



The methylome is altered for plants in a high CO₂ world: Insights into the response of a wild plant population to multigenerational exposure to elevated atmospheric [CO₂]

Jasmine M. Saban¹ | Alexander Watson-Lazowski¹ | Mark A. Chapman¹ | Gail Taylor^{1,2}

¹School of Biological Sciences, University of Southampton, Southampton, UK

²Department of Plant Sciences, University of California, Davis, Davis, CA, USA

Correspondence

Gail Taylor, Department of Plant Sciences, University of California, Davis, Davis, CA 95616, USA.

Email: gtaylor@ucdavis.edu

Present address

Alexander Watson-Lazowski, John Innes Centre, Norwich Research Park, Norwich, NR4 7UH, UK

Funding information

Natural Environment Research Council, Grant/Award Number: NE/L002531/1; University of California, Davis; The British Council

Abstract

Unravelling plant responses to rising atmospheric CO₂ concentration ([CO₂]) has largely focussed on plastic functional attributes to single generation [CO₂] exposure. Quantifying the consequences of long-term, decadal multigenerational exposure to elevated [CO₂] and the genetic changes that may underpin evolutionary mechanisms with [CO₂] as a driver remain largely unexplored. Here, we investigated both plastic and evolutionary plant responses to elevated [CO₂] by applying multi-omic technologies using populations of *Plantago lanceolata* L., grown in naturally high [CO₂] for many generations in a CO₂ spring. Seed from populations at the CO₂ spring and an adjacent control site (ambient [CO₂]) were grown in a common environment for one generation, and then offspring were grown in ambient or elevated [CO₂] growth chambers. Low overall genetic differentiation between the CO₂ spring and control site populations was found, with evidence of weak selection in exons. We identified evolutionary divergence in the DNA methylation profiles of populations derived from the spring relative to the control population, providing the first evidence that plant methylomes may respond to elevated [CO₂] over multiple generations. In contrast, growth at elevated [CO₂] for a single generation induced limited methylome remodeling (an order of magnitude fewer differential methylation events than observed between populations), although some of this appeared to be stably transgenerationally inherited. In all, 59 regions of the genome were identified where transcripts exhibiting differential expression (associated with single generation or long-term natural exposure to elevated [CO₂]) co-located with sites of differential methylation or with single nucleotide polymorphisms exhibiting significant inter-population divergence. This included genes in pathways known to respond to elevated [CO₂], such as nitrogen use efficiency and stomatal patterning. This study provides the first indication that DNA methylation may contribute to plant adaptation to future atmospheric [CO₂] and identifies several areas of the genome that are targets for future study.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2020 The Authors. *Global Change Biology* published by John Wiley & Sons Ltd

KEYWORDS

DNA methylation, elevated [CO₂] response, epigenetics, genetic variation, multi-omics, natural [CO₂] springs, *Plantago lanceolata* (ribwort plantain), plasticity

1 | INTRODUCTION

Increased atmospheric CO₂ concentration ([CO₂]) is a significant driver of climate change, with [CO₂] of 430–1,000 ppm predicted for 2100, for the first time for several million years (IPCC, 2014; Pearson & Palmer, 2000). Understanding the molecular mechanisms underpinning plant response to rising [CO₂] is critical to a wider understanding of ecosystem change (Barnaby & Ziska, 2012), for crop performance (Myers et al., 2014) and the conservation of landscapes (Monroe et al., 2018).

Plastic phenotypic responses to elevated [CO₂] within a single generation have been extensively quantified in diverse plant species using various experimental designs (e.g. growth chambers, open-topped chambers, free air CO₂ enrichment facilities; Ainsworth & Long, 2005; Loladze, Nolan, Ziska, & Knobbe, 2019). However, our understanding of the multigenerational response of plants to elevated [CO₂] has been limited by the challenges of growing plants in experimentally elevated [CO₂] for multiple generations. The association between atmospheric [CO₂] and plant morphology in the fossil record provides evidence that increases in atmospheric [CO₂] could be a driver of evolutionary change over macroevolutionary timescales (Franks & Beerling, 2009; Haworth, Elliott-Kingston, & McElwain, 2011). Additionally, selection experiments conducted in controlled environments generally suggest a role for elevated [CO₂] as a potential driver of contemporary evolution (Frenck, Linden, Mikkelsen, Brix, & Jørgensen, 2013; Ward, Antonovics, Thomas, & Strain, 2000). However, these studies are limited in number and infer evolution solely from morphological phenotypes. Furthermore, in realistic field conditions, adaptation to elevated [CO₂] is highly dependent on ecological context (Grossman & Rice, 2014; Kleynhans, Otto, Reich, & Vellend, 2016). Thus, a greater understanding of the mechanistic basis of plant adaptation to elevated [CO₂] is required.

Natural CO₂ springs provide a powerful resource for investigating plant response to elevated [CO₂] over multiple generations without the extensive labour, time and financial costs associated with other systems (Körner & Miglietta, 1994). The responses of plants exposed to elevated [CO₂] at such springs are generally consistent in direction and magnitude to those observed in single generation free air CO₂ enrichment experiments, for a range of morphological traits (Saban, Chapman, & Taylor, 2018). However, crossed factored experiments have also shown that plant responses to elevated [CO₂] at natural CO₂ springs are not solely plastic, with adaptation to elevated [CO₂] inferred in some morphological traits in these plant populations (Nakamura et al., 2011; Watson-Lazowski et al., 2016). One study identified adaptation to elevated [CO₂] in the gene expression profiles of plants at a CO₂ spring and nearby control site, but with very little genetic divergence between the two populations (Watson-Lazowski et al., 2016). This highlights a potential role

for indirect genetic effects (e.g. DNA methylation changes) in the response of plant populations to multigenerational elevated [CO₂] exposure but to date, these have yet to be investigated.

Whole-genome DNA methylation patterns have been shown to be responsive to abiotic environmental conditions, including salinity (Yaish, Al-Lawati, Al-Harrasi, & Patankar, 2018), temperature (Ma et al., 2015) and drought (Neves et al., 2017), and may coordinate adaptive phenotypes (Xia et al., 2016; Yong-Villalobos et al., 2016). Methylation of cytosine in DNA occurs more extensively in plants than animals, and with large variation in patterns between species (Niederhuth et al., 2016). In plants, DNA methylation occurs in three cytosine contexts, CG, CHH and CHG (where H is any base except G) (Henderson & Jacobsen, 2007) with different mechanisms of establishment and apparent function of methylation depending on both the cytosine sequence context and the wider genomic context (Song & Cao, 2017). Broadly, DNA methylation appears to function to silence the mobility of transposable elements, contribute to genome stability and integrity and may also play a role in gene expression regulation (Eichten, Schmitz, & Springer, 2014; Zhang, Kimatu, Xu, & Liu, 2010). Currently, there is no understanding of the relevance of this mechanism in determining plastic and adaptive responses to elevated [CO₂].

Given the potential role of DNA methylation in modulating gene expression as part of a plastic response to environmental cues (Garg, Chevala, Shankar, & Jain, 2015), coordination of some element of plant response to elevated [CO₂] by reprogramming of global DNA methylation is an attractive hypothesis to explain previous observations (Watson-Lazowski et al., 2016). Furthermore, the observation that in plants methylation can be maintained through mitotic and meiotic cell division (Quadrana & Colot, 2016; Verhoeven, Jansen, van Dijk, & Biere, 2010) has led to the hypothesis that methylation could provide transgenerational 'memory' of ancestral environment, contributing to phenotype expression in offspring (Heard & Martienssen, 2014; Quadrana & Colot, 2016). However, experimental evidence of environmentally induced methylation patterns that both influence phenotype and are inherited into the next generation is rare (Crisp, Ganguly, Eichten, Borevitz, & Pogson, 2016; Quadrana & Colot, 2016). The role of DNA methylation in coordinating plastic response to elevated [CO₂] has not been explored, despite recent progress in understanding trait responses. For example, the mechanistic basis of altered stomatal patterning (Engineer et al., 2016; Xu, Jiang, Jia, & Zhou, 2016) has identified genes regulating this pathway, while patterns of plant growth and metabolism are well established in response to exposure to elevated [CO₂] (Gamage et al., 2018). Plants growing at natural CO₂ springs provide a resource to explore the role of DNA methylation in both the plastic and evolutionary mechanisms coordinating the response to elevated [CO₂] exposure, where candidate genes involved in the plastic morphological response may have already been identified.

The aim of the research described here was to explore the mechanistic basis for the plastic and adaptive response to elevated $[\text{CO}_2]$ in the non-model plant species, *Plantago lanceolata* L., utilizing populations exposed to naturally elevated $[\text{CO}_2]$ for many generations in the Bossoleto CO_2 spring in Italy. This spring is thought to be more than 100 years old, with *P. lanceolata* populations documented since at least 1992 (Miglietta, Raschi, Bettarini, Resti, & Selvi, 1993). A crossed factored experiment was conducted with progeny of these plants and those from a nearby control site, by growing seed in a common environment for one generation before growth in either elevated or ambient $[\text{CO}_2]$ growth chambers. We combined previous phenotypic analysis of this experiment with a reanalysed RNA-Seq dataset and novel whole-genome sequencing (WGS) and whole-genome bisulphite sequencing (WGBS) datasets for 24 individuals to analyse genetic variation, DNA methylation patterns and gene expression. In applying multi-omic techniques to a crossed factored experiment, we addressed the following questions: (a) Is there evidence of genetic divergence between populations growing at elevated and ambient $[\text{CO}_2]$? (b) Does DNA methylation respond to single or multigenerational elevated $[\text{CO}_2]$ exposure? (c) What role do genetic differentiation and methylation variation play in the plant response to elevated $[\text{CO}_2]$? This study also addresses the challenge of working with a non-model organism, deploying tools developed in model species to advance genomic resources of a non-model but ecologically important species. This approach is critically needed to advance the study of plant ecological epigenetics (Richards et al., 2017).

2 | METHODS

2.1 | Plant material and sampling site

The study organism for this work was *P. lanceolata* L. (Plantaginaceae), an herbaceous perennial with widespread geographical distribution. The experimental approach has been described previously in Watson-Lazowski et al. (2016) and is summarized briefly here. Seeds were collected from nine plants growing in naturally elevated $[\text{CO}_2]$ near to the CO_2 spring at Bossoleto, Italy (Lat. $43^\circ 17'$, Long. $11^\circ 35'$), and at a nearby (ca. 200 m apart) ambient $[\text{CO}_2]$ control site in May 2008. At the elevated $[\text{CO}_2]$ site, the average daytime $[\text{CO}_2]$ is around $1,000 \mu\text{mol/mol}$, with a range of $400\text{--}1,200 \mu\text{mol}^{-1}$ (Scholefield et al., 2004). Seeds obtained from the CO_2 spring and control sites were grown for one generation in the glasshouse at the University of Southampton and then crossed within maternal families to standardize parental effects. Seeds from crosses were transferred into one of eight experimental growth chambers. Four chambers were set to ambient $[\text{CO}_2]$ (410.63 ± 33.74 ppm) and four chambers were set to elevated $[\text{CO}_2]$ (718 ± 46.81 ppm, respectively). On the 58th day, the second or third youngest leaf was harvested and stored at -80°C for RNA and DNA extractions. Further details of the experimental design (previously reported in Watson-Lazowski et al., 2016) and the analysis summarized here are available in the extended methods, Methods S1.

The bioinformatics pipeline for analysis of genetic, methylation and gene expression variation in this experiment includes the integration of multi-omic data sets for this non-model species (Figure 1).

2.2 | Genome assembly

To facilitate analyses of methylation and genetic variation in plants in the experiment, the *P. lanceolata* genome was assembled from short reads. DNA was extracted from leaf material from an individual taken as seed from the CO_2 spring site and grown in the University of Southampton glasshouse. DNA was sequenced as 150 bp paired ends with 350 bp insert size by Novogene Bioinformatics Institute (Beijing, China). Raw sequencing data were filtered for contaminants with FastQ Screen v0.11.3 (http://www.bioinformatics.babraham.ac.uk/projects/fastq_screen/) and trimmed with Trimmomatic v0.36 (Bolger, Lohse, & Usadel, 2014). Error correction was performed using SOAPec (Luo et al., 2012) and reads were assembled using ABySS (Simpson et al., 2009; Table S1). The genome was filtered to contigs larger than 2 kb for further analyses using seqtk (Li, 2016), as a trade-off to increase computational efficiency while maximizing sequence availability for downstream analysis. Genome completeness was assessed using BUSCO v3 (Simão, Waterhouse, Ioannidis, Kriventseva, & Zdobnov, 2015).

2.2.1 | Gene prediction

Genes were predicted ab initio from the 2 kb filtered genome assembly using Maker v2.31.10 (Cantarel et al., 2008). Low complexity sequences and interspersed repeats were masked using RepeatMasker with RepBase repeat library (Jurka et al., 2005). Gene prediction used two iterations of SNAP (Korf, 2004) and one of AUGUSTUS (Stanke & Waack, 2003) through BUSCO (Simão et al., 2015). SNAP algorithms were trained on *P. lanceolata* RNA-Seq evidence and UniProt SwissProt plant protein alignments. Prediction quality was assessed through AED scores, and by protein domain conservation using InterProScan v5.30 (Jones et al., 2014) with PANTHER v12.0 (Mi, Muruganujan, & Thomas, 2012) and PFAM v31.0 (Finn et al., 2015).

2.2.2 | Chloroplast assembly

The chloroplast genome was assembled to test bisulphite conversion efficiency of sample sequences in methylation analysis. Raw reads were assembled using Novoplasty v2.6.3 (Dierckxsens, Mardulyn, & Smits, 2017), with the *P. lanceolata* chloroplast gene *rbcl* as a seed sequence (Olmstead & Reeves, 1995). The chloroplast genome was annotated and visualized using the GSEq (Tillich et al., 2017) and OGdraw (Lohse, Drechsel, Kahlau, & Bock, 2013) components of CHLOROBOX (<https://www.mpimp-golm.mpg.de/chlorobox>).

