Uncovering DNA methylation landscapes to decipher evolutionary footprints of phenotypic diversity in chickpea

Anurag Daware[1*], Jitendra K. Mohanty[1*], Laxmi Narnoliya[1], Akansha Singh[1], Deepanshi Rathore[1], Virevol Thakro[1], Aleena Francis[1], Nagendra Pratap Singh[1], Philip Francis[1], Shailesh Tripathi[2], Debasis Chattopadhyay[1], Swarup K. Parida[1]

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

[2]Indian Institute of Pulses Research (IIPR), Uttar Pradesh, Kanpur 208024, India

Corresponding author

Dr. Swarup K. Parida

National Institute of Plant Genome Research (NIPGR), Aruna Asaf Ali Marg, New Delhi - 110067, India.

E-mail: swarup@nipgr.ac.in; swarupdbt@gmail.com

Phone: 91-11-26735228; Fax: 91-11-26741658

*Equal Contributions

Abstract

Genetic diversity and environmental factors are long believed to be the dominant contributor to phenotypic diversity in crop plants. However, it has been recently established that, besides genetic variation, epigenetic variation, especially variation in DNA methylation, plays a significant role in determining phenotypic diversity in crop plants. Therefore, assessing DNA methylation diversity in crop plants becomes vital, especially in the case of crops like chickpea, which has a narrow genetic base. Thus, in the present study, we employed whole-genome bisulfite sequencing to assess DNA methylation diversity in wild and cultivated (*desi* and *kabuli*) chickpea. This revealed extensive DNA methylation diversity in both wild and cultivated chickpea. Interestingly, the methylation diversity was found to be significantly higher than genetic diversity, suggesting its potential role in providing vital phenotypic diversity for the evolution and domestication of the *Cicer* gene pool. The phylogeny based on DNA methylation variation also indicates a potential complementary role of DNA methylation variation in addition to DNA sequence variation in shaping chickpea evolution. Besides, the study also identified diverse epi-alleles of many previously known genes of agronomic importance. The *Cicer* MethVarMap database developed in this study enables researchers to readily visualize methylation variation within the genes and genomic regions of their interest (http://223.31.159.7/cicer/public/). Therefore, epigenetic variation like DNA methylation variation can potentially explain the paradox of high phenotypic diversity despite the narrow genetic base in chickpea and can potentially be employed for crop improvement.

Key Words: DNA methylation, Chickpea, *Cicer*, Epi-alleles, SNPs

1. Introduction

Since the dawn of agriculture in the Neolithic Era, man has constantly sought ways to improve agriculture production. Over the years, studies focusing on genetic variation have contributed immensely to our understanding of evolution, domestication and phenotypic diversity in crop plants. Genetic variation (DNA sequence variation) is known to be the primary facilitator behind the evolution of phenotypic traits in crop plants. The tremendous success of diverse genetic methods which target DNA sequence variation for varietal improvement is testimony to the same[1,2]. However, recent studies suggest that in addition to genetic variation, epigenetic variation also plays a vital role in determining heritable trait diversity in the plants[3,4]. Therefore, epigenetic diversity could be one possible cause underlying the long-debated phenomenon of missing heritability in complex traits[5,6].

The presence/absence of cytosine methylation is one of the well-known forms epigenetic variation in the plants[6]. Further, cytosine methylation has been extensively studied, and molecular mechanisms involved in this epigenetic modification have also been deciphered in plants[6,7]. The cytosine methylation (mC) is primarily present in three distinct CG, CHG and CHH sequence contexts, where H = C, A, or T. In *Arabidopsis*, DNA methylation in CG context is maintained by DNA METHYLTRANSFERASE 1 (MET1). In contrast, methylation in the CHG context is maintained CHROMOMETHYLASE 3 (CMT3). The CHH methylation is maintained by RNA directed DNA methylation pathway, a mechanism unique to plants[6,8]. The DNA methylation in each context has diverse effects on gene expression patterns, transposable element (TE) activity, and chromatin conformation[9].

With the availability of high-quality reference genomes and the rise of bisulfite sequencing, it is now possible to produce genome-wide single-base resolution DNA methylation maps of plant genomes. Genome-wide methylation maps of many important plant genomes, including *Arabidopsis*, rice, maize, soybean etc., have been decoded so far[10-12]. These studies suggest the crucial role of DNA methylation in regulating various critical biological processes in plants. Further, comparison among methylomes of diverse natural accessions belonging to the same plant species unravelled the presence of rich methylation diversity (epi-allelic diversity) within plants species. Further, the evidence regarding the involvement of this methylation (epigenetic) diversity in evolution and domestication in plants have also been documented[13,14]. In addition to this, many

stable heritable epi-alleles regulating important agronomic traits such as flower structure, sex, vitamin E content, fruit ripening, flowering time, fruit number, plant architecture, and root architecture etc., have been discovered using a variety of approaches[3,15-19]. Among these approaches, epi-genome-wide association study (EWAS) has emerged as a promising approach to identify natural epi-alleles regulating important traits in crop plants[4]. With the recent advancement in genome editing technologies, it is also possible to specifically methylate or demethylate functionally relevant DNA sequences within the plant genome to create desirable epigenetic diversity, ultimately leading to desirable phenotypes[20]. Considering the aforementioned scenario, uncovering methylation diversity existing within natural accessions of crops species will provide an additional level of diversity to develop superior cultivars, especially for crops with a narrow genetic base.

Chickpea is one of the essential food legumes and is primarily divided into the *kabuli* (white flowers with light-coloured seed coat) and *desi* (pink flowers with dark-coloured seed coat) types. Chickpea accessions, especially *desi* type chickpea, harbour huge phenotypic diversity for various agronomically important traits, including seed size, seed/pod number, seed colour, growth habit, tolerance to biotic and abiotic stresses etc. However, despite substantial phenotypic diversity, especially between and within *desi* and *kabuli* chickpeas, the genetic diversity seems to be very limited compared to other crops[21,22]. The cause behind this paradox is still not known. Epigenetic factors are often speculated to be responsible for this paradox. Recently whole-genome bisulfite sequencing studies in many legumes, including chickpea, suggested the crucial role of DNA methylation in modulating gene expression and, therefore, may also regulate critical biological processes in the legumes[12,23,24]. Therefore, it would be interesting to explore the untapped epigenetic diversity within chickpea accessions and its contribution to phenotypic diversity and domestication.

In the present study, to get insights into levels of natural methylation diversity in the *Cicer* gene-pool, we performed whole-genome sequencing and whole-genome bisulfite sequencing of five diverse chickpea accessions, including two *kabuli*, two *desi* and one wild accession. The study revealed the presence of extensive methylation diversity in chickpea, predominantly in CG and CHG contexts. The genetic and methylation diversity levels between *desi* and *kabuli* chickpea were compared to assess the potential role of DNA methylation diversity in differentiating these two types of cultivated chickpea. The study also identified diverse epi-alleles (differential

methylation states) of many essential genes regulating important traits of agronomic importance in chickpea, which can be investigated for their functional effects on various a/biotic stress and yield-related traits in this important legume crop. Finally, a web resource (*Cicer* MethVarMap) was developed to efficiently retrieve information regarding the methylation status of different genes at a genome-wide scale across wild and cultivated *Cicer* accessions for their possible futuristic use in accelerated genetic improvement of chickpea.

2. Materials and methods

2.1 Plant materials and DNA extraction

Global methylation patterns in the young leaves (30 days old plant leaves) of four cultivated chickpea (*Cicer arietinum* L.), including two of each *kabuli* (ICC 8155 and ICC 8261) and *desi* (ICC 4958 and ICC 5590) accessions along with a wild chickpea (*C. reticulatum*) accession (ICC 17160) were analyzed in this study. For DNA isolation, all five chickpea accessions were grown under field conditions, and leaf tissue from each accession was harvested in three independent biological replicates. Genomic DNA was extracted using the DNeasy Plant Mini kit (Qiagen, USA). The purity and quantity of isolated genomic DNA were then checked using a Nanodrop Spectrophotometer (Thermo Fisher Scientific, USA) and Qubit Fluorometer (Life Technologies, USA), respectively. Further agarose gel electrophoresis was used to ensure the intactness of high molecular weight genomic DNA.

2.2 Annotation of repetitive elements in chickpea genome

To identify and annotate transposable elements in chickpea, transposable element (TE) library was primarily created for *kabuli* (CDC Frontier) reference genome v1.0 (ASM33114V1, https://www.ncbi.nlm.nih.gov/assembly/GCF_000331145.1/)[25] using RepeatModler 2[26]. RepeatMasker was further used to annotate *kabuli* genome using the generated TE library.

2.3 Whole-genome sequencing, variant calling and annotation

Five whole-genome sequencing (WGS) DNA libraries were prepared as per the manufacturer's protocol (Illumina, San Diego, CA, USA). The WGS DNA libraries were sequenced using Illumina Hiseq 2000 sequencing platform to generate 100 bp pair-end sequence reads. The quality of raw sequence reads was further assessed using FastQC. Further adapters and low-quality bases from the raw sequence reads were trimmed using the NGS QC tool kit v2.3[27]. High-quality

sequence reads were aligned and mapped to *kabuli* (CDC Frontier) reference chickpea genome v1.0 (ASM33114V1, https://www.ncbi.nlm.nih.gov/assembly/GCF_000331145.1/)[25] using BWA program v0.7.12[28]. Subsequently, Picard tools were used for format conversion and removal of duplicate reads. Genome Analysis Toolkit (GATK, v3.6.0) HaplotypeCaller[29] was used to perform variant calling, variant recalibration and variant filtration. The high-quality variants were then annotated using the snpEff program v4.3p[30].

2.4 Bisulfite sequencing, read alignment and detection of 5-methylcytosines

Bisulfite sequencing of each DNA sample was performed as described earlier[31] using Illumina HiSeq 2000 platform. The quality of raw sequence reads was first analyzed using FastQC. Further, the first nine bases, adapters and low-quality bases in raw reads were trimmed using Trim Galore[32]. Bisamrk v0.18.1[33] along with Bowtie2 was then used to align trimmed high-quality sequence reads to the pre-converted *kabuli* reference genome (V1.0, ASM33114V1) with default parameters. In addition, the first eight bases from 5' and the first four bases from 3' ends of the high-quality reads were trimmed to eliminate methylation bias from the end of the sequence reads introduced during library preparation. Further deduplication of alignments was performed to eliminate PCR artefacts, and only uniquely mapped reads were considered for further analysis. To determine the efficiency of bisulfite conversion, the chickpea chloroplast genome was used as a negative control since the chloroplast genome is known to be highly unmethylated. The alignment of sequence reads against chloroplast genome enabled us to determine the bisulfite conversion efficiency. The processed alignment files were then provided to methylKit (v0.9.2)[34] for further analysis. The methylation status of each cytosine with minimum coverage of ≥ 5 was determined by applying the binomial test (P-value ≤ 0.0001) as described earlier[35]. Subsequently, all mCs overlapping with C/T transitions identified from the WGS information were eliminated from further analysis to address the confounding caused by such transitions during the identification of authentic mCs.

