Genome-wide generation and genotyping of informative SNPs to scan molecular signatures for seed yield in chickpea

Udita Basu1, Rishi Srivastava1, Deepak Bajaj1, Virevol Thakro1, Anurag Daware4, Naveen Malik1, Hari D. Upadhyaya2 & Swarup K. Parida1

We discovered 2150 desi and 2199 kabuli accessions-derived SNPs by cultivar-wise individual assembling of sequence-reads generated through genotyping-by-sequencing of 92 chickpea accessions. Subsequent large-scale validation and genotyping of these SNPs discovered 619 desi accessions-derived (DAD) SNPs, 531 kabuli accessions-derived (KAD) SNPs, 884 multiple accessions-derived (MAD) SNPs and 1083 two accessions (desi ICC 4958 and kabuli CDC Frontier)-derived (TAD) SNPs that were mapped on eight chromosomes. These informative SNPs were annotated in coding/non-coding regulatory sequence components of genes. The MAD-SNPs were efficient to detect high intra-specific polymorphic potential and wide natural allelic diversity level including high-resolution admixed-population genetic structure and precise phylogenetic relationship among 291 desi and kabuli accessions. This signifies their effectiveness in introgression breeding and varietal improvement studies targeting useful agronomic traits of chickpea. Six trait-associated genes with SNPs including quantitative trait nucleotides (QTNs) in combination explained 27.5% phenotypic variation for seed yield per plant (SYP). A pentatricopeptide repeat (PPR) gene with a synonymous-coding SNP/QTN significantly associated with SYP trait was found most-promising in chickpea. The essential information delineated can be of immense utility in genomics-assisted breeding applications to develop high-yielding chickpea cultivars.

Chickpea (Cicer arietinum) is a self-pollinated, diploid (2n = 16) and economically important legume food crop rich in human dietary proteins1. The chickpea is primarily represented by its desi and kabuli cultivars exhibiting a distinct differentiation in both agro-morphological and architecture traits, which genomes have been sequenced recently2–5. Substantial enhancement of its yield and productivity by developing high-yielding abiotic stress tolerant and climate-ready cultivars is essential at present to sustain the global food security. To accomplish these objectives, many significant efforts involving classical genetics as well as genomics-assisted breeding strategies like quantitative trait loci (QTL) mapping, fine-mapping/map-based cloning and association analysis have been made to decipher the inheritance pattern and quantitative dissection of complex yield and stress tolerance traits for genetic improvement of chickpea6–30. However, these are constrained by narrow genetic base and low intra-specific marker polymorphism among desi and kabuli accessions due to combined strong impact of four major domestication bottlenecks in chickpea as compared to other food crop plants2,3,31–35. To overcome the aforesaid limitations, development and high-throughput genotyping of large-scale sequence-based informative markers like single nucleotide polymorphism (SNPs) exhibiting a higher potential of polymorphism among cultivated desi and kabuli accessions is a prerequisite for their use in genomics-assisted breeding applications and genetic improvement of chickpea.

The next-generation sequencing (NGS)-led genome and transcriptome sequencing of multiple cultivated (desi and kabuli) and wild accessions are found efficient in fast discovery of numerous SNPs at a genome-wide scale in chickpea3,4,13,14,30,36–41. The large-scale validation and high-throughput genotyping of these SNPs mined from limited number of sequenced accessions using diverse array-based genotyping assays commonly exhibit

1National Institute of Plant Genome Research (NIPGR), Aruna Asaf Ali Marg, New Delhi, 110067, India. 2International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru, 502324, Telangana, India. Udita Basu and Rishi Srivastava contributed equally. Correspondence and requests for materials should be addressed to S.K.P. (email: swarup@nipgr.ac.in)
less potential of inter-/intra-specific polymorphism and natural allelic diversity among cultivated and wild accesses to be deployed efficiently for genetic enhancement of chickpea.23-30,14,17-23,25-27. However, degree of these polymorphic and diversity potential can be enriched by using the genotyping information of SNPs derived from whole genome resequencing data that are generated from a large number of desi, kabuli and wild accesses of chickpea. Essentially, this approach requires huge cost, time and resources for developing and genotyping of SNPs at a genome/gene level in chickpea. In this regard, NGS-based low sequencing-depth genome coverage sequencing strategies like genotyping-by-sequencing (GBS) assay seems quite promising for rapid discovery and genotyping of genome-wide SNPs simultaneously among numerous cultivated (desi and kabuli) and wild accesses in order to drive high-throughput genetic analysis in chickpea with sub-optimal expense of cost, labour and resources.10,14,17-23,25-27. However, these low genome coverage sequencing assays often generate erroneous and non-uniform SNP genotyping data across accesses, thereby further large-scale revalidation and genotyping of GBS-derived SNPs is needed currently prior to their use in genetic improvement of chickpea. In this perspective, high-throughput validation and genotyping of numerous GBS-derived high-quality SNPs discovered by assembling all the sequence reads generated from large number of accesses belonging to each of the desi and kabuli cultivars individually (cultivar-wise) can be useful for their deployment in genomics-assisted crop improvement of chickpea. In view of aforementioned prospects, efforts were made in the current study for large-scale validation and high-throughput genotyping of informative SNPs, discovered by desi and kabuli cultivar-wise individual assembling of the sequence reads, generated from 92 accesses using GBS assay at a genome-wide scale in chickpea (Fig. 1). The efficacy of these validated whole genome/gene-derived SNPs to detect intra-specific polymorphism, molecular diversity and domestication pattern among 291 desi and kabuli cultivated chickpea accesses and for establishing marker-trait association with seed yield per plant were determined.

Results
Genomic distribution of informative SNPs in chickpea. The sequence reads generated by sequencing and genotyping of 92 desi and kabuli chickpea accesses using GBS assay were assembled cultivar-wise individually. This exertion successfully discovered various types of informative SNPs including desi accesses-derived (DAD)-, kabuli accesses-derived (KAD)- and multiple accesses-derived (MAD)-SNPs which were further compared with two accesses (ICC 4958 and CDC Frontier) derived (TAD) SNPs for evaluating the potential of these developed markers in large-scale genetic analysis in chickpea (Fig. 1).

