. 2008; Liao, et al. 2008). MYB-type transcription factors are involved in the regulation of many aspects of plant growth, development, metabolism, and stress responses. Expression levels of 43 genes were changed upon treatment with ABA, salt, drought and/or cold stress (Liao, et al. 2008). Further analysis of three *GmMYB* genes (*GmMYB76*, *GmMYB92* and *GmMYB177*) in transgenic *Arabidopsis* showed that the three *GmMYB* genes differentially affected various stress-responsive genes and commonly regulated a subset of stress-responsive genes (Liao, et al. 2008). Other soybean transcription factor genes such as *GmWRKYs* (*GmWRKY13*, *GmWRKY27* and *GmWRKY54*), *GmERFs* (*GmERF3* and *GmERF4*) and *GmSGR* also have been identified to be involved in response to abiotic stresses in soybean (Wang, et al. 2008; Zhang, et al. 2010; Zhang, et al. 2009; Zhou, et al. 2008). Stress related transcription factor genes *GmDREB* and *ABP9* have been transformed into soybean genome successfully. It is very important to use molecular approaches to improve drought, salt and cold tolerance in soybean, as these traits are complex and difficult to manipulate through traditional breeding methods (Wang, et al. 2008; Wang, et al. 2008). Enzymes, especially ubiquitin-conjugating enzyme, glutathione S-transferases, and protein kinases, also are very important for responding to abiotic stresses. The gene for a ubiquitin-conjugating enzyme, *GmUBC2*, was cloned from soybean and its functionality characterized. Detailed study of *GmUBC2* in transgenic *Arabidopsis* plants revealed it to be up-regulated by drought and salt stress and involved in the regulation of ion homeostasis, osmolyte synthesis, and oxidative stress responses (Zhou, et al. 2010). Glutathione S-transferase (GST) also plays important roles in plant abiotic stress responses. A *GST* sequence was isolated from a cDNA library constructed from *G. soja* seedlings treated with NaCl. Over-expression of *GsGST* enhances drought and salt tolerance in transgenic tobacco (Ji, et al. 2010). RING-finger proteins with E3 ubiquitin ligase activity play important roles in the regulation of plant growth and development. Isolation and characterization of *GmRFP1*, a gene encoding a novel RING-finger protein in soybean, revealed that *GmRFP1* might play unappreciated roles in ABA signaling and stress responses via mediating ubiquitination and degradation of target proteins through the ubiquitin-proteasome pathway (Du, et al. 2010). Purple acid phosphatases (PAP) are enzymes that catalyze the hydrolysis of phosphate esters and anhydrides. *GmPAP3*, a novel PAP gene identified from salt-stressed soybean using SSH techniques, could be induced by NaCl stress (Liao, et al. 2003). Transgenic analysis in *Arabidopsis* over-expressing *GmPAP3* revealed that *GmPAP3* could play a role in stress tolerance by enhancing ROS scavenging (Li, et al. 2008). S-adenosylmethionine decarboxylases (SAMDC) are key enzymes involved in polyamine biosynthesis. Cloning and characterization of *GmSAMDC1*, one of the *SAMDC* genes present in soybean, showed that it was induced by salt, drought and cold, but not by wounding, suggesting that it is involved in response to multiple-stress conditions (Tian, et al. 2004). Nine differentially expressed proteins including ferritin, 20S

proteasome  $\beta$  subunit, GST9, GST10, and seed maturation protein PM36 were identified using proteomics during soybean seed germination under NaCl treatment. These proteins could play important roles in protection of seed during germination (Xu, et al. 2006). Proteins related to low temperature stress in soybean euphylla and cotyledon such as glyceraldehyde-3-phosphate dehydrogenase and  $\text{Ca}^{2+}$ /calmodulin-dependent protein kinase also were identified in soybean cultivars with strong cold resistance. Differentially expressed proteins in aluminum resistant and sensitive soybean cultivars, including PR10-like protein, NAD(P)H-dependent 6'-deoxychalcone-synthetase reductase and 2,4-D inducible glutathione-S-transferase, also were identified (Zhen, et al. 2007) and could play important roles in improving soybean aluminum tolerance. Ion transporters are involved in abiotic stress tolerance by regulating ion homeostasis. Modulation of these transporters could be beneficial for regulation of ion homeostasis and thus plant abiotic tolerance. Using degenerate primers, putative chloride channel (*GmCLC1*) and  $\text{Na}^+/\text{H}^+$  antiporter (*GmNHX1*) genes were cloned from soybean. The expressions of *GmCLC1* and *GmNHX1* both were induced by NaCl or dehydration stress imposed by polyethylene glycol (PEG). Further analysis revealed a clear alleviation of NaCl stress in both *GmCLC1* and *GmNHX1* transgenic cells (Li, et al. 2006). Another cation/proton antiporter *GmCAX1* was also identified by Luo et al (2005). *GmCAX1* was expressed in all tissues of the soybean plants but at a lower level in roots. Its expression was induced by PEG, ABA,  $\text{Ca}^{2+}$ ,  $\text{Na}^+$  and  $\text{Li}^+$  treatments. Transgenic *Arabidopsis* plants over-expressing *GmCAX1* accumulated less  $\text{Na}^+$ ,  $\text{K}^+$ , and  $\text{Li}^+$ , and were more tolerant to elevated  $\text{Li}^+$  and  $\text{Na}^+$  levels during germination compared to control plants (Luo, et al. 2005). Other stress tolerance related genes are associated with chaperone heat shock proteins, plant defense systems, and establishing unique metabolite compositions. Characterization of *GmHsfA1* showed that over-expression of *GmHsfA1* in soybean led to the notable enhancement of heat-tolerance level of transgenic plants by possibly mediating downstream activation of transcription or improvement of expression of some *GmHSPs* (Chen, et al. 2006; Zhu, et al. 2006). *GmPti1* is another gene from soybean identified by its homology to tomato *Pti1*. RT-PCR analysis indicated that *GmPti1* was induced by salicylic acid and wounding (Tian, et al. 2004). Furthermore, *SbPRP*, which encodes a soybean proline-rich protein, also was expressed in response to salicylic acid and virus infection and also found to be regulated by circadian rhythm, salt stress, drought stress and plant hormones (He, et al. 2002). In addition, metabolism of such soybean compounds as isoflavones and saponins also plays extensive roles in interactions between plants and the environment; for example, serving as chemical defense compounds against environmental stress to plants. Furthermore, marked differences in metabolite compositions were identified between salt-sensitive and salt-tolerant soybean varieties (Wu, et al. 2008).

