



# CLAVATA signaling pathway genes modulating flowering time and flower number in chickpea

Udita Basu<sup>1</sup> · Laxmi Narnoliya<sup>1</sup> · Rishi Srivastava<sup>1</sup> · Akash Sharma<sup>1</sup> · Deepak Bajaj<sup>1</sup> · Anurag Daware<sup>1</sup> · Virevol Thakro<sup>1</sup> · Naveen Malik<sup>1</sup> · Hari D. Upadhyaya<sup>2</sup> · Shailesh Tripathi<sup>3</sup> · V. S. Hegde<sup>3</sup> · Akhilesh K. Tyagi<sup>1,4</sup> · Swarup K. Parida<sup>1</sup>

Received: 6 October 2018 / Accepted: 20 March 2019 / Published online: 30 March 2019  
© Springer-Verlag GmbH Germany, part of Springer Nature 2019

## Abstract

**Key message** A combinatorial genomic strategy delineated functionally relevant natural allele of a *CLAVATA* gene and its marker (haplotype)-assisted introgression led to development of the early-flowering chickpea cultivars with high flower number and enhanced yield/productivity.

**Abstract** Unraveling the genetic components involved in *CLAVATA* (*CLV*) signaling is crucial for modulating important shoot apical meristem (SAM) characteristics and ultimately regulating diverse SAM-regulated agromorphological traits in crop plants. A genome-wide scan identified 142 *CLV1*-, 28 *CLV2*- and 6 *CLV3*-like genes, and their comprehensive genomic constitution and phylogenetic relationships were deciphered in chickpea. The QTL/fine mapping and map-based cloning integrated with high-resolution association analysis identified SNP loci from *CaCLV3\_01* gene within a major *CaqDTF1.1/CaqFN1.1* QTL associated with DTF (days to 50% flowering) and FN (flower number) traits in chickpea, which was further ascertained by quantitative expression profiling. Molecular haplotyping of *CaCLV3\_01* gene, expressed specifically in SAM, constituted two major haplotypes that differentiated the early-DTF and high-FN chickpea accessions from late-DTF and low-FN. Enhanced accumulation of transcripts of superior *CaCLV3\_01* gene haplotype and known flowering promoting genes was observed in the corresponding haplotype-introgressed early-DTF and high-FN near-isogenic lines (NILs) with narrow SAM width. The superior haplotype-introgressed NILs exhibited early-flowering, high-FN and enhanced seed yield/productivity without compromising agronomic performance. These delineated molecular signatures can regulate DTF and FN traits through SAM proliferation and differentiation and thereby will be useful for translational genomic study to develop early-flowering cultivars with enhanced yield/productivity.

---

Communicated by Heiko C. Becker.

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s00122-019-03335-y>) contains supplementary material, which is available to authorized users.

---

✉ Swarup K. Parida  
swarup@nipgr.ac.in; swarupdbt@gmail.com

<sup>1</sup> National Institute of Plant Genome Research (NIPGR), Aruna Asaf Ali Marg, New Delhi 110067, India

<sup>2</sup> International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru, Telangana 502324, India

<sup>3</sup> Division of Genetics, Indian Agricultural Research Institute (IARI), New Delhi 110012, India

<sup>4</sup> Department of Plant Molecular Biology, University of Delhi, South Campus, New Delhi 110021, India

## Introduction

The maintenance and functioning of the shoot apical meristem (SAM) have always been a fascinating field for understanding molecular and developmental plant biology. The SAM is the source of meristematic cells which give rise to all of the aerial plant tissues including vegetative and reproductive structures. The WUSCHEL–CLAVATA signaling plays the central role in regulating stem cell proliferation and differentiation in crop plants (Schoof et al. 2000). WUSCHEL (WUS) acts as a homeodomain transcription factor responsible for stem cell identity, whereas the CLAVATA (*CLV*) peptides form the receptor–ligand complex acting upstream in the signaling pathway and restricting the uncontrolled proliferation of stem cell population (Laux et al. 1996). *CLV* mutants are among the first morphological mutants described in *Arabidopsis*, named for their short and

club-shaped siliques. These mutants are also accompanied with enlarged SAM and overproduced floral meristems. The CLV comprises a leucine-rich repeat (LRR) receptor-like kinase (RLK), namely CLV1, receptor-like protein CLV2 and the small signal peptide CLV3 which acts as a ligand to the receptor complex (Clark et al. 1997; Jeong et al. 1999).

CLV1 belongs to the LRR-RLK subfamily with 21 LRRs which is the largest group of plant transmembrane RLKs with 223 members documented in *Arabidopsis* (Li and Tax 2013). The extracellular LRR domain is involved in ligand recognition, and LRRs 6–8 of CLV1-type receptors are responsible for binding CLE peptides (Shinohara et al. 2012). The LRR domain is followed by a transmembrane domain and an intracellular kinase domain. The basic difference between CLV1 and CLV2 lies in the absence of kinase domain in CLV2. Both CLV1 and CLV2 are expressed in the L3 layer of the SAM and form a heterodimer in which CLV1 is stabilized by CLV2 (Trotochaud et al. 1999). CLV1 homodimers act in parallel to CLV2-CORYNE heterodimer, where CLV2 with its extracellular domain interacts with CORYNE which has intracellular kinase domain to transduce CLV3 signal (Muller et al. 2008). CLV3 belongs to the larger gene family of *CLAVATA 3/EMBRYO SURROUNDING REGION* (CLE)-related genes which contain a conserved 12–13-amino acid C-terminal CLE domain (Stahl et al. 2009). CLV3 encodes a 96-amino-acid-long prepropeptide with an 18-amino acid signal sequence that targets it to the extracellular space (Cock and McCormick 2001). The CLV3 signaling is conserved widely throughout the plant kingdom where a feedback loop between WUS and CLV3 stabilizes the number of stem cells in the SAM.

The *CLV* genes play a key role in regulating the population of stem cells in SAM and thereby its activity. Therefore, the role of *CLV* genes in controlling SAM-mediated regulation of diverse vital agronomic traits is well established in many crop plants. The *CLV* genes generally promote the progression of meristem cells toward organ initiation. The *clv* mutants show a gradual increase in the size of SAM dome due to the accumulation of stem cells which in turn delayed organ initiation. However, the larger SAM size is observed to have significant correlation with early flowering in natural population of maize indicating that accessions with larger SAM undergo vegetative to reproductive phase transition early (Leiboff et al. 2015). It would be thus interesting to understand the role of *CLV* genes in regulating the size of SAM in natural population as well as their effect on flowering time. In a number of plant species like *Arabidopsis*, maize and rice, mutation of *CLV* genes is accompanied by an increase in the number of all floral organs as well as number of flowers (Taguchi-Shiobara et al. 2001; Suzaki 2004; Jiang et al. 2005; Durbak and Tax 2011; Landau et al. 2015). In maize, the grain yield is highly associated with female inflorescence development and mutation in any of

the *CLV* homologs like *thick tassel dwarf1* and *fascinated ear 2 (fea2)* influence inflorescence development (Liu et al. 2012). Variations in receptor *fea2* lead to increased inflorescence meristem size and the kernel row number (Bommert et al. 2013). In *Brassica rapa*, *ml4 (clv* homolog) mutant though has a normal inflorescence and siliques but shows an increase in floral and fruit organ number. In cucumber, the *CsCLV3*, a homolog of the *Arabidopsis* gene *CLV3*, has been identified as a candidate gene for an important quality trait, carpel number through fine genetic mapping in a population and association analysis in natural germplasm accessions (Li et al. 2016). The *CLV* signaling pathway is known to be associated with a number of agronomic traits such as inflorescence architecture, flower production and fruit size in tomato (Xu et al. 2015). Recent studies reported the role of CLV-WUS signaling pathway-mediated stem cell regulation and related morpho-genes in modulation of the overall plant architecture (Whitewoods et al. 2018; Narnoliya et al. 2019). These aforesaid contextual collectively highlighted that the *CLV* genes are especially important SAM-mediated regulators of growth and development and thereby have potential to govern yield and architectural traits in multiple crop plants. In this perspective, it would be interesting to identify and study the structural organization and functional significance of these *CLV* genes at a whole-genome level in any crop plant for molecular genetic dissection of complex agronomic traits. Unfortunately, no such information is available so far in chickpea suggesting the necessity of identifying *CLV* genes and deciphering their genomic organization as well as regulatory functional perspectives at a genome-wide scale for genetic improvement of this important legume crop.

Chickpea (*Cicer arietinum*) is a self-pollinated and diploid legume food crop with a genome size of ~740 Mbp. This legume crop ranks second in consumption and third in production among pulses grown worldwide (Varshney et al. 2013a). The genomes of domesticated *desi* (cv. ICC 4958) and *kabuli* (cv. CDC Frontier) cultivars of chickpea and its related progenitor wild species accession (*C. reticulatum* cv. PI48977) are sequenced in the recent past to decode the diverse genomic features including protein coding genes (Jain et al. 2013; Varshney et al. 2013b; Parween et al. 2015; Gupta et al. 2016). With the availability of said structurally and functionally well-annotated genome sequences, the genome-wide scanning of *CLV* genes and deciphering their genomic constitution seems quite possible in chickpea. Considering the efficacy of an integrated genomics-assisted breeding strategy in complex quantitative trait dissection, its use for rapidly determining the specific role of *CLV* genes in regulating vital agronomic traits appears much promising in chickpea (Kujur et al. 2015a, b; Upadhyaya et al. 2015). To drive these, genome-wide SNPs along with SNP allelic variants mined from the diverse coding and noncoding

sequence components of *CLV* genes by their sequencing-based genotyping among phenotypically well-characterized natural germplasm accessions and mapping population may be proficient to execute an integrated genomic approach in chickpea. These exertions will assist us to identify functionally relevant alleles of *CLV* genes governing traits of agronomic importance for marker-assisted genetic improvement in chickpea.

In view of aforesaid possibilities, the major objective of the present study was to determine the genomic distribution and constitution of *CLV* and *CLV*-like genes annotated from the chickpea genome as well as to understand the macro- and micro-synthetic evolutionary perspectives of chickpea *CLVs* (paralogs/orthologs) through their comparative genome mapping among other legumes. To determine the significance of these genome-wide scanned *CLV* genes for dissection of complex agronomic traits in chickpea, the integrated genomic strategy combining high-resolution QTL/fine mapping and map-based cloning with association analysis, transcript profiling and molecular haplotyping were deployed in the constituted natural association panel (*desi* and *kabuli* germplasm accessions) and mapping population contrasting with flowering time and flower number traits. This enabled us to delineate the most promising molecular signatures (natural allelic variants and haplotypes) of *CLV* signaling pathway genes regulating flowering time as well as flower number traits through modulation of SAM morphometric characteristics for expediting genetic enhancement of chickpea. The near-isogenic lines (NILs) introgressed with superior haplotype of *CLV* gene through marker (haplotype)-assisted selection have significance to develop high-flower-number and early-flowering cultivars with enhanced yield and productivity without compromising their agronomic performance.

## Materials and methods

### Genome-wide identification and genomic constitution of *CLV* genes

The *kabuli* chickpea genome annotation information was acquired from GigaDB Dataset (<http://gigadb.org/dataset/100076>, Varshney et al. 2013b). The amino acid sequences encoded by all three *CLV* genes, namely *CLV1*, *CLV2* and *CLV3* in plant kingdom were retrieved from NCBI (<https://www.ncbi.nlm.nih.gov>) using ‘CLAVATA1,’ ‘CLAVATA2’ and ‘CLE/ESR’ as the query, respectively, on February 2, 2018. The amino acid sequences of all these *CLVs* were used to constitute a hidden Markov model (HMM). These HMM profile of *CLVs* were searched against the proteome sequences of *kabuli* chickpea via HMMER (<http://hmmerr.org>) taking an e-value cutoff of  $1e-05$ . The amino acid

sequences were further analyzed in INTERPRO (<https://www.ebi.ac.uk/interpro>) to assure the presence of domains. The organization of 14-amino-acid-long CLE motif in the *CLV3* was studied using MEME Web server (<http://meme-suite.org>). The genomic distribution of *CLV* genes across eight chromosomes of chickpea was determined in accordance with their physical positions (bp) and annotation on the pseudomolecules and unanchored scaffolds of *kabuli* genome and subsequently visualized through MapChart 2.2 (Voorrips 2002).

### Identification of paralogous *CLVs*

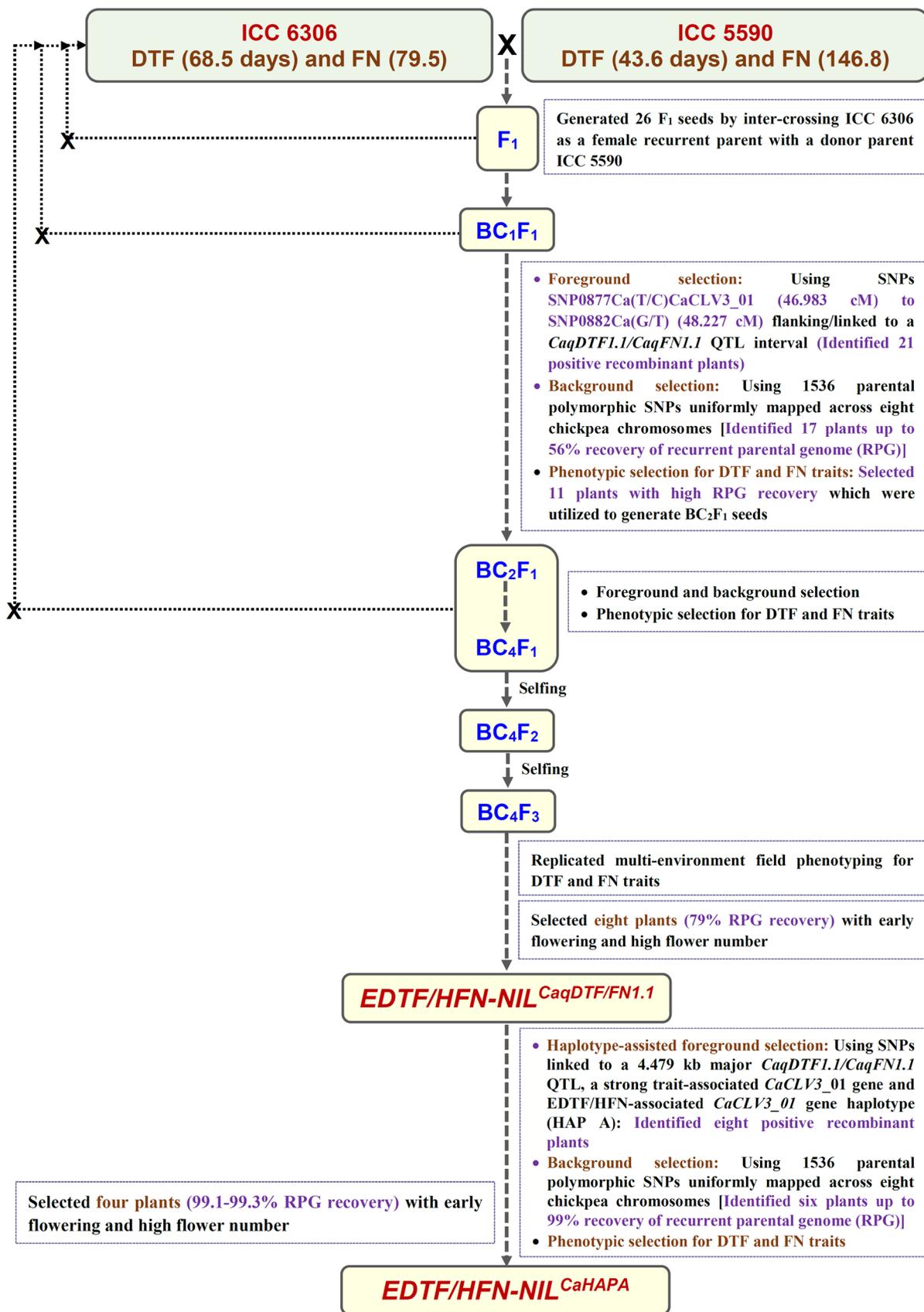
The amino acid sequences of *CLV* genes were searched against the *kabuli* chickpea protein sequence (<http://gigadb.org/dataset/100076>) using BLASTP (<ftp://ftp.ncbi.nlm.nih.gov/blast/executables/blast+/LATEST>) taking query coverage cutoff of 50% and percentage identity of > 70%. To ascertain the paralogy of *CLV* genes, reciprocal BLAST homology search was performed following Jiang et al. (2013) and Muthamilarasan et al. (2016).

