



# MutMap Approach Enables Rapid Identification of Candidate Genes and Development of Markers Associated With Early Flowering and Enhanced Seed Size in Chickpea (*Cicer arietinum* L.)

## OPEN ACCESS

### Edited by:

Patricia L. Polowick,  
National Research Council Canada  
(NRC-CNRC), Canada

### Reviewed by:

Bianyun Yu,  
Aquatic and Crop Resource  
Development Research Center,  
National Research Council Canada  
(NRC-CNRC), Canada

Harsh Raman,  
New South Wales Department  
of Primary Industries, Australia

### \*Correspondence:

Rajeev K. Varshney  
r.k.varshney@cgiar.org;  
rajeev.varshney@murdoch.edu.au  
Mahendar Thudi  
t.mahendar@cgiar.org

†These authors have contributed  
equally to this work

### Specialty section:

This article was submitted to  
Plant Breeding,  
a section of the journal  
Frontiers in Plant Science

**Received:** 31 March 2021

**Accepted:** 02 June 2021

**Published:** 12 July 2021

### Citation:

Manchikatla PK, Kalavikatte D,  
Mallikarjuna BP, Palakurthi R,  
Khan AW, Jha UC, Bajaj P, Singam P,  
Chitkineni A, Varshney RK and  
Thudi M (2021) MutMap Approach  
Enables Rapid Identification  
of Candidate Genes  
and Development of Markers  
Associated With Early Flowering  
and Enhanced Seed Size in Chickpea  
(*Cicer arietinum* L.).  
Front. Plant Sci. 12:688694.  
doi: 10.3389/fpls.2021.688694

Praveen Kumar Manchikatla<sup>1,2†</sup>, Danamma Kalavikatte<sup>1†</sup>, Bingi Pujari Mallikarjuna<sup>3</sup>,  
Ramesh Palakurthi<sup>1</sup>, Aamir W. Khan<sup>1</sup>, Uday Chand Jha<sup>4</sup>, Prasad Bajaj<sup>1</sup>,  
Prashant Singam<sup>2</sup>, Annapurna Chitkineni<sup>1</sup>, Rajeev K. Varshney<sup>1,5\*</sup> and  
Mahendar Thudi<sup>1,6\*</sup>

<sup>1</sup> Center of Excellence in Genomics and Systems Biology (CEGSB), International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Hyderabad, India, <sup>2</sup> Department of Genetics, Osmania University, Hyderabad, India, <sup>3</sup> Indian Council of Agricultural Research (ICAR)-Indian Agricultural Research Institute, Regional Research Centre, Dharwad, India, <sup>4</sup> Indian Council of Agricultural Research (ICAR)-Indian Institute of Pulses Research (IIPR), Kanpur, India, <sup>5</sup> State Agricultural Biotechnology Centre, Centre for Crop and Food Innovation, Food Futures Institute, Murdoch University, Murdoch, WA, Australia, <sup>6</sup> Centre for Crop Health, University of Southern Queensland, Toowoomba, QLD, Australia

Globally terminal drought is one of the major constraints to chickpea (*Cicer arietinum* L.) production. Early flowering genotypes escape terminal drought, and the increase in seed size compensates for yield losses arising from terminal drought. A MutMap population for early flowering and large seed size was developed by crossing the mutant line ICC4958-M3-2828 with wild-type ICC 4958. Based on the phenotyping of MutMap population, extreme bulks for days to flowering and 100-seed weight were sequenced using Hi-Seq2500 at 10X coverage. On aligning 47.41 million filtered reads to the CDC Frontier reference genome, 31.41 million reads were mapped and 332,395 single nucleotide polymorphisms (SNPs) were called. A reference genome assembly for ICC 4958 was developed replacing these SNPs in particular positions of the CDC Frontier genome. SNPs specific for each mutant bulk ranged from 3,993 to 5,771. We report a single unique genomic region on Ca6 (between 9.76 and 12.96 Mb) harboring 31, 22, 17, and 32 SNPs with a peak of SNP index = 1 for low bulk for flowering time, high bulk for flowering time, high bulk for 100-seed weight, and low bulk for 100-seed weight, respectively. Among these, 22 SNPs are present in 20 candidate genes and had a moderate allelic impact on the genes. Two markers, Ca6EF10509893 for early flowering and Ca6HSDW10099486 for 100-seed weight, were developed and validated using the candidate SNPs. Thus, the associated genes, candidate SNPs, and markers developed in this study are useful for breeding chickpea varieties that mitigate yield losses under drought stress.

**Keywords:** MutMap, early flowering, chickpea, 100 seed weight, candidate genes and SNPs

## INTRODUCTION

Chickpea (*Cicer arietinum* L.) is the second most important annual grain legume crop predominantly cultivated on residual soil moisture in the arid and semi-arid areas of the world. Global annual cultivation of chickpea is over 14.56 million ha with a total production of 14.77 million tons (FAOSTAT, 2017, accessed on January 26, 2020). Chickpea seeds are rich in protein (17–20%), minerals (phosphorus, calcium, magnesium, iron, and zinc) (Jukanti et al., 2012; Sab et al., 2020), and carotenoids; chickpea also improves soil health by adding atmospheric nitrogen (20–40 kg N ha<sup>-1</sup>) through symbiosis (Joshi et al., 2001). Climate changes during the recent past have been posing serious threats to chickpea production and causing about 19% yield losses (Kadiyala et al., 2016).

In India, chickpea is grown in a wide range of agro-climatic niches. Based on crop duration, these regions can be classified as short-duration (Southern/peninsular India), medium-duration (Central India), and long-duration (Northern India) environments. In general, chickpea matures in a wide time frame of 80–180 days. However, in 66% of chickpea-growing areas, the available crop-growing season is about 80–120 days as they are exposed to abiotic stresses such as drought and heat toward the grain-filling stage. A major shift in the chickpea area (about 3 million ha) from Northern India (cooler, long-season environment) to Southern India (warmer, short-season environment) has been observed during the past four decades. As a result, no major boost in the total production of chickpea has been substantiated.

Terminal drought is considered as one of the most important constraints to chickpea production, and almost 40–50% yield losses were observed globally (see Roorkiwal et al., 2020). The number of days to flowering is an important trait for crop adaptation and productivity, especially in arid and semi-arid regions that experience terminal drought conditions. Early phenology, an adaptation-related trait, helps in the adaptation of chickpea to short-season environments as early flowering genotypes escape terminal (end of season) stresses (drought, high/low temperature) (see Berger et al., 2006). Therefore, the ability to manipulate flowering time is an essential component of chickpea improvement. Seed size/weight is an important yield-contributing trait, and therefore, in past, major breeding emphasis was on improving this trait (Gaur et al., 2014). As a result, early flowering desi and kabuli genotypes were identified through germplasm characterization (Upadhyaya et al., 2007); the low-resolution quantitative trait loci (QTLs) have been reported for flowering time and seed size (Varshney et al., 2014c; Upadhyaya et al., 2015; Verma et al., 2015; Mallikarjuna et al., 2017). Further, efforts were also made to understand the genes and pathways involved in flower development in chickpea (Singh et al., 2013), including through a gene expression atlas (Kudapa et al., 2018). Although the QTLs mapped within large genomic intervals limit the identification of potential candidate genes and their use in marker-assisted selection, in recent years, using a marker-assisted backcrossing approach several high-yielding and drought-tolerant lines in different genetic

backgrounds of chickpea have been released for cultivation (Varshney et al., 2013a; Bharadwaj et al., 2021). Molecular breeding lines with enhanced resistance to biotic stresses were also developed (Varshney et al., 2014b; Pratap et al., 2017; Mannur et al., 2019).

The majority of the QTL mapping and gene isolation approaches using traditional approaches are time-consuming and low-throughput methods. Nevertheless, for more than a decade, the next-generation sequencing (NGS) technologies facilitated understanding of the genetics of complex traits at a faster pace in cereals and legumes (Thudi et al., 2020; Jaganathan et al., 2020). In the case of chickpea, apart from sequencing the genome (Varshney et al., 2013b) and several hundred germplasm lines (Thudi et al., 2016a,b; Varshney et al., 2019b), traits were fine-mapped (Kale et al., 2015; Singh et al., 2016). For decades, forward genetic approaches that rely on molecular characterization of altered phenotypes have been one of the driving forces for crop improvement. In the case of crops with a narrow genetic base, such as chickpea, the creation of allelic variation through mutations and the identification of causal variants will be a potential alternative that can overcome the existing production barriers. MutMap is one of the novel gene mapping approaches that allows rapid identification of causal nucleotide changes of mutants by whole-genome resequencing of pooled DNA of mutant F<sub>2</sub> progeny derived from crosses made between candidate mutants and the parental line (Abe et al., 2012; Fekih et al., 2013). This new NGS-based technique has been successfully applied in crop plants for rapid identification of the candidate gene as well as the QTL responsible for agronomically important traits (Abe et al., 2012; Megersa et al., 2015; Takagi et al., 2015; Fang et al., 2016; Klein et al., 2018; Tran et al., 2020).

