Plant Biotechnology Journal

Plant Biotechnology Journal (2013) **11**, pp. 691–701 doi: 10.1111/pbi.12059

A global view of transcriptome dynamics during flower development in chickpea by deep sequencing

Vikash K. Singh, Rohini Garg and Mukesh Jain*

National Institute of Plant Genome Research (NIPGR), Aruna Asaf Ali Marg, New Delhi 110067, India

Received 5 December 2012;
revised 22 January 2013;
accepted 29 January 2013.
**Correspondence* (fax: +91 11 26741658;
email: mjain@nipgr.ac.in)

Keywords: Chickpea, transcriptome, gene expression, flower development, RNA-seq, metabolic pathways.

Summary

Measurement of gene expression can provide important clues about gene function and molecular basis of developmental processes. Here, we have analysed the chickpea transcriptome in vegetative and flower tissues by exploiting the potential of high-throughput sequencing to measure gene expression. We mapped more than 295 million reads to quantify the transcript abundance during flower development. We detected the expression of more than 90% genes in at least one tissue analysed. We found quite a large number of genes were differentially expressed during flower development as compared to vegetative tissues. Further, we identified several genes expressed in a stage-specific manner. Various transcription factor families and metabolic pathways involved in flower development were elucidated. The members of MADS-box family were most represented among the transcription factor genes up-regulated during various stages of flower development. The abundant expression of several well-known genes implicated in flower development in chickpea flower development stages confirmed our results. In addition, we detected the expression specificities of lineage-specific genes during flower development. The expression data presented in this study is the most comprehensive dataset available for chickpea as of now and will serve as resource for unraveling the functions of many specific genes involved in flower development in chickpea and other legumes.

Introduction

Chickpea is an economically and nutritionally important grain legume especially for developing countries. Realizing its importance as crop plant, it is only recently a few genomic resources have been generated for chickpea. The transcriptomes of cultivated and wild chickpea genotypes have been sequenced for discovery of novel genes and genetic variations to facilitate genetic enhancement of chickpea (Agarwal *et al.*, 2012; Garg *et al.*, 2011; Jhanwar *et al.*, 2012). However, less genetic variability has been one of the important reasons for limited progress made in improvement of chickpea so far. Considering the limitations, it will be important to use biotechnological approaches together with breeding techniques for chickpea improvement, which will require comprehensive understanding of molecular mechanisms underlying various developmental processes.

Flower development is most important developmental event in the life cycle of higher plants. The molecular genetic control of flower development has been well studied in Arabidopsis and to some extent in few other plant species as well (Andrés and Coupland, 2012; Jack, 2004; Krizek and Fletcher, 2005; Lohmann and Weigel, 2002). The extensive genetic analyses in Arabidopsis have identified several key regulatory genes of flower development. Most of these genes encode for transcription factors, which indicates the existence of a complex gene regulatory network underlying flower development. Considering the morphological specificities possessed by different lineages of plants, it is important to study these developmental processes in other important plant species such as legumes. Based on morphological analyses, it has been proposed that flower development in legumes does not follow the ABC model of eudicots (Tucker, 1987, 2003). Although a few genes that control various aspects of flower development have been discovered in legumes especially in pea (Hecht *et al.*, 2005; Jung *et al.*, 2012; Taylor *et al.*, 2002; Tucker, 2003), a genome-wide study of molecular basis of flower development is lacking. It is imperative to gain an understanding of molecular basis of flower development, which will permit the genetic engineering of this important agronomic trait.

The whole transcriptome sequencing using next-generation sequencing technologies (RNA-seq) is the convenient and rapid means to study the gene expression at whole-genome level and defining putative gene function (Jain, 2012; Ozsolak and Milos, 2011; Wang *et al.*, 2009). Many studies have already demonstrated the power of RNA-seq in various biological contexts (Gonzalez-Ballester *et al.*, 2010; Li *et al.*, 2010; Weber *et al.*, 2007; Yang *et al.*, 2011; Zenoni *et al.*, 2010). Although rapid advances have been made towards understanding of the transcriptional programs involved in specific developmental processes of several plant species, only a few such studies have been performed in legumes (Benedito *et al.*, 2008; Libault *et al.*, 2010; Severin *et al.*, 2010). The availability of RNA-seq technique provides an excellent opportunity to perform such studies in legumes as well.

In this study, we analysed the global gene expression profiles during flower development in chickpea. High-throughput Illumina sequencing was performed from different stages of flower development and three vegetative tissues. Based on extensive data analyses, we identified the genes and pathways involved in flower development in chickpea. This data set will serve as the foundation to understand the exact regulatory mechanisms underlying flower development in legumes.

Results

Transcriptome sequencing

To generate the inventory of gene expression during flower development, we collected eight successive stages of flower development, from young flower buds to mature senescing flowers [flower buds at sizes 4 mm (FB1), 6 mm (FB2), 8 mm (FB3), 8-10 mm (FB4) and flowers with closed petals (FL1), partially opened petals (FL2), opened and faded petals (FL3) and senescing petals (FL4)], which could be clearly distinguished morphologically (Figure 1a). In addition, we included three vegetative tissues (germinating seedling, GS; young leaves, YL and shoot apical meristem, SAM) for comparative analyses. To profile the flower transcriptome, we sequenced the RNA-seq libraries from different tissues on Illumina platform. We generated about 234 million high-quality reads for eight flower development stages (ranging from 16 to 40 million reads for each stage) and 91 million high-quality reads from three vegetative tissues (Table 1). Because of nonavailability of reference genome sequence, we mapped the reads on to the chickpea transcriptome comprised of 34 760 transcripts (referred as genes hereafter) reported previously (Garg *et al.*, 2011). The mapping showed that most (91.1%) of high-quality Illumina reads from each tissue sample aligned to the chickpea genes (Table 1). Of the mapped reads, 78.6% reads were aligned uniquely to a single chickpea gene. This suggests that the majority of transcriptionally active genes have been captured in our initial transcriptome assembly (Garg *et al.*, 2011).

Global view of the chickpea floral transcriptome

For the estimation of overall transcriptional activity, we determined reads per kilobase per million of mapped reads (RPKM) for each gene in different tissues analysed. RPKM has been proposed as a normalized quantitative method for gene expression using RNA-seq data (Mortazavi *et al.*, 2008). Based on RPKM value, we determined the number of genes expressed (RPKM $\geq$ 1) in individual tissue sample (Figure 1b). Largest number of genes were transcriptionally active in vegetative tissues, GS (86.6%) and SAM (85.2%). Among the flower development stages, number of transcriptionally active genes ranged from 79.2% (FL3 stage) to 84.6% (FB1 stage). Overall, expression of at least 92.8% genes could be detected in vegetative tissues and 93.8% in flower tissues. The expression of only 1176 (3.4%) genes could not be detected in any tissue analysed. These results indicate that the expression of most of the genes could be detected in the tissues analysed and confirm that high-throughput sequencing technologies can be used to survey global gene expression in chickpea. Further, we analysed the expression profiles in more detail to identify the genes expressed in multiple/specific tissues. The comparison across the tissues identified a large fraction of genes (21 364) expressed in all the tissues analysed. About 21–28% genes exhibited transcriptional activity in more than one tissue. However, only a small fraction (0.04–1.5%) of genes were expressed in a specific tissue. Among vegetative tissues, largest number of genes were specifically expressed in GS (1.5%) and in FB1 (0.6%) stage among the flower tissues.

Different tissues of a plant can be distinguished with each other based on their transcriptional signatures (Schmid *et al.*, 2005; Sharma *et al.*, 2012). Principal component analysis (PCA) was carried out on the whole-gene expression data set to investigate the relationships between the chickpea tissues used in this study. This analysis clearly revealed the distinctness of vegetative tissues from the flower tissues (Figure S1), indicating the presence of different transcriptional programmes. In addition, among the flower tissues, the flower bud stages (FB1–FB4) clustered closely. Likewise, stages of flower development (FL1–FL4) were also clustered together. This indicates similar transcriptional programmes active within the flower bud and flower development stages, but somewhat distinct from each other.