2.5 Identification of DMRs

DMRs were identified by comparing all five accessions among each other. DMRs were categorized into "within cultivars" and "between cultivars" based on their differential methylation at least between any two accessions belonging to the same cultivar-groups and two different cultivar-groups of chickpea, respectively. The DMRs within cultivated chickpea accessions were further classified into 'within *desi*', 'within *kabuli*' and 'within wild' based on cultivar-groups

within which they are differentially methylated. DMRs between cultivars was subsequently categorised into 'wild vs *desi*, 'wild vs *kabuli*', 'wild vs cultivated', and '*desi* vs *kabuli*' based on their differential methylation at least between any two accessions belonging to two different cultivar-groups.

For identification of DMRs in all three contexts (CG, CHG and CHH), tilling window approach was employed using methyl it with a window size of 100 and a step size of 100. During this analysis, only bases with >5X sequence coverage were considered. Fisher's exact test along with Sliding Linear Model (SLIM) correction was used to identify statistically significant DMRs (q-value ≤ 0.01). A DMR is called when the bin contains a minimum of three cytosines with cumulative methylation difference $\geq$20% for CG/CHG and $\geq$25% for CHH. Customized Perl script was used for annotation of identified DMRs.

2.6 Comparison of genome-wide genetic and DNA methylation variation

Sliding window analysis (with a window size of 0.1 Mb and step size of 10 Kb) was performed to estimate genome-wide raw densities of genetic variants (SNPs/InDels) and methylation variants (DMCs). The basic densities for all eight chickpea chromosomes were further visualized as a line plot using the ggplot2 R package. However, as DNA methylation was surveyed only on cytosine bases, we further compared normalized density (normalized density = total count/total number of callable bases) for DNA methylation and sequence variants throughout the chickpea genome. Similar to raw densities, normalized densities of both genetic and methylation variants were visualized as a line plot using the ggplot2 R package.

2.7 Real-time quantitative PCR (RT-qPCR) assay of differentially methylated genes

To further investigate the impact of differential methylation on their corresponding gene expression, expression profiling of selected differentially methylated genes was performed in *kabuli* (ICC 8155 and ICC 8261), *desi* (ICC 4958 and ICC 5590) and wild (ICC 17160) *Cicer* accessions. Plants from said *Cicer* accessions were grown in the Phytotron growth chamber facility under control environmental condition at NIPGR, New Delhi. One-microgram of high-quality total RNA isolated from the 21-days old seedlings of these accessions were used for the synthesis of cDNA using Verso cDNA synthesis kit (Thermo Fisher Scientific, USA). The diluted cDNA of each accession (with 3 independent biological replicate and 3 technical replicates for each biological replicate) was used to amplify the selected differentially methylated genes using

respective gene-specific primers in CFX96 Realtime system (Bio-Rad, USA) following earlier methods[36]. The significant differential expression each gene was estimated by the $2^{-\Delta\Delta CT}$ method using *actin* as an internal control[36].

2.8 Development of a user-friendly database for methylation and sequence variation

This database was created with Laravel, which is an open-source PHP framework. This was integrated further with ggplot2 to generate methylation plots. The database was hosted on a Linux server. To integrate the Genome Browser interface to the Database, the positional methylation data generated from the Bismark methylation extractor was used to generate a bedGraph file using bismark2bedGraph application in the Bismark toolkit[33]. Subsequently, compressed binary indexed Bigwig files for respective accession representing each of 3 different methylation contexts were generated using bedGraphToBigWig command line utility tool[37]. Further, this BigWig file was used to visualize the genome-wide positional methylation status among five chickpea accessions in the JBrowse[38].

3. Results and discussion

3.1 Whole-genome sequencing of chickpea accessions reveals genetic variation in cultivated and wild Cicer gene pool

To understand the genome-wide genetic variation in chickpea, whole-genome sequencing of five chickpea accessions belonging to cultivated (*C. arietinum*) *desi* (ICC 4958 and ICC 5590) and *kabuli* (ICC 8155 and ICC 8261) as well as a wild *C. reticulatum* accession (ICC 17160) was performed. Whole-genome sequencing generated more than 250 million paired-end sequence reads for each sequenced *Cicer* accession. Further, more than 95% of these sequence reads were passed quality filtering and mapped uniquely to the chickpea genome. The high-quality uniquely mapped sequence reads provided over 85% of the chickpea genome (**Supplementary Table S1**). Further, using high-quality uniquely mapped reads of five accessions, variant calling identified more than 3.5 million high-quality sequence variants, including >3.4 million SNPs and >8 lakh InDels. Based on polymorphism patterns (SNPs and InDels) observed among chickpea accessions belonging to different/similar cultivated and wild cultivar groups, the detected SNPs were further categorized into six different classes, namely 'wild vs *desi* polymorphic', 'wild vs *kabuli* polymorphic', 'wild vs cultivated polymorphic', '*desi* vs *kabuli* polymorphic', 'within *desi* polymorphic' and 'within *kabuli* polymorphic'. Polymorphism between the wild vs *desi* cultivar

group (95.5%) was highest, followed by the wild vs. *kabuli* group (88.7%) (**Supplementary Table S2**). Little over 9.8% of total sequence variants detected polymorphic between *desi* vs. *kabuli* cultivar group. Even a lesser number of sequence variants were polymorphic within *desi* (2.74%) and within *kabuli* (0.75%), emphasizing the existence of a narrow genetic base in the cultivated *Cicer* gene pool as documented by many previous studies[39,40] (**Supplementary Table S2**). The genetic polymorphisms detected among accessions were well-distributed across eight chromosomes and unanchored scaffolds of chickpea genome varying from 8327.1 (chromosome 7) to 13993.1 (chromosome 4) sequence variants per Mb with an average density of 8826.2 variants per Mb. Further, structural annotation of identified sequence variants revealed more than 4-times abundance of these variants in the non-genic intergenic sequence components compared to genic components of the chickpea genome (**Supplementary Table S3**). This distribution pattern of sequence variants agrees well with their functional consequences and is consistent with previous studies in chickpea[39,40] .

3.2 Whole-genome bisulfite sequencing of chickpea accessions unravels the natural variation of DNA methylation in cultivated and wild Cicer gene pool

To unravel the DNA methylome variation in chickpea, whole-genome bisulfite sequencing of the five chickpea accessions belonging to *desi* (ICC4958 and ICC5590) and *kabuli* (ICC8155 and ICC8261) and wild (ICC17160) *Cicer* was carried out. Bisulfite sequencing produced more than 45.8 million sequence reads for each sequenced accession. Further, trimmed sequence reads showing no significant methylation bias were aligned to the pre-converted reference, which provided >85% coverage of the chickpea genome (**Supplementary Table S1**). The alignment files were then used to determine the methylation status of genome-wide cytosines. Subsequently, the bisulfite conversion rate was calculated based on the conversion efficiency of the chickpea chloroplast genome. All the five chickpea accessions exhibited very high conversion efficiency with a conversion rate greater than 99% **(Supplementary Table S4).**

Further, the proportion of methylated cytosines in all three contexts was analyzed. All five accessions showed similar methylation patterns, with CHH methylation (74-75%) being the highest, followed by CG methylation (12%) and CHG methylation (13-14%) (**Figure 1A**). The mean methylation level of mCs also followed a similar trend in all five *desi*, *kabuli* and wild chickpea accessions. The mCs in the CG context (78-83%) displayed the highest mean methylation

level, followed by mCs in the CHG context (65-72%). Whereas the mCs in CHH context showed a very low level of methylation (7-12%) compared to the other two contexts (**Figure 1B**). These patterns of relative DNA methylation and mean DNA methylation levels in three different contexts are comparable to other crop plants, including cereals and legumes[12,41]. Further, we compared genome-wide DNA methylation levels (independently for all three contexts) throughout the chickpea genome among five cultivated and wild chickpea accessions. There were no significant differences in CG and CHG methylation levels among five chickpea accessions. However, *kabuli* accessions showed higher CHH methylation than wild and *desi* chickpea accessions (**Figure 1B**).

3.3 Distribution of DNA methylation in diverse sequence components of the chickpea genome

To better understand the distribution of DNA methylation within the chickpea genome, we determined the methylation levels within the genic region [includes gene body along with 2 kb upstream and downstream regulatory regions] as well as in the transposable elements in all five accessions belonging to *desi*, *kabuli* and wild cultivar groups (**Figure 2**). The CG methylation levels were predominantly higher within gene-body compared to flanking upstream (URRs) and downstream (DRRs) regulatory regions. However, in contrast to CG, both CHG and CHH methylation was found significantly higher in flanking URRs and DRRs than in the gene-body (**Figure 2A**). Irrespective of the context, a sudden drop in methylation was observed near the transcriptional start site (TSS) and transcriptional termination sites (TTS). The CG methylation was considerably higher within genic regions than the other two contexts. Further, when compared among different chickpea accessions, no noticeable difference was observed for CG methylation in the genic regions, whereas a *kabuli* accession ICC 8155 showed only slightly higher CHG methylation than other *desi* and *kabuli* accessions. However, ICC 8155 was observed to have significantly higher CHH methylation within genic regions than all other accessions (**Figure 2A**). This overall trend of methylation in the genic region is comparable to previous studies in other crop plants[12,24,31,41]. The methylation of URRs is previously known to be associated with the repression of gene expression in plants, whereas gene body methylation is reported to be responsible for accurate transcription and splicing, protecting genes from TE insertion[42-45] etc. Thus, it will be interesting to investigate whether previously explained methylation has a similar effect in chickpea.

