

DNA methylation in the wild: epigenetic transgenerational inheritance can mediate adaptation in clones of wild strawberry (*Fragaria vesca*)

Iris Sammarco (✉ Iris.Sammarco@ibot.cas.cz)

Czech Academy of Sciences <https://orcid.org/0000-0002-4101-6223>

Bárbara Díez Rodríguez

Philipps-University

Dario Galanti

University of Tübingen

Adam Nunn

ecSeq Bioinformatics GmbH <https://orcid.org/0000-0002-9276-6243>

Claude Becker

Genetics, LMU Biocenter, Ludwig-Maximilians-Universität München

Oliver Bossdorf

University of Tübingen

Zuzana Münzbergová

Charles University

Vít Latzel

Institute of Botany of Czech Academy of Sciences <https://orcid.org/0000-0003-0025-5049>

Article

Keywords:

Posted Date: March 14th, 2023

DOI: <https://doi.org/10.21203/rs.3.rs-2642365/v1>

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

[Read Full License](#)

Abstract

Due to the accelerating climate change, it is crucial to understand how plants adapt to rapid environmental changes. Such adaptation may be mediated by epigenetic mechanisms like DNA methylation, which could heritably alter phenotypes without changing the DNA sequence, especially across clonal generations. However, we are still missing robust evidence of the adaptive potential of DNA methylation in wild clonal populations.

Here, we studied the methylome of *Fragaria vesca*, a predominantly clonally reproducing herb. We examined samples from 21 natural populations across three climatically distinct geographic regions, as well as clones of the same individuals grown in a common garden. We found that inherited epigenetic variation was partly associated with climate of origin, and that a subset of these epigenetic changes affected gene expression. Our results indicate that DNA methylation variation in the wild is common and may contribute to adaptation of clonal plant populations.

1. Introduction

Given their sessile nature, plants must be able to quickly adapt to changing environments. Among other mechanisms, plants can adjust their phenotypes through epigenetic alterations of gene expression, for example via DNA methylation¹. In plants, DNA methylation occurs in three sequence contexts: CG, CHG and CHH (where H is A, C or T)², which have different functions³. In all sequence contexts, DNA methylation represses transposon (TE) mobilization⁴, whereas CG methylation alone usually represses gene expression when present in gene promoters, and CHH methylation is sometimes associated with active genes^{5–7}. Importantly, part of DNA methylation is mitotically or even meiotically heritable, thus potentially affecting offspring phenotypes⁸. As inherited phenotypes are the ultimate targets of natural selection⁹, it is widely speculated that epigenetic variation, particularly that induced by environmental factors, may confer adaptive potential^{10,11}. However, solid evidence is still missing. To improve our understanding of the adaptive potential of epigenetic mechanisms, we need to assess whether epigenetic variation meets the requirements for evolution by natural selection. We thus need to quantify natural variation in DNA methylation also in plants grown in their natural environments, the extent of heritability of environmentally induced variation, and its fitness effects.

The exploration of environmentally induced epigenetic variation under natural conditions is a complex task, since epigenetic variation can have genetic, environmental and stochastic sources^{12–17}. Genetically determined DNA methylation variants are based on genetic modifications¹⁸, which can be *cis*-acting (*i.e.* when a TE inserted upstream a gene promoter drives the methylation of the promoter itself)¹⁹, or *trans*-acting (*i.e.* when genetic mutations in genes involved in the DNA methylation machinery induce overall changes in DNA methylation patterns)^{14,16,20}. Environmentally induced DNA methylation variants, in contrast, are under the sole control of environmental cues¹⁸. They can thus arise quickly in response to environmental stimuli^{12,21} and potentially contribute to rapid adaptation^{11,22}. However, most previous

studies on natural plant populations were not able or did not attempt to detect the underlying sources of epigenetic variation, because of inappropriate experimental design, such as lack of field and common garden environments, or low-resolution molecular methods^{23–25}. Clearly, the ability to discriminate between genetically determined and environmentally induced heritable epigenetic variation is key for understanding the role of epigenetic variation in the environmental adaptation of plants.

Clonality is a predominant reproductive strategy in many ecosystems, whereby genetically identical offspring are produced, resulting in a reduction in genetic diversity²⁶. Nonetheless, epigenetic variation might be particularly relevant for the success and survival of clonal species, since it could to some degree compensate for their low standing genetic variation^{27–34}. In addition, the inheritance of environmentally induced DNA methylation variation may be particularly strong across clonal generations, which lack meiosis and the associated epigenetic resetting that typically leads to erasure of many environmentally induced epigenetic variants^{35,36}. However, studies assessing the extent of epigenetic variation, and its heritability and environmental associations, in natural clonal plant populations are still scarce^{17,37–39}.

To fill this gap, we investigated the epigenetic variation, its heritability and its environmental associations in natural populations of a widespread clonal species, the wild strawberry (*Fragaria vesca*). We analyzed the methylomes of 231 plants collected from 21 natural populations across multiple geographic regions, including different natural habitats in three European countries. We also examined the DNA methylation patterns of clonal offspring grown in a common garden to determine the extent of epigenetic inheritance and its potential impact on gene expression. Specifically, we asked:

1. Do natural populations of *F. vesca* harbor DNA methylation variation in the wild? If so, is such variation associated with climate of origin?
2. Is the climate-associated epigenetic variation inherited across clonal generations?
3. Does this heritable epigenetic variation have a functional role, *i.e.* does it modulate gene expression?

2. Results

We studied 21 natural populations of *F. vesca* in Italy, Czechia and Norway (Fig. 1a), with seven populations distributed along a climatic gradient in each country. We transplanted seven plants per population of origin, *i.e.* a total of 147 plants, into a common garden (hereafter referred to as garden conditions), where we clonally propagated them for one year (Fig. S1). We then performed whole genome bisulfite sequencing (WGBS) of leaf tissue from clonal offspring of at least the third generation, as well as for a subset of their wild progenitors (hereafter referred to as *field conditions*; 4 plants per population, total = 84 plants). We used all the garden plants data for a subsequent GWA analysis, as well as for inferring single nucleotide polymorphisms (SNPs), while some other analyses were based on only those samples for which we had data from both field and garden conditions (n = 84) (see materials and methods). For a subset of the garden individuals (3 plants per population, n = 63) we also performed RNA-sequencing.

(Epi)genetic variation

To describe overall genetic, transcriptome and epigenetic variation within and among the studied populations, we performed a principal component analysis (PCA) for genetic variants (SNPs) (Fig. 1b), gene expression (Fig. 1c) and DNA methylation (Fig. 1d-f). Overall, the plants appeared to cluster more clearly by country of origin than by temperature (Fig. S2a-e) or precipitation of origin (Fig. S2f-j). For DNA methylation, these geographic clusters were strong, and almost equal in the field and garden, in the CG and CHG contexts (Fig. 1d, e), while the separation was much weaker in CHH, which on the other hand seemed more strongly influenced by the growth conditions (field vs. garden; Fig. 1f). To assess the proportions of variance explained by country, temperature and precipitation of origins, we performed redundancy analysis (RDA) for DNA methylation, SNPs and gene expression. The analysis of DNA methylation also included growth conditions as a factor (Table S1). We found that country explained the highest proportion of variance (CG: 11.9%, CHG: 9.4%, CHH: 2.5%, SNPs: 8.7%, gene expression: 9.7%), while growth conditions generally explained very little (CG: 0.4%, CHG: 1.6%, CHH: 1.2%).

To quantify methylation differences at the genomic-region level, we identified differentially methylated regions (DMRs) for all pairwise comparisons between populations from the same growth condition (*i.e.* we compared field populations to other field populations, and garden populations to other garden populations). After merging overlapping DMRs identified in each pairwise comparison, we identified over 344,000 DMRs in the field (CG = 82,675, CHG = 49,600, CHH = 211,735) and almost 249,000 in the garden (CG = 71,972, CHG = 37,925, CHH = 139,097). In both growth conditions, CG-DMRs were most frequent in gene bodies, whereas CHG- and CHH-DMRs were predominantly associated with promoters and in particular transposable elements (TEs) (Fig. 2). The numbers of CHH-DMRs in promoters and TEs were much higher in the field than in the garden, while CG- and CHG- DMRs were similarly abundant in the two datasets.

Association of epigenetic variation with genetic and climatic variation

In order to assess the association of epigenetic variation with genetic and environmental (climatic), we performed a DMR variance decomposition analysis using genetic variation in *cis* and *trans*, as well as climatic variation (mean, minimum and maximum temperatures and precipitation) as predictors. For each DMR, we ran three independent mixed models including a distance matrix derived from each of the three predictors, as in¹⁶. We classified each DMR according to what the strongest predictor of its variance was. If no predictor explained >10% of the variance, the DMR was classified as unexplained¹⁶.

