



Integrating omic approaches for abiotic stress tolerance in soybean

Rupesh Deshmukh, Humira Sonah, Gunvant Patil, Wei Chen, Silvas Prince, Raymond Mutava, Tri Vuong, Babu Valliyodan and Henry T. Nguyen*

National Center for Soybean Biotechnology and Division of Plant Sciences, University of Missouri, Columbia, MO, USA

Edited by:

Rajeev K. Varshney, International Crops Research Institute for the Semi-Arid Tropics, India

Reviewed by:

Paula Casati, Centro de Estudios Fotosintéticos-CONICET, Argentina
Iain Robert Searle, The University of Adelaide, Australia

*Correspondence:

Henry T. Nguyen, National Center for Soybean Biotechnology and Division of Plant Sciences, University of Missouri, 1-31 Agriculture Building, Columbia, MO 65211-7140, USA
e-mail: nguyenhenny@missouri.edu

Soybean production is greatly influenced by abiotic stresses imposed by environmental factors such as drought, water submergence, salt, and heavy metals. A thorough understanding of plant response to abiotic stress at the molecular level is a prerequisite for its effective management. The molecular mechanism of stress tolerance is complex and requires information at the omic level to understand it effectively. In this regard, enormous progress has been made in the omics field in the areas of genomics, transcriptomics, and proteomics. The emerging field of ionomics is also being employed for investigating abiotic stress tolerance in soybean. Omic approaches generate a huge amount of data, and adequate advancements in computational tools have been achieved for effective analysis. However, the integration of omic-scale information to address complex genetics and physiological questions is still a challenge. In this review, we have described advances in omic tools in the view of conventional and modern approaches being used to dissect abiotic stress tolerance in soybean. Emphasis was given to approaches such as quantitative trait loci (QTL) mapping, genome-wide association studies (GWAS), and genomic selection (GS). Comparative genomics and candidate gene approaches are also discussed considering identification of potential genomic loci, genes, and biochemical pathways involved in stress tolerance mechanism in soybean. This review also provides a comprehensive catalog of available online omic resources for soybean and its effective utilization. We have also addressed the significance of phenomics in the integrated approaches and recognized high-throughput multi-dimensional phenotyping as a major limiting factor for the improvement of abiotic stress tolerance in soybean.

Keywords: abiotic stress tolerance, soybean, genomics, proteomics, transcriptomics, ionomics, phenomics

INTRODUCTION

Soybean is the most important legume crop which provides sources of oil and protein for human as well as for livestock. Soybean also enhances soil fertility because of the symbiotic nitrogen fixing ability. Soybean contributed to more than 50% of globally consumed edible oil (SoyStats, 2013¹). Apart from the consumption, soybean oil is being considered as a future source of fuel and efforts are being made to improve soy-diesel production (Candeia et al., 2009). Soybean protein-based bio-degradable materials are also being considered as an alternative for plastics (Song et al., 2011). Soybean products are gaining attention because of its pharmaceutical attributes such as anti-cancerous properties (Ko et al., 2013). Such diverse uses of soybean make it a more widely desired crop plant and are rapidly increasing its demand. In this regard, soybean yield improvement has been achieved by 1.3% per year (Ray et al., 2013). However, the increasing global population will need double the current food production by the year 2050 and at the current rate it can achieve only ~55% (Ray et al., 2013). It may be more difficult to produce sufficient yield with the changing climate. Therefore soybean yield prediction must consider the ongoing challenges of extreme

weather such as drought, flood, heat, cold, frost, and possible UV stress.

Abiotic stresses are the most challenging of all major constraints in crop production. Soybean production is not only influenced by environmental factors, such as drought, water submergence, salt, and heavy metals, but it also faces challenges to get adapted in non-traditional areas. This demands extensive breeding for the development of local cultivars (Tanksley and Nelson, 1996; Grainger and Rajcan, 2013). Direct selection for yield stability based on multi-location trials has been traditionally used for the development of varieties adapted to adverse environmental conditions. This approach is more difficult for abiotic stress related traits because of low heritability and highly influenced by environmental conditions (Manavalan et al., 2009). Direct selection is also a time-consuming and labor intensive process. Strategic marker-assisted breeding can efficiently accelerate the development of tolerant cultivars; however, it also necessitates knowledge about genomic loci governing the traits and the availability of tightly linked molecular markers (Xu et al., 2012). Molecular marker development has been accelerated with the availability of sequenced genomes and organelles in crop plants (Singh et al., 2010; Sonah et al., 2011a; Tomar et al., 2014).

Marker-assisted breeding has become sophisticated with the availability of complete soybean genome sequence due to

¹ Available online at: <http://www.soystats.com> (Accessed December 10, 2013).

subsequent development of locus-specific molecular markers (Schmutz et al., 2010; Song et al., 2010). Genome-wide high density markers availability also facilitates the haplotype analysis and identification of different alleles for agronomical important traits (Tardivel et al., 2014). Marker-assisted breeding has been carried-out mostly for simple traits governed by a single, or at most a few loci (Shi et al., 2009; Jun et al., 2012). Marker-assisted breeding also suffers due to undesired genetic drag (Tanksley and Nelson, 1996; Shi et al., 2009). The genetic background of the recurrent parent also plays an important role in the phenotypic expression of newly introgressed gene(s) mostly because of the complex epistatic interaction (Palloix et al., 2009). In the case of multiple complex traits, epistatic interaction is more unpredictable and it is hard to develop a strategic breeding plan until unless solid information is available about the molecular mechanisms involved in the trait development. Recent technological development in genomics provides tremendous power to predict genetic factors, their evolution, distribution, and interactions at great extent (Morrell et al., 2011; Sonah et al., 2011b). Genetic engineering is the most advanced approach that has been used for the genetic improvement of soybean. Genetically modified (GM) soybean crops for insect-resistance and herbicide-tolerance has covered most of the cultivated area in the world (Carpenter, 2010). Although, GM soybean has proven to be very successful, it raises ethical controversies, and it is available only for few traits (Carpenter, 2010). Integration of multi-disciplinary knowledge is required to design future soybean varieties with ideal plant types providing high and stable yield in adverse climatic conditions. In this context, a detailed review was made to evaluate progress achieved in different omic approaches and to highlight future perspectives for its effective exploration toward the development of abiotic stress tolerant soybean cultivars.

OMICS APPROACHES IN THE TECHNOLOGICAL ERA

Plant molecular biology aims to study cellular processes, their genetic control, and interactions with environmental changes. Such a multi-dimensional and detailed investigation requires large-scale experiments involving entire genetic, structural, or functional components. These large scale studies are called “omics.” Major components of omics include genomics, transcriptomics, proteomics, and metabolomics (Figure 1). These omics approaches are routinely used in various research disciplines of crop plants, including soybean. Omics approaches have improved very rapidly during the last decade as technology advances. Subsequently, high-throughput data developed by omic experiments require extensive computational resources for storage and analysis. Thus, several online databases, analysis servers, and omics platforms have been developed. Omics is getting broader coverage and it is anticipated that several new omic fields will evolve in near future.

GENOMICS ADVANCES FOR ABIOTIC STRESS TOLERANCE IN SOYBEAN

MOLECULAR MARKER RESOURCES

Genomic applications in soybean have become more standard with the availability of whole genome sequence (WGS) (Schmutz et al., 2010). The WGS provided the basis for the development of

thousands of simple sequence repeat (SSR) markers and millions of single nucleotide polymorphism (SNP) markers (Song et al., 2010; Sonah et al., 2013). Recent developments in next generation sequencing (NGS) technologies make sequencing-based genotyping cost effective and efficient. Three main complexity reduction methods, namely Reduced Representation Libraries (RRLs), Restriction site Associated DNA (RAD) sequencing, and Genotyping-by-Sequencing (GBS) are being routinely used. Among these, GBS is gaining more attention because of its simplified and cost effective methodology (Elshire et al., 2011; Sonah et al., 2012). The GBS approach has been successfully used in several crop species (Poland and Rife, 2012). Recently, GBS methodology has been improved and streamlined for soybean (Sonah et al., 2013). However, sequencing-based genotyping methods require computational expertise and significant time for data analysis. This restricts its use in marker-assisted breeding where timely selection is very important. GBS will be widely used in the future with an increasing number of software packages and computational pipelines (Sonah et al., 2013).

Technological advances have also provided a high-throughput, reliable, and quick array-based genotyping platforms. The SNP array development require initial information about SNPs, fortunately, information about millions of SNPs is already available in the public domain (Table 1). The Illumina Infinium array (SoySNP50K iSelect BeadChip) for ~50,000 SNPs has been successfully developed and used for the genotyping of several soybean plant introduction (PI) lines (Song et al., 2013). Technological advances beyond this make it possible to re-sequence hundreds of lines in a cost effective manner and has started a new era of genotyping by re-sequencing (Lam et al., 2010; Li et al., 2013; Xu et al., 2013). Now, the challenge for plant biologists is how to effectively use these resources for marker-assisted applications.

QTL MAPPING FOR ABIOTIC STRESS TOLERANCE IN SOYBEAN

Genetic fingerprinting, linkage mapping, and quantitative trait loci (QTL) mapping are marker based applications that have become more sophisticated with the availability of different genotyping platforms (Table 1). Consequently, several efforts have been made to identify QTL for abiotic stress tolerance in soybean (Table S1). QTL studies have identified thousands of QTL spanning the entire genome (www.soykb.org, www.soybase.org). This is due to the complex inheritance of abiotic stress tolerance which has identified unstable QTL across different environments. Further utilization of QTL information for marker-assisted breeding or candidate gene identification has become difficult due to this complexity. Statistical tools such as “Meta-QTL analysis” have been advanced that compile QTL data from different studies together on the same linkage map for identification of precise QTL region (Deshmukh et al., 2012; Sosnowski et al., 2012). Several efforts have been performed to identify meta-QTL for different agronomical and quantitative traits in soybean (Table 2). Meta-analysis studies are still required exclusively for abiotic traits.