Transposable elements (TEs) are among the most abundant sequence elements in plant genomes and are essential drivers of genome evolution. In addition to acting as simple mutagens, TEs also alter elementary functions of the plant genome by modulating expression patterns of genes[46]. TEs and other unclassified repeats constitute approximately half (49.41%) of the 740 Mb chickpea genome, similar to other sequenced legumes[25]. Considering the crucial role of TEs in genome evolution and regulation of gene expression, we analysed methylation levels within transposable elements in the chickpea genome. Among all the chickpea accessions analyzed, the genome sequence of *kabuli* (CDC frontier) chickpea is of better quality and thus used for identification and classification of repeat elements as TEs. This TE annotation information was used further as a reference to determine the methylation status (CG, CHG and CHH) of TE-body, along with 2 kb of each URR and DRRs of protein-coding genes. The TE-body, along with flanking sequences, showed considerably higher overall methylation levels compared to the other protein-coding genes. Further, like genic regions, TE methylation in CG and CHG contexts was found to be several folds higher than that observed in the CHH context (**Figure 2B**). However, unlike typical protein-coding genes, the TSS and TTS of TEs were found to be highly methylated. These results are in accordance to previous studies in legumes and thus suggest the crucial role of DNA methylation in suppressing the transposon activity in chickpea[41] . However, it's crucial to acknowledge that the mobile nature of TEs might cause discrepancies in inferring the precise methylation status of all TEs in different accessions, especially due to rely of TE annotation in this study on the single reference genome. This highlights the limitation of the bisulfite sequencing approach for comparing the methylation status of TEs across different chickpea accessions (especially belonging to different species) in the absence of high-quality reference genome sequences in each of the accession.

3.4 Identification and distribution of DMRs in desi, kabuli and wild chickpea

Natural methylation diversity is now established as an important factor underlying vital trait variation present in natural crop germplasm accessions[4,18] (Quadrana et al., 2014; Ong-Abdullah et al., 2015). Therefore, to identify methylation variation among *desi*, *kabuli* and wild chickpea accessions, we performed differential methylation analysis by comparing methylomes of each two *desi* (ICC 4958 and ICC 5590) and *kabuli* (ICC 8155 and ICC 8261) and one wild (*Cicer reticulatum*) accession (ICC 17160). This enabled us to identify genome-wide differentially methylated cytosines (DMCs) among five said diverse chickpea accessions. However, considering

the greater biological relevance of differentially methylated regions (DMRs) over DMCs, we further identified genome-wide DMRs independently for all three contexts (CG, CHG and CHH) by calculating average methylation over non-overlapping 100 bp window as per earlier strategy[41]. Primarily, DMRs among all five accessions were identified by one-to-one comparisons (**Supplementary Table 5**). The identified DMRs were further categorized into "within cultivar groups" (within *desi* and within *kabuli*) and "between cultivar-groups" (wild vs *desi*, wild vs *kabuli*, wild vs cultivated and *desi vs kabuli*) based on their differential methylation between accessions belonging to the same cultivar-groups and two different cultivar-groups of chickpea, respectively (**Supplementary Table 5**). The highest number of DMRs were found between wild and *kabuli* (29637), followed by wild vs *desi* (28384) and wild vs cultivated (25184) chickpea. The least number of DMRs (13006) were found between *desi* and *kabuli* chickpea (**Supplementary Table 5**). This trend of DNA methylation diversity seems similar to genetic diversity observed in this study as well as those reported in previous studies in chickpea and thus aligns well with a closer evolutionary relationship of wild chickpea (*C. reticulatum*) with *desi* chickpea as compared to *kabuli* chickpea (Kujur et al., 2015; Varshney et al., 2019). Further, the identified DMRs predominantly belonged to CG (50-63%) and CHG (30-43%) contexts, whereas very few of these DMRs belong to the CHH context (5-6%). Therefore, CG and CHG methylation variation seems to account for the majority of the methylation variation existing in natural wild and cultivated chickpea accessions. This again is in accordance with the fact that most of the methylation in plants is known to exist in CG and CHG contexts[31,41,47] (Zhong et al., 2009; Garg et al., 2015; Junaid et al., 2018).

3.5 Potential role of DNA methylation variation in evolution and domestication of chickpea

To study the functional significance of DNA methylation identified variation in chickpea the identified DMRs were categorized as hypomethylated and hypermethylated DMRs based on their overall methylation level difference between different cultivated and wild chickpea cultivar groups **(Figure 3A, Supplementary Table S6)**. This revealed most CG (64-70 %) and CHG (64-76 %) DMRs show hypermethylation in cultivated (*desi* as well as *kabuli*) chickpea cultivar groups compared to wild chickpea. However, in contrast, most CHH (60-84%) DMRs showed hypomethylation in cultivated (*desi* as well as *kabuli*) chickpea compared to wild chickpea (**Figure 3B**). Previously, a similar trend had been reported in soybean where cultivated soybean (*Glycine max*) has accumulated a higher level of methylation compared to wild soybean (*G. soja*) during

domestication[47]. Therefore, we specifically focused on DMRs displaying hypermethylation in cultivated chickpea compared to wild. Many of these DMRs were found to coincide with some of the vital genes known to regulate important domestication traits in chickpea. For instance, *PHOTOPERIOD-INDEPENDENT EARLY FLOWERING 1* (*PIE1*), a gene that is known to suppressor of flowering under noninductive photoperiods[48] was found to be significantly (*p* and *q* value < *0.001*) hypermethylated in cultivated chickpea accessions compared to wild chickpea **(Supplementary Table S7)**. The wild chickpea (*Cicer reticulatum*) is well known for its photoperiod-sensitive nature, whereas flowering in most cultivated chickpea accessions is not dependent on photoperiod[49]. Interestingly, the role of DNA methylation in regulating photoperiod sensitivity has been previously reported in the case of cotton, where differential methylation of *CONSTANS LIKE* (COL) genes, *COL2A* and *COL2D* between wild and cultivated cotton accessions was shown to regulate photoperiod sensitivity[50]. Therefore, it would be interesting to further investigate the likely role of DNA methylation in chickpea domestication by modulating photoperiod sensitivity in chickpea.

As the methylation variants associated with genes can provide meaningful information on the functional impact of DNA methylation variation. Therefore, we further focused on DMRs associated with coding and upstream regulatory regions (URRs) and downstream regulatory regions (DRRs) of chickpea genes to discover differentially methylated genes among five diverse chickpea accessions. A total of 10200 genes (15758 DMRs) were found to be differentially methylated between wild and *desi* chickpea. Among these, most genes (~ 61 %) were found to have the differentially methylated gene-body and a relatively smaller number of genes were found to have differentially methylated URRs (~19 %) and DRRs (~19.5 %). Similarly, a total of 9813 differentially methylated genes (14925 DMRs) were detected between wild and *kabuli* chickpea. Again, most of these genes (~58.5 %) were found to have differentially methylated gene-body whereas comparatively fewer genes were found to have methylation variation in URRs (~20 %) and DRRs (~20.8 %). Finally, 6709 differentially methylated genes (9459 DMRs) were identified between *desi* and *kabuli* accessions. Similar to the previous two comparisons, most of these genes have differentially methylated gene-body (33-52 %) and comparatively fewer genes have differentially methylated URRs (24-43 %) and DRRs (9-10 %) (**Figure 4**). These results suggest that most of the DNA methylation variation exists within the coding region, contrary to genetic variation (SNP/InDel) which exists predominantly in URRs and DRRs. Interestingly, the genic

region displays significant methylation variation even between *desi* and *kabuli* chickpea accessions. This is particularly important as very limited genetic variation exists in the genic region, between *desi* and *kabuli* chickpea. Thus, the DNA methylation variation provides much-needed molecular diversity, which might explain the broad phenotypic diversity present, despite very little genetic diversity in cultivated (*desi* and *kabuli*) chickpea.

GO enrichment analysis was further performed to understand the biological processes associated with these differentially methylated genes. The set of genes showing differential methylation between wild and cultivated chickpea was significantly enriched for genes related to plant vegetative and reproductive development, DNA damage/repair and regulation of gene expression. (**Supplementary Figure S1**). This is by drastic differences in vegetative (plant architecture, biomass etc.) and reproductive development (flowering time, photoperiod/vernalization sensitivity, seed size etc.) between cultivated and wild chickpea. Similarly, the genes displaying differential methylation between *desi* and *kabuli* chickpea show over-representation for genes predominantly involved in response to different biotic and abiotic stresses (**Supplementary Figure S2**). This correlates well with the fact that *desi* chickpea is more tolerant to various biotic and abiotic stresses than *kabuli*[51,52]. Thus, methylation variation seems to play a potent role in modulating critical physiological/developmental processes in wild and cultivated chickpea.

To further dissect the role of methylation variation, we assessed the methylation status of previously identified genes with known functions. This revealed differential methylation in many vital genes that regulate important agronomic traits in crop plants. This includes many circadian rhythms (flowering time) regulating genes like *Cryptochrome 1*, which was found to have a differentially methylated promoter (URR) when compared between wild and cultivated chickpea (**Supplementary Table S7, Supplementary Figure S3**). Similarly, a gene encoding *METAL TOLERANCE PROTEIN 10*, which regulates seed size/weight, exhibited a highly methylated promoter (URR) in cultivated chickpea compared to wild chickpea (**Supplementary Table 7, Supplementary Figure S3**). Considering the importance of flowering time and seed weight traits in chickpea domestication, the role of methylation in *Cicer* evolution and domestication cannot be over-ruled.

Further, many genes associated with nodulation, a process specific to legumes also found to have differences in methylation levels. The *Nodulation Signaling Pathway protein 1* (*NSP1*) and *NSP2* are essential regulators of nodule development in legumes. *NSP1* was found to have a highly methylated promoter in wild and *desi* accessions compared to *kabuli* chickpea (**Figure 5A**). In addition to this, *CAROTENOID CLEAVAGE DIOXYGENASE 7*, which regulates various developmental processes, including nodulation in legumes[53] was found to be differentially methylated between *desi* and *kabuli* chickpea cultivar groups in both CG and CHG context.

Interestingly, *kabuli* accessions are known to fix significantly higher nitrogen than *desi* in soil rich in macronutrients. At the same time, the reverse is true in the case of poor-quality marginal soils, suggesting a difference in nitrogen fixation efficiency between *desi* and *kabuli*[54]. Apart from this, many genes associated with nitrogen assimilation, like *glutamate synthase* and *beta carbonic anhydrase 5*, have significantly different methylation levels in their promoter (URR), indicating an important role of methylation variation in regulating nitrogen fixation biological processes unique to legumes (**Supplementary Table S7**).