We found that in all sequence contexts and for both field and garden conditions, DMR variation was generally best predicted by genetic variation in *trans*, followed by climatic variation and *cis* genetic variation (Fig. 3a, b). The fraction of climate-predicted DMR variation gradually increased from CG to CHG and CHH in the field, but only from CG to CHG, and decreased again in CHH, in garden conditions. Overall, for both field and garden, the genomic density of predicted CG-DMRs in *trans* largely followed the distribution of genes (Fig. 3c; Fig. S3a), while *trans*- and particularly climate-predicted DMRs in CHG and

CHH mainly followed the distribution of TEs (Fig. 3d, e; Fig. S3b, c). Accordingly, the number of predicted CG-DMRs in *trans* correlated positively with the number of genes and negatively with the number of TEs, and the opposite was true for predicted CHG- and CHH-DMRs in *trans* and climate-predicted DMRs in all contexts (Table S2).

Since many DMRs overlapped with genes (Fig. S4), we then performed a Gene Ontology (GO) enrichment analysis to functionally characterize genes containing predicted DMRs in *cis*, *trans* and climate-predicted in their promoters. We found enrichment for several GO terms, but only for CHG- and CHH- but not CG-DMRs (Fig. 4; Table S3). For the predicted DMRs in *cis*, we found enrichment in two cellular component terms related to chromatin and chromatin remodeling activity ("nucleosome"; "Ino8 complex"), and several biological process terms. However, we observed only a few enriched terms for both predicted-DMRs in *trans* and climate-predicted DMRs, some of which were common between the two ("RNA-DNA hybrid ribonuclease activity"; "retrotransposon nucleocapsid"). Interestingly, we also found "retrotransposon nucleocapsid" among the unexplained DMRs.

Correlation of climate-predicted DMRs with gene expression

Climate-predicted DMRs could in principle be under *direct* or *indirect* control of the climates of origin. The strongest indication of *direct* environmental induction is in cases where climate-predicted DMRs are only associated with environmental factors but not DNA sequence. If climate-predicted DMRs are also associated with DNA sequence variation, then the link between environment and epigenetic variation could be indirect, through selection acting on genetic variants first. In this study we were interested in the potentially independent adaptive potential of epigenetic variation, and therefore mainly in the *directly* climate-associated DMRs. To identify the subset of climate-predicted DMRs with a genetic basis and exclude them from further analysis, we ran a genome-wide association (GWA) analysis. For this analysis, we used only the garden samples, to ensure that the selected climate-predicted DMRs were heritable and thus of evolutionary potential. To increase the statistical power of the analysis, we included all the 147 garden samples. We performed GWA using individual DMR-promoter methylation as phenotypes and SNPs as predictors and selected the threshold *P*-value using Bonferroni correction. Out of 2,092 CG-, 3,049 CHG- and 8,186 CHH-climate-predicted DMRs overlapping promoters, we found significant GWA hits for 62.8% of the climate-predicted DMRs in CG, 30.8% in CHG, and 18.8% in CHH (Fig. S5). We thus classified these DMRs as *indirectly* environmentally associated, as opposed to those *directly* linked to environment for which we found no significant association with SNP variation in the GWA analyses.

To explore whether the putative *directly* environmentally linked DMRs likely had a functional role, we tested whether methylation levels of individual DMRs were correlated with the expression of their overlapping genes. We found statistically significant correlations in 11.4% of the cases in CG, 10.4% in CHG and 10.4% in CHH (corresponding to 65 genes in CG, 89 genes in CHG and 411 genes in CHH). Overall, promoter methylation (Fig. 5a-c) and expression (Fig. 5d) of these genes clustered samples according to their country and climate of origin. We found both positive and negative correlations between promoter methylation of genes and their expression levels (Fig. 5e; Table S4). To assess whether

the genes with significant correlations were enriched in particular functions, we performed GO enrichment analyses for all, and also separately for those with positive or negative correlations (Table S5). In all cases, we found many enriched terms related to ribosome-related processes, especially in CG and CHG contexts, and many related to metabolic processes and recognition of pollen, especially in CHH. Among the positive correlations in the CG context, we also found enrichment for the GO term “methylation”.

3. Discussion

There is growing interest in the effects of environmental variation on plant DNA methylation, the inheritance of environmentally induced methylation variation across generations, and its effects on phenotypes and plant adaptation^{10,11,21}. However, clear evidence is still scarce, particularly from natural conditions. Here, we tested whether climate of origin is associated with DNA methylation variation in natural plant populations of wild strawberry, whether such DNA methylation variation is stable across clonal generations, and whether it correlates with gene expression.

The analyzed populations harbored comparable genetic and epigenetic geographic structure (Fig. 1; Table S1), suggesting that the observed epigenetic variation was largely genetically determined (as already reported in *F. vesca*)^{38,39}. However, in addition to the differentially methylated regions (DMRs) related to DNA sequence variation, we also identified DMRs that were related to climatic variation (Fig. 3a). Interestingly, the numbers of these climate-associated DMRs gradually increased from CG to CHG and CHH contexts, suggesting that non-CG methylation was particularly sensitive to climatic conditions (see also)^{16,17}. A GO enrichment analysis of DMR-overlapping gene promoters showed that *cis*-genetic variants induced mainly DNA methylation variants in genes related to chromatin and chromatin remodeling, telomere maintenance, and metabolic processes, while *trans*-genetic and climatic variation affected DNA methylation in genes related to RNA-DNA hybrid ribonuclease activity and retrotransposon nucleocapsid (Fig. 4). Interestingly, since RNA-DNA hybrids and retrotransposon nucleocapsid are related to retrotransposon mobilization⁴², it is likely that both *trans*-genetic and climatic variation modulate transposition and that the environment might thus control TE mobilization in wild conditions^{20,43}.

The comparison of epigenetic patterns of plants in the field with their genetically identical clonal offspring in a common garden allowed us to assess the inheritance of environmentally associated DNA methylation and its role in gene regulation. Overall, DNA methylation levels were similar between field and garden samples (Fig. S6), and the epigenetic structure was significantly different only in CHG and CHH, indicating that methylation in these contexts is more sensitive to current environmental conditions than in CG. We also found the greatest differences in the numbers of DMRs between field and garden conditions in CHG and CHH for both gene- and TE-related DMRs (and for CG TE-DMRs too), with a generally higher number of DMRs in the field than in the garden (Fig. 2). DNA methylation in all contexts plays a crucial role in TE silencing and in the establishment of heterochromatin^{44,45}, and CHH also plays an important role in gene regulation in euchromatic regions^{46–48}. The higher number of gene- and TE-related DMRs in the field suggests that natural environmental conditions likely play a role in regulating gene expression

and TE mobilization, which is in accordance with our previous finding that the environment might control TE mobilization in wild conditions (see above). This could have important evolutionary implications, since TE mobilization is known to play a key role in the evolution of plant genomes and thus speciation^{49,50}.

In order to distinguish whether the observed heritable DNA methylation variation was associated with DNA sequence variation and/or environments of origin, we performed the DMR variance decomposition analysis also for the plants grown in garden conditions (Fig. 3b). Our assumption was that climate-associated DMRs maintained under common environmental conditions are likely adaptive, because otherwise one should not find such (non-random) patterns of climate association. Interestingly, we found similar amounts of climate-associated DMRs in the field and common garden, especially in CG and CHG contexts. The CHH context, in contrast, showed decreased climate-associated variation but a large increase in unexplained variation under garden conditions. 60% of these unexplained DMRs overlapped with DMRs associated with climate in the field (data not shown), suggesting that the unexplained DMRs in CHH were due to the environmental conditions in the common garden, and that CHH methylation is the least stable across clonal generations and/or the most responsive to short-term environmental changes. On the other hand, the similar CG- and CHG-DMR variation for field and garden, indicates that the climates of origin induced DMR variation in these contexts that was heritable across clonal generations.

Our findings are corroborated by a recent study of pennycress (*Thlaspi arvense*) where similar extents of climate-predicted DMRs were observed in natural accessions grown in greenhouse conditions¹⁶. Also in pennycress the contribution of climatic variation to DMR variation increased from CG to CHG and CHH, and the CHH context harboured the greatest amount of unexplained variance. Although the pennycress study also found that *trans*-genetic variation explained most DMR variation, *cis*-genetic variation explained a greater fraction of DMR variation than in our study (~5-14% in *T. arvense*, 0.2-2% in *F. vesca*). This could be due to the higher standing genetic variation of the sexually-reproducing *T. arvense*⁵¹ in comparison to the mostly clonally-reproducing *F. vesca* in natural conditions⁵².

