

NAME: Natural Methylation Epialleles Modulating Seed Size in Chickpea

Jitendra K. Mohanty

National Institute of Plant Genome Research (NIPGR)

Antima Yadav

National Institute of Plant Genome Research (NIPGR)

Virevol Thakro

National Institute of Plant Genome Research (NIPGR)

Uday Chand Jha

Indian Institute of Pulses Research (IIPR)

Shailesh Tripathi

Indian Institute of Pulses Research (IIPR)

Pinky Agarwal

National Institute of Plant Genome Research (NIPGR)

Swarup K. Parida

`swarup@nipgr.ac.in`

National Institute of Plant Genome Research (NIPGR)

Research Article

Keywords: Chickpea; Epialleles, Methylation, NGA1, Seed size, Seed weight

Posted Date: October 26th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11312720/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Epigenetic mutations, or epimutations, involve changes in gene expression that occur without altering the underlying DNA sequence. In plants, such epimutations arise more frequently than genetic mutations and can exhibit remarkable stability, thereby playing a crucial role in generating phenotypic diversity among crops. Despite their potential significance, the limited number of documented spontaneous epimutations in plants has hindered their integration into the quantitative dissection of complex traits and genetic improvement strategies. In this context, the present study offers a proof-of-concept in *Cicer arietinum* (chickpea), a major legume crop, by demonstrating that Natural Methylation Epialleles (NAME), rather than conventional nucleotide sequence variations, can regulate key agronomic traits such as seed size and seed weight. Specifically, we identified functionally relevant epialleles in the promoter region of a B3-domain transcription factor gene, *CaNGA1*, located within a major QTL associated with seed traits. These epialleles were found to modulate the brassinosteroid (BR) signalling pathway, ultimately impacting seed development. This pioneering insight into the regulatory role of NAME in chickpea establishes a foundation for exploiting epigenetic diversity to drive trait improvement and breeding innovation in legumes and other crops.

Key Message

Natural Methylation Epialleles (NAME) modulate traits of agronomic importance in chickpea

Introduction

Epigenetic mutations, or epimutations, represent heritable changes in gene expression without altering the underlying DNA sequence. They arise more frequently than genetic mutations in plants and can exhibit remarkable stability, contributing significantly to phenotypic diversity in crops (Hofmeister et al., 2017; Kawakatsu et al., 2016). Such heritable epigenetic variants, known as epialleles, are defined by differences in DNA methylation that are maintained through mitotic and/or meiotic inheritance and represent an important, yet often underappreciated, source of variation in crop species (Hofmeister et al., 2017). DNA methylation is a key epigenetic mechanism regulating gene expression and genome stability, occurring in CG, CHG, and CHH contexts (where H = A, C, or T), each maintained by distinct pathways. Dense methylation is generally associated with transcriptional silencing and transposable element repression. These epialleles arise from spontaneous or environmentally induced changes in DNA methylation without genetic alteration. Advances in next-generation sequencing (NGS), particularly bisulfite sequencing (BS-Seq), have enabled single-base resolution mapping of DNA methylation, revealing extensive epiallelic diversity across genomes. Increasing evidence highlights the role of DNA methylation in crop domestication and trait divergence (Lister & Ecker, 2009; Song et al., 2013; Kawakatsu et al., 2016; Shen et al., 2018), with stable heritable methylation marks associated with key agronomic traits such as plant architecture, flowering time, development, and yield (Manning et al., 2006; Miura et al., 2010; Quadrana et al., 2014; Gahlaut et al., 2020). Collectively, these findings suggest that epialleles may help explain the “missing heritability” of complex traits.

While crop improvement has traditionally focused on genetic variation, epimutations, given their higher frequency and stability, represent an additional heritable layer for breeding (Richards, 2006; Slatkin, 2009). Their ability to generate phenotypic diversity, even in the absence of genetic polymorphism, underscores their important role in trait evolution and adaptation (Manning et al., 2006; Martin et al., 2009; Bondada et al., 2020). Despite this potential, epigenetic variation remains largely underutilized in crop improvement strategies.

In this context, we identify a stably inherited epiallelic variation at the *CaNGA1* locus associated with seed size and weight in chickpea, in the absence of detectable sequence polymorphism within the corresponding QTL interval. Functional analyses further suggest that *CaNGA1* modulates components of the brassinosteroid (BR) signaling pathway, linking epiallelic regulation to seed development. These findings highlight the relevance of epigenetic variation as an additional layer of heritable regulation and underscore its potential utility in epigenetics-informed crop improvement strategies in chickpea and other legumes.

Materials and Methods

Development and large-scale field phenotyping of a mapping population

To initiate QTL mapping, an intra-specific F_7 RIL (recombinant inbred lines) population comprising of 284 mapping individuals was developed through inter-crossing between high 100-seed weight *desi* ICC 4958 ($30.0 \text{ g} \pm 1.5$) and low 100-seed weight *desi* ICC 5590 ($6.0 \text{ g} \pm 1.3$) chickpea accessions and further performed their generation advancement using single seed descent method. The 284 RILs and parental accessions were planted in a single row with $35 \times 10 \text{ cm}$ spacing in the experimental field at ICRISAT, Hyderabad (latitude/longitude: $17.1^\circ\text{N}/78.9^\circ\text{E}$) and NIPGR, New Delhi ($28.6^\circ\text{N}/77.2^\circ\text{E}$) during crop season, and grown for two consecutive years following randomized complete block design (RCBD) with at least two replications. To ascertain the homogeneity of RIL population across environments, the mapping parental accessions (ICC 4958 and ICC 5590) planted at the intervals of every 10 rows of the RILs was considered as reference in the field experimental design. The RIL mapping population was phenotyped for a major yield component trait 100-seed weight (SW) across said environments (two experimental years/locations) in the field. To estimate SW, the average weight (g) of 100-matured seeds at 10% moisture content was measured by selecting 8–10 representative plants of each RIL and parental accession. Various statistical parameters, including mean, standard deviation, frequency distribution, coefficient of variation (CV) and analysis of variance (ANOVA) of SW trait were estimated in RILs using SPSS v17.0 (<http://www.spss.com/statistics>). To determine the inheritance pattern of SW trait, the effects of genotypes (G), experimental years/environments (E) and $G \times E$ interactions were estimated using the two-way ANOVA. Broad-sense heritability [$H^2 = \sigma^2_g / (\sigma^2_g + \sigma^2_{ge}/n + \sigma^2_e/nr)$] of SW trait was measured based on σ^2_g (genetic), σ^2_{ge} ($G \times E$) and σ^2_e (error) variance with n (number of environments) = 2 and r (number of replicates) = 2.

Mining and genotyping of genome-wide SNPs

For simultaneous mining and genotyping of SNPs at a genome-wide scale through genotyping-by-sequencing (GBS) assay, the genomic DNA was isolated from the leaves of 284 RIL mapping individuals using a QIAGEN DNeasy 96 Plant Kit (QIAGEN, CA, USA) in accordance with the manufacturer's instructions. The high-quality genomic DNA isolated from the RILs and parental accessions were digested with *ApeKI*, and further ligated to adapters carrying unique barcodes to construct the 3 x 96-plex GBS libraries (Elshire et al., 2011; Spindel et al., 2013). The pooled GBS libraries were sequenced (100 bp pair-end sequence reads) with the Illumina HiSeq2000 (Illumina, San Diego, CA, USA) NGS (next-generation sequencing) platform (Kujur et al. 2015a, b, c). The reproducibility of the GBS assay was evaluated using the two mapping parental accessions (ICC 4958 and ICC 5590) as biological replicates. The raw FASTQ paired-end sequence reads generated from the individual RILs were filtered and quality assessed using the recommended Illumina pipeline (FastQC) and NGS QC Toolkit v2.3 (Patel and Jain, 2012). The high-quality sequence reads with the *phred* score of ≥ 10 was de-multiplexed using the unique barcodes to extract the sequence reads of each RIL.

To mine genome-wide SNPs, the sequence reads of each RIL were mapped to the reference chickpea genome (*kabuli* CDC Frontier; Varshney et al., 2013) employing Burrows-Wheeler Aligner tool (Bowtie v2.1.0; Langmead and Salzberg, 2012), and generated the sequence alignment map (SAM) files. Further, these SAM files were analyzed with the GBS pipeline of STACKS v1.0 (Catchen et al., 2013; <http://creskolab.uoregon.edu/stacks>) as per Kujur et al. (2015a, b, c) to mine the accurate genome-wide SNPs from the RILs. The high-quality valid SNPs with base quality of 20 and supported by minimum sequence read-depth of 10 were screened using the maximum likelihood statistical model of STACKS. The SNPs were structurally and functionally annotated on the coding/non-coding sequence regions of genes and genome of chickpea (Varshney et al., 2013) using the single-nucleotide polymorphism effect predictor (Snpeff v3.1h; <http://snpeff.sourceforge.net>) and PFAM database v27.0 (<http://pfam.sanger.ac.uk>) following Kujur et al. (2015a, b).

Genetic linkage map construction and high-resolution QTL mapping

To construct a genetic linkage map, the genotyping information of genome-wide SNPs showing polymorphism between parental accessions (ICC 4958 and ICC 5590) generated from 284 RILs were analyzed using the JoinMap 4.1 (<https://www.kyazma.nl/index.php/JoinMap>) at the LOD (logarithm of odds) threshold of 10.0 with Kosambi mapping function (as per Kujur et al., 2015b). The SNPs were integrated based on their centiMorgan (cM) genetic distance and corresponding marker physical positions (bp) on eight linkage groups (LG1 to LG8)/chromosomes to generate a high-density genetic linkage map of chickpea.

For high-resolution QTL mapping, the genotyping data of parental polymorphic SNPs mapped on eight LGs/chromosomes of a high-density genetic linkage map as well as multi-environment field phenotypic data of SW trait among 284 RIL mapping individuals were analyzed with a composite interval mapping

(CIM) function ($\text{LOD} > 5.0$ with 1000 permutations and $p \leq 0.05$) of MapQTL 6 (Van Ooijen, 2009) following Das et al. (2015) and Kujur et al. (2015b). The best fit QTL model was determined using the step-wise forward selection and backward elimination method. Accordingly, PVE (phenotypic variation explained %) and positional/additive effect (parental favorable alleles) defined by each major QTL on SW trait were estimated at a significant LOD ($p \leq 0.05$) on the final selected QTL model using the ANOVA (Bajaj et al. 2015b; Saxena, et al., 2015a; Upadhyaya, et al., 2015b). The confidence interval (CI) of individual significant major QTL peak was estimated employing ± 1 -LOD support intervals (95% CI).