To evaluate the impact of differential methylation of the promoter regions of genes on their transcript expression levels, differential expression profiles of four selected differentially methylated genes were generated in the aforesaid *kabuli* (ICC 8155 and ICC 8261), *desi* (ICC 4958 and ICC 5590) and wild (ICC 17160) *Cicer* accessions **(Figure 5B, Supplementary Table S8)**. The relative gene expression levels of *NSP1* and *NSP2* were found to be higher in *kabuli* accessions compared to the *desi* and wild accessions, which may be due to the higher methylation status in the promoter regions of these genes in the corresponding *desi* and wild *Cicer* accessions. Similarly, lower gene expression of *CRY1* in the cultivated *desi* and *kabuli* accessions corresponds to its higher methylation status in the cultivated compared to the wild *Cicer* **(Figure 5B)**. However, the observed higher gene expression of *PIE1* in *kabuli* accessions in contrast to *desi* and wild accessions was unable to reveal a converse methylation pattern as witnessed in these *Cicer* accessions (**Supplementary Table S7**).

3.6 Determination of chickpea phylogeny using a genetic and epigenetic variation

To trace the contribution of genetic as well as epigenetic (DNA methylation) variation to the evolution and domestication of cultivated chickpea from its wild ancestor, phylogenetic analysis of all five accessions belonging to cultivated (*desi* and *kabuli*) as well as wild chickpea were

primarily performed using the genetic variants (SNPs and InDels) to generate a neighbour-joining (NJ)-based phylogenetic tree. Similarly, phylogenetic analysis of five chickpea accessions was also performed using the genome-wide epigenetic variants (DMCs) (**Figure 6**). Both genetic and epigenetic phylogenies broadly capture known evolutionary relationships within chickpea[2]. This suggests that DNA methylation variants, like DNA sequence variants, can be used to accurately infer phylogeny in chickpea. Furthermore, recent studies have indicated that DNA methylation patterns can evolve more rapidly than genetic mutations, offering a higher temporal resolution for phylogenetic explorations[55] . Similarly, in the current study, significant differences were found in the topologies of phylogenetic trees based on genetic variation and cytosine methylation variation. Specifically, the phylogenetic tree based on methylation variation establishes a closer evolutionary relationship between wild (*Cicer reticulatum*) and *desi* chickpea, which is not as evident in the tree based on DNA sequence variation. These results also lend support to the long-standing theory about the potential evolution of *kabuli* chickpea from desi, despite the lack of support from genetic analysis[2] . Although, due to the limited sample size, it is difficult to establish a definitive role for DNA methylation in chickpea evolution and domestication, the current study provides vital preliminary evidence for future studies aiming to investigate the potential influence of DNA methylation in chickpea evolution.

3.7 Comparison of genome-wide DNA sequence and DNA methylation variation between desi, kabuli and wild chickpea

Despite a narrow genetic base (low genetic variability), cultivated chickpea shows huge phenotypic diversity for many agronomic traits. To study the potential role of methylation diversity underlying this paradox, we compared genome-wide sequence (SNPs/InDels) and DNA methylation variation (DMCs) existing within two *desi* and two *kabuli* chickpea accessions. As DNA methylation exists only on cytosine bases, we further compared normalized density (normalized density = total count/total number of callable bases) for DNA methylation and DNA sequence variants throughout the chickpea genome. Interestingly, the normalized density for DNA methylation variants was much higher than that observed for DNA sequence variants throughout the chickpea genome, with only a few notable exceptions **(Figure 7)**. This highlights a higher DNA methylation variation rate than the genetic variation between *desi* and *kabuli* chickpea.

Further, few genomic regions were found to lack both sequences and DNA methylation variation. GO enrichment analysis of genes underlying these low-diversity genomic regions suggests these genomic regions are predominantly associated with phosphorylation, protein modification and protein transport/localization (**Supplementary Figure S1, S2)**. One of these genes, *Chalcone synthase,* encodes for a critical enzyme in the flavonoid/isoflavonoid biosynthesis pathway, known to regulate seed/flower colour in addition to biotic stresses in crop plants[56,57] .

3.8 Development of a user-friendly methylation variation database of chickpea

To make methylation variation information across cultivated *desi* and *kabuli* and wild chickpea accessions publicly accessible in a user-friendly manner, we developed the *Cicer* MethVarMap database (http://223.31.159.7/cicer/public/). This database provides researchers with a unique opportunity to search for natural DNA methylation variants (DMCs/DMRs) and sequence variants (SNPs/InDels) within genes or genomic regions of their interest. Users can query sequence-derived genetic (SNPs/InDels) and methylation variants using chickpea Gene IDs or genomic coordinates. In addition, methylation levels of chickpea genes across wild and cultivated chickpea accessions in all three contexts (CG, CHG, and CHH) can also be visualized with a few easy clicks. This database will be handy for delineating candidate genes underlying methylation QTLs detected in GWAS and bi-parental QTL mapping studies. Subsequently, a Genome Browser was integrated to the Database to visualize the genome-wide methylation level in the five said chickpea accessions across all three methyl contexts (CG, CHG and CHH). Three different color bars in varied genomic positions of the respective chickpea accessions were visualized in the Genome Browser, which represents the intensity of methylation in the 3 different methyl contexts.

3.9 Conclusion

In the present study, we conducted whole-genome bisulfite sequencing of diverse chickpea accessions to assess the DNA methylation diversity in wild (*Cicer reticulatum*) and cultivated chickpea. The study revealed extensive DNA methylation distributed within the chickpea genome. Like other legumes, most of this methylation exists in the CG context, followed by CHG and CHH context in chickpea. The comparison among diverse wild and cultivated chickpea accessions revealed the presence of extensive methylation variation in both genic and intergenic components of the chickpea genome. As expected, many DMRs were identified between wild and cultivated chickpeas, whereas a comparatively smaller number of DMRs were found between *desi* and *kabuli*

chickpea. Interestingly, wild chickpea displayed fewer methylation differences (DMRs) with *desi* compared to *kabuli*. This closer relationship between *desi* and wild chickpea is also reflected in the phylogenetic tree constituted using DMCs. This is especially important because previous studies on genetic variation failed to establish this evolutionary relationship with sufficient confidence.

The DNA methylation seems to provide an additional level of molecular diversity within chickpea. This can be the possible cause for the wide phenotypic diversity observed in cultivated chickpea accessions with extremely narrow genetic base. The differential methylation analysis between wild and cultivated chickpea also uncovered many important differentially methylated genes, suggesting a possible role of DNA methylation in regulating key chickpea domestication-related traits, including flowering duration and seed weight. In addition, methylation differences in some important genes related to abiotic stress, flowering duration and nutritional quality were also detected between *desi* and *kabuli* chickpea accessions. Thus, the study generated an extensive resource on methylation variation (epi-alleles) across wild and cultivated chickpea, which could be utilized for chickpea improvement in future. This resource could further be used to select epi-alleles for large-scale validation and establish their role in regulating important traits of agronomic importance in chickpea. Finally, this information has also been made readily accessible with a user-friendly database.

Acknowledgements

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). AD and JKM acknowledges the DBT, India for Research Fellowship Awards. We are also 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.

Availability of data and materials

All the raw sequence data submitted to NCBI with BioProject Accession Number PRJNA985750 (https://www.ncbi.nlm.nih.gov/sra/PRJNA985750). All relevant information related to DNA sequence variants and DNA methylation variants can be found in the Database (http://223.31.159.7/cicer/public/) and Supporting Materials of the manuscript.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

AD, JKM and LN performed the field/laboratory experiments and computational genomic analysis as well as drafted the manuscript. VT, AS, DR and PF assisted in drafting the manuscript and computational genomic analysis. AF and NPS helped in the devlopment of the database. ST, DC and SKP conceived and designed the research study, guided data analysis and interpretation, and participated in drafting and correcting the manuscript critically. All authors gave the final approval of the version to be published.

References

1. Shi, J., & Lai, J. (2015). Patterns of genomic changes with crop domestication and breeding. *Current Opinion in Plant Biology*, *24*, 47–53. https://doi.org/10.1016/j.pbi.2015.01.008

2. Varma Penmetsa, R., Carrasquilla-Garcia, N., Bergmann, E. M., Vance, L., Castro, B., Kassa, M. T., Sarma, B. K., Datta, S., Farmer, A. D., Baek, J. M., Coyne, C. J., Varshney, R. K., von Wettberg, E. J., & Cook, D. R. (2016). Multiple post-domestication origins of *kabuli* chickpea through allelic variation in a diversification-associated transcription factor. *The New phytologist*, *211*(4), 1440–1451. https://doi.org/10.1111/nph.14010

3. Liu, R., How-Kit, A., Stammitti, L., Teyssier, E., Rolin, D., Mortain-Bertrand, A., Halle, S., Liu, M., Kong, J., Wu, C., Degraeve-Guibault, C., Chapman, N. H., Maucourt, M., Hodgman, T. C., Tost, J., Bouzayen, M., Hong, Y., Seymour, G. B., Giovannoni, J. J., & Gallusci, P. (2015). A Demeter-like DNA demethylase governs tomato fruit ripening. *Proceedings of the National Academy of Sciences*, USA, *112*(34), 10804–10809. https://doi.org/10.1073/pnas.1503362112

4. Ong-Abdullah, M., Ordway, J. M., Jiang, N., Ooi, S.-E., Kok, S.-Y., Sarpan, N., Azimi, N., Hashim, A. T., Ishak, Z., Rosli, S. K., Malike, F. A., Bakar, N. A., Marjuni, M., Abdullah, N., Yaakub, Z., Amiruddin, M. D., Nookiah, R., Singh, R., Low, E.-T. L., … Martienssen, R. A. (2015). Loss of karma transposon methylation underlies the mantled somaclonal variant of oil palm. *Nature*, *525*(7570), 533–537. https://doi.org/10.1038/nature15365

5. Cortijo, S., Wardenaar, R., Colomé-Tatché, M., Gilly, A., Etcheverry, M., Labadie, K., Caillieux, E., Hospital, F., Aury, J.-M., Wincker, P., Roudier, F., Jansen, R. C., Colot, V., & Johannes, F. (2014). Mapping the epigenetic basis of complex traits. *Science*, *343*(6175), 1145–1148. https://doi.org/10.1126/science.1248127

