Received: 31 January 2018 | Revised: 24 March 2018 | Accepted: 26 March 2018

DOI: 10.1111/pce.13210

SPECIAL ISSUE

WILEY Plant, Cell & Environment

The RNA-Seq-based high resolution gene expression atlas of chickpea (*Cicer arietinum* L.) reveals dynamic spatio-temporal changes associated with growth and development

Himabindu Kudapa* | **Vanika Garg*** | **Annapurna Chitikineni** | **Rajeev K. Varshney**

Center of Excellence in Genomics & Systems Biology, International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru 502 324, India

Correspondence
R. K. Varshney, Center of Excellence in Genomics & Systems Biology, International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru, India.
Email: r.k.varshney@cgiar.org

Abstract

Chickpea is one of the world's largest cultivated food legumes and is an excellent source of high-quality protein to the human diet. Plant growth and development are controlled by programmed expression of a suite of genes at the given time, stage, and tissue. Understanding how the underlying genome sequence translates into specific plant phenotypes at key developmental stages, information on gene expression patterns is crucial. Here, we present a comprehensive *Cicer arietinum* Gene Expression Atlas (CaGEA) across different plant developmental stages and organs covering the entire life cycle of chickpea. One of the widely used drought tolerant cultivars, ICC 4958 has been used to generate RNA-Seq data from 27 samples at 5 major developmental stages of the plant. A total of 816 million raw reads were generated and of these, 794 million filtered reads after quality control (QC) were subjected to downstream analysis. A total of 15,947 unique number of differentially expressed genes across different pairwise tissue combinations were identified. Significant differences in gene expression patterns contributing in the process of flowering, nodulation, and seed and root development were inferred in this study. Furthermore, differentially expressed candidate genes from "*QTL-hotspot*" region associated with drought stress response in chickpea were validated.

KEYWORDS

differentially expressed genes, drought tolerance, floral development, gene atlas, nodulation, RNA-sequencing, seed and root development

1 | INTRODUCTION

Legumes are second to cereals in agricultural production accounting for approximately 27% of the world's primary production (Graham & Vance, 2003). Grain legumes alone contribute to about one third of the human dietary protein. In addition to their important contribution to human health, legumes also improve soil health by fixing atmospheric nitrogen. Chickpea (*Cicer arietinum* L.) is the second most important food legume crop, cultivated in arid and semi-arid regions of the world. It is a rich source of protein (20–30%), carbohydrates (40%), and minerals serving as an important source of nutrients in vegetarian diets of developing countries (Jukanti, Gaur, Gowda, & Chibbar, 2012). It is a self-pollinated and diploid crop (2n = 2× = 16) with a genome size of 740 Mbps (Arumuganathan & Earle, 1991).

*These authors contributed equally to this work.

During last five decades, the area under chickpea cultivation and the productivity continue to remain the same in spite of its important role in human health. Chickpea production is far below the present demand and has not achieved its potential yield owing to major constraints such as several abiotic (drought, heat, salinity, etc.) and biotic (ascochyta blight, fusarium wilt, etc.) stresses. Seed morphology in chickpea is of two types: *kabuli* (bold pale brown) and *desi* (with small dark brown) seeds. Recently reported draft genome assemblies of both the chickpea types, *kabuli* (CDC Frontier; Varshney et al., 2013) and *desi* (ICC 4958; Jain et al., 2013) provide a powerful resource for chickpea functional genomics research. Furthermore, genomic and transcriptomic resources were also rapidly developed in chickpea (Agarwal et al., 2016; Hiremath et al., 2011; Kudapa et al., 2014; Nayak et al., 2010; Thudi et al., 2011; Varshney et al., 2009). These resources have been widely used in chickpea research to investigate tolerance/resistance to abiotic (Kale et al., 2015; Pushpavalli et al., 2015; Thudi et al., 2014; Varshney et al., 2014) and biotic (Sabbavarapu et al., 2013) stresses. As part of the effort to better understand legume biology, several transcriptomic studies have been carried out on different tissues under a variety of experimental/stress conditions in chickpea and identified hundreds of differentially expressed genes (Garg, Patel, Tyagi, & Jain, 2011; Hiremath et al., 2011; Kaashyap et al., 2018; Kudapa et al., 2014; Varshney et al., 2009).

Understanding of gene expression has changed dramatically with the evolution of new technologies. High-throughput technologies such as microarrays and next-generation sequencing have been comprehensively applied to generate large amounts of genome-wide gene expression data and understand key genetic bottlenecks that limit chickpea yield under different stress conditions (Garg et al., 2016; Hiremath et al., 2011; Kudapa et al., 2014; Varshney et al., 2009). However, no single study has yet brought together genome-wide data on gene expression profiles for all major plant organs at different developmental stages of chickpea, leaving a large gap between the genome sequence and phenotype. To understand how the underlying genome sequence results in specific plant phenotypes, information on gene expression patterns is crucial. Gene expression atlases could predict cluster of genes expressed in each tissue at different plant developmental stages. Technologies such as Affymetrix GeneChip (Benedito et al., 2008) and RNA-Seq (O'Rourke et al., 2014; Severin et al., 2010) have been employed in investigating gene expression atlas in several crop species. In legumes, gene expression atlas has been developed for model legumes, *Medicago* (Benedito et al., 2008) and *Lotus* (Verdier et al., 2013), and recently in crop legumes where genome sequence is available, for example, in soybean (Libault et al., 2010; Severin et al., 2010), common bean (O'Rourke et al., 2014), pea (Alves-Carvalho et al., 2015), cowpea (Yao et al., 2016), peanut (Clevenger, Chu, Scheffler, & Ozias-Akins, 2016), and pigeonpea (Pazhamala et al., 2017). Expression profiling and systematic analysis of annotated patterns of gene expression will complement genomic information and help predict a gene's function.

In the present study, *C. arietinum* (chickpea) Gene Expression Atlas (CaGEA) has been developed from 27 chickpea tissues across five developmental stages, namely, germination, seedling, vegetative, reproductive, and senescence, of a chickpea breeding cultivar, ICC 4958. The genotype, ICC 4958, used in the present study is one of the most drought tolerant and widely used cultivar in chickpea breeding programme. The atlas data generated utilizing RNA-Seq technology were analysed to disclose expression profiles of all expressed genes in different tissues/stages of chickpea. Differential gene expressions that are accompanied by changes in the expression of key regulatory genes such as transcription factors were examined. The gene expression atlas of chickpea would not only provide a global view of gene expression patterns in all major tissues and organ systems but also will serve as valuable resource for functional genomics and accelerate gene discovery in legumes.

2 | MATERIALS AND METHODS

2.1 | Plant growth and sample collection

Drought tolerant chickpea breeding cultivar "ICC 4958" was used in the present study. Plant tissue samples were collected at five major developmental stages of the plant, namely, germination and seedling, vegetative, reproductive, and senescence. These samples were grouped into four sets for further analysis based on their developmental stages. Tissues from germination (three tissues) and seedling (two tissues) stages were grouped under Set I (total of five tissues); vegetative to Set II (four tissues); reproductive to Set III (nine tissues); and senescence to Set IV (nine tissues; Figure 1). For germination stage, the surface sterilized seeds were soaked in distilled water for 6 hr and were allowed to germinate in petriplates using Whatman No. 1 filter paper soaked with 10 ml of distilled water for 24 to 48 hr. Three petriplates with 10 seeds per replicate were used for each sampling. For seedling stage, the seeds were sown in paper cups and allowed to grow for 8–10 days after germination (DAG) under glasshouse conditions. Seeds were sown in pots and maintained in glasshouse for vegetative (25–30 DAG), reproductive (40–50 DAG), and senescence (90–110 DAG) stages. Three biological replicates were maintained at each stage for sample/tissue collection. Biological replicates were constituted by at least three plants sown at different intervals and located randomly in the glasshouse. A total of 27 samples (Table 1) from all stages (Set I–IV) were targeted for collection in three biological replications planted on separate days. After harvest, tissues were washed thoroughly with 0.1% DEPC water, flash frozen in liquid nitrogen, and stored at −80°C for RNA extraction.

2.2 | RNA extraction and cDNA library construction

The total RNA was extracted from the harvested tissues using Nucleospin RNA plant kit (Macherey-Nagel, Duren, Germany) as per manufacturer's protocol (http://www.macherey-nagel.com/). RNA was purified using the RNeasy MinElute cleanup kit (Qiagen, www.Qiagen.com) according to manufacturer's instructions. Purified RNA was quantified using a NanoDrop ND-100 spectrophotometer (Thermo Scientific), and its integrity was evaluated, using Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA). The assessment of RNA integrity is a critical first step in obtaining meaningful gene expression data. Equal amounts of RNA samples from three independent biological replicates (of each tissue) having RNA Integrity Number (algorithm for assigning integrity values to RNA quality

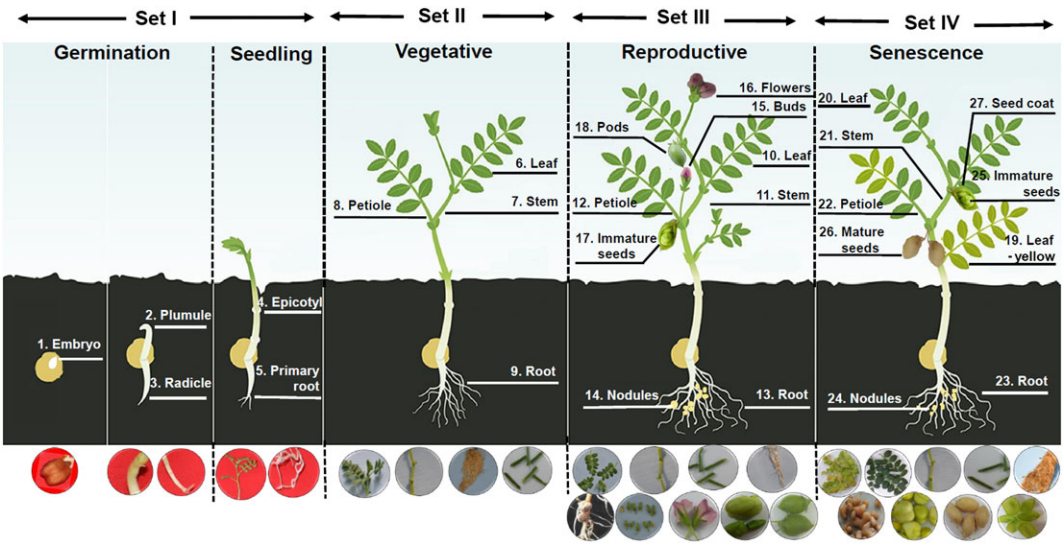

FIGURE 1 Details of samples of ICC 4958 genotype used for gene expression atlas. These 27 samples represent tissues from five developmental stages, that is, germination, seedling, vegetative, reproductive, and senescence stages of the entire life cycle of chickpea [Colour figure can be viewed at wileyonlinelibrary.com]

measurement from Agilent Technologies) value ≥8 were pooled prior to library preparation and subsequent sequencing. The cDNA libraries were prepared using Illumina TruSeq RNA library protocol outlined in "TruSeq RNA Sample Preparation Guide" (Part no. 15008136).

2.3 | Illumina sequencing and data preprocessing

Paired-end sequencing was performed in two sets: a set of 22 samples (nos.1 to 22) were sequenced using Illumina HiSeq 2000 at Genotypic Technology Pvt. Ltd., India, and a second set of 5 samples (nos. 23 to 27) were sequenced in-house (Sequencing and Informatics Services Unit, ICRISAT) using Illumina HiSeq 2500 sequencing system. The raw reads generated from all 27 samples were subjected to quality filtering using NGS-QCBOX (Katta, Khan, Doddamani, Thudi, & Varshney, 2015) and Trimmomatic v0.35 (Bolger, Lohse, & Usadel, 2014). The low quality reads (Phred score < 20; read length < 50 bases) and reads with adapter contamination were removed to generate a set of high quality reads termed as clean data thereafter. The complete downstream analysis was based on clean data.

2.4 | Global gene expression analysis, clustering, and visualization

The RNA-Seq data were analysed using Tuxedo pipeline (Trapnell et al., 2012), an open source pipeline. The filtered reads from all samples were aligned on the chickpea genome (Varshney et al., 2013) using Tophat v2.1.0 (Kim et al., 2013) with default parameters. The resulting alignment reads from each sample were then used to create a RABT (Reference Annotation Based Transcript) assembly using Cufflinks (v 2.1.1) (Trapnell et al., 2010). The assemblies generated were then merged into one consensus assembly using cuffmerge which was further used to quantify transcript abundance. Transcript abundance was estimated based on fragments per kilobase of transcript per million mapped reads (FPKM). Transcripts with abundance >1 FPKM and quantification status as "OK" were considered for further analysis. Identification of tissue-specific genes was performed on genes with FPKM ≥1 in at least one of 27 samples, by calculating tissue-specificity index (τ; Yanai et al., 2005) using the equation:

$$\tau = \frac{\sum_{i=1}^{N}(1-x_i)}{N-1},$$

where N is the number of samples and x_i is the expression value of a gene normalized by maximum value across all samples. The value of τ range from 0 to 1, where higher the value more likely the gene is specifically expressed in that stage. For this study, genes with τ = 1 were considered tissue-specific.

