



Environmental and genealogical effects on DNA methylation in a widespread apomictic dandelion lineage

Verónica Noé Ibañez, Morgane van Antro, Cristian Peña-ponton, Slavica Milanovic-ivanovic, Cornelis a M Wagemaker, Fleur Gawehns, Koen J F Verhoeven

► To cite this version:

Verónica Noé Ibañez, Morgane van Antro, Cristian Peña-ponton, Slavica Milanovic-ivanovic, Cornelis a M Wagemaker, et al.. Environmental and genealogical effects on DNA methylation in a widespread apomictic dandelion lineage. *Journal of Evolutionary Biology*, 2023, 36 (4), pp.663-674. <10.1111/jeb.14162>. <hal-04495735>

HAL Id: hal-04495735

<https://hal.inrae.fr/hal-04495735v1>

Submitted on 8 Mar 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY 4.0 - Attribution - International License

RESEARCH ARTICLE

Environmental and genealogical effects on DNA methylation in a widespread apomictic dandelion lineage

Verónica Noé Ibañez^{1,2}  | Morgane van Antrop²  | Cristian Peña-Ponton²  | Slavica Milanovic-Ivanovic² | Cornelis A. M. Wagemaker³  | Fleur Gawehns²  | Koen J. F. Verhoeven² ¹Institute of Plant Sciences Paris-Saclay (IPSS), Université Paris-Saclay, CNRS, INRAE, Université Evry, Université Paris Diderot, Gif sur Yvette, France²Department of Terrestrial Ecology, Netherlands Institute of Ecology (NIOO-KNAW), Wageningen, the Netherlands³Plant Ecology and Physiology, Radboud University, Nijmegen, the Netherlands

Correspondence

Koen J. F. Verhoeven, Department of Terrestrial Ecology, Netherlands Institute of Ecology (NIOO-KNAW), Droevendaalsesteeg 10, 6708 PB Wageningen, the Netherlands.
Email: k.verhoeven@nioo.knaw.nl

Funding information

Agencia Nacional de Promoción Científica y Tecnológica Nacional, Grant/Award Number: PICT 2020-1767; H2020 Marie Skłodowska-Curie Actions, Grant/Award Number: European Training Network grant 764965; Universidad Nacional de Cuyo, Grant/Award Number: travel grant 770/2018-C.S.

Abstract

DNA methylation in plant genomes occurs in different sequences and genomic contexts that have very different properties. DNA methylation that occurs in CG (mCG) sequence context shows transgenerational stability and high epimutation rate, and can thus provide genealogical information at short time scales. However, due to metastability and because mCG variants may arise due to other factors than epimutation, such as environmental stress exposure, it is not clear how well mCG captures genealogical information at micro-evolutionary time scales. Here, we analysed DNA methylation variation between accessions from a geographically widespread, apomictic common dandelion (*Taraxacum officinale*) lineage when grown experimentally under different light conditions. Using a reduced-representation bisulphite sequencing approach, we show that the light treatment induced differentially methylated cytosines (DMCs) in all sequence contexts, with a bias towards transposable elements. Accession differences were associated mainly with DMCs in CG context. Hierarchical clustering of samples based on total mCG profiles revealed a perfect clustering of samples by accession identity, irrespective of light conditions. Using microsatellite information as a benchmark of genetic divergence within the clonal lineage, we show that genetic divergence between accessions correlates strongly with overall mCG profiles. However, our results suggest that environmental effects that do occur in CG context may produce a heritable signal that partly dilutes the genealogical signal. Our study shows that methylation information in plants can be used to reconstruct micro-evolutionary genealogy, providing a useful tool in systems that lack genetic variation such as clonal and vegetatively propagated plants.

KEYWORDS

ecological genetics, genomics, plants

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2023 The Authors. *Journal of Evolutionary Biology* published by John Wiley & Sons Ltd on behalf of European Society for Evolutionary Biology.

1 | INTRODUCTION

DNA methylation in plants is involved in developmental and environmental plasticity (Richards et al., 2017) and in maintaining genomic integrity by suppressing transposable element activity (Zhang, Latzel, et al., 2018). Because of the meiotic stability of part of the plant methylome, DNA methylation can contribute to heritable trait variation (Cubas et al., 1999), with potential consequences for plant adaptation and evolution (Quadrana & Colot, 2016; Schmid et al., 2018; Wilschut et al., 2016).

When exploring the functions and consequences of DNA methylation, it is important to recognize that DNA methylation in plant genomes occurs at cytosines in different sequence contexts, i.e., CG, CHG and CHH (where H is C, A or T), whose dynamics in the genome are controlled by different enzymes (Cao & Jacobsen, 2002; Cokus et al., 2008; Finnegan & Kovac, 2000; Law & Jacobsen, 2010; Lindroth et al., 2001; Lister et al., 2008; Zhang, Lang, & Zhu, 2018). Methylation in these contexts shows different functions, stabilities and environmental dynamics.

Often, DNA methylation changes that are induced by exposure to specific environments occur in CHH context (Dubin et al., 2015; Galanti et al., 2022; Kawakatsu et al., 2016), which is typically transient and not transgenerationally stable (Boyko et al., 2010; Ganguly et al., 2017; Groot et al., 2016; Secco et al., 2015; Suter & Widmer, 2013a, 2013b; Van Dooren et al., 2020; Wibowo et al., 2016). In contrast, methylation in CG context shows strong transgenerational stability in plants (Zhang, Lang, & Zhu, 2018). The combination of transgenerational stability and frequent spontaneous epimutations (Johannes & Schmitz, 2019; Schmitz et al., 2011), mostly in plant gene bodies (Zhang, Lang, & Zhu, 2018) results in an increased CG methylation divergence between lineages over multiple generations (Hagmann et al., 2015; Johannes & Schmitz, 2019). Thus, environmental effects on DNA methylation occur in sequence and genomic contexts that are partly different from the contexts in which transgenerational divergence between samples occurs. Plant methylomes, therefore, contain genealogical and environmental information that can be distinguished, at least partly, by evaluating different components of the methylome.

It is an open question to what extent the genealogical information from CG methylation is diluted by frequent reversal of CG epimutations (which could uncouple the epigenetic and genetic information; van der Graaf et al., 2015) or by other effects (e.g. environmental) on CG methylation. Nevertheless, heritable CG methylation successfully captured the genetic relationships between accessions within an invasive *Arabidopsis thaliana* lineage over a one-century-old invasion in absence of strong environmental effects (Hagmann et al., 2015). As epimutations occur at a much higher rate than genetic (SNP) mutations (Ossowski et al., 2010; van der Graaf et al., 2015), isolating the genealogical signals from plant methylomes by evaluating methylation variation in CG context could provide information at shorter time scales and at higher resolutions than genetic variation does (Yao et al., 2021). This could be particularly useful for unravelling

relationships in plants with little genetic variation, such as asexual lineages and vegetatively propagated crops.

Apomictic plant lineages, which produce clonal seeds while avoiding the complicating effects of segregation and recombination, evolve via mutation accumulation and are a convenient system for characterizing DNA methylation divergence over a microevolutionary time scale in a genetically uniform background. In the common dandelion (*Taraxacum officinale* F. H.Wigg; Asteraceae family), many different triploid apomictic lineages exist, some of which show a geographically widespread distribution in Europe (Kirschner et al., 2016), likely following post-glacial expansion (Kirschner & Štěpánek, 1994). In a previous study, accessions from one such triploid apomictic dandelion lineage, *alatum*, were collected from various countries in central and north Europe and characterized using microsatellites, (i.e. SSRs from Simple Sequence Repeats). This revealed that the accessions are of uniclonal origin and that some individuals show allelic mutations at one or few SSR loci (Kirschner et al., 2016; Wilschut et al., 2016). Here, we make use of the characterized distribution and genetic diversity of the *alatum* lineage to investigate how genealogical and environmental signals are expressed in plant methylomes.