GENOME-WIDE ASSOCIATION STUDIES (GWAS) IN SOYBEAN

QTL mapping using bi-parental populations has limitations because of restricted allelic diversity and genomic resolution.

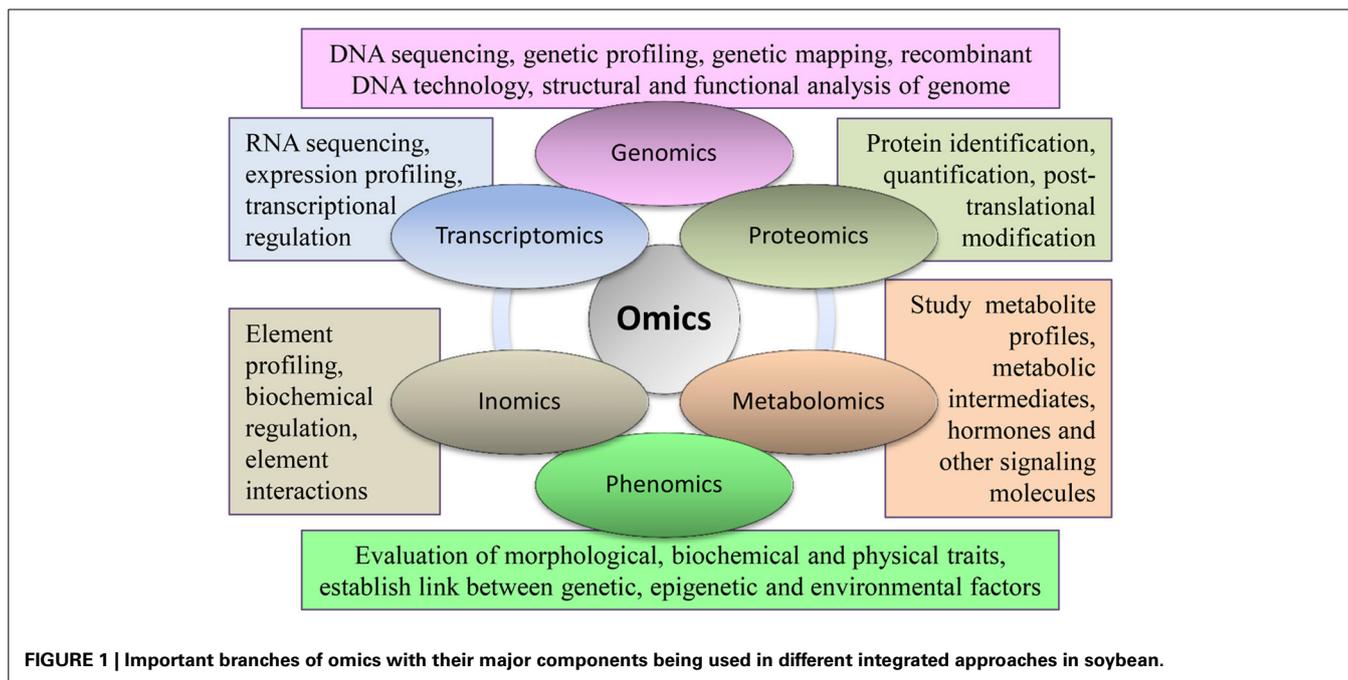


Table 1 | List of significant studies performed to develop SNP markers and subsequent genotyping using different technological platforms in soybean.

Sr. No	Genotyping platform/Approach	Genotypes	SNPs	References
1	Illumina GoldenGate assay	3 RIL mapping populations	384	Hyten et al., 2008
2	Illumina Infinium SoySNP6K BeadChip	92 RILs	5376	Akond et al., 2013
3	Illumina genome analyzer/Reduced Representation Libraries (RRLs)	5 diverse genotypes	14,550	Varala et al., 2011
4	Illumina GoldenGate assay	3 RIL mapping populations	1536	Hyten et al., 2010b; Vuong et al., 2010
5	Illumina genome analyzer /RRLs	444 RILs	25,047	Hyten et al., 2010a
6	Illumina GAIIX/Genotyping by sequencing (GBS)	8 diverse genotypes	10,120	Sonah et al., 2013
7	Illumina Genome Analyzer II/whole genome re-sequencing	17 wild and 14 cultivated	2,05,614	Lam et al., 2010
8	Illumina Genome Analyzer II/whole genome re-sequencing	25 diverse genotypes	51,02,244	Li et al., 2013
9	Illumina genome analyzer/RRLs	Parental lines of mapping population	39,022	Wu et al., 2010
10	Illumina Infinium BeadChip	96 each of landraces, elite cultivars and wild accessions	52,041	Song et al., 2013

The allelic diversity can be increased to some extent by using multi-parental crosses. Recently, Multi-parent Advanced Generation Inter-Cross populations (MAGIC) has been used to identify QTL for blast and bacterial blight resistance, salinity and submergence tolerance, and grain quality traits in rice (Bandillo et al., 2013). Such multi-parental populations has mapping resolution limitations since it depends on meiotic events (crossing-over) (Kover et al., 2009). In contrast, the genome-wide association study (GWAS) approach provides opportunities to explore the tremendous allelic diversity existing in natural soybean germplasm. Mapping resolution of GWAS is also higher since millions of crossing

events have been accumulated in the germplasm during evolution.

GWAS is routinely being used in many plant species, but only a few studies have been reported in soybean (Table S2). These studies were performed with limited markers and genotypes. GWAS in soybean is lagging behind compared to maize, mostly because of the slow linkage disequilibrium (LD) decay (Hyten et al., 2007; Mamidi et al., 2011). Another serious problem is the confounding population structure since it may cause spurious associations leading to an increased false-discovery rate (FDR). Studies that involve case-control phenotypes (binary) carefully relate the cases and controls to minimize confounding effects.

Table 2 | Meta-QTL studies performed for different traits in soybean.

Sr. No	Trait	Meta QTL	QTL compiled	Studies compiled	References
1	Soybean cyst nematode resistance	7	62	17	Guo et al., 2006
2	Soybean cyst nematode resistance	16	151	19	Zhang et al., 2010
3	Seed oil content	20	121	22	Qi et al., 2011b
4	Seed oil content	25	130	39	Qi et al., 2011a
5	100-seed weight	17	65	12	Zhao-Ming et al., 2009
6	100-seed weight	15	117	13	Sun et al., 2012a
7	Fungal disease resistance	23	107	23	Wang et al., 2010
8	Insect resistance	20	81	–	Jing et al., 2009
9	Seed protein content	23	107	29	Zhao-Ming et al., 2011
10	Plant height	12	93	13	Sun et al., 2012b
11	Phosphorus efficiency	29	96	–	Huang et al., 2011
12	Growth stages	9	98	10	Qiong et al., 2009

GWAS for quantitative traits like abiotic stress tolerance are predictable to be affected by a confounding population. Different models have been developed for population stratification and spurious allelic associations like MLM and CMLM which takes into account the population structure and kinship. Recently, GWAS for *Sclerotinia sclerotiorum* resistance was performed using 7864 SNPs in soybean (Bastien et al., 2014). The study provided details of a probable marker requirement and methodologies involving population stratification for effective GWAS (Bastien et al., 2014). Development in statistical tools, genotyping methods, and studies involving larger sets of genotypes will definitely improve GWAS power in soybean.

GENOMIC SELECTION (GS) IN SOYBEAN

Marker-assisted breeding for simple Mendelian traits are easy and effective, but it can be problematic for the complex traits such as abiotic stresses that are generally polygenic. Even major QTLs can explain only a small fraction of phenotypic variation and may show unexpected trait expression in new genetic backgrounds because of epistatic interactions. These limitations can be effectively addressed by the use of an approach called “Genomic selection” (GS). GS is relatively simple, more reliable, and a more powerful approach where breeding values of lines are predicted using their phenotypes and marker genotypes (Heffner et al., 2009). GS is more effective since it uses all marker information simultaneously to develop a prediction model avoiding biased marker effects (Heffner et al., 2009). GS captures small-effect QTL that governs most of the variation including epistatic interaction effects.