To develop these informative markers primarily, the sequencing data generated from 92 including 39 desi and 53 kabuli chickpea accesses using a GBS assay were de-multiplexed and quality filtered. As a result, 207.9 million processed high-quality sequence reads uniformly distributed across 92 chickpea accesses were obtained which altogether covered 28-fold (X) sequence depth of coverage. This exertion overall produced 3.3 to 5.7 million of high-quality sequence reads per accession covering on an average 5.4-fold (X) sequence depth of coverage. More than 85% of de-multiplexed high-quality sequence reads generated from accesses were mapped separately (desi and kabuli cultivar-wise) to unique physical locations on desi reference genome. The desi and kabuli cultivar-wise reference genome-based GBS analysis altogether detected 2150 (DAD-SNPs) and 2199 (KAD-SNPs) high-quality SNPs (supported by read-depth ≥10, SNP base quality ≥20, <1% missing data and ~2% heterozygosity in each accession) from the accesses which were further validated through PCR amlicons-based sequencing assay. For large-scale validation and high-throughput genotyping, a representative set of high-quality 672 DAD-, 576 KAD- and 1152 TAD-SNPs differentiating 39 desi, 53 kabuli and 2 (ICC 4958 and CDC Frontier) accesses, respectively, were selected based on their uniform physical distribution on eight chromosomes. These physically mapped SNPs were genotyped using the genomic DNA of 291 desi and kabuli chickpea accesses through a 34-plex high-throughput matrix assisted laser desorption ionization (MALDI-TOF) Sequenom Mass Array genotyping assay. The optimized multiplex assay successfully validated 619 DAD-, 531 KAD- and 1083 TAD-SNPs based on their homozygous and heterozygous allele discrimination as evident from the call cluster plots of low and high mass homo- and hetero-zygote yields with a 92–94% genotyping success rate (Tables 1 and S1–S4). The comparison of both DAD- and KAD-SNPs identified 884 MAD-SNPs differentiating all 92 desi and kabuli accesses together which were physically mapped on eight chickpea chromosomes (Tables 1 and S3). Therefore, all the DAD- and KAD-SNPs selected in this study were common to MAD-SNPs. However, only 266 SNPs were common between DAD- and KAD-SNPs while the remaining 353 DAD- and 265 KAD-SNPs were found to be unique. We could not obtain any DAD-, KAD and MAD-SNPs that are common to TAD-SNPs.

The detailed structural annotation of 884 MAD- and 1083 TAD-SNPs exhibited the occurrence of 664 (75.1%) and 837 (77.3%) SNPs in 436 and 461 chickpea genes, respectively (Fig. 2 and Tables S1–S4). The remaining 220 (24.9%) MAD- and 246 (22.7%) TAD-SNPs were present in the intergenic regions of chickpea genome. Both MAD- and TAD-SNPs (282–364 SNPs, 42.5–43.5%) especially discovered from the exons/CDS of chickpea genes (214–224) were found most abundant (Fig. 2 and Tables S1–S4). Among the coding MAD- and TAD-SNPs, a higher proportion (72–78.3% of SNPs (203–285) in the genes (151–189) showed synonymous substitutions while remaining (21.7–28.0%) SNPs (76–78) in the genes (61–70) revealed both missense and large-effect nonsense substitutions (Fig. 2 and Tables S1–S4). The functional annotation of 436 and 461 genes with 664 MAD- and 837 TAD-SNPs, respectively, exhibited their highest correspondence to growth, development and metabolism-related proteins (61%) followed by transcription factors (20%) and signal transduction proteins (11%). Moreover, GO enrichment analysis of these genes with MAD- and TAD-SNPs revealed a significant overrepresentation/enrichment of GO terms in the genes associated with molecular function (nucleic acid binding, 27%, P: 1.2 × 10−40) followed by biological process (metabolism, 11.5%, 1.6 × 10−50) and cellular component (macromolecular complex, 5.7%, 1.0 × 10−40). The transcription factor (TF) genes with MAD- and TAD-SNPs belonging to TF families like basic helix-loop-helix (bHLH), basic leucine zipper (bZIP) and APETALA2 ethylene-responsive element binding proteins (AP2-EREBP) were found predominant.
Natural allelic diversity and population genetic structure among desi and kabuli chickpea accessions. A set of 619 DAD-, 531 KAD-, 884 MAD- and 1083 TAD-SNPs were analysed to evaluate their potential of detecting polymorphism among 291 desi and kabuli chickpea accessions (Table 2). The 884 MAD-SNPs detected a higher degree of polymorphism as compared to that of DAD-, KAD- and TAD-SNPs among accessions which was clearly evident from their maximum level of PIC (varied from 0.10 to 0.48, mean 0.45) and high nucleotide diversity (mean θπ: 1.55) estimation (Table 2). Notably, 526 DAD-, 449 KAD-, 793 MAD- and 758 TAD-SNPs revealed high intra-specific polymorphic potential among 189 desi accessions (70–89.7%)

Figure 1. Strategy adopted in the present study to generate high-quality genotyping data of diverse kinds of informative SNPs (DAD-, KAD-, MAD- and TAD-SNPs) from 291 desi and kabuli accessions at a genome-wide scale to be deployed for multiple genomics-assisted breeding applications in chickpea.
polymorphism, mean PIC: 0.30–0.47 and mean $\theta_\pi$: 1.29–1.59) than that of 102 kabuli accessions (64.3–73.3%, 0.27–0.41 and 1.23–1.48). Notably, a higher potential of MAD-SNPs followed by DAD-SNPs and a lower potential of TAD-SNPs for detecting polymorphism among accessions belonging to desi and kabuli cultivars was observed (Table 2).

The molecular diversity and population genetic structure among 291 desi and kabuli chickpea accessions were determined using the same set of SNPs analysed earlier. High natural allelic diversity based on a broader range of genetic distance was detected especially by MAD-SNPs that varied from 0.12 to 0.79, with an average of 0.67 (Fig. 3 and Table 2). All DAD-, KAD-, MAD- and TAD-SNPs exhibited a relatively wider range of allelic diversity in 189 desi chickpea accessions than that of 102 kabuli accessions. As compared to others, MAD-SNPs had maximum potential to detect a wider range of genetic distance and thus revealed high natural allelic diversity among accessions belonging to desi (mean genetic distance: 0.69) and kabuli (0.58) chickpea (Fig. 3 and Table 2). TAD-SNPs detected low allelic diversity which is reflected from narrow range of genetic distance among accessions within desi (mean genetic distance: 0.39) and kabuli (0.31) chickpea.

