Plant Mol Biol Rep
DOI 10.1007/s11105-017-1044-0

ORIGINAL PAPER

Identifying Transcription Factor Genes Associated with Yield Traits in Chickpea

Philanim Wungmarong Shimray[1] · Deepak Bajaj[2] · Rishi Srivastava[2] · Anurag Daware[2] · Hari D. Upadhyaya[3] · Rajendra Kumar[4] · Chellapilla Bharadwaj[1] · Akhilesh K. Tyagi[2,5] · Swarup K. Parida[2]

Abstract Identification of potential transcription factor (TF) gene-derived natural SNP allelic variants regulating pod and seed yield component traits by large-scale mining and genotyping of SNPs in natural germplasm accessions coupled with high-resolution association mapping is vital for understanding the complex genetic architecture of quantitative yield traits in chickpea. In these perspectives, the current study employed a genome-wide GBS (genotyping-by-sequencing) and targeted gene amplicon resequencing-based simultaneous SNP discovery and genotyping assays, which discovered 1611 novel SNPs from 736 TF genes physically mapped on eight chromosomes and unanchored scaffolds of *kabuli* chickpea genome. These SNPs were structurally and functionally annotated in diverse synonymous and non-synonymous coding as well as non-coding regulatory and intronic sequence components of chickpea TF genes. A high-resolution genetic association analysis was performed by correlating the genotyping information of 1611 TF gene-based SNPs with multi-location/years field phenotyping data of six major pod and seed yield traits evaluated in a constituted association panel (326 *desi* and *kabuli* germplasm accessions) of chickpea. This essentially identified 27 TF gene-derived SNPs exhibiting significant association with six major yield traits, namely days to 50% flowering (DF), plant height (PH), branch number (BN), pod number (PN), seed number (SN) and seed weight (SW) in chickpea. These trait-associated SNPs individually and in combination explained 10–23% and 32% phenotypic variation respectively for the studied yield component traits. Interestingly, novel non-synonymous coding SNP allelic variants in five potential candidate TF genes encoding SBP (squamosal promoter binding protein), SNF2 (sucrose non-fermenting 2), GRAS [Gibberellic acid insensitive (GAI)-Repressor of GAI (RGA)-SCARECROW (SCR)], bZIP (basic leucine zipper) and LOB (lateral organ boundaries)-domain proteins associated strongly with DF, PH, BN, PN, SN and SW traits respectively were found most promising in chickpea. The functionally relevant molecular signatures (TFs and natural SNP alleles) delineated by us have potential to accelerate marker-assisted genetic enhancement by developing high pod and seed yielding cultivars of chickpea.

Keywords Association mapping · Chickpea · *Desi* · *Kabuli* · SNP · Transcription factor

Philanim Wungmarong Shimray and Deepak Bajaj have equal contribution

Electronic supplementary material The online version of this article (doi:10.1007/s11105-017-1044-0) contains supplementary material, which is available to authorized users.

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

[1] Division of Genetics, Indian Agricultural Research Institute (IARI), New Delhi 110012, India

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

[3] International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru, Telangana 502324, India

[4] U.P. Council of Agricultural Research, Gomati Nagar, Lucknow, Uttar Pradesh 226010, India

[5] Present address: Department of Plant Molecular Biology, University of Delhi South Campus, New Delhi 110021, India

Introduction

Chickpea (*Cicer arietinum*), represented by diverse *desi* and *kabuli* cultivars, is a legume food crop of profound economic

Published online: 22 August 2017

and nutritional value (Kumar et al. 2011; Varshney et al. 2013a). Enhancing chickpea productivity by developing multiple high seed and pod-yielding *desi* and *kabuli* cultivars is the major current genomics-assisted breeding objectives to sustain global food and nutritional security amidst climate change and huge population growth scenario. However, most of the pod and seed yield-contributing traits like flowering time, plant height, branch number, seed/pod number and seed weight are polygenic in nature and exhibit complex quantitative genetic inheritance pattern in chickpea (Gaur et al. 2014; Varshney 2016). Significant efforts have been made by several global research groups to dissect these complex yield traits by employing diverse genomics-assisted breeding strategies including high-resolution QTL (quantitative trait loci) and association mapping in chickpea (Kujur et al. 2013, 2014, 2015a, b, c, 2016; Saxena et al. 2014a; Thudi et al. 2014; Varshney et al. 2014; Bajaj et al. 2015a, b, c, 2016a, b; Das et al. 2015, 2016; Kale et al. 2015; Upadhyaya et al. 2015; Singh et al. 2016). This is primarily accomplished by high-throughput genotyping of gene/genome-based SSR (simple sequence repeat) and SNP (single nucleotide polymorphism) markers in phenotypically well-characterized natural core/mini-core germplasm accessions and advanced generation intra-/inter-specific mapping populations of chickpea. Unfortunately, very limited number of informative sequence-based markers tightly linked to the major QTLs/genes governing aforesaid complex yield traits are identified till date to be exploited for marker-assisted genetic improvement of chickpea (Varshney et al. 2013a, b, 2014; Saxena et al. 2014a; Upadhyaya et al. 2015). To drive these prime objectives, the gene-based markers essentially appear much more informative and proficient for fast delineation of potential genes and alleles underlying major QTLs regulating diverse vital agronomic traits and deciphering the complex genetic architecture of seed and pod yield component quantitative traits in chickpea (Kujur et al. 2013, 2014; Saxena et al. 2014a; Bajaj et al. 2015a, b, 2016a, b; Upadhyaya et al. 2015).

Transcription factors (TFs) controlling the expression of genes are vital for regulation of diverse biological processes including growth, development and cellular differentiation as well as responses to various environmental cues (Udvardi et al. 2007; Century et al. 2008; Libault et al. 2009; Kujur et al. 2013, 2014; Saxena et al. 2014a,b; Yu et al. 2010). The significant role of TF genes as prime regulators in governing diverse complex cellular processes like growth, development and stress responses as well as various domestication traits is well-demonstrated in legumes including chickpea (Udvardi et al. 2007; Century et al. 2008; Libault et al. 2009; Kujur et al. 2013, 2014; Saxena et al. 2014b; Yu et al. 2010). In chickpea, the draft whole genomes as well as genome/transcriptome resequencing for diverse growth/development and stress-imposed tissues/stages of multiple cultivated (*desi* and *kabuli*) and wild *Cicer* accessions are underway (Garg et al. 2011, 2014; Gaur et al. 2012, 2015; Hiremath et al. 2012; Jain et al. 2013; Singh et al. 2013; Varshney et al. 2013c; Deokar et al. 2014; Kudapa et al. 2014; Pradhan et al. 2014; Parween et al. 2015; Gupta et al. 2016). These currently serve as a useful genomic/transcriptomic resource for genome-wide identification aside structural/functional annotation of the TF-encoding genes in chickpea. Consequently, these exertions assisted in detection of diverse differentially expressed TF genes including tissue- and stage-specific TFs governing manifold stress and developmental responses in chickpea accessions with contrasting agronomic traits. Besides, these efforts are found much expedient to develop numerous genic SSR and SNP markers representing various coding and regulatory sequence components of TF genes (especially differentially expressed TFs) annotated from chickpea genome (Garg et al. 2011, 2014; Gaur et al. 2012, 2015; Hiremath et al. 2012; Jain et al. 2013; Singh et al. 2013; Varshney et al. 2013c; Deokar et al. 2014; Kudapa et al. 2014; Pradhan et al. 2014; Parween et al. 2015; Gupta et al. 2016). These TF gene-based markers thus have functional significance in establishing marker-trait linkages for driving genetic enhancement of chickpea. Essentially, this will expedite the identification and targeted mapping of potential TF genes underlying major QTLs associated with vital agronomic traits by employing QTL and association mapping as well as diverse integrated genomics-assisted breeding strategies in chickpea (Kujur et al. 2013, 2014; Saxena et al. 2014a; Upadhyaya et al. 2015; Bajaj et al. 2016a, b). These informative genic markers, therefore, have been extensively deployed in various high-throughput genetic analysis and marker-aided genetic enhancement of chickpea (Kujur et al. 2013, 2014; Saxena et al. 2014a, b; Bajaj et al. 2015a, b, 2016b; Upadhyaya et al. 2015).

Based on aforesaid perspectives, the functional significance of TF gene-derived markers in rapid genetic dissection of complex quantitative traits is apparent in chickpea. To accomplish these, large-scale genome-wide discovery and high-throughput genotyping of especially the TF gene-based markers (SNPs) among phenotypically well-characterized *desi* and *kabuli* core/mini-core germplasm accessions (association panel) is essential. Henceforth, this will be useful to delineate functionally relevant natural alleles of the TF genes governing diverse major pod and seed yield component traits through genetic association analysis in chickpea. Moreover, outcomes obtained from these exertions will essentially enrich our deep understanding on trait inheritance pattern and complex genetic architecture of quantitative yield traits for driving genomics-assisted crop improvement to develop high pod and seed yielding *desi* and *kabuli* cultivars of chickpea.