An overview of research articles regarding GS published during last decade showed exponential growth within recent years (Figure S1). The increasing popularity of GS among plant as well as animal breeders is mostly because of the reduced cost of genotyping. Currently, GS is being used for breeding in several different crops (Table S3). In soybean, efforts have been made to evaluate GS using different models. A GS study in soybean has used 126 recombinant inbred lines and 80 SSR markers to predict primary embryogenesis capacity which is a highly polygenic trait (Hu et al., 2011). In this report, high correlation ($r^2 = 0.78$) has been observed among the genomic estimated breeding value

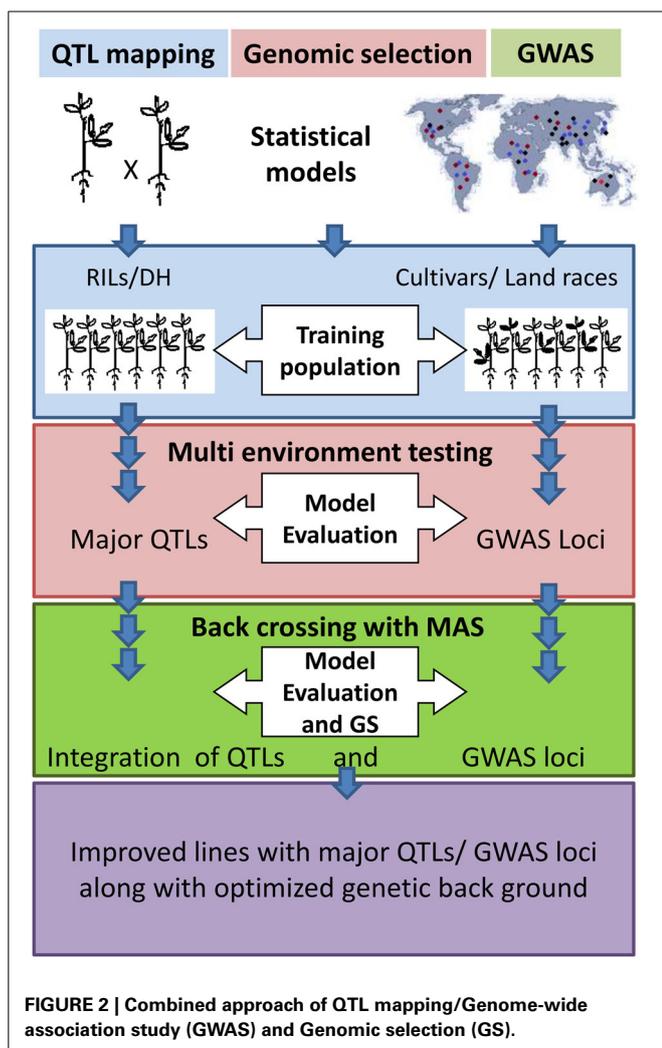
(GEBV) and the phenotypic value. Another study published recently using 288 cultivars and 79 SSR markers, found a correlation coefficient of 0.90 among the GEBV and the phenotypic value (Shu et al., 2012). Both the reports have shown high accuracy of prediction but only with a few markers and genotypes. Predicting the accuracy of GS will need more investigations involving high-throughput genotyping of larger populations evaluated across different environments.

Accuracy of GS largely depends on genetic \times environmental (G \times E) interaction but most of the studies focused only on an estimation of the main effect for each marker. These multi-environmental trials are of prime importance for plant breeding not only to study G \times E but especially to increase the number of breeding cycles per year. The challenge for GS is to get accurate GEBV in respect to the G \times E effect. Considering environmental effects is not new for plant breeders and most statistical models used for multi-location trials do reflect G \times E (Hammer et al., 2006). It is also more common in QTL mapping studies where QTL \times environment interaction evaluations were utilized to estimate QTL effect.

Improved factorial regression models have been proposed recently for GS that consider stress covariates derived from daily weather data (Heslot et al., 2014). This model has shown increased accuracy by 11.1% for predicting GEBV in unobserved environments where weather data is available (Heslot et al., 2014). This study suggests possible utilization of phenotypic data and historical data of weather conditions accumulated over decades in different soybean breeding programs. Similar information can be used for abiotic stress tolerance improvement in soybean.

COMBINING MARKER-ASSISTED BREEDING WITH GENOMIC SELECTION

Molecular marker genotyping is a common requirement for QTL mapping, GWAS, and GS and can be the basis for combining these approaches (Figure 2). Most of the GS studies have used recombinant inbred line (RIL) populations to train the prediction model (Table S3). Therefore, GS and QTL mapping can be performed simultaneously. A set of diverse cultivars can be used for GWAS and GS all together (Table S3). In the marker-assisted breeding, introgression of QTL or GWAS loci to well adapted cultivar



is performed. The donor line (for QTL or GWAS loci) may be wild or low yielding line. Therefore, several cycles of backcrossing are performed to retain the genetic background of the recipient parent (the adapted cultivar) except for the QTL/GWAS loci which represent the donor background. Nevertheless, GS does not provide control over the genetic background and this may be problematic when the donor is not an adapted line. In addition, GS cannot guarantee for major QTL which are already known. Therefore, information about QTL/GWAS loci should be incorporated with GS models so that the balance of genetic background can be made along with maximum gain of breeding value.

TRANSCRIPTOME PROFILING FOR ABIOTIC STRESS TOLERANCE

Plants, including soybean, responses to external environments is very complex. A wide range of defense mechanisms are activated that increases plant tolerance against adverse conditions in order to avoid damage imposed by abiotic stresses. The first step toward stress response is stress signal recognition and subsequent molecular, biochemical, and physiological responses activated through signal transduction (Komatsu et al., 2009; Ge et al., 2010; Le et al., 2012). Understanding such responses is very important for

effective management of abiotic stress. Transcriptome profiling provides an opportunity to investigate plant response regulation and to identify genes involved in stress tolerance mechanisms. Earlier, approaches using expressed sequence tags (ESTs) sequencing along with several techniques, such as suppression subtractive hybridization (SSH), have been extensively used for transcriptome profiling of soybean under abiotic stress conditions (Clement et al., 2008). In addition, information of ESTs have been used to develop spotted microarrays (O'Rourke et al., 2007). These techniques are efficient but do not ensure analysis of entire genes in the soybean genome. Several high-throughput techniques have been developed for transcriptome analysis due to the advancement in sequencing technology and the availability of the whole soybean genome sequence, (Libault et al., 2010; Schmutz et al., 2010; Cheng et al., 2013). These platforms have been extensively used for transcriptome profiling to uplift abiotic stress tolerance mechanisms in soybean (Table 3).

Microarray is a high-throughput technology where thousands of probes representing different genes are hybridized with RNA samples. Using the hybridization signal level, gene expression is calculated. The Affymetrix GeneChip representing 61K probe sets is routinely being used for transcriptome profiling of soybean under different abiotic stresses (Haerizadeh et al., 2011; Le et al., 2012). The normalized expression data generated using the Affymetrix GeneChip can be used to compare soybean experiments performed across the world. An expression database has been developed to globally explore public and proprietary expression data (www.geneinvestigator.com). The microarray data represents various tissues, developmental stages, and environmental conditions (Table 3). Effective analysis of such tremendous data using sequence homology and functional annotation will be helpful to understand biological processes.

RNA-Seq, AN ADVANCED APPROACH FOR TRANSCRIPTOME PROFILING

Cost effective and high-throughput sequencing technologies make it possible to analyze transcriptomes by sequencing, known as RNA-seq. The RNA-seq approach has several advances over the microarray technology where available genomic information is used to design probe sets. However, RNA-seq does not require gene information and is capable of identifying novel transcripts that were previously unknown and also provides opportunities to analyze non-coding RNAs. The relative accuracy of microarrays and RNA-Seq has been evaluated using proteomics and it has been shown that RNA-Seq provides a better estimate of absolute expression levels (Fu et al., 2009). Applications of RNA-seq can be expanded further with an increased understanding of molecular regulations. For instance, RNA-seq is being used for transcription start site mapping, strand-specific measurements, gene fusion detection, small RNA characterization, and detection of alternative splicing events (Ozsolak and Milos, 2010).

RNA-Seq has been performed to investigate seven tissues and seven stages in seed development in soybean (Severin et al., 2010). This effort has generated an expression atlas for soybean genes which serves as a useful resource. The tissue specific expression pattern of genes is helpful in understanding regulation and tissue specific function.

Table 3 | Major transcriptomic analysis for the abiotic stress tolerance in soybean using different technological platforms.

Sr. No.	Trait/tissue	Platform	DEG*	Key points	References
1	Soybean root development/root tips and non-meristematic tissue	Affymetrix chips containing 37,500 probe sets	9148	Resource of novel target genes for further studies involving root development and biology	Haerizadeh et al., 2011
2	Iron stress/root from isogenic lines	Custom array containing 9728 cDNAs	48	Genes involved in DNA repair and RNA stability were induced	O'Rourke et al., 2007
3	Drought stress at late developmental stages/V6 and R2 stages under drought and control	61 K Affymetrix Soybean Array GeneChip	3276 for V6 3270 for R2	Expression of many <i>GmNAC</i> and hormone-related genes was altered by drought in V6 and/or R2 leaves	Le et al., 2012
4	Herbicide resistance/plant under atrazine and bentazon stress	cDNA microarray with 36,760 different cDNA clones	6646	Expression of genes related to cell recovery, such ribosomal components	Zhu et al., 2009
5	Saline-alkaline stress tolerance/NaCl and NaHCO ₃ treatments	AffymetrixSoybean GeneChip	9027	Genes with altered expression regulated by alkaline stress	Ge et al., 2010
6	Flooding stress	HiCEP (29,388) high coverage expression profiling	97 genes and 34 proteins	Combined approach with proteomics	Komatsu et al., 2009

*Differentially expressed genes.

COMBINING QTL MAPPING, GWAS, AND TRANSCRIPTOME PROFILING

QTL mapping and GWAS are very effective approaches to identify chromosomal region(s) associated with a particular phenotype. However, QTL spans large segments of chromosomes and it is also the same for GWAS where LD decay is slow as in case of soybean (Hyten et al., 2007). QTL or GWAS loci possess hundreds of genes that make the identification of candidate genes difficult (Sonah et al., 2012). This is similar in transcriptome profiling where thousands of genes have been found to be differentially expressed even with genetically similar isogenic lines (Table 3). Therefore combining QTL mapping or GWAS with transcriptome profiling will complement each other. For instance, candidate genes for grain number QTL in rice have been identified using microarray based transcriptome profiling of recombinant inbred lines with contrasting phenotypes (Deshmukh et al., 2010; Sharma et al., 2011; Kadam et al., 2012). Similarly, a pair of soybean near-isogenic lines (NILs) differing in seed protein and an introgressed QTL segment (~8.4 Mb) have been used to study variation in transcript abundance in the developing seed (Bolon et al., 2010). The study identified 13 candidate genes in the QTL region using the Affymetrix Soy GeneChip and high-throughput Illumina whole transcriptome sequencing (Bolon et al., 2010). A combined approach of mapping and transcriptome profiling is based on an assumption that the quantitative trait is regulated by differential expression of candidate genes. This is not always true. Most of the time sequence variation present in candidate genes may cause defective proteins (Xu et al., 2013). Therefore, re-sequencing of QTL locus along with transcriptomics will also be a valuable approach to compliment mapping efforts.