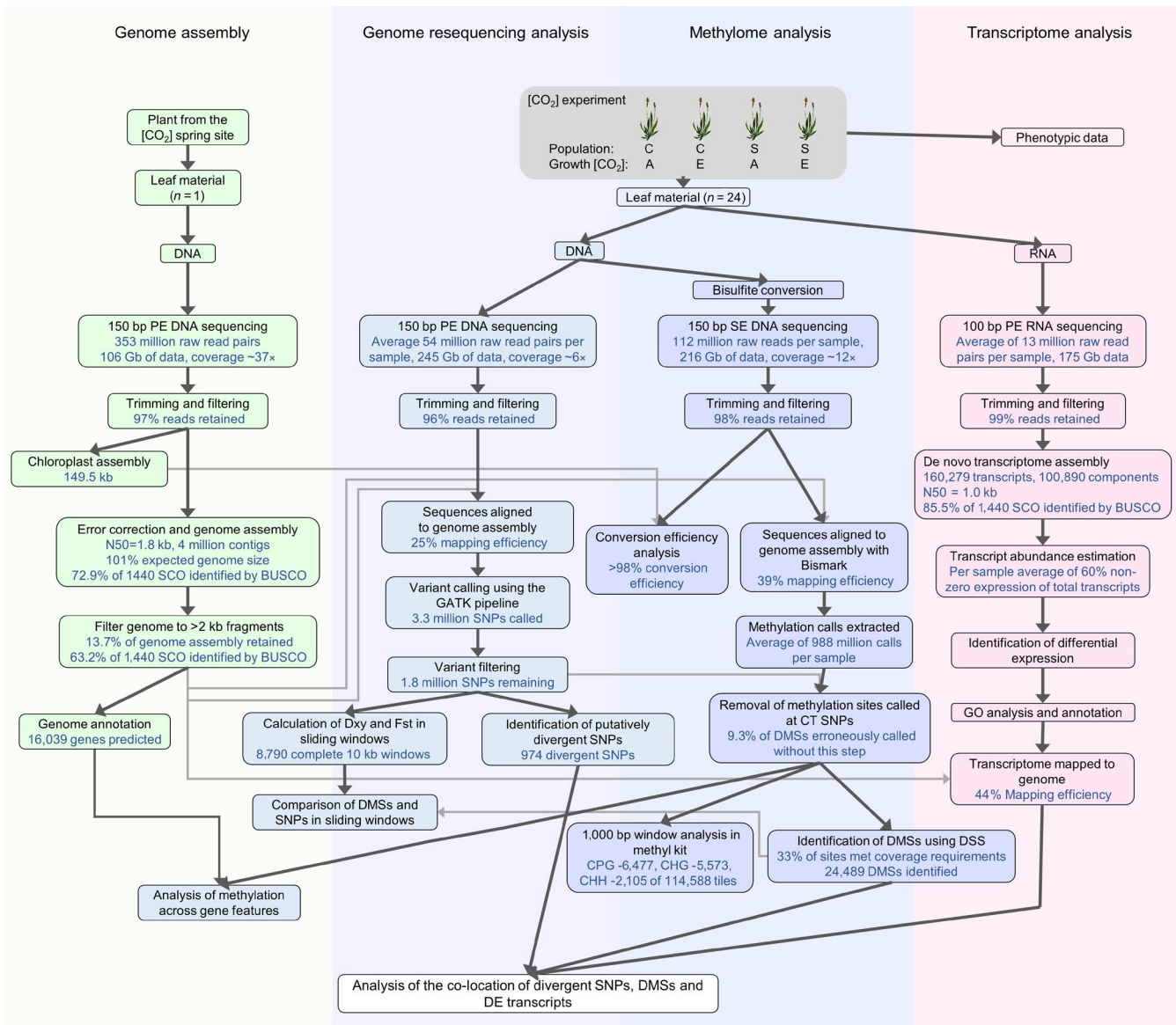


FIGURE 1 A pipeline for integrating multi-omic datasets in a non-model species. Plants derived from CO₂ spring and control sites were grown in ambient and elevated [CO₂] treatments. Leaf material was taken from six plants per treatment group ($n = 24$ total) and DNA and RNA were extracted for genome, methylome and transcriptome analysis. To facilitate these analyses, the genome of *Plantago lanceolata* (1C genome size = 1.4 Gb; Wong & Murray, 2012) was sequenced with short reads and assembled

2.3 | Population genomics

Frozen leaf samples from 24 plants in the experiment (six per site and per growth [CO₂]) were used for DNA extraction for WGS and WGBS and for RNA extraction for RNA Sequencing (Figure 1).

DNA were extracted using a modified version of the Doyle and Doyle (1987) CTAB protocol. Library preparation and sequencing were carried out by the DNA Technologies and Expression Analysis Cores at the UC Davis Genome Center (CA, USA). WGS libraries were prepared using the KAPA library preparation kit (Kapa Biosystems) and sequenced as 150 bp paired end reads. Raw reads were trimmed using Trimmomatic v0.32 (Bolger et al., 2014) and contamination was removed with FastQ Screen.

Reads were aligned to the 2 kb filtered genome assembly using Bowtie2 (Langmead & Salzberg, 2012). Genome contigs were stitched together into a 'superscaffold' using ScaffoldStitcher (Haj & O'Connor, 2016) for population genomic analysis using the Genome Analysis Tool Kit (GATK) pipeline (Van der Auwera et al., 2013). Following processing of alignment files with GATK v3.7 Genomic VCF files were generated for each sample and single nucleotide polymorphisms (SNPs) were extracted with SNP filtering expression as $MQ < 20 \parallel SOR > 3 \parallel QD < 3 \parallel FS > 60 \parallel MQRankSum < -10 \parallel ReadPosRankSum < -5$ (see Figures S1 and S2 for justification). Singleton SNPs were removed, and SNPs were filtered to include only those with identity information in at least eight of the 12 individuals per population. Variation was visualized using *smartPCA* within Eigensoft (Price et al., 2006) and basic population genetic

statistics were calculated with VCFtools (Danecek et al., 2011). D_{xy} and F_{ST} were calculated in 10 kb complete sliding windows with 1 kb step size, using publicly available python scripts (Martin, 2017).

2.3.1 | Identification of outlier F_{ST} using Bayescan

Putatively divergent SNPs were identified using Bayescan v2.1 (Foll & Gaggiotti, 2008) with q -value threshold 0.05, after running for 5,000 outputted iterations with 50,000 burn-in and retaining every 10th iteration.

2.4 | DNA methylation

2.4.1 | Whole-genome bisulphite sequencing

Bisulphite conversion of DNA was carried out using the Zymo EZ DNA Methylation-Lightning kit (Zymo Research) and bisulphite converted reads were prepared using the TruSeq DNA Methylation kit (Illumina). WGBS libraries were sequenced as 150 bp single end reads with a 20% *phiX* spike-in. Raw reads were trimmed using Trimmomatic v0.32 (Bolger et al., 2014) and *phiX* DNA and contamination was removed with FastQ Screen. Bisulphite conversion efficiency was calculated as the percentage of unmethylated cytosines that were not converted to thymine in chloroplast sequences, since the chloroplast genome is expected to be unmethylated (Fojtová, Kovařík, & Matyášek, 2001).

Due to the high level of fragmentation in the assembled *P. lanceolata* genome, all WGBS reads were aligned to the single reference rather than WGS for each individual. This approach has the advantage of being computationally tractable and is widespread in the literature (e.g. Lu et al., 2017; Yaish et al., 2018) but has the disadvantage that cytosine (C) to thymine (T) SNPs between individuals cannot be distinguished from an unmethylated C in one individual and a methylated C in another following bisulphite conversion. Therefore, sites of C-T polymorphism identified in the population were removed from the analysis of differential methylation.

To analyse methylation in the *P. lanceolata* 2 kb filtered genome, trimmed WGBS reads were aligned to the filtered assembly using Bismark v0.19.0 (Krueger & Andrews, 2011) with Bowtie2 v2.3.1 (Langmead & Salzberg, 2012). Bismark aligned the trimmed WGBS reads to the filtered genome allowing a maximum of 20 mismatches and with `--score_min L,0,-0.4` (Krueger & Andrews, 2011) and extracted methylation calls for each cytosine. Genome-wide weighted methylation was calculated as described by Schultz, Schmitz, and Ecker (2012).

Identifying differential methylation

Differential methylation was analysed through two separate approaches, first as tiles to identify genome-wide patterns of differential methylation, and second individual sites were analysed to identify treatment-associated differentially methylated sites (DMSs;

Eichten, Stuart, Srivastava, Lister, & Borevitz, 2016; Ganguly, Crisp, Eichten, & Pogson, 2017).

1,000 bp tile analysis in methylKit. Methylation call files were converted to coverage and bed files and sorted by chromosome with `bismark2bedgraph` (Krueger & Andrews, 2011) and then converted to report files with `coverage2cytosine` and read into `MethylKit` with `coverage ≥ 3` (Akalin, 2016) in R v 3.4.1 (R Core Team, 2015). 1,000 bp tiles present across all samples were identified using `MethylKit` (Akalin et al., 2012). Global analysis of methylation in unfiltered 1,000 bp tiles was conducted using the `PCAsamples` function in `MethylKit`.

To identify differentially methylated 1,000 bp tiles, a chi-squared test was implemented between pairwise treatment groups using the `calculateDiffMeth` function, using Benjamini-Hochberg correction with maximum FDR corrected $p < .00001$ and minimum methylation difference dependent on cytosine context: CG 40%, CHG 20%, CHH 10% (Akalin et al., 2012; Benjamini & Hochberg, 1995). Percentage methylation of these differentially methylated tiles was visualized with heatmap (Kolde, 2013).

DMS identification in DSS. Differentially methylated sites were identified from methylation calls using the R package DSS (Park & Wu, 2016). DSS implements algorithms for the identification of differential methylation using the dispersion shrinkage method and Wald tests assuming a beta-binomial error distribution. DMSs were identified from filtered coverage report files, including only sites with coverage $≥ 3$ and with $≥ 3$ samples per treatment group. Downstream analyses were conducted using DMSs filtered by $FDR < 0.05$. The number of DMSs was calculated in 10 kb complete sliding windows with a 1 kb step size and these were plotted against D_{xy} and F_{ST} in that region. Correlations were assessed with Spearman's rank correlation using ρ^2 to approximate the proportion of shared variance between the ranked variables.

2.5 | Gene expression

The RNA sequencing dataset previously reported in Watson-Lazowski et al. (2016) was re-analysed here utilizing increased computational resources. Protocols for RNA extraction and sequencing are described in Watson-Lazowski et al. (2016; Methods S1). Adapters and poor-quality bases were trimmed from raw reads with Trimmomatic v0.32 (Bolger et al., 2014). The reference transcriptome was assembled de novo from normalized reads of 12 samples using Trinity v2.4.0 (Haas et al., 2013). Transcripts from the assembled transcriptome were aligned to The Arabidopsis Information Resource (TAIR10; Berardini et al., 2015) using BLASTX ($e < 1 \times 10^{-10}$; Altschul, Gish, Miller, Myers, & Lipman, 1990). RNA sequencing libraries were mapped back to the assembled transcriptome and component counts were converted to TMM-normalized FPKM which standardizes counts by sequencing depth and gene length.

Transcripts that were differentially expressed (DE) were identified by implementing a generalized linear model (glm) with negative binomial error distribution in edgeR (McCarthy, Chen, & Smyth, 2012)

and transcripts were considered DE when $FDR < 0.05$. Transcripts were functionally annotated and assigned to Gene Ontology (GO) categories using DAVID (Dennis et al., 2003) and visualized using GOplot (Walter, Sánchez-Cabo, & Ricote, 2015).

2.6 | Co-location of DMSs, SNPs and DE transcripts

For analysis of the co-location of DE transcripts to regions of the genome with DMSs and SNPs, we mapped transcripts to the 2 kb filtered genome with BLAST ($e < 1 \times 10^{-4}$ and $>95\%$ similarity). The locations of mapped DE transcripts were cross-referenced with the location of DMSs and F_{ST} outlier SNPs using custom R scripts with *GenomicRanges* (Lawrence et al., 2013). LD decay was calculated using PopLDdecay (Zhang, Dong, Xu, He, & Yang, 2019) and 1 kb was used to identify DMSs and SNPs that were potentially in LD with DE transcripts, since r^2 had decreased to 0.11 at 1 kb (Figure S3). For predicted

gene features and transcripts mapping to the genome, methylation of those features was plotted against the expression of the transcript.

3 | RESULTS

The study species for this analysis, *P. lanceolata* is a non-model species with previously limited genetic resources available. Here, a genome assembly, WGS, WGBS and RNA-Seq have been integrated in parallel, to explore the mechanistic basis of plant response to elevated $[CO_2]$ (Figure 1).