In view of above prospects, the current study was undertaken with a prime objective to delineate functionally relevant natural SNP allelic variants of TF genes regulating diverse major pod and seed yield component traits by deploying high-resolution association mapping in chickpea.

Materials and Methods

Phenotyping of a Constituted Association Panel

For mining and high-throughput genotyping of TF gene-based SNPs for genetic association analysis, a trait-specific association panel comprising of 326 diverse *desi* (206) and *kabuli* (120) accessions were selected from the available chickpea core/mini-core germplasm collections following Kujur et al. (2014, 2015a, b). These accessions were grown in the field as per *alpha-lattice* design with two replications during crop season for two consecutive years (2012 and 2013) at two diverse geographical locations [National Institute of Plant Genome Research (NIPGR), New Delhi; latitude/longitude 28.4° N/77.1° E and International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru, Hyderabad; 17.1° N/78.9° E)] of India. These natural germplasm accessions were further phenotyped for six major pod and seed yield-contributing quantitative traits including flowering time, plant height, branch number per plant, pod number per plant, seed number per plant and 100-seed weight in chickpea.

Aside field phenotyping, greenhouse trail was conducted under both long (16 h light/8 h dark) and short (10 h light/14 h dark) day conditions at 22 ± 2 °C for evaluating the flowering time trait inheritance pattern in 326 *desi* and *kabuli* chickpea germplasm accessions. All the aforesaid pod and seed yield component traits were measured by estimating the average value of 10–15 representative plants selected from each germplasm accession. The days to 50% flowering time (DF) of each germplasm accession was determined by counting the number of days from sowing (first irrigation) to the stage when 50% of their plants have begun to flower in chickpea. The plant height (PH) was estimated by measuring the mean canopy height (cm) of each germplasm accession from soil surface at the time of their flower ending/pod setting initiation stage. The branch number (BN) of each germplasm accession was estimated by counting the average number of productive pods/seeds-bearing branches emerging from each plant at time of harvest. The pod number (PN) and seed number (SN) of each accession were measured by counting the average number of fully formed pods and seeds per plant at maturity, respectively. The 100-seed weight (SW) of each accession was estimated by measuring the average weight of 100-matured seeds (at 10% moisture content) of chickpea. Diverse statistical parameters including coefficient of variation (CV), frequency distribution, broad-sense heritability (H^2) and Pearson's correlation coefficient (r) of six major pod and seed yield component traits were measured using SPSSv17.0 (http://www.spss.com/statistics) and the methods of Bajaj et al. (2015a) and Kujur et al. (2015b, c).

Targeted TF Gene Amplicon Resequencing-Based SNP Mining and Genotyping

A total of 16,376 SNPs physically mapped on eight chromosomes and unanchored scaffolds of *kabuli* chickpea genome were mined from the sequenced 92 diverse *desi* and *kabuli* chickpea germplasm accessions by employing a simple, robust and economical as well as simultaneous genome-wide SNP discovery and genotyping strategy, GBS (genotyping-by-sequencing) assay (Elshire et al. 2011; Kujur et al. 2015a, b). From these, the genotyping as well as structural/functional annotation information of SNPs especially derived from TF genes mapped on chromosomes and scaffolds of *kabuli* genome were acquired. The TF genes with SNPs selected from GBS data were further resequenced using the genomic DNA of 326 diverse *desi* and *kabuli* chickpea germplasm accessions (association panel) employing the multiplexed amplicon resequencing method (TruSeq Custom Amplicon v1.5) of Illumina MiSeq next-generation sequencer (Illumina, USA). The custom oligo probes targeting the CDS (coding DNA sequences)/exons, introns, 2 kb-URRs (upstream regulatory regions) and 2 kb-DRRs (downstream regulatory regions) of TF genes were designed and synthesized using Illumina Design Studio. All the probes producing amplicons with an average size of 500 bp per reaction were pooled into a custom amplicon tube. The template libraries were constituted after adding of sample-specific indices to each library by PCR using common primers. The uniquely tagged pooled amplicon libraries were normalized and generated clusters were sequenced by Illumina MiSeq platform. The visualization/mapping of sequenced TF gene amplicons and discovery of high-quality sequence variants among accessions were performed as per Saxena et al. (2014a) and Malik et al. (2016). The pseudomolecules of *kabuli* chickpea genome (Varshney et al. 2013c) were used as a reference to map the high-quality gene amplicon sequence reads of each chickpea accession. Accordingly, the TF gene-derived high-quality SNPs were detected among chickpea accessions as per Saxena et al. (2014a) and Kujur et al. (2015a, b, c).

Genetic Association Analysis

For association mapping, we selected the chickpea germplasm accessions exhibiting high-heritability (H^2) for six major yield component traits (DF, PH, BN, PN, SN and SW) across two environments (geographical regions) and experimental years. The environment and/or replication-wise six major yield trait phenotyping data obtained from each of the selected accession was used for estimating the mean trait value in an individual accession to be employed in association analysis. For candidate gene-based association mapping, the genotyping information of TF gene-derived SNPs (≥ 5% minor allele frequency) was correlated with multi-location/years replicated

field phenotyping data of six major pod and seed yield component traits (PH, BN, DF, PN, SN and SW) and population structure (Q), kinship (K) matrix and PCA (principal component analysis) (P) data of 326 *desi* and *kabuli* accessions belonging to an association panel of chickpea. The PCA and K matrix among accessions were measured using GAPIT (Lipka et al. 2012) and SPAGeDi 1.2 (Hardy and Vekemans 2002), respectively. To perform association analysis, the CMLM (compressed mixed linear model) (P + K, K and Q + K) along with P3D [population parameters previously determined (Zhang et al. 2010; Kang et al. 2010)] interfaces of GAPIT were employed following Thudi et al. (2014); Kujur et al. (2015b) and Kumar et al. (2015). To determine the accuracy and robustness of TF gene-derived SNP marker-trait association, the quantile-quantile (Q-Q) plot-based false discovery rate (FDR cut-off ≤ 0.05) corrections (Benjamini and Hochberg 1995) for multiple comparisons between observed/expected -$\log_{10}(P)$ values and adjusted P value threshold of significance were performed as per Kujur et al. (2015b). The TF gene-derived SNP loci exhibiting significant association with DF, PH, BN, PN, SN and SW traits at a lowest FDR adjusted P values (threshold $P < 1 \times 10^{-6}$) and highest R^2 were identified in chickpea. The R^2 representing the magnitude of SNP marker-trait association was estimated based on model with the SNPs and adjusted P values following FDR-controlling method.

Results and Discussion

Discovery, Genotyping and Annotation of TF Gene-Derived SNPs for Genomics-Assisted Breeding Applications in Chickpea

A diverse array of TF-encoding genes is known to regulate multiple growth, development and yield-related traits in crop plants including chickpea (Udvardi et al. 2007; Century et al. 2008; Libault et al. 2009; Kujur et al. 2013, 2014; Saxena et al. 2014a, b; Yu et al. 2010). Therefore, it would be interesting to determine the association potential of these TF genes in governing diverse major pod and seed yield component traits in chickpea. This can be primarily achieved by large-scale genotyping of novel synonymous and non-synonymous coding and non-coding intronic and regulatory SNP allelic variants discovered from the TF genes in a diverse set of phenotypically well-characterized natural *desi* and *kabuli* germplasm accessions (association panel) of chickpea. In the current investigation, considering the pros and cons of GBS assay, a high sequencing-depth coverage targeted gene amplicon resequencing strategy coupled with GBS assay was utilized to discover and genotype TF gene-based high-quality SNPs uniformly across 326 *desi* and *kabuli* germplasm accessions (association panel) of chickpea. Primarily, the genotyping and sequencing of 92 accessions selected from 326 *desi* and *kabuli* chickpea accessions using GBS assay discovered 1029 high-quality SNPs in 736 TF genes annotated on *kabuli* reference genome (Table S1). A NGS-led GBS assay is proficient enough in fast large-scale discovery and high-throughput genotyping of SNPs simultaneously at a genome-wide scale for genomics-assisted breeding applications including association mapping to scan potential genes/QTLs regulating vital agronomic traits in chickpea (Deokar et al. 2014; Bajaj et al. 2015a, b, c, 2016a, b; Jaganathan et al. 2015; Kujur et al. 2015a, b, c, 2016; Upadhyaya et al. 2015). However, reduced potential of GBS assay to generate non-erroneous and high-quality homozygous SNP genotyping information uniformly across accessions genotyped with a high genome/gene coverage is quite apparent. This could restrain the use of GBS assay in high-resolution association mapping study in a large chickpea genome with narrow genetic base. Henceforth, revalidation of GBS-derived SNP genotyping information as well as discovery and large-scale genotyping of novel SNPs covering the whole genome/gene regions at a high-resolution scale using numerous germplasm accessions are essential prior to deployment of these markers in genomics-assisted breeding applications and genetic enhancement of chickpea.