In this study, we analysed DNA methylation in a panel of 20 accessions from the *alatum* lineage, when grown experimentally under different light conditions. By exposing the accessions to an experimental shading versus control treatment, we sought to characterize DNA methylation differences among accessions and between environmental treatments. Dandelion plants grow in open fields and are weak competitors for light in dense vegetation (Kovel et al., 1999; Molina-Montenegro et al., 2012), making the experimental shade treatment an ecologically relevant stress for this species. Specifically, we asked (1) In what sequence and genomic contexts do DNA methylation differences exist between accessions, and how well does this capture genetic divergence among the accessions? (2) In what sequence and genomic contexts does DNA methylation show a plastic response to the shading treatment? As these questions can be addressed by looking at many sites in the genome, but not necessarily entire genomes (Paun et al., 2019), we used a reduced representation bisulphite sequencing approach for methylation analysis (epiGBS; Gawehns et al., 2022; van Gorp et al., 2016). This method can be applied in absence of a reference genome for the species under study.

2 | MATERIAL AND METHODS

2.1 | Plant samples and experimental design

We selected 20 accessions with confirmed *alatum* identity, which were collected in Germany, Finland, Austria and Czech Republic (Kirschner et al., 2016) (Table S1, Figure S1). The accessions showed modest levels of genetic divergence based on genetic characterization at eight SSR loci (Table S1). All accessions were propagated under common greenhouse conditions for one generation prior to

the experiment. The haploid genome size of the species is estimated as 800Mb (Underwood et al., 2022).

Control and shading filter treatments were applied continuously from the start of seed germination until leaf tissue sampling after 4 weeks of seedling growth. In total, 40 plants were analysed, corresponding to 20 accessions by two treatments, with one replicate per accession x treatment combination. Seeds from all accessions were germinated in petri dishes on moist filter paper for 10 days. We used two petri dishes per accession for 'control' and 'shade' treatments during germination. Seedlings were transplanted individually into pots containing a 4:1 mixture of potting soil/pumice. For each accession, one plant was grown under control conditions and one plant was grown under artificial shading conditions. Artificial shading was applied with a green filter that was fixed approximately 10–20 cm above the plants (LEE Filter 122, fern green), whereas a colourless filter (LEE Filter 130, clear) was used for control plants. Light conditions and temperature during all experiments were set at 16 h–8 h of light/dark cycle at 20°C and 16°C respectively. Plants were watered 2–3 times per week to maintain well-watered growing conditions.

Leaf tissue was sampled after 4 weeks of growth in the greenhouse by taking multiple 8 mm leaf punches per plant from three young, fully developed leaves (control plants: 10 leaf punches per plant) or from all fully developed leaves (shade plants: 14 leaf punches per plant). In the shade treatment, plants produced only few but large, etiolated leaves; more leaf punches were sampled from etiolated leaves to compensate for the potentially lower number of cells per surface area in such leaves. Tissue samples were flash-frozen in liquid nitrogen and stored at –80°C until further processing. DNA was isolated using the NucleoSpin Plant II kit (Macherey-Nagel), according to the manufacturer's protocol and using the cell lysis buffer PL1 (CTAB) and sample incubation at 65°C for 30 min. An extra 30-s centrifugation step was performed after the third washing step. Then, the silica membrane was dried in a new 1.5 mL tube with an open lid for 5 min at 55°C to eliminate ethanol before continuing with eluting DNA from the column. DNA quality was checked using Nanodrop and agarose gel electrophoresis. Only samples with high gDNA integrity and purity after measuring on Nanodrop were retained (260/280 and 260/230 ratios of at least 1.8 and 1.5 respectively). DNA concentration was determined using the Qubit 2.0 Fluorometric dsDNA HS assay kit.

2.2 | epiGBS library preparation and sequencing

DNA samples were processed into a multiplexed epiGBS sequencing library (van Gurp et al., 2016) for reduced-representation bisulphite sequencing-based analysis of DNA methylation, following the library preparation protocol of epiGBS2 (Gawehns et al., 2022). Briefly, 400 ng of genomic DNA was digested using *Csp6I* and *NsiI* restriction enzymes. Hemimethylated adapter pairs were then ligated to the digested DNA, with each adapter containing a 4–6 nucleotide sample-specific barcode, followed by a string of three Unique Molecule Identifier (UMI) random nucleotides (NNN) and an unmethylated

cytosine (used to distinguish between Watson and Crick strand as required for downstream epiGBS2 data analysis, as well as to estimate the bisulphite conversion rate). Samples were then pooled in 4 aliquots of 10 samples each, concentrated and cleaned of smaller fragments (<150 bp) using the NucleoSpin Gel and PCR clean-up Kit protocol. Size selection was done using 0.8× SPRIselect magnetic beads, selecting DNA fragments of 150 bp and higher. The nicks induced by the use of unphosphorylated hemimethylated adapters were then repaired by DNA polymerase I treatment through the use of dNTPs that contain 5-meC's resulting in fully ligated and methylated adapters. The pooled samples were bisulphite converted using the EZ DNA Methylation-Lightning kits, following the manufacturer's protocol. The converted DNA was PCR amplified before a final PCR clean-up, size-selection, DNA quantification and pooling step. The final multiplexed library was sequenced paired-end (PE 2×150 bp) with a 12% phiX spike on one lane using Illumina HiSeq 2500 sequencer (Novogene, Hong Kong). In the same sequencing library, we included also a set of the same 40 plant samples after epiGBS library preparation based on enzymatic digestion with another set of restriction enzymes, specifically *Asel* and *Nsil*. This was done because it was unclear for this plant species which enzymatic restriction protocol is optimal for epiGBS library preparation; because output from reduced representation approaches depends on the interaction between genome properties and restriction enzyme recognition sites, it can be useful to test multiple enzyme combinations. The results of the *Csp6I*-*Nsil* digested samples are presented in the manuscript, as these samples yielded the highest sequencing output and thus present a more robust basis for interpreting results. The results of the *Asel*-*Nsil* digested samples were qualitatively very similar and are presented as Tables S2–S4 and Figures S2–S6. Based on the evaluation of the unmethylated cytosine nucleotide in the adapters, the bisulphite conversion rate was estimated as at least 96.3%. This estimate is a lower limit: the unmethylated cytosine appears in the first part of the reads which are prone to higher sequencing error rates, and such errors are interpreted as non-conversions.

2.3 | epiGBS data processing

Sequencing of the epiGBS library yielded 144 645 790 and 54 011 404 raw reads for the 40 *Csp6I* - *Nsil* and *Asel*-*Nsil* samples respectively. Data were processed to generate cytosine-level DNA methylation rate estimates using the epiGBS2 pipeline under the git commit number 60e5376 (Gawehns et al., 2022), with the 'de novo' option, and setting identity, min-depth and max-depth parameters as 0.95, 0 and 0 respectively. Briefly, this pipeline performs PCR duplicate removal based on the identity of the UMI sequence, sample demultiplexing based on their barcodes, de novo reference construction for the epiGBS fragments based on the combined information from Watson and Crick (bisulphite) reads in the data, adapter trimming, read alignment against the de novo reference and strand-specific variant calling. The pipeline produces a .bed file containing all cytosines in the data with,

for each sample, the total read depth and the number of methylated and unmethylated reads respectively. DNA methylation for each cytosine is subsequently calculated as the ratio of methylated/total reads. Using the de novo reference sequences, protein-coding genes were annotated with *DIAMOND* version 0.8.22 (Buchfink et al., 2015) with default parameters and considering an *e* value threshold of 1×10^{-5} . Transposable elements (TE) were annotated with *repeatmasker* version 4.1.1 with default parameters and considering Embryophyta as reference species collection (Tempel, 2012).

2.4 | Data analysis

Filtering, differential methylation analyses and visualization were performed with R (v.4.0.1, R Core Team, 2020). Filtering of the *.bed* file consisted of multiple steps. First, accessions with low read coverage were removed: for each sample, we calculated average read coverage per cytosine and accessions were removed from further analysis if their average read coverage (of one or of both the samples from that accession, shade or control) was below 50% of the grand average read coverage of all accessions. Second, we set a coverage threshold of 10 reads (turning observations with <10 reads into missing data) and we removed all cytosines from further analysis that were not present with at least 10× coverage in at least 80% of all samples. Third, we retained for further analysis only the cytosines within the 99th percentile of the coverage distribution, discarding the top 1% of cytosines with the highest sequence coverage. The filtered data set consisted of 34 samples, corresponding to 17 light-shade treatment pairs with 5385 epiGBS fragments containing 121,928 cytosines with an average coverage of 20.