For the analysis of the global gene expression patterns, hierarchical clustering was plotted using the "pheatmap" function by R software. Genes with FPKM >1 were log2 + 1 transformed and hierarchical clustering was performed. Samples were further clustered based on their pair wise correlations.

2.5 | Differential gene expression analysis

Identification of differentially expressed genes (DEGs) was done using cuffdiff (Trapnell et al., 2010). The DEGs with log2 fold change ≥2 (induced) and/ or ≤ −2 (repressed) and an FPKM ≥2 for either of the sample in each pair wise comparison were considered to be significantly differentially expressed. Log2 transformed FPKM values of the DEGs were further subjected to K-means clustering using Pearson correlation with an optimal number of clusters set at 20 in "pheatmap" of R software.

2.6 | Gene annotations, (gene ontology) GO term, and pathway analysis

The identified DEGs were subjected to Blastx similarity (E-value cut-off of ≤10−5) search against National Center for Biotechnology Information nonredundant *Viridiplantae* protein database. The Blast hits were assigned GO terms using Blast2GO v4.1.9 (Conesa et al.,

TABLE 1 Details of tissues and RNA-Seq data used for developing gene expression atlas

S.No	Tissue description	Stage	Growth stage at collection	Sample ID	Raw reads sequenced	Filtered reads used for analysis	Reads Mapped	% Alignment
Set I								
1	Embryo	Germination	24- to 48-hr imbibition	Ger_Embryo	27,659,444	26,718,608	25,813,674	96.61
2	Plumule	Germination		Ger_Plumule	24,231,898	23,224,694	22,434,357	96.60
3	Radicle	Germination		Ger_Radicle	29,822,124	28,741,462	27,702,776	96.39
4	Epicotyl	Seedling	8–10 DAG	Sed_Epicotyl	29,184,088	27,735,808	26,812,656	96.67
5	Primary root	Seedling		Sed_Primary root	68,993,868	66,759,574	64,257,507	96.25
Set II								
6	Leaf	Vegetative	25–30 DAG	Veg_Leaf	38,638,886	37,415,410	36,099,919	96.48
7	Stem	Vegetative		Veg_Stem	40,801,302	39,782,106	38,302,532	96.28
8	Petiole	Vegetative		Veg_Petiole	36,678,298	35,682,994	34,415,449	96.45
9	Root	Vegetative		Veg_Root	38,660,970	37,709,458	36,063,742	95.64
Set III								
10	Leaf	Reproductive	40–50 DAG	Rep_Leaf	34,626,638	33,752,054	32,591,449	96.56
11	Stem	Reproductive		Rep_Stem	22,110,224	21,368,210	20,589,008	96.35
12	Petiole	Reproductive		Rep_Petiole	26,531,476	25,767,598	24,792,122	96.21
13	Root	Reproductive		Rep_Root	24,720,764	23,951,266	22,847,670	95.39
14	Nodules	Reproductive		Rep_Nodules	26,475,678	25,762,234	24,220,068	94.01
15	Buds	Reproductive		Rep_Buds	30,701,604	29,858,842	28,695,237	96.10
16	Flowers	Reproductive		Rep_Flowers	31,754,896	30,755,814	29,560,588	96.11
17	Immature seeds	Reproductive		Rep_Immature seeds	24,884,960	24,192,014	23,247,012	96.09
18	Pods	Reproductive		Rep_Pods	28,958,052	28,134,276	27,038,033	96.10
Set IV								
19	Leaf-yellow	Senescence	90–110 DAG	Sen_Leaf-yellow	21,945,950	21,938,342	20,048,693	91.39
20	Leaf	Senescence		Sen_Leaf	24,375,228	24,367,808	22,356,467	91.75
21	Stem	Senescence		Sen_Stem	23,815,080	22,950,286	21,744,073	94.74
22	Petiole	Senescence		Sen_Petiole	30,434,110	29,315,484	27,773,940	94.74
23	Root	Senescence		Sen_Root	24,118,890	23,378,432	21,620,005	92.48
24	Nodules	Senescence		Sen_Nodules	39,879,110	38,492,206	35,609,494	92.51
25	Immature seeds	Senescence		Sen_Immature seeds	25,719,118	25,710,836	23,450,795	91.21
26	Mature seeds	Senescence		Sen_Mature seeds	21,495,952	21,487,056	19,054,567	88.68
27	Seed coat	Senescence		Sen_Seed coat	19,450,052	19,445,612	17,560,721	90.31
Total					816,668,660	794,398,484	754,702,554	95.00

Note. DAG = days after germination.

2005), and their enrichment was carried out using "goseq" package from R software (Young, Wakefield, Smyth, & Oshlack, 2010). The GO terms exhibiting a p value of ≤.05 were considered to be significantly enriched. Further transcription factor (TFs) encoding genes were identified from DEGs using Plant TFDB (Jin, Zhang, Kong, Gao, & Luo, 2014) with an E-value cut-off of 1E − 10. Pathway analysis was further carried out using KEGG database.

2.7 | Quantitative real-time polymerase chain reaction

Slow drought was imposed on ICC 4958 under greenhouse conditions as described by Ray and Sinclair (1998). Seeds were sown in pots, and 10 days old seedlings were subjected to slow drought stress through a dry down experiment. Three biological replicates of control and stress plants were maintained at each stage for sample/tissue collection. Pots were allowed to dry through transpirational water loss until the transpiration ratio reached 0.1 at each stage. Root samples at vegetative (25–30 DAG), reproductive (40–50 DAG), and senescence (90–110 DAG) stages under drought stress were collected along with the respective controls. RNA was isolated from the collected stress and control tissues and checked for the quality as described above in RNA extraction. First strand cDNA was synthesized from total RNA (2.5 μg) using a cDNA synthesis kit (Superscript® III, Invitrogen, CA, USA) following manufacturer's instructions.

Real-time polymerase chain reaction (RT-PCR) was performed using Applied Biosystems 7500 Real-Time PCR System with the SYBR green chemistry (Applied Biosystems, USA) according to the manufacturer's instructions. Gene-specific primers for RT-PCR were designed using primer 3 software (Rozen & Skaletsky, 2000). PCR was carried out as described by Mir et al. (2014). The data from different PCR runs or cDNA samples were compared by using the mean of the Ct values of the three biological replicates that were normalized to the mean Ct values of the endogenous gene. The relative expression ratios were determined using the $2^{-\Delta\Delta Ct}$ method and Student's t test was used to calculate significance (Livak & Schmittgen, 2001). Relative transcription levels are presented graphically.

3 | RESULTS

3.1 | Development of *C. arietinum* gene expression atlas

To document genome-wide expression profiles in chickpea, a CaGEA was generated from RNA-Seq data obtained from Illumina sequencing. Twenty-seven samples representing plant organs/tissues from five major developmental stages (germination, seedling, vegetative, reproductive, and senescence) covering the entire life cycle of chickpea were collected from ICC 4958, a drought tolerant and widely used chickpea cultivar. As mentioned above, tissue samples were grouped into four sets based on their developmental stages for further analysis. Tissues from germination and seedling stages were grouped under Set I; vegetative to Set II; reproductive to Set III, and senescence to Set IV (Table 1; Figure 1). Germinal tissues were represented by embryo, plumule, and radicle, whereas the seedling stage tissues were represented by epicotyl and primary root. Vegetative stage comprises leaf, stem, petiole, and root tissues. Majority of tissues (nine tissues) were collected from reproductive and senescence stages. The reproductive stage tissues comprise leaf, stem, petiole, root, nodules, buds, flowers, immature seeds, and pods. Senescing organs/tissues representing the developmental ageing of chickpea included leaf-yellow, leaf, stem, petiole, root, nodules, immature seeds, mature seeds, and pod wall. For convenience, the developmental stage (Ger/Sed/Veg/Rep/Sen) together with the tissue name (Radicle/Stem/Leaf/Root) has been used as sample ID in all the analysis (for example, Ger_Embryo refers to Germination stage Embryo tissue). A total of 817 million (M) raw reads were generated and subjected to QC filtering resulting in 794 M (97%) filtered reads (Table 1). These 794 M filtered reads were used for further analysis. Sequencing reads generated from each sample varied from 19 to 66 M filtered reads. Mapping of the 794 M filtered reads to the chickpea genome (Varshney et al., 2013) resulted in the alignment of 755 M reads (95% alignment) ranging from 88.68% in Sen_Mature seeds to 96.67% in Sed_Epicotyl. All sequencing data generated have been deposited in National Center for Biotechnology Information Sequence Read Archive database with the BioProject ID PRJNA413872.

3.2 | Reference guided assembly

The mapped reads (755 M) from all the 27 tissue samples (representing five developmental stages) were used to generate reference guided assembly, global gene expression profiles, and differential gene expression profiles. A reference-guided assembly of the whole dataset was generated using Cufflinks–Cuffmerge pipeline. This assembly generated 105,609 transcripts representing 32,873 gene loci. This gene loci number is considerably higher than the number of genes reported (28,269) in the chickpea genome assembly by Varshney et al. (2013). The comparison of the reference-guided assembly with the chickpea genome led to the identification of 4,604 (14%) novel genes that are not annotated in the genome. The results from reference-guided assembly indicate the potential of RNA-Seq in identification of novel genes.

3.3 | Global gene expression patterns

The normalized expression level (FPKM) of each gene was estimated in all the 27 samples analysed. To exclude genes with low confidence expression values, only those genes with an FPKM value ≥1 in at least one of the 27 tissues analysed were designated as expressed. Based on this criteria, 25,784 (78.4%) genes out of 32,873 were identified to be transcriptionally active in the present dataset. These results indicate substantial and comprehensive representation of the transcriptome. Transcriptional activity was highly variable and diverse across tissues, with the buds (Rep_Buds) expressing the largest number of genes (19,652; 70% of all genes), and the senescence mature seeds (Sen_Mature seeds) tissue expressing the smallest number of genes (15,701; 56%).

The biological identity of tissues studied was well reflected in their respective transcriptome as revealed by hierarchical clustering of Pearson's correlations in all 27 tissues (Figure 2). Most of the tissues clustered together based on morphological, physiological, and developmental attributes providing an interesting glimpse into the cell differentiation and plant tissues studied. For example, floral (Rep_Buds, Rep_Flowers) and seed (Rep_Immature seeds, Rep_Pods, Sen_Immature seeds, and Sen_Seed coat,) tissues clustered together. Similarly, leaf and stem tissues (Sed_Epicotyl, Veg_Leaf, Veg_Stem, Veg_Petiole, Rep_Leaf, Rep_Stem, Rep_Petiole, Sen_Leaf-yellow, Sen_Leaf, Sen_Stem, and Sen_Petiole) clustered together. All root tissues (Sed_Primary root, Veg_Root, Rep_Root, and Sen_Root) from different developmental stages clustered together. The nodule tissues (Rep_Nodules and Sen_Nodules) clustered together. All three germination stage tissues (Ger_Embryo, Ger_Plumule, and Ger_Radicle) clustered together. Interestingly, Sen_Mature seeds showed unique clustering pattern and did not cluster together with any of the tissues studied.