To accomplish these prime objectives, NGS-based amplicon resequencing strategy targeting TF genes (scanned previously by GBS assay) with high sequencing-depth coverage was employed in the present study for large-scale discovery and genotyping of SNPs in 326 *desi* and *kabuli* germplasm accessions belonging to an association panel of chickpea. Primarily, this strategy successfully validated all 1029 GBS-derived SNPs mined from 736 chickpea TF genes (Table S1). Further, the *kabuli* reference genome-based GBS and targeted gene amplicon resequencing-led high-throughput SNP genotyping in 326 *desi* and *kabuli* accessions (association panel) altogether discovered 1611 SNPs from 736 TF genes (representing 30 TF gene family) with an average density of 2.2 SNPs per TF (Fig. 1). Maximum of 551 and 484 SNPs were discovered from the 257 and 221 TF-encoding genes especially belonging to Zinc finger and DUF (domain of unknown function) TF gene families, respectively (Fig. 1). Of these, 1497 and 114 SNPs derived from 683 and 53 TF genes were physically mapped across eight chromosomes and unanchored scaffolds of *kabuli* genome, respectively (Fig. 2a, b). All eight chickpea chromosomes contained maximum frequency of SNPs from the TF genes belonging to Zinc finger and ARF (auxin responsive factor) TF gene families (Fig. 2a). The comprehensive structural annotation of 1611 SNPs in 736 TF genes exhibited presence of highest and lowest proportion of SNPs in the exons/CDS (58.3%, 939 SNPs) and DRRs (0.4%, 7) of 422 and 3 TFs, respectively (Fig. 2c). The coding SNPs contained 529 (56.3%) synonymous and 410 (43.7%) non-synonymous (missense and nonsense SNPs) SNPs from

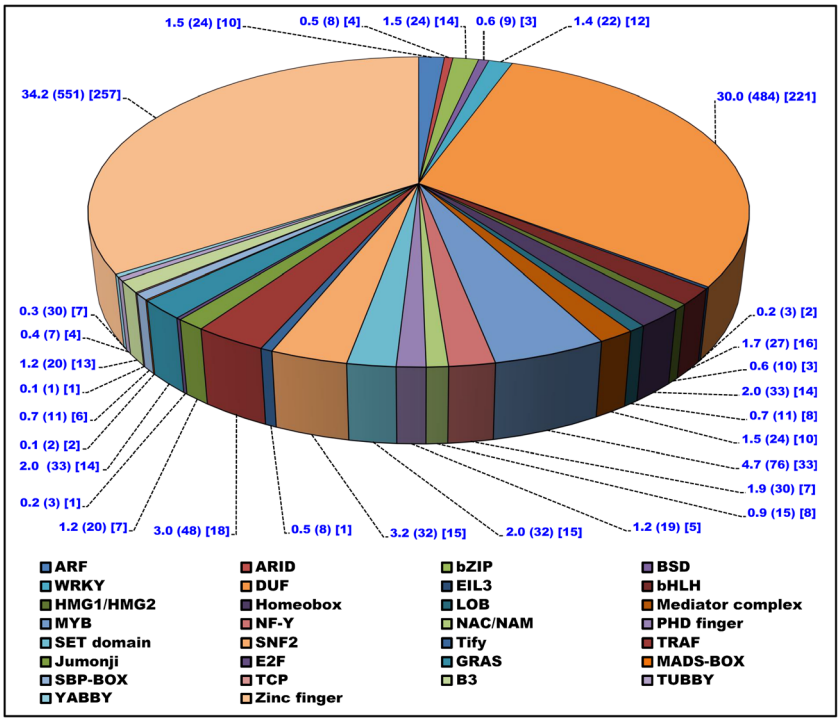

Fig. 1 Proportionate distribution of 1611 SNPs discovered from 736 TF genes representing 30 different TF gene families by their large-scale genotyping in a constituted association panel (326 *desi* and *kabuli* germplasm accessions) employing GBS and targeted gene amplicon resequencing assays. *Digits within the round and square parentheses* indicate the (number of TF gene-derived SNPs) and [number of TF genes with SNPs], respectively

204 and 218 TF genes, respectively **(**Fig. 2c**)**. The abundance of non-synonymous SNPs derived from the TF genes representing Zinc finger and ARF TF gene families was evident. The novel TF gene-based natural SNP allelic variants discovered from a diverse set of *desi* and *kabuli* chickpea germplasm accessions can be employed for multiple genomics-assisted breeding applications in chickpea. Especially, this involves marker-trait association and quick identification of functionally relevant molecular tags (markers, TFs and alleles) as well as regulatory signatures governing diverse traits of agronomic importance for marker-aided genetic enhancement of chickpea.

Association Mapping Delineates Natural SNP Allelic Variants of TF Genes Regulating Major Pod and Seed Yield Component Traits in Chickpea

The SNP markers with their desirable inherent genetic attribute of abundant genomic distribution seems to be much promising in candidate gene-based association analysis for efficient quantitative dissection of diverse complex yield component traits in chickpea (Thudi et al. 2014; Kujur et al. 2015b; Upadhyaya et al. 2015; Bajaj et al. 2016a, b). In this context, a high-resolution candidate gene-based association analysis targeting TF-encoding genes was deployed in our study to identify the potential TFs-derived natural alleles governing six major pod and seed yield component traits (DF, PH, BN, PN, SN and SW) in chickpea. For association mapping, the genotyping information of 1611 informative SNPs (differentiating the 326 *desi* and *kabuli* chickpea germplasm accessions) discovered from 736 TF genes using genome-wide GBS and targeted gene amplicon resequencing assays were utilized. The use of 1611 TF gene-based SNPs primarily in neighbour-joining phylogenetic tree and population genetic structure construction as well as principal component analysis (PCA) classified 326 *desi* and *kabuli* chickpea germplasm accessions (association panel) into two distinct population groups- POP I (173 *desi* and 29 *kabuli* accessions) and POP II (33 *desi* and 91 *kabuli* accessions), which corresponded well with our previous studies (Kujur et al. 2014, 2015a, b) (Fig. S1). A wider degree of significant population divergence (mean F_{ST} 0.47 at $P < 0.001$) was observed between POP I and POP II, while F_{ST}-led population differentiation was maximum in POP I (0.39). All the accessions with their 81% inferred ancestry were derived from one of the model-based population and rest 19% contained admixed ancestry. Maximum and minimum admixed ancestry was observed in POP II and POP I, respectively. We observed a significant deviation from population assignment of 326 germplasm accessions (representing 58 diverse geographical regions of the world) based on *desi* and *kabuli* cultivar-specific classification which was more pronounced in cultivated *kabuli* accessions (29%) belonging to POP I. This is possibly due to a greater effect of geographical origin and adaptive environment rather than cultivar-types of accessions on their assignment to a specific population group. The multiple domestication events (evolutionary bottlenecks) followed by a complex breeding history coupled with strong adaptive selection pressure might have influenced their population group assignment resulting in numerous admixtures