The population genetic structure among 291 accessions with the genotyping information of aforesaid SNPs classified all these accessions into two distinct population groups at a K (population number) value of 2 as POP and POP II (Fig. 3). The POP I included 73 kabuli and 29 desi accessions while POP II comprised of 129 desi and 60 kabuli accessions. The K value outcome with the best replicate was ascertained by estimating the average LnP(D) (log-likelihood) (exhibiting the greatest apparent inflection) and the second order statistics of $\Delta K$ (showing a sharp peak with maximum value of $\Delta K$). All these SNPs detected a higher polymorphic potential in POP II (mean PIC: 0.31–0.43) than that of POP I. A significant population divergence based on pair-wise FST polymorphism, mean PIC: 0.30–0.47 and mean $\theta_\pi$: 1.29–1.59) than that of 102 kabuli accessions (64.3–73.3%, 0.27–0.41 and 1.23–1.48). Notably, a higher potential of MAD-SNPs followed by DAD-SNPs and a lower potential of TAD-SNPs for detecting polymorphism among accessions belonging to desi and kabuli cultivars was observed (Table 2).

The molecular diversity and population genetic structure among 291 desi and kabuli chickpea accessions were determined using the same set of SNPs analysed earlier. High natural allelic diversity based on a broader range of genetic distance was detected especially by MAD-SNPs that varied from 0.12 to 0.79, with an average of 0.67 (Fig. 3 and Table 2). All DAD-, KAD-, MAD- and TAD-SNPs exhibited a relatively wider range of allelic diversity in 189 desi chickpea accessions than that of 102 kabuli accessions. As compared to others, MAD-SNPs had maximum potential to detect a wider range of genetic distance and thus revealed high natural allelic diversity among accessions belonging to desi (mean genetic distance: 0.69) and kabuli (0.58) chickpea (Fig. 3 and Table 2). TAD-SNPs detected low allelic diversity which is reflected from narrow range of genetic distance among accessions within desi (mean genetic distance: 0.39) and kabuli (0.31) chickpea.

The population genetic structure among 291 accessions with the genotyping information of aforesaid SNPs classified all these accessions into two distinct population groups at a K (population number) value of 2 as POP and POP II (Fig. 3). The POP I included 73 kabuli and 29 desi accessions while POP II comprised of 129 desi and 60 kabuli accessions. The K value outcome with the best replicate was ascertained by estimating the average LnP(D) (log-likelihood) (exhibiting the greatest apparent inflection) and the second order statistics of $\Delta K$ (showing a sharp peak with maximum value of $\Delta K$). All these SNPs detected a higher polymorphic potential in POP II (mean PIC: 0.31–0.43) than that of POP I. A significant population divergence based on pair-wise FST polymorphism, mean PIC: 0.30–0.47 and mean $\theta_\pi$: 1.29–1.59) than that of 102 kabuli accessions (64.3–73.3%, 0.27–0.41 and 1.23–1.48). Notably, a higher potential of MAD-SNPs followed by DAD-SNPs and a lower potential of TAD-SNPs for detecting polymorphism among accessions belonging to desi and kabuli cultivars was observed (Table 2).

The molecular diversity and population genetic structure among 291 desi and kabuli chickpea accessions were determined using the same set of SNPs analysed earlier. High natural allelic diversity based on a broader range of genetic distance was detected especially by MAD-SNPs that varied from 0.12 to 0.79, with an average of 0.67 (Fig. 3 and Table 2). All DAD-, KAD-, MAD- and TAD-SNPs exhibited a relatively wider range of allelic diversity in 189 desi chickpea accessions than that of 102 kabuli accessions. As compared to others, MAD-SNPs had maximum potential to detect a wider range of genetic distance and thus revealed high natural allelic diversity among accessions belonging to desi (mean genetic distance: 0.69) and kabuli (0.58) chickpea (Fig. 3 and Table 2). TAD-SNPs detected low allelic diversity which is reflected from narrow range of genetic distance among accessions within desi (mean genetic distance: 0.39) and kabuli (0.31) chickpea.

The population genetic structure among 291 accessions with the genotyping information of aforesaid SNPs classified all these accessions into two distinct population groups at a K (population number) value of 2 as POP and POP II (Fig. 3). The POP I included 73 kabuli and 29 desi accessions while POP II comprised of 129 desi and 60 kabuli accessions. The K value outcome with the best replicate was ascertained by estimating the average LnP(D) (log-likelihood) (exhibiting the greatest apparent inflection) and the second order statistics of $\Delta K$ (showing a sharp peak with maximum value of $\Delta K$). All these SNPs detected a higher polymorphic potential in POP II (mean PIC: 0.31–0.43) than that of POP I. A significant population divergence based on pair-wise FST polymorphism, mean PIC: 0.30–0.47 and mean $\theta_\pi$: 1.29–1.59) than that of 102 kabuli accessions (64.3–73.3%, 0.27–0.41 and 1.23–1.48). Notably, a higher potential of MAD-SNPs followed by DAD-SNPs and a lower potential of TAD-SNPs for detecting polymorphism among accessions belonging to desi and kabuli cultivars was observed (Table 2).

The molecular diversity and population genetic structure among 291 desi and kabuli chickpea accessions were determined using the same set of SNPs analysed earlier. High natural allelic diversity based on a broader range of genetic distance was detected especially by MAD-SNPs that varied from 0.12 to 0.79, with an average of 0.67 (Fig. 3 and Table 2). All DAD-, KAD-, MAD- and TAD-SNPs exhibited a relatively wider range of allelic diversity in 189 desi chickpea accessions than that of 102 kabuli accessions. As compared to others, MAD-SNPs had maximum potential to detect a wider range of genetic distance and thus revealed high natural allelic diversity among accessions belonging to desi (mean genetic distance: 0.69) and kabuli (0.58) chickpea (Fig. 3 and Table 2). TAD-SNPs detected low allelic diversity which is reflected from narrow range of genetic distance among accessions within desi (mean genetic distance: 0.39) and kabuli (0.31) chickpea.