Finally, part of the climate-associated DMRs identified under common garden conditions (*i.e.* inherited) were significantly correlated with gene expression. Although this correlation was significant only for a fraction of the inherited DMRs, this is still an intriguing result, as many previous studies failed to find correlations between promoter methylation and gene expression⁵³. Interestingly, our analysis found both positive and negative correlations with gene expression (Fig. 5e). Although promoter methylation is usually negatively associated with gene expression⁵, there are also some reports of positive effects of promoter methylation on expression⁵⁴, especially in the CHH context^{6,7,47,55}. Maybe we were able to capture substantially more variation than previous studies testing similar correlations among different genes in the same individual⁷, because we tested for such correlations between DNA methylation and expression of the same genes across a multitude of samples.

In a complementary study, we conducted a reciprocal transplant experiment using a subset of the same populations, which showed that DNA methylation may be involved in local adaptation³⁴. This finding supports our conclusion that the changes in gene expression observed in our study are likely to have a significant impact on plant fitness. Our results also highlight the role of environmentally induced heritable epigenetic variation in modulating gene expression and its contribution to phenotypic variation, which can serve as a substrate for natural selection.

To better understand the role of epigenetic changes in local adaptation and evolutionary processes, we need to investigate other plant species in natural populations with different life-history traits, including clonal and sexual reproduction. Direct evidence of the impact of epigenetic modifications on fitness would also provide valuable insights into the role of epigenetic variation in evolution. Such studies could expand our understanding of how environmental and genetic variation shape natural plant populations.

4. Materials And Methods

Study Species

Fragaria vesca L., Rosaceae, is an herbaceous perennial species with wide geographic distribution (Europe, northern Asia, North America, and northern Africa)⁵⁶. It reproduces both clonally through stolons and sexually through seeds, although its sexual reproduction is very rare in natural conditions⁵².

Plant collection and growth

We selected 21 natural populations of *F. vesca* from three European countries, Italy, Czechia and Norway (see Table S6 for geographic locations and climatic characteristics of the selected populations) between May and July 2018. We chose these countries as they represented the southern limit (Italy), the core (Czechia) and the northern limit (Norway) of the native range of *F. vesca* distribution in Europe. To increase the environmental difference among the populations' sites, we sampled the populations following a climatic (mostly corresponding to altitudinal) gradient within each country.

For each population, we collected mature, fully developed leaves of 4 individuals directly from the field conditions (n = 84), we dried them in silica gel and used them for whole genome bisulfite sequencing (WGBS) analysis, see later. We then dug up the same ramets plus additional three (n = 147) and planted them individually following a random block design in 70 × 40 × 20 cm pots filled with a commercial mixture of compost and sand located in the common garden of the Institute of Botany of the Czech Academy of Sciences in Průhonice, Czechia (49.994°N, 14.566°E) one to ten days after their collection (see Table S6 for the climatic characteristics of the common garden). Plants were grown under a shading coverage reducing 50% of the light to simulate natural light levels at most of the localities. We let the plants propagate clonally for one year. Then, we selected the biggest offspring ramet of at least the third generation from every clone and we collected mature, fully developed leaf samples and froze them immediately in liquid nitrogen. These samples were later used for WGBS (n = 147). From a subset of 3

plants per population ($n = 63$), we also collected mature leaf samples for RNA-sequencing in the same way as samples for WGBS.

WGBS library preparation and sequencing

We extracted genomic DNA from individual plants using the Qiagen DNeasy Plant Mini Kit, following the manufacturer's instructions with minor modifications. We prepared libraries for WGBS using the NEBNext Ultra II DNA Library Prep Kit and EZ-96 DNA Methylation-Gold MagPrep (ZYMO), and we sequenced paired-end reads on HiSeq X Ten (Illumina, San Diego, CA), using a sequencing coverage per sample of 30x. See Appendix I for more information on the DNA extraction and library preparation protocols.

All the sequencing data (WGBS and RNA-sequencing) can be found in the European Nucleotide Archive (ENA, www.ebi.ac.uk/ena/), under the project PRJEB51609.

Methylation and DMR calling

We used the EpiDiverse WGBS pipeline for bisulfite reads mapping and methylation calling (<https://github.com/EpiDiverse/wgbs>), which was specifically designed for non-model plant species (*i.e.* species that have not been extensively studied)⁵⁷. The pipeline performed quality control (FastQC), base quality and adaptor trimming (cutadapt), bisulfite-aware mapping (erne-bs5) and non-conversion rate calculation, duplicates detection (Picard MarkDuplicates), alignment statistics and methylation calling (MethylDackel). In the mapping step, we used the most recent version of the genome of *F. vesca* v4.0.a2^{58,59}. On average, the sequencing produced 97,142,389 reads per sample (see Table S7 for detailed information), of which 96% mapped successfully to the genome after retaining only uniquely-mapping reads. We calculated the bisulfite non-conversion rate using the chloroplast genome, which is naturally unmethylated⁶⁰, and we found an average bisulfite non-conversion rate among samples of 0.38% (see Table S7). We obtained individual bedGraph files of methylated positions for each sample and sequence context.

For PCA and RDA analyses, we then combined the individual bedGraph files from both field and garden conditions in a multisample unionbed file using custom scripts and bedtools⁶¹. In order to compare the field with the garden conditions, we used only the samples for which we had WGBS data for both conditions ($n = 84$ per condition). We retained all the cytosines having coverage ≥ 5 in at least 80% of the samples (total methylated positions: 1,644,729, 2,574,494 and 12,335,916, respectively for CG, CHG and CHH). We then performed PCA on Hellinger-transformed methylated positions, and using custom scripts with the R function `prcomp` in the stats package (v3.5.1)⁶², and colored the PCAs using either country of origin, mean temperature or precipitation averaged over 7 years before the sampling year (2011-2018). To test for significance of the differences among countries, growth conditions, climatic conditions of origin of the plants, we performed RDA with the RDA function in the vegan package (v2.6.4)⁶³. For DNA methylation, we performed four separate RDA analyses, using in all of them Hellinger-transformed methylated positions as independent variable and in 1) country of origin as predictor and growth

conditions (field, garden) as a covariate to account for the effect of growth condition on the plants' methylomes; 2) growth conditions as predictor and country as a covariate; 3-4) mean temperature or precipitation averaged over 7 years before the sampling year (2011-2018) as predictors and growth conditions as a covariate. We tested the statistical significance of the RDA analyses using a permutation test with 499 permutations.

We called DMRs using the EpiDiverse DMR pipeline (<https://github.com/EpiDiverse/dmr>)⁵⁷ and using the DMR caller metilene with default parameters⁶⁴. We used populations as groups, and we called DMRs separately for all the pairwise comparisons between the populations from the field, and the populations from the garden (*i.e.* we never compared a field population with a garden population). We used as input individual bedGraph files filtered for cytosine coverage ≥ 5 . Separately for field and garden conditions, we then combined the output bed files in a multisample bed file using custom scripts and bedtools (v2.27.1)⁶¹, and we merged the overlapping DMRs obtained from all the pairwise comparisons with bedtools. We used only the samples for which we had WGBS data for both conditions ($n = 84$ per condition). We obtained 82,546 CG-DMRs, 49,459 CHG-DMRs and 211,363 CHH-DMRs for field, and 71,856 CG-DMRs, 37,795 CHG-DMRs and 138,807 CHH-DMRs for garden.

To assess the number of DMRs overlapping with gene promoters, gene bodies and TEs, we intersected these bed files with gene and TE annotations. For genes, we used the gene annotations v4.0.a2 downloaded from the Genome Database for Rosaceae (GDR) (https://www.rosaceae.org/species/fragaria_vesca/genome_v4.0.a2)⁶⁵, while for TEs we used an annotation carried out using the EDTA annotation pipeline v1.9.6⁶⁶ on the substituted genome using default parameters, kindly provided by⁶⁷.

SNP calling

We inferred Single Nucleotide Polymorphisms (SNPs) from WGBS data using the EpiDiverse SNP pipeline with default parameters (<https://github.com/epidiverse/snp>)^{57,68}. For the DMR variance decomposition analysis, separately for field and garden conditions, we then combined the output individual VCF files into multisample VCF files using BCFtools (v1.9)⁶⁹. As above, we used only the samples for which we had WGBS data for both conditions ($n = 84$ per condition). Using VCFtools (v0.1.16)⁶⁹, we filtered the variants successfully genotyped in 80% of individuals, with a minimum quality score of 30 and a minimum mean depth of 3.