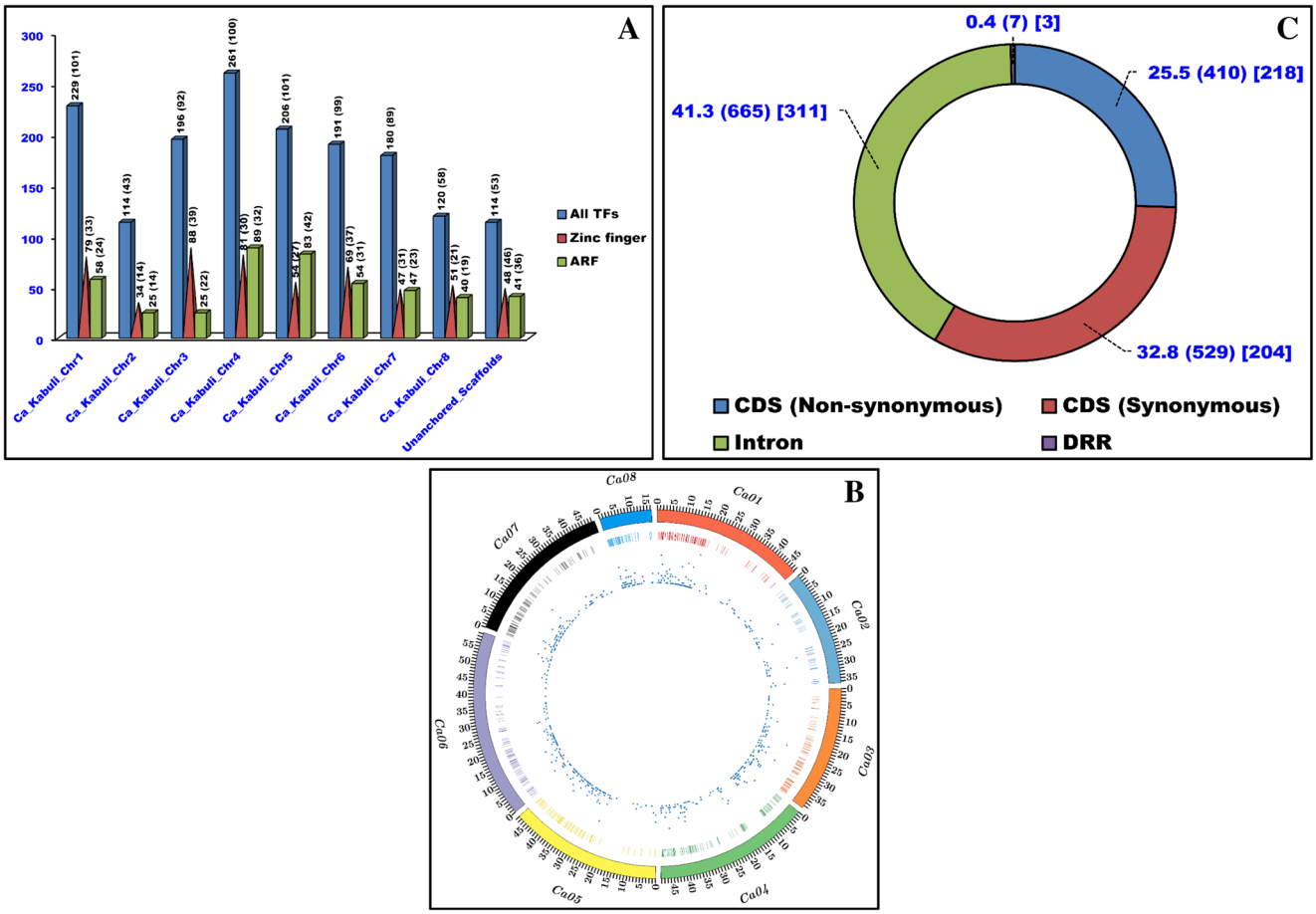

Fig. 2 **a** Frequency distribution of SNPs mined from the TF genes including two most abundant class of TF gene families [zinc finger and ARF (auxin response factor)] physically mapped on eight chromosomes and unanchored scaffolds of *kabuli* chickpea genome. Digits above the bars indicate the total number of TF gene-derived SNPs mapped. *Digits within the parentheses* represent the number of TF genes with SNPs. **b** A Circos circular ideogram depicting the genomic distribution of 1611 TF gene-derived SNPs (represented by *innermost circle*) physically mapped on eight *kabuli* chromosomes. The *outermost circles* represent the eight *kabuli* chromosomes coded with diverse colours while the inner circles denote the relative distribution of TF genes with SNPs. **c** Relative distribution and frequency of SNPs in different coding (synonymous and non-synonymous) and non-coding (intron, URR and DRR) sequence components of TF genes annotated from *kabuli* chickpea genome. The gene annotation of *kabuli* genome (Varshney et al. 2013c) was considered as a reference to deduce the CDS (coding DNA sequence)/exons, introns and DRR (downstream regulatory region) sequence components of TF genes. *Digits within the round and square parentheses* indicate the (number of TF gene-derived SNPs) and [number of TF genes with SNPs], respectively belonging to each class of coding and non-coding regions of TF genes

among accessions especially within POP II (Saxena et al. 2014b; Bajaj et al. 2015c; Das et al. 2015; Kujur et al. 2015a; Upadhyaya et al. 2015).

The comprehensive analysis of multi-location/year field phenotyping data revealed a normal frequency distribution along with a broader phenotypic variation for six major pod and seed yield component traits (DF, PH, BN, PN, SN and SW) in a constituted association panel comprising of 326 *desi* and *kabuli* chickpea germplasm accessions (Table 1). Maximum CV was observed for SN (57.2–68.2%) followed by PN (52–54%) and minimum for PH (10.2–13.9%) (Table 1). Heritability (H^2) was estimated to be ≥ 80% for all six yield traits with highest in case of SW (85–86%). This is agreed well with H^2 (80–88%) measured earlier for diverse pod and seed yield component traits using multiple geographical locations as well as several years field phenotyping data of chickpea germplasm accessions (Kujur et al. 2013, 2014, 2015a, b, c, 2016; Bajaj et al. 2015a, b, c, 2016a, b; Upadhyaya et al. 2015). A higher significant positive correlation between PN and SN ($r = 0.96$ with $P < 0.0001$) followed by SW and PH ($r = 0.61$ with $P < 0.0001$), whereas a lower negative correlation between SW and SN ($r = -0.65$ with $P < 0.0001$) followed by SN and PH ($r = -0.56$ with $P < 0.0001$) among 326 chickpea accessions based on Pearson's coefficient (r) was evident. This indicates that the 326 *desi* and *kabuli* germplasm accessions representing 58 diverse geographical regions of the world selected by us are rich in natural phenotypic diversity for all six major pod and seed yield component traits (DF, PH, BN, PN, SN and SW). Therefore, the screened germplasm lines included in a

Table 1 Diverse statistical measures estimated for six major pod and seed yield-contributing quantitative traits phenotyped in an association panel based on multi-location/years replicated field evaluation

Traits	Years	Accessions			
		Mean ± S.D.	Range	Coefficient of variation (CV %)	Heritability (%)
Days to 50% flowering (DF)	2014	56.4 ± 11.6	36.4–85.0	20.6	80
	2015	53.7 ± 9.8	35.4–82.6	18.2	80
Plant height (PH)	2014	54.7 ± 5.6	42.8–64.7	10.2	80
	2015	48.9 ± 6.8	41.7–63.5	13.9	80
Branch number (BN)	2014	19.3 ± 2.8	12.5–28.4	14.5	83
	2015	17.6 ± 3.1	13.9–29.7	17.6	82
Pod number per plant (PN)	2014	62.4 ± 33.7	21.9–204.5	54.0	81
	2015	58.5 ± 30.4	18.9–196.7	52.0	80
Seed number per plant (SN)	2014	79.0 ± 53.9	22.3–306.5	68.2	81
	2015	82.3 ± 47.1	24.7–294.7	57.2	82
100-seed weight (g) (SW)	2014	29.5 ± 13.8	5.9–70.3	46.8	86
	2015	26.4 ± 14.6	6.5–68.7	55.3	85

constituted association panel can serve as a useful genetic resource for mining novel functional allelic variants to expedite trait association mapping at a whole genome and/or gene-level in chickpea. This will essentially accelerate the delineation of functionally relevant potential genes and natural alleles regulating diverse pod and seed yield-contributing traits in domesticated *desi* and *kabuli* chickpea.

The CMLM and P3D/EMMAX-based candidate gene-based association analysis at a FDR cut-off ≤ 0.05 detected 27 TF gene-derived SNPs that were significantly associated with six major pod and seed yield component traits (DF, PH, BN, PN, SN and SW) at a $P \leq 10^{-6}$ (Fig. 3, Table 2). All these 27 trait-associated SNPs were physically mapped on seven chromosomes (except chromosome 2) and unanchored scaffolds of *kabuli* genome (Table 2). A maximum of 10 trait-associated TF gene-derived SNP loci were mapped on chromosome 3. Twenty-three and four of 27 trait-associated genic SNP loci were represented from diverse coding (20 synonymous and three non-synonymous SNPs) and non-coding (four intronic) sequence components of 16 genes, respectively (Table 2). The phenotypic variation for six major pod and seed yield component traits (DF, PH, BN, PN, SN and SW) explained by 27 TF gene-derived maximum effect SNP loci varied from 10 to 23% R^2 (P 1.7×10^{-6} to 0.3×10^{-7}) among 326 *desi* and *kabuli* chickpea germplasm accessions belonging to an association panel **(**Table 2**)**. The combined phenotypic variation for six major pod and seed yield traits explained by all significant 27 TF gene-based SNP loci was 32%. Notably, six TF gene-derived SNPs identified to be associated commonly with PN and SN traits in a constituted association panel of chickpea. One SNP derived from a TF gene encoding SNF2 (sucrose non-fermenting 2) was associated with both BN and PH traits, whereas another SNP mined from a TF gene encoding B3-domain protein was associated both SW and PH traits in chickpea. This could be due to high phenotypic correlation as observed in our study between PN vs. SN and SW vs. PH traits in a constituted association panel. Therefore, complex genetic architecture of these PH, PN, SN and SW traits was apparent, which were dissected efficiently in this study through high-resolution association mapping involving functionally relevant informative natural SNP allelic variants discovered from the TF genes of chickpea.