Here, we report for the first time in chickpea, deployment of the MutMap approach that enabled us to rapidly identify genes and single nucleotide polymorphisms (SNPs) associated with early flowering and seed size. In addition, we also report the development and validation of markers that can be used for selection in chickpea breeding programs for improving these traits.

## MATERIALS AND METHODS

### Development and Phenotyping of MutMap Population

To identify phenotypically distinct mutant lines for early flowering and larger seed size, a set of 100 mutant lines from a TILLING (target-induced local lesions in the genome) population developed through ethyl methanesulfonate (EMS) mutagenesis of desi chickpea genotype ICC 4958 (unpublished ICRISAT) was phenotyped for these traits at ICRISAT (17.5111° N, 78.2752° E).

ICC 4958 is a drought-tolerant accession available from the ICRISAT germplasm collection. It was collected from Jabalpur, Madhya Pradesh, India, in 1973, and it was among the over 1,500 germplasm accessions screened for drought resistance at ICRISAT Center between 1978 and 1983. It is being used as a donor parent for introgressing drought tolerance-related traits

and that produces high yields in low productivity, short-duration, terminal drought-prone environments, e.g., those in peninsular India (Varshney et al., 2013a; Bharadwaj et al., 2021).

A set of 45 simple sequence repeat (SSR) markers distributed equally across the genomes was used to identify the genetic similarity among the selected lines. SSR genotyping was performed as described earlier (Thudi et al., 2011). PCR products were denatured and size-fractionated using capillary electrophoresis on an ABI 3730 DNA Genetic Analyzer (Applied Biosystems, United States). Based on allelic data, the mutant line with > 95% similarity to ICC 4958 was selected as the female parent. A MutMap population was developed crossing ICC 4958-M3-2828 (with large seed size and early flowering) and ICC 4958. F<sub>1</sub>s were selfed to produce F<sub>2</sub> seeds. These F<sub>2</sub> seeds were sown in the field during crop season 2017–2018 at ICRISAT. The F<sub>2</sub> population was scored for days to flowering (DF) and 100-seed weight (SDW).

### Isolation of DNA and Sequencing of ICC 4958 (Wild Type) and Trait Bulks

Genomic DNA was extracted from the leaves of F<sub>2</sub> individuals using the NucleoSpin Plant II kit (Macherey-Nagel, Dren, Germany). An equimolar concentration of DNA from 15 F<sub>2</sub> plants with high phenotypic values was pooled together as high bulk, and similarly, DNA from low phenotypic values was pooled together as low bulk. Thus, four extreme bulks, two for each trait, were prepared for WGRS along with wild-type parent ICC 4958 separately. About 5 µg of the pooled DNA was used for the preparation of a sequencing library of average insert size 200–500 bp, according to the protocol for the Paired-End DNA Sample PrepKit (Illumina, United States). The library was sequenced to 10X of genome coverage with the Illumina HiSeq 2500 platform (Illumina, United States).

### Alignment of Short Reads to Reference Sequences and SNP Calling

Initially, a reference-based sequence of the ICC 4958 wild type was generated by aligning the sequence data generated to the CDC Frontier reference genome (Varshney et al., 2013b) as described in the study by Abe et al. (2012). In brief, 59 million paired-end short reads from ICC 4958 wild type and four mutant pools were used for the analysis. The quality checks for these reads were performed using FastQC v0.11.8 (Andrews, 2010), and Trimmomatic v0.39 (Bolger et al., 2014) was used to filter poor-quality reads and remove potential adapter contamination. For this, Illumina adapters and primers sequences were used by Trimmomatic for trimming, followed by iterative removal throughout the read length with mean base Phred qualities > 30 in 5-bp sliding windows. Remaining sequences with lengths < 35 bp after trimming were discarded as well as orphan single-end reads. These high-quality short reads were pooled and aligned with MAQ to the CDC Frontier reference sequence. Alignment files were converted to SAM or BAM files using SAMtools (Li et al., 2009) and applied to a filter pipeline (Kosugi et al., 2013) for the identification of reliable SNPs. This filter pipeline was developed to maximize

true SNP detection and minimize false SNP calling by (i) the removal of paired-end reads of insert size > 325 bp, (ii) calling SNPs only for genomic regions covered by a minimum of three reads for homozygous SNPs and five reads for heterozygous SNPs and a maximum of threefold of average read depth over the genome, and (iii) calling SNPs only on sites with an averaged Illumina Phred-like quality score ≥ 20. Using this pipeline, we identified 332,395 reliable SNPs between ICC 4958 reads and the CDC Frontier reference sequence. On the basis of this result, we generated an ICC 4958 reference sequence by replacing CDC Frontier nucleotides with those of ICC 4958 at 332,395 sites. To remove the effect of SNPs irrelevant to the mutant screen, we generated and used a reference sequence of the same wild-type ICC 4958 that was used for mutagenesis. We further refined this reference sequence by taking a consensus of cumulative genome sequences of the mutants.

Paired-end sequence reads of bulked DNA of mutant F<sub>2</sub> progeny were aligned to the ICC 4958 reference sequence, and SNPs were scored as homozygous SNPs (with SNP index ≥ 0.9) and heterozygous SNPs (with SNP index ≥ 0.3 and < 0.9). We further excluded common SNPs shared by at least two mutant lines as well as G→A or C→T transitions (as they are most frequent in EMS mutagenesis). After identifying the genomic regions harboring a cluster of SNPs with an SNP index of 1, we relaxed the condition of the filter to consider all SNPs (caused by all the transition and transversion) in the region as candidate SNPs for the causal mutation. SNP index plot regression lines were obtained by averaging SNP indices from a moving window of five consecutive SNPs and shifting the window one SNP at a time. The x-axis value of each averaged SNP index was set at a midpoint between the first and the fifth SNP.

### Primer Designing and Validation

The candidate SNPs with a SNP index = 1 were targeted for designing allele-specific markers. WASP, a web-based tool, was used for designing allele-specific primers (Wangkumhang et al., 2007)<sup>1</sup>. A total of 82 desi chickpea genotypes (47 for early flowering and 48 for seed size) were selected for marker validation. PCR was carried out in a 5 µl volume containing 10 ng of DNA, 1X buffer, 200 µM dNTP, 2.5 mM MgCl<sub>2</sub>, 1–5 picomole forward and reverse primers, and 0.1 U of *Taq* polymerase. PCR was performed using Perkin Elmer 384-well Thermal cyclers (Applied Biosystems, United States) and involved a touchdown PCR. Touch down PCR cycles involved initial denaturation at 94°C for 5 min followed by 10 cycles of denaturation at 94°C for 20 s, 60°C for 30 s, 72°C for 30 s that decreases 1°C per cycle; then 35 cycles of 94°C for 20 s, at an optimized annealing temperature of each primer pair (51–58°C) for 30 s, 72°C for 30 s; final extension of 72°C for 20 min, and hold at 4–10°C forever. The PCR products were checked on 1.2% agarose gel containing 0.5 µl/10 ml ethidium bromide (10 mg/ml) with a 100-bp DNA ladder by running it at a constant voltage of 90 V for 25 min. The amplification was visualized under UV illumination using the Uvi-Tech gel documentation system (DOL-008.XD, England).

<sup>1</sup><http://bioinfo.biotech.or.th/WASP>

## RESULTS

### MutMap Population for Early Flowering and Large Seed Size

Based on phenotyping of 100 mutant lines from the TILLING population, 25 lines that were phenotypically distinct from the wild-type ICC 4958 for flowering time and seed size were identified. In order to identify a mutant line that is > 95% similar to ICC 4958 wild type at the genome level, 25 lines along with wild type were genotyped using 25 SSR markers that are equally distributed across the genome (**Supplementary Table 1**). Based on SSR marker data, a dendrogram (**Supplementary Figure 1**) was constructed using DARWin5 (Perrier et al., 2003). The mutant ICC4958-M3-2828 with > 95% similarity to the ICC 4958 wild type and phenotypically distinct for flowering time and seed size was selected for developing a MutMap population. A total of 28 F<sub>1</sub> seeds were harvested by crossing ICC4958-M3-2828 and ICC 4958 wild type from July to September 2017 in the greenhouse at ICRISAT. During the crop season 2017–2018, F<sub>1</sub>s were advanced to F<sub>2</sub> and a total of 204 F<sub>2</sub> seeds were harvested.