To provide an overview of the genome-wide DNA methylation levels, the methylation ratio was calculated for every cytosine in each sample, and then an average was calculated across samples for each individual cytosine for CG, CHG and CHH sequence contexts. The distribution of those methylation averages was visualized with violin plots. Similarly, the average per-cytosine DNA methylation percentage within annotated genomic features (genes, repeats, TE and unclassified) were represented with heatmaps, across all individuals within each of the treatment groups. Pairwise epigenetic distances were calculated according to van Moorsel et al. (2019), as the average methylation difference across all cytosines between two samples. A non-metric MDS was used to visualize epigenetic distances with the function *isoMDS* of the R-package 'MASS' (Venables & Ripley, 2002). The statistical significance of *isoMDS* clustering was determined through a permutational multivariate analysis of variance with the *adonis* function from R-package 'vegan' (Oksanen, 2007).

A hierarchical cluster analysis was realized with *hclust* function from 'stats' R-Package Team (2020) based on the pairwise distance matrix for cytosine methylation in CG context and based on eight microsatellite loci (SSR; Kirschner et al., 2016; Wilschut et al., 2016) using Nei distance with the UPGMA linkage method. Mantel test analyses were performed to determine correlations between

SSR-based and methylation-based distance matrices. Correlations were tested separately for genomic features and cytosine methylation contexts (CG, CHG and CHH), including either all cytosines in the analysis or only those cytosines that were differentially methylated at *p*-value < 0.05 before FDR correction.

Differences in DNA methylation between accessions or light treatments were tested at each individual cytosine for statistical significance in R using the 'DSS' package (version 2.36; Park & Wu, 2016), where accession and light treatment were considered as fixed effects. DSS (without the smoothing option) fits a beta-binomial generalized linear model using an *arcsin* link function and Wald tests for determining statistical significance. Resulting *p*-values were adjusted for multiple testing using False Discovery Rate control (Benjamini & Hochberg, 1995); a cytosine was defined as differentially methylated (DMC) at an FDR threshold of 0.05. EpiGBS fragments that contained more than five DMCs between light conditions were defined as 'putative DMRs' (differentially methylated regions). Manhattan plots of epiGBS fragments were generated using 'qqman' R-package (Turner, 2021) to visualize fragments with many DMCs. The effect of light treatment for the DMCs was assessed by fitting a generalized linear fixed-effects model with a Binomial distribution with the *glm* function from 'stat' R-package. The mean post hoc comparison was done with Tukey's HSD test ($p \leq 0.05$) with the *emmeans* function from 'emmeans' R-package.

3 | RESULTS

3.1 | Variation in overall methylome patterns

The sequenced cytosines were classified according to the genomic feature in which they were found. More than half of the cytosines were unclassified. Among the annotated regions of the genome, more cytosines were found in TEs followed by genes and repeats (Figure 1a). Distributions of per-cytosine methylation levels (Figure 1b) in the CG and in CHH contexts were bimodal, with most cytosines either having high methylation (75–100%, typical for CG context) or low methylation (0–25%, typical for CHH context). DNA methylation in CHG context showed a less extreme bimodal distribution, with peaks below 25% and above 75% methylation but also with many cytosines showing intermediate methylation levels (Figure 1b). For all accessions and the two light conditions, on average DNA methylation was highest in CG context (88.4%), lower in CHG (58.8%) and lowest in CHH context (14.6%; Figure 1c). For all contexts, DNA methylation was higher in TEs and repeats than in genes (Figure 1d).

Significant clustering of samples by treatments and by accessions was observed based on DNA methylation profiles (Figure 2). Methylation in CG context was associated with the accession effect but not with the treatment effect, while methylation in CHH context was associated with the treatment effect but not with the accession effect. Methylation in CHG context was associated significantly with both accession and light condition effects.

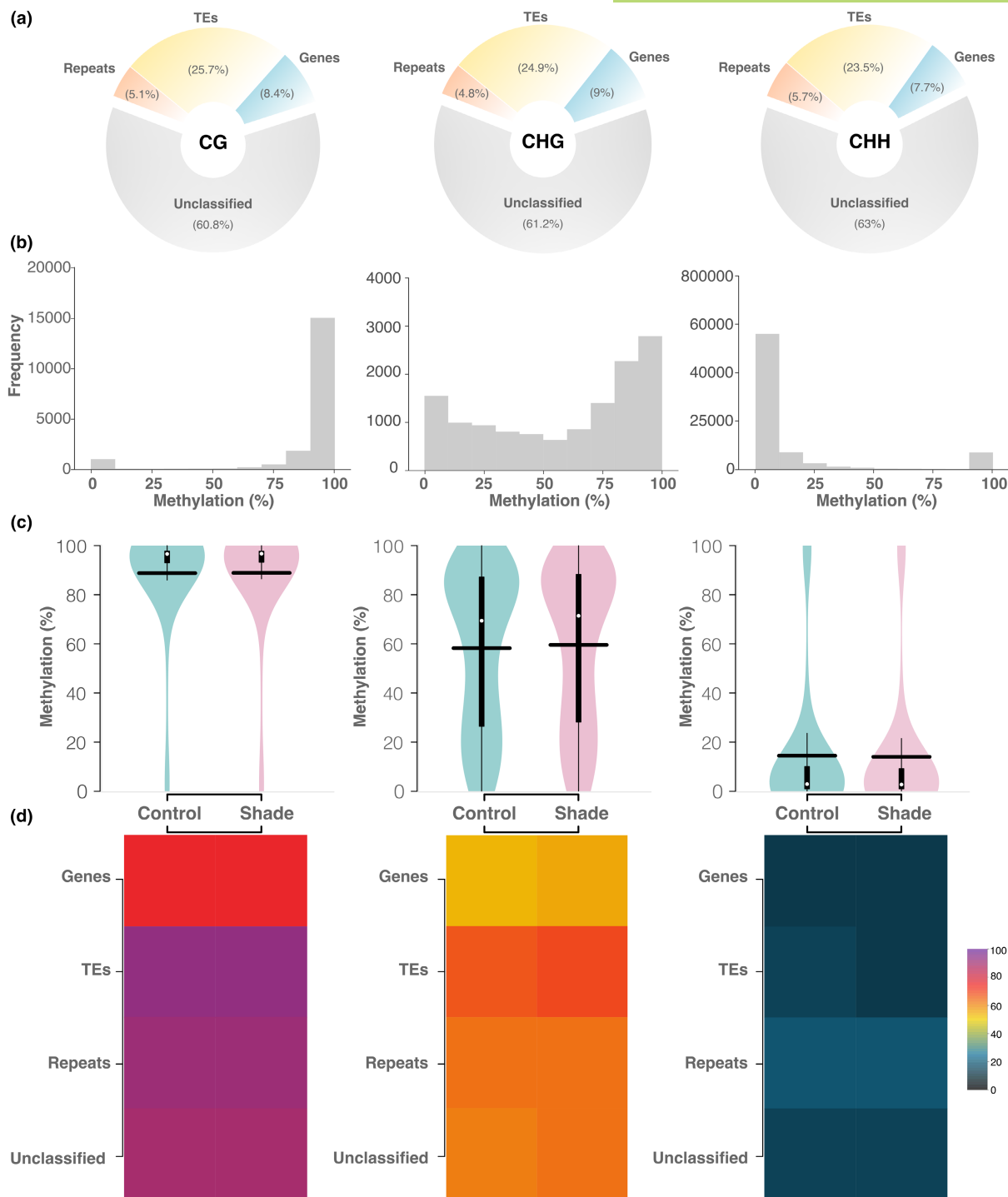


FIGURE 1 Characterization of global methylation levels, by annotation and sequence context (CG, CHG, CHH), of *Taraxacum officinale* accessions from *alatum* lineage cultivated under two light conditions. (a) percentage of sequenced cytosines classified as genes, transposable elements (TE), repeats or unclassified genomic features; (b) histograms showing the distribution of cytosine DNA methylation percentage for cytosines in different sequence contexts (CG, CHG and CHH; averaged across all accessions); (c) violin plots showing the average DNA methylation percentage for each light condition (averaged across all accessions). Means are represented with horizontal black bars; and (d) heatmap showing global DNA methylation percentage, as averaged over all cytosines per sequence context and genomic feature, for both light treatments.