Further, based on the expression levels, the genes were categorized into low/moderately expressed (FPKM ≥1 and < 10) and highly expressed (FPKM ≥10). Among 27 samples, Sen_Immature seeds followed by Sen_Pod wall, Sen_Mature seeds, and Sen_Nodules showed a maximum number of highly expressed genes (FPKM ≥10). Although a maximum number of low/moderately expressed genes (FPKM >1 and < 10) were observed in Rep_Buds followed by Rep_Root, Sen_Root, Rep_Flower, and so forth. Distribution of genes in 27 samples under these categories based on low/moderate and high expression levels has been provided in Table 2. Of the 25,784 genes (FPKM ≥1), 12,085 genes were expressed in all 27 samples studied. Analysis of 12,085 genes resulted in identification of 208 stably expressed genes in all the samples with coefficient of variation (CV = standard deviation/mean) ≤ 5% (Supporting Information Table 1). The set of 208 stably expressed genes represent a useful resource for identification of reference/housekeeping genes, which are necessary to normalize expression counts across tissues and developmental stages or due to environmental instabilities. A GO enrichment analysis was performed on these constitutively expressed genes which indicated an over-representation of the following GO terms: RNA binding (GO: 0003723), nucleobase-containing compound metabolic process (GO: 0006139), cellular component organization (GO: 0016043), transport (GO: 0006810), cellular protein

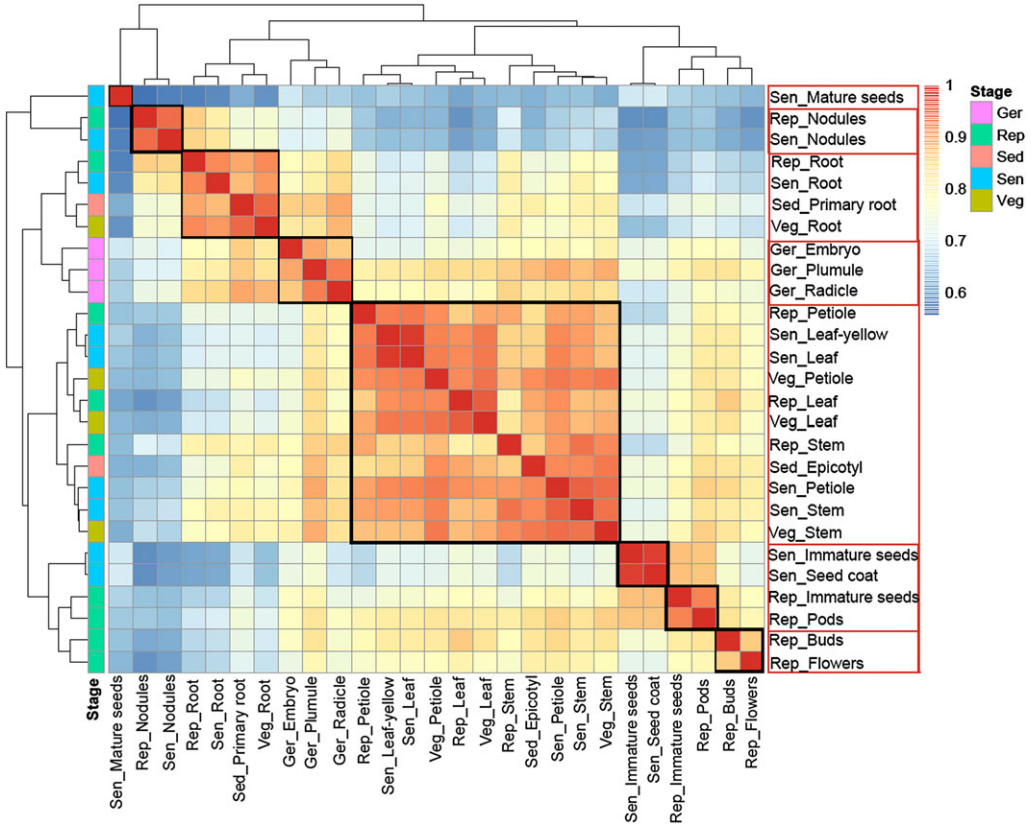

FIGURE 2 Global gene expression patterns among different chickpea tissues. Heatmap of hierarchical clustering of Pearson's correlations (*R*) for all 27 tissues included in the gene expression atlas. Genes with a normalized expression level FPKM > 1 in at least one of the 27 tissues analysed were log2, +1 transformed before analysis and were designated as expressed. The colour scale indicates the degree of correlation. Samples were clustered based on their pairwise correlations [Colour figure can be viewed at wileyonlinelibrary.com]

modification process (GO: 0006464), and translation initiation factor activity (GO: 0008135). Similar GO terms have also been found to be enriched in several other plant species adding support to the list of constitutively expressed genes identified in this study. Furthermore, of the 25,784 genes, 1,574 genes were tissue-specific showing $\tau = 1$ (Supporting Information 2). A maximum number of tissue-specific genes were observed in Rep_Buds (280) followed by Sen_Mature seeds (203), Rep_Flower (155), and Rep_Nodules (154). The lowest number of tissue-specific genes was present in Ger_Plumule (7), Sen_Leaf-yellow (7), Sen_Leaf (8), and so forth. These tissue-specific genes represented significant information for targeted gene expression and provided insights into the specialized process occurring in these tissues. In order to identify the tissue-specific biological processes, GO enrichment analysis on tissue-specific genes was carried out. As expected, over-representation of predictable specialized processes in different plant tissues such as "pollination" and "multi cellular organism development" in Rep_Buds, "reproduction" and "cell wall" in Rep_Pods, "transport" and "kinase activity" in Sen_Root, and "cell differentiation" in Veg_Root was observed. The total genes expressed were grouped into three categories based on expression in 27 samples. These include genes with low/moderate, high, and tissue-specific expression (Table 2).

3.4 | Differential gene expression

Pairwise analysis identified DEGs between any two combinations of the tissues at the same or different developmental stage studied. A total of 15,947 unique genes were identified to be significantly differentially expressed with ≤ -2 and ≥ 2 fold change in any of the combinations studied (Supporting Information—Table S3). Overall, 14,619 and 14,743 genes were induced and repressed, respectively, across all possible pairwise tissue combinations. In addition, out of the 15,947 differentially expressed genes, 2,169 genes were identified to be transcription factors. Additionally, DEGs between different combinations of similar/related tissues were identified and showed wide range in the number of DEGs. For example, the highest number of induced genes was observed in Sen_Immature seeds versus Sen_Mature seeds followed by Sen_mature seeds versus Sen_Seed coat, Rep_Pods versus Sen_Mature seeds, Rep_Immature seeds versus Sen_Mature seeds, and so forth. On the other side, the lowest number of induced genes was observed in Sen_Immature seeds versus Sen_Seed coat and Sen_Leaf-yellow versus Sen_Leaf. In the case of repressed genes, the highest number of genes was observed in Sen_Mature seeds versus Sen_Seed coat, Sen_Immature seeds versus Sen_Mature seeds, Rep_Pods versus Sen_Mature seeds, Rep_Immature seeds versus Sen_Mature seeds, and so forth. The lowest number of repressed genes was observed in Sen_Immature seeds versus Sen_Seed coat, Sen_Leaf-yellow versus Sen_Leaf, and so forth. The pairwise combinations showing the highest and lowest numbers of induced genes also showed the highest and lowest numbers of repressed genes (Figure 3). In addition, unique set of genes exhibiting expression when compared with similar/related tissues across different developmental stages were identified. For example, Ger_Plumule

TABLE 2 Distribution of expressed genes in 27 tissues of different developmental stages under low/moderate, high and tissue specific expression categories

Tissue	No. of expressed genes	Genes with low/moderate expression (FPKM ≥ 1 and ≤ 10)	Genes with high expression (FPKM ≥ 10)	No. of genes showing tissue specific expression
Ger_Embryo	17,810	17,762	48	61
Ger_Plumule	18,815	18,764	51	7
Ger_Radicle	18,586	18,530	56	25
Sed_Epicotyl	18,926	18,873	53	42
Sed_Primary root	18,409	18,352	57	26
Veg_Leaf	18,527	18,428	99	24
Veg_Stem	18,611	18,568	43	14
Veg_Petiole	18,456	18,395	61	21
Veg_Root	18,741	18,690	51	42
Rep_Leaf	19,074	18,957	117	30
Rep_Stem	18,159	18,107	52	21
Rep_Petiole	17,800	17,694	106	33
Rep_Root	19,443	19,396	47	29
Rep_Nodules	18,479	18,378	101	154
Rep_Buds	19,652	19,625	27	280
Rep_Flowers	19,209	19,130	79	155
Rep_Immature seeds	18,195	18,158	37	96
Rep_Pods	18,549	18,491	58	14
Sen_Leaf-yellow	18,292	18,201	91	7
Sen_Leaf	18,204	18,113	91	8
Sen_stem	18,703	18,665	38	20
Sen_Petiole	18,831	18,788	43	12
Sen_Root	19,357	19,318	39	89
Sen_Nodules	18,556	18,436	120	121
Sen_Immature seeds	17,947	17,657	290	28
Sen_Mature seeds	15,701	15,512	189	203
Sen_Seed coat	17,890	17,614	276	12

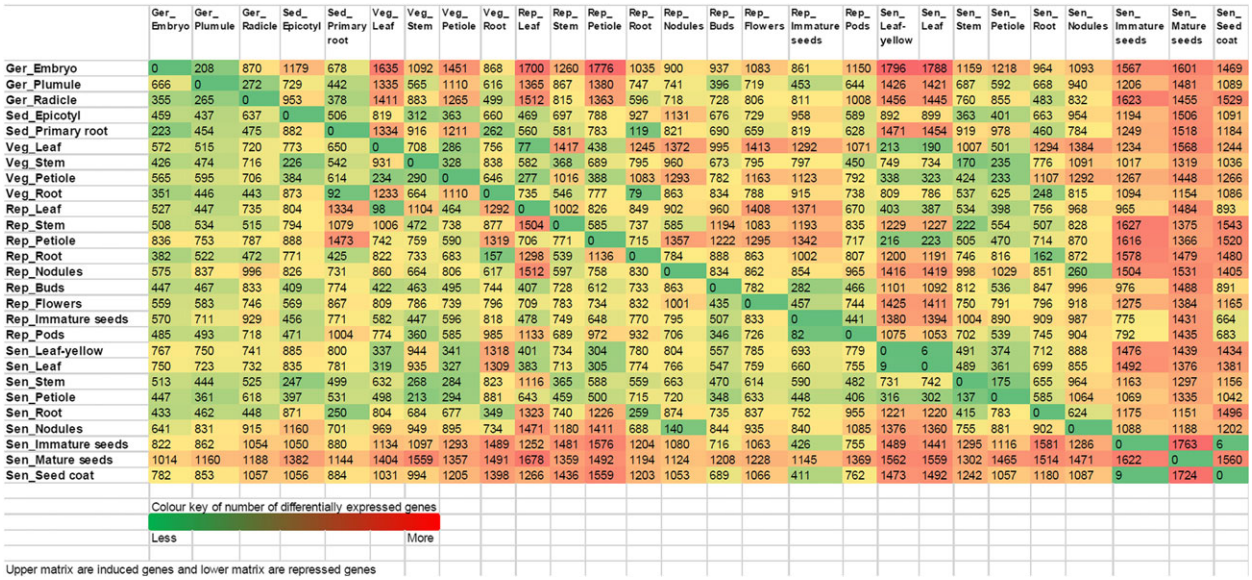

FIGURE 3 Tissue by tissue comparison of differentially expressed genes. This figure shows the total number of genes with a significant increase (upper triangular matrix) or decrease (lower triangular matrix) in gene expression between the row tissue (left) and the column tissue (top). For example, there are 208 and 666 genes which were upregulated and downregulated, respectively, when Ger_Embryo and Ger_Plumule tissues were compared. Therefore, each cell represents a different set of differentially expressed genes [Colour figure can be viewed at wileyonlinelibrary.com]

tissue when compared with Sed_Epicotyl, Veg_Leaf, Veg_Petiole, Rep_Leaf, Rep_Petiole, Sen_Leaf-yellow, Sen_Leaf, and Sen_Petiole showed 2,813 and 1,785 unique genes induced and repressed, respectively. On the other hand, 182 and 32 common genes were induced and repressed in these pairwise combinations involving Ger_Plumule. Similarly, Seed_Epicotyl (vs. Veg_Leaf, Veg_Petiole, Rep_Leaf, Rep_Petiole, Sen_Leaf-yellow, Sen_Leaf, and Sen_Petiole) showed 2,505 and 2,931 unique genes induced and repressed, respectively, whereas 87 and 53 genes were commonly induced and repressed across these combinations. Veg_Leaf (vs. Veg_Petiole, Rep_Leaf, Rep_Petiole, Sen_Leaf-yellow, Sen_Leaf, and Sen_Petiole) showed 1,267 and 1,368 unique genes and 10 and seven common genes, induced and repressed, respectively. Veg_Petiole (vs. Rep_Leaf, Rep_Petiole, Sen_Leaf-yellow, Sen_Leaf, and Sen_Petiole) showed 878 and 16 unique genes and 16 and nine common genes, induced and repressed, respectively. Rep_Leaf (vs. Rep_Petiole, Sen_Leaf-yellow, Sen_Leaf, and Sen_Petiole) showed 1,062 and 1,423 unique genes and 152 and 39 common genes, induced and repressed, respectively. Rep_Petiole (vs. Sen_Leaf yellow, Sen_Leaf, and Sen_Petiole) showed 609 and 703 unique genes and 94 and 126 common genes, induced and repressed, respectively. Sen_Leaf-yellow (vs. Sen_Leaf, Sen_Petiole) showed 375 and 324 unique genes and five and one common genes, induced and repressed, respectively. Sen_Leaf versus Sen_Petiole showed 361 and 302 induced and repressed genes. In different pairwise combinations of stem-related tissues (Ger_Plumule, Sed_Epicotyl, Veg_Stem, Rep_Stem, and Sen_Stem), 2,492 and 2,324 uniquely induced and repressed genes were identified. On the other hand, 18 and 19 commonly induced and repressed genes in stem tissues were identified in the study. Similarly, pairwise combinations of root related tissues (Ger_Radicle, Sed_Primary root, Veg_Root, Rep_Root, and Sen_Root) identified 1,553 and 1,541 uniquely induced and repressed genes, whereas, 5 and 3 common genes were induced and repressed in root tissues. Additionally, a unique set of DEGs were identified in specialized plant tissue combinations such as floral (782 induced and 435 repressed) and nodule (260 induced and 140 repressed) tissues. The DEGs identified in the present study indicate that different tissues exhibit divergent gene expression patterns across the five plant developmental stages.