### Phenotypic Diversity in MutMap Population and Preparation of Trait Bulks

A total of 204 F<sub>2</sub> plants were phenotyped for early flowering and 100-seed weight during the crop season 2018–2019 in the field. The MutMap population had high phenotypic variability for both flowering time and seed size (**Supplementary Table 2** and **Supplementary Figures 2A,B**). A negative correlation ( $R = -0.13$ ) was observed among these traits (**Supplementary Figure 2C**). DNA from 15 F<sub>2</sub> progeny that displayed early flowering (27–34 days, as EF pool) and late flowering (59–60 days, as LF pool) was combined. Similarly, we also combined DNA from 15 F<sub>2</sub> progeny that displayed high 100-seed weight (43.0–46.2 g, as HSDW pool) and low 100-seed weight (27.0–41.0 gm, as LSDW pool) and subjected them to whole-genome sequencing using the Illumina HiSeq2500 platform (**Supplementary Table 3**).

### Whole-Genome Sequencing and Alignment of Short Reads

Both wild-type ICC 4958 and four trait bulks were sequenced to ~10X of genome coverage. For each trait detail on the number of raw reads, filtered reads, mapped reads, average coverage, and average quality are presented in **Table 1**. In the case of flowering time-related pools, we obtained 44,907,030 and 37,699,572 cleaned bases for EF pool and LF pool, respectively.

Similarly, for seed size-related pools, 43,818,340 and 50,125,538 cleaned bases were obtained for the HSDW pool and LSDW pool, respectively (**Table 1**). On aligning 47.41 million filtered reads to the CDC Frontier reference genome, 31.41 million reads were mapped and 332,395 SNPs were called. These SNPs were used to develop a consensus reference genome sequence for ICC 4958 by replacing them in the particular positions of CDC Frontier. The filtered bulk sequenced paired-end reads were aligned and SNPs were called against this reference assembly that yielded alignment results as follows: 27.12 M reads for the LF pool, 31.57 M reads for the EF pool, 35.30 M reads for the HSDW pool, and 30.81 M reads for the LSDW pool (**Table 1**).

### Trait-Associated Genes and SNPs

The Illumina short reads obtained for the four bulks were separately aligned to the reference sequence of ICC 4958 and then compared to the SNPs of each mutant bulk against wild-type ICC 4958 to identify the SNPs specific for each mutant bulk as well as their distribution on each pseudomolecule (**Table 2**). The SNP index for each SNP was also calculated. The number of SNPs among the bulks ranged from 3,993 (EF pool) to 5,771 (HSDW pool). In the case of EP pool flowering bulk, MutMap revealed 3,993 and 5,081 SNPs, of which 872 and 25 were candidate sites for the EF and LF pools, respectively, with a SNP index  $\geq 0.8$  (**Table 2**). While in the case of seed size bulks, MutMap revealed 4,777–5,771 SNPs, of which 771 and 1,078 were candidate sites for the LSDW and HSDW pools, respectively, with a SNP index  $\geq 0.8$  (**Table 2**). These SNPs were presumably the candidate mutations. However, it was not possible to pinpoint causal mutations from so many candidates. Further, for each bulk, SNP index vs. SNP genomic position graphs for the eight chickpea pseudomolecules were generated as shown in **Supplementary Figures 3–6**. The SNP index plots were very similar between the mutant and wild-type bulks across the entire genome. Nevertheless, a single genomic region on Ca6, between 9.76 and 12.96 Mb with a peak of SNP index 1, was identified in all four mutant bulks overlapping in this region that is missing from the wild-type bulk (**Table 3**). As expected, the SNP index was close to 0 across the genome, but within the unique genomic region identified on Ca6, between 9.76 and 12.96 Mb, its value was greater than zero. This was the only region that exhibited a SNP index difference of > 0 that is significant between the mutant and wild-type bulks. After identifying the region specific to mutant bulk, with a SNP index = 1, the SNPs therein (Ca6, between 9.76 and 12.96 Mb) were scrutinized in detail. Accordingly, we found a total of 38, 22,17,32 SNPs with a SNP index = 1 in the case of EF (**Figure 1**), LF, HSDW (**Figure 2**), and LSDW pools,

**TABLE 1** | Summary of data generated and aligned on wild and mutant pools.

Wild/mutant pools	Raw reads	Filtered reads	Number of reads mapped	Average coverage (X)	Average mapping quality (%)
ICC 4958 (wild)	63,541,248	47,414,372	31,418,911	9.28	42.51
Days to flowering (Low; LDF)	58,196,006	44,907,030	31,571,460	9.49	42.38
Days to flowering (High; HDF)	51,747,822	37,699,572	27,123,905	7.97	42.53
100-seed weight (Low; LSDW)	58,041,216	43,818,340	30,815,440	9.19	42.37
100-seed weight (High; HSDW)	65,129,934	50,125,538	35,307,049	10.59	42.55

**TABLE 2** | Identification and distribution of associated SNPs in the genome.

Trait (bulk)	Number of SNPs	SNPs with a SNP index > 0.8	Pseudomolecules							
			Ca1	Ca2	Ca3	Ca4	Ca5	Ca6	Ca7	Ca8
Days to flowering (Low)	5,081	872	864	503	486	856	626	860	735	151
Days to flowering (High)	3,993	25	667	385	395	648	508	705	556	129
100-seed weight (High)	5,771	1078	812	491	498	775	596	817	643	145
100-seed weight (Low)	4,777	771	1,068	566	552	960	653	999	794	179

respectively in the region for the mutant bulk (Table 3). Of the 102 candidate SNPs, 41 were unique candidate SNPs and 33 were found in more than one bulk. On annotation of the 74 SNPs (41 unique and 33 in more than one bulk), 48, 16, 7, and 3 were intergenic, intronic, synonymous, and missense SNPs, respectively (Supplementary Table 4). Among 22 SNPs with a SNP index = 1, 44 were CT, and eight were GA transitions in the case of LF pool (Supplementary Table 5). We identified 31 SNPs with a SNP index = 1 in the case of the EF pool, of which 17 were CT and 14 GA transitions (Supplementary Table 6). Among 31 SNPs with a SNP index 1, a SNP (Ca6\_10099486) present in the gene Ca\_08581 that encodes putative importin beta-3 (AtKPNB1), in a previous study upregulation of AtKPNB1 led to early flowering in Arabidopsis (Luo et al., 2013). Similarly, 17 and 32 mutations were identified with a SNP index = 1, in HSDW and LSDW pools, respectively (Supplementary Tables 7, 8). Further, among 22 SNPs, two SNPs are in the gene Ca\_08530, which encodes aspartokinase homoserine dehydrogenase involved in the homoserine biosynthetic process, in the phosphorylation process, and in the oxidation–reduction process (Table 4).

### Markers for Early Flowering and Large Seed Size

Of 102 candidate SNPs on Ca6, between 9.76 and 12.96 Mb, 74 candidate SNPs were targeted to design primer pairs using WASP (see text footnote 1). A total of 12 allele-specific primer pairs were designed, and the primer sequence information is provided in Supplementary Table 9. Twelve primer pairs were initially amplified on a set of eight chickpea genotypes. Of twelve primer pairs, seven primer pairs had amplification on all eight genotypes tested. However, allele-specific amplification was obtained for Ca6HSDW10099486, Ca6HSDW9890335, Ca6HSDW9828083, and Ca6EF10509893. Hence, these 4 markers were validated on 82 select chickpea germplasm lines (47 for early flowering and 48 for seed size). As a result, one marker each for EF (Ca6EF10509893; Figure 3A) and HSDW (Ca6HSDW10099486;