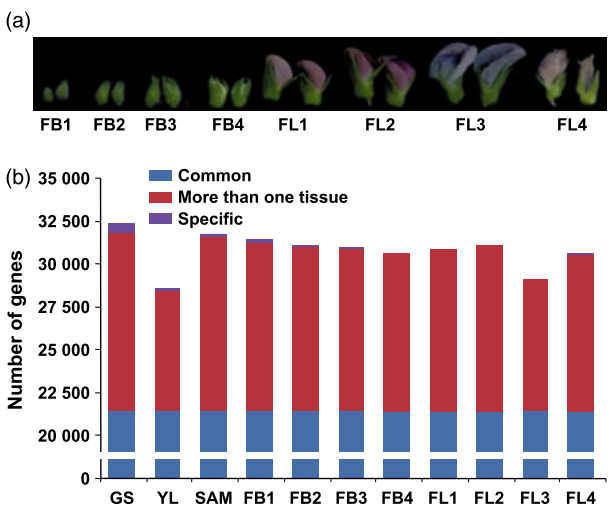

Figure 1 Global view of gene expression during flower development in chickpea. (a) The stages of flower development used in this study. (b) Overview of gene expression in different tissues. The total number of genes expressed in each tissue classified as expressed in all the tissues (common), more than one tissue and specific to each tissue are shown as stacked bar graph. GS, germinating seedling; YL, young leaf; SAM, shoot apical meristem; FB1–FB4 flower bud stages; FL1–FL4 flower stages.

Transcriptome dynamics during flower development

To study the transcriptome dynamics and identify the candidate genes involved in flower development, we followed three steps. In the first step, we compared the transcript abundance of all chickpea genes in flower development stages with individual vegetative tissue. In the second step, we identified genes differentially expressed at each stage of flower development as compared to all the three vegetative tissues. In the third step, we used a subtractive strategy to identify the genes specifically up-regulated at each stage of flower development by excluding the genes which are up-regulated at any other flower development stage.

In the first step of comparison with individual vegetative tissue, the number of down-regulated genes was higher as compared to up-regulated genes in most stages of flower development, when compared with GS (Figure 2a). This might be due to the reason that GS is comprised of dividing cells and a large fraction of genes are highly transcriptionally active. In contrast, higher number of genes exhibited up-regulation in most flower development stages when compared with YL and SAM tissues. Overall, the maximum number of genes were differentially expressed in all the flower stages as compared to GS. In the second step of analysis, the highest number of genes exhibited differential expression in FB4 followed by FL4 stage with respect to all the vegetative tissues, whereas least number of genes were differentially expressed in

Table 1 Summary of sequencing data generated and mapping to the chickpea transcriptome

Tissue	Total reads	High-quality reads	Total mapped reads (%)	Uniquely mapped reads (%)
Germinating seedling (GS)	27 889 925	25 954 346	23 927 246 (92.19)	20 569 087 (79.25)
Young leaf (YL)	32 762 959	30 339 998	27 158 340 (89.51)	23 028 092 (75.9)
Shoot apical meristem (SAM)	36 003 749	34 508 868	31 511 021 (91.31)	26 893 637 (77.93)
Flower bud (4 mm; FB1)	41 850 211	36 409 370	33 854 669 (92.98)	29 565 849 (81.2)
Flower bud (6 mm; FB2)	17 827 737	16 064 057	14 536 530 (90.49)	12 516 086 (77.91)
Flower bud (8 mm; FB3)	41 365 615	39 663 080	35 643 139 (89.86)	30 916 529 (77.95)
Flower bud opened (8 -10 mm; FB4)	35 211 613	32 605 824	29 706 745 (91.11)	25 922 104 (79.5)
Flower (unopened; FL1)	29 091 495	27 624 457	24 784 511 (89.72)	21 540 730 (77.98)
Flower (opened; FL2)	33 634 177	32 109 896	29 207 580 (90.96)	25 357 808 (78.97)
Flower (mature; FL3)	26 396 381	22 418 080	20 366 054 (90.85)	17 439 145 (77.79)
Flower (drooped; FL4)	30 245 721	26 865 870	24 950 732 (92.87)	21 523 367 (80.11)

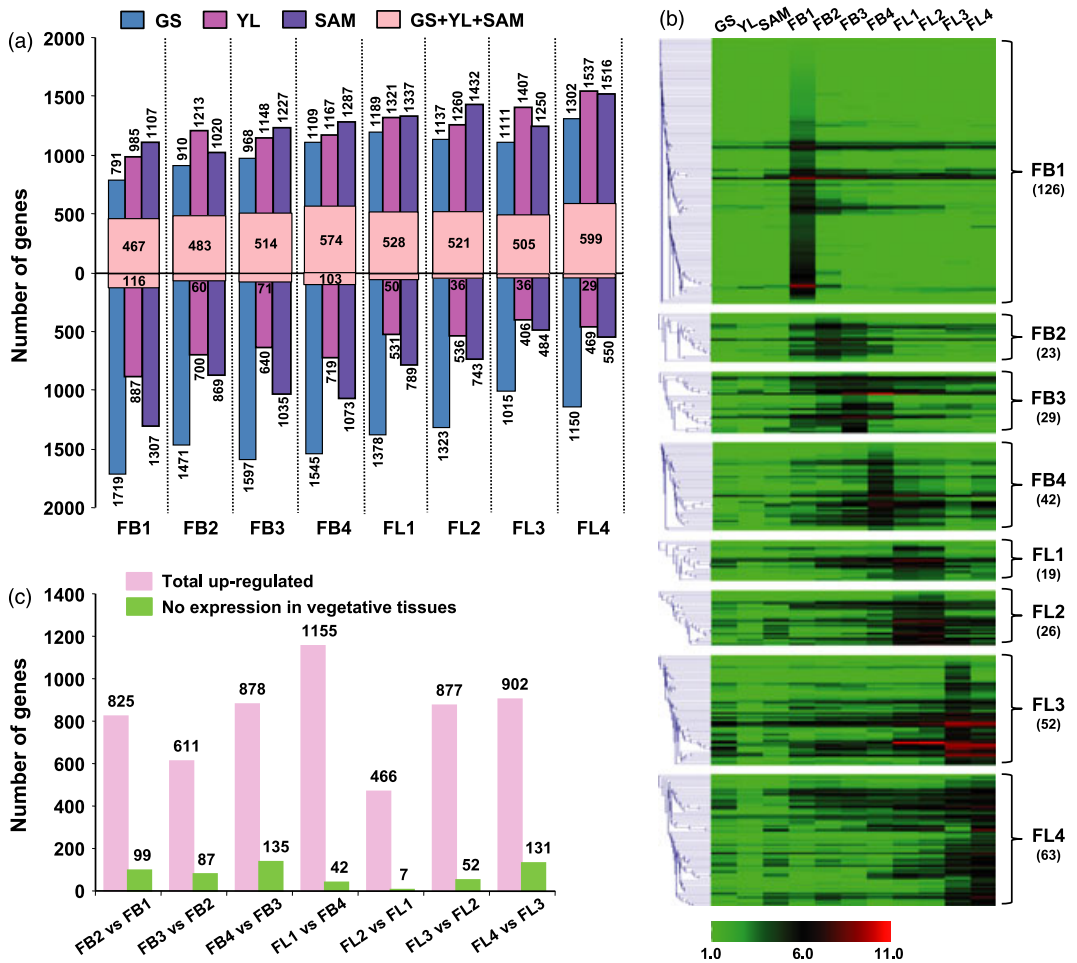

Figure 2 Gene expression dynamics during various stages of flower development. (a) Differential gene expression during flower development as compared to vegetative tissues. The number of up- (upper side) and down-regulated (lower side) genes as compared to each of three vegetative tissues (GS, YL and SAM) in each stage of flower development is shown. The number of up- (upper side) and down-regulated (lower side) genes as compared to all the three vegetative tissues (GS+YL+SAM) in each stage of flower development is also shown. FB1–FB4 flower bud stages; FL1–FL4 flower stages. (b) The expression profile of genes showing stage-specific expression is shown in the heatmap. The number of genes is given on the right side. The colour scale at the bottom represents the log-transformed RPKM value. (c) The number of up-regulated genes in various stages of flower development with respect to the preceding stage is given. In addition, the number of up-regulated genes with no expression (RPKM <1) in vegetative tissues are also shown.