3.5 | GO terms

Genes exhibiting tissue-specific expression patterns are involved in a variety of GO categories. To further validate and analyze CaGEA, GO analysis was performed. Functional annotations could be retrieved for 15,179 out of 15,947 DEGs. The 15,179 genes with annotations were functionally categorized based on GO descriptions. As a result, these genes were assigned to three principal categories: biological process (15,755), molecular function (11,050), and cellular component (14,038). The highest number fell into the following subcategories: metabolic process (5,149), followed by catalytic activity (4,908), binding (4,648), cellular process (4,360), membrane (3,622), cell (3,551), and cell part (3,477) subcategories (Figure 4). Genes related to growth and development were also identified using GO classifications. A total of 279 and 40 genes were found under the "developmental process" and "growth" subcategories.

3.6 | TF encoding genes

Transcription factors represent the key molecular switches orchestrating the regulation of plant growth and developmental processes. In the present study, a total of 2,169 TFs encoding genes were identified to be differentially expressed. These 2,169 TFs could be categorized into at least 65 families. The largest number of transcription factor encoding genes identified belong to the TF families: bHLH (157), MYB (111), WRKY (100), ERF (97), bZIP (96), MYB-like (93), NAC (88), GRAS (51), NF-Y (40), C3H (44), MIKC (33), G2-like (38), HSF (35) Trihelix (35), followed by AP2 (29), E2F-DP (29), GATA (27), TCP (25), HB (15), LOB (5), MADS (4), SET (1), and so forth. Based on significant role of TFs, heatmaps representing important TF families with relevant function in select tissues were generated. Heatmaps were generated for ERF, GRAS, NF-Y, and Zinc finger TF families using

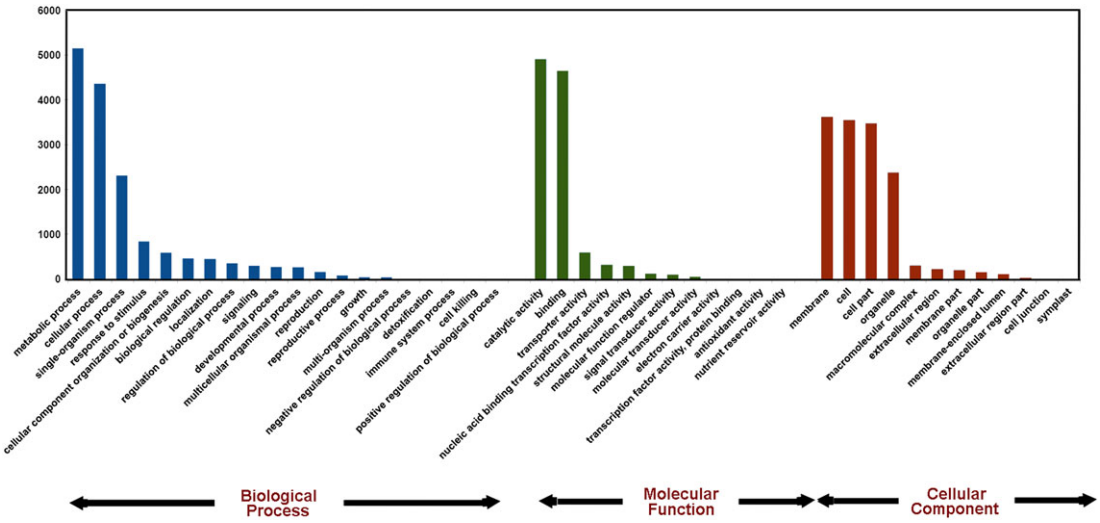

FIGURE 4 Gene ontology annotation of differentially expressed genes. Bar graph representing GO annotations of DEGs in three categories, (a) biological processes, (b) molecular functions, and (c) cellular components [Colour figure can be viewed at wileyonlinelibrary.com]

Ger_Radicle, Sed_Primary root, Veg_Root, Rep_Root, Rep_Nodules, Sen_Root, and Sen_Nodules. Similarly, heatmaps were also generated for MYB and WRKY families using Ger_Embryo, Rep_Buds, Rep_Flower, Rep_Immature seeds, Rep_Pods, Sen_Immature seeds, Sen_Mature seeds, and Sen_Seed coat. Heatmaps representing differential expression profiles of important transcription factor families in select tissues were depicted in Figure 5.

3.7 | Expression trends across the different developmental stages and identification of tissue-specific genes

To study clusters of genes with similar expression trends across different developmental stages, K-means clustering was performed on all DEGs (15,947). Based on gene expression patterns of different genes, 20 clusters (Cluster 1 to Cluster 20) were generated whose size ranged from 2,428 to 168 genes. Eleven clusters (Clusters 1, 2, 3, 4, 8, 10, 11, 15, 16, 19, and 20) exhibited unique gene expression profiles across various tissues studied. These clusters were further analysed for cluster-specific biological functions. The results demonstrated that Cluster 4 comprises 345 genes showed highly induced expression in Rep_Buds and Rep_Flowers, the representative floral tissues among 27 tissues studied indicating preferential gene expression in floral organs. This cluster consisted of several floral development-specific genes including LRR receptor-like kinase (*Ca_25115*, *Ca_17807*, *Ca_10392*, *Ca_08597*); pollen-specific LRR (*Ca_01511*); MADS-box transcription factor (*Ca_04909*, *Ca_12381*, *Ca_16424*); floral homeotic PMADS 2-like (*Ca_13111*, *Ca_03887*); pollen family allergen (*Ca_12884*); and AP2 domain class TFs (*Ca_12218*). Similarly, Cluster 19 included 279 genes showed induced expression in seed and related tissues such as Rep_Immature seeds, Rep_Pods, Sen_Immature seeds, and Sen_Seed coat. The genes clustered include auxin signalling (*Ca_03093*, *Ca_10007*, *Ca_10008*); IAA-amino acid hydrolase ILR1-like 4 (*Ca_02768*, *Ca_02769*, *Ca_14555*, *Ca_14553*); and glycosyl hydrolase family 10 (*Ca_05649*, *Ca_05648*), which are mainly involved in seed development process. This cluster of genes showed very low level of expression in Sen_Mature seeds tissue. Furthermore, Clusters 1 and 20 comprises 830 and 272 genes showed highly induced expression in root and root related tissues such as Ger_Radicle, Sed_Primary Root, Veg_Root, Rep_Root, Sen_Root, Rep_Nodules, and Sen_Nodules. These clusters consisted of several genes involved in root development process such as WRKY (*Ca_25078*, *Ca_08049*, *Ca_06163*, *Ca_10993*, *Ca_05204*, *Ca_03388*, *Ca_18932*, *Ca_17759*, *Ca_00420*) and MYB (*Ca_09400*, *Ca_15309*, *Ca_08598*, *Ca_19393*, *Ca_03266*) families; glutathione S-amino-terminal domain (*Ca_23113*, *Ca_14389*, *Ca_03442*, *Ca_08920*, *Ca_14581*, *Ca_08920*); AP2-domain class TFs (*Ca_08331*, *Ca_16397*, *Ca_06032*, *Ca_06034*, *Ca_21773*); and LRR receptor-like (*Ca_00354*, *Ca_07956*, *Ca_22522*, *Ca_18543*, *Ca_15676*, *Ca_08906*, *Ca_08620*, *Ca_05106*, *Ca_18898*, *Ca_17289*, *Ca_22105*, *Ca_22057*, *Ca_26423*). Of the two clusters, predominant gene expression patterns for root development were observed in Cluster 20. Clusters 2 and 16 comprising of 234 and 653 genes showed highly induced expression in Rep_Nodules and Sen_Nodules among all 27 tissues studied, displaying nodulation preferential gene expression. These clusters consisted of several nodulation-specific genes including early nodulin-like (*Ca_14071*, *Ca_03324*, *Ca_03325*, *Ca_27967*, *Ca_21379*, *Ca_19627*, *Ca_27967*, *Ca_21379*); leghemoglobin Lb120–1 (*Ca_16113*, *Ca_16084*, *Ca_08383*, *Ca_23593*), nodulation inception (*Ca_09832*), and nodule cysteine-rich secreted peptide (*Ca_25878*; see corresponding gene annotations in Supporting Information—Table S3). Cluster 2 showed predominant gene expression patterns for nodulation process when compared to Cluster 16. Clusters 8 and 10 comprising of 2,428 and 188 genes, respectively, showed induced expression specifically in Sen_Mature seeds. These clusters consisted of several seed development responsive genes including AP2 domain (*Ca_03239*, *Ca_20041*, *Ca_12465*); zinc finger (57) (*Ca_09517*, *Ca_12029*, *Ca_22065*, *Ca_14069*,

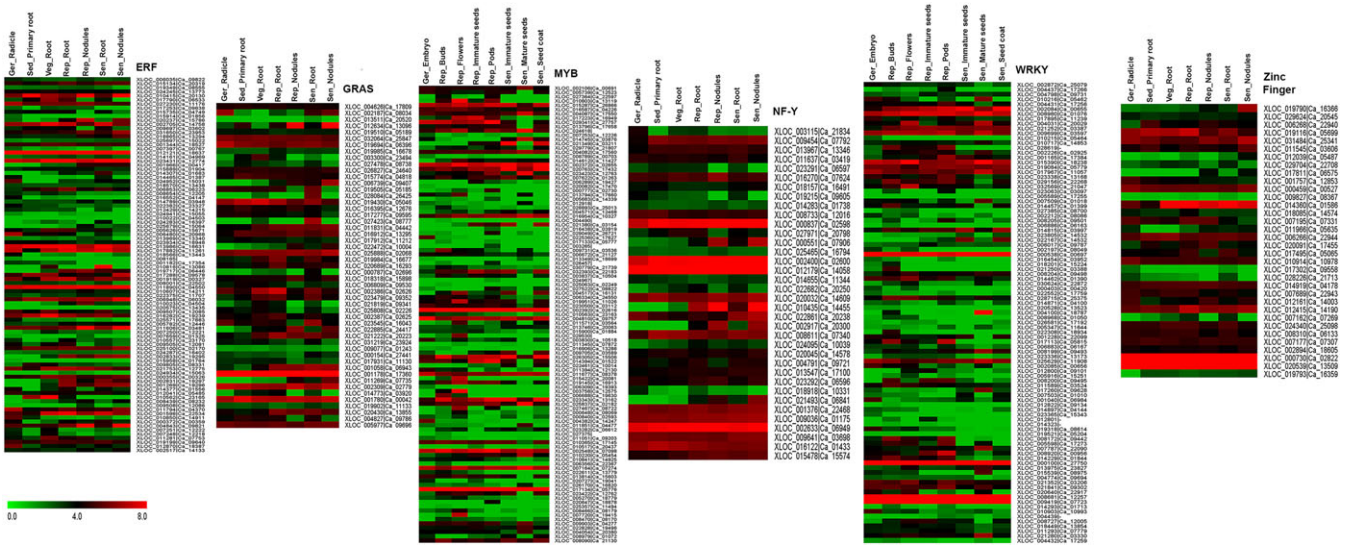

FIGURE 5 Expression profiles of important differentially expressed transcription factor (TFs) gene families in the select tissues. The unique gene identifier (suffix XLOC) and corresponding reference gene loci for each TF family have been given on right side. The colour scale at the bottom represents normalized expression values [Colour figure can be viewed at wileyonlinelibrary.com]

Ca_11259, *Ca_06604*); IAA family (*Ca_03444*, *Ca_01754*); heat shock factors (17) (*Ca_00542*, *Ca_18324*, *Ca_14692*, *Ca_12525*, *Ca_01654*, *Ca_06335*, *Ca_11271*, *Ca_16176*, *Ca_11875*); and alpha amylase (*Ca_21106*, *Ca_10465*). Cluster 11 comprising of 354 genes showed induced expression leaf and petiole tissues (Veg_Leaf, Veg_Petiole, Rep_Leaf, Rep_Petiole, Sen_Leaf-yellow, Sen_Leaf, Sen_Petiole). These tissues consisted of several genes involved in leaf developmental process such as: Auxin responsive (*Ca_06140*, *Ca_08294*, *ca_15089*, *Ca_17510*, *Ca_17504*); TCP-like (*Ca_02473*); NAC family (*Ca_18171*); bHLH (*Ca_04778*); bZIP (*Ca_22618*); ERF (*Ca_06673*). Additionally, Clusters 3 and 15 showed induced expression in aerial tissues (leaf, stem, petiole, floral tissues, etc.) and repressed expression in underground tissues such as roots and nodules across different developmental stages. In summary, these clusters contain potential genes involved in organ (leaf, floral, seed, root, and nodules) developmental process. The rest of the nine clusters, namely, 5, 6, 7, 9, 12, 13, 14, 17, and 18 did not exhibit any unique gene expression patterns across tissues. Overall, cluster analysis and selected clusters showing distinct preferential gene expression patterns with respect to specific tissues is depicted in Figure 6.