3.2 | Differentially methylated cytosines

In response to the light condition treatment, 379 differentially methylated cytosines (DMCs) were observed, mostly in CHH context

(276 DMCs), followed by CHG context (63 DMCs) and 40 DMCs observed in CG context (Figure 3, Table 1). Relative to the representation of these contexts in the epiGBS fragments, and despite low absolute numbers of treatment-responsive CG-cytosines, the

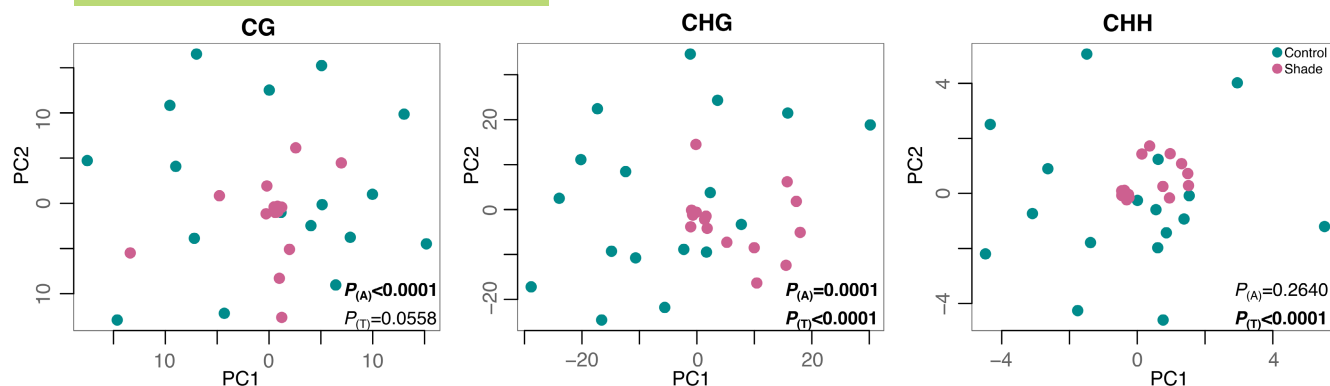


FIGURE 2 Environmental and genealogical signals are sequence context-dependent in *Taraxacum officinale* accessions from *alatum* lineage cultivated under two light conditions. Non-metric multidimensional scaling of DNA methylation for cytosines in different sequence contexts (CG, CHG and CHH) for each light condition (shade and control). Statistical significance of accession ($P_{(A)}$) and light treatment ($P_{(T)}$) effects on DNA methylation patterns was determined through permutational analysis of variance.

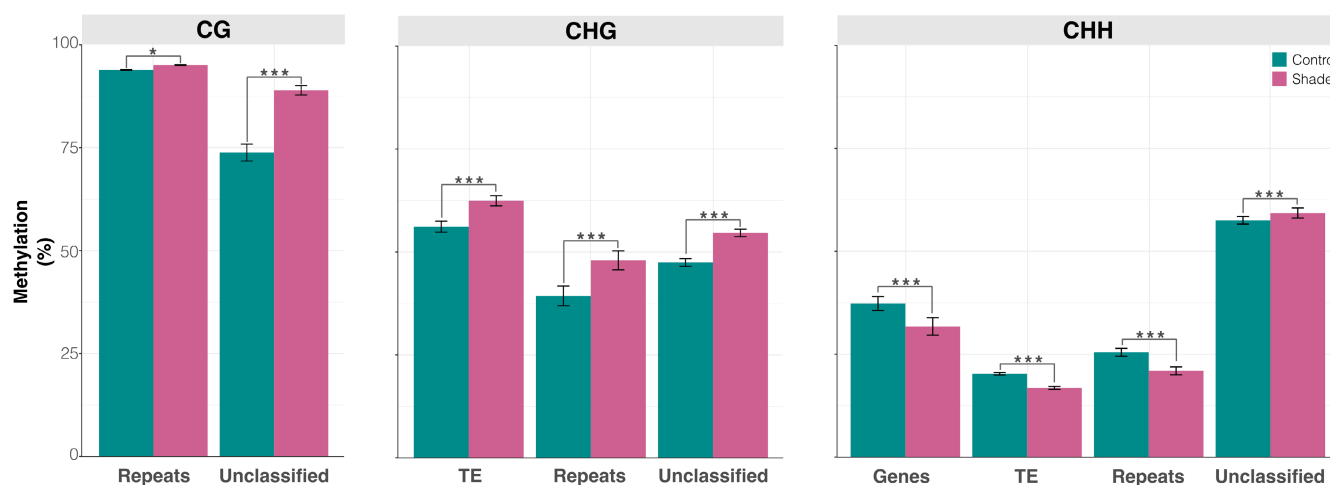


FIGURE 3 Methylation response at sequence context-dependent differentially methylated cytosines (DMCs). An overall hyper- or hypomethylation response due to shade treatment in *Taraxacum officinale* accessions from *alatum* lineage depends on the sequence context of the DMCs. Methylation levels from 40, 63 and 276 DMCs identified in CG, CHG and CHH contexts, respectively, mapped to genes, TE, repeats or unclassified genomic features. * p -value < 0.05 ; ** p -value ≤ 0.01 ; *** p -value ≤ 0.001 .

highest enrichment of DMCs was observed in CG context in repeats (3% of CG-cytosines tested responded significantly to the light treatment; Table 1). Much fewer DMCs (45) were associated with accession differences, and these occurred only in the CG context (Table 1). Putative differentially methylated regions (DMRs) were further analysed (Figure 4). Putative DMRs were observed in all sequence contexts, but the majority were found in the CHH context. From a total of 16 such fragments, seven were annotated as LTR-retrotransposon (Copia superfamily) and six as a DNA transposon (hAT superfamily; Table 2).

The large majority of DMCs observed between light conditions for the CHH context (210 of 379) were found in TEs or repeats (Table 1). In TEs and repeats, 0.2% and 0.6% of all CHH cytosines tested showed a significant treatment effect on methylation, whereas this was the case in only 0.1% of all tested CHH cytosines in genes and in unclassified genome regions (Table 1). These CHH-DMCs typically showed hypermethylation under shade conditions

(Figure 3). However, this hypomethylation trend was not observed at DMCs in un-annotated regions (Figure 3). Shade-induced hypomethylation was also not a general pattern for DMCs in CG or CHG context.

3.3 | DNA methylation correlations with genetic structure

Consistent with the observation that accession differences were observed in CG context (Table 1), hierarchical clustering of all experimental samples based on total CG methylation profiles revealed a clustering of samples by accession identity (Figure 5a). In all cases, samples were separated according to their accession identity. The two replicates of each accession always clustered closely together, even though they were grown under different light conditions. There was no clear clustering based on geographic origin, but we

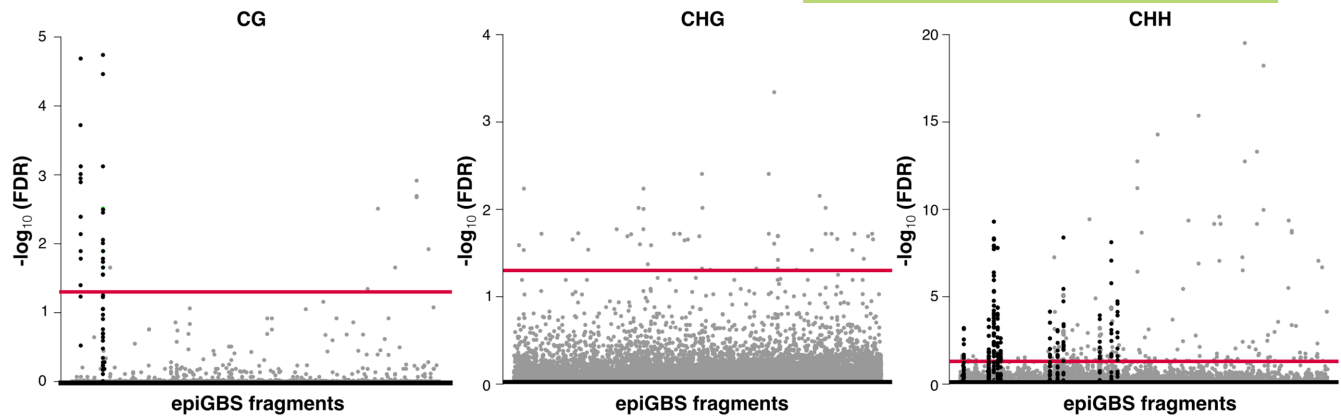


FIGURE 4 Putative differentially methylated regions (DMRs) in *Taraxacum officinale* accessions from *alatum* lineage in response to different light conditions. Manhattan plots of 5385 sequenced epiGBS fragments showing p values for tests of differential methylation at individual cytosines. The horizontal red line indicates the threshold for DMC significance at $FDR < 0.05$. Fragments were considered as putative DMRs when they have >5 DMCs and are highlighted in black. $-\log_{10}(FDR)$ values for the light treatment effect were calculated using the 'DSS' package in R with accession and light condition included as fixed effects in the statistical model.