### Phylogeny of identified orthologous *CLVs*

The protein sequence information of a selected model dicot plant species, *Arabidopsis thaliana* (<https://www.arabidopsis.org/>), as well as other legume crop plants, *Arachis duranensis* and *Arachis ipaensis* ([https://peanutbase.org/data\\_usage\\_agree\\_aradu](https://peanutbase.org/data_usage_agree_aradu)), *Cajanus cajan* (<http://gigadb.org/dataset/100028>), *Cicer arietinum* (*desi*) (<http://nipgr.res.in/CGAP/>), *Glycine max* (<https://phytozome.jgi.doe.gov/pz/portal.html>), *Lotus japonicus* ([www.kazusa.or.jp/Lotus/](http://www.kazusa.or.jp/Lotus/)), *Lupinus angustifolius*, *Medicago truncatula*, *Phaseolus vulgaris*, *Trifolium pratense*, *Vigna angularis* and *Vigna radiata* (<https://legumeinfo.org/data/public/>) for which the whole genomes have been decoded so far, were acquired from respective databases. These amino acid sequences were used for reciprocal BLAST search against aforesaid HMM profile created for *CLV* genes individually to identify different *CLVs* from each of these crop plant databases and further confirmed their signature domains by INTERPRO. Multiple sequence alignment of the amino acid sequences was performed using BLOSSUM62 scoring matrix interface of CLUSTALW, and an unrooted phylogenetic tree was constructed by a neighbor-joining (NJ) method of MEGA7 ([www.megasoftware.net/](http://www.megasoftware.net/)).

### Development of genetic resource and their phenotyping

We constituted an association panel by screening 92 accessions including 38 *desi* and 54 *kabuli* from core germplasm representing ecogeographical diversity of 20 countries



**Fig. 1** Schematic illustration of the comprehensive strategies followed to develop early-flowering and high-flower-number *CaqDTF/FNI.1* QTL- and *CaCLV3\_01* gene haplotype-introgressed near-isogenic lines (EDTF/HFN-NIL<sup>*CaqDTF/FNI.1*</sup> and EDTF/HFN-NIL<sup>*CaHAPA*</sup>) of chickpea by marker (haplotype)-assisted foreground and background selection. Similar strategy was adopted to develop late-flowering and low-flower-number near-isogenic lines (LDTF/LFN-NIL<sup>*CaqDTF/FNI.1*</sup> and LDTF/LFN-NIL<sup>*CaHAPB*</sup>) of chickpea. E/LDTF: early/late-days to 50% flowering, H/LFN: high/low-flower number and NIL: near-isogenic line

worldwide (Table S1) (Kujur et al. 2015a). Two cultivated *desi* accessions, ICC 5590 and ICC 6306 from association panel with contrasting flowering time and flower number traits, were selected to constitute a F<sub>7</sub> RIL mapping population (ICC 5590 × ICC 6306) consisting of 279 mapping individuals. ICC 5590, a traditional cultivar (landrace) originated from India, is an early-flowering and high-flower-number (FN) chickpea accession with days to 50% flowering (DTF) and FN of 43.6 days and 146.8, respectively. ICC 6306, a landrace, originated from Russian Federation, is a late-flowering and low-flower-number chickpea accession with DTF and FN of 68.5 days and 79.5, respectively. The association panel and RIL mapping population were grown and phenotyped for two successive years (2013 and 2014) in the experimental field as per RCBD (randomized complete block design) with two replications following the standard agronomic practices. Accordingly, these were advanced and phenotyped at two geographical regions, Hyderabad (latitude 17°3' N/longitude 77°2' E from October to February) and New Delhi (28°4' N/77°2' E from November to March) of India. Moreover, the association panel and mapping population were grown in the greenhouse at 22 °C/12 h and 16 °C/12 h as per Ridge et al. (2017). To evaluate the flowering time response with respect to different photoperiods, these accessions and RILs were subjected to both long (16-h light/8-h dark)- and short (10-h light/14-h dark)-day conditions in the greenhouse as per Upadhyaya et al. (2015) and Ridge et al. (2017). The DTF of 10–15 representative plants each from accession and RIL was calculated in accordance with Upadhyaya et al. (2015). The flower number (FN) of 10–12 representative plants each from accession and RIL was measured by counting the average number of fully opened mature flowers per plant. The genetic inheritance characteristics such as frequency distribution, coefficient of variation (CV) and broad-sense heritability (H<sup>2</sup>) of DTF and FN traits among accessions and RILs were determined as per Bajaj et al. (2015a).

### Large-scale SNP mining and genotyping

Genome-wide high-quality SNPs along with their structural and functional annotation information were obtained by genotyping-by-sequencing of 92 *desi* and *kabuli* chickpea

accessions (association panel) using GBS assay (Kujur et al. 2015c). The chickpea *CLV* genes scanned at a whole-genome level were sequenced using the genomic DNA of 92 accessions following the multiplexed amplicon resequencing strategy (TruSeq Custom Amplicon v1.5, Illumina, USA). The 2 kb upstream regulatory regions (URRs), exons/CDS, introns and 2 kb downstream regulatory regions (DRRs) of *CLV* genes were sequenced by Illumina MiSeq NGS assay following Saxena et al. (2014a). The generated sequences were mapped on the *kabuli* reference genome, and subsequently, high-quality *CLV* gene-derived SNPs were mined among accessions as per Saxena et al. (2014a, b) and Malik et al. (2016). The SNPs mined from *CLV* genes were structurally annotated in different coding and noncoding sequence components of genes as per Kujur et al. (2015a, b, c).

### QTL mapping

The selected genome-wide SNPs along with *CLV* gene-derived SNPs exhibiting polymorphism between two parental accessions (ICC 5590 and ICC 6306) were further genotyped using the genomic DNA of 279 mapping individuals from a RIL population (ICC 5590 × ICC 6306) by a Sequenom MALDI-TOF MassARRAY assay (<http://www.sequenom.com>) as per Saxena et al. (2014a, b). The SNP genotyping information was analyzed with Kosambi mapping function of JoinMap 4.1 ([www.kyazma.nl/index.php/mc.JoinMap](http://www.kyazma.nl/index.php/mc.JoinMap)) to measure the linkage among SNPs at a higher LOD (logarithm of odds) threshold ( $\geq 5.0$ ) following Kujur et al. (2016). The SNPs were assigned to defined linkage groups (LG1 to LG8) in accordance with their centiMorgan (cM) genetic distances and respective marker physical positions (bp) on chromosomes. Accordingly, a high-resolution genetic linkage map of chickpea was constructed and visualized by MapChart v2.2. (Voorrips 2002).

For molecular mapping of major QTLs governing DTF and FN traits, the genotyping information of SNPs genetically mapped on a high-density linkage map (eight LGs/chromosomes) as well as DTF and FN trait phenotypic data of RIL mapping individuals and parental accessions was analyzed using a CIM (composite interval mapping) function of MapQTL 6 (Van Ooijen 2009) following Bajaj et al. (2015a) and Kujur et al. (2015b). The LOD cutoff score  $> 5.0$  with 1000 permutation at a  $p < 0.05$  was considered most significant in CIM. Accordingly, the PVE and additive effect specified by each major QTL on traits at a significant LOD were determined as per Bajaj et al. (2015a).

### Fine mapping and map-based cloning of a major QTL

A major DTF/FN QTL region identified by QTL mapping (ICC 5590 × ICC 6306) was targeted for fine mapping by

developing two BC<sub>4</sub>F<sub>3</sub> NILs, one as an early-DTF with high-FN (EDTF/HFN-NIL<sup>CaqDTF/FN1.1</sup>) and another as a late-DTF with low-FN (LDTF/LFN-NIL<sup>CaqDTF/FN1.1</sup>). These NILs were developed by introgressing the said DTF/FN QTL genomic interval from the two parental accessions, ICC 5590 and ICC 6306 (serving as donor and recurrent parents) of a RIL mapping population among each other following the strategies as depicted in Fig. 1. The generated NILs were inter-crossed among each other to develop a F<sub>2</sub> mapping population (EDTF/HFN-NIL<sup>CaqDTF/FN1.1</sup> × LDTF/LFN-NIL<sup>CaqDTF/FN1.1</sup>) consisting of 190 individuals. To scale-down a QTL, SNPs showing parental polymorphism between ICC 5590 and ICC 6306 at this QTL genomic interval were genotyped in 190 mapping individuals of a F<sub>2</sub> population by the Sequenom MALDI-TOF MassARRAY assay (<http://www.sequenom.com>) as per Saxena et al. (2014a, b). Subsequently, the mapping individuals were phenotyped for DTF and FN traits in the field as well as in the greenhouse (both long- and short-day condition). QTL mapping was performed using the SNP genotyping and phenotyping data of a said population following aforementioned strategy.

For QTL region-specific association analysis, a delineated genomic interval underlying a major DTF/FN QTL, identified by high-resolution QTL mapping in NILs, was selected. This target QTL genomic region was sequenced using the genomic DNA of parental accessions and two selected homozygous individuals of a mapping population (EDTF/HFN-NIL<sup>CaqDTF/FN1.1</sup> × LDTF/LFN-NIL<sup>CaqDTF/FN1.1</sup>) contrasting with DTF and FN traits. Sequencing was performed by the multiplexed amplicon sequencing method of TruSeq Custom Amplicon v1.5 in Illumina MiSeq NGS platform as per Malik et al. (2016). Accordingly, the mapping of high-quality amplicon sequence reads onto reference *kabuli* chickpea genome as well as detection of SNPs and their structural/functional annotation were performed. The SNPs mined within the QTL interval were genotyped among 92 *desi* and *kabuli* accessions (association panel) using the MALDI-TOF assay following Saxena et al. (2014a, b). For QTL region-specific association analysis, the SNP genotyping information generated at the target QTL region was integrated with the multi-environments field/greenhouse trait phenotyping data among 92 accessions following the association mapping strategy (Saxena et al. 2014a).

For progeny analysis, the homozygous recombinant and homozygous nonrecombinant individuals derived from the NILs contrasting with DTF and FN traits were selected based on their genetic constitution and considering the recombination among SNPs flanking/tightly linked to a DTF/FN QTL and a strong trait-associated gene. The selected recombinant and nonrecombinant progenies were grown and phenotyped for the traits as per aforesaid methods. The significant variation of traits between selected recombinant and nonrecombinant progenies was evaluated by a statistical *one-tailed t test*.

## Trait association mapping

For association mapping, the population structure, PCA (principal component analysis) and LD (linkage disequilibrium) decay measured among 92 *desi* and *kabuli* chickpea accessions (association panel) were acquired (Kujar et al. 2015a, c). The *CLV* gene-derived SNP genotyping data were integrated with multi-environment field/greenhouse phenotyping information of DTF and FN traits as well as kinship (K), population structure (Q) and PCA (P) information of 92 accessions by mixed model (P + K, K and Q + K)-based CMLM strategy of GAPIT (Lipka et al. 2012) as per Bajaj et al. (2016) and Upadhyaya et al. (2015). We assured the overall authenticity of SNP marker–trait association following the quantile–quantile (Q–Q) plot-based false discovery rate (FDR) (cutoff ≤ 0.05 Benjamini and Hochberg 1995) correction for multiple comparisons between observed/expected  $-\log_{10}(P)$ -value and adjusted *P*-value threshold of significance, estimated in each trait-associated SNP. The SNP loci associated with traits at a lowest FDR adjusted *P*-value (cutoff  $P < 1 \times 10^{-9}$ ) and highest PVE (phenotypic variation explained)/*R*<sup>2</sup> were considered significant for the association mapping study.

## Differential expression profiling

For expression profiling, four *desi* and *kabuli* accessions including RIL mapping parental *desi* accessions and EDTF/HFN and LDTF/LFN gene haplotype-introgressed NILs contrasting with DTF and FN traits were grown in the greenhouse under both long- and short-day conditions following aforementioned strategy. RNA was isolated from the shoot, leaf and SAM tissues (10-day grown plants) as well as diverse flower bud (FB1, FB2, FB3 and FB4) and flower (FL1, FL2, FL3 and FL4) development stages (as classified by Singh et al. 2013) of the said accessions and NILs of chickpea. Two μg of high-quality RNA isolated from the tissues was utilized to synthesize cDNA by Applied Biosystems (ABI, USA) cDNA synthesis kit. The diluted cDNA and 2X Fast SYBR Green Master Mix (ABI) and 200 nM of forward and reverse gene/haplotype-specific primers were amplified in ABI7500 Fast Real-Time (RT)-PCR system. Biological and technical replicates were used for expression profiling in quantitative RT-PCR assay as recommended by Bajaj et al. (2015b). *Actin* was used as an internal control for expression profiling. Significant difference in gene expression was measured and differential expression profile was visualized with a heat map by the MeV (<http://www.tm4.org/mev>).

## Molecular haplotyping

The SNPs discovered from the coding and noncoding (including introns and 2 kb of each URR and DRR) sequence components of a strong DTF and FN trait-associated gene, validated by QTL/fine mapping, map-based cloning and association analysis, were used for constitution of gene haplotypes. For trait association analysis, the genotyping information of SNP haplotypes in the genes, as well as multi-environment field/green house phenotyping information DTF and FN traits evaluated in 92 desi and *kabuli* accessions (association panel), was analyzed using the aforesaid association mapping strategy.

## Development of haplotype-introgressed NILs

To perform marker (haplotype)-assisted foreground selection, the SNPs flanking/tightly linked to a major DTF/FN QTL, a strong trait-associated gene and its EDTF/HFN and LDTF/LFN gene haplotypes were genotyped among individuals of the back-cross mapping population using the MALDI-TOF assay following Saxena et al. (2014a, b). For marker (haplotype)-assisted background selection, 1536 SNPs exhibiting polymorphism between parental accessions (ICC 5590 and ICC 6306) of a RIL mapping population and mapped uniformly across eight chromosomes of *kabuli* genome were genotyped in the selected back-cross mapping individuals (screened by foreground selection) by the Illumina GoldenGate SNP genotyping assay as per Bajaj et al. (2015b) (Fig. 1). The mapping individuals of the back-cross population were phenotyped for DTF and FN traits in the multi-environment field and greenhouse conditions following the aforesaid methods.

## Histological assay

For histology, SAM samples from 10-day grown plants of two RIL mapping parental chickpea accessions (ICC 5590 and ICC 6306) as well as EDTF/HFN and LDTF/LFN gene haplotype-introgressed NILs contrasting with DTF and FN traits were prepared according to Gautam et al. (2016). Paraffin sections (8  $\mu\text{m}$ ) were cut using the rotary microtome (Leica Biosystems) and analyzed under light microscope (LMI Microscopes). The width of the SAM dome among mapping parental accessions and NILs was measured using the ImageJ (<https://imagej.net/>).