## (2) Biotic stresses

*Phytophthora* root rot of soybean (*Phytophthora sojae*) is one of most destructive diseases endangering soybean production. Molecular mechanisms of disease resistance are currently a dynamic research area. A total of 375 functional ESTs have been identified by SSH from a cDNA library of resistant soybean varieties associated with *Phytophthora* root rot (Zhang, et al. 2008). It was found that system acquired resistance (SAR) induced by *Phytophthora* in soybean primarily depended on salicylic acid mediated signaling pathways, which was different from resistance mechanisms in *Arabidopsis*. Nevertheless, the

functions of differentially expressed proteins induced by *Phytophthora* are almost the same in soybean and *Arabidopsis*, including material metabolism enzymes and various regulatory factors; this suggests that the molecular basis of system acquired resistance against *Phytophthora* is much the same in soybean and *Arabidopsis* (Sun, et al. 2008). Plant disease resistance genes usually have conserved motifs such as the nucleotide-binding site (NBS). Based on the NBS domain, two putative resistance genes (*KR1* and *KR4*) were isolated from soybean. Further study revealed that *KR1* and *KR4* were induced by exogenous salicylic acid and responded upon infection by soybean mosaic virus (He, et al. 2003; Wang, et al. 2004). Allene oxide synthase (AOS) is a major intermediate enzyme in the octadecanoid pathway for jasmonic acid biosynthesis. Over-expressing *GmAOS* in transgenic tobacco plants increased activities of allene oxide synthase, peroxidase, chymotrypsin inhibitor, and increased trichome number. Insect bioassays indicated that transgenic tobacco were more resistant to cotton worm than non-transgenic controls (Wu, et al. 2008). Moreover, a homeodomain-containing transcription factor GmHZ1 has been characterized as a transcriptional activator for response to SMV infection (Wang, et al. 2005). A resistant soybean line was used to construct a subtractive cDNA library by SSH from soybean leaves inoculated with SMV at the primary leaf stage. The EST expression profile showed that the resistance-related genes for SMV were involved in cell protection, signal transduction, restriction of pathogen growth, system acquired resistance, and house-keeping functions (Liu, et al. 2005). Differentially expressed proteins in soybean leaves infected with SMV were identified by proteomic approaches. Sixteen identified proteins were potentially involved in protein degradation, defense signal transfer, oxidative stress, cell wall reinforcement, and energy and metabolism regulation (Yang, et al. 2010). There is some evidence that insect selection of host plant is associated with plant nutritional components and secondary metabolites, and the latter have a decisive impact on insect feeding. Two soybean varieties with resistance and sensitivity to aphids were examined to study the mechanism of aphid resistance. The results showed that PAL enzyme activity was highly correlated with aphid resistance and the contents of some secondary metabolites in leaf tissue were obviously different between the two materials, including amino acids, flavonoids, alkaloids and phenols (Jiang, et al. 2009). In addition, differential expression was found by SSH in cDNA libraries from soybean varieties resistant to soybean cyst nematode (Cheng, et al. 2007; Lu and Fang 2003). In China, transgenic soybeans resistant to biotic stresses have been developed using different transformation methods in recent years. Transgenic soybeans that simultaneously express the *CryIA(c)* and *Pinellia ternata* agglutinins (*PTA*) genes were obtained by *Agrobacterium*-mediated transformation of soybean embryonic tips (Dang and Wei 2007). In addition, the target *Bt* gene has been transformed into the soybean genome successfully by both microprojectile bombardment and *Agrobacterium*-mediated somatic embryo transformation (Guo, et al. 2008; Wang, et al. 2007). Furthermore, transgenic soybeans co-integrating two anti-fungal protein genes, *chi* and *rip*, also were obtained and these transgenic lines exhibited resistance to *Phytophthora megasperma* and *Cercospora sojina* (Guo, et al. 2006).

## Soybean Developmental Biology

The growth and development of soybean, the most agronomically important member of the legume family, are unique. Soybean is one of the model plants for photoperiod and nodulation research, and many important plant photoperiodic

responses were discovered using soybean as research material. Research on mechanisms of soybean development is not only useful for acquiring a basic understanding of plant developmental biology, but also valuable for improving agronomic traits of soybean. Many developmental loci/QTLs have been identified and mapped in China, including ones associated with leaf morphology (Zhu, et al. 2001), stem form (Huang, et al. 2006), petiole (Huang, et al. 2005), fertility (Duan, et al. 2005), and leaf development (Cui, et al. 2007). In addition, progress has been made in soybean transcriptomics associated with the development research. A total of 56,147 unigenes were assembled through analysis of the 314,254 soybean ESTs from research by Tian et al (2004b) and the Genbank database. These research results also suggested that unigenes encoding cell growth and maintenance, enzyme and cell communication had low rates of gene evolution, whereas genes encoding transcription regulation, cell, binding and death related functions displayed higher evolutionary rates. The soybean-specific unigenes mainly were involved in nodule development and production of seed storage proteins (Tian, et al. 2004). Li et al. constructed a full-length cDNA library from soybean seeds at different developmental stages and obtained 27,982 unigenes from 36,656 high-quality ESTs (Li, et al. 2010). These research findings provided the basis for functional studies of the soybean developmental related genes. Classical genetic analysis has revealed that the determinate habit, which is an important agronomic trait in soybean, is controlled by a recessive allele at the determinate stem (*Dt1*) locus. Most soybean cultivars are classifiable into indeterminate and determinate growth habits, whereas *Glycine soja*, the wild progenitor of soybean, is indeterminate with a prostrate growth habit. In order to dissect the molecular basis of the determinate habit, two research groups in China collaborated with scientists from USA and Japan and independently showed evidence that *Dt1* is a homolog (designated as *GmTf11/GmTFL1b*) of *Arabidopsis* *TERMINAL FLOWRE 1 (TFL1)* (Liu, et al. 2010; Tian, et al. 2010). The *TFL1* gene of *Arabidopsis* is required for normal inflorescence meristem function and *tf1* mutations resulted in the conversion of the normally indeterminate inflorescence to a determinate inflorescence condition (Shannon and Meeks-Wagner 1991). Further analysis in transgenic soybean and *Arabidopsis* revealed that *GmTf11/GmTFL1b* can restore the indeterminate stem growth habit in a determinate soybean line or in an *Arabidopsis tf1* mutant (Liu, et al. 2010; Tian, et al. 2010). Soybean is a typical short-day crop, and its photoperiodic response of flowering is critical to growth and development. Function and expression studies of two cryptochromes in soybean, *GmCRY1a* and *GmCRY2a*, revealed that *GmCRY1a* is a major regulator of photoperiodic flowering in soybean, and photoperiod-dependent circadian rhythmic expression of the *GmCRY1a* protein correlates with flowering and latitudinal distribution of soybean cultivars (Zhang, et al. 2008). Cloning and characterization of *GmLCL1*, *GmLCL2*, and *GmTOC1* revealed that these clock-gene homologs could be central clock components in soybean (Liu, et al. 2009). *NMH7* family members encode MADS-box proteins and are unique in legume families. *In situ* expression of *GmNMH7* using soybean plants grown under different photoperiod cycles provided evidence that *GmNMH7* gene expression is closely related to floral development (Wu, et al. 2006). Rubisco activase (RCA) catalyzes the activation of Rubisco *in vivo* and plays a crucial role in photosynthesis. Two genes that encode the longer  $\alpha$ -isoform and shorter  $\beta$ -isoform of soybean RCA recently were cloned and characterized (Yin, et al. 2010). Correlation of gene expression levels of these two genes with Rubisco initial activity, photosynthetic rate, and seed yield indicated that RCA