TABLE 1 Number of differentially methylated cytosines associated with accession (Acc) and light treatment (Treat) effects ($FDR < 0.05$) for gene, transposable elements (TE), repeats and unclassified features and for each sequence context (CG, CHG and CHH).

Context	Gene		TE		Repeats		Unclassified	
	Acc (%)	Treat (%)	Acc (%)	Treat (%)	Acc (%)	Treat (%)	Acc (%)	Treat (%)
CG	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	3 (0.3)	35 (3.0)	42 (0.3)	5 (0.0)
CHG	0 (0.0)	0 (0.0)	0 (0.0)	10 (0.2)	0 (0.0)	7 (0.9)	0 (0.0)	46 (0.5)
CHH	0 (0.0)	9 (0.1)	0 (0.0)	183 (0.2)	0 (0.0)	27 (0.6)	0 (0.0)	57 (0.1)

Note: Percentages are relative to the total number of cytosines tested in that feature and context.

TABLE 2 Putative differentially methylated regions (DMR) in response to the light treatment. Putative DMRs are defined as epiGBS fragments with >5 DMCs.

Context	Number of epiGBS fragments	Total number of DMCs	Annotation
CHH	7	104	LTR retrotransposon. Copia superfamily
CHH	6	60	DNA transposon; hAT superfamily
CG	2	33	Unclassified
CHG	1	6	Unclassified

observed that different accessions that were originally collected from the same field location (which presumably share a common ancestor much more recently than accessions collected from different regions or countries do) always clustered close together (Figure 5b).

Because the genetic distances between the accessions can be estimated from independent SSR data (Kirschner et al., 2016; Wilschut et al., 2016), we explored how well methylation profiles in different sequence contexts and genome features correlate with genetic distance. Based on all cytosines, the strongest correlations were observed in CG context for methylation in repeats and in unannotated

regions and to a lesser extent with genes. Genetic variation also correlated strongly with CHG methylation in repeats (Table 3). No significant correlation was observed with any genomic feature for CHH methylation context. Correlation coefficients between genetic variation and methylation variation tended to be lower when including only the subset of cytosines in the methylation analysis that shows a significant accession effect in methylation level (Table 3).

4 | DISCUSSION

We explored how environmental and genealogical signals are detectable in different types of DNA methylation, focusing on sequence context (CG, CHG and CHH), in a geographically widespread apomictic dandelion lineage. Our results show that environmental treatment effects are detectable in all sequence contexts but are most abundant in non-CG contexts, while accession differences are expressed in CG context. The latter observation is consistent with high transgenerational stability of CG methylation (Niederhuth et al., 2016; Zhang, Lang, & Zhu, 2018). By comparing variation in methylation and microsatellites between accessions, we found that genealogical relationships are captured well by DNA methylation variation in CG context. We note that this

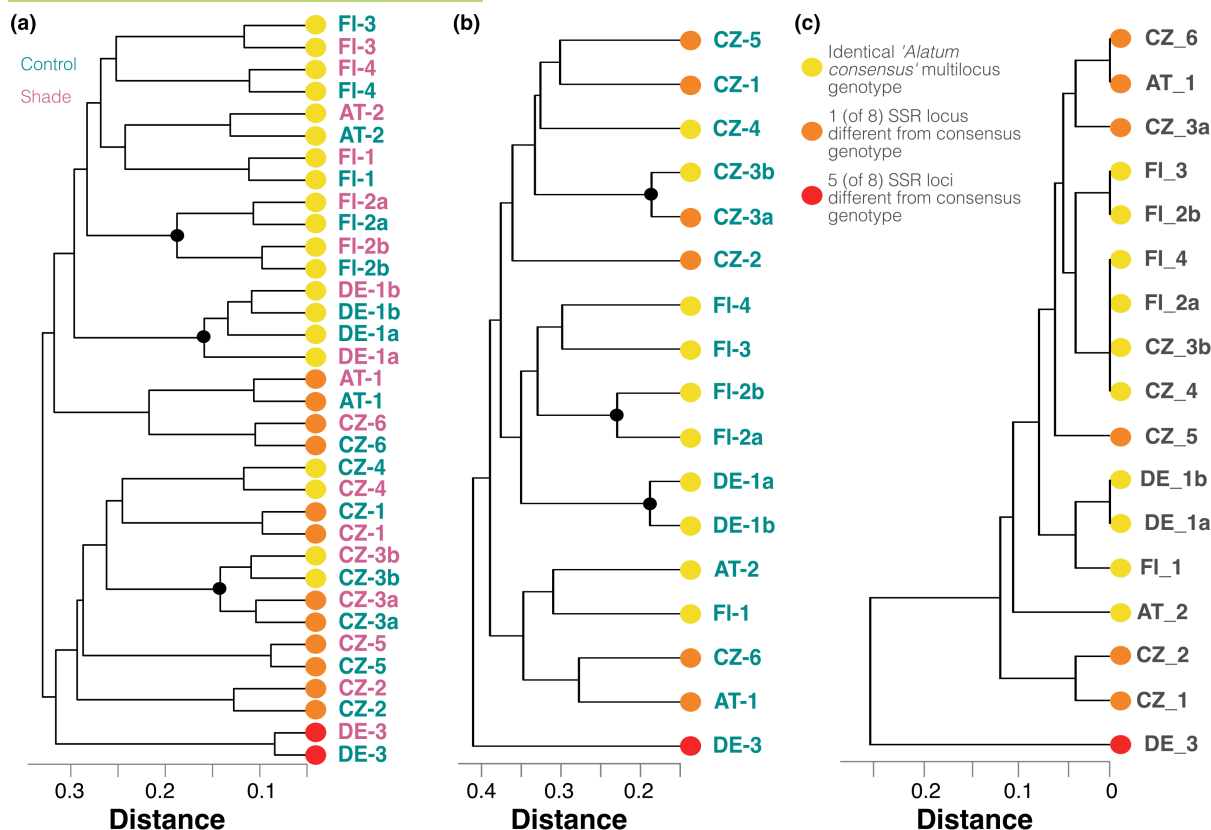


FIGURE 5 Genealogy of the *Taraxacum officinale* accessions from *alatum* lineage described by DNA methylation and genetic variation. UPGMA dendrograms based on Euclidean distances calculated from cytosine methylation levels in CG context using (a) all experimental samples or (b) using only control samples; and (c) based on Nei distances calculated from variation at eight microsatellite loci. For methylation-based dendrograms, only the subset of differentially methylated cytosines (p -value < 0.05 , before FDR correction) were included in the analysis. Black dots indicate different accessions collected from the same field location. Twenty accessions were collected in Germany (DE), Finland (FI), Austria (AT) and Czech Republic (CZ).

TABLE 3 CG methylation is correlated with genetic distance between accessions. Correlation coefficients and statistical significance are from Mantel tests between SSR-based distance matrix and methylation-based distance matrices from control samples.

Feature	CG		CHG		CHH	
	All Cs	DMCs	All Cs	DMCs	All Cs	DMCs
All features	0.61** (22868)	0.40* (261)	0.42* (16438)	0.42* (57)	0.37 (82622)	-0.03
Gene	0.43* (1932)	0.19 (12)	0.25 (1438)	- ^a	0.36 (6403)	- ^a
TE	0.48* (5877)	0.53* (5)	0.32 (4096)	- ^a	0.29 (19440)	- ^a
Repeats	0.65** (1161)	-0.16 (25)	0.49** (793)	0.32 (4)	0.37 (4745)	- ^a
Unclassified	0.62** (13898)	0.40 (219)	0.45* (10066)	0.39 (53)	0.38 (52034)	-0.03

Note: Correlations were tested for each feature and sequence context, and also considering all cytosines or only differentially methylated cytosines (DMCs at p -values < 0.05 before FDR correction). The numbers of cytosines analysed are indicated in parentheses.