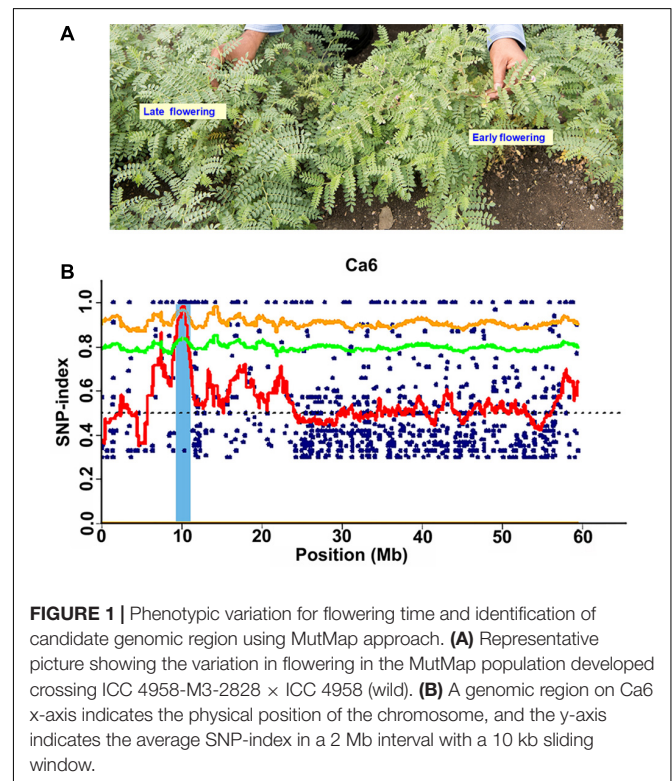
Figure 3B), with allele-specific amplification and high accuracy in the tested germplasm lines, has the potential to be used for improving the early flowering and seed size in chickpea. A clear significant difference ( $p < 0.05$ ) between the amplified and non-amplified genotypes based on their phenotypic values for Ca6EF10509893 and Ca6HSDW10099486 can be visualized in Figures 3C,D, respectively. Nevertheless, these markers need to be tested on large germplasm sets for their efficiency before being used in early-generation selection in chickpea breeding programs.

### DISCUSSION

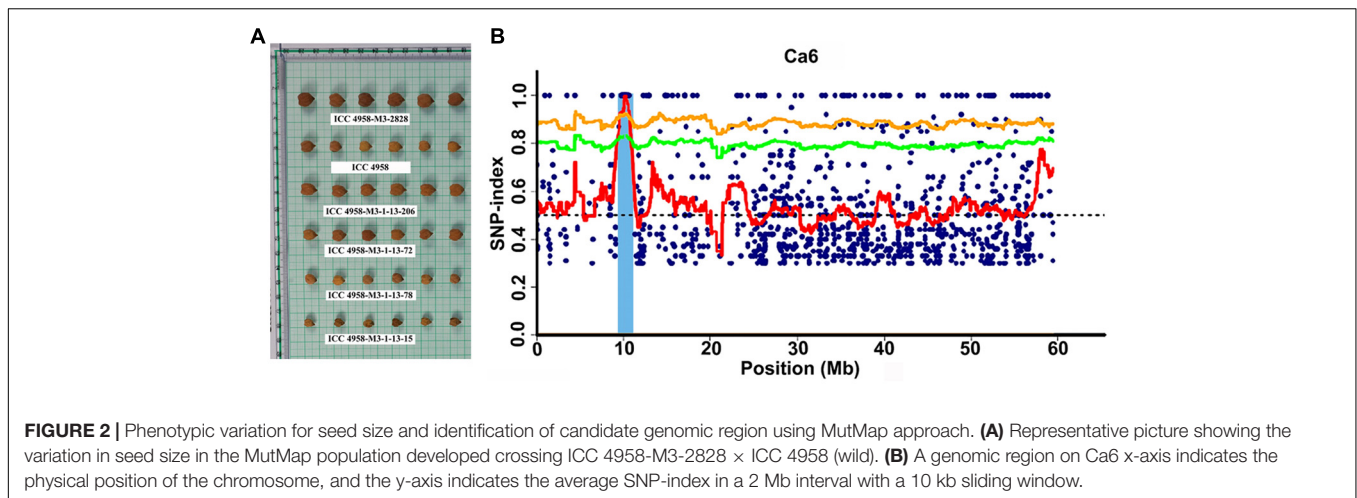
Early flowering and seed size are the two important traits in chickpea as short-duration cultivars can escape terminal drought and high/low-temperature stresses, and enhanced seed size increases the yield to compensate for yield loss due to drought stress. Although early flowering accessions of desi and

**TABLE 3** | Summary of SNPs with SNP index 1 on chromosome Ca6.

Trait	Region on Ca6 (Mb)	Total region (Mb)	SNPs with a SNP index = 1
Days to flowering (Low)	9.77–12.96	3.19	31
Days to flowering (High)	9.77–11.09	1.32	22
100-seed weight (High)	9.76–10.68	0.92	17
100-seed weight (Low)	9.82–10.68	0.86	32



**FIGURE 1** | Phenotypic variation for flowering time and identification of candidate genomic region using MutMap approach. (A) Representative picture showing the variation in flowering in the MutMap population developed crossing ICC 4958-M3-2828 × ICC 4958 (wild). (B) A genomic region on Ca6 x-axis indicates the physical position of the chromosome, and the y-axis indicates the average SNP-index in a 2 Mb interval with a 10 kb sliding window.



**FIGURE 2 |** Phenotypic variation for seed size and identification of candidate genomic region using MutMap approach. **(A)** Representative picture showing the variation in seed size in the MutMap population developed crossing ICC 4958-M3-2828 × ICC 4958 (wild). **(B)** A genomic region on Ca6 x-axis indicates the physical position of the chromosome, and the y-axis indicates the average SNP-index in a 2 Mb interval with a 10 kb sliding window.

kabuli types have been identified from germplasm collections (Upadhyaya et al., 2007) and super-early lines (ICCV 96029 and ICCV 96030) have been developed, there is a need for the identification of candidate genes and causal SNPs to accelerate the development of climate-resilient chickpea varieties. In the past, early flowering genes and their allelic relationships were reported based on the understanding of trait genetics (Gaur et al., 2014). Major QTLs for 100-seed weight were reported and were also fine-mapped (Varshney et al., 2014c; Jaganathan et al., 2015; Kale et al., 2015). However, none of the studies focused on the identification of candidate genes and causal SNPs responsible for flowering time and seed size.

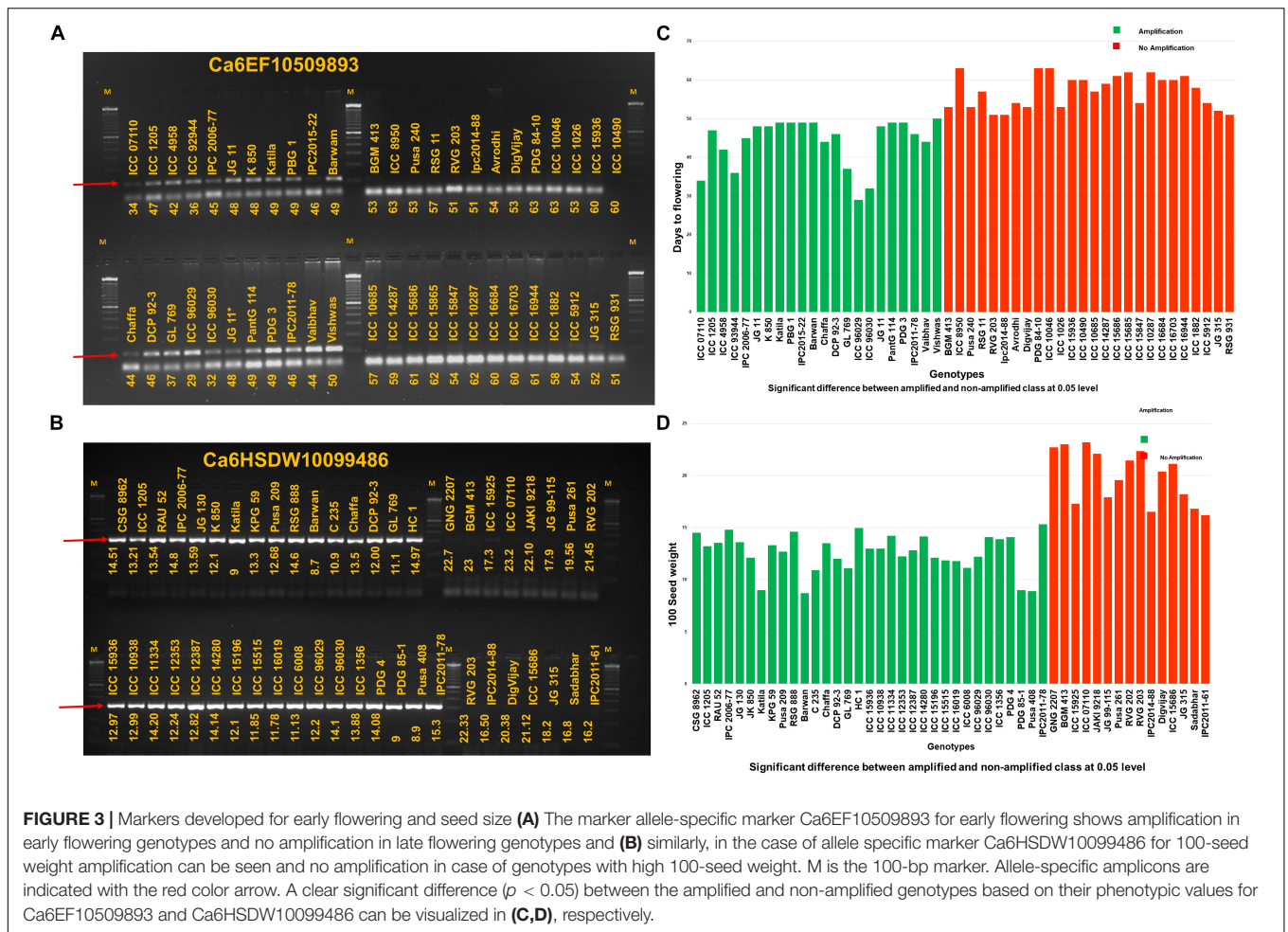
The development of climate-resilient cultivars will make small-holder agriculture profitable in the anticipated climate change scenarios. In addition to the integration of multidisciplinary approaches in breeding, the adoption of a 5Gs breeding approach will accelerate genetic gains as well as meet the future demands of nutritious food (Varshney et al., 2018, 2020). In the case of legumes, sequence-based breeding in the post-genome sequence era has improved the efficiency of developing climate-resilient cultivars considerably (Varshney et al., 2019a). In this study, we report the identification of genes and SNPs using a MutMap approach, as well as the development of markers for use in chickpea breeding programs toward the development of cultivars with early flowering and large seed size.