genes could play an important role in regulating soybean photosynthetic capacity and seed yield (Yin, et al. 2010). Furthermore, ESTs have been characterized through screening a cDNA library from soybean leaf treated with short-day light, and many are related to physiological and biochemical functions such as transcription, signal transduction and programmed cell death (Zhao, et al. 2007). Plant shoot apical meristems (SAM) define the architecture of plant leaves and flowers. The soybean *KNOX* gene *GmKNT1*, which is homologous to *Arabidopsis STM*, is expressed strongly in roots, flowers and developing seeds. Transgenic *Arabidopsis* plants over-expressing *GmKNT1* showed severe defects in leaves, internodes, and inflorescence, indicating that it affected multiple aspects of plant growth (Liu, et al. 2008). In research on mechanisms of soybean reproductive development, a number of flower developmental genes have also been identified by microarray hybridization, such as soybean aldehyde dehydrogenase gene *GmALDJ3-1* and soybean single zinc finger gene *GmZFP1* (Huang, et al. 2006; Huang, et al. 2010). Other genes related to soybean development included soybean receptor-like kinase (*RLK*) gene and *NAC* genes. Functional analysis of *GmSARK* (*Glycine max* senescence-associated receptor-like kinase) suggested that it regulated soybean leaf senescence via regulating chloroplast development and chlorophyll accumulation (Li, et al. 2006). *NAC* proteins have been considered as one of the novel classes of plant-specific transcription factors functioning in diverse and vital physiological processes during plant development. Six *NAC*-like genes from soybean, designated as *GmNAC1-GmNAC6*, were cloned and characterized. RT-PCR indicated that each *GmNAC* gene exhibited a specific expression pattern in the tissues examined (Meng, et al. 2007). Proteomics research on root, stem, leaf, seed and flower of plants provide valuable information for studying protein function in particular tissues and in important plant developmental pathways and mechanisms. Chen et al. found tissue-specific or abundant protein spots for each tissue by comparison of two-dimensional patterns of mRNA translation products (Chen, et al. 1990). A proteomic approach was used to analyze differential expression of proteins during soybean seed germination. The results showed that six proteins were related to seed development, including nucleotide diphosphate kinase, proglycinin A<sub>1a</sub>B<sub>1b</sub> subunit, thioredoxin fold, 35 ku seed maturation protein, heat shock protein, and seed maturation protein (Xu, et al. 2006; Xu, et al. 2006). Eighteen highly differentially expressed proteins were found to be involved in seed maturation, stress tolerance, cell division, and protein transportation based on comparison of protein expression at different developmental stages of soybean seed (Zheng, et al. 2008). Cytoplasmic-nuclear male sterility plays an important role in the production system of hybrid seeds. Knowledge of the mechanism of the cytoplasmic-nuclear male sterility is important for improvement of hybrid seed production. The comparative proteome analysis of different organs between the cytoplasmic-nuclear male sterile soybean lines NJCMS1A, NJCMS2A and their maintainer NJCMS1B, NJCMS2B was conducted. The results suggested that the male sterility of NJCMS1A and NJCMS2A could be related to energy metabolism turbulence, programmed cell death, starch synthesis suffocation and disturbed function of flower development related genes (Zeng, et al. 2007; Zeng, et al. 2007). Comparative proteome analysis of different organs between the cytoplasmic-nuclear male sterile soybean line W931A and its maintainer W931B also identified major differentially expressed proteins, including ADP2 glucose pyrophosphorylase, NBS-LRR disease resistance protein homologue, ATP synthase  $\beta$  subunit, thioredoxin dependent

peroxidase, and zinc finger protein (Chen, et al. 2009). Plant growth regulators (PGRs) are a class of exogenous chemicals possessing similar physiological and biological effects to plant hormones in plants. They usually are conducted to the site of action, and promote or inhibit certain aspects of life processes at very low concentrations. In order to ascertain the effects of plant growth regulators on metabolism of soybean, three different PGRs, SOD simulation material (SOD<sub>M</sub>), choline chloride and diethyl aminoethyl hexanoate (DTA-6), were used by foliar spray. The result showed that SOD<sub>M</sub> could significantly improve the accumulation and output of free amino acids in roots, while SOD<sub>M</sub> and DTA-6 could enhance nitrate, soluble sugar and soluble protein contents, and both improve assimilation metabolism and promote the normal growth of soybean roots (Zhao, et al. 2008). Research on the effects of plant growth regulator SHK-6 on nitrogen metabolism suggested that elevated protein content, improved components of amino acids, increased activity of nitrate reductase and the content of nitrate, all could be detected (Zhang, et al. 2004). In addition, there is a class of non-protein coding RNA widely found in plants, animals, bacteria and fungi (Zhang et al., 2007). These RNAs, including miRNA, play a vital role in transcription, translation, RNA splicing, DNA replication and RNA processing. Following a bioinformatics identification of miRNAs in soybean (Zhang et al., 2008), a library of small RNAs was constructed from soybean functional nodules and 32 small RNA sequences were identified as belonging to 11 miRNA families. Expression analysis and miRNA target prediction suggested that these miRNAs were involved in nutrient acquisition, and plant development (Wang, et al. 2009). Furthermore, 15 conserved miRNA and 9 new miRNA, most of which have tissue specific expression patterns, were identified from wild soybean (Chen, et al. 2009). Transgenic studies of miR482, miR1512, and miR1515 showed that over-expression of these three microRNAs could lead to significant increases in nodule numbers in soybean (Li, et al. 2010).