PROTEOMICS IN SOYBEAN

Proteomics deals with structural and functional features of all the proteins in an organism. It is important to understand

complex biological mechanisms including the plant responses to abiotic stress tolerance. Abiotic stress tolerance mechanisms involve stress perception, followed by signal transduction, which changes expression of stress-induced genes and proteins. Post-translational changes are also important in plant responses to abiotic stresses. A single gene can translate in several different proteins and a few genes can lead to a diverse proteome. Such inconsistency limits genomics and transcriptomic approaches more specifically, when post translational changes govern phenotype. Differential expression observed at the transcriptional (mRNA) level need not be translated into differential amounts of protein. To address this, several proteomic studies have been performed to understand abiotic stress tolerance mechanisms in soybean (Table S4).

Unexpected levels of changes in the soybean proteome can occur during stress response and these changes can lead to different defense mechanisms. Some common proteins involved in redox systems, carbon metabolism, photosynthesis, signaling, and amino acid metabolism have been found to be associated with various stress responses in soybean (Zhen et al., 2007; Aghaei et al., 2009; Yamaguchi et al., 2010; Qin et al., 2013). These candidate proteins can directly link to genetic regulation of stress response in soybean. Candidate protein information can be used for the functional annotation of genes present in QTL regions or found differentially expressed under stress conditions.

In the near future, various proteomics approaches will be routinely used in soybean research that will generate tremendous information regarding structural and functional attributes of proteins. A systematic cataloging of information in the form of a publically accessible database is very important. Recently, a proteome database has been developed that contains reference maps of the soybean proteome collected from several organs, tissues, and organelles (Mooney and Thelen, 2004; Brechenmacher

et al., 2009; Ohyanagi et al., 2012). Presently, these reference maps comprised information of about 3399 proteins from seven organs and 2019 proteins from four subcellular compartments that were identified using two-dimensional electrophoresis (<http://proteome.dc.affrc.go.jp/soybean/>). Volunteer deposition of proteomic information in such databases is necessary for effective utilization of available knowledge for the management of abiotic stress tolerance in soybean.

METABOLOMICS ADVANCES FOR ABIOTIC STRESS

Metabolomic studies in plants aim to identify and quantify the complete range of primary and secondary metabolites involved in biological processes. Therefore metabolomics provides a better understanding of biochemical pathways and molecular mechanisms. The knowledge of genes, transcripts and proteins involved cannot alone help to understand the biological process completely until knowledge of metabolites that are involved becomes available.

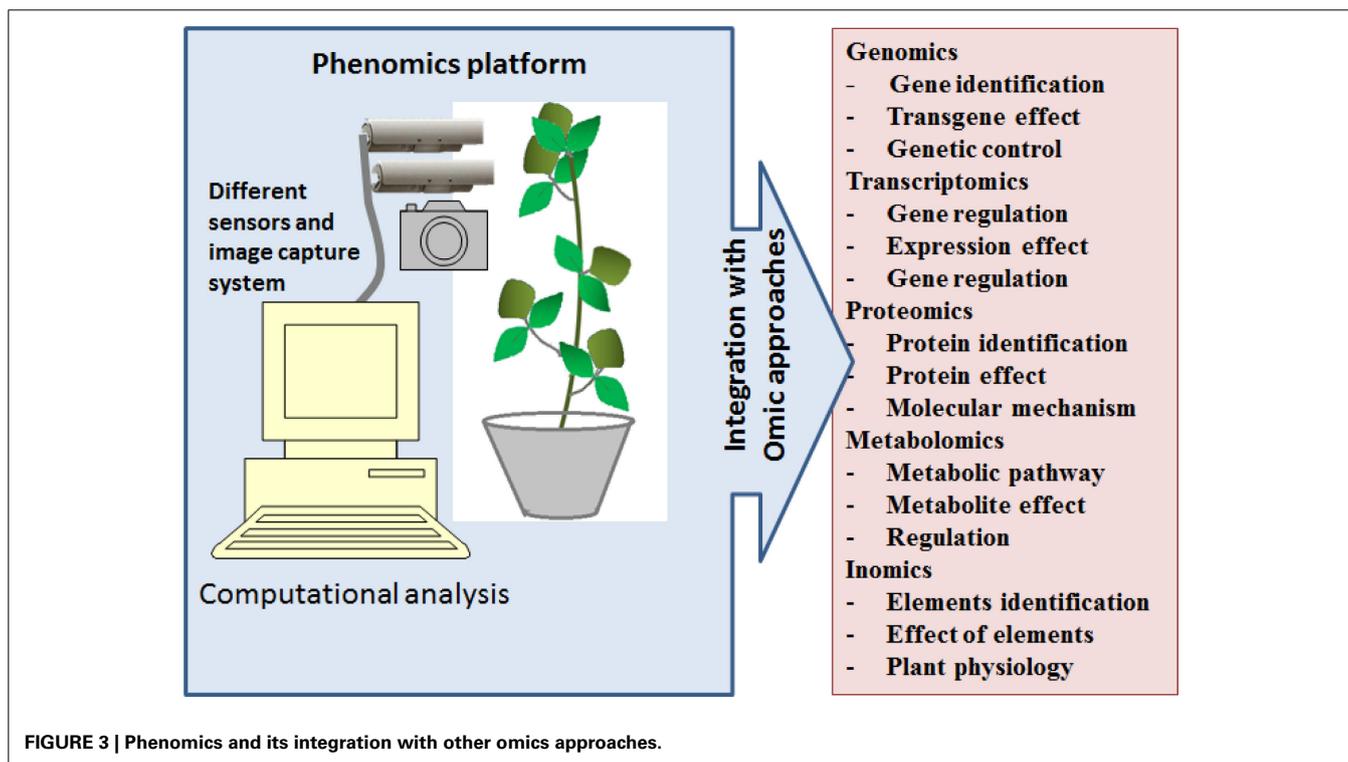
Several metabolomics studies have been performed to understand biochemical processes in soybean (Table S5). Development of new chromatographic and mass spectrometric platforms along with the enhancement of operational and analytical capabilities of existing platforms revolutionizes metabolomic investigations both in plant and animal sciences. The platforms such as gas chromatography mass spectrometry (GC-MS), fourier transform ion cyclotron resonance mass spectrometry (FT-ICR-MS), liquid chromatography mass spectrometry (LC-MS), capillary electrophoresis mass spectrometry (CE-MS), and nuclear magnetic resonance (NMR) are routinely used in plant sciences (Putri et al., 2013). Capability, limitations and specificity of these techniques has been recently reviewed in terms of effective utilization of these

advanced resources (Putri et al., 2013). In-depth accurate analyses of metabolite information including the spectral data are the major challenge for the use of high-throughput techniques. Several statistical models and bioinformatics programs have been developed to analyze the metabolome in an interactive manner (Fernie et al., 2011; Putri et al., 2013).

IONOMICS IN SOYBEAN

Ionomics is the study of elemental composition of an organism that mostly deals with high-throughput identification and quantification. Ionomics is important to understand element composition and their role in biochemical, physiological functionality and nutritional requirements of plants. Phosphorus (P) and potassium (K) are the two key elements used as macronutrients in fertilizer to ensure better crop yield. However plants require many other elements and those are not uniformly distributed among different soil types. Plants have evolved with a diverse element uptake ability at different locations because of diverse soil types (Fujita et al., 2013). This justifies the need of integrating ionomics with genomics to explore existing genetic differences. An ionic study has been performed to analyze concentrations of 17 different elements in diverse accessions and three RIL populations of *Arabidopsis thaliana* grown in several different environments (Buescher et al., 2010). Significant differences in elemental composition between the *Arabidopsis* accessions were detected and more than hundred QTL were identified for different elemental accumulation (Buescher et al., 2010). Most of the ionomics studies to date in soybean have been performed to analyze nutritive value of soybean products (Table S6).

The elemental composition of a plant is controlled by multiple factors including element availability, uptake capability of roots,



transport, and external environment which regulate physiological processes such as evapotranspiration. Because of such factors, the plant ionome has become very sensitive and specific so that the element profile reflects different physiological states. Recently a study performed in barley has analyzed ionome of wild accessions and cultivar differing in salt tolerance, grown in presence of 150 and 300 mM NaCl (Wu et al., 2013) and observed decreased amounts of K, magnesium (Mg), P and manganese (Mn) in roots and K, calcium (Ca), Mg and Sulfur (S) in shoots at the seedling stage. In addition, significant negative correlation among the amount of accumulated Na and metabolites involved in glycolysis and tricarboxylic acid (TCA) cycle have been observed (Wu et al., 2013). This ionomic study suggests the possible rearrangement of elemental profiles and metabolic processes to modify the physiological mechanisms of salinity tolerance.

