

RESEARCH

Open Access



# Genome-wide DNA methylation and their transgenerational pattern differ in *Arabidopsis thaliana* populations originated along the elevation of West Himalaya

Akanksha Singh<sup>1†</sup>, Ashwani Kumar Verma<sup>1,3†</sup>, Sunil Kumar<sup>1,3</sup>, Sumit Kumar Bag<sup>1,3,4</sup> and Sribash Roy<sup>1,2,3\*</sup>

## Abstract

Methylation at 5' cytosine of DNA molecule is an important epigenetic mark. It is known to play critical role in adaptation of organisms under different biotic and abiotic stressors via modulating gene expression and/or chromatin architecture. Plant populations evolved under variable climatic conditions may have evolved different epigenetic marks including DNA methylation. Here we, describe the genome-wide DNA methylation pattern under native field, F1 and F6 generation followed by their association with phenotypes, climate and global gene expression in the three *Arabidopsis thaliana* populations originated at different elevation ranges of Indian West Himalaya. We show that the global methyl cytosine (mC) content is more or less similar in the three populations but differ in their distribution across genome. There was an increase in differential methylation between the populations as elevation increased. The methylation divergence was the highest between the low and the high elevation populations. The high elevation populations were hypo-methylated than the low elevation population. The methylation in the genes was associated with population specific phenotypes and climate of the region. The genes which were differentially methylated as well as differentially expressed between the low and high elevation populations were mostly related to abiotic stresses. When grown under controlled condition, there was gain of differential methylation over native condition and the maximum percent changes was observed in CHH-sequence context. Further ~99.8% methylated cytosines were stably passed on from F1 to F6 generation. Overall, our data suggest that high elevation population is epigenetically more plastic under changing environmental condition.

**Background** *Arabidopsis thaliana* is the model plant species and has been extensively studied to understand plants life processes. There are numerous reports on its origin, demography, evolution, epigenomes and adaptation etc. however, Indian populations of *Arabidopsis thaliana* evolved along wide elevation ranging from ~ 700 m amsl to ~ 3400 m amsl not explored yet. Here we, describe the genome-wide DNA methylation pattern under native field, F1 and F6 generation followed by their association with phenotypes, climate and global gene expression in the three *Arabidopsis thaliana* populations originated at different elevation ranges of Indian West Himalaya.

**Results** In our study we found that total mCs percent was more or less similar in the three populations but differ in their distribution across genome. The proportion of CG-mCs was the highest, followed by CHH-mCs and CHG-mCs

<sup>†</sup>Akanksha Singh and Ashwani Kumar Verma contributed equally to this work.

\*Correspondence:

Sribash Roy

sribashroy@uohyd.ac.in

Full list of author information is available at the end of the article



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

in all the three populations. Under native field condition the methylation divergence was more prominent between low and high elevation populations and the high elevation populations were hypo-methylated than the low elevation population. The methylation in the genes was linked to population-specific phenotypes and the regional climate. The genes that showed differential methylation and expression between low and high elevation populations were primarily associated with abiotic stress responses. When grown under controlled condition, there was gain of differential methylation compared to the native condition and the maximum percent changes was observed in CHH-sequence context. Further 99.8% methylated cytosines were stably passed on from F1 to F6 generation.

**Conclusions** The populations of *A. thaliana* adapted at different climatic conditions were significantly differentially methylated both under native and controlled condition. However, the magnitude and extent of gain or loss of methylation were most significant between the low and the high elevation populations. Overall, our data suggest that high elevation population is epigenetically more plastic under changing environmental condition.

**Keywords** DNA methylation, Indian *Arabidopsis*, Transgenerational methylation, Abiotic stress, Elevation

## Background

Living organisms constantly face environmental challenges. However, being sessile plants cannot avoid the stresses of the natural environment as sensitively as animals. In order to cope with such challenges, they evolved different strategies. One such strategies are the regulation through epigenetic mechanism. Although adaptive phenotypic variation is mainly regulated by genetic architecture driven by mutation, drift and migration however, it accounts for a small proportion of the total phenotypic variation observed in many species [50, 80, 114]. A part of such variations has been attributed to heritable epigenetic variation [4, 12, 29, 66]. It is opined that due to this epigenetic variations genome-wide association studies fail to explain a substantial part of the heritable variation within species [6, 40, 48].

Epigenetic variation includes DNA methylation, histone modifications, and RNA-mediated gene silencing [76]. DNA methylation refers to the transfer of a methyl group onto the 5<sup>th</sup> position of the cytosine to form 5 methyl cytosine (5mC) [61]. In plants, DNA methylation typically occurs at cytosines in three sequence contexts: CG, CHG or CHH (where H represents a G, T or A nucleotide) [55] each of which is catalyzed by independent pathways [21, 96]. The methylation is primarily maintained during DNA replication and cell division by DNA methyltransferases, including maintenance methylases and de novo methylases [36]. In *Arabidopsis thaliana*, while CG methylation is mostly maintained by METHYLTRANSFERASE 1 (MET1), a homolog of DNA methyltransferase 1 (DNMT1) [22], CHROMOMETHYLASE 3 (CMT3) is a plant-specific DMT involves primarily in the maintenance of symmetrical CHG methylation [59] and DOMAINS REARRANGED METHYLTRANSFERASE 2 (DRM2), a homolog of DNA methyltransferase 3 (DNMT3) catalyzes asymmetric CHH methylation [15]. Since different methylation machineries are involved in methylating different sequence contexts distinguishing

these methylation pattern may provide cue to differential adaptation.

How selection pressures affect genome-wide DNA methylation levels in natural plant populations have been reported in *Arabidopsis thaliana*. Apart from genetic factors, variations in DNA methylation can arise either by altered methylation marks during mitosis or meiosis by chance, or in response to environmental changes [31, 52, 81, 109, 119]. DNA methylation varies substantially among individuals and populations [23, 32, 39, 91], Alonso, Ramos-Cruz and Becker, [2]) and often non-random and geographically structured [18, 30, 43, 60, 74]. Thus, hyper- and hypo-methylated accessions of *A. thaliana* were found in Sweden and Spain, respectively [43]. Methylation variations in response to environmental stressors, including herbivory [105], drought [27], temperature [69], salinity ([23, 118], heavy metal [20], elevated CO<sub>2</sub> [73, 86, 110] etc. have been reported which may coordinate adaptive phenotypes [117, 121]. Moreover, since rate of epi-mutation is reported to be substantially higher than genetic mutation rate [5, 37, 90], it may provide a particularly useful mechanism for generating phenotypic variation that would allow a more rapid response to environmental change than would genetic mechanisms. Besides the sequence context, DNA methylation also varies in different genomic features in which it occurs. For instances while methylation is often lower and more variable in genes and regulatory regions, the heterochromatic regions and transposable elements are usually heavily methylated [70, 91, 92]. In similar lines a report by Dubin et al., [18] on natural Swedish accessions of *Arabidopsis* revealed while CHH sequence context are primarily found in transposons and related to transposon silencing, methylation in the CG context is largely found in genes correlated with environmental factors and positively correlated with expression of those genes suggesting a possible role in local adaption by natural selection [18]. This and other studies shows while DNA

methylation is almost exclusively a repressive mark on TEs, functions of gene body methylation though lesser known but may regulate its expression [27, 110]. Nevertheless, for both influences of sequence context and genomic features on methylation variation, there appears to be high species-specificity in plants [70].

Large scale population analyses on *Arabidopsis* demonstrated association of climatic factors with DNA methylation in reference to genomic context (CG, CHG and CHH). For example, genome-wide association studies involving 122 *Arabidopsis thaliana* accessions across Eurasia has shown that climatic characteristics most abundantly co-varied with methylation in CHH context [43]. In a previous study, Dubin et al. [18] examined 150 accessions of *Arabidopsis thaliana* from Sweden and found that while temperature did not impact the genome-wide methylation pattern but had a significant effect on CHH methylation. CHH methylation was reported to be 14% higher at 16°C than at 10°C [18]. CHH methylation is largely associated with transposable elements (TEs) and their association with environmental characteristics might indicates natural selection at the level of TE-specific methyltransferase genes regulating the methylation level [47]. Similarly, CG-context methylation, particularly within or near genes related to abiotic stress responses was found to be strongly associated with climatic factor [44]. In the non-model plant, *Quercus* species the DMCs, mostly CG-contexts was shown to be associated with environmental stressors [30, 77]. These and other studies considerably increased our understanding on how epigenetic variations are associated with different climatic conditions. It will be interesting to examine how epigenetic variations in plants along a steep elevation ranging 700 m amsl to 3400 m amsl, where the interplay between migration, drift and selection can be extremely dynamic [14] shape the epigenome of the populations.

Although epigenetic variations have been shown to be associated with climatic factors, yet role of epigenetic variation in governing adaptive evolution remains controversial [17]. A growing body of literature has shown transmission of epigenetic traits over generations thus consider it as a key evolutionary force [17, 30, 58, 66, 78, 89]. Using multigenerational selection experiment demonstrated that CG-methylation in *Arabidopsis thaliana* was significantly higher after five generation and majority of differentially methylated cytosine were stably inherited for two to three generations following selection. Further, they observed decrease in epigenetic diversity and methylation levels of some CG-DMCs (Differentially Methylated Cytosines) were associated with phenotypic changes. Similarly, using *Plantago lanceolata*, Gáspár et al., [28] demonstrated that much of the environment-related epigenetic variation is maintained in a

second-generation common garden [28]. Thus, at least part of the epigenetic variation observed in the field was stable, genomically nonrandom and of ecological significance. It is evident that both under natural or simulated selection climatic association with epigenetic traits and its heritable nature is wide spread but with varied degree of magnitude.

Natural populations originated at highly heterogeneous environment covering sub-tropical to temperate climate might face considerable contrasting selection pressure. While high elevation population constantly face challenges like high light intensity, UV, low temperature, marked differences in day night temperature, limited rainfall etc., the low elevation population faces benign environment. Therefore, we aimed to explore the genome-wide DNA methylation profiling of three *Arabidopsis thaliana* populations originated at three elevations level of 700 m amsl (low), 1800 amsl (medium) and 3400 m amsl (high) of Indian west Himalaya. The populations were shown to be both phenotypically and genotypically differentiated probably driven by local topography and climatic condition (Singh et al., [93]; Tyagi, Singh, et al., [106]). While the low elevation population exhibited higher leaf count and larger siliques than higher elevation one, the high elevation population flowered late as compared to the low elevation one. Overall, the population of high elevation was more plastic as compared to the low elevation population [94]. Although, much of the trait variability is known to be genetically driven, epigenetic factors play its role in shaping trait differentiation, especially in species like *Arabidopsis thaliana* because of limited genetic diversity and clonal life cycle. We specifically asked the following questions a) What is the extent of DNA methylation variation in the populations along the elevation? b) Are differentially methylated regions enriched for ecologically significant genes? c) What is the extent of methylation plasticity under new environmental condition and subsequent generation under controlled condition?

## Methods

### Study site and sample collection

Three natural populations of *Arabidopsis thaliana* were collected from their native sites in the west Himalayas, viz. Dehradun (700 m above mean sea level (m amsl)), Munsyari (2000 m amsl) and Chitkul (3400 m amsl) by [107]. These populations are hereafter referred to as the low elevation population (Lep), medium elevation population (Mep) and high elevation population (Hep) respectively. The details of the geographic locations and sample collection of the same populations can be found elsewhere (Table S1, S2) (Singh et al., [93]; Tyagi, Singh, et al., [106]; [94]). Briefly, the three populations were referred

to as Deh, Mun, and Chit respectively, corresponding to Lep, Mep and Hep here. The habitat of the Lep population was highly disturbed open lawn, whereas the Mep population was found in un-disturbed fertile hill slope. The habitat of Hep population was along the country roadside and represented by undisturbed mountainous slopes along river-side. Field visits were conducted coinciding with the early flowering stage of the plants at each of the three sites. The early flowering stage was defined as plants with first flower open (principal growth stage 6.00) [11] to plants having not more than 10 green fruits. As the three sites were located at three different altitudes, their time of visit and habitat was also variable. The three sites were visited in the second week of February, first week of April and the third week of May. The seeds of the accessions have been deposited at National Bureau of Plant Genetic Resources, New Delhi, India (<http://www.nbpgr.ernet.in>).