FL3 followed by FB2 stage (Figure 2a). Overall, a total of 1572 genes were differentially expressed in at least one flower bud/flower stage analysed. Among the 1118 (908 up- and 210 down-regulated) and 966 (857 up- and 109 down-regulated) genes differentially expressed in flower bud and flower stages, respectively, 511 (461 up- and 50 down-regulated) genes were

commonly regulated. In the third step as described previously, we identified a total of 380 genes showing significantly higher transcript abundance at any particular stage (Figure 2b). At FB1 stage, highest (126) number of genes exhibited specific expression followed by FL4 (63) and FL3 (52) stages. A list of all the genes differentially expressed during any stage of flower development is given in Table S1.

All the stages of flower development analysed in this study show significant morphological differences, which indicate substantial differences in their transcriptional programmes. To identify the genes, which might be responsible for the regulation of specific developmental events in each stage of development, we compared the expression profiles of all the genes with respect to its preceding stage (Figure 2c). This analysis identified 825, 611 and 878 genes significantly up-regulated in FB2, FB3 and FB4 stages, respectively, as compared to their preceding stage. Likewise, 1155, 466, 877 and 902 genes were up-regulated in FL1 to FL4 stages, respectively, as compared to their preceding stage. Notably, higher number of genes was up-regulated in FL1 stage as compared to FB4 stage, which indicates a large shift in the transcriptional programmes, when flower bud opens into a flower. Further, we subtracted the genes expressed in any vegetative tissues (RPKM $\geq$ 1) and very low expression in the flower development stage (RPKM $<$ 3) to identify the gene expressed significantly only at particular stage. The filtering resulted in the set of 290 genes with a much lower number specific to each stage (Figure 2c). A list of these genes is provided in Table S2. These genes might play an important role in the developmental events specific to a particular stage of flower development.

Gene ontology enrichment and metabolic pathway analysis

To reveal the major functional categories represented in the genes involved in flower development as identified earlier, we performed gene ontology (GO) enrichment analysis. We found the genes involved in various aspects of flower development, including maintenance of floral organ identity, development of corolla, petal, androecium and gynoecium, gametophyte development, etc. significantly enriched in the genes up-regulated during flower bud stages (Figures 3a and S2a). In addition, the genes related to transport and cell wall organization were also enriched in flower bud stages. The genes involved in anther and pollen development contributed most to the flower bud up-regulated genes. The molecular function GO terms, pectinesterase activity, carboxyesterase activity and enzyme inhibitor activity were most significantly enriched in flower bud up-regulated genes (Figure S2a). The genes associated with endomembrane system and Golgi transport complex were most significant among the cellular component GO terms. Several biological process, molecular function and cellular component GO terms were common among the flower bud and flower stages. The genes involved in floral organ development were enriched in flower stages also (Figure 3b). In addition, cell differentiation, cell wall organization and biogenesis, and fatty acid transport were represented in the flower stages (Figure S2b). The molecular function term UDP-glucosyltransferase activity was more significant in flower stages as compared to flower bud stages. In addition, transferase activity and lipid transport were unique GO terms represented in flower stage gene set only.

Several genes differentially expressed during flower development encode for enzymes involved in metabolic functions. To identify the metabolic pathways involved in flower development, we integrated the up-regulated gene set to the metabolic network available in AraCyc database. The enrichment analysis revealed a few metabolic pathways significantly over-represented in flower development processes. Most evidently, genes involved in sporopollenin precursor biosynthesis, monolignol glucoside biosynthesis and cytokinin glucoside biosynthesis were significantly enriched in the data (Figure 4). In addition, the pathways for biosynthesis of colourless intermediates (leucopelargonidin, leucocyanidin and leucodelphinidin) of anthocyanin production and anthocyanin biosynthesis were enriched. The analysis revealed that the genes involved in homogalacturonan (HGA) degradation and glycerophosphodiester degradation were also up-regulated during flower development.

Further, we analysed the enrichment of GO terms and metabolic pathways in the genes up-regulated in each stage as compared to preceding stage (Figure S3). Although, many of the enriched GO terms were common, a few of them were unique in different set of genes. For example, GO terms related to energy metabolism were enriched in FB2 stage, pollen development in FB3 and FB4 stages, fatty acid and lipid transport in FL1 stage, and defence response in FL2 and FL3 stages. Likewise, a few specific metabolic pathways were also enriched significantly at each stage. Notably, metabolic pathways involved in energy generation were enriched in FB2 stage, which is in agreement with GO enrichment analysis. In FB3 and FB4 stages, homogalacturonan and sugar degradation were most enriched. At FL1 stage, pathways involved in secondary metabolite synthesis were represented. The pathways related to hormone biosynthesis were enriched in FL2 stage.

Differential expression of transcription factor genes

Transcription factors (TFs) are the key regulatory proteins, which mediate the transcriptional regulation. TFs are known to play a major role in flower development. In fact, most of key regulatory genes implicated in floral development encode for TFs. Therefore, we studied the expression dynamics of TF genes in chickpea. In total, 111 TF genes exhibited significant differential expression in flower bud and flower development stages. Many of these TFs genes were common in flower bud and flower stages. A heatmap depicting the differential expression profiles of the TF genes during flower development is presented in Figure 5. Of the 84 TF families identified previously in chickpea (Garg *et al.*, 2011), at least one member of 31 families exhibited differential expression (Figure 5). Among the various families, MADS-box family TFs were most represented. Remarkably, a total of 18 MADS-box genes showed differential expression (17 up-regulated and one down-regulated) during flower development. The MYB/MYB-related TF family members were second highest (12 up-regulated and three down-regulated) represented in these genes. Other TF families enriched in the up-regulated genes were, NAC, bZIP, bHLH, LOB, SBP and Aux/IAA. Among the down-regulated genes, AP2-EREBP family genes were most (8) represented followed by WRKY (7) family genes.

Differential expression of lineage-specific genes

Lineage-specific genes represent the genes present only in a group of related species without any significant similarity to other groups. The study of expression patterns of these genes provides insights in their putative function. Lineage-specific genes have been identified in a few plant species, including chickpea and characterized to some extent (Campbell *et al.*, 2007; Garg and Jain, 2011; Garg *et al.*, 2011; Lin *et al.*, 2010). However, the knowledge about their functions in specific