Improvement in abiotic stress tolerance with the application of several inorganic element has been observed (Liang et al., 2007; Pilon-Smits et al., 2009). For instance, silicon (Si) has shown beneficial effects against different abiotic stresses including high salinity, water stress, heavy metal stress, and UV-b (Liang et al., 2007). Previously, soybean has been considered as poor accumulator of silicon mostly because of the genetic differences existing in the germplasm and very few genotypes have been evaluated to draw this conclusion (Hodson et al., 2005). However, with the advancement in ionomics technologies, silicon transporter genes have been identified recently in soybean using the integrated omics approach (Deshmukh et al., 2013). This study has used computational genomics, transcriptomics, and ionomics information available in the model plant species such as Arabidopsis and rice. Besides this, high-throughput efforts for maximum number of elemental profiles in soybean in respective external environment are required. That will definitely improve the understanding of the soybean ionome and its subsequent utilization in the management of abiotic stress tolerance.

PHENOMICS PROSPECTIVE IN SOYBEAN

The phenotype is a physical and biochemical trait of an organism. Phenomics is a study involving high-throughput analysis of phenotype. Phenotype is the ultimate resultant from the complex interactions of genetic potential between an organism and environment. Precision phenotyping is important to understand any biological system. In plant as well as animal sciences, a particular phenotype (as symptoms) is used to understand biological status, such as disease, pest infestation or physiological disorders. With technological advances, genomic resources have been routinely used to predict phenotype based on the evaluation of genetic markers; it can be called “genetic symptoms.” The success of genomics is based on how reliable connection is there between a genetic marker and the phenotype. In plant breeding, genetic improvement through omics approaches is being conducted to achieve ideal phenotype that will ensure higher and stable yield under diverse environmental conditions. Therefore phenomics integrated with other omics approaches has the most potential in the plant breeding (Figure 3).

Phenome has a broader meaning than what is being generally considered. It is not limited to the visible morphology of an organism but expectedly larger and complex. Unlike genomics,

where the entire genome can be characterized by sequencing, the phenome cannot be characterized entirely. Therefore, the term phenomics being an analogy to genomics expected only study of particular set of phenotype at high-throughput level and not the entire set. In this regards, the technological development in image processing and the automation techniques have played important roles. Plant imaging with light sources from visible to near infrared spectrum provides an opportunity for non-destructive phenotyping. Therefore, real-time analysis of plant development became possible. Moreover, robotic technologies used in phenomic platforms have increased the precision and speed of phenotyping. This has allowed for incorporating additional aids such as precise irrigation and fertilization systems. For instance, “PHENOPSIS” an automated phenomic platform has been developed to study water stress in Arabidopsis and has a robotic arm loaded with a tube for irrigation and a camera (Granier et al., 2006). These types of advanced phenomic platforms have been developed and made available for wider range of crop plants (www.lemnatec.com). However, these platforms have not gained the expected popularity even though tremendous advancement in both imaging as well as robotic technology has been achieved.

In soybean, several phenomic efforts have been performed but most of these are pilot experiments (Table S7). Recently, a method has been developed to assess leaf growth in soybean under different environmental conditions (Mielewicz et al., 2013). This method can utilize different light sources that are available in a greenhouse as well as under field conditions. Marker tracking approaches (Martrack Leaf) have also been used to facilitate accurate analysis of two-dimensional leaf expansion with high temporal resolution (Mielewicz et al., 2013). Apart from this, phenomics has been used to facilitate efficient identification of soybean cultivars which is very important for germplasm resource management and utilization (Zhu et al., 2012). Zhu et al. (2012), used a laser light back-scattering imaging technology to analyze single seed. Images of laser light illuminated the soybean seed surface were captured by a charge-coupled device (CCD) camera. The characteristic pattern of laser luminance is analyzed by image processing technology to identify a particular cultivar. Such characteristic of laser light back-scattering can be used to assess quality and other seed characteristics as markers for selection in breeding programs.

Phenomics in soybean is lagging far behind genomics because hundreds of genomes and many genetic populations are re-sequenced. One best example is the 1000 genome re-sequencing project at the University of Missouri, MO, USA (<http://soybeangenomics.missouri.edu/news2012.php>). The 1000 genome project will generate a huge amount of genomic information which will require utilization of comparable phenomic data. This will be helpful to accelerate soybean research in many ways.

ROLE OF ONLINE DATABASES FOR EFFECTIVE INTEGRATION OF OMICS PLATFORMS

The recent advancement in the omic platforms has generated tremendous information which has been used to promote research activities in all possible dimensions. Utilization of available information has become possible because of computational resources that helps to catalog, store, and analyze available

Table 4 | Online databases exclusively developed to host soybean research data generated from different omics platforms.

Sr. No	Database	Features	Tools
1	SoyBase SoyBase and the Soybean Breeder's Toolbox, USDA and Iowa University, http://soybase.org/	Genetic and physical maps, QTL, Genome sequence, Transposable elements, Annotations, Graphical chromosome visualizer	BLAST search, ESTs search, SoyChip Annotation Search, Potential Haplotype (pHap) and Contig Search, Soybean Metabolic Pathways, Fast Neutron Mutants Search, RNA-Seq Atlas
2	SoyKB Soybean Knowledge Base, University of Missouri, Columbia, http://soykb.org/	Multi-omics datasets, Genes/proteins, miRNAs/sRNAs, Metabolite profiling, Molecular markers, information about plant introduction lines and traits, Graphical chromosome visualizer	Germplasm browser, QTL and Trait browser, Fast neutron mutant data, Differential expression analysis, Phosphorylation data, Phylogeny, Protein BioViewer, Heatmap and hierarchical clustering, PI and trait search, FTP/data download capabilities
3	SoyDB Soybean transcription factors database, Missouri University, http://casp.rnet.missouri.edu/soydb/	Protein sequences, Predicted tertiary structures, Putative DNA binding sites, Protein Data Bank (PDB), Protein family classifications	PSI-BLAST, Browse database, Family Prediction by HMM, FTP data retriever
4	SGMD The Soybean Genomics and Microarray Database, http://bioinformatics.towson.edu/SGMD/	Integrated view genomic, EST and microarray data	Analytical tools allowing correlation of soybean ESTs with their gene expression profiles
5	Deltasoy An Internet-Based Soybean Database for Official Variety Trials, http://msucares.com/deltasoy/testlocationmap.htm	Official variety trial (OVT) information in soybean, Mississippi OVT data, including yield, location, and disease information	Comparison tools for variety trial data, phenotypic data and disease related data
6	DaizuBase An integrated soybean genome database including BAC-based physical maps, http://daizu.dna.affrc.go.jp/	BAC-based physical map, Linkage map and DNA markers, BAC-end, BAC contigs, ESTs, full-length cDNAs	Gbrowse, Unified Map, Gene viewer, BLAST
7	SoyMetDB The soybean metabolome database, http://soymetdb.org	Soybean metabolomic data	Pathway Viewer
9	SoyProDB Soybean proteins database, http://bioinformatics.towson.edu	Several 2D Gel images showing isolated soybean seed proteins	Search tool for 2D spots, Navigation tools for protein data
10	SoyGD The Soybean GBrowse Database, Southern Illinois University, http://soybeangenome.siu.edu/	Physical map and genetic map, Bacterial artificial chromosome (BAC) fingerprint database, Associated genomic data	Sequence data retrieval tools, Navigation tool for sequence information of different builds
11	SoyTEdb Soybean transposable elements database, www.soybase.org/soytedb/	Williams 82 transposable element database	Browse for Repetitive elements, Transposable Element and Map position, Data retrieval tools
12	SoyXpress Soybean transcriptome database, http://soyexpress2.agrenv.mcgill.ca	Soybean ESTs, Metabolic pathways, Gene Ontology terms, Swiss-prot Identifiers and Affymetrix gene expression data	BLAST search, Microarray experiments, Pathway search etc

data and make it easily accessible through user friendly interfaces so called “databases.” In this regard, several databases have been developed for soybean (Table 4). Among these, Soybean Knowledge Base (SKB, <http://soykb.org>) is a very useful database that provides a comprehensive web resource for omics data from several different platforms (Joshi et al., 2012). The SKB resources are helpful for bridging soybean translational genomics and molecular breeding research. It contains information of genes, proteins, microRNAs, sRNAs, metabolites, molecular markers, and phenomic information of soybean plant introductions (PI). It also provides interference to integrate multi-omics datasets and because of this, a galaxy of information becomes comparable and more useful. For instance, genes in the QTL region can be retrieved very easily along with the functional annotations, associated protein information in respect of structure and functional features, syntenic information with other model plants, sequence variation among different cultivars, gene expression data including tissue specific variations and many other types of information for soybean.

GENERAL CONCLUSION

Different omics tools have been employed to understand how soybean plants respond to abiotic stress conditions. We realize that the studies to integrate multiple omics approaches are limiting in soybean due to the increased cost and potential challenging integrated omic scale analysis. Recent developments in computational resources, statistical tools, and instrumentation have lowered the cost of omics in many folds but integrated analysis needs novel tools and technical wizards. The comprehensive nature of multi-omic studies provides an entirely new avenue and future research programs should plan to adapt accordingly. In soybean, genomics and transcriptomics have progressed as expected but the other major omic branches like proteomics, metabolomics, and phenomics are still lagging behind. These omic branches are equally important to get clear picture of the biological system. Notably, phenomic studies need to be extensively employed along with the other omics approaches. Desired phenotype is ultimate aim of crop sciences; therefore it needs to be understood intensely. Different omic tools and integrated approaches discussed in the present review will provide glimpses of current scenarios and future perspectives for the effective management of abiotic stress tolerance in soybean.