We developed a MutMap population by crossing ICC 4958-M3-2828 to ICC 4958 (wild type) to identify the candidate genes and causal SNPs for early flowering and large seed size. A previous MutMap approach and its variants were successfully deployed to localize the position of genes for agronomically important genes in cereals such as rice (Abe et al., 2012; Hu et al., 2016; Deng et al., 2017; Yuan et al., 2017; Nakata et al., 2018), sorghum (Jiao et al., 2018), and also in cucumber (Xu et al., 2018). This is the first study deploying a MutMap approach in chickpea. In this study, we demonstrated that MutMap is a powerful approach to identify causal homozygous mutation in bulked F<sub>2</sub> plants selected for a phenotype of interest. Although the MutMap method was initially considered for mapping of monogenic recessive (mutations) gene-controlled traits, it is now

possible to map dominant mutations through progeny testing and bulking of homozygous dominant F<sub>2</sub> individuals. In total 3,32,395 SNPs were used to develop a consensus reference genome sequence by replacing them at a particular position on the CDC Frontier genome. The sliding window analysis was applied to identify the trait linked SNPs with a SNP index value = 1. The extreme bulks sequenced reads were aligned to a consensus reference genome through a sliding window approach (moving averages). The SNPs in the sequenced F<sub>2</sub> population were in heterozygous state, show a 1:1 segregation pattern, 50% SNPs were mutant type, and the remaining 50% SNPs were wild type represented by a SNP index value of 0.5. If the F<sub>2</sub> population SNPs are in homozygous condition, then these are linked to the mutant phenotype (100% mutant reads, 0% wild-type reads represented as the SNP index value = 1). A SNP index value of 1 or near to 1 indicates the causal mutant SNP linked to the trait of interest, whereas a value of 0.5 indicates SNP not linked to the trait. The SNPs possess a SNP index value of 1 or near to 1 can be successfully targeted for marker development that can potentially be used in breeding. The number of SNPs among the bulks ranged from 3,993 (EF pool) to 5,771 (HSDW pool). Nevertheless, we identified only 102 candidate SNPs with a SNP index = 1. Interestingly, a genomic region harboring the candidate SNPs for all four bulks was on Ca6. In previous reports on chickpea, major QTLs for flowering time were reported on CaLG04 in the genomic region referred to as “QTL-hotspot.” Nevertheless, major QTLs corresponding to flowering time genes *efl-1* from ICCV 96029, *efl-3* from BGD 132, and *efl-4* from ICC 16641 were mapped on CaLG04, CaLG08, and CaLG06, respectively (Mallikarjuna et al., 2017). This indicates that flowering is a complex process coordinated by environmental and endogenous factors to ensure plant reproduction in appropriate conditions (Kumar et al., 2016). The “QTL-hotspot” was reported on Ca4 from 9.1 to 16.1 Mb (Varshney et al., 2014a). The relative positions of *efl1* and *efl2* genes mapped by Mallikarjuna et al. (2017) to “QTL-hotspot” were determined using the primer sequence of flanking markers NCPGR21 and GAA47 using *blastn-task* “*blastn-short*.” The marker for *efl1* (NCPGR21) was found to be present inside

**TABLE 4** | Summary of candidate genes in the genomic region on chromosome Ca6 harboring candidate SNPs with SNP-index = 1.

Reference/consensus base	SNP position	Gene	Sequence description	Biological process	Cellular component	Molecular function
C/T	10,685,694	Ca_08529	Subtilisin-like protease SDD1-like protein	Proteolysis	–	Serine-type endopeptidase activity
C/T	10,670,959	Ca_08530	Aspartokinase-homoserine dehydrogenase	Homoserine biosynthetic process, phosphorylation, oxidation–reduction process	–	Aspartate kinase activity, homoserine dehydrogenase activity, NADP binding
C/T	10,678,606	Ca_08530	Aspartokinase-homoserine dehydrogenase	Homoserine biosynthetic process, phosphorylation, oxidation–reduction process	–	Aspartate kinase activity, homoserine dehydrogenase activity, NADP binding
G/A	10,509,893	Ca_08544	Probable magnesium transporter NIPA1	Magnesium ion transmembrane transport	Early endosome, plasma membrane, integral component of membrane	Magnesium ion transmembrane transporter activity
G/A	10,498,876	Ca_08547	Beta-galactosidase 9	Carbohydrate metabolic process	Vacuolar membrane, plant-type cell wall, integral component of membrane	Beta-galactosidase activity, carbohydrate binding
C/T	10,432,389	Ca_08552	TBC1 domain family member 13-like	Intracellular protein transport, activation of GTPase activity	Intracellular	GTPase activator activity, Rab GTPase binding
G/A	10,394,606	Ca_08553	UV radiation resistance-associated-like protein	Protein targeting to vacuole, SNARE complex assembly, multivesicular body sorting pathway	Endosome, cytosol, integral component of membrane	SNARE binding
G/A	10,259,516	Ca_08563	Mediator of RNA polymerase II transcription subunit 10b-like	Regulation of transcription by RNA polymerase II	Mediator complex	Transcription coregulator activity
G/A	10,226,669	Ca_08566	Serine/threonine-protein phosphatase PP2A catalytic subunit isoform X2	Protein dephosphorylation	–	Phosphoprotein phosphatase activity
C/T	10,190,456	Ca_08570	Acyl-activating enzyme 17, peroxisomal protein, putative	–	Integral component of membrane	Catalytic activity
G/A	10,099,486	Ca_08581	Importin beta-3, putative	NLS-bearing protein import into nucleus, ribosomal protein import into nucleus	Nuclear membrane, nuclear periphery	Nuclear localization sequence binding, Ran GTPase binding
G/A	10,040,897	Ca_08587	Acyl-activating enzyme 17, peroxisomal protein, putative	–	Integral component of membrane	Catalytic activity
C/T	9,828,083	Ca_08609	Hypothetical protein glysoja_010758	–	–	–
G/A	10,599,353	Ca_08537	S-Adenosyl-L-homocysteine hydrolase	One-carbon metabolic process, S-adenosylhomocysteine catabolic process	Cytosol	Adenosylhomocysteinase activity, NAD binding
G/A	10,498,876	Ca_08547	Beta-galactosidase 9	Carbohydrate metabolic process	Vacuolar membrane, plant-type cell wall, integral component of membrane	Beta-galactosidase activity, carbohydrate binding
G/A	10,259,516	Ca_08563	Mediator of RNA polymerase II transcription subunit 10b-like	Regulation of transcription by RNA polymerase II	Mediator complex	Transcription coregulator activity
C/T	10,190,456	Ca_08570	Acyl-activating enzyme 17, peroxisomal protein, putative	–	Integral component of membrane	Catalytic activity
G/A	10,099,486	Ca_08581	Importin beta-3, putative	NLS-bearing protein import into nucleus, ribosomal protein import into nucleus	Nuclear membrane, nuclear periphery	Nuclear localization sequence binding, Ran GTPase binding
G/A	10,040,897	Ca_08587	Acyl-activating enzyme 17, peroxisomal protein, putative	–	Integral component of membrane	Catalytic activity
C/T	10,025,005	Ca_08590	Receptor-like protein kinase	Protein phosphorylation	Integral component of membrane	Protein serine/threonine kinase activity, ATP binding
G/A	9,890,335	Ca_08601	LRR receptor-like kinase family protein	Protein phosphorylation	Integral component of membrane	Protein kinase activity, ATP binding
C/T	9,828,083	Ca_08609	Hypothetical protein glysoja_010758	–	–	–



the “QTL-hotspot” at ~10 Mb on Ca4, whereas the marker for *efl2* (GAA47) was found present ~818.1kb upstream of the “QTL-hotspot” on the genome at ~8.3 Mb on Ca4. Further, the flowering time genes are distributed throughout the genome and are dependent on the genetic background. Genome-wide distribution of flowering time genes is not uncommon and was recently also reported in cucurbits (Yi et al., 2020).