6. Chan, S. W.-L., Henderson, I. R., & Jacobsen, S. E. (2005). Gardening the genome: DNA methylation in *Arabidopsis thaliana*. *Nature Reviews Genetics*, *6*(5), 351–360. https://doi.org/10.1038/nrg1601

7. Matzke, M. A., & Mosher, R. A. (2014). RNA-directed DNA methylation: An epigenetic pathway of increasing complexity. *Nature Reviews Genetics*, *15*(6), 394–408. https://doi.org/10.1038/nrg3683

8. Law, J. A., & Jacobsen, S. E. (2010). Establishing, maintaining, and modifying DNA methylation patterns in plants and animals. *Nature Reviews Genetics*, *11*(3), 204–220. https://doi.org/10.1038/nrg2719

9. Niederhuth, C. E., & Schmitz, R. J. (2017). Putting DNA methylation in context: From genomes to gene expression in plants. *Biochimica Et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, *1860*(1), 149–156. https://doi.org/10.1016/j.bbagrm.2016.08.009

10. Eichten, S. R., Briskine, R., Song, J., Li, Q., Swanson-Wagner, R., Hermanson, P. J., Waters, A. J., Starr, E., West, P. T., Tiffin, P., Myers, C. L., Vaughn, M. W., & Springer, N. M. (2013). Epigenetic and genetic influences on DNA methylation variation in maize populations. *The Plant Cell*, *25*(8), 2783–2797. https://doi.org/10.1105/tpc.113.114793

11. Schöb, H., & Grossniklaus, U. (2006). The first high-resolution DNA "methylome." *Cell*, *126*(6), 1025–1028. https://doi.org/10.1016/j.cell.2006.09.002

12. Song, Q.-X., Lu, X., Li, Q.-T., Chen, H., Hu, X.-Y., Ma, B., Zhang, W.-K., Chen, S.-Y., & Zhang, J.-S. (2013). Genome-wide analysis of DNA methylation in soybean. *Molecular Plant*, *6*(6), 1961–1974. https://doi.org/10.1093/mp/sst123

13. Kawakatsu, T., Huang, S.-shan C., Jupe, F., Sasaki, E., Schmitz, R. J., Urich, M. A., Castanon, R., Nery, J. R., Barragan, C., He, Y., Chen, H., Dubin, M., Lee, C.-R., Wang, C., Bemm, F., Becker, C., O'Neil, R., O'Malley, R. C., Quarless, D. X., … Zhou, X. (2016). Epigenomic diversity in a global collection of *Arabidopsis thaliana* accessions. *Cell*, *166*(2), 492–505. https://doi.org/10.1016/j.cell.2016.06.044

14. Shen, Y., Zhang, J., Liu, Y., Liu, S., Liu, Z., Duan, Z., Wang, Z., Zhu, B., Guo, Y.-L., & Tian, Z. (2018). DNA methylation footprints during soybean domestication and improvement. *Genome Biology*, *19*(1), 128. https://doi.org/10.1186/s13059-018-1516-z

15. Bossdorf, O., Arcuri, D., Richards, C. L., & Pigliucci, M. (2010). Experimental alteration of DNAmethylation affects the phenotypic plasticity of ecologically relevant traits in *Arabidopsis thaliana*. *Evolutionary Ecology*, *24*(3), 541–553. https://doi.org/10.1007/s10682-010-9372-7

16. Lang, Z., Wang, Y., Tang, K., Tang, D., Datsenka, T., Cheng, J., Zhang, Y., Handa, A. K., & Zhu, J.-K. (2017). Critical roles of DNA demethylation in the activation of ripening-induced

genes and inhibition of ripening-repressed genes in tomato fruit. *Proceedings of the National Academy of Sciences*, USA, *114*(22), E4511–E4519. https://doi.org/10.1073/pnas.1705233114

17. Martin, A., Troadec, C., Boualem, A., Rajab, M., Fernandez, R., Morin, H., Pitrat, M., Dogimont, C., & Bendahmane, A. (2009). A transposon-induced epigenetic change leads to sex determination in melon. *Nature*, *461*(7267), 1135–1138. https://doi.org/10.1038/nature08498

18. Quadrana, L., Almeida, J., Asís, R., Duffy, T., Dominguez, P. G., Bermúdez, L., Conti, G., Corrêa da Silva, J. V., Peralta, I. E., Colot, V., Asurmendi, S., Fernie, A. R., Rossi, M., & Carrari, F. (2014). Natural occurring epialleles determine vitamin E accumulation in tomato fruits. *Nature Communications*, *5*(1), 4027. https://doi.org/10.1038/ncomms5027

19. Yang, H., Chang, F., You, C., Cui, J., Zhu, G., Wang, L., Zheng, Y., Qi, J., & Ma, H. (2014). Whole-genome DNA methylation patterns and complex associations with gene structure and expression during flower development in *Arabidopsis*. *The Plant Journal*, *81*(2), 268–281. https://doi.org/10.1111/tpj.12726

20. Gallego-Bartolomé, J., Gardiner, J., Liu, W., Papikian, A., Ghoshal, B., Kuo, H. Y., Zhao, J. M.-C., Segal, D. J., & Jacobsen, S. E. (2018). Targeted DNA demethylation of the Arabidopsis genome using the human TET1 catalytic domain. *Proceedings of the National Academy of Sciences*, USA, *115*(9), E2125–E2134. https://doi.org/10.1073/pnas.1716945115

21. Kujur, A., Bajaj, D., Upadhyaya, H. D., Das, S., Ranjan, R., Shree, T., Saxena, M. S., Badoni, S., Kumar, V., Tripathi, S., Gowda, C. L., Sharma, S., Singh, S., Tyagi, A. K., & Parida, S. K. (2015). Employing genome-wide SNP discovery and genotyping strategy to extrapolate the natural allelic diversity and domestication patterns in chickpea. *Frontiers in Plant Science*, *6*. https://doi.org/10.3389/fpls.2015.00162

22. Gupta, S., Nawaz, K., Parween, S., Roy, R., Sahu, K., Kumar Pole, A., Khandal, H., Srivastava, R., Kumar Parida, S., & Chattopadhyay, D. (2016). Draft genome sequence of *Cicer reticulatum* L., the wild progenitor of chickpea provides a resource for agronomic trait improvement. *DNA Research*, *24*(1), 1–10. https://doi.org/10.1093/dnares/dsw042

23. Hirayama, T., & Alonso, J. M. (2000). Ethylene captures a metal! metal ions are involved in ethylene perception and signal transduction. *Plant and Cell Physiology*, *41*(5), 548–555. https://doi.org/10.1093/pcp/41.5.548

24. Bhatia, H., Khemka, N., Jain, M., & Garg, R. (2018). Genome-wide bisulphite-sequencing revealsOrgan-specific methylation patterns in chickpea. *Scientific Reports*, *8*(1), 9704.https://doi.org/10.1038/s41598-018-27979-w

25. Varshney, R. K., Song, C., Saxena, R. K., Azam, S., Yu, S., Sharpe, A. G., Cannon, S., Baek, J., Rosen, B. D., Tar'an, B., Millan, T., Zhang, X., Ramsay, L. D., Iwata, A., Wang, Y., Nelson, W., Farmer, A. D., Gaur, P. M., Soderlund, C., … Cook, D. R. (2013). Draft genome sequence of Chickpea (cicer arietinum) provides a resource for trait improvement. *Nature Biotechnology*, *31*(3), 240–246. https://doi.org/10.1038/nbt.2491

26. Flynn, J. M., Hubley, R., Goubert, C., Rosen, J., Clark, A. G., Feschotte, C., & Smit, A. F. (2020). RepeatModeler2 for automated genomic discovery of transposable element families. Proceedings of the National Academy of Sciences of the United States of America, USA, 117(17), 9451–9457. https://doi.org/10.1073/pnas.1921046117

27. Patel, R. K., Jain, M. (2012). NGS QC Toolkit: a toolkit for quality control of next generation sequencing data. *PloS One*, *7*(2), e30619. https://doi.org/10.1371/journal.pone.0030619

28. Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, *25*(14), 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

29. DePristo, M. A., Banks, E., Poplin, R., Garimella, K. V., Maguire, J. R., Hartl, C., Philippakis, A. A., del Angel, G., Rivas, M. A., Hanna, M., McKenna, A., Fennell, T. J., Kernytsky, A. M., Sivachenko, A. Y., Cibulskis, K., Gabriel, S. B., Altshuler, D., & Daly, M. J. (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature Genetics*, *43*(5), 491–498. https://doi.org/10.1038/ng.806

30. Cingolani, P., Platts, A., Wang, leL., Coon, M., Nguyen, T., Wang, L., Land, S. J., Lu, X., & Ruden, D. M. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly*, *6*(2), 80–92. https://doi.org/10.4161/fly.19695

31. Garg, R., Narayana Chevala, V., Shankar, R., & Jain, M. (2015). Divergent DNA methylation patterns associated with gene expression in rice cultivars with contrasting drought and salinity stress response. *Scientific Reports*, 5, 14922. https://doi.org/10.1038/srep14922

32. Lindgreen S. (2012). AdapterRemoval: easy cleaning of next-generation sequencing reads. *BMC Research Notes*, *5*, 337. https://doi.org/10.1186/1756-0500-5-337

33. Krueger, F., & Andrews, S. R. (2011). Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics*, *27*(11), 1571-1572. https://doi.org/10.1093/bioinformatics/btr167

34. Akalin, A., Kormaksson, M., Li, S. *et al.* methylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles. *Genome Biology*, 13, R87 (2012). https://doi.org/10.1186/gb-2012-13-10-r87

35. 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*, USA, *109*(9), 3570–3575. https://doi.org/10.1073/pnas.1201043109

36. Malik, N., Dwivedi, N., Singh, A., 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*, 23253. https://doi.org/10.1038/srep23253

37. Kent, W. J., Zweig, A. S., Barber, G., Hinrichs, A. S., & Karolchik, D. (2010). BigWig and BigBed: enabling browsing of large distributed datasets. *Bioinformatics*, *26*(17), 2204-2207.

38. Buels, R., Yao, E., Diesh, C. M., Hayes, R. D., Munoz-Torres, M., Helt, G., ... & Holmes, I. H. (2016). JBrowse: a dynamic web platform for genome visualization and analysis. *Genome Biology*, *17*(1), 1-12.