3.8 | Pathways triggered during plant growth and development

Pathway analysis was performed to investigate pathways activated during chickpea growth and developmental process. A total of 121 pathways representing 640 enzymes involved in biosynthesis, metabolism, signalling, and cell differentiation were found to be activated. Major pathways activated are alanine, aspartate and glutamate metabolism (17 genes), amino sugar and nucleotide sugar metabolism (32), biosynthesis of antibiotics (133), carbon fixation in photosynthetic organisms (30), cysteine and methionine metabolism (37), galactose metabolism (26), glycine, serine, and threonine metabolism (20), glycolysis/gluconeogenesis (31), glyoxylate and dicarboxylate metabolism (19), phenylalanine, tyrosine, and tryptophan biosynthesis (23), purine metabolism (91), pyrimidine metabolism (35), pyruvate metabolism (32), starch and sucrose metabolism (55), and valine, leucine, and isoleucine degradation (19).

Glutamate is of central importance in plant N metabolism because all other amino acids biosynthesis requires this compound. Alanine, aspartate, and glutamate metabolism pathway involved glutamine-hydrolysing synthase (EC 6.3.5.4, EC 6.3.5.5) and glutaminase (EC 3.5.1.2). Similarly, the sulphur-containing amino acids are methionine and cysteine. Cysteine and methionine metabolism includes adenosyltransferase (EC 2.5.1.6), S-methyltransferase (EC 2.1.1.14), and S-adenosylhomocysteine synthase (EC 3.3.1.1). These pathways strongly indicated changes in gene expression during plant growth and development. Especially, during early stages such as germination stage, starch, and sucrose metabolism plays important role. In the present study, starch and sucrose metabolism pathway was identified as one of the major pathway involving endo-1,4-beta-D-glucanase (EC 3.2.1.4), glycogenase (EC 3.2.1.1), endo-1,3-beta-D-glucosidase (EC 3.2.1.39), phosphorylase (EC 2.4.1.1), branching enzyme (EC 2.4.1.18), and alpha-glucosidase (EC 3.2.1.48).

The importance of sulphate for plant growth, vigour leading to increased crop yield and nutritional quality has clearly been recognized, and the major sites of sulphur metabolism are the photosynthetic organs. In the present study, pathway analysis further

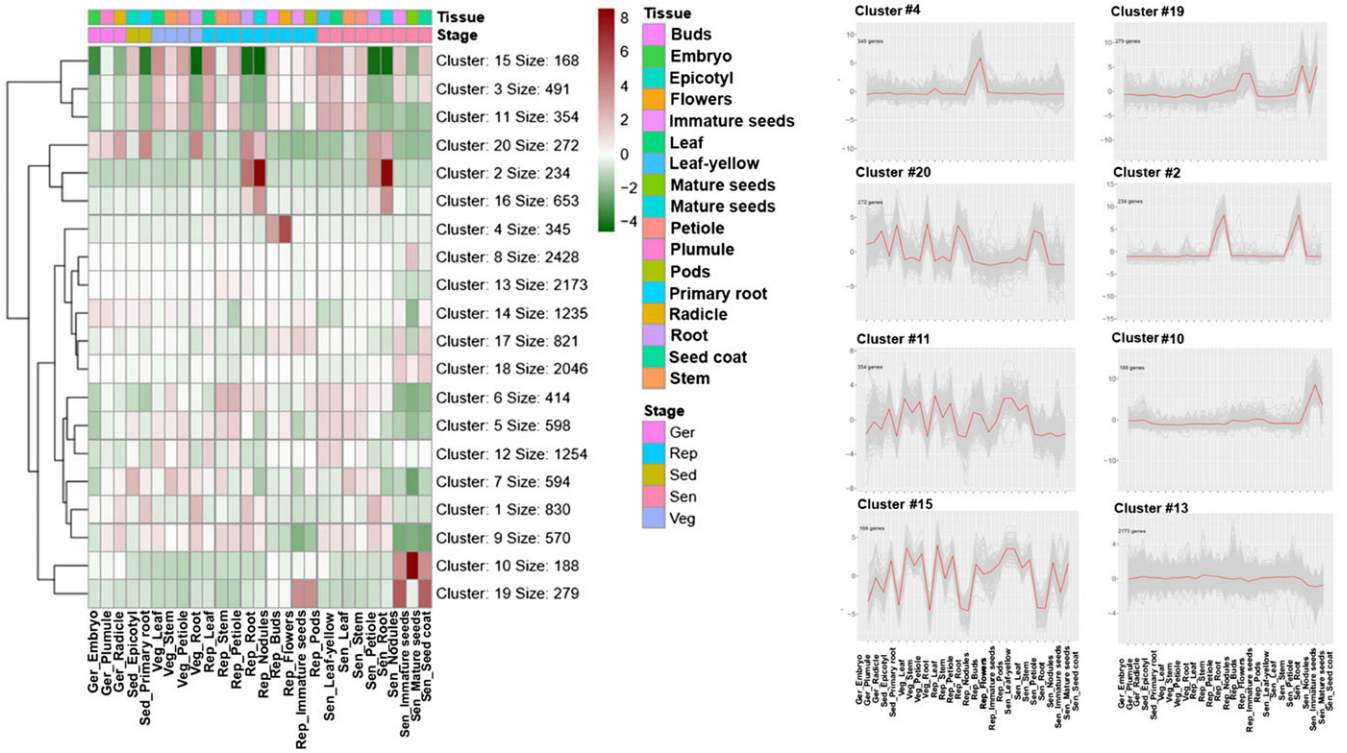

FIGURE 6 Cluster analysis depicting tissue-specific gene expression across different developmental stages [Colour figure can be viewed at wileyonlinelibrary.com]

revealed involvement of sulphur metabolism pathway that could be due to presence of photosynthetic organs such as leaf tissues, stems, and petioles. This pathway involves nucleotidase (EC 3.1.3.7), sulphurtransferase (EC 2.8.1.1), synthase (EC 2.5.1.47), and kinase (EC 2.7.1.25). Details of pathways and the corresponding genes identified are given in Supporting Information—Table 4 and Figure 7.

3.9 | Expression trends in the drought "*QTL-hotspot*" region

To investigate the overall patterns of gene expression in the "*QTL-hotspot*" region for drought tolerance present on *Ca4* pseudomolecule in chickpea, 26 genes (15 "*QTL-hotspot_a*" and 11 "*QTL-hotspot_b*") identified in the region from a different study were considered for analysis (Kale et al., 2015). Major QTLs in this region were identified for 11 traits such as root length density (cm/cm^3), root dry weight/total plant dry weight ratio (%), shoot dry weight (g), plant height (cm), primary branches, days to 50% flowering, days to maturity, pods/plant, 100-seed weight (g), harvest index (%), and delta carbon ratio. Gene expression patterns of the 26 candidate genes (identified by Kale et al., 2015) in select tissues specific for the traits of interest were analysed to ensure the role of key genes responsible for regulation of drought tolerance in chickpea. These tissues include all root, flower, and seed tissues (Ger_Radicle, Sed_Primary Root, Veg_Root, Rep_Root, Rep_Buds, Rep_Flower, Rep_Immature seeds, Rep_Pods, Sen_Root, Sen_Immature seeds, Sen_Mature seeds) at different developmental stages. Of the 26 genes described above, 17 genes (*Ca_04550*, *Ca_04551*, *Ca_04552*, *Ca_04555*, *Ca_04556*, *Ca_04557*, *Ca_04558*, *Ca_04560*, *Ca_04561*, *Ca_04563*, *Ca_04564*, *Ca_04566*, *Ca_04567*, *Ca_04568*, *Ca_04569*, *Ca_04570*, and *Ca_04571*) showed differential gene expression in tissues studied across five developmental stages. Majority of these genes were found to be induced in root, flower, and seed tissues. Kale et al. (2015) identified a subset of 12 common genes out of 26 genes using two different analysis, namely, high density QTL mapping and gene enrichment analysis. In the present study, nine genes (*Ca_04561*, *Ca_04563*, *Ca_04564*, *Ca_04566*, *Ca_04567*, *Ca_04568*, *Ca_04569*, *Ca_04570*, and *Ca_04571*) among 17 differentially expressed genes represent the subset of 12 genes. Among these nine, five genes encoding, E3 ubiquitin-protein ligase RNF128-like (*Ca_04561*), Calmodulin-binding motif (*Ca_04563*), LRR extension 2 (*Ca_04564*), kinase interacting (KIP1-like) family (*Ca_04566*), and homocysteine S-methyltransferase (*Ca_04569*) are highly expressed in at least one of the five root tissues studied. Similarly, six genes encoding E3 ubiquitin-protein ligase RNF128-like (*Ca_04561*), calmodulin-binding motif (*Ca_04563*), kinase interacting (KIP1-like) family (*Ca_04566*), serine threonine-kinase HT1-like (*Ca_04567*), homocysteine S-methyltransferase (*Ca_04568*), and homocysteine S-methyltransferase (Ca_04569) are highly expressed in both the flower tissues studied. Furthermore, six genes encoding E3 ubiquitin-protein ligase RNF128-like (*Ca_04561*), kinase interacting (KIP1-like) family (*Ca_04566*), serine threonine-kinase HT1-like (*Ca_04567*), homocysteine S-methyltransferase (*Ca_04568*), and vicilin 47 kDa (*Ca_04570*, *Ca_04571*) are highly expressed in at least one of the four seed tissues studied.

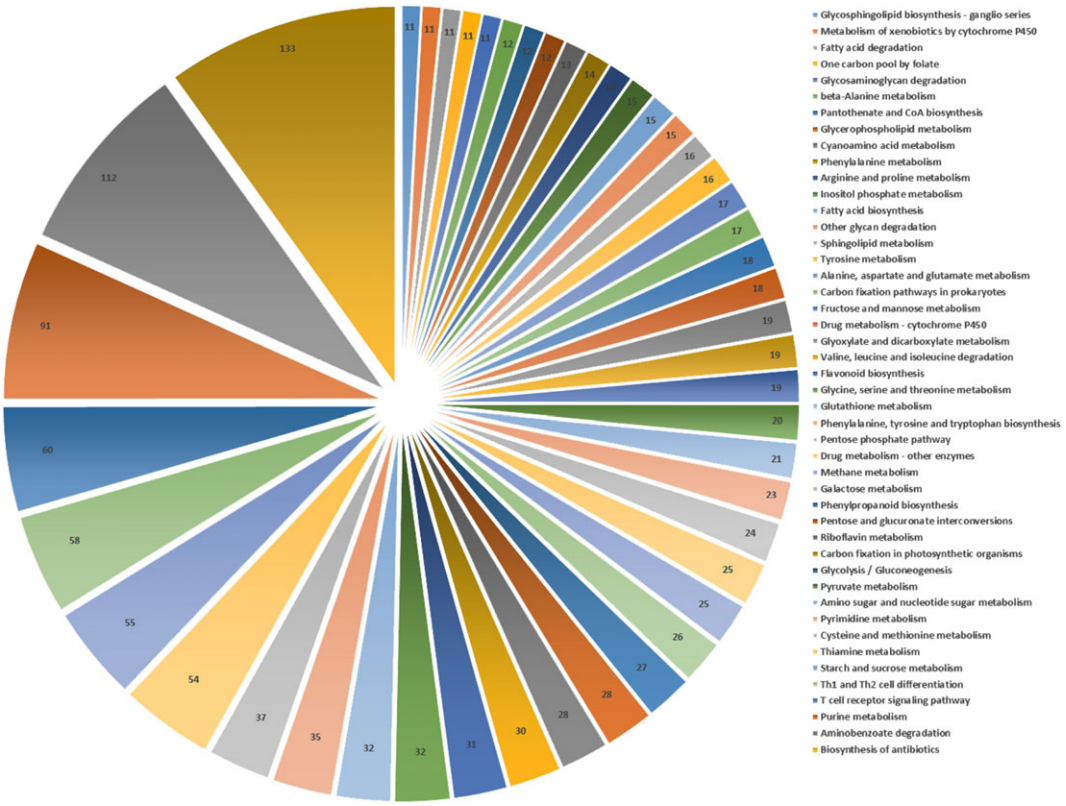

FIGURE 7 Major pathways for the differentially expressed genes. The numbers on the chart indicates the total number of genes falling within each pathway [Colour figure can be viewed at wileyonlinelibrary.com]

Of the nine genes, three genes (*Ca_04561*, *Ca_04566*, and *Ca_04568*) were common for all the tissues studied. Upregulation of these genes in selected tissues (mentioned above) which were specific for drought tolerance related traits revealed the important role of key candidates in flower, root, and seed developmental process under drought stress. An overview of gene expression profiles across different developmental stages highlighting *Ca4* containing "*QTL-hotspot*" has been depicted in Circos plot and heatmap (Figure 8).