Association Mapping Identifies TF Gene-Based Natural Allelic Variants Governing Flowering Time in Chickpea

Four SNPs derived from the diverse coding (two non-synonymous SNPs) and intronic (two SNPs) sequence components of four TF genes exhibited significant association (13–23% R^2 with P 2.0×10^{-6} to 0.3×10^{-7}) with DF trait (Fig. 3a, Table 2). Four DF-associated TF gene-based SNPs were physically mapped on two *kabuli* chickpea chromosomes (3 and 4) with a maximum of 3 SNPs on the chromosome 3. The proportion of DF phenotypic variation explained by four SNP loci derived from four TF genes [encoding bZIP (basic leucine zipper), SBP (squamosal promoter binding protein), bHLH (basic helix-loop-helix) and Myb (myeloblastosis) TFs] in an association panel (326 *desi* and *kabuli* accessions) varied from 13 to 23% R^2 (Fig. 3a, Table 2). All four significant SNP loci in combination explained 29% DF phenotypic variation. One intronic SNP (T/A) in a bZIP TF gene and another non-synonymous SNP (A/T) causing amino acid substitutions from cysteine (TGT) to serine (AGT) in a SBP TF gene (20–23% R^2 with P 0.3–1.1×10^{-7}) exhibited strong association with DF trait in chickpea (Fig. 3a, Table 2).

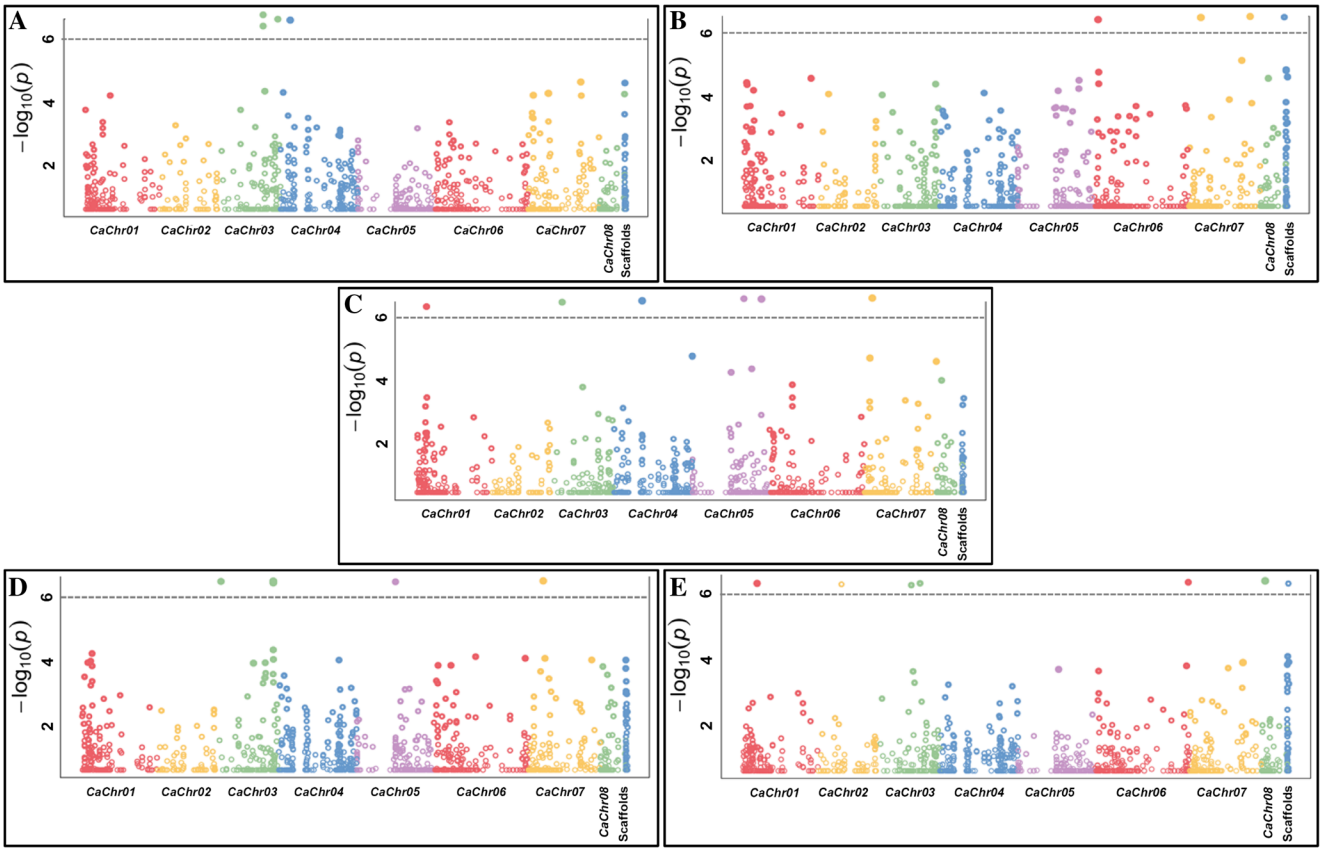

Fig. 3 GWAS-led Manhattan plot depicting significant *P* values associated with six major pod and seed yield component traits including DF (**a**), PH (**b**), BN (**c**), PN/SN (**d**) and SW (**e**) using 1611 TF gene-derived SNPs in chickpea. The genomic distribution of SNPs physically mapped on eight chromosomes and unanchored scaffolds of *kabuli* genome are denoted by the *x-axis*. The *y-axis* indicates the -$\log_{10}$ (*P*) value for significant association of SNP loci with six major yield traits. The SNPs revealing significant association with target yield traits at a cut-off *P* value $\leq 1 \times 10^{-6}$ are demarcated with a *dotted line*

Association Mapping Identifies TF Gene-Based Natural Allelic Variants Governing Plant Height in Chickpea

Four non-synonymous SNPs derived from the diverse coding sequence components of four TF genes revealing significant association (10–20% R^2 with *P* 2.3 × 10^{-6} to 1.2 × 10^{-7}) with PH trait were detected (Fig. 3b, Table 2). The PH-associated four TF gene-based SNPs were physically mapped on two *kabuli* chickpea chromosomes (6 and 7) with a maximum of 2 SNPs on the chromosome 6. The proportion of PH phenotypic variation explained by four SNP loci derived from four TF genes [encoding SNF2, WD40 (Trp-Asp 40), B3 and Tify-domain proteins] in an association panel varied from 10 to 20% R^2 (Fig. 3b, Table 2). All significant four SNP loci in combination explained 25% PH phenotypic variation. Strong association of one non-synonymous SNP (A/C) causing amino acid substitutions from histidine (CAC) to proline (CCC) in a WD40-domain protein-coding TF gene (20% R^2 with *P* 1.2 × 10^{-7}) with PH trait was observed in chickpea (Fig. 3b, Table 2).

Association Mapping Identifies TF Gene-Based Natural Allelic Variants Governing Branch Number in Chickpea

Six SNPs derived from the diverse coding (five non-synonymous SNPs) and intronic (one SNP) sequence components of six TF genes exhibited significant association (12–20% R^2 with *P* 2.4 × 10^{-6} to 0.7 × 10^{-7}) with BN trait (Fig. 3c, Table 2). Six BN-associated TF gene-based SNPs were physically mapped on five *kabuli* chickpea chromosomes (1, 3, 4, 5 and 7) with a maximum of 2 SNPs on the chromosome 5. The proportion of DF phenotypic variation explained by six SNP loci derived from six TF genes [encoding GRAS [Gibberellic acid insensitive (GAI)-Repressor of GAI (RGA)-SCARECROW (SCR)], RING-type Zinc finger, DUF827 (domain of unknown function 827), SNF2, C2H2-Zinc finger and SANT [switching-defective protein 3 (Swi3)-adaptor 2 (Ada2)-nuclear receptor co-repressor (N-CoR), transcription factor (TF)IIIB]-domain proteins] in an association panel varied from 12 to 20% R^2 (Fig. 3c, Table 2). All significant six SNP loci in combination explained 26% BN phenotypic variation. Two non-synonymous SNPs- (C/A) and (C/T) causing amino acid

Table 2 TF gene-derived SNPs significantly associated with six major pod and seed yield component traits in chickpea