ACKNOWLEDGMENTS

The authors are thankful to Theresa Musket and Michelle Keough for their insight, critical reviews and language improvement. This research was supported by grants from the United Soybean Board, USA.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://www.frontiersin.org/journal/10.3389/fpls.2014.00244/abstract>

REFERENCES

- Aghaei, K., Ehsanpour, A., Shah, A., and Komatsu, S. (2009). Proteome analysis of soybean hypocotyl and root under salt stress. *Amino Acids* 36, 91–98. doi: 10.1007/s00726-008-0036-7
- Akond, M., Schoener, L., Kantartzi, S., Meksem, K., Song, Q., Wang, D., et al. (2013). A SNP-based genetic linkage map of soybean using the SoySNP6K Illumina Infinium BeadChip genotyping array. *J. Plant Genome Sci.* 1, 80–89. doi: 10.5147/jpgs.2013.0090
- Bandillo, N., Raghavan, C., Muyco, P. A., Sevilla, M. A. L., Lobina, I. T., Dilla-Ermita, C. J., et al. (2013). Multi-parent advanced generation inter-cross (MAGIC) populations in rice: progress and potential for genetics research and breeding. *Rice* 6, 1–15. doi: 10.1186/1939-8433-6-11
- Bastien, M., Sonah, H., and Belzile, F. (2014). Genome wide association mapping of *Sclerotinia sclerotiorum* resistance in soybean with a genotyping by sequencing approach. *Plant Genome* 7, 1–13. doi: 10.3835/plantgenome2013.10.0030
- Bolon, Y.-T., Joseph, B., Cannon, S. B., Graham, M. A., Diers, B. W., Farmer, A. D., et al. (2010). Complementary genetic and genomic approaches help characterize the linkage group I seed protein QTL in soybean. *BMC Plant Biol.* 10:41. doi: 10.1186/1471-2229-10-41
- Brechenmacher, L., Lee, J., Sachdev, S., Song, Z., Nguyen, T. H. N., Joshi, T., et al. (2009). Establishment of a protein reference map for soybean root hair cells. *Plant Physiol.* 149, 670–682. doi: 10.1104/pp.108.131649
- Buescher, E., Achberger, T., Amusan, I., Giannini, A., Ochsenfeld, C., Rus, A., et al. (2010). Natural genetic variation in selected populations of *Arabidopsis thaliana* is associated with ionic differences. *PLoS ONE* 5:e11081. doi: 10.1371/journal.pone.0011081
- Candeia, R., Silva, M., Carvalho Filho, J., Brasilino, M., Bicudo, T., Santos, I., et al. (2009). Influence of soybean biodiesel content on basic properties of biodiesel–diesel blends. *Fuel* 88, 738–743. doi: 10.1016/j.fuel.2008.10.015
- Carpenter, J. E. (2010). Peer-reviewed surveys indicate positive impact of commercialized GM crops. *Nat. Biotech.* 28, 319–321. doi: 10.1038/nbt0410-319
- Cheng, Y.-Q., Liu, J.-F., Yang, X., Ma, R., Liu, C., and Liu, Q. (2013). RNA-seq analysis reveals ethylene-mediated reproductive organ development and abscission in soybean (*Glycine max* L. Merr.). *Plant Mol. Biol. Rep.* 31, 607–619. doi: 10.1007/s11105-012-0533-4
- Clement, M., Lambert, A., Herouart, D., and Boncompagni, E. (2008). Identification of new up-regulated genes under drought stress in soybean nodules. *Gene* 426, 15–22. doi: 10.1016/j.gene.2008.08.016
- Deshmukh, R., Singh, A., Jain, N., Anand, S., Gacche, R., Singh, A., et al. (2010). Identification of candidate genes for grain number in rice (*Oryza sativa* L.). *Funct. Integr. Genomics* 10, 339–347. doi: 10.1007/s10142-010-0167-2
- Deshmukh, R. K., Sonah, H., Kondawar, V., Tomar, R. S. S., and Deshmukh, N. K. (2012). Identification of meta quantitative trait loci for agronomical traits in rice (*Oryza sativa*). *Ind. J. Genet. Plant Breed.* 72, 264–270.
- Deshmukh, R. K., Vivancos, J., Guérin, V., Sonah, H., Labbé, C., Belzile, F., et al. (2013). Identification and functional characterization of silicon transporters in soybean using comparative genomics of major intrinsic proteins in *Arabidopsis* and rice. *Plant Mol. Biol.* 83, 303–315. doi: 10.1007/s11103-013-0087-3
- Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., et al. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS ONE* 6:e19379. doi: 10.1371/journal.pone.0019379
- Fernie, A. R., Aharoni, A., Willmitzer, L., Stitt, M., Tohge, T., Kopka, J., et al. (2011). Recommendations for reporting metabolite data. *Plant Cell* 23, 2477–2482. doi: 10.1105/tpc.111.086272
- Fu, X., Fu, N., Guo, S., Yan, Z., Xu, Y., Hu, H., et al. (2009). Estimating accuracy of RNA-Seq and microarrays with proteomics. *BMC Genomics* 10:161. doi: 10.1186/1471-2164-10-161
- Fujita, Y., Venterink, H. O., van Bodegom, P. M., Douma, J. C., Heil, G. W., Hölzel, N., et al. (2013). Low investment in sexual reproduction threatens plants adapted to phosphorus limitation. *Nature* 505, 82–86. doi: 10.1038/nature12733
- Ge, Y., Li, Y., Zhu, Y. M., Bai, X., Lv, D. K., Guo, D., et al. (2010). Global transcriptome profiling of wild soybean (*Glycine soja*) roots under NaHCO₃ treatment. *BMC Plant Biol.* 10:153. doi: 10.1186/1471-2229-10-153
- Grainger, C. M., and Rajcan, I. (2013). Characterization of the genetic changes in a multi-generational pedigree of an elite Canadian soybean cultivar. *Theor. Appl. Genet.* 1–19. doi: 10.1007/s00122-013-2211-9
- Granier, C., Aguirrezabal, L., Chenu, K., Cookson, S. J., Dauzat, M., Hamard, P., et al. (2006). PHENOPSIS, an automated platform for reproducible phenotyping of plant responses to soil water deficit in *Arabidopsis thaliana* permitted the identification of an accession with low sensitivity to soil water deficit. *New Phytol.* 169, 623–635. doi: 10.1111/j.1469-8137.2005.01609.x