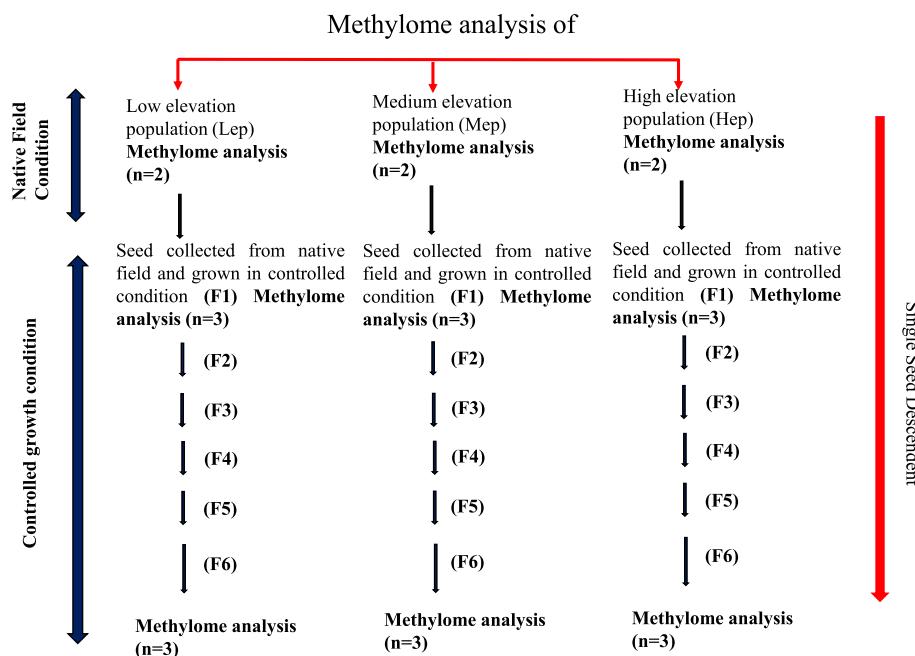
## Whole genome bisulfite sequencing

A total of 24 samples, including plants grown in the field and controlled environments (F1 and F6 generations), were used for bisulfite sequencing. For the analysis of DNA methylation pattern under the native field conditions, two individuals per population were randomly collected coinciding with the mid flowering stage of the plants as described above in sample collection section (Fig. 1). The seeds collected from the native field were

grown under controlled conditions (F1 generation) and propagated up to six generations through single seed descent method (F6 generation). Nine individuals representing three populations (three accessions per population) were collected from the F1 generation and the F6 progenies of the same nine parents were used for bisulfite sequencing. All the samples corresponded to the mid flowering stage of the plant. DNA was isolated from the leaves of these plants using innuPREP Plant DNA Kit (Analytik Jena, Germany) following manufacturer's protocol. The quality of DNA was checked using 0.8% agarose gel electrophoresis and quantified using QIAxpert<sup>TM</sup> Nanodrop (Qiagen, USA). About 1ug of DNA from each sample was bisulfite converted using Zymo EZ DNA methylation kit (Zymo research, USA) using manufacturer's protocol. The bisulfite converted DNA was purified and used for library preparation using TrueSeq DNA sample prep kit using manufacturer's protocol. Paired end sequencing of these 24 libraries were sequenced on Illumina HiSeq 1000 platform with 100 bp reads. The list of accession sequenced and the corresponding sequence SRA ID is given in table S1 and S3.

## Processing of bisulfite sequencing reads and identification of methylation

The raw paired end reads were filtered for low quality reads ( $Q < 30$ ) and adapter using Trimmomatic [10]. The quality filtered reads of each individual were then aligned



**Fig. 1** Schematic representation of experimental design and sample collection for methylome sequencing (n, represents number of samples used for methylome sequencing)

to its population specific pseudo-reference genome using bowtie 2 [54]. The reference genome was first converted in-silico into bisulfite converted genome using Bismark tool [51]. To maintain stringency only uniquely mapped reads were considered for further analysis. The reads mapping to the same start and end coordinates were considered as duplicates and were removed using perl script deduplicate\_bismark provide with Bismark tool. The mapping file was examined to identify the presence of methylation bias (M-bias). If M-bias was detected, reads were trimmed for biased sequences (~10 bases from 5' end) to eliminate M-bias. After mapping of the reads the absolute methylation profile in each population was obtained using Bismark Methyl Extractor and whole genome cytosine report was obtained using option *-cytosine\_report*. A binomial test was executed on the comprehensive genome cytosine report, derived from the bismark methyl extractor, to identify cytosines that were significantly methylated across the entire genome ( $P < 0.05$ ) for each sample. This analysis took into account the computed non-conversion rate and a 5% false-positive sequencing error rate. The mCs represent the absolute profile of methylation per sample. The methylation level of each cytosine was defined as percentage of reads supporting methylation call to total number of reads covering that cytosine. These methylated cytosines were identified as cytosine having methylation percentage of at least 25% and covered by at least three reads. These methylated cytosines were further annotated against the *Arabidopsis thaliana* gene annotation file using BEDTOOLS [79]. The detailed pipeline of bisulfite sequencing is shown in figure S1a. The summary statistics of methylation data are shown in table S3.

#### Identification of differential methylation

The mapped files generated using Bismark were used to call for differential profiles both in terms of differentially methylated positions (DMPs) and differentially methylated regions (DMRs). To identify differential methylation profile between the populations, pair wise comparison was performed using three sets—Lep-Mep, Mep-Hep and Lep-Hep. The individuals of each population were regarded as biological replicates. Only those mCs which were covered by a minimum of 5 reads with  $Q > 30$  in each comparison set were considered. The coverage of each cytosine was normalized using option '*normalizeCoverage*' to ensure accuracy in comparisons. The differential methylation was then called using logistic regression controlling for overdispersion (variability among samples) [1]. The significance of the differential methylation was assessed using Fisher's exact test ( $P < 0.05$ ) and corrected for multiple comparisons. Methylation percentage was estimated as:  $100 * (\text{methylated cytosine} / (\text{methylated cytosine} + \text{unmethylated cytosine}))$  for each cytosine in each of the three contexts. The sites that had methylation percentage difference greater than 25% and an FDR  $< 0.05$  were considered as differentially methylated.

Cytosine + unmethylated cytosine)) for each cytosine in each of the three contexts. The sites that had methylation percentage difference greater than 25% and an FDR  $< 0.05$  were considered as differentially methylated.

The differentially methylated positions (DMPs) were further expressed in terms of differentially methylated regions (DMRs). The whole genome was divided into a tiling window with a window size of 100 bases and step size of five. Differential methylation was called for each region as methylation percentage average over all the cytosines covered in that region. Similar statistical analysis as done for the DMPs was conducted on DMRs to assess the significance of differential methylation. The regions were considered as differentially methylated regions (DMRs), if they had methylation percentage difference greater than 25% and an FDR  $< 0.05$ . Further the overlapping adjacent DMRs were merged into one larger DMRs. Both DMPs and DMRs were identified and analyzed separately for each of the three sequence contexts (CG, CHG and CHH). The detailed pipeline of identification of DMPs and DMRs are shown in figure S1b. The DMRs and DMPs were identified between the three populations growing under native field of controlled growth condition (F1 generation), or within populations growing under two different conditions (Field Vs controlled (F1)) and over generations (F1 Vs F6) as described above. The methylation variations identified between field and controlled grown (F1) plants for a population were regarded as 'environment driven methylation variations' and those identified between F1 and F6 plants were called as 'transgenerational methylation variations'.

#### Annotation and gene ontology analysis

The identified DMPs and DMRs were annotated against the available annotation of genes and transposable elements (TAIR10 annotation (gtf) file) for *A. thaliana* (<https://plants.ensembl.org/info/website/ftp/index.html>) using BEDTOOLS [79]. To further find the biological function of these genes lying in the differentially methylated sites Gene Ontology (GO) analysis was conducted using DAVIDGO [34, 35]. The GO terms were analyzed using *Arabidopsis thaliana* (TAIR 10) gene ontology terms as background. The GO terms were considered to be significantly enriched when false detection rate (FDR)  $< 0.05$ .

#### Analysis of association between phenotype and DNA methylation

To identify the specific single nucleotide methylation significantly associated with the phenotype (Table S4), we employed a linear mixed model method implemented in PyLMM (<https://github.com/nickFurlotte/pylmm>) [42]. The pair wise relatedness among the samples was

adjusted using a correlation structure among the samples in the form of kinship matrix. The kinship matrix was estimated using the genetic variation (SNPs) data derived from whole genome sequences of these populations (PRJNA564218). Individual sites of methylation were analyzed for association with each of the previously published phenotypic trait (Singh et al., [93]). Significant DMPs identified as described above were used in the analysis to minimize the false positive associations. Nineteen phenotypic traits that were shown to vary significantly between the populations under native field conditions were used for the analysis (Singh et al., [93]) (Table S4). The significance values (*P*-values) obtained from the analysis were further adjusted to *Q*-values using FDR method to account of multiple testing using R package QVALUE2 (Dabney, Storey and Warnes, [16]). Though the sample size was relatively less as required for this type of analysis (EGWAS), but a relatively lower significance values ( $P < 1E-07$ ) was considered to minimize false associations. The associated DMPs were also annotated to identify the site they corresponded to and to identify their role. The annotation was done using BEDTOOLS as described above.

#### Climate-methylation association analysis

A significant association of the climate with the methylation variations could be an indicative of either natural selection acting on genetic variants or strong influence of climate in guiding natural epigenetic variation at that locus [120]. The methylation association with climatic factors were analyzed as described above. The bioclimatic data for each site were downloaded from WordClim2 database. All 23 bioclimatic variables (as described in table S4) were used individually for the analysis. Specific sites were considered to be significantly associated with climate if the *Q*-value was less than 0.001. The significantly associated sites were annotated using BEDTOOLS [79].

Additionally, a multivariate analysis was performed to assess the overall association between methylation variations and climatic variations, using redundancy analysis (RDA). RDA uses a constrained ordination method similar to linear regression designed to check association in the presence of multiple dependent and independent variables. It was performed using Vegan package implemented in R version 2.4–3 (Oksanen et al., [72]). This analysis was conducted for methylation in each of the three sequence contexts separately using only uncorrelated (Pearson's Correlation  $r \geq 0.08$ ) bioclimatic variables. The significance of association was obtained using permutation test with 999 permutations. A stepwise selection analysis (using *forward.sel* function of Adegenet package of R version 2.1.10) was conducted to

find the most significant bioclimatic variable that was able to explain the majority of the methylation variation observed in these populations. *Varpart* function was used to partition the amount of variation in DNA methylation explained by each bioclimatic variable.

#### Clustering analysis

To identify the differentiation of the individual samples based on methylation, clustering analysis was conducted using R. The distance metric used for clustering was correlation along with 'Ward's method of agglomeration in the hierarchical clustering algorithm'. Only the cytosines that were covered by at least five reads with a quality score of more than 30 in all samples were considered for this analysis. This analysis was conducted individually for the cytosines in the three different sequence contexts (CG, CHG and CHH).

#### 5-azacytidine treatment

5-azacytidine (5-azaC) is a DNA demethylating agent that inhibits methyltransferase activity thus blocking methylation of newly synthesized DNA. This makes the global DNA hypo-methylated compared to the non-treated one. For induction of global hypo-methylation we essentially followed Kottler et al., 2018 (Kottler, Van-Wallendael and Franks, [49] [110],). Briefly, seeds of F1 generation were placed on a petri plate containing filter paper either soaked with 1.4 ml of 50  $\mu$ mol 5-azaC solution or water as control. After four days of cold-stratification, both the plates were transferred to controlled growth condition for germination. Two days post-germination, seedlings were transferred to pot filled with SoilriteMix. Thereafter, at weekly intervals a spray of 50  $\mu$ mol 5-azaC solution were given on the leaves of treated plants and the control plants were sprayed with distilled water till samples were harvested.

#### qRT-PCR of differentially methylated genes

For validation, three differentially methylated genes were randomly selected. The primer sequences are listed in table S5. For each comparison, three biological and two technical replicates were considered. Individual libraries were prepared from 2  $\mu$ g of total RNA. cDNA was synthesized using Go Script™ Reverse Transcription System (Promega, USA) following manufacturer's protocol. The amplification reaction was performed using ABI7300 real-time PCR system (Applied Biosystems). The threshold cycle (*C<sub>t</sub>*) value for each gene was quantified and normalized by *C<sub>t</sub>* value of internal control gene, Actin. The relative expression was calculated using  $2 - \Delta\Delta C_t$  method converted to log2 fold. Student's t-test was used as test of significance.