For PCA and RDA analyses, we combined the individual VCF files from both field and garden ($n = 84$ per condition) into a multisample VCF file and performed the same filtering steps as above. We also filtered for Minor Allele Frequency (MAF) ≥ 0.05 , and pruned for Linkage Disequilibrium (LD) with an LD threshold (r^2) of 0.2 for SNP pairs in a sliding window of 50 SNPs, sliding by 5. After filtering, we were able to retain 76 669 SNPs. We plotted the PCAs with custom scripts in R, using Hellinger-transformed SNPs. We performed RDA analysis similar to methylation (see above), and using Hellinger-transformed SNPs as independent variables.

DMR variance decomposition analysis and GO enrichment

To assess the amount of methylation variance explained by *cis*-variants, *trans*-variants and climatic variation, we ran three mixed models for each individual DMR for both field and garden conditions, as in¹⁶. Briefly, for *cis*-variants, we used an Isolation-By-State (IBS) matrix generated with PLINK (v1.90b6.12) (<http://pngu.mgh.harvard.edu/purcell/plink/>)⁷⁰ using variants within 50kb from the DMR middle point. For *trans*-variants, we used an IBS matrix obtained from variants filtered for Minor Allele Frequency (MAF) ≥ 0.01 and pruned for Linkage Disequilibrium (LD) with an LD threshold (r^2) of 0.8 for SNP pairs in a sliding window of 50 SNPs, sliding by 5. For climatic variation, for both field and garden conditions, we calculated a Euclidean distance matrix between climatic data from all the field sites, which we reversed and normalized to obtain a similarity matrix in a 0 to 1 range. The climatic data included mean, maximum and minimum temperature, and precipitation, all averaged over 7 years before the sampling year (2011-2018). We sourced climatic data at a resolution of $0.1 \times 0.1^\circ$ (v20.0e) from the European gridded dataset E-OBS, available through the C3S Climate Data Store (CDS) website (<https://cds.climate.copernicus.eu/cdsapp#!/home>)⁷¹. For methylation variants, we use the merged DMRs obtained from all the pairwise comparisons with bedtools, separately for field and garden (see above). As above, we used only the samples for which we had WGBS data for both conditions (n = 84 per condition). We extracted average methylation of the resulting DMRs from all the samples with the function *regionCounts* from the R package methylKit (v1.16.1)⁷², using a minimum cytosine coverage of 5.

We plotted circos plots using the R package circlize (v0.4.9)⁷³. We performed the correlation analysis between number of *cis*-, *trans*-, climate-, unexplained-DMRs and number of genes and TEs using the Pearson correlation method. For this analysis, we calculated the DMR, gene and TE counts assigned to 1-kb genomic bins, and performed the correlation between DMR count and gene or TE count. We then ran a GO enrichment analysis for *cis*-, *trans*-, climate- and unexplained-predicted DMRs, separately for each sequence context and for field and garden conditions. We extracted DMR-related gene promoters with bedtools, and performed a GO enrichment analysis using the R package clusterProfiler (v3.18.1)⁷⁴ with an FDR-adjusted P value < 0.05.

Genome-wide association (GWA) analysis

To assess the putative genetic basis of the climate-predicted DMRs, we ran GWA analysis for the garden conditions, including all the available samples (7 plants per population, n = 147) to increase the statistical power of the analysis. We ran GWA analysis as described in¹⁶, using the R package rrBLUP (4.6.1)⁷⁵. For genetic variants, we imputed the missing genotype calls with BEAGLE 5.2 (Browning et al., 2018) and filtered for MAF > 0.04. After filtering, we were able to retain 83,095 SNPs. We corrected for population structure using an IBS matrix obtained from variants filtered for MAF ≥ 0.01 and pruned for Linkage Disequilibrium (LD) with an LD threshold (r^2) of 0.8 for SNP pairs in a sliding window of 50 SNPs, sliding by 5. We used individual average DMR-promoter methylation for each sequence context as

phenotype, calculated with the *regionCounts* methylKit function (v1.16.1)⁷², using a minimum cytosine coverage of 5. We performed GWA for all the climate-predicted DMRs overlapping promoters, and we used the Bonferroni correction method to select the threshold P value. We then excluded the climate-predicted DMRs overlapping promoters with significant GWA peaks from the analysis testing for correlation between DMR methylation and gene expression.

RNA-sequencing and correlation of climate-predicted DMRs with gene expression

We collected mature leaf samples from 3 randomly selected plants per population from garden condition (total samples: n = 63), and we snap-froze them in liquid nitrogen. We extracted mRNA using the Nucleospin RNA Plus kit (Macherey Nagel), following the manufacturer's instructions with minor modifications. To improve RNA quality and yield from *F. vesca*, we used an increased amount of lysis buffer (500 µl) together with 100 µl of EDTA (0.5 M, pH=8) and PVPP (polyvinylpolypyrrolidone). The cDNA library and sequencing (PE150, 6 Gb per sample of raw data) were performed by Novogene Co., Ltd, Cambridge, using an Illumina NovaSeq 6000 platform. On average, we obtained 22,252,025 raw reads. We trimmed adaptors with cutadapt (v1.16) and assessed sequencing quality with MultiQC (v1.10.1)⁷⁶. We aligned the reads to the *Fragaria vesca* genome (v4.0.a2) using STAR (Spliced Transcripts Alignment to a Reference) (v2.7.1a)⁷⁷, assembled them into transcripts and quantified using StringTie (v2.1.5)⁷⁸.

For PCA and RDA analyses, we normalized the read counts with the R function DESeq in the DESeq2 package (v1.30.1)⁷⁹ and applied a variance stabilizing transformation with the function vst from the same package. We performed PCA with custom scripts using the function plotPCA in the DESeq2 package using Hellinger-transformed read counts, and RDA with the RDA function in the vegan package, using Hellinger-transformed read counts as independent variable and country of origin as predictor. As for methylation and SNP data, we tested the statistical significance of the RDA analyses using a permutation test with 499 permutations.

For the correlation analysis between methylation and gene expression, for each sample, we normalized the Fragments Per Kilobase of transcript per Million mapped reads (FPKM) values and extracted the genes adjacent to the *directly* environmentally induced DMR-related promoters. For each of these DMRs, we combined DMR-promoter methylation, expression of the adjacent gene, sample ID and gene ID in the same file. We removed lines containing missing values in the gene expression data, and performed correlation analysis with the remaining genes (572, 856 and 3,955 genes in the CG, CHG and CHH contexts, respectively). We performed Spearman correlation between methylation and expression of each individual gene, and selected only the genes with a statistically significant correlation ($P < 0.05$). We considered these correlations statistically significant, as they were twofold more than the ones expected by chance (at P -value = 0.05, corresponding to 29, 43 and 198 random significant correlations in the CG, CHG and CHH contexts, respectively). We plotted heatmaps using the R function pheatmap in the pheatmap package (v1.0.12)⁸⁰. For methylation, we used the methylation values of the climate-predicted DMRs presenting statistically significant correlation with gene expression, and for expression we used the normalized FPKM values of those genes.

Declarations

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

FUNDING

The study was supported by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 764965, the Czech Science Foundation (GACR 20-00871S), the Austrian Academy of Sciences (ÖAW), and partly by institutional research project RVO 67985939.

ACKNOWLEDGMENTS

We thank members of the EpiDiverse consortium (www.epidiverse.eu) for valuable inputs during preparation and execution of the study, and Katharina Jandrasits (GMI) and Daniela Ramos Cruz for help with the WGBS library preparation. Computational resources were supplied by the project "e-Infrastruktura CZ" (e-INFRA CZ LM2018140) supported by the Ministry of Education, Youth and Sports of the Czech Republic.

References

1. Riggs, A. D. & Porter, T. N. Overview of Epigenetic Mechanisms. *Cold Spring Harb. Monogr. Arch.* **32**, 29–45 (1996).
2. Finnegan, E. J., Genger, R. K., Peacock, W. J. & Dennis, E. S. DNA METHYLATION IN PLANTS. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **49**, 223–247 (1998).
3. Niederhuth, C. E. & Schmitz, R. J. Putting DNA methylation in context: from genomes to gene expression in plants. *Biochim. Biophys. Acta* **1860**, 149 (2017).
4. Zemach, A. & Zilberman, D. Evolution of Eukaryotic DNA Methylation and the Pursuit of Safer Sex. *Curr. Biol.* **20**, R780–R785 (2010).
5. Li, X. *et al.* Single-base resolution maps of cultivated and wild rice methylomes and regulatory roles of DNA methylation in plant gene expression. *BMC Genomics* **13**, 1–15 (2012).
6. Rajkumar, M. S., Shankar, R., Garg, R. & Jain, M. Bisulphite sequencing reveals dynamic DNA methylation under desiccation and salinity stresses in rice cultivars. *Genomics* **112**, 3537–3548 (2020).
7. Wang, G. *et al.* Analysis of Global Methylome and Gene Expression during Carbon Reserve Mobilization in Stems under Soil Drying. *Plant Physiol.* **183**, 1809–1824 (2020).
8. Niederhuth, C. E. & Schmitz, R. J. Covering Your Bases: Inheritance of DNA Methylation in Plant Genomes. *Mol. Plant* **7**, 472–480 (2014).