3.10 | Validation of candidate genes identified from drought tolerance "*QTL-hotspot*" region

For investigating expression patterns of the nine genes identified within "*QTL-hotspot*" region, gene expression profiling was done in root tissues collected from ICC 4958 under stress and control conditions at vegetative, reproductive, and senescence stages of the plant using quantitative RT-PCR (qRT-PCR). Among the nine genes analysed, six genes, *Ca_04561*, *Ca_04564*, *Ca_04566*, *Ca_04567*, *Ca_04568*, and *Ca_04569* showed induced expression under stress conditions (Figure 8). Of all, *Ca_04568*, encoding for homocysteine S-methyltransferase, showed high induced expression in all the three root stress tissues (3.47, 7.12, and 11.15 fold change in vegetative, reproductive, and senescence stages, respectively) when compared with control tissues. Similarly, highly induced expression patterns in one or more root stress tissues when compared with control tissues at different developmental stages with respect to the genes studied were observed as follows: *Ca_04564* encoding for leucine-rich repeat extensin (LRX) 2 (0.91, 4.22, and 9.44 fold change), *Ca_04561* encoding E3 ubiquitin-protein ligase (1.44, 7.51, and 4.88 fold change), *Ca_04567* encoding serine threonine-kinase HT1-like (2.09, 5.08, and

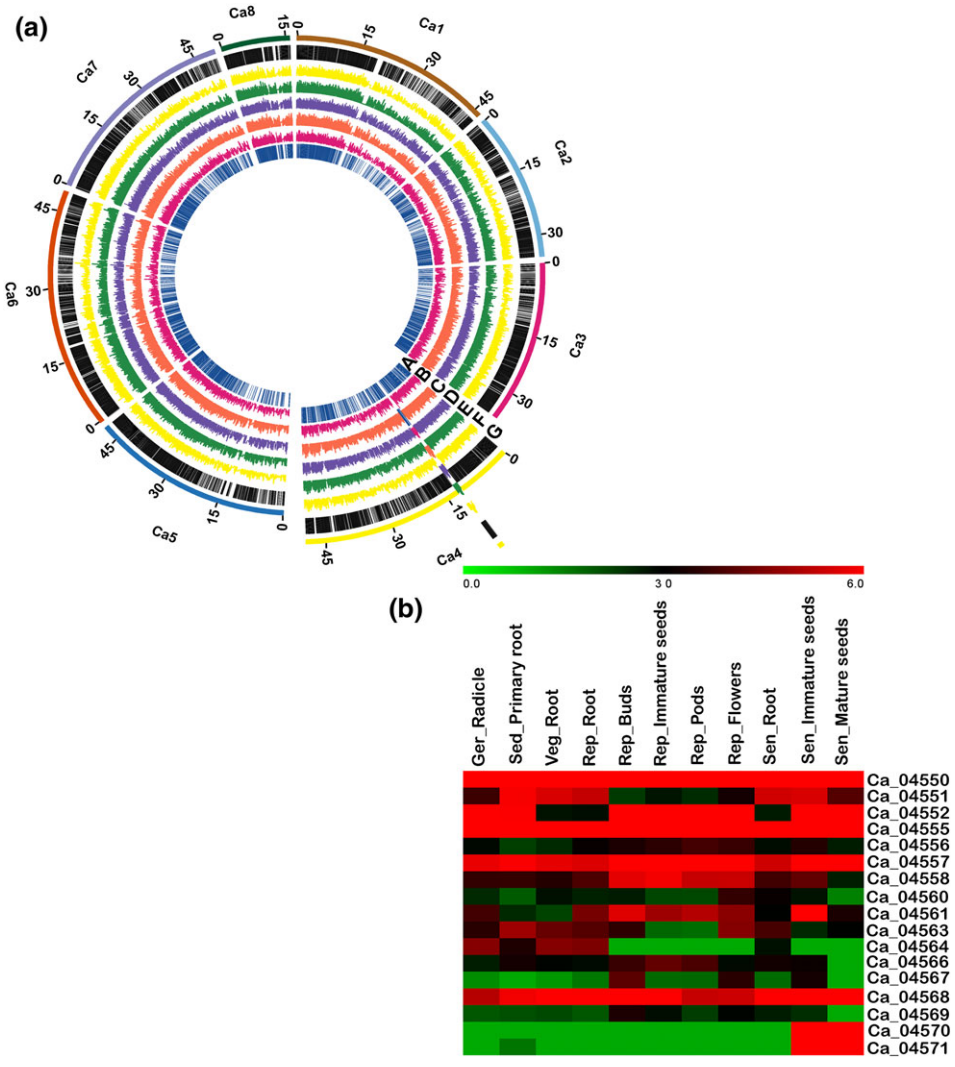

FIGURE 8 Expression profiles of genes studied under different developmental stages and differential expression of key candidate genes identified in the "*QTL-hotspot*" region. (a) Concentric circles in the Circos plot from outside to inside represent A = blue colour shows the transcription factors, B = pink colour shows the expression of genes across germination stage, C = orange colour shows the expression of genes across senescence stage, D = purple colour shows the expression of genes across reproductive stage, E = green colour—expression of genes across vegetative stage, F = yellow colour—expression of genes across seedling stage, G = black colour—denotes DEGs across all possible tissue combinations; (b) heatmap showing differential expression of 17 genes in the "*QTL-hotspot*" region for drought tolerance on *Ca4* pseudomolecule of chickpea. The gene IDs are given on the right side and tissue ID is represented on the top. The colour scale at the top represents normalized expression values of these genes in select tissues [Colour figure can be viewed at wileyonlinelibrary.com]

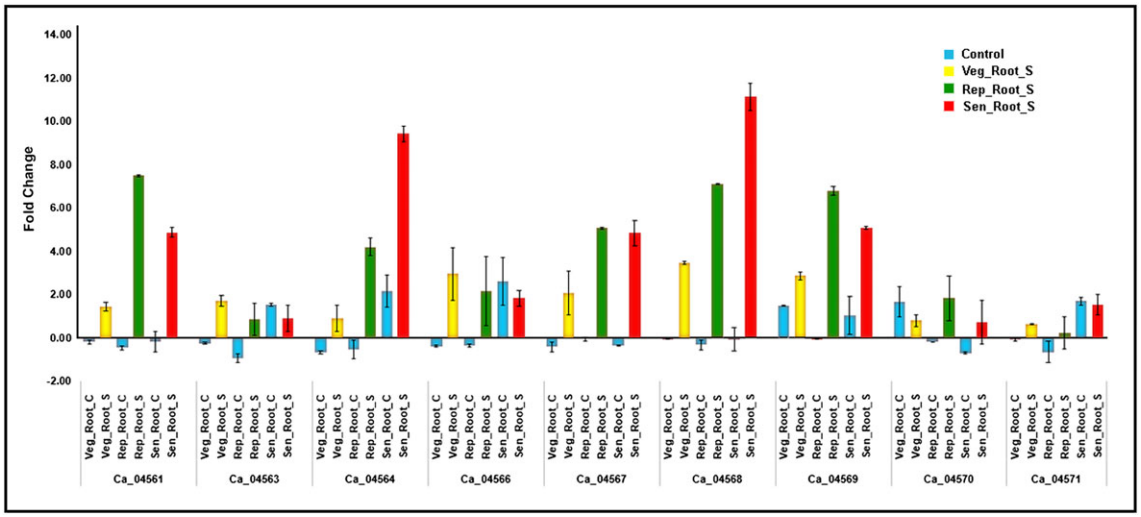

FIGURE 9 Validation of gene expression patterns of 12 candidate genes from the "*QTL-hotspot*" region in control and stressed root tissues of ICC 4958 at vegetative, reproductive, and senescence stages using quantitative real time-polymerase chain reaction [Colour figure can be viewed at wileyonlinelibrary.com]

4.85 fold change) and *Ca_04569* encoding type I inositol 1,4,5-trisphosphate 5-phosphatase CVP2 (2.88, 6.81, and 5.09 fold change), *Ca_04567* encoding serine threonine-kinase HT1-like (2.09, 5.08, and 4.85 fold change), and *Ca_04566* encoding kinase interacting (KIP1-like) family (2.98, 2.19, and 2 fold change). These six genes showed induced expression in vegetative, reproductive, and senescence stages root tissues under drought stress (Figure 9). Among these six, four genes encoding E3 ubiquitin-protein ligase (*Ca_04561*), LRX 2 (*Ca_04564*), kinase interacting (KIP1-like) family (*Ca_04566*), and homocysteine S-methyltransferase (*Ca_04568*) showed similar expression patterns both in the present RNA-Seq analysis and validation using qRT-PCR.

4 | DISCUSSION

Patterns of gene expression provide an excellent opportunity to understand the molecular mechanisms controlling chickpea growth and development. In this study, a comprehensive CaGEA was generated using RNA-Seq approach to study gene expression patterns across different tissues (27) and developmental stages (5) covering entire life cycle of chickpea. This comprehensive dataset provided valuable insights into gene expression and the functional relatedness of genes enhancing the present understanding of various genes involved in regulatory and metabolic process. Expression profiles from these different tissues facilitated (a) mining of gene expression data of genes of interest, (b) identify expression patterns of genes exhibiting in specific tissues, and (c) providing gene expression data on important biological process for legumes such as floral development, seed development, root development, and nodulation process. Advances in genomics research with advent of new tools and technologies facilitated genomics assisted breeding (GAB) in many crops including legumes (Varshney et al., 2013). Deployment of the CaGEA together with the developed genomic resources including chickpea genome sequence in crop improvement programmes would enhance the precision and efficiency in breeding programs in chickpea.

The RNA-Seq data were analysed for global gene expression patterns, pairwise comparisons for identifying DEGs in all possible combinations, gene clustering, and pathway analysis and also to identify key candidates from drought "*QTL-hotspot*" region. The 32,873 genes identified in the current global gene expression analysis is considerably higher than the number of genes reported (28,269) in the chickpea genome assembly by Varshney et al. (2013). This may be due to incomplete (approximately 74%) chickpea genome sequence available. Significantly higher number of genes than the number of genes annotated in the chickpea genome were also observed in earlier studies (Garg et al., 2016). The comparison of the reference-guided assembly with the available chickpea genome led to identification of 4,604 (14%) novel gene loci that are not yet annotated in the chickpea genome. Overall, these results demonstrate the potential use of RNA-Seq approach in discovery of novel genes/transcripts in the sequenced genomes as well. Similarly, novel genes were identified using RNA-Seq in other crops such as maize (Stelpflug et al., 2016) and potato (Shan et al., 2013). Furthermore, these observations confirm that the tissues and developmental stages used in this study triggered a majority of genes in the genome and provided a comprehensive resource of genome-wide gene expression patterns in chickpea. Interestingly, hierarchical clustering of all expressed genes indicated shared transcriptional patterns among spatially related tissues across different developmental stages. This could be due to shared genetic reinforcements and morphological origins (Eveland et al., 2014; Thompson & Hake, 2009; Vollbrecht & Schmidt, 2009). Altogether, the data indicate that different samples exhibit divergent gene expression patterns at different developmental stages of the plant.

The RNA-Seq data were further analysed for DEGs by pairwise combinations in different tissues across five developmental stages. This facilitated identification of gene expression patterns in similar/different tissues in one or more developmental stages of the plant. Several novel gene loci and clusters of genes identified in this study demonstrated and reinforced the utility of RNA-Seq in genomics research. The tissue-specific genes involved in several biological

processes, metabolic and cellular functions were found to be altered in different tissues studied. Most of the genes exhibited tissue or developmental stage-specific response. On the other hand, the highest number of differentially expressed genes was observed in different combinations of seed tissues. These results suggested the higher level of gene expression during seed developmental process at maturity. However, a significant number of differentially expressed genes were also identified in flower tissue followed by roots and nodules. Transcriptome analysis in soybean for seed set and size demonstrated induced expression patterns at early seed developmental stages (Du et al., 2017) supporting the results of present study where induced gene expression patterns were observed during reproductive and senescence stage tissues of the plant. Transcriptome studies conducted in diverse plant species have noted the enrichment of GO terms and metabolic pathways related to growth and development. In the present study, accumulation of genes encoding enzymes involved in growth and development was observed. For instance, the genes involved in various metabolic pathways, cellular process, biological regulation, reproduction, transcription factor activity, cell, cell part, and organelle were significantly differentially expressed. In addition, a significantly large number of genes falling under the subcategory "response to stimulus" were observed. Several studies have reported the regulation of growth and development-related genes under various developmental stages of the plant (Benedito et al., 2008; Pazhamala et al., 2017; Severin et al., 2010; Stelpflug et al., 2016). TFs such as protein kinases have been found to regulate the photosynthetic machinery and associated metabolic pathways during plant development (Gururani, Venkatesh, & Tran, 2015; Saibo, Lourenço, & Oliveira, 2008). In the current dataset, significant number of TFs were identified to be differentially expressed. It has been previously reported that TFs of various gene families perform a crucial function in growth and development via gene regulatory networks (Garg, Bhattacharjee, & Jain, 2015; Shinozaki & Yamaguchi-Shinozaki, 2007). The members of several well-known TF families implicated in growth and development such as bZIP, GRF, and MYB were significantly differentially expressed suggesting an important role of plant TFs in growth and development of the plant. The TFs involved in various developmental processes, such as MADS-box, ARF, homeobox, and TCP were also found in the present study. Upregulation/downregulation of these TFs suggests their role in developmental stage-specific regulation of the plant. The role of homeobox TF family in regulation of stage-specific tissues has been suggested in previous studies (Kanrar, Bhattacharya, Arthur, Courtier, & Smith, 2008; Jain, Tyagi, & Khurana, 2008).