### Status and Perspective

Significant progress has been made in soybean omics and biotechnology research in China during recent years. Chinese scientists in the field of functional genomics have discovered many soybean genes associated with seed quality attributes, stress resistance, and growth and development. The functions of some of these genes have been confirmed using transgenic soybean or model plants. At the same time, soybean transcriptomics, proteomics, and metabolomics research also are well under way in China. The results of these studies provide valuable information on the molecular bases of important agronomic traits in soybean cultivar varieties, and also a basis for improving important soybean traits using biotechnology. Although a number of genes/QTLs have been identified, few have been used in marker assisted selection (MAS) because only a few genes have been fine mapped compared to initial course mapping of the majority of genes. This also has negatively impacted map-based cloning. In China most soybean genes obtained to date have been identified from homology-based gene cloning approach. Functional verification of many of these genes has been limited to the model plants *Arabidopsis* and tobacco, while only a very small number of gene functions have been confirmed in soybean (Li, et al. 2006; Liu, et al. 2010; Zhu, et al. 2006). Meanwhile genotyping soybean germplasms and cultivars still are limited to using cloned genes (Tian, et al. 2010) that could influence an accelerated pace of molecular breeding (Qiu, et al. 2007). China's research on soybean transcriptomics, proteomics, and

metabolomics still is in its initial stage. The advancement of biotechnology will facilitate discovery of new genes and how they function to confer important agronomic traits. In addition, emerging new biotechnologies will further promote future soybean research in China.

### Acknowledgements

This work was supported by the State High-tech R&D Program of China (863 Program) (No.2006AA100104, No.2006AA10A111, No.2006AA10B1Z3, No.2006AA10Z1F1, No.2007AA10Z189, and No.2006AA10Z164), National Natural Science Foundation of China (No.30490251, No.30671310, No.30490250, and No.30471096), National Basic Research Program of China (973 Program) (No.2004CB117203, No.2009CB118400, and No.2010CB-125900), National Key Technologies R&D Program in the 11th Five-Year Plan (No.2006BAD13B05), National Genetic Modified Organism Program, International collaboration program (948) (No.2006-G1) and Crop Germplasm Conservation Program. We thank Dr. Pengyin Chen (University of Arkansas, USA) for critical reading and commenting on the manuscript.

### References

- Bo YP, Wang GK, Hu GW, Sun HY, Ren Y, Li H, Li MC, Xing LJ (2003) Introduction of  $\Delta^6$ -fatty acid desaturase gene from *Mortierella isabellina* into soybeans by *Agrobacterium* infection. *Chin J Biotechnol* 13:6-8
- Chen DG, Sun WL, Fan YL (1990) 2D-PAGE comparison of *in vitro* translation products of mRNAs from different tissues of soybean. *Chin J Biotechnol* 6:307-310
- Chen Y, Wang QY, Cheng XG, Xu ZS, Li LC, Ye XG, Xia LQ, Ma YZ (2007) GmDREB2, a soybean DRE-binding transcription factor, conferred drought and high-salt tolerance in transgenic plants. *Biochem Biophys Res Commun* 353:299-305
- Chen M, Xu Z, Xia L, Li L, Cheng X, Dong J, Wang Q, Ma Y (2009) Cold-induced modulation and functional analyses of the DRE-binding transcription factor gene, *GmDREB3*, in soybean (*Glycine max* L.). *J Exp Bot* 60:121-135
- Chen P, Zhang L, Qiu LJ, Lu B (2009) Proteomic analysis of differentiation between cytoplasmic-nuclear male-sterile line W931A and its maintainer in soybean. *Crops* 6:17-21
- Chen R, Hu Z, Zhang H (2009) Identification of microRNAs in wild soybean (*Glycine soja*). *J Integr Plant Biol* 51:1071-1079
- Chen XJ, Ye CJ, Lu HY, Xu MX, Li W, Zhang LM, Wang C, Luo SP, Zhu BG (2006) Cloning of *GmHSFA1* gene and its overexpression leading to enhancement of heat tolerance in transgenic soybean. *Hereditas (Beijing)* 28:1411-1420
- Chen YD, Yu DY (2009) Construction of mutant pools for soybean "Nannong 94-16" induced by EMS and analysis of SSR marker on several mutants. *Soybean Sci* 28:574-577
- Cheng JP, Chen B, Xu YM, Zhou XA, Lu CM, Wei WH (2007) Construction of suppression subtractive cDNA library and analysis of differentially expressed genes during maturing period seed of high and low oil content soybean variety. *Chin J Oil Crop Sci* 29:365-371
- Cui SY, Chen HT, Yu DY (2007) Mapping of QTL associated with leaf senescence in soybean. *Scientia Agri Sin* 40:2103-2108
- Cui SY, Geng LY, Meng QC, Yu DY (2007) Mapping of phosphorus deficiency tolerance in soybean (*Glycine max* L.) during seedling stage. *Acta Agron Sin* 33:378-388

