Published August 5, 2013

Transcriptome Analyses in Legumes: A Resource for Functional Genomics

Rohini Garg and Mukesh Jain*

Abstract

Legumes represent an important family of flowering plants in terms of providing human nutrition and capacity to fix atmospheric N for agricultural sustainability. The recent availability of genome sequence of several legume plants has helped boosting genomics research. Study of the transcriptome at a global level can provide insights into the gene space, gene function, transcriptional programs, and molecular basis of various cellular processes in legumes, even in the absence of genome sequence. Transcriptome analysis has been realized as an essential step for basic and applied research in any organism. Considering the importance of transcriptome analyses, a few studies have been performed in legumes, such as soybean [*Glycine max* (L.) Merr.], *Medicago truncatula* Gaertn., *Lotus corniculatus* L. var. *japonicus* Regel [syn. *Lotus japonicus* (Regel) K. Larsen], and chickpea (*Cicer arietinum* L.), to uncover the overall and specific transcriptional activity of genes across various tissues and/or organs and developmental stages. Several candidate genes putatively involved in important agronomic traits, such as nodule, flower, and seed development, have been identified. The availability of these transcriptome data and future investigations will enable a variety of functional genomic studies to characterize these genes and define their function in legumes.

Published in The Plant Genome 6
doi: 10.3835/plantgenome2013.04.0011

5585 Guilford Rd., Madison, WI 53711 USA
An open-access publication

LEGUMES (Leguminosae) are the third largest family of flowering plants. They are important to humans not only as food and fodder but also for oil, fiber, and green manure. Legumes are the important source of dietary proteins especially in developing countries, which complement cereals. Legumes also play an important role in sustainable agricultural practices, because of their ability to fix atmospheric N in symbiotic association with rhizobia. The legume family includes important food plants, such as common bean (*Phaseolus vulgaris* L.), soybean [*Glycine max* (L.) Merr.], pea (*Pisum sativum* L.), chickpea (*Cicer arietinum* L.), and pigeonpea [*Cajanus cajan* (L.) Huth], and important forage species, such as alfalfa (*Medicago sativa* L.) and clover (*Trifolium* L.). Despite their importance, crop improvement programs in legumes lagged behind those of cereals (Graham and Vance, 2003). One of the major reasons for this lag has been the nonavailability of enough genomic resources. However, with the recent availability of cost-effective and high-throughput technologies, the scenario has changed to much extent and legumes are also the focus of numerous genomics projects now (Cannon et al., 2009; Varshney et al., 2009).

Currently, the genome sequence of five legume plants, including soybean, *Lotus japonicus*, *Medicago truncatula*, pigeonpea, and chickpea, are available (Sato et al., 2008; Schmutz et al., 2010; Young et al., 2011; Varshney et al., 2012, 2013; Jain et al., 2013). However, the next question that needs to be addressed is "how to make sense of the genome sequence data to translate it into meaningful information for crop improvement?" There is a great need to understand the relationship between genome sequence and molecular

National Institute of Plant Genome Research (NIPGR), Aruna Asaf Ali Marg, New Delhi- 110067, India. Received 26 Apr. 2013. *Corresponding author (mjain@nipgr.ac.in).

Abbreviations: DAP, days after pollination; LjGEA, *Lotus japonicus* Gene Expression Atlas; MtGEA, *Medicago truncatula* Gene Expression Atlas; NGS, next generation sequencing; RNA-seq, ribonucleic acid sequencing.

mechanisms underlying various cellular processes that are important for crop improvement. This can be achieved by coupling genome sequence with global transcriptome analyses to understand the relationship between gene function and regulation of gene expression (Jackson et al., 2011). Despite having a similar genetic content, different cells in a multicellular organism show different patterns of gene expression that underlie wide range of physical, biochemical, and developmental differences among them. Thus, through comparative transcriptome analyses among different tissues and/or cell types and developmental stages, the way in which transcriptional changes contribute to various cellular processes may be understood. This will also enable an understanding of gene interactions and regulation. In fact, transcriptome analyses has laid the foundation of functional and applied genomics research in various plants (Hansen et al., 2008; Jackson et al., 2011). In this review, we briefly discuss various technologies available for transcriptome analyses and their usage in legume biology. The major focus is on the progress made to date on transcriptome studies in legumes to gain knowledge on the important developmental processes related to sustainability of crop production, including nodule, flower, and seed development.

Technologies for Transcriptome Analysis

With the availability of high-throughput methods for studying global gene expression, an unbiased view of the transcriptional activity within genomes can be obtained. Many hybridization- and sequencing-based technologies are available for studying gene expression at whole-genome level. Among the hybridization-based methods, DNA microarray has been the most popular technology for large-scale gene expression studies. However, such methods are limited to interrogation of expression patterns of only known set of genes represented on the array. Nonetheless, microarrays have been most widely used to study gene expression patterns at whole-genome level or a set of genes in many plant species, addressing a wide range of biological problems (Rensink and Buell, 2005; Garg et al., 2010; Sharma et al., 2012). Microarrays have been used to study transcriptome dynamics during various developmental processes, such as seed and nodule development, in legumes as well (Le et al., 2007; He et al., 2009; Verdier et al., 2013b).

The sequencing-based approaches provide a better alternative to quantify gene expression ('t Hoen et al., 2008; Marioni et al., 2008). Previously, sequencing of expressed sequence tags was used to analyze the transcriptome and identify differentially expressed transcripts. However, this method was very costly, time consuming, and low throughput. Later, serial analysis of gene expression and massively parallel signature sequencing methods were adopted for studying gene expression. Despite also being high throughput, these methods were not as widely implemented as microarray. With the advent of next generation sequencing (NGS) technologies, the potential of sequencing-based methods have been highlighted. Whole transcriptome sequencing using NGS technologies (ribonucleic acid sequencing [RNA-seq]) is a more convenient and rapid method to study gene expression at whole-genome level and can be used to predict putative gene function (Wang et al., 2009; Ozsolak and Milos, 2011; Jain, 2012). RNA-seq can also be used for gene discovery by de novo constructing the complete transcriptome of several organisms including nonmodel species (Jain, 2011; Martin and Wang, 2011). RNA-seq provides a better alternative over microarray for gene expression studies in terms of robustness, resolution, and reproducibility ('t Hoen et al., 2008; Marioni et al., 2008; Ozsolak and Milos, 2011). Using RNA-seq, it is possible to gain insights into the gene expression without prior knowledge of sequence. Many studies have already demonstrated the power of RNA-seq in various biological contexts (Li et al., 2010; Zenoni et al., 2010; Garg et al., 2011; Yang et al., 2011; Singh et al., 2013). Few such studies have been performed in legumes too, such as development of transcription atlas in soybean and study of transcriptome dynamics during flower development in chickpea (Libault et al., 2010; Severin et al., 2010; Singh et al., 2013). Despite several advantages offered by RNA-seq, its use for gene expression studies has been limited in legumes so far. It may be due to the bioinformatics challenges associated with data analysis, such as accurate quality control, read mapping, read count normalization, and implementation of suitable statistical algorithm for differential gene expression analysis (Ozsolak and Milos, 2011; Jain, 2012). However, with the increased availability of user-friendly data analysis tools, use of RNA-seq in more studies in coming years is highly anticipated.