The population genetic structure among 291 accessions with the genotyping information of aforesaid SNPs classified all these accessions into two distinct population groups at a K (population number) value of 2 as POP and POP II (Fig. 3). The POP I included 73 kabuli and 29 desi accessions while POP II comprised of 129 desi and 60 kabuli accessions. The K value outcome with the best replicate was ascertained by estimating the average LnP(D) (log-likelihood) (exhibiting the greatest apparent inflection) and the second order statistics of $\Delta K$ (showing a sharp peak with maximum value of $\Delta K$). All these SNPs detected a higher polymorphic potential in POP II (mean PIC: 0.31–0.43) than that of POP I. A significant population divergence based on pair-wise FST...
(P < 0.001) was estimated as 0.69 between POP I and POP II. Highest $F_{ST}$-based population differentiation of 0.30 was observed within POP II. All 291 desi and kabuli chickpea accessions clearly belonged to a structured population in which 69% of their inferred ancestry was derived from one model-based population and the remaining 31% had an admixed ancestry (Fig. 3). Maximum admixed ancestry was observed in POP II (28%) and minimum in POP I (17%). Admixed ancestry was estimated as 39% between POP I and POP II. We could not detect admixture in 19.2% of chickpea accessions while 23% of accessions exhibited up to 20% admixed ancestry.

**Polymorphism and molecular diversity potential of genome-wide informative SNPs estimated in 291 chickpea accessions using diverse statistical measures.** $\theta_{\pi}$: Average pair-wise nucleotide diversity. PIC: Polymorphism information content.

| Accessions       | DAD-SNPs | KAD-SNPs | MAD-SNPs | TAD-SNPs |
|------------------|----------|----------|----------|----------|
|                  | Percent polymorphism | PIC | $\theta_{\pi}$ | Genetic distance | Percent polymorphism | PIC | $\theta_{\pi}$ | Genetic distance | Percent polymorphism | PIC | $\theta_{\pi}$ | Genetic distance | Percent polymorphism | PIC | $\theta_{\pi}$ | Genetic distance |
| All 291 accessions | 619 | 0.10–0.41 (0.35) | 1.36 | 0.13–0.64 (0.55) | 531 | 0.10–0.38 (0.32) | 1.32 | 0.15–0.62 (0.46) | 884 | 0.10–0.48 (0.45) | 1.55 | 0.12–0.79 (0.67) | 1083 | 0.11–0.33 (0.29) | 1.25 | 0.10–0.40 (0.37) |
| 189 desi accessions | 526 (85.0) | 0.10–0.40 (0.36) | 1.45 | 0.13–0.69 (0.58) | 449 (84.5) | 0.10–0.38 (0.32) | 1.42 | 0.15–0.63 (0.52) | 793 (89.7) | 0.10–0.48 (0.47) | 1.59 | 0.12–0.78 (0.69) | 758 (70.0) | 0.11–0.32 (0.30) | 1.29 | 0.11–0.41 (0.39) |
| 102 kabuli accessions | 376 (60.7) | 0.10–0.37 (0.32) | 1.33 | 0.14–0.62 (0.49) | 334 (62.9) | 0.10–0.31 (0.20) | 1.27 | 0.15–0.59 (0.44) | 648 (73.3) | 0.10–0.43 (0.41) | 1.48 | 0.12–0.73 (0.58) | 697 (64.3) | 0.11–0.29 (0.27) | 1.23 | 0.10–0.35 (0.31) |

**Figure 3.** Unrooted phylogenetic tree and population structure (with population number K = 2) illustrating the evolutionary relationships and best possible population genetic structure, respectively, among 291 desi and kabuli chickpea accessions using 619 DAD (A) 531 KAD (B) 884 MAD (C) and 1083 TAD (D) SNPs. The mapped markers assigned accessions into two populations (POP I and POP II) which essentially corresponded to their geographical as well as cultivar-specific origination. Red and green colours correspond to POP I and POP II, respectively, as defined by both phylogenetic tree and population genetic structure. The accessions represented by vertical bars along the horizontal axis were classified into K colour segments as per their estimated membership fraction in each K cluster.
The genetic association analysis at a false discovery rate (FDR) cut-off ≤0.05 using the genotyping data of 1967 SNPs identified six genomic loci exhibiting significant association with SYP trait at a P ≤ 10^{-6} (Fig. 5 and Table 3). Interestingly, the DAD-, KAD- and MAD-SNPs (six loci), as compared to TAD-SNPs (three loci), detected a higher number of significant SNP loci and therefore, these informative markers had a greater potential for SYP trait association in chickpea (Fig. 5). Considering a greater efficiency of MAD-SNPs in detecting high intra-specific polymorphism and natural allelic diversity among accessions belonging to an association panel, a strong potential of these marker types for precise trait association mapping is expected at a genome-wide scale. These six SYP trait-associated SNPs were physically mapped on four chromosomes (2, 4, 6 and 7) of desi chickpea genome (Fig. 5 and Table 3). A maximum of three SYP-associated genomic SNP loci were mapped on chromosome 4. The identification of more than one SYP trait-associated genomic loci reflects complex genetic architecture of the studied trait, which was further dissected in our study through high-resolution association mapping involving four various kinds of informative natural SNP allelic variants (DAD-, KAD-, MAD- and TAD-SNPs). The detailed structural annotation of six SYP-associated genomic loci revealed the presence of five SNPs in the coding (two synonymous and three non-synonymous SNPs) and one SNP in the non-coding (intron) sequence components of six genes (Table 3). The SYP trait phenotypic variation explained (PVE) by six gene-derived maximum effect SNP loci varied from 11.2 to 24.2% R² (P: 2.3 \times 10^{-6} to 1.2 \times 10^{-8}) among desi and kabuli accessions belonging to an association panel. The combined SYP trait phenotypic variation explained by all significant six gene-based SNP loci was 27.5%. Notably, one synonymous SNP derived from a gene encoding pentatricopeptide repeat (PPR)-containing protein exhibited strong association (24.2% PVE and 1.2 \times 10^{-8} P) with SYP trait in chickpea (Table 3). The outcome of aforesaid association study was further revalidated by a multi-locus association mapping which identified 6 quantitative trait nucleotides (QTNs) for SYP trait varying from 4.8 to 10.3 LOD (logarithm of odds) and 3.5 to 6.9% R². To ascertain the trait association, SNPs identified in the 100 kb sequences flanking either side of these six SYP-associated QTNs were used for association mapping. None of the SNPs beside trait-associated SNPs was identified to be significantly associated with SYP trait in chickpea.