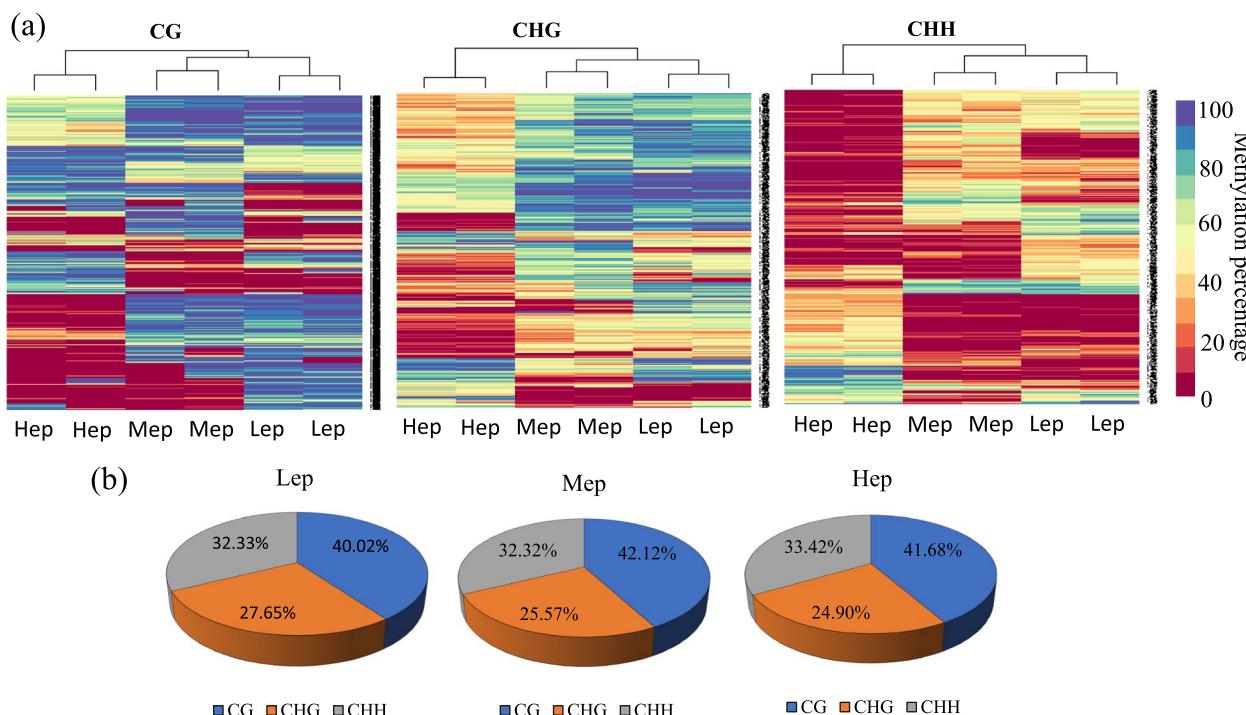
## Results

### Methylation landscape in the native field populations

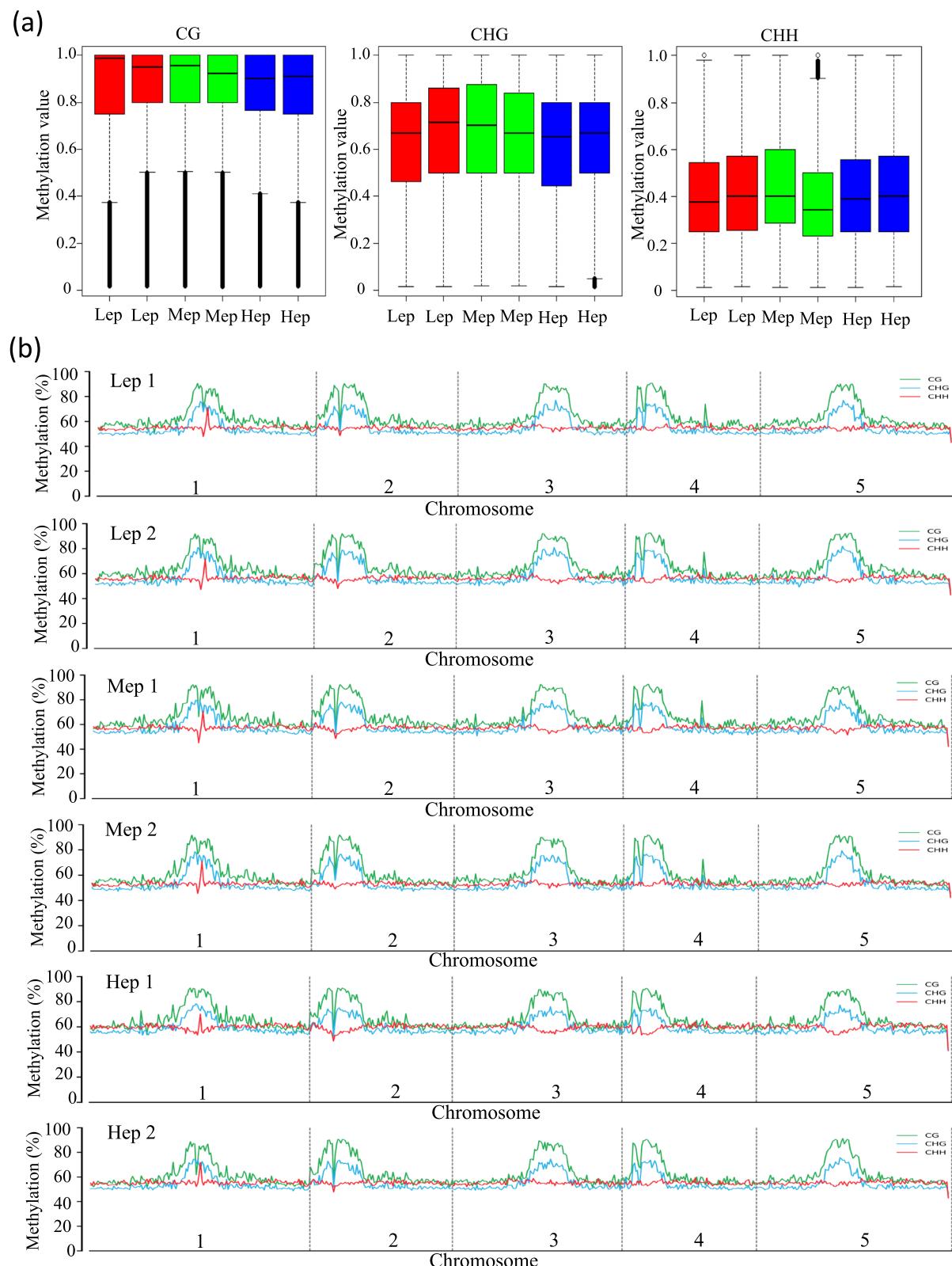
The total methyl cytosines (mCs) percent was more or less similar in the three populations (Lep-5.75%; Mep-6.22% and Hep-5.89%) (Table S6). However, the distribution of mCs varied in the three population (Fig. 2a). The proportion of CG-mCs (as a percent of total mCs) was the highest ( $41.27\% \pm 1.1$ ), followed by CHH-mCs ( $32.69\% \pm 0.63$ ) and CHG-mCs ( $26.04\% \pm 1.43$ ) in all the three populations (Fig. 2b). The average methylation level was also higher in CG context followed by CHG and CHH (Fig. 3a). As reported earlier, extensive DNA methylation was observed in the pericentromeric regions (Fig. 3b). More than half of the mCs in all the three sequence contexts were located in genic region and the rest were in promoter and intergenic region in almost equal proportion in the three populations. Further, about half of the genic CG-mCs and ~ 85% of genic CHG-mCs and CHH-mCs were distributed in transposons (Figure S2a, b). Notably, the accessions were clustered according to the population based on CG-mCs and CHG-mCs correlation analysis but not with CHH-mCs. This may suggest CG and CHG methylation were more of population specific than the CHH-mCs (Figure S3).

### Methylation divergence is more pronounced between the low and the high elevation populations under native field condition

Although genome-wide mCs frequency didn't vary significantly among the three populations, a comparative analysis between any two populations showed significant differentially methylated positions (DMPs). We considered sites having methylation percentage  $> 25\%$  and FDR  $< 0.05$  as differentially methylated. A total of 10,662 (9160 CG, 1007 CHG and 495 CHH), 15,945 (10,566 CG, 4106 CHG and 1273 CHH) and 20,390 (14,784 CG, 3897 CHG and 1709 CHH) DMPs were identified in Mep/Lep, Hep/Mep and Hep/Lep, respectively indicating a trend of increasing DMPs as elevation difference increased (Fig. 4a). At sequence context level, CG- and CHH-DMPs but not the CHG- DMPs showed increasing trend with increase in the elevation. Further, with respect to CG-mCs, higher elevation population were more hypo-methylated as compared to the respective low elevation population (Fig. 4a). For example, 63.56% of DMPs were hypo-methylated in Hep as compared to the Mep, 61.22% were hypo-methylated in Mep as compared to the Lep and 69.45% of DMPs were hypo-methylated in Hep as compared to the Lep (Figure S3).



**Fig. 2** Methylation profiling of the three populations under native field condition. (a) Heatmap showing differences in methylation values of the cytosines between the populations in three sequence contexts under native field condition. The scale shows the methylation level of the cytosine in percentage ranging from 0% (un-methylated) to 100% (complete methylation). (b) Pie-chart showing proportion of significantly methylated cytosines identified in the three sequence contexts in the three populations



**Fig. 3** Global methylation level and its distribution pattern of the native field samples in three sequence contexts. **(a)** The global methylation level (calculated as number of methylated cytosines/total number of cytosine) of the cytosine in the three populations (x-axis, individuals of three populations, Lep, Mep, and Hep); y-axis, methylation level, **(b)** Chromosome-wise methylation distribution level of three different sequence contexts

Next, we identified mCs coverage spanning a window-size of 100 bp with an overlap of five bp in order to identify the differentially methylated regions (DMRs) over the entire genome. A window was considered to be a DMR if methylation was  $> 25\%$  and an FDR  $< 0.05$ . A total of 8005 (6865 CG, 881 CHG and 259 CHH DMRs), 10,013 (6862 CG, 2862 CHG and 323 CHH DMRs) and 12,891 DMRs (9886 CG, 2593 CHG and 412 CHH DMRs) were identified between Mep/Lep, Hep/Mep and Hep/Lep, respectively. Like DMPs, number of DMRs increased with increase in elevation difference (Fig. 4b). Further, CG-DMRs were mostly located in genic region (gene body) (75.6%–84.8%) followed by promoter (10%–12.2%) and intergenic region (5.09%–12.06%). The CHG-DMRs were also mostly located in genic region (51.3%–54.08%) but followed by intergenic (27.07%–28.3%) and promoter region (18.8%–20.2%) and CHH-DMRs were located mostly in promoter region (42.7%–44.4%) followed by genic (27.9%–34.05%) and intergenic region (23.2%–27.6%) (Fig. 4c). The CG- and CHH-DMRs were mostly located in protein coding genes followed by TEs while the CHG-DMRs were mostly present in TEs in all three comparisons (Fig. 4d). Like DMPs, DMRs of higher elevation population were hypo-methylated with respect to the corresponding lower elevation one. Notably, a higher proportion of all the DMRs was protein coding (64.2%), representing a total of 16,847 DMGs (differentially methylated genes including TEs) of which 1451 were common among the three populations (Fig. 4e, f). The maximum DMGs were found between the Lep and the Hep (6485).

#### Biological functions of DMGs in the three populations under native field condition

In order to find out the functional role of the differentially methylated genes (DMGs), gene ontology (GO) term enrichment analysis of biological process was performed using DAVIDGO (Table S7). Different classes of genes were enriched in the three different comparisons in addition to enrichment of some common genes those are mostly related to regulation of gene expression, positive regulation of macromolecule metabolic process etc. While the unique gene enrichment between the Hep/Lep included the genes related to developmental growth (137 genes), seed germination (30 genes), calcium ion

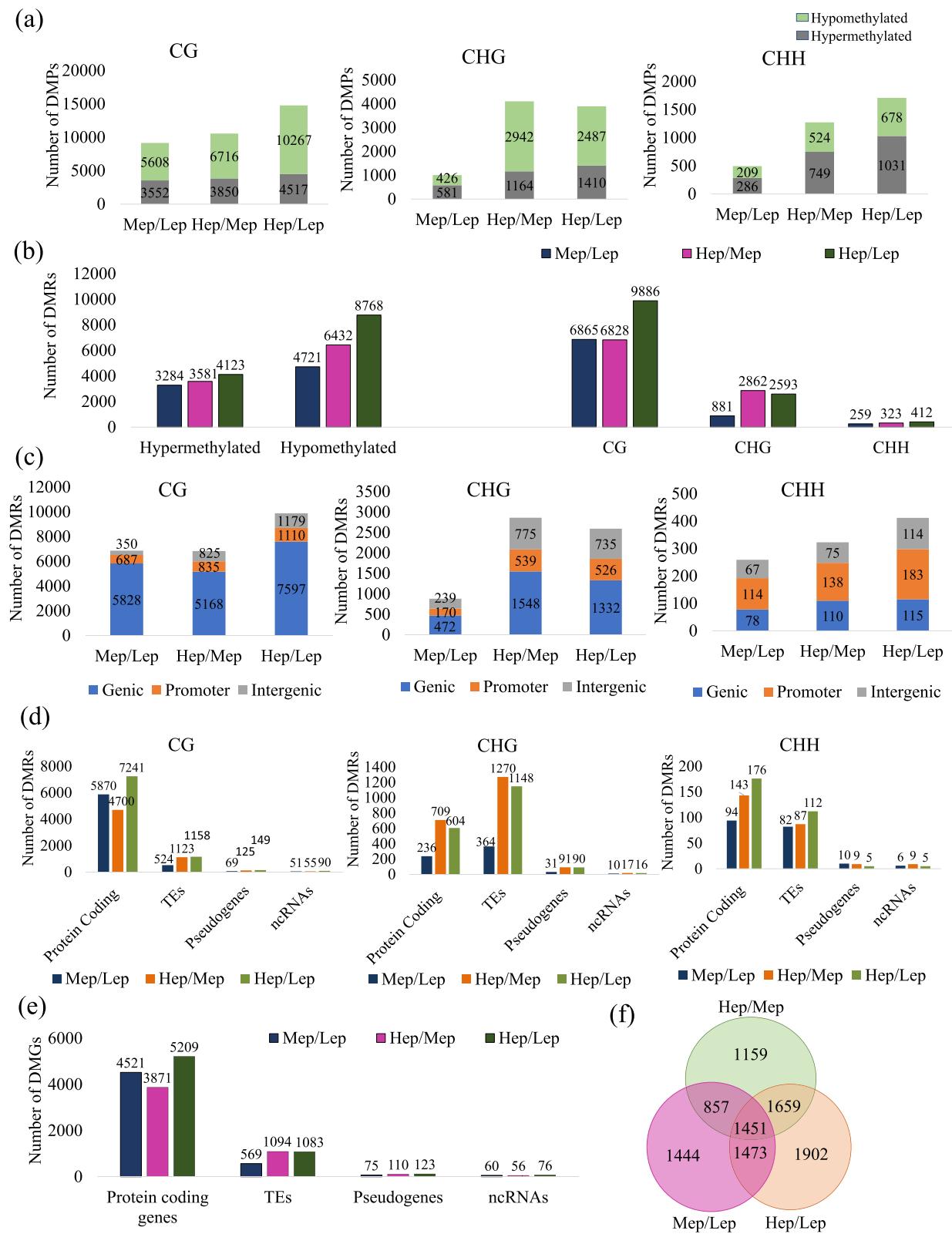
transport (17 genes), histone de-acetylation (13 genes), cellular response to hormone stimulus (4) etc., the Hep/Mep included genes related to intracellular signal transduction (41 genes), response to glucose (16 genes), response to salt (11 genes), regulation of DNA methylation (9 genes), etc. and the Mep/Lep exhibited enrichment in genes related to root development (115 genes), ER to Golgi vesicle-mediated transport (27 genes), chlorophyll biosynthetic process (22 genes), immune response (14 genes), starch biosynthesis process (13 genes) etc. (Table S7).