Transcriptome Data Resources in Legumes

As mentioned earlier, only a handful of genomewide transcriptome analyses studies have been performed in legumes. Both microarrays and RNA-seq technologies have been adopted to perform these studies. The transcriptome and gene expression data for these legumes have also been made publicly available via various databases or web servers (Table 1), which can help in identification of a target gene or genes of interest for functional analysis. For instance, SoySeq is a utility of SoyBase (http://soybase.org/; accessed 15 Apr. 2013), which provides gene expression information in a diverse set of 14 tissues from soybean based on RNA-seq data (Severin et al., 2010). Based on analysis of these data, over 2000 genes with preferential expression in seed and more than 177 genes involved in the seed filling process have been reported. This database provides data for tissue-by-tissue comparison and list of differentially expressed genes. SoyPLEX (http://www.plexdb.org/plex.php?database=Soybean; accessed 15 Apr. 2013) is another web server, which hosts gene expression data for soybean derived from Affymetrix Soybean GeneChip (Le et al., 2007; Dash et al., 2012). Microarray data from a total of 41 experiments (more than 1000 samples) representing various tissues, organs, and/or cell types, developmental

Table 1. Web resources for transcriptome data in legumes.

Web resource[†]	URL	Legume	Data type
SoySeq	http://soybase.org/soyseq/ (accessed 15 Apr. 2013)	Soybean	RNA-seq[‡]
SoyPLEX	http://www.plexdb.org/plex.php?database=Soybean (accessed 15 Apr. 2013)	Soybean	Microarray
MtGEA	http://mtgea.noble.org/v3/ (accessed 15 Apr. 2013)	*Medicago truncatula*	Microarray
CTDB	http://www.nipgr.res.in/ctdb.html (accessed 15 Apr. 2013)	Chickpea	RNA-seq
LjGEA	http://ljgea.noble.org/v2/ (accessed 15 Apr. 2013)	Lotus corniculatus	Microarray
LegumeIP	http://plantgrn.noble.org/LegumeIP/ (accessed 24 May 2013)	Soybean, *M. truncatula*, and *L. corniculatus*	RNA-seq and microarray

[†]MtGEA, *Medicago truncatula* Gene Expression Atlas; CTDB, Chickpea Transcriptome Database; LjGEA, *Lotus japonicus* Gene Expression Atlas.

[‡]RNA-seq, ribonucleic acid sequencing.

stages (seed development and vegetative to floral transition), and effect of various treatments, such as salinity, heat, Fe, and Al in soybean, are available in this web server. The transcriptome data from six different tissue samples of chickpea is also available through Chickpea Transcriptome Database (CTDB) (http://www.nipgr.res.in/ctdb.html; accessed 15 Apr. 2013). This data has been used for identification of tissue-specific genes in chickpea (Garg et al., 2011). The expression data for a single gene or batch of genes can be downloaded from the database.

Medicago truncatula Gene Expression Atlas (MtGEA) (http://mtgea.noble.org/v3/; accessed 15 Apr. 2013) web server includes a large amount of publicly available data derived from Affymetrix Medicago GeneChip (He et al., 2009). This web server hosts data from 254 experiments (670 GeneChips) covering all its major organ systems (roots, nodules, stems, petioles, leaves, vegetative buds, flowers, seeds, and pods) with detailed developmental time series for nodules and seeds. In addition, transcriptome data from plants subjected to various abiotic and biotic stresses and data from specific cell and tissue types are also included in MtGEA. This transcriptome data has been used to study nodule and seed development in *M. truncatula* (Benedito et al., 2008). Recently, *Lotus japonicus* Gene Expression Atlas (LjGEA) (http://ljgea.noble.org/v2/; accessed 15 Apr. 2013) has been developed from 83 experiments to present a global view of gene expression in all major organs (leaves, petioles, stem, roots, nodules, flowers, seeds, and pods) using Affymetrix Lotus GeneChip (Verdier et al., 2013b). The power of LjGEA has been demonstrated by exploring the transcriptome of developing seed (Verdier et al., 2013b). Both MtGEA and LjGEA allow analysis and visualization of gene expression for one or more genes of interest using probe set identifier, gene sequence, or gene annotation information as query. These web servers provide the utility of co-expression and differential gene expression analyses too. The availability of these public web resources for gene expression data will aid gene function discovery and molecular breeding efforts in legumes.

Transcriptome Analysis of Nodule Development

Legumes possess the unique ability to perform N fixation in endosymbiotic interaction with certain microorganisms. Plants develop specific organs known as root nodules that differentiate from root tissues. These highly specialized plant structures provide C source and appropriate cellular environment to the microorganisms allowing them to fix atmospheric N. Various events involved in this symbiosis process, such as signal perception (perception of nodulation factors from bacteria), infection (rhizobial infection via infection threads), organogenesis (formation of root nodules), and functioning (N fixation) are highly controlled (Schultze and Kondorosi, 1998). The identification of genes involved in root nodule formation and N fixation process has been the focus of research in past decades (Stacey et al., 2006). Studies have shown that not only the well-known nodulin genes but also many novel genes are involved in nodulation process (Colebatch et al., 2002; El Yahyaoui et al., 2004; Moreau et al., 2011). Using macroarrays comprising a few thousand genes, efforts have been made to identify nodule-induced genes, genes accompanying arbuscular mycorrhiza symbiosis, and genes involved in symbiotic program (Fedorova et al., 2002; Journet et al., 2002; El Yahyaoui et al., 2004). Colebatch et al. (2002) integrated transcriptome data with metabolite profiling and revealed several metabolic pathways upregulated in nodules. In a small-scale study, a total of 756 genes differentially regulated at different stages of nodulation process were identified by macroarray analysis of 6000 probes in wild-type and symbiotic mutant in *M. truncatula* and their possible involvement in various aspects of symbiotic interaction was proposed (El Yahyaoui et al., 2004). Several genes involved in hormone (auxin, cytokinin, and gibberellin) metabolism, cell wall proteins (proline-rich proteins, xyloglucans, and cellulose synthase), protein synthesis and degradation (ubiquitin pathway and microsomal signal peptidase), and genes of unknown function were found to be associated with nodule formation. Genes involved in C metabolism, N fixation, and membrane transport were associated with nodule functioning. In addition, several signal transduction genes and transcription factors were recognized as regulatory genes controlling infection, nodule organogenesis, and functioning (Colebatch et al., 2002; El Yahyaoui et al., 2004).