3.1 | Genome assembly and feature prediction

Short read genome assembly produced a genome of 1,425,357 kb of sequence, which corresponds closely to the 1.4 Gb estimated

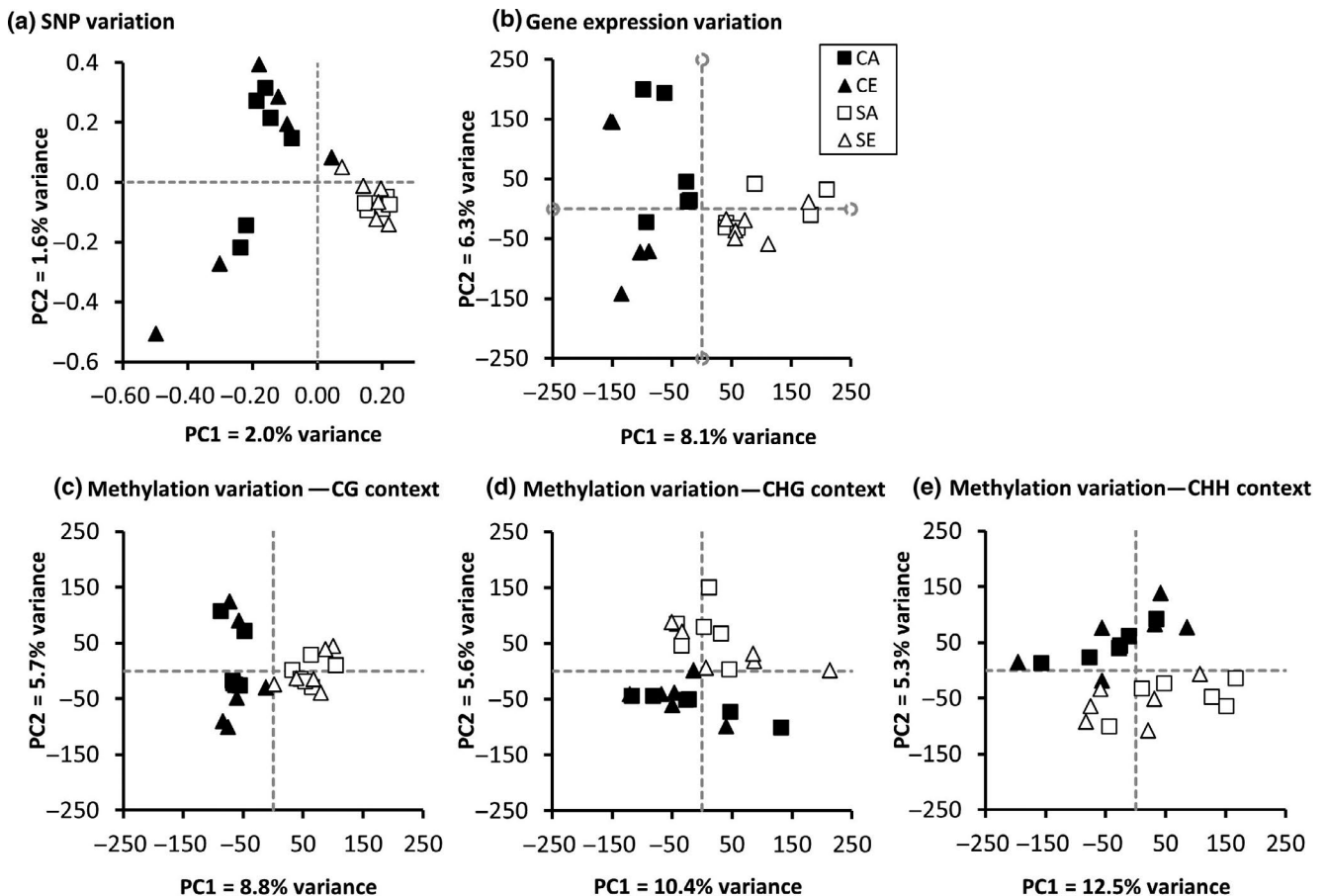


FIGURE 2 Principal component analysis of variation in omic datasets of *Plantago lanceolata* plants derived from CO_2 spring or control site populations, and grown under elevated or ambient $[CO_2]$ ($n = 24$, six reps per treatment group). Principal components one and two are given with percentage of variance explained by the component. Treatment groups; CA, Control Ambient; CE, Control Elevated; SA, Spring Ambient; SE, Spring Elevated. All analyses show clustering by population site of origin (spring vs. control) in principal component 1 or principle component 2. (a) Single nucleotide polymorphism (SNP) variation from 8 million SNPs, with singletons removed and SNPs filtered to those that contain information for at least 8/12 individuals per population; (b) variation in gene expression profiles; (c) variation in methylation at cytosines in the CG context across complete 1 kb tiles; (d) variation in methylation in the CHG context; and (e) variation in methylation in the CHH context. Note that principal component analysis of SNP variation was conducted using smartPCA (Price et al., 2006), while *prcomp()* in R was used for gene expression and methylation variation. The scale of the principal components differs between these two methods but this does not affect interpretation of clustering

genome size of *P. lanceolata* (Wong & Murray, 2012). The assembly was highly fragmented, consisting of 4,075,744 contigs with N50 of 1.8 kb (Table S1). When the genome was filtered to sequences >2 kb, 13.7% of the full sequence was represented, with 63.2% of single-copy orthologs (SCO) from *Arabidopsis thaliana* identified with BUSCO (Simão et al., 2015) as being complete and a further 7.2% present but fragmented (Table S2; Figure S4). The Maker annotation pipeline (Cantarel et al., 2008) predicted 16,039 genes with 26% of predicted genes having AED < 0.5 (90% is considered well-annotated; Campbell, Holt, Moore, & Yandell, 2014). Gene prediction was generally poor, likely because of the high fragmentation of the contig assembly and lack of protein sequence evidence from *Plantago* or closely related species for training, an issue common to genetic analyses in non-model species. Genes with AED < 0.5 were considered sufficiently supported for analysis of general trends of methylation across gene features. The chloroplast was assembled as two alternative sequences of 149.6 kb, differing in the orientation of an inverted repeat sequence (Figure S5).

3.2 | Population genomics

3.2.1 | Low overall genomic differentiation between spring- and control-derived populations

WGS produced 57 million paired end reads per sample (15× coverage). A total of 1.8 million SNPs were identified across the 24 individuals after excluding SNPs for which less than eight individuals per population had a base call. PCA analysis of the SNPs showed the

component with the highest variance (2.0%) separated the samples by site of origin (Figure 2a).

Population genomic analysis of SNPs supported previous findings of limited genetic differentiation between spring and control site populations (Watson-Lazowski et al., 2016), with $F_{ST} = 0.050 \pm 0.125$ (mean \pm SD) across the genome and similar values in exon regions (Table 1). Nucleotide diversity was significantly higher in the control population (Table 1). The exon regions of the spring population had a negative Tajima's *D*, indicating some purifying selection. Within each population, Tajima's *D* was positive across the whole genome, signifying low levels of extreme frequency polymorphisms characteristic of balancing selection or a decrease in population size. The most parsimonious explanation for this pattern of differentiation is that the spring population originated from the control population and underwent weak purifying selection in exon regions (Figure S6). Of the SNPs, 974 (0.05%) were identified as putative targets of divergent selection between the spring and control populations as F_{ST} outliers ($q < 0.05$).

3.3 | DNA methylation

3.3.1 | Divergence in methylation profiles between control and spring populations, with limited methylome remodelling under single generation exposure to elevated [CO₂]

Whole-genome bisulphite sequencing produced 113 million single end reads per sample (12× coverage). Conversion efficiency was >98.5% for all samples (Table S3). Methylation of cytosines in all contexts was

TABLE 1 Population genomic analysis statistics calculated from SNPs for whole-genome and exon-only regions

Population statistic	Region	Site of origin	Estimate	t test	
				t statistic	p value
F_{ST}	Whole genome		0.050 ± 0.125		
	Exon regions		0.048 ± 0.112		
Nucleotide diversity (π) in windows	Whole genome	Control	0.0059 ± 0.0056	19.63	<.0001
		Spring	0.0057 ± 0.0056		
	Exon regions	Control	0.0045 ± 0.0040	7.01	<.0001
		Spring	0.0041 ± 0.0038		
Per site nucleotide diversity (π)	Whole genome	Control	0.298 ± 0.152	4.26	<.0001
		Spring	0.297 ± 0.150		
	Exon regions	Control	0.280 ± 0.141	11.14	<.0001
		Spring	0.264 ± 0.144		
Tajima's <i>D</i>	Whole genome	Control	0.200 ± 0.948	14.7	<.0001
		Spring	0.174 ± 0.989		
	Exon regions	Control	0.053 ± 0.925	10.24	<.0001
		Spring	-0.073 ± 0.952		

Note: Estimates are provided as mean \pm SD. Tajima's *D* and nucleotide diversity were calculated in 100 bp windows. *T* tests were used to determine whether nucleotide diversity and Tajima's *D* significantly differed between the population derived from the spring site and from the control site. Abbreviation: SNP, single nucleotide polymorphism.

relatively high with 83%, 70% and 15% of cytosines methylated in the CG, CHG and CHH contexts, respectively. As for other plant species, methylation was depleted at the transcription start site of predicted genes and CHG and CHH methylation were additionally depleted across the gene body (Figure 3d; Figure S7). Methylation in the CG context increased to near non-genic levels across the gene body and was depleted again at the transcription end site.

Differences in methylation between the two populations dominated the methylation variation identified in this experiment, both in unfiltered tile analysis and in differential methylation analysis. In global methylation patterns in unfiltered 1 kb tiles, PCs dividing the individuals by site of origin explained 8.8% of the variation in CG methylation (PC1), 5.6% of the CHG methylation (PC2) and 5.3% of the CHH methylation (PC2) (Figure 2c–e). This is similar in magnitude

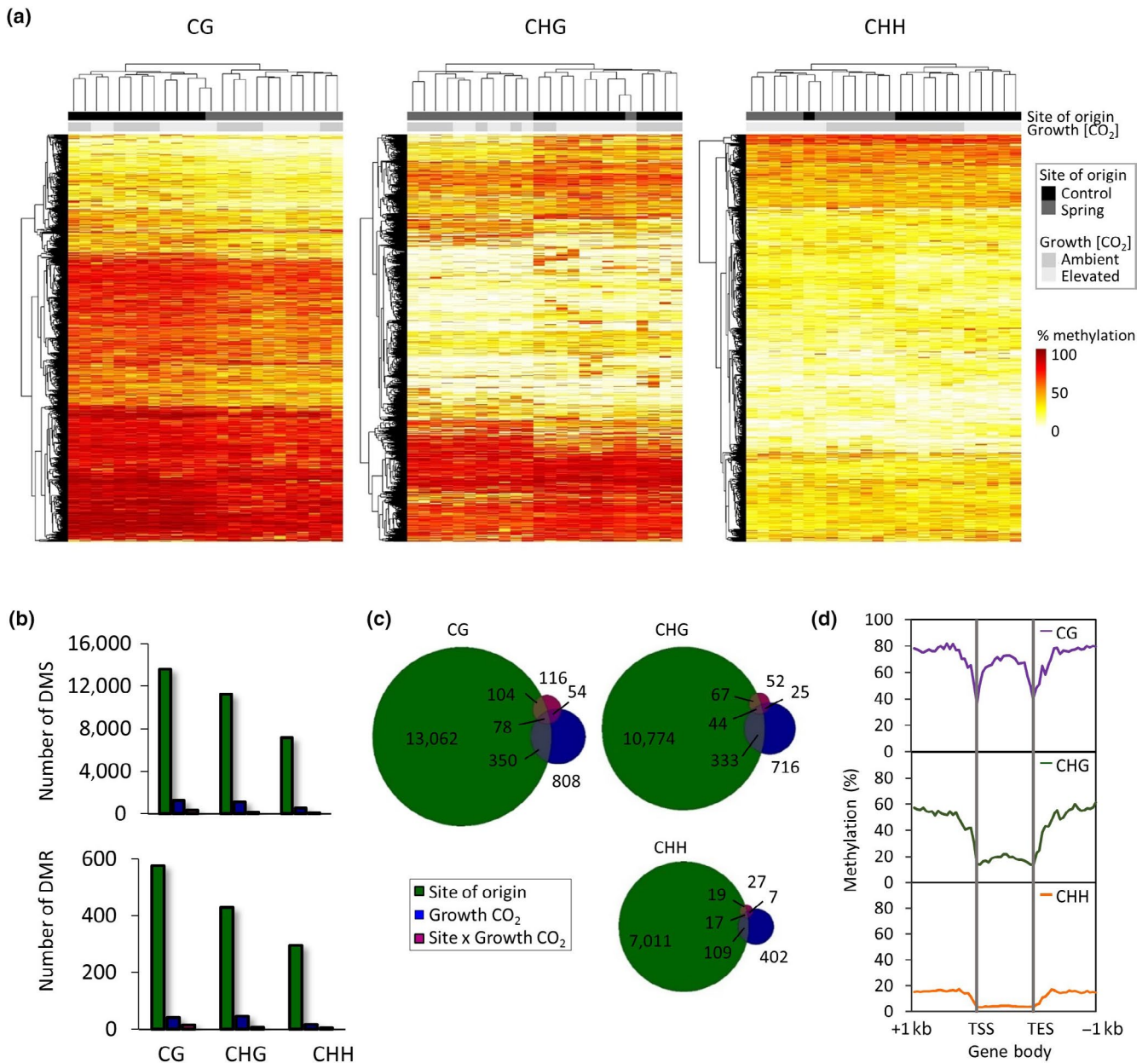


FIGURE 3 There were more differentially methylated sites (DMSs) and regions (DMRs) associated with individuals that originate from different sites (spring or control) than there were associated with growth [CO₂] treatment (ambient or elevated). (a) Heatmaps of average methylation of differentially methylated pre-specified 1,000 bp tiles identified by pairwise comparisons of the four treatment groups. Columns are clustered by similarity and colour coded, CA, Control Ambient; CE, Control Elevated; SA, Spring Ambient; SE, Spring Elevated. (b) Numbers of DMSs and regions DMRs as identified by R package DSS. (c) The overlap of DMS that were significantly differentially methylated in more than one category. Numbers describe transcripts belonging only to the category indicated. (d) Mean methylation of predicted genes in each methylation context. Percentage methylation 1 kb upstream and 1 kb downstream of the genes were calculated in 50 bp intervals, and across the gene body 5% intervals. TES, transcription end site; TSS, transcription start site

to the variance explained by PCs dividing gene expression variation (8.1%; Figure 2b).