9. Darwin, C. & Wallace, A. On the Tendency of Species to form Varieties; and on the Perpetuation of Varieties and Species by Natural Means of Selection. *J. Proc. Linn. Soc. London. Zool.* **3**, 45–62 (1858).
10. Jablonka, E. V. A. & Raz, G. A. L. Transgenerational epigenetic inheritance: prevalence, mechanisms, and implications for the study of heredity and evolution. *Q. Rev. Biol.* **84**, 131–176 (2009).
11. Ashe, A., Colot, V. & Oldroyd, B. P. How does epigenetics influence the course of evolution? *Philos. Trans. R. Soc. B* **376**, (2021).
12. Zhang, Y.-Y., Fischer, M., Colot, V. & Bossdorf, O. Epigenetic variation creates potential for evolution of plant phenotypic plasticity. *New Phytol.* **197**, 314–322 (2013).
13. Zhang, Y.-Y., Latzel, V., Fischer, M. & Bossdorf, O. Understanding the evolutionary potential of epigenetic variation: a comparison of heritable phenotypic variation in epiRILs, RILs, and natural ecotypes of *Arabidopsis thaliana*. *Heredity (Edinb.)*. **121**, 257–265 (2018).
14. Dubin, M. J. *et al.* DNA methylation in *Arabidopsis* has a genetic basis and shows evidence of local adaptation. *Elife* **4**, (2015).
15. Johannes, F. & Schmitz, R. J. Spontaneous epimutations in plants. *New Phytol.* **221**, 1253–1259 (2019).
16. Galanti, D. *et al.* Genetic and environmental drivers of large-scale epigenetic variation in *Thlaspi arvense*. *PLOS Genet.* **18**, e1010452 (2022).
17. Díez Rodríguez, B. *et al.* Epigenetic variation in the Lombardy poplar along climatic gradients is independent of genetic structure and persists across clonal reproduction. *bioRxiv* 2022.11.17.516862 (2022) doi:10.1101/2022.11.17.516862.
18. Richards, E. J. Inherited epigenetic variation — revisiting soft inheritance. *Nat. Rev. Genet.* **2006** **757**, 395–401 (2006).
19. Martin, A. *et al.* A transposon-induced epigenetic change leads to sex determination in melon. *Nature* **461**, 1135–1138 (2009).
20. Baduel, P. *et al.* Genetic and environmental modulation of transposition shapes the evolutionary potential of *Arabidopsis thaliana*. *Genome Biol.* **22**, 1–26 (2021).
21. Thiebaut, F., Hemerly, A. S. & Ferreira, P. C. G. A role for epigenetic regulation in the adaptation and stress responses of non-model plants. *Frontiers in Plant Science* vol. 10 (2019).
22. Miryeganeh, M. & Saze, H. Epigenetic inheritance and plant evolution. *Popul. Ecol.* **62**, 17–27 (2020).
23. Zoldoš, V. *et al.* Epigenetic Differentiation of Natural Populations of *Lilium bosniacum* Associated with Contrasting Habitat Conditions. *Genome Biol. Evol.* **10**, 291–303 (2018).
24. Medrano, M., Alonso, C., Bazaga, P., López, E. & Herrera, C. M. Comparative genetic and epigenetic diversity in pairs of sympatric, closely related plants with contrasting distribution ranges in south-eastern Iberian mountains. *AoB Plants* **12**, (2020).
25. Miryeganeh, M., Marlétaz, F., Gavriouchkina, D. & Saze, H. De novo genome assembly and in natura epigenomics reveal salinity-induced DNA methylation in the mangrove tree *Bruguiera gymnorhiza*.

New Phytol. **233**, 2094–2110 (2022).

26. Klimeš, L., Klimešová, J., Hendriks, R. J. J. & Groenendaal, J. Clonal plant architecture A comparative analysis of form and function. *Ecol. Evol. clonal plants* (1997).
27. Latzel, V. & Klimešová, J. Transgenerational plasticity in clonal plants. *Evol. Ecol.* **24**, 1537–1543 (2010).
28. Verhoeven, K. J. F. & Preite, V. Epigenetic variation in asexually reproducing organisms. *Evolution (N. Y.)*. **68**, 644–655 (2014).
29. Dodd, R. S. & Douhovnikoff, V. Adjusting to Global Change through Clonal Growth and Epigenetic Variation. *Front. Ecol. Evol.* **4**, (2016).
30. Latzel, V., Rendina González, A. P. & Rosenthal, J. Epigenetic Memory as a Basis for Intelligent Behavior in Clonal Plants. *Front. Plant Sci.* **7**, 1354 (2016).
31. Rendina González, A. P., Preite, V., Verhoeven, K. J. F. & Latzel, V. Transgenerational Effects and Epigenetic Memory in the Clonal Plant *Trifolium repens*. *Front. Plant Sci.* **9**, 1677 (2018).
32. Münzbergová, Z., Latzel, V., Šurinová, M. & Hadincová, V. DNA methylation as a possible mechanism affecting ability of natural populations to adapt to changing climate. *Oikos* **128**, 124–134 (2019).
33. Shi, W. *et al.* Transient Stability of Epigenetic Population Differentiation in a Clonal Invader. *Front. Plant Sci.* **9**, 1851 (2019).
34. Sammarco, I., Münzbergová, Z. & Latzel, V. DNA Methylation Can Mediate Local Adaptation and Response to Climate Change in the Clonal Plant *Fragaria vesca*: Evidence From a European-Scale Reciprocal Transplant Experiment. *Front. Plant Sci.* **13**, 435 (2022).
35. Feng, S., Jacobsen, S. E. & Reik, W. Epigenetic reprogramming in plant and animal development. *Science* **330**, 622–627 (2010).
36. Anastasiadi, D., Venney, C. J., Bernatchez, L. & Wellenreuther, M. Epigenetic inheritance and reproductive mode in plants and animals. *Trends Ecol. Evol.* **36**, 1124–1140 (2021).
37. Richards, C. L., Schrey, A. W. & Pigliucci, M. Invasion of diverse habitats by few Japanese knotweed genotypes is correlated with epigenetic differentiation. (2012) doi:10.1111/j.1461-0248.2012.01824.x.
38. De Kort, H. *et al.* Pre-adaptation to climate change through topography-driven phenotypic plasticity. *J. Ecol.* **108**, 1465–1474 (2020).
39. De Kort, H. *et al.* Signatures of polygenic adaptation align with genome-wide methylation patterns in wild strawberry plants. *New Phytol.* (2022) doi:10.1111/NPH.18225.
40. Pebesma, E. Simple features for R: Standardized support for spatial vector data. *R J.* **10**, 439–446 (2018).
41. South, A. naturalearth: World Map Data from Natural Earth. (2017).
42. Todd, L. A. *et al.* RNA-cDNA hybrids mediate transposition via different mechanisms. *Sci. Reports* **2020 101** **10**, 1–12 (2020).