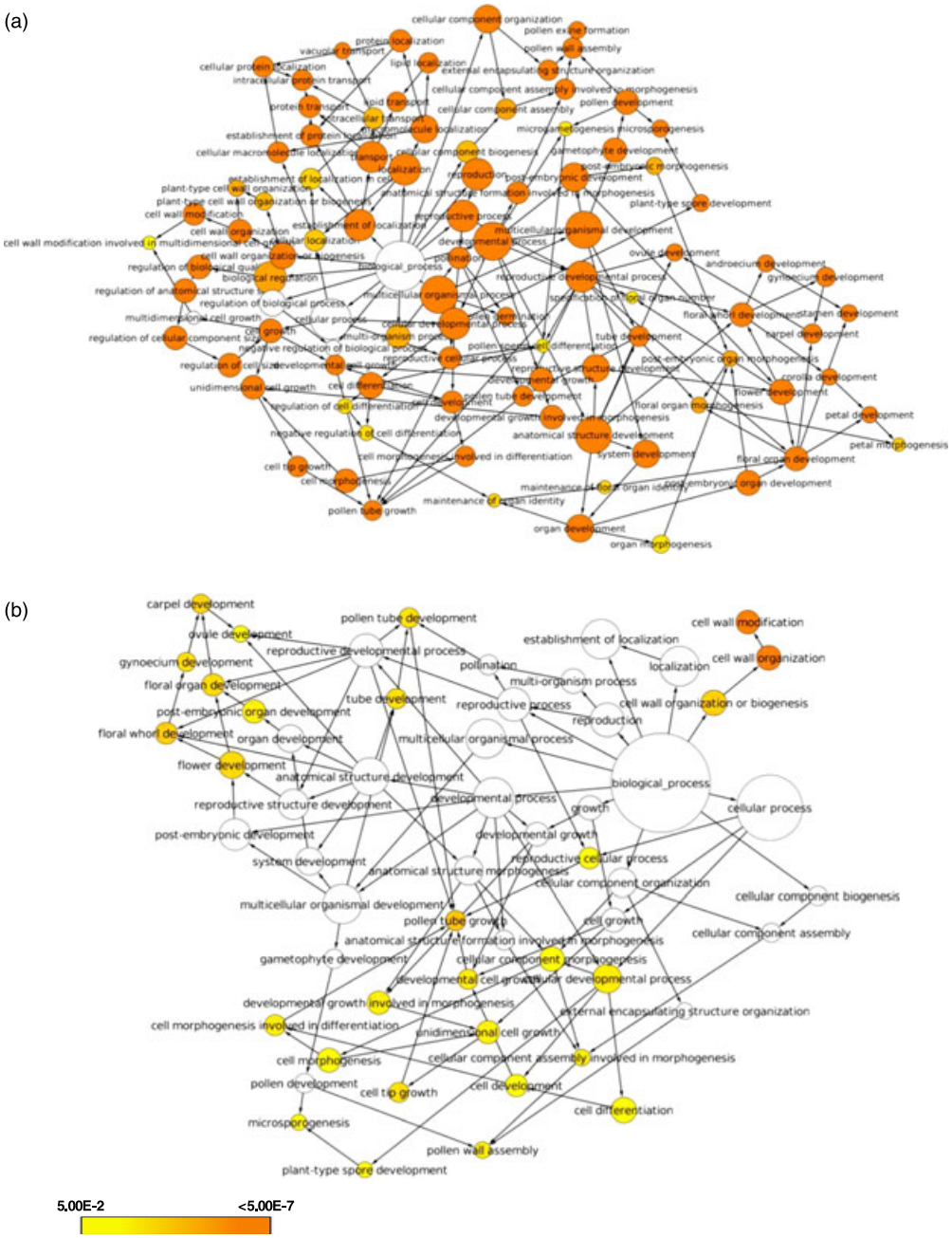

Figure 3 Gene ontology enrichment during flower development. Significantly, enriched GO categories in the up-regulated chickpea genes in (a) flower bud and (b) flower development stages. The genes showing up-regulation in flower bud and flower development stages were analysed using BiNGO and the biological process terms showing significant enrichment are shown. Node size is proportional to the number of transcripts in each category and colours shaded according to the significance level (white—no significant difference; colour scale, yellow—*P*-value = 0.05, orange—*P*-value <0.0000005).

biological processes is still very limited. To gain insights into the putative role of lineage-specific genes in chickpea, we studied their expression profiles during flower development. The analysis identified a significant fraction (38; 5.1%) of legume-specific genes exhibiting differential expression during flower development (Figure S4a, Table S1). Likewise, 96 (2.6%) chickpea-specific genes were also differentially expressed in flower development stages (Figure S4b, Table S1). Most of these genes are uncharacterized and have been annotated as expressed protein without any significant similarity to known plant proteins/domains. These results provide a clue towards the potential function of lineage-specific genes and it will be very interesting to elucidate their role in flower development.

Validation of differential gene expression results

To validate the results of differential gene expression analysis obtained using RNA-seq data, we followed two approaches. In the first approach, we performed quantitative RT-PCR analysis for at least 25 randomly selected genes, which showed differential expression during flower development, in all the tissues analysed. The outcome of this analysis was in very good agreement (correlation coefficient of 0.71) with RNA-seq results (Figure 6a).

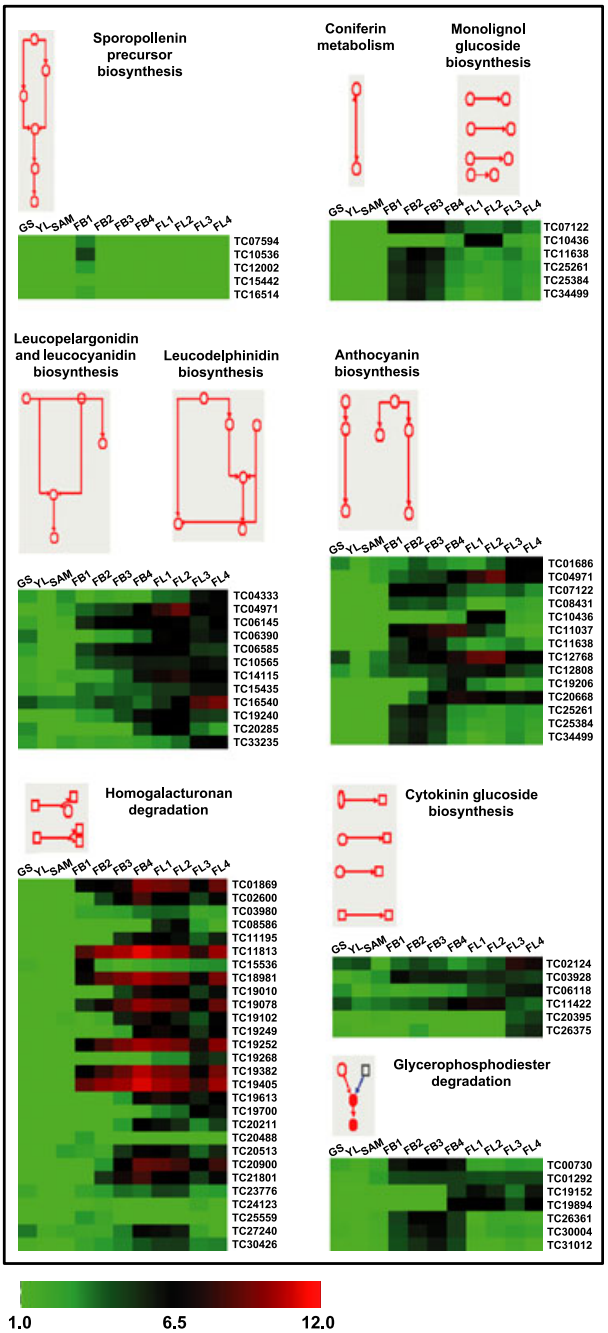

Figure 4 Regulation of metabolic pathways during flower development. The metabolic pathways enriched in differentially expressed genes during flower development are shown. Heatmaps showing the expression profiles of genes involved in these pathways are also shown. The colour scale at the bottom represents the log-transformed RPKM value.

All the genes exhibited similar expression patterns in quantitative RT-PCR analysis as observed from RNA-seq data (Figure S5).

As mentioned earlier, several genes involved in flower development have been identified in model plants, including class A, B, C, D and E genes of flower development model and many other genes (Andrés and Coupland, 2012; Jack, 2004; Krizek and Fletcher, 2005; Lohmann and Weigel, 2002). In the second approach, we identified the putative orthologues of class A [*APETALA1* (*AP1*) and *AP2*], B [*AP3* and *PISTILLATA* (*PI*)], C [*AGAMOUS* (*AG*)], D [*SEEDSTICK* (*STK*), *SHATTERPROOF1* (*SHP1*) and *SHP2*] and E [*SEPALLATA1* (*SEP1*), *SEP2*, *SEP3* and *SEP4*] genes and few other genes [*CRC* (*CRABS CLAW*), *FRUITFUL* (*FUL*), *LEAFY* (*LFY*), *SUPPRESSOR OF CONSTANS1* (*SOC1*), *SUPERMAN* (*SUP*), *SHORT VEGETATIVE PHASE1* (*SVP1*), *TERMINAL FLOWER1* (*TFL1*), *WUSCHEL* (*WUS*)] of flower development model (Figure 6b) in chickpea by BLAST searches (*E*-value <1e-50) and analysed their expression patterns in various tissues. Most of these genes exhibited higher expression in one or more flower development tissues as expected (Figure 6c). For example, the expression of *AP1* and *AP2* orthologues was significantly higher in most of flower tissues as compared to vegetative tissues. The expression of *PI* was most predominant during initial stages of flower development. *STK* was expressed more predominantly during later stages of flower development. *CRC* orthologue was expressed preferentially during all flower bud stages. The expression of *WUS* orthologue was restricted to FB1 stage in chickpea. The *LFY* orthologue was expressed in SAM and narrow window of early stages of flower development. The expression profiles of these genes further validate our differential gene expression results and suggest the involvement of similar genes in flower development in legumes as well.