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

A major SW QTL (*CaqSW3.1*) genomic region identified by high-resolution QTL mapping in a RIL population (ICC 4958 $\times$ ICC 5590) was targeted for fine-mapping. For this, GBS-SNPs within *CaqSW3.1* genomic interval showing polymorphism between parental genotypes (ICC 4958 and ICC 5590) were genotyped among 764 mapping individuals of the same F_7 RIL population using the Sequenom MALDI-TOF MassARRAY (Saxena et al., 2014a, b). To perform high-resolution QTL mapping, the SNP genotyping information at the *CaqSW3.1* genomic region was further correlated with the SW field phenotypes of RILs following the aforesaid strategy.

To further fine-map the *CaqSW3.1* QTL genomic interval, two BC_3F_3 chickpea NILs contrasting with high (HSW-NIL^{*CaqSW3.1*})- and low (LSW-NIL^{*CaqSW3.1*})-SW phenotypes were generated using ICC 5590 and ICC 4958 as the recurrent parents, respectively in the back-cross mapping population development breeding program. For this, primarily the marker-assisted foreground selection was performed for which the SNPs flanking/tightly-linked to the *CaqSW3.1* major QTL region were genotyped among mapping individuals of the back-cross population by the MALDI-TOF MassARRAY (Saxena et al., 2014a, b). Subsequently, marker-assisted background selection through large-scale genotyping of 960 SNPs mapped uniformly across eight chickpea chromosomes was performed among the selected recombinants (back-cross mapping individuals) using Illumina GoldenGate assay (Bajaj et al., 2015b). The back-cross mapping individuals were further phenotyped comprehensively for 100-SW in the field across multiple environments (two locations/years) following RCBD with at least three replications (as per aforesaid phenotyping methods). The high and low SW BC_3F_3 NILs were inter-crossed among each other to develop a F_2 mapping population, (HSW-NIL^{*CaqSW3.1*} $\times$ LSW-NIL^{*CaqSW3.1*}) consisting of 345 individuals. The parental polymorphic SNPs within the *CaqSW3.1* major QTL genomic interval were genotyped among 345 mapping individuals using the MALDI-TOF assay (Saxena et al., 2014a, b). Multi-environment field phenotyping of these mapping individuals for the SW trait, and subsequently fine-mapping of SW QTL was performed following aforesaid strategy. To perform progeny testing, both homozygous recombinant and non-recombinant individuals from low and high SW NILs were screened according to their overall genetic constitution including recombination among SNPs flanking/tightly-linked to the *CaqSW3.1*. The recombinant and non-recombinant progenies were field phenotyped (RCBD with three replications) individually for SW trait across environments (two years). Significant SW trait variations between recombinant and non-recombinant progenies were assessed using a statistical *t*-test.

Genomic constitution and phylogeny of B3-domain containing NGA1 gene in chickpea

To determine the genomic organization and identity of B3-domain containing transcription factor gene *NGA1* in the chickpea genome (Varshney et al., 2013), a genome-wide scan (*E-value* cutoff of $1e-10$) was carried out with HMMER (<http://hmmer.org>). The amino acid sequences of *NGA1* protein accessible at a global database (<https://www.ncbi.nlm.nih.gov>) were used to constitute a hidden markov model (HMM) profile. Subsequently, the amino acid sequences of output HMM were examined with INTERPRO (<https://www.ebi.ac.uk/interpro>) to assure the occurrence of functional B3 domains in the identified *NGA1* genes.

The amino acid sequence of a B3-domain containing chickpea *NGA1* gene (*CaNGA1*) was used for reciprocal BLAST search against that from the genomes of sequenced legumes including *Cajanus cajan*, *Cicer arietinum*, *Glycine max*, *Pisum sativum*, *Medicago truncatula*, *Lotus japonicus*, and *Trifolium pretense*, (Phytozome; <https://phytozome.jgi.doe.gov/pz/portal.html>, and Legume Information System; <https://legumeinfo.org/>) as well as with the model dicot, *Arabidopsis thaliana* (<https://www.arabidopsis.org/>) and monocot *Oryza sativa*. INTERPRO was used to ascertain the occurrence of valid signature B3-domains in the orthologous *NGA1* gene sequences. Multiple sequence alignment of the B3-domain containing amino acid sequence encoded by a chickpea *CaNGA1* gene with its counterparts *Arabidopsis* and legumes orthologs was performed using the CLUSTALW (BLOSSUM62). Correspondingly, an unrooted Neighbour-Joining (NJ) phylogenetic tree was constructed and visualized with MEGA7 (www.megasoftware.net/, Kumar et al. 2016).

Targeted amplicon resequencing

The *CaNGA1* gene region was sequenced with the genomic DNA of high and low SW NILs and RIL mapping parental genotypes (ICC 4958 and ICC 5590) through the multiplexed amplicon-sequencing strategies (TruSeq Custom Amplicon v1.5 as per the manufacturer's instructions) using the Illumina MiSeq NGS platform (Saxena et al. 2014 a,b). The CDSs/exons, introns, and 2-kb upstream/downstream regulatory regions (URRs/DRRs) of *CaNGA1* gene were targeted to design and synthesize the custom oligo probes. The synthesized oligo probes were further pooled into the custom amplicon tube to constitute the template libraries. The individual libraries were tagged with the sample-specific unique indices and pooled. The clusters of normalized uniquely tagged pooled amplicon libraries were sequenced with the Illumina MiSeq platform. The mapping of high-quality amplicon sequence reads onto the reference chickpea genome and detection/annotation of sequence variants were carried out as per aforesaid methods.

Real-time quantitative PCR (RT-qPCR) assay

To determine the differential gene regulatory function of a *CaNGA1*, the expression profiling of said gene was performed using the semi-quantitative and quantitative RT-PCR assays. Total high-quality RNA was isolated from both vegetative tissues (root, shoot, and leaf) and reproductive tissues (flower, pod,

embryo, cotyledon, and mature seed), including specific seed developmental stages: DS1 (0–10 DAP), DS2 (11–20 DAP), DS3 (21–30 DAP), and DS4 (31–40 DAP). These samples were collected from the parental lines of the RIL mapping population (ICC 5590 and ICC 4958) as well as from the low- and high-seed weight NILs.

Although previous studies by Seita et al. (1985), Pradhan et al. (2014) and Garg et al. (2017) have proposed different classification systems for chickpea seed development, the present study adopts a consensus framework that aligns with and integrates both classifications. Specifically, the selected stages correspond to key developmental phases: 0–10 DAP represents the early embryonic stage (cell division and morphogenesis), 10–20 DAP represents late embryonic stage (Organogenesis and late embryogenesis), 20–30 DAP marks the mid-stage (cell expansion and storage accumulation), and 40 DAP reflects the maturation phase. This tailored classification was designed to match the growth dynamics of the accessions used in our study. The RNA isolation was performed using the RNeasy MinElute Cleanup Kit (QIAGEN, USA) and the LiCl method for tender and seed tissue, respectively (Pradhan et al., 2014). About 1 µg of high-quality RNA isolated from the tissues/stages of NILs/genotypes was used to synthesize cDNA with the Verso cDNA synthesis kit (ThermoFisher Scientific, USA). The diluted cDNAs (three biological replicates each with three technical replicates) were amplified using *CaNGA1* gene-specific forward and reverse primers (**Table S2**) and iTaq Universal SYBR Green Supermix (Biorad, USA) on a CFX96 Real-Time PCR system (Biorad, USA) to estimate relative gene expression level as per Bajaj et al. (2015a) and Upadhyaya et al. (2015). To normalize expression, a housekeeping gene *ACTIN* was utilized as an endogenous control in RT-qPCR assay (Garg et al. 2010). The $2^{-\Delta CT}$ method and *two-tailed t test* were used to determine the significant differences in gene expression level among analysed samples (Malik et al. 2016; Mukhopadhyay and Tyagi 2015). The significant differential gene expression level was estimated across samples and further visualized with a heat map using the MultiExperiment Viewer (MeV, <http://www.tm4.org/mev>).

Treatment of NILs with 5-aza-2'deoxyctidine (5-aza-dC)

To validate the correlation between differential DNA methylation states of *CaNGA1* gene with its expression profile, the soaked chickpea seeds of high- and low-SW NILs were incubated for 16 h at 37°C. Subsequently, the chickpea seeds were submerged in 20 mM Tris-HCl (pH 7.5) with or without 0.3 mM 5-aza-dC (Sigma, USA), an inhibitor of DNA methylation (Chang and Pikaard, 2005) for 48 h at room temperature in the dark. Further, chickpea seeds after washing were planted in the pots filled with soil and grown at 22–24°C in the greenhouse. The developing seeds from both control and treated plants were harvested at developing stage3 (DS3) and subsequently the expression of *CaNGA1* along with the BR responsive genes (*CPD1*, *BRI1*, *BRZ1*) and seed size responsive genes (*CaSHB1*, *CaIKU1*, *CaMINI3*, *CaIKU2*, *CaKLU*) were quantified using gene specific primers (**Table S2**) through quantitative real-time PCR (qRT-PCR) assay.

To validate the epiallelic regulation of *CaNGA1*, bisulfite-PCR (BS-PCR) followed by Sanger sequencing was performed at the DS3 stage. Genomic DNA was extracted from contrasting NILs under both control and treatment conditions at the DS3 developmental stage. Bisulfite conversion of genomic DNA was carried out using the EZ DNA Methylation-Gold™ Kit (Zymo Research). Before performing BS-PCR, we assessed bisulfite conversion efficiency and potential DNA degradation by amplifying both BS-treated and untreated (non-converted) genomic DNA using a gene-specific marker for LOC101513448 designed from available genomic sequence. As expected, the marker amplified only the non-converted DNA, while no product was obtained from the BS-treated DNA. This confirmed successful bisulfite conversion of the treated samples. Bisulfite-specific primers, for the targeted *CaNGA1* promoter, were designed using MethPrimer 2.0 (Li et al., 2002). PCR amplification was conducted as per the protocol described by (Ding et al., 2018). Amplified products were separated on a 2% agarose gel stained with ethidium bromide. Bands corresponding to the expected product size were excised and purified using the GeneJET PCR Purification Kit (Thermo Scientific). The purified clones from each sample were subjected to Sanger sequencing. Resulting sequences were analyzed using BioEdit v7.7.1 and aligned with Clustal Omega for methylation pattern assessment.