Differential methylation in 1 kb tiles (identified using pairwise chi-squared tests and a minimum methylation difference dependent on context, CG 40%, CHG 20%, CHH 10%) clustered according to site of origin in all three contexts (Figure 3). Methylation in the CHH context was the only context in which methylation variation then clustered according to growth [CO₂], highlighting responsiveness to single generation exposure to elevated [CO₂]. In the per site analysis, there were >10-fold more DMSs associated with population site of origin than there were associated with growth [CO₂] in all three methylation contexts (CG 10-fold; CHG 10-fold; CHH 13-fold; Figure 3a). In all, 34,175 DMSs were identified as associated with growth [CO₂], population site of origin, an interaction between these effects or any combination of these. There was some evidence for methylome remodelling in response to elevated [CO₂] with 2,939 sites differentially methylated by growth [CO₂] (8.6%, $n = 34,175$), but the majority were associated with site of origin (93.5%; Figure 3b). Of the DMSs associated with growth [CO₂], 26% were also associated with population site of origin, significantly more than would be expected by chance (Fisher's exact test; $p < .0001$, odds ratio = 88.3; Table S4).

Differential methylation was disproportionately represented among cytosine contexts in DMSs analysis, with CG and CHG DMS enriched relative to the distribution of total cytosine sites at which methylation status was called (Figure 3c; Table S5).

It is estimated from this analysis that 9.3% of identified DMS would have been called erroneously if the genome for each individual had not been sequenced to identify C->T variants in the population (Table S6). This exemplifies that alignment of WGBS reads to a reference genome other than the genome of the individual can erroneously inflate the number of DMSs called and serves to caution the interpretation of DMSs in the absence of WGS for each individual. This approach cannot identify differential methylation at sites if there is also C-T polymorphism, but exclusion of these sites at least gives a conservative estimate of differential methylation.

3.3.2 | Correlations between genetic and methylome variation account for approximately 1% of differences in methylation between populations

When the genome was analysed in complete 10 kb tiles with a 1 kb step size, there were small but significant positive correlations between the number of DMSs associated with population of origin and both absolute (D_{xy}) and relative (F_{ST}) sequence divergence across the tile (Figure 4). Correlation with sequence divergence explained approximately 1% of the variation in number of DMSs in a 10 kb window across the three sequence contexts. This suggests that to some extent, regions of the genome with high sequence divergence are also more likely to harbour more DMSs (Table S7).

3.4 | Gene expression

3.4.1 | Divergence in the gene expression response of spring- and control-derived plants to elevated [CO₂]

A total of 160,279 transcripts were de novo assembled in 100,890 components (loosely comparable to genes) in Trinity, and 86% of transcripts had a BLASTN match with $e < 1 \times 10^{-10}$ to the transcriptome assembled in Watson-Lazowski et al. (2016). 44% of the transcripts mapped to the 2 kb filtered genome with $e < 1 \times 10^{-4}$ and >95% similarity, and BUSCO analysis of transcriptome completeness identified 85.5% of SCOs from *A. thaliana* in this de novo assembly (Figure S8).

PCA evidenced clustering of gene expression profiles by population site of origin but not growth [CO₂] (Figure 2b). Additionally, in the differential expression (DE) analysis where there were 1.8-fold more DE transcripts between site of origin than there were with growth [CO₂] (Figure 5).

There was an enrichment in the overlap between genes DE by growth [CO₂] and by population of origin with 40 genes (14%) DE between elevated [CO₂] and ambient [CO₂] also being DE by site of origin (Fisher's exact test; $p < .0001$, odds ratio = 539; Table S4). All of these transcripts showed the same directional response in expression in elevated versus ambient [CO₂] and in spring versus control site populations.

The same two GO categories, *Photosynthesis* and *Photosynthesis light reaction*, were the most represented for both the growth [CO₂] and population site of origin transcript response. 50% and 55% of transcripts in these GO categories were DE both by site of origin and by growth [CO₂]. In contrast, DE transcripts involved in *response to stimuli* and *response to stress* GO categories were only found to be overrepresented in the transcripts DE by site of origin (Figure 6).

Only 20% of the transcripts identified as DE in this analysis mapped with BLASTN ($e < 1 \times 10^{-10}$) to components DE in the analysis of this data by Watson-Lazowski et al. (2016), likely reflecting the different statistical approaches to defining DE (a glm used in this analysis and pairwise t tests between treatment groups in Watson-Lazowski et al., 2016). Nevertheless, broad trends presented here were supported in both studies, including evidence of more differential expression between populations relative to growth [CO₂] and enriched GO categories. Transcripts described in more detail below were DE in both analyses.

3.5 | Co-location of DMSs, SNPs and DE transcripts

3.5.1 | Co-location of DE transcripts, DMSs and high divergence SNPs highlight potential targets for differential methylation and selection

In total, 7,089 transcripts (corresponding to 4,205 components) mapped to within 1 kb of a DMS (86%), an outlier SNP (10%) or both

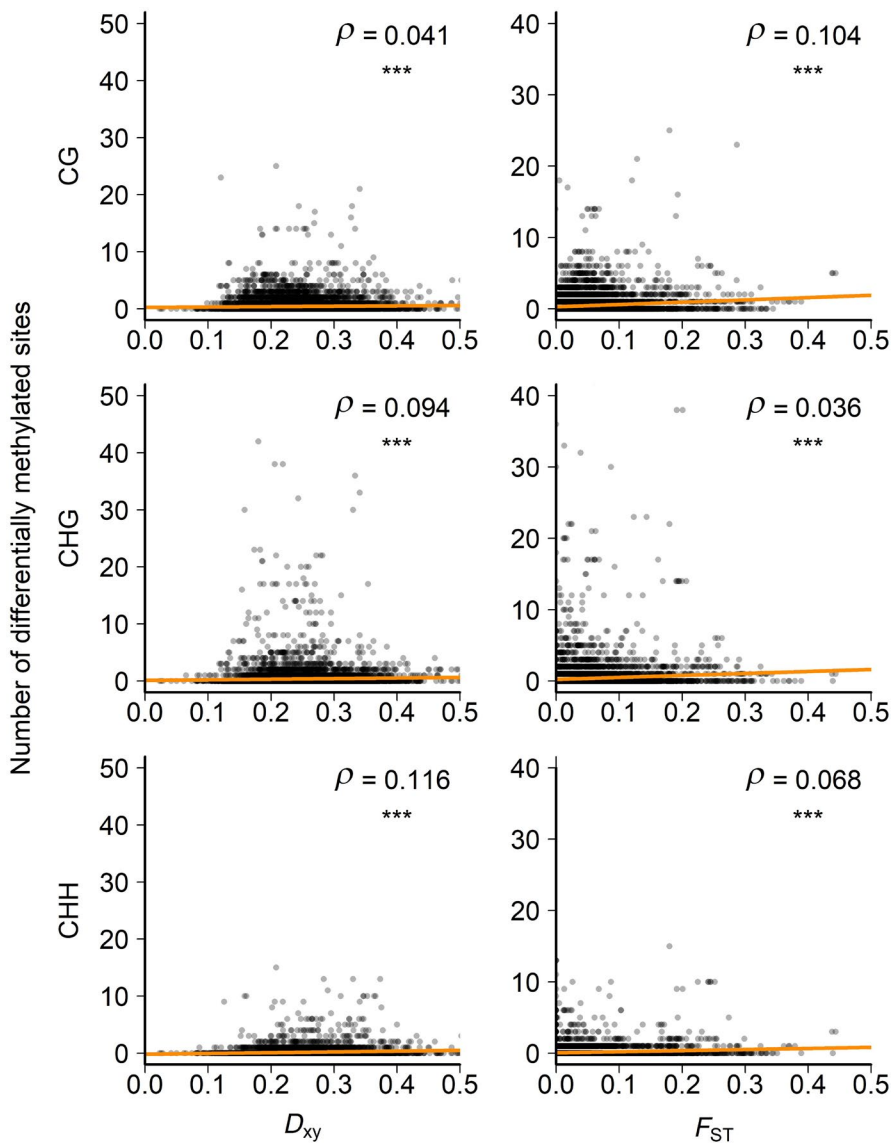


FIGURE 4 The relationship between the number of differentially methylated sites (DMSs) between spring- and control-originating individuals in complete 10 kb windows, and an absolute (D_{xy} ; left) or a relative (F_{ST} ; right) measure of sequence divergence for that window. Spearman's rank correlation coefficient is reported and p value indicated as; *** $p < .001$

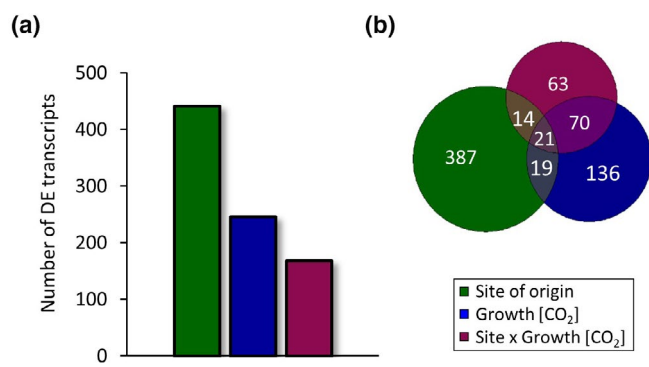


FIGURE 5 Differential expression of transcripts in individuals from four treatment groups. (a) Number of differentially expressed (DE) transcripts (FDR < 0.05) associated with population site of origin (spring vs. control), growth [CO₂] (ambient vs. elevated) and the interaction between the two. (b) Overlap between the number of DE transcripts associated with population site of origin, growth [CO₂] or an interaction between the two. Numbers describe transcripts belonging only to the category indicated

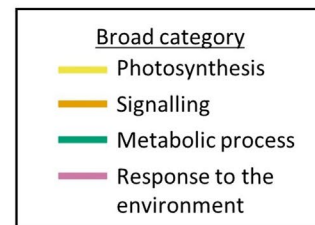
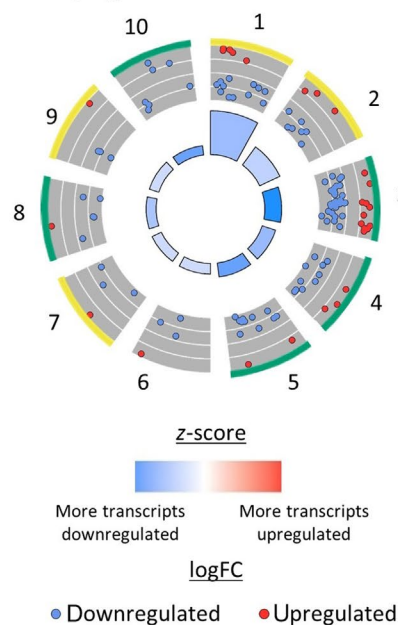
(4%). Of these, 64 DE transcripts corresponding to 59 components mapped to within 1 kb of at least one DMS (48 transcripts) or outlier SNP (21 transcripts), with five DE transcripts within 1 kb of both DMSs and outlier SNPs (Table S8). All 64 of the transcripts identified in this analysis were also DE in the Watson-Lazowski et al. (2016) analysis of this data. Transcripts mapping to within 1 kb of a DMS were enriched for those that were DE (Fisher's exact test; $p < .001$, odds ratio = 1.8) as were those mapping to within 1 kb of an SNP (Fisher's exact test; $p < .0001$, odds ratio = 4.9). Eight DMSs within the exon region of a gene were identified, as well as a weak association between CHG methylation across the exon region and the expression of that region (Figure S9; Table S9).

Among the 64 transcripts identified as co-locating to within 1 kb of a DMS or SNP, many were annotated as relevant to plant physiological response to elevated [CO₂] including abscisic acid response (three transcripts), stomatal movement (two transcripts), photosynthesis (two transcripts) and carbohydrate metabolism (three transcripts).

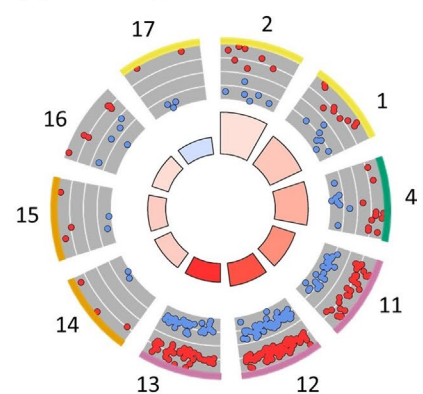
Three examples outlined here demonstrate the potential role of selection on DNA methylation and sequence polymorphisms in the

FIGURE 6 Gene ontology (GO) analysis of differentially expressed transcripts identifies key biological processes that show differential expression under single generation or multigenerational growth under elevated $[\text{CO}_2]$. (a) Circular plot highlighting differential transcript expression in the 10 GO categories that encompass the most differentially expressed genes between growth $[\text{CO}_2]$ treatments. Outer circle shows a scatter plot of each gene associated with the GO term and its log-fold change in expression with red and blue upregulated and downregulated in elevated $[\text{CO}_2]$ relative to ambient $[\text{CO}_2]$, respectively. The inner circle quadrilateral size is scaled by adjusted p -value of the GO term and is coloured by z -score, a crude measurement of up or downregulation of the category based on the number of genes that were up or downregulated. The colour of the outermost lines indicate broad categorizations of the GO terms. (b) Differentially expressed GO ontology categories in plants originating from spring versus control sites, with red and blue upregulated and downregulated in spring relative to control respectively. GO term identifiers for those categories visualized in (a) and (b) are given with significance of adjusted p -value indicated as; $**p < .01$; $***p < .001$

Num.	GO category ID	Adj. p -value	
		Growth $[\text{CO}_2]$	Site of origin
1	Photosynthesis	***	***
2	Photosynthesis, light reaction	***	***
3	Organonitrogen compound metabolic process	***	
4	Generation of precursor metabolites and energy	***	***
5	Cofactor metabolic process	***	
6	Reductive pentose-phosphate cycle	***	
7	Photosynthesis, dark reaction	***	
8	Chlorophyll metabolic process	***	
9	Photosynthesis, light harvesting photosystem I	***	
10	Pyridine-containing compound metabolic process	**	
11	Response to abiotic stimulus		***
12	Response to stimulus		***
13	Response to stress		**
14	Carbohydrate-mediated signalling		**
15	Sugar-mediated signalling pathway		**
16	Cofactor biosynthetic process		**
17	Photosynthesis, light harvesting		**