- Guo, B., Sleper, D., Lu, P., Shannon, J., Nguyen, H., and Arelli, P. (2006). QTLs associated with resistance to soybean cyst nematode in soybean: meta-analysis of QTL locations. *Crop Sci.* 46, 595–602. doi: 10.2135/cropsci2005.04-0036-2
- Haerizadeh, F., Singh, M. B., and Bhalla, P. L. (2011). Transcriptome profiling of soybean root tips. *Funct. Plant Biol.* 38, 451–461. doi: 10.1071/FP10230
- Hammer, G., Cooper, M., Tardieu, F., Welch, S., Walsh, B., van Eeuwijk, F., et al. (2006). Models for navigating biological complexity in breeding improved crop plants. *Trends Plant Sci.* 11, 587–593. doi: 10.1016/j.tplants.2006.10.006
- Heffner, E. L., Sorrells, M. E., and Jannink, J. L. (2009). Genomic selection for crop improvement. *Crop Sci.* 49, 1–12. doi: 10.2135/cropsci2008.08.0512
- Heslot, N., Akdemir, D., Sorrells, M. E., and Jannink, J. L. (2014). Integrating environmental covariates and crop modeling into the genomic selection framework to predict genotype by environment interactions. *Theor. Appl. Genet.* 127, 463–480. doi: 10.1007/s00122-013-2231-5
- Hodson, M., White, P., Mead, A., and Broadley, M. (2005). Phylogenetic variation in the silicon composition of plants. *Ann. Bot.* 96, 1027–1046. doi: 10.1093/aob/mci255
- Hu, Z., Li, Y., Song, X., Han, Y., Cai, X., Xu, S., et al. (2011). Genomic value prediction for quantitative traits under the epistatic model. *BMC Genet.* 12:15. doi: 10.1186/1471-2156-12-15
- Huang, L. L., Zhong, K. Z., Ma, Q. B., Nian, H., and Yang, C. Y. (2011). Integrated QTLs map of phosphorus efficiency in soybean by Meta-analysis. *Chin. J. Oil Crop Sci.* 33, 25–32.
- Hyten, D. L., Cannon, S. B., Song, Q., Weeks, N., Fickus, E. W., Shoemaker, R. C., et al. (2010a). High-throughput SNP discovery through deep resequencing of a reduced representation library to anchor and orient scaffolds in the soybean whole genome sequence. *BMC Genomics* 11:38. doi: 10.1186/1471-2164-11-38
- Hyten, D. L., Choi, I. Y., Song, Q., Shoemaker, R. C., Nelson, R. L., Costa, J. M., et al. (2007). Highly variable patterns of linkage disequilibrium in multiple soybean populations. *Genetics* 175, 1937–1944. doi: 10.1534/genetics.106.069740
- Hyten, D. L., Choi, I. Y., Song, Q., Specht, J. E., Carter, T. E., Shoemaker, R. C., et al. (2010b). A high density integrated genetic linkage map of soybean and the development of a 1536 universal soy linkage panel for quantitative trait locus mapping. *Crop Sci.* 50, 960–968. doi: 10.2135/cropsci2009.06.0360
- Hyten, D. L., Song, Q., Choi, I. Y., Yoon, M. S., Specht, J. E., Matukumalli, L. K., et al. (2008). High-throughput genotyping with the GoldenGate assay in the complex genome of soybean. *Theor. Appl. Genet.* 116, 945–952. doi: 10.1007/s00122-008-0726-2
- Jing, W., Wankun, S., Wenbo, Z., Chunyan, L., Guohua, H., and Qingshan, C. (2009). Meta-analysis of insect-resistance QTLs in soybean. *Hereditas (Beijing)* 31, 953–961. doi: 10.3724/SPJ.1005.2009.00953
- Joshi, T., Patil, K., Fitzpatrick, M. R., Franklin, L. D., Yao, Q., Cook, J. R., et al. (2012). Soybean Knowledge Base (SoyKB): a web resource for soybean translational genomics. *BMC Genomics* 13:S15. doi: 10.1186/1471-2164-13-S1-S15
- Jun, T. H., Mian, M. R., Kang, S. T., and Michel, A. P. (2012). Genetic mapping of the powdery mildew resistance gene in soybean PI 567301B. *Theor. Appl. Genet.* 125, 1159–1168. doi: 10.1007/s00122-012-1902-y
- Kadam, S., Singh, K., Shukla, S., Goel, S., Vikram, P., Pawar, V., et al. (2012). Genomic associations for drought tolerance on the short arm of wheat chromosome 4B. *Funct. Integr. Genomics* 12, 447–464. doi: 10.1007/s10142-012-0276-1
- Ko, K. P., Park, S. K., Yang, J. J., Ma, S. H., Gwack, J., Shin, A., et al. (2013). Intake of soy products and other foods and gastric cancer risk: a prospective study. *J. Epidemiol.* 23, 337. doi: 10.2188/jea.JE20120232
- Komatsu, S., Yamamoto, R., Nanjo, Y., Mikami, Y., Yunokawa, H., and Sakata, K. (2009). A comprehensive analysis of the soybean genes and proteins expressed under flooding stress using transcriptome and proteome techniques. *J. Proteome Res.* 8, 4766–4778. doi: 10.1021/pr900460x
- Kover, P. X., Valdar, W., Trakalo, J., Scarcelli, N., Ehrenreich, I. M., Purugganan, M. D., et al. (2009). A multiparent advanced generation inter-cross to fine-map quantitative traits in *Arabidopsis thaliana*. *PLoS Genet.* 5:e1000551. doi: 10.1371/journal.pgen.1000551
- Lam, H. M., Xu, X., Liu, X., Chen, W., Yang, G., Wong, F. L., et al. (2010). Resequencing of 31 wild and cultivated soybean genomes identifies patterns of genetic diversity and selection. *Nat. Genet.* 42, 1053–1059. doi: 10.1038/ng.715
- Le, D. T., Nishiyama, R., Watanabe, Y., Tanaka, M., Seki, M., Yamaguchi-Shinozaki, K., et al. (2012). Differential gene expression in soybean leaf tissues at late developmental stages under drought stress revealed by genome-wide transcriptome analysis. *PLoS ONE* 7:e49522. doi: 10.1371/journal.pone.0049522
- Li, Y. H., Zhao, S. C., Ma, J. X., Li, D., Yan, L., Li, J., et al. (2013). Molecular footprints of domestication and improvement in soybean revealed by whole genome re-sequencing. *BMC Genomics* 14:579. doi: 10.1186/1471-2164-14-579
- Liang, Y., Sun, W., Zhu, Y. G., and Christie, P. (2007). Mechanisms of silicon-mediated alleviation of abiotic stresses in higher plants: a review. *Environ. Pollut.* 147, 422–428. doi: 10.1016/j.envpol.2006.06.008
- Libault, M., Farmer, A., Joshi, T., Takahashi, K., Langley, R. J., Franklin, L. D., et al. (2010). An integrated transcriptome atlas of the crop model *Glycine max*, and its use in comparative analyses in plants. *Plant J.* 63, 86–99. doi: 10.1111/j.1365-313X.2010.04222.x
- Mamidi, S., Chikara, S., Goos, R. J., Hyten, D. L., Annam, D., Moghaddam, S. M., et al. (2011). Genome-wide association analysis identifies candidate genes associated with iron deficiency chlorosis in soybean. *Plant Genome* 4, 154–164. doi: 10.3835/plantgenome2011.04.0011
- Manavalan, L. P., Guttikonda, S. K., Tran, L. S. P., and Nguyen, H. T. (2009). Physiological and molecular approaches to improve drought resistance in soybean. *Plant Cell Physiol.* 50, 1260–1276. doi: 10.1093/pcp/pcp082
- Mielewicz, M., Friedli, M., Kirchgessner, N., and Walter, A. (2013). Diel leaf growth of soybean: a novel method to analyze two-dimensional leaf expansion in high temporal resolution based on a marker tracking approach (Martrack Leaf). *Plant Methods* 9, 30. doi: 10.1186/1746-4811-9-30
- Mooney, B. P., and Thelen, J. J. (2004). High-throughput peptide mass fingerprinting of soybean seed proteins: automated workflow and utility of UniGene expressed sequence tag databases for protein identification. *Phytochemistry* 65, 1733–1744. doi: 10.1016/j.phytochem.2004.04.011
- Morrell, P. L., Buckler, E. S., and Ross-Ibarra, J. (2011). Crop genomics: advances and applications. *Nat. Rev. Genet.* 13, 85–96. doi: 10.1038/nrg3097
- Ohyanagi, H., Sakata, K., and Komatsu, S. (2012). Soybean Proteome Database 2012: update on the comprehensive data repository for soybean proteomics. *Front. Plant Sci.* 3:110. doi: 10.3389/fpls.2012.00110
- O'Rourke, J., Charlson, D., Gonzalez, D., Vodkin, L., Graham, M., Cianzio, S., et al. (2007). Microarray analysis of iron deficiency chlorosis in near-isogenic soybean lines. *BMC Genomics* 8:476. doi: 10.1186/1471-2164-8-476
- Ozsolak, F., and Milos, P. M. (2010). RNA sequencing: advances, challenges and opportunities. *Nat. Rev. Genet.* 12, 87–98. doi: 10.1038/nrg2934
- Palloix, A., Ayme, V., and Moury, B. (2009). Durability of plant major resistance genes to pathogens depends on the genetic background, experimental evidence and consequences for breeding strategies. *New Phytol.* 183, 190–199. doi: 10.1111/j.1469-8137.2009.02827.x
- Pilon-Smits, E. A., Quinn, C. F., Tapken, W., Malagoli, M., and Schiavon, M. (2009). Physiological functions of beneficial elements. *Curr. Opin. Plant Biol.* 12, 267–274. doi: 10.1016/j.pbi.2009.04.009
- Poland, J. A., and Rife, T. W. (2012). Genotyping-by-sequencing for plant breeding and genetics. *Plant Genome* 5, 92–102. doi: 10.3835/plantgenome2012.05.0005
- Putri, S. P., Yamamoto, S., Tsugawa, H., and Fukusaki, E. (2013). Current metabolomics: technological advances. *J. Biosci. Bioeng.* 116, 9–16. doi: 10.1016/j.jbiosc.2013.01.004
- Qi, Z. M., Han, X., Sun, Y. N., Wu, Q., Shan, D. P., Du, X. Y., et al. (2011a). An integrated quantitative trait locus map of oil content in soybean, (*Glycine max* L.) Merr., generated using a meta-analysis method for mining genes. *Agric. Sci. China* 10, 1681–1692. doi: 10.1016/S1671-2927(11)60166-1
- Qi, Z. M., Wu, Q., Han, X., Sun, Y. N., Du, X. Y., Liu, C. Y., et al. (2011b). Soybean oil content QTL mapping and integrating with meta-analysis method for mining genes. *Euphytica* 179, 499–514. doi: 10.1007/s10681-011-0386-1
- Qin, J., Gu, F., Liu, D., Yin, C., Zhao, S., Chen, H., et al. (2013). Proteomic analysis of elite soybean Jidou17 and its parents using iTRAQ-based quantitative approaches. *Proteome Sci.* 11, 12. doi: 10.1186/1477-5956-11-12
- Qiong, W., Zhaoming, Q., Chunyan, L., Guohua, H., and Qingshan, C. (2009). An integrated QTL map of growth stage in soybean [*Glycine max* (L.) Merr.]: constructed through meta-analysis. *Acta Agronomica Sinica* 35, 1418–1424. doi: 10.3724/SPJ.1006.2009.01418
- Ray, D. K., Mueller, N. D., West, P. C., and Foley, J. A. (2013). Yield trends are insufficient to double global crop production by 2050. *PLoS ONE* 8:e66428. doi: 10.1371/journal.pone.0066428
- Schmutz, J., Cannon, S. B., Schlueter, J., Ma, J., Mitros, T., Nelson, W., et al. (2010). Genome sequence of the palaeopolyploid soybean. *Nature* 463, 178–183. doi: 10.1038/nature08670