Discussion

The temporal regulation of gene expression plays an important role in plant growth and development. Detailed information about gene expression is crucial for understanding the molecular mechanisms underlying any developmental process. Flower development, a key feature of higher plants, represents the reproductive phase of plant development. A few genome-wide studies have been performed to dissect the transcriptional programmes operative during reproductive development in model plants (Fujita *et al.*, 2010; Hennig *et al.*, 2004; Laitinen *et al.*, 2005; Sharma *et al.*, 2012; Wang *et al.*, 2010; Wellmer *et al.*, 2006). Fabaceae is a large family comprised of three subfamilies, Caesalpinioideae, Mimosoideae and Papilionoideae. These subfamilies differ substantially in floral symmetry (Tucker, 2003). Chickpea belongs to Papilionoideae subfamily of plants, in which inflorescence is usually raceme. A gene expression atlas covering various tissues for model legumes, soybean and *Medicago* has also been reported (Benedito *et al.*, 2008; Libault *et al.*, 2010; Severin *et al.*, 2010). Based on similarity search, most of Arabidopsis genes involved in flower development were found to be conserved in legumes, including soybean, *Medicago* and *Lotus* (Hecht *et al.*, 2005). Recently, a comparative genomic analysis identified key flowering genes in soybean (Jung *et al.*, 2012). However, a genome-wide comprehensive analysis of gene expression during flower development is lacking in legumes. Here, we describe a global view of gene expression dynamics during flower development in chickpea, an important legume crop plant.

We used high-throughput Illumina technology to generate a compendium of gene expression of various chickpea tissues, including eight stages of flower development and three vegetative tissues. We detected the expression of more than 96% genes in the tissue samples analysed, indicating the power and sensitivity of RNA-seq technology. In the PCA analysis, we found that the tissues analysed in this study, based on the transcriptional activity of chickpea genes, fell into three groups, vegetative, flower bud and flower stages. The tissues included within these groups represent closely related plant structures. The gene expression profiles of all the chickpea genes during each flower development stage was compared with vegetative tissues and

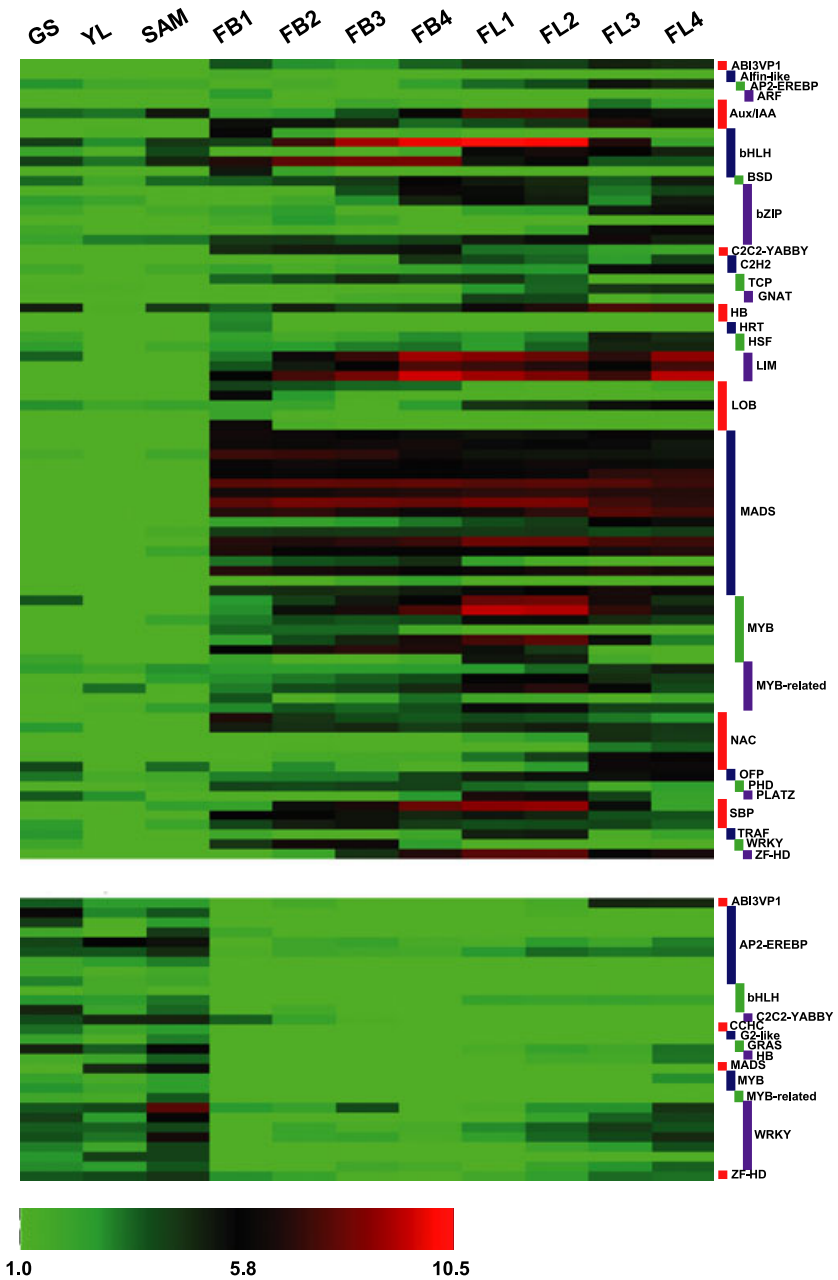

Figure 5 Differential expression of transcription factor genes during flower development. Various transcription factors families showing differential expression (up- and down-regulated in upper and lower panel, respectively) is shown on the right side. The colour scale at the bottom represents the log-transformed RPKM value.

preceding stage of development to study the stage-specific regulation of gene expression and identify genes involved in flower development. The homologues of many of the stage-specific gene identified in our study are well known to regulate aspects of flower development in plants.

Several molecular genetic studies have demonstrated the crucial role of transcription factors in reproductive development of plants. We also found the biological process, regulation of transcription, significantly enriched in preferentially expressed genes in chickpea flower stages. Among the various transcription factor families, MADS-box, MYB, NAC, bZIP and bHLH were found to be particularly important during flower development in chickpea. The members of these families have been reported to be involved in reproductive development in other plant species too (Hennig *et al.*, 2004; Sharma *et al.*, 2012). The crucial role of MADS-box TFs in orchestrating floral organ specification and development is well demonstrated in several studies (Kater *et al.*, 2006; Ng and Yanofsky, 2001; Urbanus *et al.*, 2010). The orthologues of well-known MADS-box genes (e.g. *AP1*, *AP3*, *PI* and *AG*) implicated in flower development exhibited preferential expression in chickpea flowers. The conservation of these genes in chickpea (this study) and other legumes (Hecht *et al.*, 2005; Jung *et al.*, 2012) together with their expression profiles suggests similar transcription regulatory network operative during flower development in plants. The significant enrichment of other transcription factor families highlights their important roles during flower development too.

As expected, we observed the abundance of genes involved in various processes related to flower development, such as maintenance of floral organ identity, development of corolla, petal, androecium and gynoecium and gametophyte development, during chickpea flower development. In addition, we found