43. Rey, O., Danchin, E., Mirouze, M., Loot, C. & Blanchet, S. Adaptation to Global Change: A Transposable Element-Epigenetics Perspective. *Trends in Ecology and Evolution* vol. 31 514–526 (2016).
44. Slotkin, R. K. & Martienssen, R. Transposable elements and the epigenetic regulation of the genome. *Nature Reviews Genetics* vol. 8 272–285 (2007).
45. Fultz, D., Choudury, S. G. & Slotkin, R. K. Silencing of active transposable elements in plants. *Curr. Opin. Plant Biol.* **27**, 67–76 (2015).
46. Zemach, A., McDaniel, I. E., Silva, P. & Zilberman, D. Genome-wide evolutionary analysis of eukaryotic DNA methylation. *Science* **328**, 916–919 (2010).
47. Gent, J. I. *et al.* CHH islands: de novo DNA methylation in near-gene chromatin regulation in maize. *Genome Res.* **23**, 628–637 (2013).
48. Martin, G. T., Seymour, D. K. & Gaut, B. S. CHH Methylation Islands: A Nonconserved Feature of Grass Genomes That Is Positively Associated with Transposable Elements but Negatively Associated with Gene-Body Methylation. *Genome Biol. Evol.* **13**, (2021).
49. Schmidt, A. L. & Anderson, L. M. Repetitive DNA elements as mediators of genomic change in response to environmental cues. *Biol. Rev. Camb. Philos. Soc.* **81**, 531–543 (2006).
50. Oliver, K. R. & Greene, W. K. Transposable elements: powerful facilitators of evolution. *Bioessays* **31**, 703–714 (2009).
51. Frels, K. *et al.* Genetic Diversity of Field Pennycress (*Thlaspi arvense*) Reveals Untapped Variability and Paths Toward Selection for Domestication. *Agron. 2019, Vol. 9, Page 302* **9**, 302 (2019).
52. Schulze, J., Rufener, R., Erhardt, A. & Stoll, P. The relative importance of sexual and clonal reproduction for population growth in the perennial herb *Fragaria vesca*. *Popul. Ecol. 2012 543* **54**, 369–380 (2012).
53. Ganguly, D. R., Crisp, P. A., Eichten, S. R. & Pogson, B. J. The Arabidopsis DNA Methylome Is Stable under Transgenerational Drought Stress. *Plant Physiol.* **175**, 1893–1912 (2017).
54. 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. U. S. A.* **114**, E4511–E4519 (2017).
55. Xu, J. *et al.* Single-base methylome analysis reveals dynamic epigenomic differences associated with water deficit in apple. *Plant Biotechnol. J.* **16**, 672–687 (2018).
56. Darrow, G. M. The strawberry. History, breeding and physiology. *strawberry. Hist. Breed. Physiol.* (1966).
57. Nunn, A. *et al.* EpiDiverse Toolkit: a pipeline suite for the analysis of bisulfite sequencing data in ecological plant epigenetics. *NAR Genomics Bioinforma.* **3**, (2021).
58. Edger, P. P. *et al.* Single-molecule sequencing and optical mapping yields an improved genome of woodland strawberry (*Fragaria vesca*) with chromosome-scale contiguity. *Gigascience* **7**, (2018).
59. Li, Y., Pi, M., Gao, Q., Liu, Z. & Kang, C. Updated annotation of the wild strawberry *Fragaria vesca* V4 genome. *Hortic. Res.* **6**, (2019).

60. Fojtová, M., Kovařík, A. & Matyášek, R. Cytosine methylation of plastid genome in higher plants. Fact or artefact? *Plant Sci.* **160**, 585–593 (2001).
61. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
62. Sigg, C. D. & Buhmann, J. M. Expectation-maximization for sparse and non-negative PCA. *Proc. 25th Int. Conf. Mach. Learn.* 960–967 (2008) doi:10.1145/1390156.1390277.
63. Oksanen Jari *et al.* CRAN - Package vegan. <https://cran.r-project.org/web/packages/vegan/index.html> (2020).
64. Jühling, F. *et al.* metilene: fast and sensitive calling of differentially methylated regions from bisulfite sequencing data. *Genome Res.* **26**, 256–262 (2016).
65. Jung, S. *et al.* 15 years of GDR: New data and functionality in the Genome Database for Rosaceae. *Nucleic Acids Res.* **47**, D1137 (2019).
66. Ou, S. *et al.* Benchmarking transposable element annotation methods for creation of a streamlined, comprehensive pipeline. *Genome Biol.* **20**, (2019).
67. López, M.-E., Roquis, D., Becker, C., Denoyes, B. & Bucher, E. DNA methylation dynamics during stress response in woodland strawberry (*Fragaria vesca*). *Hortic. Res.* **9**, (2022).
68. Nunn, A., Otto, C., Fasold, M., Stadler, P. F. & Langenberger, D. Manipulating base quality scores enables variant calling from bisulfite sequencing alignments using conventional bayesian approaches. *BMC Genomics* **23**, (2022).
69. Danecek, P. *et al.* Twelve years of SAMtools and BCFtools. *Gigascience* **10**, 1–4 (2021).
70. Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
71. Cornes, R. C., van der Schrier, G., van den Besselaar, E. J. M. & Jones, P. D. An Ensemble Version of the E-OBS Temperature and Precipitation Data Sets. *J. Geophys. Res. Atmos.* **123**, 9391–9409 (2018).
72. Akalin, A. *et al.* MethylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles. *Genome Biol.* **13**, 1–9 (2012).
73. Gu, Z., Gu, L., Eils, R., Schlesner, M. & Brors, B. circlize Implements and enhances circular visualization in R. *Bioinformatics* **30**, 2811–2812 (2014).
74. Yu, G., Wang, L. G., Han, Y. & He, Q. Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* **16**, 284–287 (2012).
75. Endelman, J. B. Ridge Regression and Other Kernels for Genomic Selection with R Package rrBLUP. *Plant Genome* **4**, 250–255 (2011).
76. Ewels, P., Magnusson, M., Lundin, S. & Käller, M. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* **32**, 3047–3048 (2016).
77. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
78. Kovaka, S. *et al.* Transcriptome assembly from long-read RNA-seq alignments with StringTie2. *Genome Biol.* **20**, 1–13 (2019).

79. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 1–21 (2014).
80. Kolde, R. Pheatmap: pretty heatmaps. *R Packag. version 1*, 726 (2012).

Figures

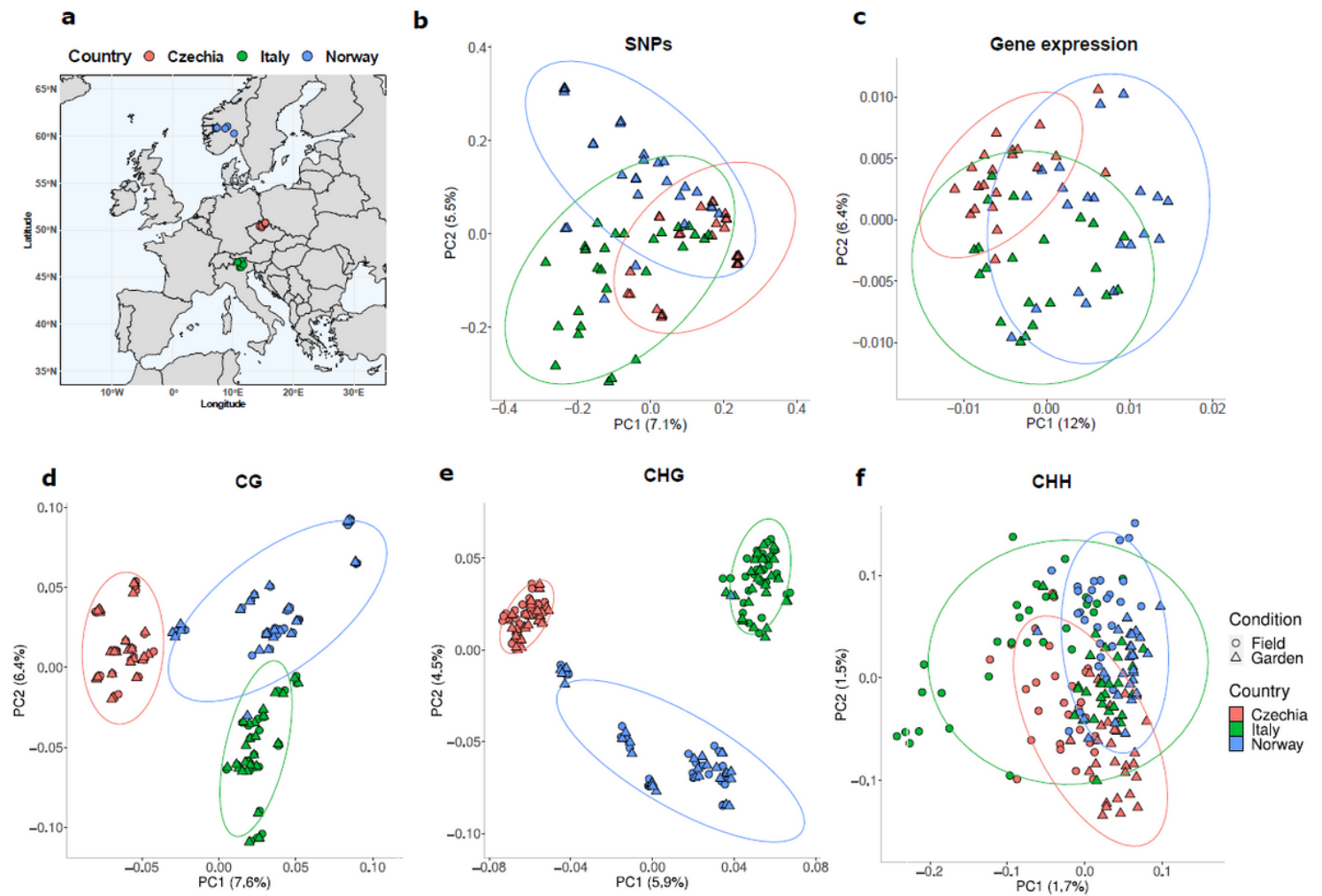


Figure 1

Map of the sampling locations and principal component analyses (PCA) for DNA methylation, genetic variants (SNPs) and gene expression. Sampling locations (R packages: sf (v1.0.9)⁴⁰ and rnaturalearth (v0.1.0)⁴¹) (a), SNPs (b), gene expression (c) and CG (d), CHG (e) and CHH (f) methylated positions. For SNPs we show only the garden plants, as these are clones of the field ones and thus are genetically identical. For methylated positions, field plants: n = 84, garden plants: n = 84. For SNPs, garden plants: n = 84. For gene expression, garden plants: n = 63.