Dual Luciferase Reporter Effector Assay

The promoter region of *CaBZR1* derived from the HSW chickpea genotype ICC 4958 was initially cloned into the pENTR entry vector and subsequently recombined into the p635nRRF destination vector using Gateway LR Clonase II (Kumar et al. 2018). The resulting reporter constructs harboured Firefly luciferase (LUC) driven by the *CaBZR1* promoter, along with Renilla luciferase (REN) under the control of the 35S promoter, which served as an internal reference. For effector construction, the full-length coding sequence of *CaNGA1* was cloned into the pMDC32 vector downstream of a dual CaMV 35S promoter. Both reporter and effector plasmids were transformed into *Agrobacterium tumefaciens* and co-infiltrated into fully expanded leaves of *Nicotiana benthamiana*. Leaf tissues were harvested 48 h post-infiltration, and luciferase activities were quantified using the Dual-Luciferase Reporter Assay System (Promega, Madison, WI, USA) with a POLARstar Omega microplate reader (BMG Labtech, Ortenberg, Germany). Firefly luciferase activity was normalised against Renilla luciferase to account for variation in infiltration efficiency. Infiltrated leaf areas were additionally visualised under low-light conditions using a CCD-based imaging system (Bio-Rad).

Treatment of NILs with 24-epibrassinolide (BR) and brassinazole (BRZ)

To decipher the gene regulatory role of *CaNGA1* epialleles in modulating SW in chickpea through Brassinosteroid (BR) biosynthesis and signalling, the early maturation (embryo) seed development stage (DS1) of high and low SW NILs were treated individually with or without 10 µm of 24-epibrassinolide (BR) and 1 µm of brassinazole (BRZ), an inhibitor of brassinolide biosynthesis for 3–4 h (Asami et al. 2001; Jiang et al. 2013). Subsequently, the expression profile of BR responsive (*CaBZR1*, *CaCPD1*, *CaBRI1*) and

seed size responsive genes (*CaSHB1*, *CaIKU1*, *CaMINI3*, *CaIKU2*, *CaKLU*) were quantified using quantitative real-time PCR (qRT-PCR) assay. To estimate SW, the mean weight (g) of 100-matured seeds of control untreated as well as BR- and BRZ-treated low and high SW NILs (5 biological/technical replicates) was measured at 10% moisture content.

Expression profiling of known BR-signalling and seed size-regulating genes

Total RNA was extracted from developing seed tissues of the two contrasting parental accessions (ICC 5590 and ICC 4958) as well as from the NIL lines (HSW-NIL and LSW-NIL), both under treated (5-aza-dC, BR, BRZ) and control untreated conditions following the LiCl method (Pradhan et al. 2014). The extracted RNA was quantified using a Nanodrop™ 8000 spectrophotometer (Thermo Fisher Scientific, USA). For cDNA synthesis, 1 microgram of total RNA was used with the Verso cDNA synthesis kit (ThermoFisher Scientific, USA). The diluted cDNA from each sample (with three biological and three technical replicates) was amplified using 200 nM of forward and reverse gene-specific primers along with iTaq Universal SYBR Green Supermix (Biorad, USA) on a CFX96 Real-Time PCR system (Biorad, USA). The housekeeping gene *ACTIN* was used as an internal reference to normalize the gene expression pattern (Garg et al. 2010). The relative fold expression levels of all the specified genes in the treated compared to the control untreated tissues, were determined using the comparative *Ct* method and expressed as $2^{-\Delta\Delta C_t}$ values (Bajaj et al. 2015a; Malik et al. 2020).

Results

Identification and fine-mapping of a major QTL controlling seed weight in chickpea

To identify the genetic region regulating seed size and weight (SW) in chickpea, we used high-resolution genetic mapping and fine-mapping. An F_7 recombinant inbred line (RIL) population ($n = 284$) derived from contrasting *desi* accessions ICC 4958 (100-SW: 30 g) and ICC 5590 (100-SW: 6 g) was used for QTL mapping (Fig. 1A). The RILs exhibited a wider phenotypic variation with normal frequency distribution of seed weight, indicating quantitative inheritance of the trait (Fig. 1B). This intra-specific RIL mapping population was further used to construct an ultra-high-density genetic linkage map by assigning 61415 whole-genome resequencing-based SNPs across eight chickpea chromosomes (**Table S1**), with an average map density of 0.08 cM (5.654 kb). By integrating genome-wide genotyping data of 61415 SNPs mapped on an ultra-high-density genetic linkage map with multi-environment 100-SW field phenotypes from 284 RILs, a major robust QTL (*CaqSW3.1*) genomic region (1100 kb) governing 100-SW was identified on chromosome 3 of chickpea (Fig. 1C). This QTL explained 43.4% of the phenotypic variation in seed weight and exhibited a LOD score of 15.2, indicating its robustness and significance. To refine the QTL interval, polymorphic SNPs within the *CaqSW3.1* region were further genotyped in an expanded set of 764 F_7 RILs (ICC 4958 $\times$ ICC 5590) and 345 F_2 individuals from a NIL mapping population (HSW-

NIL^{CaqSW3.1} × LSW-NIL^{CaqSW3.1}). This process narrowed-down the initial 1100 kb *CaqSW3.1* QTL region into a smaller 5.92 kb genomic interval, with 15–27 recombinants (Fig. 1D). Further, the comprehensive 100-SW phenotyping following progeny testing of all 27 recombinants identified a 5.92 kb *CaqSW3.1* QTL genomic region in the five promising recombinant lines of NILs where tightly-linked SNPs showed no recombination with the *CaNGA1* gene (Fig. 1E). Therefore, a B3-domain containing *CaNGA1* (gene accession ID *Ca_01368*), delineated within a major *CaqSW3.1* QTL, was identified as the causal gene responsible for controlling 100-SW in chickpea. The genomic structure and extensive phylogenetic analysis of *CaNGA1* gene orthologs, annotated from the genomes of various legumes and model crops, revealed a high degree of conservation in the B3 domains encoded by this gene across both legumes and model crops (Fig. 2A, 2B). Differential gene expression profiling of *CaNGA1* exhibited its seed-specific expression, which steadily increased with the progression of seed development from the developmental stage DS1 (0–10 days after pollination) to DS4 (31–40 days) (Fig. 2C). A significant increase (> 10.5-fold) in *CaNGA1* gene transcript accumulation was observed during the mid to later stages of seed development (DS3 and DS4) including cotyledons and mature seeds of high-SW lines (ICC 4958 and HSW-NIL^{CaqSW3.1}), compared to the low-SW lines (ICC 5590 and LSW-NIL^{CaqSW3.1}). This enhanced expression pattern aligns with physiological changes, as the increase in dry seed weight becomes markedly more pronounced around 21 days after pollination (DAP), corresponding to the DS3 stage of seed development (Pradhan et al., 2014).

Epiallelic regulation of *CaNGA1* through promoter methylation controls seed size variation

To understand the molecular basis of differential *CaNGA1* expression associated with seed size variation, the complete gene sequence, including 3-kb regulatory regions, exons, and introns, was compared between high-SW (ICC 4958 and HSW-NIL^{CaqSW3.1}) and low-SW lines (ICC 5590 and LSW-NIL^{CaqSW3.1}). Notably, no nucleotide sequence variations were detected in the *CaNGA1* gene between the high-SW and low-SW lines, suggesting that differential expression may be attributed to epigenetic regulation rather than genetic polymorphism. Given the absence of genetic variation, the methylation status of the *CaNGA1* promoter was examined by bisulfite-PCR (BS-PCR) using methylation-specific primers (**Table S2**). Analysis of the DS3 seed developmental stage revealed four differentially methylated CG sites (Ca3_39814439 bp, Ca3_39814444 bp, Ca3_39814454 bp, Ca3_39814495 bp) between NILs, where LSW-NIL showed the highest methylation (C/T ratio = 0.17) compared to HSW-NIL (C/T ratio = 0.05). (Fig. 2I) This is clearly evident from the presence of non-converted cytosines in the bisulfite-treated DNA of LSW-NIL. Overall, the average CG methylation rate in the promoter of *CaNGA1* was approximately 0.315 in LSW-NIL versus 0.10 in HSW-NIL, supporting the hypothesis that promoter methylation level regulates *CaNGA1* expression.

Treatment of germinating seeds with 5-aza-dC has been previously shown in multiple crops to reduce methylation, with effects persisting into later developmental stages and sometimes being transmitted to the next generation (Jabre et al., 2022; Akimoto et al., 2007; Sano et al., 1990; Fieldes et al., 1994; Katoh

et al., 1993). To examine whether promoter methylation influenced *CaNGA1* expression, germinating seeds were treated with the DNA methylation inhibitor 5-aza-2'-deoxycytidine (5-aza-dC). Bisulfite sequencing of DS3-stage seeds from treated plants showed extensive C-to-T conversion at previously methylated cytosine residues in LSW-NIL, indicating effective demethylation of the *CaNGA1* promoter. In contrast, the HSW-NIL showed similar C-to-T conversion even without 5-aza-dC treatment, consistent with naturally hypomethylated promoter (Fig. 2I). To further confirm the role of DNA methylation in regulating *CaNGA1* expression, we observed a significant difference in *CaNGA1* transcript levels between 5-aza-dC (DNA methylation inhibitor)-treated and untreated LSW-NILs whereas only marginal changes were observed between HSW-NIL (Fig. 2F). These results demonstrate a strong association between promoter CG methylation and *CaNGA1* transcription, identifying naturally occurring methylation epialleles as the primary source of expression variation at the *CaNGA1* locus that consequently determine seed size in chickpea. Based on the findings from the above experiments, the contrasting NILs harboring functional epialleles of the differentially methylated *CaNGA1* promoters displayed marked differences in seed size and weight, strongly indicating a regulatory role for *CaNGA1* in chickpea seed development.