39. Varshney, R. K., Thudi, M., Roorkiwal, M., He, W., Upadhyaya, H. D., Yang, W., Bajaj, P., Cubry, P., Rathore, A., Jian, J., Doddamani, D., Khan, A. W., Garg, V., Chitikineni, A., Xu, D., Gaur, P. M., Singh, N. P., Chaturvedi, S. K., Nadigatla, G. V., … Liu, X. (2019). Resequencing of 429 chickpea accessions from 45 countries provides insights into genome diversity, domestication and agronomic traits. *Nature Genetics*, *51*(5), 857–864. https://doi.org/10.1038/s41588-019-0401-3

40. Varshney, R. K., Roorkiwal, M., Sun, S., Bajaj, P., Chitikineni, A., Thudi, M., Singh, N. P., Du, X., Upadhyaya, H. D., Khan, A. W., Wang, Y., Garg, V., Fan, G., Cowling, W. A., Crossa, J., Gentzbittel, L., Voss-Fels, K. P., Valluri, V. K., Sinha, P., … Liu, X. (2021). A chickpea genetic variation map based on the sequencing of 3,366 genomes. *Nature*, *599*(7886), 622–627. https://doi.org/10.1038/s41586-021-04066-1

41. Junaid, A., Kumar, H., Rao, A. R., Patil, A. N., Singh, N. K., & Gaikwad, K. (2018). Unravelling the epigenomic interactions between parental inbreds resulting in an altered hybrid methylome in pigeonpea. *DNA Research*, *25*(4), 361–373. https://doi.org/10.1093/dnares/dsy008

42. Jeltsch, A. (2010). Phylogeny of methylomes. *Science*, *328*(5980), 837–838. https://doi.org/10.1126/science.1190738

43. Zemach, A., & Zilberman, D. (2010). Evolution of eukaryotic DNA methylation and the pursuit of safer sex. *Current Biology*, *20*(17), R780–R785. https://doi.org/10.1016/j.cub.2010.07.007

44. Takuno, S., & Gaut, B. S. (2011). Body-methylated genes in Arabidopsis thaliana are functionally important and evolve slowly. *Molecular Biology and Evolution*, *29*(1), 219–227. https://doi.org/10.1093/molbev/msr188

45. To, T. K., Nakaminami, K., Kim, J.-M., Morosawa, T., Ishida, J., Tanaka, M., Yokoyama, S., Shinozaki, K., & Seki, M. (2011). *Arabidopsis HDA6* is required for freezing tolerance. *Biochemical and Biophysical Research Communications*, *406*(3), 414–419. https://doi.org/10.1016/j.bbrc.2011.02.058

46. Sahebi, M., Hanafi, M. M., van Wijnen, A. J., Rice, D., Rafii, M. Y., Azizi, P., Osman, M., Taheri, S., Bakar, M. F., Isa, M. N., & Noor, Y. M. (2018). Contribution of transposable elements in the plant's genome. *Gene*, *665*, 155–166. https://doi.org/10.1016/j.gene.2018.04.050

47. Zhong, X., Wang, Y., Liu, X., Gong, L., Ma, Y., Qi, B., Dong, Y., & Liu, B. (2009). DNA methylation polymorphism in annual wild soybean (*Glycine soja sieb.* et zucc.) and cultivated soybean (*G. max* L. Merr.). *Canadian Journal of Plant Science*, *89*(5), 851–863. https://doi.org/10.4141/cjps08215

48. Noh, Y.-S., & Amasino, R. M. (2003). Pie1, an ISWI family gene, is required for FLC activation and floral repression in *Arabidopsis*. *The Plant Cell*, *15*(7), 1671–1682. https://doi.org/10.1105/tpc.012161

49. Sharma, S., & Upadhyaya, H. D. (2019). Photoperiod response of annual wild *cicer* species and cultivated chickpea on phenology, growth, and yield traits. *Crop Science*, *59*(2), 632–639. https://doi.org/10.2135/cropsci2018.07.0438

50. Song, Q., Zhang, T., Stelly, D. M., & Chen, Z. J. (2017). Epigenomic and functional analyses reveal roles of epialleles in the loss of photoperiod sensitivity during domestication of allotetraploid cottons. *Genome Biology*, *18*(1). https://doi.org/10.1186/s13059-017-1229-8

51. Yadav, S. S., Kumar, J., Yadav, S. K., Singh, S., Yadav, V. S., Turner, N. C., & Redden, R. (2006). Evaluation of *Helicoverpa* and drought resistance in *desi* and *kabuli* Chickpea. *Plant Genetic Resources*, *4*(3), 198–203. https://doi.org/10.1079/pgr2006123

52. Hh Farooq, M., Ullah, A., Lee, D.-J., Alghamdi, S. S., & Siddique, K. H. M. (2018). Desi chickpea genotypes tolerate drought stress better than *kabuli* types by modulating germination metabolism, trehalose accumulation, and carbon assimilation. *Plant Physiology and Biochemistry*, *126*, 47–54. https://doi.org/10.1016/j.plaphy.2018.02.020

53. Liu, J., Novero, M., Charnikhova, T., Ferrandino, A., Schubert, A., Ruyter-Spira, C., Bonfante, P., Lovisolo, C., Bouwmeester, H. J., & Cardinale, F. (2013). Carotenoid cleavage dioxygenase 7 modulates plant growth, reproduction, senescence, and determinate nodulation in the model legume *Lotus japonicus*. *Journal of Experimental Botany*, *64*(7), 1967–1981. https://doi.org/10.1093/jxb/ert056

54. Imran, A., Mirza, M. S., Shah, T. M., Malik, K. A., & Hafeez, F. Y. (2015). Differential response of *kabuli* and *desi* chickpea genotypes toward inoculation with PGPR in different soils. *Frontiers in Microbiology*, *6*. https://doi.org/10.3389/fmicb.2015.00859

55. Yao, N., Zhang, Z., Yu, L., Hazarika, R., Yu, C., Jang, H., Smith, L. M., Ton, J., Liu, L., Stachowicz, J. J., Reusch, T. B. H., Schmitz, R. J., & Johannes, F. (2023) An evolutionary epigenetic clock in plants. *Science*, *381*(6665):1440-1445. https://doi.org/10.1126/science.adh9443

56. Dao, T. T., Linthorst, H. J., & Verpoorte, R. (2011). Chalcone synthase and its functions in plant resistance. *Phytochemistry Reviews*, *10*(3), 397–412. https://doi.org/10.1007/s11101-011-9211-7

57. Liu, X.-J., Chuang, Y.-N., Chiou, C.-Y., Chin, D.-C., Shen, F.-Q., & Yeh, K.-W. (2012). Methylation effect on chalcone synthase gene expression determines anthocyanin pigmentation in floral tissues of two oncidium orchid cultivars. *Planta*, *236*(2), 401–409. https://doi.org/10.1007/s00425-012-1616-z

Figure legends

Figure 1. **Overview of cytosine methylation levels in cultivated and wild *Cicer* gene pool. A)** Relative proportion of methyl cytosines detected in each of the three methyl contexts (CG, CHG and CHH) among five accessions belonging to *desi*, *kabuli* and wild chickpea. The proportion is estimated as the percentage of methyl cytosines representing the individual methyl contexts of CG, CHG and CHH detected from all these three combined methyl contexts. **B)** Global methylation levels of CG, CHG and CHH methyl cytosines in *desi*, *kabuli* and wild chickpea.

Figure 2. Genomic constitution of methyl cytosines in cultivated and wild chickpea. Distribution of methyl cytosines in gene-body as well as in 2-kb upstream and downstream genomic regions of **(A)** genes and **(B)** transposable elements, in five accessions belonging to cultivated and wild chickpea.

Figure 3. Diversity and dynamics of methylation in cultivated and wild chickpea. A) A heatmap depicting methylation diversity (DMRs) among different wild and cultivated chickpea accessions. **B)** Hypermethylated (blue) and hypomethylated (red) differentially methylated regions (DMRs) in three different methyl sequence contexts identified among five *desi, kabuli* and wild chickpea accessions.

Figure 4. Distribution of differentially methylated regions (DMRs) identified using five diverse cultivated and wild chickpea accessions within different sequence components of the chickpea genome.

Figure 5. Differential methylation and their correspondence with expression of known genes regulating traits of agronomic importance in chickpea. A) Differential methylation (CG) status of two Nodulation-Signaling Pathway (*NSP1* and *NSP2*) genes known to regulate nodulation efficiency in legumes. Locus IDs of genes correspond to NCBI *kabuli* (CDC Frontier) reference genome (v1.0, ASM33114V1). The highlighted (dark grey) regions correspond to genebody (exon) region whereas non-highlighted regions correspond to upstream (URR) and downstream (DRR) regulatory regions. **B)** Relative expression profiling of differentially methylated *NSPs*, Cryptochrome 1 (*CRY1*), *PHOTOPERIOD-INDEPENDENT EARLY FLOWERING 1* (*PIE1*) genes in *desi* (ICC 4958, ICC 5590), *kabuli* (ICC 8155, ICC 8260), and wild (ICC 17160) *Cicer* accessions. Data shown here, represents the relative expression changes in the corresponding

differentially methylated genes with respect to the internal control (*actin*). Values represent the mean ± SE with 3 biological and 3 technical replicates for each sample used in the RT-PCR assay.

Figure 6. **Contribution of genetic and epigenetic variations to chickpea evolution and domestication. A)** Phylogenetic (NJ) tree constructed based on genome-wide genetic polymorphism (SNPs/InDels), **B)** Phylogenetic (NJ) tree constructed based on genome-wide differential methylated cytosine in CG, CHG and CHH contexts.

Figure 7. **Relative abundance of identified DNA sequence variation and DNA methylation variation throughout the chickpea genome. A)** Density plots of DNA sequence variants (red line) and DNA methylation variants (green line) identified by comparison between wild and cultivated chickpea accessions. **B)** Density plots of DNA sequence variants (red line) and DNA methylation variants (green line) identified by comparison between *desi* and *kabuli* cultivated chickpea accessions.