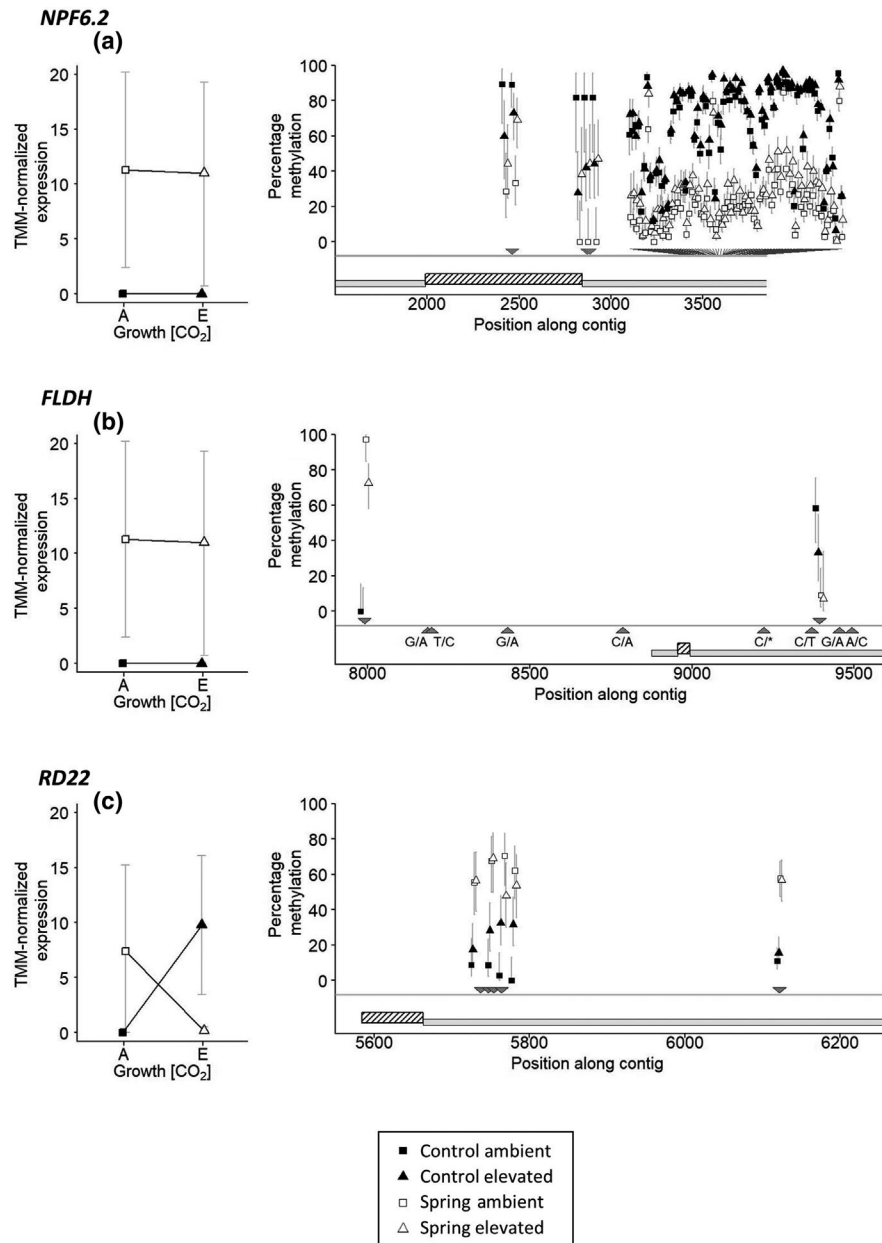
(a) Growth $[\text{CO}_2]$ effect

(b) Site of origin effect



multigenerational response of plants to elevated $[\text{CO}_2]$. One transcript co-located to a region with many DMSs. This was annotated as a nitrate transporter gene (*NPF6.2*) that plays a critical role in regulating leaf nitrate homeostasis in *A. thaliana* (Iqbal et al., 2020; Figure 7a). The potential *NPF6.2* orthologue was upregulated in the plants from the CO_2 spring, and this corresponded to a region of extensive demethylation upstream as well as along the transcript match. Another transcript (a putative orthologue of *FLDH*, encoding an NAD^+ -dependent dehydrogenase that regulates ABA signalling) had

a similar pattern of expression and co-located to a region of the genome containing several outlier SNPs, suggesting divergent selection at the sequence level (Figure 7b). A third example (a putative orthologue of *RD22*) is DE by site of origin, growth $[\text{CO}_2]$ and site \times growth $[\text{CO}_2]$, and exhibits a plastic increase in expression when grown in the non-ancestral $[\text{CO}_2]$ environment. *RD22* is a protein implicated in resistance to drought and salt stress is responsive to light and is induced by ABA (Goh, Nam, & Park, 2003; Iwasaki, Yamaguchi-Shinozaki, & Shinozaki, 1995). This component also co-locates to DMSs that were



hypermethylated in plants derived from the spring site population, but this differential methylation does not correlate with the observed expression differences (Figure 7c).

4 | DISCUSSION

This is the first study, to our knowledge, to capture genome, methylome and transcriptome responses in plants (analysed in half of the gene space) following multigenerational exposure to atmospheric $[\text{CO}_2]$ predicted over the coming decades. There was evidence of weak selection in exons of the population of *P. lanceolata* derived from the CO_2 spring compared to a nearby control site population, despite low overall differentiation. In methylome analysis, significant differences between the methylation profiles of spring and control plant populations were observed, with limited differential methylation induced by

single generation exposure to elevated $[\text{CO}_2]$. The identification of 59 regions of the genome where DE genes co-locate with regions of differential methylation and/or significant genetic divergence highlights both mechanisms may contribute to the adaptation to elevated $[\text{CO}_2]$. It also identifies candidate genes with a putative role in adaptation to future $[\text{CO}_2]$.

4.1 | Population genomics

4.1.1 | Low overall genomic differentiation between the spring- and control-derived populations

The low sequence divergence between spring- and control-derived populations calculated here is supported by previous calculations from transcriptome data (Watson-Lazowski et al., 2016) and

FIGURE 7 Selected differentially expressed transcripts that map to within 1 kb of a differentially methylated site (DMS) or single nucleotide polymorphism (SNP). These examples illustrate the potential role of methylation in plant adaptation to elevated $[\text{CO}_2]$ the potential for adaptation through selection on the genetic sequence and the potential for variation in methylation and expression to be a component of broader processes in the plant response to elevated $[\text{CO}_2]$. TMM-normalized expression of the four treatment groups based on site of origin (spring or control) and growth $[\text{CO}_2]$ (ambient or elevated; left-hand side) and a schematic of the location of DMS or SNPs within 1 kb of the region to which the differentially expressed transcripts map to in the fragmented genome (right-hand side). The location to which the transcript mapped is indicated with a striped box, but note that this represents the length of the match, not the length of the whole transcript (white boxes). The position of a DMS or SNP is indicated by a black triangle, with the reference/alternate nucleotide given for SNPs and the average percentage methylation with 95% confidence intervals (calculated with the adjusted Wald method) given for DMS: (a) putative nitrate transporter (NPF6.2), (b) NAD(P)-binding Rossmann-fold superfamily protein (FLDH) and (c) BURP domain-containing protein (RD22)

between populations of *P. lanceolata* in other studies (Bos, Harmens, & Vrieling, 1986; Tonsor, Kalisz, Fisher, & Holtsford, 1993; Van Dijk, Wolff, & De Vries, 1988). Limited pollen dispersal distance (1.5 m; Bos et al., 1986), passive seed dispersal (0.08 m; Bos et al., 1986) and obligate outcrossing likely results in high within population genetic diversity and high genomic heterogeneity (Gáspár, Bossdorf, & Durka, 2018), and therefore interpopulation divergence is generally low. This is especially likely given the proximity (200 m) of the in situ spring and control populations used in this study. Lower genetic diversity in the spring population suggests that the spring population originated from the control population.

4.2 | DNA methylation

4.2.1 | Divergence in methylation profiles between control and spring populations, with limited methylome remodelling under single generation exposure to elevated [CO₂]

Both methods identified significant methylation differences between control and spring populations and these differences were 10-fold greater than those induced by exposure to elevated [CO₂] for a single generation. The divergence in CG and CHG methylation profiles between the control and spring populations was stable even after at least one generation of growth at ambient [CO₂], highlighting that a significant proportion of the methylation differences are not (or no longer) responsive to [CO₂]. Furthermore, since these methylation differences occur at high enough frequency to distinguish the spring and control population, the causative underlying genetic or epigenetic variation must be subject to either selection or drift.

Some remodelling of the plant methylome in response to elevated [CO₂] was evident in this analysis. Environmentally induced methylation changes in wild non-model plants have previously been associated with transgenerational light environment (Baker, Berg, & Sultan, 2018) and exposure to transgenerational drought treatment (Alsdurf, Anderson, & Siemens, 2016). DNA methylation changes facilitated a trade-off between drought tolerance and defence (glucosinolate content) in a study by Alsdurf et al. (2016) that was postulated to allow range-shift in *Boechera stricta*. However, this is the first time the methylome has been shown to be responsive to elevated [CO₂]. Methylation in the CHH context was the most responsive to growth at elevated [CO₂]. Since methylation in the CHH context is erased and re-established de novo after replication (Gehring, 2019), CHH sites differentially methylated by site of origin are probably maintained due to association with genetic polymorphisms.

Around a quarter of DMSs associated with growth at elevated [CO₂] were also differentially methylated by site of origin (spring vs. control), highlighting the potential for environmentally induced CG and CHG methylation changes to be transgenerationally stable. However, this overlap accounted for just 1%–3% of the DMSs

associated with site of origin and there was relatively little response of CG and CHG methylation to growth [CO₂]. Furthermore, none of the DE transcripts (DE by any factor) co-located to sites that were differentially methylated by both growth [CO₂] and site of origin. The contribution of transgenerational inheritance of elevated [CO₂] induced methylation changes to adaptation to elevated [CO₂] might therefore be expected to be small in comparison to other mechanisms.

4.2.2 | Correlations between genetic and methylome variation only account for approximately 1% of the large differences in methylation between populations

Three possible processes could contribute to the large variation in methylation (particularly CG) between the two populations. Methylation variation may be associated with genetic variation either (a) in *cis* or (b) in *trans* and this is subject to selection or drift. Alternatively, (c) spontaneous epimutations may arise and reach high frequency in the population as a result of selection or drift.

We found that only 1% of the variation in number of DMSs in a 10 kb window was explained by a positive correlation with sequence divergence of that window, and that DE transcripts more frequently co-located to DMSs than to SNPs. This implies that genetic variation associated with methylation variation in *cis* has a relatively small contribution to these observed differences. The amount of DNA methylation variation explained by genetic variation has previously been estimated in *P. lanceolata* at 2%–3% from MSAP and AFLP markers but with limited resolution (Gáspár et al., 2018). However, these analyses (including this one) do not consider the relationship between larger structural variants with variation in methylation profiles, which may provide further insight (Eichten et al., 2016; Kawakatsu et al., 2016; Schmitz et al., 2013). This may be particularly important considering the role of methylation in silencing transposons and the link between transposon activity and mutation rates (Wicker et al., 2016).

Since only 13% of the *P. lanceolata* genome is analysed here it has not been possible to characterize the association between genetic variation and methylation variation acting in *trans*. A study in *Arabidopsis* revealed gene body CG methylation variation in populations along a longitudinal gradient was linked to trans-acting polymorphisms subject to selection (Dubin et al., 2015). However, a review by Vidalis et al. (2016) argues that over short time scales methylome evolution is more likely to be driven by spontaneous epimutations while long-term methylome evolution is driven by genomic changes. Establishing the relative contribution of each of these processes to the observed methylation differences between populations would facilitate a greater understanding of the evolutionary processes driving this variation. This could be achieved with the existing methylation dataset but will require a more complete genome assembly from long read sequencing. Whether these epimutations arise 'spontaneously' or whether they arise due to genetic variation, they could play a significant role in

coordinating plant response to elevated $[\text{CO}_2]$ if they impact the regulation of genes (Chinnusamy & Zhu, 2009).

Epigenetic mechanisms more broadly (including DNA methylation, histone modification or RNA interference) have well-documented roles in response to multigenerational abiotic stresses in some species (Lämke & Bäurle, 2017). Eriksson, Szukala, Tian, and Paun (2020) propose that epigenetic responses to such abiotic stimuli are complex with multiple pathways to modulate gene expression and are diffused across the genome rather than highly localized. This is consistent with our findings of DNA methylation changes in plant response to elevated $[\text{CO}_2]$. Furthermore, these studies imply that research into the roles of histone modification and small RNAs as well as DNA methylation in elevated $[\text{CO}_2]$ response is warranted, perhaps focussed initially around a panel of $[\text{CO}_2]$ responsive candidate genes (e.g. Johansson, El-Soda, Nilsson, Andersson, & Fredin, 2015; Rae, Ferris, Tallis, & Taylor, 2006).

4.3 | Gene expression and morphology

4.3.1 | Divergence between the response of spring- and control-derived plants to elevated $[\text{CO}_2]$: Evidence for adaptation

The overlap of genes that were DE between $[\text{CO}_2]$ treatments and between population sites of origin (14% and significantly more than expected by chance) suggests that at least some of the fixed differences in gene expression in the spring plant population were a result of adaptation to the elevated $[\text{CO}_2]$ at the spring site. Furthermore, since the direction of expression differences between control versus spring populations and between ambient versus elevated $[\text{CO}_2]$ is consistent for all 40 transcripts that were DE in both categories, the regulation of these transcripts may have been a target for selection.