CaNGA1 encodes a B3-domain transcription factor associated with seed development

The fine-mapped candidate gene, *CaNGA1*, encodes a B3-domain-containing transcription factor. The B3 transcription factor superfamily, to which *CaNGA1* belongs, is broadly classified into four families: ARF, LAV, RAV, and REM. Among these, the RAV family is further subdivided into three subfamilies: RAV, TEMPRANILLO (TEM), and NGATHA (NGA). *CaNGA1*, a member of the NGA subfamily, specifically binds to the 5'-CACCTG-3' DNA motif and has been implicated in several key developmental processes, including seed development (Zhang et al., 2015; Sasnauskas et al., 2018; Guo et al., 2020; Salava et al., 2022). Structural annotation of *CaNGA1* confirms the presence of a single B3 DNA-binding domain (Fig. 2A). Comparative phylogenetic analysis grouped *CaNGA1* with NGA1 homologs from legumes and model plant species, with the closest relationship observed with *Medicago truncatula* NGA1 (MtNGA1), indicating a high degree of evolutionary conservation (Fig. 2B). Notably, the *Arabidopsis* ortholog *AtNGA1* has been experimentally validated to possess transcriptional activation potential, directly regulating downstream gene expression (Sato et al., 2018). Moreover, the RAV family of transcription factors, to which *CaNGA1* belongs, is well known to participate in brassinosteroid (BR) signaling pathways, modulating key developmental traits such as leaf angle and seed size (Zhang et al., 2015). In this context, the potential role of *CaNGA1* in BR signalling and its contribution to seed size regulation in chickpea warrants investigation.

Epiallelic regulation of CaNGA1 is associated with brassinosteroid signaling

BRs are plant steroid hormones synthesized via multiple biosynthetic genes (e.g., D2, D11, BRD1) and perceived by the BRI1 (Brassinosteroid Insensitive 1) receptor kinase located at the cell surface (Je et al., 2010). The BR signal is subsequently transduced via key transcription factors BZR1/BES1, which regulate a suite of BR-responsive genes. Thus, as a member of the RAV transcription factor family,

CaNGA1 may act downstream or in parallel to these signalling components to influence seed size through transcriptional regulation. To explore the regulatory role of *CaNGA1* epialleles in BR biosynthesis and signaling, the binding site of *CaNGA1* was scanned in the 3-kb *cis*-regulatory elements of chickpea homologs of several BR biosynthetic (*CaCPD1*) and BR-signalling (*CaBRI1*, *CaBZR1*) genes. The *in-silico* mining for binding site prediction revealed several *cis*-regulatory elements for the *CaNGA1* in the promoter region of aforesaid BR biosynthetic/signaling genes. To experimentally validate this prediction, a dual-luciferase reporter-effector assay was performed in *Nicotiana benthamiana* leaves using the *proCaBZR1:LUC* reporter constructs together with a double 35S::*CaNGA1* effector construct (Fig. 2E). The relative LUC activity derived from the *CaBZR1* promoter was upregulated when *proCaBZR1:LUC* was co-infiltrated with double 35S: *CaNGA1*, but not with empty vector control. This high LUC/REN ratio (0.448) in the presence of *CaNGA1* compared to that in the control (0.076) suggested *CaNGA1*-mediated transcriptional activation of *CaBZR1*.

To further assess the influence of promoter methylation on BR signaling, transcript levels of these genes were analysed following treatment with the DNA methylation inhibitor 5-aza-2'-deoxycytidine (5-aza-dC). In LSW-NIL, demethylation induced a significant increase in *CaNGA1* expression, accompanied by coordinated upregulation of *CaCPD1*, *CaBRI1*, and *CaBZR1* (Fig. 2F). In contrast, only minor changes in transcript abundance were observed in HSW-NIL following treatment. Collectively, these results demonstrate that promoter methylation of *CaNGA1* is associated with coordinated changes in the expression of key components of the brassinosteroid biosynthesis and signalling pathway.

CaNGA1 -mediated BR signaling regulates downstream seed size genes and phenotypic variation

The role of brassinosteroids (BRs) in determining seed size and shape has been well established, primarily through their ability to transcriptionally regulate developmental pathways involved in integument, endosperm, and embryo formation (Jiang et al., 2013). Manipulation of endogenous BR levels, either by enhancing biosynthesis or modulating signaling components, has been shown to improve seed yield. Notably, the BR-activated transcription factor BZR1 directly regulates a set of well-characterized seed size genes in *Arabidopsis*, including *SHB1*, *IKU1*, *MINI3*, *IKU2*, *HSF15*, and *KLU* (Jiang et al., 2013; Zhu et al., 2015). To determine whether differential *CaNGA1* expression influences downstream regulators of seed development, the transcript abundance of chickpea homologs of known seed size-related genes (*CaSHB1*, *CaIKU1*, *CaMINI3*, *CaIKU2*, and *CaKLU*) was examined in the contrasting NILs. (Fig. 2F). All five genes exhibited significantly higher expression in HSW-NIL compared to LSW-NILs, indicating their likely contribution to enhanced seed size. Importantly, the expression patterns of these genes were closely aligned with that of *CaNGA1* and *CaBZR1* in both NILs, further supporting the model where *CaNGA1* modulates BR signaling to regulate the expression of downstream seed size regulators. This co-expression strongly implicates a *CaNGA1*–BR–seed development regulatory axis in governing seed size variation in chickpea.

To further validate the role of *CaNGA1*-derived epialleles in regulating seed size, we examined the expression patterns of *CaNGA1*, key BR-responsive genes (*CaCPD1*, *CaBRI1*, *CaBZR1*), and known seed

size-responsive genes (*CaSHB1*, *CaIKU1*, *CaMINI3*, *CaIKU2*, *CaKLU*) in both untreated (control) and 5-aza-dC-treated NILs (Fig. 2F). In the LSW-NILs, treatment with 5-aza-dC led to a distinct co-expression shift, with upregulated expression of *CaNGA1* followed by a transcriptional upsurge of BR pathway and seed size-related genes, indicating a cascade effect triggered by the removal of promoter methylation. In contrast, no significant expression changes were detected in the HSW-NILs upon treatment, likely due to their naturally hypomethylated *CaNGA1* promoter, which limits further response to the demethylating agent. These findings underscore that *CaNGA1* methylation status directly influences its transcriptional activity, which in turn modulates downstream gene expression through the BR signalling pathway, ultimately affecting seed size.

To further reinforce these results, we explored the BR-mediated regulation of seed size by treating HSW- and LSW-NILs with 24-epibrassinolide (BR) and Brassinazole (BRZ), an inhibitor of BR biosynthesis. BR treatment led to a significant increase in seed size/weight in LSW-NILs, while the effect in HSW-NILs was minimal, suggesting that LSW-NILs are more responsive to exogenous BR, possibly due to lower endogenous BR levels. Conversely, BRZ treatment caused a marked reduction in seed size/weight in both NILs, confirming the critical role of BR in seed development (Fig. 2G). Corroborating these phenotypic observations, genes involved in BR biosynthesis (*CaCPD1*), BR perception (*CaBR11*), BR-responsive transcription (*CaBZR1*), and seed size regulation (*CaSHB1*, *CaMINI3*, *CaIKU1*, *CaIKU2*, *CaKLU*) exhibited upregulated expression in BR-treated NILs, with greater fold-changes in LSW-NILs compared to HSW-NILs (Fig. 2H). This suggests that BR supplementation compensates for the epigenetically silenced *CaNGA1* in LSW-NILs, whereas HSW-NILs, with optimal endogenous BR synthesis and active *CaNGA1*, show limited additional response. As expected, BRZ-treated samples displayed downregulation of the same gene set, reinforcing the role of BR in regulating seed size pathways. Together, these data offer a compelling demonstration of epiallelic regulation, where natural methylation variation at the *CaNGA1* promoter governs its expression and downstream BR-mediated signaling, ultimately determining seed size and weight in chickpea (Fig. 3).

Discussion

Natural epigenetic variation has emerged as an important contributor to phenotypic diversity and adaptation in plants, yet only a limited number of naturally occurring epialleles have been causally linked to complex agronomic traits in crop species (Richards, 2006; Hofmeister et al., 2017; Kawakatsu et al., 2016). In the present study, we demonstrate that a natural methylation epiallele (NAME), rather than nucleotide sequence polymorphism, underlies a major seed-weight QTL in chickpea. Fine mapping localized the *CaqSW3.1* locus to a 5.92 kb interval containing a single candidate gene, *CaNGA1*, while targeted resequencing revealed complete sequence conservation between contrasting NILs. Instead, differential CG methylation within the *CaNGA1* promoter was associated with contrasting transcript abundance and seed-weight phenotypes, and chemical demethylation restored *CaNGA1* expression in the low-seed-weight genotype. These findings provide strong evidence that naturally occurring promoter methylation acts as the functional regulatory variation governing *CaNGA1* expression and highlight the

importance of epigenetic diversity as an additional layer of heritable variation for quantitative trait regulation.

Our findings are consistent with previous methylome studies in chickpea showing that differentially methylated regions (DMRs) associated with differentially expressed genes (DEGs) are predominantly enriched in the CG sequence context (Rajkumar et al., 2020). Such CG methylation is generally associated with stable transcriptional regulation and can generate heritable epialleles without altering the underlying DNA sequence (Lister and Ecker, 2009; Hofmeister et al., 2017). The differential CG methylation detected in the *CaNGA1* promoter between contrasting NILs, together with the increased transcript abundance following 5-aza-2'-deoxycytidine treatment, indicates that promoter methylation directly modulates *CaNGA1* expression. Similar regulatory mechanisms have been reported for naturally occurring epialleles controlling important developmental traits in other plant species, including flowering time and fruit ripening (Manning et al., 2006; Miura et al., 2010; Quadrana et al., 2014). Our study extends these observations by demonstrating that stable methylation epialleles can also regulate a major yield-related trait in chickpea.

Another important finding is the developmental regulation of *CaNGA1* during seed formation. *CaNGA1* expression increased progressively during seed development and reached its highest level during the DS3 stage, coinciding with the onset of rapid dry matter accumulation and seed biomass expansion around 21 days after pollination (Pradhan et al., 2014). This temporal expression pattern suggests that *CaNGA1* functions during the critical transition from embryogenesis to seed filling, when cell proliferation, cell expansion and storage reserve accumulation collectively determine final seed size. Previous genome-wide methylome and transcriptome analyses similarly demonstrated that the DS3 stage contains the highest proportion of DMR-associated DEGs involved in cell cycle regulation, cell expansion and grain filling, many of which are localized within major seed-size QTLs (Rajkumar et al., 2020). The identification of *CaNGA1* as a DMR-associated DEG located within a major seed-weight QTL therefore provides an important mechanistic link between developmental DNA methylation dynamics and quantitative variation in seed weight.