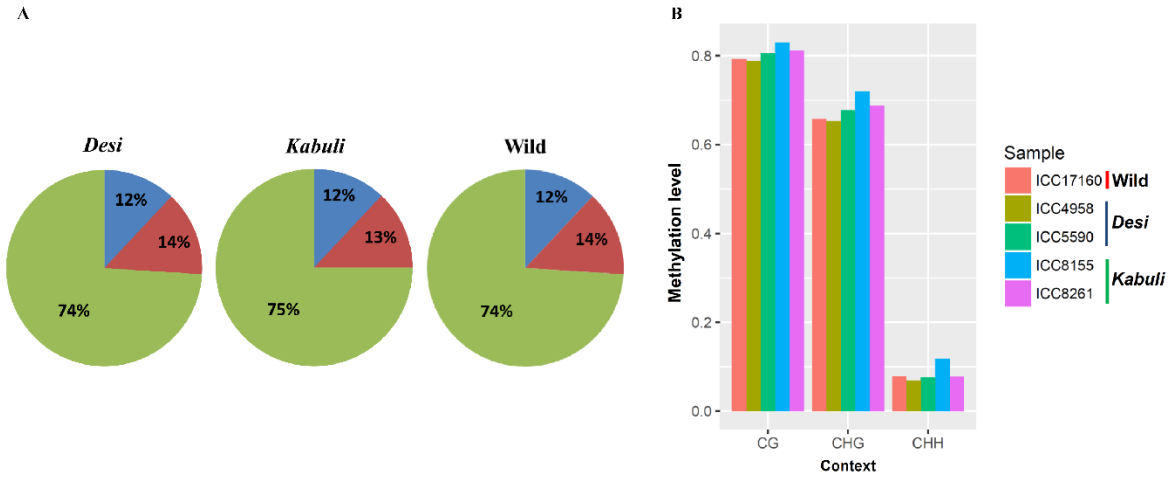

Figure 1. Overview of cytosine methylation levels in cultivated and wild *Cicer* gene pool. A) Relative proportion of methyl cytosines detected in each of the three methyl contexts (CG, CHG and CHH) among five accessions belonging to *desi*, *kabuli* and wild chickpea. The proportion is estimated as the percentage of methyl cytosines representing the individual methyl contexts of CG, CHG and CHH detected from all these three combined methyl contexts. **B)** Global methylation levels of CG, CHG and CHH methyl cytosines in *desi*, *kabuli* and wild chickpea.

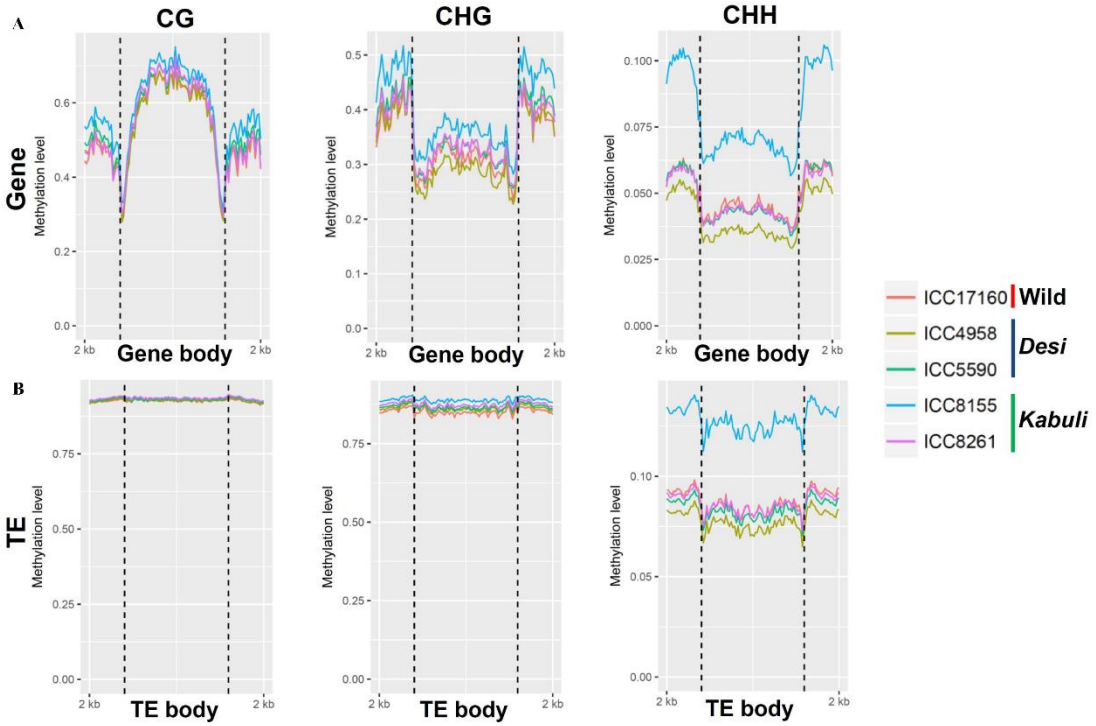

Figure 2. Genomic constitution of methyl cytosines in cultivated and wild chickpea. Distribution of methyl cytosines in gene-body as well as in 2-kb upstream and downstream genomic regions of **(A)** genes and **(B)** transposable elements, in five accessions belonging to cultivated and wild chickpea.

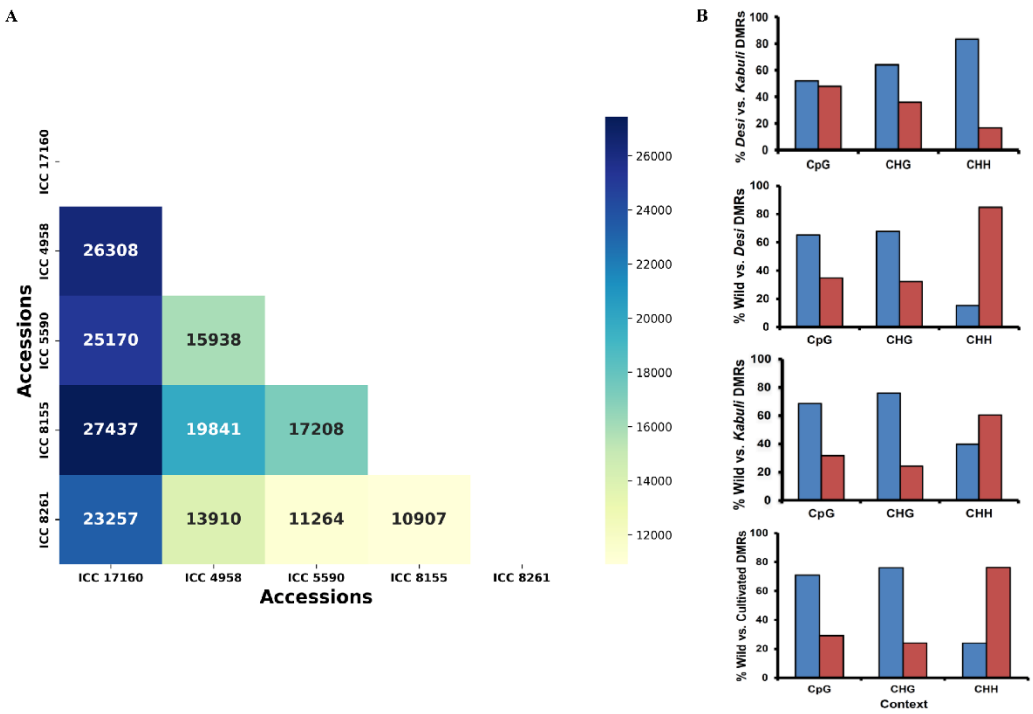

Figure 3. Diversity and dynamics of methylation in cultivated and wild chickpea. A) A heatmap depicting methylation diversity (DMRs) among different wild and cultivated chickpea accessions. **B)** Hypermethylated (blue) and hypomethylated (red) differentially methylated regions (DMRs) in three different methyl sequence contexts identified among five *desi, kabuli* and wild chickpea accessions.

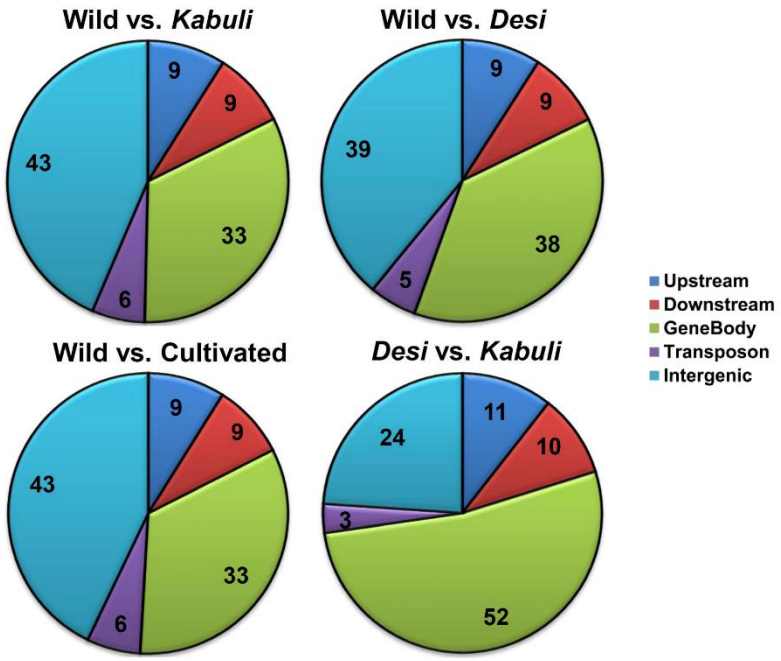

Figure 4. Distribution of differentially methylated regions (DMRs) identified using five diverse cultivated and wild chickpea accessions within different sequence components of the chickpea genome.