Previous analysis of phenotypes in this set of plants identified a highly plastic response to elevated $[\text{CO}_2]$ in two of eight traits (epidermal cell number and above ground biomass; Watson-Lazowski et al., 2016), with growth at elevated $[\text{CO}_2]$ resulting in generally larger plants. A further three (stomatal index, single leaf dry mass and area) were significantly affected by population site of origin and this could be the result of adaptation to elevated $[\text{CO}_2]$ in the CO_2 spring population. In spring relative to control-site-derived plants; stomatal index was higher under elevated $[\text{CO}_2]$ treatment, single leaf area was higher in ambient $[\text{CO}_2]$ treatment, and single leaf dry weight was higher in both $[\text{CO}_2]$ treatments.

Adaptation to elevated $[\text{CO}_2]$ has previously been inferred in growth, biomass and gas exchange traits in several other plant species at natural CO_2 spring sites (Barnes et al., 1997; Nakamura et al., 2011). Although the mechanistic basis of this adaptation has not been well studied, a loss of plasticity in selected traits that are initially responsive to elevated $[\text{CO}_2]$ has been observed in CO_2 spring plant populations, in support of our own findings

(Barnes et al., 1997; Nakamura et al., 2011). Furthermore, a study of an annual plant species in a 7-year experiment identified genetic assimilation as the mechanism for this reduction in trait plasticity after multigenerational exposure (Grossman & Rice, 2014). However, the causal genetic changes responsible were not identified in this study, and it is possible that epimutations (which our data suggest are more common than genetic polymorphisms) could also facilitate a loss of trait plasticity, especially over such a small timeframe. The significant overlap between genes DE by site of origin and growth $[\text{CO}_2]$, and the reduced responsiveness of gene expression to growth $[\text{CO}_2]$ in spring plants (Watson-Lazowski et al., 2016), supports a role for a reduction in gene expression plasticity in the multigenerational plant response to elevated $[\text{CO}_2]$.

A larger proportion of transcripts were DE by population site of origin than growth $[\text{CO}_2]$. Although we postulate that these fixed differences in gene expression may have arisen as spontaneous genetic or epigenetic mutations to gene regulatory elements, it is important to consider that this could be the result of local adaptation conditioned by other environmental differences between sites, for example, pH differences of the soil (Körner & Miglietta, 1994) or altered microbial activity (Šibanc et al., 2018).

4.4 | Co-location of DMSs, SNPs and DE transcripts

4.4.1 | Co-location of DE transcripts to DMSs and high divergence SNPs highlight the potential role of differential methylation and selection in plant adaptation to elevated $[\text{CO}_2]$

Through the identification of DE transcripts and their co-location to DMSs and outlier SNPs, we highlight a potential role for methylation and genetic variation underlying differential expression, and potentially in plant adaptation to elevated $[\text{CO}_2]$. In all, 59 regions of the genome where DE transcripts and either DMS or F_{ST} outlier SNPs co-located were identified, irrespective of site of origin and $[\text{CO}_2]$ treatment. Examples include transcripts putatively orthologous to *FLDH* and *NPF6.2*. *FLDH* had fixed expression differences between spring and control plant populations such that gene expression was always higher for this gene in spring plants and mapped to a region rich in F_{ST} outlier SNPs, suggesting that this region could be under selection. It has been shown that the action of elevated $[\text{CO}_2]$ on both stomatal development and control of aperture may require an increase in ABA biosynthesis (Chater et al., 2015). *FLDH*, as a negative regulator of ABA signalling (Bhandari, Fitzpatrick, & Crowell, 2010), could represent a control point for stomatal responses to elevated $[\text{CO}_2]$, but this remains to be investigated. Two other transcripts that putatively function in the regulation of stomatal movement were DE by population site of origin and co-located to DMSs. In the context of increased stomatal index in plants from the spring population (Watson-Lazowski et al., 2016), this highlights potential adaptation

in pathways coordinating stomatal patterning and function in the spring plant population.

The putative orthologue of nitrate transporter *NPF6.2* had higher expression in spring-derived plants, corresponding to extensive demethylation upstream and across the mapped transcript, suggesting that demethylation (in spring plants) promotes transcription of the gene through chromatin conformational changes. NPF genes are a large family of nitrate transporters (mostly studied in *Arabidopsis*) with diverse functions in nitrate homeostasis (Buchner & Hawkesford, 2014). Since increased nitrogen use efficiency of plants grown under elevated $[\text{CO}_2]$ is widely observed in the literature (Ainsworth & Long, 2005), our data present an intriguing possibility that this gene could be part of a nitrogen transport response to future $[\text{CO}_2]$. For other transcripts, such as *RD22*, there is no straightforward correlation between differential expression and differential methylation, highlighting that they are both likely a product of other processes within the wider context of the genomic region.

Differentially Expressed transcripts identified as co-locating to DMSs and outlier SNPs in this study also included other examples with annotations relevant to plant response to elevated $[\text{CO}_2]$. More detailed characterization of these regions and their role in gene regulation is required to link changes in methylation to adaptive changes in gene expression, as well as the functional characterization of candidate genes for their role in elevated $[\text{CO}_2]$ response (Tirado-Magallanes, Rebbani, Lim, Pradhan, & Benoukraf, 2017).

This analysis would greatly benefit from an improved reference assembly, which would also facilitate an analysis of larger-scale genetic variation and provide greater resolution for identifying potential candidates. The current genome assembly facilitated the analysis of a proportion of the genome (13%, with 44% of the transcriptome mapping) where co-location of DMSs/outlier SNPs and DE genes will be missed if they map to different contigs that are, in reality, adjacent or in linkage disequilibrium. The co-location approach is also limited in that it can only detect where methylation could impact gene regulation through a *cis* acting mechanism and not those that act in *trans* (Niederhuth & Schmitz, 2017). Furthermore, identification of co-locating DE transcripts and DMSs/outlier SNPs is limited by estimates for LD, which is known to be highly heterogeneous across the genome (Gupta, Rustgi, & Kulwal, 2005). As genomic resources are further developed for non-model species found at CO_2 spring sites, we anticipate the identification of candidate genes whose expression are impacted by *cis* or *trans* methylation differences.

This study provides insight into the potential mechanisms coordinating plant plastic and evolutionary response to growth at elevated $[\text{CO}_2]$ over multiple generations. A crossed factored experiment was combined with multi-omic approaches to study these evolutionary mechanisms in a high CO_2 spring system. The results highlight a role for both genetic variation and methylome variation in plant adaptation to elevated $[\text{CO}_2]$. We found substantial divergence in methylation profiles between populations naturally exposed to elevated $[\text{CO}_2]$ at a CO_2 spring relative to a control population. This variation may be attributable to genetic variation in association with methylation variation in *cis* or *trans* or to evolutionary processes acting on

spontaneous epimutations. In addition, there was some responsivity of the methylome to growth at elevated $[\text{CO}_2]$, a proportion of which may be stably transgenerationally inherited. Taken together, these data, alongside plant phenotypic responses and altered gene expression profiles, suggest there is local adaptation to elevated $[\text{CO}_2]$ in these CO_2 spring populations. The co-location of DE genes to differential methylation and/or genetic polymorphisms provide examples of the potential role of these mechanisms in coordinating adaptation. A critical implication of this is that DNA methylation changes could facilitate adaptation to rising $[\text{CO}_2]$ even in the absence of high in-trapopulation genetic variation, with implications to conservation and crop-breeding. Of the 59 areas of the genome where DE genes co-locate to differential methylation or putatively divergent sequences, many were annotated as components of pathways responsive to elevated $[\text{CO}_2]$. For example, two DE transcripts in the stomatal patterning pathway co-located to these regions. Stomatal development and function is one of the few plant phenotypic traits where there is some evidence for adaptation to elevated $[\text{CO}_2]$ in both the fossil record (Franks & Beerling, 2009; Hetherington & Woodward, 2003) and studies of plants at natural CO_2 springs (Saban et al., 2018). Since changes in stomatal function and patterning have wider consequences for crop water use and hydrological cycling (Franks, Berry, Lombardozzi, & Bonan, 2017), these candidate genomic regions and evolutionary mechanisms are worthy of further exploration.

5 | CONCLUSION

Utilizing *P. lanceolata* seed from a natural CO_2 spring in a crossed factored experiment, combined with multi-omic technologies, we provide critical insight into the mechanisms coordinating the plastic and evolutionary response to elevated $[\text{CO}_2]$. Populations of plants growing at natural CO_2 springs show evidence of weak selection in exons but low overall divergence when compared to populations growing in nearby ambient $[\text{CO}_2]$ control sites. CO_2 spring plants also exhibit DNA methylation profiles that have substantially diverged from plants in the control population, as a result of three possible processes. In contrast, there was limited responsivity of the *P. lanceolata* methylome to growth at elevated $[\text{CO}_2]$, but some of the elevated $[\text{CO}_2]$ induced changes may be transgenerationally inherited. The co-location of transcripts that were DE in this experiment, with differential methylation or genetic polymorphisms provide specific examples that support a potential role of these mechanisms in coordinating adaptation to elevated $[\text{CO}_2]$.

ACKNOWLEDGEMENTS

This work was supported by the use of the IRIDIS High Performance Computing Facility and associated support services at the University of Southampton. We thank F. Miglietta for assisting with access to the CO_2 spring site and Y. Lin for experimental work. We also thank three anonymous reviewers whose insightful comments helped to substantially improve and clarify this manuscript. This work was

funded by the Natural Environmental Research Council (NERC) through a SPITFIRE DTP PhD studentship to J.M.S. (grant number NE/L002531/1), and by funding to G.T. from University of California, Davis and The British Council.

DATA AVAILABILITY STATEMENT

The raw whole genome sequencing and whole genome bisulfite sequencing data have been uploaded to the NCBI short read archive as BioProject PRJNA649873, alongside the contig genome and chloroplast assemblies. The corresponding RNA Seq raw reads are available from the NCBI short read archive BioProject PRJNA338760.

ORCID

Jasmine M. Saban  <https://orcid.org/0000-0001-6268-3925>

Alexander Watson-Lazowski  <https://orcid.org/0000-0002-2323-8287>

Mark A. Chapman  <https://orcid.org/0000-0002-7151-723X>

Gail Taylor  <https://orcid.org/0000-0001-8470-6390>

REFERENCES

- Ainsworth, E. A., & Long, S. P. (2005). What have we learned from 15 years of free-air CO₂ enrichment (FACE)? A meta-analytic review of the responses of photosynthesis, canopy properties and plant production to rising CO₂. *New Phytologist*, *165*(2), 351–372. <https://doi.org/10.1111/j.1469-8137.2004.01224.x>
- Akalin, A. (2016). *readBismarkFiles()* function. GitHub repository. Retrieved from <https://gist.github.com/al2na/4839e615e2401d73fe51>
- Akalin, A., Kormaksson, M., Li, S., Garrett-Bakelman, F. E., Figueroa, M. E., Melnick, A., & Mason, C. E. (2012). methylKit: A comprehensive R package for the analysis of genome-wide DNA methylation profiles. *Genome Biology*, *13*(10), R87. <https://doi.org/10.1186/gb-2012-13-10-r87>
- Alsdurf, J., Anderson, C., & Siemens, D. H. (2016). Epigenetics of drought-induced trans-generational plasticity: Consequences for range limit development. *AoB Plants*, *8*(146). <https://doi.org/10.1093/aobpla/plv146>
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, *215*(3), 403–410. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2)
- Baker, B. H., Berg, L. J., & Sultan, S. E. (2018). Context-dependent developmental effects of parental shade versus sun are mediated by DNA methylation. *Frontiers in Plant Science*, *9*(1251). <https://doi.org/10.3389/fpls.2018.01251>
- Barnaby, J. Y., & Ziska, L. H. (2012). Plant responses to elevated CO₂. *eLS*. <https://doi.org/10.1002/9780470015902.a0023718>
- Barnes, J. D., Bettarini, I., Polle, A., Snee, N., Raines, C., Miglietta, F., ... Fordham, M. (1997). The impact of elevated CO₂ on growth and photosynthesis in *Agrostis canina* L. ssp. *monteluccii* adapted to contrasting atmospheric CO₂ concentrations. *Oecologia*, *110*(2), 169–178. <https://doi.org/10.1007/s004420050146>
- 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 (Methodological)*, *57*(1), 289–300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>
- Berardini, T. Z., Reiser, L., Li, D., Mezheritsky, Y., Muller, R., Strait, E., & Huala, E. (2015). The *Arabidopsis* information resource: Making and mining the “gold standard” annotated reference plant genome. *Genesis*, *53*(8), 474–485. <https://doi.org/10.1002/dvg.22877>
- Bhandari, J., Fitzpatrick, A. H., & Crowell, D. N. (2010). Identification of a novel abscisic acid-regulated farnesol dehydrogenase from *Arabidopsis*. *Plant Physiology*, *154*(3), 1116–1127. <https://doi.org/10.1104/pp.110.157784>
- Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>
- Bos, M., Harmens, H., & Vrieling, K. (1986). Gene flow in *Plantago* l. Gene flow and neighbourhood size in *P. lanceolata*. *Heredity*, *56*(1), 43–54. <https://doi.org/10.1038/hdy.1986.7>
- Buchner, P., & Hawkesford, M. J. (2014). Complex phylogeny and gene expression patterns of members of the NITRATE TRANSPORTER 1/PEPTIDE TRANSPORTER family (NPF) in wheat. *Journal of Experimental Botany*, *65*(19), 5697–5710. <https://doi.org/10.1093/jxb/eru231>
- Campbell, M. S., Holt, C., Moore, B., & Yandell, M. (2014). Genome annotation and curation using MAKER and MAKER-P. *Current Protocols in Bioinformatics*, *48*(1), 4–11. <https://doi.org/10.1002/0471250953.bi0411s48>
- Cantarel, B. L., Korf, I., Robb, S. M., Parra, G., Ross, E., Moore, B., ... Yandell, M. (2008). MAKER: An easy-to-use annotation pipeline designed for emerging model organism genomes. *Genome Research*, *18*(1), 188–196. <https://doi.org/10.1101/gr.6743907>
- Chater, C., Peng, K., Movahedi, M., Dunn, J. A., Walker, H. J., Liang, Y.-K., ... Hetherington, A. M. (2015). Elevated CO₂-induced responses in stomata require ABA and ABA signaling. *Current Biology*, *25*(20), 2709–2716. <https://doi.org/10.1016/j.cub.2015.09.013>
- Chinnusamy, V., & Zhu, J. K. (2009). Epigenetic regulation of stress responses in plants. *Current Opinion in Plant Biology*, *12*(2), 133–139. <https://doi.org/10.1016/j.pbi.2008.12.006>
- Crisp, P. A., Ganguly, D., Eichten, S. R., Borevitz, J. O., & Pogson, B. J. (2016). Reconsidering plant memory: Intersections between stress recovery, RNA turnover, and epigenetics. *Science Advances*, *2*(2), e1501340. <https://doi.org/10.1126/sciadv.1501340>
- Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., ... Durbin, R. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*(15), 2156–2158. <https://doi.org/10.1093/bioinformatics/btr330>
- Dennis, G., Sherman, B. T., Hosack, D. A., Yang, J., Gao, W., Lane, H. C., & Lempicki, R. A. (2003). DAVID: Database for annotation, visualization, and integrated discovery. *Genome Biology*, *4*(9), R60. <https://doi.org/10.1186/gb-2003-4-9-r60>
- Dierckxnsens, N., Mardulyn, P., & Smits, G. (2017). NOVOPlasty: De novo assembly of organelle genomes from whole genome data. *Nucleic Acids Research*, *45*(4), e18.
- Doyle, J. J., & Doyle, J. L. (1987). CTAB DNA extraction in plants. *Phytochemical Bulletin*, *19*, 11–15.
- Dubin, M. J., Zhang, P., Meng, D., Remigereau, M., Osborne, E. J., Casale, F. P., ... Jean, G. (2015). DNA methylation in *Arabidopsis* has a genetic basis and shows evidence of local adaptation. *eLife*, *4*, e05255.
- Eichten, S. R., Schmitz, R. J., & Springer, N. M. (2014). Epigenetics: Beyond chromatin modifications and complex genetic regulation. *Plant Physiology*, *165*(3), 933–947. <https://doi.org/10.1104/pp.113.234211>
- Eichten, S. R., Stuart, T., Srivastava, A., Lister, R., & Borevitz, J. O. (2016). DNA methylation profiles of diverse *Brachypodium distachyon* aligns with underlying genetic diversity. *Genome Research*, *26*(11), 1520–1531. <https://doi.org/10.1101/gr.205468.116>
- Engineer, C. B., Hashimoto-Sugimoto, M., Negi, J., Israelsson-Nordström, M., Azoulay-Shemer, T., Rappel, W. J., ... Schroeder, J. I. (2016). CO₂ sensing and CO₂ regulation of stomatal conductance: Advances and open questions. *Trends in Plant Science*, *21*(1), 16–30. <https://doi.org/10.1016/j.tplants.2015.08.014>
- Eriksson, M. C., Szukala, A., Tian, B., & Paun, O. (2020). Current research frontiers in plant epigenetics: An introduction to a Virtual Issue.