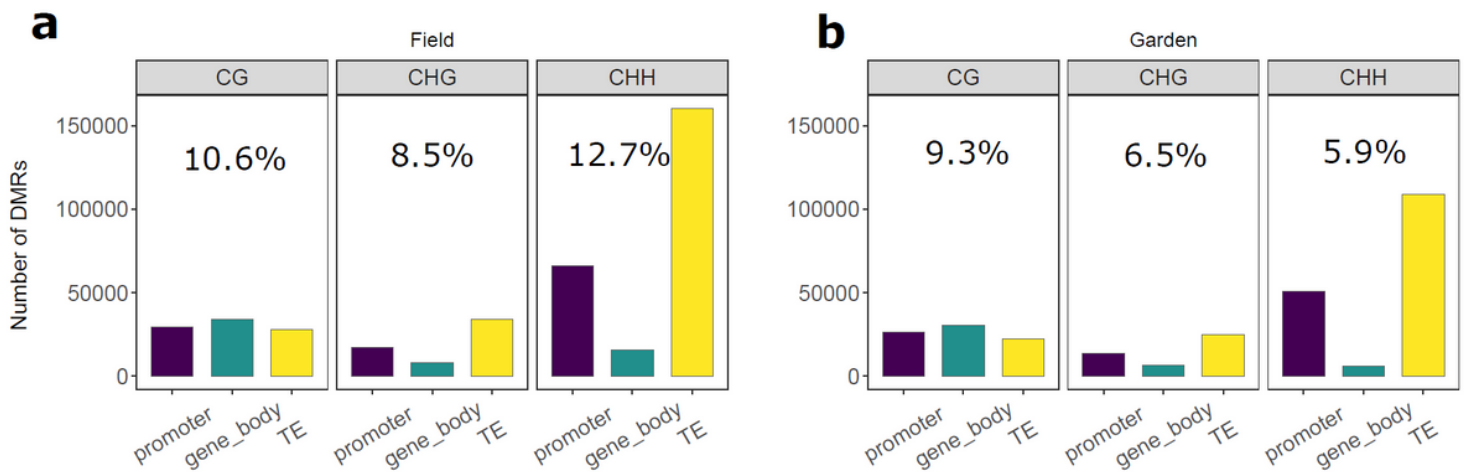


Figure 2

Total numbers of differentially methylated regions (DMRs) identified in different genomic contexts. (a) Plants in the field ($n = 84$). **(b)** Plants in the garden ($n = 84$). The number in each panel shows the frequencies of DMRs in the different sequence contexts.

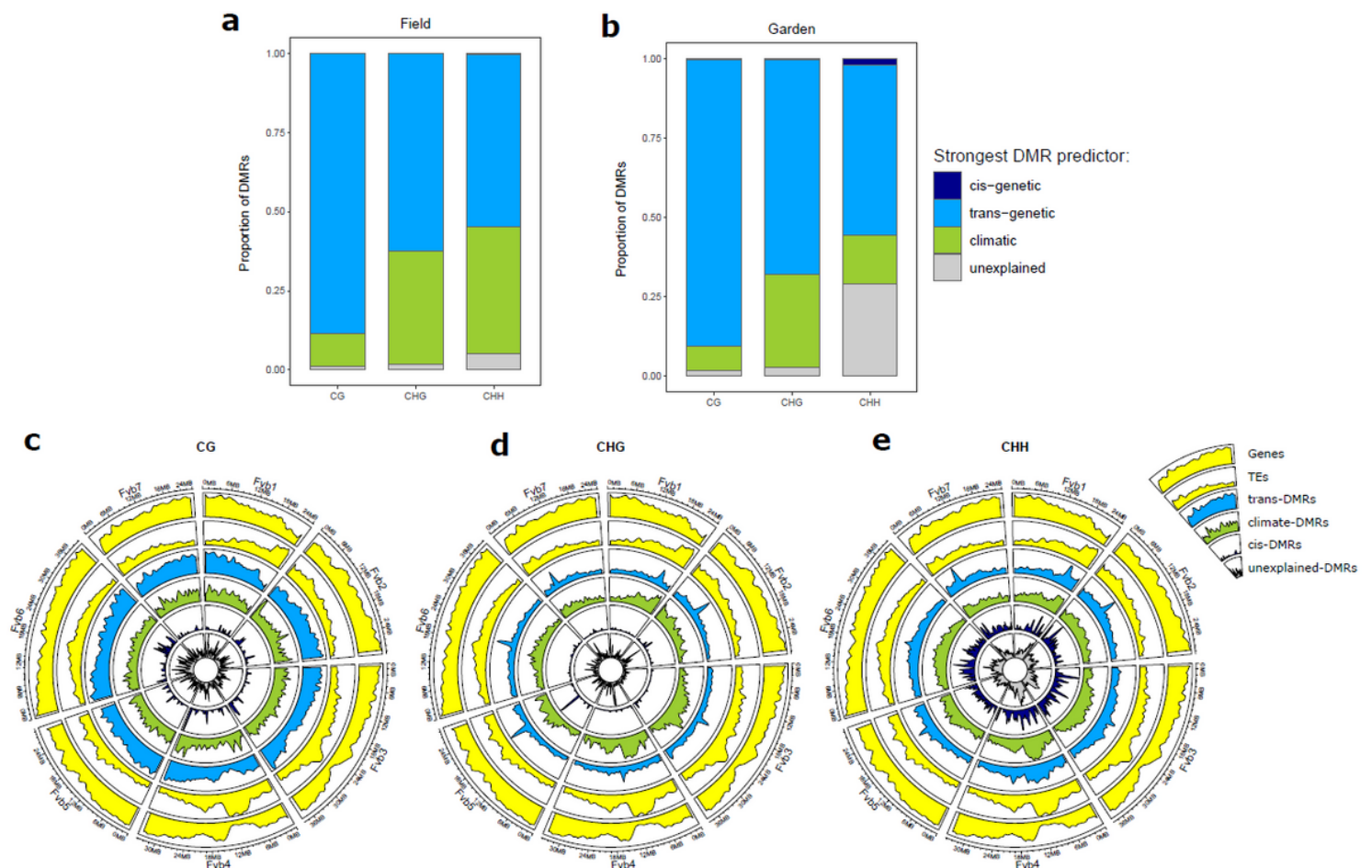


Figure 3

Strongest predictors of differentially methylated regions (DMRs). The proportions of DMRs in different sequence contexts that best predicted by either *cis*-genetic, *trans*-genetic or climatic variation. DMRs with <10% variance explained by all three predictors are classified as “unexplained”. **(a)** Field, total DMRs: 82,546 CG, 49,459 CHG and 211,363 CHH (n = 84 plants). **(b)** Garden, total DMRs: 71,856 CG, 37,795 CHG and 138,807 CHH (n = 84 plants). **(c-e)** Circos plots visualising the densities of field-DMRs, genes and TEs for all chromosomes (Fvb1-7) in CG **(c)**, CHG **(d)** and CHH **(e)** contexts. From outer to inner circles: genes and TE annotations (yellow), predicted DMRs in *trans*(light blue), climate-predicted DMRs (green), predicted DMRs in *cis* (dark blue) and unexplained DMRs (grey).