- Dang W, Wei ZM (2007) Efficient *Agrobacterium*-mediated transformation of soybean. *J Mol Cell Biol* 40:185-195
- Dong X, Sun Q, Wei D, Li J, Tang B, Jia Q, Hu W, Zhao Y, Hua ZC (2007) A novel ferritin gene, *SferH-5*, reveals heterogeneity of the 26.5-kDa subunit of soybean (*Glycine max*) seed ferritin. *FEBS Lett* 581:5796-5802
- Dong ZM, Li YH, Zhang BS, Guan RX, Chang RZ, Qiu LJ (2006) Construction and characterization of cDNA library from leaf of soybean. *Crops* 5:1-4
- Du QL, Cui WZ, Zhang CH, Yu DY (2010) *GmRFP1* encodes a previously unknown RING-type E3 ubiquitin ligase in soybean (*Glycine max*). *Mol Biol Rep* 37:685-693
- Du SW, Liu YL, Yao BC, Bai C, Liu CY, Chen QS, Hu GH (2010) Optimization of soybean transformation system and transferring *Dof4* gene in soybean. *Soybean Sci* 29:398-402
- Duan MP, Yang SP, Yu DY, Gai JY (2005) Genetic analysis and SSR markers of the male sterile gene of the mutant NJ89-1 in soybeans. *Jiangsu Agri Sci* 5:19-21
- Frankel OH, Brown AHD (1984) Current plant genetic resources: a critical appraisal. In: *Genetics: New Frontiers (Vol IV)* Oxford and IBH publishing, New Delhi, India
- Gong XC, Ji J, Kang YH, Wang G, Wu Y, Wang P (2005) Genetic transformation of phytoene synthase gene (*PSY*) into soybean. *Soybean Sci* 24:30-33
- Gong XC, Ji J, Wang G, Wang P (2006) Genetic transformation of *PSY* into soybean by particle bombardment. *Soybean Sci* 25:137-140
- Guo B, Qiu LJ, Shao GH, Chang RZ, Liu LH, Xu ZY, Li XH, Sun JY (2000) Tagging salt tolerant gene using PCR markers in soybean. *Scientia Agri Sin* 33:10-16
- Guo DQ, Yang XD, Bao SJ, Guo SD, Kang JS, Yin AP (2008) Synchronous expression of *CryIA* and *CpTI* genes in soybean and analysis of their resistance to insect pests. *Scientia Agri Sin* 41:2957-2962
- Guo YS, Zhang ZY, Zhu YM, Li J, Bo X, Zhang SZ, Wu SY, Li HY (2006) Obtainment of transgenic soybean plants with chitinase and ribosome inactivating protein genes and their resistance identification. *Acta Agron Sin* 32:1841-1847
- Han SY, Yang ML, Chen YD, Yu JJ, Zhao TJ, Gai JY, Yu DY (2008) Construction of mutant library for soybean "Nannong 94-16" and analysis some characters. *J Nucl Agric Sci* 22:131-135
- Han SY, Zhang HY, Yang ML, Zhao TJ, Gai JY, Yu DY (2007) Screening of mutants and construction of mutant population in soybean "Nannong 86-4". *Acta Agron Sin* 33:2059-2062
- Han YP, Li WB, Yu K, Anderson TR, Poysa V, Weng JZ, Gao JG (2006) Mapping quantitative trait loci influencing tolerance to *Phytophthora* root rot in soybean. *Soybean Sci* 25:23-27
- He CY, Tian AG, Zhang JS, Zhang ZY, Gai JY, Chen SY (2003) Isolation and characterization of a full-length resistance gene homolog from soybean. *Theor Appl Genet* 106:786-793
- He CY, Zhang JS, Chen SY (2002) A soybean gene encoding a proline-rich protein is regulated by salicylic acid, an endogenous circadian rhythm and by various stresses. *Theor Appl Genet* 104:1125-1131
- Huang F, Chi Y, Meng Q, Gai J, Yu D (2006) *GmZFPI* encoding a single zinc finger protein is expressed with enhancement in reproductive organs and late seed development in soybean (*Glycine max*). *Mol Biol Rep* 33:279-285
- Huang F, Chi YJ, He H, Yu DY (2010) Cloning and characterization of an aldehyde dehydrogenase gene *GmALDH3-1* from soybean. *Hereditas (Beijing)* 32:492-497
- Huang F, Zhao TJ, Meng QC, Zhang YM, Gai JY, Yu DY (2006) Genetic analysis and RAPD marker of the genes for brachytic stem trait in soybean. *Hereditas (Beijing)* 28:55-56
- Huang G, Meng QC, Zhao TJ, Gai JY, Yu DY (2005) Genetic analysis and RAPD markers of genes related to short petiole in soybean. *Acta Agron Sin* 31:818-820
- Huang ZW, Zhao TJ, Yu DY, Chen SY, Gai JY (2008) Lodging resistance indices and related QTLs in soybean. *Acta Agron Sin* 34:605-611
- Ji W, Li Y, Li J, Dai CH, Wang X, Bai X, Cai H, Yang L, Zhu YM (2006) Generation and analysis of expressed sequence tags from NaCl-treated *Glycine soja*. *BMC Plant Biol* 6:4
- Ji W, Zhu Y, Li Y, Yang L, Zhao X, Cai H, Bai X (2010) Over-expression of a glutathione S-transferase gene, *GsGST*, from wild soybean (*Glycine soja*) enhances drought and salt tolerance in transgenic tobacco. *Biotechnol Lett* 32:1173-1179
- Jiang YN, Wang B, Wu TL (2009) Response of enzyme activity and secondary metabolites of different soybean genotypes to *Aphis glycines* matsumura invasion. *Soybean Sci* 28:103-107
- Li C, Yan XH, Zhou XA, Sha AH, Shan ZH, Zhou R, Wei WH (2010) Construction of a normalized full-length cDNA library of soybean seed at different developmental stages. *Scientia Agri Sin* 43:462-467
- Li GJ, Li HN, Cheng LG, Zhang YM (2009) Mapping quantitative trait loci for resistance in soybean to bean pyralid (*Lamprosema indicata* Fabricius). *Chin J Oil Crop Sci* 31:365-369
- Li H, Deng Y, Wu T, Subramanian S, Yu O (2010) Misexpression of miR482, miR1512, and miR1515 increases soybean nodulation. *Plant Physiol* 153:1759-1770
- Li L, Wang X, Gai J, Yu D (2007) Molecular cloning and characterization of a novel microsomal oleate desaturase gene from soybean. *J Plant Physiol* 164:1516-1526
- Li L, Wang X, Gai J, Yu D (2008) Isolation and characterization of a seed-specific isoform of microsomal omega-6 fatty acid desaturase gene (*FAD2-1B*) from soybean. *DNA Seq* 19:28-36
- Li WY, Shao G, Lam HM (2008) Ectopic expression of *GmPAP3* alleviates oxidative damage caused by salinity and osmotic stresses. *New Phytol* 178:80-91
- Li WY, Wong FL, Tsai SN, Phang TH, Shao G, Lam HM (2006) Tonoplast-located *GmCLC1* and *GmNHX1* from soybean enhance NaCl tolerance in transgenic bright yellow (BY)-2 cells. *Plant Cell Environ* 29:1122-1137
- Li XP, Gan R, Li PL, Ma YY, Zhang LW, Zhang R, Wang Y, Wang NN (2006) Identification and functional characterization of a leucine-rich repeat receptor-like kinase gene that is involved in regulation of soybean leaf senescence. *Plant Mol Biol* 61:829-844
- Li XP, Tian AG, Luo GZ, Gong ZZ, Zhang JS, Chen SY (2005) Soybean DRE-binding transcription factors that are responsive to abiotic stresses. *Theor Appl Genet* 110:1355-1362
- Liao H, Wong FL, Phang TH, Cheung MY, Li WY, Shao G, Yan X, Lam HM (2003) *GmPAP3*, a novel purple acid phosphatase-like gene in soybean induced by NaCl stress but not phosphorus deficiency. *Gene* 318:103-111
- Liao Y, Zhang JS, Chen SY, Zhang WK (2008) Role of soybean *GmbZIP132* under abscisic acid and salt stresses. *J Integr Plant Biol* 50:221-230
- Liao Y, Zou HF, Wang HW, Zhang WK, Ma B, Zhang JS, Chen SY (2008) Soybean *GmMYB76*, *GmMYB92*, and *GmMYB177* genes confer stress tolerance in transgenic *Arabidopsis* plants. *Cell Res* 18:1047-1060
- Liao Y, Zou HF, Wei W, Hao YJ, Tian AG, Huang J, Liu YF, Zhang JS, Chen SY (2008) Soybean *GmbZIP44*, *GmbZIP62* and *GmbZIP78* genes function as negative regulator of ABA signaling and confer salt and freezing tolerance in transgenic *Arabidopsis*. *Planta* 228:225-240