The predicted function of *CaNGA1* further supports its regulatory role during seed development. *CaNGA1* belongs to the highly conserved NGATHA (NGA) subfamily of B3-domain transcription factors, members of which regulate diverse developmental processes through sequence-specific DNA binding (Zhang et al., 2015; Sasnauskas et al., 2018; Guo et al., 2020; Salava et al., 2022). Although NGA proteins have been implicated in organ growth and developmental patterning in several plant species, their contribution to seed development in legumes has remained largely unexplored. The developmental expression pattern observed here, together with the localization of *CaNGA1* within the major **CaqSW3.1** QTL, identifies this transcription factor as a previously uncharacterized regulator of seed development in chickpea.

Our results further establish a mechanistic connection between *CaNGA1* and brassinosteroid (BR) signalling. Brassinosteroids are central regulators of plant growth that control cell proliferation, cell expansion and reproductive development through the BRI1 receptor kinase and the downstream

transcription factors BZR1/BES1 (Je et al., 2010). In chickpea, global transcriptome analyses across different seed developmental stages of contrasting seed-size accessions have revealed that BR biosynthesis and metabolism-related genes are significantly enriched in gene ontology (GO) terms and exhibit higher activity at the S3 stage, where they promote cell expansion by regulating DNA replication and endoreduplication levels (Garg et al., 2017). Furthermore, higher transcript abundance of BR signaling genes in the large-seeded cultivar JGK 3 compared to the small-seeded Himchana 1 supports the critical role of BR signaling in determining seed size and weight in chickpea. Building upon these observations, Rajkumar et al. (2022) analyzed DMR-associated DEGs between these contrasting lines and found significant enrichment for BR stimulus and cell growth-related processes, indicating that epiallelic marks may influence seed size through BR-mediated regulation, a phenomenon particularly prominent at the S3 stage of the high-seed-weight genotype. Building upon these observations, our findings identify *CaNGA1* as a potential upstream regulator of BR signalling during chickpea seed development. Consistent with this hypothesis, we identified multiple *CaNGA1*-binding motifs within the promoters of BR-related genes and demonstrated that *CaNGA1* transcriptionally activates *CaBZR1* using a dual-luciferase assay.

Brassinosteroid signalling has previously been shown to regulate seed size by controlling the expression of genes involved in integument development, endosperm proliferation and embryo growth (Jiang et al., 2013). In *Arabidopsis*, the BR-responsive transcription factor BZR1 directly regulates several key seed-size genes, including *SHB1*, *IKU1*, *MINI3*, *IKU2* and *KLU*, thereby promoting seed growth (Jiang et al., 2013; Zhu et al., 2015). Consistent with this regulatory framework, all corresponding chickpea homologues displayed coordinated upregulation in the high-seed-weight NILs and following demethylation or exogenous BR treatment, whereas inhibition of BR biosynthesis produced the opposite response. Together with the observed changes in seed weight, these findings support a model in which promoter hypomethylation enhances *CaNGA1* expression, leading to activation of BR signalling and downstream seed-development genes that collectively promote seed growth.

In summary, this study identifies *CaNGA1* as a naturally regulated epiallelic locus underlying seed-size variation in chickpea and provides mechanistic evidence linking promoter methylation with BR-mediated developmental regulation. These findings broaden our understanding of how stable epigenetic variation contributes to complex agronomic traits and highlight naturally occurring methylation epialleles as an additional source of selectable variation for genomics-assisted and epigenetics-assisted crop improvement.

Declarations

Conflict of Interest

The authors declare that the study was carried out without any financial or commercial ties that could be constructed as potential conflict of interest.

Author Contribution

JKM performed experiments, performed data analysis and drafted the manuscript. AY assisted in experiments and data analysis. VT helped with data analysis. ST helped with field phenotypic evaluation. PA and SKP conceived the idea, and guided, participated in drafting, and correcting the manuscript critically and all authors gave the final approval of the version to be published.

Acknowledgement

We are thankful to central instrumentation facility (CIF) and DBT-eLibrary Consortium (DeLCON) of NIPGR, New Delhi for providing timely support and access to E-resources for this research work. Financial support for this study was provided by a research grant from the Department of Biotechnology (DBT), Government of India (BT/HRD/NBA-NWB/39/2020-21). JKM acknowledges the DBT, India for Research Fellowship Award.

Data Availability

The whole-genome sequencing files for both parental genotypes submitted to NCBI with BioProject Accession Number PRJNA985750, where VGS1 code for (ICC 4958; accession SRR24988769) and LGS1 code for (ICC 5590; accession SRR24988771). All relevant information related to BS-PCR and Amplicon sequences can be found with the NCBI GenBank accession (PX842563-PX842580)

References

1. Akalin, A., Kormaksson, M., Li, S., Garrett-Bakelman, F. E., Figueroa, M. E., Melnick, A., & Mason, C. E. (2012). methylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles. *Genome Biology*, 13, 1–9.
2. Akimoto, K., Katakami, H., Kim, H. J., Ogawa, E., Sano, C. M., Wada, Y., & Sano, H. (2007). Epigenetic inheritance in rice plants. *Annals of botany*, 100(2), 205–217.
3. Asami, T., Mizutani, M., Fujioka, S., Goda, H., Min, Y. K., Shimada, Y., ... Yoshida, S. (2001). Selective interaction of triazole derivatives with DWF4, a cytochrome P450 monooxygenase of the brassinosteroid biosynthetic pathway, correlates with brassinosteroid deficiency in planta. *Journal of Biological Chemistry*, 276(28), 25687–25691.
4. Baduel, P., & Colot, V. (2021). The epiallelic potential of transposable elements and its evolutionary significance in plants. *Philosophical Transactions of the Royal Society B*, 376(1826), 20200123.
5. Bajaj, D., Saxena, M. S., Kujur, A., Das, S., Badoni, S., Tripathi, S., ... Parida, S. K. (2015a). Genome-wide conserved non-coding microsatellite (CNMS) marker-based integrative genetical genomics for quantitative dissection of seed weight in chickpea. *Journal of Experimental Botany*, 66(5), 1271–1290.

6. Bajaj, D., Upadhyaya, H. D., Khan, Y., Das, S., Badoni, S., Shree, T., ... Parida, S. K. (2015b). A combinatorial approach of comprehensive QTL-based comparative genome mapping and transcript profiling identified a seed weight-regulating candidate gene in chickpea. *Scientific Reports*, 5(1), 9264.
7. Bondada, R., Somasundaram, S., Marimuthu, M. P., Badarudeen, M. A., Puthiyaveedu, V. K., & Maruthachalam, R. (2020). Natural epialleles of Arabidopsis SUPERMAN display superwoman phenotypes. *Communications Biology*, 3(1), 772.
8. Bosssdorf, O., Arcuri, D., Richards, C. L., & Pigliucci, M. (2010). Experimental alteration of DNA methylation affects the phenotypic plasticity of ecologically relevant traits in *Arabidopsis thaliana*. *Evolutionary Ecology*, 24, 541–553.
9. Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., & Cresko, W. A. (2013). Stacks: an analysis tool set for population genomics. *Molecular Ecology*, 22(11), 3124–3140.
10. Chang, S., & Pikaard, C. S. (2005). Transcript Profiling in Arabidopsis Reveals Complex Responses to Global Inhibition of DNA Methylation and Histone Deacetylation box. *Journal of Biological Chemistry*, 280(1), 796–804.
11. Das, S., Upadhyaya, H. D., Srivastava, R., Bajaj, D., Gowda, C. L. L., Sharma, S., ... Parida, S. K. (2015). Genome-wide insertion–deletion (InDel) marker discovery and genotyping for genomics-assisted breeding applications in chickpea. *DNA Research*, 22(5), 377–386.
12. Daware, A., Mohanty, J. K., Narnoliya, L., Singh, A., Rathore, D., Thakro, V., ... Parida, S. K. (2024). Uncovering DNA methylation landscapes to decipher evolutionary footprints of phenotypic diversity in chickpea. *DNA Research*, 31(3).
13. Ding, C. J., Liang, L. X., Diao, S., Su, X. H., & Zhang, B. Y. (2018). Genome-wide analysis of day/night DNA methylation differences in *Populus nigra*. *PLoS One*, 13(1), e0190299.
14. Eichten, S. R., Briskine, R., Song, J., Li, Q., Swanson-Wagner, R., Hermanson, P. J., ... Springer, N. M. (2013). Epigenetic and genetic influences on DNA methylation variation in maize populations. *The Plant Cell*, 25(8), 2783–2797.
15. Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., & Mitchell, S. E. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PloS One*, 6(5), e19379.
16. Fieldes, M. A. (1994). Heritable effects of 5-azacytidine treatments on the growth and development of flax (*Linum usitatissimum*) genotrophs and genotypes. *Genome*, 37(1), 1–11.
17. Gahlaut, V., Zinta, G., Jaiswal, V., & Kumar, S. (2020). Quantitative epigenetics: a new avenue for crop improvement. *Epigenomes*, 4(4), 25.
18. Garg, R., Sahoo, A., Tyagi, A. K., & Jain, M. (2010). Validation of internal control genes for quantitative gene expression studies in chickpea (*Cicer arietinum* L.). *Biochemical and Biophysical Research Communications*, 396(2), 283–288.
19. Garg, R., Singh, V. K., Rajkumar, M. S., Kumar, V., & Jain, M. (2017). Global transcriptome and coexpression network analyses reveal cultivar-specific molecular signatures associated with seed