On annotating the candidate SNPs in this genome region on Ca6, we identified that these SNPs are located within candidate genes that are involved in flowering time as well as in seed development. For instance, the candidate SNPs, Ca\_10137361 and Ca6\_11657245, present in Ca\_08579 and Ca\_25060 genes, respectively, are associated with calmodulin sensing Ca<sup>2+</sup> signals and are reported to be involved in flowering time (Kumar et al., 2016). The SNP Ca6\_10099486, present in gene Ca\_08581 that encodes putative importin beta-3, was reported to play a key role in drought tolerance in *Solanum tuberosum* (Xu et al., 2020). AtKPNB1, which is a member of the Arabidopsis importin family, was reported to be a gene associated with ABA sensitivity at germination, early seedling development, drought tolerance, and stomatal closure regulation; it is expressed in various organs and any specific tissues in listed organs such as leaves, roots, and flowers (Luo et al., 2013). Further, a

SNP Ca6\_10685694, present in gene Ca\_08529, encodes for subtilisin-like protease SDD1 (STOMATAL DENSITY AND DISTRIBUTION-1) and SDD1-like transcripts in *Solanum chilense* and *Solanum lycopersicum*. SDD1 is also known to play an important role in early leaf and flower development in both tomato species (Morales-Navarro et al., 2018). Similarly, a SNP Ca6\_10498876, present in gene Ca\_08547 that encodes beta-galactosidase9 to be expressed during fruit ripening, plays a major role in abscission, early onset of growth, and development processes in flowers and fruitlets (Wu and Burns, 2004). The genes Ca\_08587 and Ca\_08570 encode for the acyl-activating enzyme 17 (AAE17) reported to having a functional role in seed development. The Ca\_08530 gene encodes aspartokinase-homoserine dehydrogenase (AK/HSD) enzyme involved in aspartate kinase activity, homoserine dehydrogenase activity, and NADP binding activity. AK/HSD-GUS gene has been reported to be expressed in actively growing tissues and seed development. A SNP Ca6\_10685694, present in the gene Ca\_08529 located on Ca6, encodes subtilisin-like protease (SBT) SDD1-like protein. SBTs have been shown to control diverse developmental processes like stomatal distribution and density (Berger and Altmann, 2000; Von Groll et al., 2002). Two allele-specific markers, Ca6EF10509893 for early flowering and



Ca6HSDW10099486 for 100-seed weight, developed in this study were also validated on a select set of chickpea germplasm lines. These markers can be further tested on a larger germplasm panel with the potential to be converted to high-throughput assays for early-generation selection in chickpea breeding programs.

## CONCLUSION

MutMap has the advantage of both bulk segregant analysis and WGRS (whole-genome resequencing) approaches and enables the identification of candidate genes and causal SNPs. In the present study, we report 102 candidate SNPs in 22 candidate genes. The candidate genes identified in this study are involved in early flowering as well as enhanced seed size. Further, we also report the development and validation of markers for use with chickpea. Testing of these markers on a large and diverse panel of genotypes will be required prior to use in breeding programs for improving these traits.

## DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found at: <https://www.ncbi.nlm.nih.gov/bioproject/PRJNA715624>.

## AUTHOR CONTRIBUTIONS

MT conceived the project and secured funding. PM performed the experiments and prepared the first draft. BPM developed the MutMap population. DK, PB, and AWK performed the bioinformatics analysis. RP performed the experiments and phenotyped the population. UCJ contributed germplasm for marker validation. RKV, AC, and PS contributed the resources. MT and RKV wrote the review, edited manuscript and finalized the manuscript. All authors read and approved the manuscript.

## FUNDING

MT acknowledges financial support from the Science Engineering Research Board (SERB), Department of Science and Technology, Government of India (Grant ID: EEQ/2016/00348; Dated 20 February 2017). RKV is grateful to Bill and Melinda Gates Foundation for sponsoring the Tropical Legumes Project that partially supported this study (Grant ID: OPP1114827). This work has been carried out as part of the CGIAR Research

Program on Grain Legumes and Dryland Cereals. ICRISAT is a member of the CGIAR System Organization. PKM is grateful to Council for Scientific and Industrial Research (CSIR), Ministry of Science and Technology, Government of India for awarding him fellowship during his Ph.D. research work.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2021.688694/full#supplementary-material>

**Supplementary Figure 1** | Dendrogram constructed based on the 25 SSR markers distributed across chickpea genome indicates more than 95% similarity between ICC 4958 (wild type) and ICC 4958-M3-2828.

**Supplementary Figure 2** | Frequency distribution and correlation among days to flowering and 100-seed weight. Histogram showing the frequency distribution of days to flowering (A) and 100-seed weight (B). A negative correlation of 0.13 was observed between these traits.

**Supplementary Figure 3** | Single nucleotide polymorphism (SNP) index plots for all eight pseudomolecules for early flowering pool. The red line indicates the sliding window average of 2-Mb interval with an increment of 10 kb for SNP index.

**Supplementary Figure 4** | Single nucleotide polymorphism index plots for all eight pseudomolecules for late flowering pool. The red line indicates the sliding window average of 2 Mb interval with an increment of 10 kb for SNP index.

**Supplementary Figure 5** | Single nucleotide polymorphism index plots for all eight pseudomolecules for large seed size pool. The red line indicates the sliding window average of 2 Mb interval with an increment of 10 kb for SNP index.

**Supplementary Figure 6** | Single nucleotide polymorphism index plots for all eight pseudomolecules for small seed size pool. The red line indicates the sliding window average of 2 Mb interval with an increment of 10 kb for the SNP index.

**Supplementary Table 1** | Allelic data of 25 SSR markers distributed on the genome generated on mutant and wild lines.

**Supplementary Table 2** | Phenotyping data generated on MutMap population derived from ICC 4958-M3-2828 × ICC 4958.

**Supplementary Table 3** | Summary of the extreme bulks prepared for flowering time and seed size mutant based on phenotyping data generated on MutMap population derived from ICC 4958-M3-2828 × ICC 4958.

**Supplementary Table 4** | Summary of non-redundant candidate SNPs in the genomic region of Ca6 and their annotations.

**Supplementary Table 5** | Summary of SNPs identified in late flowering pool.

**Supplementary Table 6** | Summary of SNPs identified in early flowering pool.

**Supplementary Table 7** | Summary of SNPs identified in large seed size pool.

**Supplementary Table 8** | Summary of SNPs identified in small seed size pool.

**Supplementary Table 9** | Details of primer pairs designed for candidate SNPs.

## REFERENCES

- Abe, A., Kosugi, S., Yoshida, K., Natsume, S., Takagi, H., Kanzaki, H., et al. (2012). Genome sequencing reveals agronomically important loci in rice using MutMap. *Nat. Biotechnol.* 30:174. doi: 10.1038/nbt.2095
- Andrews, S. (2010). *FastQC—a Quality Control Tool for High Throughput Sequence Data*. Available online at: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/> (accessed January 20, 2019).
- Berger, D., and Altmann, T. (2000). A subtilisin-like serine protease involved in the regulation of stomatal density and distribution in *Arabidopsis thaliana*. *Genes Dev.* 14, 1119–1131.
- Berger, J. D., Ali, M., Basu, P. S., Chaudhary, B. D., Chaturvedi, S. K., Deshmukh, P. S., et al. (2006). Genotype by environment studies demonstrates the critical role of phenology in adaptation of chickpea (*Cicer arietinum* L.) to high and low yielding environments of India. *Field Crop Res.* 98, 230–244. doi: 10.1016/j.fcr.2006.02.007