In a recent large-scale approach using 16,400 microarrays, Moreau et al. (2011) identified more than 3400 differentially regulated genes and associated regulators during nodule development in *M. truncatula*. Their analyses defined four distinct phases of transcriptional programming during nodule differentiation, early signaling, and/or bacterial infection,

root cell differentiation independent or dependent of bacteriod differentiation, and N fixation. This study implicated cytokinin biosynthesis and jasmonate pathways in nodule development. Interestingly, many legume-specific genes (genes that appear to be absent in nonlegumes) were found to be preferentially expressed in root nodules, indicating that these genes were endowed to perform a specific role or roles in this important process during evolution (Benedito et al., 2008). The overlay of gene expression data onto metabolic maps using MapMan (Thimm et al., 2004) showed induction of genes involved in glycolysis, N metabolism, and secondary metabolism, such as terpenoid and flavonoid pathways, which corroborated the results of previous studies (Colebatch et al., 2002; El Yahyaoui et al., 2004). Nodule-specific genes have been identified in the transcriptome atlas studies in soybean and *L. japonicus* as well (Severin et al., 2010; Verdier et al., 2013b). Overall, these studies have provided candidate genes involved in nodule development and associated events. However, much remains to be discovered and we are still far away from the complete molecular understanding of the nodulation process. These gaps can be filled with more detailed and genome-level transcriptome analyses of different tissues and/or parts of nodule at various stages of development and integration with proteomic and metabolite datasets.

Transcriptome Analysis of Flower Development

In higher plants, transition from vegetative growth to flowering is indicative of reproductive development. Flower development is the most important event in the life cycle of higher plants and has vital importance in agriculture. The molecular basis of flower development has been extensively studied in model plants such as *Arabidopsis thaliana* (L.) Heynh. and rice (*Oryza sativa* L.) (Jack, 2004; Krizek and Fletcher, 2005; Andrés and Coupland, 2012). Genomewide transcriptome investigations complemented with molecular genetic analyses have identified several key regulators of flower development and proposed a flower development model. A gene regulatory network controlling flowering time has been successfully established via genomewide gene expression analyses in *A. thaliana* (Keurentjes et al., 2007). Although such detailed studies have not been performed in legumes as of now, similarity search has revealed that most of known genes involved in flower development are conserved in soybean, chickpea, *M. truncatula*, and *L. japonicus* (Dong et al., 2005; Hecht et al., 2005; Jung et al., 2012; Singh et al., 2013). However, the morphological analyses of legume flowers have shown that flower development is somewhat dissimilar in legumes and does not follow the ABC model of eudicots (Tucker, 1987, 2003). At molecular level also, a few studies have shown that certain flower-related genes, such as *UNIFOLIATA* and *STAMINA PISTILLOIDA*, have acquired additional functions in legumes (Dong et al., 2005; Hofer et al., 1997; Taylor et al., 2001). Therefore, it is imperative to study this important developmental process in legumes as well to find the key regulators and establish the molecular basis of morphological peculiarities. Wong et al. (2009) studied molecular processes underlying floral transition in the soybean shoot apical meristem via microarray analysis. This study identified a total of 331 transcripts with significant differential expression during floral transition and implicated sugar, auxin, and abscisic acid in this process. An evidence of overlap of abiotic stress and floral signaling pathways was also highlighted.

In a recent RNA-seq based transcriptome analyses of three vegetative tissues and eight stages of flower development, a global view of transcriptome dynamics in chickpea has been presented (Singh et al., 2013). This study identified several genes expressed preferentially and specifically during different stages of flower development. Many of these genes were found to encode for transcription factors. A large number of MADS-box transcription factors were found upregulated during stages of flower development. MADS-box transcription factors are well known key regulators of floral organ development (Kater et al., 2006; Ng and Yanofsky, 2001). The overlay of transcriptome data with known metabolic pathways using AraCyc database of Gramene (http://pathway.gramene.org/gramene/aracyc.shtml; accessed 15 Apr. 2013) implicated several pathways, such as sporopollenin precursor biosynthesis, anthocyanin biosynthesis, homogalacturonan degradation, and cytokinin glucoside biosynthesis, etc., in flower development (Singh et al., 2013). This study provided the molecular basis of several physiological and morphological changes that occur during flower development. The flower transcriptome have been analyzed in other legumes as well, as a part of gene expression atlas (Severin et al., 2010; Verdier et al., 2013b). There are evidences that show the role of lineage-specific genes in flower development, which might be responsible, at least partly, for distinct floral symmetry in legumes (Garg and Jain, 2011; Singh et al., 2013). The functional analysis of these genes can provide novel insights into flower development process in legumes and other plants.

Transcriptome Analysis of Seed Development

Legumes are important for their seeds (fresh or desiccated) that have high seed storage protein content. Seed development is a complex developmental process that requires coordinated expression and regulation of several genes. Being the most important agronomic trait, seed development has been the major focus of research in legumes. In fact, many of whole transcriptome studies performed to date have focused their analysis on the process of seed and/or pod development (Severin et al., 2010; Verdier et al., 2013a, 2013b). The transcriptome dynamics at spatial and temporal levels during seed development have been analyzed and compared with the proteome in *M. truncatula* (Gallardo et al., 2007). This study revealed the compartmentalization of methionine biosynthesis

enzymes between seed tissues that may regulate protein synthesis during seed filling. This observation is of agronomic importance and can be used to devise strategies for modifying nutritional value of legume seeds. Furthermore, the comparative analysis of transcript and protein profiles suggested the posttranscriptional regulation of a significant fraction of genes during seed development (Gallardo et al., 2007). The analysis of seven stages of seed development in soybean identified over 2000 genes with preferential expression in seed and many of them were implicated in seed filling process (Severin et al., 2010). The genes involved in cellulose synthase activity, nutrient reservoir activity, and urease activity were found to be associated with seed development. Pod-specific genes have been identified in chickpea as well, which also showed nutrient reservoir activity as most enriched gene ontology term (Garg et al., 2011). Furthermore, several inter- and intraspecific genetic variations were detected in the pod-specific genes (Agarwal et al., 2012; Jhanwar et al., 2012), which can be used for high-throughput genotyping and association mapping studies.

Recently, a detailed time-series (10 to 20 d after pollination [DAP]) transcriptome study of the seed maturation has been performed in *L. japonicus* as a part of gene expression atlas (Verdier et al., 2013b). A total of 30% of all genes, including 190 legume-specific, 624 transcription factor, and 293 transporters genes, were identified to be differentially expressed during seed development in *L. japonicus*. The seed storage protein genes such as *VICILIN*, *LEGUMIN*, and *CONVICILIN* were among the most highly expressed genes in seeds. This study identified several clusters of genes with distinct expression patterns that correlated with the biological processes during different stages of seed development. For example, genes involved in ligand-receptor interaction and cell cycle were overrepresented during 10 DAP, where embryogenesis occurs. Likewise, genes involved in starch and sucrose metabolism were overrepresented during seed filling stage (16 DAP), and genes involved in protein folding and degradation were enriched at seed desiccation and maturation stage (20 DAP). In addition, this study identified 558 pod wall–enhanced genes, including 114 pod-specific genes. Pod wall and/or seed coat feeds and protects the developing seeds and coordinate grain filling by regulating the partitioning of nutrients from plants to seeds. In a recent study, a combined histology and transcriptome analysis has been undertaken to dissect the cellular and molecular events taking place during seed coat development in *M. truncatula* (Verdier et al., 2013a). This study showed a complex gene regulation of seed coat development and identified genes related to nucleotide metabolism, suberin biosynthesis, and hormonal regulation associated with seed coat. The seed coat-specific genes are potential targets for manipulation of seed traits, such as seed size, seed permeability, and resource partitioning and assimilation into seeds. In another study, the transcript profiling of *M. truncatula* plants ectopically expressing *TRANSPARENT TESTA 2* in hairy roots identified a glucosyltransferase gene expressed in the seed coat and involved in proanthocyanidin biosynthesis (Pang et al., 2008). Chen and colleagues (2013) recently studied the transcriptome of aerial and subterranean young pods in peanut (*Arachis hypogaea* L.) and identified candidate genes (senescence-associated genes) that potentially contribute to early embryo abortion in aerial pods. This study highlighted the evolutionary implications resulting from pod development under light and dark conditions. Overall, it will be fascinating to dissect the precise role of candidate genes involved in seed development processes identified from transcriptome studies for engineering of this important agronomic trait.