TF gene identity	TF gene accession IDs	SNP IDs	*Kabuli* chromosomes	SNP physical positions (bp)	SNPs	Structural annotation	Encoded amino acids	Association analysis		
								Pod and seed yield traits associated	*P*	PVE (%)
bZIP (basic leucine zipper)	Ca07292	CakSNP439[a]	*Ca_Kabuli_Chr03*	29,844,509	[T/A]	Intron	–	DF	0.3×10^{-7}	23
SBP (squamosal promoter binding protein)	Ca01036	CakSNP510[a]	*Ca_Kabuli_Chr03*	37,101,339	[T/A]	CDS (non--synonymous)	TGT (cysteine)-AGT (serine)	DF	1.1×10^{-7}	20
bHLH (basic helix-loop-helix)	Ca07844	CakSNP559[a]	*Ca_Kabuli_Chr04*	2,213,710	[G/A]	CDS (non--synonymous)	GGA (glycine)-(AGA) arginine	DF	1.0×10^{-7}	19
Myb (myeloblastosis)	Ca07304	CakSNP443	*Ca_Kabuli_Chr03*	29,932,504	[T/A]	Intron	–	DF	2.0×10^{-6}	13
SNF2 (sucrose non-fermenting 2)	Ca13243	CakSNP1255	*Ca_Kabuli_Chr07*	7,756,455	[C/T]	CDS (non--synonymous)	CTT (leucine)-TTT (phenylalanine)	PH	1.9×10^{-6}	15
WD 40 (Trp-Asp40)	Ca17996	CakSNP1372	*Ca_Kabuli_Chr07*	41,387,475	[A/C]	CDS (non--synonymous)	CAC (histidine)-CCC (proline)	PH	1.2×10^{-7}	20
B3-domain protein	Ca23185	CakSNP1582	*Ca_Kabuli_US*	157,238	[G/C]	CDS (non--synonymous)	AGC (serine)-ACC (threonine)	PH	2.3×10^{-6}	10
Tify	Ca10431	CakSNP1031	*Ca_Kabuli_Chr06*	2,927,386	[C/G]	CDS (non--synonymous)	CTG (leucine)-GTG (valine)	PH	2.8×10^{-6}	13
GRAS [Gibberellic acid insensitive (GAI)- Repressor of GAI (RGA)- SCARECROW (SCR)]	Ca02625	CakSNP116[a]	*Ca_Kabuli_Chr01*	11,285,291	[G/T]	CDS (non--synonymous)	GCA (alanine)-TCA (serine)	BN	0.7×10^{-7}	20
RING-type Zinc finger	Ca22505	CakSNP359[a]	*Ca_Kabuli_Chr03*	14,653,193	[T/G]	Intron	–	BN	1.9×10^{-6}	15
DUF827 (domain of unknown function 827)	Ca13354	CakSNP873[a]	*Ca_Kabuli_Chr05*	28,757,467	[T/A]	CDS (non--synonymous)	GTT (valine)-GAT (aspartic acid)	BN	1.6×10^{-6}	14
SNF2 (sucrose non-fermenting 2)	Ca13243	CakSNP1255[a]	*Ca_Kabuli_Chr07*	7,756,455	[C/T]	CDS (non--synonymous)	CTT (leucine)-TTT (phenylalanine)	BN	1.0×10^{-7}	20
C2H2-Zinc finger	Ca05430	CakSNP641	*Ca_Kabuli_Chr04*	16,926,614	[T/A]	CDS (non--synonymous)	CAT (histidine)-CAA (glutamine)	BN	2.0×10^{-6}	13
SANT [switching-defective protein 3 (*Swi3*)-adaptor 2 (*Ada2*)-nuclear receptor co-repressor (*N-CoR*), transcription factor (TF) IIIB]	Ca12641	CakSNP977	*Ca_Kabuli_Chr05*	44,159,237	[C/G]	CDS (non--synonymous)	CCA (proline)-CGA (arginine)	BN	2.4×10^{-6}	12
DUF3437 (domain of unknown function 3437)	Ca17652	CakSNP847	*Ca_Kabuli_Chr05*	24,166,095	[G/A]	Intron	–	PN/SN	2.2×10^{-6}	10

Table 2 (continued)

TF gene identity	TF gene accession IDs	SNP IDs	*Kabuli* chromosomes	SNP physical positions (bp)	SNPs	Structural annotation	Encoded amino acids	Association analysis		
								Pod and seed yield traits associated	*P*	PVE (%)
LOB (lateral organ boundaries)-domain protein	Ca09342	CakSNP1282	*Ca_Kabuli_Chr07*	11,545,846	[A/G]	CDS (non--synonymous)	AAC (asparagine)-G-AC (aspartic acid)	PN/SN	0.3×10^{-7}	21
C2H2-Zinc finger	Ca00943	CakSNP498	*Ca_Kabuli_Chr03*	36,345,158	[T/C]	CDS (non--synonymous)	ATG (methionine)-G-TG (valine)	PN/SN	1.8×10^{-6}	15
WD 40 (Trp-Asp40)	Ca00947	CakSNP501	*Ca_Kabuli_Chr03*	36,387,718	[C/T]	Synonymous		PN/SN	2.0×10^{-6}	12
ZF (zinc finger)-HD (homeodomain) homeobox protein	Ca21346	CakSNP345	*Ca_Kabuli_Chr03*	3,617,555	[C/T]	CDS (non--synonymous)	GCC (alanine)-GTC (valine)	PN/SN	1.1×10^{-7}	17
Homeobox	Ca00942	CakSNP496	*Ca_Kabuli_Chr03*	36,330,750	[A/G]	CDS (non--synonymous)	AAC (asparagine)-G-AC (aspartic acid)	PN/SN	1.0×10^{-7}	16
bZIP (basic leucine zipper)	Ca02472	CakSNP124[a]	*Ca_Kabuli_Chr01*	12,794,147	[C/A]	CDS (non--synonymous)	ACG (threonine)-AA-G (lysine)	SW	0.8×10^{-7}	20
SBP (squamosal promoter binding protein)	Ca18591	CakSNP207[a]	*Ca_Kabuli_Chr01*	26,026,817	[C/T]	Synonymous	–	SW	1.1×10^{-6}	11
Jumonji	Ca11471	CakSNP1482	*Ca_Kabuli_Chr08*	8,389,117	[C/T]	CDS (non--synonymous)	GCG (alanine)-GTG (valine)	SW	1.2×10^{-7}	16
RING-type Zinc finger	Ca13705	CakSNP1179	*Ca_Kabuli_Chr06*	57,382,220	[G/A]	CDS (non--synonymous)	CGA (arginine)-CAA (glutamine)	SW	1.0×10^{-7}	15
Med12 (mediator complex)	Ca06206	CakSNP390	*Ca_Kabuli_Chr03*	23,042,249	[A/C]	CDS (non--synonymous)	TAT (tyrosine)-GAT (aspartic acid)	SW	1.5×10^{-6}	13
B3-domain protein	Ca23185	CakSNP1582	*Ca_Kabuli_US*	157,238	[G/C]	CDS (non--synonymous)	AGC (serine)-ACC (threonine)	SW	2.1×10^{-6}	11
WRKY	Ca06124	CakSNP406	*Ca_Kabuli_Chr03*	23,966,110	[G/A]	Synonymous	–	SW	1.7×10^{-6}	10

US unanchored scaffolds, *CDS* coding DNA sequence

[a] Validated by previous studies (Kujur et al. 2013, 2014, 2015b, c; Bajaj et al. 2015a, b, 2016a, b; Upadhyaya et al. 2015) using an integrated genomic strategy

substitutions from alanine (GCA) to serine (TCA) and leucine (CTT) to phenylalanine (TTT) in GRAS and SNF2 TF genes (20% R^2 with P 0.7–1.0 × 10^{-7}), respectively, revealed strong association with BN trait in chickpea (Fig. 3c, Table 2).

Association Mapping Identifies TF Gene-Based Natural Allelic Variants Governing Pod and Seed Number in Chickpea

Six SNPs derived from the diverse coding (four non-synonymous and one synonymous SNPs) and intronic (one SNP) sequence components of six TF genes revealed significant association (10–21% R^2 with P 2.2 × 10^{-6} to 0.3 × 10^{-7}) with both PN and SN traits (Fig. 3d, Table 2). Six PN/SN-associated TF gene-based SNPs were physically mapped on three *kabuli* chickpea chromosomes (3, 5 and 7) with a maximum of 4 SNPs on the chromosome 3. The proportion of PN and SN phenotypic variation explained by six SNP loci derived from six TF genes [encoding DUF3437, LOB (lateral organ boundaries)-domain protein, C2H2-Zinc finger, WD40, ZF (zinc finger)-HD (homeodomain) homeobox protein and homeobox TFs] in an association panel varied from 10 to 21% (Fig. 3d, Table 2). All significant six SNP loci in combination explained 25% PN and SN phenotypic variation. One non-synonymous SNP (A/G) causing amino acid substitutions from asparagine (AAC) to aspartic acid (GAC) in a LOB-domain protein-coding TF gene (21% R^2 with P 0.3 × 10^{-7}) exhibited strong association with PN and SN traits in chickpea (Fig. 3d, Table 2).

Association Mapping Identifies TF Gene-Based Natural Allelic Variants Governing Seed Weight in Chickpea

Seven SNPs derived from the diverse coding (five non-synonymous and two synonymous SNPs) sequence components of seven TF genes revealed significant association (10–20% R^2 with P 1.7 × 10^{-6} to 0.8 × 10^{-7}) with SW trait (Fig. 3e, Table 2). Six SW-associated TF gene-based SNPs were physically mapped on four *kabuli* chickpea chromosomes (1, 3, 6 and 8) with a maximum of 2 SNPs each on the chromosomes 1 and 3. The proportion of SW phenotypic variation explained by seven SNP loci derived from seven TF genes [encoding bZIP, SBP, Jumonji, RING-type Zinc finger, Med12 (mediator complex), B3 and WRKY-domain protein] in an association panel varied from 10 to 20% (Fig. 3e, Table 2). All significant seven SNP loci in combination explained 30% SW phenotypic variation. One non-synonymous SNP (G/T) causing amino acid substitutions from threonine (ACG) to lysine (AAG) in a bZIP TF gene (20% R^2 with P 0.8 × 10^{-7}) exhibited strong association with SW trait in chickpea (Fig. 3e, Table 2).