- Severin, A. J., Woody, J. L., Bolon, Y. T., Joseph, B., Diers, B. W., Farmer, A. D., et al. (2010). RNA-Seq Atlas of *Glycine max*: a guide to the soybean transcriptome. *BMC Plant Biol.* 10:160. doi: 10.1186/1471-2229-10-160
- Sharma, A., Deshmukh, R. K., Jain, N., and Singh, N. K. (2011). Combining QTL mapping and transcriptome profiling for an insight into genes for grain number in rice (*Oryza sativa* L.). *Ind. J. Genet. Plant Breed.* 71, 115–119.
- Shi, A., Chen, P., Li, D., Zheng, C., Zhang, B., and Hou, A. (2009). Pyramiding multiple genes for resistance to soybean mosaic virus in soybean using molecular markers. *Mol. Breed.* 23, 113–124. doi: 10.1007/s11032-008-9219-x
- Shu, Y., Yu, D., Wang, D., Bai, X., Zhu, Y., and Guo, C. (2012). Genomic selection of seed weight based on low-density SCAR markers in soybean. *Genet. Mol. Res.* 12, 2178–2188. doi: 10.4238/2013.July.3.2
- Singh, H., Deshmukh, R. K., Singh, A., Singh, A. K., Gaikwad, K., Sharma, T. R., et al. (2010). Highly variable SSR markers suitable for rice genotyping using agarose gels. *Mol. Breed.* 25, 359–364. doi: 10.1007/s11032-009-9328-1
- Sonah, H., Bastien, M., Iquiria, E., Tardivel, A., Légaré, G., Boyle, B., et al. (2013). An improved genotyping by sequencing (GBS) approach offering increased versatility and efficiency of SNP discovery and genotyping. *PLoS ONE* 8:e54603. doi: 10.1371/journal.pone.0054603
- Sonah, H., Deshmukh, R. K., Chand, S., Srinivasprasad, M., Rao, G. J., Upreti, H. C., et al. (2012). Molecular mapping of quantitative trait loci for flag leaf length and other agronomic traits in rice (*Oryza sativa*). *Cereal Res. Commun.* 40, 362–372. doi: 10.1556/CRC.40.2012.3.5
- Sonah, H., Deshmukh, R. K., Sharma, A., Singh, V. P., Gupta, D. K., Gacche, R. N., et al. (2011a). Genome-wide distribution and organization of microsatellites in plants: an insight into marker development in *Brachypodium*. *PLoS ONE* 6:e21298. doi: 10.1371/journal.pone.0021298
- Sonah, H., Deshmukh, R. K., Singh, V. P., Gupta, D. K., Singh, N. K., and Sharma, T. R. (2011b). Genomic resources in horticultural crops: status, utility and challenges. *Biotechnol. Adv.* 29, 199–209. doi: 10.1016/j.biotechadv.2010.11.002
- Song, F., Tang, D. L., Wang, X. L., and Wang, Y. Z. (2011). Biodegradable soy protein isolate-based materials: a review. *Biomacromolecules* 12, 3369–3380. doi: 10.1021/bm200904x
- Song, Q., Hyten, D. L., Jia, G., Quigley, C. V., Fickus, E. W., Nelson, R. L., et al. (2013). Development and evaluation of SoySNP50K, a high-density genotyping array for soybean. *PLoS ONE* 8:e54985. doi: 10.1371/journal.pone.0054985
- Song, Q., Jia, G., Zhu, Y., Grant, D., Nelson, R. T., Hwang, E. Y., et al. (2010). Abundance of SSR motifs and development of candidate polymorphic SSR markers (BARCSOYSSR_1.0) in soybean. *Crop Sci.* 50, 1950–1960. doi: 10.2135/cropsci2009.10.0607
- Sosnowski, O., Charcosset, A., and Joets, J. (2012). BioMercator V3: an upgrade of genetic map compilation and quantitative trait loci meta-analysis algorithms. *Bioinformatics* 28, 2082–2083. doi: 10.1093/bioinformatics/bts313
- Sun, Y. N., Luan, H., Qi, Z., Shan, D., Liu, C., Hu, G., et al. (2012b). Mapping and meta-analysis of height QTLs in soybean. *Legume Genomics Genet.* 3, 1–7. doi: 10.5376/lgg.2012.03.0001
- Sun, Y. N., Pan, J.-B., Shi, X. L., Du, X. Y., Wu, Q., Qi, Z. M., et al. (2012a). Multi-environment mapping and meta-analysis of 100-seed weight in soybean. *Mol. Biol. Rep.* 39, 9435–9443. doi: 10.1007/s11033-012-1808-4
- Tanksley, S., and Nelson, J. (1996). Advanced backcross QTL analysis: a method for the simultaneous discovery and transfer of valuable QTLs from unadapted germplasm into elite breeding lines. *Theor. Appl. Genet.* 92, 191–203. doi: 10.1007/BF00223376
- Tardivel, A., Sonah, H., Belzile, F., and O'Donoghue, L. S. (2014). Rapid identification of alleles at the soybean maturity gene E3 using genotyping by sequencing and a haplotype-based approach. *Plant Genome* 7, 1–9. doi: 10.3835/plantgenome2013.10.0034
- Tomar, R. S. S., Deshmukh, R. K., Naik, K., Tomar, S. M. S., and Vinod (2014). Development of chloroplast-specific microsatellite markers for molecular characterization of alloplasmic lines and phylogenetic analysis in wheat. *Plant Breed.* 133, 12–18. doi: 10.1111/pbr.12116
- Varala, K., Swaminathan, K., Li, Y., and Hudson, M. E. (2011). Rapid genotyping of soybean cultivars using high throughput sequencing. *PLoS ONE* 6:e24811. doi: 10.1371/journal.pone.0024811
- Vuong, T. D., Sleper, D. A., Shannon, J. G., and Nguyen, H. T. (2010). Novel quantitative trait loci for broad-based resistance to soybean cyst nematode (*Heterodera glycines* Ichinohe) in soybean PI 567516C. *Theor. Appl. Genet.* 121, 1253–1266. doi: 10.1007/s00122-010-1385-7
- Wang, J. L., Liu, C. Y., Wang, J., Qi, Z. M., Li, H., Hu, G. H., et al. (2010). An integrated QTL map of fungal disease resistance in soybean (*Glycine max* L. Merr): a method of meta-analysis for mining R genes. *Agric. Sci. China* 9, 223–232. doi: 10.1016/S1671-2927(09)60087-0
- Wu, D., Shen, Q., Cai, S., Chen, Z. H., Dai, F., and Zhang, G. (2013). Ionic responses and correlations between elements and metabolites under salt stress in wild and cultivated barley. *Plant Cell Physiol.* 54, 1976–1988. doi: 10.1093/pcp/pct134
- Wu, X., Ren, C., Joshi, T., Vuong, T., Xu, D., and Nguyen, H. (2010). SNP discovery by high-throughput sequencing in soybean. *BMC Genomics* 11:469. doi: 10.1186/1471-2164-11-469
- Xu, X., Zeng, L., Tao, Y., Vuong, T., Wan, J., Boerma, R., et al. (2013). Pinpointing genes underlying the quantitative trait loci for root-knot nematode resistance in paleopolyploid soybean by whole genome resequencing. *Proc. Natl. Acad. Sci. U.S.A.* 110, 13469–13474. doi: 10.1073/pnas.1222368110
- Xu, Y., Lu, Y., Xie, C., Gao, S., Wan, J., and Prasanna, B. M. (2012). Whole-genome strategies for marker-assisted plant breeding. *Mol. Breed.* 29, 833–854. doi: 10.1007/s11032-012-9699-6
- Yamaguchi, M., Valliyodan, B., Zhang, J., Lenoble, M. E., Yu, O., Rogers, E. E., et al. (2010). Regulation of growth response to water stress in the soybean primary root. I. Proteomic analysis reveals region-specific regulation of phenylpropanoid metabolism and control of free iron in the elongation zone. *Plant Cell Environ.* 33, 223–243. doi: 10.1111/j.1365-3040.2009.02073.x
- Zhang, W. B., Jiang, H. W., Li, C. D., Qiu, P. C., Qi, Z. M., Liu, C. Y., et al. (2010). Integration of QTLs related to soybean cyst nematode resistance based on meta-analysis. *Chin. J. Oil Crop Sci.* 32, 104–109.
- Zhao-Ming, Q., Yanan, S., Lijun, C., Qiang, G., Chunyan, L., Guohua, H., et al. (2009). Meta-analysis of 100-seed weight QTLs in soybean. *Scientia Agricultura Sinica* 42, 3795–3803.
- Zhao-Ming, Q., Ya-Nan, S., Qiong, W., Chun-Yan, L., Guo-Hua, H., and Qing-Shan, C. (2011). A meta-analysis of seed protein concentration QTL in soybean. *Can. J. Plant Sci.* 91, 221–230. doi: 10.4141/cjps09193
- Zhen, Y., Qi, J. L., Wang, S. S., Su, J., Xu, G. H., Zhang, M. S., et al. (2007). Comparative proteome analysis of differentially expressed proteins induced by Al toxicity in soybean. *Physiol. Plant.* 131, 542–554. doi: 10.1111/j.1399-3054.2007.00979.x
- Zhu, D., Li, Y., Wang, D., Wu, Q., Zhang, D., and Wang, C. (2012). The identification of single soybean seed variety by laser light backscattering imaging. *Sensor Lett.* 10, 1–2. doi: 10.1155/2012/539095
- Zhu, J., Patzoldt, W. L., Radwan, O., Tranel, P. J., and Clough, S. J. (2009). Effects of photosystem-II-interfering herbicides atrazine and bentazon on the soybean transcriptome. *Plant Genome* 2, 191–205. doi: 10.3835/plantgenome2009.02.0010

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 10 March 2014; accepted: 13 May 2014; published online: 03 June 2014.
 Citation: Deshmukh R, Sonah H, Patil G, Chen W, Prince S, Mutava R, Vuong T, Valliyodan B and Nguyen HT (2014) Integrating omic approaches for abiotic stress tolerance in soybean. *Front. Plant Sci.* 5:244. doi: 10.3389/fpls.2014.00244
 This article was submitted to *Plant Genetics and Genomics*, a section of the journal *Frontiers in Plant Science*.
 Copyright © 2014 Deshmukh, Sonah, Patil, Chen, Prince, Mutava, Vuong, Valliyodan and Nguyen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.