Comparative Transcriptome Analysis

To gain insights into the possible function of a gene in a plant species, researchers often look into the function of its ortholog (if known) in a model species. The assumption is that orthologs in different species, more likely, perform similar function, which may not be true. However, this assumption can be strengthened by looking into the expression pattern of orthologous genes in the two species. The comparative transcriptome analysis can provide insights into the conservation or divergence of gene function across different plant species. Such studies have been performed to study orthologs and specific biological processes, such as hypoxia response, light regulation, and leaf development in plants (Jiao et al., 2005; Street et al., 2008; Mustroph et al., 2010). The comparative transcriptome analysis of *M. truncatula* and *A. thaliana* revealed significant differences in the developmental expression profiles of orthologous genes, indicating their functional divergence in these two species (Benedito et al., 2008). The comparison of soybean transcriptome (genes expressed specifically in only one tissue) with *M. truncatula* and *L. japonicus* revealed similar tissue specificity of only 45% of the genes analyzed (Libault et al., 2010). The divergent expression patterns of the remaining genes reflected sub- or neofunctionalization events. Likewise, conserved and divergent expression patterns of known flower development genes were observed in chickpea (Singh et al., 2013). An integrative database, LegumeIP (http://plantgrn.noble.org/LegumeIP/; accessed 24 May 2013), is available for comparative genomics and transcriptomics of model legumes to study gene function and evolution (Table 1; Li et al., 2012). This database contains gene annotations, synteny, phylogeny, and gene family information for three legumes, soybean, *M. truncatula*, and *L. japonicus*, and two reference plant species, *A. thaliana* and *Populus trichocarpa* Torr. & A. Gray. In addition, large-scale microarray and RNA-seq gene expression data for above three legumes have also been integrated. The members of a gene family within and across the legume species display conserved and divergent expression patterns as revealed from LegumeIP database. For example, most of the members of leghemoglobin family in soybean exhibited highest expression in nodule and root whereas members from *M. truncatula* were expressed in nodules and members from

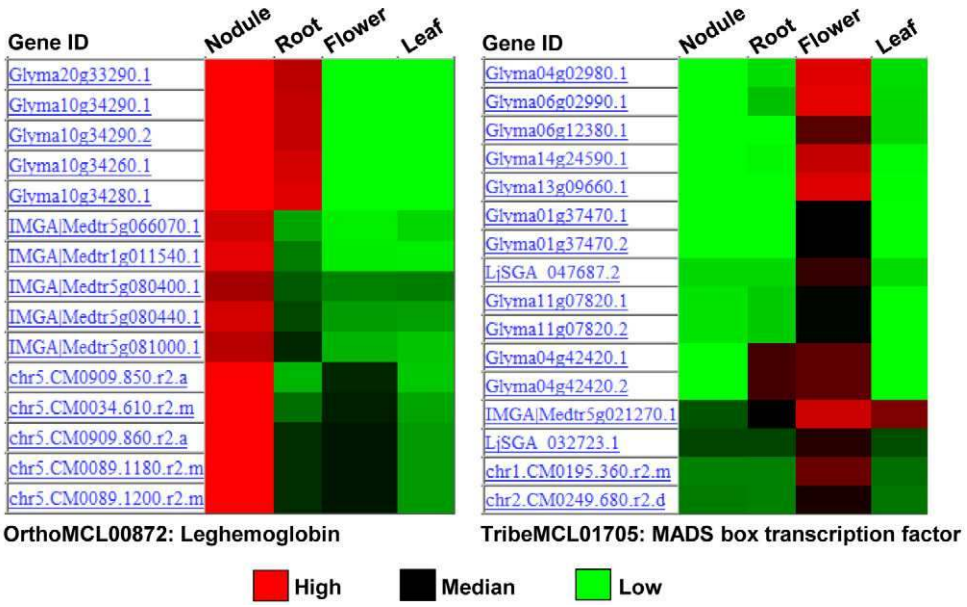

Figure 1. Heat map displaying the expression profile of all expressed genes in OrthoMCL00872 for leghemoglobins and TribeMCL01705 for MADS-box transcription factors. LegumeIP database (http://plantgrn.noble.org/LegumeIP/; accessed 24 May 2013) was searched for "leghemoglobin" and "MADS box" keywords and the representative clusters displaying conserved and divergent expression profiles of orthologous genes among the three legumes (soybean, *M. truncatula*, and *L. japonicus*) are shown.

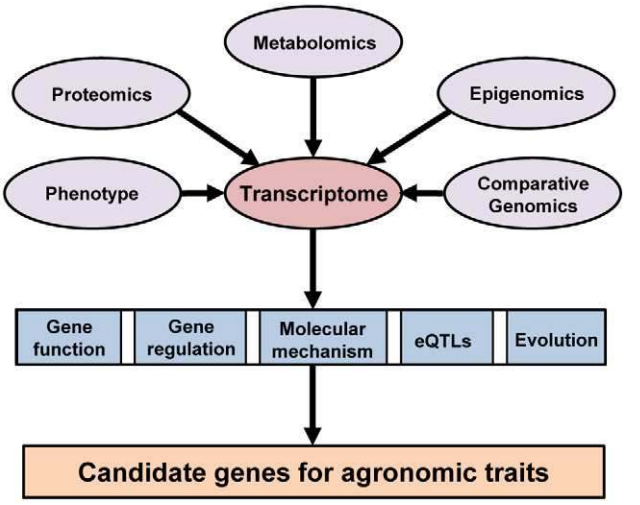

Figure 2. An integrated view of the transcriptome analysis for functional genomics and identification of candidate genes. The integration of global transcriptome data with proteomics, metabolomics, and epigenomics data along with phenotype information and comparative genomics can lead to elucidation of gene function and regulation, molecular mechanisms, evolution, and expression quantitative trait loci (eQTLs). This information can be used to select target candidate genes for manipulation of agronomic traits.

L. japonicus were expressed in nodule and flower (Fig. 1). Likewise, most of the MADS-box transcription factors from the three legumes included in an orthologous group were expressed specifically in flower with the exception of few genes expressed in root and leaf as well (Fig. 1).