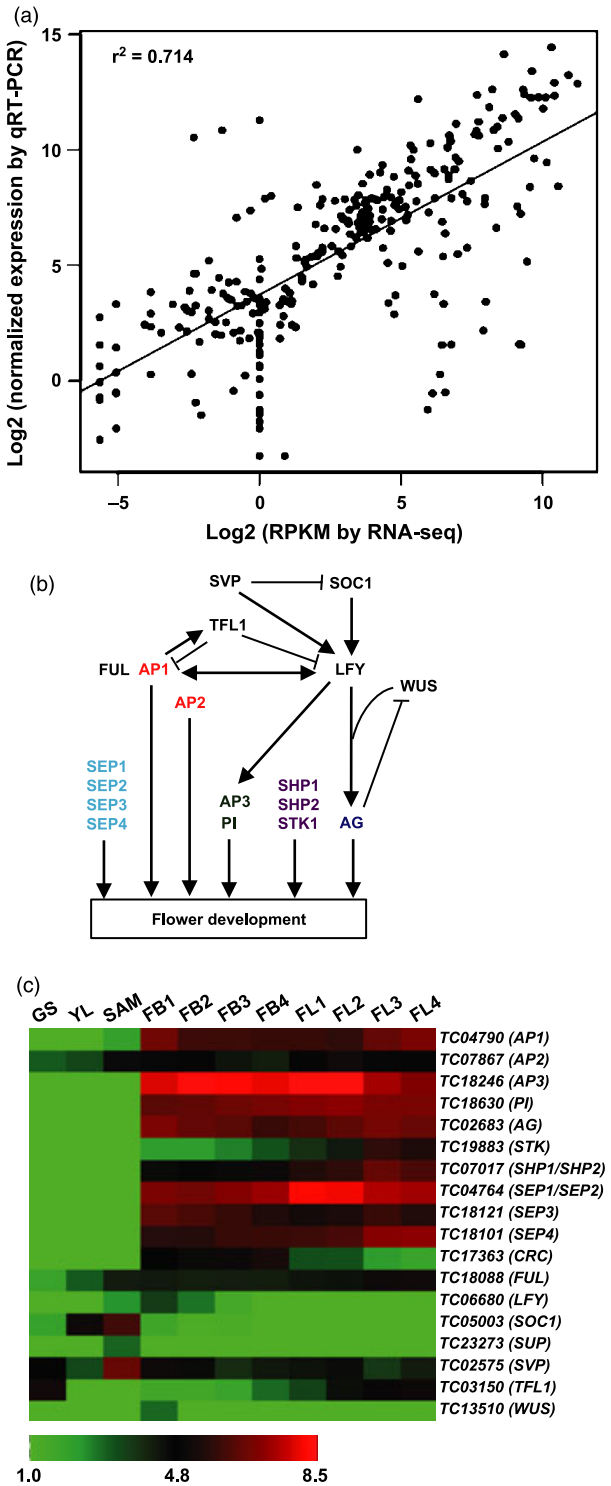

Figure 6 Validation of differential expression results obtained by RNA-seq. (a) Correlation of gene expression results obtained from real-time PCR analysis and RNA-seq for 25 selected genes in 11 tissue samples. (b) An outline of flower development pathway depicting genes for which orthologues have been identified in chickpea. Class A (*AP1* and *AP2*), B (*AP3* and *PI*), C (*AG*), D (*SHP1*, *SHP2* and *STK*) and E (*SEP1*, *SEP2*, *SEP3* and *SEP4*) genes have been indicated in different colours. (c) Heatmap showing differential expression of orthologues of known genes implicated in flower development. The colour scale at the bottom represents the log-transformed RPKM value.

enrichment of fatty acid/lipid transport and defence response–related genes enriched in few stages. The abundance of pathogenesis-related genes during stamen development and lipid transfer proteins (LTPs) during late petal development has been reported in *Gerbera hybrida* of Asteraceae family (Laitinen *et al.*, 2005). LTPs have been implicated in defence against pathogens and environmental changes (Kader, 1997; Maldonado *et al.*, 2002). These results suggest that floral organs like petals and stamens may have a role in defence mechanism against pathogens. We found the genes implicated in jasmonic acid, phenylpropanoid, cytokinin glucoside biosynthesis and ethylene biosynthesis to be up-regulated during flower development. A few previous studies have also documented the involvement of these pathways in various aspects of flower development, such as floral induction, bolting, pollen development, etc. (Hu *et al.*, 2008; Ma, 2005; Sharma *et al.*, 2012; Wilson and Zhang, 2009). The genes regulating other metabolic pathways, including hormone biosynthesis, sugar degradation, flavonoid biosynthesis, anthocyanin biosynthesis and their intermediates also exhibited higher expression during flower development stages, suggesting their involvement in some developmental events occurring in various stages. In a recent study, many of these metabolic pathways have been found to be enriched during panicle development in rice as well (Sharma *et al.*, 2012). The pathways involved in biosynthesis of secondary metabolites in floral organs have been observed in many plant species (Hennig *et al.*, 2004; Laitinen *et al.*, 2005; Sharma *et al.*, 2012). The synthesis of flavonoids in several floral organs is common, which is required for pollen function (Xie *et al.*, 2003). Likewise, synthesis of anthocyanins in pistils after pollination has also been reported. These observations provide a molecular basis of earlier results showing the significance of flavonoid and anthocyanin biosynthesis during flower development. The pathway responsible for biosynthesis of sporopollenin polymer was also found enriched, which constitutes the outer layer of spores and pollen wall (Dobritsa *et al.*, 2009). The enrichment of HGA degradation pathway was also found in identified flower development genes. HGA accounts for a significant fraction of the pectin found in plant primary cell walls and its degradation is required during anther dehiscence and pollen grain maturation with the help of methylesterase and polygalacturonase enzymes (Francis *et al.*, 2006; Rhee *et al.*, 2003). The representation of sugar biosynthesis and degradation pathways suggests their role in flower development. Previously, the role of sugar in floral organ has also been documented (van Doorn, 2004; Eveland and Jackson, 2012). Overall, it will be interesting to dissect the precise role of genes involved in various biological processes and metabolic pathways enriched during flower development.

Although the exact series of events occurring during flower development in legumes has not been studied so far, a few reports have described these processes to some extent (Tucker, 1987, 2003). It has been found that flower development in legumes is quite different than model plants like Arabidopsis, which follow ABC model of development (Tucker, 2003). In Arabidopsis, the development of floral organs is sequential without any overlap in the timing of initiation and a set of genes determine the initiation of each organ. However, in Papilionoideae subfamily legumes, the simultaneous initiation of more than one type of floral organ has been reported. It has been shown that a common primordium initiates two or three organ primordia (Tucker, 1987, 2003), which suggests that similar set of genes controlling the development of multiple floral organs. We

also observed quite a large number of preferentially expressed genes overlapping among different stages of flower development, which might be responsible for initiation and development of different floral whorls simultaneously. In addition, the genes uniquely regulated in a particular developmental stage might control specific developmental events. These speculations need to be substantiated by further experimentation. We observed the preferential/specific expression of many lineage-specific genes during flower development as well, which might be responsible, at least in part, for the distinct floral symmetry in legumes. In addition, although all the transcription factors families and distribution of genes in them were found to be conserved in legume and non-legume plants (Garg *et al.*, 2011; Libault *et al.*, 2009), diversity in their gene expression patterns might impart phenotypic specificities to some extent in legumes.

To our knowledge, this study is the first report showing genome-wide transcriptome dynamics during flower development in legumes. This analysis has revealed a large set of candidate genes, which apparently play important roles in flower development in legumes. The detailed investigations of the novel pathways and candidate genes identified in this study would be very useful to understand the molecular mechanism involved in flower development. The identification of genes involved in reproductive development has economic importance, which provides potential targets for crop improvement via breeding and reverse genetics approaches.

Experimental procedures

Plant material collection and RNA extraction

We collected 11 tissue samples, including three vegetative tissues (GS, YL and SAM) and eight stages of flower development (flower buds and flowers) from chickpea (*Cicer arietinum* genotype ICC4958) plants. At least three biological replicates were collected for each tissue sample. All the tissue samples except GS and SAM were collected from the field-grown plants. For GS and SAM tissues, chickpea seedlings were grown in a culture room as described (Garg *et al.*, 2010). Germinating seedlings were collected after 5-days of seed germination on wet Whatman paper sheets in Petri dishes. Young leaves were collected from the mature plants surrounding the SAM. SAM was dissected from the 21-day-old young plants under a stereo zoom microscope. Four stages each of flowed bud (FB1–FB4) and flower (FL1–FL4) development were collected. The four stages of flower buds were collected based on their size (4 mm, FB1; 6 mm, FB2; 8 mm, FB3 and 8–10 mm, FB4). The four stages of flower development included young flower with closed petals (FL1), mature flower with partially opened petals (FL2), mature flower with opened and faded petals (FL3) and drooped flower with senescing petals (FL4).

Total RNA was extracted using TRI reagent (Sigma Life Science, St. Louis, MO) following the manufacturer's protocol. About 100–200 mg of tissue was used for total RNA isolation. RNA quality and quantity was determined using Nanodrop 1000 spectrophotometer (Thermo Fisher Scientific, Wilmington, DE) and Bioanalyzer RNA nano chip (Agilent Technologies, Singapore). Only the RNA samples with 260/280 ratio between 1.8 to 2.1, 260/230 ratio between 2.0 to 2.5 and RIN (RNA integrity number) more than 8.0, were used for sequencing.