* p -value < 0.05 ; ** p -value ≤ 0.01 ; *** p -value ≤ 0.001 .

^aNo or few C's were differentially methylated, thus, they were not analysed.

does not require that the observed CG methylation variants are caused by genetic differences between the accessions. While genetic differences between accessions (such as novel transposable element inserts) can be a potent source of DNA methylation differences between accessions, there are also frequent CG epimutations that arise spontaneously (that is, without requiring a change in DNA sequence) and that are stable over generations

(Shahryary et al., 2020; Yao et al., 2021). The patterns of DNA methylation variation that such epimutations produce between clonal accessions are expected to correlate with genetic mutations that also arise in the system. However, the correlation between CG methylation-based distances and SSR-based distances was far from perfect. Possibly, and notwithstanding that most environmental effects on DNA methylation are observed in non-CG

contexts, the environmental effects that do occur in CG context produce a heritable signal in CG methylation profiles that partly dilutes the genealogical signal. Our study shows that methylation information can be used to reconstruct micro-evolutionary and genealogical relationships in plants (Figure 5), providing a useful tool in systems that lack genetic variation such as clonal lineages and crops.

Our observation that current environmental effects on DNA methylation are detectable in both CG and in non-CG methylation and are biased towards transposable elements and repeats, adds to growing evidence from different plant species that the hypo- or hyper-methylation of transposable element (TE) is a key feature of plant stress responses (e.g. Dowen et al., 2012; Eichten & Springer, 2015; Niederhuth et al., 2016; Secco et al., 2015; Van Dooren et al., 2020; Wibowo et al., 2016). TE methylation is associated with the RdDM pathway (RNA-dependent DNA methylation), which affects cytosine methylation in all sequence contexts (Deniz et al., 2019; Sigman & Slotkin, 2016). However, most of the TE-associated methylation changes in response to environmental stress are observed specifically in CHH contexts (Dubin et al., 2015; Niederhuth et al., 2016; Van Dooren et al., 2020; Wibowo et al., 2016). Methylation in this context is not stable; it may capture recent or current environmental differences (Van Dooren et al., 2020; Wibowo et al., 2016) but it is not expected to be confounded with historical or genealogical effects. Notably, however, the shading treatment also triggered DMCs in CG context (see Table 1). In fact, the strongest enrichment of DMCs in response to the treatment was observed in CG context in repeats (see Table 1). Because methylation in this context is transgenerationally stable, such environment-induced CG methylation modifications may leave a long-term signature in the methylome of the apomictic lineage. This suggests that CG methylation profiles in plant genomes are shaped by accumulation of heritable epimutations (Yao et al., 2021), as well as, by persistent effects of historical environmental exposures (Kawakatsu et al., 2016). Unlike methylation in CHH context, which lacks an inherited component, current environments are not a dominant determinant of observed patterns of CG methylation, leading to the absence of a significant treatment effect on the overall CG methylome (Figure 2).

A clear result of our study is that the information contained in overall CG methylomes accurately reconstructs the genealogical structure of the dandelion accessions within the apomictic lineage. This is possible due to the high transgenerational stability of CG methylation (Yao et al., 2021). Both spontaneous epimutations that are transgenerationally stable and also epigenetic variants that are downstream consequences of genetic variants between samples (such as transpositions) should correlate well with genealogy, just like stable genetic variants do (Hagmann et al., 2015). However, the transgenerational stability of CG methylation is not absolute: epimutations revert frequently, and the dynamics of epimutation and reversal can uncouple methylation variation from patterns of underlying genetic variation in sexual organisms (Shahryari et al., 2020; van der Graaf et al., 2015). In addition,

in agreement with other recent studies (Galanti et al., 2022), we also detected environmental effects on CG methylation (Table 1). These might leave an imprint of historical environments, not only current environments, in CG methylomes, which could obscure the genealogical signal in CG methylomes. It is interesting that overall CG methylomes performed better in capturing the genealogical structure of the dandelion accessions (as benchmarked against microsatellite information) than the subset of CG-cytosines that showed significant between-accession methylation differences. This might be because pairwise similarities are based on fewer cytosines, potentially introducing more noise. However, the result is unexpected, because the subset of DMCs emphasizes methylation information that differentiates between the accessions. A possible explanation for this could be that CG-DMCs contain heritable methylation variants that were induced by local environments in the natural fields from which the accessions were collected, which are confounded with accessions (but not necessarily with genealogy) because accessions were collected from very different localities. In that case, the cumulative effect of many spontaneous epimutations could provide a better reflection of genealogy than the subset of DMCs with between-accession effects.

One caveat with the interpretation of CHH methylation results is that accurate estimation of methylation levels is prone to error and bias when methylation levels are low, which is typical for CHH methylation (Niederhuth et al., 2016). Accurate methylation estimation is hampered in such cases by low sequencing coverage. In our study, after applying the 10x coverage threshold, mean and median cytosine coverage for CHH context were 65x and 34x, respectively, which is relatively high and suitable for methylation estimation also when methylation levels are low. Accordingly, the CHH methylation responses to treatment that we revealed (specifically: the overall methylation level, a bias in the response to transposable elements, and clear environmental effects but no accession effects) are consistent with reports in other species (Niederhuth et al., 2016), indicating that the epiGBS approach is suitable for capturing patterns of DNA methylation variation also in CHH context.

In addition to the possible contribution of epigenetic variability in phenotypic plasticity and/or variation in heritable traits, new insights into the use of epigenetics are emerging. Given that the epigenetic mutation rate is higher than the genetic one (Ossowski et al., 2010; van der Graaf et al., 2015), it could provide phylogenetic resolution over time scales that cannot be easily tracked using genetic variants, mostly systems with low genetic diversity. As proposed by Yao et al. (2021), the neutral nature of spontaneous epimutations could be used as a molecular clock. This could be especially useful in species that are propagated clonally, and/or in which the phylogenetic relationships are difficult to infer with classical neutral genetic markers. For instance, a clonally propagated sterile crop such as saffron (*Crocus sativus*), in which the origin and relationships between close species have recently been discovered (Nemati et al., 2019; Schmidt et al., 2019), could benefit from the use of this molecular clock, as differences between accessions could be best explained by DNA methylation (Busconi et al., 2021).

AUTHOR CONTRIBUTIONS

Verónica Noé Ibañez: Conceptualization (supporting); data curation (lead); formal analysis (lead); project administration (lead); visualization (lead); writing – original draft (lead). **Morgane van Antro:** Conceptualization (supporting); methodology (supporting); writing – review and editing (supporting). **Cristian Peña-Ponton:** Conceptualization (supporting); methodology (supporting); writing – review and editing (supporting). **Slavica Milanovic-Ivanovic:** Investigation (lead); methodology (supporting); writing – review and editing (supporting). **Cornelis A. M. Wagemaker:** Investigation (supporting); writing – review and editing (supporting). **Fleur Gawehns:** Methodology (supporting); software (supporting); writing – review and editing (supporting). **Koen J. F. Verhoeven:** Conceptualization (lead); formal analysis (supporting); methodology (supporting); supervision (lead); writing – review and editing (lead).

ACKNOWLEDGEMENTS

The authors thank Mark Schmid for his contribution to the data analysis, Alejandro Edera for his technical assistance and Jan Kirschner for collecting the plant samples. This work was partially supported by the Universidad Nacional de Cuyo under the travel grant number 770/2018-C.S and Agencia Nacional de Promoción Científica y Tecnológica Nacional (PICT 2020-1767). The work was also supported by the European Training Network “EpiDiverse” which received funding from the EU Horizon 2020 program under Marie Skłodowska-Curie grant agreement No 764965.

CONFLICT OF INTEREST STATEMENT

The authors have no conflict of interest to declare.

PEER REVIEW

The peer review history for this article is available at <https://publons.com/publon/10.1111/jeb.14162>.