## Results

### The CLV genes in chickpea genome

Comprehensive search of HMM profile (constituted from available plant species) and further validation for the

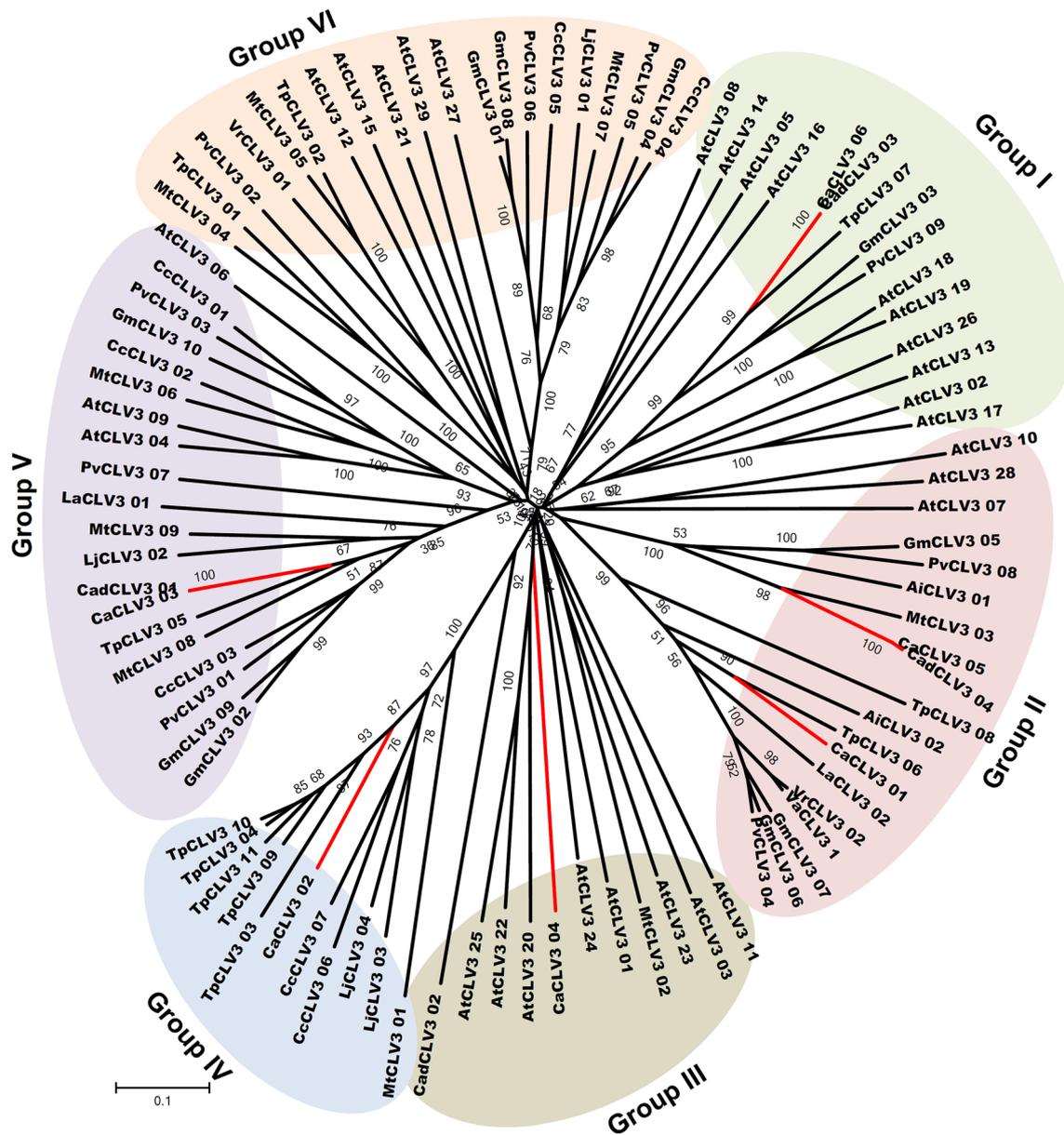
presence of signature domains altogether identified 143 *CLV1*-like genes (*CaCLV1\_001* to *CaCLV1\_143*) as well as 28 *CLV2*-like (*CaCLV2\_01* to *CaCLV2\_28*) and 6 *CLV3*-like (*CaCLV3\_01* to *CaCLV3\_06*) genes. The identities of these *CLV* genes were assigned based on their corresponding physical positions (bp) and structural annotation on chickpea genome (Table S2). Following structural annotation and organization of signature domains in the CDS of *CLV* genes, the detail gene structure of *CLV1*, *CLV2* and *CLV3* was made and illustrated in Figures S1 and S2. Essentially, the *CLV1* genes consist of LRR, TMhelix and Kinase domains, whereas *CLV2* included LRR and TMhelix domains lacking the kinase activity and *CLV3* had only CLE domain. The presence of signal peptide (SP) coding region was observed in all *CLV1*, *CLV2* and *CLV3* genes. Among all six *CLV3* genes identified, *CaCLV3\_02* possesses multiple CLE motifs (RKVPTGPNPAQSP) throughout the gene at regular intervals (Figure S2B). The identified *CLV1* and *CLV2* genes were physically mapped and distributed randomly throughout the eight chickpea chromosome pseudomolecules, while *CLV3* genes were localized solely on the chromosomes 1, 3 and 6 (Figure S3). The number of *CLV3* genes identified in chickpea is quite less than that documented in *Glycine max* (84) at a genome-wide scale (Hastwell et al. 2015). This is expected considering the involvement of ancient and recent whole-genome duplication events for the evolution of *Glycine max* in contrast to single recent whole-genome duplication event in chickpea (Jain et al. 2013; Varshney et al. 2013b).

### The phylogeny and evolutionary significance of CLVs in chickpea

The analysis of CLV protein sequences for possible paralogous relationship among each other exhibited 80–82% intra- and inter-chromosomal paralogy of *CLV1* and *CLV2* (Table S3, Table S4). However, we could not find any paralogous relationship among *CLV3* under an applied condition of reciprocal sequence homology search analysis. The comprehensive analysis of orthologous relationship of chickpea CLV proteins with that of eleven legumes (*Arachis duranensis*, *Arachis ipaensis*, *Cajanus cajan*, *Glycine max*, *Lotus japonicus*, *Lupinus angustifolius*, *Medicago truncatula*, *Phaseolus vulgaris*, *Trifolium pratense*, *Vigna angularis* and *Vigna radiata*) and *Arabidopsis* identified a total of 2981, 287 and 96 *CLV1*, *CLV2* and *CLV3* sequences, respectively (Table S4, Table S5). Interestingly, about 45.4 and 10.7% of the *CLV1* and *CLV2* sequences, respectively, exhibited paralogy within the chickpea genome as well as orthology with other legumes. The use of 67 *CLV3* orthologous sequences among chickpea and legumes in neighbor-joining phylogenetic tree construction grouped these orthologs into six distinct clusters (Fig. 2). Based on this

analysis, CaCLV3 sequences were found to exhibit more close phylogenetic relationship with MtCLV3 and TpCLV3 sequences belonging to *Medicago truncatula* and *Trifolium pratense*, respectively. The clustering and grouping of CLV3 sequences among chickpea and legumes were in close agreement with the genomic constitution of their CLE motifs. The observed outcomes are also consistent with the earlier study in which CLV3 sequences of *Glycine max* are clustered and categorized into seven groups based on their conserved CLE domain (Hastwell et al. 2015). Notably, CaCLV3\_01 and CaCLV3\_05 with similar CLE motif sequences were

clustered together with GmCLV3\_05, GmCLV3\_06 and GmCLV3\_07, respectively, in group II (Fig. 2). All three of these well-characterized GmCLV3 play a critical role in regulation of meristematic cell niche in the roots and shoots of soybean (Grienberger and Fletcher 2015). CaCLV3\_03 contained the conserved CLE motifs similar to GmCLV3\_01, GmCLV3\_02, GmCLV3\_08, GmCLV3\_09 and GmCLV3\_10 and thus were clustered together in group V (Fig. 2). The CLV3 sequences with multiple CLE domains, namely LjCLV3\_03, LjCLV3\_04, MtCLV3\_01, CcCLV3\_06, CcCLV3\_07, TpCLV3\_03, TpCLV3\_04,



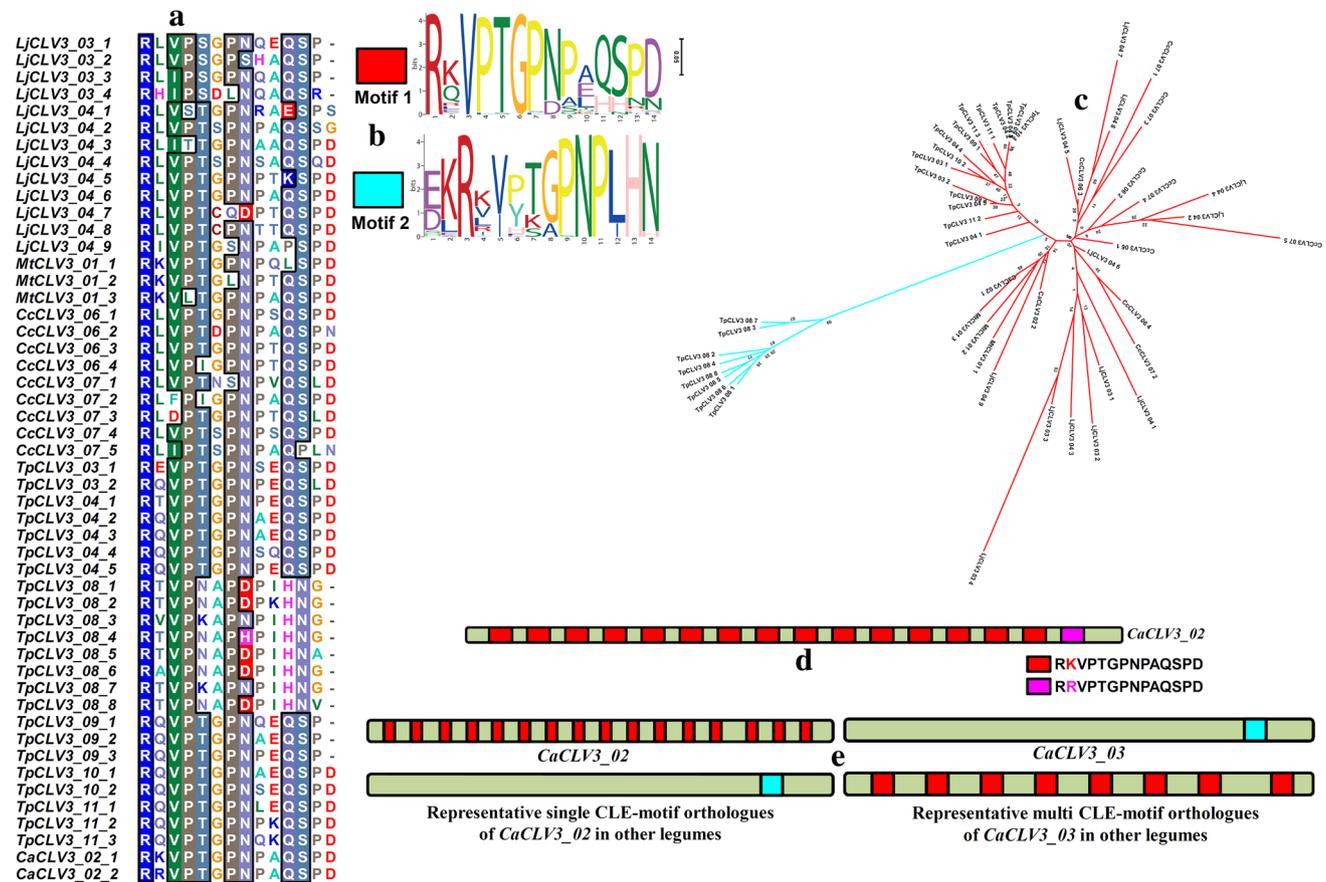
**Fig. 2** Phylogenetic tree depicting the evolutionary relationship among the different CLV3 genes identified in chickpea genome. The CLV3 genes could be grouped into six major groups among which

*CaCLV3\_01* and *CaCLV3\_054* clustered together in Group II. The Group IV comprised of all the CLV3 genes with two or more CLE domains

TpCLV3\_09, TpCLV3\_10, TpCLV3\_11 and CaCLV3\_02, were clustered together in group IV (Fig. 2).

The CLE motifs of these multi-CLE CLV3 sequences were classified and grouped in accordance with sequence context of two CLE variants (Motif 1 and Motif 2) (Fig. 3a, b). Among the multiple-CLE-motifs-containing CLV3 sequences, highest number of 24 CLE motifs was obtained in TpCLV3\_08 followed by 16 CLE motifs in CaCLV3\_02. All other CLV3 sequences had CLE motifs less than 10 in occurrence, usually ranging from 2 to 5 in number. The CLE motif of CaCLV3\_02 contained a repetition of same sequence, RKVPTGPNPAQSP, with a minor change of one amino acid residue at a single place (Fig. 3c). In case of *Trifolium*, TpCLV3\_08 contained eight different variants of CLE motifs, and all others are repetition of one of the

eight variants. The evolutionary close relative of chickpea, *Medicago* CLV3 ortholog MtCLV3\_01 was found to contain three CLE motifs and exhibited utmost similarity with a CLE motif sequence of CaCLV3\_02 (Fig. 3a, b). We observed an expansion of the CLE motif both in chickpea and *Trifolium*. As compared to 2–5 CLE motifs occurrence per gene in other studied legumes, there was an expansion of CLE motifs with 16 and 24 occurrences, respectively, in *Trifolium* and chickpea. The occurrence of tandem CLE domain repeat in CLV3 has been reported earlier in rice and wheat (Oelkers et al. 2008). The chickpea CaCLV3\_02 orthologs consisted of both single-motif and multiple-CLE-motif sequences (Fig. 3d). However, we observed a distinct expansion of CLE motifs in chickpea CLV3 and no other legume orthologs of this protein showed motifs as many as 16.



**Fig. 3** The multi-CLE motif sequences of *CLV3* genes identified across studied legumes including chickpea. **a** Multiple sequence alignment of the different motif sequences in multi-CLE motif *CLV3* like proteins identified. The number alongside the protein name corresponds to the sequence variants of CLE motif in that particular protein. The first amino acid residue is highly conserved among all these sequences. **b** Two different types of motifs identified through MEME denoted by red and blue color, respectively. **c** The neighbor-joining phylogenetic tree of the varied CLE-motif sequences of *CLV3* genes exhibits two distinct clusters. One with the Motif 1 denoted

by red-colored branch and the blue-colored branch represents the Motif 2 sequences belonging to *TpCLV3\_08*. **d** *CaCLV3\_02* with 16 CLE motifs exhibits actually the duplication of the same sequence (Motif 1) with only a substitution of second amino acid residue (K to R) at one position. **e** The orthologs of *CaCLV3\_02* were observed to be possessing Motif 2 instead of Motif 1. Most of the multi-CLE motifs contain Motif 1 (with an exception of *TpCLV3\_08*) that shows homology with *CaCLV3\_03* comprising a single CLE motif sequence (Motif 2) (color figure online)

The *Trifolium* CLV3, TpCLV3\_08, containing CLE motifs with 24 repeats (Motif 2) did not exhibit homology with CaCLV3\_02 CLE motif. The multi-CLE-motif sequences of *Cajanus*, *Lotus*, *Medicago*, and *Trifolium*, except for LjCLV3\_03 and TpCLV3\_08, were found to be homologous with CaCLV3\_03 (Motif 2). However, the CLE motif of these sequences was similar to that of CaCLV3\_02 (Motif 1) (Fig. 3e).