Although comparative expression analysis seems to be very straightforward, the nonavailability of compatible expression datasets that cover equivalent tissues and/or conditions for all species makes it very difficult. In fact, only a small fraction of all available data in different species can be used for comparative analysis (Tirosh et al., 2007; Libault et al., 2010). To overcome this problem, the comparison of co-expression data instead of individual expression values has been proposed to identify the co-expressed genes across the species (Stuart et al., 2003; Ficklin and Feltus, 2011; Movahedi et al., 2011). However, only a few such co-expression studies across the species have been performed in plants, mainly in cereals and model plants (Ficklin and Feltus, 2011; Mutwil et al., 2011). Mutwil et al. (2011) included two legumes, soybean and *M. truncatula*, for co-expression studies along with five other plant species (*A. thaliana*, barley [*Hordeum vulgare* L.], poplar (*P. trichocarpa*), rice, and wheat [*Triticum aestivum* L.]) and developed a platform, PlaNet (http://aranet.mpimp-golm.mpg.de; accessed 24 May 2013), for identification of functional homologs across the species. However, more such studies are required to discover co-expression network, identify functional homologs, and understand complexity of transcriptional programs within and across the legumes.

Concluding Remarks

Transcriptome analyses at whole-genome level enable us not only to measure the activity of majority of genes but also gene regulation in different biological contexts. Although only a few transcriptome analyses studies have been performed in legumes so far, several key genes involved in important developmental processes have been identified. However, more in-depth investigations are required to understand the molecular mechanisms and gene regulatory network underlying various developmental processes in these interesting crops. Combining

transcriptomics with proteomics and metabolomics can provide correlative insights into the global molecular changes occurring during development and enhance our understanding about legume biology (Fig. 2). Furthermore, the integration of global transcriptome data with epigenetic modifications (DNA and histone modifications) and small RNA can elucidate another layer of gene regulation mechanism. The detailed analysis of changes in transcript abundance can lead to identification of expression quantitative trait loci, which, coupled with phenotypic information, can reveal the function of allelic variations. It is also anticipated that transcriptome analyses can help uncover the interactions of biochemical pathways and quantitative trait loci. The comparative transcriptome analyses can provide insights into conserved and divergent transcriptional programs among legumes and/or other plant species. The candidate genes identified from transcriptome analyses can be targeted for improvement of legume crops using reverse genetics or breeding approaches (Fig. 2). Overall, the transcriptome studies will surely accelerate functional and applied genomics research in legumes for their genetic enhancement.

Acknowledgments

The work on chickpea genomics is funded by the Department of Biotechnology, Government of India, under the Next Generation Challenge Programme on Chickpea Genomics (grant number BT/PR12919/AGR/02/676/2009 from 2009-2014). RG acknowledges INSPIRE Faculty Award from the Department of Science and Technology, Government of India.

References

Agarwal, G., S. Jhanwar, P. Priya, V.K. Singh, M.S. Saxena, S.K. Parida, R. Garg, A.K. Tyagi, and M. Jain. 2012. Comparative analysis of kabuli chickpea transcriptome with desi and wild chickpea provides a rich resource for development of functional markers. PLoS ONE 7:e52443. doi:10.1371/journal.pone.0052443

Andrés, F., and G. Coupland. 2012. The genetic basis of flowering responses to seasonal cues. Nat. Rev. Genet. 13:627–639. doi:10.1038/nrg3291

Benedito, V.A., I. Torres-Jerez, J.D. Murray, A. Andriankaja, S. Allen, K. Kakar, M. Wandrey, J. Verdier, H. Zuber, T. Ott, S. Moreau, A. Niebel, T. Frickey, G. Weiller, J. He, X. Dai, P.X. Zhao, Y. Tang, and M.K. Udvardi. 2008. A gene expression atlas of the model legume *Medicago truncatula*. Plant J. 55:504–513. doi:10.1111/j.1365-313X.2008.03519.x

Cannon, S.B., G.D. May, and S.A. Jackson. 2009. Three sequenced legume genomes and many crop species: Rich opportunities for translational genomics. Plant Physiol. 151:970–977. doi:10.1104/pp.109.144659

Chen, X., W. Zhu, S. Azam, H. Li, F. Zhu, H. Li, Y. Hong, H. Liu, E. Zhang, H. Wu, S. Yu, G. Zhou, S. Li, N. Zhong, S. Wen, X. Li, S.J. Knapp, P. Ozias-Akins, R.K. Varshney, and X. Liang. 2013. Deep sequencing analysis of the transcriptomes of peanut aerial and subterranean young pods identifies candidate genes related to early embryo abortion. Plant Biotechnol. J. 11:115–127. doi:10.1111/pbi.12018

Colebatch, G., S. Kloska, B. Trevaskis, S. Freund, T. Altmann, and M.K. Udvardi. 2002. Novel aspects of symbiotic nitrogen fixation uncovered by transcript profiling with cDNA arrays. Mol. Plant Microbe Interact. 15:411–420. doi:10.1094/MPMI.2002.15.5.411

Dash, S., J. Van Hemert, L. Hong, R.P. Wise, and J.A. Dickerson. 2012. PLEXdb: Gene expression resources for plants and plant pathogens. Nucleic Acids Research 40 (D1):D1194-D1201.

Dong, Z.C., Z. Zhao, C.W. Liu, J.H. Luo, J. Yang, W.H. Huang, X.H. Hu, T.L. Wang, and D. Luo. 2005. Floral patterning in *Lotus japonicus*. Plant Physiol. 137:1272–1282. doi:10.1104/pp.104.054288

El Yahyaoui, F., H. Küster, B. Ben Amor, N. Hohnjec, A. Pühler, A. Becker, J. Gouzy, T. Vernié, C. Gough, A. Niebel, L. Godiard, and P. Gamas. 2004. Expression profiling in *Medicago truncatula* identifies more than 750 genes differentially expressed during nodulation, including many potential regulators of the symbiotic program. Plant Physiol. 136:3159–3176. doi:10.1104/pp.104.043612

Fedorova, M., J. van de Mortel, P.A. Matsumoto, J. Cho, C.D. Town, K.A. VandenBosch, J.S. Gantt, and C.P. Vance. 2002. Genome-wide identification of nodule-specific transcripts in the model legume *Medicago truncatula*. Plant Physiol. 130:519–537. doi:10.1104/pp.006833

Ficklin, S.P., and F.A. Feltus. 2011. Gene coexpression network alignment and conservation of gene modules between two grass species: Maize and rice. Plant Physiol. 156:1244–1256. doi:10.1104/pp.111.173047

Gallardo, K., C. Firnhaber, H. Zuber, D. Héricher, M. Belghazi, C. Henry, H. Küster, and R. Thompson. 2007. A combined proteome and transcriptome analysis of developing *Medicago truncatula* seeds: Evidence for metabolic specialization of maternal and filial tissues. Mol. Cell. Proteomics 6:2165–2179. doi:10.1074/mcp.M700171-MCP200