DATA AVAILABILITY STATEMENT

The demultiplexed sequencing data generated in this work can be accessed through the European Nucleotide Archive (ENA) under project PRJEB56325. The raw sequencing data, R scripts and configuration file used for the epiGBS pipeline, as well as the output generated from it, are deposited into Zenodo (<https://doi.org/10.5281/zenodo.7566483>). The source code and details used are freely available at <https://zenodo.org/badge/latestdoi/10.5281/zenodo.7566483>.

ORCID

Verónica Noé Ibañez  <https://orcid.org/0000-0001-6153-9286>

Morgane van Antro  <https://orcid.org/0000-0002-7788-4866>

Cristian Peña-Ponton  <https://orcid.org/0000-0003-0671-2124>

Cornelis A. M. Wagemaker  <https://orcid.org/0000-0002-8968-4546>

Fleur Gawehns  <https://orcid.org/0000-0002-9236-966X>

Koen J. F. Verhoeven  <https://orcid.org/0000-0003-3002-4102>

REFERENCES

- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B*, 57, 289–300.
- Boyko, A., Blevins, T., Yao, Y., Golubov, A., Bilichak, A., Ilnytsky, Y., Hollander, J., Meins, F., & Kovalchuk, I. (2010). Transgenerational adaptation of Arabidopsis to stress requires DNA methylation and the function of dicer-like proteins. *PLoS One*, 5, e9514.
- Buchfink, B., Xie, C., & Huson, D. H. (2015). Fast and sensitive protein alignment using DIAMOND. *Nature Methods*, 12, 59–60.
- Busconi, M., Wischnitzki, E., Del Corvo, M., Colli, L., Soffritti, G., Stagnati, L., Fluch, S., Sehr, E. M., de Los Mozos Pascual, M., & Fernández, J. A. (2021). Epigenetic variability among saffron crocus (*Crocus sativus* L.) accessions characterized by different phenotypes. *Frontiers in Plant Science*, 12, 1–4.
- Cao, X., & Jacobsen, S. E. (2002). Locus-specific control of asymmetric and CpNpG methylation by the DRM and CMT3 methyltransferase genes. *Proceedings of the National Academy of Sciences*, 99, 16491–16498.
- Cokus, S. J., Feng, S., Zhang, X., Chen, Z., Merriman, B., Haudenschild, C. D., Pradhan, S., Nelson, S. F., Pellegrini, M., & Jacobsen, S. E. (2008). Shotgun bisulphite sequencing of the Arabidopsis genome reveals DNA methylation patterning. *Nature*, 452, 215–219.
- Cubas, P., Vincent, C., & Coen, E. (1999). An epigenetic mutation responsible for natural variation in floral symmetry. *Nature*, 401, 157–161.
- Deniz, Ö., Frost, J. M., & Branco, M. R. (2019). Regulation of transposable elements by DNA modifications. *Nature Reviews Genetics*, 20, 417–431.
- Dowen, R. H., Pelizzola, M., Schmitz, R. J., Lister, R., Dowen, J. M., Nery, J. R., Dixon, J. E., & Ecker, J. R. (2012). Widespread dynamic DNA methylation in response to biotic stress. *Proceedings of the National Academy of Sciences*, 109, 2183–2191.
- Dubin, M. J., Zhang, P., Meng, D., Remigereau, M.-S., Osborne, E. J., Casale, F. P., Drewe, P., Kahles, A., Jean, G., Vilhjálmsson, B., Jagoda, J., Irez, S., Voronin, V., Song, Q., Long, Q., Ratsch, G., Stegle, O., Clark, R. M., & Nordborg, M. (2015). DNA methylation in Arabidopsis has a genetic basis and shows evidence of local adaptation. *eLife*, 4, e05255.
- Eichten, S. R., & Springer, N. M. (2015). Minimal evidence for consistent changes in maize DNA methylation patterns following environmental stress. *Frontiers in Plant Science*, 6, 308.
- Finnegan, E. J., & Kovac, K. A. (2000). Plant DNA methyltransferases. In M. A. Matzke & A. J. M. Matzke (Eds.), *Plant gene silencing* (pp. 69–81). Springer.
- Galanti, D., Ramos-Cruz, D., Nunn, A., Rodríguez-Arévalo, I., Scheepens, J. F., Becker, C., & Bosse, O. (2022). Genetic and environmental drivers of large-scale epigenetic variation in *Thlaspi arvense*. *PLoS genetics*, 18(10), e1010452. <https://doi.org/10.1371/journal.pgen.1010452>
- Ganguly, D. R., Crisp, P. A., Eichten, S. R., & Pogson, B. J. (2017). The Arabidopsis DNA methylome is stable under transgenerational drought stress. *Plant Physiology*, 175, 1893–1912.
- Gawehns, F., Postuma, M., van Antro, M., Nunn, A., Sepers, B., Fatma, S., van Gurp, T. P., Wagemaker, N. C. A. M., Mateman, A. C., Milanovic-Ivanovic, S., Große, I., van Oers, K., Vergeer, P., & Verhoeven, K. J. F. (2022). epiGBS2: Improvements and evaluation of highly multiplexed, epiGBS-based reduced representation bisulfite sequencing. *Molecular Ecology Resources*, 22, 2087–2104.
- Groot, M. P., Kooke, R., Knob, N., Vergeer, P., Keurentjes, J. J. B., Joop Ouborg, N., & Verhoeven, K. (2016). Effects of multi-generational stress exposure and offspring environment on the expression and persistence of transgenerational effects in *Arabidopsis thaliana*. *PLoS One*, 11, e0151566.
- Hagmann, J., Becker, C., Müller, J., Stegle, O., Meyer, R. C., Wang, G., Schneeberger, K., Fitz, J., Altmann, T., Bergelson, J., Borgwardt, K.,