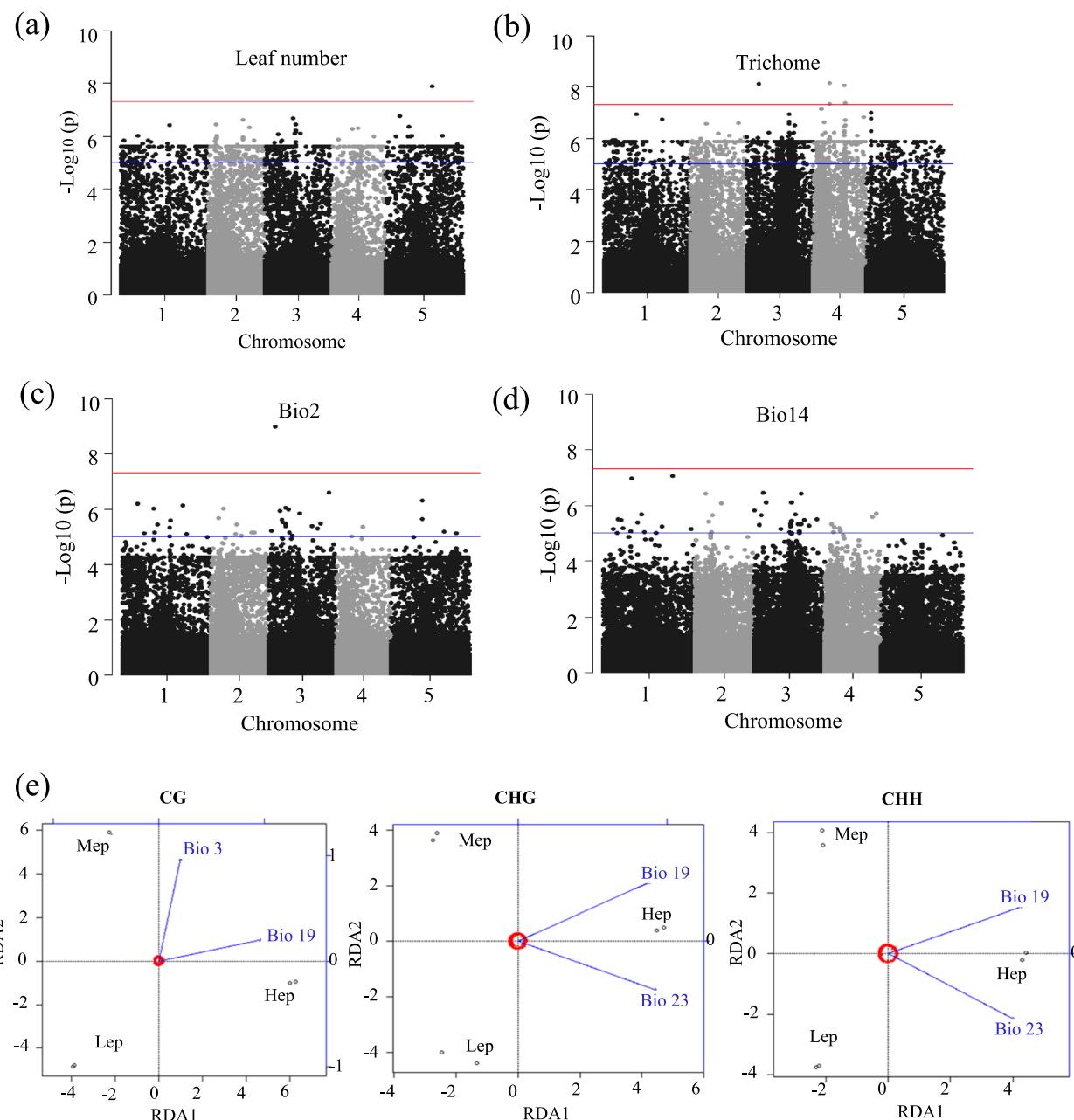
#### Phenotypic and climatic association with methylation

Association of phenotypes with differentially methylated positions (DMPs) under native field condition was examined using linear mixed model and a kinship matrix of SNPs as covariant to rule out contribution of genetic variations and corrected using multiple testing. A total of 36,134 mCs (CG-20072, CHG-1691, CHH-14371 mCs) mostly located in and near the genic regions were found to be associated with various phenotypes, including trichome density, stomatal density, number of stems, rosette area, number of leaves, pedicel of siliques, no of inflorescence and length of pedicel of flower (Table S4). The DMPs were associated with the genes governing different traits (Fig. 5a, b). For example, genes like *STICHEL*, *NEK6*, *ZWICHEL* known to control trichome density were hypo-methylated in the Hep as compared to the Lep and the Mep (Figure S4). It is to be noted here that trichome density of the Hep was higher than the Lep and the Mep (Singh et al., [93]). Similarly, genes those are known to govern leaf related traits like *ANGUSTIFOLIA*, *PHABULOSA*, *LONGIFOLIA*, *ALTERED MERISTERM PROGRAM 1*, *EIN2* were hypo-methylated in the Hep and the Mep compared to the Lep. This was again consistent with the morphological observation of variation in leaf number between the three populations. Similarly, the gene *ALTERED MERISTERM PROGRAM 1* governs multiple traits like increased rate of leaf initiation, photomorphogenesis, enhanced hypocotyl elongation and increased cytokinesis biosynthesis [63, 71] was hypo-methylated in high elevation population. These data suggest methylation status of these genes might influence the differential trait manifestations in the three populations.

(See figure on next page.)

**Fig. 4** Differentially methylated positions, regions, genes and their distribution in different comparisons of the native field samples. (a) Number of differentially methylated positions (DMPs) and percentage of hyper- or hypo-methylated sites, (b) differentially methylated regions (DMRs, hyper- and hypo-methylated) in three sequence contexts (CG, CHG, and CHH), (c) whole genome distribution of DMRs at three different genomic feature contexts. (d) Distribution of DMRs in different genomic regions. (e) Number of differentially methylated genes at different genomic feature contexts (protein coding genes, TEs, pseudogenes and non-coding RNA) and (f) number of common genes and unique differentially methylated genes annotated in the three populations

**Fig. 4** (See legend on previous page.)



**Fig. 5** Phenotypic and climatic association with methylation under native field condition. Manhattan plots showing epigenome-wide association with phenotypic traits, (a) leaf number, (b) trichome density, and climatic factors, (c) bioclimatic variable 2 (mean monthly temperature), (d) bioclimatic variable 14, (precipitation of driest month). Blue line indicates the significance cutoff line and red line indicates highly significantly associated sites. (e) Biplot for redundancy analysis conducted using differentially methylated sites and uncorrelated climatic variables. The analysis was performed separately for the three sequence contexts of methylation

Climatic association with methylation may provide cue for adaptive divergence of the populations. Following similar approach as phenotypic association, climatic association analysis showed 17 of the 23 analyzed climatic variables were associated with at least one mCs (Table S4). In CG-mCs context, a total of 2888 mCs

covering 1384 genes, 2244 mCs covering 1575 genes and 16 mCs representing 14 genes were found to be associated with precipitation of the driest month (bioclimatic variable 14), mean monthly temperature range (bioclimatic variable 2) and with temperature seasonality (bioclimatic variable 4), respectively. On the other hand,

CHH-mCs were mostly found to be associated with mean monthly temperature range (1795 CHH-mCs, covering 898 genes) followed radiation seasonality (7 genes). Genes associated with precipitation included among others, *SUPER SENSITIVE TO ABA* and *DROUGHT 2*, *RESPONSE REGULATOR 1*, which are known to be abiotic stress responsive (Fig. 5c, d) and were found to be un-methylated in high elevation population but methylated in low and medium elevation populations corresponding to prevailing precipitation regime in the areas. Further, redundancy analysis suggested, about 87.19% of the total variation in CG-mCs was correlated with the climate of the native sites ( $P\text{-value}=0.0667$ ). In CHG- and CHH- mCs, 88.89% and 92.17% of the variation were associated with the climate, respectively. Step-wise selection analysis showed 55.7% ( $P\text{-value}=0.005$ ) and 31.5% ( $P\text{-value}=0.047$ ) of the CG variation could be explained by bioclimatic variable 19 (precipitation of coldest month) and bioclimatic variable 3 (iso-thermality), respectively. On the other hand, Bioclimatic variable 23 (radiation seasonality) explained 60.6% of variation ( $P\text{-value}=0.011$ ) and bioclimatic variable 19 explained 28.3% of the variation in the CHG-mCs ( $P\text{-value}=0.029$ ) and 59.4% and 32.8% of the CHH-mCs variations were explained by bioclimatic variable 19 and 23 respectively (Fig. 5e). Thus, CG-mCs were mainly associated with precipitation and temperature and CHG- and CHH- mCs were associated with precipitation and radiation.

#### **Gain of differentially methylated cytosine between populations under controlled condition (F1)**

Methylation status is known to be affected by environmental condition and epigenetic memories. We evaluated the genome-wide comparative methylation pattern of three populations grown under controlled condition. There was gain of 0.71, 1.45 and 1.62-fold of DMPs in Mep/Lep, Hep/Mep and Hep/Lep, respectively under controlled condition over the native field condition. A total of 18,328, 39,158 and 53,471 DMPs corresponding 7910, 15,143 and 19,372 DMRs were identified between the Mep/Lep, Hep/Mep and Hep/Lep, respectively (Figure S5a, b).

The overall methylation pattern was similar as that of native condition including higher DMPs and DMRs as elevation increased. The maximum number of DMRs was observed between the Lep and Hep with hypo-methylation prevalent in the higher elevation population compared to the corresponding lower elevation one. However approximately equal proportion of hypo- and hyper-methylation was observed between the Lep and the Mep (Figure S5b, c). Although there was no change in the trends of the above features however, there were local level differential changes in methylation and thus giving

rise to different classes of DMGs as compared to the native condition (Figure S5d). Importantly, CHH-DMRs which are known to be influenced by environmental conditions was found to be much higher (1.28 times) than CHG- (0.73 times) and CG- DMGs (0.24 times) under control condition than native field condition (Figure S5). Interestingly, a few genes known to be involved in DNA methylation were found to be differentially expressed between the populations. For example, *ROS1* and *CMT2* were up-regulated in the Hep as compare to the Lep, and *SUVH9*, *FDM2* and *ROS1* were up-regulated in the Mep as compared to the Lep while *SUVH9* and *FDM2* were down-regulated in the Hep as compared to the Mep. Differential expression of these genes, might trigger differential methylation in the populations (Table S8). Moreover, as expected CHH-DMRs changes was the highest in the Hep/Lep than the other two comparisons and those genes were enriched in protein ubiquitination (14 genes), DNA integration (6 genes), plant cell wall type organization (6 genes), etc., (Table S9). Interestingly all the six genes that are known to be involved in plant cell wall type organization like *EXTENSIN* and flowering (*ARP6*) were hypo-methylated (CHH-context) in the Hep whereas, photo protection gene (*SPC1*) and *PYL1* known to involve in drought stress by ABA signaling were hyper-methylated in promoter region of the Hep (Table S10).

#### **Differentially methylated as well as differentially expressed genes are more between the Hep and the Lep**

Gene expression is known to be affected by both promoter and gene body methylation (GB). To examine the effect of methylation on gene expression, we analyzed only the common DMGs (both GB and promoter methylated genes) and DEGs ( $P<0.05$ ) of our previously published data by (Tyagi, et al., [107]) in the three populations under native field condition. Only 0.021% of total 16,847 DMGs, (135 Hep/Lep, 112 Hep/Mep, 101 Mep/Lep) were differentially expressed as well, (here after termed as DMDEGs, differentially methylated as well as differentially expressed genes) (Figure S6). Most of the DMDEGs were gene-body methylated and those were mostly hypo-methylated and down-regulated in the three populations. Interestingly the gene-body methylated up-regulated genes were also mostly hypo-methylated. The Hep/Lep exhibited the maximum number of genes having gene-body methylation and were hypo-methylated. These hypo-methylated genes were both up- and down-regulated. On the other hand, the promoter DMDEGs were mostly down-regulated and the proportion of hyper- and hypo-methylated genes were almost equal. However, the up-regulated genes were mostly hypo-methylated in all the three populations (Table S11).

The biological role of the DMDEGs of Hep/Lep were found to be mostly in abiotic stress related processes like response to water deprivation (16 genes), signal transduction (12 genes), response to cold (9 genes), circadian rhythm (8 genes), response to blue light (4 genes) etc. Among the water and light stress related genes, *AKR4C9* (AT2G3770) and *LNK1* (AT5G64170) were found to be hyper-methylated and up-regulated in the Hep as compare to the Lep. These genes are known to be involved in drought and high light stress condition, respectively [84, 87]. On the other hand, *MAX2* (AT2G42620) and *BZIP62* (AT1G19490) were hypo-methylated and up-regulated in the Hep as compare to the Lep. These genes are reported to be up-regulated in different abiotic stress conditions like drought and osmotic stress [13, 83]. Besides, *LHY* (AT1G01060) reported to be involved in circadian rhythm was found to be hypo-methylated and up-regulated in the Hep as compared to the Lep [57]. The prevalent of large number of DMDEGs related to abiotic stress between the Hep and the Lep was quiet expected, considering the prevalent climatic condition of high elevation zone dominated by high light intensity, less precipitation and large variation in day-night temperature etc. (Singh et al., [93]; Tyagi, et al., [107]).

On the other hand, unlike Hep/Lep, where climatic conditions were most contrasting, we did not find many such abiotic stress related DMDEGs between the other two comparisons (Mep/Lep and Hep-Mep), except a few. The DMDEGs of Mep/Lep were mostly involved in different biological processes like organic cycle compound biosynthetic process (9 genes), defense response to bacterium (9 genes), transmembrane transport (7 genes). Amongst the defense response genes, *AtWRKY52* (AT5G45260) and *FLS2* (AT5G46330) were hypo-methylated and down-regulated in the Mep as compared to the Lep. These genes were reported to play role in biotic stress condition [45, 112]. A few of the abiotic stress-related DMDEGs were *DIN11* (AT3G49620), *NDPK2* (AT5G63310) and *EGR1* (AT3G05640) which are reported to be up-regulated in dark and senescence [24], oxidative stress [46] and drought stress [56] respectively, were found to be hyper-methylated and up-regulated in the Mep as compared to the Lep.