Illumina sequencing and data preprocessing

Approximately, 5 μg of total RNA (pooled in equal amount from three biological replicates) for each tissue sample was used for the construction of libraries using mRNA-Seq Sample Prep kit (Illumina Inc., San Diego, CA) according to the manufacturer's protocol. Single-end sequencing of each library was performed using Genome Analyzer IIx (Illumina Technologies) platform following the manufacturer's instructions. After base calling via Illumina pipeline, the filtered data (about 18–42 million reads) for all the samples were obtained in the Fastq file format. These Fastq files were further subjected to stringent sequence quality control, using NGS QC Toolkit (v2.2.3) with default parameters (Patel and Jain, 2012). This led to the removal of 4–15% of the reads (low-quality reads and reads containing primer/adaptor sequences) from different tissue samples. The remaining high-quality filtered reads were used for data analysis. Sequence data generated in this study have been deposited in the Gene Expression Omnibus (GEO) database under the series accession number GSE42679.

Differential gene expression and tissue specificity analysis

To estimate the expression levels, the mapping of high-quality reads from each sample was performed on the chickpea transcriptome assembly (34760 transcripts; Garg *et al.*, 2011) using CLC Genomics Workbench (v4.7.2). A maximum of two mismatches were allowed for the alignments. The high-quality 54 bp long (except for FB1, for which read length was 100 bp) reads were aligned to the chickpea transcripts in the first round, which resulted in mapping of more than 88% reads from each tissue except for YL (66%). In the second round, we trimmed the remaining unmapped reads (14 bases from the 3′-end were trimmed after optimization of trimming one base recursively) and aligned again to the transcriptome sequence. This increased the total mapping percentage of reads to 90–93% for all the samples. Read counts were normalized by calculating the read per kilobase per million (RPKM) value for each transcript in different tissue samples. The genes represented by at least one RPKM value were considered to be expressed in a particular tissue. PCA analysis was performed using MultiExperiment Viewer (MeV, v4.6.2). The differential gene expression analysis was performed using DESeq software (Anders and Huber, 2010) with *p*-value cut-off of <0.05 after adjustment for multiple testing with the Benjamini Hochberg method. The heatmaps showing expression profiles were generated based on the $\log_2$-transformed RPKM values using MeV.

Quantitative PCR analysis

To validate the expression profiles of selected genes, the gene-specific primers were designed using Primer Express (v3.0) software (Applied Biosystems, Foster City, CA) and real-time PCRs performed using the 7500 Sequence Detection System (Applied Biosystems) as described previously (Garg *et al.*, 2010). At least two independent biological replicates and three technical replicates of each biological replicates for each tissue sample were analysed in real-time PCR analysis. The primers sequences used for real-time PCR analysis in this study are given in the Table S3. The correlation between expression profiles of 25 selected genes measured by qRT-PCR and RNA-seq was determined using R package.

GO enrichment and pathway analysis

For GO enrichment and pathway analysis, the best Arabidopsis hit corresponding to each chickpea transcript was identified using BLAST search. GO enrichment of various sets of genes was performed using BiNGO tool (Maere *et al.*, 2005). The metabolic

network data available in AraCyc database of Gramene (Liang *et al.*, 2008) was analysed to identify the enriched metabolic pathways in different gene sets. The enrichment analyses were performed with *P*-value cut-off of <0.05 after applying Benjamini Hochberg correction.

Acknowledgements

This work was 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). VKS acknowledges the award of research fellowship from the Department of Biotechnology. RG acknowledges INSPIRE Faculty Award from the Department of Science and Technology, Government of India.

References

Agarwal, G., Jhanwar, S., Priya, P., Singh, V.K., Saxena, M.S., Parida, S.K., Garg, R., Tyagi, A.K. and Jain, M. (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.

Anders, S. and Huber, W. (2010) Differential expression analysis for sequence count data. *Genome Biol.* **11**, R106.

Andrés, F. and Coupland, G. (2012) The genetic basis of flowering responses to seasonal cues. *Nat. Rev. Genet.* **13**, 627–639.

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

Campbell, M.A., Zhu, W., Jiang, N., Lin, H., Ouyang, S., Childs, K.L., Haas, B.J., Hamilton, J.P. and Buell, C.R. (2007) Identification and characterization of lineage-specific genes within the Poaceae. *Plant Physiol.* **145**, 1311–1322.

Dobritsa, A.A., Shrestha, J., Morant, M., Pinot, F., Matsuno, M., Swanson, R., Møller, B.L. and Preuss, D. (2009) CYP704B1 is a long-chain fatty acid omega-hydroxylase essential for sporopollenin synthesis in pollen of Arabidopsis. *Plant Physiol.* **151**, 574–589.

van Doorn, W.G. (2004) Is petal senescence due to sugar starvation? *Plant Physiol.* **134**, 35–42.

Eveland, A.L. and Jackson, D.P. (2012) Sugars, signalling, and plant development. *J. Exp. Bot.* **63**, 3367–3377.

Francis, K.E., Lam, S.Y. and Copenhaver, G.P. (2006) Separation of Arabidopsis pollen tetrads is regulated by *QUARTET1*, a pectin methylesterase gene. *Plant Physiol.* **142**, 1004–1013.

Fujita, M., Horiuchi, Y., Ueda, Y., Mizuta, Y., Kubo, T., Yano, K., Yamaki, S., Tsuda, K., Nagata, T., Niihama, M., Kato, H., Kikuchi, S., Hamada, K., Mochizuki, T., Ishimizu, T., Iwai, H., Tsutsumi, N. and Kurata, N. (2010) Rice expression atlas in reproductive development. *Plant Cell Physiol.* **51**, 2060–2081.

Garg, R. and Jain, M. (2011) Pyrosequencing data reveals tissue-specific expression of lineage-specific transcripts in chickpea. *Plant Signal. Behav.* **6**, 1868–1870.

Garg, R., Sahoo, A., Tyagi, A.K. and Jain, M. (2010) Validation of internal control genes for quantitative gene expression studies in chickpea (*Cicer arietinum* L.). *Biochem. Biophys. Res. Commun.* **396**, 283–288.

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

Gonzalez-Ballester, D., Casero, D., Cokus, S., Pellegrini, M., Merchant, S.S. and Grossman, A.R. (2010) RNA-seq analysis of sulfur-deprived *Chlamydomonas* cells reveals aspects of acclimation critical for cell survival. *Plant Cell* **2**, 2058–2084.

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

Hennig, L., Gruissem, W., Grossniklaus, U. and Köhler, C. (2004) Transcriptional programs of early reproductive stages in Arabidopsis. *Plant Physiol.* **135**, 1765–1775.

Hu, J., Mitchum, M.G., Barnaby, N., Ayele, B.T., Ogawa, M., Nam, E., Lai, W.C., Hanada, A., Alonso, J.M., Ecker, J.R., Swain, S.M., Yamaguchi, S., Kamiya, Y. and Sun, T.P. (2008) Potential sites of bioactive gibberellin production during reproductive growth in Arabidopsis. *Plant Cell* **20**, 320–336.

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

Jain, M. (2012) Next-generation sequencing technologies for gene expression profiling in plants. *Brief. Funct. Genomics* **11**, 63–70.

Jhanwar, S., Priya, P., Garg, R., Parida, S.K., Tyagi, A.K. and Jain, M. (2012) Transcriptome sequencing of wild chickpea as a rich resource for marker development. *Plant Biotechnol. J.* **10**, 690–702.

Jung, C.H., Wong, C.E., Singh, M.B. and Bhalla, P.L. (2012) Comparative genomic analysis of soybean flowering genes. *PLoS One* **7**, e38250.

Kader, J.-C. (1997) Lipis transfer proteins: a puzzling family of plant proteins. *Trends Plant Sci.* **2**, 66–70.

Kater, M.M., Dreni, L. and Colombo, L. (2006) Functional conservation of MADS-box factors controlling floral organ identity in rice and Arabidopsis. *J. Exp. Bot.* **57**, 3433–3444.

Krizek, B.A. and Fletcher, J.C. (2005) Molecular mechanisms of flower development: an armchair guide. *Nat. Rev. Genet.* **6**, 688–698.