Gene cluster analysis revealed coregulated and functionally related genes. Interestingly, gene clusters responsible for floral development, leaf development, root development, nodulation process, and immature and mature seed development were identified. Each of these clusters contain potential genes that govern organogenesis at specific developmental stages of the plant. In addition, genes showing clear demarcation in aerial and underground tissues were also highlighted in the study. Similar approach was successfully demonstrated to uncover genes involved in complex networks such as floral development and seed development in legumes (Du et al., 2017; Pazhamala et al., 2017). Studying cluster of genes involved in growth and developmental process would be more informative rather than studying individual gene because such clusters collectively share a generalized function and are often located close to each other. For example, Cluster 4 in the present study comprises 345 genes and are highly expressed in floral tissues. These include floral development-specific genes such as pollen-specific LRR, MADS-box transcription factor, floral homeotic PMADS 2-like, pollen family allergen, and AP2 domain class TFs. AP2 domains class TFs play an important role in flower development including ovule development, floral organ initiation and growth, and petal development (Krizek, Prost, & Macias, 2000; Nole-Wilson & Krizek, 2000). Similarly, genes in Cluster 2 are highly induced in nodule tissues. This cluster consisted of genes encoding for early nodulation such as leghemoglobin, nodulation inception, and nodule cysteine-rice secreted peptide. Early nodulation such as transcripts were reported to be involved in nodulation events of *Medicago* (Lohar et al., 2006). Similarly, major increase in expression patterns of leghemoglobin gene family was observed in immature N-fixing nodules. These genes were predicted to regulate other genes in the network (Moreau et al., 2011). Similarly, overall analysis of RNA-Seq data revealed a complex transcriptional network governing growth and development in chickpea. Extensive number of genes clustered together were observed in leaf tissues, stem, root, and immature and mature seed tissues at different developmental stages of the plant.

Pathways involved in biosynthesis, metabolism, and cell signalling play a crucial role in plant growth and development. Enzymes involved in alanine, aspartate, and glutamate metabolism pathway identified in the present study play a central role in N metabolism. Similar results were observed in *Arabidopsis* and demonstrated the key role of Glutamate in N metabolism (Miyashita & Good, 2008). Cysteine and methionine metabolism regulates plant growth and development and are also involved in seed development process. Methionine is glucogenic and an essential amino acid. It serves as a precursor for the synthesis of cystine and cysteine which are non-essential amino acids. Cystine and cysteine are interconvertible. S-adenosylmethionine is the precursor for the synthesis of plant hormone, ethylene, which regulates plant growth and development. It is reported that a member of this pathway, S-adenosylmethionine, is the precursor for the synthesis of plant hormone and ethylene, which is involved in ripening of fruits (Tan, Li, & Wang, 2013). The acquisition of sulphur by plants has become a major concern for the agriculture due to the decreasing trends of S-emissions from industrial sources and the consequent limitation of inputs from deposition. Cysteine is the first committed molecule in plant metabolism that contains both sulphur and nitrogen, and the regulation of its biosynthesis is of very important for the synthesis of a number of essential metabolites in plant metabolic pathways. Several studies reported that plants adapt inorganic sulphur and metabolize it further to organic sulphur compounds essential for plant growth, development, and stress mitigation (Kopriva et al., 2016). In the present study, sulphur metabolism has been identified as one of the important metabolic pathways triggered involving several enzymes such as nucleotidase, sulphur transferase, synthase, and kinase. Overall, pathway analysis results suggest the involvement of complex transcriptional regulation of various pathways governing many aspects of growth and development of the plant. Integration of these transcriptome data with other omics and genetics data can help in further

selection and pin down the important candidate genes for functional analysis.

In chickpea, drought tolerance mechanism and yield enhancement under drought conditions has been well studied (Kale et al., 2015; Varshney et al., 2014). A large number of QTLs for several drought tolerance related traits (e.g., root depth, root biomass, and root length density) and yield-related traits (e.g., 100-seed weight, pods per plant, and seeds per pod) were identified by Varshney et al. (2014). In this context, identification and validation of expression patterns of genes involved in root and seed development was considered in the present study. WRKY (*Ca_00420*) and MYB (*Ca_19393*, *Ca_03266*); glutathione S-amino-terminal domain (*Ca_03442*, *Ca_08920*, *Ca_14581*, *Ca_08920*); AP2-domain class TFs (*Ca_06032*, *Ca_06034*, *Ca_21773*); and LRR receptor-like (*Ca_22105*, *Ca_22057*, *Ca_26423*) showed increased expression in root tissues across different developmental stages—Ger_Radicle, Sed_Primary Root, Veg_Root, Rep_Root, Sen_Root, Rep_Nodules, and Sen_Nodules. These genes were identified to be involved in root developmental process. Similar studies involving LRR and AP2-domain class have been reported earlier (see Mai et al., 2014). In addition, 26 candidate genes reported from the refined "*QTL-hotspot*" region responsible for drought tolerance using high density QTL mapping and gene enrichment analysis by Kale et al. (2015) were also considered in the present study to observe expression patterns across different developmental stages. QTLs identified in the earlier study represent root and yield related traits and hence tissues representing flower, bud, seed, and root at different developmental stages of the plant were studied. A subset of 12 genes from these 26 genes that are in common from high density QTL mapping and gene enrichment analysis were prioritized (Kale et al., 2015). Among the subset of 12 genes, nine genes showing high differential expression in the present study were validated using qRT-PCR. Four genes, E3 ubiquitin-protein ligase, LRX 2, kinase interacting (KIP1-like) family, and homocysteine S-methyltransferase showed induced expression in root tissues under stress conditions at different developmental stages (vegetative, reproductive, and senescence) thus validating the RNA-Seq results emphasizing their important role in drought tolerance mechanism. Similar expression patterns of genes were evidenced by Kale et al. (2015). The *Ca_04561* gene encoding E3 ubiquitin-protein ligase is known to be important post translational modification to regulate growth and development in all eukaryotes (Pokhilko et al., 2011; Thomann et al., 2005). LRX proteins are known to modify cell wall composition and influence plant growth (Draeger et al., 2015). Earlier studies demonstrated that protein kinases play a crucial role in abiotic stresses such as drought and salinity (Yang et al., 2008). Induced expression and the role of homocysteine S-methyltransferase under drought stress is well studied (Kale et al., 2015; Mohammadi, Moieni, Hiraga, & Komatsu, 2012; Wang, Cai, Xu, Wang, & Dai, 2016). The role of key genes in growth and development identified in the present study such as E3 ubiquitin-protein ligase, LRX, and protein kinases were supported by earlier studies in different plant species such as *Arabidopsis* and rice (Draeger et al., 2015; Pokhilko et al., 2011; Wang et al., 2016).

It has long been understood that cell growth and differentiation in plants are under genetic control. The availability of new tools and technologies allowed biologists to recognize genes that control various aspects of growth and development. In summary, the gene expression atlas provides a robust assessment of gene expression patterns across a wide developmental series of chickpea. The CaGEA facilitated identification of candidate genes of agronomic importance for possible deployment in chickpea breeding programme. This has been demonstrated by identification of key candidates involved in floral development, seed development, root development, and nodulation process providing potential genes for future studies. Furthermore, this resource could also be utilized to identify and validate drought tolerance responsive genes crucial for GAB in chickpea. Thus, this could be a valuable resource for basic and applied research for crop improvement in chickpea and other food legumes to understand key genes and genetic regulatory mechanisms involved in growth and development.

ACKNOWLEDGEMENTS

The authors would like to thank Mr. Jaipal Goud and Poornima B for extending their support in this work. Thanks are due to Genotypic Technology Pvt Ltd for generating partial sequence data in the present study. Authors are grateful to the Bill & Melinda Gates Foundation for supporting this study. VG acknowledges the award of research fellowship from Department of Science and Technology, Government of India. This work has been undertaken as part of the CGIAR Research Program on Grain Legumes & Dryland Cereals. ICRISAT is a member of CGIAR Consortium.

ORCID

Rajeev K. Varshney http://orcid.org/0000-0002-4562-9131

REFERENCES

Agarwal, G., Garg, V., Kudapa, H., Doddamani, D., Pazhamala, L. T., Khan, A. W., … Varshney, R. K. (2016). Genome-wide dissection of AP2/ERF and HSP90 gene families in five legumes and expression profiles in chickpea and pigeonpea. *Plant Biotechnology Journal*, *14*, 1563–1577.

Alves-Carvalho, S., Aubert, G., Carrère, S., Cruaud, C., Brochot, A. L., Jacquin, F., … Burstin, J. (2015). Full-length de novo assembly of RNA-Seq data in pea (*Pisum sativum* L.) provides a gene expression atlas and gives insights into root nodulation in this species. *The Plant Journal*, *84*, 1–19.

Arumuganathan, K., & Earle, E. D. (1991). Nuclear DNA content of some important plant species. *Plant Molecular Biology Reporter*, *9*, 208–218.

Benedito, V. A., Torres-Jerez, I., Murray, J. D., Andriankaja, A., Allen, S., Kakar, K., … Moreau, S. (2008). A gene expression atlas of the model legume *Medicago truncatula*. *The Plant Journal*, *55*, 504–513.

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*, 2114–2120.

Clevenger, J., Chu, Y., Scheffler, B., & Ozias-Akins, P. (2016). A developmental transcriptome map for allotetraploid *Arachis hypogaea*. *Frontiers in Plant Science*, *7*, 1446.

Conesa, A., Götz, S., García-Gómez, J. M., Terol, J., Talón, M., & Robles, M. (2005). Blast2GO: A universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics*, *21*, 3674–3676.

Draeger, C., Fabrice, T. N., Gineau, E., Mouille, G., Kuhn, B., Moller, I., … Ringli, C. (2015). Arabidopsis leucine-rich repeat extensin (LRX) proteins modify cell wall composition and influence plant growth. *BMC Plant Biology*, *15*, 155.

Du, J., Wang, S., He, C., Zhou, B., Ruan, Y. L., & Shou, H. (2017). Identification of regulatory networks and hub genes controlling soybean seed

set and size using RNA sequencing analysis. *Journal of Experimental Botany*, *68*, 1955–1972.

Eveland, A. L., Goldshmidt, A., Pautler, M., Morohashi, K., Liseron-Monfils, C., Lewis, M. W., … Olson, A. (2014). Regulatory modules controlling maize inflorescence architecture. *Genome Research*, *24*, 431–443.

Garg, R., Bhattacharjee, A., & Jain, M. (2015). Genome-scale transcriptomic insights into molecular aspects of abiotic stress responses in chickpea. *Plant Molecular Biology Reporter*, *33*, 388–400.

Garg, R., Patel, R. K., Tyagi, A. K., & Jain, M. (2011). De novo assembly of chickpea transcriptome using short reads for gene discovery and marker identification. *DNA Research*, *18*, 53–63.

Garg, R., Shankar, R., Thakkar, B., Kudapa, H., Krishnamurthy, L., Mantri, N., & Jain, M. (2016). Transcriptome analyses reveal genotype-and developmental stage-specific molecular responses to drought and salinity stresses in chickpea. *Scientific Reports*, *6*, 19228.

Graham, P. H., & Vance, C. P. (2003). Legumes: Importance and constraints to greater use. *Plant Physiology*, *131*, 872–877.

Gururani, M. A., Venkatesh, J., & Tran, L. S. P. (2015). Regulation of photosynthesis during abiotic stress-induced photoinhibition. *Molecular Plant*, *8*, 1304–1320.

Hiremath, P. J., Farmer, A., Cannon, S. B., Woodward, J., Kudapa, H., Tuteja, R., … Varshney, R. K. (2011). Large-scale transcriptome analysis in chickpea (*Cicer arietinum* L.), an orphan legume crop of the semi-arid tropics of Asia and Africa. *Plant Biotechnology Journal*, *9*, 922–931.

Jain, M., Misra, G., Patel, R. K., Priya, P., Jhanwar, S., Khan, A. W., … Chattopadhyay, D. (2013). A draft genome sequence of the pulse crop chickpea (*Cicer arietinum* L.). *The Plant Journal*, *74*, 715–729.