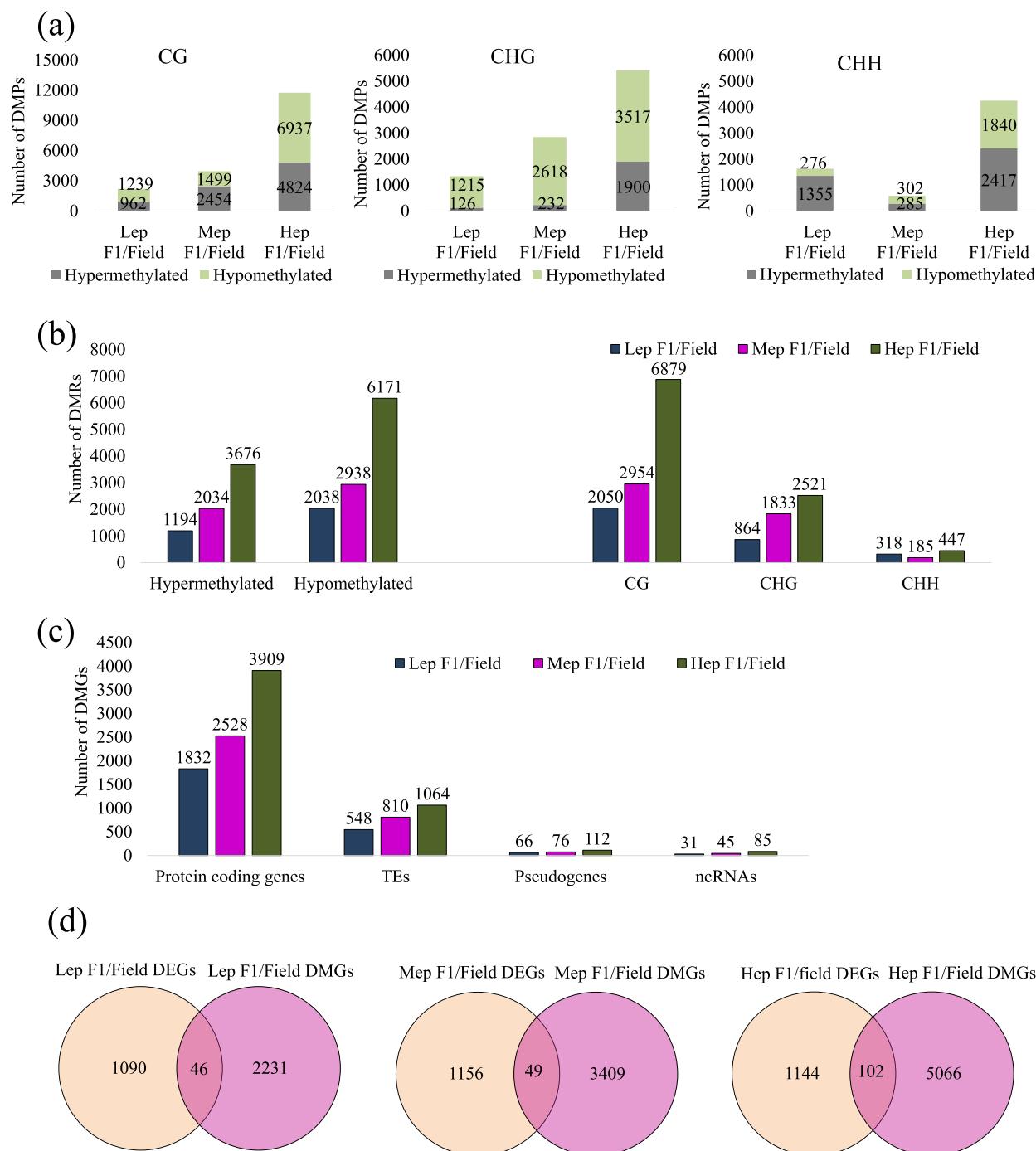
The DMDEGs of Hep/Mep were found to be mostly involved in different biological processes like starch catabolic process (3 genes), protein refolding (3 genes), glycogen catabolic process (2 genes) (Figure S7, Table S12). The *ISOMYLASE3* (AT4G09020) involves in starch degradation under chilling condition [68] and *DPE2* (AT2G40840) involves in starch degradation [85] were found to be hypo-methylated and down-regulated in the Hep as compared to the Mep. The DMDEGs *HSP70-2* (AT5G02490), *REV1* (AT5G44750) and *MAPKKK18*

(AT1G05100) were found to be hypo-methylated and up-regulated in the Hep as compared to the Mep. These genes are known to be up-regulated under abiotic stress including heat [104], UV stress [98] and drought condition (Zhao et al., [124]) respectively. Analysis of the DMDEGs of the three populations grown under control conditions retrieved similar trend as that of native conditions (Figure S8, Table S13).

#### Methylation plasticity is more in the high elevation population in new condition

To examine the plasticity of methylation we compared the genome-wide methylation between native field and control condition grown plants of the same population. Around 97.0%—99.5% mCs were passed on to the control condition grown plants from the native field plants. There was both gain or loss of methylation in the control grown plants. A loss of 0.17%, 0.12% and 0.80% and a gain of 0.14%, 0.32%, and 0.59% mCs was observed in the Lep, Mep and Hep respectively, clearly indicating both loss and gain of mCs were more in the Hep. A total of 5173, 7390 and 21,435 DMPs corresponding to 3232, 4972 and 9847 DMRs and 2477, 3458, 5170 DMGs were found between the native and the control plants of the Lep, Mep and Hep, respectively (Fig. 6a-c). As expected, the DMPs, DMRs and DMGs were higher in the Hep than the other two populations. Further, both the CG-DMRs and CHG-DMRs followed an increasing trend with an increase in the elevation (Fig. 6b). Interestingly, a higher proportion of DMRs were found to be hypo-methylated (66.44%, 58.4%, and 60.9%, in the Lep, Mep and Hep, respectively), although it didn't follow a particular trend with the elevation. Notably, on an average 69.7%-72.1% DMGs were associated with protein coding genes followed by transposable elements (23.6%-26.3%), pseudogenes (2.6%-2.9%) and non-coding RNA (1.3%-1.4%) (Fig. 6c).

There was, on an average only 0.012% of DMGs between native field and controlled plants those were DMDEGs (197/15284) in the three populations. The maximum DMDEGs were in the Hep (102) followed by the Lep (46) and the Mep (49) (Fig. 6d). Interestingly, a significant proportion of the DMDEGs exhibited gene-body methylation (86.3%, 91.3% and 80.8%, respectively) which were mostly hypo-methylated (64.6%, 64.28% and 57.8% respectively) and were both up-and down-regulated. However, proportion of up-regulated genes were more in the Hep (67.9%) than the Lep (48.14%) and the Mep (42.8%). Similarly, the promoter methylated DMDEGs were mostly hypo-methylated and associated genes were both up- and down-regulated (Table S14). Gene ontology analysis of DMDEGs of the three populations suggested the maximum gene enrichment was observed in the Hep



**Fig. 6** Differentially methylation between native field vs controlled conditions of the respective population. **(a)** Number of differentially methylated positions (DMPs) (hyper- and hypomethylated), **(b)** differentially methylated regions (DMRs) (hyper- and hypomethylated) at different sequence context, **(c)** number of differentially methylated genes at different genomic contexts (protein coding genes, TEs, pseudogenes and non-coding RNA), **(d)** common differentially methylated and differentially expressed genes between field vs control grown plants, F1

and those were mostly related to protein phosphorylation (9 genes), transmembrane transport (8 genes), response to red or far red light (7 genes), response to cold (7 genes) etc. and in the Mep enrichment was in the genes related

to response to temperature stimulus (4 genes), small molecule biosynthetic process (3 genes) and in the Lep enrichment was in the genes related to defense response

(5 genes) circadian rhythm (3), and pentose phosphate shunt (2 genes) etc. (Table S15).

#### Trans-generational methylation pattern of the populations

To determine the transgenerational methylation pattern, we sequenced methylome of the three populations after growing them for six generations (F6) under control conditions and compared with the first generation (F1, plants grown from seeds of native field condition)). On an average 99.8% of the mCs was stably passed on from F1 to F6 under the controlled condition in the three populations and it was maximum in the Lep (99.81) followed by the Hep (99.80) and the Mep (99.79). Further, contrary to native field vs F1, the proportion of hyper-methylated states was higher than hypo-methylated states in all the three populations over generations, suggesting a gain of methylation in the F6 as compared to F1 (Fig. 7).

Next, we were interested to examine the DMPs between F1 and F6 arose due to loss or gain of methylation over generations. Around 1892, 1402 and 664 DMPs corresponding to 457, 447 and 248 DMRs and 451, 327 and 179 DMGs were identified in the Lep, Mep and Hep, respectively (Fig. 7a, b, c). Importantly, only 6.3% DMPs (0.3% CG, 2.4% CHG and 3.6% CHH) between F1 and F6 was common among the three populations suggesting a majority of DMPs were genotype dependent (Fig. 7d). GO analysis of the genes having DMPs between F6/F1 of the Lep suggested that they are mostly involved in different biological processes (33) and DNA degradation (3). For example, Subtilisin-like serine endopeptidase family protein (AT5G58820) play a role in abiotic stress condition [97] and *BCHC1* (AT3G60920) involves in biotic stress condition [102] both were hyper-methylated at CHH-context in the promoter region in F6 generation as compared to F1 of Lep. On the other hand, *YSL7* (AT1G48370) reported to be involved in biotic stress [33] was found to be hypo-methylated at CHH-context in the promoter region in F6 generation as compared to the F1 of the Lep.

The genes having DMPs between F6/F1 of Mep are mainly involved in biological process (39) and DNA integration (3) etc. A few of them were *MSIL4* (AT3G07810) known to be involved in cell wall formation [41] was hypo-methylated in genic CG- context of F6 generation as compare to the F1 generation of the Mep and *ATDOA5* (AT1G68960) plays role in potassium stress [3] was hyper-methylated in genic CHH-context of F6 generation as compare to the F1 generation of the Mep. Only a few of the genes having DMPs between F6/F1 of Hep were identified and these are involved in cellulose catabolic process (2). For example, *GLYCOSYL HYDROLASE 9C3* (AT4G11050) involves in cellulose catabolic process [108] was hyper-methylated in CHH-context

located in promoter region of F6 generation as compared to the F1 generation (Breeanna R. Urbanowicz, 2007). On the other hand, *GLYCOSYL HYDROLASE 9B14* (AT4G09740) which is known to be involved in cellulose catabolic processes [108] was hyper-methylated in genic region CG-context of F6 generation as compared to the F1 generation (Table S16).

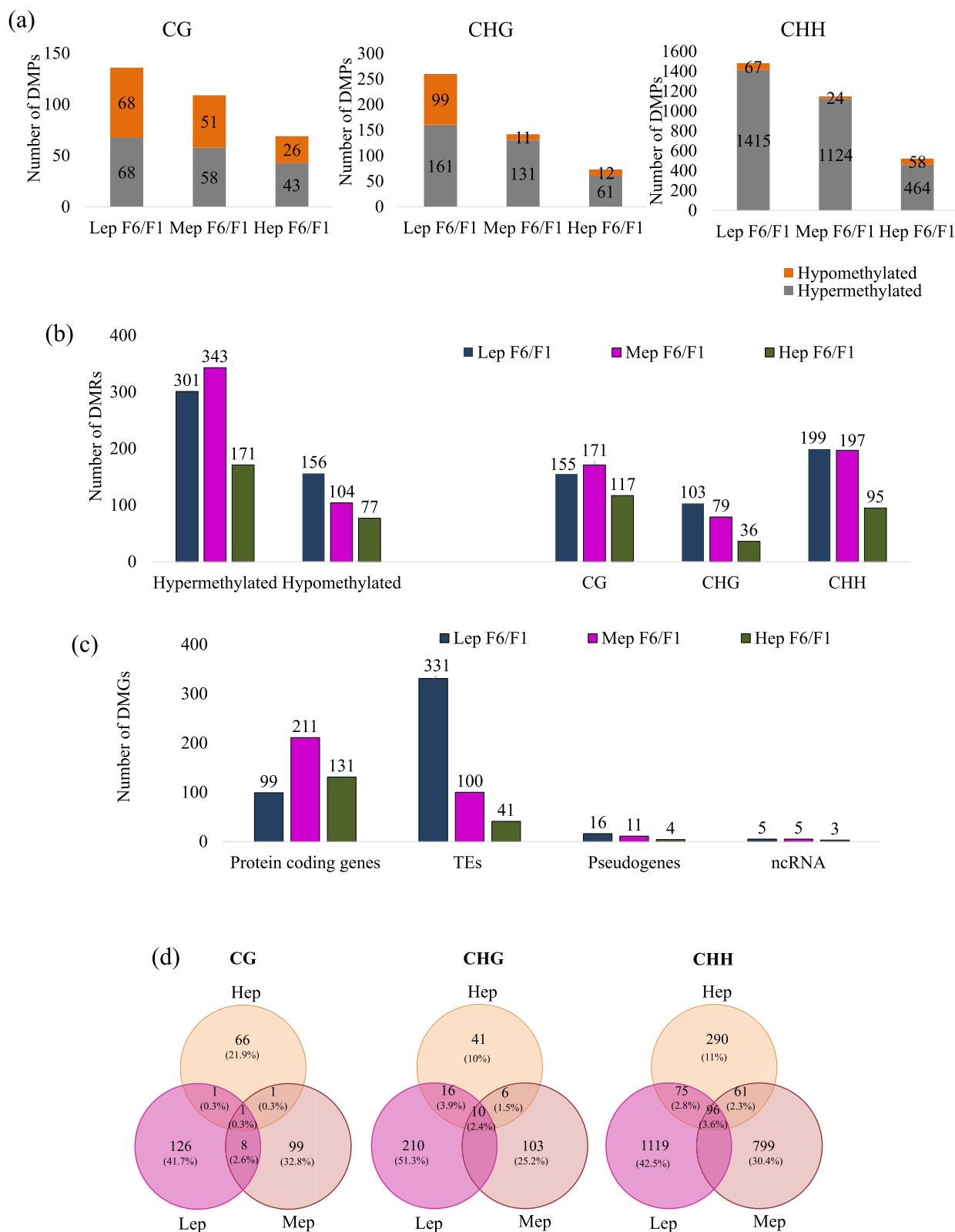
A correlation based hierarchical clustering analysis based on methylation states of common cytosine in all the three populations of native field, F1 and F6 generations showed that in CG context, besides the three population being clustered independently as observed earlier, the Hep made two independent cluster of native fields and F1/F6 but the Lep and the Mep didn't exhibit such clustering pattern (Figure S9). However, CHG- and CHH-mCs clearly differentiate the native field and F1/ F6 of the three populations. This data presumably suggests CHH and CHG methylation is mostly environment driven and genome-wide transgenerational methylation under control condition is insignificant to distinguish between generations.