Discussion

Genomics-assisted crop improvement of chickpea is usually hindered due to its narrow genetic base and low intra-specific polymorphism among cultivated desi and kabuli accessions. To overcome these, large-scale discovery and genotyping of informative sequence-based markers like SNPs differentiating the maximum number of desi and kabuli accessions together and exhibiting high intra-specific potential among cultivated accessions using a user-friendly, rapid and cost-effective genotyping assay is essential at a genome-wide scale. This will essentially accelerate the generation of genotyping information of numerous genome-wide SNP markers differentiating large-scale accessions for high-resolution association mapping and efficient dissection of complex quantitative traits in order to drive genetic enhancement of chickpea. In this perspective, validation and genotyping of numerous non-erroneous and robust SNPs that are discovered by desi and kabuli cultivar-wise individual assembling of low-sequence depth-coverage sequence reads generated from numerous germplasm accessions using GBS assay, can be an efficient alternative genome-wide strategy in chickpea. Adopting this strategy, the current study has developed diverse kinds of genome-wide 1967 SNPs including 884 DAD-, KAD-, MAD-SNPs and 1083 TAD-SNPs to evaluate their potential for discriminating cultivated desi and kabuli accessions and exhibiting association with SYP traits in chickpea. The structurally and functionally annotated informative natural SNP allelic variants discovered by us can be utilized for rapid identification of functionally relevant molecular tags associated with multiple agronomic traits in chickpea. These developed SNPs will further serve as a useful genomic resource for cultivar-wise selection of informative SNPs differentiating desi and kabuli accessions at a genome-wide scale to expedite high-throughput genetic analysis in chickpea.

The intra-specific polymorphic potential especially detected by SNPs among 291 cultivated desi and kabuli chickpea germplasm accessions was much higher as compared to that reported in earlier studies with SNP markers. Nevertheless, the nucleotide diversity measured by DAD-, KAD- and MAD-SNPs among 189 desi and 102 kabuli accessions was almost comparable to that documented using the genome-wide resequencing- and GBS-derived SNP markers. The level of molecular diversity detected especially by MAD-SNPs among 291...
cultivated \emph{desi} and \emph{kabuli} chickpea accessions is much higher as compared to that documented earlier with SNP markers\cite{3,4,5,46,47,48}. Distinct differentiation and clustering of 291 accessions into two major population groups are in accordance with their cultivar-specific origin, pedigree relationships and parentage that well-corresponded with previous studies of Upadhyaya \textit{et al.}\cite{49} and Kujur \textit{et al.}\cite{45}.

The clustering of accessions into a specific population group was strongly influenced by geographical origin as well as adaptive environment rather than their known parentage and pedigree. The observed admixed population genetic structure inferred the occurrence of multiple domestication events along with strong evolutionary bottleneck among \emph{desi} and \emph{kabuli} accessions. The continuous crop genetic improvement breeding efforts with an aim to develop high-yielding stress tolerant \emph{desi} and \emph{kabuli} cultivars may influence their domestication pattern and genetic structure immensely therefore, detection of numerous admixture traces among these cultivated 291 \emph{desi} and \emph{kabuli} chickpea accessions is expected. The determination of high-resolution population genetic structure among accessions is vital to assess the accuracy (non-spurious) of robust SNPs associated with a particular agronomic trait in crop plants. In this context, high-resolution admixed population genetic structure can essentially be effective in detecting non-spurious genomic loci associated significantly with target traits in chickpea.

While trying to evaluate the potential of the developed markers in explaining molecular diversity and population genetic structure in chickpea it was observed that large-scale genome-wide MAD-SNPs discovered by especially comparing the low-sequencing depth-coverage sequence-reads of multiple accessions exhibited a higher potential of intra-specific polymorphism as well as a wider natural allelic diversity than that of two accessions (TAD-SNPs). Therefore, these informative MAD-SNPs with the requisite potential to establish precise phylogenetic relationships and determine admixed population genetic structure among 291 \emph{desi} and \emph{kabuli} chickpea

---

**Figure 5.** GWAS-derived Manhattan plot depicting significant P-values associated with seed yield per plant (SYP) using 619 DAD-, 531 KAD-, 884 MAD- and 1083 TAD-SNPs in chickpea. The genomic distribution of SNPs physically mapped on eight chromosomes of chickpea genome are designated by the x-axis. The y-axis specifies the $-\log_{10}(P)$-value for significant association of SNP loci with SYP trait. The SNPs significantly associated with SYP trait at a cut-off P value $\leq 10^{-6}$ are highlighted with dotted lines.
accessions can be deployed in marker-aided introgression breeding for varietal improvement targeting multiple traits of agronomic importance in chickpea.

The current investigation developed numerous informative DAD-, KAD-, MAD- and TAD-SNPs differentiating 291 precisely field-phenotyped desi and kabuli chickpea accessions at a genome-wide scale to evaluate their potential for marker-trait association and identify potential genomic loci and natural allelic variants governing complex seed yield quantitative trait in chickpea. The high-resolution association study altogether detected six genomic SNP loci including QTNs significantly associated with SYP traits. None of these six major SYP trait-associated QTNs identified has been documented by earlier association and QTL mapping studies in chickpea. This infers the novelty of the detected gene-based natural SNP allelic variants in governing seed yield trait in chickpea. Essentially, the five protein-coding genes with synonymous and non-synonymous SNP loci associated with SYP trait delineated by high-resolution association mapping have functional relevance for quantitative dissection of complex seed yield traits in chickpea. These SYP trait-associated SNPs will be of immense use in establishing marker-trait linkages and identification of natural allelic variants in the genes regulating SYP trait, thereby can accelerate the marker-assisted selection to develop desirable cultivars with high seed yield.

The delineated genes encoding response regulator, arginine decarboxylase, Mediator transcription factor and two-component system associated with SYP trait in chickpea are known to control gene expression dynamics and transcriptional regulatory pathways underlying growth, development and yield traits in multiple crop plants. Essentially, the role of delineated most-promising SYP trait-associated PPR genes with synonymous SNPs in regulating fertility restoration, heterosis breeding and hybrid seed production is well-documented in crop plants including chickpea. This infers the novelty of the detected gene-based natural SNP allelic variants in governing seed yield trait in chickpea. Essentially, the five protein-coding genes with synonymous and non-synonymous SNP loci associated with SYP trait delineated by high-resolution association mapping have functional relevance for quantitative dissection of complex seed yield traits in chickpea. These SYP trait-associated SNPs will be of immense use in establishing marker-trait linkages and identification of natural allelic variants in the genes regulating SYP trait, thereby can accelerate the marker-assisted selection to develop desirable cultivars with high seed yield.