Laitinen, R.A., Immanen, J., Auvinen, P., Rudd, S., Alatalo, E., Paulin, L., Ainasoja, M., Kotilainen, M., Koskela, S., Teeri, T.H. and Elomaa, P. (2005) Analysis of the floral transcriptome uncovers new regulators of organ determination and gene families related to flower organ differentiation in *Gerbera hybrida* (Asteraceae). *Genome Res.* **15**, 475–486.

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

Liang, C., Jaiswal, P., Hebbard, C., Avraham, S., Buckler, E.S., Casstevens, T., Hurwitz, B., McCouch, S., Ni, J., Pujar, A., Ravenscroft, D., Ren, L., Spooner, W., Tecle, I., Thomason, J., Tung, C.W., Wei, X., Yap, I., Youens-Clark, K., Ware, D. and Stein, L. (2008) Gramene: a growing plant comparative genomics resource. *Nucl. Acids Res.* **36**(Database issue), D947–D953.

Libault, M., Joshi, T., Benedito, V.A., Xu, D., Udvardi, M.K. and Stacey, G. (2009) Legume transcription factor genes: what makes legumes so special? *Plant Physiol.* **151**, 991–1001.

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

Lin, H., Moghe, G., Ouyang, S., Iezzoni, A., Shiu, S.H., Gu, X. and Buell, C.R. (2010) Comparative analyses reveal distinct sets of lineage-specific genes within *Arabidopsis thaliana*. *BMC Evol. Biol.* **10**, 41.

Lohmann, J.U. and Weigel, D. (2002) Building beauty: the genetic control of floral patterning. *Dev. Cell* **2**, 135–142.

Ma, H. (2005) Molecular genetic analyses of microsporogenesis and microgametogenesis in flowering plants. *Annu. Rev. Plant Biol.* **56**, 393–434.

Maere, S., Heymans, K. and Kuiper, M. (2005) BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. *Bioinformatics* **21**, 3448–3449.

Maldonado, A.M., Doerner, P., Dixon, R.A., Lamb, C.J. and Cameron, R.K. (2002) A putative lipid transfer protein involved in systemic resistance signalling in Arabidopsis. *Nature* **419**, 399–403.

Mortazavi, A., Williams, B.A., McCue, K., Schaeffer, L. and Wold, B. (2008) Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat. Methods* **5**, 621–628.

Ng, M. and Yanofsky, M.F. (2001) Function and evolution of the plant MADS-box gene family. *Nat. Rev. Genet.* **2**, 186–195.

Ozsolak, F. and Milos, P.M. (2011) RNA sequencing: advances, challenges and opportunities. *Nat. Rev. Genet.* **12**, 87–98.

Patel, R.K. and Jain, M. (2012) NGS QC Toolkit: a toolkit for quality control of next generation sequencing data. *PLoS One* **7**, e30619.

Rhee, S.Y., Osborne, E., Poindexter, P.D. and Somerville, C.R. (2003) Microspore separation in the *quartet 3* mutants of Arabidopsis is impaired by a defect in a developmentally regulated polygalacturonase required for pollen mother cell wall degradation. *Plant Physiol.* **133**, 1170–1180.

Schmid, M., Davison, T.S., Henz, S.R., Pape, U.J., Demar, M., Vingron, M., Schölkopf, B., Weigel, D. and Lohmann, J.U. (2005) A gene expression map of *Arabidopsis thaliana* development. *Nat. Genet.* **37**, 501–506.

Severin, A.J., Woody, J.L., Bolon, Y.T., Joseph, B., Diers, B.W., Farmer, A.D., Muehlbauer, G.J., Nelson, R.T., Grant, D., Specht, J.E., Graham, M.A., Cannon, S.B., May, G.D., Vance, C.P. and Shoemaker, R.C. (2010) RNA-Seq Atlas of *Glycine max*: a guide to the soybean transcriptome. *BMC Plant Biol.* **10**, 160.

Sharma, R., Agarwal, P., Ray, S., Deveshwar, P., Sharma, P., Sharma, N., Nijhawan, A., Jain, M., Singh, A.K., Singh, V.P., Khurana, J.P., Tyagi, A.K. and Kapoor, S. (2012) Expression dynamics of metabolic and regulatory components across stages of panicle and seed development in indica rice. *Funct. Integr. Genomics* **12**, 229–248.

Taylor, S.A., Hofer, J.M., Murfet, I.C., Sollinger, J.D., Singer, S.R., Knox, M.R. and Ellis, T.H. (2002) *PROLIFERATING INFLORESCENCE MERISTEM*, a MADS-box gene that regulates floral meristem identity in pea. *Plant Physiol.* **129**, 1150–1159.

Tucker, S.C. (1987) Floral initiation and development in legumes. In *Advances in Legume Systematics Part 3*. (Stirton, C.H., ed), pp. 183–239. Kew, UK: Royal Botanic Gardens.

Tucker, S.C. (2003) Floral development in legumes. *Plant Physiol.* **131**, 911–926.

Urbanus, S.L., Dinh, Q.D., Angenent, G.C. and Immink, R.G. (2010) Investigation of MADS domain transcription factor dynamics in the floral meristem. *Plant Signal Behav.* **5**, 1260–1262.

Wang, Z., Gerstein, M. and Snyder, M. (2009) RNA-Seq: a revolutionary tool for transcriptomics. *Nat. Rev. Genet.* **10**, 57–63.

Wang, L., Xie, W., Chen, Y., Tang, W., Yang, J., Ye, R., Liu, L., Lin, Y., Xu, C., Xiao, J. and Zhang, Q. (2010) A dynamic gene expression atlas covering the entire life cycle of rice. *Plant J.* **61**, 752–766.

Weber, A.P., Weber, K.L., Carr, K., Wilkerson, C. and Ohlrogge, J.B. (2007) Sampling the *Arabidopsis* transcriptome with massively parallel pyrosequencing. *Plant Physiol.* **144**, 32–42.

Wellmer, F., Alves-Ferreira, M., Dubois, A., Riechmann, J.L. and Meyerowitz, E.M. (2006) Genome-wide analysis of gene expression during early Arabidopsis flower development. *PLoS Genet.* **2**, e117.

Wilson, Z.A. and Zhang, D.B. (2009) From Arabidopsis to rice: pathways in pollen development. *J. Exp. Bot.* **60**, 1479–1492.

Xie, D.Y., Sharma, S.B., Paiva, N.L., Ferreira, D. and Dixon, R.A. (2003) Role of anthocyanidin reductase, encoded by *BANYULS* in plant flavonoid biosynthesis. *Science* **299**, 96–99.

Yang, H., Lu, P., Wang, Y. and Ma, H. (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.

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

Supporting information

Additional Supporting information may be found in the online version of this article:

Figure S1 Principle component analysis of various tissues analysed. Each point represents independent tissue sample.

Figure S2 Gene ontology enrichment during flower development. Top 20 (if available) significantly enriched GO (biological process, molecular function and cellular component) categories in the up-regulated chickpea genes in (a) flower bud and (b) flower stages are given in the bar graphs.

Figure S3 Gene ontology and metabolic pathway enrichment in various stages of flower development with respect to the preceding stage. Top five significantly enriched biological process GO terms (a) and metabolic pathways (b) in the up-regulated chickpea genes in various stages of flower development with respect to the preceding stage are given in the bar graphs.

Figure S4 Differential expression of lineage-specific genes. Heatmaps display the differential expression profiles of legume- (a) and chickpea-specific (b) genes in various tissues (up- and down-regulated in upper and lower panel, respectively). The colour scale at the bottom represents the log transformed RPKM value.

Figure S5 Quantitative RT-PCR validation of differential gene expression. Relative transcript abundance of representative genes validated by real-time PCR analysis is shown. The fold change in all tissues/stages for each gene was calculated with respect to GS sample. GS, germinating seedling; YL, young leaf; SAM, shoot apical meristem; FB1–FB4 flower bud stages; FL1–FL4 flower stages.

Table S1 List of genes differentially expressed during flower development. The gene ID, description, RPKM values in all the tissues analysed, differential expression in flower bud and flower stages, stage specificity, TF family and lineage-specificity information are given.

Table S2 List of up-regulated genes in various stages of flower development with respect to the preceding stage and no expression in vegetative tissues.

Table S3 Primer sequences used for real-time PCR analysis in this study.