Jain, M., Tyagi, A. K., & Khurana, J. P. (2008). Genome-wide identification, classification, evolutionary expansion and expression analyses of homeobox genes in rice. *The FEBS Journal*, *275*, 2845–2861.

Jin, J. P., Zhang, H., Kong, L., Gao, G., & Luo, J. C. (2014). PlantTFDB 3.0: A portal for the functional and evolutionary study of plant transcription factors. *Nucleic Acids Research*, *42*, 1182–1187.

Jukanti, A. K., Gaur, P. M., Gowda, C. L. L., & Chibbar, R. N. (2012). Nutritional quality and health benefits of chickpea (*Cicer arietinum* L.): A review. *British Journal of Nutrition*, *108*, 11–26.

Kaashyap, M., Ford, R., Kudapa, H., Jain, M., Edwards, D., Varshney, R. K., & Mantri, N. (2018). Differential regulation of genes involved in root morphogenesis and cell wall modification is associated with salinity tolerance in chickpea. *Scientific Reports*, *19*, 4855.

Kale, S. M., Jaganathan, D., Ruperao, P., Chen, C., Punna, R., Kudapa, H., … Varshney, R. K. (2015). Prioritization of candidate genes in "*QTL-hotspot*" region for drought tolerance in chickpea (*Cicer arietinum* L.). *Scientific Reports*, *5*, 15296.

Kanrar, S., Bhattacharya, M., Arthur, B., Courtier, J., & Smith, H. (2008). Regulatory networks that function to specify flower meristems require the function of homeobox genes PENNYWISE and POUND-FOOLISH in Arabidopsis. *The Plant Journal*, *54*, 924–937.

Katta, M. A., Khan, A. W., Doddamani, D., Thudi, M., & Varshney, R. K. (2015). NGS-QCbox and raspberry for parallel, automated and rapid quality control analysis of large-scale next generation sequencing (Illumina) data. *PLoS one*, *10*, e0139868.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., & Salzberg, S. L. (2013). TopHat2: Accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biology*, *14*, 36.

Kopriva, S., Talukdar, D., Takahashi, H., Hell, R., Sirko, A., D'Souza, S. F., & Talukdar, T. (2016). Frontiers of sulfur metabolism in plant growth, development, and stress response. *Frontiers in Plant Science*, *6*, 1220.

Krizek, B. A., Prost, V., & Macias, A. (2000). AINTEGUMENTA promotes petal identity and acts as a negative regulator of AGAMOUS. *The Plant Cell*, *12*, 1357–1366.

Kudapa, H., Azam, S., Sharpe, A. G., Taran, B., Li, R., Deonovic, B., … Varshney, R. K. (2014). Comprehensive transcriptome assembly of chickpea (*Cicer arietinum* L.) using Sanger and next generation sequencing platforms: development and applications. *PLoS One*, *9*, e86039.

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

Livak, K. J., & Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(−ΔΔCT) method. *Methods*, *25*, 402–408.

Lohar, D. P., Sharopova, N., Endre, G., Penuela, S., Samac, D., Town, C., … VandenBosch, K. A. (2006). Transcript analysis of early nodulation events in *Medicago truncatula*. *Plant Physiology*, *140*, 221–234.

Mai, C. D., Phung, N. T., To H.T, Gonin, M., Hoang, G. T., Nguyen, K. L., … Gantet, P. (2014). Genes controlling root development in rice. *Rice*, *7*, 30.

Mir, R. R., Kudapa, H., Srikanth, S., Saxena, R. K., Sharma, A., Azam, S., … Varshney, R. K. (2014). Candidate gene analysis for determinacy in pigeonpea (*Cajanus* spp.). *Theoretical and Applied Genetics*, *127*, 2663–2678.

Miyashita, Y., & Good, A. G. (2008). NAD (H)-dependent glutamate dehydrogenase is essential for the survival of *Arabidopsis thaliana* during dark-induced carbon starvation. *Journal of Experimental Botany*, *59*, 667–680.

Mohammadi, P. P., Moieni, A., Hiraga, S., & Komatsu, S. (2012). Organ-specific proteomic analysis of drought-stressed soybean seedlings. *Journal of Proteomics*, *75*, 1906–1923.

Moreau, S., Verdenaud, M., Ott, T., Letort, S., De Billy, F., Niebel, A., … Gamas, P. (2011). Transcription reprogramming during root nodule development in *Medicago truncatula*. *PLoS One*, *6*, e16463.

Nayak, S. N., Zhu, H., Varghese, N., Datta, S., Choi, H. K., Horres, R., … Hoisington, D. A. (2010). Integration of novel SSR and gene-based SNP marker loci in the chickpea genetic map and establishment of new anchor points with *Medicago truncatula* genome. *Theoretical and Applied Genetics*, *120*, 1415–1441.

Nole-Wilson, S., & Krizek, B. A. (2000). DNA binding properties of the Arabidopsis floral development protein AINTEGUMENTA. *Nucleic Acids Research*, *28*, 4076–4082.

O'Rourke, J. A., Iniguez, L. P., Fu, F., Bucciarelli, B., Miller, S. S., Jackson, S. A., … Hernandez, G. (2014). An RNA-Seq based gene expression atlas of the common bean. *BMC Genomics*, *15*, 1.

Pazhamala, L. T., Shilp, S., Saxena, R. K., Garg, V., Krishnamurthy, L., Verdier, J., & Varshney, R. K. (2017). Gene expression atlas of pigeonpea and its application to gain insights into genes associated with pollen fertility implicated in seed formation. *Journal of Experimental Botany*, *68*, 2037–2054.

Pokhilko, A., Ramos, J. A., Holtan, H., Maszle, D. R., Khanna, R., & Millar, A. J. (2011). Ubiquitin ligase switch in plant photomorphogenesis: A hypothesis. *Journal of Theoretical Biology*, *270*, 31–41.

Pushpavalli, R., Krishnamurthy, L., Thudi, M., Gaur, P. M., Rao, M. V., Siddique, K. H., … Vadez, V. (2015). Two key genomic regions harbour QTLs for salinity tolerance in ICCV 2×JG 11 derived chickpea (*Cicer arietinum* L.) recombinant inbred lines. *BMC Plant Biology*, *15*, 124.

Ray, J. D., & Sinclair, T. R. (1998). The effect of pot size on growth and transpiration of maize and soybean during water deficit stress. *Journal of Experimental Botany*, *49*, 1381–1386.

Rozen, S., & Skaletsky, H. (2000). Primer3 on the WWW for general users and for biologist programmers. In *Bioinformatics methods and protocols* (pp. 365–386). Totowa, NJ: Humana Press.

Sabbavarapu, M. M., Sharma, M., Chamarthi, S. K., Swapna, N., Rathore, A., Thudi, M., … Varshney, R. K. (2013). Molecular mapping of QTLs for resistance to Fusarium wilt (race 1) and Ascochyta blight in chickpea (*Cicer arietinum* L.). *Euphytica*, *193*, 121–133.

Saibo, N. J., Lourenço, T., & Oliveira, M. M. (2008). Transcription factors and regulation of photosynthetic and related metabolism under environmental stresses. *Annals of Botany*, *103*, 609–623.

Severin, A. J., Woody, J. L., Bolon, Y. T., Joseph, B., Diers, B. W., Farmer, A. D., … Graham, M. A. (2010). RNA-Seq Atlas of *Glycine max*: A guide to the soybean transcriptome. *BMC Plant Biology*, *10*, 1.

Shan, J., Song, W., Zhou, J., Wang, X., Xie, C., Gao, X., … Liu, J. (2013). Transcriptome analysis reveals novel genes potentially involved in photoperiodic tuberization in potato. *Genomics*, *102*, 388–396.

Shinozaki, K., & Yamaguchi-Shinozaki, K. (2007). Gene networks involved in drought stress response and tolerance. *Journal of Experimental Botany*, *58*, 221–227.

Stelpflug, S. C., Sekhon, R. S., Vaillancourt, B., Hirsch, C. N., Buell, C. R., de Leon, N., & Kaeppler, S. M. (2016). An expanded maize gene expression atlas based on RNA sequencing and its use to explore root development. *The Plant Genome*, *9*(1).

Tan, D., Li, T., & Wang, A. (2013). Apple 1-aminocyclopropane-1-carboxylic acid synthase genes, MdACS1 and MdACS3a, are expressed in different systems of ethylene biosynthesis. *Plant Molecular Biology Reporter*, *31*, 204–209.

Thomann, A., Brukhin, V., Dieterle, M., Gheyeselinck, J., Vantard, M., Grossniklaus, U., & Genschik, P. (2005). Arabidopsis CUL3A and CUL3B genes are essential for normal embryogenesis. *The Plant Journal*, *43*, 437–448.

Thompson, B. E., & Hake, S. (2009). Translational biology: From Arabidopsis flowers to grass inflorescence architecture. *Plant Physiology*, *149*, 38–45.

Thudi, M., Bohra, A., Nayak, S. N., Varghese, N., Shah, T. M., Penmetsa, R. V. ...., & Varshney, R. K. (2011). Novel SSR markers from BAC-end sequences, DArT arrays and a comprehensive genetic map with 1,291 marker loci for chickpea (*Cicer arietinum* L.). *PLoS One*, *6*, e27275.

Thudi, M., Upadhyaya, H. D., Rathore, A., Gaur, P. M., Krishnamurthy, L., Roorkiwal, M., … Varshney, R. K. (2014). Genetic dissection of drought and heat tolerance in chickpea through genome-wide and candidate gene-based association mapping approaches. *PLoS One*, *9*, e96758.

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D. R., … Pachter, L. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature Protocols*, *7*, 562–578.

Trapnell, C., Williams, B. A., Pertea, G., Mortazavi, A., Kwan, G., Van Baren, M. J., … Pachter, L. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nature Biotechnology*, *28*, 511–515.

Varshney, R. K., Hiremath, P., Lekha, P., Kashiwagi, J., Balaji, J., Deokar, A. A., … Hoisington, D. A. (2009). A comprehensive resource of drought- and salinity-responsive ESTs for gene discovery and marker development in chickpea (*Cicer arietinum* L.). *BMC Genomics*, *10*, 523.

Varshney, R. K., Song, C., Saxena, R. K., Azam, S., Yu, S., Sharpe, A. G., … Cook, D. R. (2013). Draft genome sequence of chickpea (*Cicer arietinum*) provides a resource for trait improvement. *Nature Biotechnology*, *31*, 240–246.

Varshney, R. K., Thudi, M., Nayak, S. N., Gaur, P. M., Kashiwagi, J., Krishnamurthy, L., … Rathore, A. (2014). Genetic dissection of drought tolerance in chickpea (*Cicer arietinum* L.). *Theoretical and Applied Genetics*, *127*, 445–462.

Verdier, J., Torres-Jerez, I., Wang, M., Andriankaja, A., Allen, S. N., He, J., … Udvardi, M. K. (2013). Establishment of the *Lotus japonicus* Gene Expression Atlas (LjGEA) and its use to explore legume seed maturation. *The Plant Journal*, *74*, 351–362.

Vollbrecht, E., & Schmidt, R. J. (2009). Development of the inflorescences. In *Handbook of maize: Its biology* (pp. 13–40). New York, NY: Springer.

Wang, X., Cai, X., Xu, C., Wang, Q., & Dai, S. (2016). Drought-responsive mechanisms in plant leaves revealed by proteomics. *International Journal of Molecular Sciences*, *17*, 1706.

Yanai, I., Benjamin, H., Shmoish, M., Chalifa-Caspi, V., Shklar, M., Ophir, R., … Shmueli, O. (2005). Genome-wide midrange transcription profiles reveal expression level relationships in human tissue specification. *Bioinformatics*, *21*, 650–659.

Yang, W., Kong, Z., Omo-Ikerodah, E., Xu, W., Li, Q., & Xue, Y. (2008). Calcineurin B-like interacting protein kinase OsCIPK23 functions in pollination and drought stress responses in rice (*Oryza sativa* L.). *Journal of Genetics Genomics*, *35*, 531–543.

Yao, S., Jiang, C., Huang, Z., Torres-Jerez, I., Chang, J., Zhang, H., … Verdier, J. (2016). The *Vigna unguiculata* Gene Expression Atlas (VuGEA) from de novo assembly and quantification of RNA-seq data provides insights into seed maturation mechanisms. *The Plant Journal*, *88*, 318–327.

Young, M. D., Wakefield, M. J., Smyth, G. K., & Oshlack, A. (2010). Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biology*, *11*, R14.

How to cite this article: Kudapa H, Garg V, Chitikineni A, Varshney RK. The RNA-Seq-based high resolution gene expression atlas of chickpea (*Cicer arietinum* L.) reveals dynamic spatio-temporal changes associated with growth and development. *Plant Cell Environ*. 2018;1–17. https://doi.org/10.1111/pce.13210