The delineated genes encoding response regulator, arginine decarboxylase, Mediator transcription factor and two-component system associated with SYP trait in chickpea are known to control gene expression dynamics and transcriptional regulatory pathways underlying growth, development and yield traits in multiple crop plants. Essentially, the role of delineated most-promising SYP trait-associated PPR genes with synonymous SNPs in regulating fertility restoration, heterosis breeding and hybrid seed production is well-documented in crop plants including chickpea. This infers the novelty of the detected gene-based natural SNP allelic variants in governing seed yield trait in chickpea.

Methods

Chickpea genetic resources. A total of 291 phenotypically and genotypically diverse accessions (representing diverse eco-geographical locations of the world) belonging to *C. arietinum* desi (189 accessions) and *kabuli* (102) cultivars were selected from available chickpea reference core/mini-core germplasm collections as per Kujur et al. (Table S5). Of these, 92 including 39 desi and 53 kabuli accessions were chosen as a representative set for large-scale mining and genotyping of genome-wide SNPs employing GBS assay in chickpea (Table S6). The genomic DNA was isolated from the young leaves of aforesaid chickpea accessions using a DNasey 96 Plant Kit (QIAGEN, CA, USA) as per the manufacturer’s instructions.

Selection and annotation of genome-wide SNPs. The raw FASTQ sequence reads generated by sequencing of 92 desi and kabuli chickpea accessions using GBS assay were acquired from NCBI SRA database (Accession ID SRX845396) and de-multiplexed into individual sequences by using their unique barcodes in accordance with Kujur et al. The high-quality sequence reads of 92 accessions after filtering with NGS QC Toolkit, were grouped into desi (39 accessions) and kabuli (53) cultivar-wise separately and further mapped individually on desi chickpea genome sequences using BWA. The SNPs were mined by desi and kabuli cultivar-wise comparison of the mapped sequences among accessions using Samtools (http://www.htslib.org) and subsequently high-quality SNPs supported by at least 10 sequence reads (with SNP base quality ≥20 and ≤1% missing data) were screened following Srivastava et al. A set of chromosome-wise uniformly mapped high-quality SNPs exhibiting differentiation only between two desi (ICC 4958) and kabuli (CDC Frontier) cultivars (which draft genomes are sequenced) were screened for further analysis. The structural and functional annotation of SNPs

| SNP IDs | SNPs | SNP physical positions (bp) | Desi chromosomes | Gene accession IDs | Structural annotation | Encoded amino acids | Putative functions | P | PVE (R²%) | LOD | Effect | MAF | r² (%) |
|--------|------|----------------------------|------------------|------------------|----------------------|---------------------|--------------------|---|-----------|-----|--------|-----|--------|
| MAD_SNP155 | [G/A] | 2285980 | Ca_Des Chr02 | Ca_03044 | CDS (SYNONYMOUS) | — | Pentatricopeptide repeat (PPR)-containing protein | 1.2 × 10^-8 | 24.2 | 10.3 | -0.310 | 0.24 | 6.9 |
| MAD_SNP415 | [A/G] | 16444752 | Ca_Des Chr04 | Ca_09334 | CDS (NON-SYNONYMOUS) | TGT (cysteine) - CGT (arginine) | Two-component response regulator (ARR2) | 2.1 × 10^-7 | 13.4 | 6.7 | -0.253 | 0.27 | 4.3 |
| MAD_SNP522 | [G/C] | 27958812 | Ca_Des Chr04 | Ca_10271 | CDS (NON-SYNONYMOUS) | ACT (threonine) - AGT (serine) | Acyl-[acyl-carrier-protein] desaturase, chloroplastic | 2.5 × 10^-7 | 12.4 | 5.5 | -0.217 | 0.20 | 3.9 |
| MAD_SNP522 | [A/T] | 34596600 | Ca_Des Chr04 | Ca_10851 | INTRON | — | Mediator transcription factor | 1.5 × 10^-9 | 20.1 | 8.7 | -0.243 | 0.21 | 5.8 |
| MAD_SNP547 | [C/G] | 52329312 | Ca_Des Chr06 | Ca_18831 | CDS (SYNONYMOUS) | — | Arginine decarboxylase (ADC) | 1.5 × 10^-7 | 14.7 | 7.2 | -0.231 | 0.23 | 5.2 |
| MAD_SNP572 | [G/C] | 1103940 | Ca_Des Chr07 | Ca_19072 | CDS (NON-SYNONYMOUS) | GAC (alanine) - GAG (glutamine) | Unknown expressed protein | 1.4 × 10^-7 | 11.2 | 4.8 | -0.211 | 0.20 | 3.5 |

Table 3. Genomic SNP loci significantly associated with seed yield per plant in chickpea.
in diverse coding (synonymous and non-synonymous substitutions) and non-coding (intron, upstream/downstream regulatory and intergenic regions) sequence components of genes and genomes (chromosomes/pseudomolecules and unanchored scaffolds) were performed as per the desi (CGAP v2.0) genome annotation.

Large-scale genotyping of SNPs. For genotyping of SNPs, desi accessions (39)-derived (DAD) SNPs, kabuli accessions (53)-derived (KAD) SNPs, multiple accessions (92 desi and kabuli)-derived (MAD) SNPs and two accessions (ICC 4958 and CDC Frontier)-derived (TAD) SNPs were genotyped in the genomic DNA of 291 chickpea accessions using Sequenom MALDI-TOF MassARRAY assay (http://www.sequenom.com). The SNP loci along with their flanking 100-bp sequences were curated and a 34-plex MassARRAY multiplex iPLEX array including adapter-ligated pre-amplification and unextended/nucleotide extension PCR primers targeting each SNP were designed by MassARRAY Assay Design software v3.1 following the criteria of Saxena et al.60,61. The multiplex PCR amplification, incubation, primer extension, resin clean-up and mass spectrometry-based SNP genotyping were carried out as per the manufacturer’s instructions of Sequenom iPLEX Gold amplification kit. The SNP genotyping data detected in iPLEX spectrochip bio-arrays were analysed by a MassARRAY Typer 3.4. The genotyping information of SNPs based on allele-specific differences in mass between extension products was visualized and documented among 291 chickpea accessions following Saxena et al.60,61.