- Bharadwaj, C., Tripathi, S., Soren, K. R., Thudi, M., Singh, R. K., Sheoran, S., et al. (2021). Introgression of “*QTL-hotspot*” region enhances drought tolerance and grain yield in three elite chickpea cultivars. *Plant Genome* 14:e20076. doi: 10.1002/tpg2.20076
- Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. doi: 10.1093/bioinformatics/btu170
- Deng, L., Qin, P., Liu, Z., Wang, G., Chen, W., Tong, J., et al. (2017). Characterization and fine-mapping of a novel premature leaf senescence mutant yellow leaf and dwarf 1 in rice. *Plant Physiol. Biochem.* 111, 50–58. doi: 10.1016/j.plaphy.2016.11.012
- Fang, N., Xu, R., Huang, L., Zhang, B., Duan, P., Li, N., et al. (2016). *Small grain 11* controls grain size, grain number and grain yield in rice. *Rice* 9:64. doi: 10.1186/s12284-016-0136-z
- FAOSTAT (2017). Available online at: <https://www.fao.org/faostat/>
- Fekih, R., Takagi, H., Tamiru, M., Abe, A., Natsume, S., Yaegashi, H., et al. (2013). MutMap+: genetic mapping and mutant identification without crossing in rice. *PLoS One* 8:e68529. doi: 10.1371/journal.pone.0068529
- Gaur, P. M., Samineni, S., and Varshney, R. K. (2014). Drought and heat tolerance in chickpea. *Legume Perspect.* 3:15.
- Hu, Y., Guo, L. A., Yang, G., Qin, P., Fan, C., Peng, Y., et al. (2016). Genetic analysis of dense and erect panicle 2 allele *DEP2-1388* and its application in hybrid rice breeding. *Yi Chuan Hereditas* 38, 72–81. doi: 10.16288/j.ycz.15-158
- Jaganathan, D., Bohra, A., Thudi, M., and Varshney, R. K. (2020). Fine mapping and gene cloning in the post-NGS era: advances and prospects. *Theor. Appl. Genet.* 133, 1791–1810.
- Jaganathan, D., Thudi, M., Kale, S., Azam, S., Roorkiwal, M., Gaur, P. M., et al. (2015). Genotyping-by-sequencing based intra-specific genetic map refines a “*QTL-hotspot*” region for drought tolerance in chickpea. *Mol. Genet. Genom.* 290, 559–571. doi: 10.1007/s00438-014-0932-3
- Jiao, Y., Burrow, G., Gladman, N., Acosta-Martinez, V., Chen, J., Burke, J., et al. (2018). Efficient identification of causal mutations through sequencing of bulked F<sub>2</sub> from two allelic bloomless mutants of *Sorghum bicolor*. *Front. Plant Sci.* 8:2267. doi: 10.3389/fpls.2017.02267
- Joshi, P. K., Parthasarathy Rao, P., Gowda, C. L. L., Jones, R. B., Silim, S. N., Saxena, K. B., et al. (2001). *The World Chickpea and Pigeonpea Economies Facts, Trends, and Outlook*. Patancheru: International Crops Research Institute for the Semi-Arid Tropics.
- Jukanti, A. K., Gaur, P. M., Gowda, C. L. L., and Chibbar, R. N. (2012). Nutritional quality and health benefits of chickpea (*Cicer arietinum* L.): a review. *Br. J. Nut.* 108, S11–S26. doi: 10.1017/S0007114512000797
- Kadiyala, M. D. M., Kumara Charyulu, D., Nedumaran, S., Moses Shyam, D., Gumma, M. K., and Bantilan, M. C. S. (2016). Agronomic management options for sustaining chickpea yield under climate change scenario. *J. Agrometeorol.* 18, 41–47.
- Kale, S. M., Jaganathan, D., Ruperao, P., Chen, C., Punna, R., Kudapa, H., et al. (2015). Prioritization of candidate genes in “*QTL-hotspot*” region for drought tolerance in chickpea (*Cicer arietinum* L.). *Sci. Rep.* 5, 1–14. doi: 10.1038/srep15296
- Klein, H., Xiao, Y., Conklin, P. A., Govindarajulu, R., Kelly, J. A., Scanlon, M. J., et al. (2018). Bulk-segregant analysis coupled to whole genome sequencing (BSA-Seq) for rapid gene cloning in maize. *G3 Genes Genom. Genet.* 8, 3583–3592. doi: 10.1534/g3.118.200499
- Kosugi, S., Natsume, S., Yoshida, K., MacLean, D., Cano, L., Kamoun, S., et al. (2013). Coval: improving alignment quality and variant calling accuracy for next-generation sequencing data. *PLoS One* 8:e75402. doi: 10.1371/journal.pone.0075402
- Kudapa, H., Garg, V., Chitkineni, A., and Varshney, R. K. (2018). The RNA-Seq-based high resolution gene expression atlas of chickpea (*Cicer arietinum* L.) reveals dynamic spatio-temporal changes associated with growth and development. *Plant Cell Environ.* 41, 2209–2225. doi: 10.1111/pce.13210
- Kumar, J., Kant, R., Kumar, S., Basu, P. S., Sarker, A., and Singh, N. P. (2016). Heat tolerance in lentil under field conditions. *Legume Genom. Genet.* 7, 1–11. doi: 10.5376/lgg.2016.07.0001
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics* 25, 2078–2079. doi: 10.1093/bioinformatics/btp352
- Luo, Y., Wang, Z., Ji, H., Fang, H., Wang, S., Tian, L., et al. (2013). An Arabidopsis homolog of importin  $\beta 1$  is required for ABA response and drought tolerance. *Plant J.* 75, 377–389. doi: 10.1111/tpj.12207
- Mallikarjuna, B. P., Samineni, S., Thudi, M., Sajja, S. B., Khan, A. W., Patil, A., et al. (2017). Molecular mapping of flowering time major genes and QTLs in chickpea (*Cicer arietinum* L.). *Front. Plant Sci.* 8:1140. doi: 10.3389/fpls.2017.01140
- Mannur, D. M., Babbar, A., Thudi, M., Sabbavarapu, M. M., Roorkiwal, M., Sharanabasappa, B. Y., et al. (2019). Super Annigeri 1 and improved JG 74: two Fusarium wilt-resistant introgression lines developed using marker-assisted backcrossing approach in chickpea (*Cicer arietinum* L.). *Mol. Breed.* 39, 1–13. doi: 10.1007/s11032-018-0908-9
- Megersa, A., Lee, D., Park, J., and Koh, H. J. (2015). Genetic mapping of a rice loose upper panicle mutant. *Plant Breed. Biotech.* 3, 366–375. doi: 10.9787/PBB.2015.3.4.366
- Morales-Navarro, S., Pérez-Díaz, R., Ortega, A., de Marcos, A., Mena, M., Fenoll, C., et al. (2018). Overexpression of a *SDD1-like* gene from wild tomato decreases stomatal density and enhances dehydration avoidance in arabidopsis and cultivated tomato. *Front. Plant Sci.* 9:940. doi: 10.3389/fpls.2018.00940
- Nakata, M., Miyashita, T., Kimura, R., Nakata, Y., Takagi, H., Kuroda, M., et al. (2018). MutMapPlus identified novel mutant alleles of a rice starch branching enzyme II b gene for fine-tuning of cooked rice texture. *Plant Biotechnol. J.* 16, 111–123. doi: 10.1111/pbi.12753
- Perrier, X., Flori, A., and Bonnot, F. (2003). “Data analysis methods,” in *Genetic Diversity of Cultivated Tropical Plants*, eds P. Hamon, M. Seguin, X. Perrier, and J. C. Glaszmann (Montpellier: Enfield), 43–76.
- Pratap, A., Chaturvedi, S. K., Tomar, R., Rajan, N., Malviya, N., Thudi, M., et al. (2017). Marker assisted introgression of resistance to fusarium wilt race 2 in Pusa 256, an elite cultivar of desi chickpea. *Mol. Genet. Genom.* 292, 1237–1245. doi: 10.1007/s00438-017-1343-z
- Roorkiwal, M., Bharadwaj, C., Barmukh, R., Dixit, G. P., Thudi, M., Gaur, P. M., et al. (2020). Integrating genomics for chickpea improvement: achievements and opportunities. *Theor. Appl. Genet.* 133, 1703–1720. doi: 10.1007/s00122-020-03584-2
- Sab, S., Lokesh, R., Mannur, D. M., Somasekhar, Jadhav, K., Mallikarjuna, B. P., et al. (2020). Genome-wide SNP discovery and mapping QTLs for seed iron and zinc concentrations in chickpea (*Cicer arietinum* L.). *Front. Nutr.* 7:559120. doi: 10.3389/fnut.2020.559120
- Singh, V. K., Garg, R., and Jain, M. (2013). A global view of transcriptome dynamics during flower development in chickpea by deep sequencing. *Plant Biotechnol. J.* 11, 691–701. doi: 10.1111/pbi.12059
- Singh, V. K., Khan, A. W., Jaganathan, D., Thudi, M., Roorkiwal, M., Takagi, H., et al. (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. doi: 10.1111/pbi.12567
- Takagi, H., Tamiru, M., Abe, A., Yoshida, K., Uemura, A., Yaegashi, H., et al. (2015). MutMap accelerates breeding of a salt-tolerant rice cultivar. *Nat. Biotechnol.* 33, 445–449. doi: 10.1038/nbt.3188
- Thudi, M., Bohra, A., Nayak, S. N., Varghese, N., Shah, T. M., Penmetta, R. V., et al. (2011). Novel SSR markers from BAC-end sequences, DArT arrays and a comprehensive genetic map with 1,291 marker loci for chickpea (*Cicer arietinum* L.). *PLoS One* 6:e27275. doi: 10.1371/journal.pone.0027275
- Thudi, M., Chitkineni, A., Liu, X., He, W., Roorkiwal, M., Yang, W., et al. (2016a). Recent breeding programs enhanced genetic diversity in both desi and kabuli varieties of chickpea (*Cicer arietinum* L.). *Sci. Rep.* 6, 1–10. doi: 10.1038/srep38636
- Thudi, M., Khan, A. W., Kumar, V., Gaur, P. M., Katta, K., Garg, V., et al. (2016b). Whole genome re-sequencing reveals genome-wide variations among parental lines of 16 mapping populations in chickpea (*Cicer arietinum* L.). *BMC Plant Biol.* 16:10. doi: 10.1186/s12870-015-0690-3
- Thudi, M., Palakurthi, R., Schnable, J. C., Chitkineni, A., Dreisigacker, S., Mace, E., et al. (2020). Genomic resources in plant breeding for sustainable agriculture. *J. Plant Physiol.* 257:153351. doi: 10.1016/j.jplph.2020.153351
- Tran, Q. H., Bui, N. H., Kappel, C., Dau, N. T. N., Nguyen, L. T., Tran, T. T., et al. (2020). Mapping-by-sequencing via MutMap identifies a mutation in ZmCLE7