- development and seed size/weight determination in chickpea. *The Plant journal* 91(6), 1088–1107.
20. Gehring, M. (2019). Epigenetic dynamics during flowering plant reproduction: evidence for reprogramming? *New Phytologist*, 224(1), 91–96.
21. Greaves, I. K., Groszmann, M., Ying, H., Taylor, J. M., Peacock, W. J., & Dennis, E. S. (2012). Trans chromosomal methylation in *Arabidopsis* hybrids. *Proceedings of the National Academy of Sciences*, 109(9), 3570–3575.
22. Guo, T., Wang, S., Li, Y., Yuan, J., Xu, L., Zhang, T., ... Han, L. (2020). Expression of a *ngatha1* gene from *Medicago truncatula* delays flowering time and enhances stress tolerance. *International Journal of Molecular Sciences*, 21(7), 2384.
23. Jabre, I., Chaudhary, S., Wilson, C. M., Staiger, D., & Syed, N. (2022). Stochastic variation in DNA methylation modulates nucleosome occupancy and alternative splicing in *Arabidopsis thaliana*. *Plants*, 11(9), 1105.
24. Jacobsen, S. E., & Meyerowitz, E. M. (1997). Hypermethylated SUPERMAN epigenetic alleles in *Arabidopsis*. *Science*, 277(5329), 1100–1103.
25. Je, B. I., Piao, H. L., Park, S. J., Park, S. H., Kim, C. M., Xuan, Y. H., ... Han, C. D. (2010). RAV-Like1 maintains brassinosteroid homeostasis via the coordinated activation of BRI1 and biosynthetic genes in rice. *The Plant Cell*, 22(6), 1777–1791.
26. Jiang, W. B., Huang, H. Y., Hu, Y. W., Zhu, S. W., Wang, Z. Y., & Lin, W. H. (2013). Brassinosteroid regulates seed size and shape in *Arabidopsis*. *Plant Physiology*, 162(4), 1965–1977.
27. Jiang, W. B., Huang, H. Y., Hu, Y. W., Zhu, S. W., Wang, Z. Y., & Lin, W. H. (2013). Brassinosteroid regulates seed size and shape in *Arabidopsis*. *Plant Physiology*, 162(4), 1965–1977.
28. Katoh, Y., Maekawa, M., & Sano, Y. (1993). Effects of 5-azacytidine on somatic mutation in a soybean test system. *Mutation Research/Genetic Toxicology*, 300(1), 49–55.
29. Kawakatsu, T., Huang, S. S. C., Jupe, F., Sasaki, E., Schmitz, R. J., Urich, M. A., ... Schork, N. J. (2016). Epigenomic diversity in a global collection of *Arabidopsis thaliana* accessions. *Cell*, 166(2), 492–505.
30. Krueger, F., & Andrews, S. R. (2011). Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics*, 27(11), 1571–1572.
31. Kujur, A., Bajaj, D., Upadhyaya, H. D., Das, S., Ranjan, R., Shree, T., ... Parida, S. K. (2015a). A genome-wide SNP scan accelerates trait-regulatory genomic loci identification in chickpea. *Scientific Reports*, 5(1), 11166.
32. Kujur, A., Bajaj, D., Upadhyaya, H. D., Das, S., Ranjan, R., Shree, T., ... Parida, S. K. (2015c). Employing genome-wide SNP discovery and genotyping strategy to extrapolate the natural allelic diversity and domestication patterns in chickpea. *Frontiers in Plant Science*, 6, 162.
33. Kujur, A., Upadhyaya, H. D., Shree, T., Bajaj, D., Das, S., Saxena, M. S., ... Parida, S. K. (2015b). Ultra-high density intra-specific genetic linkage maps accelerate identification of functionally relevant molecular tags governing important agronomic traits in chickpea. *Scientific Reports*, 5(1), 9468.

34. Kumar, K., Purayannur, S., Kaladhar, V. C., Parida, S. K., & Verma, P. K. (2018). mQTL-seq and classical mapping implicates the role of an AT-HOOK MOTIF CONTAINING NUCLEAR LOCALIZED (AHL) family gene in A scochyta blight resistance of chickpea. *Plant, cell & environment*, 41(9), 2128–2140.
35. Kumar, S., Stecher, G., & Tamura, K. (2016). MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution*, 33(7), 1870–1874.
36. Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, 9(4), 357–359.
37. Li, L. C., & Dahiya, R. (2002). MethPrimer: designing primers for methylation PCRs. *Bioinformatics*, 18(11), 1427–1431.
38. Lister, R., & Ecker, J. R. (2009). Finding the fifth base: genome-wide sequencing of cytosine methylation. *Genome Research*, 19(6), 959–966.
39. Malik, N., Dwivedi, N., Singh, A. K., Parida, S. K., Agarwal, P., Thakur, J. K., & Tyagi, A. K. (2016). An integrated genomic strategy delineates candidate mediator genes regulating grain size and weight in rice. *Scientific Reports*, 6(1), 23253.
40. Manning, K., Tör, M., Poole, M., Hong, Y., Thompson, A. J., King, G. J., ... Seymour, G. B. (2006). A naturally occurring epigenetic mutation in a gene encoding an SBP-box transcription factor inhibits tomato fruit ripening. *Nature Genetics*, 38(8), 948–952.
41. Manning, K., Tör, M., Poole, M., Hong, Y., Thompson, A. J., King, G. J., ... Seymour, G. B. (2006). A naturally occurring epigenetic mutation in a gene encoding an SBP-box transcription factor inhibits tomato fruit ripening. *Nature Genetics*, 38(8), 948–952.
42. Martin, A., Troadec, C., Boualem, A., Rajab, M., Fernandez, R., Morin, H., ... Bendahmane, A. (2009). A transposon-induced epigenetic change leads to sex determination in melon. *Nature*, 461(7267), 1135–1138.
43. Miura, K., Ikeda, M., Matsubara, A., Song, X. J., Ito, M., Asano, K., ... Ashikari, M. (2010). OsSPL14 promotes panicle branching and higher grain productivity in rice. *Nature Genetics*, 42(6), 545–549.
44. Mukhopadhyay, P., & Tyagi, A. K. (2015). OsTCP19 influences developmental and abiotic stress signaling by modulating ABI4-mediated pathways. *Scientific Reports*, 5(1), 9998.
45. Mukhopadhyay, P., & Tyagi, A. K. (2015). OsTCP19 influences developmental and abiotic stress signaling by modulating ABI4-mediated pathways. *Scientific Reports*, 5(1), 9998.
46. Ong-Abdullah, M., Ordway, J. M., Jiang, N., Ooi, S. E., Kok, S. Y., Sarpan, N., ... Martienssen, R. A. (2015). Loss of Karma transposon methylation underlies the mantled somaclonal variant of oil palm. *Nature*, 525(7570), 533–537.
47. Ooijen, V. (2009). MapQTL6, Software for the mapping of quantitative trait loci in experimental population of diploid species. *Wageningen, Netherlands, Kyazma BV*.
48. Patel, R. K., & Jain, M. (2012). NGS QC Toolkit: a toolkit for quality control of next generation sequencing data. *PloS One*, 7(2), e30619.

49. Pradhan, S., Bandhiwal, N., Shah, N., Kant, C., Gaur, R., & Bhatia, S. (2014). Global transcriptome analysis of developing chickpea (*Cicer arietinum* L.) seeds. *Frontiers in Plant Science*, 5, 698.
50. Quadrana, L., Almeida, J., Asís, R., Duffy, T., Dominguez, P. G., Bermúdez, L., ... Carrari, F. (2014). Natural occurring epialleles determine vitamin E accumulation in tomato fruits. *Nature Communications*, 5(1), 4027.
51. Quadrana, L., Almeida, J., Asís, R., Duffy, T., Dominguez, P. G., Bermúdez, L., ... Carrari, F. (2014). Natural occurring epialleles determine vitamin E accumulation in tomato fruits. *Nature Communications*, 5(1), 4027.
52. Rajkumar, M. S., Gupta, K., Khemka, N. K., Garg, R., & Jain, M. (2020). DNA methylation reprogramming during seed development and its functional relevance in seed size/weight determination in chickpea. *Communications Biology*, 3(1), 340.
53. Richards, E. J. (2006). Inherited epigenetic variation—revisiting soft inheritance. *Nature Reviews Genetics*, 7(5), 395–401.
54. Salava, H., Thula, S., Sánchez, A. S., Nodzyński, T., & Maghuly, F. (2022). Genome wide identification and annotation of NGATHA transcription factor family in crop plants. *International Journal of Molecular Sciences*, 23(13), 7063.
55. Sano, H., Kamada, I., Youssefian, S., Katsumi, M., & Wabiko, H. (1990). A single treatment of rice seedlings with 5-azacytidine induces heritable dwarfism and undermethylation of genomic DNA. *Molecular and General Genetics MGG*, 220, 441–447.
56. Sasnauskas, G., Manakova, E., Lapėnas, K., Kauneckaitė, K., & Siksnyš, V. (2018). DNA recognition by Arabidopsis transcription factors ABI 3 and NGA 1. *The FEBS Journal*, 285(21), 4041–4059.
57. Sato, H., Takasaki, H., Takahashi, F., Suzuki, T., Iuchi, S., Mitsuda, N., ... Shinozaki, K. (2018). Arabidopsis thaliana NGATHA1 transcription factor induces ABA biosynthesis by activating NCED3 gene during dehydration stress. *Proceedings of the National Academy of Sciences*, 115(47), E11178-E11187.
58. Saxena, M. S., Bajaj, D., Das, S., Kujur, A., Kumar, V., Singh, M., ... Parida, S. K. (2014a). An integrated genomic approach for rapid delineation of candidate genes regulating agro-morphological traits in chickpea. *DNA Research*, 21(6), 695–710.
59. Saxena, M. S., Bajaj, D., Kujur, A., Das, S., Badoni, S., Kumar, V., ... Parida, S. K. (2014b). Natural allelic diversity, genetic structure and linkage disequilibrium pattern in wild chickpea. *PLoS One*, 9(9), e107484.
60. Setia, N., and Malik, C. P. (1985). Distribution of mineral nutrients in developing fruits of chickpea (*Cicer arietinum*). *Phyton (Austria)* 25, 117–121
61. Shen Y., Zhang J., Liu Y., Liu S., Liu Z., Duan Z., et al. (2018). DNA methylation footprints during soybean domestication and improvement. *Genome Biol.* 19, 128. 10.1186/s13059-018-1516-z
62. Slatkin, M. (2009). Epigenetic inheritance and the missing heritability problem. *Genetics*, 182(3), 845–850.