For validation of high-quality SNPs, primer-pairs were designed targeting 100-bp either side flanking sequence model of 192 representative SNP loci and PCR amplified using the genomic DNA of 48 selected desi and kabuli chickpea accessions. The amplicons were sequenced and high-quality sequences were aligned and compared among accessions to detect SNPs following Kujur et al.8 and Saxena et al.90.

Assessment of polymorphic potential. The genotyping data of DAD-, KAD-, MAD- and TAD-SNPs were analysed with PowerMarker v3.5162 to estimate the per cent polymorphism and PIC (polymorphism information content) among 291 desi and kabuli chickpea accessions. The SNP genotyping information was further analysed on a 100-kb non-overlapping sliding window approach of TASSEL v5.063 (http://www.maizegenetics.net) to measure one of the major nucleotide diversity parameter θπ among chickpea accessions as per Kujur et al.45 and Bajaj et al.18.

Determination of molecular diversity and population genetic structure. The genotyping data of SNPs were used to assess the molecular diversity and for establishing phylogenetic relationships among 291 desi and kabuli chickpea accessions. The cluster analysis among accessions was carried out based on Nei’s genetic distance44 using the neighbour-joining (NJ) method (with 1000 bootstrap replicates) of PowerMarker, and unrooted phylogenetic tree was constructed employing MEGA v7.065.

To determine the population structure among 291 chickpea accessions, the genotyping data of SNPs were analysed in a model-based program STRUCTURE v2.3.466 as per Kujur et al.45. The optimum population number (K) was determined following the ad hoc66 and delta K87 methods. Using the optimum K, the population structure model representing better relationships among chickpea accessions was determined. Diverse population genetic parameters including genetic divergence (FST) and degree of admixture among different population groups were estimated.

Phenotyping of a constituted association panel. For large-scale phenotyping, an association panel consisting of 291 diverse desi (189) and kabuli (102) germplasm accessions were planted in the experimental field with 35 × 10 cm row-to-row and plant-to-plant spacing, respectively, following standard agronomic practices. These accessions were grown in the field as per ‘alpha-lattice’ design with two replications during crop season for two consecutive years at two diverse geographical locations (New Delhi latitude/longitude: 28.4°N/77.1°E and Hyderabad; 17.1°N/78.9°E) of India. Four selected desi (ICCV 10 and ICCV 93954) and kabuli (ICC 12968 and Annigeri) chickpea accessions sown after every 10 individual plants were served as reference in the field experimental design to ascertain the homogeneity of association panel across two environments. Seed yield per plant (SYP) was estimated by measuring the average weight (g) of fully matured dried seeds (at 10% moisture content) harvested from 10–15 representative plants of each accession. The genetic inheritance pattern of SYP trait based on diverse statistical parameters including coefficient of variation (CV), broad-sense heritability (H2) and frequency distribution were estimated in an association panel as per Bajaj et al.18 and Das et al.23.

Association analysis. For efficient trait association mapping, we employed the environment-wise SYP phenotyping data measured across two diverse geographical locations and experimental years from each of the selected accession for estimating the mean SYP trait value in an individual accession. The SNP genotyping data was correlated with multi- environments replicated field phenotyping data of SYP trait as well as population structure (Q), kinship (K) matrix and principal component analysis (PCA) (P) information of 291 desi and kabuli accessions belonging to an association panel of chickpea. To perform association analysis, the compressed mixed linear model (CMLM) (P + K, K and Q + K) and population parameters previously determined (P3D) interfaces of GIPIT were utilized as per Thudi et al.14, Kujur et al.18 and Kumar et al.85. To determine the accuracy of SNP marker-trait association, we performed the quantile-quantile (Q-Q) plot-based false discovery rate (FDR) cut-off ≤0.05) corrections for multiple comparisons between observed/expected −log10(P)-values and adjusted P-value threshold of significance in accordance with Benjamini and Hochberg68. The SNP loci associated with SYP trait at a lowest FDR adjusted P-value (threshold P < 1 × 10−6) and highest R2 were identified to be highly significant in chickpea. Following FDR-controlling method of model with the SNPs and adjusted P-value, R2 representing the magnitude of SNP marker-SYP trait association was estimated. The trait association outcomes were reevaluated through a fast multi-locus random-SNP-effect EMMA (FASTmrEMMA) model of GWAS as per Wen et al.69. The QTNs with LOD >3.0 and P-value < 0.005 were identified to be significantly associated with studied trait.
References