#### Validation of methylation regulated gene expression

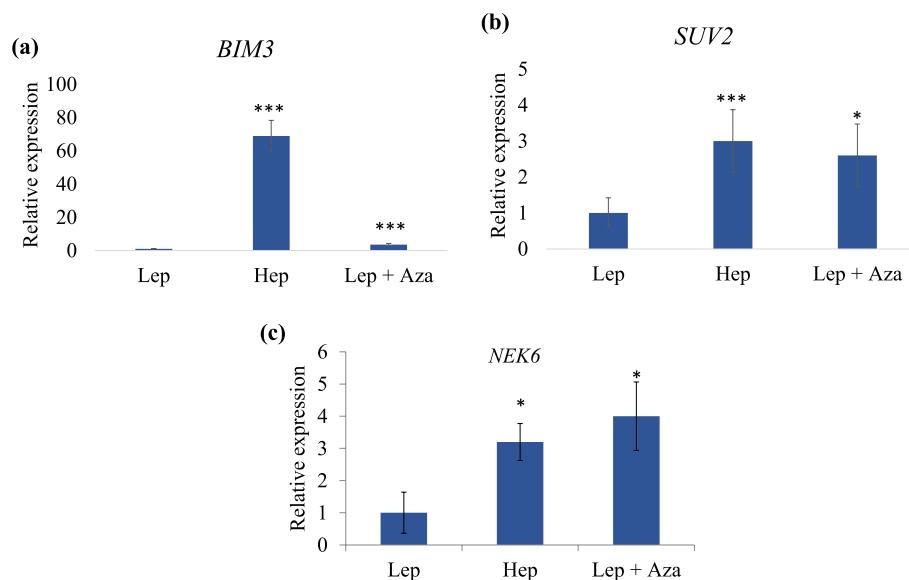
Finally, we validated the expression pattern of a few randomly selected differentially methylated genes after treatment with 5-azacytidine, a global demethylating agent. It was expected that hyper-methylated promoter linked gene will exhibit higher expression level after 5-azacytidine treatment. Accordingly, the *BIM3* known to play a critical role in plant growth and development which was hypo- and hyper-methylated in the promoter region of the Hep and the Lep respectively, exhibited significantly higher expression in the Lep after treatment, while its expression was higher in the Hep before treatment. Similarly, *SUV2* was hypo-methylated at the 3'- downstream in the Hep and showed higher expression as compared to the hyper-methylated Lep, but expressed higher level in the Lep after treatment. This gene is known to play a role in DNA damage repair and in response to UV. The gene body hyper-methylated *NEK6* of the Lep which otherwise showed lower expression as compared the hypo-methylated Hep, showed higher level of expression after 5-azacytidine treatment (Fig. 8, Figure S10).

#### Discussion

Epigenetic variations including DNA methylation is an alternative strategy for survival adopted by plants, particularly under challenging environmental conditions. Inspite of exhibiting more or less similar mC frequency, the two populations had a large number of DMPs suggesting rearrangement in mC pattern via gain or loss in methylation at population level probably to better cope with prevalent environmental conditions. The magnitude



**Fig. 7** Transgenerational methylation pattern under controlled condition. **(a)** Number of differentially methylated positions (DMPs), **(b)** differentially methylated regions (DMRs) and **(c)** differentially methylated genes between F1 and F6. **(d)** Also shown the common and unique DMPs among the three populations over the generations



**Fig. 8** Validation of effect of methylation on gene expression at three different genomic regions between the Lep and Hep under controlled glass house condition. Expression levels of differentially methylated genes ((a) *BIM3*, (b) *SUV2* and (c) *NEK6*) as estimated by qRT-PCR in the Lep and in the Hep after treatment with DNA methyltransferase inhibitor (5-azacytidine (Aza)). Vertical bars represent standard deviation (SD) of the mean of three biological and two technical replicates. \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ ; ns, not significant

of this molecular remodeling was the maximum between the highest and the lowest elevation populations where, the climatic and topographical variability were the most prominent. Moreover, the gradual increase in DMPs, except CHH-DMPs along the elevation, particularly of CG mC is of interest as it may control gene expression in response to environmental variability. Similarly, increase in CG mC divergence along elevation was reported in strawberry [17]. Although variations in methylation in different populations have primarily been linked to underlying genetic differences [91] but growing body of literatures have also shown methylation variations exclusively driven by environment [19, 25, 26, 60, 62, 74, 82, 88, 115]. In *Fragaria vesca*, Sammarco et al. [88] showed that DNA methylation enables clonal plant populations to adapt to environmental conditions independent of genetic changes. Another report by Lira-Medeiros et al. [60] suggested that the two populations of mangrove originating from contrasting natural environment possess divergent epigenetic profile with similar genetic background that were characteristic of the population in a particular environment [60]. Similarly, the genetically identical orchid accessions collected from different regions also showed variation in DNA methylation that might be the cause of ecological differentiation among them [74]. Besides, the geographic origin of the accessions has also been shown to be a major predictor of global methylation pattern [43]. Although it was not possible to completely detangle genetic contribution

to methylation in our study however, by eliminating genome-wide SNPs the effect could be minimized. Nevertheless, the prevalent climatic differences between the sites might have shaped differential methylation of the populations aiding in better adaptability.

The individual DMPs however, have been suggested to contribute little to gene expression [75, 103], while a differentially methylated genomic regions provide relevant biological information. Our data, in this context, suggested trend of increasing DMRs/DMGs along the elevation as observed with DMPs. Importantly, higher enrichment of CG-mCs in genic regions and protein coding genes along the elevation underscore the biological significance of the gene body methylation, particularly in *A. thaliana* where only approximately 5% of the genes are methylated in promoter regions [123]. While methylation in promoter predominantly inhibit transcriptional activity barring a few where it promotes transcription [53, 116], the functional relevance of gene-body methylation remains enigmatic [9, 126]. However, gene-body methylation driven higher expression [122], Zilberman et al., [125]; [99] to no effect on gene expression [7, 8] and even low level of gene expression [64], Zilberman et al., [125]; [67] has been reported. Nevertheless, correlation analysis of global DNA methylation with our previous global gene expression data of the same populations suggested there was no correlation between global gene expression and gene body methylation ( $R^2 = 0.0025$ ,  $P = 0.82$ ). However, a few differentially methylated (mostly hypo-methylated

gene-body) genes were both up- and down-regulated, albeit the maximum differentiation was between the low and the high elevation populations. The observed increasing trend of DMRs/DMGs particularly of higher enrichment of CG-mCs with elevation presumably suggest differential regulation of gene expression via methylation along the elevation. Indeed, various unique gene ontology terms were identified in each comparison for example, photo inhibition, red or far-red light signaling pathway, photosystem I etc., between low and high elevation population. Notable here, the high elevation zone is dominated by high light intensity, UV, low precipitation and extreme day-night temperature fluctuation etc. (Singh et al., [93]; Tyagi, Singh, et al., [106]). Moreover, the common differentially methylated and differentially expressed genes (DMDEGs) were also different between populations and the high and the low elevation populations showed enrichment of abiotic stress related genes like response to water deprivation (16 genes), signal transduction (12 genes), response to cold (9 genes), circadian rhythm (8 genes), response to blue light (4 genes), etc. It intuitively suggests the differential gene-body methylation may contribute to better adaptability under extreme environmental condition.

There were a few interesting observations: i) there was net gain in DMPs between the populations when grown under controlled condition, ii) methylation heritability varied in the three populations and the heritability was less from native field to F1 (environment driven heritability) as compared to F1 to F6 plants grown under controlled condition iii) the environment driven gain and loss of methylation was more in the high elevation population and iv) Most of the protein coding genes are hypo-methylated and are down-regulated in the high elevation population as compared to the low elevation population under native condition. These observations intuitively suggest if the hyper-methylation status under benign environment in controlled condition is a fine-tuning mechanism via switching-on/off of some cellular processes not required under benign condition but under harsh environmental condition. Thus, genes related to light, cold and osmotic stresses were hyper-methylated under control condition over the generations. This was further supported by the fact that the average changes in CHH-DMGs which are known to be influenced by environmental conditions was found to be much higher under control condition than native field condition. Thus, after initial fine tuning there is overall minimum adjustments in the global methylation status over generations under controlled condition. This was further exemplified by the fact that gain and loss in methylation was more in the field to F1 plants than F1 to F6.

Conversely, the hypo-methylation status may be a strategy adopted by high elevation population to cope with such conditions. The hypo-methylation status has often been associated with different environmental stressors like drought [100], high light [101], salt tolerance [111]. Indeed, the hypo-methylated genes of the high elevation populations were enriched in plant growth and development including root and shoot development, flower development, auxin response, light response, leaf senescence processes etc. (Table S7). For example, *ALTERED MERISTEM PROGRAM 1* which governs multiple traits like increased rate of leaf initiation, photomorphogenesis, enhanced hypocotyl elongation and increased cytokinesis biosynthesis, *AUXIN RESPONSE FACTOR (ARF1 and ARF2)*, *ETHYLENE INSENSITIVE 2 (EIN2)*, *RESPONSE REGULATOR 2 (RR2)*, those play critical role in leaf senescence and plant growth etc., were found to be hypo-methylated with reduced expression level in high elevation population. On the contrary genes like *SPA1-RELATED 3 (SPA3)* which is responsible for suppression of photo morphogenesis in seedlings and elongation growth in adult plants and *ANGUSTIFOLIA* controlling polar cell expansion in leaf width direction were hypo-methylated but showed higher expression in high altitude population. Notably, previously this high elevation population was reported to exhibit reduced growth and biomass compared to the low elevation one (Singh et al., [93]). The loss of methylation has been shown to be associated with increased transcriptional regulation (through alternate splicing) and better adaptability and phenotypic plasticity [38, 113] and attributed primarily to a consequence of variation in expression of methylating enzymes. For instance, cold stress was found to decrease the expression of methyltransferases in *Zea mays* that was correlated with decreased methylation [95]. Similarly, genome-wide hypo-methylation was observed in the anthers of *Gossypium hirsutum* upon exposure to heat that was due to decreased expression of DNA methyltransferases (*DRM1* and *DRM3*) [65]. Thus, the hypo-methylation of the high elevation populations may be an adaptation strategy to counter harsh environmental stresses prevailed in those areas. Indeed, the genes which were both differentially methylated and differentially expressed were mostly related to abiotic stress like cold, water deprivation etc. (Table S7). Further, the observation that most of the gain or loss of methylation occurred in first generation when plants were grown first time at common condition might indicate maximum changes in methylation contexts occur in the first encounter of new environment. Only a few loss or gain occurred over the generations up to F6, both in CHH-context of promoter region, especially in the Lep and CG-contexts mostly in Mep. It remains to be seen how the loss or gain

in methylation in these genes affect gene expression and advantages provided by them. Nevertheless, the loss and gain in important genes like *YSL7*, *BCHC1*, *ATDOA5* etc., those are reported to be involved in biotic and abiotic stresses and cellulose catabolism like *GLYCOSYL HYDROLASE 9C3* and *GLYCOSYL HYDROLASE 9B14* might indicate tight regulation of these gene via methylation and/or demethylation depending on the environmental condition.

## Conclusions

The populations of *A. thaliana* adapted at different climatic conditions were significantly differentially methylated both under native and controlled condition. However, the magnitude and direction were most significant between the low and the high elevation populations. More number of protein coding genes were hypo-methylated in the high elevation population than the low elevation one under native field condition. This may indicate evolutionary significance of methylation in adaptation of the high elevation population. Although the observed variations in methylation may partly be driven by genetic differences of the populations but spontaneous local variations in DNA methylation and their heritability was significant and might have evolved independently to cope with the respective prevalent condition along the elevation.

## Abbreviations

mCs	Methyl Cytosines
Lep	Low Elevation Population
Mep	Medium Elevation Population
Hep	High Elevation Population
MET1	METHYLTRANSFERASE 1
CMT2	CHROMOMETHYLASE 2
CMT3	CHROMOMETHYLASE 3
DRM2	DOMAINS REARRANGED METHYLTRANSFERASE 2
DMPs	Differentially Methylated Positions
DMRs	Differentially Methylated Regions
DMGs	Differentially Methylated Genes
DMDEGs	Differentially Methylated and Differentially Expressed Genes
m amsl	Meter Above Mean Sea Level
TEs	Transposable Elements
GO	Gene Ontology

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12870-024-05641-0>.