63. Song, C. X., Szulwach, K. E., Dai, Q., Fu, Y., Mao, S. Q., Lin, L., ... He, C. (2013). Genome-wide profiling of 5-formylcytosine reveals its roles in epigenetic priming. *Cell*, 153(3), 678–691.
64. Spindel, J., Wright, M., Chen, C., Cobb, J., Gage, J., Harrington, S., ... McCouch, S. (2013). Bridging the genotyping gap: using genotyping by sequencing (GBS) to add high-density SNP markers and new value to traditional bi-parental mapping and breeding populations. *Theoretical and Applied Genetics*, 126, 2699–2716.
65. Upadhyaya, H. D., Bajaj, D., Das, S., Saxena, M. S., Badoni, S., Kumar, V., ... Parida, S. K. (2015). A genome-scale integrated approach aids in genetic dissection of complex flowering time trait in chickpea. *Plant Molecular Biology*, 89, 403–420.
66. Varshney, R. K., Song, C., Saxena, R. K., Azam, S., Yu, S., Sharpe, A. G., ... Cook, D. R. (2013). Draft genome sequence of chickpea (*Cicer arietinum*) provides a resource for trait improvement. *Nature Biotechnology*, 31(3), 240–246.
67. Zhang, Y., Du, L., Xu, R., Cui, R., Hao, J., Sun, C., & Li, Y. (2015). Transcription factors SOD7/NGAL2 and DPA4/NGAL3 act redundantly to regulate seed size by directly repressing KLU expression in *Arabidopsis thaliana*. *The Plant Cell*, 27(3), 620–632.
68. Zhang, X., Sun, J., Cao, X., & Song, X. (2015). Epigenetic mutation of RAV6 affects leaf angle and seed size in rice. *Plant Physiology*, 169(3), 2118–2128.
69. Zhu, X., Liang, W., Cui, X., Chen, M., Yin, C., Luo, Z., ... Zhang, D. (2015). Brassinosteroids promote development of rice pollen grains and seeds by triggering expression of Carbon Starved Anther, a MYB domain protein. *The Plant Journal*, 82(4), 570–581.

Figures

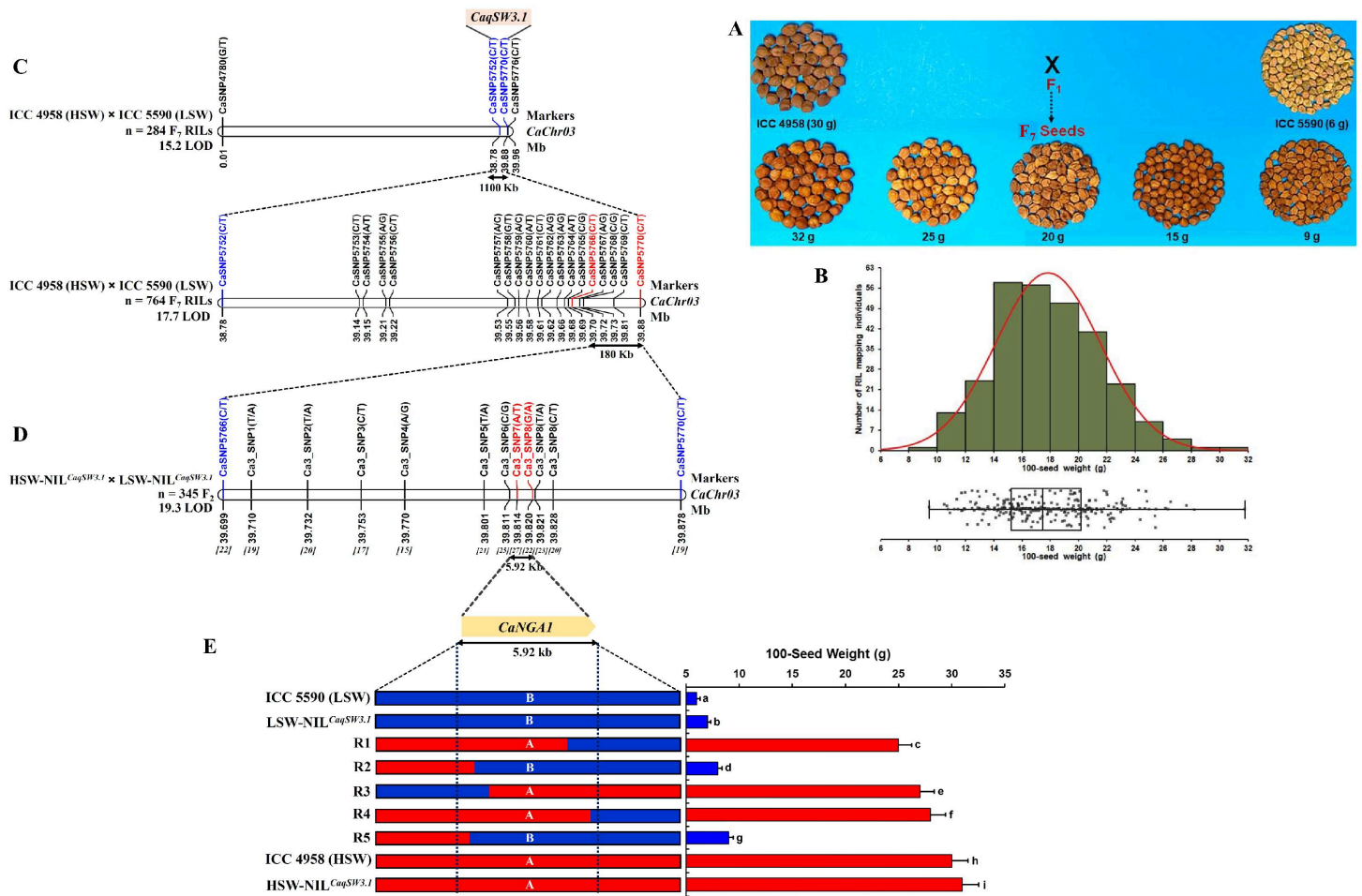


Figure 1

Figure 1

High-resolution molecular mapping, fine-mapping, and map-based cloning delineated a causal gene *CaNGA1* of a major QTL genomic region controlling 100-seed weight and seed size in chickpea. (A) Generation of a chickpea recombinant inbred line (RIL) mapping population for 100 seed weight (100-SW) trait. **(B)** A broad phenotypic variation observed among 287 F₇ RIL mapping individuals for 100-SW. Frequency distribution and box-plots represent the phenotypic variations for 100-SW where the edges of the boxes represent the upper and lower quantiles and the median values depict in the middle of the box. The black circles constituting the box-plots denote the phenotypic variations of 287 individual RILs for the 100-SW. **(C)** Molecular mapping of a major QTL *CaqSW3.1* on a genetic linkage map (ICC 4958 × ICC 5590) of chromosome 3 in chickpea. **(D)** Subsequently, map-based cloning and fine-mapping the *CaqSW3.1* with parental polymorphic SNPs in a 764 RILs followed by fine-mapping in a NIL mapping population (HSW-NIL^{*CaqSW3.1*} × LSW-NIL^{*CaqSW3.1*}; 345 individuals) scaled-down the QTL interval from 1100 to 5.92 kb genomic region with a single protein coding gene *CaNGA1*. **(E)** Progeny testing by integrating the multi-environmental replicated phenotyping of the homozygous recombinant and non-recombinant

progenies derived from mapping population (HSW-NIL^{CaqSW3.1} × LSW-NIL^{CaqSW3.1}) along with their mapping parental lines and contrasting NILs for 100 SW. At this QTL interval, *Ca3_SNP7*(A/T) and *Ca3_SNP8*(G/A) are tightly linked to *CaNGA1* and demonstrate zero recombination with the target loci in the five recombinant NILs. The genomic constitution of HSW-NIL and LSW-NIL was represented with ‘A’ and ‘B’ respectively, whereas the horizontal lines indicate the average phenotypic constitution for the 100 SW (n = 5 to 8).