Garg, R., and M. Jain. 2011. Pyrosequencing data reveals tissue-specific expression of lineage-specific transcripts in chickpea. Plant Signal. Behav. 6:1868–1870. doi:10.4161/psb.6.11.17879

Garg, R., S. Jhanwar, A.K. Tyagi, and M. Jain. 2010. Genome-wide survey and expression analysis suggest diverse roles of glutaredoxin gene family members during development and response to various stimuli in rice. DNA Res. 17:353–367. doi:10.1093/dnares/dsq023

Garg, R., R.K. Patel, S. Jhanwar, P. Priya, A. Bhattacharjee, G. Yadav, S. Bhatia, D. Chattopadhyay, A.K. Tyagi, and M. Jain. 2011. Gene discovery and tissue-specific transcriptome analysis in chickpea with massively parallel pyrosequencing and web resource development. Plant Physiol. 156:1661–1678. doi:10.1104/pp.111.178616

Graham, P.H., and C.P. Vance. 2003. Legumes: Importance and constraints to greater use. Plant Physiol. 131:872–877. doi:10.1104/pp.017004

Hansen, B.G., B.A. Halkier, and D.J. Kliebenstein. 2008. Identifying the molecular basis of QTLs: EQTLs add a new dimension. Trends Plant Sci. 13:72–77. doi:10.1016/j.tplants.2007.11.008

He, J., V.A. Benedito, M. Wang, J.D. Murray, P.X. Zhao, Y. Tang, and M.K. Udvardi. 2009. The *Medicago truncatula* gene expression atlas web server. BMC Bioinf. 10:441. doi:10.1186/1471-2105-10-441

Hecht, V., F. Foucher, C. Ferrándiz, R. Macknight, C. Navarro, J. Morin, M.E. Vardy, N. Ellis, J.P. Beltrán, C. Rameau, and J.L. Weller. 2005. Conservation of Arabidopsis flowering genes in model legumes. Plant Physiol. 137:1420–1434. doi:10.1104/pp.104.057018

't Hoen, P.A., Y. Ariyurek, H.H. Thygesen, E. Vreugdenhil, R.H. Vossen, R.X. de Menezes, J.M. Boer, G.J. van Ommen, and J.T. den Dunnen. 2008. Deep sequencing-based expression analysis shows major advances in robustness, resolution and inter-lab portability over five microarray platforms. Nucleic Acids Res. 36:e141. doi:10.1093/nar/gkn705

Hofer, J., L. Turner, R. Hellens, M. Ambrose, P. Matthews, A. Michael, and N. Ellis. 1997. UNIFOLIATA regulates leaf and flower morphogenesis in pea. Curr. Biol. 7:581–587. doi:10.1016/S0960-9822(06)00257-0

Jack, T. 2004. Molecular and genetic mechanisms of floral control. Plant Cell 16(Suppl.):S1–S17.

Jackson, S.A., A. Iwata, S.H. Lee, J. Schmutz, and R. Shoemaker. 2011. Sequencing crop genomes: Approaches and applications. New Phytol. 191:915–925. doi:10.1111/j.1469-8137.2011.03804.x

Jain, M. 2011. A next-generation approach to the characterization of a non-model plant transcriptome. Curr. Sci. 101:1435–1439.

Jain, M. 2012. Next-generation sequencing technologies for gene expression profiling in plants. Brief. Funct. Genomics 11:63–70. doi:10.1093/bfgp/elr038

Jain, M., G. Misra, R.K. Patel, P. Priya, S. Jhanwar, A.W. Khan, N. Shah, V.K. Singh, R. Garg, G. Jeena, M. Yadav, C. Kant, P. Sharma, G. Yadav, S. Bhatia, A.K. Tyagi, and D. Chattopadhyay. 2013. A draft genome sequence of the pulse crop chickpea (*Cicer arietinum* L.). Plant J. 74:715–729. doi:10.1111/tpj.12173

Jhanwar, S., P. Priya, R. Garg, S.K. Parida, A.K. Tyagi, and M. Jain. 2012. Transcriptome sequencing of wild chickpea as a rich resource for marker development. Plant Biotechnol. J. 10:690–702. doi:10.1111/j.1467-7652.2012.00712.x

Jiao, Y., L. Ma, E. Strickland, and X.W. Deng. 2005. Conservation and divergence of light-regulated genome expression patterns during seedling development in rice and Arabidopsis. Plant Cell 17:3239–3256. doi:10.1105/tpc.105.035840

Journet, E.P., D. van Tuinen, J. Gouzy, H. Crespeau, V. Carreau, M.J. Farmer, A. Niebel, T. Schiex, O. Jaillon, O. Chatagnier, L. Godiard, F. Micheli, D. Kahn, V. Gianinazzi-Pearson, and P. Gamas. 2002. Exploring root symbiotic programs in the model legume *Medicago truncatula* using EST analysis. Nucleic Acids Res. 30:5579–5592. doi:10.1093/nar/gkf685

Jung, C.H., C.E. Wong, M.B. Singh, and P.L. Bhalla. 2012. Comparative genomic analysis of soybean flowering genes. PLoS ONE 7:e38250. doi:10.1371/journal.pone.0038250

Kater, M.M., L. Dreni, and L. Colombo. 2006. Functional conservation of MADS-box factors controlling floral organ identity in rice and Arabidopsis. J. Exp. Bot. 57:3433–3444. doi:10.1093/jxb/erl097

Keurentjes, J.J., J. Fu, I.R. Terpstra, J.M. Garcia, G. van den Ackerveken, L.B. Snoek, A.J. Peeters, D. Vreugdenhil, M. Koornneef, and R.C. Jansen. 2007. Regulatory network construction in Arabidopsis by using genome-wide gene expression quantitative trait loci. Proc. Natl. Acad. Sci. USA 104:1708–1713. doi:10.1073/pnas.0610429104

Krizek, B.A., and J.C. Fletcher. 2005. Molecular mechanisms of flower development: An armchair guide. Nat. Rev. Genet. 6:688–698. doi:10.1038/nrg1675

Le, B.H., J.A. Wagmaister, T. Kawashima, A.Q. Bui, J.J. Harada, and R.B. Goldberg. 2007. Using genomics to study legume seed development. Plant Physiol. 144:562–574. doi:10.1104/pp.107.100362

Li, J., X. Dai, T. Liu, and P.X. Zhao. 2012. LegumeIP: An integrative database for comparative genomics and transcriptomics of model legumes. Nucleic Acids Res. 40(Database issue):D1221–D1229. doi:10.1093/nar/gkr939

Li, P., L. Ponnala, N. Gandotra, L. Wang, Y. Si, S.L. Tausta, T.H. Kebrom, N. Provart, R. Patel, C.R. Myers, E.J. Reidel, R. Turgeon, P. Liu, Q. Sun, T. Nelson, and T.P. Brutnell. 2010. The developmental dynamics of the maize leaf transcriptome. Nat. Genet. 42:1060–1067. doi:10.1038/ng.703

Libault, M., A. Farmer, T. Joshi, K. Takahashi, R.J. Langley, L.D. Franklin, J. He, D. Xu, G. May, and G. Stacey. 2010. An integrated transcriptome atlas of the crop model *Glycine max*, and its use in comparative analyses in plants. Plant J. 63:86–99.