Especially, the non-synonymous coding SNPs in the TF genes associated with six major pod and seed yield component traits (DF, PH, BN, PN, SN and SW) delineated in our study, using high-resolution candidate gene-based association mapping, have functional significance for quantitative dissection of these complex yield traits in chickpea. This information can be useful for establishing rapid marker-trait linkages and efficient identification of potential TFs and natural allelic variants governing diverse pod and seed yield component traits in chickpea. Among six major yield trait-associated 27 TF gene-derived SNPs, merely two, four and three TFs-based SNP allelic variants governing SW, BN and DF traits, respectively delineated by us, have also been documented in our earlier studies for similar pod and yield component traits through integrating association analysis with QTL mapping, expression profiling and gene-based molecular haplotyping in chickpea (Kujur et al. 2013, 2014, 2015a, b, c, 2016; Bajaj et al. 2015a, b, c, 2016a, b; Upadhyaya et al. 2015). The validation of these molecular tags in two of our independent studies infers the functional relevance and robustness of the identified TF gene-based natural SNP allelic variants in governing major pod and seed yield traits in chickpea. The six major pod and seed yield trait-associated 27 TF genes with SNPs delineated by association mapping in chickpea are reported to be involved in transcriptional regulation of growth, development and yield traits in multiple crop plants (Manning et al. 2006; Agarwal et al. 2007, 2011; Udvardi et al. 2007; Nijhawan et al. 2008; Libault et al. 2009; Wang et al. 2011; Heang and Sassa 2012; Martínez-Andújar et al. 2012; Jones and Vodkin 2013; Hudson and Hudson 2015; Wang et al. 2015; Wang and Wang 2015; Zhang et al. 2015). Therefore, functionally relevant novel as well as earlier documented TF gene-based molecular signatures (TFs and natural SNP alleles) regulating six major yield traits delineated in our study, once comprehensively validated and characterized, will essentially be employed for marker-assisted genetic enhancement to develop high pod and seed yielding cultivars of chickpea.

Acknowledgments The authors gratefully acknowledge the financial support by the core grant of National Institute of Plant Genome Research (NIPGR), New Delhi, India.

References

Agarwal P, Arora R, Ray S, Singh AK, Singh VP, Takatsuji H, Kapoor S, Tyagi AK (2007) Genome-wide identification of C2H2 zinc-finger gene family in rice and their phylogeny and expression analysis. Plant Mol Biol 65:467–485

Agarwal P, Kapoor S, Tyagi AK (2011) Transcription factors regulating the progression of monocot and dicot seed development. BioEssays 33:189–202

Bajaj D, Upadhyaya HD, Khan Y, Das S, Badoni S, Shree T, Kumar V, Tripathy S, Gowda CLL, Singh S, Sharma S, Tyagi AK, Chattopadhyay D, Parida SK (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, Das S, Badoni S, Tripathi S, Upadhyaya HD, Gowda CLL, Sharma S, Singh S, Tyagi AK, Parida SK (2015b) Genome-wide conserved non-coding microsatellite (CNMS) marker-based integrative genetical genomics for quantitative dissection of seed weight in chickpea. J Exp Bot 66:1271–1290

Bajaj D, Das S, Badoni S, Kumar V, Singh M, Bansal KC, Tyagi AK, Parida SK (2015c) Genome-wide high-throughput SNP discovery and genotyping for understanding natural (functional) allelic diversity and domestication patterns in wild chickpea. Sci Rep 5:11627

Bajaj D, Upadhyaya HD, Das S, Kumar V, Gowda CLL, Sharma S, Sharma S, Tyagi AK, Parida SK (2016a) Identification of candidate genes for dissecting complex branch number trait in chickpea. Plant Sci 245:61–70

Bajaj D, Srivastava R, Tripathi S, Bharadwaj C, Upadhyaya HD, Tyagi AK, Parida SK (2016b) EcoTILLING-based association mapping efficiently delineates functionally relevant natural allelic variants of candidate genes governing agronomic traits in chickpea. Front Plant Sci 7:450

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc B 57:289–300

Century K, Reuber TL, Ratcliffe OJ (2008) Regulating the regulators: the future prospects for transcription-factor-based agricultural biotechnology products. Plant Physiol 147:20–29

Das S, Upadhyaya HD, Bajaj D, Kujur A, Badoni S, Laxmi, Kumar V, Tripathi S, Gowda CLL, Sharma S, Singh S, Tyagi AK, Parida SK (2015) Deploying QTL-seq for rapid delineation of a potential candidate gene underlying major trait-associated QTL in chickpea. DNA Res 22:193–203

Das S, Singh M, Srivastava R, Bajaj D, Saxena MS, Rana JC, Bansal KC, Tyagi AK, Parida SK (2016) mQTL-seq delineates functionally relevant candidate gene harbouring a major QTL regulating pod number in chickpea. DNA Res 23:53–65

Deokar AA, Ramsay L, Sharpe AG, Diapari M, Sindhu A, Bett K, Warkentin TD, Tar'an B (2014) Genome wide SNP identification in chickpea for use in development of a high density genetic map and improvement of chickpea reference genome assembly. BMC Genomics 15:708

Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, Mitchell SE (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PLoS One 6:e19379

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

Garg R, Bhattacharjee A, Jain M (2014) Genome-scale transcriptomic insights into molecular aspects of abiotic stress responses in chickpea. Plant Mol Biol Rep 33:388–400

Gaur R, Azam S, Jeena G, Khan AW, Choudhary S, Jain M, Yadav G, Tyagi AK, Chattopadhyay D, Bhatia S (2012) High-throughput SNP discovery and genotyping for constructing a saturated linkage map of chickpea (*Cicer arietinum* L.) DNA Res 19:357–373

Gaur PM, Samineni S, Tripathi S, Varshney RK, Gowda CLL (2014) Allelic relationships of flowering time genes in chickpea. Euphytica 203:295–308

Gaur R, Jeena G, Shah N, Gupta S, Pradhan S, Tyagi AK, Jain M, Chattopadhyay D, Bhatia S (2015) High density linkage mapping of genomic and transcriptomic SNPs for synteny analysis and anchoring the genome sequence of chickpea. Sci Rep 5:13387

Gupta S, Nawaz K, Parween S, Roy R, Sahu K, Pole AK, Khandal H, Srivastava R, Kumar 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

Hardy OJ, Vekemans X (2002) SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. Mol Ecol Res 2:618–620

Heang D, Sassa H (2012) Antagonistic actions of HLH/bHLH proteins are involved in grain length and weight in rice. PLoS One 7:e31325

Hiremath PJ, Kumar A, Penmetsa RV, Farmer A, Schlueter JA, Chamarthi SK, Whaley AM, Carrasquilla-Garcia N, Gaur PM, Upadhyaya HD, Kavi Kishor PB, Shah TM, Cook DR, Varshney RK (2012) Large-scale development of cost-effective SNP marker assays for diversity assessment and genetic mapping in chickpea and comparative mapping in legumes. Plant Biotechnol J 10:716–732

Hudson KA, Hudson ME (2015) A classification of basic helix-loop-helix transcription factors of soybean. Int J Genomics 2015:603182

Jaganathan D, Thudi M, Kale S, Azam S, Roorkiwal M, Gaur PM, Kishor PB, Nguyen H, Sutton T, Varshney RK (2015) Genotyping-by-sequencing based intra-specific genetic map refines a "QTL-hotspot" region for drought tolerance in chickpea. Mol Gen Genomics 290:559–571

Jain M, Misra G, Patel RK, Priya P, Jhanwar S, Khan AW, Shah N, Singh VK, Garg R, Jeena G, Yadav M, Kant C, Sharma P, Yadav G, Bhatia S, Tyagi AK, Chattopadhyay D (2013) A draft genome sequence of the pulse crop chickpea (*Cicer arietinum* L.) Plant J 74:715–729

Jones SI, Vodkin LO (2013) Using RNA-Seq to profile soybean seed development from fertilization to maturity. PLoS One 8:e59270

Kale SM, Jaganathan D, Ruperao P, Chen C, Punna R, Kudapa H, Thudi M, Roorkiwal M, Katta MA, Doddamani D, Garg V, Kishor PB, Gaur PM, Nguyen HT, Batley J, Edwards D, Sutton T, Varshney RK (2015) Prioritization of candidate genes in "QTL-hotspot" region for drought tolerance in chickpea (*Cicer arietinum* L.) Sci Rep 5:15296