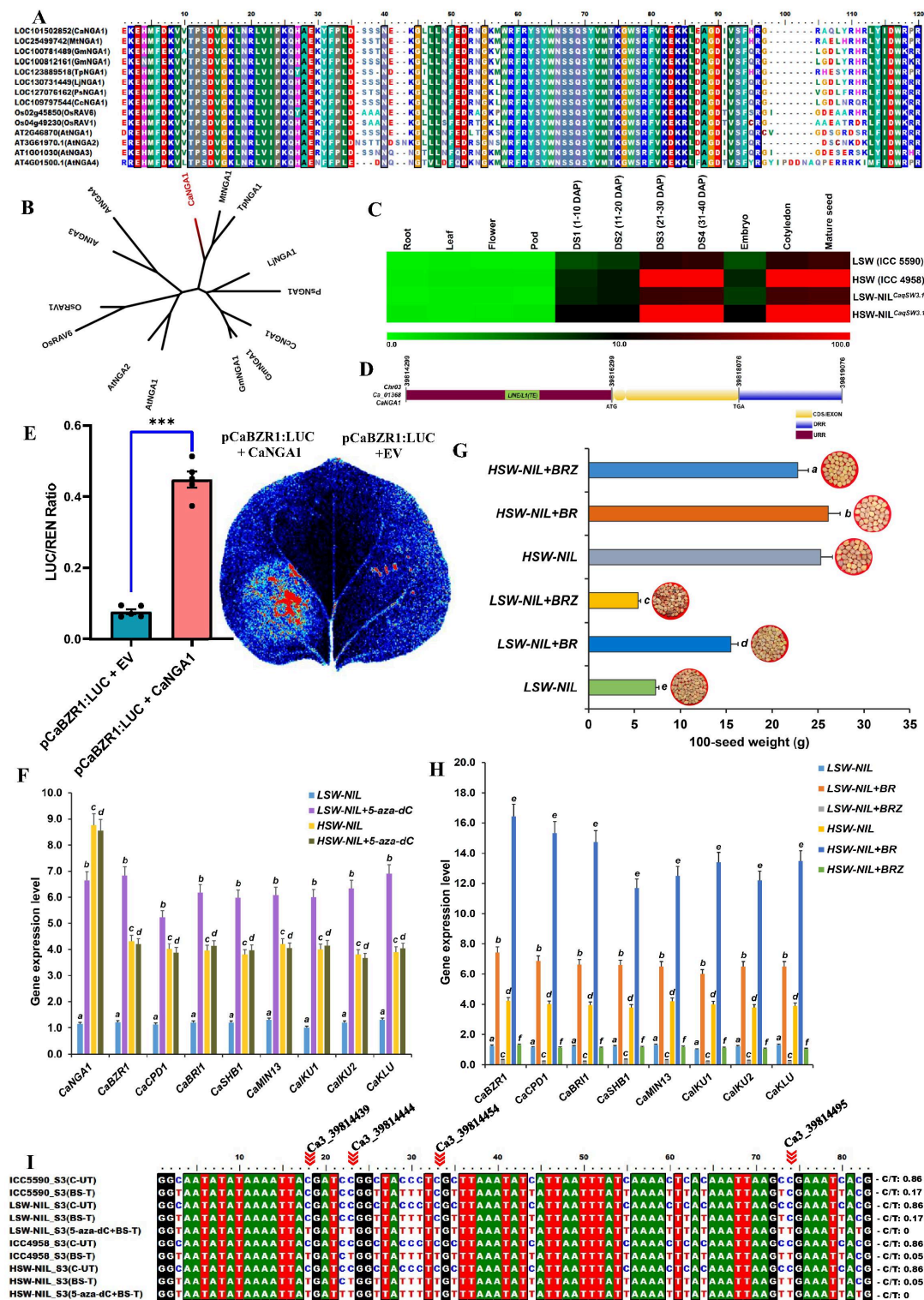


Figure 2

Figure 2

The epialleles of *CaNGA1* regulates chickpea seed size and seed weight by modulating the expression of brassinosteroid- and seed size-responsive genes. (A) An unrooted phylogenetic tree based on protein sequence diversity, illustrates the evolutionary relationships of chickpea *CaNGA1* with other legume species and model crop plants. The sequence IDs begin with two-letter abbreviations for the plant species: *Ca* (*Cicer arietinum*), *Mt* (*Medicago truncatula*), *Cc* (*Cajanus cajan*), *Gm* (*Glycine max*), *At* (*Arabidopsis thaliana*), *Os* (*Oryza sativa*), *Tp* (*Trifolium pratense*), *Lj* (*Lotus japonicus*) and *Ps* (*Pisum sativum*). (B) Represents multiple sequence alignments of amino acids from the functional B3 domain of *CaNGA1* with that of several related legumes and model crops. (C) Depicts the gene expression across different tissue types and developmental stages. The heat-map shows the expression pattern of *CaNGA1* in various vegetative and reproductive tissues, including four stages of seed development (DS1–DS4), in contrasting parental accessions and QTL-introgressed near-isogenic lines (NILs) (HSW-NIL^{CaqSW3.1} and LSW-NIL^{CaqSW3.1}). Expression values were obtained using qRT-PCR ($\Delta\Delta C_t$ method) and subsequently min–max normalized to a 0–100 scale for visualization. (D) Illustrate the gene structure of *CaNGA1*, including its exonic sequence as well as upstream and downstream regulatory regions. A transposable element in the upstream regulatory (promoter) region is highlighted in light green, marking the position where differential methylation observed. (E) Relative Firefly luciferase (LUC) to Renilla luciferase (REN) activity in *Nicotiana benthamiana* leaves co-infiltrated with the reporter construct (ProCaBZR1:LUC) and the indicated effector constructs (empty vector or 2×35S:CaNGA1). Renilla luciferase driven by the 35S promoter served as an internal normalization control. Data are presented as mean $\pm$ SE from five biological replicates. Statistical significance was determined using Student's *t*-test ($***p \leq 0.001$). The right side of the image is a representative CCD images showing luminescence signals from *N. benthamiana* leaves used in the dual-luciferase reporter assay. (F) Demonstrates the fold expression of *CaNGA1*, BR-responsive genes (*CaBZR1*, *CaCPD1*, *CaBRI1*), and seed size-responsive genes (*CaSHB1*, *CaMINI3*, *CaIKU1*, *CaIKU2*, *CaKLU*) in contrasting NILs following treatment with a methylation inhibitor (5-aza-dC), in comparison to their respective untreated counterparts. The expression demonstrates here is the cumulative mean of 3 biological and 3 technical replicates where *ACTIN* was used as an internal control for normalization of expression data. (G) Represents the 100-seed weight of contrasting NILs after treatment with exogenous BR (24-epibrassinolide) and BRZ (brassinazole), with untreated NILs serving as the control group. The 100-seed weight was measured using fully matured seeds with a moisture content of 10%, and the experiment included 5 biological and technical replicates for accuracy. (H) Shows the fold expression of BR-responsive and seed size-responsive genes in contrasting NILs upon exposure to exogenous BR and BRZ, compared to their respective untreated counterparts. The expression data presented is the cumulative mean of three biological and three technical replicates, with *ACTIN* used as an internal control for normalization. (I) BS-PCR followed by amplicon sequencing of control and 5-aza-dC treated NILs highlighting the methylation status of the *CaNGA1* promoter sequence. The 4 red arrows indicate the functional epialleles in the CG context whose differential methylation determines the contrasting seed weight trait

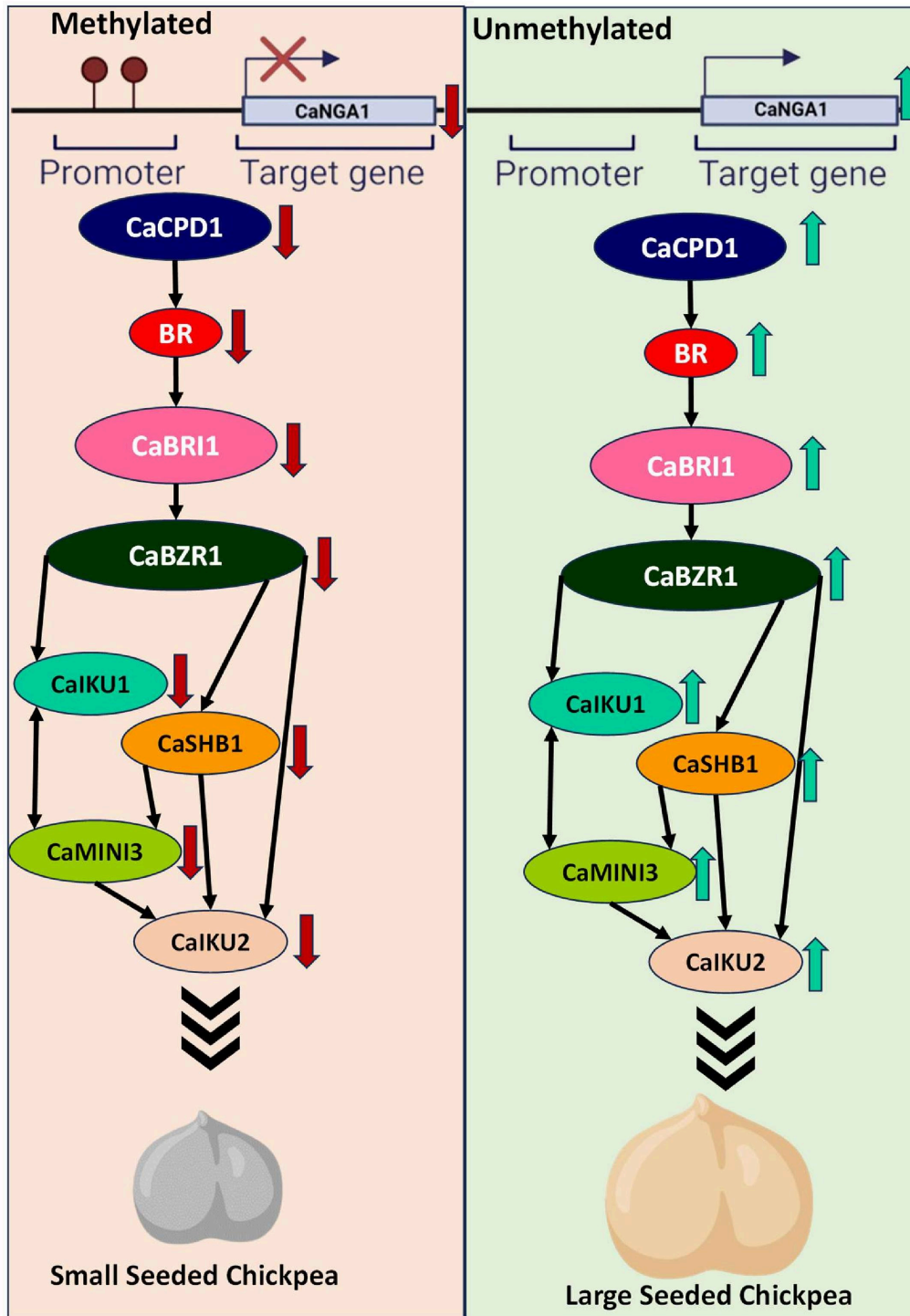


Figure 3

Figure 3

A schematic representation of epiallelic mediated transactivation of *CaNGA1* igniting the BR signaling cascade for a better seed size in chickpea. The diagram illustrates the differing transactivation of *CaNGA1* between high-seed weight (SW) (ICC 4958 and HSW-NIL^{CaqSW3.1}) and low-SW (ICC 5590 and LSW-NIL^{CaqSW3.1}) accessions/NILs. In the low-SW lines, such as ICC 5590 and LSW-NIL^{CaqSW3.1}, the hypermethylated promoter of *CaNGA1* limits its transactivation by various transcription factors, which

subsequently reduces the expression of several genes involved in BR biosynthesis (*CaCPD1*), BR receptors (*CaBRI1*), and BR signaling (*CaBZR1*). This diminished BR signaling negatively impacts the expression of downstream seed size-responsive genes (*CaIKU1*, *CaMINI3*, *CaIKU2*, *CaSHB1*) known to be regulated by BR signaling, ultimately impairing the signaling cascade responsible for seed size and resulting in smaller seeds. In contrast, the high-SW lines (ICC 4958 and HSW-NIL^{CaqSW3.1}) exhibit higher expression levels of *CaNGA1* due to a hypomethylated promoter, which effectively stimulates BR signaling. This induced BR signaling enhances the transcriptional abundance of seed size-responsive genes, ultimately leading to larger seed size and increased seed weight.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [MohantyetalTableS2010426.pdf](#)
- [MohantyetalTableS1010426.xlsx](#)