### High-throughput SNP genotyping identifies genome-wide including *CLV* gene-derived SNP alleles in chickpea

The sequencing and genotyping of 92 *desi* and *kabuli* accessions (association panel) simultaneously using a GBS assay altogether detected 14115 high-quality SNPs at a genome-wide scale that were mapped on eight chickpea chromosomes. Additionally, large-scale genotyping of SNPs mined from 138 *CLV1*, 25 *CLV2* and 5 *CLV3* genes was performed among 92 accessions. For this, the coding and noncoding intronic and regulatory regions (including 2 kb of each URR and DRR) of all these *CLV* genes were sequenced and genotyped among 92 accessions through a targeted high-resolution (~110×) multiplex-amplicon resequencing assay. This identified a total of 8262 high-quality SNPs from 168 *CLV* genes with a mean frequency of 49.2 SNPs per gene (Table S6). Of these, 8250 SNPs including 7381 *CLV1*, 702 *CLV2* and 183 *CLV3* gene-derived SNPs were mapped on eight chromosomes, whereas 12 *CLV1* gene-based SNPs were localized on unanchored scaffolds of *kabuli* chickpea genome. These SNPs were structurally annotated in diverse coding (synonymous and non-synonymous) as well as noncoding intronic and regulatory sequence components of *CLV1*, *CLV2* and *CLV3* genes. The *CLV* gene-derived informative SNPs (non-synonymous and regulatory SNPs) differentiating domesticated *desi* and *kabuli* accessions discovered by us have proficiency to expedite genomics-assisted

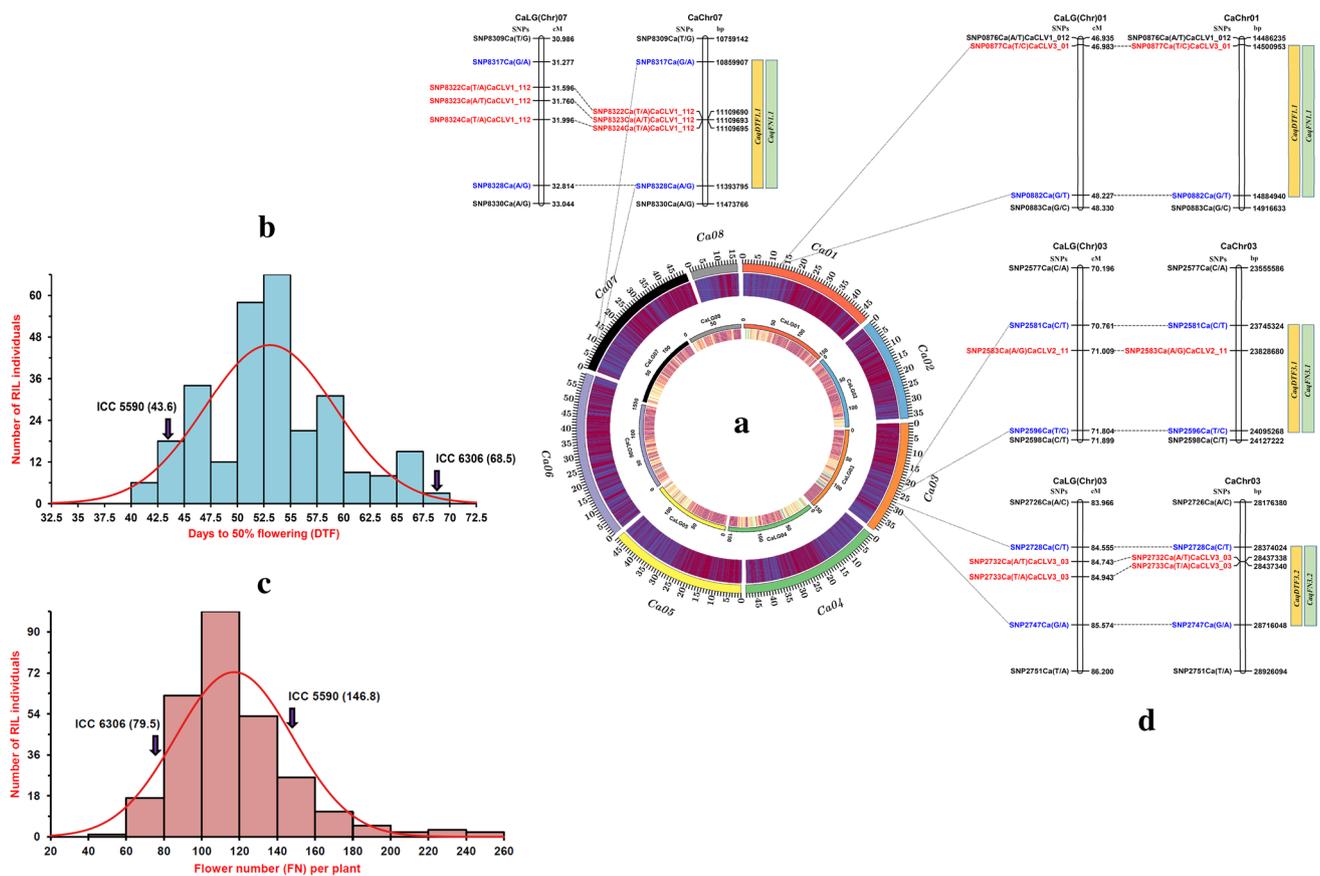
breeding applications including genetic/QTL and association mapping for crop improvement in chickpea.

### Molecular mapping of major QTLs scans potential *CLV* gene-derived SNPs governing DTF and FN traits in chickpea

We screened 9320 and 165 of the total identified 14115 genome-wide and 8262 *CLV* gene-derived SNPs, respectively, exhibiting polymorphism between parental accessions (ICC 5590 and ICC 6306) of a RIL mapping population (ICC 5590×ICC 6306). Accordingly, a high-density genetic linkage map (ICC 5590×ICC 6306) was constructed by integrating 9485 SNPs including 9320 genome-wide as well as 135 *CLV1*, 24 *CLV2* and 6 *CLV3* gene-derived SNPs across eight LGs (LG1-LG8) of chickpea. The genetic linkage map covered a total map length of 1085.982 cM with a map density (mean inter-marker distance) of 0.114 cM (Table 1, Table S7, Fig. 4a). The constructed high-density genetic linkage map can serve as a suitable reference for high-resolution QTL mapping of agronomic traits in chickpea. The generated RIL mapping population (ICC 5590×ICC 6306) was phenotyped in the field at two diverse geographical locations as well as in the green house (long- and short-day photoperiod conditions) for two consequent years. A broader level of significant ( $P < 0.0001$ ) DTF and FN trait variation (8.3–26.9% CV) and high  $H^2$  (80–88%) was observed among 279 mapping individuals and parental accessions across the multiple environments (Table S8, Fig. 4b, c). The normal frequency distribution as well as bidirectional transgressive segregation of DTF and FN traits in RILs underlines a greater efficiency of a generated RIL mapping population (ICC 5590×ICC 6306) for high-resolution QTL mapping study in chickpea (Fig. 4b, c). We observed a significant ( $P < 0.001$ ) positive correlation between DTF and FN traits (31%) in a RIL mapping population.

**Table 1** Genome-wide and *CLV* gene-derived SNPs mapped on eight chromosomes of a chickpea genetic linkage map

| Linkage groups (LGs)/ chromosomes (Chr) | SNPs mapped |             |             |             |       | Map length covered (cM) | Map density [mean inter-marker distance (cM)] |
|---|-------------|-------------|-------------|-------------|-------|-------------------------|---|
|   | Genome wide | <i>CLV1</i> | <i>CLV2</i> | <i>CLV3</i> | Total |                         |   |
| <i>CaLG(Chr)01</i>                      | 1348        | 19          | 4           | 2           | 1373  | 156.406                 | 0.114   |
| <i>CaLG(Chr)02</i>                      | 789         | 16          | 6           | 0           | 811   | 134.448                 | 0.166   |
| <i>CaLG(Chr)03</i>                      | 1147        | 16          | 2           | 2           | 1167  | 119.024                 | 0.102   |
| <i>CaLG(Chr)04</i>                      | 1719        | 14          | 4           | 0           | 1737  | 140.176                 | 0.081   |
| <i>CaLG(Chr)05</i>                      | 1213        | 17          | 1           | 0           | 1231  | 138.992                 | 0.113   |
| <i>CaLG(Chr)06</i>                      | 1470        | 21          | 5           | 2           | 1498  | 177.660                 | 0.119   |
| <i>CaLG(Chr)07</i>                      | 1058        | 17          | 1           | 0           | 1076  | 139.707                 | 0.130   |
| <i>CaLG(Chr)08</i>                      | 576         | 15          | 1           | 0           | 592   | 79.569                  | 0.134   |
| All                                     | 9320        | 135         | 24          | 6           | 9485  | 1085.982                | 0.114   |



**Fig. 4** A Circos ideogram **a** represents the inner high-density genetic map and outer physical map of the SNPs employed in high-resolution QTL mapping. **d** QTLs associated with DTF and FN traits identified and mapped on three LGs (linkage groups)/chromosomes using a RIL mapping population (ICC 5590×ICC 6306). Frequency distribution of mean **b** days to 50% flowering time (DTF) and **c** flower number (FN) trait variation measured among 279 mapping individuals and parental accessions of a F<sub>7</sub> RIL population (ICC 5590×ICC 6306).

Arrows indicate the identities of mapping parental accessions contrasting with DTF and FN traits. The genetic distance (cM) and the physical position (bp) of the associated SNPs are indicated on right side of the LGs/chromosomes, and the identity of the marker loci integrated on the LGs/chromosomes is indicated on the left side of chromosomes. SNPs flanking and tightly linked to four *CLV* genes harboring eight robust QTLs associated strongly with DTF and FN traits are coded with blue and red color (color figure online)

For molecular mapping of major QTLs, the genotyping data of 9485 genetically mapped SNPs and multi-environment DTF and FN trait phenotyping information of RIL mapping individuals and parental accessions were correlated. This detected eight major QTLs (8.1–11.5 LOD) harboring four genomic regions governing both DTF and FN traits that were mapped on three chickpea chromosomes (*Ca01*, *Ca03* and *Ca07*) (Table 2, Fig. 4d). All these eight QTLs were validated successfully across multiple environments in the field and greenhouse (long- and short-day photoperiod conditions). The PVE determined by individual QTL for DTF and FN traits ranged from 18.4 to 27.5% *R*<sup>2</sup>. The PVE measured for all 8 QTLs in combination was 30.6%. All these QTLs demonstrated maximum positive additive effect and thus highest positive interaction of

QTLs/alleles governing both desirable DTF and FN traits derived from an early-DTF and high-FN mapping parental accession ICC 5590 was observed (Table 2). Seven regulatory and intronic SNPs derived from the four candidate *CLV* genes (*CaCLV3\_01*, *CaCLV2\_11*, *CaCLV3\_03* and *CaCLV1\_112*) were found to be linked with the eight major QTLs governing DTF and FN traits in chickpea (Table 2, Table S7). Of these, informative SNPs in a *CaCLV3\_01* gene mapped on the genomic interval (1.244 cM corresponding to 383.987 kb) of a *CaqDTF1.1/CaqFN1.1* QTL explained maximum (22.5–27.5% PVE with 9.8–11.5 LOD) phenotypic variation for both DTF and FN traits observed in a RIL mapping population (Table 2) and thus this major QTL was selected for further fine mapping and map-based cloning study.

**Table 2** Molecular mapping of major DTF and FN QTLs in chickpea

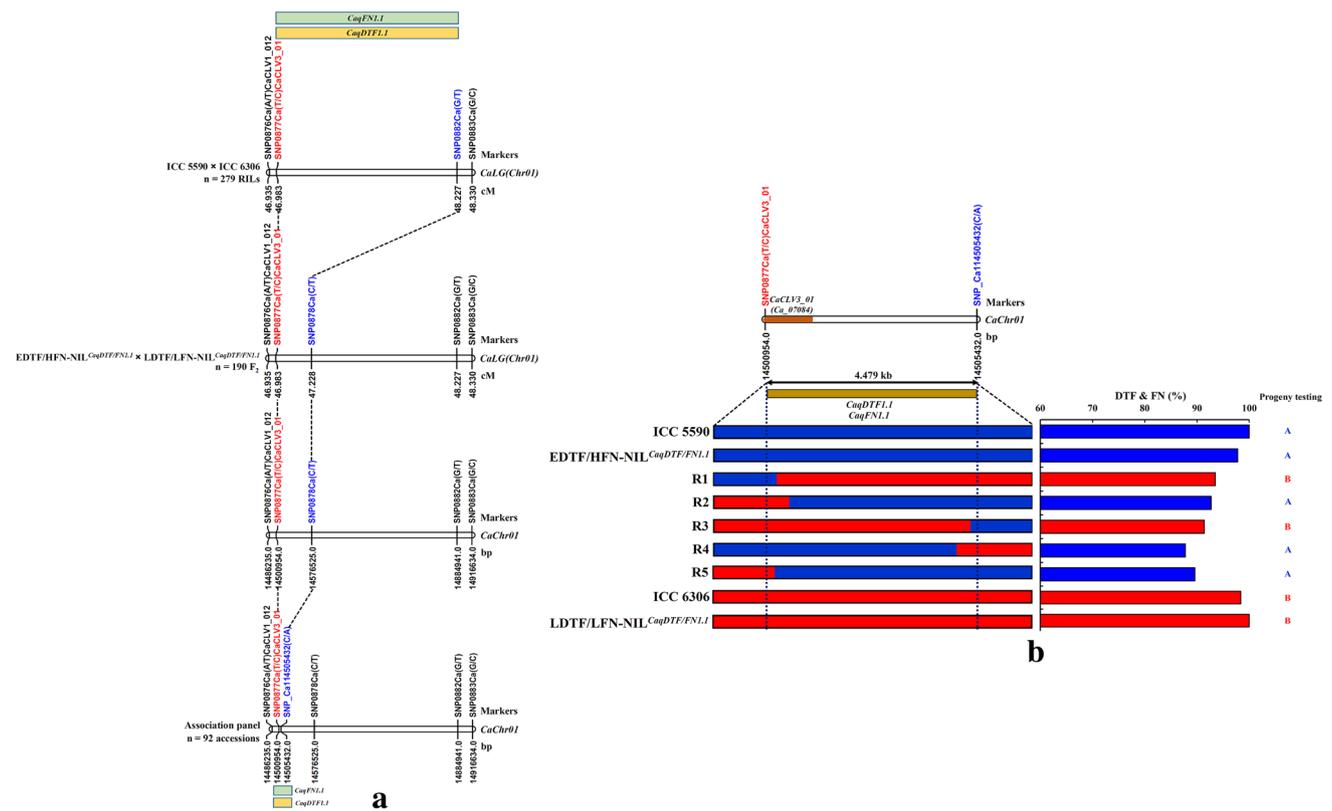
| QTLs <sup>a</sup>                   | LGs/Chromosomes   | Marker intervals with genetic positions (cM)   | QTL physical intervals (bp)  | Markers tightly linked to QTLs with genetic (cM) and physical positions (bp)   | Structural and functional annotation of QTL-linked markers | LOD  | PVE (R <sup>2</sup> %) | A   |
|-------------------------------------|-------------------|--|--|--|--|------|------------------------|-----|
| <i>CaqDTF1.1</i><br><i>CaqFNI.1</i> | <i>CaLG(Chr)1</i> | 1.244: <i>SNP0877Ca(T/C)</i><br><i>CaCLV3_01</i> (46.983) to<br><i>SNP0882Ca(G/T)</i> (48.227) | 383987: <i>SNP0877Ca(T/C)</i><br><i>CaCLV3_01</i> (14500953)<br>to <i>SNP0882Ca(G/T)</i><br>(14884940) | <i>SNP0877Ca(T/C)</i><br>(46.983) and (14500953)   | Intron and <i>CaCLV3_01</i>                                | 10.5 | 23.4                   | 3.2 |
| <i>CaqDTF3.1</i><br><i>CaqFN3.1</i> | <i>CaLG(Chr)3</i> | 1.043: <i>SNP2581Ca(C/T)</i><br>(70.761) to <i>SNP2596Ca(T/C)</i><br>(71.804)                  | 349944: <i>SNP2581Ca(C/T)</i><br>(23745324) to<br><i>SNP2596Ca(T/C)</i> (24095268)                     | <i>SNP2583Ca(A/G)</i> <i>CaCLV2_11</i><br>(71.009) and (23828679)  | URR and <i>CaCLV2_11</i>                                   | 10.7 | 25.4                   | 2.7 |
| <i>CaqDTF3.2</i><br><i>CaqFN3.2</i> | <i>CaLG(Chr)3</i> | 1.019: <i>SNP2728Ca(C/T)</i><br>(84.555) to <i>SNP2747Ca(G/A)</i><br>(85.574)                  | 342023: <i>SNP2728Ca(C/T)</i><br>(28374024) to<br><i>SNP2747Ca(G/A)</i><br>(28716047)                  | <i>SNP2733Ca(T/A)</i> <i>CaCLV3_03</i><br>(84.743) and (28437338)<br><i>SNP2733Ca(T/A)</i> <i>CaCLV3_03</i><br>(84.943) and (28437340)   | URR and <i>CaCLV3_03</i>                                   | 11.5 | 27.5                   | 4.1 |
| <i>CaqDTF7.1</i><br><i>CaqFN7.1</i> | <i>CaLG(Chr)7</i> | 1.537: <i>SNP8317Ca(G/A)</i><br>(31.277) to <i>SNP8328Ca(A/G)</i><br>(32.814)                  | 533888: <i>SNP8317Ca(G/A)</i><br>(10859907) to<br><i>SNP8328Ca(A/G)</i><br>(11393795)                  | <i>SNP8322Ca(T/A)</i> <i>CaCLV1_112</i><br>(31.596) and (11109690)<br><i>SNP8323Ca(A/T)</i> <i>CaCLV1_112</i><br>(31.760) and (11109693)<br><i>SNP8324Ca(T/A)</i> <i>CaCLV1_112</i><br>(31.996) and (11109695) | URR and <i>CaCLV1_112</i>                                  | 8.6  | 19.8                   | 2.1 |
|                                     |                   |  |  |  |  | 8.1  | 18.4                   | 4.3 |