- Liu B, Watanabe S, Uchiyama T, Kong F, Kanazawa A, Xia Z, Nagamatsu A, Arai M, Yamada T, Kitamura K, Masuta C, Harada K, Abe J (2010) The soybean stem growth habit gene *Dt1* is an ortholog of *Arabidopsis TERMINAL FLOWER1*. *Plant Physiol* 153:198-210
- Liu CY, Wang WQ, Chen QS, Yang CP, Li WB, Xin DW, Jin ZG, Song YB (2005) Construction and primary analysis of subtractive library induced by soybean mosaic virus (SMV). *Chin J Biotechnol* 21:320-322
- Liu H, Wang H, Gao P, Xu J, Xu T, Wang J, Wang B, Lin C, Fu YF (2009) Analysis of clock gene homologs using unifoliolates as target organs in soybean (*Glycine max*). *J Plant Physiol* 166:278-289
- Liu H, Wang H, Li Q, Xu P, Gai JY, Yu DY (2005) Inheritance analysis and mapping QTLs related to cotton worm resistance in soybeans. *Scientia Agri Sin* 38:1369-1372
- Liu J, Ha D, Xie Z, Wang C, Wang H, Zhang W, Zhang J, Chen S (2008) Ectopic expression of soybean *GmKNT1* in *Arabidopsis* results in altered leaf morphology and flower identity. *J Genet Genomics* 35:441-449
- Liu SH, Zhou RB, Yu DY, Chen SY, Gai JY (2009) QTL mapping of protein related traits in soybean (*Glycine max* (L.) Merr.). *Acta Agron Sin* 35:2139-2149
- Liu Y, Gai JY, Lv HN, Wang YJ, Chen SY (2005) Identification of drought tolerant germplasm and inheritance and QTL mapping of related root traits in soybean (*Glycine max* (L.) Merr.). *Acta Genet Sin* 32:855-863
- Lu B, Fang XJ (2003) Express analysis of cDNA induced by infecting roots of soybean cyst nematode race 4 in soybean ZD2315. *Mol Plant Breed* 1:193-200
- Lu WG, Gai JY, Zheng YZ, Li WD (2008) Construction of a soybean genetic linkage map and mapping QTLs resistant to soybean cyst nematode (*Heterodera glycines* Ichinohe). *Acta Agron Sin* 32:1272-1279
- Luo GZ, Wang HW, Huang J, Tian AG, Wang YJ, Zhang JS, Chen SY (2005) A putative plasma membrane cation/proton antiporter from soybean confers salt tolerance in *Arabidopsis*. *Plant Mol Biol* 59:809-820
- Ma J, Zhang J, Qu J, Wang YP, Wei YF, Wang PW (2009) Development of novel soybean germplasms with low activity of lipoxygenases by RNAi method. *Scientia Agri Sin* 42:3804-3811
- Ma JL, Li C, Wei Y, Zhao Y (2007) The pathway of isoflavone biosynthesis and its regulation. *J Northeast Agri Univ* 38:692-696
- Meng Q, Zhang C, Gai J, Yu D (2007) Molecular cloning, sequence characterization and tissue-specific expression of six NAC-like genes in soybean (*Glycine max* (L.) Merr.). *J Plant Physiol* 164:1002-1012
- Ning H, Zhang C, Yao Y, Yu D (2010) Overexpression of a soybean O-acetylserine (thiol) lyase-encoding gene *GmOASTLA* in tobacco increases cysteine levels and enhances tolerance to cadmium stress. *Biotechnol Lett* 32:557-564
- Nishiyama T, Fujita T, Shin IT, Seki M, Nishide H, Uchiyama I, Kamiya A, Carninci P, Hayashizaki Y, Shinozaki K, Kohara Y, Hasebe M (2003) Comparative genomics of *Physcomitrella patens* gametophytic transcriptome and *Arabidopsis thaliana*: implication for land plant evolution. *Proc Natl Acad Sci U S A* 100:8007-8012
- Qiu LJ, Chang RZ (2010) The origin and history of soybean. In: Singh G (ed) *Soybean: Botany, Production and Uses* CAB International Publishing, UK, p 1-13:
- Qiu LJ, Chen PY, Liu Z, X., Li YH, Guan RX, Wang LH, Chang RZ (2010) Soybean germplasm collection and exploration in China and their utilization in the world. *Plant Genet Resour Special Issue (Legum)* In press:
- Qiu LJ, Li YH, Guan RX, Liu ZX, Wang LX, Chang RZ (2009) Establishment, representative testing and research progress of soybean core collection and mini core collection. *Acta Agron Sin* 35:571-579
- Qiu LJ, Wang CL, Zhou GA, Chen SY, Chang RZ (2007) Soybean molecular breeding. *Scientia Agri Sin* 40:2418-2436
- Schmutz J, Cannon SB, Schlueter J, Ma J, Mitros T, Nelson W, Hyten DL, Song Q, Thelen JJ, Cheng J, Xu D, Hellsten U, May GD, Yu Y, Sakurai T, Umezawa T, Bhattacharyya MK, Sandhu D, Valliyodan B, Lindquist E, Peto M, Grant D, Shu S, Goodstein D, Barry K, Futrell-Griggs M, Abernathy B, Du J, Tian Z, Zhu L, Gill N, Joshi T, Libault M, Sethuraman A, Zhang XC, Shinozaki K, Nguyen HT, Wing RA, Cregan P, Specht J, Grimwood J, Rokhsar D, Stacey G, Shoemaker RC, Jackson SA (2010) Genome sequence of the palaeopolyploid soybean. *Nature* 463:178-183
- Shan DP, Zhu RS, Chen LJ, Qi ZM, Liu CY, Hu GH, Chen QS (2009) Epistatic effects and QE interaction effects of QTLs for protein content in soybean. *Acta Agron Sin* 35:45-47
- Shannon S, Meeks-Wagner DR (1991) A mutation in the *Arabidopsis TFL1* gene affects inflorescence meristem development. *Plant Cell* 3:877-892
- Shih CH, Chen Y, Wang M, Chu IK, Lo C (2008) Accumulation of isoflavone genistin in transgenic tomato plants overexpressing a soybean isoflavone synthase gene. *J Agric Food Chem* 56:5655-5661
- Singh RJ, Hymowitz T (1999) Soybean genetic resources and crop improvement. *Genome* 42:605-616
- Song WW, Liu BH, Yang ML, Zhang Y, Wang JA (2008) Transformation of *SPS* gene into soybean via *Agrobacterium*-mediated method. *Soybean Sci* 27:387-390
- Sun GZ, Zhu ZD, Wu XF, Cui YL, Wang XM (2008) Comparative proteomics of *Arabidopsis* induced by salicylic acid, wound and *Phytophthora sojae*. *Scientia Agri Sin* 4:1030-1039
- Sun JM, Wu SM, Tao WJ, Ding AL, Han FX, Jia SR (2004) Detection of RAPD markers linked to gene *lx\_1* in soybean. *Scientia Agri Sin* 37:170-174
- Tian AG, Luo GZ, Wang YJ, Zhang JS, Gai JY, Chen SY (2004) Isolation and characterization of a *Pti1* homologue from soybean. *J Exp Bot* 55:535-537
- Tian AG, Wang J, Cui P, Han YJ, Xu H, Cong LJ, Huang XG, Wang XL, Jiao YZ, Wang BJ, Wang YJ, Zhang JS, Chen SY (2004) Characterization of soybean genomic features by analysis of its expressed sequence tags. *Theor Appl Genet* 108:903-913
- Tian AG, Zhao JY, Zhang JS, Gai JY, Chen SY (2004) Genomic characterization of the S-adenosylmethionine secarboxylase genes from soybean. *Theor Appl Genet* 108:842-850
- Tian Z, Wang X, Lee R, Li Y, Specht JE, Nelson RL, McClean PE, Qiu L, Ma J (2010) Artificial selection for determinate growth habit in soybean. *Proc Natl Acad Sci U S A* 107:8563-8568
- Wang B, Wang Y, Wang Q, Luo G, Zhang Z, He C, He SJ, Zhang J, Gai J, Chen S (2004) Characterization of an NBS-LRR resistance gene homologue from soybean. *J Plant Physiol* 161:815-822
- Wang C, Wang H, Zhang J, Chen S (2008) A seed-specific AP2-domain transcription factor from soybean plays a certain role in regulation of seed germination. *Sci China C Life Sci* 51:336-345
- Wang CL, Zhao J, Li YH, Fan YL, Zhang LJ, Liu ZX, Guan RX, Lv SX, Chang RZ, Qiu LJ (2008) Transforming transcription factor *ABP9* into soybean and optimization of the transformation system. *Scientia Agri Sin* 41:1908-1916