Additional file 1: Supplementary Figure S1. (a) Flow chart showing method followed for processing of bisulfite converted sequenced reads and (b) identification of differentially methylated positions (DMPs) and differentially methylated regions (DMRs). Supplementary Figure S2. (a) Global distribution of mC patterns in CG-, CHG- and CHH-sequence contexts vis-a-vis-genomic features contexts of genic, promoter and intergenic in the three populations form the native field condition, (b) Distribution pattern of genic mCs in protein coding and transposable element region. Supplementary Figure S3. Dendograms showing correlation based clustering based on common mCs of the three populations under native

field condition. Hep, Mep and Lep clustered separately on CG and CHG context while showing mixed clustering on CHH context. Supplementary Figure S4. Methylation status of trichome development related genes in Lep, Mep and Hep and trichome density under native field condition. Supplementary Figure S5. (a) Bar chart showing number of differentially methylated positions (DMPs) (hyper- and hypomethylated), (b) differentially methylated regions (DMRs) (hyper- and hypomethylated), (c) whole genome distribution of DMRs, and (d) number of differentially methylated genes at different genomic contexts (protein coding genes, TEs, pseudogenes and non-coding RNA). Supplementary Figure S6. Venn diagram showing common differentially methylated and differentially expressed genes between the populations under native field condition. Supplementary Figure S7. Bar diagram shows the gene ontology enrichment term of commonly differentially methylated as well as differentially expressed genes between the populations ((a) Mep/Lep, (b) Hep/Mep, and (c) Hep/ Lep) under native field condition. Supplementary Figure S8. Venn diagram showing common differentially methylated and differentially expressed genes between the populations grown under common controlled condition (F1 generation). Supplementary Figure S9. Correlation based clustering of the accessions of the three populations on the basis of methylation status of common cytosine positions among field, F1 and F6 samples. Supplementary Figure S10. Methylation status of differentially methylated genes BIM3 (a), SUV2 (b), and NEK6 (c). The region represents 100 bp differentially methylated region between the Hep and the Lep under controlled grown plants. Supplementary Table S1. Details of the sample collection sites: The geographical co-ordinates (degree°-min'-sec"), altitude (Meters Above Mean Sea Level) and growing season of the populations. Supplementary Table S2. Climatic variables of the three sample collection sites in west Himalayan region Supplementary Table S3. Summary statistics of bisulfite sequencing of DNA methylation data. Supplementary Table S4. Number of methylated cytosine associated with the significantly variable traits under native field conditions as identified by the epi-genome-wide association analysis. Supplementary Table S5. Primer used in this study for the validation of DNA methylation regulated gene expression. Supplementary Table S6. Whole genome cytosine methylation statistics for three populations (average of two individuals per population  $\pm$  standard deviation (SD)) under native field condition. Supplementary Table S7. Gene ontology term enrichment of biological process of DMG under native field condition. Supplementary Table S8. Differentially expressed genes involved in DNA methylation and demethylation between three population under common glasshouse condition. Supplementary Table S9. Gene ontology term enrichment of biological process of DMGs of CHH context between Hep and Lep under controlled growth condition. Supplementary Table S10. Differentially methylated genes between Hep and Lep CHH context of genic region and promoter region. Supplementary Table S11. Common DMGs and DEGs and their methylation and expression status between the populations under native field condition. Supplementary Table S12. Gene ontology term enrichment of biological process of differentially methylated as well as differentially expressed genes between the populations under native field condition. Supplementary Table S13. Gene ontology term enrichment of biological process of differentially methylated as well as differentially expressed genes between the populations under controlled growth condition in f1 generation. Supplementary Table S14. Common DMG and DEGs their methylation and expression between Field vs F1 condition. Supplementary Table S15. Gene ontology term enrichment of biological process of differentially methylated as well as differentially expressed genes between field and F1 condition of the three populations. Supplementary Table S16. Gene ontology term enrichment of biological process of differentially methylated genes between F1 and F6 condition of the three populations.

## Acknowledgements

The authors acknowledge the financial support by the Department of Biotechnology, Government of India, New Delhi (BT/PR11963/BPA/118/9/2014) and Council of Scientific and Industrial Research, New Delhi. AS acknowledges Council of Scientific and Industrial Research, New Delhi, AKV and SK acknowledges the University Grant Commission (UGC), New Delhi for providing fellowship.

## Authors' contributions

AS conducted experiments and analyzed the sequence data with AKV, SK and SKB. SR designed experiments, critically analyzed the data and prepared the manuscript with AKV. All the authors read and approved the manuscript for publication.

## Funding

This research work was supported by the Department of Biotechnology, Government of India, New Delhi (BT/PR11963/BPA/118/9/2014) and Council of Scientific and Industrial Research, New Delhi.

## Availability of data and materials

The sequencing data are available at NCBI Sequence Read Archive under Bio project: PRJNA1072841. The link of NCBI Sequence Read Archive is <https://www.ncbi.nlm.nih.gov/sra>.

## Declarations

### Ethics approval and consent to participate

Not applicable.

### Consent for publication

All authors read the MS and agreed to publish.

### Competing interests

The authors declare no competing interests.

### Author details

<sup>1</sup>Plant Molecular Biology and Biotechnology Division, CSIR-National Botanical Research Institute, Lucknow 226001, India. <sup>2</sup>Department of Plant Sciences, Central University of Hyderabad, Hyderabad, Telangana, India. <sup>3</sup>Academy of Scientific and Innovative Research (AcSIR), Ghaziabad 201002, India.

<sup>4</sup>Computational Biology, CSIR-National Botanical Research Institute, Lucknow 226001, India.

Received: 14 June 2024 Accepted: 26 September 2024

Published online: 09 October 2024

## References

- Akalin A, et al. methylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles. *Genome Biol.* 2012;13(10):1–9.
- Alonso C, Ramos-Cruz D, Becker C. The role of plant epigenetics in biotic interactions. *New Phytol.* 2019;221(2):731–7.
- Armengaud P, Breitling R, Amtmann A. Coronatine-insensitive 1 (COI1) mediates transcriptional responses of *Arabidopsis thaliana* to external potassium supply. *Mol Plant.* 2010;3(2):390–405.
- Banta JA, Richards CL. Quantitative epigenetics and evolution. *Heredity.* 2018;121(3):210–24.
- Becker C, et al. Spontaneous epigenetic variation in the *Arabidopsis thaliana* methylome. *Nature.* 2011;480(7376):245–9.
- Bergelson J, Roux F. Towards identifying genes underlying ecologically relevant traits in *Arabidopsis thaliana*. *Nat Rev Genet.* 2010;11(12):867–79.
- Bewick AJ, et al. On the origin and evolutionary consequences of gene body DNA methylation. *Proc Natl Acad Sci.* 2016;113(32):9111–6.
- Bewick AJ, et al. 'Evolutionary and experimental loss of gene body methylation and its consequence to gene expression', *G3: Genes. Genomes, Genetics.* 2019;9(8):2441–5.
- Bewick AJ, Schmitz RJ. Gene body DNA methylation in plants. *Curr Opin Plant Biol.* 2017;36:103–10.
- Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics.* 2014;30(15):2114–20.
- Boyes DC, et al. Growth stage-based phenotypic analysis of *Arabidopsis*: a model for high throughput functional genomics in plants. *Plant Cell.* 2001;13(7):1499–510.
- Brachi B, Morris GP, Borevitz JO. Genome-wide association studies in plants: the missing heritability is in the field. *Genome Biol.* 2011;12(10):1–8.
- Bu Q, et al. Regulation of drought tolerance by the F-box protein MAX2 in *Arabidopsis*. *Plant Physiol.* 2014;164(1):424–39.
- Burrows MT, et al. Geographical limits to species-range shifts are suggested by climate velocity. *Nature.* 2014;507(7493):492–5.
- Cao X, Jacobsen SE. Role of the *Arabidopsis* DRM methyltransferases in de novo DNA methylation and gene silencing. *Curr Biol.* 2002;12(13):1138–44.
- Dabney A, Storey JD, Warnes G.R. (2010) 'qvalue: Q-value estimation for false discovery rate control', *R package version*, 1(0).
- De Kort H, et al. Ecological divergence of wild strawberry DNA methylation patterns at distinct spatial scales. *Mol Ecol.* 2020;29(24):4871–81. >
- Dubin, M.J. et al. (2015) 'DNA methylation in *Arabidopsis* has a genetic basis and shows evidence of local adaptation', *elife*, 4, p. e05255.
- Eichten SR, Springer NM. Minimal evidence for consistent changes in maize DNA methylation patterns following environmental stress. *Front Plant Sci.* 2015;6:308.
- Feng SJ, et al. Variation of DNA methylation patterns associated with gene expression in rice (*Oryza sativa*) exposed to cadmium. *Plant, Cell Environ.* 2016;39(12):2629–49.
- Finnegan EJ, et al. DNA methylation in plants. *Annu Rev Plant Biol.* 1998;49(1):223–47.
- Finnegan EJ, and Kovac, K.A. (2000) 'Plant DNA methyltransferases', *Plant gene silencing*, pp. 69–81.
- Foust CM, et al. Genetic and epigenetic differences associated with environmental gradients in replicate populations of two salt marsh perennials. *Mol Ecol.* 2016;25(8):1639–52.
- Fujiki Y, et al. Dark-inducible genes from *Arabidopsis thaliana* are associated with leaf senescence and repressed by sugars. *Physiol Plant.* 2001;111(3):345–52.
- Galanti D, et al. Genetic and environmental drivers of large-scale epigenetic variation in *Thlaspi arvense*. *PLoS Genet.* 2022;18(10):e1010452.
- Ganguly DR, et al. The *Arabidopsis* DNA methylome is stable under transgenerational drought stress. *Plant Physiol.* 2017;175(4):1893–912.
- Garg R, et al. Divergent DNA methylation patterns associated with gene expression in rice cultivars with contrasting drought and salinity stress response. *Sci Rep.* 2015;5(1):14922.
- Gáspár B, Bossdorf O, Durka W. Structure, stability and ecological significance of natural epigenetic variation: a large-scale survey in *Plantago lanceolata*. *New Phytol.* 2019;221(3):1585–96.
- Gienapp P, et al. Genomic quantitative genetics to study evolution in the wild. *Trends Ecol Evol.* 2017;32(12):897–908.
- Gugger PF, et al. Species-wide patterns of DNA methylation variation in *Quercus lobata* and their association with climate gradients. *Mol Ecol.* 2016;25(8):1665–80.
- He Y, Li Z. Epigenetic environmental memories in plants: establishment, maintenance, and reprogramming. *Trends Genet.* 2018;34(11):856–66.
- Herrera CM, Bazaga P. Epigenetic differentiation and relationship to adaptive genetic divergence in discrete populations of the violet *Viola cazelensis*. *New Phytol.* 2010;187(3):867–76.
- Hofstetter SS, et al. *Arabidopsis* YELLOW STRIPE-LIKE7 (YSL7) and YSL8 transporters mediate uptake of *Pseudomonas* virulence factor syringolin A into plant cells. *Mol Plant Microbe Interact.* 2013;26(11):1302–11.
- Huang DW, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* 2009;37(1):1–13.
- Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc.* 2009;4(1):44–57.
- Huang J, et al. Roles of DNA methyltransferases in *Arabidopsis* development. *Afr J Biotech.* 2010;9(50):8506–14.
- Jablonska E, Raz G. Transgenerational epigenetic inheritance: prevalence, mechanisms, and implications for the study of heredity and evolution. *Q Rev Biol.* 2009;84(2):131–76.
- Jabre I, et al. Does co-transcriptional regulation of alternative splicing mediate plant stress responses? *Nucleic Acids Res.* 2019;47(6):2716–26.
- Jeremias G, et al. Synthesizing the role of epigenetics in the response and adaptation of species to climate change in freshwater ecosystems. *Mol Ecol.* 2018;27(13):2790–806.