<sup>a</sup>*CaqDTF1.1* (*Cicer arietinum* QTL for days to 50% flowering time on chromosome 1 number 1) and *CaqFNI.1* (*Cicer arietinum* QTL for flower number on chromosome 1 number 1) PVE, proportion of phenotypic variation explained by QTLs; A, additive effect; DTF- and FN-positive additive effect inbred alleles from early-DTF and high-FN mapping parental chickpea accession ICC 5590

### Fine mapping and map-based cloning ascertains the role of a *CaCLV3\_01* gene harboring a major QTL in regulating flowering time and flower number in chickpea

One strong trait-associated *CaqDTF1.1/CaqFN1.1* QTL interval [SNP0877Ca(T/C)CaCLV3\_01 (46.983 cM) to SNP0882Ca(G/T) (48.227 cM)] harboring the *CaCLV3\_01* gene, mapped on a chromosome 1 of an intra-specific RIL genetic linkage map (ICC 5590×ICC 6306), was selected for fine mapping to ascertain the potential of this gene in governing flowering time and flower number in chickpea. For this, a 1.244-cM genomic region corresponding to the said *CaqDTF1.1/CaqFN1.1* QTL interval (383.987 kb) from the recurrent and/or donor parental accessions (ICC 5590 and ICC 6306) was backcrossed four times between each other to produce two BC<sub>4</sub>F<sub>3</sub> NILs, EDTF/HFN-NIL-*CaqDTF/FN1.1* and LDTF/LFN-NIL-*CaqDTF/FN1.1* with about

79% recovery of recurrent parental genome. Using 190 mapping individuals of a F<sub>2</sub> population (EDTF/HFN-NIL-*CaqDTF/FN1.1* × LDTF/LFN-NIL-*CaqDTF/FN1.1*), 23 recombinants were detected between SNP0877Ca(T/C)CaCLV3\_01 (46.983 cM) to SNP0878Ca(C/T) (47.228 cM) SNPs at a 0.245 cM QTL interval (Fig. 5a). This delineated genomic interval underlying *CaqDTF1.1/CaqFN1.1* QTL mapped on chromosome 1 which was defined to 75.572 kb region by integrating its genetic linkage map information with that of physical map of *kabuli* chickpea genome (Fig. 5a). The high-coverage (~ 130×) multiplex amplicon resequencing of a 75.572 kb *CaqDTF1.1/CaqFN1.1* QTL genomic region in parental accessions and two selected EDTF/HFN and LDTF/LFN homozygous individuals of a mapping population (EDTF/HFN-NIL-*CaqDTF/FN1.1* × LDTF/LFN-NIL-*CaqDTF/FN1.1*) altogether detected 480 SNPs (Table S9). The QTL region-specific association analysis at a *CaqDTF1.1/CaqFN1.1* QTL genomic interval by integrating the genotyping



**Fig. 5 a** Map-based cloning of a major *CaqDTF1.1/CaqFN1.1* QTL by fine mapping in a mapping population (EDTF/HFN-NIL-*CaqDTF/FN1.1* × LDTF/LFN-NIL-*CaqDTF/FN1.1*) following QTL region-specific association analysis in 92 *desi* and *kabuli* accessions contrasting with DTF and FN traits. The identity of the markers mapped on the linkage groups (LGs)/chromosomes and their genetic (cM)/physical (bp) positions is indicated above and below the chromosomes, respectively. The SNPs flanking and tightly linked with the QTL and *CaCLV3\_01* gene are indicated with blue and red color, respectively. **b** Progeny testing-based individual phenotyping of five selected

recombinants as well as contrasting NILs and mapping parental accessions (ICC 5590 and ICC 6306) to deduce the genotype of a delineated 4.479 kb major *CaqDTF1.1/CaqFN1.1* QTL in chickpea. A informative SNP, SNP0877Ca(T/C)CaCLV3\_01 in a *CaCLV3\_01* gene exhibiting zero recombination in the selected recombinants is associated strongly with DTF and FN traits in chickpea. The genomic constitution of EDTF/HFN and LDTF/LFN parental accessions/NILs is represented by ‘A’ and ‘B’, respectively. E/LDTF: early/late days to 50% flowering time. H/LFN: high/low-flower number (color figure online)

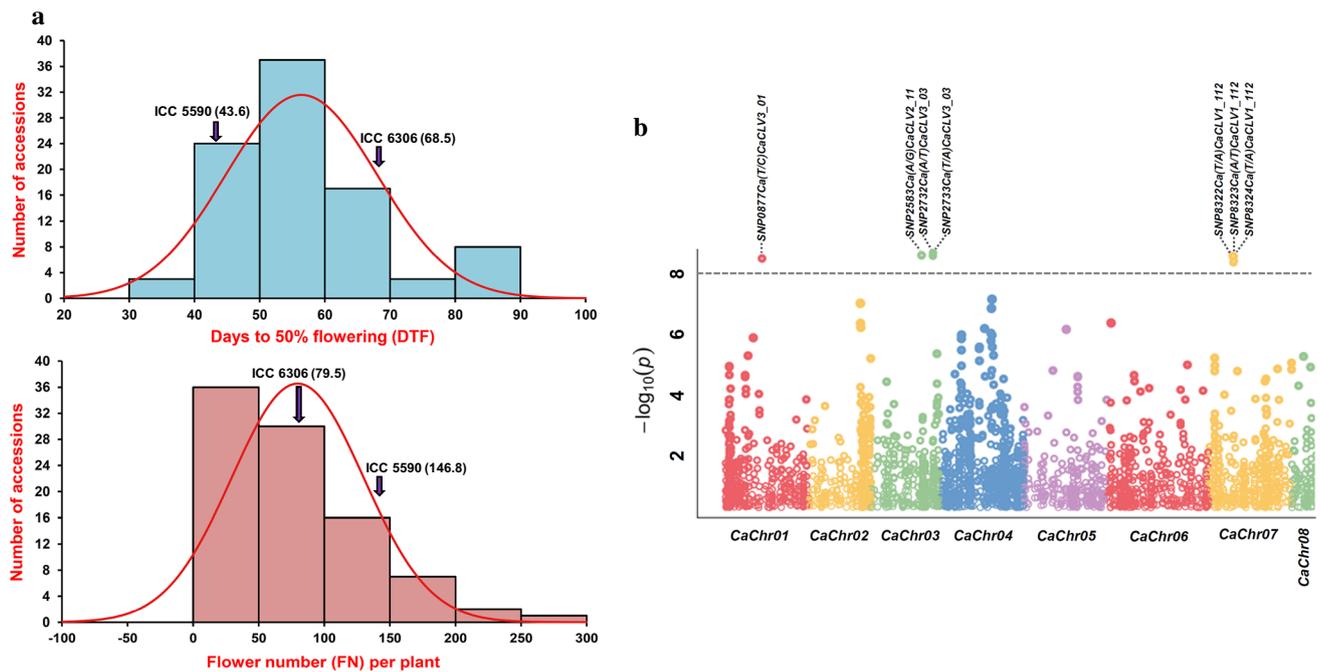
information of 480 SNPs with the multi-environments DTF and FN phenotyping data among 92 germplasm accessions (association panel) was performed. This detected a *CaCLV3\_01* gene exhibiting strong association with DTF and FN traits that was mapped in the 4.479-kb genomic interval between SNP0877Ca(T/C)CaCLV3\_01 and SNP\_Ca114505432(C/A) SNPs. The high-resolution QTL mapping in RILs and QTL-introgressed NILs integrated with QTL region-specific association analysis overall scaled-down a 75.572-kb *CaqDTF1.1/CaFN1.1* QTL interval into a 4.479-kb genomic region (Fig. 5b). The comprehensive phenotyping of 23 recombinants following progeny testing, a 75.572-kb QTL genomic interval, was resolved into a 4.479-kb region between SNP0877Ca(T/C)CaCLV3\_01 and SNP\_Ca114505432(C/A) SNPs in five selected most promising recombinants of NILs (Fig. 5b). The structural and functional annotation of this 4.479 kb *CaqDTF/FN1.1* QTL genomic interval with *kabuli* genome identified one gene *CaCLV3\_01* with accession ID, *Ca\_07084* (Fig. 5b). The same gene was also identified through QTL mapping and QTL region-specific association analysis and now has been validated through fine mapping and map-based cloning in NILs. Therefore, *CaCLV3\_01* gene harboring a major *CaqDTF/FN1.1* QTL indeed comes up as the most potential candidate regulating flowering time and flower number traits in chickpea.

### Potential *CLV* gene-derived SNP alleles influencing DTF and FN traits in chickpea

To ascertain the association potential of *CLV* genes governing DTF and FN traits in chickpea, candidate gene-based association mapping was performed. For this, high-throughput genotyping data of 8262 SNPs including 7393 *CLV1*, 702 *CLV2* and 183 *CLV3* gene-based SNPs assayed in 92 chickpea accessions (association panel) were integrated with their genetic relatedness and population diversity as well as multi-environment DTF and FN phenotyping information. The constituted association panel was phenotyped in the field at two diverse geographical locations as well as in the green house (long- and short-day photoperiod conditions) for two successive years. A wider level of phenotypic variation (20.6–63.6% CV), including normal frequency distribution of DTF and FN traits among 92 *desi* and *kabuli* chickpea accessions representing two population groups based on the multi-environment phenotyping data, was evident (Table S10, Fig. 6a). All the accessions belonging to an association panel exhibited high  $H^2$  (ranged from 80 to 86%) for the studied DTF and FN traits across environments. A significant ( $P < 0.001$ ) positive correlation between DTF and FN traits (34%) was observed in an association panel. The 8250 SNPs mined from *CLV* genes mapped on eight chromosomes were

employed in determining high-resolution population structure and PCA which differentiated all 92 *desi* and *kabuli* accessions from each other and overall grouped into two distinct populations, POP I and POP II. The LD estimate in an association panel of 92 accessions using 8250 SNPs mapped across eight chromosomes demonstrated overall LD decay ( $r^2$  reduced to half of its maximum value) nearly at a 200–300 kb physical distance of chromosomes.

The CMLM-based association mapping study at a significant FDR cutoff  $\leq 0.05$  detected seven SNPs from diverse coding and noncoding sequence components of four *CLV* genes exhibiting high association potential for both DTF and FN traits at a  $P \leq 10^{-9}$  (Table 3, Fig. 6b). All these trait-associated SNPs were validated efficiently across multiple environments in the field and greenhouse (long- and short-day photoperiod conditions). The seven trait-associated SNPs were mapped on three chromosomes (*Ca1*, *Ca3* and *Ca7*) of chickpea genome (Table 3, Fig. 6b). Six of seven trait-associated loci were derived from the URRs of three *CLV* genes, namely *CaCLV1\_112* (3 SNPs), *CaCLV2\_11* (1) and *CaCLV3\_03* (2), whereas remaining one intronic SNP was identified from *CaCLV3\_01* gene (Table 3, Fig. 6b). The PVE determined by 7 DTF and FN trait-associated individual SNPs of four *CLV* genes in a constituted association panel varied from 20.7 to 30.7%  $R^2$ . All these trait-associated SNPs derived from *CLV* genes in combination gave 35.6% PVE. As compared to four other *CLV* gene-derived SNPs, we observed a strong association of two regulatory SNPs in a *CaCLV3\_03* gene ( $10^{-10}$  P with 29.5–30.7% PVE) and an intronic SNP from a *CaCLV3\_01* gene ( $10^{-10}$  P with 29.2% PVE) with DTF and FN traits and thus considered as promising candidates for their immense use in genetic enhancement studies of chickpea (Table 3, Fig. 6b). The integration of association and QTL mapping outcomes collectively exhibited significant association of seven SNPs derived from the four candidate *CLV* genes (*CaCLV3\_01*, *CaCLV2\_11*, *CaCLV3\_03* and *CaCLV1\_112*) linked to the eight major QTLs governing DTF and FN traits in chickpea (Tables 2, 3). Notably, URR-derived and intronic SNPs in the *CaCLV3\_01* and *CaCLV3\_03* genes mapped on the genomic intervals (1.019–1.244 cM corresponding to 342023–383987 bp) of *CaqDTF1.1*, *CaqFN1.1*, *CaqDTF3.2* and *CaqFN3.2* QTLs exhibited strong association potential for both DTF and FN traits based on QTL mapping (9.8–11.5 LOD with 22.5–27.5% PVE), fine mapping and map-based cloning as well as high-resolution association analysis ( $10^{-10}$  P with 29.2–30.7% PVE) (Table 2). Collectively, SNP allelic variants scanned from the *CLV1*, *CLV2* and *CLV3* genes associated with DTF and FN traits can essentially be deployed in quick identification of promising molecular signatures (genes/QTLs and alleles) governing traits of agronomic importance in chickpea.