- Wang F, Zhao TJ, Yu DY, Chen SY, Gai JY (2008) Inheritance and QTL analysis of submergence tolerance at seedling stage in soybean [*Glycine max* (L.) Merr.]. *Acta Agron Sin* 34:748-753
- Wang HW, Zhang JS, Gai JY, Chen SY (2006) Cloning and comparative analysis of the gene encoding diacylglycerol acyltransferase from wild type and cultivated soybean. *Theor Appl Genet* 112:1086-1097
- Wang LX, Guan Y, Guan RX, Li YH, Ma YS, Dong ZM, Liu X, Zhang HY, Zhang YQ, Liu ZX, Chang RZ, Xu HM, Li LH, Lin FY, Luan WJ, Ya NZ, Ning XC, Zhu L, Cui YH, Piao RH, Liu Y, Chen PY, Qiu LJ (2006) Establishment of Chinese soybean (*Glycine max*) core collections with agronomic traits and SSR markers. *Euphytica* 151: 215-223
- Wang M, Liu P, Shi MW, Wang QL (2005) Construction and characterization of cDNA library from seeds of wild soybean (*Glycine Soja*). *Lett Biotechnol* 16:509-511
- Wang P, Gao SQ, Guo YL, Cheng XG, Wang G, Ma YZ, Ji J (2008) Transformation of stress resistance related gene *GmDREB* into soybean via *Agrobacterium*-mediation. *Soybean Sci* 27:47-51
- Wang XC, Wang G, Ji J, Wang P (2007) A study of *Bt* gene transformation into smatic microprojectile bombardment. *Soybean Sci* 26:140-143
- Wang Y, Li P, Cao X, Wang X, Zhang A, Li X (2009) Identification and expression analysis of miRNAs from nitrogen-fixing soybean nodules. *Biochem Biophys Res Commun* 378:799-803
- Wang YJ, Li YD, Luo GZ, Tian AG, Wang HW, Zhang JS, Chen SY (2005) Cloning and characterization of an HDZip I gene *GmHZI* from soybean. *Planta* 221:831-843
- Wang ZX, Zhou R, Zhang XJ, Shan ZH, Sha AH, Chen HF, Qiu DZ, Li PW, Zhou XA (2008) Map construction and QTL analysis of oil content in soybean. *Chin J Oil Crop Sci* 30:272-278
- Wei L, Sun HY, Tang XF, Yang Z, Gao MJ, Liu LJ (2009) Selection of genetic population and obtaining of the quality QTLs in soybean. *Mol Plant Breed* 4:727-735
- Wei W, Huang J, Hao YJ, Zou HF, Wang HW, Zhao JY, Liu XY, Zhang WK, Ma B, Zhang JS, Chen SY (2009) Soybean GmPHD-type transcription regulators improve stress tolerance in transgenic *Arabidopsis* plants. *PLoS One* 4:e7209
- Wu C, Ma Q, Yam KM, Cheung MY, Xu Y, Han T, Lam HM, Chong K (2006) *In situ* expression of the *GmNMH7* gene is photoperiod-dependent in a unique soybean [*Glycine max* (L.) Merr.] flowering reversion system. *Planta* 223:725-735
- Wu J, Wu Q, Gai J, Yu D (2008) Constitutive overexpression of AOS-like gene from soybean enhanced tolerance to insect attack in transgenic tobacco. *Biotechnol Lett* 30:1693-1698
- Wu W, Zhang Q, Zhu Y, Lam HM, Cai Z, Guo D (2008) Comparative metabolic profiling reveals secondary metabolites correlated with soybean salt tolerance. *J Agric Food Chem* 56:11132-11138
- Wu XL, He CY, Wang YJ, Zhang ZY, Dongfang Y, Zhang JS, Chen SY, Gai JY (2001) Construction and analysis of a genetic linkage map of soybean. *Acta Genet Sin* 28:1051-1061
- Xie ZM, Zou HF, Lei G, Wei W, Zhou QY, Niu CF, Liao Y, Tian AG, Ma B, Zhang WK, Zhang JS, Chen SY (2009) Soybean trihelix transcription factors GmGT-2A and GmGT-2B improve plant tolerance to abiotic stresses in transgenic *Arabidopsis*. *PLoS One* 4:e6898
- Xiong DJ, Zhao TJ, Gai JY (2008) Geographical sources of germplasm and their nuclear and cytoplasmic contribution to soybean cultivars released during 1923 to 2005 in China. *Acta Agron Sin* 34: 175-183
- Xu P, Wang H, Li Q, Gai JY, Yu DY (2007) Mapping QTLs related to oil content of soybean. *Hereditas* (Beijing) 29:92-96
- Xu XY, Zheng R, Li CM, Gai JY, Yu DY (2006) 2-DE analysis of proteins during seed germination in soybean. *Soybean Sci* 25:355-359
- Xu XY, Zheng R, Li CM, Gai JY, Yu DY (2006) Differential proteomic analysis of seed germination in soybean. *Prog Biochem Biophys* 33:1106-1111
- Yang GC, Liu HL, Wang SF, Liu SL, Wei ZY, Cai QA, Xing SC (2009) Construction and identification of cDNA libraries of soybean treated by drought and low temperature. *Heilongjiang Agri Sci* 2:1-3
- Yang H, Huang Y, Zhi H, Yu D (2010) Proteomics-based analysis of novel genes involved in response toward soybean mosaic virus infection. *Mol Biol Rep* 38:511-521
- Yang Z, Guan RX, Wang YQ, Liu ZX, Chang RZ, Wang SM, Qiu LJ (2004) Construction of genetic map and QTL analysis for some agronomic traits in soybean. *J Plant Genet Resour* 5:309-314
- Yin Z, Meng F, Song H, Wang X, Xu X, Yu D (2010) Expression quantitative trait loci analysis of two genes encoding rubisco activase in soybean. *Plant Physiol* 152:1625-1637
- Zeng WY, Yang SP, Yu DY, Gai JY (2007) A comparative study on anther proteomics between cytoplasmic-nuclear male-sterile line NJCMS2A and its maintainer of soybean. *Acta Agron Sin* 33:1637-1643
- Zeng WY, Yang SP, Yu DY, Gai JY (2007) Proteomic study of anther differentiation between cytoplasmic-nuclear male-sterile line NJCMS1A and its maintainer in soybean [*Glycine max* (L.) Merr.]. *Scientia Agric Sin* 40:2679-2687
- Zhang BH, Wang QL, Wang KB, Pan XP, Liu F, Guo TL, Cobb GP, Anderson TA (2007) Identification of cotton microRNAs and their targets. *Gene* 397:26-37
- Zhang BH, Pan XP, Stellwag EJ (2008) Identification of soybean microRNAs and their targets. *Planta* 229:161-182
- Zhang BH, Wang QL, Pan XP (2007) MicroRNAs and their regulatory roles in animals and plants. *J Cell Biochem* 210:279-289
- Zhang C, Meng Q, Gai J, Yu D (2008) Cloning and functional characterization of an O-acetylserine(thiol)lyase-encoding gene in wild soybean (*Glycine soja*). *Mol Biol Rep* 35:527-534
- Zhang G, Chen M, Chen X, Xu Z, Li L, Guo J, Ma Y (2010) Isolation and characterization of a novel EAR-motif-containing gene *GmERF4* from soybean (*Glycine max* L.). *Mol Biol Rep* 37:809-818
- Zhang G, Chen M, Li L, Xu Z, Chen X, Guo J, Ma Y (2009) Overexpression of the soybean *GmERF3* gene, an AP2/ERF type transcription factor for increased tolerances to salt, drought, and diseases in transgenic tobacco. *J Exp Bot* 60:3781-3796
- Zhang M, Li K, Zhang C, Gai J, Yu D (2009) Identification and characterization of class I *DXS* gene encoding 1-deoxy-D-xylulose-5-phosphate synthase, the first committed enzyme of the MEP pathway from soybean. *Mol Biol Rep* 36:879-887
- Zhang MC, Li Z, H., Tian XL, Duan LS, Wang BM, Zhai ZX, He ZP (2004) Effect of plant growth regulator SHK-6 on nitrogen metabolism of soybean leaf. *Soybean Sci* 23:15-20
- Zhang Q, Li H, Li R, Hu R, Fan C, Chen F, Wang Z, Liu X, Fu Y, Lin C (2008) Association of the circadian rhythmic expression of *GmCRY1a* with a latitudinal cline in photoperiodic flowering of soybean. *Proc Natl Acad Sci U S A* 105:21028-21033
- Zhang SZ, Xu PF, Wu JJ, Li WB, Chen WY, Chen C, Yu AL,