- & Weigel, D. (2015). Century-scale methylome stability in a recently diverged *Arabidopsis thaliana* lineage. *PLoS Genetics*, 11, e1004920.
- Johannes, F., & Schmitz, R. J. (2019). Spontaneous epimutations in plants. *New Phytologist*, 221, 1253–1259.
- Kawakatsu, T., Huang, S.-S. C., Jupe, F., Sasaki, E., Schmitz, R. J., Urlich, M. A., Castanon, R., Nery, J. R., Barragan, C., He, Y., Chen, H., Dubin, M., Lee, C. R., Wang, C., Bemm, F., Becker, C., O'Neil, R., O'Malley, R. C., Quarless, D. X., ... Zhou, X. (2016). Epigenomic diversity in a global collection of *Arabidopsis thaliana* accessions. *Cell*, 166, 492–505.
- Kirschner, J., Oplaat, C., Verhoeven, K. J. F., Zeisek, V., Uhleman, I., Trávníček, B., Räsänen, J., Wilschut, R. A., & Štěpánek, J. (2016). Identification of oligoclonal agamosperous microspecies: Taxonomic specialists versus microsatellites. *Preslia*, 88, 1–17.
- Kirschner, J., & Štěpánek, J. (1994). Clonality as a part of the evolution process in *Taraxacum*. *Folia Geobotanica et Phytotaxonomica*, 29, 265–275.
- Kovel, C. G. F., Jong, G., & Kovel, C. G. F. (1999). Responses of sexual and apomictic genotypes of *Taraxacum officinale* to variation in light. *Plant Biology*, 1, 541–546.
- Law, J. A., & Jacobsen, S. E. (2010). Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nature Reviews Genetics*, 11, 204–220.
- Lindroth, A. M., Cao, X., Jackson, J. P., Zilberman, D., McCallum, C. M., Henikoff, S., & Jacobsen, S. E. (2001). Requirement of CHROMOMETHYLASE3 for maintenance of CpXpG methylation. *Science*, 292, 2077–2080.
- Lister, R., O'Malley, R. C., Tonti-Filippini, J., Gregory, B. D., Berry, C. C., Millar, A. H., & Ecker, J. R. (2008). Highly integrated single-base resolution maps of the epigenome in *Arabidopsis*. *Cell*, 133, 523–536.
- Molina-Montenegro, M. A., Peñuelas, J., Munné-Bosch, S., & Sardans, J. (2012). Higher plasticity in ecophysiological traits enhances the performance and invasion success of *Taraxacum officinale* (dandelion) in alpine environments. *Biological Invasions*, 14, 21–33.
- Nemati, Z., Harpke, D., Gemicioglu, A., Kerndorff, H., & Blattner, F. R. (2019). Saffron (*Crocus sativus*) is an autotriploid that evolved in Attica (Greece) from wild *Crocus cartwrightianus*. *Molecular Phylogenetics and Evolution*, 136, 14–20.
- Niederhuth, C. E., Bewick, A. J., Ji, L., Alabady, M. S., Kim, K. D., Li, Q., Rohr, N. A., Rambani, A., Burke, J. M., Udall, J. A., Egesi, C., Schmutz, J., Grimwood, J., Jackson, S. A., Springer, N. M., & Schmitz, R. J. (2016). Widespread natural variation of DNA methylation within angiosperms. *Genome Biology*, 17, 194.
- Oksanen, J. (2007). Vegan: Community ecology package. R package version 1.8-5. <http://www.cran.r-project.org>
- Ossowski, S., Schneeberger, K., Lucas-Lledó, J. I., Warthmann, N., Clark, R. M., Shaw, R. G., Weigel, D., & Lynch, M. (2010). The rate and molecular spectrum of spontaneous mutations in *Arabidopsis thaliana*. *Science*, 327, 92–94.
- Park, Y., & Wu, H. (2016). Differential methylation analysis for BS-seq data under general experimental design. *Bioinformatics*, 32, 1446–1453.
- Paun, O., Verhoeven, K. J. F., & Richards, C. L. (2019). Opportunities and limitations of reduced representation bisulfite sequencing in plant ecological epigenomics. *New Phytologist*, 221, 738–742.
- Quadrana, L., & Colot, V. (2016). Plant transgenerational epigenetics. *Annual Review of Genetics*, 50, 467–491.
- R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.
- Richards, C. L., Alonso, C., Becker, C., Bossdorf, O., Bucher, E., Colomé-Tatché, M., Durka, W., Engelhardt, J., Gaspar, B., Gogol-Döring, A., Grosse, I., van Gurp, T. P., Heer, K., Kronholm, I., Lampe, C., Latzel, V., Mirouze, M., Opgenoorth, L., Paun, O., ... Verhoeven, K. J. F. (2017). Ecological plant epigenetics: Evidence from model and non-model species, and the way forward. *Ecology Letters*, 20, 1576–1590.
- Schmid, M. W., Heichinger, C., Coman Schmid, D., Guthörl, D., Gagliardini, V., Bruggmann, R., Aluri, S., Aquino, C., Schmid, B., Turnbull, L. A., & Grossniklaus, U. (2018). Contribution of epigenetic variation to adaptation in *Arabidopsis*. *Nature Communications*, 9, 4446.
- Schmidt, T., Heitkam, T., Liedtke, S., Schubert, V., & Menzel, G. (2019). Adding color to a century-old enigma: Multi-color chromosome identification unravels the autotriploid nature of saffron (*Crocus sativus*) as a hybrid of wild *Crocus cartwrightianus* cytotypes. *New Phytologist*, 222, 1965–1980.
- Schmitz, R. J., Schultz, M. D., Lewsey, M. G., O'Malley, R. C., Urlich, M. A., Libiger, O., Schork, N. J., & Ecker, J. R. (2011). Transgenerational epigenetic instability is a source of novel methylation variants. *Science*, 334, 369–373.
- Secco, D., Wang, C., Shou, H., Schultz, M. D., Chiarenza, S., Nussbaum, L., Ecker, J. R., Whelan, J., & Lister, R. (2015). Stress induced gene expression drives transient DNA methylation changes at adjacent repetitive elements. *eLife*, 4, 1–26.
- Shahryari, Y., Symeonidi, A., Hazarika, R. R., Denkena, J., Mubeen, T., Hofmeister, B., van Gurp, T., Colomé-Tatché, M., Verhoeven, K. J. F., Tuskan, G., Schmitz, R. J., & Johannes, F. (2020). AlphaBeta: Computational inference of epimutation rates and spectra from high-throughput DNA methylation data in plants. *Genome Biology*, 21, 260.
- Sigman, M. J., & Slotkin, R. K. (2016). The first rule of plant transposable element silencing: Location, location, location. *Plant Cell*, 28, 304–313.
- Suter, L., & Widmer, A. (2013a). Environmental heat and salt stress induce transgenerational phenotypic changes in *Arabidopsis thaliana*. *PLoS One*, 8, e60364.
- Suter, L., & Widmer, A. (2013b). Phenotypic effects of salt and heat stress over three generations in *Arabidopsis thaliana*. *PLoS One*, 8, e80819.
- Tempel, S. (2012). Using and understanding RepeatMasker. *Methods in Molecular Biology*, 859, 29–51.
- Turner, S. (2021). Package “qqman”: QQ and Manhattan plots for GWAS data. R Package Version 0.1.
- Underwood, C. J., Vijverberg, K., Rigola, D., Okamoto, S., Oplaat, C., Camp, R. H. M., Radoeva, T., Schauer, S. E., Fierens, J., Jansen, K., Mansveld, S., Busscher, M., Xiong, W., Datema, E., Nijbroek, K., Blom, E.-J., Bicknell, R., Catanach, A., Erasmus, S., ... Pri, M. (2022). A PARTHENOGENESIS allele from apomictic dandelion can induce egg cell division without fertilization in lettuce. *Nature Genetics*, 54, 84–93.
- van der Graaf, A., Wardenaar, R., Neumann, D. A., Taudt, A., Shaw, R. G., Jansen, R. C., Schmitz, R. J., Colomé-Tatché, M., & Johannes, F. (2015). Rate, spectrum, and evolutionary dynamics of spontaneous epimutations. *Proceedings of the National Academy of Sciences*, 112, 6676–6681.
- van Dooren, T. J. M., Silveira, A. B., Gilbault, E., Jiménez-Gómez, J. M., Martin, A., Bach, L., Tisné, S., Quadrana, L., Loudet, O., & Colot, V. (2020). Mild drought in the vegetative stage induces phenotypic, gene expression, and DNA methylation plasticity in *Arabidopsis* but no transgenerational effects. *Journal of Experimental Botany*, 71, 3588–3602.
- van Gurp, T. P., Wagemaker, N. C. A., Wouters, B., Vergeer, P., Ouborg, J. N. J., & Verhoeven, K. J. F. (2016). epiGBS: Reference-free reduced representation bisulfite sequencing. *Nature Methods*, 13, 322–324.
- van Moorsel, S. J., Schmid, M. W., Wagemaker, N. C. A., Gurp, T., Schmid, B., & Vergeer, P. (2019). Evidence for rapid evolution in a grassland biodiversity experiment. *Molecular Ecology*, 28, 4097–4117.
- Venables, W. N., & Ripley, B. D. (2002). *Modern applied statistics with S*. Springer Science & Business Media.
- Wibowo, A., Becker, C., Marconi, G., Durr, J., Price, J., Hagmann, J., Papareddy, R., Putra, H., Kageyama, J., Becker, J., Weigel, D., & Gutierrez-Marcos, J. (2016). 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*, 5, 1–27.

- Wilschut, R. A., Oplaat, C., Basten Snoek, L., Kirschner, J., & Verhoeven, K. J. F. (2016). Natural epigenetic variation contributes to heritable flowering divergence in a widespread asexual dandelion lineage. *Molecular Ecology*, 25, 1759–1768.
- Yao, N., Schmitz, R. J., & Johannes, F. (2021). Epimutations define a fast-ticking molecular clock in plants. *Trends in Genetics*, 37, 699–710.
- Zhang, H., Lang, Z., & Zhu, J.-K. (2018). Dynamics and function of DNA methylation in plants. *Nature Reviews Molecular Cell Biology*, 19, 489–506.
- Zhang, Y.-Y., Latzel, V., Fischer, M., & Bossdorf, O. (2018). Understanding the evolutionary potential of epigenetic variation: A comparison of heritable phenotypic variation in epiRILs, RILs, and natural ecotypes of *Arabidopsis thaliana*. *Heredity*, 121, 257–265.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Ibañez, V N., van Antro, M., Peña-Ponton, C., Milanovic-Ivanovic, S., Wagemaker, C A M., Gawehns, F., & Verhoeven, K J F. (2023). Environmental and genealogical effects on DNA methylation in a widespread apomictic dandelion lineage. *Journal of Evolutionary Biology*, 36, 663–674. <https://doi.org/10.1111/jeb.14162>