**Fig. 6** **a** Frequency distribution of mean days to 50% flowering time (DTF) and flower number (FN) trait variation measured in an association panel consisting of 92 *desi* and *kabuli* chickpea accessions. Arrows indicate the identities of mapping parental accessions contrasting with DTF and FN traits. **b** GWAS-led Manhattan plot illustrating the significant  $P$  value (estimated by CMLM model) associated with DTF (days to 50% flowering time) and FN (flower number)

employing 8250 *CLV* gene-derived SNPs. The x-axis depicts the relative density of reference genome-based SNPs physically mapped on eight chromosomes of *kabuli* genome. The y-axis represents the  $-\log_{10}(P)$  value for significant association of SNP loci with DTF and FN traits in chickpea. The SNPs exhibiting significant association with traits at cutoff  $P$  value  $\leq 1 \times 10^{-8}$  are indicated with dotted lines

**Table 3** *CLV* gene-derived SNPs significantly associated with days to 50% flowering time (DTF) and flower number (FN) in chickpea

| SNP IDs                         | SNPs  | SNP physical positions (bp) | <i>Kabuli</i> chromosomes | Gene accession IDs | Structural annotation | P                     | Combined PVE ( $R^2$ %) | Associated traits |
|---------------------------------|-------|-----------------------------|---------------------------|--------------------|-----------------------|-----------------------|-------------------------|-------------------|
| <i>SNP0877Ca(T/C)CaCLV3_01</i>  | [T/C] | 14500953                    | <i>Ca01</i>               | <i>Ca_07084</i>    | Intron                | $1.0 \times 10^{-10}$ | 29.2                    | DTF and FN        |
| <i>SNP2583Ca(A/G)CaCLV2_11</i>  | [A/G] | 23828679                    | <i>Ca03</i>               | <i>Ca_06142</i>    | URR                   | $1.0 \times 10^{-10}$ | 27.8                    | DTF and FN        |
| <i>SNP2732Ca(A/T)CaCLV3_03</i>  | [A/T] | 28437338                    | <i>Ca03</i>               | <i>Ca_07156</i>    | URR                   | $1.2 \times 10^{-10}$ | 30.7                    | DTF and FN        |
| <i>SNP2733Ca(T/A)CaCLV3_03</i>  | [T/A] | 28437340                    | <i>Ca03</i>               | <i>Ca_07156</i>    | URR                   | $1.0 \times 10^{-10}$ | 29.5                    | DTF and FN        |
| <i>SNP8322Ca(T/A)CaCLV1_112</i> | [T/A] | 11109690                    | <i>Ca07</i>               | <i>Ca_09375</i>    | URR                   | $1.0 \times 10^{-9}$  | 23.4                    | DTF and FN        |
| <i>SNP8323Ca(A/T)CaCLV1_112</i> | [A/T] | 11109693                    | <i>Ca07</i>               | <i>Ca_09375</i>    | URR                   | $1.2 \times 10^{-9}$  | 21.5                    | DTF and FN        |
| <i>8324Ca(T/A)CaCLV1_112</i>    | [T/A] | 11109695                    | <i>Ca07</i>               | <i>Ca_09375</i>    | URR                   | $1.0 \times 10^{-9}$  | 20.7                    | DTF and FN        |

### Differential expression profiling defines regulatory function of associated *CLV* genes for DTF and FN traits in chickpea

For differential expression profiling, four DTF and FN trait-associated genes identified through QTL/association mapping including one strong DTF/FN trait-associated gene validated through fine mapping and map-based cloning were selected. These genes were amplified using the RNA isolated

from the shoot, young leaf and SAM of two parental *desi* accessions (ICC 5590 and ICC 6306) of a RIL mapping population (ICC 5590  $\times$  ICC 6306) as well as two additional *kabuli* accessions (ICC 12968 and ICC 16814) contrasting with DTF and FN traits through quantitative RT-PCR assay. Two of four DTF and FN trait-associated genes (*CaCLV3\_01* and *CaCLV3\_03*) with regulatory and intronic SNPs exhibiting tissue-specific expression were significantly up-regulated ( $\geq$  fourfold) in SAM as compared to respective shoot and

young leaf tissues of accessions contrasting with DTF and FN traits (Figure S4). Of these, one strong DTF and FN trait-associated gene (*CaCLV3\_01*) gave pronounced up-regulation ( $\geq$  threefold) in the SAM of early-DTF and high-FN *desi* mapping parental accession ICC 5590 as compared to that of late-DTF and low-FN *desi* mapping parental accession ICC 6306 (Figure S4). However, up-regulation ( $\sim$  twofold) of *CaCLV3\_01* gene was observed in early-DTF and high-FN *kabuli* accession ICC 12968 than that of its counterpart *kabuli* accession ICC 16814. Such significant differential expression of *CLV3* gene observed among accessions within *desi* and *kabuli* did not hold true when we compared that between *desi* and *kabuli* accessions contrasting with DTF and FN traits.

### Natural haplotypes of *CLV* genes regulating DTF and FN traits in chickpea

For molecular haplotyping, a SAM-specific differentially up-regulated *CaCLV3\_01* gene exhibiting strong association with DTF and FN traits validated by QTL/fine mapping, map-based cloning, association analysis and expression profiling was sequenced among 92 *desi* and *kabuli* accessions (association panel) of chickpea. This identified a total of five coding and noncoding SNPs which constituted two major haplotypes (HAP A and HAP B) in a *CaCLV3\_01* gene (Fig. 7a). Haplotype-based association analysis by integrating the genotyping information of SNP haplotypes in a *CaCLV3\_01* gene with multi-environment DTF and FN phenotyping data of 92 accessions demonstrated strong association of haplotype HAP A ( $10^{-12}$  P with 31.5% PVE) with early-DTF (EDTF: 36.3–56.3 days represented by 52 accessions) and high-FN (HFN: 80.6–251.2 by 40 accessions), whereas haplotype HAP B ( $10^{-12}$  P with 30.6% PVE) with late-DTF (LDTF: 56.8–85.0 days by 58 accessions) and low-FN (LFN: 23.0–78.9 by 36 accessions) in chickpea (Fig. 7b). The use of an integrated genomics-assisted breeding strategy by combining QTL/fine mapping, map-based cloning and association analysis with differential expression profiling and molecular haplotyping delineated functionally relevant novel alleles and natural haplotypes of a *CaCLV3\_01* gene regulating DTF and FN traits in chickpea.

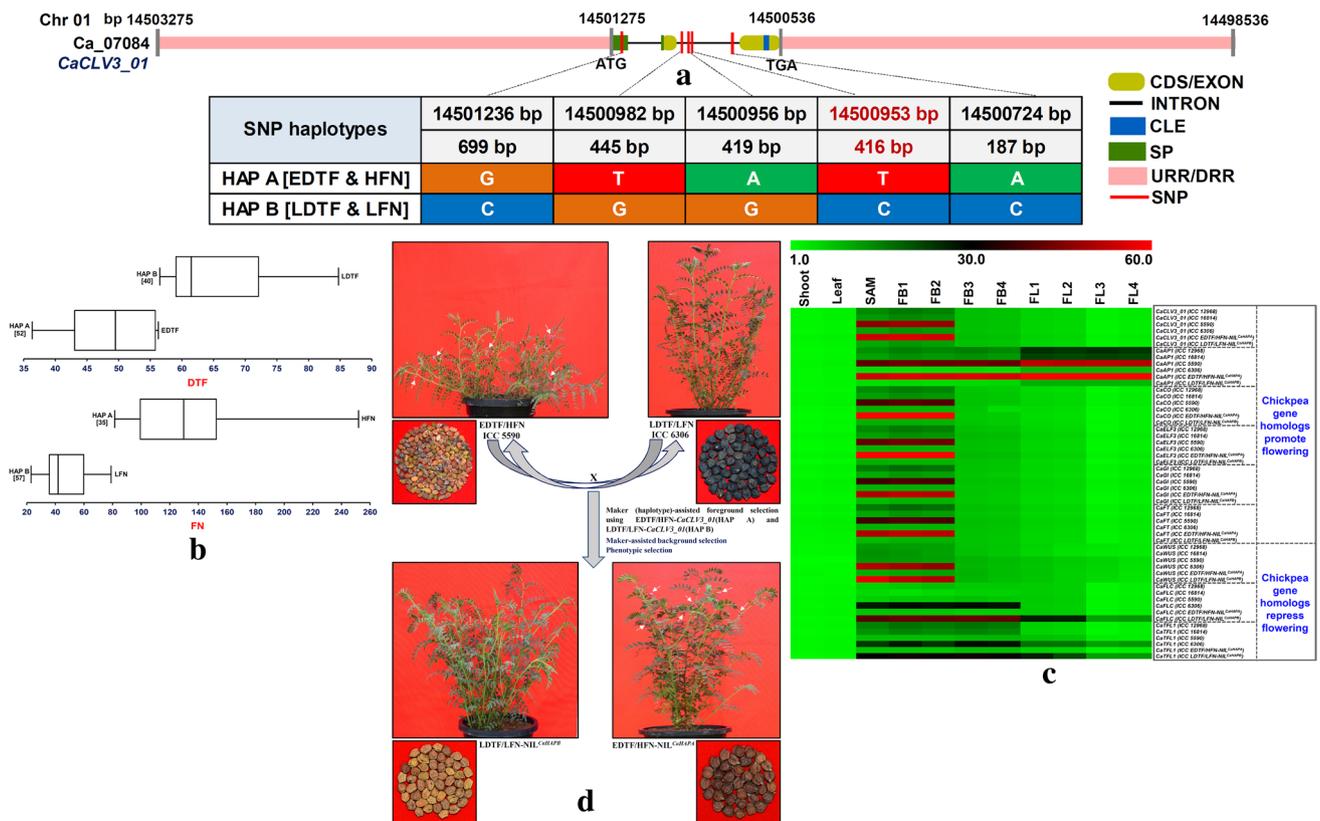
### Marker (haplotype)-assisted selection using a superior haplotype delineated from a *CaCLV3\_01* gene regulating early-DTF and high-FN traits drives genetic enhancement of chickpea

For haplotype-assisted foreground selection, the SNPs linked/flanking to a major *CaqDTF/FN1.1* QTL and a strong DTF/FN trait-associated *CaCLV3\_01* gene followed by its constituted EDTF/HFN (HAP A) as well as LDTF/LFN (HAP B) gene haplotypes were used (Fig. 1). Subsequently,

this was integrated with 1536 SNPs-based background selection that enhanced recovery of parental recurrent genome up to 99.1–99.3% in the constituted early-flowering and high-flower-number (*EDTF/HFN-NIL<sup>CaHAPA</sup>*) as well as late-flowering and low-flower-number (*LDTF/LFN-NIL<sup>CaHAPB</sup>*) *CaCLV3\_01* gene haplotypes-introgressed NILs of chickpea. This infers added-advantage and wider practical applicability of gene haplotype-assisted foreground/background selection as compared to most commonly used marker-assisted selection (79%) in recovery of parental recurrent genome as well as precise selection of most promising recombinants for further genetic enhancement studies in chickpea.

The *CaCLV3\_01* gene haplotype-specific differential expression profiling depicted higher expression ( $\geq$  five-fold) and enhanced accumulation of transcript of EDTF/HFN gene haplotype especially in the SAM and two flower development stages (FB1 and FB2) of corresponding haplotype-introgressed NILs (*EDTF/HFN-NIL<sup>CaHAPA</sup>*) of chickpea (Fig. 7c). However, their counterpart late-flowering and low-flower-number NILs (*LDTF/LFN-NIL<sup>CaHAPB</sup>*) had reduced accumulation of transcript of LDTF/LFN gene haplotype. The effect of *CaCLV3\_01* gene haplotypes on flowering time and flower number traits was compared among four selected individual NILs representing each *EDTF/HFN-NIL<sup>CaHAPA</sup>* and *LDTF/LFN-NIL<sup>CaHAPB</sup>* lines along with mapping parental accessions (ICC 5590 and ICC 6306) using multi-environment replicated field as well as greenhouse phenotyping data. The *EDTF/HFN-NIL<sup>CaHAPA</sup>* line had early-flowering (an average of 42.1 DTF) as well as enhanced flower number (228.6 FN), yield per plant (20.0 g YP) and productivity (1935.0 kg/ha YH) as compared to its counterpart *LDTF/LFN-NIL<sup>CaHAPB</sup>* line (81.0 DTF, 68.0 FN, 12.0 YP and 1252.2 kg/ha YH) (Table S11, Fig. 7D). This implicates the greater efficacy of superior EDTF/HFN haplotype delineated from a *CaCLV3\_01* gene employing an integrated genomics-assisted breeding and functional genomic strategy in developing early-flowering/maturing chickpea cultivars with enhanced seed yield and productivity.

To determine the role of constituted *CaCLV3\_01* gene haplotypes for regulating DTF and FN traits in their introgressed NILs, we performed expression profiling of downstream homeobox TF *CaWUS* gene and chickpea gene homologs that are reported to promote and repress flowering in *Arabidopsis* (Fig. 7c, Table S12). The *CLV* signaling pathway down-regulated the expression of *WUS* as evident from the expression profiles generated. The *WUS* gene showed higher expression level in the late-flowering and low-flower-number accessions and NILs with broad SAM width (ICC 6306 and *LDTF/LFN-NIL<sup>CaHAPB</sup>*) as compared to early-flowering and high-flower-number accessions and NILs with narrow SAM width (ICC 5590 and *EDTF/HFN-NIL<sup>CaHAPA</sup>*). Thus, converse expression pattern of *CaWUS* with *CaCLV3\_01* especially in the SAM was



**Fig. 7** **a** The genotyping of five SNPs of *CaCLV3\_01* among association panel (92 chickpea accessions) constituted two major haplotypes, HAP A (EDTF and HFN) and HAP B (LDTF and LFN). **b** Boxplots illustrating the variation of DTF (days to 50% flowering) and FN (flower number) traits among 92 *desi* and *kabuli* chickpea accessions. Box edges denote the lower and upper quantiles with median values in the middle of the box. Digits within the square parentheses indicate the number of *desi* and *kabuli* chickpea accessions represented by individual haplotypes, HAP A and HAP B constituted. **c** Heat map depicting the differential expression profile of *CaCLV3\_01* gene and eight other flowering-related genes in *CaCLV3\_01* gene haplotypes-introgressed NILs (*EDTF/HFN-NIL<sup>CaHAPA</sup>* and *LDTF/LFN-NIL<sup>CaHAPB</sup>*) as well as four *desi* and *kabuli* chickpea accessions including two RIL mapping parental *desi* accessions (ICC 5590 and

ICC 6306) contrasting with DTF and FN traits. The average log signal expression value of genes is represented at the top with a color scale, in which green, black and red colors denote low, medium and high level of expression, respectively. The tissues and genes selected for expression profiling are mentioned on the upper and right side of expression map, respectively. **d** Development of *EDTF/HFN-NIL<sup>CaHAPA</sup>* and *LDTF/LFN-NIL<sup>CaHAPB</sup>* lines by marker (haplotype)-assisted foreground and background selection using two mapping parental accessions (ICC 5590 and ICC 6306) contrasting with DTF and FN traits in chickpea. E/LDTF: early/late days to 50% flowering. H/LFN: high/low-flower number. CDS: coding DNA sequence, CLE: (CLVATA3/ESR-related), SP: signal peptide and URR/DRR: upstream/downstream regulatory region (color figure online)

evident suggesting their active involvement in controlling the DTF and FN traits through modulation of a vital SAM morphometric trait (SAM width) in chickpea. The chickpea gene homologs (*CaAPI1*, *CaCO*, *CaELF3*, *CaGI* and *CaFT*) that promote flowering in *Arabidopsis* exhibited their pronounced higher expression and transcript accumulation in the SAM and flower development stages especially of *EDTF/HFN-NIL<sup>CaHAPA</sup>* line. In contrast, chickpea gene homologs (*CaFLC* and *CaTFL1*) that repress flowering in *Arabidopsis* showed their reduced expression and transcript accumulation in SAM and flower development stages of *EDTF/HFN-NIL<sup>CaHAPA</sup>* line (Fig. 7c). Therefore, differential accumulation of transcripts of known flowering promoting and repressing genes in *CaCLV3\_01* gene haplotype-introgressed

NILs might regulate flowering time and flower number in chickpea.