40. Johannes F, Colot V, Jansen RC. Epigenome dynamics: a quantitative genetics perspective. *Nat Rev Genet.* 2008;9(11):883–90.
41. Kairouani, A. et al. (2023) 'Cell-type-specific control of secondary cell wall formation by Musashi-type translational regulators in *Arabidopsis*', *Elife*, 12, p. RP88207.
42. Kang HM, et al. Variance component model to account for sample structure in genome-wide association studies. *Nat Genet.* 2010;42(4):348–54.
43. Kawakatsu T, et al. Epigenomic diversity in a global collection of *Arabidopsis thaliana* accessions. *Cell.* 2016;166(2):492–505.
44. Keller TE, Lasky JR, Yi SV. The multivariate association between genome-wide DNA methylation and climate across the range of *Arabidopsis thaliana*. *Mol Ecol.* 2016;25(8):1823–37.
45. Kim H, et al. Molecular basis for the interference of the *Arabidopsis* WRKY54-mediated immune response by two sequence-unrelated bacterial effectors. *Plant J.* 2024;118(3):839–55.
46. Kim Y, et al. Transgenic poplar expressing *Arabidopsis* NDK2 enhances growth as well as oxidative stress tolerance. *Plant Biotechnol J.* 2011;9(3):334–47.
47. Klein SP, Anderson SN. The evolution and function of transposons in epigenetic regulation in response to the environment. *Curr Opin Plant Biol.* 2022;69: 102277.
48. Korte A, Farlow A. The advantages and limitations of trait analysis with GWAS: a review. *Plant Methods.* 2013;9(1):1–9.
49. Kottler, E.J., VanWallendael, A. and Franks, S.J. (2018) 'Experimental Treatment with a Hypomethylating Agent Alters Life History Traits and Fitness in *Brassica rapa*', *Journal of Botany* [Preprint].
50. Krishna Kumar S, et al. Limitations of GCTA as a solution to the missing heritability problem. *Proc Natl Acad Sci.* 2016;113(1):E61–70.
51. Krueger, F. and Andrews, S.R. (2011) 'Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications', *bioinformatics*, 27(11), pp. 1571–1572.
52. Lämke J, Bäurle I. Epigenetic and chromatin-based mechanisms in environmental stress adaptation and stress memory in plants. *Genome Biol.* 2017;18(1):1–11.
53. Lang Z, et al. Critical roles of DNA demethylation in the activation of ripening-induced genes and inhibition of ripening-repressed genes in tomato fruit. *Proc Natl Acad Sci.* 2017;114(22):E4511–9.
54. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods.* 2012;9(4):357–9.
55. Law JA, Jacobsen SE. Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nat Rev Genet.* 2010;11(3):204–20.
56. Li C, et al. EGR1 and EGR2 positively regulate plant ABA signaling by modulating the phosphorylation of SnRK2. 2. *New Phytol.* 2024;241(4):1492–509.
57. Liang T, et al. The interplay between the circadian clock and abiotic stress responses mediated by ABF3 and CCA1/LHY. *Proc Natl Acad Sci.* 2024;121(7): e23168252121.
58. Lind MI, Spagopoulou F. Evolutionary consequences of epigenetic inheritance. *Heredity.* 2018;121(3):205–9.
59. Lindroth AM, et al. Requirement of CHROMOMETHYLASE3 for maintenance of CpXpG methylation. *Science.* 2001;292(5524):2077–80.
60. Lira-Medeiros CF, et al. Epigenetic variation in mangrove plants occurring in contrasting natural environment. *PLoS ONE.* 2010;5(4): e10326.
61. Lister R, et al. Highly integrated single-base resolution maps of the epigenome in *Arabidopsis*. *Cell.* 2008;133(3):523–36.
62. Liu T, et al. Cold acclimation alters DNA methylation patterns and confers tolerance to heat and increases growth rate in *Brassica rapa*. *J Exp Bot.* 2017;68(5):1213–24.
63. López-García CM, et al. ALTERED MERISTEM PROGRAM 1 promotes growth and biomass accumulation influencing guard cell aperture and photosynthetic efficiency in *Arabidopsis*. *Protoplasma.* 2020;257:573–82.
64. Lorincz MC, et al. Infragenic DNA methylation alters chromatin structure and elongation efficiency in mammalian cells. *Nat Struct Mol Biol.* 2004;11(11):1068–75.
65. Min L, et al. Sugar and auxin signaling pathways respond to high-temperature stress during anther development as revealed by transcript profiling analysis in cotton. *Plant Physiol.* 2014;164(3):1293–308.
66. Miska EA, Ferguson-Smith AC. Transgenerational inheritance: Models and mechanisms of non-DNA sequence-based inheritance. *Science.* 2016;354(6308):59–63.
67. Muyle A, Gaut BS. Loss of gene body methylation in *Eutrema salsugineum* is associated with reduced gene expression. *Mol Biol Evol.* 2019;36(1):155–8.
68. Nagler M, et al. Integrative molecular profiling indicates a central role of transitory starch breakdown in establishing a stable C/N homeostasis during cold acclimation in two natural accessions of *Arabidopsis thaliana*. *BMC Plant Biol.* 2015;15:1–19.
69. Naydenov M, et al. High-temperature effect on genes engaged in DNA methylation and affected by DNA methylation in *Arabidopsis*. *Plant Physiol Biochem.* 2015;87:102–8.
70. Niederhuth CE, et al. Widespread natural variation of DNA methylation within angiosperms. *Genome Biol.* 2016;17(1):1–19.
71. Nobusawa T, et al. Highly pleiotropic functions of CYP78As and AMP1 are regulated in non-cell-autonomous/organ-specific manners. *Plant Physiol.* 2021;186(1):767–81.
72. Oksanen, J. et al. (2017) 'Ordination methods, diversity analysis and other functions for community and vegetation ecologists', *Vegan: Community Ecol Package*, pp. 05–26.
73. Panda K, et al. The plant response to high CO<sub>2</sub> levels is heritable and orchestrated by DNA methylation. *New Phytol.* 2023;238(6):2427–39.
74. Paun O, et al. Stable epigenetic effects impact adaptation in allopolyploid orchids (*Dactylorhiza*: *Orchidaceae*). *Mol Biol Evol.* 2010;27(11):2465–73.
75. Paun O, Verhoeven KJ, Richards CL. Opportunities and limitations of reduced representation bisulfite sequencing in plant ecological epigenomics. *New Phytol.* 2019;221(2):738–42.
76. Pikaard CS, Scheid OM. Epigenetic regulation in plants. *Cold Spring Harb Perspect Biol.* 2014;6(12): a019315.
77. Platt A, et al. Genome-wide signature of local adaptation linked to variable C p G methylation in oak populations. *Mol Ecol.* 2015;24(15):3823–30.
78. Quadrana L, Colot V. Plant transgenerational epigenetics. *Annu Rev Genet.* 2016;50:467–91.
79. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics.* 2010;26(6):841–2.
80. Resende RT, et al. Regional heritability mapping and genome-wide association identify loci for complex growth, wood and disease resistance traits in *Eucalyptus*. *New Phytol.* 2017;213(3):1287–300.
81. Rico L, et al. Changes in DNA methylation fingerprint of *Q uercus ilex* trees in response to experimental field drought simulating projected climate change. *Plant Biol.* 2014;16(2):419–27.
82. Rodríguez, B.D. et al. (2022) 'Epigenetic variation in the Lombardy poplar along climatic gradients is independent of genetic structure and persists across clonal reproduction', *bioRxiv*, p. 2022.11. 17.516862.
83. Rolly NK, Mun B-G, Yun B-W. Insights into the transcriptional regulation of branching hormonal signaling pathways genes under drought stress in *arabidopsis*. *Genes.* 2021;12(2):298.
84. Rungnone ML, et al. LNK genes integrate light and clock signaling networks at the core of the *Arabidopsis* oscillator. *Proc Natl Acad Sci.* 2013;110(29):12120–5.
85. Ruiz-Gayoso A, et al. Phosphorylation of DPE2 at S786 partially regulates starch degradation. *Plant Physiol Biochem.* 2022;193:70–7.
86. Saban JM, et al. The methylome is altered for plants in a high CO<sub>2</sub> world: insights into the response of a wild plant population to multi-generational exposure to elevated atmospheric [CO<sub>2</sub>]. *Glob Change Biol.* 2020;26(11):6474–92.
87. Saito R, et al. Functional analysis of the AKR4C subfamily of *Arabidopsis thaliana*: model structures, substrate specificity, acrolein toxicity, and responses to light and [CO<sub>2</sub>]. *Biosci Biotechnol Biochem.* 2013;77(10):2038–45.
88. Sammarco I, et al. DNA methylation in the wild: epigenetic transgenerational inheritance can mediate adaptation in clones of wild strawberry (*Fragaria vesca*). *New Phytol.* 2024;241(4):1621–35.
89. Schmid MW, et al. Contribution of epigenetic variation to adaptation in *Arabidopsis*. *Nat Commun.* 2018;9(1):4446.
90. Schmitz RJ, et al. Transgenerational epigenetic instability is a source of novel methylation variants. *Science.* 2011;334(6054):369–73.

91. Schmitz RJ, et al. Patterns of population epigenomic diversity. *Nature*. 2013;495(7440):193–8.
92. Seymour DK, et al. Evolution of DNA methylation patterns in the Brassicaceae is driven by differences in genome organization. *PLoS Genet*. 2014;10(11): e1004785.
93. Singh, A. et al. (2015) 'Morphological trait variations in the west Himalayan (India) populations of *Arabidopsis thaliana* along altitudinal gradients', *Current Science*, pp. 2213–2222.
94. Singh A, Roy S. High altitude population of *Arabidopsis thaliana* is more plastic and adaptive under common garden than controlled condition. *BMC Ecol*. 2017;17:1–16.
95. Steward N, Kusano T, Sano H. Expression of ZmMET1, a gene encoding a DNA methyltransferase from maize, is associated not only with DNA replication in actively proliferating cells, but also with altered DNA methylation status in cold-stressed quiescent cells. *Nucleic Acids Res*. 2000;28(17):3250–9.
96. Stroud H, et al. Non-CG methylation patterns shape the epigenetic landscape in *Arabidopsis*. *Nat Struct Mol Biol*. 2014;21(1):64–72.
97. Swindell WR. The association among gene expression responses to nine abiotic stress treatments in *Arabidopsis thaliana*. *Genetics*. 2006;174(4):1811–24.
98. Takahashi S, et al. Roles of *Arabidopsis* AtREV1 and AtREV7 in translesion synthesis. *Plant Physiol*. 2005;138(2):870–81.
99. Takuno S, Gaut BS. Body-methylated genes in *Arabidopsis thaliana* are functionally important and evolve slowly. *Mol Biol Evol*. 2012;29(1):219–27.
100. Tang X-M, et al. Analysis of DNA methylation of perennial ryegrass under drought using the methylation-sensitive amplification polymorphism (MSAP) technique. *Mol Genet Genomics*. 2014;289:1075–84.
101. Tatra GS, et al. Effect of light quality and 5-azacytidine on genomic methylation and stem elongation in two ecotypes of *Stellaria longipes*. *Physiol Plant*. 2000;109(3):313–21.
102. Teh O, et al. BEACH-domain proteins act together in a cascade to mediate vacuolar protein trafficking and disease resistance in *Arabidopsis*. *Mol Plant*. 2015;8(3):389–98.
103. Teschendorff AE, Relton CL. Statistical and integrative system-level analysis of DNA methylation data. *Nat Rev Genet*. 2018;19(3):129–47.
104. Tiwari LD, Khungar L, Grover A. Athsc70-1 negatively regulates the basal heat tolerance in *Arabidopsis thaliana* through affecting the activity of HsfAs and Hsp101. *Plant J*. 2020;103(6):2069–83.
105. Troyee AN, et al. Variation in DNA methylation and response to short-term herbivory in *Thlaspi arvense*. *Flora*. 2022;293: 152106.
106. Tyagi, A., Singh, S., et al. (2016) 'Genetic diversity and population structure of *Arabidopsis thaliana* along an altitudinal gradient', *AoB Plants*, 8.
107. Tyagi A, Yadav A, et al. High light intensity plays a major role in emergence of population level variation in *Arabidopsis thaliana* along an altitudinal gradient. *Sci Rep*. 2016;6(1):26160.
108. Urbanowicz BR, et al. Structural organization and a standardized nomenclature for plant endo-1,4- $\beta$ -glucanases (cellulases) of glycosyl hydrolase family 9. *Plant Physiol*. 2007;144(4):1693–6.
109. Verhoeven KJ, et al. Stress-induced DNA methylation changes and their heritability in asexual dandelions. *New Phytol*. 2010;185(4):1108–18.
110. Verma AK, et al. Population specific methyloyme remodeling in high and low elevation populations of Indian west Himalayan *Arabidopsis thaliana* in response to elevated CO<sub>2</sub>. *Environ Exp Bot*. 2022;203: 105074.
111. Wang B, et al. Analysis of methylation-sensitive amplified polymorphism in different cotton accessions under salt stress based on capillary electrophoresis. *Genes & Genomics*. 2015;37:713–24.
112. Wang B, et al. Myosin XI-mediated BIK1 recruitment to nanodomains facilitates FLS2–BIK1 complex formation during innate immunity in *Arabidopsis*. *Proc Natl Acad Sci*. 2024;121(25): e2312415121.
113. Wang X, et al. DNA methylation affects gene alternative splicing in plants: an example from rice. *Mol Plant*. 2016;9(2):305–7.
114. Wellenreuther M, Hansson B. Detecting polygenic evolution: problems, pitfalls, and promises. *Trends Genet*. 2016;32(3):155–64.
115. Wibowo A, et al. Hyperosmotic stress memory in *Arabidopsis* is mediated by distinct epigenetically labile sites in the genome and is restricted in the male germline by DNA glycosylase activity. *Elife*. 2016;5: e13546.
116. Williams BP, et al. Methylation-sensitive expression of a DNA demethylase gene serves as an epigenetic rheostat. *PLoS Genet*. 2015;11(3): e1005142.
117. Xia H, et al. Adaptive epigenetic differentiation between upland and lowland rice ecotypes revealed by methylation-sensitive amplified polymorphism. *PLoS ONE*. 2016;11(7): e0157810.
118. Yaish MW, et al. Genome-wide DNA Methylation analysis in response to salinity in the model plant caliph medic (*Medicago truncatula*). *BMC Genomics*. 2018;19(1):1–17.
119. Yakovlev IA, et al. Temperature-dependent differential transcriptomes during formation of an epigenetic memory in Norway spruce embryogenesis. *Tree Genet Genomes*. 2014;10:355–66.
120. Yoder JB, et al. Genomic signature of adaptation to climate in *Medicago truncatula*. *Genetics*. 2014;196(4):1263–75.
121. Yong-Villalobos L, et al. Phosphate starvation induces DNA methylation in the vicinity of cis-acting elements known to regulate the expression of phosphate-responsive genes. *Plant Signal Behav*. 2016;11(5): e1173300.
122. Zhang X, et al. Genome-wide high-resolution mapping and functional analysis of DNA methylation in *Arabidopsis*. *Cell*. 2006;126(6):1189–201.
123. Zhang X, et al. Global analysis of genetic, epigenetic and transcriptional polymorphisms in *Arabidopsis thaliana* using whole genome tiling arrays. *PLoS Genet*. 2008;4(3): e1000032.
124. Zhao, G. et al. (2023) 'The abscisic acid-responsive element binding factors MAPKKK18 module regulates abscisic acid-induced leaf senescence in *Arabidopsis*', *Journal of Biological Chemistry*, 299(4).
125. Zilberman D, et al. Genome-wide analysis of *Arabidopsis thaliana* DNA methylation uncovers an interdependence between methylation and transcription. *Nat Genet*. 2007;39(1):61–9.
126. Zilberman D. An evolutionary case for functional gene body methylation in plants and animals. *Genome Biol*. 2017;18(1):1–3.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.