Marioni, J.C., C.E. Mason, S.M. Mane, M. Stephens, and Y. Gilad. 2008. RNA-seq: An assessment of technical reproducibility and comparison with gene expression arrays. Genome Res. 18:1509–1517. doi:10.1101/gr.079558.108

Martin, J.A., and Z. Wang. 2011. Next-generation transcriptome assembly. Nat. Rev. Genet. 12:671–682. doi:10.1038/nrg3068

Moreau, S., M. Verdenaud, T. Ott, S. Letort, F. de Billy, A. Niebel, J. Gouzy, F. de Carvalho-Niebel, and P. Gamas. 2011. Transcription reprogramming during root nodule development in *Medicago truncatula*. PLoS ONE 6:e16463. doi:10.1371/journal.pone.0016463

Movahedi, S., Y. Van de Peer, and K. Vandepoele. 2011. Comparative network analysis reveals that tissue specificity and gene function are important factors influencing the mode of expression evolution in Arabidopsis and rice. Plant Physiol. 156:1316–1330. doi:10.1104/pp.111.177865

Mustroph, A., S.C. Lee, T. Oosumi, M.E. Zanetti, H. Yang, K. Ma, A. Yaghoubi-Masihi, T. Fukao, and J. Bailey-Serres. 2010. Cross-kingdom comparison of transcriptomic adjustments to low-oxygen stress highlights conserved and plant-specific responses. Plant Physiol. 152:1484–1500. doi:10.1104/pp.109.151845

Mutwil, M., S. Klie, T. Tohge, F.M. Giorgi, O. Wilkins, M.M. Campbell, A.R. Fernie, B. Usadel, Z. Nikoloski, and S. Persson. 2011. PlaNet: Combined sequence and expression comparisons across plant networks derived from seven species. Plant Cell 23:895–910. doi:10.1105/tpc.111.083667

Ng, M., and M.F. Yanofsky. 2001. Function and evolution of the plant MADS-box gene family. Nat. Rev. Genet. 2:186–195. doi:10.1038/35056041

Ozsolak, F., and P.M. Milos. 2011. RNA sequencing: Advances, challenges and opportunities. Nat. Rev. Genet. 12:87–98. doi:10.1038/nrg2934

Pang, Y., G.J. Peel, S.B. Sharma, Y. Tang, and R.A. Dixon. 2008. A transcript profiling approach reveals an epicatechin-specific glucosyltransferase expressed in the seed coat of *Medicago truncatula*. Proc. Natl. Acad. Sci. USA 105:14210–14215. doi:10.1073/pnas.0805954105

Rensink, W.A., and C.R. Buell. 2005. Microarray expression profiling resources for plant genomics. Trends Plant Sci. 10:603–609. doi:10.1016/j.tplants.2005.10.003

Sato, S., Y. Nakamura, T. Kaneko, E. Asamizu, T. Kato, M. Nakao, S. Sasamoto, A. Watanabe, A. Ono, K. Kawashima, T. Fujishiro, M. Katoh, M. Kohara, Y. Kishida, C. Minami, S. Nakayama, N. Nakazaki, Y. Shimizu, S. Shinpo, C. Takahashi, T. Wada, M. Yamada, N. Ohmido, M. Hayashi, K. Fukui, T. Baba, T. Nakamichi, H. Mori, and S. Tabata. 2008. Genome structure of the legume, *Lotus japonicus*. DNA Res. 15:227–239. doi:10.1093/dnares/dsn008

Schmutz, J., S.B. Cannon, J. Schlueter, J. Ma, T. Mitros, W. Nelson, D.L. Hyten, Q. Song, J.J. Thelen, J. Cheng, D. Xu, U. Hellsten, G.D. May, Y. Yu, T. Sakurai, T. Umezawa, M.K. Bhattacharyya, D. Sandhu, B. Valliyodan, E. Lindquist, M. Peto, D. Grant, S. Shu, D. Goodstein, K. Barry, M. Futrell-Griggs, B. Abernathy, J. Du, Z. Tian, L. Zhu, N. Gill, T. Joshi, M. Libault, A. Sethuraman, X.C. Zhang, K. Shinozaki, H.T. Nguyen, R.A. Wing, P. Cregan, J. Specht, J. Grimwood, D. Rokhsar, G. Stacey, R.C. Shoemaker, and S.A. Jackson. 2010. Genome sequence of the palaeopolyploid soybean. Nature 463:178–183. doi:10.1038/nature08670

Schultze, M., and A. Kondorosi. 1998. Regulation of symbiotic root nodule development. Annu. Rev. Genet. 32:33–57. doi:10.1146/annurev.genet.32.1.33

Severin, A.J., J.L. Woody, Y.T. Bolon, B. Joseph, B.W. Diers, A.D. Farmer, G.J. Muehlbauer, R.T. Nelson, D. Grant, J.E. Specht, M.A. Graham, S.B. Cannon, G.D. May, C.P. Vance, and R.C. Shoemaker. 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, R., P. Agarwal, S. Ray, P. Deveshwar, P. Sharma, N. Sharma, A. Nijhawan, M. Jain, A.K. Singh, V.P. Singh, J.P. Khurana, A.K. Tyagi, and S. Kapoor. 2012. Expression dynamics of metabolic and regulatory components across stages of panicle and seed development in *indica* rice. Funct. Integr. Genomics 12:229–248. doi:10.1007/s10142-012-0274-3

Singh, V.K., R. Garg, and M. Jain. 2013. A global view of transcriptome dynamics during flower development in chickpea by deep sequencing. Plant Biotechnol. J. 10.1111/pbi.12059.

Stacey, G., M. Libault, L. Brechenmacher, J. Wan, and G.D. May. 2006. Genetics and functional genomics of legume nodulation. Curr. Opin. Plant Biol. 9:110–121. doi:10.1016/j.pbi.2006.01.005

Street, N.R., A. Sjödin, M. Bylesjö, P. Gustafsson, J. Trygg, and S. Jansson. 2008. A cross-species transcriptomics approach to identify genes involved in leaf development. BMC Genomics 9:589. doi:10.1186/1471-2164-9-589

Stuart, J.M., E. Segal, D. Koller, and S.K. Kim. 2003. A gene-coexpression network for global discovery of conserved genetic modules. Science 302:249–255. doi:10.1126/science.1087447

Taylor, S., J. Hofer, and I. Murfet. 2001. *Stamina pistilloida*, the pea ortholog of Fim and UFO, is required for normal development of flowers, inflorescences, and leaves. Plant Cell 13:31–46.