- underlying fasciation in a newly developed EMS mutant population in an elite tropical maize inbred. *Genes* 11:281. doi: 10.3390/genes11030281
- Upadhyaya, H. D., Bajaj, D., Das, S., Saxena, M. S., Badoni, S., Kumar, V., et al. (2015). A genome-scale integrated approach aids in genetic dissection of complex flowering time trait in chickpea. *Plant Mol. Biol.* 89, 403–420. doi: 10.1007/s11103-015-0377-z
- Upadhyaya, H. D., Salimath, P. M., Gowda, C. L. L., and Singh, S. (2007). New early-maturing germplasm lines for utilization in chickpea improvement. *Euphytica* 157, 195–208. doi: 10.1007/s10681-007-9411-9
- Varshney, R. K., Gaur, P. M., Chamarthi, S. K., Krishnamurthy, L., Tripathi, S., Kashiwagi, J., et al. (2013a). 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* 6:3. doi: 10.3835/plantgenome2013.07.0022
- Varshney, R. K., Song, C., Saxena, R. K., Azam, S., Yu, S., Sharpe, A. G., et al. (2013b). Draft genome sequence of chickpea (*Cicer arietinum*) provides a resource for trait improvement. *Nat. Biotechnol.* 31, 240–246. doi: 10.1038/nbt.2491
- Varshney, R. K., Sinha, P., Singh, V., Kumar, A., Zhang, Q., and Bennetzen, J. L. (2020). 5Gs for crop genetic improvement. *Curr. Opin. Plant Biol.* 56, 190–196.
- Varshney, R. K., Mir, R. R., Bhatia, S., Thudi, M., Hu, Y., Azam, S., et al. (2014a). Integrated physical map with the genetic maps and reference genome sequence for chickpea (*Cicer arietinum* L.) improvement. *Funct. Integr. Genom.* 14, 59–73. doi: 10.1007/s10142-014-0363-6
- Varshney, R. K., Mohan, S. M., Gaur, P. M., Chamarthi, S. K., Singh, V. K., Srinivasan, S., et al. (2014b). Marker-assisted backcrossing to introgress resistance to fusarium wilt race 1 and ascochyta blight in C 214, an elite cultivar of chickpea. *Plant Genome* 7, 1–11. doi: 10.3835/plantgenome2013.10.0035
- Varshney, R. K., Thudi, M., Nayak, S. N., Gaur, P. M., Kashiwagi, J., Krishnamurthy, L., et al. (2014c). Genetic dissection of drought tolerance in chickpea (*Cicer arietinum* L.). *Theor. Appl. Genet.* 127, 445–462. doi: 10.1007/s00122-013-2230-6
- Varshney, R. K., Thudi, M., Pandey, M. K., Tardieu, F., Ojiewo, C., Vadez, V., et al. (2018). Accelerating genetic gains in legumes for the development of prosperous smallholder agriculture: integrating genomics, phenotyping, systems modelling and agronomy. *J. Exp. Bot.* 69, 3293–3312.
- Varshney, R. K., Pandey, M. K., Bohra, A., Singh, V. K., Thudi, M., and Saxena, R. K. (2019a). Toward the sequence-based breeding in legumes in the post-genome sequencing era. *Theor. Appl. Genet.* 132, 797–816.
- Varshney, R. K., Thudi, M., Roorkiwal, M., He, W., Upadhyaya, H. D., Yang, W., et al. (2019b). Resequencing of 429 chickpea accessions from 45 countries provides insights into genome diversity, domestication and agronomic traits. *Nat. Genet.* 51, 857–864. doi: 10.1038/s41588-019-0401-3
- Verma, S., Gupta, S., Bandhiwal, N., Kumar, T., Bharadwaj, C., and Bhatia, S. (2015). High-density linkage map construction and mapping of seed trait QTLs in chickpea (*Cicer arietinum* L.) using genotyping-by-sequencing (GBS). *Sci. Rep.* 5:17512. doi: 10.1038/srep17512
- Von Groll, U., Berger, D., and Altmann, T. (2002). The subtilisin-like serine protease SDD1 mediates cell-to-cell signaling during *Arabidopsis* stomatal development. *Plant Cell* 14, 1527–1539.
- Wangkumhang, P., Chaichoompu, K., Ngamphiw, C., Ruangrit, U., Chanprasert, J., Assawamakin, A., et al. (2007). WASP: a Web-based allele-specific PCR assay designing tool for detecting SNPs and mutations. *BMC Genom.* 8:275. doi: 10.1186/1471-2164-8-275
- Wu, Z., and Burns, J. K. (2004). A  $\beta$ -galactosidase gene is expressed during mature fruit abscission of ‘Valencia’ orange (*Citrus sinensis*). *J. Exp. Bot.* 55, 1483–1490. doi: 10.1093/jxb/erh163
- Xu, L., Wang, C., Cao, W., Zhou, S., and Wu, T. (2018). CLAVATA1-type receptor-like kinase *CsCLAVATA1* is a putative candidate gene for dwarf mutation in cucumber. *Mol. Genet. Genom.* 293, 1393–1405. doi: 10.1007/s00438-018-1467-9
- Xu, Y., Liu, L., Zhao, P., Tong, J., Zhong, N., Zhang, H., et al. (2020). Genome-wide identification, expression profile and evolution analysis of karyopherin  $\beta$  gene family in *Solanum tuberosum* group Phureja DM1-3 reveals its roles in abiotic stresses. *Int. J. Mol. Sci.* 21:931. doi: 10.3390/ijms21030931
- Yi, L., Wang, Y., Huang, X., Gong, Y., Wang, S., and Dai, Z. (2020). Genome-wide identification of flowering time genes in cucurbit plants and revealed a gene *CIGA2/KS* associate with adaption and flowering of watermelon. *Mol. Biol. Rep.* 47, 1057–1065. doi: 10.1007/s11033-019-05200-z
- Yuan, H., Fan, S., Huang, J., Zhan, S., Wang, S., Gao, P., et al. (2017). *OSG2/OsBAK1* regulates grain size and number, and functions differently in Indica and Japonica backgrounds in rice. *Rice* 10, 1–12. doi: 10.1186/s12284-017-0165-2

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

Copyright © 2021 Manchikatta, Kalavikatte, Mallikarjuna, Palakurthi, Khan, Jha, Bajaj, Singam, Chitikineni, Varshney and Thudi. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.