## Discussion

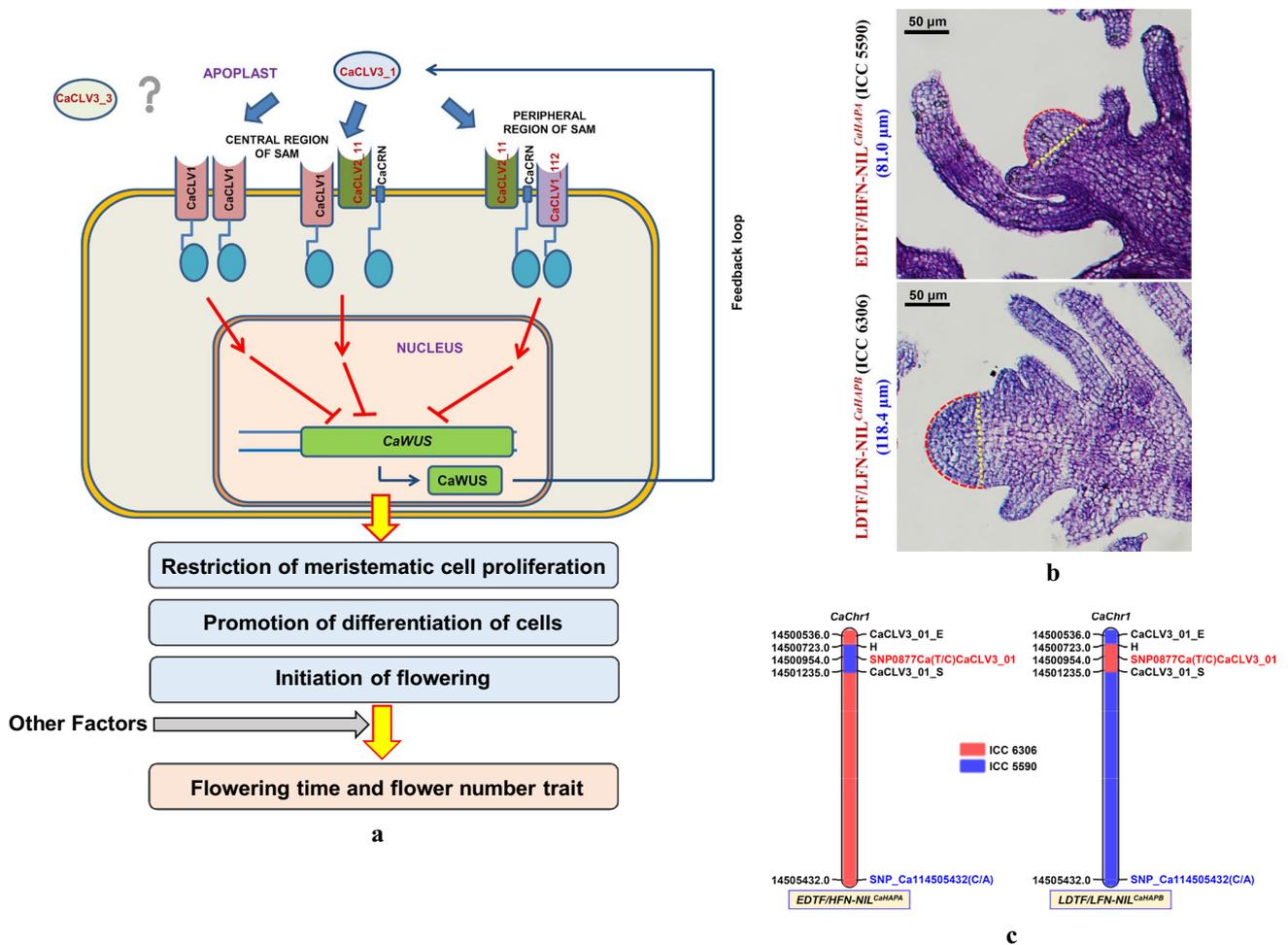
The *CLV* genes in *Arabidopsis* primarily correspond to three major classes of genes (*CLV1*, *CLV2* and *CLV3*) and related *CLV* sequences functioning in the multiple vegetative and reproductive tissues including SAM during plant growth and development (Somssich et al. 2016). The *CLV* genes in chickpea genome are randomly distributed across all the eight chromosomes and expressed in different vegetative and reproductive tissues. The *CLV3* genes characterized by

their CLE motifs have been reported earlier to be involved in diverse biological functions. One interesting observation about this group of genes is the presence of multi-CLE motif sequences. Sixteen possible CLE motifs were identified in *CaCLV3\_02* of chickpea which is significantly higher than earlier reported multi-CLE peptide-encoding genes in *Glycine max* (six CLE domains from each of two genes), *Medicago truncatula* (three–eight CLE domains in four genes) and *Lotus japonicus* (seven–ten CLE domains in four genes) (Oelkers et al. 2008; Mortier et al. 2011; Hastwell et al. 2015, 2017). Multi-CLE motif sequences were also identified in *Cajanus cajan* and *Trifolium pratense*. Among the studied legumes, two variants of CLE motifs were detected (Motif 1 and Motif 2). The CLE motif 1 exhibited greater expansion (42.3%) across the legumes, both galegoids and phaseoloides, while CLE motif 2, with an exception in *TpCLV3\_08*, was retained as one motif per *CLV3* gene sequence. The higher expansion of CLE motifs may be a result of extensive genetic recombination and/or tandem duplication of the domains especially in the *CLV3* of chickpea. In multicellular eukaryotes, there have been several reports on tandem repeats of protein domains and are majorly related to the process of stress and development pathways (Sharma and Pandey 2016). Being a small stretch of amino acid, domains are mostly unable to function as an independent structural domain. However, *CLV3* proteins with CLE domains mainly function as ligands in signaling pathways; therefore, it would be interesting to know whether the expansion of CLE domains leads to sub- and/or neo-functionalization of the proteins in controlling useful agronomic traits particularly in chickpea as compared to other legumes in which no such expansion of CLE was observed. The recent segmental gene duplication and divergence among chickpea and other legumes belonging to Papilionoideae, Galegoid and Millettoid family have led to formation of varied number of repeating CLE peptides in *CLV3* which might be essential for plant growth, development and stress response in legumes (Sawa et al. 2006; Meng et al. 2010). A comprehensive understanding of the structural, functional and evolutionary prospects of occurrence of these multiple CLE motifs in *CLVs* is required to decipher their precise role in governing useful traits of agronomic importance in crop plants.

The present study identified four *CLV* (*CaCLV1\_112*, *CaCLV2\_11*, *CaCLV3\_01* and *CaCLV3\_03*) genes that were highly associated with DTF and FN traits in chickpea (Fig. 8a). *AtBAM1* (*Barely Any Meristem 1*), *AtCLV2*, *AtCLV3* and *AtCLE16* were found to be true homologs of the *CaCLV1\_112*, *CaCLV2\_11*, *CaCLV3\_01* and *CaCLV3\_03* chickpea genes in *Arabidopsis*, respectively. The well-characterized *AtCLV2* and *AtCLV3* genes are the key players in regulation of meristematic stem cell population in *Arabidopsis* (Schoof et al. 2000). The *CLV*-*WUS* feedback

loop is essentially responsible for transition of SAM from vegetative to reproductive stage in crop plants (Clark et al. 1993). The receptor–ligand interaction of *CLV1*–*CLV2* with *CLV3* initiates a downstream signaling cascade to down-regulate the expression of *WUS* and thereby restricts the size of SAM dome. This is clearly apparent from the converse expression profiles between *CaCLV3\_01* and *CaWUS* genes of which *CaWUS* is transcriptionally down-regulated especially in the SAM of early-flowering and high-flower-number NILs with narrow SAM width (Fig. 7c). All the *clv* mutants are characterized with enlarged SAM and floral meristem volume as well as increase in floral organs. The *AtBAM1* although exhibits close homology with *AtCLV1*, but both these genes perform opposite functions in development. The *bam1* mutants show not only reduction in SAM, but also are accompanied with reduced number of floral organs (De Young et al. 2006). However, both *CLV1* and *BAM1* can complement each other partially due to overlapping biochemical function. The SAM size has been correlated with flowering time in a natural population of maize (Leiboff et al. 2015). The early-flowering accessions exhibit increased SAM volume compared to late-flowering ones at a given time point. However, this correlation was not solely due to changing levels of *CLV3* expression. The *Arabidopsis clv* mutants with enlarged SAM size are characterized by delayed flowering as *CLV3* genes restrict the meristematic stem cell population and promote its differentiation (Fig. 8a).

To ascertain this hypothesis, the SAM size of two contrasting RIL mapping parental accessions (ICC 5590 and ICC 6306) as well as *EDTF/HFN-NIL<sup>CaHAPA</sup>* and *LDTF/LFN-NIL<sup>CaHAPB</sup>* lines was correlated with differential expression of their corresponding gene haplotypes and subsequently with early-/late-DTF and low-/high-FN traits in chickpea. An enlarged SAM size (SAM width) in a late-DTF and low-FN mapping parental accession, ICC 6306 and *LDTF/LFN-NIL<sup>CaHAPB</sup>* (120.2  $\mu\text{m}$ ) line as compared to that of an early-DTF and high-FN mapping parental accession, ICC 5590 and *EDTF/HFN-NIL<sup>CaHAPA</sup>* (90.3  $\mu\text{m}$ ) line was observed (Fig. 8b, c). This was an interesting observation which hinted at the possible mechanism through which the CLAVATA signaling pathway regulates the flowering time and flower number trait in chickpea (Fig. 8a). To further validate this correlation, the SAM width was measured in seven other developed NILs with contrasting DTF trait (Table S11, Fig. 9). A direct positive correlation was observed between DTF and SAM width in these NILs where flowering time increases with gradual enlargement of SAM width. This opens up novel aspects of utilization of SAM morphometric features including SAM width as a key parameter for dissection of SAM-regulated agronomic traits in crop plants such as chickpea. Deployment of SAM morphometric characteristics in QTL mapping will target master regulators influencing simultaneously a number of correlated agronomic traits

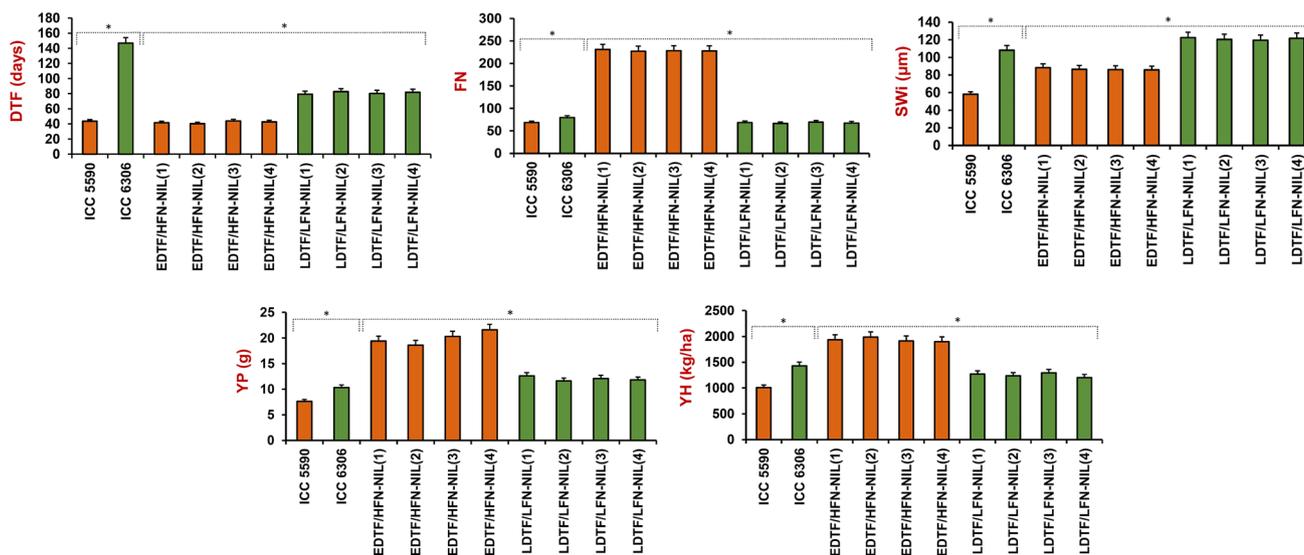


**Fig. 8** **a** Model depicting known interaction between CLAVATA and WUSCHEL signaling pathway in *Arabidopsis*. *CaCLV3\_01*, an ortholog of *AtCLV3*, is a small ligand peptide which initiates a downstream signaling cascade through a number of receptor homodimer/heterodimer complexes. The key transmembrane receptors in central portion of SAM are CLV1, CLV2 (*CaCLV2\_11*) and CRN (*CORYNE*) while in the peripheral region *BAM1* (*CaCLV1\_112*) forms hetero-complex with *CaCLV2\_11* and CRN to ultimately repress the *WUS* expression. Transcriptional down-regulation of *WUS* promotes cell differentiation by restricting the stem cell proliferation. The *CLV* signaling pathway along with other factors like photoperiod regulates flowering time and flower number in chickpea. **b** SAM width measured in the *CaCLV3\_01* gene haplotypes-introgressed NILs (*EDTF/HFN-NIL<sup>CaHAPA</sup>* and *LDTF/LFN-NIL<sup>CaHAPB</sup>*)

in crops. Correlation between SAM morphometrics and *WUS* gene expression in governing vital agromorphological traits is recently reported in chickpea (Narnoliya et al. 2019). The differential expression analysis of the *CaCLV3\_01* gene and its haplotypes depicted a significant increase in its expression in early-flowering *desi* mapping parental accession, ICC 5590 and *EDTF/HFN-NIL<sup>CaHAPA</sup>* line as compared to its *desi* counterpart ICC 6306 and *LDTF/LFN-NIL<sup>CaHAPB</sup>* line. Therefore, enhanced accumulation of transcript of a

and parental accessions (ICC 5590 and ICC 6306) of a RIL mapping population contrasting with flowering time and flower number traits. The average of SAM width ( $\mu\text{m}$ ) estimated in the parental accessions and NILs as per Table S11 is mentioned. **c** Physical maps of chromosome 1 illustrating the genomic constitution of *EDTF/HFN* (HAP A from ICC 5590, coded with blue color) and *LDTF/LFN* (HAP B from ICC 6306, coded with red color) *CaCLV3\_01* gene haplotypes-introgressed NILs (*EDTF/HFN-NIL<sup>CaHAPA</sup>* and *LDTF/LFN-NIL<sup>CaHAPB</sup>*) developed through marker (haplotype)-assisted foreground and background selection. The SNPs tightly linked with a *CaCLV3\_01* gene mapped on chromosome 1 are indicated with red color. The physical (bp) distance and identity of the markers mapped on the chromosomes are indicated on the left and right sides of the chromosomes, respectively (color figure online)

*CaCLV3\_01* gene and its superior *EDTF/HFN* haplotypes in the *desi* accession and NILs with narrow SAM size might lead to its early-DTF and high-FN characteristics (Figs. 7, 8a). Contrary, the late-DTF and low-FN *desi* accession and NILs with broader SAM size exhibited decreased accumulation of transcript of a said *CLV* gene and its *LDTF/LFN* haplotypes (Figs. 7, 8a). Within *kabuli*, ICC 12968 is an early-flowering accession which demonstrated a slight up-regulation of a strong DTF-associated *CaCLV3\_01* gene as



**Fig. 9** Graphical representation comparing five different traits, namely DTF (days to 50% flowering), FN (flower number), SWi (SAM width), YP (yield per plant) and YH (yield per hectare) among two mapping parental accessions (ICC 5590 and ICC 6306) and four lines representing each from HAP A (EDTF/HFN) and HAP B

(LDTF/LFN) introgressed NILs. Error bars represent standard error ( $n = 10$ ). \* $P < 0.01$   $t$ -test significance. EDTF/HFN-NIL and LDTF/LFN-NIL are corresponded to EDTF/HFN-NIL<sup>CaHAPA</sup> and LDTF/LFN-NIL<sup>CaHAPB</sup>, respectively. E/LDTF: early/late days to 50% flowering, H/LFN: high/low flower number and NIL: near-isogenic line

compared to late-flowering accession ICC 16814 (Figure S4). However, the definite correlation between *CLV3* gene expression level and flowering time was not valid when *kabuli* accessions were compared with *desi* accessions. This is possibly due to difference in the genetic background and distinct variation of agromorphological characteristics between *desi* and *kabuli* cultivars. The *CLV* pathway plays an important role in regulation of flowering time alongside the photoperiod pathway in crop plants (Jeong and Clark 2005). The effect of photoperiod on flower meristem development is masked by the *CLV*-*WUS* signaling pathway in *Arabidopsis* and *clv* mutants show defects in flower development in long-day conditions. The transcript profiling implicates the tissue-specific expression of *CaCLV3\_01* gene especially in the vegetative SAM of both early-/late-flowering and high-/low-flower-number accessions and NILs. Therefore, *CaCLV3\_01* gene instead of directly being regulated by photoperiod possibly acts as factor that affects photoperiod-regulated flowering time genes in chickpea. The high expression of a strong FN-associated *CLV3* (*CaCLV3\_01*) gene and its superior HFN haplotype in EDTF/HFN-NIL<sup>CaHAPA</sup> line, as well as *desi* (ICC 5590) and *kabuli* (ICC 12968) accessions with more FN, was evident. The early transition from vegetative to flowering stages in chickpea due to enhanced accumulation of transcript of superior HFN gene (*CaCLV3\_01*) haplotypes leads to production of more number of flowers per plant ultimately affecting the yield. Novel allelic variants mined from three of four *CLV* genes associated with DTF and FN traits belong to the *CLV* signaling pathway

which ascertains a major role of this pathway in regulation of these said traits. *CaCLV3\_01*, the homolog of *AtCLV3*, is the master regulator of this downstream cascade and thus comes up as a potential gene which can be modulated to achieve desired traits.