Thimm, O., O. Blasing, Y. Gibon, A. Nagel, S. Meyer, P. Kruger, J. Selbig, L.A. Muller, S.Y. Rhee, and M. Stitt. 2004. MAPMAN: A user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. Plant J. 37:914-939. doi:10.1111/j.1365-313X.2004.02016.x

Tirosh, I., Y. Bilu, and N. Barkai. 2007. Comparative biology: Beyond sequence analysis. Curr. Opin. Biotechnol. 18:371–377. doi:10.1016/j.copbio.2007.07.003

Tucker, S.C. 1987. Floral initiation and development in legumes. In: C.H. Stirton, editor, Advances in legume systematics part 3. Royal Botanic Gardens, Kew, UK. p. 183–239.

Tucker, S.C. 2003. Floral development in legumes. Plant Physiol. 131:911–926. doi:10.1104/pp.102.017459

Varshney, R.K., W. Chen, Y. Li, A.K. Bharti, R.K. Saxena, J.A. Schlueter, M.T. Donoghue, S. Azam, G. Fan, A.M. Whaley, A.D. Farmer, J. Sheridan, A. Iwata, R. Tuteja, R.V. Penmetsa, W. Wu, H.D. Upadhyaya, S.P. Yang, T. Shah, K.B. Saxena, T. Michael, W.R. McCombie, B. Yang, G. Zhang, H. Yang, J. Wang, C. Spillane, D.R. Cook, G.D. May, X. Xu, and S.A. Jackson. 2012. Draft genome sequence of pigeonpea (*Cajanus cajan*), an orphan legume crop of resource-poor farmers. Nat. Biotechnol. 30:83–89. doi:10.1038/nbt.2022

Varshney, R.K., T.J. Close, N.K. Singh, D.A. Hoisington, and D.R. Cook. 2009. Orphan legume crops enter the genomics era! Curr. Opin. Plant Biol. 12:202–210. doi:10.1016/j.pbi.2008.12.004

Varshney, R.K., C. Song, R.K. Saxena, S. Azam, S. Yu, A.G. Sharpe, S. Cannon, J. Baek, B.D. Rosen, B. Tar'an, T. Millan, X. Zhang, L.D. Ramsay, A. Iwata, Y. Wang, W. Nelson, A.D. Farmer, P.M. Gaur, C. Soderlund, R.V. Penmetsa, C. Xu, A.K. Bharti, W. He, P. Winter, S. Zhao, J.K. Hane, N. Carrasquilla-Garcia, J.A. Condie, H.D. Upadhyaya, M.C. Luo, M. Thudi, C.L. Gowda, N.P. Singh, J. Lichtenzveig, K.K. Gali, J. Rubio, N. Nadarajan, J. Dolezel, K.C. Bansal, X. Xu, D. Edwards, G. Zhang, G. Kahl, J. Gil, K.B. Singh, S.K. Datta, S.A. Jackson, J. Wang, and D.R. Cook. 2013. Draft genome sequence of chickpea (*Cicer arietinum*) provides a resource for trait improvement. Nat. Biotechnol. 31:240–246. doi:10.1038/nbt.2491

Verdier, J., F. Dessaint, C. Schneider, and M. Abirached-Darmency. 2013a. A combined histology and transcriptome analysis unravels novel questions on *Medicago truncatula* seed coat. J. Exp. Bot. 64:459–470. doi:10.1093/jxb/ers304

Verdier, J., I. Torres-Jerez, M. Wang, A. Andriankaja, S.N. Allen, J. He, Y. Tang, J.D. Murray, and M.K. Udvardi. 2013b. Establishment of the *Lotus japonicus* gene expression atlas (LjGEA) and its use to explore legume seed maturation. Plant J. 74:351–362. doi:10.1111/tpj.12119

Wang, Z., M. Gerstein, and M. Snyder. 2009. RNA-Seq: A revolutionary tool for transcriptomics. Nat. Rev. Genet. 10:57–63. doi:10.1038/nrg2484

Wong, C.E., M.B. Singh, and P.L. Bhalla. 2009. Molecular processes underlying the floral transition in the soybean shoot apical meristem. Plant J. 57:832–845. doi:10.1111/j.1365-313X.2008.03730.x

Yang, H., P. Lu, Y. Wang, and H. Ma. 2011. The transcriptome landscape of Arabidopsis male meiocytes from high-throughput sequencing: The complexity and evolution of the meiotic process. Plant J. 65:503–516. doi:10.1111/j.1365-313X.2010.04439.x

Young, N.D., F. Debellé, G.E. Oldroyd, R. Geurts, S.B. Cannon, M.K. Udvardi, V.A. Benedito, K.F. Mayer, J. Gouzy, H. Schoof, Y. Van de Peer, S. Proost, D.R. Cook, B.C. Meyers, M. Spannagl, F. Cheung, S. De Mita, V. Krishnakumar, H. Gundlach, S. Zhou, J. Mudge, A.K. Bharti, J.D. Murray, M.A. Naoumkina, B. Rosen, K.A. Silverstein, H. Tang, S. Rombauts, P.X. Zhao, P. Zhou, V. Barbe, P. Bardou, M. Bechner, A. Bellec, A. Berger, H. Bergès, S. Bidwell, T. Bisseling, N. Choisne, A. Couloux, R. Denny, S. Deshpande, X. Dai, J.J. Doyle, A.M. Dudez, A.D. Farmer, S. Fouteau, C. Franken, C. Gibelin, J. Gish, S. Goldstein, A.J. González, P.J. Green, A. Hallab, M. Hartog, A. Hua, S.J. Humphray, D.H. Jeong, Y. Jing, A. Jöcker, S.M. Kenton, D.J. Kim, K. Klee, H. Lai, C. Lang, S. Lin, S.L. Macmil, G. Magdelenat, L. Matthews, J. McCorrison, E.L. Monaghan, J.H. Mun, F.Z. Najar, C. Nicholson, C. Noirot, M. O'Bleness, C.R. Paule, J. Poulain, F. Prion, B. Qin, C. Qu, E.F. Retzel, C. Riddle, E. Sallet, S. Samain, N. Samson, I. Sanders, O. Saurat, C. Scarpelli, T. Schiex, B. Segurens, A.J. Severin, D.J. Sherrier, R. Shi, S. Sims, S.R. Singer, S. Sinharoy, L. Sterck, A. Viollet, B.B. Wang, K. Wang, M. Wang, X. Wang, J. Warfsmann, J. Weissenbach, D.D. White, J.D. White, G.B. Wiley, P. Wincker, Y. Xing, L. Yang, Z. Yao, F. Ying, J. Zhai, L. Zhou, A. Zuber, J. Dénarié, R.A. Dixon, G.D. May, D.C. Schwartz, J. Rogers, F. Quétier, C.D. Town, and B.A. Roe. 2011. The *Medicago* genome provides insight into the evolution of rhizobial symbioses. Nature 480:520–524. doi:10.1038/480162a

Zenoni, S., A. Ferrarini, E. Giacomelli, L. Xumerle, M. Fasoli, G. Malerba, D. Bellin, M. Pezzotti, and M. Delledonne. 2010. Characterization of transcriptional complexity during berry development in *Vitis vinifera* using RNA-Seq. Plant Physiol. 152:1787–1795. doi:10.1104/pp.109.149716