- Wang JS, Jin LM, Li C (2008) Construction and analysis of a SSH library of soybean upon infection with *Phytophthora sojae*. *Soybean Sci* 27:543-545
- Zhang ZC, Zhan XL, Chen QS, Teng WL, Yang QK, Li WB (2004) QTL mapping of seed oil and protein content of soybean. *Soybean Sci* 23:81-85
- Zhao L, Luo QL, Yang CL, Han YP, Li WB (2007) Construction and analysis of differentially expressed cDNA library from soybean induced by darkness. *Soybean Sci* 26:134-139
- Zhao LM, Zheng DF, Feng NJ, Du JD, Zhang YX (2008) Effects of different growth regulators (PGRs) on metabolism in soybean roots. *Soybean Sci* 27:242-246
- Zhen Y, Qi JL, Wang SS, Su J, Xu GH, Zhang MS, Miao L, Peng XX, Tian D, Yang YH (2007) Comparative proteome analysis of differentially expressed proteins induced by Al toxicity in soybean. *Physiol Plant* 4:542-554
- Zheng CM, Chang RZ, Qiu LJ (2001) Inheritance of resistance to SMV3 and identification of RAPD marker linked to the resistance gene in soybean. *Scientia Agri Sin* 34:1-4
- Zheng R, Xu XY, Li CM, Gai JY, Yu DY (2008) Proteomic analysis of differentially expressed proteins in developmental soybean [*Glycine max* (L.) Merr.] seed. *Soybean Sci* 27:556-563
- Zheng YZ, Gai JY, Lu WG, Li WD, Zhou RB, Tian SJ (2006) QTL mapping for fat and fatty acid composition contents in soybean. *Acta Agron Sin* 12:1823-1830
- Zhou GA, Chang RZ, Qiu LJ (2010) Overexpression of soybean ubiquitin-conjugating enzyme gene *GmUBC2* confers enhanced drought and salt tolerance through modulating abiotic stress-responsive gene expression in *Arabidopsis*. *Plant Mol Biol* 72:357-367
- Zhou QY, Tian AG, Zou HF, Xie ZM, Lei G, Huang J, Wang CM, Wang HW, Zhang JS, Chen SY (2008) Soybean WRKY-type transcription factor genes, *GmWRKY13*, *GmWRKY21*, and *GmWRKY54*, confer differential tolerance to abiotic stresses in transgenic *Arabidopsis* plants. *Plant Biotechnol J* 6:486-503
- Zhu B, Ye C, Lu H, Chen X, Chai G, Chen J, Wang C (2006) Identification and characterization of a novel heat shock transcription factor gene, *GmHsfA1*, in soybeans (*Glycine max*). *J Plant Res* 119:247-256
- Zhu BG, Bai HX, Zhang Y (2001) Genetic analysis of mutant genes and RAPD marker of mutant gene of narrow leaflet in soybean. *Acta Agron Sin* 28:64-68