The current study employed an integrated genomic-assisted breeding strategy involving QTL/fine mapping and map-based cloning with high-resolution association analysis and molecular haplotyping along with functional genomic strategy that delineates natural alleles and haplotypes from a *CaCLV3\_01* gene harboring a major *CaqDTF/FN3.1* QTL regulating flowering time and flower number in chickpea (Figure S5). The enhanced accumulation of transcripts of EDTF/HFN *CaCLV3\_01* gene haplotypes including a diverse array of known flowering promoting genes in its corresponding haplotype-introgressed NILs and *desi* accession with reduced SAM width might regulate SAM differentiation vis-à-vis early flowering and high flower number in chickpea. The NILs introgressed with a superior EDTF/HFN *CaCLV3\_01* gene haplotypes exhibited early-flowering and high-flower number as well as enhanced yield and productivity suggesting their use in further genetic enhancement studies of chickpea. A much comprehensive analysis of transcriptional regulation of four aforesaid *CLV* genes including the EDTF/HFN trait-associated haplotypes of *CLV3* (*CaCLV3\_01*) and understanding their subsequent role in control of SAM development for regulating flowering time and flower number as well as enhancing seed yield is essential to further deploy these molecular tags in

genomics-assisted crop improvement in chickpea. The promising molecular signatures (SNPs, QTLs, genes, alleles and haplotypes) regulating DTF and FN traits delineated by us deploying an integrated genomic strategy will be useful to develop early-flowering and high-flower-/seed-yielding cultivars in chickpea.

**Author contribution statement** UB, LN and NM conducted all experiments and drafted the manuscript. RS, AS, DB, AD and VT involved in data analysis. HD, VSH and ST helped in the constitution of association panel and performed phenotyping. AKT and SKP conceived and designed the study, guided data analysis and interpretation and participated in drafting and correcting the manuscript critically, and all authors gave the final approval of the version to be published.

**Acknowledgments** The financial support provided by (Grant No. 102/IFD/SAN/2161/2013-14) the Department of Biotechnology (DBT), Ministry of Science and Technology, Government of India, is acknowledged. UB, LN and VT acknowledge the UGC (University Grants Commission) and DBT for Senior/Junior Research Fellowship awards.

## Compliance with ethical standards

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

## References

- Bajaj D, Upadhyaya HD, Khan Y et al (2015a) A combinatorial approach of comprehensive QTL-based comparative genome mapping and transcript profiling identified a seed weight-regulating candidate gene in chickpea. *Sci Rep* 5:9264
- Bajaj D, Saxena MS, Kujur A et al (2015b) Genome-wide conserved non-coding microsatellite (CNMS) marker-based integrative genomics for quantitative dissection of seed weight in chickpea. *J Exp Bot* 66:1271–1290
- Bajaj D, Upadhyaya HD, Das S, Kumar V, Gowda CL, Sharma S, Tyagi AK, Parida SK (2016) Identification of candidate genes for dissecting complex branch number trait in chickpea. *Plant Sci* 245:61–70
- Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Ser B Stat Methodol* 57:289–300
- Bommert P, Nagasawa NS, Jackson D (2013) Quantitative variation in maize kernel row number is controlled by the FASCIATED EAR2 locus. *Nat Genet* 45:334–337
- Clark SE, Running MP, Meyerowitz EM (1993) CLAVATA1, a regulator of meristem and flower development in *Arabidopsis*. *Development* 119:397–418
- Clark SE, Williams RW, Meyerowitz EM (1997) The CLAVATA1 gene encodes a putative receptor-kinase that controls shoot and floral meristem size in *Arabidopsis*. *Cell* 89:575–585
- Cock JM, McCormick S (2001) A large family of genes that share homology with CLAVATA3. *Plant Physiol* 126:939–942
- De Young BJ, Bickle KL, Schrage KJ, Muskett P, Patel K, Clark SE (2006) The CLAVATA1-related BAM1, BAM2 and BAM3 receptor kinase-like proteins are required for meristem function in *Arabidopsis*. *Plant J* 45:1–16
- Durbak AR, Tax FE (2011) CLAVATA signaling pathway receptors of *Arabidopsis* regulate cell proliferation in fruit organ formation as well as in meristems. *Genetics* 189:177–194
- Gautam V, Singh A, Singh S, Sarkar AK (2016) An efficient LCM-based method for tissue specific expression analysis of genes and miRNAs. *Sci Rep* 6:21577
- Grienenberger E, Fletcher JC (2015) Polypeptide signalling molecules in plant development. *Curr Opin Plant Biol* 23:8–14
- Gupta S, Nawaz K, Parween S, Roy R, Sahu K, Pole A, Khandal H, Srivastava R, Parida S, Chattopadhyay D (2016) Draft genome sequence of *Cicer reticulatum* L., the wild progenitor of chickpea provides a resource for agronomic trait improvement. *DNA Res* 24:1–10
- Hastwell AH, Gresshoff PM, Ferguson BJ (2015) Genome-wide annotation and characterization of CLAVATA/ESR (CLE) peptide hormones of soybean (*Glycine max*) and common bean (*Phaseolus vulgaris*), and their orthologues of *Arabidopsis thaliana*. *J Exp Bot* 66:5271–5287
- Hastwell AH, De Bang TC, Gresshoff PM, Ferguson BJ (2017) CLE peptide-encoding gene families in *Medicago truncatula* and *Lotus japonicus*, compared with those of soybean, common bean and *Arabidopsis*. *Sci Rep* 7:1–13
- Jain M, Misra G, Patel RK et al (2013) A draft genome sequence of the pulse crop chickpea (*Cicer arietinum* L.). *Plant J* 74:715–729
- Jeong S, Clark SE (2005) Photoperiod regulates flower meristem development in *Arabidopsis thaliana*. *Genetics* 169:907–915
- Jeong S, Trotochaud AE, Clark SE (1999) The *Arabidopsis* CLAVATA2 gene encodes a receptor-like protein required for the stability of the CLAVATA1 receptor-like kinase. *Plant Cell* 11:1925–1934
- Jiang L, Qian Q, Mao L, Zhou QY, Zhai WX (2005) Characterization of the rice floral organ number mutant fon3. *J Integr Plant Biol* 47:100–106
- Jiang SY, González JM, Ramachandran S (2013) Comparative genomic and transcriptomic analysis of tandemly and segmentally duplicated genes in rice. *PLoS ONE* 8:e63551
- Kujur A, Bajaj D, Upadhyaya HD et al (2015a) A genome-wide SNP scan accelerates trait-regulatory genomic loci identification in chickpea. *Sci Rep* 5:11166
- Kujur A, Upadhyaya HD, Shree T et al (2015b) Ultra-high density intra-specific genetic linkage maps accelerate identification of functionally relevant molecular tags governing important agronomic traits in chickpea. *Sci Rep* 5:9468
- Kujur A, Bajaj D, Upadhyaya HD et al (2015c) Employing genome-wide SNP discovery and genotyping strategy to extrapolate the natural allelic diversity and domestication patterns in chickpea. *Front Plant Sci* 6:162
- Kujur A, Upadhyaya HD, Bajaj D, Gowda CL, Sharma S, Tyagi AK, Parida SK (2016) Identification of candidate genes and natural allelic variants for QTLs governing plant height in chickpea. *Sci Rep* 6:27968
- Landau U, Asis L, Eshed Williams L (2015) The ERECTA, CLAVATA and class III HD-ZIP pathways display synergistic interactions in regulating floral meristem activities. *PLoS ONE* 10:e0125408
- Laux T, Mayer KF, Berger J, Jürgens G (1996) The WUSCHEL gene is required for shoot and floral meristem integrity in *Arabidopsis*. *Development* 122:87–96
- Leiboff S, Li X, Hu HC et al (2015) Genetic control of morphometric diversity in the maize shoot apical meristem. *Nat Commun* 6:8974
- Li J, Tax FE (2013) Receptor-like kinases: key regulators of plant development and defense. *J Integr Plant Biol* 55:1184–1187
- Li S, Pan Y, Wen C, Li Y, Liu X, Zhang X, Behera TK, Xing G, Weng Y (2016) Integrated analysis in bi-parental and natural populations

- reveals *CsCLAVATA3* (*CsCLV3*) underlying carpel number variations in cucumber. *Theor Appl Genet* 129:1007–1022
- Lipka AE, Tian F, Wang Q, Peiffer J, Li M, Bradbury PJ, Gore MA, Buckler ES, Zhang Z (2012) GAPIT: genome association and prediction integrated tool. *Bioinformatics* 2:2397–2399
- Liu R, Jia H, Cao X, Huang J, Li F, Tao Y, Qiu F, Zheng Y, Zhang Z (2012) Fine mapping and candidate gene prediction of a pleiotropic quantitative trait locus for yield-related trait in *Zea mays*. *PLoS ONE* 7:e49836
- Malik N, Dwivedi N, Singh AK, Parida SK, Agarwal P, Thakur JK, Tyagi AK (2016) An integrated genomic strategy delineates candidate mediator genes regulating grain size and weight in rice. *Sci Rep* 6:23253
- Meng L, Buchanan BB, Feldman LJ, Luan S (2010) A putative nuclear CLE-Like (CLEL) peptide precursor regulates root growth in *Arabidopsis*. *Mol Plant* 5:955–957
- Mortier V, Fenta B, Martens C, Rombauts S, Holsters M, Kunert K, Goormachtig S (2011) Search for nodulation-related CLE genes in the genome of *Glycine max*. *J Exp Bot* 62:2571–2583
- Muller R, Bleckmann A, Simon R (2008) The receptor kinase CORYNE of *Arabidopsis* transmits the stem cell-limiting signal CLAVATA3 independently of CLAVATA1. *Plant Cell* 20:934–946
- Muthamilarasan M, Mangu VR, Zandkarimi H, Prasad M, Baisakh N (2016) Structure, organization and evolution of ADP-ribosylation factors in rice and foxtail millet, and their expression in rice. *Sci Rep* 6:24008
- Narnoliya L, Basu U, Bajaj D et al (2019) Transcriptional signatures modulating SAM morphometric and plant architectural traits enhance yield and productivity in chickpea. *Plant J*. <https://doi.org/10.1111/tpj.14284>
- Oelkers K, Goffard N, Weiller GF, Gresshoff PM, Mathesius U, Frickey T (2008) Bioinformatic analysis of the CLE signalling peptide family. *BMC Plant Biol* 8:1
- Parween S, Nawaz K, Roy R et al (2015) An advanced draft genome assembly of a *desi* type chickpea (*Cicer arietinum* L.). *Sci Rep* 5:12806
- Ridge S, Deokar A, Lee R, Daba K, Macknight RC, Weller JL, Tar'an B (2017) The chickpea Early Flowering 1 (*Efl1*) locus is an ortholog of *Arabidopsis* *ELF3*. *Plant Physiol* 175:802–815
- Sawa S, Kinoshita A, Nakanomyo I, Fukuda H (2006) CLV3/ESR related (CLE) peptides as intercellular signalling molecules in plants. *Chem Rec* 6:303–310
- Saxena MS, Bajaj D, Das S, Kujur A, Kumar V, Singh M, Bansal KC, Tyagi AK, Parida SK (2014a) An integrated genomic approach for rapid delineation of candidate genes regulating agro-morphological traits in chickpea. *DNA Res* 21:695–710
- Saxena MS, Bajaj D, Kujur A, Das S, Badoni S, Kumar V, Singh M, Bansal KC, Tyagi AK, Parida SK (2014b) Natural allelic diversity, genetic structure and linkage disequilibrium pattern in wild chickpea. *PLoS ONE* 9:e107484
- Schoof H, Lenhard M, Haecker A, Mayer KF, Jürgens G, Laux T (2000) The stem cell population of *Arabidopsis* shoot meristems is maintained by a regulatory loop between the CLAVATA and WUSCHEL genes. *Cell* 100:635–644
- Sharma M, Pandey GK (2016) Expansion and function of repeat domain proteins during stress and development in plants. *Front Plant Sci* 6:1218
- Shinohara H, Moriyama Y, Ohyama K, Matsubayashi Y (2012) Biochemical mapping of a ligand-binding domain within *Arabidopsis* BAM1 reveals diversified ligand recognition mechanisms of plant LRR-RKs. *Plant J* 70:845–854
- Singh VK, Garg R, Jain M (2013) A global view of transcriptome dynamics during flower development in chickpea by deep sequencing. *Plant Biotechnol J* 11:691–701
- Somssich M, Bi Je, Simon R, Jackson D (2016) CLAVATA-WUSCHEL signaling in the shoot meristem. *Development* 143:3238–3248
- Stahl Y, Wink RH, Ingram GC, Simon R (2009) A signaling module controlling the stem cell niche in *Arabidopsis* root meristems. *Curr Biol* 19:909–914
- Suzaki T (2004) The gene FLORAL ORGAN NUMBER1 regulates floral meristem size in rice and encodes a leucine-rich repeat receptor kinase orthologous to *Arabidopsis* CLAVATA1. *Development* 131:5649–5657
- Taguchi-Shiobara F, Yuan Z, Hake S, Jackson D (2001) The *fasciated ear2* gene encodes a leucine-rich repeat receptor-like protein that regulates shoot meristem proliferation in maize. *Genes Dev* 15:2755–2766
- Trotochaud AE, Hao T, Wu G, Yang Z, Clark SE (1999) The CLAVATA1 receptor-like kinase requires CLAVATA3 for its assembly into a signaling complex that includes KAPP and a Rho-related protein. *Plant Cell* 11:393–406
- Upadhyaya HD, Bajaj D, Das S et al (2015) A genome-scale integrated approach aids in genetic dissection of complex flowering time trait in chickpea. *Plant Mol Biol* 89:403–420
- Van Ooijen JW (2009) MapQTL 6: software for the mapping of quantitative trait loci in experimental populations of diploid species. B.V. Kyazma, Wageningen
- Varshney RK, Mohan SM, Gaur PM et al (2013a) Achievements and prospects of genomics-assisted breeding in three legume crops of the semi-arid tropics. *Biotechnol Adv* 31:1120–1134
- Varshney RK, Song C, Saxena RK et al (2013b) Draft genome sequence of chickpea (*Cicer arietinum*) provides a resource for trait improvement. *Nat Biotechnol* 31:240–246
- Voorrips RE (2002) MapChart: software for the graphical presentation of linkage maps and QTLs. *J Hered* 93:77–78
- Whitewoods CD, Cammarata J, Nemeček V, Z et al (2018) CLAVATA was a genetic novelty for the morphological innovation of 3D growth in land plants. *Curr Biol* 28:2365–2376.e5
- Xu C, Liberatore KL, Macalister CA et al (2015) A cascade of arabinosyltransferases controls shoot meristem size in tomato. *Nat Genet* 47:784–792

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.