1. Varshney, R. K. et al. Translational genomics in agriculture: some examples in grain legumes. Crit. Rev. Plant Sci. 34, 169–194 (2015).
2. Jain, M. et al. A draft genome sequence of the pulse crop chickpea (Cicer arietinum L.). Plant J. 74, 715–729 (2013).
3. Varshney, R. K. et al. Draft genome sequence of chickpea (Cicer arietinum) provides a resource for trait improvement. Nat. Biotechnol. 31, 240–246 (2013).
4. Parween, S. et al. An advanced draft genome assembly of a desi type chickpea (Cicer arietinum L.). Sci. Rep. 5, 12806 (2015).
5. Gupta, S. et al. Draft genome sequence of Cicer reticulatum L., the wild progenitor of chickpea provides a resource for agronomic trait improvement. DNA Res. 24, 1–10 (2016).
6. Upadhyaya, H. D. et al. Genomic tools and germplasm diversity for chickpea improvement. Plant Genet. Resour. 9, 45–58 (2011).
7. Upadhyaya, H. D. et al. A genome-scale integrated approach aids in genetic dissection of complex flowering time trait in chickpea. Plant Mol. Biol. 89, 403–420 (2015).
8. Kujar, A. et al. Functionally relevant microsatellite markers from chickpea transcription factor genes for efficient genotyping applications and trait association mapping. DNA Res. 20, 355–374 (2013).
9. Kujar, A. et al. 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 (2014).
10. Kujar, A. et al. A genome-wide SNP scan accelerates trait-regulatory genomic loci identification in chickpea. Sci. Rep. 5, 11166 (2015).
11. Kujar, A. et al. 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 (2015).
12. Kujar, A. et al. Identification of candidate genes and natural allelic variants for QTLs governing plant height in chickpea. Sci. Rep. 6, 27968 (2016).
13. Deokar, A. A. et al. 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 (2014).
14. Thudi, M. et al. Genetic dissection of drought and heat tolerance in chickpea through genome-wide and candidate gene-based association mapping approaches. PloS One 9, e96758 (2014).
15. Varshney, R. K. et al. Genetic dissection of drought tolerance in chickpea (Cicer arietinum L.). Theor. Appl. Genet. 127, 445–462 (2014).
16. Bajaj, D. et al. 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 (2015).
17. Bajaj, D. et al. Genome-wide high-throughput SNP discovery and genotyping for understanding natural (functional) allelic diversity and domestication patterns in wild chickpea. Sci. Rep. 5, 11627 (2015).
18. Bajaj, D. et al. 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 (2015).
19. Bajaj, D. et al. Identification of candidate genes for dissecting complex branch number trait in chickpea. Plant Sci. 245, 61–70 (2016).
20. Bajaj, D. et al. EcoTILLING-based association mapping efficiently delineates functionally relevant natural allelic variants of candidate genes governing agronomic traits in chickpea. Front. Plant Sci. 7, 450 (2016).
21. Das, S. et al. Genome-wide insertion-deletion (InDel) marker discovery and genotyping for genomics-assisted breeding applications in chickpea. DNA Res. 22, 377–386 (2015).
22. Das, S. et al. Deploying QTL-seq for rapid delineation of a potential candidate gene underlying major trait-associated QTL in chickpea. DNA Res. 22, 193–203 (2015).
23. Das, S. et al. mQTL-seq delineates functionally relevant candidate gene harbouring a major QTL regulating pod number in chickpea. DNA Res. 23, 53–65 (2016).
24. Gaur, P. M., Samineni, S., Tripathi, S., Varshney, R. K. & Gowda, C. L. L. Allelic relationships of flowering time genes in chickpea. Euphytica 203, 295–308 (2014).
25. Jaganathan, D. et al. Genotyping-by-sequencing based intra-specific genetic map refines a “QTL-hotspot” region for drought tolerance in chickpea. Mol. Genet. Genomics 290, 559–571 (2015).
26. Kale, S. M. et al. Prioritization of candidate genes in “QTL-hotspot” region for drought tolerance in chickpea (Cicer arietinum L.). Sci. Rep. 5, 15296 (2015).
27. Singh, V. K. et al. 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 (2016).
28. Srivastava, R., Singh, M., Bajaj, D. & Parida, S. A high-resolution InDel (Insertion-Deletion) markers-anchored consensus genetic map identifies major QTLs governing pod number and seed yield in chickpea. Front. Plant Sci. 7, 1362 (2016).
29. Srivastava, R. et al. Genome-wide development and deployment of informative intron-spanning and intron-length polymorphism markers for genetics-assisted breeding applications in chickpea. Plant Sci. 252, 374–387 (2016).
30. Srivastava, R., Bajaj, D., Malik, A., Singh, M. & Parida, S. K. Transcriptome landscape of perennial wild Cicer microphyllum uncovers functionally relevant molecular tags regulating agronomic traits in chickpea. Sci. Rep. 6, 33616 (2016).
31. Abbo, S., Berger, J. & Turner, N. C. Evolution of cultivated chickpea: four bottlenecks limit diversity and constrain adaptation. Funct. Plant Biol. 30, 1081–1087 (2003).
32. Abbo, G., Gopher, A., Rubin, B. & Lev-Yadun, S. On the origin of near Eastern founder crops and the ‘dump-heap’ hypothesis. Genet. Res. Crop Ecol. 52, 491–495 (2005).
33. Berger, J. D., Abbo, S. & Turner, N. C. Ecogeography of annual wild Cicer species: the poor state of the world collection. Crop Sci. 43, 1076–1090 (2003).
34. Berger, J. D., Buck, R., Henzell, J. M. & Turner, N. C. Evolution in the genus Cicer vernalisation response and low temperature pod set in chickpea (C. arietinum L.) and its annual wild relatives. Aust. J. Agric. Res. 56, 1191–1208 (2005).
35. Toker, J. C. A note on the evolution of kabuli chickpeas as shown by induced mutations in Cicer reticulatum Ladzinskys. Genet. Resour. Crop Evol. 56, 7–12 (2009).
36. Garg, R., Patel, R. K., Tyagi, A. K. & Jain, M. De novo assembly of chickpea transcriptome using short reads for gene discovery and marker identification. DNA Res. 18, 53–63 (2011).
37. Garg, R., Bhattacharjee, A. & Jain, M. Genome-scale transcriptomic insights into molecular aspects of abiotic stress responses in chickpea. Plant Mol. Biol. Rep. 33, 388–400 (2015).
38. Garg, R. et al. Transcriptome analyses reveal geno-type- and developmental stage-specific molecular responses to drought and salinity stresses in chickpea. Sci. Rep. 6, 19228 (2016).
39. Agarwal, G. et al. Comparative analysis of kabuli chickpea transcriptome with desi and wild chickpea provides a rich resource for development of functional markers. PloS One 7, e52443 (2012).
40. Hiremath, P. et al. 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 (2012).
41. Jhanwar, S. et al. Transcriptome sequencing of wild chickpea as a rich resource for marker development. Plant Biotechnol. J. 10, 690–702 (2012).
42. Gaur, R. et al. High-throughput SNP discovery and genotyping for constructing a saturated linkage map of chickpea (Cicer arietinum L.). DNA Res. 19, 357–373 (2012).
43. Gaur, R. et al. High density linkage mapping of genomic and transcriptomic SNPs for synteny analysis and anchoring the genome sequence of chickpea. Sci. Rep. 5, 13387 (2015).
Acknowledgements
The authors gratefully acknowledge the financial support by the core grant of National Institute of Plant Genome Research (NIPGR), New Delhi, India.

Author Contributions
U.B. and R.S. conducted all experiments and drafted the manuscript. U.B., R.S., D.B., V.T., A.D. and N.M. involved in genotyping-by-sequencing and data analysis. H.D.U. assisted in constitution of an association panel and its field phenotyping. S.K.P. conceived and designed the study, guided data analysis and interpretation, participated in genotyping-by-sequencing and data analysis. H.D.U. assisted in constitution of an association panel and its field phenotyping. S.K.P. conceived and designed the study, guided data analysis and interpretation, participated in genotyping-by-sequencing and data analysis.

Additional Information
Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-018-29926-1.

Competing Interests: The authors declare no competing interests.

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s) 2018