Kang J, Hwang JU, Lee M, Kim YY, Assmann SM, Martinoia E, Lee Y (2010) PDR-type ABC transporter mediates cellular uptake of the phytohormone abscisic acid. Proc Natl Acad Sci U S A 107:2355–2360

Kudapa H, Azam S, Sharpe AG, Tar'an B, Li R, Deonovic B, Cameron C, Farmer AD, Cannon SB, Varshney RK (2014) Comprehensive transcriptome assembly of chickpea (*Cicer arietinum* L.) using sanger and next generation sequencing platforms: development and applications. PLoS One 9:e86039

Kujur A, Bajaj D, Saxena MS, Tripathi S, Upadhyaya HD, Gowda CLL, Singh S, Jain M, Tyagi AK, Parida SK (2013) Functionally relevant microsatellite markers from chickpea transcription factor genes for efficient genotyping applications and trait association mapping. DNA Res 20:355–373

Kujur A, Bajaj D, Saxena MS, Tripathi S, Upadhyaya HD, Gowda CLL, Singh S, Tyagi AK, Jain M, Parida SK (2014) An efficient and cost-effective approach for genic microsatellite marker-based large-scale trait association mapping: identification of candidate genes for seed weight in chickpea. Mol Breed 34:241–265

Kujur A, Bajaj D, Upadhyaya HD, Das S, Ranjan R, Shree T, Saxena MS, Badoni S, Kumar V, Tripathi S, Gowda CLL, Sharma S, Singh S, Tyagi AK, Parida SK (2015a) 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, Shree T, Bajaj D, Das S, Saxena MS, Badoni S, Kumar V, Tripathy S, Gowda CLL, Sharma S, Singh S, Tyagi AK, Parida SK (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, Das S, Ranjan R, Shree T, Saxena MS, Badoni S, Kumar V, Tripathi S, Gowda CLL, Sharma S, Singh S, Tyagi AK, Parida SK (2015c) A genome-wide SNP scan accelerates trait-regulatory genomic loci identification in chickpea. Sci Rep 5:11166

Kujur A, Upadhyaya HD, Bajaj D, Gowda CLL, 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

Kumar A, Choudhary AK, Solanki RK, Pratap A (2011) Towards marker-assisted selection in pulses: a review. Plant Breed 130: 297–313

Kumar V, Singh A, Amitha Mithra SV, Krishnamurthy SL, Parida SK, Jain S, Tiwari KK, Kumar P, Rao AR, Sharma SK, Khurana JP, Singh NK, Mohapatra T (2015) Genome-wide association mapping of salinity tolerance in rice (*Oryza sativa*). DNA Res 22:133–145

Libault M, Joshi T, Benedito VA, Xu D, Udvardi MK, Stacey G (2009) Legume transcription factor genes: what makes legumes so special? Plant Physiol 151:991–1001

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

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

Manning K, Tör M, Poole M, Hong Y, Thompson AJ, King GJ, Giovannoni JJ, Seymour GB (2006) A naturally occurring epigenetic mutation in a gene encoding an SBP-box transcription factor inhibits tomato fruit ripening. Nature Genet 38:948–952

Martínez-Andújar C, Martin RC, Nonogaki H (2012) Seed traits and genes important for translational biology-highlights from recent discoveries. Plant Cell Physiol 53:5–15

Nijhawan A, Jain M, Tyagi AK, Khurana JP (2008) Genomic survey and gene expression analysis of the basic leucine zipper transcription factor family in rice. Plant Physiol 146:333–350

Parween S, Nawaz K, Roy R, Pole AK, Venkata Suresh B, Misra G, Venkata Suresh B, Misra G, Jain M, Yadav G, Parida SK, Tyagi AK, Bhatia S, Chattopadhyay D (2015) An advanced draft genome assembly of a *desi* type chickpea (*Cicer arietinum* L.) Sci Rep 5:12806

Pradhan S, Bandhiwal N, Shah N, Kant C, Gaur R, Bhatia S (2014) Global transcriptome analysis of developing chickpea (*Cicer arietinum* L.) seeds. Front Plant Sci 5:698

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

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

Singh VK, Khan AW, Jaganathan D, Thudi M, Roorkiwal M, Takagi H, Garg V, Kumar V, Chitikineni A, Gaur PM, Sutton T, Terauchi R, Varshney RK (2016) QTL-seq for rapid identification of candidate genes for 100-seed weight and root/total plant dry weight ratio under rainfed conditions in chickpea. Plant Biotechnol J 14:2110–2119

Thudi M, Upadhyaya HD, Rathore A, Gaur PM, Krishnamurthy L, Roorkiwal M, Nayak SN, Chaturvedi SK, Basu PS, Gangarao NV, Fikre A, Kimurto P, Sharma PC, Sheshashayee MS, Tobita S, Kashiwagi J, Ito O, Killian A, Varshney RK (2014) Genetic dissection of drought and heat tolerance in chickpea through genome-wide and candidate gene-based association mapping approaches. PLoS One 9:e96758

Udvardi MK, Kakar K, Wandrey M, Montanari O, Murray J, Andriankaja A, Zhang JY, Benedito V, Hofer JM, Chueng F, Town CD (2007) Legume transcription factors: global regulators of plant development and response to the environment. Plant Physiol 144:538–549

Upadhyaya HD, Bajaj D, Das S, Saxena MS, Badoni S, Kumar V, Tripathi S, Sharma S, Tyagi AK, Parida SK (2015) A genome-scale integrated approach aids in genetic dissection of complex flowering time trait in chickpea. Plant Mol Biol 89:403–420

Varshney RK (2016) Exciting journey of 10 years from genomes to fields and markets: some success stories of genomics-assisted breeding in chickpea, pigeon pea and groundnut. Plant Sci 242:98–107

Varshney RK, Mohan SM, Gaur PM, Gangarao NV, Pandey MK, Bohra A, Sawargaonkar SL, Chitikineni A, Kimurto PK, Janila P, Saxena KB, Fikre A, Sharma M, Rathore A, Pratap A, Tripathi S, Datta S, Chaturvedi SK, Mallikarjuna N, Anuradha G, Babbar A, Choudhary AK, Mhase MB, Bharadwaj C, Mannur DM, Harer PN, Guo B, Liang X, Nadarajan N, Gowda CLL (2013a) Achievements and prospects of genomics-assisted breeding in three legume crops of the semi-arid tropics. Biotechnol Adv 31:1120–1134

Varshney RK, Gaur PM, Chamarthi SK, Krishnamurthy L, Tripathi S, Kashiwagi J, Samineni S, Singh VK, Thudi M, Jaganthan D (2013b) Fast-track introgression of "QTL-hotspot" for root traits and other drought tolerance traits in JG 11, an elite and leading variety of chickpea. Plant Genome 6:1–26

Varshney RK, Song C, Saxena RK et al (2013c) Draft genome sequence of chickpea (*Cicer arietinum*) provides a resource for trait improvement. Nat Biotechnol 31:240–246

Varshney RK, Thudi M, Nayak SN, Gaur PM, Kashiwagi J, Krishnamurthy L, Jaganathan D, Koppolu J, Bohra A, Tripathi S, Rathore A, Jukanti AK, Jayalakshmi V, Vemula A, Singh SJ, Yasin M, Sheshshayee MS, Viswanatha KP (2014) Genetic dissection of drought tolerance in chickpea (*Cicer arietinum* L.) Theor Appl Genet 127:445–462

Wang H, Wang H (2015) The miR156/SPL module, a regulatory hub and versatile toolbox, gears up crops for enhanced agronomic traits. Mol Plant 8:677–688

Wang J, Zhou J, Zhang B, Vanitha J, Ramachandran S, Jiang SY (2011) Genome-wide expansion and expression divergence of the basic leucine zipper transcription factors in higher plants with an emphasis on sorghum. J Integr Plant Biol 53:212–231

Wang Z, Cheng K, Wan L, Yan L, Jiang H, Liu S, Lei Y, Liao B (2015) Genome-wide analysis of the basic leucine zipper (bZIP) transcription factor gene family in six legume genomes. BMC Genomics 16:1053

Yu JK, Paik H, Choi JP, Han JH, Choe JK, Hur CG (2010) Functional domain marker (FDM): an in silico demonstration in *Solanaceae* using simple sequence repeats (SSRs). Plant Mol Biol Rep 28: 352–356

Zhang Z, Ersoz E, Lai CQ, Todhunter RJ, Tiwari HK, Gore MA, Bradbury PJ, Yu J, Arnett DK, Ordovas JM, Buckler ES (2010) Mixed linear model approach adapted for genome-wide association studies. Nature Genet 42:355–368

Zhang SD, Ling LZ, Yi TS (2015) Evolution and divergence of SBP-box genes in land plants. BMC Genomics 16:787