- The New Phytologist*, 226(2), 285–288. <https://doi.org/10.1111/nph.16493>
- Finn, R. D., Coghill, P., Eberhardt, R. Y., Eddy, S. R., Mistry, J., Mitchell, A. L., ... Salazar, G. A. (2015). The Pfam protein families database: Towards a more sustainable future. *Nucleic Acids Research*, 44(1), 279–285. <https://doi.org/10.1093/nar/gkv1344>
- Fojtová, M., Kovařík, A., & Matyášek, R. (2001). Cytosine methylation of plastid genome in higher plants. Fact or artefact? *Plant Science*, 160(4), 585–593. [https://doi.org/10.1016/s0168-9452\(00\)00411-8](https://doi.org/10.1016/s0168-9452(00)00411-8)
- Foll, M., & Gaggiotti, O. (2008). A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. *Genetics*, 180(2), 977–993. <https://doi.org/10.1534/genetics.108.092221>
- Franks, P. J., & Beerling, D. J. (2009). Maximum leaf conductance driven by CO₂ effects on stomatal size and density over geologic time. *Proceedings of the National Academy of Sciences of the United States of America*, 106(25), 10343–10347. <https://doi.org/10.1073/pnas.0904209106>
- Franks, P. J., Berry, J. A., Lombardozi, D. L., & Bonan, G. B. (2017). Stomatal function across temporal and spatial scales: Deep-time trends, land-atmosphere coupling and global models. *Plant Physiology*, 174(2), 583–602. <https://doi.org/10.1104/pp.17.00287>
- Frenck, G., Linden, L., Mikkelsen, T. N., Brix, H., & Jørgensen, R. B. (2013). Response to multi-generational selection under elevated [CO₂] in two temperature regimes suggests enhanced carbon assimilation and increased reproductive output in *Brassica napus* L. *Ecology and Evolution*, 3(5), 1163–1172. <https://doi.org/10.1002/ece3.523>
- Gamage, D., Thompson, M., Sutherland, M., Hirotsu, N., Makino, A., & Seneweera, S. (2018). New insights into the cellular mechanisms of plant growth at elevated atmospheric carbon dioxide concentrations. *Plant, Cell and Environment*, 41(6), 1233–1246. <https://doi.org/10.1111/pce.13206>
- Ganguly, D., Crisp, P. A., Eichten, S. R., & Pogson, B. J. (2017). The *Arabidopsis* DNA methylome is stable under transgenerational drought stress. *Plant Physiology*, 175(4), 1893–2017. <https://doi.org/10.1104/pp.17.00744>
- Garg, R., Chevala, V. N., Shankar, R., & Jain, M. (2015). Divergent DNA methylation patterns associated with gene expression in rice cultivars with contrasting drought and salinity stress response. *Scientific Reports*, 5, 14922. <https://doi.org/10.1038/srep14922>
- Gáspár, B., Bossdorf, O., & Durka, W. (2018). Structure, stability and ecological significance of natural epigenetic variation: A large-scale survey in *Plantago lanceolata*. *New Phytologist*, 221, 1585–1596. <https://doi.org/10.1111/nph.15487>
- Gehring, M. (2019). Epigenetic dynamics during flowering plant reproduction: Evidence for reprogramming? *New Phytologist*, 224(1), 91–96. <https://doi.org/10.1111/nph.15856>
- Goh, C. H., Nam, H. G., & Park, Y. S. (2003). Stress memory in plants: A negative regulation of stomatal response and transient induction of *rd22* gene to light in abscisic acid-entrained *Arabidopsis* plants. *The Plant Journal*, 36(2), 240–255. <https://doi.org/10.1046/j.1365-313x.2003.01872.x>
- Grossman, J. D., & Rice, K. J. (2014). Contemporary evolution of an invasive grass in response to elevated atmospheric CO₂ at a Mojave Desert FACE site. *Ecology Letters*, 17(6), 710–716. <https://doi.org/10.1111/ele.12274>
- Gupta, P. K., Rustgi, S., & Kulwal, P. L. (2005). Linkage disequilibrium and association studies in higher plants: Present status and future prospects. *Plant Molecular Biology*, 57(4), 461–485. <https://doi.org/10.1007/s11103-005-0257-z>
- Haas, B. J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P. D., Bowden, J., ... Regev, A. (2013). *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature Protocols*, 8(8), 1494–1512. <https://doi.org/10.1038/nprot.2013.084>
- Haj, A., & O'Connor, D. H. (2016). *Scaffold Stitcher*. Bitbucket repository. Retrieved from <https://bitbucket.org/dholab/scaffoldstitcher.git>
- Haworth, M., Elliott-Kingston, C., & McElwain, J. C. (2011). Stomatal control as a driver of plant evolution. *Journal of Experimental Botany*, 62(8), 2419–2423. <https://doi.org/10.1093/jxb/err086>
- Heard, E., & Martienssen, R. A. (2014). Transgenerational epigenetic inheritance: Myths and mechanisms. *Cell*, 157(1), 95–109. <https://doi.org/10.1016/j.cell.2014.02.045>
- Henderson, I. R., & Jacobsen, S. E. (2007). Epigenetic inheritance in plants. *Nature*, 447(7143), 418–424. <https://doi.org/10.1038/nature05917>
- Hetherington, A. M., & Woodward, F. I. (2003). The role of stomata in sensing and driving environmental change. *Nature*, 424(6951), 901–908. <https://doi.org/10.1038/nature01843>
- IPCC. (2014). Summary for policymakers. In C. B. Field, V. R. Barros, D. J. Dokken, K. J. Mach, M. D. Mastrandrea, T. E. Bilir, M. Chatterjee, K. L. Ebi, Y. O. Estrada, R. C. Genova, B. Girma, E. S. Kissel, A. N. Levy, S. MacCracken, P. R. Mastrandrea, & L. L. White (Eds.), *Climate change 2014: Impacts, adaptation, and vulnerability. Part A: Global and sectoral aspects. Contribution of working group II to the fifth assessment report of the Intergovernmental Panel on Climate Change* (pp. 1–1127). Cambridge, UK, and New York, NY: Cambridge University Press.
- Iqbal, A., Qiang, D., Alamzeb, M., Xiangru, W., Huiping, G., Hengheng, Z., ... Meizhen, S. (2020). Untangling the molecular mechanisms and functions of nitrate to improve nitrogen use efficiency. *Journal of the Science of Food and Agriculture*, 100(3), 904–914. <https://doi.org/10.1002/jsfa.10085>
- Iwasaki, T., Yamaguchi-Shinozaki, K., & Shinozaki, K. (1995). Identification of a cis-regulatory region of a gene in *Arabidopsis thaliana* whose induction by dehydration is mediated by abscisic acid and requires protein synthesis. *Molecular and General Genetics*, 247(4), 391–398. <https://doi.org/10.1007/BF00293139>
- Johansson, K. S., El-Soda, M., Nilsson, A. K., Andersson, M. X., & Fredin, J. U. (2015). Molecular basis of variation in stomatal responsiveness to elevated CO₂. *Procedia Environmental Sciences*, 29, 210. <https://doi.org/10.1016/j.proenv.2015.07.266>
- Jones, P., Binns, D., Chang, H.-Y., Fraser, M., Li, W., McAnulla, C., ... Hunter, S. (2014). InterProScan 5: Genome-scale protein function classification. *Bioinformatics*, 30(9), 1236–1240. <https://doi.org/10.1093/bioinformatics/btu031>
- Jurka, J., Kapitonov, V. V., Pavlicek, A., Klonowski, P., Kohany, O., & Walichewicz, J. (2005). RepBase update, a database of eukaryotic repetitive elements. *Cytogenetic and Genome Research*, 110(1), 462–467. <https://doi.org/10.1159/000084979>
- Kawakatsu, T., Stuart, T., Valdes, M., Breakfield, N., Schmitz, R. J., Nery, J. R., ... Ecker, J. R. (2016). Unique cell-type-specific patterns of DNA methylation in the root meristem. *Nature Plants*, 2(5), 16058. <https://doi.org/10.1038/nplants.2016.58>
- Kleynhans, E. J., Otto, S. P., Reich, P. B., & Vellend, M. (2016). Adaptation to elevated CO₂ in different biodiversity contexts. *Nature Communications*, 7(1), 1–8. <https://doi.org/10.1038/ncomm12358>
- Kolde, R. (2013). *heatmap: Pretty heatmaps*. Retrieved from <http://CRAN.R-project.org/package=heatmap>
- Korf, I. (2004). Gene finding in novel genomes. *BMC Bioinformatics*, 5(1), 59. <https://doi.org/10.1186/1471-2105-5-59>
- Körner, C., & Miglietta, F. (1994). Long term effects of naturally elevated CO₂ on Mediterranean grassland and forest trees. *Oecologia*, 99(3), 343–351. <https://doi.org/10.1007/BF00627748>
- Krueger, F., & Andrews, S. R. (2011). Bismark: A flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics*, 27(11), 1571–1572. <https://doi.org/10.1093/bioinformatics/btr167>
- Lämke, J., & Bäurle, I. (2017). Epigenetic and chromatin-based mechanisms in environmental stress adaptation and stress memory in plants. *Genome Biology*, 18(124). <https://doi.org/10.1186/s13059-017-1263-6>
- Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, 9(4), 357–359. <https://doi.org/10.1038/nmeth.1923>