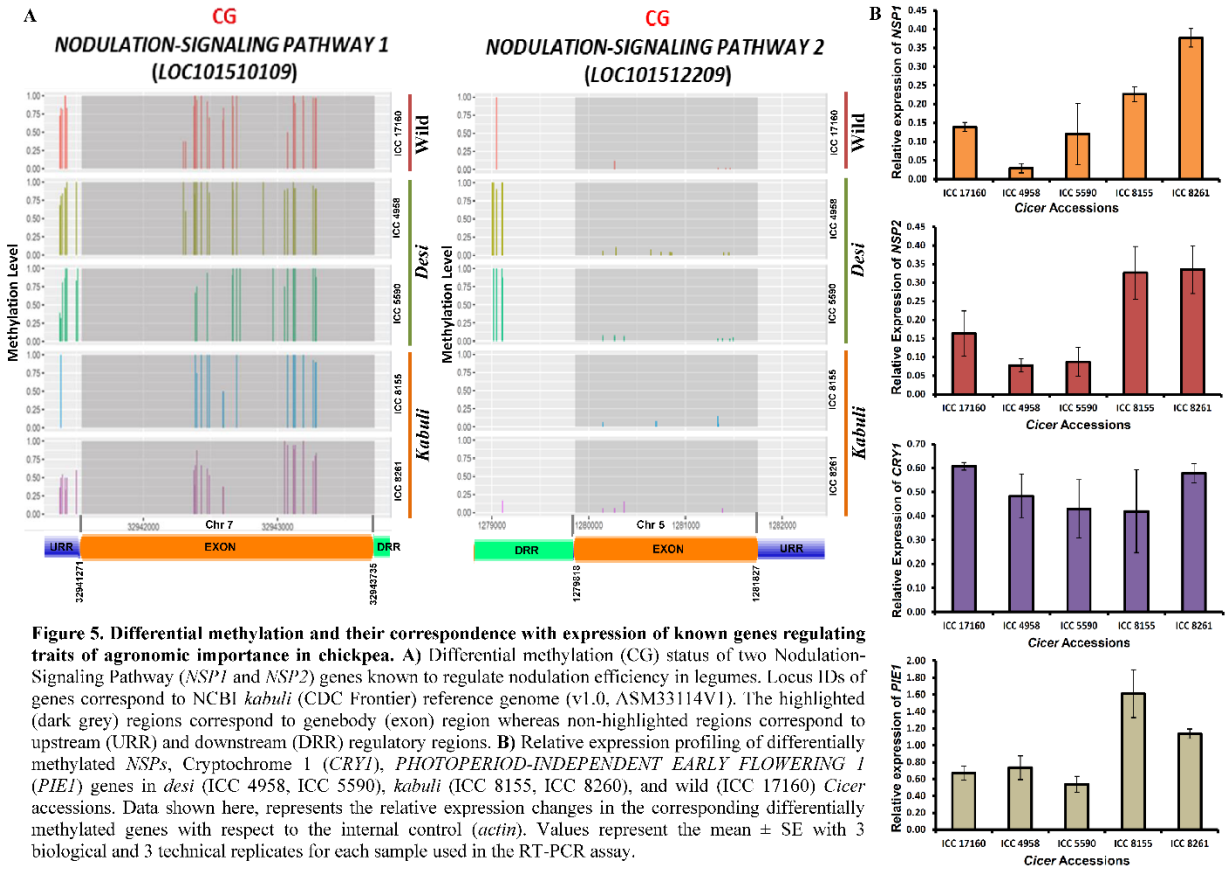

Figure 5. Differential methylation and their correspondence with expression of known genes regulating traits of agronomic importance in chickpea. A) Differential methylation (CG) status of two Nodulation-Signaling Pathway (*NSP1* and *NSP2*) genes known to regulate nodulation efficiency in legumes. Locus IDs of genes correspond to NCBI *kabuli* (CDC Frontier) reference genome (v1.0, ASM33114V1). The highlighted (dark grey) regions correspond to genebody (exon) region whereas non-highlighted regions correspond to upstream (URR) and downstream (DRR) regulatory regions. **B)** Relative expression profiling of differentially methylated *NSPs*, Cryptochrome 1 (*CRY1*), *PHOTOPERIOD-INDEPENDENT EARLY FLOWERING 1* (*PIE1*) genes in *desi* (ICC 4958, ICC 5590), *kabuli* (ICC 8155, ICC 8260), and wild (ICC 17160) *Cicer* accessions. Data shown here, represents the relative expression changes in the corresponding differentially methylated genes with respect to the internal control (*actin*). Values represent the mean ± SE with 3 biological and 3 technical replicates for each sample used in the RT-PCR assay.

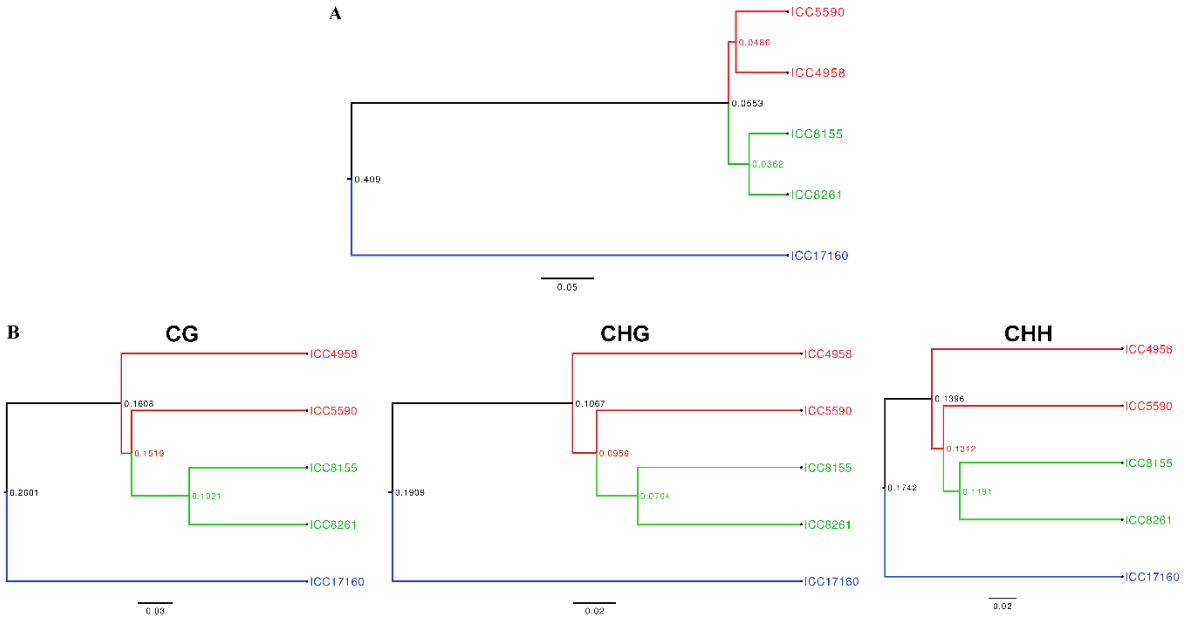

Figure 6. Contribution of genetic and epigenetic variations to chickpea evolution and domestication. A) Phylogenetic (NJ) tree constructed based on genome-wide genetic polymorphism (SNPs/InDels), **B)** Phylogenetic (NJ) tree constructed based on genome-wide differential methylated cytosine in CG, CHG and CHH contexts.

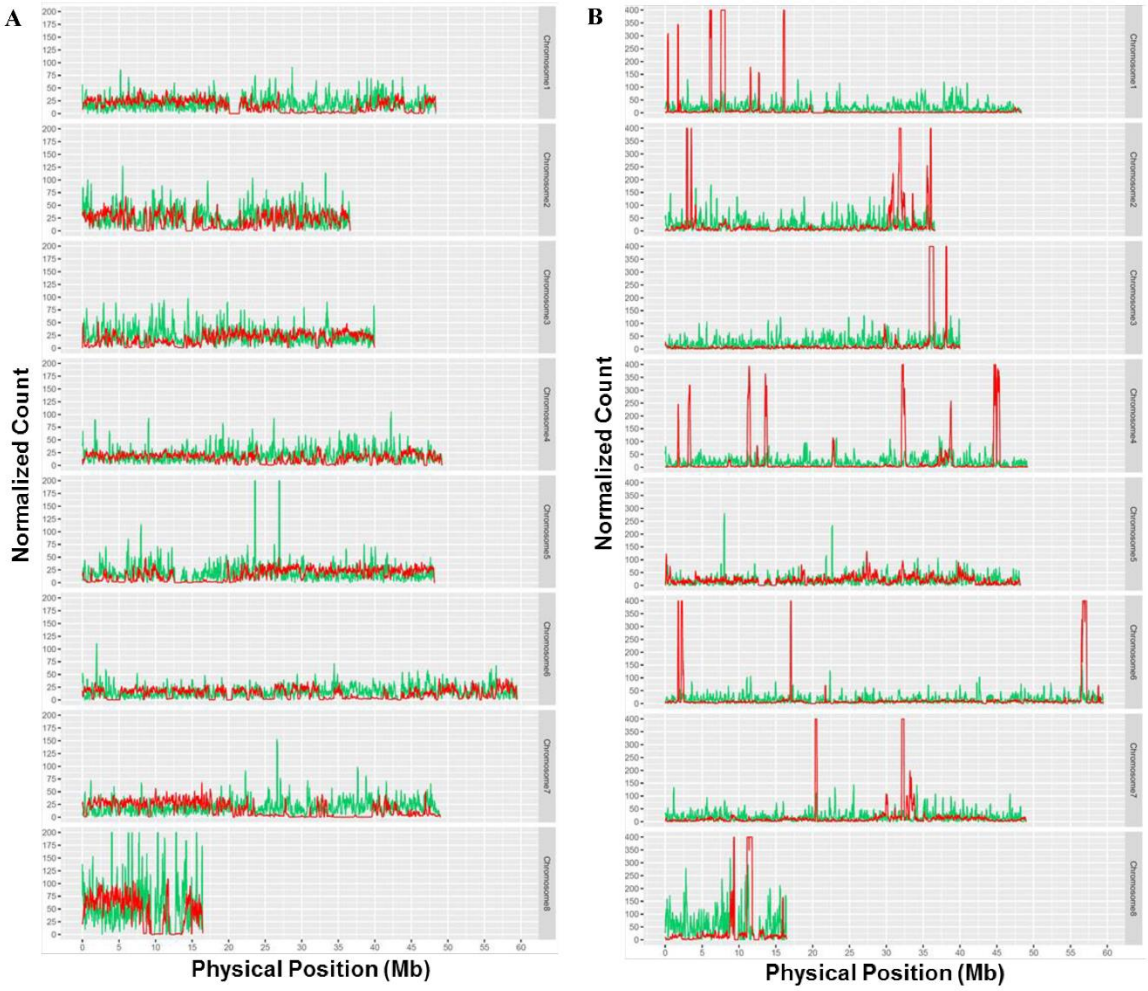

Figure 7. Relative abundance of identified DNA sequence variation and DNA methylation variation throughout the chickpea genome. A) Density plots of DNA sequence variants (red line) and DNA methylation variants (green line) identified by comparison between wild and cultivated chickpea accessions. **B)** Density plots of DNA sequence variants (red line) and DNA methylation variants (green line) identified by comparison between *desi* and *kabuli* cultivated chickpea accessions.