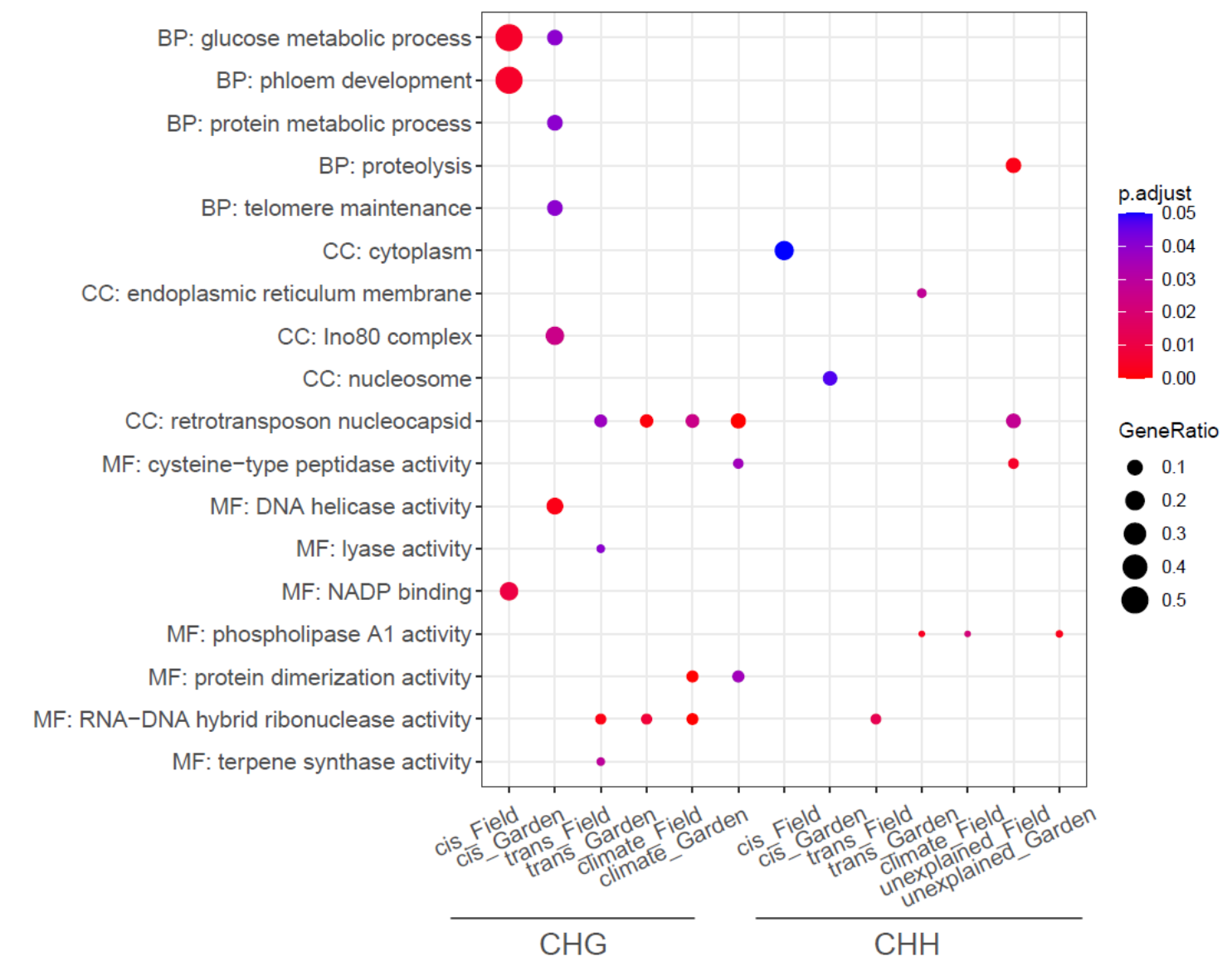


Figure 4

Gene Ontology (GO) enrichment analysis of genes containing *cis*, *trans*, and climate-predicted DMRs, as well as unexplained DMRs, in their promoters. Only GO terms with an adjusted *P*-value < 0.05 are shown. BP: biological process, CC: cellular component, MF: molecular function.

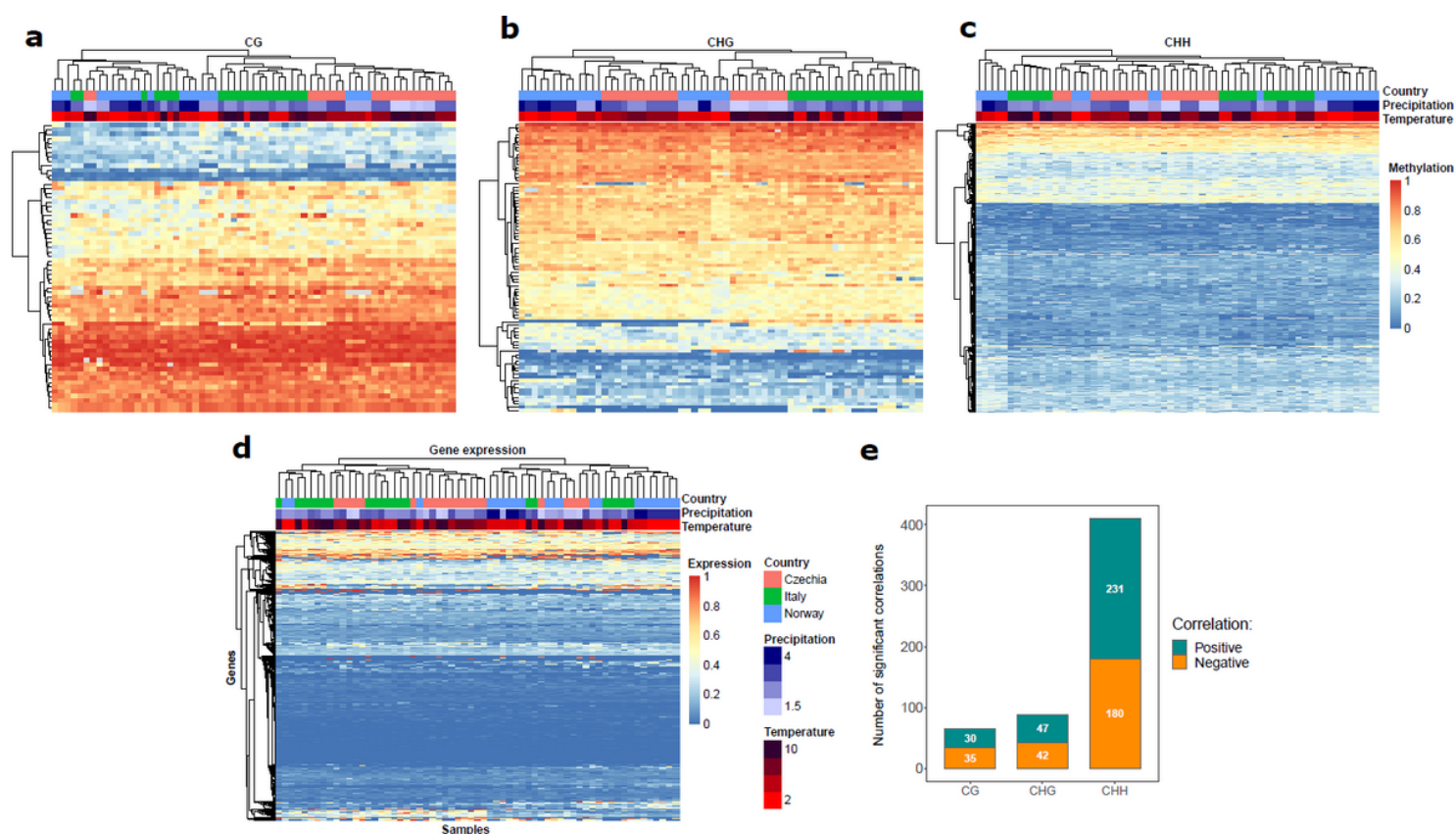


Figure 5

Climate-associated DMRs that were significantly correlated with gene expression. (a-d) Heatmaps showing methylation in the CG (a), CHG (b) and CHH (c) contexts and expression (for all the sequence contexts) (d) of the genes presenting significant correlation between promoter methylation and gene expression. Rows show clustered genes and columns show samples. The heatmap color scale on the right shows higher and lower methylation or expression in red and blue, respectively. The other color scales on the right show the clustering of the samples according to the country, precipitation (mm/day), and temperature (°C) of origin of each sample. (e) Number of genes showing positive or negative correlation between promoter methylation of climate-predicted DMRs and expression of the adjacent genes. Plants: $n = 63$.

Supplementary Files

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

- [AppendixI.docx](#)
- [Descriptionofsupplementarytables.docx](#)
- [Supplementaryfigures.docx](#)
- [Supplementarytable1.xlsx](#)
- [Supplementarytable2.xlsx](#)

- [Supplementarytable3.xlsx](#)
- [Supplementarytable4.xlsx](#)
- [Supplementarytable5.xlsx](#)
- [Supplementarytable6.xlsx](#)
- [Supplementarytable7.xlsx](#)