- Lawrence, M., Huber, W., Pages, H., Aboyoun, P., Carlson, M., Gentleman, R., ... Carey, V. J. (2013). Software for computing and annotating genomic ranges. *PLoS Computational Biology*, 9(8), e1003118. <https://doi.org/10.1371/journal.pcbi.1003118>
- Li, H. (2016). *seqtk: A fast and lightweight tool for processing sequences*. GitHub repository. Retrieved from <https://github.com/lh3/seqtk>
- Lohse, M., Drechsel, O., Kahlau, S., & Bock, R. (2013). Organellar GenomeDRAW – A suite of tools for generating physical maps of plastid and mitochondrial genomes and visualizing expression data sets. *Nucleic Acids Research*, 41(1), 575–581. <https://doi.org/10.1093/nar/gkt289>
- Loladze, I., Nolan, J. M., Ziska, L. H., & Knobbe, A. R. (2019). Rising atmospheric CO₂ lowers concentrations of plant carotenoids essential to human health: A meta-analysis. *Molecular Nutrition & Food Research*, 63(15), 1801047.
- Lu, X., Wang, X., Chen, X., Shu, N., Wang, J., Wang, D., ... Ye, W. (2017). Single-base resolution methylomes of upland cotton (*Gossypium hirsutum* L.) reveal epigenome modifications in response to drought stress. *BMC Genomics*, 18(1), 1–14. <https://doi.org/10.1186/s12864-017-3681-y>
- Luo, R., Liu, B., Xie, Y., Li, Z., Huang, W., Yuan, J., ... Wang, J. (2012). SOAPdenovo2: An empirically improved memory-efficient short-read *de novo* assembler. *Gigascience*, 1(1), 18. <https://doi.org/10.1186/2047-217X-1-18>
- Ma, N., Chen, W., Fan, T., Tian, Y., Zhang, S., Zeng, D., & Li, Y. (2015). Low temperature-induced DNA hypermethylation attenuates expression of RhAG, an AGAMOUS homolog, and increases petal number in rose (*Rosa hybrida*). *BMC Plant Biology*, 15(1). <https://doi.org/10.1186/s12870-015-0623-1>
- Martin, S. H. (2017). *VCF_processing*. GitHub repository. Retrieved from https://github.com/simonhmartin/genomics_general
- McCarthy, J. D., Chen, Y., & Smyth, K. G. (2012). Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Research*, 40(10), 4288–4297. <https://doi.org/10.1093/nar/gks042>
- Mi, H., Muruganujan, A., & Thomas, P. D. (2012). PANTHER in 2013: Modeling the evolution of gene function, and other gene attributes, in the context of phylogenetic trees. *Nucleic Acids Research*, 41(1), 377–386. <https://doi.org/10.1093/nar/gks1118>
- Miglietta, F., Raschi, A., Bettarini, I., Resti, R., & Selvi, F. (1993). Natural CO₂ springs in Italy: A resource for examining long-term response of vegetation to rising atmospheric CO₂ concentrations. *Plant, Cell and Environment*, 16(7), 873–878. <https://doi.org/10.1111/j.1365-3040.1993.tb00510.x>
- Monroe, J. G., Markman, D. W., Beck, W. S., Felton, A. J., Vahsen, M. L., & Pressler, Y. (2018). Ecoevolutionary dynamics of carbon cycling in the anthropocene. *Trends in Ecology & Evolution*, 33(3), 213–225. <https://doi.org/10.1016/j.tree.2017.12.006>
- Myers, S. S., Zanobetti, A., Kloog, I., Huybers, P., Leakey, A. D., Bloom, A. J., ... Holbrook, N. M. (2014). Increasing CO₂ threatens human nutrition. *Nature*, 510(7503), 139–142. <https://doi.org/10.1038/nature13179>
- Nakamura, I., Onoda, Y., Matsushima, N., Yokoyama, J., Kawata, M., & Hikosaka, K. (2011). Phenotypic and genetic differences in a perennial herb across a natural gradient of CO₂ concentration. *Oecologia*, 165(3), 809–818. <https://doi.org/10.1007/s00442-010-1900-1>
- Neves, D. M., Almeida, L. A. D. H., Santana-Vieira, D. D. S., Freschi, L., Ferreira, C. F., Soares Filho, W. D. S., ... Gesteira, A. D. S. (2017). Recurrent water deficit causes epigenetic and hormonal changes in citrus plants. *Scientific Reports*, 7(1), 13684. <https://doi.org/10.1038/s41598-017-14161-x>
- Niederhuth, C. E., Bewick, A. J., Ji, L., Alabady, M. S., Kim, K. D., Li, Q., ... Schmitz, R. J. (2016). Widespread natural variation of DNA methylation within angiosperms. *Genome Biology*, 17(1). <https://doi.org/10.1186/s13059-016-1059-0>
- Niederhuth, C. E., & Schmitz, R. J. (2017). Putting DNA methylation in context: From genomes to gene expression in plants. *Biochimica et Biophysica Acta (BBA) – Gene Regulatory Mechanisms*, 1860(1), 149–156. <https://doi.org/10.1016/j.bbarm.2016.08.009>
- Olmstead, R. G., & Reeves, P. A. (1995). Evidence for the polyphyly of the Scrophulariaceae based on chloroplast *rbcl* and *ndhF* sequences. *Annals of the Missouri Botanical Garden*, 82(2), 176–193. <https://doi.org/10.2307/2399876>
- Park, Y., & Wu, H. (2016). Differential methylation analysis for BS-seq data under general experimental design. *Bioinformatics*, 32(10), 1446–1453. <https://doi.org/10.1093/bioinformatics/btw026>
- Pearson, P. N., & Palmer, M. R. (2000). Atmospheric carbon dioxide concentrations over the past 60 million years. *Nature*, 406(6797), 695–699. <https://doi.org/10.1038/35021000>
- Price, A. L., Patterson, N. J., Plenge, R. M., Weinblatt, M. E., Shadick, N. A., & Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. *Nature Genetics*, 38(8), 904–909. <https://doi.org/10.1038/ng1847>
- Quadana, L., & Colot, V. (2016). Plant transgenerational epigenetics. *Annual Review of Genetics*, 50(2016), 467–491. <https://doi.org/10.1146/annurev-genet-120215-035254>
- R Core Team. (2015). *R: A language and environment for statistical computing* [Online]. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <https://www.R-project.org/>
- Rae, A. M., Ferris, R., Tallis, M. J., & Taylor, G. (2006). Elucidating genomic regions determining enhanced leaf growth and delayed senescence in elevated CO₂. *Plant, Cell and Environment*, 29(9), 1730–1741. <https://doi.org/10.1111/j.1365-3040.2006.01545.x>
- Richards, C. L., Alonso, C., Becker, C., Bossdorf, O., Bucher, E., Colomé-Tatché, M., ... Verhoeven, K. J. F. (2017). Ecological plant epigenetics: Evidence from model and non-model species, and the way forward. *Ecology Letters*, 20(12), 1576–1590. <https://doi.org/10.1111/ele.12858>
- Saban, J. M., Chapman, M. A., & Taylor, G. (2018). Face facts hold for multiple generations; Evidence from natural CO₂ springs. *Global Change Biology*, 25(1), 1–11.
- Schmitz, R. J., Schultz, M. D., Urlich, M. A., Nery, J. R., Pelizzola, M., Libiger, O., ... Ecker, J. R. (2013). Patterns of population epigenomic diversity. *Nature*, 495(7440), 193–198. <https://doi.org/10.1038/nature11968>
- Scholefield, P. A., Doick, K. J., Herbert, B. M. J., Hewitt, C. S., Schnitzler, J. P., Pinelli, P., & Loreto, F. (2004). Impact of rising CO₂ on emissions of volatile organic compounds: Isoprene emission from *Phragmites australis* growing at elevated CO₂ in a natural carbon dioxide spring. *Plant, Cell and Environment*, 27, 393–401. <https://doi.org/10.1111/j.1365-3040.2003.01155.x>
- Schultz, M. D., Schmitz, R. J., & Ecker, J. R. (2012). 'Leveling' the playing field for analyses of single-base resolution DNA methylomes. *Trends in Genetics*, 28(12), 583–585. <https://doi.org/10.1016/j.tig.2012.10.012>
- Šibanc, N., Zalar, P., Schroers, H. J., Zajc, J., Pontes, A., Sampaio, J. P., & Macek, I. (2018). *Occlutifur mephitis* fa, sp. nov. and other yeast species from hypoxic and elevated CO₂. *International Journal of Systematic and Evolutionary Microbiology*, 68(7), 2285–2298. <https://doi.org/10.1099/ijsem.0.002824>
- Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., & Zdobnov, E. M. (2015). BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*, 31(19), 3210–3212. <https://doi.org/10.1093/bioinformatics/btv351>
- Simpson, J. T., Wong, K., Jackman, S. D., Schein, J. E., Jones, S. J., & Birol, I. (2009). ABySS: A parallel assembler for short read sequence data. *Genome Research*, 19(6), 1117–1123. <https://doi.org/10.1101/gr.089532.108>
- Song, X., & Cao, X. (2017). Context and complexity: Analyzing methylation in trinucleotide sequences. *Trends in Plant Science*, 22(5), 351–353. <https://doi.org/10.1016/j.tplants.2017.03.013>

- Stanke, M., & Waack, S. (2003). Gene prediction with a hidden Markov model and a new intron submodel. *Bioinformatics*, *19*(2), 215–225. <https://doi.org/10.1093/bioinformatics/btg1080>
- Tillich, M., Lehwarck, P., Pellizzer, T., Ulbricht-Jones, E. S., Fischer, A., Bock, R., & Greiner, S. (2017). GeSeq – Versatile and accurate annotation of organelle genomes. *Nucleic Acids Research*, *45*(1), 6–11. <https://doi.org/10.1093/nar/gkx391>
- Tirado-Magallanes, R., Rebbani, K., Lim, R., Pradhan, S., & Benoukraf, T. (2017). Whole genome DNA methylation: Beyond genes silencing. *Oncotarget*, *8*(3), 5629–5637. <https://doi.org/10.18632/oncotarget.13562>
- Tonsor, S. J., Kalisz, S., Fisher, J., & Holtsford, T. P. (1993). A life-history based study of population genetic structure: Seed bank to adults in *Plantago lanceolata*. *Evolution*, *47*(3), 833–843. <https://doi.org/10.1111/j.1558-5646.1993.tb01237.x>
- Van der Auwera, G. A., Carneiro, M. O., Hartl, C., Poplin, R., Del Angel, G., Levy-Moonshine, A., ... Banks, E. (2013). From FastQ data to high-confidence variant calls: The genome analysis toolkit best practices pipeline. *Current Protocols in Bioinformatics*, *43*(1), 11.10.1–11.10.33. <https://doi.org/10.1002/0471250953.bi1110s43>
- Van Dijk, H., Wolff, K., & De Vries, A. N. D. A. (1988). Genetic variability in *Plantago* species in relation to their ecology. *Theoretical and Applied Genetics*, *75*(1), 518–528. <https://doi.org/10.1007/BF00276760>
- Verhoeven, K. J., Jansen, J. J., van Dijk, P. J., & Biere, A. (2010). Stress-induced DNA methylation changes and their heritability in asexual dandelions. *New Phytologist*, *185*(4), 1108–1118. <https://doi.org/10.1111/j.1469-8137.2009.03121.x>
- Vidalis, A., Živković, D., Wardenaar, R., Roquis, D., Tellier, A., & Johannes, F. (2016). Methylome evolution in plants. *Genome Biology*, *17*(1), 264. <https://doi.org/10.1186/s13059-016-1127-5>
- Walter, W., Sánchez-Cabo, F., & Ricote, M. (2015). GOplot: An R package for visually combining expression data with functional analysis. *Bioinformatics*, *31*(17), 2912–2914. <https://doi.org/10.1093/bioinformatics/btv300>
- Ward, J. K., Antonovics, J., Thomas, R. B., & Strain, B. R. (2000). Is atmospheric CO₂ a selective agent on model C₃ annuals? *Oecologia*, *123*(3), 330–341. <https://doi.org/10.1007/s004420051019>
- Watson-Lazowski, A., Lin, Y., Miglietta, F., Edwards, R. J., Chapman, M. A., & Taylor, G. (2016). Plant adaptation or acclimation to rising CO₂? Insight from first multigenerational RNA-Seq transcriptome. *Global Change Biology*, *22*(11), 3760–3773. <https://doi.org/10.1111/gcb.13322>
- Wicker, T., Yu, Y., Haberer, G., Mayer, K. F., Marri, P. R., Rounsley, S., ... Roffler, S. (2016). DNA transposon activity is associated with increased mutation rates in genes of rice and other grasses. *Nature Communications*, *7*(1), 1–9. <https://doi.org/10.1038/ncomms12790>
- Wong, C., & Murray, B. G. (2012). Variable changes in genome size associated with different polyploid events in *Plantago* (Plantaginaceae). *Journal of Heredity*, *103*(5), 711–719. <https://doi.org/10.1093/jhered/ess049>
- Xia, H., Huang, W., Xiong, J., Tao, T., Zheng, X., Wei, H., ... Luo, L. (2016). Adaptive epigenetic differentiation between upland and lowland rice ecotypes revealed by methylation-sensitive amplified polymorphism. *PLoS One*, *11*(17), e0157810. <https://doi.org/10.1371/journal.pone.0157810>
- Xu, Z., Jiang, Y., Jia, B., & Zhou, G. (2016). Elevated-CO₂ response of stomata and its dependence on environmental factors. *Frontiers in Plant Science*, *7*, 657. <https://doi.org/10.3389/fpls.2016.00657>
- Yaish, M. W., Al-Lawati, A., Al-Harrasi, I., & Patankar, H. V. (2018). Genome-wide DNA methylation analysis in response to salinity in the model plant *Medicago truncatula*. *BMC Genomics*, *19*, 78. <https://doi.org/10.1186/s12864-018-4484-5>
- Yong-Villalobos, L., Cervantes-Pérez, S. A., Gutiérrez-Alanis, D., González-Morales, S., Martínez, O., & Herrera-Estrella, L. (2016). Phosphate starvation induces DNA methylation in the vicinity of cis-acting elements known to regulate the expression of phosphate-responsive genes. *Plant Signalling & Behaviour*, *11*(5), e1173300. <https://doi.org/10.1080/15592324.2016.1173300>
- Zhang, C., Dong, S. S., Xu, J. Y., He, W. M., & Yang, T. L. (2019). PopLDdecay: A fast and effective tool for linkage disequilibrium decay analysis based on variant call format files. *Bioinformatics*, *35*(10), 1786–1788. <https://doi.org/10.1093/bioinformatics/bty875>
- Zhang, M., Kimatu, J. N., Xu, K., & Liu, B. (2010). DNA cytosine methylation in plant development. *Journal of Genetics and Genomics*, *38*(3), 1–12. [https://doi.org/10.1016/S1673-8527\(09\)60020-5](https://doi.org/10.1016/S1673-8527(09)60020-5)

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Saban JM, Watson-Lazowski A, Chapman MA, Taylor G. The methylome is altered for plants in a high CO₂ world: Insights into the response of a wild plant population to multigenerational exposure to elevated atmospheric [CO₂]. *Glob Change Biol*. 2020;00:1–19. <https://doi.org/10.1111/gcb.15249>